



**UNIVERSIDAD PONTIFICIA COMILLAS
MADRID**

**Escuela Técnica Superior de Ingenieros Industriales (ICAI)
Departamento de Electrónica y Automática**

**APLICACION DE
LA TEORIA DE CONJUNTOS BORROSOS
AL DIAGNOSTICO DE PROCESOS INDUSTRIALES**

Tesis para la obtención del doctorado

Director: Dr. D. Miguel Angel Sanz Bobi

Autor: Ing. D. José Villar Collado

Madrid 1997

Resumen de la tesis

La tesis doctoral que aquí se presenta se enmarca dentro de la teoría de conjuntos borrosos por un lado, y dentro del diagnóstico y mantenimiento de procesos industriales por otro.

En primer lugar se hace un análisis detallado de la inferencia en el marco de la lógica borrosa, con vistas a caracterizar desde un punto de vista semántico los distintos operadores necesarios para inferir conclusiones. Esta caracterización permite modelar distintas relaciones semánticas entre los antecedentes y consecuentes de una regla de conocimiento proporcionando una mayor riqueza y flexibilidad a la representación del conocimiento.

El estudio de la inferencia ha permitido además diseñar una estrategia de razonamiento en dos niveles. El primero permite obtener conclusiones con grados de certidumbre cualitativamente altos, lo que a su vez exige observaciones o hechos precisos y con grados de certidumbre también elevados. Cuando los hechos son por algún motivo incompletos o con grados de certidumbre bajos, el primer nivel no permite inferir ningún tipo de conclusión. Por ello se recurre al segundo nivel, capaz de inferir a partir de observaciones incompletas, pero por contra sus conclusiones tienen grados de certidumbre cualitativamente menores, relacionados con el grado de contradicción entre las hipótesis de las reglas y las observaciones disponibles. Estos dos niveles se basan en dos modos de razonamiento distintos pero complementarios, que hemos denominado respectivamente razonamiento en el dominio de lo necesario y razonamiento en el dominio de lo posible.

En segundo lugar se propone una metodología para el desarrollo de un sistema experto de diagnóstico basado en reglas con conocimiento borroso. Para ello se obtienen las reglas que describen el conocimiento disponible, asociando los síntomas que puedan aparecer en el proceso con las causas que los hayan originado. El conjunto de reglas debe ser lo más completo posible, de forma que contemple el mayor número de condiciones anómalas que el proceso pueda presentar, lo que a su vez facilita la determinación sistemática de las relaciones semánticas existentes entre cada antecedente y consecuente de cada una de las reglas. El sistema de reglas obtenido puede entonces ser utilizado para inferir tanto en el dominio de lo necesario como en el dominio de lo posible, recurriendo a este último cuando las conclusiones del primero no sean suficientemente satisfactorias. Además el cálculo de las conclusiones se lleva a cabo mediante un mecanismo de inferencia simplificado, basado en dos índices de similitud entre hipótesis y observaciones, que resulta ser más rápido y fácil de interpretar que la tradicional regla composicional de inferencia.

Agradecimientos

Quisiera expresar en esta página mi más sincero agradecimiento a todas aquellas personas que de uno u otro modo han contribuido a la realización de esta tesis.

En primer lugar mi especial gratitud a mis padres por todos aquellos sacrificios que sin duda mi formación ha supuesto para ellos, y a toda mi familia y a Lourdes por el cariño y apoyo moral que han hecho más llevadero este trabajo.

A mi director de tesis y amigo, Miguel Angel Sanz Bobi, por su confianza y su ayuda incondicional, por haber soportado estoicamente mi carácter en ocasiones difícil, y por haber sabido guiar mi trabajo a lo largo de todo su desarrollo.

A mi inseparable amigo Antonio Muñoz por ayudarme, aguantarme y darme ánimos en muchos de los momentos delicados.

A todos mis amigos y compañeros del IIT (cuya enumeración sería afortunadamente demasiado larga) por el apoyo y amistad que siempre me han prestado.

Índice general

1. INTRODUCCION	1
1.1 Descripción del problema	3
1.2 Planteamiento de la tesis	6
1.3 Organización de la exposición.....	7
2. SISTEMAS DE DIAGNOSTICO: CONOCIMIENTO, INCERTIDUMBRE E INFERENCIA	11
2.1 Organización del capítulo.....	13
2.2 Introducción a los sistemas de diagnóstico.....	14
2.3 Conocimiento e incertidumbre en un sistema de diagnóstico.....	20
2.3.1 El conocimiento requerido para diagnosticar.....	21
2.3.2 La incertidumbre del conocimiento.....	27
2.3.2.1 Fuentes de incertidumbre	27
2.3.2.2 Tipos de incertidumbre	29
2.4 Estrategias de diagnóstico.....	32
2.4.1 Estrategias de diagnóstico basadas en la naturaleza del razonamiento causal llevado a cabo...32	
2.4.2 Diagnóstico en discreto y en continuo	33
2.4.3 Abducción y deducción.....	35
2.4.4 Tipo de conocimiento y estrategias de diagnóstico	36
2.5 Técnicas de representación del conocimiento.....	39
2.5.1 Lógica formal.....	40
2.5.2 Esquemas estructurados: Frames, scripts y objetos.....	45
2.5.3 Redes semánticas	47
2.5.4 Reglas de producción	47
2.5.5 Modelos.....	50
2.6 Técnicas de representación de incertidumbre e inferencia.....	51
2.6.1 Las medidas de incertidumbre	52
2.6.2 Las medidas de probabilidad	53
2.6.2.1 Definición.....	53
2.6.2.2 Aplicación de la probabilidad a la inferencia: el enfoque Bayesiano	54
2.6.3 Los factores de certeza	58
2.6.4 Las medidas de credibilidad y plausibilidad.....	60
2.6.4.1 Definiciones.....	60
2.6.4.2 Aplicación de la teoría de la evidencia a la inferencia	64
2.6.5 Las medidas de posibilidad y necesidad, la lógica borrosa y las distribuciones de posibilidad 66	
2.7 Sistemas expertos de diagnóstico.....	67

3. TEORIA DE CONJUNTOS BORROSOS Y DISTRIBUCIONES DE POSIBILIDAD	73
3.1 Organización del capítulo.....	75
3.2 Breve recorrido histórico.....	76
3.3 Conjuntos borrosos y definiciones básicas.....	78
3.3.1.1 Igualdad de conjuntos.....	79
3.3.1.2 Inclusión entre conjuntos.....	80
3.3.1.3 Núcleo y soporte de un conjunto borroso	80
3.3.1.4 α -cortes de un conjunto borroso.....	81
3.3.1.5 Altura de un conjunto borroso.....	81
3.3.1.6 Cardinal de un conjunto.....	82
3.3.1.7 Conjuntos borrosos convexos.....	82
3.3.1.8 Números borrosos.....	82
3.3.1.9 Conjuntos borrosos paramétricos.....	82
3.4 Operaciones con conjuntos borrosos.....	85
3.4.1 La negación de conjuntos borrosos.....	85
3.4.1.1 Negación estricta.....	85
3.4.1.2 Negaciones fuertes	86
3.4.1.3 Ejemplos de operadores de negación.....	87
3.4.2 La intersección y la unión de conjuntos borrosos.....	87
3.4.2.1 Propiedades.....	88
3.4.2.2 Las t-normas y t-conormas	89
3.4.2.3 Las pseudo-conjunciones y pseudo-disyunciones.....	95
3.4.2.4 Otras uniones e intersecciones	97
3.4.2.5 Resumen.....	99
3.4.3 Modificadores lingüísticos.....	101
3.4.4 El principio de extensión	102
3.4.4.1 El álgebra borrosa.....	103
3.5 Relaciones borrosas.....	105
3.5.1 Definición	105
3.5.2 Propiedades.....	105
3.5.2.1 Reflexividad.....	105
3.5.2.2 Simetría	105
3.5.2.3 Antisimetría	105
3.5.2.4 A-transitividad.....	106
3.5.2.5 Comparabilidad.....	106
3.5.2.6 Clasificación de relaciones según sus propiedades.....	106
3.5.3 Proyección de una relación.....	106
3.5.4 Extensión cilíndrica de una relación	107
3.5.5 Composición de relaciones borrosas	107
3.5.6 Algunas relaciones borrosas comunes.....	108
3.6 Distribuciones de posibilidad.....	111
3.6.1 Definición	111
3.6.2 Especificación de una distribución de posibilidad.....	111
3.6.3 Agregación de distribuciones de posibilidad.....	115
3.6.3.1 Distribuciones definidas en un único universo.....	115
3.6.3.2 Distribuciones definidas en universos distintos.....	119

3.6.4 Modelado de la verdad, imprecisión, vaguedad e incertidumbre mediante distribuciones de posibilidad y conjuntos borrosos	120
3.6.5 Algunas distribuciones de posibilidad particulares	124
3.6.6 Valor de verdad de una afirmación cuando los hechos vienen modelados con una distribución de posibilidad	126
3.6.6.1 Valor de verdad borroso de una afirmación	127
3.6.6.2 Medidas de necesidad y posibilidad	128
3.6.6.3 Valores de verdad borrosos y medidas de necesidad y posibilidad	130
3.7 Razonamiento aproximado en lógica borrosa con distribuciones de posibilidad.....	136
3.7.1 Propositiones vagas o borrosas	136
3.7.2 Reglas de inferencia en lógica borrosa.....	137
3.8 El estado del arte en la aplicación de la lógica borrosa a los sistemas expertos de diagnóstico	145
3.8.1 Ecuaciones relacionales borrosas	146
3.8.2 Enfoques lógicos y modelado de valores de verdad con números borrosos	148
3.8.3 Cálculo de significados con reglas borrosas, el enfoque numérico.....	149
3.8.4 Medidas de incertidumbre y reglas simbólicas, el enfoque simbólico	150
4. ESTUDIO DE LA INFERENCIA EN LOGICA BORROSA.....	153
4.1 Organización del capítulo.....	155
4.2 La regla del modus ponens	157
4.3 La regla composicional de inferencia	160
4.3.1 Condicional borroso basado en la RCI: modelos de reglas	160
4.3.2 Funciones de implicación y funciones generadoras de modus ponens	162
4.3.2.1 Implicaciones clásicas.....	162
4.3.2.2 Implicaciones borrosas y funciones generadoras de modus ponens.....	163
4.3.2.3 Relación de inclusión entre la conclusión y el consecuente.....	165
4.3.3 Generación de funciones de implicación.....	166
4.3.3.1 Principales funciones de implicación.....	166
4.3.3.2 Propiedades	168
4.3.3.3 Otras funciones de implicación	172
4.3.4 Reformulación de la RCI mediante la compatibilidad borrosa.....	173
4.4 La modificación por similitud entre proposiciones	174
4.5 La modificación del consecuente ante observación puntual	176
4.5.1 n-implicaciones y observación puntual.....	177
4.5.1.1 r-implicaciones.	177
4.5.1.2 s-implicaciones.	178
4.5.1.3 La implicación de Yager.....	179
4.5.1.4 Las implicaciones cuánticas.....	180
4.5.2 p-implicaciones y observación puntual.....	181
4.5.2.1 t-normas	181
4.5.2.2 pseudo-conjunciones	182
4.5.2.3 pseudo-conjunción de Yager.....	183
4.5.3 Resumen.....	184
4.6 Condiciones para la equivalencia entre la RCI y la MS	185
4.6.1 n-implicaciones borrosas	185
4.6.1.1 El modelo (M_R^T, I_R^T)	185
4.6.1.2 El modelo (M_S^T, I_S^T)	186

4.6.1.3 El modelo de Yager (M_Y, I_Y).....	187
4.6.2 p-implicaciones borrosas	188
4.6.2.1 El modelo (M_R^T, M_R^T).....	188
4.6.2.2 El modelo (M_S^T, M_S^T).....	189
4.6.2.3 El modelo (M_R^T, M_S^T).....	190
4.6.2.4 El modelo (M_S^T, M_R^T).....	192
4.6.2.5 El modelo de Yager (M_Y, M_Y).....	193
4.6.3 Influencia de la FGMP	194
4.6.4 Resumen	195
4.6.4.1 n-modelos.....	195
4.6.4.2 p-modelos	196
4.6.5 Nueva familia de modelos de regla	199
4.7 Núcleo y soporte de las conclusiones según el tipo de modelo.....	201
4.7.1 Núcleos de los modelos basados en n-implicaciones	201
4.7.1.1 El modelo (M_R^T, I_R^T).....	201
4.7.1.2 El modelo (M_S^T, I_S^T).....	202
4.7.2 Soportes de los modelos basados en p-implicaciones	203
4.7.2.1 El modelo (M_R^T, M_R^T).....	203
4.7.2.2 El modelo (M_S^T, M_S^T).....	203
4.7.2.3 El modelo (M_R^T, M_S^T).....	203
4.7.2.4 El modelo (M_S^T, M_R^T).....	204
4.7.3 Resumen	205
4.8 Condiciones de activación de modelos de reglas bajo la RCI.....	207
4.8.1 La implicación material clásica y la implicación material clásica restringida o producto cartesiano	207
4.8.2 La conclusión trivial de p y n-implicaciones	209
4.8.3 Condición suficiente para la no activación de un modelo	209
4.8.4 n-implicaciones borrosas	210
4.8.4.1 El modelo (M_R^T, I_R^T).....	211
4.8.4.2 El modelo (M_S^T, I_S^T).....	214
4.8.5 p-implicaciones borrosas	216
4.8.5.1 El modelo (M_R^T, M_R^T).....	217
4.8.5.2 El modelo (M_S^T, M_S^T).....	218
4.8.5.3 El modelo (M_R^T, M_S^T).....	220
4.8.5.4 El modelo (M_S^T, M_R^T).....	221
4.8.6 Resumen	221
4.9 Interpretación del comportamiento global de los modelos analizados	223
4.9.1 Los modelos de necesidad.....	225
4.9.1.1 Grado de verdad y grado de certidumbre	226
4.9.1.2 Modelos de necesidad graduales	228
4.9.1.3 Modelos de necesidad de certidumbre	230
4.9.1.4 Modelos de necesidad mixtos	231
4.9.1.5 Modificación de las condiciones de activación:.....	233
4.9.2 Los modelos de posibilidad.....	234
4.9.2.1 Posible grado de verdad y posible grado de certidumbre	235
4.9.2.2 Modelos de posibilidad graduales	240
4.9.2.3 Modelos de posibilidad de certidumbre.....	241
4.9.2.4 Modelos de posibilidad mixtos.....	243
4.9.2.5 Modificación de las condiciones de activación.....	244

4.10 Medidas generalizadas de posibilidad y necesidad	245
4.10.1 Definiciones.....	245
4.10.2 Dualidad.....	247
4.11 Simplificación de la RCI para su aplicación práctica a sistemas de reglas borrosas: la MS de dos índices.....	252
4.11.1 Modelo general de necesidad.....	252
4.11.2 Modelo general de posibilidad.....	255
4.11.3 Representantes semánticos	258
4.11.3.1 Representantes semánticos de necesidad	258
4.11.3.2 Representantes semánticos de posibilidad.....	260
5. SISTEMAS DE REGLAS BORROSAS	265
5.1 Organización del capítulo.....	267
5.2 Tipos de modelos de reglas y tipos de distribuciones de posibilidad.....	269
5.2.1 Razonamiento con una regla	269
5.2.2 Razonamiento con reglas encadenadas	272
5.2.3 Razonamiento con información incompleta.....	275
5.3 Agregación de varias observaciones relativas a una misma variable.....	278
5.3.1 Agregación de varias observaciones	278
5.3.2 Resolución de incoherencias	292
5.4 Agregación de varios antecedentes en una misma regla.....	294
5.4.1 Reglas con múltiples antecedentes.....	294
5.4.2 Agregación de antecedentes y condiciones de activación	296
5.4.3 La MS con múltiples antecedentes	298
5.5 La agregación conjuntiva de la observación y el condicional de una regla borrosa.....	303
5.6 Sistemas de reglas paralelas y particiones borrosas.....	307
5.6.1 Inferencia con reglas paralelas y la RCI: inferencias FATI y FITA	307
5.6.2 Inferencia con reglas paralelas y la MS.....	314
5.6.3 Particiones, completitud y coherencia de reglas paralelas	314
5.6.3.1 Particiones borrosas	314
5.6.3.2 Particiones borrosas, verificación del MPE y completitud del sistema de reglas.....	316
5.6.3.3 Coherencia de un sistema de reglas paralelas	325
5.7 Modelos de necesidad y posibilidad semánticamente duales	328
5.7.1 Modelos graduales.....	328
5.7.2 Modelos de certidumbre.....	332
5.7.3 Modelos mixtos.....	335
5.8 Sistemas de reglas encadenadas	337
5.8.1 Encadenamiento de reglas y transitividad de condicionales	337
5.8.2 Orden de activación de reglas en un sistema borroso.....	338

6. VALORES DE VERDAD BORROSOS	341
6.1 Organización del capítulo.....	343
6.2 Definición.....	344
6.3 Compatibilidad borrosa, valores borrosos de necesidad y posibilidad.....	346
6.4 Interpretación de la compatibilidad borrosa a través de su relación con las medidas de grado de verdad y de certidumbre	349
6.5 Relación de la compatibilidad borrosa con las medidas de posibilidad y necesidad generalizadas	353
6.6 Valores de posibilidad y necesidad borrosos e inferencia.....	355
6.7 Familia general de valores de verdad borrosos	365
6.8 Agregación de valores de verdad borrosos.....	370
6.8.1 Agregación de valores de necesidad borrosos	370
6.8.2 Agregación de valores de posibilidad borrosos	372
7. METODOLOGIA DE DISEÑO DE UN SISTEMA DE DIAGNOSTICO BASADO EN LA TEORIA DE CONJUNTOS BORROSOS	375
7.1 Organización del capítulo.....	377
7.2 Pasos necesarios para la construcción de un sistema de diagnóstico.....	378
7.3 Identificación de componentes y medidas del proceso.....	381
7.3.1 Componentes del proceso	381
7.3.2 Medidas del proceso	384
7.3.2.1 Descripción del valor de una medida.....	387
7.3.2.2 Estados de una medida.....	391
7.4 Módulo de detección de anomalías y su integración en el sistema de diagnóstico.....	395
7.4.1 Obtención del grado de anomalía de un componente	396
7.4.2 Obtención de la certidumbre sobre el grado de anomalía de un componente	399
7.5 El módulo de identificación de anomalías: sistema experto de diagnóstico y reglas de conocimiento	403
7.5.1 El sistema experto de diagnóstico.....	403
7.5.2 Estrategias de diagnóstico.....	404
7.5.2.1 Estrategia de diagnóstico en continuo	404
7.5.2.2 Conclusiones necesarias y conclusiones posibles	404
7.5.2.3 Relaciones semánticas entre consecuentes y antecedentes	404
7.5.3 El motor de inferencias	405
7.5.3.1 Obtención de conclusiones necesarias y conclusiones posibles.....	405
7.5.3.2 Orden de disparo de las reglas del sistema experto.....	407
7.5.4 Estructura, formato y disparo de las reglas del sistema experto	410
7.5.4.1 Estructura general de una regla de conocimiento.....	410
7.5.4.2 Regla de inferencia, modelos de reglas y relaciones semánticas	412
7.5.4.3 Determinación de las conclusiones de una regla.....	424
7.5.4.4 Mecanismo propuesto para la determinación de los valores de verdad y certidumbre de un consecuente: esquema general de inferencia simplificado.....	428
7.5.4.5 Tipos de antecedentes.....	431
7.5.4.6 Tipos de consecuentes	437
7.5.4.7 Diagnósticos.....	438
7.5.4.8 Tipos de acciones.....	439

7.6 Caso ejemplo: diagnóstico de un transformador de potencia mediante el análisis de los gases disueltos en el aceite	441
7.6.1 Principales fallos detectables mediante el análisis de los gases disueltos en el aceite	442
7.6.2 Las medidas utilizadas y sus estados.....	446
7.6.3 Reglas de conocimiento.....	450
7.6.4 Determinación de los modelos de las reglas de conocimiento	453
7.6.5 Inferencia	472
8. CONCLUSIONES, APORTACIONES Y LINEAS DE FUTUROS DESARROLLOS	487
8.1 Conclusiones	489
8.2 Aportaciones de la tesis	493
8.3 Líneas de futuros desarrollos.....	497
A. LOS φ-CONDICIONALES	501
A.1 φ -implicaciones y φ -condicionales.....	503
A.1.1 Definiciones	503
A.1.2 Generación de φ -implicaciones.....	503
A.2 Función generadora de modus ponens de un φ -condicional y φ -modelos de reglas	505
A.3 Condiciones de activación de φ -condicionales	506
A.4 Agregación de φ -condicionales	509
A.5 Relajación de hipótesis en sistemas de reglas.....	512
BIBLIOGRAFIA.....	517

1. Introducción

1.1 Descripción del problema

Las importantes inversiones que hay que realizar para la construcción, puesta en marcha y operación de nuevas instalaciones industriales hacen que se aproveche cada vez con mayor intensidad las ya existentes. Este uso intensivo puede propiciar la aparición imprevista de averías o indisponibilidades.

Las averías detectadas tardíamente suelen generar unos altos costes de reparación, pérdida de producción y a su vez grandes deficiencias en la calidad del servicio que la empresa presta a sus clientes. Para evitar esta situación se han de desarrollar adecuados planes de mantenimiento que conserven en buen estado de salud los equipos para que operen fiel y eficientemente cuando les sea requerido.

Por otro lado, la necesidad de una intensiva utilización de la instalación obliga a veces a sobrecargar ciertos equipos dentro del proceso. Estas sobrecargas pueden ser el germen a su vez de averías a medio y largo plazo, agravando el problema que se pretendía resolver a corto plazo.

De toda esta problemática se desprende las grandes ventajas que puede aportar el conocimiento en todo momento del estado de salud de los equipos industriales involucrados en el proceso así como sus condiciones de operación, con vistas a detectar de forma incipiente las posibles averías que se puedan presentar, y de este modo, poder planificar de forma óptima la operación y el mantenimiento de la manera más económica posible.

Una primera aproximación para la vigilancia del estado de salud de los equipos o componentes de un proceso es la instalación de sistemas de seguimiento (*[Pau, 81]*), que constan básicamente de un conjunto de sensores destinados a medir las variables más representativas del funcionamiento del proceso, un sistema de adquisición de datos para recoger los valores de esas variables, y algún sistema de almacenamiento, típicamente un ordenador, en el que se va acumulando la información que se recoge. Estos sistemas suelen llevar a cabo la detección de las indisponibilidades mediante comparación de los valores de la variables medidas con unos umbrales máximos y mínimos. Suministran una información valiosa desde dos puntos de vista. Por un lado muestran una panorámica de lo que ocurre en cada momento en el proceso industrial. Por otro permiten obtener una perspectiva histórica de la evolución de las variables fundamentales. Son sistemas que no se anticipan a la indisponibilidad, sino que únicamente la detectan, y es la experiencia humana la que analiza la situación para encontrar las causas del problema. La realidad suele ser que al técnico que convive con estos sistemas le resulta difícil procesar toda la información

instantánea que llega, y más aún procesarla desde una perspectiva histórica con rigor y minuciosidad.

La insuficiencia de estos sistemas queda por tanto patente por dos razones principales. Por un lado, y como ya hemos señalado, estos sistemas detectan la indisponibilidad cuando ya se ha producido, siendo incapaces de anticiparse a la misma. Por otro, son incapaces de analizar los datos de que disponen no pudiendo así realizar un diagnóstico de forma automática, sino es mediante la intervención de un experto, al que se exige una gran capacidad para procesar la información en continuo y para, al mismo tiempo, conjuntarla con una perspectiva histórica global de la dinámica del proceso. Para mejorar estos inconvenientes aparecen los sistemas expertos de diagnóstico, en los que el equipo investigador al que pertenezco y yo mismo hemos trabajado durante varios años en aplicaciones de diagnóstico en continuo, de marcado carácter industrial ([Sanz et al, 93], [Sanz et al, 94], [Sanz, Villar, Muñoz, 94]).

Los sistemas expertos de diagnóstico pretenden, por un lado, detectar averías incipientes anticipándose a las indisponibilidades, y por otro, emitir el diagnóstico resultante de la investigación de las causas del problema. La arquitectura típica de un sistema experto de diagnóstico consta de tres módulos. El módulo de adquisición de datos es el encargado de recoger las señales enviadas por los sensores colocados en puntos representativos del comportamiento de los distintos equipos del proceso. El módulo de detección de anomalías es el encargado de detectar anomalías correspondientes a posibles degradaciones o averías de los equipos. Por último el módulo de diagnóstico es el encargado de analizar las anomalías detectadas para de este modo identificar sus causas y recomendar acciones a realizar, con vistas a planificar de forma óptima la operación y el mantenimiento.

Estos sistemas aunque consiguen alcanzar sus objetivos en buena medida, no siempre consiguen simular el comportamiento de un experto humano ante un problema, ya que no siempre ofrecen mecanismos de representación del conocimiento suficientemente versátiles. En efecto, obligan a introducir en estructuras lingüísticas y matemáticas, a veces no suficientemente flexibles, la forma de actuar de una persona humana, que aunque en general es incierta, se caracteriza por ser extremadamente flexible y capaz de adaptarse con rapidez a distintas situaciones.

La teoría de conjuntos borrosos (TCB) introducida por Zadeh en el año 1965 ([Zadeh, 65]) se ha ido poco a poco mostrando como una de las mejores herramientas de representación del conocimiento humano, debido a su gran riqueza semántica y flexibilidad para expresar conceptos del lenguaje hablado, de forma próxima a como lo hacen las personas. Por otro lado su soporte matemático hace que, tal y como el propio Zadeh

afirma ([Zadeh, 95], [Zadeh, 96]), la teoría de conjuntos borrosos permita calcular o computar con palabras. Es pues uno de los mejores lazos de unión entre las técnicas numéricas de representación del conocimiento y la incertidumbre, y las técnicas simbólicas de representación del conocimiento humano, lo que la hace idónea para enriquecer y mejorar las técnicas más tradicionales de desarrollo de sistemas expertos. Aunque por lógica borrosa se designa habitualmente el estudio de los conjuntos borrosos, la aritmética borrosa y el estudio de la inferencia con conjuntos borrosos, nosotros intentaremos restringir su uso a aquella parte que se dedica únicamente al estudio de la inferencia, utilizando en otros casos la denominación más general de teoría de conjuntos borrosos.

La lógica borrosa ha sido estudiada desde numerosos puntos de vista, pero sólo recientemente se ha hecho un mayor esfuerzo e hincapié desde el punto de vista de interpretación semántica ([Dubois, Prade, 93], [Dubois, Prade, Ughetto, 96], [Yager, 96]), en el que todavía creemos que existen muchas lagunas. El diseño de un sistema borroso exige al diseñador llevar a cabo múltiples elecciones para escoger unos u otros operadores, unas u otras técnicas de inferencia, etc. Estas elecciones, en general complejas debido a las diferencias a menudo sutiles entre unas y otras opciones, junto con la ausencia de criterios semánticos, ha hecho que a menudo se recurriese a criterios fundamentalmente empíricos, no siempre justificables, lo que ha complicado la interpretación de los resultados que estos sistemas producen. Esta falta de criterios es uno de los argumentos que han justificado la investigación llevada a cabo en esta tesis, y es donde fundamentalmente se centra el trabajo que en ella se expone.

Esta tesis pretende estudiar detalladamente varias de las técnicas tradicionalmente utilizadas en los sistemas basados en la teoría de conjuntos borrosos, principalmente en lo que respecta a las técnicas de inferencia, para aportar una interpretación semántica que permita orientar la tarea del diseñador de sistemas borrosos. Utilizando los criterios que vayan siendo establecidos a lo largo de la tesis, al final de la misma se propondrá una estructura de sistema experto de diagnóstico basada en el uso de distintos tipos de reglas borrosas cada una de las cuales permitirá modelar una relación semántica diferente entre la parte antecedente y la parte consecuente de la regla.

1.2 Planteamiento de la tesis

La tesis se centra en la aplicación de la teoría de conjuntos borrosos al diagnóstico de procesos industriales. El objetivo es establecer una metodología de diseño de sistemas expertos de diagnóstico basados en el uso de la teoría de conjuntos borrosos como principal herramienta de representación del conocimiento y de su incertidumbre. La metodología está orientada al diagnóstico en continuo (independientemente de la frecuencia de recogida de datos) por lo que solamente se contemplan esquemas de razonamiento hacia adelante (modus ponens generalizado), aunque la mayoría de los resultados podrían generalizarse a esquemas de razonamiento hacia atrás.

No se ha tratado de proponer una metodología general de diagnóstico, puesto que existen muchos posibles enfoques para abordar el problema, sino flexibilizar y enriquecer los sistemas expertos de diagnóstico basados en reglas (uno de los sistemas de diagnóstico más extendidos y exitosos en aplicaciones reales) mediante la incorporación de algunas de las herramientas que ofrece la teoría de conjuntos borrosos. De ahí que no se hayan abordado directamente otras estrategias de diagnóstico como por ejemplo el razonamiento basado en modelos, el basado en casos, técnicas de reconocimiento de patrones, simulación cualitativa, etc. Sin embargo y dada la generalidad de la teoría de conjuntos borrosos es posible extender algunos de los resultados obtenidos en este trabajo a estas otras estrategias, como de hecho ya se viene haciendo.

Dado que una de las principales lagunas encontradas en la teoría de conjuntos borrosos ha sido la falta de criterios claros para la elección de los distintos operadores que intervienen en un sistema borroso, gran parte de la tesis está dedicada a un análisis semántico de la representación del conocimiento basado en distribuciones de posibilidad, haciendo especial hincapié en el estudio de las distribuciones de posibilidad condicionadas, su aplicación a la inferencia, y la interpretación de sus resultados. Se ha intentado profundizar en el significado de la regla composicional de inferencia, el mecanismo más utilizado para la aplicación del modus ponens en lógica borrosa, estudiando muchas de sus posibles variantes e interpretando en la medida de lo posible su comportamiento. Se ha tratado también de llegar a resultados finales prácticos y por tanto aplicables a sistemas reales.

1.3 Organización de la exposición

El objetivo de esta tesis es profundizar en la teoría de conjuntos borrosos, y en especial en los aspectos relacionados con las distribuciones de posibilidad y la inferencia en lógica borrosa, de cara a su posterior aplicación al desarrollo de sistemas expertos para el diagnóstico de procesos industriales.

El punto de partida de la exposición (capítulo 2) comienza con una breve introducción a los sistemas de diagnóstico. A continuación se repasan la principales parcelas del conocimiento necesario para diagnosticar, el origen de los distintos tipos de incertidumbre que dicho conocimiento lleva aparejado, las estrategias de diagnóstico más comunes, y por último las principales técnicas de representación del conocimiento y de representación de la incertidumbre así como los mecanismos de inferencia a los que dan lugar.

Los capítulos 3 a 6 son una descripción de las herramientas de la teoría de conjuntos borrosos que serán empleadas en el capítulo 7 para la construcción de sistemas de diagnóstico.

El capítulo 3 es un repaso general de la teoría de conjuntos borrosos, en el que además de repasar algunas de las definiciones más básicas (conjuntos borrosos, relaciones borrosas, operaciones entre conjuntos, etc.) se introducen por primera vez dos tipos de distribuciones de posibilidad con diferente interpretación semántica, distribuciones que hemos denominado de posible y de necesaria posibilidad, y que a nuestro entender son claves para la comprensión de los desarrollos que se presentan en los capítulos 4, 5 y 6. En este capítulo también se proponen reglas de inferencia para cada uno de los dos tipos de distribuciones, y finalmente se da un breve repaso de lo que ha sido la aplicación de la lógica borrosa al desarrollo de sistemas de diagnóstico.

El capítulo 4 está dedicado por completo al estudio de la inferencia en lógica borrosa, en concreto a través de dos de los principales mecanismos de inferencia para la aplicación de la regla del modus ponens generalizado, la regla composicional de inferencia (RCI) y la modificación por similitud (MS). En este capítulo se analiza la inferencia desde un punto de vista semántico, se estudian las diferencias de comportamiento entre unos y otros operadores, permitiendo distinguir dos tipos de condicionales claramente diferenciados, los condicionales de necesidad y los de posibilidad, que permiten modelar respectivamente relaciones necesarias y relaciones posibles. Finalmente se propone una simplificación de la regla composicional de inferencia que hemos llamado la MS de dos índices, que permite garantizar la riqueza semántica de la RCI pero con las ventajas computacionales y de interpretación de la MS.

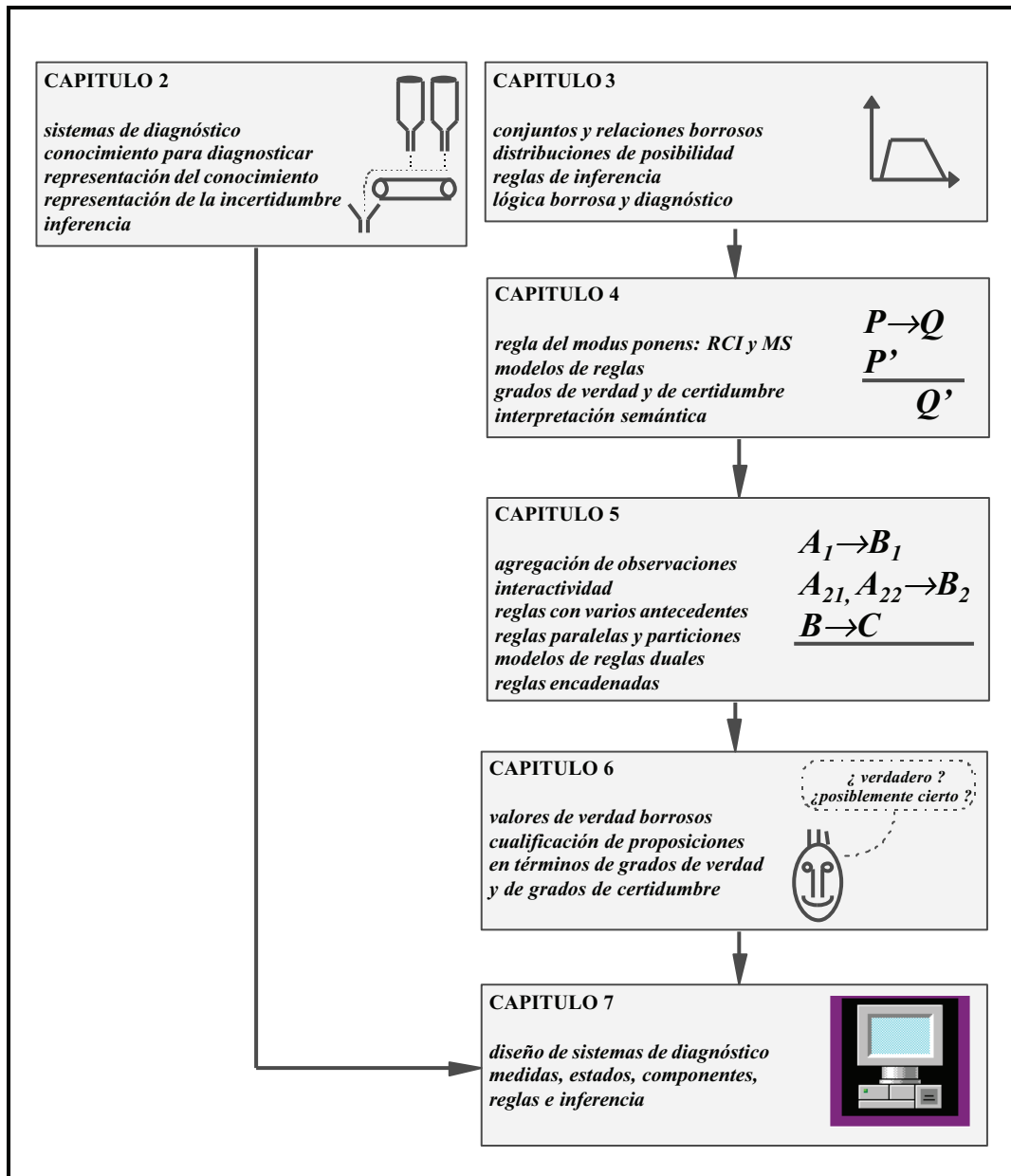
El capítulo 5 es la continuación natural del capítulo 4, y está dedicado al análisis de la agregación de observaciones e interactividad de variables, a las reglas con más de un antecedente, y en general a la problemática que se deriva del uso simultáneo de más de una regla en el mismo sistema. Se tratan aspectos importantes como los sistemas de reglas paralelas y el diseño de particiones borrosas, o los sistemas de reglas encadenadas, y la coexistencia de distribuciones de posible y necesaria posibilidad así como de reglas basadas en ambos tipos de distribuciones.

El capítulo 6 está destinado al análisis de los valores de verdad borrosos entendidos como distribuciones de posible y necesaria posibilidad, y basándose en la inferencia planteada en términos de valores de verdad borrosos se propone una familia general de los mismos, especialmente diseñada para cualificar proposiciones en términos de grados de verdad y grados de certidumbre parciales, conceptos que se derivan del análisis llevado a cabo a lo largo del capítulo 4.

Por último el capítulo 7 es una aplicación de las técnicas descritas en los capítulos anteriores, en el que se propone una metodología de diseño de sistemas de diagnóstico. Este capítulo consta básicamente de cuatro partes. En primer lugar se repasan algunos puntos de carácter general relacionados con la descomposición de los procesos en componentes, la elección de medidas representativas, y estructuras de representación del conocimiento para su codificación. En segundo lugar se propone un mecanismo de integración, dentro del sistema de diagnóstico global, para que el sistema de detección de anomalías propuesto en [Muñoz, 96] produzca hechos directamente interpretables por el sistema de identificación de anomalías. Este se describe a continuación, detallándose la estrategia de diagnóstico escogida, la estructura de las reglas de conocimiento y la forma de llevar a cabo la inferencia según las opciones de diseño que se adopten. Por último se propone un caso ejemplo que pretende ilustrar la utilización de las técnicas propuestas a lo largo de este último capítulo al diagnóstico de un transformador eléctrico de potencia.

A modo de apéndice se propone un tipo general de condicional que engloba los condicionales de necesidad y los condicionales de posibilidad estudiados en el capítulo 4, y que a nuestro entender puede ser una interesante herramienta para seguir profundizando en el estudio de las semánticas de los condicionales, dando lugar a nuevos condicionales y nuevos tipos de distribuciones de posibilidad con nuevas y diferentes interpretaciones.

La siguiente figura resume gráficamente el contenido de los capítulos principales de la tesis:



2. Sistemas de diagnóstico: conocimiento, incertidumbre e inferencia

Para construir un sistema de diagnóstico es necesario, primero, determinar con precisión el alcance del dominio de trabajo, de cara a recopilar el conocimiento que abarca el problema y determinar la estrategia a seguir, para después, a partir de la observación de unos determinados síntomas, deducir las causas que los han provocado. Desde el punto de vista de la Ingeniería del Conocimiento, los sistemas basados en conocimiento y en particular los aplicados al diagnóstico se asientan en tres grandes pilares. El primero son las técnicas de adquisición del conocimiento, el segundo las técnicas de representación del conocimiento y de su incertidumbre, y el tercero las técnicas de razonamiento que permiten utilizar de forma eficiente el conocimiento almacenado. Este capítulo es una introducción general a los sistemas de diagnóstico, a las técnicas de representación del conocimiento y de su incertidumbre, y a los esquemas de inferencia a los que dan lugar.

2.1 Organización del capítulo

Este capítulo se organiza en seis puntos principales, y es una introducción de carácter general a los sistemas de diagnóstico, donde se analiza el conocimiento a emplear, las posibles estrategias de diagnóstico, y algunas de las principales técnicas de representación del conocimiento y de la incertidumbre así como sus mecanismos de inferencia asociados.

En primer lugar se introducen los sistemas de diagnóstico en continuo de procesos industriales, se describen las principales tareas que estos sistemas deben ser capaces de realizar, y la estructura modular que típicamente adoptan para llevar a buen término su cometido.

En segundo lugar se analiza el conocimiento necesario para que un sistema de diagnóstico pueda desarrollar correctamente su tarea. Se diferencian los distintos tipos de conocimiento de que se dispone, la estructura del proceso por un lado, su funcionamiento por otro, y por último su punto de operación. La correcta identificación de estos tipos facilita la elección de las técnicas de representación del conocimiento más apropiadas. Se analizan también las principales fuentes de incertidumbre del conocimiento, lo que da lugar a su clasificación en tres grandes categorías cuyo significado debe ser tenido en cuenta para una elección apropiada del mecanismo de representación de incertidumbre.

En tercer lugar se repasan algunas de las principales estrategias de diagnóstico, esto es, aquellos modelos de razonamiento que pueden utilizarse para, a partir del conocimiento disponible, llegar a la determinación de las causas de las anomalías detectadas en el proceso.

En cuarto lugar se abordan algunas de las técnicas de representación del conocimiento más habituales en el entorno de los sistemas basados en conocimiento, desde la perspectiva de la Ingeniería del Conocimiento.

A continuación se repasan las principales técnicas de representación de la incertidumbre y sus mecanismos de inferencia asociados, señalando algunas de las principales ventajas y desventajas que unos y otros presentan.

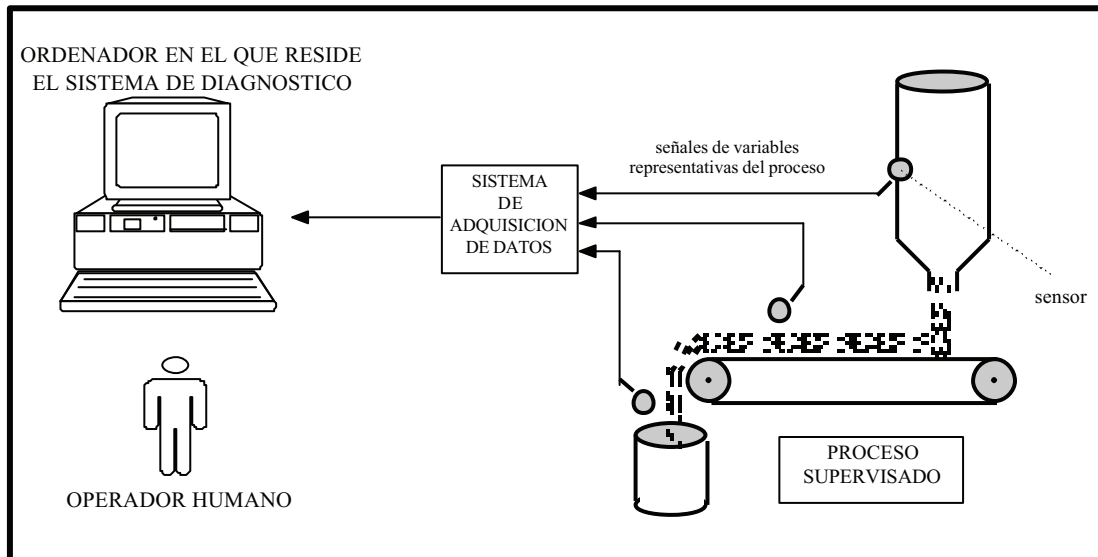
En sexto y último lugar se describen los sistemas expertos de diagnóstico basados en reglas, sobre los que se basará la metodología propuesta en esta tesis para el desarrollo de sistemas de diagnóstico.

2.2 Introducción a los sistemas de diagnóstico

Un sistema de diagnóstico es una aplicación informática que utiliza señales procedentes de variables representativas del funcionamiento de un proceso para determinar si está funcionando de forma correcta, y en caso contrario proponer explicaciones que permitan justificar las anomalías detectadas. Los sistemas más habituales en entornos industriales trabajan en continuo, lo que significa que los valores de las variables del proceso llegan de forma periódica al sistema, y a intervalos de tiempo generalmente pequeños, lo que hace casi indispensable el uso de algún procedimiento automático de recogida y presentación de datos.

Los sistemas más elementales se basan en la monitorización en continuo de las variables más representativas del proceso, cuyos valores son procesados y presentados por algún procedimiento automático (como un ordenador) al operador encargado de vigilar el proceso. El operador es el encargado de analizarlos para detectar e identificar en su caso las posibles anomalías que puedan presentarse. Estos sistemas pueden tener también mecanismos automáticos de comprobación de umbrales, de forma que cuando alguna variable supera unos umbrales preestablecidos el sistema avisa al operador de la existencia de una anomalía o alarma. Sin embargo es tarea del operador determinar la causa de la anomalía a través del análisis de las alarmas y los datos que suministra el sistema de monitorización ([Pau, 81]).

El gran volumen de información que debe ser analizado (tanto datos como alarmas simultáneas) es una de las principales razones que justifica el uso de sistemas de diagnóstico automáticos. Estos se basan en el uso de un ordenador, y son capaces de procesar de forma inteligente (es decir, como lo haría el operador humano) la información disponible, detectando e identificando las anomalías del proceso, y presentando, como resultados finales al operador, las causas que explican las anomalías detectadas, facilitando así su tarea de vigilancia y control. La siguiente figura representa esquemáticamente los principales elementos que intervienen en un sistema de diagnóstico de este tipo:



El punto de partida para desarrollar un sistema de diagnóstico en continuo para la detección e identificación de anomalías en un proceso industrial pasa por la recogida de las señales de sensores instalados en distintos puntos de la planta, que deben medir los valores de variables representativas del funcionamiento del proceso. Como ya se ha dicho estas señales suelen recogerse en continuo, a intervalos de tiempo regulares, generalmente pequeños, para ser posteriormente procesadas y analizadas por el sistema de diagnóstico.

Para la instalación de sensores es conveniente, en primer lugar, llevar a cabo una descomposición de la planta o proceso monitorizado en unidades menores, llamadas componentes, que a su vez pueden descomponerse nuevamente en otros componentes, y que generalmente se organizan en una estructura jerárquica o árbol. El grado de descomposición que se alcance depende al menos parcialmente del grado de detalle con que el que se desee llevar a cabo la tarea de diagnóstico. Identificados los componentes, para cada uno de ellos debe encontrarse un conjunto de variables suficientemente representativas de su estado, económica y físicamente medibles, y determinar la localización y tipo de los sensores que se utilizarán para su medición.

Las señales procedentes de los sensores instalados se llevan, mediante un sistema de adquisición de datos, a un ordenador en el que reside el sistema de diagnóstico, y donde los datos son recogidos, almacenados, y suministrados a aquellos módulos del sistema que los necesiten.

El sistema de diagnóstico debe ser capaz de interpretar esta información procedente de la monitorización del proceso para llevar a cabo la tarea de diagnóstico, que puede condensarse en cuatro tareas fundamentales ([Patton, Chen, 91], [Muñoz, 96]):

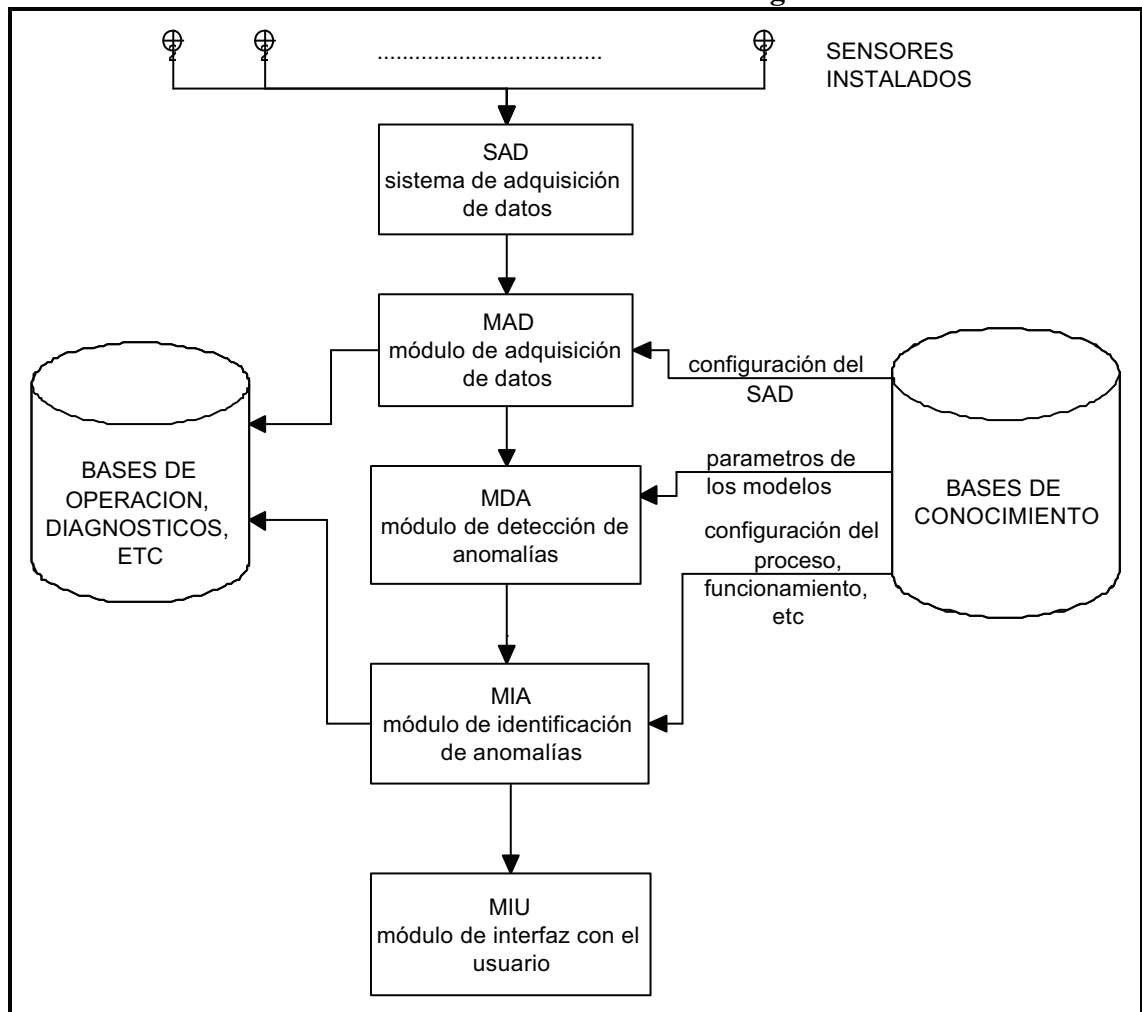
- la detección de anomalías, que consiste en determinar la presencia de faltas en los componentes que integran el proceso. Falta o anomalía se interpreta como la incapacidad parcial o total de un componente para llevar a cabo la tarea que le corresponde dentro del proceso.
- el aislamiento de anomalías, que estrechamente relacionada con el proceso de detección de las mismas, consiste en determinar qué componentes son los causantes de las anomalías o fallos detectados en el proceso o componentes del mismo.
- la identificación de anomalías, tarea que trata de asignar causas que expliquen el origen de las anomalías detectadas en el proceso.
- la corrección de anomalías, que coopera con las tareas de mantenimiento, y consiste en proponer soluciones que permitan corregir las anomalías detectadas e identificadas. En el caso de sistemas capaces de actuar sobre el proceso una posible alternativa es por ejemplo modificar su estructura o punto de operación. En otros casos las alternativas posibles para la corrección pasan desde la simple emisión de mensajes de aviso al usuario hasta la replanificación de la propia política de mantenimiento.

Según la metodología que se emplee para la construcción del sistema de diagnóstico algunas de las tareas pueden realizarse conjuntamente. Por ejemplo en los enfoques basados en ecuaciones de paridad ([Patton, Chen, 91]) los sistemas de diagnóstico agrupan la detección y aislamiento bajo una misma tarea. En otros enfoques como los basados en sistemas expertos ([Sanz et al, 93], [Sanz, Villar, Muñoz, 94]), la propia identificación de anomalías conlleva la identificación de los componentes afectados.

Además de las tareas propias del diagnóstico, suele ser conveniente llevar a cabo otras tareas auxiliares, como la recogida de las señales procedentes de los sensores, su almacenamiento para un uso posterior, la presentación en pantalla de los resultados que el sistema genera, etc.

Para llevar a cabo todas estas tareas el sistema de diagnóstico se estructura normalmente en módulos, dando lugar a una estructura modular como la que se propone por ejemplo en [Sanz et al,93] o [Sanz, Villar, Muñoz, 94], y que queda representada en figura:

Estructura modular de un sistema de diagnóstico



El sistema de adquisición de datos (SAD) es el equipo electrónico que recoge las señales procedentes de los sensores. El módulo de adquisición de datos (MAD) es el módulo encargado de pedir esas señales al sistema de adquisición de datos y ponerlas a disposición del resto de los módulos del sistema de diagnóstico. En general pueden existir otras fuentes de datos no recogidos en continuo (y no representados en la figura anterior), que el sistema debe ser capaz de utilizar, y que pueden provenir por ejemplo de observaciones directas hechas por los operadores, de análisis realizados en laboratorios, etc. Estos datos suelen introducirse en el sistema a través del módulo de interfaz con el usuario, y pasan a disposición del resto de los módulos del sistema al igual que los datos recogidos en continuo, para poder ser utilizados tanto desde el módulo de detección como desde el módulo de identificación de anomalías.

La primera tarea a realizar, llevada a cabo por el módulo de detección de anomalías (MDA) consiste precisamente en detectar la presencia de anomalías del proceso. Si la detección es positiva el módulo de identificación de anomalías (MIA) procede a su análisis, utilizando la información proveniente tanto de los sensores como del módulo de detección, para tratar de identificar las causas origen de las anomalías detectadas. La información sobre el proceso de diagnóstico, el estado de la planta, las anomalías detectadas e identificadas, etc., es suministrada al usuario por el módulo de interfaz con el usuario (MIU).

Además existen otros mecanismos de interacción entre varios de los módulos del sistema y las bases de datos de entrada al mismo, en las que reside por ejemplo la configuración del sistema de adquisición de datos, la configuración del proceso (componentes y medidas) y el conocimiento disponible para el diagnóstico. Del mismo modo los módulos del sistema se comunican con las bases de datos de salida en las que almacenan los datos relevantes del proceso de diagnóstico, como los diagnósticos emitidos, los datos de operación del proceso, etc. Separar el conocimiento sobre la configuración del proceso y el conocimiento para diagnóstico, en ficheros externos al sistema de diagnóstico, facilita por un lado la permanente modificación por los propios usuarios del conocimiento de que dispone el sistema (en base a su propia experiencia), y permite por otro transportarlo de forma flexible para que el sistema sea fácilmente adaptable a otros procesos similares (adaptando adecuadamente el conocimiento almacenado en los ficheros). Dado que las variables que se miden varían de una planta a otra, también suele ser necesario utilizar distintos sistemas de adquisición de datos lo que obliga en muchos casos a modificar el módulo de adquisición de datos, limitando parcialmente su transportabilidad.

El sistema de detección de anomalías se basa normalmente en la utilización de modelos de funcionamiento normal que estiman los valores de las variables representativas del estado de los componentes, y que son comparados con los valores reales medidos a través de los sensores instalados. Cuando las diferencias entre ambos son suficientemente significativas el sistema de detección considera que hay una anomalía.

El sistema de identificación de anomalías utiliza normalmente conocimiento previamente almacenado sobre el funcionamiento del proceso, que consiste básicamente en los fallos o anomalías y su relación con los síntomas detectables (aunque depende del tipo de estrategia de diagnóstico escogida). Este conocimiento se utiliza comparando los síntomas observados en el proceso real con los síntomas de los distintos fallos que puede presentar el proceso para así identificar las causas que mejor expliquen las anomalías detectadas. Algunos sistemas son capaces de deducir en continuo nuevos diagnósticos y sus síntomas asociados mediante simulación del funcionamiento del proceso, utilizando modelos previamente almacenados en el sistema.

En general todo esto hace que para construir un sistema de diagnóstico, y más concretamente los módulos de detección e identificación, sea necesario manejar un gran conjunto de técnicas entre las que podemos destacar por ejemplo técnicas de aproximación funcional y optimización para el módulo de detección de anomalías ([Muñoz, 96]) y técnicas de Inteligencia Artificial (o Ingeniería del Conocimiento), especialmente en lo que concierne a sistemas basados en conocimiento, para el módulo de identificación de anomalías, en las que principalmente se centrará esta tesis.

2.3 Conocimiento e incertidumbre en un sistema de diagnóstico

El conocimiento se puede definir como una colección de hechos y relaciones entre hechos, que adecuadamente interpretados y utilizados, permiten deducir nuevos hechos y/o llevar a cabo acciones inteligentes ([Popovic, Bhatkar, 94]). Por hecho entendemos aquellos fragmentos de conocimiento más elementales que permiten definir algo sobre los objetos que constituyen el mundo en estudio o mundo real, es decir aquél cuyo conocimiento debe ser representado.

Para llevar a cabo la tarea de diagnóstico un experto dispone por lo general de varios tipos de conocimiento que maneja de forma conjunta y eficaz, para a partir de las señales o alarmas que hayan provocado su atención, llegar a concluir cuales son las causas que explican de la forma más satisfactoria posible el o los problemas detectados en el funcionamiento del proceso industrial vigilado. Básicamente el conocimiento que se requiere para esta tarea es el conjunto de fallos que el proceso puede presentar, y una asociación entre los distintos fallos y aquellos síntomas u observaciones que permiten identificar cada uno de los citados fallos ([Milne,87]). Para que este conocimiento pueda ser utilizado y gestionado por un sistema de diagnóstico deben encontrarse estructuras para su representación y almacenamiento, de forma que algún mecanismo de razonamiento automático pueda posteriormente utilizarlo para llevar a cabo la tarea de diagnóstico. Esto hace que los sistemas de diagnóstico se engloben dentro de una categoría más general de sistemas informáticos, los sistemas basados en conocimiento, que constituyen una importante parcela dentro de las técnicas de la Ingeniería del Conocimiento o Inteligencia Artificial. Sobre este tipo de técnicas y los sistemas basados en el conocimiento puede consultarse por ejemplo [Popovic, Bhatkar, 94] y [Hopgood, 93]. Otra referencia interesante es [Dillon, Lee, 93] donde se dedica un capítulo a la descripción, modelado y diseño de este tipo de sistemas desde la perspectiva de la orientación a objetos.

Al igual que otros muchos sistemas basados en conocimiento, el éxito de los sistemas de diagnóstico pasa por la correcta adquisición del conocimiento disponible sobre un proceso, por su conveniente estructuración y almacenamiento, y por la selección de mecanismos de razonamiento automático similares a los llevados a cabo por los expertos, adaptados a la utilización eficiente del conocimiento almacenado. Dado que una de las características del conocimiento humano es su capacidad de razonar en entornos inciertos (en su sentido más amplio), será necesario optar por algún esquema de representación de incertidumbre que permita modelar la existencia de verdades parciales, de explicaciones más o menos ciertas/posibles/plausibles/ probables/ etc. para unos determinados síntomas, de un

conocimiento del proceso generalmente incompleto, vago, impreciso e incierto. Estos términos serán aclarados en el apartado 2.3.2 de éste mismo capítulo.

2.3.1 El conocimiento requerido para diagnosticar

Para clarificar de qué conocimiento debe disponerse para el diagnóstico y analizar las diferentes estrategias disponibles conviene clasificar el conocimiento en distintos tipos, atendiendo por ejemplo a lo que describe, a la profundidad con que lo describe, o a su procedencia, lo que generalmente condiciona el tipo de estrategia de diagnóstico a emplear.

Algunos autores dividen el conocimiento en niveles, cada uno de los cuales se corresponde con un determinado nivel de profundidad en la descripción de la planta o proceso industrial bajo estudio. Así por ejemplo [Milne,87] divide el conocimiento en cuatro niveles. El nivel más bajo es el nivel estructural que refleja los distintos componentes del sistema y sus interconexiones, es decir la estructura física del proceso. La información estructural es especialmente útil en la tarea de aislamiento de aquellos componentes que están fallando, atendiendo principalmente a las conexiones entre los mismos. El siguiente nivel es el nivel de comportamiento, que pretende recoger el conocimiento disponible sobre el funcionamiento interno de los componentes. El tercer nivel es el nivel funcional, relacionado con los objetivos que el componente debe cumplir dentro del proceso industrial, y que se explican en base al conocimiento del nivel anterior, es decir su comportamiento interno. El cuarto y último nivel está basado en patrones de comportamiento. En este nivel está todo aquel conocimiento que se basa en asociaciones concretas de síntomas-fallos. Esto puede hacerse mediante bibliotecas de fallos, como en los sistemas que razonan basándose en casos reales previamente recopilados y almacenados en el sistema ("case-based-reasoning"), o casos simulados previamente mediante modelos de comportamiento. Otra posible alternativa es la utilización de reglas de producción (que se describen en el apartado 2.5.4), como se hace en la mayoría de sistemas de diagnóstico. Los dos primeros niveles formarían parte del llamado conocimiento profundo o "deep knowledge" mientras que los dos últimos lo harían del llamado conocimiento superficial o "shallow knowledge", o más comúnmente conocimiento heurístico.

Un enfoque similar puede encontrarse en [Fink, Lusth, 87] donde se divide el conocimiento en el conocimiento de sentido común, el conocimiento profundo y el conocimiento superficial. El conocimiento profundo se define como aquél que el experto comienza a adquirir durante sus primeros años de formación, basado en la física de los componentes y descripciones matemáticas precisas de los mismos, mientras que el conocimiento superficial se define como aquél que se adquiere con la experiencia,

conocimiento de carácter fundamentalmente empírico y por tanto muy práctico, lo que le hace más eficaz para diagnosticar con rapidez y precisión.

El conocimiento puede ser clasificado también atendiendo a su procedencia. Así por ejemplo en [Ulieru, Isermann, 93] se divide el conocimiento en tres categorías contrapuestas. Por un lado el conocimiento heurístico basado en la experiencia de los expertos y que se expresa en forma de relaciones síntomas/fallos. Por otro, un conocimiento de naturaleza estadística, basado en la historia de operación del proceso, formado fundamentalmente por el conjunto de las probabilidades de unos y otros fallos, pero en general difícil de obtener y susceptible de no reflejar correctamente las relaciones causa efecto entre fallos y síntomas. Por último, un tercer tipo obtenido a partir de simulaciones realizadas utilizando modelos analíticos de los componentes del proceso. En base a la clasificación propuesta por [Milne,87], los dos primeros tipos estaría en el cuarto nivel, mientras que el último estaría en el segundo o tercero, estando los tres destinados a describir el comportamiento del proceso y componentes, y no su estructura. Nótese que la procedencia del conocimiento determina en gran medida la técnica de representación del conocimiento a emplear (por ejemplo el conocimiento estadístico sugiere claramente el uso de probabilidades).

El conocimiento necesario para las tareas de diagnóstico suele estructurarse entorno a la figura del componente, entendido desde un punto de vista abstracto (pudiendo o no corresponder con componentes físicos). Para cada componente (o para aquellos componentes más relevantes en el proceso) hay que caracterizar su comportamiento normal para poder detectar posibles anomalías, así como las principales causas de fallo que pueden dar lugar a las anomalías detectadas. El comportamiento normal suele ser más fácilmente caracterizable debido en general a la existencia de datos en las bases de datos de operación del proceso, o a la facilidad de recogerlos, ya que los componentes están, la mayor parte del tiempo, funcionando correctamente. La caracterización de los fallos suele ser mucho más compleja debido precisamente a la escasez de datos, lo que en general requiere la intervención de expertos, de operadores con experiencia, de simulaciones del funcionamiento del proceso, etc.

En general puede decirse que el conocimiento se divide en dos grandes bloques, por un lado el conocimiento relativo a la estructura del proceso, sus componentes interconexiones y relaciones, y por otro el conocimiento relativo al funcionamiento del proceso, donde de algún modo se describe el comportamiento de los componentes que lo integran. En los sistemas expertos de diagnóstico el primero de ellos suele denominarse base de hechos, mientras que el segundo se denomina base de conocimiento. Sin embargo a estos dos tipos hemos añadido otro bloque de conocimiento que aunque implícitamente presente en cualquier sistema de diagnóstico, no se le suele atribuir categoría propia. Este tercer tipo sería el conocimiento del punto de operación del proceso en cada instante, que no

necesariamente tiene que ser expresado como la enumeración exhaustiva de los valores de sus medidas, ya que es susceptible de ser representado mediante técnicas más compactas, que aunque menos precisas, sean suficientemente informativas. Siguiendo estas divisiones definimos los siguientes conceptos:

- **conocimiento estructural:** es el conocimiento sobre la estructura física del proceso. En todo sistema de diagnóstico uno de los primeros pasos suele consistir en la representación del proceso como un conjunto jerarquizado de unidades funcionales o componentes. En general cada componente estará constituido a su vez por un conjunto de componentes formando así una estructura jerárquica o árbol. Los componentes de las hojas del árbol corresponderán con equipos físicos más o menos complejos, según el nivel de detalle con el que se quiera o se necesite supervisar y diagnosticar la instalación. La estructura física del proceso puede enunciarse como el conjunto de sus componentes, las interrelaciones existentes entre los mismos (que pueden ser de pertenencia, de localización en el proceso, de conexión, etc.), y el conjunto de medidas asociadas a cada uno de los componentes del proceso. Para describir la estructura física del proceso definimos los siguientes objetos:
 - **componentes:** son unidades funcionales del proceso que pueden contener a su vez a otros componentes y que en general llevan asociados una serie de atributos que los describen estática y dinámicamente, y que denominaremos de forma genérica **medidas**. Los componentes pueden corresponderse con componentes físicos del proceso o ser componentes abstractos compuestos por otros componentes físicos o abstractos. Para poder expresar de forma suficientemente flexible la estructura del proceso (tal y como lo haría un experto humano) es necesario poder definir **relaciones** entre los distintos objetos que lo configuran (como las relaciones jerárquicas “pertenece_a”, o relaciones de conexión “conectado_con”, etc.). Las relaciones del tipo “es_un” pueden ser de especial interés para modelar lo que se conoce como herencia, que permite que un componente, por ser de un determinado tipo, herede por defecto todas las propiedades y características del tipo al que pertenece.
 - **medidas:** son atributos asociados a componentes que describen las características estáticas o dinámicas de los mismos. A cada medida se le asocia un conjunto de atributos que la describe. El principal atributo de una medida es lógicamente su valor, aunque también resultan de especial interés sus estados (que definimos en el siguiente punto), así como su tipo de sensor, su procedimiento de medida, sus unidades, etc. Suele también ser interesante disponer de algún mecanismo que permita agrupar medidas bajo un nombre común, para poder referirse a ellas de forma compacta y no mediante su enumeración exhaustiva. Una forma de agruparlas es a través de sus atributos,

como por ejemplo el tipo de sensor, y otra disponer de un atributo especial, que denominaremos **propiedades**, cuyos valores pueden utilizarse para agrupar las medidas bajo valores de propiedades comunes. Según su procedencia pueden distinguirse dos tipos de medidas:

- **medidas estáticas**: son atributos asociados a componentes que describen las características estáticas de los mismos, que se fijan durante el diseño e instalación de la planta industrial, y que por tanto no cambian con el punto de operación del proceso (por ejemplo la potencia nominal de un motor, el volumen de una cuba, etc.).
- **medidas dinámicas**: son atributos asociados a componentes que describen las características dinámicas de los mismos, y que por tanto cambian con el punto de operación del proceso (por ejemplo la potencia consumida por un motor, o el nivel de líquido en una cuba, etc.). Como ya ha sido señalado, después de la descomposición del proceso en componentes debe escogerse para cada uno de ellos un grupo de medidas dinámicas que caractericen su comportamiento. Dentro de las medidas dinámicas podemos distinguir básicamente dos tipos según su procedencia, que se corresponden con dos modos de monitorización distintos (monitorización en continuo y monitorización discreta). Las medidas dinámicas automáticas, o **medidas automáticas**, proceden de los sensores instalados en la planta y son recogidos por uno o varios sistemas de adquisición de datos, conectados a su vez al sistema de diagnóstico. Estas medidas se toman en continuo y a intervalos de tiempo regulares, aunque no necesariamente con la misma frecuencia de muestreo para todas. Por otro lado están las medidas dinámicas manuales o **medidas manuales**, suministradas por algún tipo de procedimiento generalmente aperiódico, y que proceden en general de análisis realizados en algún laboratorio, inspecciones visuales de características difícilmente medibles, etc. Los sistemas de diagnóstico en continuo basan su funcionamiento principalmente en la información procedente de las medidas automáticas, aunque deben ser también capaces de utilizar la información que suministran las medidas manuales.
- **estados**: uno de los atributos más importantes de una medida (además de su valor) es lo que hemos denominado estado, que permite describir cualitativa o cuantitativamente los posibles valores que una variable estática o dinámica puede tomar. Un estado no es el valor de una medida (aunque podría serlo) sino una categoría en la que el valor de la medida puede ser clasificado. Así por

ejemplo valores cualitativos típicos en el lenguaje hablado son “alto”, “medio” y “bajo”, aplicables a muchos tipos de variables, mientras que valores de la medida podrían ser “12.5”, “entorno_a_10”, etc. Aunque los estados de las medidas pueden incluirse en la descripción estructural del proceso, están íntimamente relacionadas con la descripción funcional y de operación del mismo (que se definen a continuación).

- **conocimiento funcional:** es el conocimiento sobre el funcionamiento del proceso. La estructura física por sí sola es insuficiente para diagnosticar posibles anomalías o fallos en el funcionamiento del proceso. Por ejemplo durante la fase de detección de anomalías se utilizan típicamente relaciones entre medidas que deben ajustarse a ciertos patrones predeterminados cuando el sistema funciona de forma correcta. Por otro lado durante la fase de diagnóstico es necesario en general comprobar una serie de relaciones adicionales para poder identificar las causas de las anomalías detectadas. Estas relaciones pueden ser de carácter numérico (como habitualmente ocurre en la fase de detección de anomalías), en cuyo caso se expresan generalmente mediante el uso de modelos matemáticos, o de carácter cualitativo (en cuyo caso pueden expresarse mediante relaciones entre los estados de las medidas). La tarea de diagnóstico requiere típicamente el conocimiento relativo al funcionamiento normal y funcionamiento anómalo de un proceso. Como ya se ha visto, a menudo se distingue, dentro del conocimiento funcional, entre conocimiento profundo y conocimiento superficial. Además cada tipo de conocimiento suele estar asociado con una técnica de representación distinta. Así por ejemplo el conocimiento superficial suele representarse principalmente con reglas de producción o ejemplos (también llamados casos) organizados en bibliotecas, mientras que el conocimiento profundo suele asociarse principalmente con modelos matemáticos (esta asociación puede llegar a ser arbitraria, como se comenta en [Hopgood, 93] ya que por ejemplo se puede representar conocimiento profundo mediante reglas de producción). Otra clasificación de especial interés práctico en la realización de sistemas de diagnóstico consiste en diferenciar entre el conocimiento sobre el funcionamiento normal y el funcionamiento anómalo del proceso:

- conocimiento sobre el funcionamiento normal del proceso: son las relaciones que cumplen las medidas del proceso cuando éste está funcionando correctamente, y que cuando dejan de cumplirse permiten apuntar a la existencia de algún tipo de fallo o malfunción en algún componente del proceso.
- conocimiento sobre el funcionamiento anómalo del proceso: este conocimiento se expresa mediante relaciones entre estados de medidas y diagnósticos, que permiten confirmar la existencia de anomalías en el funcionamiento del proceso, y catalogarlas dentro de un conjunto de causas o diagnósticos. El conocimiento del funcionamiento anómalo suele ser más difícil de enunciar ya que pueden

existir múltiples modos de fallo diferentes, requiere un buen conocimiento del proceso, y en general no suele disponerse de suficientes datos correspondientes a las situaciones de fallo que pueden darse. Sin embargo el conocimiento del funcionamiento normal puede obtenerse a partir de los datos de operación del proceso, que suelen proceder de situaciones de funcionamiento normal, lo que facilita su adquisición.

- **conocimiento de operación:** es el conocimiento sobre el estado actual del proceso, dependiente de su punto de operación. El conocimiento de operación del proceso está estrechamente relacionado con el conocimiento estructural ya que se expresa asignando estados a las medidas dinámicas de cada uno de los componentes, a partir de sus valores medidos. En un sistema que trabaje en continuo, este conocimiento debe ser actualizado constantemente, en base a la información suministrada por el sistema de adquisición de datos en el caso de las medidas automáticas, y en base a la información adicional que puede proporcionar el usuario en el caso de las medidas manuales (en [Sanz et al, 93] y [Sanz et al, 94] se describen dos sistemas que combinan el uso de medidas manuales y automáticas). El estado actual del proceso queda determinado tanto por los valores de las medidas como por el estado en que están, así como posibles funciones aplicadas a los valores de las medidas como son tendencias, correlaciones, etc., que puedan caracterizar mejor su dinámica. Según cómo esté expresado o codificado el conocimiento funcional, este hará uso de los valores de las medidas (como es el caso de los modelos matemáticos del funcionamiento de un proceso) o de los estados de las mismas (como es el caso por ejemplo de las reglas de conocimiento expresadas más en términos cualitativos que numéricos). Este conocimiento, aunque presente en cualquier sistema de diagnóstico, no siempre se trata explícitamente. Sin embargo esto puede ser importante a la hora de almacenar las distintas situaciones de fallo que el sistema de diagnóstico pueda ir detectando de cara a crear una base de diagnósticos para su uso posterior (por ejemplo para el razonamiento basado en casos, o para el refinamiento del conocimiento inicial del sistema de diagnóstico).

El conocimiento estructural, debido a su carácter estructurado, se presta especialmente al uso de técnicas de representación estructuradas, como pueden ser los objetos o los frames, que se describen en el apartado 2.5.2. En este sentido puede destacarse por ejemplo [Bécher, 94], donde se aborda el problema de la construcción sistemática de sistemas de diagnósticos desde una perspectiva modular orientada a objetos. En este trabajo se propone la descomposición de la planta en un conjunto de componentes abstractos, caracterizados por sus variables de estado, entradas y salidas, formando una estructura jerárquica, y utilizando el concepto de relación para expresar relaciones adicionales entre componentes, como por ejemplo “comunica_con”, “físicamente_incluido”, etc. Asociado a cada componente se definen elementos de

conocimiento adicionales necesarios para el diagnóstico. Entre ellos cabe destacar por ejemplo una completa descripción que comprende un modelado numérico o cualitativo de su comportamiento que puede hacerse mediante modelos matemáticos, reglas de producción, umbrales, correlaciones entre variables, retardos entre dinámicas de variables, etc. También se les asocia un conjunto de leyes, cada una de las cuales describe un tipo de anomalía (que se expresan en forma de restricciones), y un contexto en el que se describen aquellas variables externas que puedan tener una influencia significativa en su comportamiento (y pueden ser causa de comportamientos excepcionales).

Por otro lado el conocimiento funcional que describe el funcionamiento normal de los componentes suele representarse mediante modelos matemáticos ajustados previamente con datos reales de operación (véase por ejemplo [*Sanz, Villar, Muñoz, 94*], [*Muñoz, 96*]), mientras que el conocimiento sobre el funcionamiento anómalo del proceso suele expresarse mediante algún tipo de asociación síntomas-causas que puede hacerse como ya se ha comentado mediante reglas de producción, casos (es decir situaciones reales o simuladas, previamente almacenadas), etc.

La representación del conocimiento de operación, que incluye tanto los valores de las medidas (numéricos o cualitativos) como su dinámica, no es una parcela muy explorada en estos sistemas. La alternativa más inmediata consiste en utilizar directamente los valores numéricos de las medidas del sistema, de forma que no se produzca pérdida de información, pero esto puede conducir a bases de datos demasiado voluminosas. Otra posible alternativa consiste, por un lado en utilizar los estados de las medidas (que suelen ser sólo unos pocos) en lugar de sus valores, con lo que se produce una cierta compresión aunque también una pérdida de información, que sólo en el caso en que los estados hayan sido bien escogidos, podrá considerarse despreciable.

2.3.2 La incertidumbre del conocimiento

2.3.2.1 Fuentes de incertidumbre

A la hora de conjuntar la información de se que dispone, un experto lleva a cabo complejos procesos de razonamiento que le permiten llegar a conclusiones más o menos seguras, destinadas a encontrar las causas de las anomalías detectadas. A lo largo de este proceso pueden aparecer varias fuentes de incertidumbre (entendida en su sentido más amplio y que iremos aclarando en este apartado) que pueden localizarse básicamente en los siguientes puntos (véase [*Hopgood, 93*]). Por un lado la información disponible puede no ser del todo fiable o cierta, por otro el conocimiento puede ser incompleto e incierto debido a un desconocimiento parcial del proceso o a una falta de experiencia del que lo

enuncia. Por último el conocimiento puede ser vago e impreciso por estar a menudo formulado en base a conceptos vagos que son los que habitualmente se utilizan en el lenguaje hablado, pero que resultan en la mayoría de los casos suficientes para ser correctamente entendidos y utilizados en el razonamiento humano, y que permiten abstraerse de detalles irrelevantes que no harían sino complicarlo innecesariamente.

En el ámbito del diagnóstico la incertidumbre de la información suele estar motivada porque el conocimiento de operación del proceso es por alguna razón incierto o incompleto. Incompleto es efectivamente en la mayoría de los casos puesto que a menudo resulta cara, compleja o incluso imposible la instrumentación necesaria para obtener todas las medidas de todas las variables que podrían resultar de interés de cara a conocer con precisión el punto de trabajo de un componente o su estado de salud. La incertidumbre por el contrario puede deberse a procedimientos de medida poco fiables, a medidas realizadas por los operadores que pueden involucrar fuertes apreciaciones subjetivas, y a que en general es imposible afirmar con total precisión cuál es el valor de una determinada medida.

Por otro lado, la incertidumbre en el conocimiento funcional del proceso se debe habitualmente al desconocimiento parcial que siempre se tiene de los procesos reales, que tienden a apartarse por varios factores de su descripción teórica idealizada (que, indispensable durante las fases de diseño, puede conducir en algunos casos a conclusiones engañosas cuando se utilizan para describir el proceso durante su funcionamiento real). Uno de estos factores es a menudo la falta de datos de diseño o especificaciones exhaustivas de los componentes que configuran el proceso, y que imposibilitan el uso de los modelos físicos idealizados utilizados durante la fase de diseño. Esto hace que la experiencia subjetiva de un operador, o la experiencia objetiva extraída de unos datos numéricos del proceso, tenga una fuerte componente de incertidumbre que debe ser tenida en cuenta por el sistema de diagnóstico. Nótese que además en algunos casos la inexperiencia del operador o del experto puede introducir nuevas fuentes de incertidumbre.

Cuando el conocimiento no se basa en precisos modelos matemáticos o en el almacenamiento de datos numéricos recogidos durante la operación del proceso, se suele expresar en términos principalmente cualitativos que establecen relaciones entre los valores también de tipo cualitativo entre las medidas involucradas. En este caso el conocimiento es vago debido a la existencia de fronteras poco definidas entre unos conceptos y otros (por ejemplo entre temperatura alta y media), pero sin embargo son esa vaguedad e imprecisión del lenguaje las que permiten una abstracción de los detalles irrelevantes para concentrarse en lo que realmente es importante para la resolución de problemas, facilitando y simplificando el proceso completo de razonamiento.

Otra fuente de incertidumbre característica de los sistemas de diagnóstico viene motivada por el hecho de que en general un mismo síntoma puede estar provocado por varias causas diferentes. Dado que no suele ser viable medir todas las variables necesarias para identificar con precisión la o las anomalías detectadas, la técnica de representación del conocimiento que se utilice para representar la asociación de síntomas con fallos debe ser capaz de admitir grados de incertidumbre. La incertidumbre será tanto mayor cuanto mayor sea el número de diagnósticos capaces de explicar un mismo síntoma.

Parece crucial en este punto clarificar con precisión los distintos tipos de incertidumbre señalados, para poder referirnos de forma unívoca a todos ellos en capítulos posteriores. Para ello definimos a continuación lo que para nosotros van a significar los términos de verdad, imprecisión, vaguedad e incertidumbre (en este caso en su sentido más restringido). Nótese que en muchas ocasiones utilizaremos el término incertidumbre (en su sentido amplio) para englobarlos a todos ellos. Estos conceptos se definirán en el capítulo 3 dedicado a la teoría de conjuntos borrosos, utilizando como herramienta básica las distribuciones de posibilidad.

2.3.2.2 Tipos de incertidumbre

La incertidumbre en su sentido más amplio engloba distintos tipos de incertidumbre, de los que podemos distinguir principalmente tres: la existencia de **verdades parciales** que da lugar a la vaguedad en los conceptos, la **incertidumbre** entendida como el desconocimiento parcial de los grados de verdad, y la **imprecisión** de la información.

La verdad se entiende generalmente como la conformidad, compatibilidad o parecido entre una afirmación y lo que efectivamente se conoce de la realidad, que en nuestro caso viene descrito por un conjunto de medidas y relaciones que describen el conocimiento sobre el proceso industrial bajo estudio. La existencia de parecidos parciales genera la necesidad de manejar grados de verdad parciales, que se expresan normalmente con un cifra en el intervalo $[0,1]$. Nótese que tanto el conocimiento disponible como la afirmación que se contrasta pueden ser precisos, imprecisos, vagos y/o inciertos.

Los grados de verdad parciales dan origen a la vaguedad de las propiedades o conceptos. La vaguedad de una propiedad se debe a la posibilidad de que dicha propiedad pueda cumplirse de forma gradual, y está presente en la mayoría de los conceptos o propiedades de uso común en el lenguaje hablado, tanto en la vida cotidiana como en ambientes más técnicos o científicos (considerados más precisos). Por ejemplo en las afirmaciones "*ha llovido mucho*" o "*la presión es alta*", los conceptos *mucha lluvia*, *presión alta*, etc., tienen (en cualquier contexto) un amplio contenido en vaguedad. En el caso de la lluvia resulta imposible determinar con precisión cuando se considera que ha llovido mucho y en

general existirá una zona de transición entre los conceptos *poca lluvia* y *mucha lluvia*, que además dependerá de la zona geográfica considerada. Del mismo modo no parece posible afirmar con precisión cuales son los límites que permiten considerar a un valor de presión alto y a otro no. Utilizar un límite rígido no permite describir con suficiente fidelidad la mayoría de estos conceptos, ya que en la vida real las transiciones no suelen ser bruscas sino graduales. Solo algunas propiedades como las relaciones matemáticas o las de carácter discreto admiten fronteras bruscas, al no existir grados de cumplimientos parciales (por ejemplo *mayor_que_10*, *estar_casado*, etc.).

La incertidumbre (o certidumbre) de una afirmación permite modelar el desconocimiento que sobre su grado de verdad se tiene. Independientemente del grado de verdad de una afirmación (que puede como hemos señalado ser parcial), su incertidumbre (respectivamente certidumbre) será tanto menor (respectivamente mayor) cuanto mejor se conozca su grado de verdad y viceversa. Así cuando se conoce con total exactitud podemos decir que se tiene una total certidumbre, o que la certidumbre (también denominada certeza y necesidad) es 1, mientras que un total desconocimiento lleva a una total incertidumbre o certidumbre 0. Existen varias formas más o menos habituales de expresar la incertidumbre en el lenguaje hablado. Una muy común es hablar de probabilidades (que en general suelen ser subjetivas) como por ejemplo cuando se afirma "*es probable que mañana llueva*". Otras formas alternativas pueden apreciarse en las afirmaciones "*es posible que mañana llueva*", y "*es cierto que mañana lloverá*", y en todas ellas hay diferencias semánticas notables que se caracterizan por distintas formas de entender la incertidumbre. Mientras la primera refleja un creencia basada en que cuando el cielo está obscuro es frecuente que llueva, la segunda puede reflejar falta de información (no se conoce el aspecto del cielo) que impide afirmar o contradecir la predicción de lluvia (o la inexistencia de impedimento conocido a que llueva), y la última refleja un total convencimiento o seguridad en lo que se afirma. Lo posible se asocia habitualmente con lo factible, lo viable o practicable, lo probable con lo frecuente, y lo necesario con lo fiable, lo seguro o lo conocido.

A menudo se confunde grado de verdad con probabilidad, así como certidumbre con probabilidad, pero son conceptos distintos que conviene diferenciar. En primer lugar no debe confundirse probabilidad con grado parcial de verdad. Un ejemplo típico es el de la botella que contiene un cierto líquido. No es lo mismo afirmar "*el líquido es venenoso en grado 0.3*" que "*el líquido es venenoso con probabilidad 0.3*". En el primero de los casos se tiene la certidumbre de que la ingestión del líquido producirá ciertas molestias no excesivamente graves al que lo ingiera, ya que el líquido es ligeramente venenoso (siendo el valor 0.3 interpretable como su grado de peligrosidad). En el segundo, aquel que ingiera el líquido o bien sufre graves trastornos, o bien no le pasa nada, aunque es más probable que no le pase nada. En este caso el líquido es o bien venenoso o bien inocuo. Además una vez

ingerido es posible eliminar la incertidumbre inicial, de forma que la probabilidad pasa a ser 1 si el individuo sufre graves trastornos, o 0 en caso contrario.

Por otro lado si la certidumbre (o necesidad) de la afirmación “*el líquido es venenoso*” es 1 entonces estamos seguros de que cualquiera que lo ingiera sufrirá graves trastornos, y en este caso la probabilidad de que sea venenoso ha de ser también 1. Si por el contrario no tenemos ninguna información sobre la procedencia de la botella, la certidumbre de “*el líquido es venenoso*” es 0, aunque se puede afirmar que es posible que el líquido sea venenoso, por lo que su posibilidad valdría 1, y ninguna persona sensata osaría beberlo. En este caso la escasez de información no permite afirmar nada sobre la probabilidad de que el líquido sea venenoso y resulta imposible especificar su valor. Si por último un especialista analiza el líquido y concluye que no es venenoso, entonces tanto la certidumbre (o necesidad) como la posibilidad de “*el líquido es venenoso*” son ambas 0, y también lo es su probabilidad. Este ejemplo evidencia las diferencias entre verdad y certidumbre, y la existencia de distintas formas alternativas de representar la incertidumbre.

Hay que señalar que certidumbre (se represente o no mediante probabilidades) y vaguedad no son conceptos contradictorios sino complementarios, que se dan simultáneamente en el lenguaje hablado por lo que es en muchos casos necesario representarlos conjuntamente, lo que significa que un concepto vago puede también ser parcialmente incierto.

La imprecisión por último es una característica de la información que hace que ésta no sea suficientemente específica. Por ejemplo afirmar “*la presión atmosférica está en [0 atm, 10 atm]*” es una afirmación con total certidumbre, pero tan imprecisa que carece de información. Una afirmación más precisa es por ejemplo “*la presión está en [0.9 atm, 1.1 atm]*”. En este caso la información, aunque sigue siendo imprecisa lo es menos, y tiene mayor contenido informativo. Nótese que una afirmación imprecisa puede además tener cierta incertidumbre, y de hecho cuanto más imprecisa es, menor es su incertidumbre. La imprecisión en el valor de una medida se modela generalmente mediante un sobredimensionamiento del conjunto en el que se dice que toma su valor. Cuanto mayor es este conjunto mayor es la imprecisión de la información pero también por ello mayor es su certidumbre.

Señalar por último que algunos autores (véase por ejemplo [Prade, 85], [Buisson, Farreny, Prade, 85]) agrupan los conceptos de verdad, imprecisión y vaguedad, íntimamente relacionados, bajo el término único imprecisión, para diferenciarlos de la incertidumbre.

2.4 Estrategias de diagnóstico

Por estrategias de diagnóstico entendemos aquellos modelos de razonamiento que pueden utilizarse para realizar la tarea de diagnóstico, es decir los pasos que deben seguirse para a partir del conocimiento disponible llegar a la determinación de las causas que provocaron las anomalías detectadas. Las estrategias de diagnóstico pueden ser clasificadas desde muchos puntos de vista, y aquí solo abordaremos algunas de las más relevantes.

2.4.1 Estrategias de diagnóstico basadas en la naturaleza del razonamiento causal llevado a cabo

Una de las formas de identificar distintas estrategias de razonamiento ([Rasmussen, 93]) consiste en analizar la naturaleza del razonamiento causal llevado a cabo. Así pueden distinguirse por ejemplo estrategias, que partiendo de la situación normal buscan la aparición de anomalías para identificar el origen del fallo final, otras que partiendo de descripciones generales intentan su particularización a la situación específica que se presente, y otras, por último, que partiendo de casos particulares intentan generalizarlos para obtener conocimiento aplicable a nuevos casos particulares.

- Diagnóstico basado en la detección de anomalías:

Una estrategia de diagnóstico común en procesos industriales consiste en encontrar las causas de un fallo identificando anomalías en sus componentes mediante comparación entre su comportamiento actual y su comportamiento esperado o comportamiento normal (similar a como pueda trabajar el módulo de detección de anomalías). Así un experto en busca de explicaciones ante un determinado fallo puede analizar el comportamiento del proceso hacia atrás en el tiempo, en busca de un instante en el que alguno de sus componentes presente un comportamiento anormal, o alguno de los operadores haya realizado una acción inapropiada. El razonamiento se basa por tanto en partir de la situación de funcionamiento normal hasta llegar al punto de detección de alguna anomalía que en el mejor de los casos estará localizada y por tanto será fácil de identificar e interpretar. Esta estrategia no requiere conocer situaciones de fallo anteriores, pero si requiere un buen conocimiento del funcionamiento normal del proceso, y se basa en la comparación del funcionamiento medido con el funcionamiento normal esperado.

- Diagnóstico basado en casos de fallos:

Cuando no se tiene un buen conocimiento del funcionamiento normal del proceso, pero se dispone de información empírica de situaciones de fallo anteriores, es posible asociar el

fallo a una colección de síntomas que se repiten cuando dicho fallo se presenta, de forma que la observación de los síntomas permite determinar el fallo que los origina. Esta estrategia no requiere un conocimiento profundo del funcionamiento del proceso, no siempre disponible (como ocurre por ejemplo en los procesos biológicos), y se basa en la búsqueda del conjunto de síntomas más parecidos a aquellos observados entre los casos conocidos para encontrar así su causa.

- Diagnóstico basado en la generalización:

Esta estrategia, más que una estrategia de diagnóstico es una estrategia para la mejora del diagnóstico mediante un refinamiento del conocimiento disponible, y es especialmente útil de cara a evitar fallos considerados inaceptables, pero ya detectados con anterioridad. Consiste básicamente en analizar de forma temporal la evolución del proceso en situaciones de fallo anteriores, para identificar un posible cambio en el mismo que pudiese impedir o romper la secuencia de eventos conducentes al fallo. El principal problema reside en conseguir soluciones de carácter suficientemente general, y no particulares a la situación de fallo analizada, es decir en conseguir una buena generalización a partir de los casos particulares disponibles.

2.4.2 Diagnóstico en discreto y en continuo

Una importante distinción entre unos sistemas de diagnóstico y otros es su forma de trabajar en lo que respecta a su interacción con el objeto de estudio y el usuario, lo que condiciona la estrategia de diagnóstico a seguir. Los sistemas de diagnóstico pueden trabajar en continuo (sistemas “on-line”), o trabajar de forma discreta, a petición del usuario, respondiendo preguntas concretas a medida que estas son formuladas (sistemas “off-line”).

En general los sistemas de diagnóstico aplicados a procesos industriales complejos trabajan en continuo ya que a menudo estos procesos disponen de sistemas de monitorización en continuo previamente instalados, utilizados para controlar y vigilar el proceso, cuyas variables pueden ser utilizadas por el sistema de diagnóstico. Esto les permite llevar a cabo un seguimiento permanente del buen funcionamiento del proceso, vigilando constantemente el proceso a intervalos de tiempo reducidos, analizando los valores que les llegan de forma periódica de los sensores instalados ([Sanz et al, 93], [Sanz et al, 94], [Sanz, Villar, Muñoz, 94], [Dounias, Tsourveloudis, 95]).

Además el funcionamiento en continuo en los sistemas de diagnóstico facilita la predicción de fallos en su fase incipiente, antes de que su evolución los convierta en catastróficos, lo

que podría llevar al proceso a puntos de operación peligrosos tanto desde un punto de vista material, produciendo así nuevos fallos consecuencia del primero con su correspondiente coste económico, como desde un punto de vista humano.

Los sistemas de diagnóstico que trabajan en discreto pueden encontrarse por ejemplo en entornos médicos e industriales. Los sistemas de diagnóstico médico suelen trabajar a través de la interacción directa con el usuario, el médico, que a partir de su observación del paciente y de los análisis que solicita, actúa como fuente de datos para el sistema de diagnóstico. Normalmente estos sistemas van guiando el proceso de diagnóstico sugiriendo en cada paso qué medida u observación puede resultar más beneficiosa, en base al conocimiento disponible hasta ese momento y a las deducciones realizadas, para identificar las causas de las anomalías detectadas ([Shortliffe, 76], [Fieschi et al., 83], [Buisson, Farreny, Prade, 85]).

Los sistemas de diagnóstico en discreto suelen aplicarse a procesos industriales cuando su tarea se limita a diagnosticar sobre algún tipo de fallo de evolución lenta, para el que resulte suficiente tomar y analizar los datos procedentes de sus medidas a intervalos de tiempo muy espaciados ([Guth, 87], [Cantú-Ortiz, García-Espinosa, 92], [Tomsovic, Tapper, Ingvarsson, 93]), o bien cuando es necesario modificar substancialmente el estado de operación del proceso (por ejemplo llegando a su desconexión) para tomar datos representativos de su estado de salud. También se utilizan en algunos procesos en los que el diagnóstico se lleva a cabo a partir de las medidas y comprobaciones que paso a paso el sistema sugiere al usuario, para llegar al diagnóstico de la forma más rápida y directa posible, tal y como lo indique el conocimiento almacenado (como ocurre en el diagnóstico médico).

Normalmente los sistemas en continuo y en discreto se caracterizan a su vez por estrategias de razonamiento distintas. Así por ejemplo los sistemas que trabajan en continuo suelen razonar hacia adelante (véase el apartado 2.5.4). Esto significa que si el conocimiento está formulado como las posibles causas que pueden dar lugar a un síntoma determinado, requieren la observación de dicho síntoma para deducir sus causas, pero son incapaces de deducir los síntomas asociados a una causa determinada. Este tipo de estrategia se traduce generalmente en la utilización de la regla del modus ponens como principal mecanismo de inferencia (véase 2.5.1 y 2.5.4). Se dice también que este tipo de sistemas están guiados por hechos.

Por el contrario los sistemas en discreto, hacen también uso del razonamiento hacia detrás (véase el apartado 2.5.4), lo que les permite determinar qué síntomas deberían observarse para una determinada causa, o por ejemplo, en el ámbito de la toma de decisiones, les permite determinar qué condiciones deberían darse para alcanzar el objetivo marcado. Este tipo de estrategia suele traducirse en la utilización de la regla del modus tollens (véase

2.5.1 y 2.5.4). Estos sistemas son en general capaces de responder a las hipótesis formuladas por el usuario, lo que exige una mayor flexibilidad en las estrategias de razonamiento. Se dice que estos sistemas están guiados por objetivos.

2.4.3 Abducción y deducción

Una característica común a la mayoría de los sistemas de diagnóstico ([Hopgood, 93]) es el uso de la abducción (determinación de las causas dados los efectos) frente a la deducción (determinación de los efectos dadas las causas). En efecto, en los sistemas de diagnóstico se observan los efectos o síntomas producidos por unas determinadas causas, a partir de los cuales es necesario identificar de entre las posibles causas aquellas que son más probables/posibles/creíbles/etc. Esto hace que el conocimiento se formule normalmente asociando un conjunto de diagnósticos a uno o unos determinados síntomas, en lugar de como se haría si se razonara deductivamente, donde el conocimiento se formularía asociando unos determinados síntomas con una causa o fallo concreto. Mientras que en el primer caso la existencia de varias posibles causas para uno o varios síntomas introduce una cierta incertidumbre que hace que el razonamiento sea inexacto, en el segundo suele ser posible determinar casi unívocamente qué síntomas pueden observarse cuando se produce un fallo concreto.

Si el conocimiento se formula con reglas de producción (2.5.4) un ejemplo de las dos alternativas mencionadas, para un motor eléctrico, podría ser:

si
temperatura motor alta
entonces
cortocircuito en espiras del estator con certeza α_1
cortocircuito entre chapas magnéticas con certeza α_2

donde es necesario cuantificar de algún modo la certidumbre de cada uno de los posibles fallos, mientras que en la regla:

si
cortocircuito en espiras del estator
entonces
temperatura motor alta

puede considerarse casi seguro que el fallo dará lugar al síntoma señalado, por lo que la regla se puede formular sin incertidumbre. La incertidumbre en la primera de las reglas solo

podría ser eliminada si fuese posible analizar exhaustivamente todos los síntomas característicos de un determinado fallo, lo que en general no resulta ni económica ni técnicamente viable. De hecho el empleo de la abducción es una de las fuentes de incertidumbre más importante en los sistemas de diagnóstico, y suele ser casi siempre ineludible.

2.4.4 Tipo de conocimiento y estrategias de diagnóstico

El tipo de conocimiento empleado para diagnosticar condiciona la estrategia a seguir. Así por ejemplo en [Milne,87] para cada uno de los niveles de conocimiento propuesto se asocia una estrategia distinta. Al primer nivel le corresponde una estrategia basada en suponer que algunos componentes no funcionan, o que ciertas conexiones se han roto, y analizar como se ve afectado el resto de la estructura a la que pertenecen, lo que permite una primera localización de posibles componentes o grupos de componentes anómalos. La tarea de diagnóstico se completa con otras estrategias asociadas al resto de los niveles. Asociados al segundo y tercer nivel están los modelos numéricos y cualitativos del comportamiento de los componentes que permiten determinar las salidas de los mismos ante las entradas medidas, lo que permite una identificación más precisa de los componentes anómalos. La combinación de modelos de funcionamiento normal con modelos de fallos permite no solo la identificación de los componentes que fallan sino también de las causas de dichos fallos. El uso de modelos de comportamiento presenta varias limitaciones como por ejemplo la imposibilidad de conocer con suficiente precisión el modelo que explica un cierto fenómeno o comportamiento, o la imposibilidad práctica de manejar el modelo por su complejidad matemática. Esto da pie al uso de la información del cuarto y último nivel, basada principalmente en la experiencia de los operadores, y que puede codificarse eficientemente mediante el uso de reglas de producción, o basada en la experiencia directa de operación del proceso, que puede haberse ido almacenando en forma de bibliotecas de fallos.

En [Fink, Lusth, 87] el conocimiento profundo se representa mediante modelos basados en componentes genéricos (tanques, conductos, etc.) que se utilizan para simular de forma cualitativa los comportamientos normal y anómalos del proceso (varios enfoques diferentes para abordar la simulación cualitativa pueden encontrarse en [Forbus, 84], [Kleer, Brown, 84] y [Kuipers, 86], mientras que en [Shen, Leitch, 93] puede encontrarse una aproximación desde la teoría de conjuntos borrosos). Además mediante unas reglas de carácter general el uso de los modelos permite guiar el proceso de diagnóstico sugiriendo qué nuevas medidas o comprobaciones deben hacerse, y qué componentes deben analizarse por ser susceptibles de funcionar incorrectamente. Por otro lado el conocimiento superficial se representa mediante una red semántica subdividida en tres niveles, el primero

para el conocimiento operativo (medidas, observaciones, etc.), el segundo para las hipótesis sobre el estado o anomalías de los componentes, y el tercero para los síntomas observables asociados a las distintas hipótesis, comprobaciones de síntomas y soluciones a los problemas observados. Los dos tipos de conocimiento se integran en un único sistema que permite compartir las hipótesis que provienen del uso de cada uno de los tipos de conocimiento. La tarea de diagnóstico se aborda primero utilizando el conocimiento superficial, generalmente más rápido, y después el conocimiento profundo, cuando los resultados del superficial no son suficientemente satisfactorios.

En [Béchir, 94] la estrategia de control se articula mediante las llamadas “reglas del dedo gordo”, los procedimientos de diagnóstico y los heurísticos. Los heurísticos, de carácter muy general, permiten escoger de entre una lista de componentes susceptibles de estar fallando, cuáles van a ser analizados primero. Los procedimientos de diagnóstico son procedimientos que permiten determinar qué componentes de una lista de componentes son los causantes del fallo o anomalía del componente al que pertenecen. Por último las reglas del dedo gordo son un último paso para determinar, a partir de observaciones cualitativas localizadas, qué componentes del paso anterior, causantes del fallo, son los que realmente están fallando.

En resumen, según el tipo de conocimiento empleado, la estrategia de diagnóstico queda fuertemente condicionada, y podemos distinguir las siguientes:

- razonamiento basado en heurísticos: este conocimiento, en la mayoría de los casos superficial, se representa generalmente mediante reglas de producción que asocian causas y síntomas, y metarreglas de diagnóstico que permiten determinar el orden de actuación del sistema (como qué componentes analizar primero, etc.). Las reglas de producción y metarreglas se describen en el apartado 2.5.4. Los sistemas basados en reglas entran dentro de los llamados sistemas expertos, y se describen en el apartado 2.7. El orden de disparo de las reglas del sistema admite distintos tipos de estrategias que también se comentarán más adelante.
- razonamiento basado en modelos: este conocimiento, generalmente profundo, se representa mediante modelos, formulados en términos de restricciones, modelos que pueden ser cualitativos o cuantitativos. Se utilizan tanto a nivel del proceso en su conjunto, para modelar la planta como una interrelación de componentes que son vistos como cajas negras con entradas y salidas, como a nivel de componentes, con los que puede describirse tanto su comportamiento normal como su comportamiento en presencia de algún tipo de anomalía. El uso de modelos da lugar a varios tipos de estrategias significativamente diferentes de las que destacamos dos. Por un lado están las que se basan en modelos de paridad que utilizan los modelos para estimar medidas de componentes que caracterizan su comportamiento normal, que son comparadas con

las medidas reales para detectar diferencias significativas, reflejo de algún tipo de anomalía. Por otro lado están los que se basan en modelos de comportamiento global o de fallo del proceso y componentes, basados generalmente en la física que los gobierna, que se utilizan normalmente para simular situaciones de fallo que posteriormente son también comparadas con el comportamiento medido.

- razonamiento basado en casos: este conocimiento, generalmente superficial, se representa mediante una descripción cualitativa o numérica del estado de los componentes bajo distintos tipos de comportamientos anómalos, previamente detectados e identificados. El diagnóstico se hace comparando en cada caso el estado real (medido) del componente con alguno de los casos almacenados previamente en el sistema. Los casos, aunque normalmente se obtienen de ejemplos de funcionamiento real del proceso, pueden también ser obtenidos mediante simulación exhaustiva de fallos a partir de los modelos de comportamiento del proceso y de los componentes.

La asociación de técnicas de representación de conocimiento (véase el apartado 2.5) con tipos de conocimiento representados puede en ocasiones conducir a confusiones, ya que mediante reglas puede modelarse conocimiento profundo, como por ejemplo el caso de los modelos borrosos, mientras que los modelos de comportamiento normal podrían englobarse en el conocimiento superficial puesto que su objetivo no es profundizar en la física del comportamiento interno de los componentes del proceso, sino únicamente en la manifestación externa de su comportamiento en ausencia de fallos.

2.5 Técnicas de representación del conocimiento

El objetivo de las técnicas de representación del conocimiento consiste en proporcionar mecanismos que permitan almacenar informáticamente el conocimiento de que se dispone para el problema concreto que se pretende resolver. Este conocimiento debe poder ser utilizado con eficacia por algún mecanismo de razonamiento con objeto de inferir conclusiones lógicas y/o llevar a cabo acciones inteligentes. En general las técnicas de representación del conocimiento condicionan la técnica de razonamiento a emplear, estando ambas íntimamente ligadas.

Tradicionalmente se ha dividido el conocimiento en declarativo y procedural, división que se muestra cada vez más artificial. El conocimiento declarativo se representa normalmente mediante la lógica formal, los frames, los scripts y las redes semánticas. Por declarativo se entiende un conocimiento que se enuncia principalmente como una colección de hechos más o menos estructurados, al que se suman ciertos métodos que explicitan cómo dicho conocimiento debe ser usado. El conocimiento procedural por otro lado, se representa típicamente mediante las llamadas reglas de producción, que llevan directamente incluido cómo proceder.

Desde una perspectiva más actual las técnicas de representación del conocimiento se dividen en cuatro categorías principales ([Popovic, Bhatkar, 94]), los esquemas lógicos que se basan en la utilización de la lógica formal, los esquemas procedurales basados en las reglas de producción, los esquemas en forma de redes como son las redes semánticas, y los esquemas estructurados entre los que se cuentan principalmente los frames, scripts y objetos. Muchos de estos esquemas se apoyan directa o indirectamente en un esquema más elemental, el esquema (atributo, objeto, valor), con el que se suele representar los fragmentos de conocimiento más elementales o hechos que permiten definir algo sobre los objetos que constituyen el mundo real.

Queremos mencionar también el uso de modelos (ya sean matemáticos o cualitativos) como un componente fundamental de muchos sistemas de diagnóstico. Los modelos se estudian generalmente desde disciplinas diferentes de la Ingeniería del Conocimiento, pero se utilizan con frecuencia en la tarea de detección de anomalías gracias a la redundancia analítica que proporcionan ([Gertler, 91], [Muñoz, 96]), o bien para la fase de identificación de anomalías ([Fink, Lusth, 87]) donde se utilizan para extraerles conocimiento, dando lugar a los esquemas de razonamiento basados en modelos.

Aunque la mayor parte de los sistemas de diagnóstico son sistemas basados en reglas (entre los que destacan algunos ya clásicos como son MYCIN, [Shortliffe, 76], PROSPECTOR, [Duda, Gaschnig, Hart, 81] y DENDRAL, [Buchanan, Feigenbaum,

78]), casi todos utilizan conjuntamente varios esquemas de representación del conocimiento según su tipo. Así por ejemplo, como ya hemos comentado, el conocimiento estructural y de operación se presta especialmente al uso de esquemas estructurados mientras que el conocimiento funcional se presta más al uso de modelos matemáticos y/o reglas de producción.

2.5.1 Lógica formal

Un sistema lógico, desde el punto de vista clásico, consiste en ([Kleene, 67], [Dubois, Prade, 91b], [Popovic, Bhatkar, 94]):

- un lenguaje o gramática que permite generar fórmulas bien formuladas, en inglés “well-formed-formulae” y que abreviaremos con WFF. Está compuesto, entre otros, de símbolos que representan átomos de información, operadores lógicos que permiten combinar de forma lógica los átomos y las WFF, cuantificadores lógicos, leyes de formación de WFF a partir de los átomos, etc.
- una semántica, definida mediante interpretaciones que asocian a cada átomo y a cada WFF un valor de verdad perteneciente a un conjunto L . Dentro de L se escoge un subconjunto de valores de verdad tales que cualquier WFF cuyo valor de verdad pertenezca a dicho subconjunto se dice que es válida. Si una WFF es válida independientemente de la interpretación, se denomina tautología. Normalmente el subconjunto escogido suele ser el máximo valor de L . En lógica clásica $L = \{0, 1\}$ donde 0 representa el valor de verdad *falso* y 1 representa al valor de verdad *verdadero*, y son tautologías todas aquellas WFF con valor de verdad 1 independientemente del valor de verdad de los átomos que la componen.
- una sintaxis que se compone de:
 - axiomas (o premisas)
 - reglas de inferencia

Los axiomas son los fragmentos de conocimiento fundamental, que se suponen válidos, y que se utilizan junto con las reglas de inferencia para obtener nuevas fórmulas válidas denominadas *teoremas*, siendo el procedimiento de obtención su *prueba*. La prueba consiste en una sucesión de WFF válidas que conducen al teorema, de forma que una WFF puede suceder a otra si es un axioma o si es válida. Las reglas de inferencia permiten obtener WFF válidas a partir de otras WFF también válidas. La sintaxis se

dice consistente (“sound”) si todos los teoremas son tautologías y completa si todas las tautologías son teoremas.

El proceso de razonamiento lógico consiste en, a partir de los axiomas o premisas que suelen corresponder con observaciones realizadas, hipótesis de trabajo, etc., aplicar las reglas de inferencia para obtener nuevos teoremas o conclusiones.

Desde el punto de vista del conocimiento empleado, los axiomas o premisas son el conocimiento sobre el dominio en el que se razona, que permiten describirlo, conocimiento que se supone verdadero, mientras que las reglas de inferencia son el metaconocimiento, que es el conocimiento sobre cómo debe razonarse, y que es independiente del dominio.

- Lógica proposicional:

La lógica proposicional ([Kleene, 67]) parte de un conjunto de frases o sentencias, cuya estructura interna es indiferente, denominadas átomos, y que se representan por un conjunto de símbolos p_1, p_2, \dots, p_n . Las WFF se obtienen combinando átomos y WFF entre sí mediante un conjunto de operadores lógicos, {‘.’, ‘+’, ‘-’, ‘→’, ‘≡’}, que se interpretan del siguiente modo:

- ‘ $p_1.p_2$ ’ como ‘ p_1 y p_2 ’
- ‘ p_1+p_2 ’ como ‘ p_1 o p_2 ’
- ‘ $\neg p_1$ ’ como ‘no p_1 ’
- ‘ $p_1 \rightarrow p_2$ ’ como ‘ p_1 implica a p_2 ’ o bien ‘si p_1 entonces p_2 ’
- ‘ $p_1 \equiv p_2$ ’ como ‘ p_1 equivale a p_2 ’ o bien ‘si p_1 entonces p_2 y si p_2 entonces p_1 ’

Estas son las WFF más sencillas que pueden construirse a partir de los átomos, además de los propios átomos del lenguaje (que obviamente también son WFF). Se denomina proposiciones a los significados de los átomos y WFF, expresados en algún otro lenguaje (por ejemplo el lenguaje natural).

Cada uno de los operadores lógicos se define mediante una tabla de verdad que permite obtener el valor de verdad de la WFF a la que da lugar, a partir del valor de verdad de sus argumentos, obteniéndose las bien conocidas tablas de verdad, que quedan resumidas en el siguiente cuadro:

p_1	p_2	$p_1 \cdot p_2$	$p_1 + p_2$	$\neg p_1$	$p_1 \rightarrow p_2$	$p_1 \equiv p_2$
0	0	0	0	1	1	1
0	1	0	1	1	1	0
1	0	0	1	0	0	0
1	1	1	1	0	1	1

Para evitar la ambigüedad que puede aparecer en fórmulas compuestas por varios operadores, respecto al orden en que deben ser evaluados, se define la precedencia entre operadores mediante el siguiente orden:

$$\equiv, \rightarrow, \cdot, +, \neg$$

y a menudo se añaden dos símbolos adicionales al lenguaje, como son los paréntesis { ‘(’, ‘)’ } que permiten hacer explícito el orden de evaluación cuando la precedencia no resulta suficientemente clara.

Los elementos anteriores (tablas de verdad y orden de evaluación) permiten obtener el valor de verdad de cualquier WFF. Las tautologías son todas aquellas WFF que tienen el valor de verdad 1 independientemente de los valores de verdad de los átomos y WFF que las componen.

Se dice que una WFF P es válida si su valor de verdad es 1 y normalmente se escribe ‘ $\models P$ ’. Nótese que ‘ \models ’ no es un símbolo del lenguaje del sistema lógico, sino sencillamente una forma abreviada de escribir *es válido* o *es verdadero*. Cuando a partir de un conjunto de WFF P_1, P_2, \dots, P_n se prueba un teorema P se suele escribir de forma abreviada $P_1, P_2, \dots, P_n \models P$. Al igual que antes el símbolo ‘ \models ’ no es un símbolo del lenguaje del sistema lógico. Los símbolos ‘ \models ’ y ‘ \square ’ (entre otros) forman parte del llamado metalenguaje que es el lenguaje que utiliza el observador para describir el proceso deductivo (o pruebas) de unas WFF válidas a otras.

Las reglas de inferencia que forman parte del metaconocimiento se expresan mediante el metalenguaje, y las WFF que en ellas aparecen son WFF que se suponen válidas, denominadas axiomas lógicos. Algunas de las reglas de inferencia más comunes son:

- regla del modus ponens:

$$P, P \rightarrow Q \models Q$$

que se interpreta como si P es válida y $P \rightarrow Q$ es válida entonces Q también es válida, donde P y $P \rightarrow Q$ se denominan premisas y Q conclusión (por comodidad hemos omitido el símbolo \models para indicar que las WFF son válidas). Otros símbolos habituales

en el metalenguaje permiten expresar las reglas de inferencia con una apariencia distinta (que utilizaremos a menudo en esta tesis):

$$\frac{P \quad P \rightarrow Q}{Q}$$

- regla del modus tollens:

$$\neg Q, P \rightarrow Q \quad \square \quad \neg P$$

- regla de reducción:

$$P, \neg P \quad \text{es una contradicción}$$

Esta regla permite probar teoremas mediante reducción al absurdo.

- regla de sustitución:

Sea F una WFF que contiene únicamente los átomos p_1, p_2, \dots, p_n , y F^* la WFF que se obtiene a partir de F reemplazando simultáneamente los átomos p_1, p_2, \dots, p_n por las WFF P_1, P_2, \dots, P_n .

La regla de sustitución dice que si $\models F$ entonces $\models F^*$.

Nótese que la regla recíproca no es cierta.

- regla de encadenamiento:

$$P_1 \rightarrow P_2, P_2 \rightarrow P_3 \quad \square \quad P_1 \rightarrow P_3$$

- regla de resolución:

$$(P_1 + P_2), (\neg P_2 + P_3) \quad \square \quad (P_1 + P_3)$$

Esta regla es especialmente interesante de cara a la automatización del razonamiento en lógica formal, y es la base del lenguaje PROLOG en el que está implantada como principal mecanismo de inferencia.

- Lógica de predicados:

La lógica de predicados es una extensión de la lógica proposicional a la que se añade la posibilidad de utilizar variables. Cuando en los átomos del lenguaje se contempla también

el uso de predicados, es decir expresiones del tipo $f(x_1, x_2, \dots, x_m)$ donde x_1, x_2, \dots, x_m son variables, la lógica resultante se denomina lógica de predicados de primer orden. Nótese que las variables son nuevos símbolos del lenguaje de la lógica de predicados. Cuando $m=0$ entonces se recupera la lógica proposicional. Cada función f establece una relación entre las variables x_1, x_2, \dots, x_m , y un predicado $f(x_1, x_2, \dots, x_m)$ no es una proposición, pero da lugar a proposiciones cuando a cada una de las variables x_1, x_2, \dots, x_m se les asigna un valor. Los átomos del lenguaje están ahora compuestos por los de la lógica proposicional p_1, p_2, \dots, p_n y todos los que se pueden formar mediante las funciones o predicados f_i que se definan, después de reemplazar sus variables por alguna combinación obtenida a partir de las proposiciones p_1, p_2, \dots, p_n , pudiendo cualquiera de ellas aparecer más de una vez como argumentos de los predicados f_i .

Además de los operadores de la lógica proposicional se definen otros operadores, los cuantificadores existencial \exists y universal \forall . La gramática para la formación de las WFF es la misma que en la lógica proposicional, combinando WFF mediante los operadores lógicos antes definidos, añadiendo las fórmulas obtenidas a partir de una WFF P con alguno de los cuantificadores \exists o \forall , esto es $\exists P$ y $\forall P$. La precedencia de \forall y \exists es menor que la del operador unario \neg . Estos operadores se interpretan del siguiente modo:

- ‘ $\forall x$ ’ como ‘para todo x ’ y expresa que una WFF es válida para cualquiera de los valores que puedan tomar sus variables.
- ‘ $\exists x$ ’ como ‘existe x ’ y expresa que una WFF es válida para al menos una combinación de valores de sus variables.

Mediante la lógica de predicados es posible expresar por ejemplo:

$$\begin{aligned} & \text{caliente}(\text{motor}) \\ & \exists x \text{caliente}(x) \rightarrow \text{avería}(x) \\ & \forall x \text{avería}(x) \rightarrow \text{desconectar}(x) \end{aligned}$$

donde los predicados son $\text{caliente}(x)$, $\text{avería}(x)$, $\text{desconectar}(x)$ y motor es una constante. De las WFF anteriores, mediante la aplicación de reglas de inferencia, podría deducirse la conclusión $\text{desconectar}(\text{motor})$. Una forma más precisa de representar lo mismo podría ser:

$$\begin{aligned} & \text{tener}(\text{motor}, \text{alta}(\text{temperatura})) \\ & \exists x \text{tener}(x, \text{alta}(\text{temperatura})) \rightarrow \text{avería}(x) \\ & \forall x \text{avería}(x) \rightarrow \text{desconectar}(x) \end{aligned}$$

La lógica formal o parte de ella está en la base de la mayoría de las estrategias de razonamiento de los sistemas basados en conocimiento. En particular los sistemas basados

en reglas guardan un estrecho paralelismo con la lógica de predicados, siendo la regla del modus ponens su principal mecanismo de inferencia. Sin embargo se hace necesario el uso de técnicas de representación de la incertidumbre, no disponibles en lógica clásica, lo que da lugar a distintas modificaciones del modus ponens y a la aparición de nuevas lógicas llamadas comúnmente no clásicas.

2.5.2 Esquemas estructurados: frames, scripts y objetos

Estos esquemas de representación del conocimiento son unos de los más utilizados en el ámbito de los sistemas basados en conocimiento. De hecho los objetos, actualmente muy presentes en los lenguajes de programación modernos, provienen de la Inteligencia Artificial, en particular de los frames y estructuras similares, de donde han heredado muchas de sus características principales.

Estos esquemas de representación de conocimiento se basan en el uso de las tripletas (atributo, objeto, valor), que es una forma elemental de formular los hechos que describen el mundo real. Para ello el mundo real se divide en objetos o entes conceptuales que pueden corresponder con objetos físicos (por ejemplo componentes) o conceptos. A cada objeto del mundo real se le asocia una serie de atributos que pueden admitir un conjunto finito o infinito de posibles valores. El conocimiento se expresa por tanto mediante tripletas (atributo, objeto, valor), como por ejemplo (*temperatura, motor, alta*) que se interpreta como “*la temperatura del motor es alta*”. Este esquema permite formular el conocimiento sobre un determinado objeto enumerando todos sus atributos con sus respectivos valores. Nótese el paralelismo existente con *tener(motor,alta(temperatura))* que expresa el mismo conocimiento desde la perspectiva de la lógica de predicados.

Los valores que el atributo de un objeto puede tomar pueden ser en general de múltiples tipos, numéricos o simbólicos, un único valor o todo un conjunto de valores, y cada valor puede tener asociado un cierto contenido de incertidumbre en sentido amplio, para lo que es necesario utilizar alguna técnica de representación de incertidumbre como las que se comentan en el apartado 2.6. Esto hace que el valor del atributo en ciertos casos sea relativamente complejo (un valor numérico y un coeficiente de certeza, una función de densidad de probabilidad, una distribución de posibilidad, etc.) dependiendo de la técnica empleada.

Cuando en un mismo esquema de representación se agrupan todos los atributos de un objeto o tipo de objeto se obtienen los esquemas de representación de conocimiento estructurados, como son los frames, los scripts o los objetos.

Los frames ([Popovic, Bhatkar, 94], [Dillon, Lee, 93]) son estructuras de conocimiento que agrupan a una serie de campos o “slots” que describen el estereotipo de algún objeto, evento, etc. Los campos pueden ser de distintos tipos (numéricos, simbólicos, etc.) y contener datos que son su o sus valores o “fillers”. Los frames pueden disponer además de mecanismos adicionales llamados “facets” (que forman parte de una categoría más general de funciones denominadas “daemons”, ejecutadas únicamente cuando ocurre algún evento particular y no en un punto concreto de un programa). Estos mecanismos permiten por ejemplo delimitar o restringir los valores de un campo por sí mismo o en relación con los valores de otros (los “value-facets”), proporcionar valores por defecto a un campo (los “default-facets”) o llamar a procedimientos o funciones especiales cuando se accede al valor de un campo (los “if-needed-facets”). Además los frames pueden heredar campos de otros frames y pueden contener a su vez otros frames, lo que los convierte en una técnica estructurada y jerárquica de representación del conocimiento muy potente.

Los scripts son estructuras de conocimiento similares a los frames que se utilizan para representar una secuencia estereotipada de acciones o eventos, como por ejemplo los eventos y/o acciones que se dan cuando se repara un determinado tipo de componente, etc.

Los objetos (desde la perspectiva de la orientación a objetos, que no deben ser confundidos con los objetos del mundo real) son estructuras de conocimiento que permiten agrupar bajo una misma entidad un conjunto de datos relacionados entre sí y las funciones que operan con dichos datos. Estos objetos llamémosles informáticos están a su vez asociados con objetos del mundo real fácilmente detectables e identificables, con comportamientos claramente diferenciados, lo que hace el diseño orientado a objetos especialmente atractivo. Aunque el estudio de la representación del conocimiento mediante objetos se aborda más en otras disciplinas informáticas, a nuestro entender pueden considerarse como otra forma más de representar el conocimiento alternativa por ejemplo a los frames. De carácter mucho más general que éstos, los objetos permiten la creación de nuevos tipos de datos para los que puede definirse un determinado comportamiento o funcionalidad. La representación del conocimiento en forma de objetos se basa en la creación de clases para definir tipos de datos, haciendo uso de características particulares de entre las cuales podemos destacar el encapsulamiento, consistente en la ocultación de aquellos datos y funciones internas del objeto que no deben ser ni vistas ni utilizadas desde ningún otro objeto, y la herencia, que permite que una clase derivada de otra herede todas sus características, de forma que la clase derivada contiene automáticamente todos los datos y funciones de la clase padre. En [Dillon, Lee, 93] puede encontrarse una buena introducción a los sistemas orientados a objetos. En este trabajo se señalan también las principales diferencias entre frames y objetos, debidas principalmente a la mayor funcionalidad de los frames (gracias a los daemons y a las reglas asociadas a campos), y la

mayor flexibilidad de los objetos (para los que se puede crear complejos métodos de acceso y control de sus campos, que además gozan de la encapsulación).

Los esquemas estructurados son especialmente útiles para la representación del conocimiento en aquellos dominios en los que éste está muy estructurado. Este es el caso por ejemplo, en los sistemas de diagnóstico, del conocimiento estructural del proceso bajo vigilancia, que suele representarse típicamente mediante frames, o de forma más general y flexible mediante objetos.

2.5.3 Redes semánticas

Las redes semánticas permiten representar gráficamente interrelaciones entre objetos. Estas relaciones se representan mediante grafos orientados, cuyos nudos son los objetos y cuyos arcos representan relaciones entre objetos ([Quillian, 68]). La denominación de semánticas viene del hecho de que cada arco se corresponden con una relación que tiene una interpretación semántica. El hecho de que los nudos sean objetos permite combinar fácilmente este esquema de representación con los esquemas estructurados.

2.5.4 Reglas de producción

Las reglas de producción son una de las técnicas de representación del conocimiento más extendidas en los sistemas basados en conocimiento, y en la que principalmente se basan la mayoría de sistemas de diagnóstico, y tienen su origen en los sistemas de producción utilizados originalmente en el ámbito económico ([Post, 43]). Una regla está compuesta por dos partes, los antecedentes, hipótesis o condiciones, y los consecuentes (nótese el estrecho paralelismo entre una regla de producción y la regla de inferencia del modus ponens en lógica formal.):

si
 condición P_1
 condición P_2
 ...
 condición P_n
entonces
 conclusión Q_1
 conclusión Q_2
 ...

conclusión Q_m

o de forma abreviada:

$$P_1, P_2, \dots, P_n \rightarrow Q_1, Q_2, \dots, Q_m$$

Los consecuentes de la regla pueden ser a su vez antecedentes de otras reglas, o acciones que deban ser llevadas a cabo cuando la regla se dispara. Cuando todas las condiciones de la regla son verdaderas o se cumplen (interpretación conjuntiva del conjunto de condiciones) diremos que la regla está **activada**. Cuando sus conclusiones estén disponibles para activar a otras reglas, o sus acciones sean ejecutadas, diremos que la regla se ha **disparado**.

El disparo de una regla se basa en la aplicación del modus ponens, y suele escribirse:

$$\frac{P_1, P_2, \dots, P_n \quad P_1, P_2, \dots, P_n \rightarrow Q_1, Q_2, \dots, Q_m}{Q_1, Q_2, \dots, Q_m}$$

donde $P_1, P_2, \dots, P_n \rightarrow Q_1, Q_2, \dots, Q_m$ es la regla de producción o simplemente regla, P_1, P_2, \dots, P_n es el conjunto de observaciones y Q_1, Q_2, \dots, Q_m es el conjunto de conclusiones. Es importante en este punto distinguir entre antecedente y observación. Los primeros son aquellos que deben cumplirse u observarse para que la regla se dispare, mientras que los segundos son aquellos hechos que se sabe que son ciertos o que han sido observados. Del mismo modo los consecuentes son aquellas conclusiones que se deducirían de la regla si las observaciones disponibles permitiesen confirmar el total cumplimiento de sus antecedentes. Esta distinción es especialmente importante en el ámbito de la lógica borrosa, donde el modus ponens generalizado permite concluir conclusiones diferentes de los consecuentes, cuando las observaciones son diferentes de los antecedentes, es decir cuando se produce un cumplimiento parcial de los antecedentes de la regla (el capítulo 4 de la tesis se dedica al análisis del modus ponens en el ámbito de la lógica borrosa, que se introduce en el capítulo 3).

Podría también suponerse que el conjunto de antecedentes de la regla se interpreta disyuntivamente, de forma que el cumplimiento de uno solo de ellos fuese suficiente para activar la regla. Sin embargo este tipo de reglas es equivalente a tantas reglas como antecedentes tenga la regla original, todas ellas con los mismos consecuentes, iguales a los de la regla original, por lo que de ahora en adelante supondremos siempre que el conjunto de antecedentes se interpreta conjuntivamente.

Cuando se utiliza el modus ponens para inferir se dice que el razonamiento es hacia adelante, también llamado “forward reasoning” o “forward chaining”. Como ya hemos señalado este tipo de razonamiento suele utilizarse en los sistemas que trabajan en continuo, a los que constantemente llegan un conjunto de observaciones susceptibles de disparar las reglas del sistema.

Sin embargo también es posible utilizar la regla del modus tollens, en cuyo caso la observación que dispara la regla viene dada por $\neg Q_1, \neg Q_2, \dots, \neg Q_m$ y la conclusión es entonces $\neg P_1, \neg P_2, \dots, \neg P_n$. En este caso queda patente la diferencia entre observación y antecedente por un lado, y conclusión y consecuente por otro. Este tipo de razonamiento se denomina razonamiento hacia atrás, también llamado “backward reasoning” o “backward chaining”, que como hemos comentado se utiliza a menudo en sistemas orientados a responder consultas o a la toma de decisiones, a los que a menudo se les presenta una serie de objetivos en espera de la respuesta de cómo lograrlos.

Los antecedentes P_i de las reglas pueden ser a su vez unidades de conocimiento que generalmente se expresan mediante el esquema (atributo, objeto, valor), lo que hace perfectamente complementables los esquemas de representación de conocimiento estructurados con las reglas de producción. También es posible utilizar variables en la formulación de las reglas (que en el ámbito de los sistemas basados en reglas suelen denominarse comodines). Así por ejemplo del conocimiento:

temperatura-motor-alta

*si temperatura-x-alta
entonces estado-x-averiado*

*si estado-x-averiado
entonces desconectar-x*

se deduce que el motor debe ser desconectado, mediante la sustitución de la variable x por el valor *motor* y la aplicación sucesiva del modus ponens, lo que recuerda estrechamente a la lógica de predicados.

Las reglas son uno de los principales esquemas de representación del conocimiento funcional en los sistemas de diagnóstico. Aunque el conocimiento funcional codificado en forma de reglas suele corresponder con el conocimiento superficial, es posible también representar conocimiento profundo. Los sistemas que utilizan reglas de producción se denominan sistemas expertos basados en reglas, y se describen en el apartado 2.7

Existe un tipo especial de reglas, utilizadas para expresar conocimiento, no sobre el dominio sobre el que se diagnostica o se tiene que tomar decisiones, sino sobre cómo debe llevarse a cabo el proceso de razonamiento, como por ejemplo qué reglas deben ser disparadas primero ante unas determinadas observaciones. Estas reglas se denominan metarreglas, y el conocimiento que en ellas se enuncia se denomina metaconocimiento.

2.5.5 Modelos

Aunque los modelos generalmente no se estudian en el ámbito de la Inteligencia Artificial, dada su importancia en los sistemas de diagnóstico, nosotros los consideraremos como una técnica más de representación del conocimiento, y como veremos serán de especial utilidad para la fase de detección de anomalías.

Por modelo se entiende un sistema con unas determinadas entradas y salidas, de forma que a partir de los valores de las entradas, el modelo es capaz de calcular los valores de sus salidas. Los valores de las variables deben ser entendidos en sentido amplio, como ocurría con el esquema (atributo, objeto, valor) pudiendo ser numéricos, simbólicos, distribuciones de densidad de probabilidad, con o sin incertidumbre, etc., dependiendo del tipo de modelo empleado.

En los sistemas de diagnóstico los modelos se utilizan habitualmente para modelar alguna pauta de comportamiento de un componente, ya sea su comportamiento normal o su comportamiento en alguno de sus modos de fallo. Las variables de los modelos se escogen entre las entradas y variables de estado de los componentes, o variables contextuales (por ejemplo temperatura ambiente) que puedan afectar al funcionamiento del componente, y la salida suele ser la estimación de alguna medida dinámica significativa del funcionamiento del componente.

Existen multitud de técnicas para obtener modelos, desde la utilización del conocimiento profundo para la elaboración de modelos físicos, en los que tanto las variables como los parámetros del modelo tienen una interpretación en términos físicos, hasta la utilización de técnicas estadísticas para la obtención de modelos de caja negra, donde los parámetros son difícilmente interpretables, pasando por los modelos cualitativos, modelos basados en aproximadores funcionales ajustados a partir de datos de operación, modelos borrosos, etc.

2.6 Técnicas de representación de incertidumbre e inferencia

Como se vio en el apartado 2.3.2 la incertidumbre es una característica del conocimiento humano siempre presente en mayor o menor medida. La incertidumbre en su acepción más amplia incluye como ya hemos comentado imprecisión, vaguedad e incertidumbre (entendida en este caso en su acepción más restringida).

La mayoría de las técnicas de representación de incertidumbre están enfocadas a su representación en su sentido más restringido, es decir a evaluar en qué medida se conoce el grado de verdad de una proposición. Este es el enfoque que se adopta tanto desde el punto de vista probabilístico, como de los coeficientes de certeza o desde el punto de vista de la teoría de la evidencia.

Los enfoques probabilísticos (principalmente basados en el teorema de Bayes) tienen la gran desventaja de que no pueden ser utilizados para representar la ausencia total de información o desconocimiento. En efecto, el mayor desconocimiento que sobre la ocurrencia de un suceso P puede ser modelado viene dado por la probabilidad $\mathcal{P}(P)=0.5$, pero esto implica automáticamente $\mathcal{P}(\neg P)=0.5$, haciendo a P y a $\neg P$ equiprobables, lo que tiene un gran contenido informativo. Un mayor desconocimiento exige por ejemplo no aportar evidencia en favor de P sin aportarla en favor de $\neg P$. Por otro lado, en entornos reales donde la información suele ser a menudo incompleta, resulta difícil estimar las probabilidades de unos u otros sucesos, por lo que se recurre a las llamadas probabilidades subjetivas, debiendo cuidarse la posible aparición de inconsistencias. Por último, aunque tiene la ventaja de tener un sólido fundamento teórico, requiere demasiadas hipótesis para que muchas de sus herramientas tengan validez (como por ejemplo las hipótesis de independencia) lo que, o bien dificulta considerablemente su aplicación práctica, o bien reduce substancialmente su validez teórica cuando se aplica. Otros enfoques más recientes utilizan probabilidades intervalo valoradas ([Walley, 96]), basados en la estimación de una cota inferior y otra superior para el valor de probabilidad real pero desconocido, lo que los asemeja mucho a la teoría de la evidencia, flexibilizando así las restrictivas hipótesis de validez que exige el uso de probabilidades.

Los enfoques basados en coeficientes de certeza y teoría de la evidencia sí permiten modelar el desconocimiento sobre la ocurrencia de un suceso P . Los coeficientes de certeza han sido muy utilizados pero también muy criticados por su escaso fundamento teórico, aunque como se señala en [Prade, 85] son parcialmente consistentes con la teoría de la posibilidad, y presentan fuertes semejanzas con algunos aspectos de la teoría de la evidencia. La teoría de la evidencia por su parte presenta un marco teórico muy sólido e interesante ya que permite representar el total desconocimiento, define y utiliza dos medidas que pueden interpretarse como cotas inferior y superior de una probabilidad

desconocida, y además incluye (bajo ciertas hipótesis) a las teorías de la probabilidad y de la posibilidad como casos particulares. Todos estos enfoques sin embargo sólo permiten modelar la incertidumbre pero no la vaguedad de las proposiciones, es decir la existencia de verdades parciales conocidas con precisión.

La teoría de conjuntos borrosos, combina la extensión de las lógicas multivaluadas a los conjuntos borrosos con la teoría de la posibilidad, lo que permite manejar en un único marco teórico la incertidumbre y la vaguedad, y modelar el desconocimiento total de forma similar a como se hace en la teoría de la evidencia. Estas características la convierten en un herramienta especialmente atractiva y flexible para la representación del conocimiento tanto incierto como vago, y será la técnica de representación de incertidumbre escogida en esta tesis.

Dado que las principales técnicas de representación de incertidumbre se fundamentan en el uso de medidas de incertidumbre, estas serán el punto de partida de esta exposición.

2.6.1 Las medidas de incertidumbre

En la base de la mayoría de las teorías de representación de incertidumbre están las medidas de incertidumbre ([Prade, 85], [Klir, Folger, 88], [Alvarez, Castillo, 94]), también llamadas medidas borrosas por el profesor Sugeno. Una medida de **incertidumbre** g dado un universo de discurso U es una función de conjunto que verifica las siguientes propiedades:

$$\begin{aligned} g: \mathfrak{S}(U) &\rightarrow [0,1] && \text{(o bien } \mathcal{P}(U) \rightarrow [0,1] \text{ para conjuntos clásicos)} \\ g(\emptyset) &= 0 \\ g(U) &= 1 \\ \text{si } P_1 \subset P_2 &\text{ entonces } g(P_1) \leq g(P_2) && \forall P_1, P_2 \in \mathfrak{S}(U) \end{aligned}$$

donde $\mathfrak{S}(U)$ es el conjunto de todos los conjuntos borrosos definidos en U . Es necesario añadir una propiedad adicional de forma que si $P_1 \subset P_2 \subset P_3 \subset \dots \subset P_n$ entonces se verifique:

$$\lim_{i \rightarrow \infty} g(P_i) = g(\lim P_i)$$

Las medidas de incertidumbre más comunes y que repasaremos brevemente en este capítulo son las medidas de probabilidad, de plausibilidad y credibilidad, y de posibilidad y necesidad.

2.6.2 Las medidas de probabilidad

2.6.2.1 Definición

Una medida de probabilidad es una medida de incertidumbre $\mathcal{P}r$, definida para conjuntos clásicos. Dados dos conjuntos P_1 y P_2 definidos en U una medida de probabilidad verifica:

$$\text{si } P_1 \cap P_2 = \emptyset \text{ entonces } \mathcal{P}r(P_1 \cup P_2) = \mathcal{P}r(P_1) + \mathcal{P}r(P_2)$$

y en general cuando la intersección de P_1 y P_2 no es nula se tiene:

$$\mathcal{P}r(P_1 \cap P_2) = \mathcal{P}r(P_1) + \mathcal{P}r(P_2) - \mathcal{P}r(P_1 \cup P_2)$$

Una medida de probabilidad puede definirse a través de una distribución de probabilidad en el caso en que el dominio U sea discreto, o una función de densidad de probabilidad $f_{dp}(u)$ cuando el dominio es continuo. Una distribución de probabilidad $pr(u)$ es una función que verifica:

$$pr: U \rightarrow \mathfrak{R} \text{ tal que } \sum_{u \in \mathfrak{R}} pr(u) = 1$$

y en su versión continua una función de densidad de probabilidad $f_{dp}(u)$ verifica:

$$f_{dp}: U \rightarrow \mathfrak{R} \text{ tal que } \int_{u \in \mathfrak{R}} f_{dp}(u) du = 1$$

Una medida de probabilidad de un conjunto P puede obtenerse a partir de una distribución de probabilidad pr como:

$$\mathcal{P}r(P) = \sum_{u \in P} pr(u)$$

o en su versión continua:

$$\mathcal{P}r(P) = \int_{u \in P} f_{dp}(u) du$$

Aunque las medidas de probabilidad se definen para conjuntos clásicos existen algunas extensiones a conjuntos borrosos (véase por ejemplo [Prade, 85], [Hoffman, 94]) de las que podemos destacar por ejemplo:

$$\mathcal{P}r(P) = \sum P(u) \cdot pr(u)$$

donde $P(u)$ es el grado de pertenencia de u al conjunto borroso P (los conjuntos borrosos se definen en el capítulo 3).

2.6.2.2 Aplicación de la probabilidad a la inferencia: el enfoque Bayesiano

El enfoque Bayesiano se basa en utilizar el teorema de Bayes para actualizar la probabilidad de las posibles explicaciones a un conjunto de síntomas, cuando alguno de los síntomas ha sido observado. El objetivo es calcular las nuevas probabilidades de las posibles explicaciones a partir de sus probabilidades a priori (antes de conocer la observación), y de la probabilidad de la observación. El teorema de Bayes puede formularse como sigue:

$$\mathcal{P}r(Q_i / P) = \frac{\mathcal{P}r(P / Q_i) \cdot \mathcal{P}r(Q_i)}{\sum_{j=1}^n (\mathcal{P}r(P / Q_j) \mathcal{P}r(Q_j))}$$

donde $\mathcal{P}r$ es una medida de probabilidad, P puede interpretarse como la observación, y $\{Q_1, Q_2, \dots, Q_n\}$ es el conjunto de posibles explicaciones para la observación P (que en el caso del diagnóstico podrían ser las posibles causas que expliquen el síntoma P). Para que la aplicación del teorema de Bayes sea matemáticamente correcta, deben cumplirse las siguientes hipótesis:

- Las explicaciones Q_i del conjunto $\{Q_1, Q_2, \dots, Q_n\}$ son todas incompatibles.
- El conjunto $\{Q_1, Q_2, \dots, Q_n\}$ contiene a todas las posibles explicaciones para la observación P .

El teorema de Bayes permite obtener la probabilidad de cada posible explicación Q_i cuando la observación es P . Sin embargo sus hipótesis de validez, junto con la necesidad de conocer todas las probabilidades $\mathcal{P}r(Q_i)$ y $\mathcal{P}r(P/Q_i)$ hace que su aplicación sea en general difícil, y que se requieran grandes bases de datos para poder estimar las mencionadas probabilidades.

Es muy frecuente particularizar el teorema de Bayes al conjunto de posibles observaciones incompatibles dado por $\{Q, \neg Q\}$ de forma que se obtiene la expresión simplificada:

$$\mathcal{P}r(Q / P) = \frac{\mathcal{P}r(P / Q) \cdot \mathcal{P}r(Q)}{\mathcal{P}r(P / Q) \mathcal{P}r(Q) + \mathcal{P}r(P / \neg Q) \mathcal{P}r(\neg Q)}$$

que proporciona la probabilidad de Q cuando se observa P . La incertidumbre asociada con la observación se puede modelar suponiendo que se observa P' , distinto de P y de $\neg P$, pero del que se conoce su relación con P mediante $\mathcal{P}r(P/P')$ (y $\mathcal{P}r(\neg P/P')=1-\mathcal{P}r(P/P')$). Aplicando entonces el teorema de la probabilidad total se tiene:

$$\mathcal{P}r(Q/P') = \mathcal{P}r(Q/P \text{ y } P') \cdot \mathcal{P}r(P/P') + \mathcal{P}r(Q/\neg P \text{ y } P') \cdot \mathcal{P}r(\neg P/P')$$

Si además se supone que cuando P o $\neg P$ son ciertos entonces P' no aporta ninguna información sobre Q , se tiene:

$$\begin{aligned} \mathcal{P}r(Q/P \text{ y } P') &= \mathcal{P}r(Q/P) \\ \mathcal{P}r(Q/\neg P \text{ y } P') &= \mathcal{P}r(Q/\neg P) \end{aligned}$$

y puede escribirse:

$$\mathcal{P}r(Q/P') = \mathcal{P}r(Q/P) \cdot \mathcal{P}r(P/P') + \mathcal{P}r(Q/\neg P) \cdot \mathcal{P}r(\neg P/P')$$

con lo que se puede obtener la probabilidad de Q cuando P' ha sido observado.

Una forma de aplicar el teorema de Bayes a reglas de producción, (tal y como se hace en el sistema experto PROSPECTOR, [Duda, Gaschnig, Hart, 81]), consiste en utilizar las llamadas "odds" en lugar de las probabilidades, asignando a cada regla una pareja de coeficientes. De este modo a cada regla del sistema:

$$R: P \rightarrow Q$$

se le asignan los coeficientes (LS, LN) que viene dados por ([Dillon, Lee, 93]):

$$LS = \frac{\mathcal{P}r(P/Q)}{\mathcal{P}r(P/\neg Q)} \quad LN = \frac{\mathcal{P}r(\neg P/Q)}{\mathcal{P}r(\neg P/\neg Q)}$$

Un valor alto de LS se corresponde con una probabilidad $\mathcal{P}r(P/Q)$ alta y una probabilidad $\mathcal{P}r(P/\neg Q)$ baja, lo que significa que la explicación Q explica mejor a P que $\neg Q$. Vista la influencia de LS en el cálculo $\mathcal{P}r(Q/P)$, LS se interpreta como una medida de suficiencia de la regla, es decir, cómo de suficiente es que la probabilidad de P sea alta para que también lo sea la de Q , ya que valores grandes de LS garantizan valores grandes de $\mathcal{P}r(Q/P)$.

Por otro lado un valor bajo de LN se corresponde con una probabilidad $\mathcal{P}r(\neg P/Q)$ alta y una probabilidad $\mathcal{P}r(\neg P/\neg Q)$ baja, lo que significa que la explicación Q explica mejor a la

observación $\neg P$ que lo que lo hace $\neg Q$. A partir de su influencia en el cálculo de $Pr(Q/\neg P)$ un valor de LN alto hace que también sea alto el valor de $Pr(\neg Q/\neg P)$ con lo que la no observación de P hace más probable a la explicación $\neg Q$ y por tanto se interpreta como una medida de necesidad de la regla, es decir, cómo de necesario es que la probabilidad de P sea alta para que también lo sea la de Q .

Si LS y LN son ambos grandes entonces Q explica bien a P y no explica a $\neg P$, y el grado de implicación $P \rightarrow Q$ es grande. Si LS y LN son ambos pequeños entonces Q explica bien a $\neg P$ y no explica a P , por lo que el grado de implicación de $\neg P \rightarrow Q$ es grande. Si LS es grande y LN es pequeño, entonces Q explica bien a P pero también explica bien a $\neg P$, por lo que los dos grados de implicación $P \rightarrow Q$ y $\neg P \rightarrow Q$ son pequeños. Por último si LS es pequeño y LN es grande, entonces Q no explica bien ni a P ni a $\neg P$, por lo que los dos grados de implicación $P \rightarrow Q$ y $\neg P \rightarrow Q$ son también pequeños.

Estos dos coeficientes son suministrados por el experto que está enunciando las reglas de conocimiento, pero no son independientes el uno del otro, por lo que deben determinarse conjuntamente. La relación que guardan viene dada por:

$$LN = \frac{1 - LS \cdot Pr(P / \neg Q)}{1 - Pr(P / \neg Q)}$$

La aplicación del teorema de Bayes, siendo Q y $\neg Q$ todas las posibles explicaciones incompatibles para la observación P , permite obtener las llamadas odds:

$$O(Q / P) = \frac{Pr(Q / P)}{Pr(\neg Q / P)} = LS \cdot O(Q) = LS \frac{Pr(Q)}{Pr(\neg Q)}$$

$$O(Q / \neg P) = \frac{Pr(Q / \neg P)}{Pr(\neg Q / \neg P)} = LN \cdot O(Q) = LN \frac{Pr(Q)}{Pr(\neg Q)}$$

Cuando la medida $O(Q)$ vale 1 no se tiene evidencia ni en contra ni a favor de Q . Cuando la medida es menor que 1 se tiene evidencia a favor de $\neg Q$ mientras que cuando es mayor que 1 se tiene evidencia a favor de Q . Las expresiones anteriores permiten obtener cómo se modifica $O(Q)$ cuando se observa o bien P o bien $\neg P$. Dado que $Pr(Q/P) = 1 - Pr(\neg Q/P)$, $Pr(Q/\neg P) = 1 - Pr(\neg Q/\neg P)$ y $Pr(Q) = 1 - Pr(\neg Q)$, conocidos LS , LN y $Pr(Q)$ se puede obtener la nueva probabilidad de Q cuando se observa P o $\neg P$.

La probabilidad de $\neg P$ dada la de P se obtiene, como es bien sabido, como:

$$Pr(\neg P) = 1 - Pr(P)$$

Cuando se quiere evaluar la probabilidad de una observación con varias premisas, por ejemplo ' P_1 y P_2 ' o bien ' P_1 o P_2 ' suele utilizarse las siguientes expresiones (*Valverde, García, Godó, 87*), [*Hopgood, 93*):

$$\begin{aligned} \mathcal{P}r(P_1 \cap P_2) &= \min(\mathcal{P}r(P_1), \mathcal{P}r(P_2)) \\ \mathcal{P}r(P_1 \cup P_2) &= \max(\mathcal{P}r(P_1), \mathcal{P}r(P_2)) \end{aligned}$$

aunque desde la teoría de la probabilidad no es correcto ya que dados dos conjuntos P_1 y P_2 solo puede escribirse:

$$\mathcal{P}r(P_1 \cap P_2) = \mathcal{P}r(P_1) + \mathcal{P}r(P_2) - \mathcal{P}r(P_1 \cup P_2)$$

En el caso de que sean incompatibles se tiene:

$$\begin{aligned} \mathcal{P}r(P_1 \cap P_2) &= 0 \\ \mathcal{P}r(P_1 \cup P_2) &= \mathcal{P}r(P_1) + \mathcal{P}r(P_2) \end{aligned}$$

mientras que si son compatibles e independientes:

$$\begin{aligned} \mathcal{P}r(P_1 \cap P_2) &= \mathcal{P}r(P_1) \cdot \mathcal{P}r(P_2) \leq \min(\mathcal{P}r(P_1), \mathcal{P}r(P_2)) \\ \mathcal{P}r(P_1 \cup P_2) &= \mathcal{P}r(P_1) + \mathcal{P}r(P_2) - \mathcal{P}r(P_1) \cdot \mathcal{P}r(P_2) \geq \max(\mathcal{P}r(P_1), \mathcal{P}r(P_2)) \end{aligned}$$

En el caso en que coexistan varias reglas del tipo:

$$\begin{aligned} R_1: P_1 \rightarrow Q \\ R_l: P_l \rightarrow Q \\ \dots \\ R_m: P_m \rightarrow Q \end{aligned}$$

a cada regla R_i se le asigna una pareja de coeficientes (LS_i, LP_i) y puede establecerse un procedimiento de cálculo para actualizar las probabilidades de la conclusión en base a las observaciones disponibles.

Si todos las posibles observaciones P_i son independientes, entonces se tiene directamente:

$$O(Q / P_1, P_2, \dots, P_m) = \left(\prod_{i=1}^m LS_i \right) O(Q)$$

$$O(Q / \neg P_1, \neg P_2, \dots, \neg P_m) = \left(\prod_{i=1}^m LN_i \right) O(Q)$$

Si se modela la incertidumbre de las observaciones suponiendo que lo que realmente se observa es P'_i , entonces puede calcularse:

$$LS'_i = \frac{O(Q / P'_i)}{O(Q)}$$

y suponiendo que todas las observaciones P'_i son independientes se tiene:

$$O(Q / P'_1, P'_2, \dots, P'_m) = \left(\prod_{i=1}^m LS'_i \right) O(Q)$$

2.6.3 Los factores de certeza

Los factores de certeza utilizados por primera vez en el sistema experto MYCIN ([Shortliffe, 76], [Buchanan, Shortliffe, 84]) se construyen a partir de consideraciones heurísticas, pero basándose en el uso de probabilidades y el teorema de Bayes, y presentan también ciertas similitudes con la teoría de la evidencia. Una de las críticas que a menudo se les ha hecho es su frágil fundamento teórico.

A una regla:

$$R: P \rightarrow Q$$

se le asocia la probabilidad $Pr(P/Q)$. A partir de esta probabilidad se calculan dos coeficientes, llamados grado de credibilidad (“measure of belief” o MB) y grado de no credibilidad (“measure of disbelief” o MD) de Q dado P :

$$MB(Q / P) = \begin{cases} \frac{\max(Pr(Q / P), Pr(Q)) - Pr(Q)}{1 - Pr(Q)} & \text{si } Pr(Q) \neq 1 \\ 1 & \text{si } Pr(Q) = 1 \end{cases}$$

$$MD(Q / P) = \begin{cases} \frac{Pr(Q) - \min(Pr(Q / P), Pr(Q))}{Pr(Q)} & \text{si } Pr(Q) \neq 0 \\ 1 & \text{si } Pr(Q) = 0 \end{cases}$$

Estas dos medidas verifican:

$$\text{si } MB(Q/P) > 0 \text{ entonces } MD(Q/P) = 0$$

$$\text{si } MD(Q/P) > 0 \text{ entonces } MB(Q/P) = 0$$

y ambas pertenecen al intervalo $[0, 1]$. Se define entonces el factor de certeza de Q dado P como:

$$CF(Q/P) = MB(Q/P) - MD(Q/P)$$

que pertenece al intervalo $[-1, 1]$. Puede comprobarse que cuanto más próximo es CF de 1 mayor es la evidencia a favor de Q y cuanto más próximo es de -1 mayor es la evidencia en contra de Q . Cuando CF está próximo a 0 no existe evidencia ni a favor ni en contra de Q . Más concretamente pueden establecerse las siguientes relaciones:

- $CF(Q/P) = 1$ si $Pr(Q) = 1$
- $CF(Q/P) = MB(Q/P)$ si $Pr(Q/P) > Pr(Q)$
- $CF(Q/P) = 0$ si $Pr(Q/P) = Pr(Q)$
- $CF(Q/P) = MD(Q/P)$ si $Pr(Q/P) < Pr(Q)$
- $CF(Q/P) = -1$ si $Pr(Q) = 0$

En general no se parte de $Pr(P/Q)$ para obtener el factor de certeza de una regla sino que suele ser suministrado directamente por el experto que enuncia las reglas de conocimiento.

El factor de certeza de la conclusión Q obtenido a partir de la regla R se obtiene como:

$$CF(Q) = CF(P) \cdot CF(Q/P)$$

El factor de certeza de observaciones con varias premisas, por ejemplo ' P_1 y P_2 ' o bien ' P_1 o P_2 ' se obtiene mediante las siguientes expresiones:

$$CF(P_1 \cap P_2) = \min(CF(P_1), CF(P_2))$$

$$CF(P_1 \cup P_2) = \max(CF(P_1), CF(P_2))$$

y el factor de certeza de $\neg P$ dado el de P se obtiene como:

$$CF(\neg P) = 1 - CF(P)$$

Cuando mediante dos reglas se obtienen dos factores de certeza para una misma conclusión Q , entonces, denotando por $CF_i(Q)$ el factor de certeza de Q concluido por cada regla R_i , el factor de certeza resultante de la combinación de ambos viene dado por:

$$\begin{aligned}
 CF(Q) &= CF_1(Q) + CF_2(Q) - CF_1(Q) \cdot CF_2(Q) && \text{si } CF_1(Q) > 0 \text{ y } CF_2(Q) > 0 \\
 CF(Q) &= CF_1(Q) + CF_2(Q) + CF_1(Q) \cdot CF_2(Q) && \text{si } CF_1(Q) < 0 \text{ y } CF_2(Q) < 0 \\
 CF(Q) &= \frac{CF_1(Q) + CF_2(Q)}{1 - \min(|CF_1(Q)|, |CF_2(Q)|)} && \text{en otro caso} \\
 CF(Q) & \text{ indefinido} && \text{si } CF_1(Q) \cdot CF_2(Q) = 1
 \end{aligned}$$

que es una combinación asociativa, lo que permite extender su uso a más de dos reglas.

2.6.4 Las medidas de credibilidad y plausibilidad

2.6.4.1 Definiciones

Las medidas de credibilidad y plausibilidad se definen en la teoría de la evidencia desarrollada principalmente por los profesores Dempster y Shaffer ([Dempster, 67], [Shaffer, 76]), y se basan en asignar a cada proposición P dos coeficientes en $[0,1]$ que miden respectivamente cómo de plausible y cómo de creíble es una afirmación. Estos dos coeficientes se denominan plausibilidad (“plausibility”) y credibilidad (“belief”) de P , y se representan por $Pl(P)$ y $Bel(P)$ respectivamente.

Los dos coeficientes $Pl(P)$ y $Bel(P)$ son medidas de incertidumbre que además verifican:

$$\begin{aligned}
 Bel(P_1 \cup P_2) &\geq Bel(P_1) + Bel(P_2) - Bel(P_1 \cap P_2) \\
 Pl(P_1 \cap P_2) &\geq Pl(P_1) + Pl(P_2) - Pl(P_1 \cup P_2)
 \end{aligned}$$

A cada medida de credibilidad $Bel(P)$ se le puede asociar una medida de plausibilidad $Pl(P)$ definida como:

$$Pl(P) = 1 - Bel(\neg P)$$

y se dice que son duales.

Las medidas de plausibilidad y credibilidad pueden definirse a partir de una función de conjunto $m(P)$ que se denomina asignación básica de probabilidad y que verifica:

$$m: \mathcal{P}(U) \rightarrow [0,1] \quad \text{tal que} \quad m(\emptyset) = 0 \quad \text{y} \quad \sum_{P \in 2^U} m(P) = 1$$

La función de asignación básica de probabilidad es una asignación de probabilidad subjetiva a conjuntos de $\mathcal{P}(U)$, y por tanto no es ni una función de distribución de probabilidad en U , ni una medida de incertidumbre. Por ejemplo:

- $m(P)$ es una función de conjunto y no de los elementos de U
- no se verifica $m(U)=1$
- no se verifica $m(P_1) \leq m(P_2)$ si $P_1 \subset P_2$
- no hay relación entre $m(P)$ y $m(\neg P)$

Sin embargo la función de asignación de probabilidad sí es una distribución de probabilidad en $\mathcal{P}(U)$. Esto permite fundamentar la teoría de la evidencia desde la teoría de la probabilidad, considerando conjuntos en lugar de elementos, lo que le confiere un sólido fundamento teórico.

Las medidas de credibilidad y plausibilidad de un conjunto P se obtienen a partir de m mediante las siguientes expresiones:

$$Bel(P) = \sum_{P_i \subseteq P} m(P_i)$$

$$Pl(P) = \sum_{P_i \cap P \neq \emptyset} m(P_i)$$

La medida $m(P_i)$ se interpreta como la creencia que se tiene únicamente en el conjunto P_i , y es su probabilidad en relación con todos los demás conjuntos de $\mathcal{P}(U)$, $Bel(P)$ se interpreta como la creencia total que se tiene en P y en cualquier subconjunto de P , y por último $Pl(P)$ se interpreta como la creencia en P , en cualquier subconjunto de P y en cualquier otro conjunto que tenga una intersección no nula con P . Esto hace que siempre se verifique:

$$Pl(P) \geq Bel(P)$$

Dada una asignación básica de probabilidad m , se dice que un conjunto P es un elemento focal si $m(P) > 0$. Los elementos focales son todos aquellos conjuntos en los que se tiene una creencia no nula. El par compuesto por una asignación m y el conjunto $\{P / m(P) > 0\}$ se denomina cuerpo de evidencia (“body of evidence”).

A partir de una medida de credibilidad se puede recuperar la asignación básica de probabilidad mediante la siguiente expresión:

$$\forall P_i \in \mathcal{P}(U) \quad m(P_i) = \sum_{P \subseteq P_i} (-1)^{|P_i - P|} Bel(P)$$

donde $|P_i - P|$ denota al cardinal del conjunto $P_i - P$. Teniendo en cuenta la relación $Pl(P) = 1 - Bel(\neg P)$ también se puede recuperar la asignación básica de probabilidad a partir de una medida de plausibilidad.

En la teoría de la evidencia la ignorancia o total desconocimiento sobre la verdad de las posibles afirmaciones que puedan hacerse se expresa mediante una asignación básica de probabilidad dada por $m(U) = 1$ y $m(P) = 0$ para cualquier conjunto P distinto de U . La plausibilidad y credibilidad para cualquier conjunto P vienen entonces dadas por:

$$\begin{aligned} Bel(P) &= 0 \\ Pl(P) &= 1 \end{aligned}$$

lo que parece acorde con la intuición, cuando se modelan por ejemplo afirmaciones del tipo “*es posible que llueva pero no estoy seguro*”, afirmaciones que no pueden modelarse bajo un marco estrictamente probabilístico. La cantidad de ignorancia o desconocimiento sobre una proposición P puede evaluarse como $Pl(B) - Bel(P)$ ([Hopgood, 93]) o bien como $1 - 2Bel(P)$ ([Dubois, Prade, 88b]).

La teoría de la evidencia resulta especialmente atractiva en el sentido en que contiene como casos particulares la teoría de la probabilidad y la teoría de la posibilidad. En [Prade, 85] o [Klir, Folger, 88] se describe la teoría de la evidencia y su estrecha relación con las medidas de probabilidad y posibilidad.

- Teoría de la evidencia y probabilidad

Cuando los elementos focales son conjuntos de un único elemento, es decir $m(\{u\}) = Bel(\{u\})$, y $m(P) = 0$ si P no es un conjunto unitario, entonces:

$$Bel(\{u\}) = Pl(\{u\}) = Pr(\{u\})$$

donde Pr es una medida de probabilidad. Además se tiene, para cualquier conjunto P en U :

$$Pr(P) = Bel(P) = Pl(P) = \sum_{u \in U} m(\{u\})$$

En este caso $m(\{u\})$ es equivalente a una distribución de probabilidad y se suele escribir:

$$m(\{u\}) = pr(u)$$

En general cuando no se verifica $Bel(\{u\}) = Pl(\{u\}) = Pr(\{u\})$ las medidas de plausibilidad y credibilidad de un conjunto P pueden interpretarse como cotas superior e inferior respectivamente de la probabilidad de P .

- Teoría de la evidencia y teoría de la posibilidad

Cuando por el contrario las medidas de plausibilidad y credibilidad son tales que verifican:

$$\begin{aligned} Bel(P_1 \cap P_2) &= \min(Bel(P_1), Bel(P_2)) \\ Pl(P_1 \cup P_2) &= \max(Pl(P_1), Pl(P_2)) \end{aligned}$$

entonces son respectivamente medidas de necesidad y posibilidad, y se denotan por:

$$\begin{aligned} N(P) &= Bel(P) \\ \Pi(P) &= Pl(P) \end{aligned}$$

Para que las medidas de credibilidad y plausibilidad degeneren en medidas de posibilidad y necesidad los elementos focales $\{ P_i \}$ deben estar anidados, esto es:

$$P_1 \subset P_2 \subset \dots \subset P_n$$

Las medidas de posibilidad pueden definirse a partir de una función de distribución de posibilidad para la variable u que denotaremos por $\pi(u)$:

$$\Pi(P) = \sup_{u \in P} \pi(u)$$

La medida de necesidad de P viene entonces dada por:

$$N(P) = \inf_{u \in -P} (1 - \pi(u))$$

Una distribución de posibilidad juega en la teoría de la posibilidad un papel similar al que juega una distribución de probabilidad en la teoría de la probabilidad, pero son conceptos diferentes, con propiedades diferentes. Por ejemplo una distribución de posibilidad en general verifica:

$$\sum_{u \in P} \pi(u) \neq 1$$

Dado un conjunto borroso P y siendo P^α sus α -cortes, es decir $P^\alpha = \{ u / P(u) \geq \alpha \}$ el conjunto de α -cortes, entonces $\{ P^\alpha / \alpha \in]0, 1] \}$ es el conjunto de elementos focales de la medida de posibilidad $\Pi(P)$. Además la asignación básica $m(P^\alpha)$ puede obtenerse de forma sencilla cuando $P(u)$ toma valores discretos. Si los grados de pertenencia de P para cualquier u se ordenan de forma decreciente, $\alpha_1 = 1 > \alpha_2 > \dots > \alpha_r$ y $\alpha_{r+1} = 0$, entonces se tiene ([Dubois, Prade, 89], [Klir, Folger, 88]):

$$m(P^\alpha) = \alpha_i - \alpha_{i+1} \quad i=1, r$$

$$m(P)=0 \text{ si } P \neq P^{\alpha i}$$

2.6.4.2 Aplicación de la teoría de la evidencia a la inferencia

Al igual que con otros mecanismos de representación de incertidumbre la inferencia requiere una regla de actualización de incertidumbres cuando nueva información está disponible, así como una regla de combinación de incertidumbres procedentes de distintas fuentes (o reglas).

Cuando se dispone de una observación sobre una variable dada por un conjunto P definido en U , se puede aplicar la regla condicional de Dempster, de forma análoga a como se hace en probabilidad con el teorema de Bayes, para modificar la creencia en una hipótesis Q . Interpretando la medida de credibilidad como una cota superior de probabilidad, y la de plausibilidad como una cota inferior, y partiendo de una asignación básica $m(P)=1$, que expresa que se ha observado P , la regla de condicional de Dempster da lugar a las expresiones condicionales ([Walley, 96]):

$$\begin{aligned} Bel(Q/P) &= Bel(Q \cap P) / Bel(P) \text{ definida para } Bel(P) > 0 \\ Pl(Q/P) &= 1 - Bel(-Q/P) \end{aligned}$$

que permiten actualizar la credibilidad de Q a partir de la credibilidad de P y la credibilidad de $P \cap Q$.

Cuando se dispone de una nueva observación P la creencia en cada conjunto Q debe revisarse mediante la regla condicional de Dempster:

$$Bel(Q / P) = \frac{Bel(Q \cup -P) - Bel(-P)}{Bel(U) - Bel(-P)}$$

Esto permite también formular reglas de inferencia paralelas a la regla de inferencia Bayesiana pero utilizando medidas de credibilidad y plausibilidad.

La regla de combinación de Dempster permite combinar asignaciones básicas de probabilidad provenientes de fuentes independientes. Dadas dos asignaciones básicas m_1 y m_2 independientes, la asignación básica resultante de la combinación de ambas viene dada por:

$$\left\{ \begin{array}{l} m(P) = \frac{\sum_{P_1 \cap P_2 = P} m_1(P_1).m_2(P_2)}{1 - \sum_{P_1 \cap P_2 = \emptyset} m_1(P_1).m_2(P_2)} \quad \text{para } P \neq \emptyset \\ m(\emptyset) = 0 \end{array} \right.$$

Esta regla ha sido aplicada parcialmente en otros sistemas de representación y razonamiento con incertidumbre no basados en la teoría de la evidencia, en concreto para la combinación de los coeficientes de certeza en el sistema MYCIN cuando ambos tienen mismo signo.

A partir de dos asignaciones básicas la regla de combinación permite también obtener directamente la credibilidad en cada conjunto, que se deduce de las dos asignaciones mediante la expresión:

$$\left\{ \begin{array}{l} Bel(P) = \frac{\sum_{P_1 \cap P_2 \subseteq P} m_1(P_1).m_2(P_2)}{1 - \sum_{P_1 \cap P_2 = \emptyset} m_1(P_1).m_2(P_2)} \quad \text{para } P \neq \emptyset \\ Bel(\emptyset) = 0 \end{array} \right.$$

Así una forma de aplicar la teoría de la evidencia a un sistema basado en reglas consiste en suponer que cada regla $P \rightarrow Q$ se dispara cuando se observa P y asigna a su conclusión una cierta credibilidad $Bel(Q)$, que mide la creencia en Q cuando la hipótesis P es verdadera. A partir de la medida $Bel(Q)$ se puede entonces recuperar la asignación básica de probabilidad que la genera.

Suponiendo entonces que se dispone de dos reglas diferentes que se disparan a un tiempo y que ambas concluyen una determinada creencia para un conjunto, a partir de cada una de ellas puede obtenerse una asignación básica de probabilidad. Estas asignaciones de probabilidad, supuestas independientes, pueden agregarse mediante la regla de combinación de Dempster, con lo que finalmente se puede calcular la creencia final en un determinado conjunto Q .

2.6.5 Las medidas de posibilidad y necesidad, la lógica borrosa y las distribuciones de posibilidad

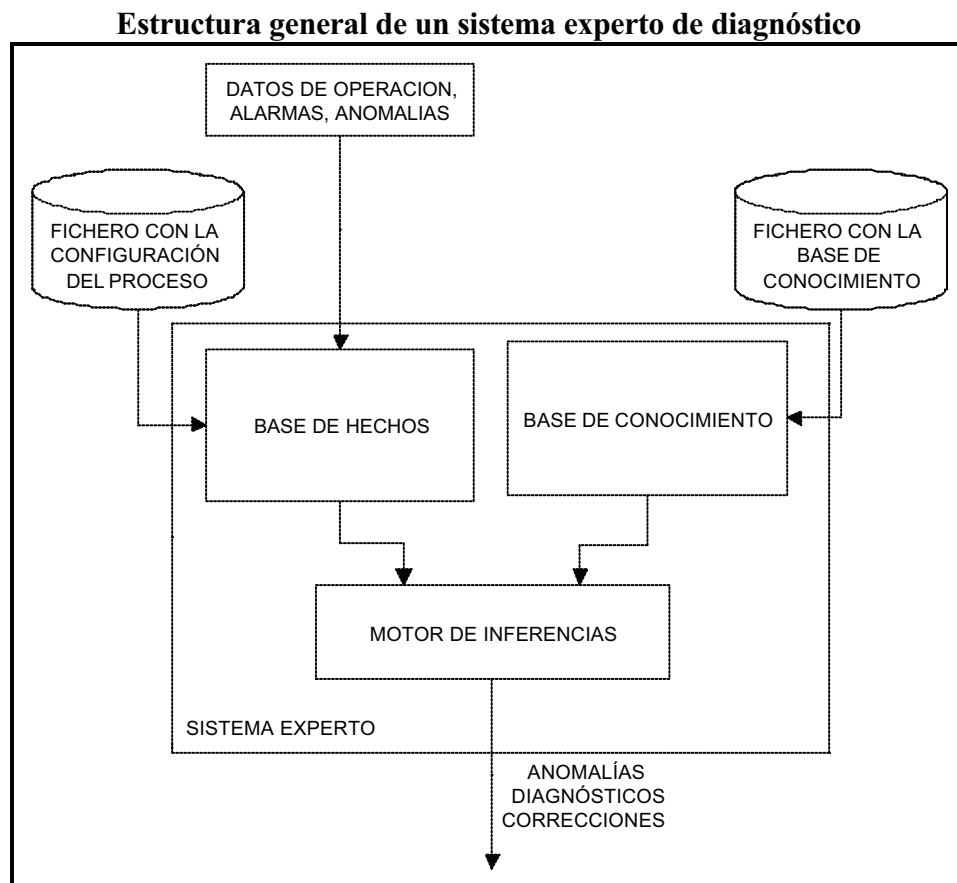
La representación del conocimiento mediante la lógica borrosa y las distribuciones de posibilidad, así como su aplicación a la inferencia en sistemas de diagnóstico es el tema central de esta tesis, por lo que se irá tratando a lo largo de los capítulos siguientes. Los fundamentos teóricos así como algunas de las aplicaciones más importantes al diagnóstico se abordan en el capítulo 3, la inferencia con una regla en el capítulo 4 y los sistemas de reglas se estudian en el capítulo 5.

2.7 Sistemas expertos de diagnóstico

- Estructura de un sistema experto:

Un sistema experto ([Sanz , 92]) es un sistema informático que trata de reproducir el razonamiento humano ante un problema específico en campo técnico muy especializado. Los sistemas expertos han sido aplicados con éxito por ejemplo en el ámbito de la toma de decisiones o en el ámbito del diagnóstico.

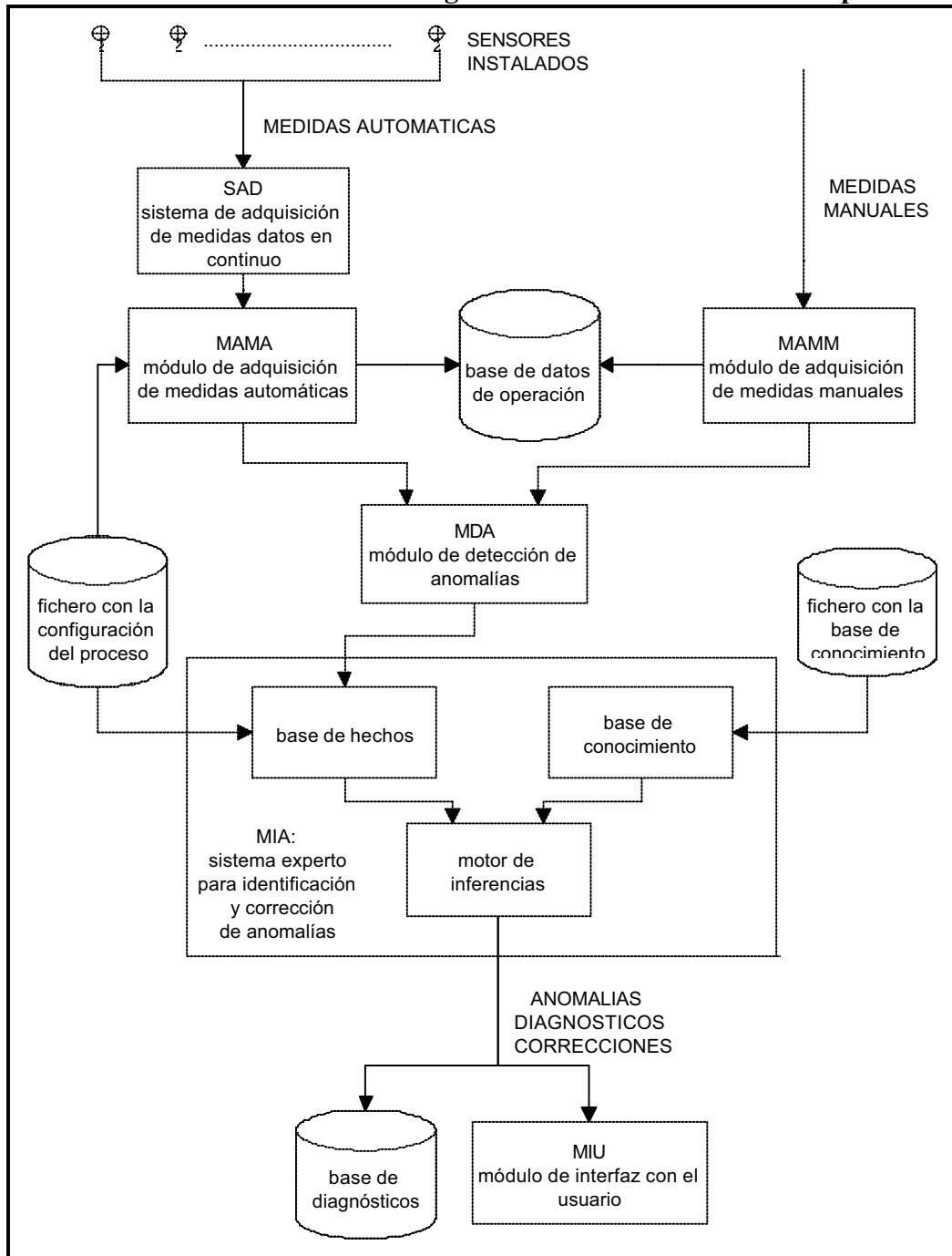
Un sistema experto puede descomponerse en tres componentes principales, que son la base de conocimiento, la base de hechos y el motor de inferencias. La siguiente figura muestra gráficamente esta estructura:



La estructura completa típica de un sistema de diagnóstico queda representada en la siguiente figura, donde aparece de forma explícita el sistema experto como módulo de

identificación (y corrección de anomalías), integrado en el sistema de diagnóstico. Algunos de los módulos ya fueron descritos en la introducción a los sistemas de diagnóstico, en este mismo capítulo, mientras que las distintas bases de datos y ficheros del sistema están algo más detallados y se describen a continuación:

Estructura de un sistema de diagnóstico basado en un sistema experto



- La base de conocimiento:

La base de conocimiento contiene aquel conocimiento que el sistema experto tiene almacenado para llevar a cabo su tarea, conocimiento relativo a la estructura del proceso y a su funcionamiento, y que se corresponde con el conocimiento funcional. Este conocimiento puede obtenerse de los expertos humanos relacionados con el proceso, a partir de su propia experiencia acerca de su funcionamiento. Puede también obtenerse a partir del análisis del comportamiento del proceso en base a sus propias especificaciones, criterios de diseño, tecnologías empleadas, fundamentos físicos sobre su comportamiento, etc. Una tercera fuente viene de la experiencia sobre la operación del proceso de los operadores directamente relacionados con el mismo, o a través de datos de operación de funcionamiento normal o bajo distintas condiciones de fallo, previamente almacenadas.

En general en la base de conocimiento se encuentra la descripción de las posibles anomalías que el proceso puede presentar, los síntomas a los que dan lugar y los diagnósticos que explican la aparición de las citadas anomalías.

La forma más común de enunciar este conocimiento es mediante el uso de reglas de producción. La procedencia del conocimiento da lugar a reglas con distinta naturaleza tal y como se señala en [Sanz, 92], de las que destacamos algunas de las más importantes:

- reglas heurísticas obtenidas a partir de entrevistas con expertos humanos.
- reglas obtenidas del estudio de mecanismos de fallo.
- reglas heurísticas obtenidas a partir de criterios estadísticos.
- reglas heurísticas obtenidas a partir de la experiencia de operación en el propio proceso.

Además del conocimiento mencionado, la base de conocimiento puede también contener otro tipo de conocimiento, el metaconocimiento, que indica al motor de inferencias estrategias a seguir durante el proceso de razonamiento, como por ejemplo el orden de disparo de las reglas. Así, ante un determinado síntoma, puede ser conveniente comprobar primero las reglas cuyas conclusiones se corresponden con fallos especialmente peligrosos, de cara a detectarlos o descartarlos cuanto antes. Otras veces interesa descartar ciertos grupos de reglas ante la aparición de un determinado síntoma que indique que los diagnósticos que proponen dichas reglas no van a tener relación con el fallo real, evitando así comprobaciones innecesarias ([Sanz, Villar, Palacios, 95]). Todo este tipo de conocimiento puede enunciarse también en forma de reglas, y las reglas utilizadas para enunciarlo se denominan metarreglas. Un ejemplo de metarregla es el siguiente:

*si motor caliente
activar grupo de reglas térmicas*

La regla anterior dice al sistema que si el motor está caliente, entonces debe activarse el grupo de reglas *térmicas* que lógicamente serán reglas dedicadas a analizar problemas relacionados con calentamientos excesivos en el motor. En este caso un grupo de reglas permanece desactivado, en el sentido en que no son tenidas en cuenta, hasta que se verifique el hecho que las activa.

El conocimiento de la base de conocimiento suele estar contenido en un fichero para facilitar su modificación, fichero que a menudo se denomina también base de conocimiento.

- La base de hechos:

La base de hechos contiene la descripción del entorno físico donde debe trabajar el sistema experto, relativa por ejemplo a la estructura concreta (y no al conocimiento que de una estructura determinada pueda derivarse, ya que este último se englobaría dentro de la base de conocimiento) del proceso en cuestión, a sus componentes e interrelaciones entre componentes, a sus datos de diseño, a las medidas que están siendo recogidas, al tipo de sensores empleados, al tipo y configuración del sistema de adquisición de datos, etc. Todos estos datos suelen residir en un fichero de configuración (llamado también base de hechos) que el sistema lee al arrancar. Además de todos estos datos descriptivos de las características estáticas del proceso, la base de hechos también contiene información dinámica correspondiente al conocimiento de operación del proceso. La base de hechos engloba por tanto el conocimiento estructural y el conocimiento de operación del proceso bajo vigilancia.

El conocimiento sobre la estructura del proceso, sus componentes y medidas, suele estar contenido en un fichero de configuración para facilitar su modificación. Por extensión a este fichero se le denomina a menudo base de hechos.

- El motor de inferencias:

El motor de inferencias es el encargado de conjuntar la información contenida en la base de conocimiento con la contenida en la base de hechos para concluir nuevos hechos, proponer diagnósticos y correcciones, o llevar a cabo acciones, etc. Dependiendo de cómo esté representado el conocimiento las estrategias a seguir pueden ser muy distintas por lo que el motor de inferencias puede variar substancialmente de unos sistemas a otros.

- La base de datos de operación:

Aunque no es un componente esencial del sistema de diagnóstico, suele ser conveniente idear algún tipo de mecanismo capaz de almacenar el estado del proceso en distintos puntos de operación, de cara a un posterior refinamiento del conocimiento de

funcionamiento normal ([Muñoz, 96]), o poder comprobar el funcionamiento del sistema de forma off-line, es decir lejos del proceso real, pero con datos reales de operación previamente almacenados.

- La base de diagnósticos:

Cuando el sistema de diagnóstico detecta alguna anomalía es conveniente almacenar, por un lado todos aquellos datos de operación que describen el estado del proceso durante la anomalía, y por otro los diagnósticos emitidos por el sistema experto, todo ello de cara a un posterior refinamiento del conocimiento funcional y en concreto del funcionamiento anómalo del proceso. Los ficheros utilizados para almacenar esta información componen la base de diagnósticos del sistema.

Los sistemas expertos de diagnóstico son un núcleo fundamental de muchos sistemas de diagnóstico, y en ellos nos apoyaremos para proponer una estructura general de sistema de diagnóstico, aspecto al que está dedicado el capítulo 7 de esta tesis.

3. Teoría de conjuntos borrosos y distribuciones de posibilidad

Este capítulo presenta una descripción de algunos de los aspectos de la teoría de conjuntos borrosos más importantes para la comprensión de los siguientes capítulos de esta tesis. Se pretende dar una visión enfocada a su aplicación a la construcción de sistemas basados en conocimiento, describiendo las herramientas que serán utilizadas a lo largo de la metodología propuesta en esta tesis. La exposición no se ha limitado a una revisión de los principales fundamentos teóricos sino que ha pretendido dar una visión personal de los mismos, introduciendo por primera vez nuevos conceptos que serán clave en el desarrollo de esta tesis, como son las distribuciones de posible posibilidad y sus reglas de inferencia asociadas.

3.1 Organización del capítulo

En primer lugar y a modo de introducción se propone un breve recorrido histórico desde los orígenes de la teoría de conjuntos borrosos hasta nuestros días, en el que únicamente se señalan algunas de las contribuciones más importantes de entre la gran multitud de trabajos que han proliferado desde sus inicios.

En segundo lugar se repasan algunos de los principales conceptos y definiciones relacionados con los conjuntos clásicos y borrosos, sobre los que se asienta la teoría que se expone en este capítulo, y a los que se irá haciendo continua referencia a lo largo de la tesis.

A continuación se estudian las principales operaciones entre conjuntos borrosos como por ejemplo las negaciones, las t -normas y t -conormas, las pseudo-conjunciones y pseudo-disyunciones, haciendo hincapié en su representación funcional mediante los llamados generadores aditivos, representación que será frecuentemente utilizada durante el capítulo 4, cuando se analice en profundidad el modus ponens generalizado en lógica borrosa.

En el siguiente apartado se abordan las relaciones borrosas, concretamente aquellas obtenidas a partir de las operaciones lógicas entre conjuntos borrosos, en particular de las conjunciones e implicaciones definidas en el punto anterior, que dan lugar a los condicionales borrosos más empleados.

Se repasan después algunos de los principales aspectos de la teoría de la posibilidad necesarios para el desarrollo de esta tesis, donde se introducen las distribuciones de necesaria y posible posibilidad, aportación original de esta tesis, que permiten modelar en un único marco teórico vaguedad, imprecisión e incertidumbre, tanto en el dominio de lo seguro o necesario, como en el dominio de lo posible.

A continuación se formaliza de forma somera el razonamiento aproximado en lógica borrosa y sus principales reglas de inferencia, tanto para distribuciones de necesaria posibilidad, como su contrapartida para distribuciones de posible posibilidad. Las reglas de inferencia utilizadas con distribuciones de necesaria posibilidad se corresponden con las reglas de inferencia clásicas, pero para las distribuciones de posible posibilidad es necesario recurrir a reglas ya aplicadas pero no formalizadas con anterioridad.

En el último apartado se presenta el estado del arte de la aplicación de la lógica borrosa a los sistemas expertos de diagnóstico, repasando brevemente algunas de las principales técnicas y enfoques empleados.

3.2 Breve recorrido histórico

La teoría de conjuntos borrosos aparece por primera vez con el artículo del profesor Zadeh titulado “*Fuzzy Sets*” ([Zadeh, 65]), publicado en 1965 en la revista “*Information and Control*”. Zadeh conocido internacionalmente en el campo del control clásico y la teoría de sistemas se embarca a partir de entonces en una nueva línea de investigación. Inicialmente muy contestada por muchos investigadores, la teoría propuesta por Zadeh se ha ido poco a poco mostrando como una poderosa herramienta para la resolución de cierto tipo de problemas difícilmente abordables con herramientas clásicas, especialmente en el campo de la inteligencia artificial. Como el propio profesor Zadeh afirma (en su prólogo a [Cox, 94]), el artículo “*Fuzzy Sets*” estuvo motivado por su convicción de que los métodos tradicionales de análisis de sistemas no eran los más apropiados para aquellos sistemas en los que las relaciones entre variables no pueden formalizarse mediante ecuaciones diferenciales o mediante ecuaciones en diferencias. El modelado del razonamiento humano es un claro ejemplo de este tipo de sistemas.

Hacia finales de los 60 y principios de los 70 la teoría de conjuntos borrosos se extiende a Japón, donde comienzan a oírse nombres como los de los profesores Asai, Tanaka y Terano, cuyas primeras contribuciones estaban enfocadas al estudio de autómatas borrosos y sistemas de aprendizaje.

Durante los años 70 hay importantes contribuciones teóricas, pero es a finales de los 70 y durante los 80 cuando las aportaciones teóricas, que se siguen multiplicando, ven su aplicación en la industria.

Por ejemplo en Europa una de las contribuciones iniciales más importantes fue la del profesor Mamdani, investigador británico, que a principios de los 70 consigue la primera aplicación industrial importante, el control basado en lógica borrosa de un horno de cemento. Su aportación al control borroso ([Mamdani, 75] y [Mamdani, 76]) se consolida rápidamente y ha tenido un espectacular auge desde entonces, siendo uno de los campos donde la aplicación de la teoría de conjuntos borrosos ha sido más exitosa.

En 1978 el profesor Zadeh publica “*PRUF, a meaning representation language for natural languages*” ([Zadeh, 78]) donde se propone un lenguaje de representación del conocimiento basado en los conjuntos y relaciones borrosas, y en [Zadeh, 78b] se establecen las bases de la teoría de la posibilidad para el modelado de la vaguedad e incertidumbre mediante el uso de conjuntos borrosos. Esta teoría ha sido extensamente estudiada y desarrollada posteriormente por los profesores Dubois y Prade ([Dubois, Prade, 88], [Dubois, Prade, 90]), que han colaborado significativamente en su aplicación

a la construcción de sistemas expertos ([*Martin-Clouaire, Prade, 84*], [*Buisson, Farreny, Prade, 85*], [*Lebailly, Martin-Clouaire, Prade, 86*]).

En España aparecen importantes contribuciones teóricas sobre el estudio de implicaciones y operadores borrosos, principalmente de la mano de los profesores Trillas y Valverde ([*Trillas, Valverde, 85*], [*Trillas, Valverde, 85b*], [*Valverde, 85*]), y se desarrolla MILORD, un sistema experto borroso aplicado al diagnóstico médico (descrito en [*Mántaras, et al., 91*] y [*Mántaras, Sierra, Agustí, 93*]). También durante estos años nuevas aplicaciones hacen cada vez más sugerente el uso de sistemas borrosos para el control de sistemas complejos orientados a tareas. A este respecto cabe destacar, por ejemplo, el trabajo del profesor Sugeno para el aparcamiento automático de vehículos ([*Sugeno, Murkami, 84*]) y una aplicación ya clásica en el terreno del control, del año 1987, consistente en un sistema de frenado para el metro de Sendai, en Japón, considerada como una de las aplicaciones más exitosas de los sistemas borrosos.

Hacia final de los 80 y ya en los 90 se consolidan las aportaciones teóricas. Cabe destacar algunos trabajos de carácter general como por ejemplo [*Dubois, Prade, 91*] y [*Dubois, Lang, Prade, 91*], referencias casi obligadas, donde se da un interesante repaso a la teoría de conjuntos borrosos aplicada al campo de la lógica y el razonamiento aproximado, con multitud de referencias a otros trabajos más específicos para cada uno de los temas que tratan, o también [*Trillas, 93*] donde el autor presenta una visión matemática moderna de las lógicas clásica y borrosa y el razonamiento aproximado. Aparecen también con fuerza nuevos enfoques para llevar a cabo la inferencia como por ejemplo [*Magrez, Smets, 89*], [*Cross, Sudkamp, 93*], [*Turksen, Zhong, 90*], trabajos orientados a aportar una interpretación semántica en términos de distancias y relaciones de similitud [*Ruspini, 91*], etc.

También se consolidan las contribuciones en el terreno del control ([*Lee, 90*], [*Lee, 90b*], [*Mizumoto, 95*], [*Zheng, 92*], [*Pedrycz, 93*]) y de la aproximación funcional ([*Wang, 92*], [*Wang, 92b*], [*Nakanishi, Turksen, Sugeno, 93*], [*Pedrycz, 96*]), y comienzan a abordarse problemas muy diversos tanto teóricos como industriales mediante técnicas borrosas, por ejemplo en el ámbito de la simulación cualitativa ([*Shen, Leitch, 93*], [*Hajek, Harmancova, 95*]), la fiabilidad ([*Cai, 96*], [*Bowles, Pelaez, 95*], [*Bazu, 95*]), los sistemas eléctricos de potencia ([*Hsu, Ho, 91*], [*Dash et al, 93*], [*Kenarangui, Seifi, 94*]), etc. También se establecen las bases de la estrecha relación existente entre los sistemas borrosos y algunos tipos de redes neuronales ([*Jang, Sun, 93*]) y se consolida el estudio de sistemas que combinan la teoría de conjuntos borrosos y la de las redes neuronales en entornos comunes ([*Buckley, Hayashi, 94*], [*Gupta, Rao, 94*]), un campo muy prometedor y en pleno auge.

3.3 Conjuntos borrosos y definiciones básicas

Dado un universo de discurso U , un predicado p , nombre de una propiedad (por ejemplo en el universo de los números reales “*mayor que 0.5*”, “*par*”, etc.), permite definir en U un conjunto P de elementos de U , tales que verifican la propiedad expresada por p , es decir:

$$P = \{ u \in U / u \text{ es } p \text{ es verdadero} \}$$

de forma que “*u es p*”, “*u cumple p*” y “*u está en P*” son proposiciones equivalentes ([Trillas, 94]). Se dice que P es la extensión conjuntista del predicado p en U .

Conocer P es equivalente a conocer su **función característica** definida como:

$$P(u) = \begin{cases} 1 & \text{si } u \in P \\ 0 & \text{si } u \notin P \end{cases}$$

y se dice también que $P(u)$ es la extensión funcional del predicado p en U .

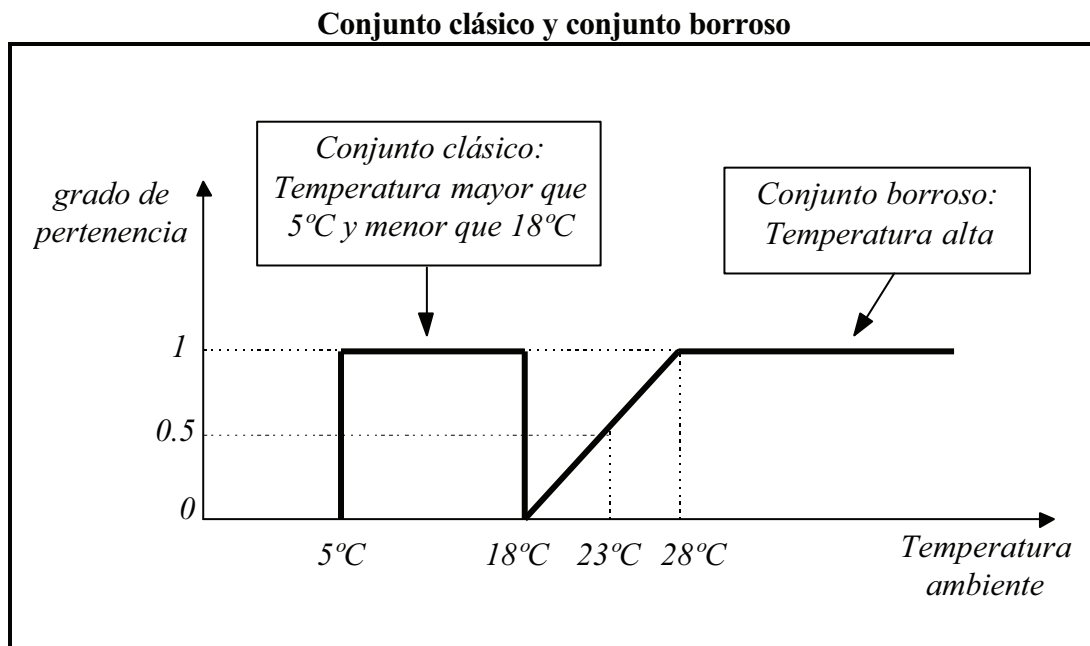
Cuando el predicado p expresa una propiedad graduada (por ejemplo “*mucho mayor que 0.5*”, “*parecido a 0.5*”, etc.) se dice que es un predicado vago en U , y las proposiciones “*u es p*” ya no son sólo ciertas o falsas. Cada proposición “*u es p*” tiene asociado un grado de verdad $t \in [0,1]$, lo que lleva a considerar la función de compatibilidad o **función de pertenencia** $P(u)$ (que es una generalización de la función característica) definida como:

$$P(u) = t \quad \text{ssi} \quad \text{“} u \text{ es } p \text{” es verdadero en grado } t$$

que puede ser interpretada como el grado en que los elementos u de U pertenecen a P , siendo P un conjunto borroso definido en U , con nombre p , y siendo $P(u)$ la función de pertenencia de dicho conjunto. Un conjunto borroso P en U queda por tanto definido por su función de pertenencia $P(u): U \rightarrow [0,1]$, y constituye una forma natural de expresar la vaguedad inherente en el lenguaje. En este contexto $P(u)$ será interpretado como el parecido o la similitud de u con el concepto o propiedad expresado por P . A es respecto puede consultarse [Bilgiç, Türksen, 96], donde se ofrece un repaso de distintas formas de interpretar los conjuntos borrosos así como algunos métodos de carácter práctico para la obtención de sus funciones de pertenencia.

La proposición “ u es p ” la denotaremos también por “ u es P ” donde P es la extensión conjuntista de la propiedad p . Utilizaremos la minúscula cuando queramos hacer referencia explícita a la propiedad para diferenciarla del conjunto de los elementos que la verifican. El grado de pertenencia de un elemento u de un universo U a una conjunto P lo denotaremos como $P(u)$, y cuando no se haga referencia a ningún conjunto concreto se utilizara la notación $\mu(u)$, que debe leerse sencillamente como grado de pertenencia. La notación $\mu_P(u)$ se utilizará también algunas veces como forma alternativa de expresar el grado de pertenencia de u a P .

La siguiente figura muestra un conjunto clásico y un conjunto borroso. Nótese que el conjunto clásico contiene a aquellos elementos que verifican una propiedad no gradual, mientras que el borroso a aquellos que verifican una propiedad gradual, dando lugar a grados de pertenencia intermedios entre 0 y 1, y a transiciones suaves, de total pertenencia a la no pertenencia.



3.3.1.1 Igualdad de conjuntos

Dos conjuntos borrosos P_1 y P_2 definidos en U son iguales cuando verifican:

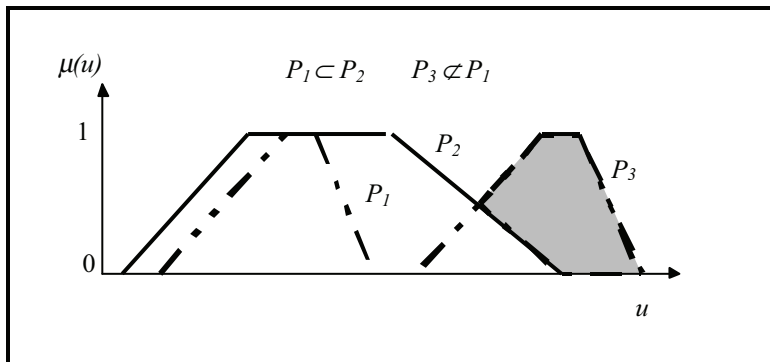
$$P_1(u) = P_2(u) \quad \forall u \in U$$

3.3.1.2 Inclusión entre conjuntos

Dados dos conjuntos borrosos P_1 y P_2 definidos en U se dice que P_1 está incluido en P_2 cuando se verifica:

$$P_1(u) \leq P_2(u) \quad \forall u \in U$$

lo que denotaremos por $P_1 \subseteq P_2$. La siguiente figura muestra gráficamente la inclusión entre conjuntos:



A veces es necesario distinguir entre inclusión e inclusión estricta, para lo que se usan los símbolos \subseteq y \subset de forma que:

$$P_1 \subset P_2 \text{ ssi } P_1 \subseteq P_2 \text{ y además } P_1 \neq P_2$$

Nótese que no debe confundirse $P_1 \subset P_2$ con $P_1(u) < P_2(u)$, ya que aunque se dé la igualdad de grados de pertenencia para algún valor de u , puede seguir verificándose $P_1 \subset P_2$.

3.3.1.3 Núcleo y soporte de un conjunto borroso

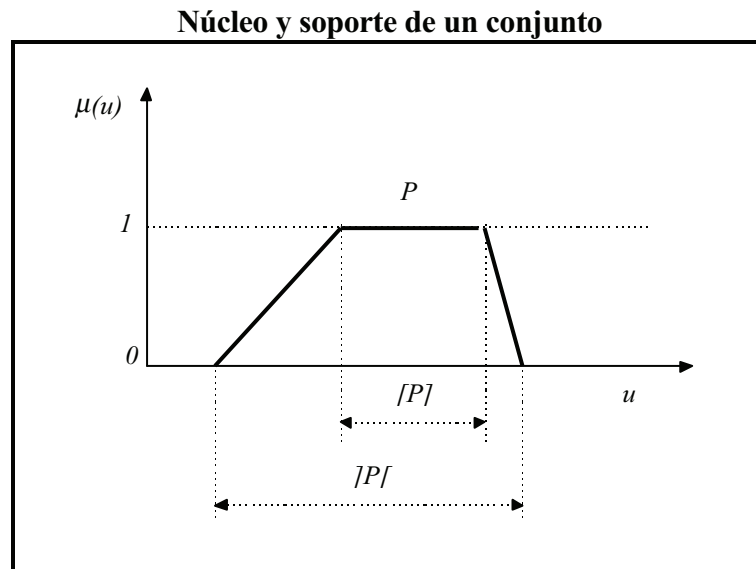
El núcleo de un conjunto borroso P definido en U es el conjunto clásico en U , que denotaremos por $[P]$, definido por:

$$[P] = \{ u \in U / P(u) = 1 \}$$

El soporte de un conjunto borroso P definido en U es el conjunto clásico en U , que denotaremos por $]P[$, definido por:

$$]P[= \{ u \in U / P(u) > 0 \}$$

Nótese que tanto el soporte como el núcleo de un conjunto borroso pueden estar vacíos. En la siguiente figura puede verse el núcleo y el soporte de un conjunto borroso con forma trapezoidal:



3.3.1.4 α -cortes de un conjunto borroso

Los α -cortes de un conjunto borroso son los conjuntos clásicos definidos como:

$$P^\alpha = \{ u / P(u) \geq \alpha \}$$

y los α -cortes estrictos como:

$$P^\alpha = \{ u / P(u) > \alpha \}$$

A modo de ejemplo el núcleo de P es el α -corte (no estricto) obtenido para $\alpha=1$, es decir $P^1 = [P]$.

3.3.1.5 Altura de un conjunto borroso

La altura de un conjunto borroso P definido en U , que denotaremos por $Alt(P)$, es el máximo grado de pertenencia de P , y viene dado por:

$$Alt(P) = \sup_{u \in U} P(u)$$

Cuando la altura de P sea $Alt(P)=1$ diremos que P es normal o que está normalizado.

3.3.1.6 Cardinal de un conjunto

El cardinal de un conjunto P es una medida del número de elementos que contiene, y se define para universos discretos como:

$$card(P) = \sum_{u \in U} P(u)$$

definición que también puede extenderse a conjuntos continuos mediante la expresión:

$$card(P) = \int_U P(u) du$$

3.3.1.7 Conjuntos borrosos convexos

Diremos que un conjunto borroso P es convexo cuando se cumple:

$$\forall u, u_1 y u_2 \in U \text{ si } u_1 \leq u \leq u_2 \text{ entonces } P(u) \geq \min(P(u_1), P(u_2))$$

o si $U \subseteq \mathfrak{R}$:

$$\forall u_1 y u_2 \in U, \forall \lambda \in [0, 1], P(\lambda u_1 + (1-\lambda)u_2) \geq \min(P(u_1), P(u_2))$$

Nótese que para poder hablar de convexidad el universo U tiene que estar ordenado.

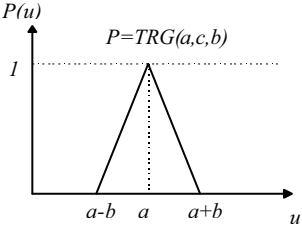
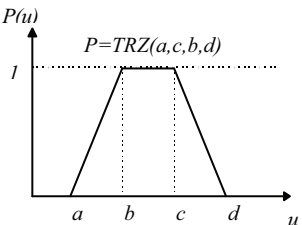
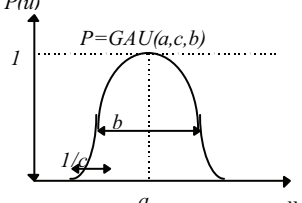
3.3.1.8 Números borrosos

Se dice que un conjunto borroso es un número borroso cuando está definido en \mathfrak{R} , es normal, convexo, tiene soporte acotado, y su función de pertenencia es continua. En algunos textos cuando el núcleo tiene más de un elemento se le llama intervalo borroso en lugar de número borroso. En cualquier caso la definición varía ligeramente de unos autores a otros.

3.3.1.9 Conjuntos borrosos paramétricos

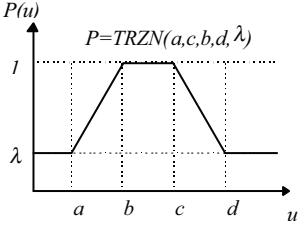
La definición de conjunto borroso permite en principio el uso de cualquier tipo de función de pertenencia. Resulta cómodo sin embargo utilizar tipos paramétricos, más sencillos de definir y manejar ya que de este modo la función de pertenencia queda caracterizada por su tipo y un número reducido de parámetros.

Algunos de los más utilizados son los que aparecen en la siguiente tabla. En ella P denota el conjunto borroso que se define, $P(u)$ es el grado de pertenencia de un elemento u de U al conjunto P , y a, b, c, d y λ son los parámetros de la expresión paramétrica de la función de pertenencia de P .

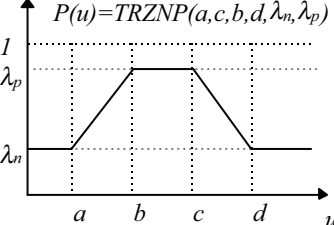
<p>Triangular (simétrico)</p>	$P(u) = \begin{cases} 0 & \text{si } u < a - b \text{ o } u > a + b \\ \frac{ u - a }{b} & \text{si } u \in [a - b, a + b] \end{cases}$	
<p>Trapezoidal</p>	$P(u) = \begin{cases} 0 & \text{si } u < a \\ \frac{u - a}{b - a} & \text{si } u \in [a, b] \\ 1 & \text{si } u \in [b, c] \\ \frac{d - u}{d - c} & \text{si } u \in [c, d] \\ 0 & \text{si } u > d \end{cases}$	
<p>Campana</p>	$P(u) = e^{-\left(\frac{u-a}{b}\right)^{2c}}$	

Uno de los tipos más empleados en el ámbito de la lógica son los conjuntos trapezoidales por su sencillez y riqueza semántica, mientras que los triangulares y de campana lo son más en ámbitos numéricos como el control o la aproximación funcional. En [Martin-Clouaire,

89] se extiende la definición de conjunto trapezoidal a conjunto λ -trapezoidal, para poder modelar los conjuntos que resultan de inferir conclusiones cuando la hipótesis y la observación de una regla borrosa se modelan con conjuntos trapezoidales. Su forma y expresión matemática queda reflejada en la siguiente tabla:

λ -Trapezoidal	$P(u) = \begin{cases} \lambda & \text{si } u < a \\ \lambda + \frac{\lambda - 1}{a - b}(u - a) & \text{si } u \in [a, b] \\ 1 & \text{si } u \in [b, c] \\ \lambda + \frac{\lambda - 1}{d - c}(u - d) & \text{si } u \in [c, d] \\ \lambda & \text{si } u > d \end{cases}$	
------------------------	--	--

Veremos sin embargo que el uso de distribuciones de posible posibilidad, que se definen en el apartado 3 de este capítulo, hace necesario recurrir a conjuntos no normalizados, por lo que se ha extendido la definición de conjuntos borrosos λ -trapezoidales a una familia más general que también los engloba, que hemos denominado conjuntos λ_{np} -trapezoidales, y que definimos en la siguiente tabla:

λ_{np} -Trapezoidal	$P(u) = \begin{cases} \lambda_n & \text{si } u < a \\ \lambda_n + \frac{\lambda_n - \lambda_p}{a - b}(u - a) & \text{si } u \in [a, b] \\ \lambda_p & \text{si } u \in [b, c] \\ \lambda_n + \frac{\lambda_n - \lambda_p}{d - c}(u - d) & \text{si } u \in [c, d] \\ \lambda_n & \text{si } u > d \end{cases}$	
-----------------------------	--	--

3.4 Operaciones con conjuntos borrosos

En la teoría de conjuntos borrosos pueden definirse múltiples operaciones entre conjuntos, algunas de las cuales son extensiones directas de operaciones clásicas, como por ejemplo las intersecciones, uniones y negaciones, y por tanto particularizadas a conjuntos clásicos permiten recuperar las operaciones clásicas, mientras que otras sólo tienen sentido en la teoría de conjuntos borrosos. Este es el caso de las medias (calculadas como medias entre grados de pertenencia de los conjuntos), que por ejemplo aplicadas a conjuntos clásicos producen en general conjuntos borrosos. Por otro lado el principio de extensión permite extender a conjuntos borrosos funciones ordinarias, y en concreto funciones reales. Así, además de las operaciones básicas como la suma, resta, etc., pueden definirse otras operaciones de interés, como por ejemplo las medias, obtenidas a partir de la extensión a números borrosos de medias reales (calculadas en este caso como medias entre elementos del universo de definición de los conjuntos y no entre grados de pertenencia).

3.4.1 La negación de conjuntos borrosos

La negación de un conjunto borroso P es otro conjunto borroso $\neg P$ con función de pertenencia dada por:

$$\neg P(u) = \Gamma(P(u))$$

donde $\Gamma(x)$ es un operador de negación. Diremos que una función $\Gamma(x):[0,1] \rightarrow [0,1]$ es un operador de negación, o simplemente una negación cuando cumpla:

$$\begin{aligned} \Gamma(0) &= 1 \text{ y } \Gamma(1) = 0 \\ \Gamma(x) &\text{ no creciente con } x \end{aligned}$$

La mayor negación es $\Gamma_{max}(x) = \begin{cases} 0 & \text{si } x = 1 \\ 1 & \text{otro caso} \end{cases}$ y la menor $\Gamma_{min}(x) = \begin{cases} 1 & \text{si } x = 0 \\ 0 & \text{otro caso} \end{cases}$

Existen algunas negaciones que por sus propiedades son más comúnmente utilizadas ([Alsina, Trillas, Valverde, 83], [Dubois, Prade, 85], [Dubois, Prade, 84]) y que repasamos a continuación.

3.4.1.1 Negación estricta

Una negación Γ es una negación estricta cuando es continua y estrictamente decreciente con x .

3.4.1.2 Negaciones fuertes

Una negación fuerte es una negación estricta $\Gamma(x)$ que además cumple:

$$\Gamma(\Gamma(x))=x \quad (\text{involutiva})$$

Una negación fuerte puede ponerse en función de un generador g :

$$\Gamma(x)=g^{[-1]}(g(1)-g(x))$$

donde $g^{[-1]}$ es la pseudo-inversa de g , y ambas verifican:

$g: \begin{cases} [0,1] \rightarrow [0,\infty[\\ \text{continua y estrictamente creciente} \\ g(0)=0 \\ g(1)=1 \end{cases}$	$g^{[-1]}: \begin{cases} [0,1] \rightarrow [0,\infty[\\ g^{[-1]}(x) = \begin{cases} g^{-1}(x) & \text{si } x \in [0,g(1)] \\ 0 & \text{si } x \in [g(1),\infty[\end{cases} \\ g^{[-1]} \text{ continua y creciente} \end{cases}$
--	--

Si k es una constante positiva, $k.g(x)$ genera la misma negación que $g(x)$, por lo que puede suponerse $g(1)=1$ cuando $g(1)$ es finito, sin pérdida de generalidad.

Puede estudiarse el caso en que $g(1)$ no es finito, obteniéndose la negación Γ_{max} , que no es una negación fuerte. Como veremos estas son las negaciones generadas por las r-implicaciones residuadas de t-normas positivas haciendo $\Gamma(x)=I_R^T(0,x)$, donde $I_R^T(x,y)$ es una función de implicación residual (véase el apartado 3.5.6). Del mismo modo las implicaciones fuertes permiten obtener la negación fuerte utilizada para generarlas, ya que $\Gamma(x)=I_S^T(0,x)$.

En lo que sigue utilizaremos principalmente la negación $\Gamma(x)=1-x$, que llamaremos negación habitual.

3.4.1.3 Ejemplos de operadores de negación

Zadeh	<i>fuerte</i>	$g(x)=x$	$\Gamma(x)=1-x$
Redondeada	<i>fuerte</i>	$g(x)=x^2$	$\Gamma(x) = \sqrt{1-x^2}$
Sugeno	<i>fuerte</i>	$g(x) = \frac{1}{s} \log(1+sx)$	$\Gamma_s(x) = \frac{1-x}{1+sx}$ con $s > -1$
Yager	<i>fuerte</i>	$g(x) = x^w$	$\Gamma(x) = (1-x^w)^{\frac{1}{w}}$
Klir	<i>no involutiva</i>	<i>sin generador</i>	$\Gamma(x)=0.5(1+\cos(\pi x))$
Crisp	<i>no involutiva discontinua</i>	<i>sin generador</i>	$\Gamma(x) = \begin{cases} 1 & \text{si } x \leq t \\ 0 & \text{si } x > t \end{cases}$ con $t \in [0,1)$
Maxima negación	<i>no involutiva discontinua</i>	<i>sin generador</i>	$\Gamma_{max}(x) = \begin{cases} 0 & \text{si } x = 1 \\ 1 & \text{en otro caso} \end{cases}$
Mínima negación	<i>no involutiva discontinua</i>	<i>sin generador</i>	$\Gamma_{min}(x) = \begin{cases} 1 & \text{si } x = 0 \\ 0 & \text{en otro caso} \end{cases}$

3.4.2 La intersección y la unión de conjuntos borrosos

La intersección y la unión de dos conjuntos borrosos, P_1 y P_2 definidos en U son dos conjuntos borrosos $P_1 \cap P_2$ y $P_1 \cup P_2$ definidos en U con funciones de pertenencia:

$$(P_1 \cap P_2)(u) = A(P_1(u), P_2(u))$$

$$(P_1 \cup P_2)(u) = O(P_1(u), P_2(u))$$

donde A es un operador de agregación conjuntivo y O es un operador de agregación disyuntivo.

Diremos que una función $A(x,y):[0,1] \times [0,1] \rightarrow [0,1]$ es un operador de agregación conjuntivo cuando cumpla las siguientes propiedades:

$$A(0,1) = A(1,0) = A(0,0) = 0$$

$$A(1,1) = 1$$

$$A(x,y) \text{ no decreciente ni con } x \text{ ni con } y$$

Del mismo modo diremos que una función $O(x,y):[0,1] \times [0,1] \rightarrow [0,1]$ es un operador de agregación disyuntivo cuando verifique:

$$\begin{aligned} O(0,1) &= O(1,0) = O(1,1) = 1 \\ O(0,0) &= 0 \\ O(x,y) & \text{ no decreciente ni con } x \text{ ni con } y \end{aligned}$$

Los mayores y menores operadores conjuntivos y disyuntivos aparecen en la siguiente tabla:

$A_{max}(x,y) = \begin{cases} 0 & \text{si } x = 0 \text{ o bien } y = 0 \\ 1 & \text{en otro caso} \end{cases}$	$A_{min}(x,y) = \begin{cases} 1 & \text{si } x = y = 1 \\ 0 & \text{en otro caso} \end{cases}$
$O_{max}(x,y) = \begin{cases} 0 & \text{si } x = y = 0 \\ 1 & \text{en otro caso} \end{cases}$	$O_{min}(x,y) = \begin{cases} 1 & \text{si } x = 1 \text{ o bien } y = 1 \\ 0 & \text{en otro caso} \end{cases}$

Diremos que un operador conjuntivo $A(x,y)$ y uno disyuntivo $O(x,y)$ son duales respecto a una negación $\Gamma(x)$, o que son Γ -duals, cuando se verifique:

$$\begin{aligned} A(x,y) &= \Gamma(O(\Gamma(x), \Gamma(y))) \\ O(x,y) &= \Gamma(A(\Gamma(x), \Gamma(y))) \end{aligned}$$

Cuando Γ sea la negación habitual diremos que operadores Γ -duals son simplemente duals.

Entre los operadores conjuntivos y disyuntivos destacan por su extendido uso y sus buenas propiedades matemáticas las conocidas t-normas y t-conormas que se describen en el apartado 3.4.2.2. También, aunque de un uso más restringido, pueden destacarse las pseudo-conjunciones que resultan de obtener la máxima función generadora de modus ponens de las s-implicaciones, y sus duals, las pseudo-disyunciones, descritas en el apartado 3.4.2.3.

3.4.2.1 Propiedades

Los operadores de unión, intersección y negación en la teoría clásica de conjuntos verifican un gran número de propiedades interesantes. Los operadores borrosos deben verificar estas propiedades en los contornos de sus dominios de definición para generalizar correctamente a los operadores clásicos. Puede también ser interesante preservar algunas de estas propiedades en el resto del dominio de definición, mientras que otras pueden ser consideradas innecesarias en el resto del dominio.

Dados un operador conjuntivo $A(x,y)$, un operador disyuntivo $O(x,y)$, una negación $\Gamma(x)$, y los conjuntos P, P_1, P_2 y P_3 definidos en U , la siguiente tabla resume algunas de las citadas propiedades, expresadas como uniones, intersecciones y negaciones de conjuntos, y también en función de los operadores A, O y Γ utilizados para modelar las operaciones conjuntistas.

CONMUTATIVIDAD	$(P_1 \cap P_2) = (P_2 \cap P_1)$ $(P_1 \cup P_2) = (P_2 \cup P_1)$	$A(x,y) = A(y,x)$ $O(x,y) = O(y,x)$
ASOCIATIVIDAD	$P_1 \cup (P_2 \cup P_3) = (P_1 \cup P_2) \cup P_3$ $P_1 \cap (P_2 \cap P_3) = (P_1 \cap P_2) \cap P_3$	$O(x, O(y,z)) = O(O(x,y), z)$ $A(x, A(y,z)) = A(A(x,y), z)$
INVOLUCIÓN	$\neg(\neg P) = P$	$\Gamma(\Gamma(x)) = x$
LEYES DE DE MORGAN	$\neg(P_1 \cap P_2) = \neg P_1 \cup \neg P_2$ $\neg(P_1 \cup P_2) = \neg P_1 \cap \neg P_2$	$A(x,y) = \Gamma(O(\Gamma(x), \Gamma(y)))$ $O(x,y) = \Gamma(A(\Gamma(x), \Gamma(y)))$
DISTRIBUTIVIDAD	$P_1 \cap (P_2 \cup P_3) = (P_1 \cap P_2) \cup (P_1 \cap P_3)$ $P_1 \cup (P_2 \cap P_3) = (P_1 \cup P_2) \cap (P_1 \cup P_3)$	$A(x, O(y,z)) = O(A(x,y), A(x,z))$ $O(x, A(y,z)) = A(O(x,y), O(x,z))$
ABSORCIÓN	$(P_1 \cap P_2) \cup P_1 = P_1$ $(P_1 \cup P_2) \cap P_1 = P_1$	$O(A(x,y), x) = x$ $A(O(x,y), x) = x$
IDEMPOTENCIA	$P \cup P = P$ $P \cap P = P$	$O(x,x) = x$ $A(x,x) = x$
TERCERO EXCLUIDO	$P \cup \neg P = U$	$O(x, \Gamma(x)) = 1$
NO CONTRADICCIÓN	$P \cap \neg P = \emptyset$	$A(x, \Gamma(x)) = 0$
ABSORCIÓN CON U	$P \cup U = U$	$O(x, 1) = 1$
ABSORCIÓN CON \emptyset	$P \cap \emptyset = \emptyset$	$A(x, 0) = 0$
IDENTIDAD	$P \cap U = P$ $P \cup \emptyset = P$	$A(x, 1) = x$ $O(x, 0) = x$

A continuación repasaremos algunos de las propiedades más importantes de las familias de operadores conjuntivos y disyuntivos más utilizadas.

3.4.2.2 Las t-normas y t-conormas

Una t-norma $T(x,y):[0,1] \times [0,1] \rightarrow [0,1]$ (véase por ejemplo [Weber, 83], [Alsina, Trillas, Valverde, 83], [Alsina, Trillas, 84], [Dubois, Prade, 85] o [Yager, 91]) es un operador de agregación conjuntivo que cumple:

$$\begin{aligned} \text{conmutatividad: } T(x,y) &= T(y,x) \\ \text{asociatividad: } T(x, T(y,z)) &= T(T(x,y), z) \end{aligned}$$

$$T(x,1)=x, T(x,0)=0$$

Una t-conorma $S(x,y):[0,1] \times [0,1] \rightarrow [0,1]$ es un operador de agregación disyuntivo que cumple:

$$\begin{aligned} \text{conmutatividad: } S(x,y) &= S(y,x) \\ \text{asociatividad: } S(x, T(y,z)) &= S(T(x,y), z) \\ S(x,1) &= 1, S(x,0) = x \end{aligned}$$

A partir de una t-norma T cualquiera se puede obtener una t-conorma S como su Γ -dual respecto a una negación fuerte Γ , es decir:

$$S(x,y) = \Gamma(T(\Gamma(x), \Gamma(y)))$$

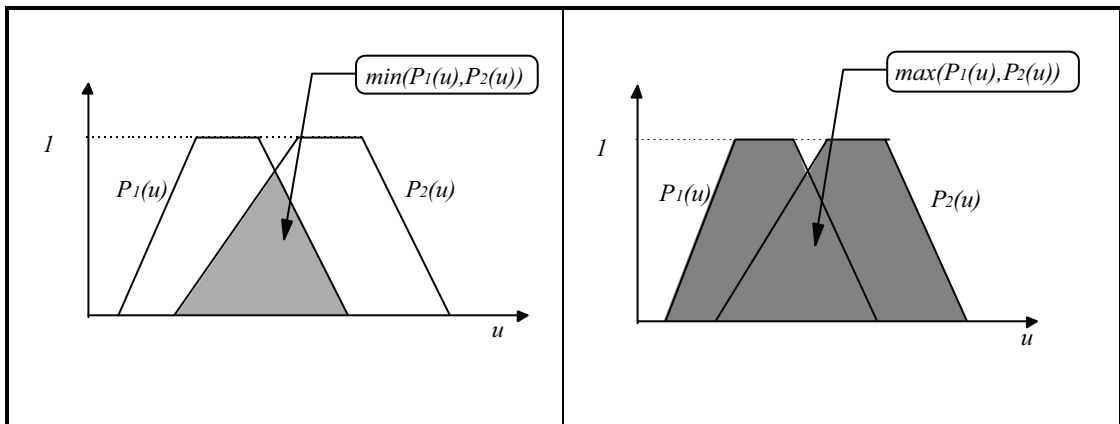
La menor t-norma y mayor t-conorma son discontinuas, duales y vienen dadas por:

$$T_{min}(x,y) = \begin{cases} x & \text{si } y = 1 \\ y & \text{si } x = 1 \\ 0 & \text{en otro caso} \end{cases} \quad \text{y} \quad S_{max}(x,y) = \begin{cases} x & \text{si } y = 0 \\ y & \text{si } x = 0 \\ 1 & \text{en otro caso} \end{cases}$$

Entre las t-normas y t-conormas continuas pueden distinguirse básicamente tres tipos de operadores distintos. Los operadores idempotentes, los operadores estrictos (o positivos) y los operadores nilpotentes (o negativos).

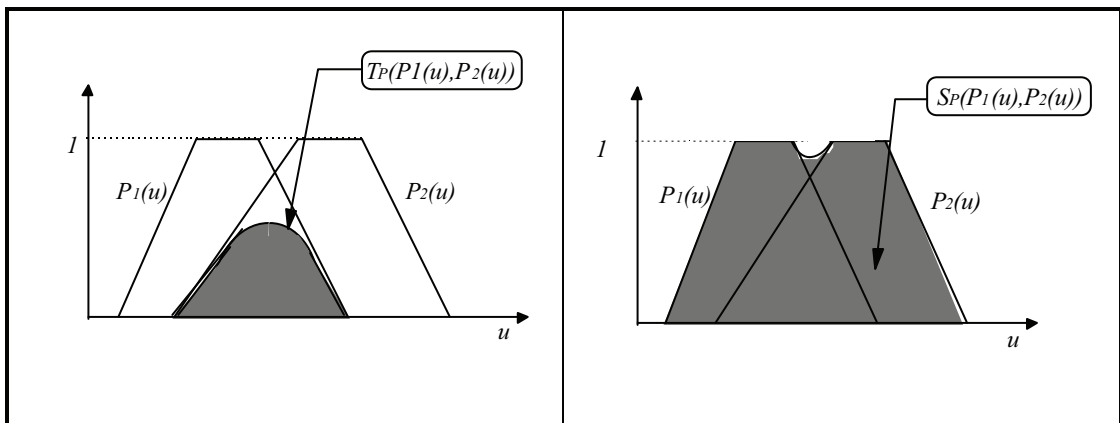
- t-normas y t-conormas idempotentes:

Se dice que un operador $H(x,y)$ es **idempotente** si cumple $H(x,x)=x$. Las únicas t-norma y t-conorma continuas e idempotentes son $T(x,y)=\min(x,y)$ y $S(x,y)=\max(x,y)$. La t-norma $\min(x,y)$ es la mayor de todas las t-normas, mientras que la t-conorma $\max(x,y)$ es la menor de todas las t-conormas. Ambas son duales respecto a la negación habitual, y son los operadores de intersección y unión más utilizados. La siguiente figura ilustra su comportamiento:



- t-normas y t-conormas estrictas (positivas):

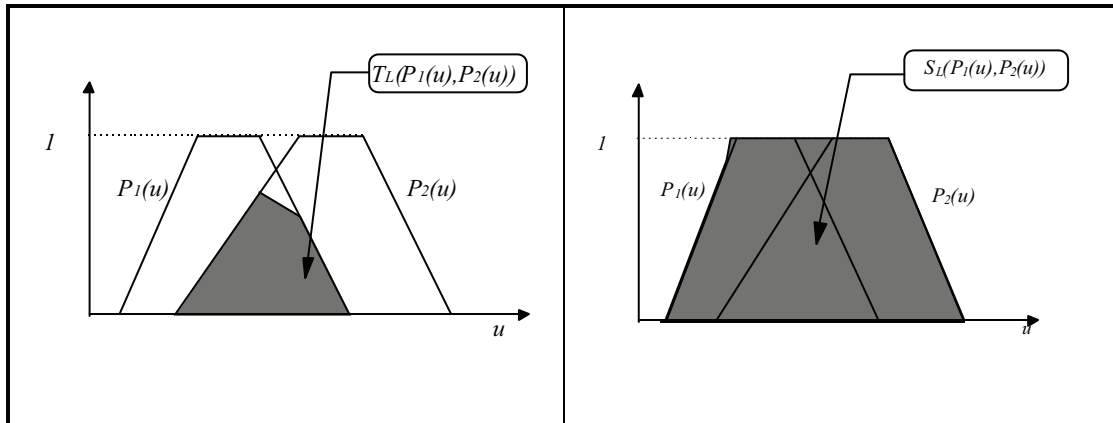
Una t-norma o t-conorma $H(x,y)$ continua se dice **estricta** cuando H es estrictamente creciente tanto con x como con y cuando $x,y \in]0,1[$. Estas t-normas cumplen $T(x,y)=0$ ssi $x=0$ o bien $y=0$, mientras que las t-conormas cumplen $S(x,y)=1$ ssi $x=1$ o bien $y=1$, y se dicen **operadores positivos**. Estos operadores nunca son mutuamente distributivos ni idempotentes, ni satisfacen la no contradicción ni la ley del tercero excluido para ninguna negación. Las t-normas positivas las denotaremos por T^+ , las t-conormas por S^+ . Un ejemplo típico son los operadores probabilísticos, $T_P(x,y)=p(x,y)=x \cdot y$ y $S_P(x,y)=x+y-x \cdot y$, duales respecto a la negación habitual. La siguiente figura ilustra de forma aproximada su comportamiento:



- t-normas y t-conormas nilpotentes (negativas):

Una t-norma o t-conorma continua se dice nilpotente cuando no es estricta, y se les dice también **operadores negativos**. Para las t-normas negativas se verifica $T(x,y)=0$ con $x \neq 0$ e $y \neq 0$, y para las t-conormas negativas $S(x,y)=1$ con $x \neq 1$ e $y \neq 1$. Las t-normas negativas

las denotaremos por T , las t-conormas por S . Un típico ejemplo son los operadores de Lukasiewicz, $T_L(x,y)=L(x,y)=\max(0,x+y-1)$ y $S_L(x,y)=\min(1,x+y)$, duales respecto a la negación habitual. La siguiente figura ilustra su comportamiento:



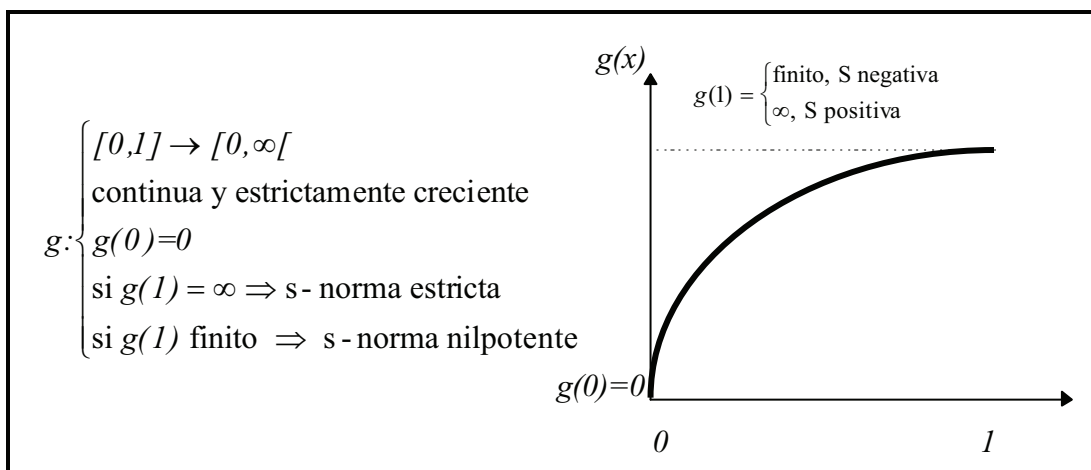
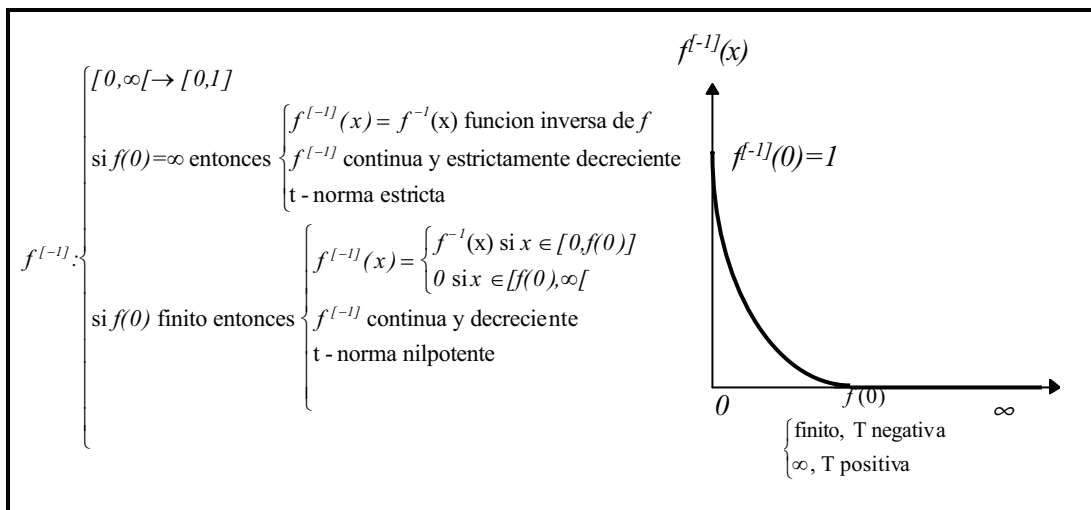
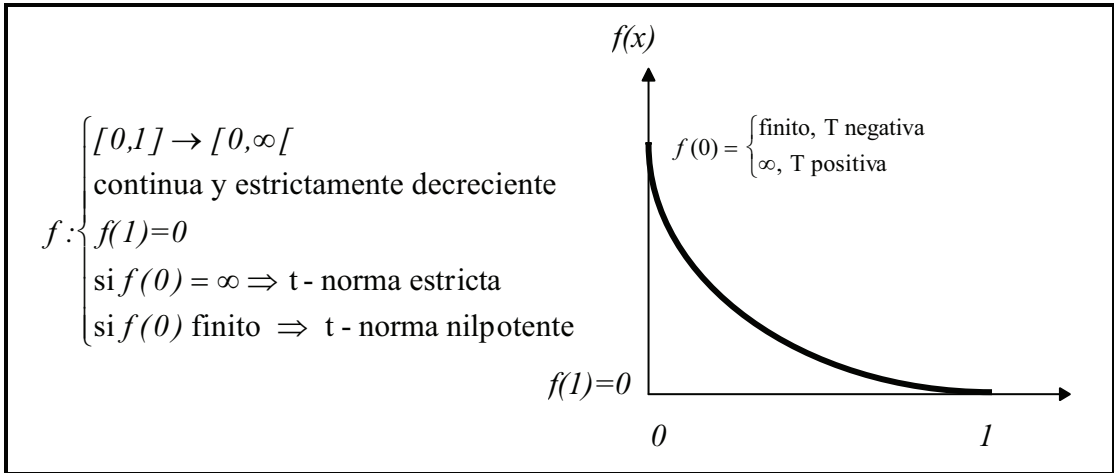
- t-normas y t-conormas arquímedeas y generadores aditivos

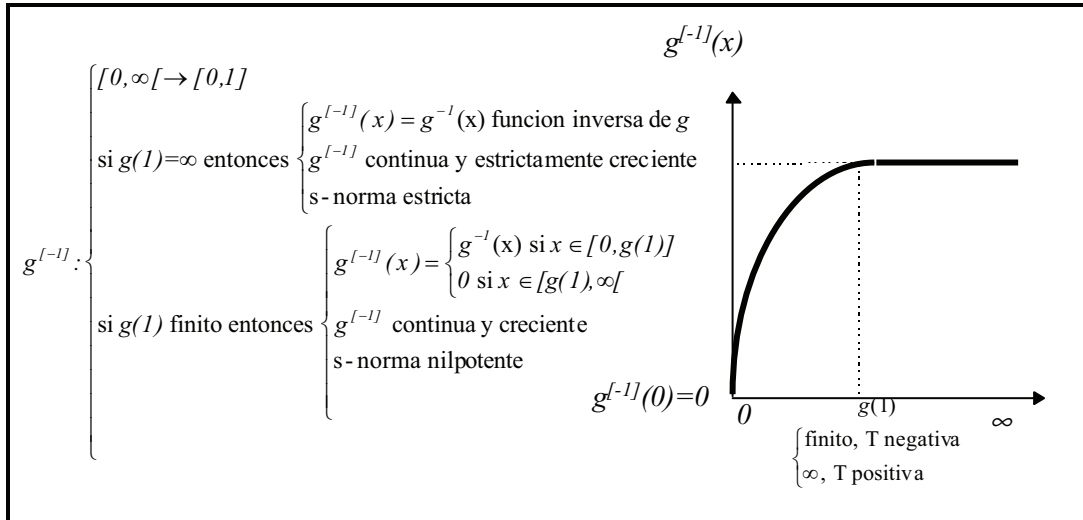
Cualquier t-norma o t-conorma continua se dice que es **arquímedea** si verifica $T(x,x) < x$ para todo $x \in (0,1)$. Son por tanto arquímedeas las t-normas y t-conormas estrictas y nilpotentes, y todas ellas pueden ponerse en función de un generador aditivo f o g :

$$T(x,y)=f^{-1}(f(x)+f(y))$$

$$S(x,y)=g^{-1}(g(x)+g(y))$$

donde f^{-1} y g^{-1} son las pseudo-inversas de f y g . Las propiedades de los generadores f y g así como las de sus pseudo-inversas quedan representadas en las siguientes figuras:





Los operadores *min* y *max* no son arquímdeos y no tienen generadores aditivos. Existen pocos tipos más de t-normas (o t-conormas) continuas, que se obtienen principalmente como operaciones arquímdeas en ciertas regiones de su dominio de definición y *min* (o *max*) en el resto.

Es interesante también señalar que el carácter positivo o negativo de una t-norma permite establecer una cierta relación de orden entre unas y otras. Así mientras que el operador *min* es la mayor t-norma, las t-normas positivas vendrían a continuación y las negativas pueden considerarse las más pequeñas de todas, a excepción de la mínima t-norma que no es continua. Lo mismo puede hacerse en orden inverso para las t-conormas, siendo *max* la menor de todas.

Para las t-normas y t-conormas nilpotentes, para las que se tiene $f(0)$ y $g(1)$ finitos, dado que los generadores son únicos salvo una constante positiva (si $f(x)$ es generador de T también lo es $k \cdot f(x)$ siendo $k > 0$) puede suponerse sin pérdida de generalidad que $f(0) = g(1) = 1$.

Nótese que en general se tiene ([Weber, 83]):

$$f^{[-1]}(f(x)) = x$$

$$f(f^{[-1]}(y)) = \min(y, f(0)) \quad (f(f^{[-1]}(y)) = y \text{ para el caso estricto})$$

$$g^{[-1]}(g(x)) = x$$

$$g(g^{[-1]}(y)) = \min(y, g(1)) \quad (g(g^{[-1]}(y)) = y \text{ para el caso estricto})$$

Dado un generador aditivo f para una t-norma T , y dada una negación fuerte Γ , la función $g(x) = \Gamma(f(x))$ genera aditivamente una t-conorma S , Γ -dual de T . Del mismo modo dado un generador aditivo g para una t-conorma S , y dada una negación fuerte Γ , la función $f(x) = \Gamma(g(x))$ genera aditivamente una t-norma T , Γ -dual de S .

Por último señalar que para parejas de t-normas y t-conormas duales respecto a $\Gamma(x) = 1-x$ se tiene:

$$\begin{aligned} g(x) &= f(1-x) \\ g^{\Gamma^{-1}}(x) &= 1-f^{\Gamma^{-1}}(x) \end{aligned}$$

- distributividad:

Una propiedad de especial importancia para ciertos análisis teóricos es la distributividad entre operadores lógicos. Recordemos que la distributividad de una t-norma T respecto a una t-conorma S se formula como:

$$T(x, S(y, z)) = S(T(x, y), T(x, z))$$

y del mismo modo la de una t-conorma S respecto a una t-norma T se formula como:

$$S(x, T(y, z)) = T(S(x, y), S(x, z))$$

Si nos restringimos al análisis de t-normas y t-conormas continuas entonces (véase por ejemplo [Weber, 83]) se obtienen dos propiedades de especial interés:

a partir de la distributividad, $T(x, S(y, z)) = S(T(x, y), T(x, z))$, tomando $z=1$ se tienen $S(y, 1) = 1$ y $T(x, 1) = 1$, que es la propiedad de absorción, $x = S(T(x, y), x)$ y tomando $y=1$ se obtiene la idempotencia, $S(x, x) = x$, resultando dichas propiedades equivalentes, solamente verificadas por la t-conorma $S = \max$.

Del mismo modo,

a partir de la distributividad, $S(x, T(y, z)) = T(S(x, y), S(x, z))$, tomando $z=0$ se tiene $T(x, 0) = 0$ y $S(x, 0) = x$ obteniéndose la propiedad de absorción, $x = T(S(x, y), x)$, y tomando $y=1$ se obtiene la idempotencia, $S(x, x) = x$, resultando dichas propiedades equivalentes, solamente verificada por la t-norma $T = \min$.

Por tanto la distributividad implica a la absorción que implica a su vez a la idempotencia, y cualquier t-norma continua es distributiva respecto a la t-conorma \max , y cualquier t-conorma continua es distributiva respecto a la t-norma \min .

3.4.2.3 Las pseudo-conjunciones y pseudo-disyunciones

El estudio de otros operadores de unión e intersección distintos de las clásicas t-normas y t-conormas ha sido en general bastante limitado ([Dubois, Prade, 84b], [Fodor, Keresztfalvi, 93] o [Fodor, Keresztfalvi, 93b]), probablemente debido a las buenas propiedades de éstas, y al gran abanico de posibilidades que ofrecen. Sin embargo otros operadores conjuntivos aparecen de forma natural cuando se busca la función generadora de modus ponens de s-implicaciones y de otro tipo de implicaciones no residuales (véase el apartado - de este mismo capítulo o bien el capítulo 4), operadores que presentan algunas propiedades notablemente distintas de las de las t-normas (como por ejemplo la no-conmutatividad), y que, deseables o no, no son reproducibles con ninguna de estas.

Definimos una pseudo-conjunción $C(x,y)$ como:

$$C(x,y) = \inf\{c \in [0,1] / S(\Gamma(x),c) \geq y\}$$

donde $S(x,y)$ es una t-conorma y $\Gamma(x)$ es una negación fuerte ([Dubois, Prade, 84b], [Fodor, Keresztfalvi, 93] y [Fodor, Keresztfalvi, 93b]). Como se verá en el capítulo 4, $C(x,y)$ así definido resulta ser la máxima función generadora de modus ponens de la s-implicación $I_S(x,y) = S(\Gamma(x),y)$. En lo que sigue nos limitaremos principalmente a la negación habitual. En este caso, la pseudo-conjunción $C(x,y)$ puesta en función de la t-norma de la que procede queda:

$$C(x,y) = \inf\{c \in [0,1] / T(x,1-c) \leq 1-y\} = 1 - I_R^T(x,1-y)$$

donde $I_R^T(x,y)$ es la r-implicación residuada de la t-norma $T(x,y)$. Puede comprobarse que una pseudo-conjunción $C(x,y)$ es efectivamente un operador conjuntivo. De la definición de C se tiene:

$$\begin{aligned} C(x,y) &\text{ no decreciente con } x \text{ o con } y \\ C(0,y) &= \inf\{c \in [0,1] / S(1,c) \geq y\} = 0 \\ C(x,0) &= \inf\{c \in [0,1] / S(\Gamma(x),c) \geq 0\} = 0 \\ C(1,1) &= \inf\{c \in [0,1] / S(0,c) \geq 1\} = 1 \end{aligned}$$

Dada una pseudo-conjunción $C(x,y)$ se puede definir un operador de disyunción que llamaremos pseudo-disyunción $D(x,y)$, como el operador Γ -dual de C siendo Γ una negación fuerte:

$$D(x,y) = \Gamma(C(\Gamma(x), \Gamma(y)))$$

por lo que $D(x,y)$ también procede de una t-norma (o de su t-conorma Γ -dual).

- Generadores “aditivos”:

Cuando los operadores $C(x,y)$ y $D(x,y)$ provienen de una t-norma y una t-conorma arquímedeas respectivamente, y por tanto con generadores aditivos respectivos $f(x)$ y $g(x)$, pueden ponerse a su vez en función de dichos generadores, con lo que se obtiene:

$$C(x,y) = g^{f^{-1}}(g(y) - g(1-x))$$

$$D(x,y) = 1 - C(1-x, 1-y) = 1 - g^{f^{-1}}(g(1-y) - g(x))$$

y tomando $f(x) = g(1-x)$ y $f^{f^{-1}}(x) = 1 - g^{f^{-1}}(x)$ se tiene:

$$C(x,y) = 1 - f^{f^{-1}}(f(1-y) - f(x))$$

$$D(x,y) = f^{f^{-1}}(f(y) - f(1-x))$$

- Conmutatividad y asociatividad:

Es interesante señalar que las pseudo-conjunciones no son operadores conmutativos o asociativos. Sin embargo se pueden obtener dos propiedades que resultarán de interés en desarrollos posteriores. Si por ejemplo a partir de la expresión $M(x, M(y,z))$ se intenta agrupar los términos x y z se obtiene:

$$M(x, M(y,z)) = g^{f^{-1}}(g(M(y,z)) - g(1-x)) =$$

$$g^{f^{-1}}(g(g^{f^{-1}}(g(z) - g(1-y))) - g(1-x)) =$$

$$= g^{f^{-1}}(g(z) - g(1-y) - g(1-x)) =$$

$$g^{f^{-1}}(g(g^{f^{-1}}(g(z) - g(1-x))) - g(1-x)) =$$

$$M(M(x,z), y)$$

si por el contrario se agrupan los términos x e y se tiene:

$$M(x, M(y,z)) = g^{f^{-1}}(g(M(y,z)) - g(1-x)) =$$

$$g^{f^{-1}}(g(z) - g(1-y) - g(1-x)) =$$

$$g^{f^{-1}}(g(z) - (g(1-y) + g(1-x))) =$$

$$g^{f^{-1}}(g(z) - g(g^{f^{-1}}(g(1-y) + g(1-x)))) =$$

$$g^{f^{-1}}(g(z) - g(S(1-x, 1-y)))$$

Y como $g^{f^{-1}}(g(b) - g(a)) = g^{f^{-1}}(g(b) - g(1 - (1-a))) = M(1-a, b)$ queda por tanto:

$$M(x, M(y,z)) = M(1 - S(1-x, 1-y), z) = M(T(x,y), z)$$

3.4.2.4 Otras uniones e intersecciones

Pueden definirse otros muchos operadores conjuntivos y disyuntivos. Por ejemplo los que resultan de la implicación de Yager:

$$A_Y(x, y) = \begin{cases} 0 & \text{si } x=0 \\ \frac{1}{y^x} & \text{otro caso} \end{cases} \quad O_Y(x, y) = \begin{cases} 1 & \text{si } x=1 \\ 1 - (1-y)^{\frac{1}{1-x}} & \text{otro caso} \end{cases}$$

o los que propone se propone en [Trillas et al., 92]:

$$A_T(x, y) = \sqrt{x \cdot y} \quad O_T(x, y) = 1 - \sqrt{(1-x)(1-y)}$$

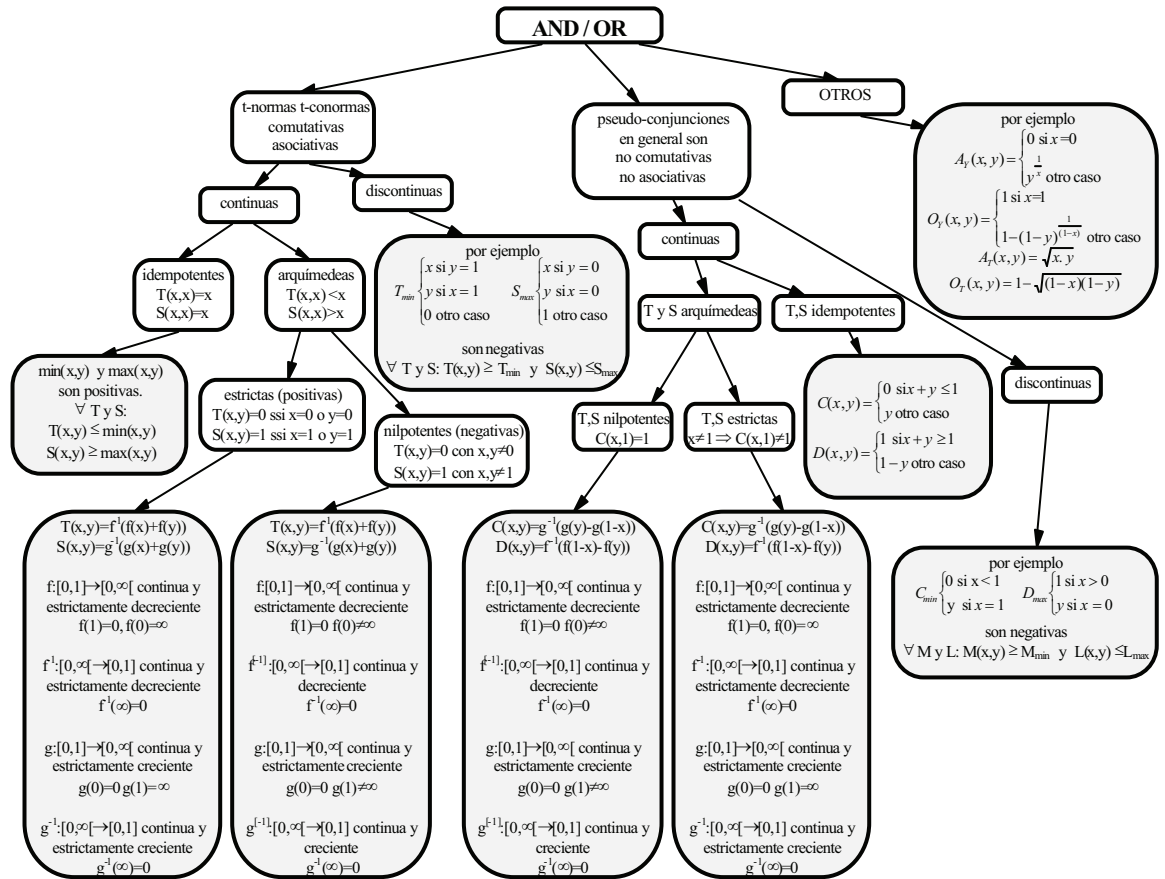
De hecho en trabajos clásicos como por ejemplo [Dubois, Prade, 84b], [Dubois, Prade, 84b], [Trillas, Valverde, 85] y [Bouchon, 87] y en otros más recientes como [Fodor, Keresztfalvi, 93], [Fodor, Keresztfalvi, 93b], [Trillas et al., 92], y [Trillas, Cubillo, Castro, 95]) se ha puesto de manifiesto la existencia de nuevas familias de operadores conjuntivos, $A(x, y)$, y de funciones de implicación $I(x, y)$, distintos de los digamos tradicionales dentro de la teoría de conjuntos borrosos (como son las t-normas, pseudo-conjunciones, r-implicaciones y s-implicaciones). En estos trabajos también se pone de manifiesto las ventajas de definir ambos tipos de operadores de forma conjunta, a través de la desigualdad del modus ponens:

$$A(x, I(x, y)) \leq y$$

de forma que a partir de una implicación se obtiene una conjunción o viceversa, y definido uno de ellos el otro queda determinado, lo que abre más aún el amplio abanico de posibilidades.

3.4.2.5 Resumen

El esquema de la siguiente tabla resume la clasificación que se ha hecho de los operadores conjuntivos y disyuntivos.



La siguiente tabla es un resumen de las principales t-normas y t-conormas, sus generadores aditivos en caso de que existan, y en algunos casos aparecen también las pseudo-conjunciones y pseudo-disyunciones que dichas t-normas generan.

$T(x,y)$ $S(x,y)$	$f(x)$ $g(x)$	$C(x,y)$ $D(x,y)$
$T(x,y) = \min(x,y)$ $S(x,y) = \max(x,y)$	sin generadores positivas	$C(x,y) = \begin{cases} 0 & \text{si } x + y \leq 1 \\ y & \text{otro caso} \end{cases}$ $D(x,y) = \begin{cases} 1 & \text{si } x + y \geq 1 \\ 1 - y & \text{otro caso} \end{cases}$

$T_{min}(x, y) = \begin{cases} x & \text{si } y = 1 \\ y & \text{si } x = 1 \\ 0 & \text{en otro caso} \end{cases}$ $S_{max}(x, y) = \begin{cases} x & \text{si } y = 0 \\ y & \text{si } x = 0 \\ 1 & \text{en otro caso} \end{cases}$	<p>sin generadores negativas</p>	$C(x, y) = \begin{cases} 0 & \text{si } x < 1 \\ y & \text{si } x = 1 \end{cases}$ $D(x, y) = \begin{cases} 1 & \text{si } x > 0 \\ y & \text{si } x = 0 \end{cases}$
$T(x, y) = x \cdot y$ $S(x, y) = x + y - xy$	<p>$-\ln(x)$ $-\ln(1-x)$ positivas</p>	$C(x, y) = \begin{cases} 0 & \text{si } x = 0 \\ \max\left(0, \frac{x+y-1}{x}\right) \end{cases}$ $D(x, y) = \begin{cases} 1 & \text{si } x = 1 \\ \min\left(1, \frac{y}{1-x}\right) \end{cases}$
<p>Lukasiewicz</p> $T(x, y) = \max(0, x+y-1)$ $S(x, y) = \min(1, x+y)$	<p>$1-x$ x negativas</p>	$C(x, y) = \max(0, x+y-1)$ $D(x, y) = \min(1, x+y)$
<p>Yager</p> $T(x, y) = 1 - \min\left(1, \left((1-x)^p + (1-y)^p\right)^{\frac{1}{p}}\right)$ $S(x, y) = \min\left(1, \left(x^p + y^p\right)^{\frac{1}{p}}\right)$	<p>$(1-x)^p$ x^p $p > 0$ negativas</p>	$C(x, y) = \max\left(0, \left(y^p - (1-x)^p\right)^{\frac{1}{p}}\right)$ $D(x, y) = 1 - \max\left(0, \left((1-y)^p - x^p\right)^{\frac{1}{p}}\right)$
<p>Hamacher</p> $T(x, y) = \frac{xy}{x+y-xy}$ $S(x, y) = \frac{x+y-2xy}{1-xy}$	<p>$\frac{1-x}{x}$ $\frac{x}{1-x}$ positivas</p>	
<p>Hamacher</p> $T(x, y) = \frac{xy}{\gamma + (1-\gamma)(x+y-xy)}$ $S(x, y) = \frac{x+y-(2-\gamma)xy}{1-(1-\gamma)xy}$	<p>$\log\left(\frac{\gamma - (\gamma-1)x}{x}\right)$ $\log\left(\frac{(1-\gamma)x-1}{1-x}\right)$ $\gamma > 0$ positivas</p>	
<p>Frank</p> $T(x, y) = \log_s\left(1 + \frac{(s^x-1)(s^y-1)}{s-1}\right)$ $S(x, y) = 1 - \log_s\left(1 + \frac{(s^{1-x}-1)(s^{1-y}-1)}{s-1}\right)$	<p>$\log_s\left(\frac{s-1}{s^x-1}\right)$ $1 - \log_s\left(\frac{s-1}{s^{1-x}-1}\right)s$ > 0 negativas</p>	

<p style="text-align: center;"><i>Schweizer & Sklar</i></p> $T(x, y) = \max(0, x^{-p} + y^{-p} - 1)^{\frac{-1}{p}}$ $S(x, y) = 1 - \max(0, (1-x)^{-p} + (1-y)^{-p} - 1)^{\frac{-1}{p}}$	$\frac{1}{p} (x^{-p} - 1)$ $1 - \frac{1}{p} ((1-x)^{-p} - 1)$ <p>$p \geq 0$ positivas $p < 0$ negativas</p>	
<p style="text-align: center;"><i>Dubois & Prade</i></p> $T(x, y) = \frac{xy}{\max(x, y, \alpha)}$ $S(x, y) = \frac{x + y - xy - \min(x, y, 1 - \alpha)}{\max(1 - x, 1 - y, \alpha)}$	<p>sin generador $\alpha \in (0, 1)$ no arquimedeas</p>	
<p style="text-align: center;"><i>Dombi</i></p> $T(x, y) = \frac{1}{1 + \left(\left(\frac{1}{x} - 1 \right)^\lambda + \left(\frac{1}{y} - 1 \right)^\lambda \right)^{\frac{1}{\lambda}}}$ $S(x, y) = \frac{1}{1 + \left(\left(\frac{1}{x} - 1 \right)^{-\lambda} + \left(\frac{1}{y} - 1 \right)^{-\lambda} \right)^{\frac{-1}{\lambda}}}$	<p>$\lambda \in (0, 1)$ positivas</p>	

3.4.3 Modificadores lingüísticos

En la teoría de conjuntos borrosos a menudo se definen otros operadores que se suelen englobar en los llamados modificadores lingüísticos. Los principales son (véase por ejemplo [Zadeh, 73], [Pedrycz, 93]):

- Operadores de concentración

Estos operadores se caracterizan por reducir el cardinal del conjunto borroso al que se aplican, especialmente en lo referente a elementos con pertenencia parcial. Un ejemplo típico es la familia:

$$(CON(P))(u) = P(u)^\alpha \text{ con } \alpha > 1$$

- Operadores de dilatación

Estos operadores se caracterizan por aumentar el cardinal del conjunto borroso al que se aplican, especialmente en lo referente a elementos con pertenencia parcial. Un ejemplo típico es la familia:

$$(DIL(P))(u) = P(u)^\alpha \text{ con } \alpha < 1$$

- Operadores de intensificación

Estos operadores se caracterizan por aumentar los grados de pertenencia elevados y reducir los grados de pertenencia pequeños, por lo que tienden a generar un conjunto menos borroso que el argumento al que se aplican. Un ejemplo es:

$$(INT(P))(u) \begin{cases} 2P^2(u) & \text{si } P(u) < 0.5 \\ 1 - 2(1 - P(u))^2 & \text{otro caso} \end{cases}$$

- Antónimo

El antónimo puede definirse como:

$$(ANT(P))(u) = P(u_0 - u)$$

siendo u_0 el elemento central del conjunto U cuando este es finito, o un elemento que puede considerarse como punto de simetría cuando el conjunto U no es finito. Nótese que una definición precisa de antónimo plantea problemas no triviales como puede verse en [Rodríguez, 95].

3.4.4 El principio de extensión

Las operaciones que hemos visto hasta ahora eran generalizaciones de las operaciones clásicas entre conjuntos. El principio de extensión es por el contrario un procedimiento general que permite extender funciones $\varphi: U_1 \times U_2 \times \dots \times U_n \rightarrow V$ a funciones con argumentos borrosos, $\Phi: \mathfrak{F}(U_1) \times \mathfrak{F}(U_2) \times \dots \times \mathfrak{F}(U_n) \rightarrow \mathfrak{F}(V)$, donde $\mathfrak{F}(U)$ representa en general el conjunto de todos los conjuntos borrosos que pueden definirse en U , denominado a menudo partes borrosas de U . Recordemos que en contrapartida en la teoría clásica de conjuntos se denomina partes de U y se denota por $\mathcal{P}(U)$.

El principio de extensión se enuncia del siguiente modo. Dada una función $\varphi:U^n \rightarrow V$ la extensión Φ de φ viene dada por:

$$\Phi: \mathcal{I}(U_1) \times \mathcal{I}(U_2) \times \dots \times \mathcal{I}(U_n) \rightarrow \mathcal{I}(V)$$

$$\Phi(P_1, P_2, \dots, P_n)(v) = \sup \{ T(P(u_1), P(u_2), \dots, P(u_n)) \mid (u_1, u_2, \dots, u_n) \in U_1 \times U_2 \times \dots \times U_n \text{ y } v = \varphi(u_1, u_2, \dots, u_n) \}$$

$T(P(u_1), P(u_2), \dots, P(u_n))$ representa el grado conjunto de pertenencia de (u_1, u_2, \dots, u_n) a (P_1, P_2, \dots, P_n) , por lo que el principio de extensión puede entenderse como la expresión matemática de “*existe al menos un conjunto de valores (u_1, u_2, \dots, u_n) de $U_1 \times U_2 \times \dots \times U_n$ con grado de pertenencia conjunto $T(P(u_1), P(u_2), \dots, P(u_n))$ a (P_1, P_2, \dots, P_n) , tal que v es función de (u_1, u_2, \dots, u_n) según φ , es decir $v = \varphi(u_1, u_2, \dots, u_n)$ ”.*

Dicho de otro modo, si la imagen v por la función φ tiene más de un antecedente (u_1, u_2, \dots, u_n) , entonces el máximo grado de pertenencia de estos antecedentes al antecedente borroso (P_1, P_2, \dots, P_n) es el grado de pertenencia de v a la imagen borrosa de (P_1, P_2, \dots, P_n) por la función φ . El máximo (o supremo) en este contexto se interpreta como el cuantificador existencial \exists . Sin embargo el grado de pertenencia de cada tupla (u_1, u_2, \dots, u_n) a (P_1, P_2, \dots, P_n) se obtiene como una conjunción de los grados de pertenencia individuales de cada u_i al correspondiente conjunto P_i .

Una aplicación típica del principio de extensión es la extensión del álgebra ordinaria al álgebra borrosa. Otra aplicación importante es la extensión de las medias entre números reales para la obtención de nuevos operadores de agregación entre conjuntos borrosos, como puede verse en [Castro, Trillas, Zurita, 95], donde se utilizan estos operadores para proponer un sistema de razonamiento no monótono.

3.4.4.1 El álgebra borrosa

El álgebra borrosa es tanto una generalización del álgebra de números reales como del álgebra de intervalos, y permite extender, mediante el principio de extensión, las operaciones ordinarias entre números reales a los números borrosos. Por ejemplo las cuatro operaciones típicas de suma, resta, producto y cociente pueden definirse del siguiente modo:

$$\begin{aligned} \text{suma:} & \quad (P_1 + P_2)(v) = \sup_{v / v = u_1 + u_2} T(P_1(u_1), P_2(u_2)) \\ \text{resta:} & \quad (P_1 - P_2)(v) = \sup_{v / v = u_1 - u_2} T(P_1(u_1), P_2(u_2)) \\ \text{producto:} & \quad (P_1 \times P_2)(v) = \sup_{v / v = u_1 \times u_2} T(P_1(u_1), P_2(u_2)) \\ \text{cociente:} & \quad (P_1 \div P_2)(v) = \sup_{v / v = u_1 \div u_2} T(P_1(u_1), P_2(u_2)) \end{aligned}$$

donde P_1 y P_2 son conjuntos borrosos definidos en U y u_1, u_2 y v pertenecen a U .

El álgebra borrosa está siendo aplicada principalmente en temas de optimización borrosa e investigación operativa, análisis borroso de datos y modelado borroso. En estos trabajos suelen usarse funciones paramétricas sencillas, por ejemplo triangulares, de cara a obtener expresiones paramétricas exactas o aproximadas de los resultados de las operaciones borrosas, sin una excesiva carga computacional.

3.5 Relaciones borrosas

3.5.1 Definición

Dada una variable vectorial $\vec{u} = (u_1, u_2, \dots, u_n)$ definida en $U_1 \times U_2 \times \dots \times U_n$, una relación borrosa $R(u_1, u_2, \dots, u_n) = R(\vec{u})$ es un conjunto borroso caracterizado por su función de pertenencia:

$$R(u_1, u_2, \dots, u_n) = R(\vec{u}): U_1 \times U_2 \times \dots \times U_n \rightarrow [0, 1]$$

que expresa para cada valor \vec{u} de $U_1 \times U_2 \times \dots \times U_n$ el grado en que sus componentes están relacionadas. Cuando la relación se define como:

$$R(u_1, u_2): U \times U \rightarrow [0, 1]$$

se dice que R es una relación borrosa binaria en U .

Dado que una relación borrosa es un conjunto borroso definido sobre una variable vectorial, todas las definiciones y operaciones enunciadas para conjuntos borrosos siguen siendo aplicables a las relaciones borrosas.

3.5.2 Propiedades

En lo que sigue R es una relación borrosa binaria definida en $U \times U$.

3.5.2.1 Reflexividad

Se dice que R es reflexiva si $R(u, u) = 1 \quad \forall u \in U$.

3.5.2.2 Simetría

Se dice que R es simétrica si $R(u_1, u_2) = R(u_2, u_1) \quad \forall u_1, u_2 \in U$.

3.5.2.3 Antisimetría

Se dice que R es antisimétrica si $\forall u_1, u_2 \in U, u_1 \neq u_2$, entonces si $R(u_1, u_2) > 0$ se tiene $R(u_2, u_1) = 0$.

3.5.2.4 A-transitividad

Se dice que R es A-transitiva si $\forall u_1, u_2, u_3 \in U$, se cumple $A(R(u_1, u_2), R(u_2, u_3)) \leq R(u_1, u_3)$, y donde A es en general un operador conjuntivo.

3.5.2.5 Comparabilidad

Se dice que R verifica la propiedad de comparabilidad si $\forall u_1, u_2 \in U$, $u_1 \neq u_2$, entonces o bien $R(u_1, u_2) > 0$ o bien $R(u_2, u_1) > 0$.

3.5.2.6 Clasificación de relaciones según sus propiedades

La siguiente tabla muestra una clasificación habitual de las relaciones borrosas en función de las propiedades enunciadas hasta ahora:

	reflexividad	simetría	antisimetría	A-transitividad	comparabilidad
equivalencia, similitud o indistinguibilidad	sí	sí		sí	
compatibilidad o tolerancia	sí	sí			
orden parcial	sí		sí	sí	
orden lineal o total	sí		sí	sí	sí
orden parcial débil				sí	sí
A-preorden	sí			sí	

3.5.3 Proyección de una relación

Dada una relación borrosa $R(u_1, u_2, \dots, u_{j-1}, u_j, u_{j+1}, \dots, u_n)$ definida en $U_1 \times U_2 \times \dots \times U_{j-1} \times U_j \times U_{j+1} \times \dots \times U_n$ se define su proyección sobre el universo U_j como la relación borrosa R_{U_j} dada por:

$$R_{U_j}(u_1, u_2, \dots, u_{j-1}, u_{j+1}, \dots, u_n) = \sup_{u_j} R(u_1, u_2, \dots, u_{j-1}, u_j, u_{j+1}, \dots, u_n)$$

La proyección permite eliminar aquellas variables de una relación borrosa que no interesan de cara a obtener en qué grado están relacionadas el resto de variables de la relación. El supremo puede nuevamente interpretarse como el cuantificador existencial \exists , de forma que la proyección $R_W(u, v)$ de una relación $R(u, v, w)$ sobre el universo W se interpreta como que al menos existe un valor de w tal que $R_W(u, v) = R(u, v, w)$.

3.5.4 Extensión cilíndrica de una relación

Dada una relación $R: U_2 \times \dots \times U_n \rightarrow [0, 1]$ definida en el universo producto $U_2 \times \dots \times U_n$, a menudo interesa obtener una relación equivalente pero extendida a un universo producto mayor, como por ejemplo $U = U_1 \times U_2 \times \dots \times U_n$. Esta nueva relación se denomina extensión cilíndrica de R a U_1 , se denota como $R \uparrow_{U_1}$, y viene dada por:

$$R \uparrow_{U_1}: U_1 \times U_2 \times \dots \times U_n$$

$$R \uparrow_{U_1}(u_1, u_2, \dots, u_n) = R(u_2, \dots, u_n)$$

La relación $R \uparrow_{U_1}$ obtenida como extensión cilíndrica de R a U_1 es la mayor relación en U tal que su proyección sobre U_1 es igual a R .

3.5.5 Composición de relaciones borrosas

Dadas dos relaciones borrosas binarias R_1 definida en $U_1 \times U_2$ y R_2 definida en $U_2 \times U_3$, la composición de R_1 y R_2 permite conocer el grado de relación entre los elementos de U_1 y los elementos de U_3 . El tipo de composición viene determinado por el criterio con el que se quieran combinar los grados de relación de las distintas variables. Dos de los principales tipos de composición son la composición *sup-A* y la composición *inf-O*, donde A es un operador conjuntivo y O un operador disyuntivo. La relación R resultante de la composición según los tipos indicados tiene como función de pertenencia:

$$R(u_1, u_3) = \sup_{u_2} A(R_1(u_1, u_2), R_2(u_2, u_3))$$

$$R(u_1, u_3) = \inf_{u_2} O(R_1(u_1, u_2), R_2(u_2, u_3))$$

Entre estos dos tipos destacan por su uso más extendido las composiciones *sup-min* e *inf-max*. El supremo puede interpretarse en este contexto como el cuantificador existencial \exists , el ínfimo como el cuantificador universal \forall , y los operadores A y O como la conjunción y disyunción lógicas.

Existen otros tipos de composiciones. Por ejemplo podemos citar una poco conocida que sin embargo puede aparecer en el ámbito del razonamiento borroso:

$$R(u_1, u_3) = \inf_{u_2} I(R_1(u_1, u_2), R_2(u_2, u_3))$$

donde $I(x,y)$ es una función de implicación que generaliza la implicación material clásica. En [Dubois, Prade, 92b] se analiza la semántica de este tipo de implicaciones y se proponen posibles aplicaciones, por ejemplo al diagnóstico.

La composición de relaciones se generaliza sin dificultad a relaciones no binarias.

3.5.6 Algunas relaciones borrosas comunes

Algunas de las relaciones borrosas más utilizadas se obtienen combinando dos o más conjuntos borrosos mediante algún operador como son los operadores conjuntivos, disyuntivos, etc. Entre las relaciones así obtenidas cabe destacar las siguientes.

- Condicionales borrosos y funciones de implicación

Un condicional borroso $R(u,v)$ es una relación borrosa que pretende expresar el grado con que u implica a v para cada par (u,v) del universo $U \times V$. Los condicionales borrosos pueden obtenerse combinando dos conjuntos borrosos $P(u)$ y $Q(v)$ mediante una función $I(x,y): [0,1] \times [0,1] \rightarrow [0,1]$, llamada función de implicación, de modo que el condicional viene dado por:

$$R(u,v) = I(P(u), Q(v))$$

Un condicional así definido expresa de forma matemática la relación condicional que se genera entre u y v al enunciar la regla ‘si u es P entonces v es Q ’ o abreviadamente ‘ $P \rightarrow Q$ ’. Las funciones de implicación que se utilizan para obtener el condicional no son funciones cualesquiera ya que deben reunir un conjunto mínimo de propiedades, que serán analizadas en detalle en el capítulo 4. Señalaremos aquí las principales formas de generar funciones de implicación, las implicaciones residuales o r-implicaciones (r de residual) y las implicaciones fuertes o s-implicaciones (s procede de ‘strong’). Una r-implicación se define como:

$$I_R^T(x,y) = \sup\{c \in [0,1] / T(x,c) \leq y\}$$

donde T es una t-norma. Una implicación fuerte o s-implicación se define como:

$$I_S^T(x,y) = S(\Gamma(x), y) = \Gamma(T(x, \Gamma(y)))$$

donde S es la t-norma Γ -dual de T y Γ es una negación fuerte.

En aplicaciones numéricas y en concreto en el campo del control las funciones de implicación escogidas suelen ser operadores conjuntivos, y en particular t-normas, de forma que se tiene:

$$I(x,y)=T(x,y)$$

La siguiente tabla presenta algunas de las funciones de implicación más comunes, con el o los nombres con los que normalmente se conocen, y las abreviaturas que se han utilizado en esta tesis:

Nombre/s de la implicación	I(x,y)
Zadeh o Willmot	$\max(1-x, \min(x,y))$
Lukasiewicz	$I_S^L(x,y)=\min(1, 1-x+y)$
Godel-Brower o Godel	$I_R^{min}(x,y)=\begin{cases} 1 & \text{si } x \leq y \\ y & \text{en otro caso} \end{cases}$
Kleene-Dienes	$I_S^{min}(x,y)=\max(1-x,y)$
Reichenbach	$I_S^P(x,y)=1-x+x.y$
Menguer-Goguen o Gaines	$I_R^P(x,y)=\begin{cases} 1 & \text{si } x \leq y \\ \frac{y}{x} & \text{en otro caso} \end{cases}$
Yager	$I_Y(x,y)=y^x$
Mamdani	$T_m(x,y)=\min(x,y)$
Larsen	$T_P(x,y)=\text{prod}(x,y)=x.y$

- Producto cartesiano y operadores conjuntivos

El producto cartesiano entre dos conjuntos $P(u)$ y $Q(v)$ es la relación borrosa $R(u,v)$ definida como:

$$R(u,v)=(P \times Q)(u,v)=T(P(u),Q(v))$$

donde T es una t-norma. Cuando se quiere obtener el producto cartesiano de varios conjuntos P_1, P_2, \dots, P_n , la relación borrosa resultante viene dada por:

$$R(u_1, u_2, \dots, u_n) = T(P_1(u_1), P_2(u_2), \dots, P_n(u_n))$$

donde $T(x,y,z)$ se define a partir de la asociatividad de las t-normas (a partir de $T(x,T(y,z))$ la definición de T puede extenderse por recursividad a más argumentos). Como ya hemos mencionado el producto cartesiano (que no es un preorden al no ser reflexivo) es utilizado

frecuentemente como función de implicación en aplicaciones numéricas y de control. Su comportamiento será analizado en el capítulo 4.

Las relaciones borrosas basadas en el producto cartesiano surgen de forma natural cuando se estudian reglas con varios antecedentes, ya que estos se suelen agregar mediante un operador conjuntivo, generalmente una t-norma (la conmutatividad y asociatividad se asumen en este contexto como propiedades deseables) de forma que dada la regla “ $(P_1, P_2, \dots, P_n) \rightarrow Q$ ” el condicional que la modela matemáticamente viene dado por:

$$R(u_1, u_2, \dots, u_n, v) = I(T(P_1(u_1), P_2(u_2), \dots, P_n(u_n)), Q(v))$$

donde T es una t-norma.

- Equivalencias

Las equivalencias o relaciones de similitud o indistinguibilidad son otras de las relaciones borrosas más empleadas. Se pueden obtener como bi-implicaciones ($P \rightarrow Q$ y $Q \rightarrow P$) cuando la implicación a partir de la que se generan cumple ciertos requisitos básicos. También pueden obtenerse a partir de métricas generalizadas obtenidas a su vez a partir de métricas ordinarias. Las equivalencias han sido ampliamente estudiadas en [Trillas, Valverde, 85b], [Valverde, 85] y [Rodríguez, 95]. Son particularmente interesantes, desde un punto de vista teórico, para generar familias de conjuntos borrosos (por ejemplo los trapezoidales) como clases de equivalencia de la relación, de tal forma que el grado de pertenencia de un elemento al conjunto borroso de la familia generada se interpreta como el grado en que es indistinguible de los elementos del núcleo.

Una relación $E(x,y)$ es una relación de indistinguibilidad o equivalencia ([Trillas, Valverde, 85b], [Valverde, 85]) cuando verifica las propiedades:

$$\begin{aligned} E(x,x) &= 1 \\ E(x,y) &= E(y,x) \\ T(E(x,y), E(y,z)) &\leq E(x,z) \end{aligned}$$

Aunque las equivalencias se pueden obtener a partir de dos implicaciones, y por tanto pueden ser relevantes para ciertos esquemas de inferencia, en esta tesis nos limitaremos al estudio de las implicaciones.

3.6 Distribuciones de posibilidad

Una aplicación importante de la teoría de conjuntos borrosos consiste en la teoría de la posibilidad, donde el valor que toma una variable se modela mediante un conjunto borroso que permite expresar simultáneamente imprecisión, vaguedad e incertidumbre, tal y como se definieron estos conceptos en el capítulo 2.

3.6.1 Definición

Sea una variable u definida sobre un universo de discurso U . Una distribución de posibilidad π asociada a la variable u es una función $\pi: U \rightarrow [0,1]$, que expresa el conocimiento que se tiene de lo factible, admisible, creíble, posible (término que utilizaremos a lo largo de la tesis), etc., que es el que la variable u tome un determinado valor de U ([Dubois, Prade, 93]). Por tanto para cada valor u_0 de U el grado de posibilidad de que u tome dicho valor u_0 viene dado por $\pi(u_0)$.

Cuando el conocimiento que se tiene sobre los valores que la variable u puede tomar se expresa mediante un predicado vago del tipo “ u es P ” donde P es un conjunto borroso, entonces en este contexto la distribución de posibilidad coincide con P , de forma que la posibilidad de que u tome el valor u_0 viene dada por la compatibilidad del valor u_0 con el concepto expresado por P ([Zadeh, 78b]), es decir por la pertenencia de u_0 a P :

$$\pi(u_0) = P(u_0)$$

Nótese que en lenguaje natural se usan constantemente predicados vagos, “*temperaturas altas*”, “*gente joven*”, “*coche rápido*”, etc., expresados a través de propiedades que no admiten una transición brusca del cumplimiento al no cumplimiento de la misma. Por ello una forma natural de modelar estos predicados es mediante el uso de conjuntos borrosos. Cuando estos predicados se utilizan para describir el valor que toma una variable entonces se interpretan como distribuciones de posibilidad. Esta es una de las principales razones por las que los conjuntos borrosos constituyen una herramienta especialmente sugerente para la representación del conocimiento.

3.6.2 Especificación de una distribución de posibilidad

La información que se tiene de una variable no siempre se puede expresar como “ u es $\pi(u)$ ”, ya que en muchos casos la información de que se dispone es insuficiente y no

se conoce con precisión la distribución de posibilidad de la variable (esto es especialmente cierto en el ámbito del razonamiento aproximado). Por ello, como se muestra en [Dubois, Prade, 93], la distribución de posibilidad de una variable puede ser parcialmente especificada mediante la cualificación de un conjunto borroso $P(u)$ en términos de posibilidad o en términos de necesidad, como a continuación se describe. Esta cualificación en términos de posibilidad o necesidad permite distinguir dos tipos de distribuciones notablemente distintas, aportación original de esta tesis, y que hemos denominado distribuciones de posible posibilidad y distribuciones de necesaria posibilidad:

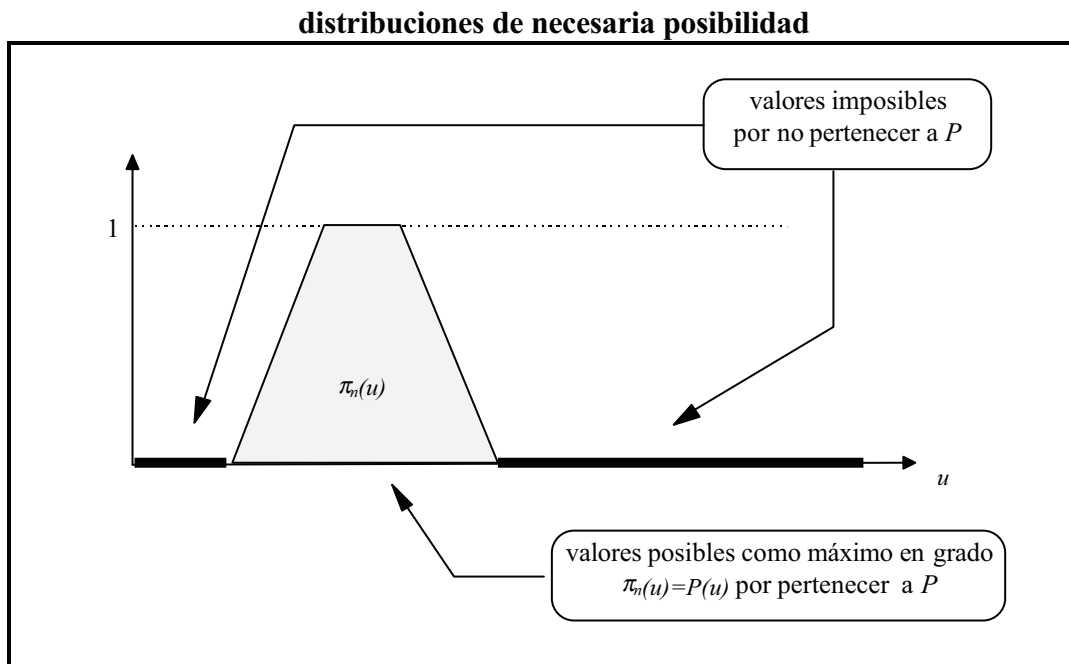
- Cualificación en términos de necesidad

“es cierto/seguro/necesario que u esté en P ”, o bien “ $P(u)$ es un rango cierto/seguro/necesario para u ” se interpreta como que el conjunto P excluye sólo a valores imposibles para u , por lo que $P(u)$ resulta ser una cota superior de $\pi(u)$:

$$\pi(u) \leq P(u)$$

En este caso diremos que $P(u)$ es una **distribución de necesaria posibilidad**, que denotaremos como $\pi_n(u)=P(u)$, y $\pi_n(u_0)$ lo interpretaremos como el grado máximo de posibilidad de que la variable u tome el valor u_0 .

La siguiente figura ilustra el significado de este tipo de distribuciones:



- Cualificación en términos de posibilidad

“es posible que u esté en P ”, o bien “ $P(u)$ es un rango posible para u ” se interpreta como que el conjunto P incluye sólo a valores considerados posibles para u , pero no se descarta que otros elementos de U sean también posibles, o que incluso elementos de $P(u)$ puedan ser imposibles. En [Dubois, Prade, 93] se dice que un rango posible para la variable u verifica, respecto de la distribución de posibilidad $\pi(u)$ de la variable, la siguiente desigualdad:

$$P(u) \leq \pi(u)$$

con lo que $P(u)$ se interpreta como una cota inferior de la distribución de posibilidad de u . Para nosotros $P(u_0)$ será interpretado de una forma ligeramente diferente, esto es, como un posible grado de posibilidad de que u tome el valor u_0 , y estrictamente hablando no será considerado ni como una cota inferior ni como una cota superior. Su interpretación dependerá del resto de información de que se disponga sobre la misma variable u . En este caso diremos que $P(u)$ es una **distribución de posible posibilidad** para la variable u que denotaremos por $\pi_p(u)=P(u)$.

En la desigualdad anterior la distribución de posibilidad $\pi(u)$ es en general desconocida, y $\pi_p(u)$ sólo intenta ser una aproximación razonable a la misma. Como ya hemos señalado para nosotros $\pi_p(u)$ no será una cota inferior de $\pi(u)$ ya que la distribución de posibilidad ideal que mejor restringe el valor que la variable u puede tomar es un conjunto unitario $\{u_0\}$ donde u_0 es el verdadero valor que toma la variable u , por lo que la desigualdad $P(u) \leq \pi(u)$ podría no verificarse nunca, a menos que P coincidiese con $\{u_0\}$.

Por otro lado la existencia de una distribución de necesaria posibilidad para la misma variable obliga a que se verifique la desigualdad:

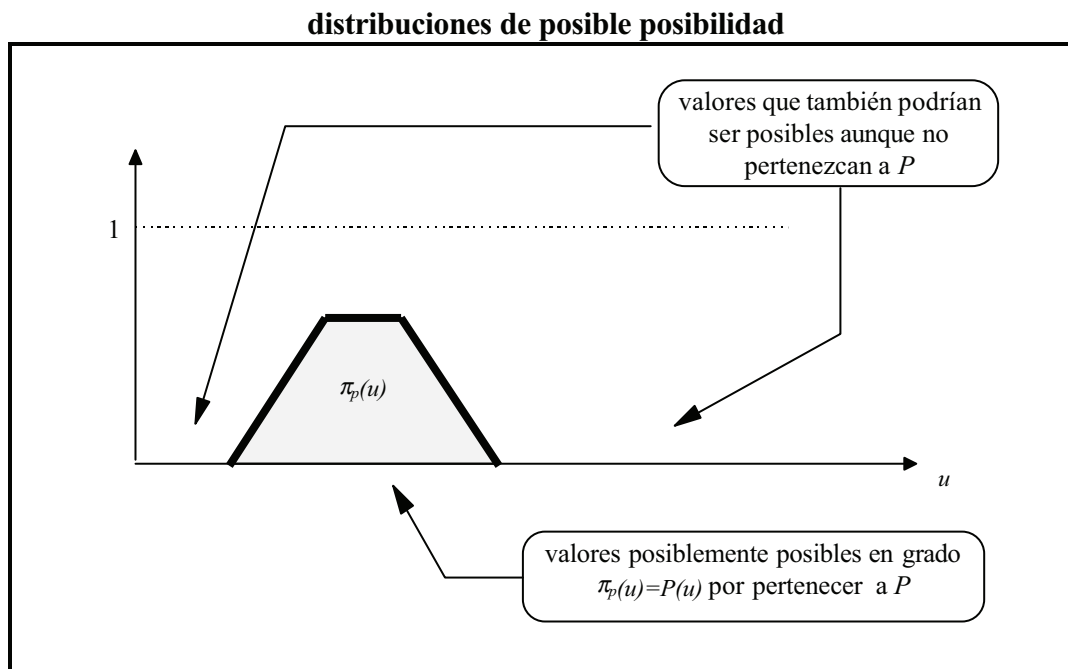
$$\pi_p(u) \leq \pi_n(u)$$

ya que $\pi_n(u)$ sí es el máximo grado de posibilidad de u , y es por tanto un límite rígido para el grado de posibilidad de u que no puede ser transgredido. Esta desigualdad será para nosotros la restricción fundamental a la que una distribución de posible posibilidad π_p debe estar sometida.

En definitiva, aunque grado de posibilidad de u venga dado por $\pi_p(u)$ también es posible que su grado de posibilidad final sea mayor o menor. Será menor si una distribución de necesaria posibilidad lo limita por arriba, y como veremos a continuación, será mayor si otra distribución de posible posibilidad lo considera más posible que la primera. Mientras que la restricción impuesta por $\pi_n(u)$ es una restricción cierta y por tanto definitiva, $\pi_p(u)$

es una restricción débil, que intenta aproximar en base a criterios subjetivos, el grado de posibilidad de cada elemento u de U . Debido a su certidumbre cualitativamente baja, no puede considerarse como una cota superior o inferior de posibilidad.

La siguiente figura ilustra gráficamente la interpretación de este tipo de distribuciones:



- Cualificación en términos de posibilidad y necesidad

Si “ $P(u)$ es un rango posible y necesario para u ”, o bien “ u es P ”, entonces el conjunto P sólo incluye a valor posibles y sólo excluye a valores imposibles, y coincide con la distribución de posibilidad de u :

$$\pi(u)=P(u)$$

En esta caso las distribuciones de posible y necesaria posibilidad coinciden y son a su vez iguales a la **distribución de posibilidad** de la variable u . Es necesario resaltar que afirmar que el conjunto P sólo incluye a valor posibles y sólo excluye a valores imposibles es una afirmación contextual que depende fuertemente de la información que se tenga. Así por ejemplo en una situación de total información, P tiene que reducirse a un conjunto con un único elemento, ya que son imposibles todos aquellos valores distintos del valor real de la variable, que en este caso se supone conocido.

Se dice a menudo que si π es una distribución de posibilidad para la variable u , también lo es cualquier distribución π_n tal que $\pi_n(u) \geq \pi(u)$ aunque sea menos específica que π y por tanto menos informativa (véase por ejemplo [Dubois, Prade, 91]). Según la distinción que hemos hecho en esta tesis entre las distribuciones de posible y necesaria posibilidad, deberíamos decir que, dada una distribución de posibilidad $\pi(u)$ cualquier distribución que verifique $\pi_n(u) \geq \pi(u)$ es una distribución de necesaria posibilidad para la variable u . De forma más general puede decirse que si $\pi_n(u)$ es una distribución de necesaria posibilidad para u entonces cualquier distribución π_n' que verifique $\pi_n'(u) \geq \pi_n(u)$ también lo es. Esta desigualdad es la restricción fundamental a la que está sometida cualquier distribución de necesaria posibilidad para la variable u , respecto a una distribución de necesaria posibilidad $\pi_n(u)$ tomada como referencia.

Del mismo modo cualquier distribución π_p tal que $\pi_p(u) \leq \pi_n(u)$ es una distribución de posible posibilidad para dicha variable, siendo esta desigualdad la restricción fundamental impuesta a las distribuciones de posible posibilidad respecto a una distribución de necesaria posibilidad $\pi_n(u)$ tomada como referencia. La coexistencia de distribuciones de necesaria y posible posibilidad en un mismo entorno se tratará en detalle en el capítulo 5.

3.6.3 Agregación de distribuciones de posibilidad

3.6.3.1 Distribuciones definidas en un único universo

- Agregación de distribuciones:

Dadas dos distribuciones de necesaria posibilidad π_{n1} y π_{n2} , dado que ambas pueden ser interpretadas como cotas superiores del grado de posibilidad real de la variable u , se verifica:

$$\pi_{n1}(u) \geq \pi_n(u) \quad \text{y} \quad \pi_{n2}(u) \geq \pi_n(u)$$

respecto a una distribución de necesaria posibilidad de referencia, por lo que parece natural combinarlas con el operador *min*:

$$\pi_{n12}(u) = \min(\pi_{n1}(u), \pi_{n2}(u)) \geq \pi_n(u)$$

Así se obtiene la distribución de necesaria posibilidad más específica posible, pero no contradictoria con la información de partida ya que verifica también la restricción fundamental respecto a la distribución de referencia $\pi_n(u)$. Nótese que si se dispone de información adicional que permita relacionar $\pi_{n1}(u)$ y $\pi_{n2}(u)$ entonces es posible que otros

operadores conjuntivos menores sean más apropiados. Este punto se tratará en el apartado 5.3.

Por otro lado, dadas dos distribuciones de posible posibilidad π_{p1} y π_{p2} para una variable u , dado que ambas verifican la restricción fundamental respecto a una distribución de necesaria posibilidad $\pi_n(u)$ tomada como referencia, se verifica:

$$\pi_{p1}(u) \leq \pi_n(u) \quad \text{y} \quad \pi_{p2}(u) \leq \pi_n(u)$$

por lo que parece natural combinarlas con el operador *max*:

$$\pi_{p12}(u) = \max(\pi_{p1}(u), \pi_{p2}(u)) \leq \pi_n(u)$$

obteniéndose de este modo la mayor distribución de posible posibilidad compatible con cada una de las restricciones. Al igual que en el caso anterior, información adicional que permita relacionar $\pi_{p1}(u)$ y $\pi_{p2}(u)$ puede aconsejar el uso de otros operadores disyuntivos mayores (véase el apartado 5.3 del capítulo 5).

Esta distinción formal entre los distintos tipos de distribuciones de posibilidad permite explicar por qué las distribuciones de posibilidad (englobando en este caso a ambos tipos) se agregan algunas veces con operadores conjuntivos, otras con operadores disyuntivos, y otras con algún tipo de media u operador de agregación intermedio entre los conjuntivos y disyuntivos. A este respecto, en [Dubois, Prade, 88b] se estudian varios tipos de agregación para dos distribuciones, ambas igualmente fiables (o con igual certidumbre), para una misma variable, dando lugar a los siguientes casos:

- si no son fiables deben agregarse mediante un operador disyuntivo. En este caso, según nuestra interpretación, las distribuciones de posibilidad pueden interpretarse como distribuciones de posible posibilidad (el término posible indica ciertamente la escasa fiabilidad o certidumbre de la información).
- si son fiables deben agregarse mediante un operador conjuntivo. En este caso las distribuciones de posibilidad pueden interpretarse como distribuciones de necesaria posibilidad. Nótese la intuitiva proximidad entre el concepto de necesario y fiable.
- cuando se realiza una media, se asume que son más posibles aquellos elementos que pertenecen a la intersección de las dos distribuciones originales, que aquellos que pertenecen o bien a una o bien a otra, pero sin rechazar por completo estos últimos. En este caso las distribuciones de posibilidad podrían interpretarse como distribuciones de ϕ -posibilidad, tal y como se describen en el apéndice de la tesis, donde se presentan las ϕ -implicaciones y los ϕ -condicionales, y donde se propone un nuevo tipo de distribuciones de posibilidad, las distribuciones de ϕ -posibilidad intermedias entre las de

necesaria y posible posibilidad. En dicho apéndice se sugiere que estas distribuciones deben combinarse mediante operadores de agregación $H_\varphi(x,y)$ tales que $H_\varphi(1,1)=1$, $H_\varphi(0,1)=H_\varphi(1,0)=H_\varphi(\varphi,\varphi)=\varphi$, y $H_\varphi(0,0)=0$, entre los que la media aritmética resulta ser un caso particular para $\varphi=0.5$.

- Tipos de distribuciones y principios de agregación

El principio de mínima especificidad afirma que dadas varias distribuciones de posibilidad todas fiables, y en ausencia de información adicional, la distribución de posibilidad resultante se obtiene agregándolas todas ellas mediante el operador conjuntivo $A(x,y)=\min(x,y)$ ([Dubois, Prade, 91]).

Como hemos visto, las distribuciones de necesaria posibilidad están sujetas al mismo principio, mientras que las distribuciones de posible posibilidad (definidas en un mismo universo) lo contradicen puesto que por defecto se agregan mediante el operador max .

Esto sugiere la necesidad de encontrar un principio más general que englobe al anterior, pero que sea aplicable a cualquier tipo de distribución de posibilidad. Haciendo uso del sentido común es posible enunciar un principio básico y general para la agregación de piezas de información. En efecto la agregación de dos piezas de información debe someterse a dos restricciones simultáneas, por un lado debe conducir a la obtención de una nueva pieza de información que sea lo más informativa posible, es decir que aproveche lo mejor posible toda la información disponible, y por otro, la nueva información así obtenida debe tener la mayor certidumbre posible, compatible con las certidumbres de las distintas piezas de información de partida. Este principio podría denominarse principio de máxima información y máxima certidumbre, y es por su generalidad aplicable a cualquier tipo de distribución. A este respecto puede consultarse [Turksen, Tian, 93] donde se propone como principio alternativo al de mínima especificidad el principio de máxima información, aplicable a lo que nosotros hemos denominado distribuciones de posible posibilidad.

En el caso de las distribuciones de necesaria posibilidad la distribución más informativa que no conlleve una pérdida de certidumbre viene dada por:

$$\pi_{n12}(u) = \min(\pi_{n1}(u), \pi_{n2}(u))$$

debido a que estamos agregando dos cotas superiores ambas totalmente fiables. El uso de operadores mayores que min supondría una pérdida de información al obtenerse una distribución menos específica. Por el contrario el uso de un operador menor no daría lugar a una distribución de necesaria posibilidad con el consiguiente decremento cualitativo de la certidumbre de la información final obtenida, dada la información contenida en las

distribuciones π_{n1} y π_{n2} de partida. El uso de un operador menor solo se podría justificar en base a alguna información adicional que permitiese restringir aún más los grados de posibilidad, información que no se deriva de π_{n1} y π_{n2} por sí solas (véase el apartado 5.3 del capítulo 5). Nótese que el nombre de principio de mínima especificidad indica que la agregación con el operador *min* es la agregación conjuntiva menos específica, pero bajo el supuesto de que la agregación debe hacerse conjuntivamente. Por tanto el principio que estamos enunciando es compatible con el de mínima especificidad.

En el caso de las distribuciones de posible posibilidad, la distribución más informativa que sin pérdida de certidumbre puede obtenerse viene dada por:

$$\pi_{p12}(u) = \max(\pi_{p1}(u), \pi_{p2}(u))$$

Mantener la certidumbre de las distribuciones de partida exige obtener como resultado de la agregación una nueva distribución de posibilidad, lo que obliga a utilizar un operador menor o igual que el operador *max*. Por otro lado la distribución de posible posibilidad más informativa es en este caso la mayor, puesto que la distribución de posible posibilidad trivial es el conjunto vacío. Visto de otro modo, se obtiene la mayor información posible incluyendo en la distribución final a todos aquellos elementos considerados posibles por las distribuciones de partida, y excluir a alguno de ellos sería ignorar parte de la información disponible. El uso de un operador mayor que *max* solo se justifica, como en el caso anterior, por la existencia de información adicional. Puesto que la nueva distribución obtenida debe ser una distribución de posible posibilidad debe verificarse la restricción fundamental:

$$\pi_{p12}(u) = \max(\pi_{p1}(u), \pi_{p2}(u)) \leq \pi_n(u)$$

que está garantizada, puesto que de forma individual también se verifica:

$$\begin{aligned} \pi_{p1}(u) &\leq \pi_n(u) \\ \pi_{p2}(u) &\leq \pi_n(u) \end{aligned}$$

Es posible que en algunos casos particulares en los que las distribuciones de posible y necesaria posibilidad se obtengan por caminos diferentes, se llegue a un punto en el que no se verifique la desigualdad $\pi_p(u) \leq \pi_n(u)$. En este caso parece razonable que la distribución $\pi_p(u)$ sea corregida del siguiente modo:

$$\pi_{p-final}(u) = \max(\pi_p(u), \pi_n(u))$$

que sencillamente garantiza que ningún elemento u de U pueda tener un posible grado de posibilidad mayor que su cota máxima. Como ya hemos comentado el uso de otras

t-normas o t-conormas no son una violación del principio enunciado sino precisamente el resultado de su aplicación al tener en cuenta la existencia de información adicional.

3.6.3.2 Distribuciones definidas en universos distintos

Cuando las distribuciones a agregar están definidas para variables distintas, y por tanto definidas sobre universos también distintos, el resultado de la agregación no es un conjunto borroso sino una relación borrosa. En este caso la agregación no depende del tipo de distribución ya que la información sobre una de las variables no aporta ninguna información sobre la otra, y es el carácter conjuntivo de la variable conjunta (u, v) el que hace que su distribución deba obtenerse siempre de forma conjuntiva. La distribución conjunta es la distribución de posible o necesaria posibilidad que mejor describe a la variable (u, v) , y esto en general supone siempre una pérdida de información para alguna de las dos variables.

Independientemente del carácter de las distribuciones que en este caso denotaremos por $P(u)$ y $Q(v)$, la distribución de posibilidad conjunta $\pi(u, v)$ es el grado de posibilidad del par $\pi(u, v)$, lo que significa que $\pi(u_0, v_0)$ es el grado de posibilidad (máximo o posible) de que u tome el valor u_0 y v tome el valor v_0 , ambas simultáneamente.

Si estamos tratando con distribuciones de necesaria posibilidad, $\pi_n(u_0, v_0)$ es el máximo grado de posibilidad de que la variable u tome el valor u_0 y simultáneamente la variable v tome el valor v_0 . Por tanto el valor (u_0, v_0) no puede ser más posible que u_0 o que v_0 por lo que la agregación lógica a emplear es el mínimo de las dos distribuciones iniciales, obteniéndose:

$$\pi_n(u, v) = \min(P(u), Q(v)).$$

Del mismo modo cuando P y Q son distribuciones de posible posibilidad, $\pi_p(u_0, v_0)$ es una estimación del grado de posibilidad de que la variable (u, v) tome el valor (u_0, v_0) , y la estimación más razonable para (u_0, v_0) viene otra vez dada por:

$$\pi_p(u, v) = \min(P(u), Q(v)).$$

La obtención de la distribución conjunta puede justificarse a partir del mismo principio de máxima información y máxima certidumbre. Tanto la agregación \min como \max suponen una pérdida de información puesto que en general no es posible recuperar las distribuciones iniciales de las variables u y v . Sin embargo la máxima certidumbre obliga a realizar una agregación conjuntiva, ya que la ocurrencia de dos hechos simultáneos no puede ser más posible o necesaria (o incluso probable) que la ocurrencia de cada uno de

ellos por separado, puesto que conlleva un mayor nivel de exigencia. Esto hace que la agregación deba ser siempre de tipo conjuntiva, y como ya ha sido visto, la agregación conjuntiva de grados de posibilidad, en ausencia de información adicional, debe hacerse con el operador *min*.

Nótese que también sería posible obtener la posibilidad de que la variable u tome el valor u_0 o de que la variable v tome el valor v_0 y vendría dada por $\max(P(u_0), Q(v_0))$ independientemente del tipo de distribuciones involucradas, pero la distribución así obtenida no puede calificarse de distribución conjunta.

3.6.4 Modelado de la verdad, imprecisión, vaguedad e incertidumbre mediante distribuciones de posibilidad y conjuntos borrosos

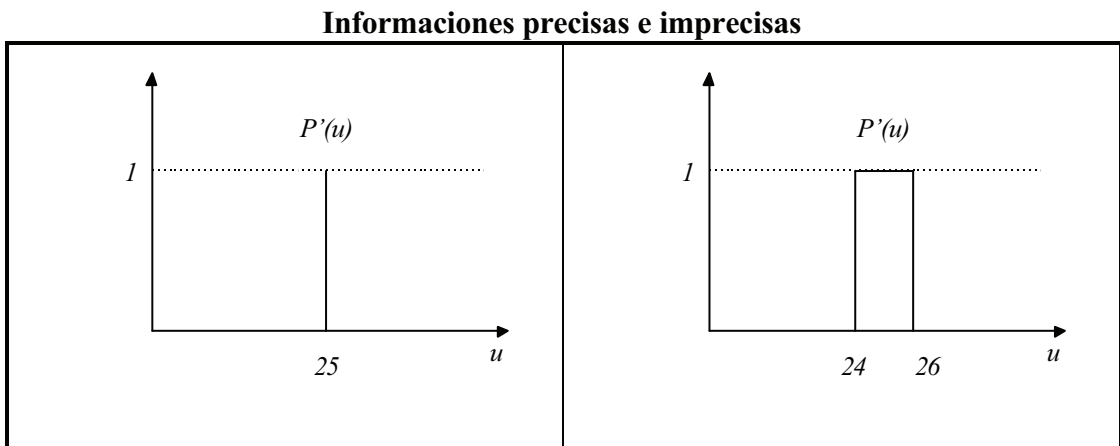
La distinción conceptual existente entre los distintos tipos de certidumbre (o incertidumbre), en su sentido más amplio, ya ha sido comentada en el capítulo 2, en el apartado 2.3.2. Resumiendo lo allí comentado, la verdad de una afirmación puede entenderse como la conformidad, compatibilidad o parecido (que podría ser parcial) entre su significado y la información disponible, su imprecisión es la falta de especificidad en la información que contiene y su incertidumbre es el grado de desconocimiento de su grado de verdad (aunque sea parcial). Así cuanto más imprecisa es una afirmación, menor será la incertidumbre que se tiene de que sea cierta. Por otro lado una información es vaga cuando lo que afirma sobre un conjunto de objetos puede ser verificado parcialmente por algunos de ellos, y se debe a la imposibilidad de establecer fronteras o distinciones rígidas entre categorías diferentes. La vaguedad recordemos proviene de la existencia de grados de verdad parciales.

Una de las importantes ventajas de los conjuntos borrosos y las distribuciones de posibilidad es que permiten manejar en un mismo marco teórico todos los tipos de incertidumbre mencionados. A continuación describimos las distintas situaciones que pueden presentarse y cómo cada una de ellas puede modelarse con una distribución de posibilidad.

- Imprecisión

La imprecisión de una observación se traduce en la existencia de más de un único valor posible para la medida que se observa. Ello hace que se pueda modelar mediante un conjunto con un cardinal mayor que 1, es decir con más de un elemento (lógicamente esto

sólo es válido para variables disyuntivas, que toman un único valor simultáneamente). Nótese que tanto la precisión como la imprecisión pueden expresarse en el ámbito clásico, puesto que la información precisa se corresponde con un conjunto clásico con un único elemento, mientras que la imprecisa con un conjunto clásico con más de un elemento (un conjunto borroso modela tanto la imprecisión como la vaguedad, como veremos a continuación). La siguiente figura ilustra la diferencia entre una información precisa e imprecisa:

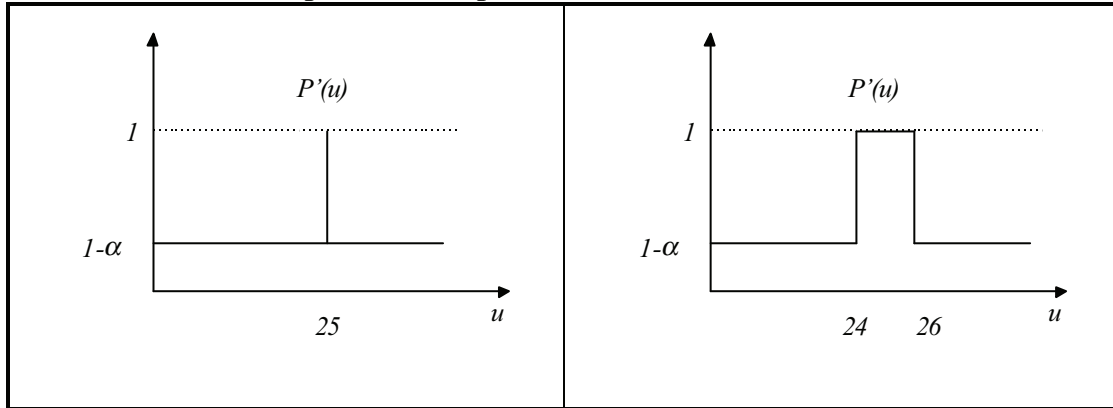


- Incertidumbre en los dominios de lo posible y lo necesario

La incertidumbre es el desconocimiento del grado de verdad de una afirmación. Las distribuciones de posibilidad permiten modelar dos tipos de incertidumbre: incertidumbre sobre la verdad de una afirmación (incertidumbre en el dominio de lo necesario), y la incertidumbre sobre la posibilidad de que una afirmación sea verdadera (incertidumbre en el dominio de lo posible). Una afirmación es verdadera cuando no existe posibilidad de que su negación lo sea, y se entiende que una afirmación es posible que sea verdadera, o sencillamente posible, cuando no existen contradicciones entre la afirmación y los hechos observados.

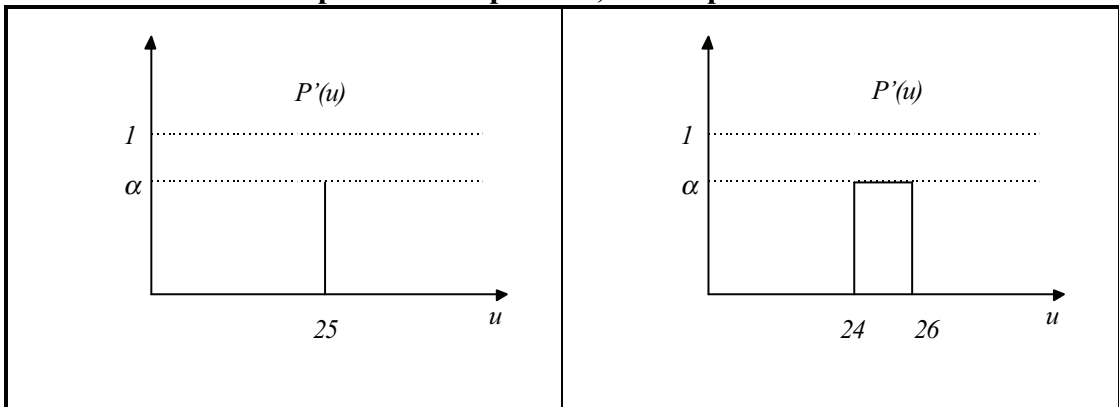
El primer tipo de incertidumbre se puede modelar con una distribución de necesaria posibilidad con un nivel de indeterminación distinto de 0. De este modo, aunque haya elementos del universo de discurso totalmente posibles, cualquier elemento del universo de discurso tiene un grado de posibilidad no nulo. La afirmación “la variable u toma sus valores en P con certidumbre α ” se modela habitualmente con la distribución de posibilidad $P'(u)=\max(P(u), 1-\alpha)$ (véase por ejemplo [Dubois, Prade, 91]), de forma que ningún valor del universo queda excluido con total certidumbre, y el máximo grado de exclusión es $1-\alpha$ (nótese que este tipo de incertidumbre no puede modelarse en el ámbito clásico). La siguiente figura ilustra el modelado de este tipo de incertidumbre:

Informaciones precisas e imprecisas, ambas necesariamente inciertas



El segundo tipo de incertidumbre puede modelarse mediante distribuciones de posible posibilidad con una altura inferior a 1. Esto significa que la información disponible no permite afirmar siquiera que existan valores del todo posibles, o lo que es lo mismo que hay un cierto grado de contradicción entre la afirmación y las observaciones, por lo que no es del todo posible que la afirmación sea verdadera. Sin embargo no debe interpretarse como la inexistencia de valores totalmente posibles para la variable en cuestión, sino que el grado de contradicción existente no permite identificarlos. Sin embargo ningún elemento del universo de discurso es rechazado definitivamente. La afirmación “*es posible en grado α que la variable u tome sus valores en P* ” se modela habitualmente con la distribución de posibilidad $P'(u)=\min(P(u),\alpha)$ ([Dubois, Prade, 91]). Este segundo tipo de incertidumbre lo denominaremos incertidumbre posibilística o incertidumbre en el dominio de lo posible para diferenciarlo del primero, y da lugar a distribuciones de posible posibilidad. Nótese además que dada una distribución de posible posibilidad para una variable u , aun cuando exista algún valor de la variable no del todo posible, este podrá posteriormente llegar a serlo si aparece un nuevo aporte de información en forma de nuevas distribuciones de posible posibilidad. La siguiente figura ilustra el modelado de este tipo de incertidumbre.

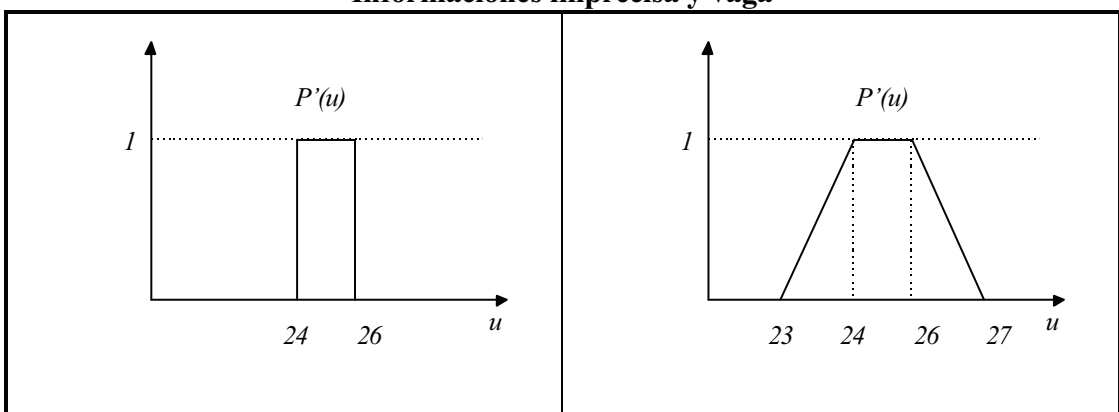
Informaciones precisas e imprecisas, ambas posiblemente inciertas



- Vaguedad y verdad parcial

La vaguedad de una observación radica en la existencia de elementos del universo de discurso de la variable observada sobre los cuales no se puede afirmar con exactitud si son totalmente posibles o totalmente imposibles. Esto da lugar a que existan, además de grados de posibilidad pertenecientes a $\{0,1\}$, grados de posibilidad intermedios en $[0,1]$. La observación no se define de forma categórica sino matizada, de forma que suele existir una transición gradual entre los elementos del universo de discurso totalmente posibles y los totalmente imposibles. La vaguedad conlleva siempre un cierto grado de imprecisión puesto que supone la existencia de más de un elemento del universo de discurso como posible candidato (en un grado de posibilidad que puede ser parcial) para el valor de la variable. La siguiente figura ilustra las diferencias entre una afirmación imprecisa y una afirmación vaga (y por tanto también imprecisa):

Informaciones imprecisa y vaga



La vaguedad conlleva la existencia de cierta incertidumbre e imprecisión. La imprecisión por un lado está principalmente asociada con la existencia de varias alternativas todas ellas con grados de posibilidad altos y es modelable en el ámbito clásico. La incertidumbre por otro lado está relacionada, en el dominio de lo necesario, con la imposibilidad de excluir totalmente a la mayoría de las alternativas del universo de discurso, o en el dominio de lo posible, con la imposibilidad de obtener alternativas totalmente posibles, debido a la existencia de contradicciones parciales. Otros tipos de incertidumbre como la incertidumbre probabilística pueden basarse en otras consideraciones como la frecuencia con que un determinado elemento ha sido observado con anterioridad, pero en cualquiera de los casos ninguna de ellas puede modelarse con conjuntos clásicos. Pueden existir afirmaciones vagas con un cierto contenido de incertidumbre, ya sea en el dominio de lo seguro o necesario como en el dominio de lo posible (y también de lo probable). Mientras que en la teoría de conjuntos borrosos la incertidumbre puede modelarse con conjuntos borrosos, la inclusión de probabilidades u otras técnicas de representación de incertidumbre requiere en general utilizar otro formalismo matemático adicional (que aunque con distinta interpretación y posiblemente más apropiado para algunos problemas concretos, supone un nivel de complejidad adicional).

La distinción propuesta para distribuciones de posible y necesaria posibilidad no permite combinar simultáneamente los dos tipos de incertidumbre señalados, en el dominio de lo necesario y en el dominio de lo posible. Una introducción a un enfoque propuesto en esta tesis para la combinación simultánea de ambas incertidumbres en una misma distribución se aborda desde un punto de vista teórico en el apéndice dedicado al estudio de los φ -condicionales.

3.6.5 Algunas distribuciones de posibilidad particulares

Cuando por algún tipo de procedimiento, por ejemplo mediante la inferencia a partir de un conjunto de reglas, obtenemos una distribución de necesaria posibilidad π_n para una variable objeto de estudio u , pueden presentarse tres situaciones cualitativamente distintas:

- $\pi_n(u)=U$: la distribución de necesaria posibilidad es igual al universo de discurso. Eso implica una falta absoluta de información sobre la variable en cuestión, y es equivalente a una distribución de necesaria posibilidad cualquiera con un grado de certidumbre nulo.
- $\pi_n(u)\neq U$, y $\exists u/\pi_n(u)=1$: la distribución de necesaria posibilidad está normalizada pero no es igual al universo de discurso de la variable. Esta situación se corresponde con la más habitual, en la que se tiene una información fiable coherente, que permite graduar el

grado máximo en que son posibles las alternativas de que consta el universo de discurso, entre las que deben siempre existir algunas totalmente posibles.

- $\forall u \pi_n(u) < 1$: la distribución de necesaria posibilidad no está normalizada. Este caso se corresponde con una inconsistencia de la información utilizada para obtener la distribución. En efecto la variable tiene necesariamente que tomar algún valor del universo de discurso, por lo que no es admisible que todos tengan un máximo grado de posibilidad inferior a 1. Una distribución de necesaria posibilidad no normalizada significa que todos los valores del universo de discurso son parcialmente imposibles, lo que obviamente es una contradicción. En el caso por ejemplo de un sistema de reglas, una inconsistencia de este tipo se arrastra a lo largo de todo el razonamiento, ya que si una regla determina que ningún valor es del todo posible, ninguna otra regla, debido a la forma en que estas distribuciones se agregan, puede alterar este resultado.

Cuando de lo que se trata es de una distribución de posible posibilidad π_p la situación es distinta:

- $\pi_p(u) = 0 \forall u$: la distribución de posible posibilidad es el conjunto vacío. Esta distribución refleja la existencia de una inconsistencia transitoria, en el sentido en que no parece lógico que todas las alternativas del universo de discurso estén en contradicción con la información de que se dispone, pero transitoria, ya que nueva información puede aumentar cualquiera de los grados de posibilidad de los elementos de u . Significa por tanto que la información no aporta ninguna evidencia en favor de una u otra alternativa, y que no se puede concluir nada, o dicho de otro modo, todas las alternativas del universo de discurso se consideran a priori contradictorias con la información de que se dispone siendo imposible escoger entre ellas. Esto evidencia lagunas en el conocimiento, en el sentido en que o bien el conjunto de alternativas disponible para la variable u no es suficiente, o bien el conocimiento disponible es incompleto, puesto que es insuficiente para obtener al menos una alternativa no contradictoria con los hechos.
- $\pi_p(u) \neq U$ y $\exists u / \pi_p(u) > 0$: la distribución de posible posibilidad tiene una altura no nula, pero es diferente del universo de discurso. Es la situación más habitual y se corresponde con la existencia de cierta evidencia en favor de algunas de las alternativas (o ciertas contradicciones que limitan su grado de posibilidad). Nótese que cuando alguna alternativa tenga un grado de posibilidad unitario lo mantendrá aunque nuevas informaciones permitan obtener distribuciones de posible posibilidad adicionales para la misma variable, debido a la forma en que se combinan. Una distribución no normalizada refleja la existencia de evidencia parcial en favor de determinadas alternativas lo que permite al menos establecer preferencias, pero como en el caso anterior, también refleja lagunas parciales en el conocimiento disponible.

- $\pi_p(u)=U$: es un caso extremadamente raro. Significa que en general el conocimiento que se tiene considera totalmente posibles a todos los elementos de U . En sistemas de reglas por ejemplo, esto podría darse si se activasen simultáneamente varias reglas tales que la unión de sus conclusiones fuese igual al universo de discurso. La adición de nuevas reglas no modificaría la situación, y la información que aporta una distribución de este tipo significa que la evidencia disponible apoya a todas las alternativas del universo de discurso por igual. En cierto modo, visto en el dominio de lo posible, significa que cualquier valor es una opción posiblemente admisible para la variable u .

Nótese por tanto que inconsistencias lógicas sólo se detectan con distribuciones de necesaria posibilidad. Estas inconsistencias se corresponden con contradicciones en el conocimiento, ya que el aporte de nuevo conocimiento no es capaz de eliminarlas. Esto ya ha sido señalado por ejemplo en [Dubois, Prade, Ughetto, 96] donde se sugiere la necesidad de modelar las reglas con implicaciones obtenidas como generalización del condicional material, si se quiere poder comprobar la consistencia de una base de conocimiento. Sin embargo hemos visto que es posible detectar otros problemas del conocimiento disponible a partir de las distribuciones de posible posibilidad. El hecho de que a partir de unas observaciones y un determinado conocimiento (expresado por ejemplo en forma de reglas) se obtengan distribuciones de posible posibilidad no normalizadas refleja lagunas en dicho conocimiento, posiblemente subsanables con la aportación de nuevo conocimiento que complementa al inicial. Esto apunta a dos tipos de defectos diferentes en el conocimiento que un sistema pueda tener almacenado, las contradicciones lógicas por un lado, y su insuficiencia o falta de completitud por otro.

3.6.6 Valor de verdad de una afirmación cuando los hechos vienen modelados con una distribución de posibilidad

Cuando se dispone de una observación de la realidad dada por una distribución de posibilidad P' sobre una variable u , " u es P' ", y se quiere contrastar una hipótesis P efectuada sobre esa misma variable, " u es P ", se necesita disponer de algún mecanismo de evaluación de la veracidad de la afirmación P dado P' . Uno de los procedimientos más generales consiste en calcular el valor de verdad borroso de P dado P' . Aunque existen otros tipos de medidas menos complejas también son mucho menos informativas. Este es el caso de las medidas de necesidad y posibilidad, que permiten evaluar en qué medida es necesario que P sea verdadero dado P' , o bien en qué medida es posible que P sea verdadero dado P' . En este apartado revisaremos la obtención de los valores de verdad borrosos, así como de las medidas de necesidad y posibilidad más comunes, y veremos algunos de los inconvenientes que unos y otros plantean.

3.6.6.1 Valor de verdad borroso de una afirmación

La verdad de una afirmación se entiende como su conformidad con los hechos que describen el mundo conocido. Por tanto la evaluación del grado de verdad de una proposición involucra dos elementos, por un lado los hechos, y por otro la afirmación que se contrasta. En lógica multivaluada la verdad de una afirmación se mide con un número en el intervalo $[0, 1]$, mientras que en lógica borrosa se generaliza a un conjunto borroso definido en $[0, 1]$, llamado valor de verdad borroso (véase por ejemplo [Godo, Jacas, Valverde, 91]).

El valor de verdad borroso de P dado P' es la distribución de posibilidad del grado de verdad de la afirmación P , de forma que además de aceptarse la existencia de grados de verdad parciales, se admite que estos no sean conocidos con precisión. En [Dubois, Lang, Prade, 91] o [Dubois, Prade, 88] se hace un repaso de las distintas situaciones que pueden presentarse a la hora de evaluar el grado (o el valor borroso) de verdad de una afirmación dado un hecho u observación, que depende de la imprecisión, vaguedad e incertidumbre tanto de la afirmación como del hecho en cuestión.

Cuando el hecho es preciso, modelado por tanto mediante un conjunto unitario $P' = \{u_0\}$, el grado de verdad de una afirmación “ u es P ” se obtiene sencillamente como $P(u_0)$, valor también preciso, cuya distribución de posibilidad asociada es un conjunto clásico unitario.

En el caso general, cuando la observación viene dada por un conjunto cualquiera $P'(u)$, la expresión anterior puede generalizarse mediante el principio de extensión obteniéndose el valor de verdad borroso como la pertenencia de la observación P' en la hipótesis P , $P(P')$. Formalmente el **valor de verdad borroso** de una afirmación P dado un hecho u observación P' que describe el conocimiento que sobre el mundo real se tiene, y que denotaremos por $\tau_{PP'}$, se obtiene como:

$$\tau_{PP'}(x) = \begin{cases} \sup_{u/P(u)=x} P'(u) \\ 0 \text{ if } P^{-1}(x) = \emptyset \end{cases} = P(P')$$

que equivale a:

$$\tau_{PP'}(x) = \begin{cases} \sup\{P'(u) / u \in P^{-1}(\{x\})\} & \text{if } P^{-1}(\{x\}) \neq \emptyset \\ 0 & \text{en otro caso} \end{cases}$$

Esta medida también se denomina comúnmente compatibilidad borrosa ([Zadeh, 78]) y como iremos viendo resulta extraordinariamente informativa, a diferencia de otras medidas

con valores reales utilizadas para propósitos similares, que solo pueden describir aspectos parciales sobre la veracidad de una afirmación.

Al tratarse de distribuciones de posibilidad, y según lo visto hasta ahora, debería hacerse explícita la distinción entre distribuciones de posible y necesaria posibilidad en el universo de grados de verdad, que recordemos tienen interpretaciones significativamente distintas. Esto hace que, junto con la gran cantidad de información que contienen, los valores de verdad borrosos sean en general difíciles de interpretar. Por ello el capítulo 6 de esta tesis estará dedicado a su estudio y simplificación, de cara a facilitar tanto su interpretación como su manipulación.

3.6.6.2 Medidas de necesidad y posibilidad

Como se vio en el capítulo 2, las medidas de necesidad y posibilidad, al igual que las medidas de probabilidad, forman parte de una familia más general, llamadas medidas de confianza, medidas de incertidumbre o medidas borrosas.

Como ya se dijo una **medida de necesidad** es una medida de incertidumbre que verifica la propiedad adicional:

$$N(P_1 \cap P_2) = \min(N(P_1), N(P_2))$$

mientras que una **medida de posibilidad** es una medida de incertidumbre que verifica la propiedad adicional:

$$\Pi(P_1 \cup P_2) = \max(\Pi(P_1), \Pi(P_2))$$

Nótese sin embargo que cuando intercambiamos el tipo de agregación lógica en las expresiones anteriores, sólo se obtienen las desigualdades:

$$\begin{aligned} N(P_1 \cup P_2) &\geq \max(N(P_1), N(P_2)) \\ \Pi(P_1 \cap P_2) &\leq \min(\Pi(P_1), \Pi(P_2)) \end{aligned}$$

Las medidas de necesidad y posibilidad (que llamaremos habituales) de una hipótesis P dada una observación P' vienen definidas como:

$$\begin{aligned} N(P / P') &= \inf_u (\max(P(u), 1 - P'(u))) \\ \Pi(P / P') &= \sup_u (\min(P(u), P'(u))) \end{aligned}$$

y son duales respecto a la negación habitual:

$$N(P / P') = \inf_u (\max(P(u), 1 - P'(u))) = 1 - \Pi(\neg P / P') = 1 - \Pi(P' / \neg P) = \sup_u (\min(P'(u), 1 - P(u)))$$

lo que significa que una proposición P es más necesariamente verdadera cuanto más imposible sea su negación $\neg P$, lo que resulta acorde con el sentido común.

Las medidas de necesidad y posibilidad permiten estimar, dado un hecho P' , cómo de necesario es que una proposición P sea verdadera, o cómo de posible es que P sea verdadera. La medida de necesidad de P dado P' es una medida de la inclusión de P en P' , y generalmente se interpreta como una medida de la certidumbre que se tiene en P conocido P' . Por otro lado la medida de posibilidad de P dado P' es una medida de la intersección de P con P' y se interpreta como una medida de lo posible que es que P sea cierto dado P' , o de la certidumbre en el dominio de lo posible que se tiene en P . Sin embargo como veremos a lo largo del capítulo 4 pueden encontrarse otras medidas de posibilidad y necesidad que permitan diferenciar claramente entre grados de verdad parciales y grados de certidumbre parciales.

Las medidas de necesidad y posibilidad habituales pueden ser generalizadas reemplazando la implicación $\max(1-x, y)$ por otra implicación, y sustituyendo la t-norma \min por otra t-norma distinta:

$$N(P / P') = \inf_u (I(P(u), P'(u))) \\ \Pi(P / P') = \sup_u T(P(u), P'(u))$$

aunque en este caso las medidas dejan de ser en general duales. Es interesante señalar que, como ocurre con otros muchos operadores en la teoría de conjuntos borrosos, no existen criterios claros que permitan escoger entre unos u otros tipos de medidas de necesidad y posibilidad, y que de hecho como en otros problemas similares en el ámbito de la teoría de conjuntos borrosos, la elección suele apoyarse en criterios de sencillez más que en criterios semánticos, generalmente más complejos, pero que son los que a nuestro entender deberían primar.

Las medidas de necesidad y posibilidad duales aplicadas a proposiciones clásicas (es decir modeladas con conjuntos clásicos) presentan algunas propiedades interesantes que pueden facilitar su interpretación semántica. Así por ejemplo se tiene ([Dubois, Lang, Prade, 91], [Palmer, 91]):

- $N(P/P') > 0 \Rightarrow \Pi(P/P')=1$: para que una afirmación P sea necesaria en grado no nulo debe ser primero totalmente posible.

- $\Pi(P/P') < 1 \Rightarrow N(P/P')=0$: si una afirmación P no es totalmente posible no puede ser necesaria en ningún grado.
- $N(P/P') = 1 \Rightarrow \Pi(\neg P/P')=0$: si una afirmación P es totalmente necesaria su negación $\neg P$ debe ser imposible.
- $\Pi(P/P') = 0 \Rightarrow N(\neg P/P') = 1$: si una afirmación P es imposible su negación $\neg P$ debe ser necesaria.
- $\Pi(P/P') < 1 \Rightarrow \Pi(\neg P/P') = 1$: dada una afirmación P y su negación $\neg P$, alguna de las dos debe ser totalmente posible: $\max(\Pi(P/P'), \Pi(\neg P/P'))=1$.
- $N(P/P') > 0 \Rightarrow N(\neg P/P') = 0$: dada una afirmación P y su negación $\neg P$, no pueden ser ambas necesarias en grado no nulo: $\min(N(P/P'), N(\neg P/P'))=0$.

Sin embargo cuando las proposiciones son borrosas estas propiedades no se verifican. Baste para ello comprobar, por ejemplo, que cuando P es borroso y P' es un conjunto unitario entonces $N(P/P')=\Pi(P/P')$. De hecho las dos últimas propiedades sólo se verifican si se verifica la ley del tercero excluido. También se tiene en general:

$$0.5 \leq \max(\Pi(P/P'), \Pi(\neg P/P')) \leq 1$$

$$0 \leq \min(N(P/P'), N(\neg P/P')) \leq 0.5.$$

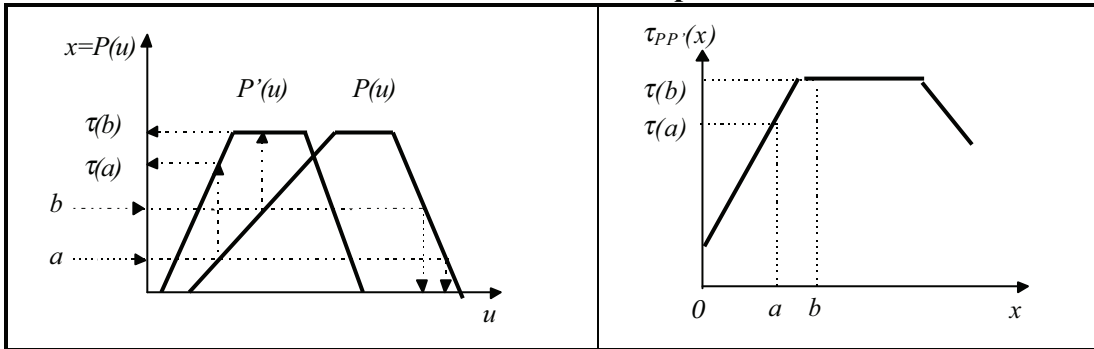
Veremos a continuación (como ya se señala por ejemplo en [Dubois, Prade, 90] o [Magrez, Smets, 89b]) que la medida de necesidad habitual, aplicada a proposiciones borrosas, presenta algunas propiedades que chocan con la intuición, y hacen que su uso sea en muchas ocasiones desaconsejable.

3.6.6.3 Valores de verdad borrosos y medidas de necesidad y posibilidad

Resulta interesante, para su mejor comprensión, detallar el procedimiento de obtención de la compatibilidad borrosa entre una hipótesis P y un hecho P' de cara a introducir y facilitar el análisis cualitativo los distintos tipos de valores de verdad borrosos que pueden obtenerse ante varias combinaciones distintas, suficientemente representativas, de hechos e hipótesis. Por simplicidad nos limitaremos a modelar P y P' como conjuntos borrosos trapezoidales.

El procedimiento de obtención de la compatibilidad borrosa se ilustra en la siguiente figura:

Procedimiento de obtención de la compatibilidad borrosa



donde como puede verse, para cada grado de pertenencia de P' se buscan aquellos valores de u que lo verifican, y se obtiene su máximo grado de pertenencia a P .

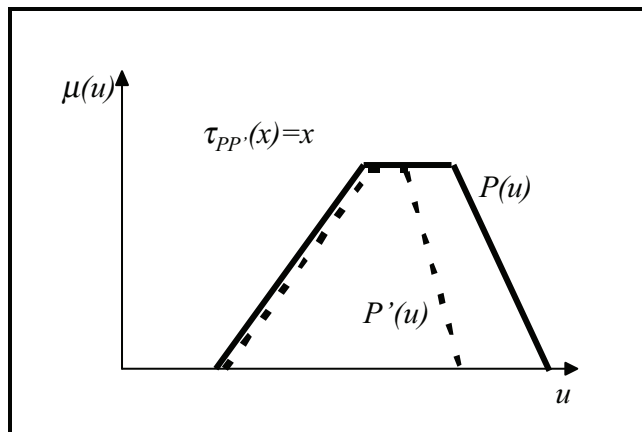
Uno de los primeros puntos que conviene señalar es que la proposición “ u es P' ” no es del todo equivalente a “ u es p es τ_{PP} ”, cuya distribución de posibilidad asociada viene dada por:

$$P'(u) = \tau_{PP}(P(u))$$

ya que en efecto solamente se cumple:

$$P'(u) \leq \tau_{PP}(P(u)), \text{ o lo que es lo mismo: } P' \subset \tau_{PP}(P)$$

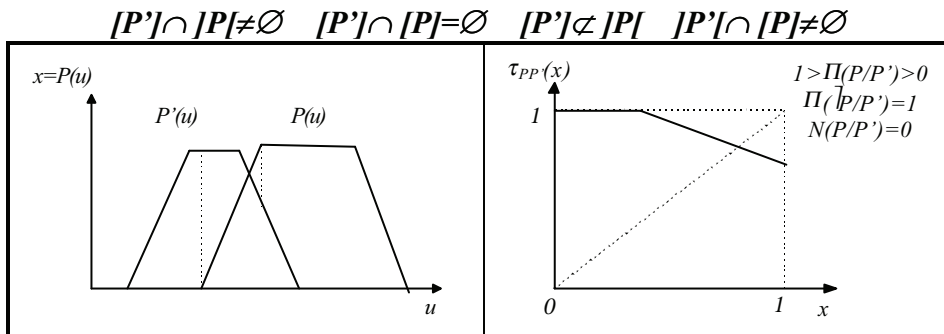
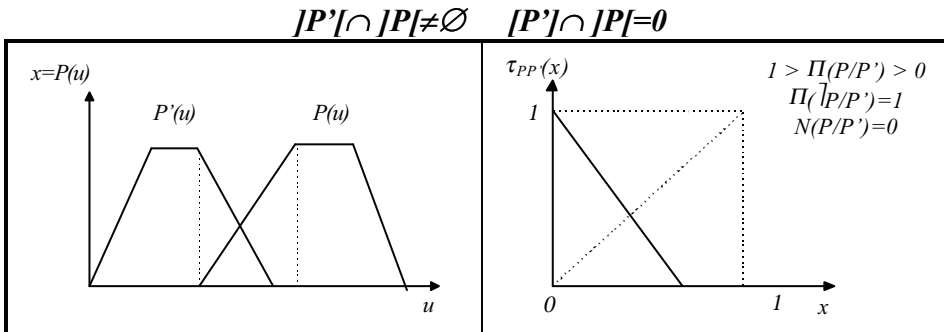
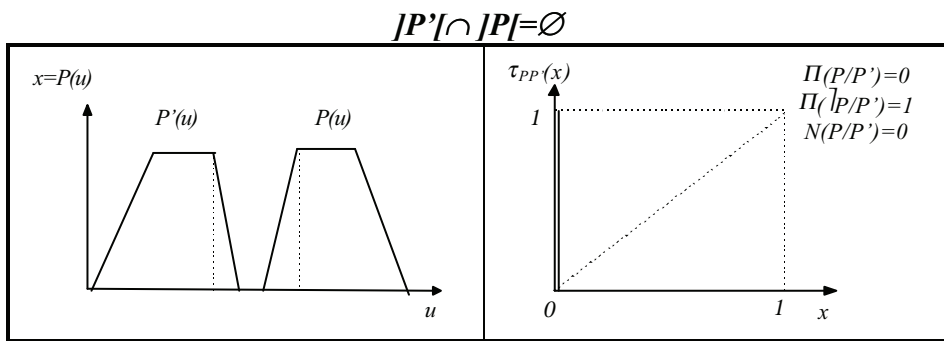
Baste para ello ver que cuando P' y P son tal y como aparecen en la siguiente figura:



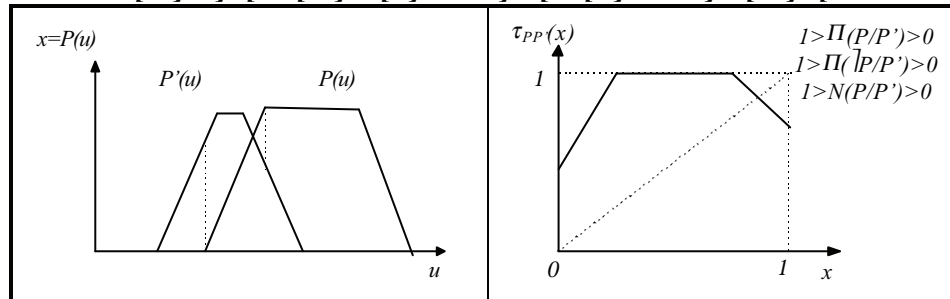
se obtiene $\tau_{PP'}(x) = x$ y obviamente se cumple $P' \subset \tau_{PP'}(P) = P$, pero no $P' = P$.

Por otro lado la compatibilidad borrosa es una medida totalmente asimétrica, lo que resulta intuitivamente lógico, ya que el grado de verdad de P dado P' está íntimamente ligado con la inclusión de P' en P que es una operación de naturaleza totalmente asimétrica.

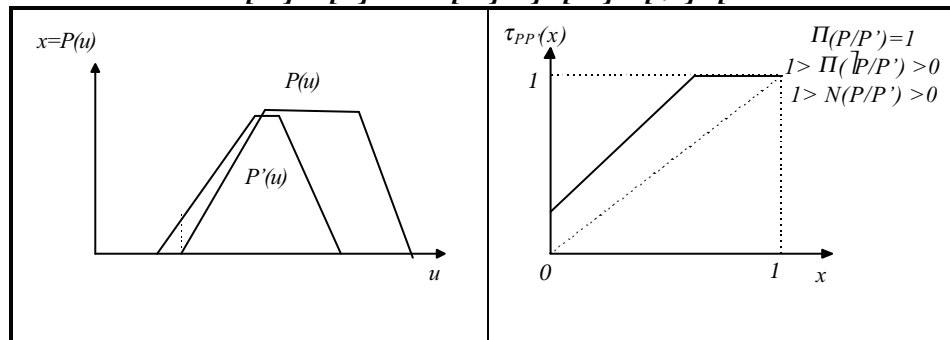
Cuando tanto P como P' se representan mediante conjuntos borrosos trapezoidales resulta sencillo obtener de forma aproximada la compatibilidad borrosa entre ambos conjuntos. Las siguientes figuras pretenden mostrar algunas de las situaciones que pueden darse entre P y P' , la compatibilidad borrosa que resulta entre ambos, y los valores cualitativos de las medidas de posibilidad y necesidad habituales de P dado P' .



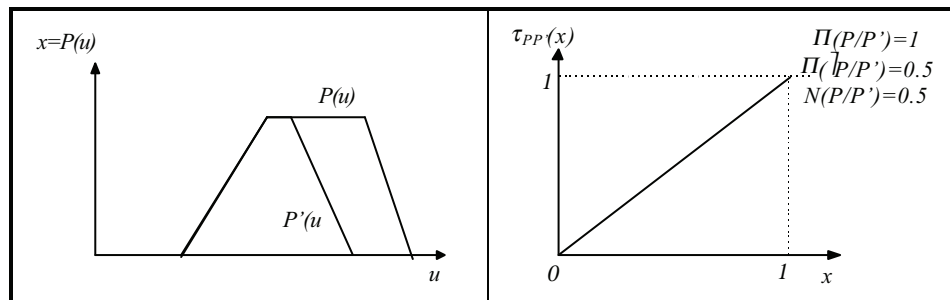
$[P'] \subset [P]$ $[P'] \cap [P] = \emptyset$ $[P'] \cap [P] \neq \emptyset$ $[P'] \not\subset [P]$



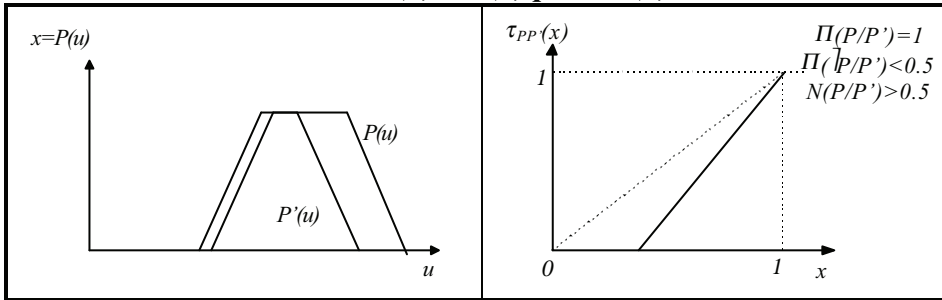
$[P'] \cap [P] \neq \emptyset$ $[P'] \subset [P]$ $[P'] \not\subset [P]$



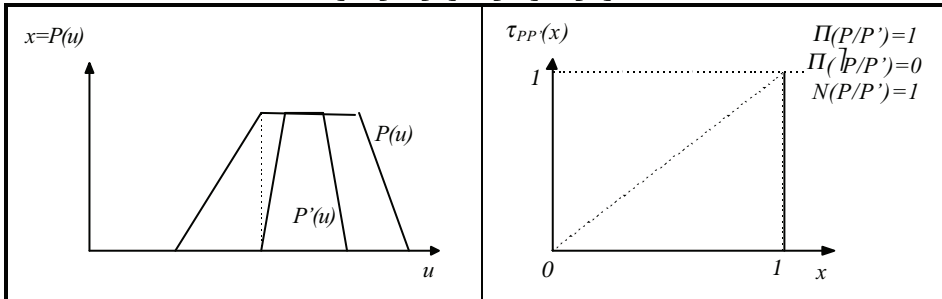
$P' \subset P$



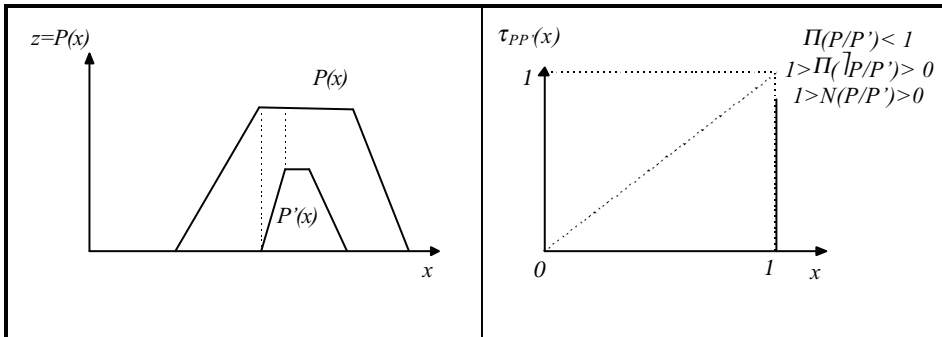
$P' \subset P \quad P'(u) < P(u) \text{ para } P'(u) \neq 1$



$]P'[\subset [P] \quad [P']\subset [P]$



$P' \subset P \quad P' \text{ no normal}$



Del análisis del comportamiento de las medidas anteriores podemos destacar algunos aspectos de interés. Por un lado la medida de necesidad estándar $N(P/P')$ puede considerarse como una medida muy estricta de la inclusión de P' en P . En efecto, supuesto que tanto P' como P están normalizados, la inclusión de P' en P (recordemos que se define por $P'(u) \leq P(u)$) sólo garantiza $N(P/P')=0.5$ y por la misma razón se tiene $N(P/P)=0.5$ lo que no parece intuitivamente razonable. Para tener $N(P/P')=1$ debe cumplirse $P' \subset [P]$ que es una condición de inclusión excesivamente estricta. Sin embargo desde un punto de vista semántico parece razonable suponer que $P' \subset P$ debe ser leído como ‘ P es cierto/necesario dado P' ’. Por otro lado $]P'[\subset [P]$ que es un tipo de

inclusión bastante restrictivo solo garantiza $N(P/P') > 0$. Esto pone de manifiesto la parcial inadecuación de la medida de necesidad habitual para caracterizar con suficiente aproximación el cumplimiento de una hipótesis P ante un hecho P' , y sugiere la conveniencia de combinar simultáneamente más de una medida de necesidad para dicho fin.

Por otro lado parece razonable pensar que $[P'] \cap [P] \neq \emptyset$ sea leído como que es del todo posible que P sea cierto dado P' , y en este caso en efecto se tiene $\Pi(P/P') = 1$. Cuando sólo se tiene $]P'] \cap]P[\neq \emptyset$ se verifica $\Pi(P/P') > 0$ lo que parece concordar con el sentido común ya que en efecto la intersección de ambos soportes parece indicar que P no está en total contradicción con P' y por tanto es parcialmente posible que P sea cierto dado P' . Esta medida parece presentar un comportamiento más satisfactorio, lo que evidencia una contradicción entre ambas, puesto que a pesar de su dualidad matemática presentan comportamientos asimétricos que chocan con lo intuitivamente deseable.

Otro punto importante que cabe señalar es que, como ya hemos indicado, si bien el valor de verdad borroso de una afirmación P contiene una gran cantidad de información, esto le confiere un mayor grado de complejidad a la hora de su interpretación. A menudo se tiene una cierta tendencia a escoger de forma arbitraria una familia base de valores de verdad borrosos, y a cada uno de ellos asociarles nombres como *verdadero*, *poco_verdadero*, etc. Sin embargo la existencia de distintos tipos de incertidumbre (ya comentada en secciones anteriores) pone de manifiesto la conveniencia asociar tipos de distribuciones con tipos de incertidumbre, y de investigar con precisión la interpretación de estos tipos de distribuciones de cara a asignarles nombres que describan con suficiente exactitud la información que contienen. Por ello, y como ya hemos comentado, el capítulo 6 está exclusivamente dedicado al estudio de los valores de verdad borrosos.

3.7 Razonamiento aproximado en lógica borrosa con distribuciones de posibilidad

La lógica borrosa pretende a partir de un conjunto de proposiciones borrosas verdaderas, y de un conjunto de reglas de inferencia, deducir nuevas proposiciones borrosas verdaderas, que sean conclusiones lógicas de las proposiciones del conjunto de partida, proceso que se denomina **inferencia**. A continuación analizaremos lo que se entiende por proposición borrosa, para posteriormente enunciar algunas de las reglas de inferencia más utilizadas.

3.7.1 Proposiciones vagas o borrosas

Sea un conjunto finito de variables, u_1, u_2, \dots, u_n , definidas en sus respectivos universos de discurso U_1, U_2, \dots, U_n .

Una proposición atómica borrosa r ([Castro, Trillas, Zurita, 95]) es una proposición del tipo “ $(u_{i1}, u_{i2}, \dots, u_{im})$ es R ”, siendo R una relación borrosa definida en $U_{i1} \times U_{i2} \times \dots \times U_{im}$, con $U_{ij} \in \{U_1, U_2, \dots, U_n\}$. Las proposiciones atómicas más sencillas son del tipo “ u_i es P_i ” siendo P_i un conjunto borroso definido en el universo U_i .

Diremos que la proposición está en su forma normal cuando sea del tipo “ $u=(u_1, u_2, \dots, u_n)$ es R ” con R definida en $U_1 \times U_2 \times \dots \times U_n$. Cualquier proposición atómica se puede poner en su forma normal obteniendo su extensión cilíndrica a aquellos universos sobre los que no está definida.

Para cualquier proposición borrosa “ u es R ”, R será interpretado como una distribución de posibilidad, de forma que cuando R es una distribución de posible posibilidad diremos que “ u es R ” es una **p-proposición** borrosa o proposición borrosa de tipo p, mientras que cuando R es una distribución de necesaria posibilidad diremos que “ u es R ” es una **n-proposición** borrosa o proposición borrosa de tipo n.

Una proposición borrosa se define recursivamente como sigue:

- cualquier proposición atómica es una proposición borrosa.
- dadas dos proposiciones borrosas r_1 y r_2 del mismo tipo, son proposiciones borrosas del mismo tipo las obtenidas como $(r_1 \text{ y } r_2)$, $(r_1 \text{ o } r_2)$, (si r_1 entonces r_2) y $(\neg r_1)$
- cualquier proposición borrosa se obtiene aplicando las dos reglas anteriores un número finito de veces.

Las reglas anteriores serán válidas sólo cuando se combinen proposiciones borrosas del mismo tipo. Cualquier proposición borrosa se puede poner en forma de proposición atómica, reemplazando los conectivos lógicos por sus correspondientes operadores borrosos. Formalmente resulta más sencillo expresarlo en función de sus formas normales: dadas dos proposiciones borrosas r_v y r_w con expresiones atómicas normales respectivas “ v es R_v ” y “ w es R_w ” se obtienen las siguientes formas normales atómicas para las proposiciones borrosas antes enunciadas:

- (r_v y r_w) es “ u es $(R_v \cap R_w)$ ” con $(R_v \cap R_w)(u) = T(R_v(u), R_w(u))$
- (r_v o r_w) es “ u es $(R_v \cup R_w)$ ” con $(R_v \cup R_w)(u) = S(R_v(u), R_w(u))$
- (si r_v entonces r_w) es “ u es $(R_v \rightarrow R_w)$ ” con $(R_v \rightarrow R_w)(u) = I(R_v(u), R_w(u))$
- ($\neg r_v$) es “ u es $(\neg R_v)$ ” con $(\neg R_v)(u) = \Gamma(R_v(u))$

Nótese que para poder obtener las expresiones anteriores es necesario escoger los operadores de unión, intersección, implicación y negación. Este conjunto de operadores se denomina a menudo interpretación. La elección del operador de implicación será detalladamente analizada a lo largo del capítulo 4, ya que es uno de los aspectos cruciales en la inferencia en lógica borrosa. La elección de los operadores de unión e intersección se abordará en el capítulo 5.

3.7.2 Reglas de inferencia en lógica borrosa

Consideraremos como ya se ha mencionado en el apartado anterior dos tipos de proposiciones, de forma similar a como se hace en las lógicas modales (véase por ejemplo [Dubois, Prade, 88b]), por un lado proposiciones de tipo n o n-proposiciones, “ u es R_n ”, que pueden leerse como “es necesario que u esté en R_n ” y donde por tanto R_n se interpreta como una distribución de necesaria posibilidad, y por otro lado proposiciones de tipo p o p-proposiciones, “es posible que u esté en R_p ”, donde R_p debe interpretarse como una distribución de posible posibilidad. Este último tipo de proposiciones las denotaremos también por “ u es R_p ”. El concepto de conclusión lógica varía entonces según sea el tipo de proposiciones de partida. De un conjunto de proposiciones necesarias serán conclusiones lógicas aquellas que también sean necesarias, mientras que de un conjunto de proposiciones posibles serán conclusiones lógicas aquellas que también sean posibles.

Las reglas de inferencia tal y como se enuncian habitualmente sólo se aplican a distribuciones de necesaria posibilidad, pero no son en general aplicables a las distribuciones de posible posibilidad, introducidas en esta tesis, al igual que el principio de

mínima especificidad no dice como deben agregarse las distribuciones de posible posibilidad.

Dadas dos variables vectoriales u y v definidas respectivamente en los universos $U=U_1 \times U_2 \times \dots \times U_n$ y $V=V_1 \times V_2 \times \dots \times V_n$, vamos a enunciar algunas de las principales reglas de inferencia utilizadas en lógica borrosa, cuando los valores de las variables vienen caracterizados mediante distribuciones de posible y necesaria posibilidad. Nótese que las reglas para distribuciones de necesaria posibilidad son las que se pueden encontrar en otros trabajos (véase por ejemplo [Castro, Trillas, Zurita, 95], [Yager, 84]) pero no así las reglas para distribuciones de posible posibilidad, formalmente introducidas en esta tesis. En lo que sigue los subíndices n y p indican si la distribución a la que se aplican es una distribución de necesaria o posible posibilidad respectivamente, P y P_i son conjuntos borrosos o distribuciones de posibilidad definidos en U , Q y Q_i son conjuntos borrosos o distribuciones de posibilidad definidos en V y R es una relación borrosa definida en $U \times V$.

- inclusión

$$\begin{array}{l} u \text{ es } P_{1n} \\ P_{1n} \subseteq P_{2n} \end{array}$$

$$u \text{ es } P_{2n}$$

$$\begin{array}{l} u \text{ es } P_{1p} \\ P_{1p} \subseteq P_{2p} \end{array}$$

$$u \text{ es } P_{2p}$$

Estas reglas expresan el hecho de que dada una distribución de necesaria posibilidad para una variable, cualquier distribución mayor también lo es. Y del mismo modo dada una distribución de posible posibilidad para una variable, cualquier distribución menor también lo es. En efecto se tiene, respecto a una distribución de necesaria posibilidad de $\pi_n(u)$ referencia:

$$\pi_n(u) \leq P_{1n}(u) \text{ y } P_{1n}(u) \leq P_{2n}(u) \Rightarrow \pi_n(u) \leq P_{2n}(u)$$

y

$$P_{1p}(u) \leq \pi_n(u) \text{ y } P_{2p}(u) \leq P_{1p}(u) \Rightarrow P_{2p}(u) \leq \pi_n(u)$$

Es importante señalar que cuando no se tiene ninguna información sobre el valor que puede tomar una variable, es admisible suponer:

$$\pi(u) = \pi_n(u) = \pi_p(u) = 1$$

ya que por un lado $\pi_n(u) = 1$ es la única distribución de necesaria posibilidad que no impone restricciones arbitrarias a la variable u y por otro lado $\pi_p(u) = 1$ es compatible con la restricción fundamental $\pi_n(u) \geq \pi_p(u)$ por lo que es una distribución de posible posibilidad válida. En definitiva estas dos suposiciones son equivalentes a la suposición razonable de que la distribución de posibilidad viene dada por $\pi(u) = 1$.

La restricción fundamental aplicable a las distribuciones de posible posibilidad permite enunciar otra regla de inferencia más general para este tipo de distribuciones:

$$\begin{array}{l} u \text{ es } P_n \\ P_p \subseteq P_n \end{array}$$

$$u \text{ es } P_p$$

que viene a expresar el hecho de que cualquier conjunto borroso menor que la distribución de necesaria posibilidad de la variable u es una distribución de posible posibilidad válida.

Sin embargo ninguna de estas reglas debe interpretarse como un procedimiento de obtención de distribuciones puesto que todas ellas suponen una pérdida de información innecesaria, por lo que no creemos que se puedan considerar como reglas de inferencia propiamente dichas.

- Agregación de distribuciones en un mismo universo

$$\begin{array}{l} u \text{ es } P_{1n} \\ u \text{ es } P_{2n} \end{array}$$

$$u \text{ es } P_{1n} \cap P_{2n}$$

donde $(P_{1n} \cap P_{2n})(u) = T(P_{1n}(u), P_{2n}(u))$ y T es una t-norma.

$$\begin{array}{l} u \text{ es } P_{1p} \\ u \text{ es } P_{2p} \end{array}$$

u es $P_{1p} \cup P_{2p}$

donde $(P_{1p} \cup P_{2p})(u) = S(P_{1p}(u), P_{2p}(u))$ y S es una t-conorma.

Como ya hemos visto estas dos reglas de inferencia se deducen del principio de máxima información y máxima certidumbre anteriormente enunciado. En ausencia de información adicional la agregación conjuntiva debe hacerse con el operador *min* y la disyuntiva con el operador *max*. Ambas reglas expresan el hecho de que dadas dos distribuciones de posibilidad para una misma variable, su agregación depende del tipo de distribución, como ya ha sido visto en el apartado -. Mientras que estas reglas son directamente válidas para los operadores *min* y *max* puesto que se tiene:

$$\pi_n(u) \leq P_{1n}(u) \text{ y } \pi_n(u) \leq P_{2n}(u) \Rightarrow \pi_n(u) \leq \min(P_{1n}(u), P_{2n}(u))$$

y

$$\pi_n(u) \geq P_{1p}(u) \text{ y } \pi_n(u) \geq P_{2p}(u) \Rightarrow \pi_n(u) \geq \max(P_{1p}(u), P_{2p}(u))$$

y su validez e interpretación para otros operadores de intersección y unión será analizada en el capítulo 5. Cuando se agregan dos distribuciones de posible posibilidad sobre un mismo universo, el uso de un operador disyuntivo, tal y como se vio en el apartado 3.6.3, no modifica la semántica conjuntiva inducida por la existencia simultánea de dos distribuciones.

- Agregación de distribuciones en universos distintos

Debido a consideraciones semánticas la agregación de distribuciones de posibilidad definidas sobre universos de discurso diferentes, y que por tanto restringen los valores de variables diferentes, debe hacerse mediante una agregación conjuntiva, independientemente del tipo de distribución considerado (a diferencia de lo que ocurría con distribuciones sobre un mismo universo). Esta forma de proceder garantiza que se verifican las restricciones fundamentales para ambos tipos de distribuciones. En efecto para dos distribuciones de necesaria posibilidad $P_n(u)$ y $Q_n(v)$ se tiene, respecto a las distribuciones de referencia $\pi_{u-n}(u)$ y $\pi_{v-n}(v)$:

$$P_n(u) \geq \pi_{u-n}(u) \text{ y } Q_n(v) \geq \pi_{v-n}(v) \Rightarrow \\ \pi_{uv-n}(u, v) = \min(P_n(u), Q_n(v)) \geq \min(\pi_{u-n}(u), \pi_{v-n}(v))$$

Y para dos distribuciones de posible posibilidad $P_p(u)$ y $Q_p(v)$ la distribución obtenida mediante la agregación *min* es también una distribución de posible posibilidad válida:

$$P_p(u) \leq \pi_{u-n}(u) \quad y \quad Q_p(v) \leq \pi_{v-n}(v) \Rightarrow \\ \pi_{uv-p}(u,v) = \min(P_p(u), Q_p(v)) \leq \pi_{uv-n}(u,v) = \min(\pi_{u-n}(u), \pi_{v-n}(v))$$

mientras que no lo es la obtenida con la agregación *max*, ya que la distribución $\pi_{uv-p}(u,v) = \max(P_p(u), Q_p(v))$ no es en menor que la distribución de necesaria posibilidad $\pi_{uv-n}(u,v) = \min(\pi_{u-n}(u), \pi_{v-n}(v))$ tomada como referencia.

Por tanto la regla de inferencia para la agregación de distribuciones de posibilidad de un mismo tipo de variables en universos distintos es única y viene dada por:

u es P
 v es Q

(u,v) es $P \times Q$

donde $(P \cap Q)(u,v) = T(P(u), Q(v))$ y T es un operador conjuntivo.

y en general en ausencia de información adicional el operador a emplear debe ser la t -norma *min*. Las distribuciones involucradas deben tener el mismo tipo, que será el que tenga de la distribución resultante.

- proyección

(u,v) es R

u es R_V

donde $R_V(u) = \sup_v R(u,v)$

En este caso esta regla de inferencia es válida tanto si R es una distribución de posible como de necesaria posibilidad para la variable (u,v) y R_V tiene entonces el mismo tipo que R . Esta regla permite eliminar la influencia de una de las variables en la distribución conjunta para obtener una aproximación para la distribución de las restantes variables.

- regla composicional de inferencia (RCI)

$$\begin{array}{l} u \text{ es } P' \\ (u,v) \text{ es } R \end{array}$$

$$v \text{ es } P' \circ R$$

donde $(P' \circ R)(v) = \sup_u A(P'(u), R(u, v))$ y debe leerse P' compuesto con R .

P' debe tener también el mismo carácter que R , de forma que la combinación de distribuciones de posible posibilidad genere una distribución de posible posibilidad, y la combinación de distribuciones de necesaria posibilidad genere una distribución de necesaria posibilidad, es decir que el carácter de $P' \circ R$ es el mismo que el de R . La combinación conjuntiva mediante el operador A expresa la necesidad de que se verifiquen simultáneamente la relación R y la observación P' , y por tanto es independiente del carácter de las distribuciones involucradas, al igual que ocurre con la agregación de distribuciones de posibilidad sobre universos distintos. De hecho el uso de un operador disyuntivo daría en general lugar a distribuciones carentes de contenido informativo. Por otro lado el no escoger una norma se debe a que la asociatividad y conmutatividad no parecen en este caso estrictamente necesarias. Sin embargo el uso de unos y otros operadores conjuntivos debe estar basado en la existencia de información adicional sobre el tipo de implicación existente, como se verá en el capítulo 5, y parece razonable seguir manteniendo el criterio de utilizar el operador \min cuando no se disponga de dicha información adicional. En el capítulo 5 veremos cuales son operadores por defecto que se deducen del principio más general de máxima información y máxima certidumbre, para distribuciones de posible y necesaria posibilidad.

La regla composicional de inferencia es una generalización del principio de extensión cuando la relación existente entre las variables u y v no es una relación clásica (como la que define una función ordinaria $f(u, v) = 0$) sino una relación borrosa. Por tanto lo que aquí se comenta es también aplicable al principio de extensión, que a su vez puede considerarse como otra regla de inferencia, al ser un caso particular de la regla composicional de inferencia.

- **regla del modus ponens generalizado**

si u es P entonces v es Q (o bien $P \rightarrow Q$)
 u es P'

v es Q'

donde $Q'(v) = \sup_u M(P'(u), I(P(u), Q(v)))$ y (M, I) es un modelo de $P \rightarrow Q$

La proposición “si u es P entonces v es Q ” (o bien $P \rightarrow Q$) se denomina regla (véase el apartado 2.5.4 del capítulo 2), siendo “ u es P ” el antecedente o hipótesis de la regla, y “ v es Q ” su consecuente. La proposición “ u es P' ” es la observación, y “ v es Q' ” es la conclusión que se obtiene a partir de la regla y la observación. Diremos que un valor u_0 de la variable u ha sido observado cuando se cumpla $P'(u_0) > 0$, y diremos que la hipótesis contempla (o tiene en cuenta) a un valor u_0 de u cuando se cumpla $P(u_0) > 0$.

Esta regla de inferencia es válida para distribuciones de posible y necesaria posibilidad. En aplicación de las dos reglas anteriores se deduce que tanto $P'(u)$ como $I(P(u), Q(v))$ deben tener el mismo carácter, por lo que es importante determinar en función de $I(x, y)$ cuándo la distribución $I(P(u), Q(v))$ puede interpretarse como distribución de posible o necesaria posibilidad, o directamente cómo debe interpretarse $Q'(v)$ en función de I , aspecto éste analizado en el capítulo 4.

Por otro lado el cumplimiento del modus ponens clásico parece una propiedad intuitivamente exigible. Esta propiedad se traduce en que de la observación “ u es P ” debería deducirse la conclusión “ v es Q ”, propiedad que hemos denominado **modus ponens estricto** (MPE), y que impone un serie de restricciones al par (M, I) . Por ello el estudio del modus ponens generalizado debe basarse en un estudio en profundidad de lo que hemos denominado en esta tesis **modelos de regla**, esto es el conjunto de pares (M, I) utilizables para modelar matemáticamente la regla. Los modelos de reglas serán estudiados en el capítulo 4, dedicado íntegramente a la inferencia en lógica borrosa. Podemos sin embargo anticipar que un modelo está constituido por una función de implicación I (que puede ser por ejemplo una implicación residual o fuerte o un operador conjuntivo) y una función generadora de modus ponens M que debe ser un operador conjuntivo (asociatividad y conmutatividad no parecen necesarias), y que ambos deben cumplir la **desigualdad del modus ponens** (DMP), es decir ([Trillas, 93]):

$$M(x, I(x, y)) \leq y$$

Además creemos necesario exigir la verificación del modus ponens estricto, que matemáticamente se traduce en:

$$Q(v) = \sup_u M(P(u), I(P(u), Q(v)))$$

con lo que se preserva en el ámbito borroso el modus ponens clásico.

En general suele argumentarse que la función de implicación debe ser un M -preorden, es decir reflexiva y M -transitiva. La M -transitividad está relacionada con el encadenamiento (véase el apartado 5.8 del capítulo 5), mientras que la reflexividad asegura que de la observación y la regla se deduce siempre la observación. Sobre la reflexividad se hacen algunos comentarios y críticas en el capítulo 4, en el apartado 4.3.1, mientras que la transitividad se trata someramente cuando se analizan sistemas de reglas en el apartado 5.8 del capítulo 5. Veremos que ambas propiedades no son tan necesarias como en una primera aproximación podría pensarse.

La regla composicional de inferencia es la principal regla de inferencia de los sistemas basados en reglas. En ella aparecen dos operadores, la función de implicación y la función generadora de modus ponens, a los que hemos llamado conjuntamente modelo de regla. Existen multitud de posibles opciones a la hora de seleccionar estos operadores, pero todavía no se tienen criterios claros de cuando deben ser escogidos unos y otros. Existen muchos estudios sobre propiedades básicas que deben verificar (aunque no hay un consenso claro entre distintos autores) y aquellas que pueden o no ser de interés. Sin embargo no hay demasiados estudios semánticos, que intenten dilucidar el significado que realmente tiene la elección de un operador frente a otro, y que tengan en cuenta la existencia de distribuciones de posibilidad distintas, las de posible y las de necesaria posibilidad, aspectos básicos que creemos justifican gran parte del trabajo llevado a cabo en esta tesis.

3.8 El estado del arte en la aplicación de la lógica borrosa a los sistemas expertos de diagnóstico

Uno de los principales problemas que plantea la aplicación de la teoría de conjuntos borrosos a aplicaciones reales es la elección de la técnica concreta a emplear y de los operadores necesarios para las distintas combinaciones, agregaciones, etc., que el enfoque escogido requiera.

La complejidad del diseño de sistemas borrosos se acrecienta debido a tres factores principales. Por un lado está, al igual que en muchas otras disciplinas y en especial aquellas menos consolidadas, una mala o insuficiente comprensión de la teoría empleada, lo que puede dar lugar a confusiones entre verdad parcial e incertidumbre, como se señala en [Dubois, Lang, Prade, 91], malas interpretaciones del comportamiento de algunas implicaciones, como se apunta en [Dubois, Prade, 96], etc., dando finalmente lugar a elecciones arbitrarias difícilmente justificadas, y por lo tanto no siempre adecuadas.

Por otro lado está la existencia de demasiados grados de libertad a la hora de escoger las herramientas teóricas (como es el caso de los operadores lógicos), problema característico de la teoría de conjuntos borrosos, lo que complica la fase de diseño de un sistema borroso y a menudo desanima a llevar un análisis en profundidad de las distintas posibilidades.

Esto da lugar a una gran escasez de estudios semánticos que puedan guiar las distintas elecciones que deben hacerse, haciendo que al final casi siempre hayan primado criterios de sencillez o criterios empíricos, no siempre justificables. Esta escasez de criterios hace aun más ardua la comprensión y el diseño de sistemas borrosos.

Por otro lado unas de las características que hacen especialmente atractivos a los sistemas borrosos son por un lado la riqueza que ofrecen para la representación del conocimiento, y por otro su robustez frente a pequeñas variaciones numéricas, lo que hace que la influencia de optar por unos u otros operadores no sea tan determinante en el funcionamiento final del sistema diseñado. Esto compensa en gran medida las debilidades antes apuntadas, y es probablemente una de las claves de su éxito.

En los últimos años han aparecido varios trabajos orientados a analizar desde un punto de vista semántico las consecuencias derivadas de escoger unos u otros operadores ([Dubois, Prade, 92], [Dubois, Prade, 92b], [Dubois, Prade, 93], [Yager, 96], [Grabot, Caillaud, 96]), pero todavía faltan criterios que permitan diferenciar claramente

muchas de las opciones que la teoría ofrece, y es en gran medida en esta parcela donde se enmarca la contribución que esta tesis aporta.

Existen varios enfoques diferentes con los que se han aplicado técnicas basadas en la teoría de conjuntos borrosos al diagnóstico, enfoques entre los cuales pueden destacarse las siguientes técnicas:

- ecuaciones relacionales borrosas
- lógica multivaluada
- cálculo de significados con reglas borrosas (enfoque numérico)
- medidas de incertidumbre y reglas simbólicas (enfoque simbólico)

Repasaremos muy brevemente las principales características de cada uno de estos enfoques.

3.8.1 Ecuaciones relacionales borrosas

Las ecuaciones relacionales borrosas ([Terano, Asai, Sugeno, 91], [Di Nola e tal., 89]) tienen la siguiente expresión general:

$$B = A \circ R$$

que representa la composición del conjunto A con la relación R para obtener el conjunto B . Por simplicidad estas ecuaciones suelen plantearse únicamente en universos discretos, con lo que A y B pueden expresarse como vectores, y R puede expresarse como una matriz cuyas dimensiones deben ser compatibles con las de A y B . La composición más empleada suele ser la composición *sup-min*.

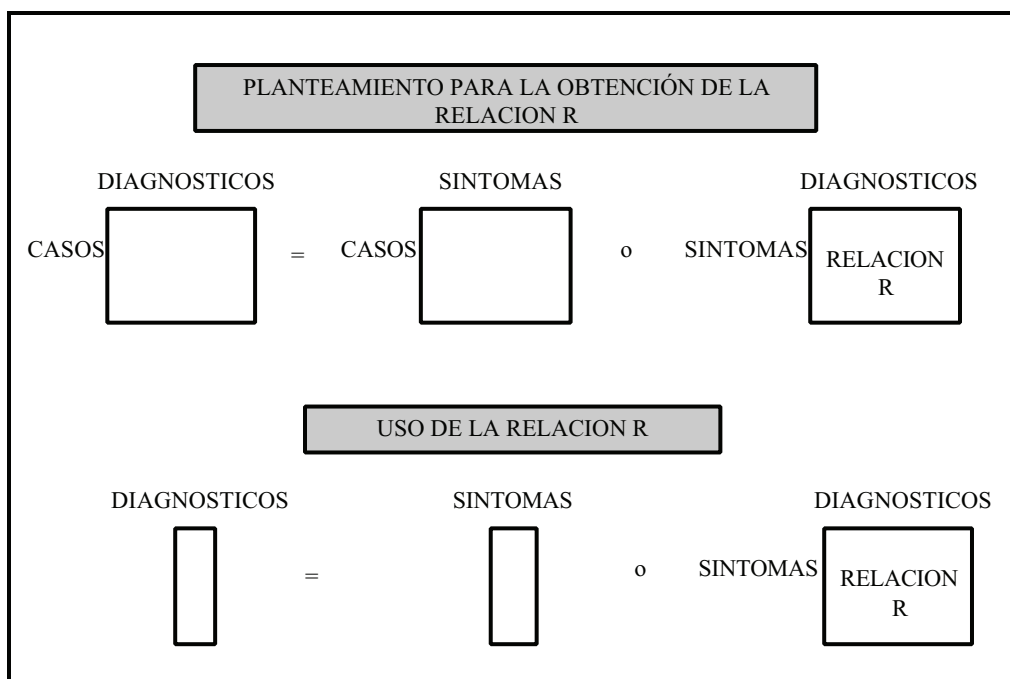
Su aplicación al diagnóstico se basa normalmente en codificar el conocimiento disponible en forma de relación borrosa R . Para ello se definen los universos *Síntomas* y *Diagnósticos*, y se construye la relación R en el universo $Síntomas \times Diagnósticos$, de forma que la ecuación relacional queda:

$$Diagnósticos = Síntomas \circ Relación_síntomas-diagnósticos$$

Los síntomas observados se representan en forma de vector, que a cada síntoma del universo *Síntomas* hace corresponder un valor en $[0,1]$ que indica en qué medida el síntoma ha sido observado. La composición de los síntomas con la relación permite obtener el vector de diagnósticos, que a cada diagnóstico del universo *Diagnósticos* hace

corresponder un número en $[0, 1]$ que es una medida de la posibilidad, creencia, etc., de la validez del diagnóstico. Este enfoque ha sido utilizado por ejemplo en [Kitoswski, Bargiel, 75], [Gazdik, 85], [Kitoswski, Bargiel, 86], [Lin et al., 91], [Ulieru, Isermann, 93], etc.

La relación R que liga los síntomas con los diagnósticos puede construirse a partir de datos empíricos o simulados, que asocien a cada caso o ejemplo disponible los síntomas observados y los diagnósticos identificados. Se plantea entonces una ecuación relacional borrosa con incógnita R que debe ser resuelta. La siguiente figura ilustra gráficamente este proceso:



Un inconveniente de este enfoque es la interpretación de los conjuntos borrosos que representan los síntomas observados y los diagnósticos deducidos. En efecto, no queda suficientemente claro si los grados de pertenencia expresan el grado en que el diagnóstico está presente, o la incertidumbre sobre la existencia de dicho diagnóstico. Esto hace que al final el conjunto borroso que sobre el universo de diagnósticos se obtiene se interprete como un orden de preferencia entre los diagnósticos disponibles, de forma que se escogen aquellos cuyos grados de pertenencia son mayores. Una correcta interpretación del resultado pasaría por una análisis semántico del tipo de composición empleado, como el realizado en el capítulo 4 de esta tesis. Esta limitación se debe en parte a que en estos enfoques no se hace una distinción formal entre los distintos tipos de incertidumbre existentes, lo que dificulta su interpretación, y empobrece parcialmente su capacidad de representación del conocimiento.

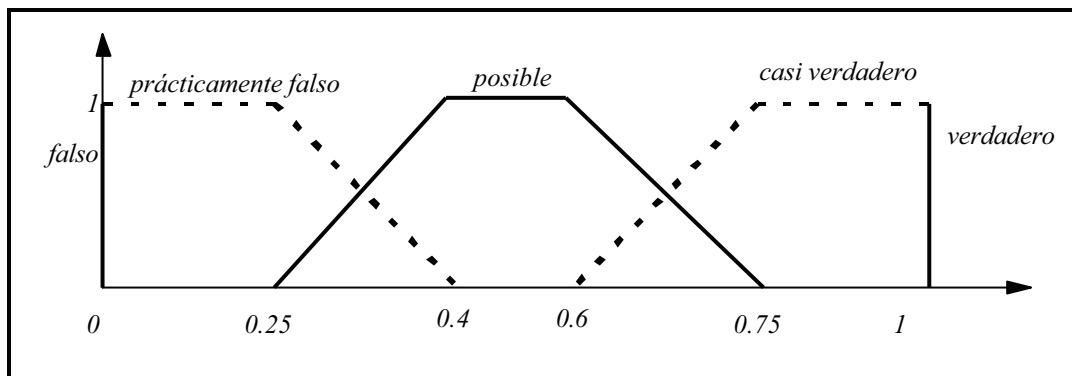
3.8.2 Enfoques lógicos y modelado de valores de verdad con números borrosos

Estos enfoques se basan en utilizar una lógica multivaluada o alguna extensión de la misma para llevar a cabo la inferencia con valores de verdad/incertidumbre, véase por ejemplo [Eklund, Forsstrom, 91].

Dentro de este tipo de enfoques cabe destacar el trabajo desarrollado para el sistema experto MILORD, descrito por ejemplo en [Mántaras, et al., 91], [Mántaras, Sierra, Agustí, 93], donde los valores de certidumbre son “borrosificados” y representados con un número borroso definido en el universo de valores de certidumbre $[0,1]$. Este número borroso no debe confundirse con los valores de verdad borrosos definidos en este capítulo, ya que por ejemplo no pueden ser utilizados como modificadores lingüísticos.

En MILORD el conocimiento se expresa mediante reglas a las que se asocia una certeza lingüística que se modela con un número borroso en el universo $[0,1]$ de certezas. La inferencia se basa en combinar los valores borrosos de certeza con los mismos operadores que se combinarían si fuesen numéricos, mediante la aplicación del principio de extensión a las lógicas multivaluadas, para lo que se usan las t-normas en las agregaciones conjuntivas, las t-conormas en las disyuntivas y la negación habitual.

Si por ejemplo se trabaja con 5 valores de certeza basados en conjuntos trapezoidales se puede plantear una partición borrosa (véase el apartado 5.6.3 del capítulo 5) como la que aparece en la siguiente figura (nótese las diferencias entre los valores de certezas así representados y los valores de verdad borrosos introducidos en este capítulo, y analizados en el capítulo 6):



La aplicación del principio de extensión a conjuntos trapezoidales permite deducir expresiones sencillas para el cálculo de operaciones básicas como sumas, restas, máximos y mínimos. Así por ejemplo dados dos conjuntos $P=(a,b,c,d)$ y $P'=(a',b',c',d')$ se tiene:

$$\begin{aligned} P+P' &= (a+a', b+b', c+c', d+d') \\ P-P' &= (a-a', b-b', c-c', d-d') \\ P.P' &= (a.a', b.b', c.c', d.d') \\ \min(P,P') &= (\min(a,a'), \min(b,b'), \min(c,c'), \min(d,d')) \\ \max(P,P') &= (\max(a,a'), \max(b,b'), \max(c,c'), \max(d,d')) \end{aligned}$$

expresiones con las que se puede extender el cálculo de las principales t-normas y t-conormas utilizando numerosos borrosos como argumentos. Para obtener la certeza de la conclusión de una regla se combina la certeza de las premisas con la certeza de la regla mediante la extensión borrosa de alguna función generadora de modus ponens asociada a alguna de las implicaciones de la lógica multivaluada. Por último para mantener cerrado el conjunto de valores de certeza se reducen los valores que resulten de los cálculos a alguno de los valores definidos en la partición.

La distinción entre distintos tipos de incertidumbre, en concreto entre la incertidumbre en su sentido más restringido y la verdad gradual no queda clara en los artículos consultados, por lo que no parece que exista una distinción suficientemente formal entre ambos.

3.8.3 Cálculo de significados con reglas borrosas, el enfoque numérico

Uno de los enfoques más extendidos de la aplicación de la lógica borrosa a la inferencia y en concreto al diagnóstico es el basado en el cálculo del significado (“*meaning computation*”) de la conclusión. Este enfoque propuesto originalmente por Zadeh, ha sido adoptado por ejemplo en [Martin-Clouaire, Prade, 84], [Buisson, Farreny, Prade, 85], [Lebailly, Martin-Clouaire, Prade, 86], [Sanz et al, 93], [Dounias, Tsourveloudis, 95], [Grabot, Caillaud, 96]. Este enfoque se caracteriza por el hecho de que para cada regla, el significado de la conclusión se calcula mediante la aplicación de la regla composicional de inferencia o algún mecanismo de inferencia similar. El conocimiento se formula en forma de reglas, por ejemplo:

$$\begin{aligned} &si \\ & \quad s_1 \text{ es } P_1, \end{aligned}$$

s_2 es P_2 ,
 ...
 entonces
 d_1 es Q

donde s_1, s_2, \dots pueden corresponder con síntomas observables, y d_1 puede corresponderse con alguno de los diagnósticos posibles, o con alguna variable del sistema cuyo valor debe ser estimado. Como ya hemos dicho el valor Q' que se deduce de unas observaciones P_1', P_2', \dots se obtiene calculándolo generalmente mediante la aplicación de la regla composicional de inferencia. Si d_1 es una variable más del sistema entonces Q' será una distribución de posibilidad en el universo de dicha variable que permite tener una estimación de su o sus posibles valores. Si d_1 es un diagnóstico entonces normalmente Q' será una distribución de posibilidad en el universo de valores de verdad $[0, 1]$.

Este es el enfoque que se ha adoptado en esta tesis, y que se desarrolla en profundidad a lo largo de los siguientes capítulos, ya que su riqueza semántica hace que puedan expresarse y distinguirse con claridad los distintos tipos de incertidumbre comentados en el capítulo 2. Una correcta interpretación permite obtener una mejor capacidad de representación del conocimiento que la de los otros enfoques.

3.8.4 Medidas de incertidumbre y reglas simbólicas, el enfoque simbólico

Un enfoque alternativo al anterior se basa en complementar reglas de producción simbólicas mediante medidas de incertidumbre como se ha hecho por ejemplo en el sistema MYCIN utilizando los coeficientes de certeza.

A cada regla se le asignan uno o más coeficientes que miden la certeza o confianza que se tiene en dicha regla. En el ámbito de la lógica borrosa estos coeficientes suelen ser o suelen estar directamente relacionados con medidas de posibilidad y necesidad. Así por ejemplo en [Lebailly, Martin-Clouaire, Prade, 86] a cada regla se le asigna una pareja de coeficientes (s, n) que miden respectivamente en qué medida es suficiente que las hipótesis sean ciertas para que la conclusión también lo sea, y en qué medida es necesario que las hipótesis sean ciertas para que la conclusión también lo sea.

La teoría de la posibilidad permite interpretar l -s como una medida de posibilidad de que la conclusión sea falsa siendo verdadera la hipótesis, y n se interpreta directamente como una medida de necesidad. Así a partir de la regla y de la posibilidad y necesidad de las

premisas puede deducirse la posibilidad y necesidad de la conclusión según la regla de inferencia:

$$\begin{aligned} & \text{si } p \text{ entonces } q \text{ con } (s,n) \\ & \text{necesidad de que } p \text{ sea verdadero} = N(p) \\ & \text{posibilidad de que } p \text{ sea verdadero} = \Pi(p) \quad (\text{con } \max(\Pi(p), 1-N(p))=1) \end{aligned}$$

$$\begin{aligned} & \text{necesidad de que } q \text{ sea verdadero} = N(q) = \min(N(p), s) \\ & \text{posibilidad de que } q \text{ sea verdadero} = \Pi(q) = \max(\Pi(p), 1-n) \end{aligned}$$

El esquema de inferencia anterior se complica cuando las proposiciones involucradas son borrosas, dando lugar a fórmulas aproximadas, que se deducen de garantizar para proposiciones borrosas la condición $\max(\Pi(p), \Pi(\neg p))=1$, que sólo se verifica con proposiciones clásicas.

El enfoque que se propondrá en esta tesis en el capítulo 7 permite combinar los enfoques numéricos y simbólicos aplicados a sistemas de reglas, y permite hacer una distinción formal de certidumbre y grados de verdad parciales, distinción implícita en la regla composicional de inferencia cuando se utilizan los modelos de reglas habituales. El análisis de los principales modelos de reglas se llevará a cabo en el capítulo 4, los sistemas de reglas en el capítulo 5, y el sistema de diagnóstico propuesto en la tesis en el capítulo 7.

4. Estudio de la inferencia en lógica borrosa

Una de las ventajas de la teoría de conjuntos borrosos es por un lado su riqueza para la representación del conocimiento, especialmente flexible en comparación con los conjuntos clásicos, y por otro las distintas posibilidades que ofrece para el modelado de esquemas de razonamiento aproximado, capaces de razonar simultáneamente con la vaguedad y la incertidumbre inherentes en el conocimiento humano.

En este capítulo se hace un análisis de la regla de inferencia del Modus Ponens Generalizado, basada en la regla composicional de inferencia (RCI), procurando ofrecer un marco suficientemente general, capaz de englobar muchas de las distintas formas en las que dicha regla ha sido aplicada. Se estudia la RCI desde un punto de vista semántico, analizando varias de las posibles parejas de operadores que en dicha regla intervienen (implicaciones borrosas y funciones generadoras de modus ponens) a las que hemos llamado modelos de reglas. Se profundiza en las diferencias existentes entre unos y otros modelos, para así poder encontrar criterios para la interpretación y simplificación de cada uno de ellos, y de este modo decidir en aplicaciones prácticas para cada tipo de regla qué modelo debe ser aplicado.

4.1 Organización del capítulo

El trabajo que se presenta en este capítulo es parte de la aportación original de esta tesis y en él se estudian los principales mecanismos de inferencia para aplicar la regla del modus ponens generalizado al ámbito de la lógica borrosa. Sólo se ha optado por el estudio de la regla del modus ponens, ya que en sistemas de diagnóstico de procesos industriales, a los que se pretende aplicar los resultados obtenidos, y en general en sistemas de diagnóstico en continuo, el razonamiento es hacia adelante, como ya se comentó en el capítulo 2, lo que implica que la principal regla de inferencia sea la regla del modus ponens (aunque se podría hacer un estudio paralelo de la regla del modus tollens). A lo largo de este capítulo se revisarán por tanto los conceptos de condicionales, implicaciones e inferencia en lógica borrosa. Nos restringiremos al estudio de una regla con un antecedente y un consecuente, ya que en el siguiente capítulo analizaremos el comportamiento de sistemas más complejos, estudiando reglas con varios antecedentes, y abordando el comportamiento de grupos de reglas paralelas o encadenadas.

En primer lugar se introduce la regla del modus ponens o modus ponens generalizado tal y como se conoce en lógica borrosa, y se repasan brevemente algunos de los estudios que sobre su comportamiento se han hecho desde distintas perspectivas.

Los siguientes dos apartados del capítulo están dedicados a una presentación formal de los dos mecanismos más habituales de aplicar el modus ponens generalizado, la regla composicional de inferencia (RCI) y la modificación por compatibilidad o similitud (MS), y se repasan los principales operadores que intervienen en ambos. Se introduce además el concepto de modelo de regla, que es el par de operadores (M,I) que deben ser escogidos para poder calcular la RCI, o bien el par de operadores (S,I) para la aplicación de la MS.

A continuación, a lo largo de cuatro apartados, se analiza el comportamiento de los modelos de reglas propuestos bajo distintas hipótesis. En primer lugar se observa la modificación del consecuente cuando la observación es puntual. En segundo lugar se analizan las condiciones que deben darse para que la RCI se reduzca a la MS. En tercer lugar se calculan los núcleos de las conclusiones para los modelos basados en n-implicaciones y los soportes para los basados en p-implicaciones. En cuarto y último se analiza la condición de activación de cada uno de los modelos y el nivel de indeterminación que aparece en la conclusión. Por un lado, los resultados obtenidos pueden ser utilizados como una guía para la selección de operadores en la inferencia borrosa. Por otro, el procedimiento seguido es una metodología a seguir para el análisis semántico de modelos de regla para la RCI.

En el siguiente apartado, y tras un breve resumen del comportamiento de los principales modelos de regla analizados, donde se resaltan sus principales diferencias semánticas, se propone una interpretación semántica para el comportamiento de cada uno de ellos y se clasifican en distintas categorías claramente diferenciadas, basadas en sus características semánticas.

A continuación se propone una generalización de las medidas de posibilidad y necesidad habituales (introducidas en los capítulos 2 y 3), basadas en los modelos de regla, que además de mejorar algunas de las propiedades poco deseables de estas últimas pueden interpretarse como un grado de activación de las reglas, ya que miden la incertidumbre de la conclusión obtenida.

Finalmente, en base a los resultados obtenidos durante el análisis del comportamiento de los modelos, se propone un esquema de inferencia simplificado, del tipo MS pero con dos índices de similitud, que permite conjuntar la riqueza semántica de la RCI con las ventajas computacionales de la MS, de cara a su aplicación práctica a un sistema experto de diagnóstico basado en reglas.

4.2 La regla del Modus Ponens Generalizado

Una de las principales reglas de inferencia en lógica borrosa es el modus ponens generalizado, al que llamaremos de forma abreviada MPG, y que permite inferir una conclusión aproximada a partir de una regla expresada con proposiciones vagas o borrosas, y una observación también expresada mediante una proposición vaga.

La regla del MPG puede formularse del siguiente modo. Sean U y V dos universos de discurso en los que toman sus valores dos variables u y v respectivamente. Sean P y Q dos conjuntos borrosos definidos en U y V con funciones de pertenencia $P(u)$ y $Q(v)$. A partir de la regla “si u es P entonces v es Q ” (que no debe ser confundida con una regla de inferencia), expresada de forma abreviada como $P(u) \rightarrow Q(v)$, y a partir de una observación de la variable u dada por “ u es $P'(u)$ ” siendo $P'(u)$ un conjunto borroso, la regla del MPG pretende inferir “ v es $Q'(v)$ ” donde $Q'(v)$ es un conjunto borroso que restringe los valores que la variable v puede tomar. Cuando el MPG se basa en la regla composicional de inferencia (RCI), $Q'(v)$ se calcula a partir de $P'(u)$ y de una relación borrosa $R(v/u)$, función de las variables u y v , que denotaremos normalmente como $R(u,v)$. De forma abreviada podemos escribir:

$$\frac{\begin{array}{l} \text{si } u \text{ es } P \\ u \text{ es } P' \end{array} \quad \text{entonces} \quad v \text{ es } Q}{v \text{ es } Q'}$$

o bien:

$$\frac{\begin{array}{l} P(u) \\ P'(u) \end{array} \quad \longrightarrow \quad Q(v)}{Q'(v)}$$

Dados dos universos de discurso U y V y dos variables u y v que toman sus valores en U y V respectivamente, un condicional borroso $R(u,v)$ es una relación borrosa $R(u,v):U \times V \rightarrow [0,1]$ que representa el grado de implicación $u \rightarrow v$ para cada valor de u y v , o lo que es lo mismo el grado en que u implica a v para cada par (u,v) . Cuando los valores que pueden tomar las variables del problema se modelan mediante distribuciones de posibilidad, un condicional borroso entre dos variables u y v puede interpretarse como la distribución de posibilidad condicional de v condicionada a u , denotada también como $\pi(v/u)$. A toda regla $P(u) \rightarrow Q(v)$ (o de forma abreviada $P \rightarrow Q$) se le asocia un condicional borroso que la modela y que no es más que la expresión numérica de su significado, al igual que al concepto “*temperatura alta*” se le puede asociar como

significado un conjunto borroso definido en el universo de las temperaturas a las que dicho concepto se refiere.

Uno de los principales problemas que la lógica borrosa aborda consiste precisamente en definir el o los posibles condicionales asociados a la regla, a partir de los grados de verdad o posibilidad de u y v (que para la regla $P(u) \rightarrow Q(v)$ vienen dados por $P(u)$ y $Q(v)$ respectivamente), de forma que la regla de inferencia escogida presente un comportamiento suficientemente bueno de cara a obtener conclusiones llamémoslas lógicas o razonables, acordes a lo que se desea expresar. Para nosotros una conclusión será lógica cuando sea coherente con la intuición y sea por tanto interpretable mediante el sentido común. Por otro lado, debido a su dependencia con el mecanismo de inferencia empleado, veremos que no se puede desligar la definición de condicional del procedimiento utilizado para inferir.

Dentro de los distintos esquemas de inferencia que han sido estudiados y propuestos, dos de ellos destacan por la atención que se les ha dedicado en multitud de trabajos en el ámbito de la lógica y el control borrosos. Por un lado está la llamada regla composicional de inferencia, que llamaremos RCI, introducida por Zadeh en sus primeros trabajos. A este respecto puede por ejemplo consultarse [Zadeh, 65] donde se introduce la composición de relaciones, de la que la RCI es un caso particular, o [Zadeh, 73] donde ya se propone la RCI como una forma de aplicar el MPG. En la aplicación de la RCI intervienen por un lado una función de implicación que permite obtener el condicional borroso $R(u,v)$ con el que modelar la regla, a partir de $P(u)$ y $Q(v)$, y por otro una función de agregación de tipo conjuntivo, que permite combinar el condicional $R(u,v)$ con la observación $P'(u)$. Al par compuesto por la función de implicación y la función de agregación lo hemos denominado modelo para la regla $P \rightarrow Q$. La regla de inferencia RCI es la de más aceptación en la mayor parte de trabajos y aplicaciones de lógica borrosa para la aplicación del MPG.

Por otro lado está el razonamiento por analogía ([Turksen, Zhong, 88] y [Turksen, Zhong, 90]), que da lugar a una regla denominada en algunos trabajos modificación por compatibilidad o similitud ([Cross, Sudkamp, 93], [Cross, Sudkamp, 94]), que abreviaremos con MS, y que es una regla de inferencia más sencilla que la RCI, pero como se mostrará en este capítulo, menos rica en matices. De hecho probaremos que bajo ciertas condiciones la RCI se reduce a la MS, lo que permite obtener un mejor conocimiento del comportamiento de la RCI. Para aplicar la MS se definen por un lado una medida de similitud entre el antecedente de la regla y la observación, y por otro, y al igual que con la RCI, una función de implicación que se utiliza para modificar el consecuente en base a la medida de similitud entre observación y antecedente, obteniéndose de este modo la conclusión.

De los numerosos estudios que del MPG se han hecho gran parte se han orientado al estudio teórico de las propiedades matemáticas básicas exigibles al mecanismo de inferencia basado en la RCI, para a partir de ellas, encontrar familias de operadores apropiadas para su cálculo (véase por ejemplo [Dubois, Prade, 84], [Trillas, Valverde, 85], [Trillas, Valverde, 85b], [Bouchon, 87]).

Estos trabajos han dado poco a poco lugar, por un lado, a una teoría matemática de la lógica borrosa, los condicionales borrosos y la inferencia, cada vez más consolidada, como puede apreciarse en [Trillas, 93] o en [Rodríguez-Soto, 95], tesis doctoral dirigida por el mismo profesor Trillas, en la que se trata tanto el problema de la inferencia como el de la representación de predicados vagos, siempre desde una óptica muy teórica, dejando el terreno de las interpretaciones por explorar.

Por otro lado otros autores han profundizado más en el estudio, desde un punto de vista semántico, de algunas de las funciones de implicación más comunes, intentando de este modo explicar las diferencias semánticas existentes entre las mismas, asociando a cada uno de los principales tipos de implicaciones una interpretación semántica en términos de cualificación veritativa, posibilista o en términos de certidumbre. En esta línea de trabajo figuran principalmente [Dubois, Prade, 92] y [Dubois, Prade, 93]. Otro trabajo interesante es [Martin-Clouaire, 89] donde se hace un estudio del comportamiento semántico, analizando principalmente el núcleo y el soporte de la conclusión que se obtiene a partir de varios modelos de reglas. Son estos trabajos los que principalmente han inspirado y guiado el trabajo de este capítulo.

Existen también muchos enfoques más prácticos entre los que podemos por ejemplo señalar [Gupta, Qi, 91] o [Mizumoto, Zimmermann, 80]. El buen comportamiento numérico de los sistemas borrosos, como se ha demostrado en numerosos trabajos y aplicaciones de aproximación funcional ([Nakanishi, Turksen, Sugeno, 93], [Wang, 92]), predicción ([Jou, 93], [Wang, 92b]), y control borroso ([Lee, 90], [Lee, 90b]), ha propiciado el que muchos autores hayan optado por un enfoque principalmente numérico, alejándose de los postulados axiomáticos, y estudiado el comportamiento de la inferencia buscando mejorar sus capacidades numéricas aproximadoras o interpoladoras. Esto ha dado lugar a mecanismos de inferencia no monótonos y en muchas ocasiones muy próximos a los procedimientos de evaluación de algunas estructuras de redes neuronales (véase por ejemplo [Mizumoto, 96] donde se propone agregar conclusiones con una suma ordinaria, o [Jang, Sun, 93] y [Wienholt, 93] donde se demuestra la equivalencia entre dos tipos muy concretos de estructura de red neuronal y sistema borroso).

Sin embargo se echa en falta un análisis comparativo sistemático, desde un punto de vista semántico, de los posibles modelos para una regla borrosa, tanto para la aplicación de la RCI como la de la MS, que permita comprender sus diferencias y por tanto orientar su

elección, lo que ha motivado gran parte del trabajo desarrollado en esta tesis, que se expone en este capítulo.

4.3 La regla composicional de inferencia

La RCI introducida por Zadeh ([Zadeh, 73]) es una regla de inferencia que permite aplicar el MPG, una de las principales reglas de inferencia en el ámbito de la lógica borrosa. El MPG permite inferir una conclusión $Q'(v)$ a partir de una regla $P(u) \rightarrow Q(v)$ y una observación $P'(u)$, en general distinta de $P(u)$. Según la RCI la conclusión $Q'(v)$ viene dada por:

$$Q'(v) = \sup_u M(P'(u), R(u, v))$$

donde $R(u, v)$ es el condicional que modela a la regla y M es una función de agregación (de tipo conjuntivo). Gran parte de la polémica surgida entorno a la aplicación de esta regla y a la definición del condicional R y de la función M implicados, radica en el hecho de que no todas las combinaciones M - R permiten inferir $Q(v)$ cuando la observación es $P(u)$, propiedad que hemos llamado modus ponens estricto, de forma abreviada MPE, y que parece lógicamente exigible a una regla de MP (o MPG). Bajo este punto de vista la RCI ha sido estudiada por ejemplo en [Dubois, Prade, 84], [Trillas, Valverde, 85], y [Trillas, Valverde, 87] donde se establecen condiciones básicas para el cumplimiento del MPE.

4.3.1 Condicional borroso basado en la RCI: modelos de reglas

Diremos que una relación borrosa $R(u, v): U \times V \rightarrow [0, 1]$ es un **M-condicional** para la regla $P(u) \rightarrow Q(v)$ si verifica:

$$M(P(u), R(u, v)) \leq Q(v)$$

donde M es un operador de agregación de tipo conjuntivo, que cumple:

$$\begin{aligned} M(1, 1) &= 1 \\ M(0, 1) &= M(1, 0) = 0 \\ M(x, y) &\text{ no decreciente ni con } x \text{ ni con } y \end{aligned}$$

A la desigualdad anterior la llamaremos desigualdad del modus ponens, DMP, y expresa el hecho de que cuanto mayor sea el grado de verdad de u y mayor el grado en que u implique a v , mayor debe ser el grado de verdad de v . Esta definición es prácticamente idéntica a la que se propone en [Trillas, Alsina, 92] o [Trillas, 93] para definir un condicional, donde se dice que R es un M - P - Q -condicional resaltándose de este modo la

dependencia del condicional con los conjuntos P y Q que intervienen en la regla y con la función de agregación M .

Sin embargo el cumplimiento de la DMP no garantiza el cumplimiento del MPE por lo que es necesario exigir algún tipo de restricción adicional. En lo que sigue exigiremos que además se cumpla:

$$M(1,y)=y \quad y \quad R(u,v)=1 \text{ si } u \in [P]$$

lo que garantiza el MPE bajo la RCI, ya que el supremo de la expresión $M(P(u), R(u,v))$, que es $Q(v)$, se alcanza para $u \in [P]$.

Diremos que un par (M,R) es un **modelo** para la regla $P(u) \rightarrow Q(v)$ si R es un M -condicional para la regla y el par (M,R) verifica el MPE. En este caso diremos también que M es una **función generadora de modus ponens** (FGMP) para R ([Trillas, Valverde, 85]).

En algunas de las referencias señaladas (por ejemplo [Trillas, 93]) se exige la reflexividad o f -reflexividad del condicional. Cuando los universos U y V coinciden, esta propiedad se expresa como $R(u,v)=1$, mientras que cuando son distintos se puede formular mediante una función biyectiva f de U a V como $R(u,f(u))=1$. Sin embargo esta propiedad nos parece excesivamente estricta y excluye a numerosos condicionales, muchos de los cuales han sido utilizados con éxito en numerosas aplicaciones. Cuando se estudia la inferencia dentro de un mismo universo el mecanismo de inferencia tiene que ser lógicamente reflexivo ya que de la regla $P(u) \rightarrow Q(v)$ y de la observación $P'(u)$ debe poder inferirse tanto $Q'(v)$ como $P'(u)$. Sin embargo en el caso por ejemplo de un sistema experto basta con que la gestión de la base de hechos se encargue de añadir $Q'(v)$ sin eliminar $P'(u)$, siendo innecesario que la relación condicional sea ella misma reflexiva, aunque el mecanismo global de inferencia sí lo sea. Por otro lado, cuando se infiere en universos distintos, si se garantiza que la base de hechos va a contener tanto a $Q'(v)$ como a $P'(u)$ después de aplicar la regla del modus ponens, se está verificando también una cierta reflexividad. Ahora bien exigir $R(u,f(u))=1$ implica la existencia de una función f que permita determinar qué elementos de v están asociados con los elementos de u , que es precisamente lo que la regla pretende hacer a nivel local. Creemos que la f -reflexividad está implícita en la regla, debido al desconocimiento de la verdadera función f , y el hacerla explícita significaría conocer f con lo que la regla sería innecesaria.

4.3.2 Funciones de implicación y funciones generadoras de modus ponens

4.3.2.1 Implicaciones clásicas

En lógicas clásicas multivaluadas se asume que el grado de verdad de la implicación $u \rightarrow v$ depende únicamente de los grados de verdad de u y v (véase [Trillas, Valverde, 85] o [Dubois, Prade, 84]) es decir que:

$$R(v/u) = R(u, v) = t(u \rightarrow v) = I(t(u), t(v))$$

donde I se denomina función de implicación.

En el marco de la lógica clásica, donde no existen grados intermedios de verdad o pertenencia, la implicación comúnmente aceptada desde un punto de vista formal es la llamada implicación material ([Kleene, 67]), dada por la función booleana $I(x, y) = \neg x + y$ cuya tabla de verdad es:

$t(u)$	$t(v)$	$t(u \rightarrow v)$
0	0	1
0	1	1
1	0	0
1	1	1

Como veremos a lo largo de este capítulo en la tabla anterior, dada una proposición cualquiera w , $t(w)$ puede interpretarse como el grado de verdad de w , o lo que es lo mismo, cómo de necesario es que w sea verdadero. El condicional asociado a una regla $P(u) \rightarrow Q(v)$ obtenido a partir de la implicación material viene dado por:

$$R(v / u) = (P(u) \cap Q(v)) \cup (\neg P(u) \cap V(v))$$

Aunque en lógica clásica no ha sido muy utilizado, y sí criticado su uso (véase por ejemplo [Kleene, 67] en su primer capítulo donde se afirma que la implicación material es la única definición razonable de implicación) algunos autores han propuesto la intersección como implicación lógica, con función booleana $I(x, y) = x.y$, y cuya tabla de verdad es:

$t(u)$	$t(v)$	$t(u \rightarrow v)$
0	0	0
0	1	0
1	0	0
1	1	1

A diferencia del condicional material, veremos a lo largo de este capítulo que en este caso $t(w)$ debe interpretarse de distinto modo, ya que expresa el grado de posibilidad de w , o lo que es lo mismo cómo de posible es que w sea verdadero. Esta implicación da lugar al denominado condicional material restringido ([Trillas, 92]):

$$R(v / u) = P(u) \cap Q(v)$$

que es el producto cartesiano de P y Q . La generalización de las dos implicaciones clásicas al ámbito borroso permite obtener dos grandes familias de funciones de implicación borrosas, aquellas que generalizan la implicación material, y que llamaremos n-implicaciones o implicaciones de necesidad, tratadas principalmente en estudios con clara orientación teórica y matemática (implicaciones de Lukasiewicz, Kleene, etc.), y aquellas que generalizan el producto cartesiano, y que llamaremos p-implicaciones o implicaciones de posibilidad, principalmente utilizadas en aplicaciones prácticas de sistemas borrosos con marcado carácter numérico (por ejemplo las implicaciones de Mandani o Larsen, definidas en el capítulo anterior, ampliamente utilizadas en sistemas de control borroso), a menudo estudiadas desde un punto de vista teórico después de comprobadas sus buenas cualidades numéricas.

Además de coincidir con las implicaciones clásicas en los contornos del dominio de definición, parece razonable exigir a una función de implicación $I(x,y)$ que sea no decreciente con y , de forma que si la conclusión es más verdadera, la implicación sea al menos igual de verdadera, pero en ningún caso más falsa. En estas consideraciones nos basaremos para definir lo que para nosotros será una función de implicación borrosa.

4.3.2.2 Implicaciones borrosas y funciones generadoras de modus ponens

Diremos que una función $I(x,y)$ es una función de implicación borrosa, o simplemente una **función de implicación**, o implicación, cuando verifique:

$$\begin{cases} I: [0,1] \times [0,1] \rightarrow [0,1] \\ I(1,y) = y \\ I(x,y) \text{ no decreciente con } y \end{cases}$$

Diremos que una implicación I es una **n-implicación** cuando generalice la implicación material, es decir $I(0,y)=1$, y cuando además $I(x,y)$ sea no creciente con x .

Así mismo diremos que una implicación I es una **p-implicación** cuando generalice el producto cartesiano, es decir $I(0,y)=0$, y cuando además $I(x,y)$ sea no decreciente con x .

Utilizaremos también la expresión **n-modelo** para referirnos a los modelos cuya función de implicación sea una n-implicación, y **p-modelos** cuando sea una p-implicación.

Cuando se admiten grados intermedios de verdad pueden definirse múltiples funciones de agregación conjuntivas (véase el capítulo 2). Hemos visto con anterioridad que el MPE queda garantizado si se cumplen la DMP, $M(1,y)=y$ y $R(u,v)=1$ si $u \in [P]$, condición ésta última equivalente a $I(1,y)=y$. Para garantizar el cumplimiento de la DMP puede obtenerse la máxima función de agregación que para una implicación dada cumple el modus ponens ([Trillas, Valverde, 85]), resolviendo la desigualdad:

$$M_{max}(x,y) = \inf \{ c \in [0,1] / I(x,c) \geq y \}$$

Obviamente para cualquier operador de agregación de tipo conjuntivo M tal que $M(x,y) \leq M_{max}(x,y)$ se cumple que la relación dada por $R(u,v)=I(P(u),Q(v))$ es un M -condicional para la regla $P(u) \rightarrow Q(v)$, o dicho de otro modo el par (M,R) es un modelo para la regla $P(u) \rightarrow Q(v)$. Recordemos que a M la llamaremos función generadora de modus ponens para la implicación I . Por extensión, si (M,R) es un modelo para la regla $P(u) \rightarrow Q(v)$, también diremos que (M,I) es un modelo para dicha regla, entendiéndose que I se utiliza para generar el condicional de la regla.

Es interesante señalar que si M es una FGMP para una n-implicación I_n cualquiera, entonces es FGMP para cualquier p-implicación I_p . En efecto al ser $I_n(x,y)$ no decreciente con x e $I_n(1,y)=1$ se tiene $I_n(x,y) \geq y$. Por otro lado $I_p(x,y)$ es no decreciente con ninguno de sus argumentos, y como $I_p(1,y)=y$ se tiene $I_p(x,y) \leq y$. Por tanto:

$$I_n(x,y) \geq I_p(x,y)$$

de donde:

$$M(x, I_p(x, y)) \leq M(x, I_n(x, y)) \leq y$$

y por tanto M es también FGMP para I_p .

Si $R(u, v) = I(P(u), Q(v))$ entonces $R(u, v)$ verifica:

$$\begin{cases} R: U \times V \rightarrow [0, 1] \\ R(u, v) = 1 \text{ si } u \in [P], v \in [Q] \\ R(u, v) = 0 \text{ si } u \in [P], v \in [1-Q] \\ R(u, v) \text{ no decreciente con } Q(v) \end{cases}$$

4.3.2.3 Relación de inclusión entre la conclusión y el consecuente

Es interesante analizar en este punto la relación de inclusión existente entre la conclusión y el consecuente de una regla, según se emplee un p-modelo o un n-modelo para la aplicación de la RCI, ya que una primera caracterización radica precisamente en su diferente comportamiento a este respecto.

Para un modelo de regla compuesto por una n-implicación y una FGMP se tiene siempre:

$$Q'(v) \geq Q(v)$$

Efectivamente una n-implicación $I_n(x, y)$ es una función no creciente con x , y al ser $I_n(1, y) = y$ se tiene $I_n(x, y) \geq y$. Como además $M(1, y) = y$ y $M(x, y)$ es creciente con su segundo argumento se tiene:

$$Q'(v) = \sup_u M(P'(u), I_n(P(u), Q(v))) \geq \sup_u M(P'(u), Q(v)) = Q(v)$$

Este tipo de implicaciones han sido denominadas implicaciones expansivas en algunos trabajos, por ejemplo en [Turksen, Tian, 93] donde se diferencia entre implicaciones expansivas y reductoras (que se corresponden respectivamente con las n y p-implicaciones definidas en esta tesis). Veremos que el comportamiento de las implicaciones expansivas consiste en modificar $Q(v)$ de forma que cuanto menor es la intersección de P' en P mayor es la conclusión $Q'(v)$, hasta llegar a ser igual a V cuando no se produce la activación de la regla. Veremos que esto hace que las conclusiones de ambos tipos de implicaciones puedan y deban interpretarse de distinto modo.

Por otro lado, para un modelo compuesto por una p-implicación I_p y una FGMP M se tiene siempre:

$$Q'(v) \leq Q(v)$$

En efecto una p-implicación $I_p(x,y)$ es una función no decreciente con x , y al ser $I_p(1,y)=y$ se tiene $I_p(x,y) \leq y$. Como además $M(1,y)=y$ y $M(x,y)$ es creciente con su segundo argumento se tiene:

$$Q'(v)=\sup_u M(P'(u), I_p(P(u), Q(v))) \leq \sup_u M(P'(u), Q(v))=Q(y)$$

Como ya hemos señalado en [Turksen, Tian, 93] este tipo de implicaciones se denominan reductoras. Veremos que su comportamiento consiste en modificar $Q(v)$ de forma que cuanto menor es la intersección de P' y P menor es la conclusión $Q'(v)$, hasta llegar a ser igual a \emptyset cuando no se produce la activación de la regla.

A lo largo de este capítulo propondremos una interpretación de la conclusión obtenida con una n-implicación en términos de distribución de necesaria posibilidad, lo que resulta en cierto modo coherente con la desigualdad $Q'(v) \geq Q(v)$ (donde la distribución de referencia es $Q(v)$) de forma que cuanto menos se parece la observación a la hipótesis mayor es $Q'(v)$ y por tanto menos restringe el rango de valores que puede tomar la variable v . De hecho veremos cuando analicemos la condición de activación de los modelos de reglas, que para este tipo de implicaciones la conclusión trivial, es decir cuando la observación y la hipótesis no se parecen en nada, es la distribución de necesaria posibilidad menos restrictiva, esto es el universo de discurso V de la variable v .

Del mismo modo veremos cómo las conclusiones obtenidas con p-implicaciones se interpretan como distribuciones de posible posibilidad, lo que resulta coherente con la desigualdad $Q'(v) \leq Q(v)$ (donde la distribución de referencia es $Q(v)$), de forma que cuanto menos se parezcan la hipótesis y la observación menor es la conclusión $Q'(v)$ obtenida, siendo el caso límite la conclusión trivial $Q'(v)=\emptyset$.

4.3.3 Generación de funciones de implicación

4.3.3.1 Principales funciones de implicación

En [Trillas, Valverde, 85b] se proponen como funciones de implicación que generalizan la implicación material a las r-implicaciones y s-implicaciones, y se analiza exhaustivamente gran parte de sus propiedades. Como se introdujo en el capítulo 3, una r-implicación (r proviene de ‘residual’) se define como:

$$I_R^T(x,y)=\sup\{c \in [0,1] / T(x,c) \leq y\}$$

donde T es una t-norma. Una implicación fuerte o s-implicación (la 's' viene de 'strong') se define como:

$$I_S^T(x,y) = S(1-x,y) = 1 - T(x, 1-y)$$

donde S es la t-norma dual de T . En [Trillas, Valverde, 85] y [Dubois, Prade, 84] se obtienen las máximas funciones generadoras de modus ponens de estas implicaciones. Para las r-implicaciones son la t-norma de la que provienen:

$$M_R^T(x,y) = T(x,y)$$

y para las implicaciones las pseudo-conjunciones definidas, para cada s-implicación, por:

$$M_S^T(x,y) = \inf\{c \in [0,1] / S(1-x,c) \geq y\} = \inf\{c \in [0,1] / 1 - T(x, 1-c) \geq y\}$$

Como puede apreciarse las cuatro funciones provienen de una t-norma T . La notación utilizada pretende resaltar el tipo de implicación (R o S , residual o fuerte) así como la t-norma de la que la implicación proviene. Recordemos que las t-normas positivas y negativas serán denotadas por T^+ y T^- respectivamente, las t-conormas positivas y negativas por S^+ y S^- .

Cuando la t-norma (respectivamente t-conorma) admite un generador aditivo $f(x)$ (respectivamente $g(x)$) tal y como se describe en el capítulo 3, los cuatro tipos de implicaciones pueden ponerse en función de estos generadores aditivos como sigue:

$$M_R^T(x,y) = T(x,y) = f^{[-1]}(f(x) + f(y))$$

$$\begin{aligned} M_S^T(x,y) &= C(x,y) = g^{[-1]}(\max(g(y) - g(1-x), 0)) \\ &= \begin{cases} g^{[-1]}(g(y) - g(1-x)) & \text{si } x + y > 1 \\ 0 & \text{otro caso} \end{cases} \\ &= \begin{cases} 1 - f^{[-1]}(f(1-y) - f(x)) & \text{si } x + y > 1 \\ 0 & \text{otro caso} \end{cases} \end{aligned}$$

$$\begin{aligned} I_R^T(x,y) &= f^{[-1]}(\max(f(y) - f(x), 0)) \\ &= \begin{cases} f^{[-1]}(f(y) - f(x)) & \text{si } x > y \\ 1 & \text{otro caso} \end{cases} \end{aligned}$$

$$I_S^T(x,y) = g^{f^{-1}}(g(1-x) + g(y)) = 1 - f^{f^{-1}}(f(x) + f(1-y))$$

donde las funciones f , g , $f^{f^{-1}}$ y $g^{f^{-1}}$ son las definidas en el apartado 3.4.2.2 del capítulo 3. Puede fácilmente comprobarse que tanto las r-implicaciones como las s-implicaciones son ejemplos de n-implicaciones, mientras que las t-normas y pseudo-conjunciones lo son de las p-implicaciones, y estos cuatro tipos serán ampliamente analizados a lo largo de esta tesis. En los casos de $M_S^T(x,y)$ y $I_R^T(x,y)$ la definición por trozos debe hacerse para tener en cuenta los casos en que los argumentos de $f^{f^{-1}}$ y $g^{f^{-1}}$ se hacen negativos (esto es debido a que en rigor $f(f^{f^{-1}}(x)) = \min(x, f(0))$). En lo que sigue, y a menos que se especifique lo contrario sólo se considerarán t-normas continuas, y en concreto el operador \min y las t-normas arquímedeas.

4.3.3.2 Propiedades

En este apartado mencionaremos brevemente lo que entendemos por condicionales duales, y luego determinaremos el comportamiento de las cuatro implicaciones cuando sus argumentos o sus resultados toman valores en $\{0,1\}$.

- Dualidad n-implicación/p-implicación:

Diremos que una n-implicación I_n es dual de una p-implicación I_p cuando ambas verifiquen:

$$I_p(x,y) = 1 - I_n(x, 1-y)$$

Puede comprobarse fácilmente que las implicaciones M_S^T e I_R^T y M_R^T e I_S^T son duales:

$$M_S^T(x,y) = 1 - I_R^T(x, 1-y)$$

$$M_R^T(x,y) = 1 - I_S^T(x, 1-y)$$

- r-implicaciones:

En el caso de una r-implicación procedente de una t-norma arquímdea se tiene:

$$f^{f^{-1}}(f(y) - f(x)) = 1 \quad \text{ssi} \quad f(y) - f(x) \leq 0$$

Y dado el carácter de f (monótona decreciente) se tiene $I_R^T(x,y) = 1$ ssi $x \leq y$. Si la r-implicación proviene de la t-norma \min se puede comprobar fácilmente que se obtiene el mismo resultado. Por tanto para cualquier r-implicación:

$$I_R^T(x,y) = 1 \quad \text{ssi} \quad x \leq y$$

Si queremos determinar cuando la implicación toma el valor 0 debe resolverse la ecuación $I_R^T(x,y)=0$. Si T es arquímedea equivale a resolver $f^{[-1]}(f(y)-f(x))=0$. Si además es positiva entonces $f^{[-1]}$ vale 0 cuando su argumento se hace infinito, es decir para $f(y)-f(x)=\infty$ lo que se cumple para $y=0$ y $x \neq 0$. En el caso de la t-norma *min* puede comprobarse fácilmente que se obtiene el mismo resultado. Por tanto:

$$I_R^{T+}(x,y)=0 \text{ ssi } y=0 \text{ y } x \neq 0$$

Si la t-norma es negativa la solución viene dada por $f(y)-f(x) \geq f(0)$. Ahora bien, dado el carácter estrictamente decreciente el caso más favorable se da para y mínimo y x máximo, es decir $y=0$ y $x=1$, situación para la que justo se da la igualdad. Por tanto:

$$I_R^{T-}(x,y)=0 \text{ ssi } y=0 \text{ y } x=1$$

Por otro lado es inmediato comprobar que $I_R^T(0,y)=1$, $I_R^T(1,y)=y$ e $I_R^T(x,1)=1$. Para calcular $I_R^T(x,0)$ debemos distinguir entre t-normas positivas y negativas. Si T es positiva entonces en el caso continuo se tiene $I_R^{T+}(x,0)=f^{[-1]}(f(0)-f(x))$ con $f(0)=\infty$ por lo que $I_R^{T+}(x,0)=0$ salvo si $x=0$ en cuyo caso $I_R^{T+}(x,0)=1$. Puede comprobarse que la implicación residuada de la t-norma *min* se comporta igual. Si por el contrario T es negativa entonces $I_R^{T+}(x,0)=f^{[-1]}(f(0)-f(x))$ expresión que en general depende de f .

- s-implicaciones:

En el caso de una s-implicación procedente de una t-conorma arquímedea positiva se tiene:

$$g^{[-1]}(g(1-x)+g(y))=1 \text{ ssi } g(1-x)+g(y)=\infty$$

La igualdad anterior se cumple para $1-x=1$ o $y=1$, es decir:

$$I_S^{T+}(x,y)=1 \text{ para } x=0 \text{ ó } y=1$$

Si la t-conorma hubiese sido negativa se tendría $g^{[-1]}(g(1-x)+g(y))=1$ ssi $g(1-x)+g(y) \geq g(1)$ y por tanto:

$$I_S^{T-}(x,y)=1 \text{ ssi } g(1-x)+g(y) \geq g(1)$$

condición que en general depende de g . Por ejemplo para la familia paramétrica de Yager, la igualdad queda:

$$(1-x)^p + y^p \geq 1$$

Para $p=1$ la condición equivale a $x \leq y$ (como era de esperar ya que se recupera la implicación de Lukasiewicz). Para $p < 1$ puede tenerse $I(x,y)=1$ para parejas (x,y) tales

que $x > y$, mientras que para $p > 1$ no es suficiente con $x \leq y$. Por tanto y comparando con las r-implicaciones podemos decir que cuanto menor es p menos restrictiva es la condición impuesta a x e y para obtener $I(x,y)=1$.

Por otro lado la resolución de $I_S^T(x,y)=0$ da directamente $1-x=0$ o bien $y=0$:

$$I_S^T(x,y)=0 \text{ ssi } x=1 \text{ o bien } y=0$$

Es también inmediato comprobar que $I_S^T(0,y)=1$, $I_S^T(1,y)=y$ e $I_S^T(x,1)=1$ así como $I_S^T(x,0)=1-x$.

- t-normas:

Dada la definición de $T(x,y)$, la solución a $T(x,y)=1$ es $x=y=1$. Por tanto:

$$M_R^T(x,y)=1 \text{ ssi } x=y=1$$

Para determinar cuándo $T(x,y)=0$ debemos distinguir según sea T positiva o negativa. Si T es positiva, entonces $T(x,y)=0$ para $x=0$ o $y=0$. Por tanto:

$$M_R^{T^+}(x,y)=0 \text{ ssi } x=0 \text{ o bien } y=0$$

Si T es negativa entonces $T(x,y)=0$ para $f(x)+f(y) \geq f(0)$ que en general depende de f :

$$M_R^{T^-}(x,y)=0 \text{ ssi } f(x)+f(y) \geq f(0)$$

En el caso de la familia de Yager para la que la t-norma viene dada por:

$$T(x,y) = 1 - \min\left(1, \left((1-x)^p + (1-y)^p\right)^{\frac{1}{p}}\right)$$

se tiene:

$$(1-x)^p + (1-y)^p \geq 1$$

Cuanto mayor es p mayores pueden ser los valores de x y de y que verifiquen $T(x,y)=0$. Para $p=1$ se tiene la condición $x+y \leq 1$ y cuanto menor es p más próximos a cero deben de estar x e y para que se cumpla $T(x,y)=0$.

Es inmediato comprobar que $T(0,y)=T(x,0)=0$, mientras que $T(1,y)=y$ y $T(x,1)=x$.

- pseudo-conjunciones:

Si $C(x,y)$ proviene de una t-norma arquimedea tenemos $C(x,y)=g^{f-1}(g(y)-g(1-x))$, y $C(x,y)=1$ equivale a:

$$g^{f-1}(g(y)-g(1-x))=1$$

Si S es positiva entonces $g^{f-1}(x)=1$ para $x=\infty$. Por tanto se tendrá $C(x,y)=1$ si $y=1$ y $x \neq 0$:

$$M_S^{T+}(x,y)=1 \text{ ssi } y=1 \text{ y } x \neq 0$$

Si S es negativa entonces $g^{f-1}(x)=1$ para $x \geq g(1)$. En este caso debe resolverse la ecuación $g(y)-g(1-x) \geq g(1)$. Ahora bien, al ser g estrictamente creciente y positiva, y definida en $[0,1] \rightarrow [0,g(1)]$ para que la igualdad anterior se cumpla, dado que el caso más favorable es $g(y)$ máximo y $g(1-x)$ mínimo, debe tenerse $y=1$ y $1-x=0$ simultáneamente. Por tanto:

$$M_S^{T-}(x,y)=1 \text{ ssi } y=1 \text{ y } x=1$$

Por otro lado, por definición de $C(x,y)$ se tiene que $C(x,y)=0$ para $x+y \leq 1$. Por tanto:

$$M_S^T(x,y)=0 \text{ ssi } y+x \leq 1$$

Por último señalemos que $M(0,y)=0$, $M(1,y)=y$ y $M(x,0)=0$. En cuanto a $M(x,1)$ es necesario distinguir entre el tipo de t-norma. Si la t-norma es arquimedea positiva se tiene $M(x,1)=g^{f-1}(g(1)-g(x))$ y por tanto al ser $g(1)=\infty$ se verifica $M(x,1)=1$ a menos que x valga 0 , en cuyo caso $M(0,1)=0$. Puede comprobarse que se obtiene el mismo resultado para la pseudo-conjunción obtenida a partir de la t-norma *min*. Si por el contrario la t-norma es negativa el resultado es $M(x,1)=g^{f-1}(g(1)-g(1-x))$ que en general depende de g .

- Resumen

La siguiente tabla resume los resultados de este apartado. En la primera columna aparece el tipo de implicación analizada, que hemos denominado de forma genérica $H(x,y)$. La segunda y la tercera columna resumen las condiciones que x e y deben cumplir para que $H(x,y)$ valga 1 o 0 . Las últimas cuatro columnas son el valor de la implicación cuando sus argumentos son los que se especifican en la primera línea de dichas columnas.

$H(x,y)$	$H(x,y)=1$	$H(x,y)=0$	$H(1,y)$	$H(0,y)$	$H(x,1)$	$H(x,0)$
$I_R^{T+}(x,y)$	$x \leq y$	$y=0$ y $x \neq 0$	y	1	1	0 si $x \neq 0$ 1 si $x=0$
$I_R^T(x,y)$	$x \leq y$	$f(y)-f(x) \geq f(0)$	y	1	1	$f^{-1}(f(0)-f(x))$
$I_S^{T+}(x,y)$	$x=0$ o bien $y=1$	$x=1$ o bien $y=0$	y	1	1	$1-x$
$I_S^T(x,y)$	$g(1-x)+g(y) \geq g(1)$	$x=1$ o bien $y=0$	y	1	1	$1-x$
$M_R^{T+}(x,y)$	$x=y=1$	$x=0$ o bien $y=0$	y	0	x	0
$M_R^T(x,y)$	$x=y=1$	$f(x)+f(y) \leq f(0)$	y	0	x	0
$M_S^{T+}(x,y)$	$y=1$ y $x \neq 0$	$y+x \leq 1$	y	0	1 si $x \neq 0$	0
$M_S^T(x,y)$	$y=1$ y $x=1$	$y+x \leq 1$	y	0	$g^{1-1}(g(1)-g(1-x))$	0

4.3.3.3 Otras funciones de implicación

En algunos casos utilizaremos también alguna otra implicación, como por ejemplo la n -implicación de Yager y su máxima FGMP dadas por:

$$I_Y(x,y)=y^x$$

$$M_Y(x,y) = \begin{cases} y^{\frac{1}{x}} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

Existen sin embargo otras muchas familias de implicaciones. A este respecto puede consultarse por ejemplo [Dubois, Prade, 84] donde se repasan la mayoría de los métodos de las lógicas multivaluadas para generar implicaciones. En [Trillas, 92] se propone una expresión general de función de implicación obtenida a partir de la DMP:

$$I(x,y) \leq x.y + \neg x.y + \neg x.\neg y$$

Esta familia es muy general y engloba entre otras a las r y s -implicaciones y a las t -normas y pseudo-conjunciones, pero algunas de las implicaciones resultantes no siempre presentan un carácter monótono respecto a su primer argumento por lo que en general no son

catalogables como n o p-implicaciones. En el mismo trabajo se estudia también en qué medida algunas de las familias de implicaciones que se pueden obtener generalizan mejor o peor la implicación material clásica.

Un familia que a menudo aparece en trabajos afines es la de las llamadas implicaciones cuánticas, dadas por $I(x,y)=S(\Gamma(x),T(x,y))$ (véase por ejemplo [Dubois, Prade, 91b] y [Trillas, 92]) donde S y T son una t-conorma y una t-norma respectivamente, y Γ una negación fuerte. Estas implicaciones, crecientes con y no son monótonas con x , por lo que no pueden catalogarse como n o p-implicaciones.

En [Alsina, Trillas, 92] y [Trillas et al., 92] se estudia la obtención de implicaciones a partir de medias cuasi-aritméticas, y por ejemplo puede obtenerse el par (M_{Tr}, I_{Tr}) dado por:

$$M_{Tr}(x, y) = \sqrt{x \cdot y}$$

$$I_{Tr}(x, y) = \begin{cases} 1 & \text{si } x = 0 \\ \min\left(1, \frac{y^2}{x}\right) & \text{si } x > 0 \end{cases}$$

Nótese que la función de agregación M_{Tr} verifica $M_{Tr}(x,y) \geq \min(x,y)$ y $M_{Tr}(1,y) \neq y$. Por otro lado aunque la implicación es creciente con y y decreciente con x , sin embargo no verifica $I(1,y)=y$. Pero el par (M_{Tr}, I_{Tr}) sí podría considerarse como modelo válido para una regla ya que utilizado conjuntamente se puede comprobar que sí verifica el MPE. Además utilizando la metodología desarrollada a lo largo de este capítulo se podría fácilmente comprobar que su comportamiento es parecido al del modelo (M_R^{prod}, I_R^{prod}) , generado por la t-norma $T(x,y)=xy$.

Es interesante notar que la agregación con una pseudo-conjunción o una pseudo-uni6n de dos n-implicaciones genera una n-implicaci6n, mientras que de dos p-implicaciones genera una p-implicaci6n.

En [Dubois, Prade, 91b] y [Ruan, Kerre, 93] se recoge un amplio conjunto de propiedades que podrían ser exigidas a una funci6n de implicaci6n para su definici6n desde un punto de vista axiomático, o cuyo estudio podría resultar de interés.

En esta tesis se propone en el apéndice un nuevo tipo implicaciones llamadas ϕ -implicaciones, familia general que engloba tanto a las n como a las p-implicaciones, que permiten por ejemplo justificar el uso de operadores de agregaci6n diferentes de los conjuntivos y disyuntivos. Sin embargo en lo que sigue nos limitaremos principalmente al

estudio de los \mathfrak{n} -modelos (M_R^T, I_R^T) , (M_S^T, I_S^T) , y a los \mathfrak{p} -modelos (M_R^T, M_R^T) , (M_S^T, M_S^T) , (M_R^T, M_S^T) y (M_S^T, M_R^T) .

4.3.4 Reformulación de la RCI mediante la compatibilidad borrosa

En [Baldwin, 79] se propone una reformulación de la RCI utilizando la compatibilidad borrosa o valores de verdad borrosos entre predicados definidos en el apartado 3.6.6.1 del capítulo 3:

$$Q'(v) = \tau_{Q'}(Q(v)) = \sup_x \{M(\tau_{PP'}(x), I(x, Q(v)))\}$$

En [Tong, Efstathiou, 82] se demuestra la equivalencia entre ambas expresiones de la RCI, cuando la función de agregación M es la \mathfrak{t} -norma \min . Sin embargo es sencillo extender esta demostración a las definiciones aquí propuestas para función de implicación y función generadora de modus ponens, de forma que ambas formulaciones de la RCI resultan ser totalmente equivalentes.

4.4 La modificación por similitud entre proposiciones

La MS es, al igual que la RCI, un procedimiento que permite aplicar el MPG. De este modo a partir de una regla $P(u) \rightarrow Q(v)$ y de una observación $P'(u)$, en general distinta de $P(u)$, la MS permite inferir la conclusión $Q'(v)$ dada por:

$$Q'(v) = I(S(P'(u), P(u)), Q(v))$$

donde $S(P'(u), P(u))$ es una medida de similitud entre la observación $P'(u)$ y la hipótesis de la regla $P(u)$, e $I(x, y)$ es una función de implicación según han sido definidas en el apartado anterior. Puede comprobarse cómo no existe en este caso un condicional explícito $R(u, v)$ que se combine con la observación $P'(u)$, sino que la función de implicación es utilizada directamente, en base al parecido entre P' y P calculado con S , para modificar el consecuente Q y así obtener la conclusión. Como mostraremos en este capítulo la RCI degenera en la MS cuando la hipótesis y la observación cumplen ciertos requisitos.

La MS se introduce en [Turksen, Zhong, 88] y [Turksen, Zhong, 90] donde se la denomina razonamiento aproximado por analogía. También denominada razonamiento plausible en otros trabajos, y modificación por compatibilidad o similitud en otros como [Cross, Sudkamp, 93] o [Cross, Sudkamp, 94], se la presenta generalmente como un mecanismo de inferencia alternativo a la RCI pero más eficiente, ya que la comprobación del cumplimiento de la hipótesis está desligada de la obtención de la conclusión, con lo que se controla mejor el disparo de las reglas analizando únicamente la semejanza entre hipótesis y observación. Aunque es cierto que esta regla de inferencia puede presentar ventajas computacionales frente a la RCI, como iremos viendo este mecanismo de inferencia basado en un sólo índice de similitud entre hipótesis y observación es mucho menos rico en matices que la RCI y sus ventajas no parecen suficientes para sustituir a la RCI en un sistema de reglas borrosas.

En [Magrez, Smets, 89] donde también se propone una versión de la MS como método alternativo a la RCI se critican algunos aspectos de ésta, en concreto el que algunos modelos (M, I) no verifican el MPE, lo que obviamente se debe a una mala elección del modelo de la regla, y el que la forma del consecuente dependa no solo del parecido entre la hipótesis P y la observación P' , sino también de la forma de P' .

En la MS el único requisito necesario para que se verifique el MPE es que la medida de similitud sea reflexiva en el sentido en que $S(P(u), P(u)) = 1$, siempre que la función de implicación se atenga a la definición propuesta y cumpla por tanto $I(1, y) = y$.

Aunque en la MS no puede hablarse de un par (M,I) como modelo de regla por no existir una FGMP, sí puede definirse un modelo como el par (S,I) . De hecho como se mostrará en sucesivos apartados ambos mecanismos de inferencia están bajo ciertas condiciones estrechamente relacionados, llegando incluso a ser equivalentes.

En los siguientes apartados analizaremos el comportamiento de algunos de los principales modelos de reglas obtenidos a partir de las implicaciones y funciones generadoras de modus ponens introducidas en la sección 4.3.2.2. En lo que respecta a las \mathfrak{n} -implicaciones analizaremos el comportamiento de los modelos con las máximas FGMP de cada implicación, esto es (M_R^T, I_R^T) , (M_S^T, I_S^T) , y veremos también la influencia de una FGMP menor. En lo que respecta a las \mathfrak{p} -implicaciones la variedad de combinaciones es más amplia, ya que cualquier FGMP de una \mathfrak{n} -implicación lo es de cualquier \mathfrak{p} -implicación. Esto puede dar lugar por ejemplo a los modelos (M_R^T, M_R^T) , (M_S^T, M_S^T) , pero también a otros como (M_R^T, M_S^T) , (M_R^T, M_S^T) , que también serán analizados.

4.5 La modificación del consecuente ante observación puntual

Una de las situaciones más sencillas que pueden presentarse a la hora de realizar la inferencia con la RCI se da cuando la observación es puntual. Diremos que una observación es puntual cuando se puede modelar mediante un conjunto con un único elemento, $P'(u)=\{u_0\}$. En este caso la aplicación de la RCI se simplifica notablemente quedando:

$$Q'(v) = M(1, I(P(u_0), Q(v))) = I(\alpha, Q(v))$$

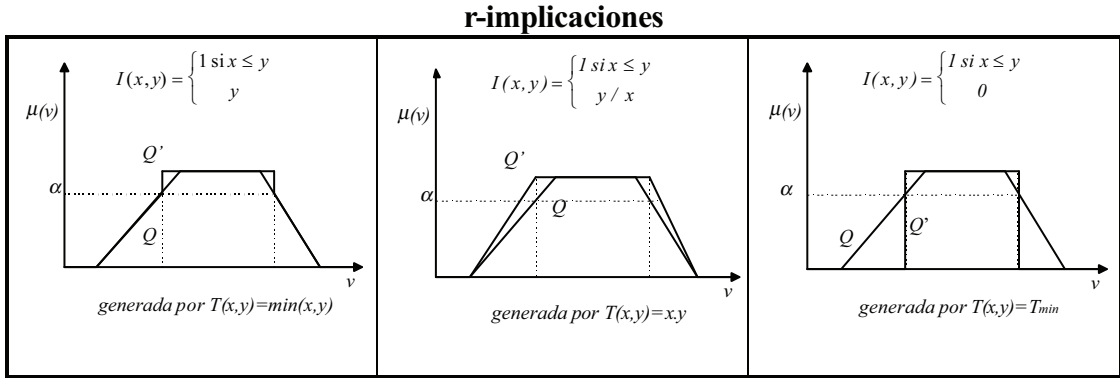
siendo $\alpha=P(u_0)$. Vemos en este caso sencillo que la RCI se reduce a la MS donde el índice de similitud empleado es directamente la compatibilidad de la hipótesis dada la observación, $\alpha=P(u_0)$. En esta situación resulta sencillo analizar la modificación que sobre el consecuente se produce para obtener la conclusión de la regla. Nótese que la expresión de la conclusión de la regla no depende en absoluto de la FGMP del modelo de la regla.

Aunque existen algunos trabajos que abordan también el estudio de la inferencia cuando la observación es puntual, éstos suelen limitarse al estudio de ejemplos concretos de funciones de implicación. En este apartado sin embargo hemos optado por ofrecer un análisis sistemático basado en las propiedades de las implicaciones, más que en ejemplos particulares, obteniéndose de este modo resultados de carácter general. Este estudio permite poner de manifiesto de forma sencilla que ciertas propiedades de las familias de funciones de implicación inciden directamente en su comportamiento en la inferencia, haciéndolas susceptibles de ser interpretadas de distinta forma desde un punto de vista semántico.

En las figuras utilizadas para ilustrar este apartado están representados el consecuente de la regla $Q(v)$ y la conclusión obtenida $Q'(v)$ cuando la observación es puntual, dada por $P'(u)=\{u_0\}$, siendo $\alpha=P(u_0)$. En el eje de abscisas está por tanto la variable v y en el eje de ordenadas el grado de pertenencia de v ya sea a Q , ya sea a Q' , y que hemos denominado genéricamente $\mu(v)$.

4.5.1 n-implicaciones y observación puntual

4.5.1.1 r-implicaciones:



Cuando la t-norma de la que procede la r-implicación es una t-norma continua la modificación de la conclusión se traduce siempre en un aumento del núcleo del consecuente, debido a que $I_R^{T^+}(\alpha, y) = 1$ para $\alpha < y$, obteniéndose $Q'(v) = 1$ para $Q(v) \geq \alpha$:

$$[Q'] = \{ v / Q(v) \geq \alpha \}$$

Cuando la t-norma es positiva el soporte de la conclusión se mantiene ya que $I_R^{T^+}(\alpha, y) = 0$ ssi $y = 0$ y $\alpha \neq 0$, y obviamente el caso $\alpha = 0$ queda descartado por corresponder a la total indeterminación:

$$]Q'[=]Q[$$

Cuando la t-norma es negativa el soporte varía. En efecto $I_R^{T^-}(\alpha, y) = 0$ para $y = 0$ o $x = 1$, es decir que el consecuente sólo tiene soporte cuando $\alpha = 1$, situación de total cumplimiento de la hipótesis, y para la que $Q' = Q$. En cualquier otro caso puede calcularse cuánto vale el mínimo grado de pertenencia del consecuente como $I(\alpha, 0) = f^{-1}(f(0) - f(\alpha))$ con lo que:

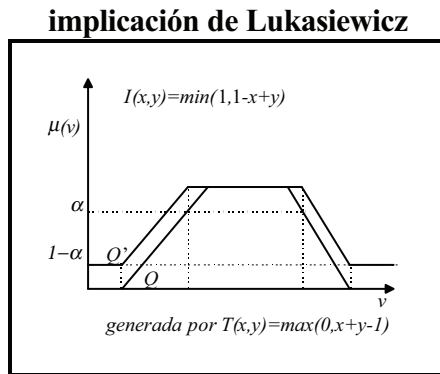
$$Q'(v) \geq f^{-1}(f(0) - f(\alpha))$$

que en general depende de f . Por ejemplo para la familia paramétrica de Yager (que no debe confundirse con la implicación de Yager I_Y antes mencionada) se tiene:

$$I_R^Y(\alpha, 0) = \min(1, 1 - (1 - (1 - \alpha)^p)^{\frac{1}{p}})$$

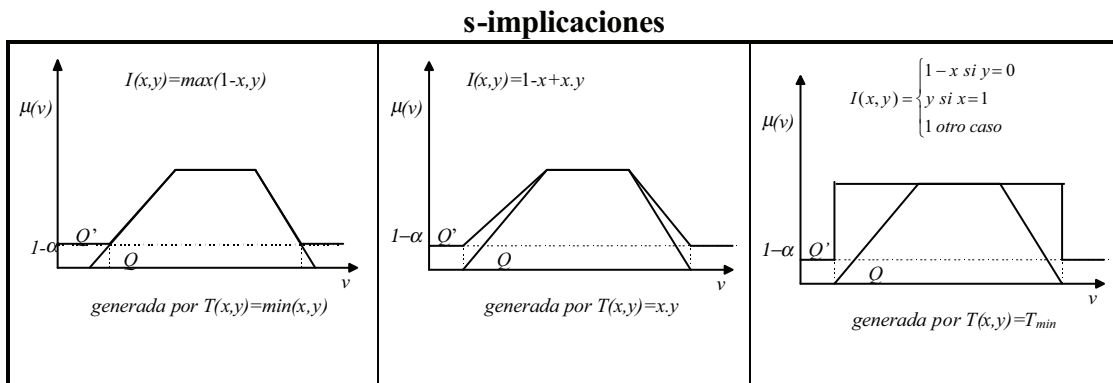
que resulta ser 1 para $p \rightarrow \infty$, 0 para $p \rightarrow 0$, y $1-\alpha$ para $p=1$, caso este último en el que la implicación se hace igual a la de Lukasiewicz. Por tanto el nivel de indeterminación que aparece en el consecuente es variable y dependiente de p .

En el caso de la implicación de Lukasiewicz la modificación del consecuente queda representada en la siguiente figura:



Esta implicación es obviamente un caso particular ya que es tanto una r-implicación como una s-implicación, generada a partir de una t-norma negativa, por lo que se produce un aumento del núcleo del consecuente y la aparición de un nivel de indeterminación dado por $1-\alpha$.

4.5.1.2 s-implicaciones:



Cuando la t-norma de la que procede la s-implicación es una t-norma continua la modificación siempre se traduce en la aparición de un nivel de indeterminación, que hace

que el soporte quede vacío. En el límite puede interpretarse como un agrandamiento infinito del soporte, lo que supone un cambio cualitativo. Esto es debido a que $I_R^T(\alpha, 0) = 0$ sólo para $\alpha = 1$ e $y = 0$. El nivel de indeterminación que aparece puede calcularse de forma sencilla ya que $I_R^T(\alpha, 0) = 1 - \alpha$:

$$Q'(v) \geq 1 - \alpha$$

Por otro lado el núcleo, cuando la t-norma es positiva, no varía ya que $I_R^{T+}(\alpha, y) = 1$ para $\alpha = 0$ o $y = 1$, es decir que a menos que la hipótesis se cumpla en grado 0, el núcleo de Q' coincide con el de Q :

$$[Q'] = [Q]$$

Sin embargo cuando la t-norma es negativa el núcleo se ve modificado ya que $I_R^{T-}(\alpha, y) = 1$ para $g(1 - \alpha) + g(y) \geq g(1)$, y por tanto depende de g , y se tiene:

$$[Q'] = \{ v / Q(v) \geq g^{1-1}(g(1) - g(1 - \alpha)) \}$$

Por ejemplo para la familia paramétrica de Yager la s-implicación implicación resultante es:

$$I_S^Y(x, y) = \min\left(1, \left((1 - x)^p + y^p\right)^{\frac{1}{p}}\right)$$

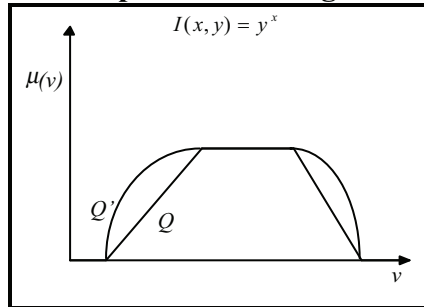
y el núcleo que se obtiene para el consecuente es variable con p . La solución a $I(\alpha, y) = 1$ es $(1 - \alpha)^p + y^p \geq 1$, es decir el núcleo viene dado por todos aquellos valores $v \in V$ tales que se verifica $Q(v) \geq (1 - (1 - \alpha)^p)^{1/p}$. Cuanto mayor es p mayor es el núcleo resultante, y para $p = 1$ se recupera la implicación de Lukasiewicz, ya analizada en el apartado anterior, cuyo núcleo como ya hemos indicado es:

$$[Q'] = \{ v / Q(v) \geq \alpha \}$$

4.5.1.3 La implicación de Yager:

Otro ejemplo interesante dentro de la familia de las n-implicaciones es el de la implicación I_Y de Yager (que recordemos no pertenece ni a la familia de implicaciones residuales ni a la de implicaciones fuertes):

implicación de Yager

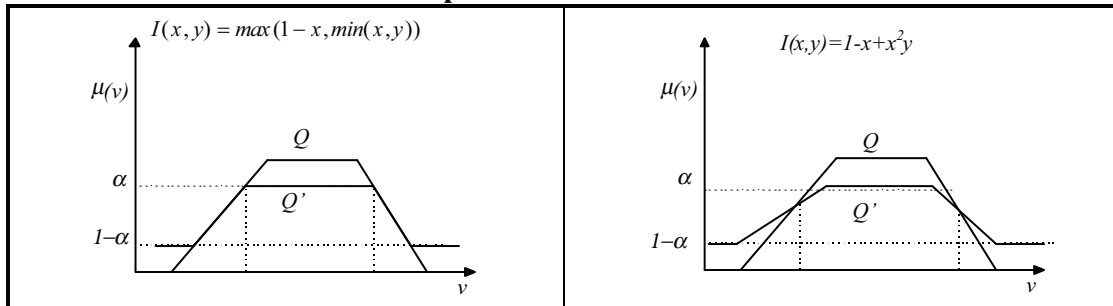


En este caso ni el núcleo ni el soporte del consecuente se ven modificados. Únicamente los grados de pertenencia intermedios aumentan su valor.

4.5.1.4 Las implicaciones cuánticas:

Haremos un último comentario sobre las implicaciones llamadas cuánticas, para descartarlas en los sucesivos análisis. Aunque estrictamente hablando no se corresponden con una n-implicación, cuando sus argumentos pertenecen a $\{0,1\}$ sí se recupera la implicación material clásica. Dos ejemplos de implicaciones de este tipo y de la modificación del consecuente que producen vienen representados en las siguientes figuras:

implicaciones cuánticas

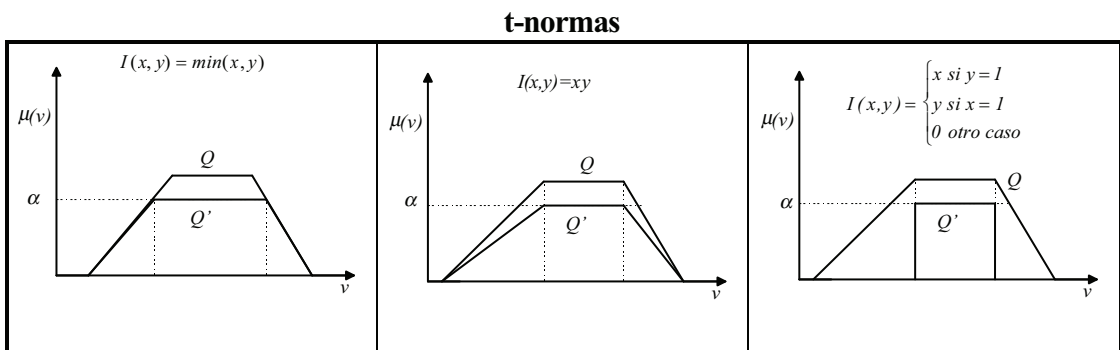


Una característica de estas implicaciones es su carácter no monótono con x , lo que hace que para $x \leq y$ son decrecientes con x mientras que para $x \geq y$ son crecientes con x . En definitiva los grados de pertenencia altos se ven reducidos mientras que valores de pertenencia bajos se ven aumentados, hasta que todos valen 1 cuando $\alpha=0$. No pueden por tanto catalogarse como n-implicaciones (estas implicaciones suelen descartarse también en otros trabajos debido a su comportamiento atípico, véase por ejemplo las llamadas implicaciones expansivas en [Turksen, Tian, 93], similares a las n-implicación

aquí propuestas). Semánticamente este tipo de implicaciones aunque generalizan el condicional material clásico presentan unos ciertos tintes de equivalencia, ya que a la vez que aparece un nivel de indeterminación en la conclusión la altura de la misma también se hace menor a 1 hasta que se llega a la indeterminación total que sigue siendo $Q'(v)=V$. Esto significa que para ciertos valores de la hipótesis se obtienen distribuciones de posibilidad no normalizadas, que como ya se ha señalado, en el caso de que fuesen interpretables como distribuciones de necesaria posibilidad reflejarían la inconsistencia del conocimiento empleado, en este caso de la propia implicación. Una verdadera equivalencia daría como conclusión, para $\alpha=0$, algún tipo de negación de Q .

4.5.2 p-implicaciones y observación puntual

4.5.2.1 t-normas:



Si la t-norma es continua la modificación siempre se traduce en un nivel de indeterminación que se manifiesta truncando la altura del consecuente. En efecto $M_R^T(\alpha,y)=1$ solo se cumple para $\alpha=y=1$:

$$Q'(v) \leq \alpha$$

Si además es positiva el soporte del consecuente no se ve modificado, ya que en efecto $M_R^{T+}(\alpha,y)=0$ solo se cumple para $\alpha=0$, que se corresponde con la total indeterminación, o bien $y=0$, que corresponde con el soporte de Q . Si por contra es negativa se tiene $M_R^{T-}(\alpha,y)=0$ para $f(x)+f(y) \geq f(0)$, condición que en general depende de f , con lo que:

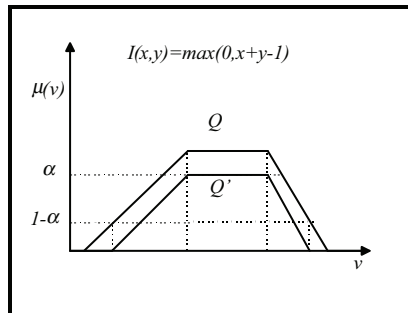
$$]Q[= \{ v / Q(v) \leq f^{t-1}(f(0)-f(\alpha)) \}$$

Por ejemplo para la familia paramétrica de Yager la t-norma viene dada por:

$$T(x, y) = 1 - \min\left(1, \left((1-x)^p + (1-y)^p\right)^{\frac{1}{p}}\right)$$

y el soporte de la conclusión viene dado por aquellos valores v tales que $Q(v) \leq 1 - (1 - (1 - \alpha)^p)^{1/p}$. En el caso en que $p=1$ se recupera la t-norma de Lukasiewicz para la que el soporte viene dado por $Q(v) \leq 1 - \alpha$, y que queda representada en la siguiente figura:

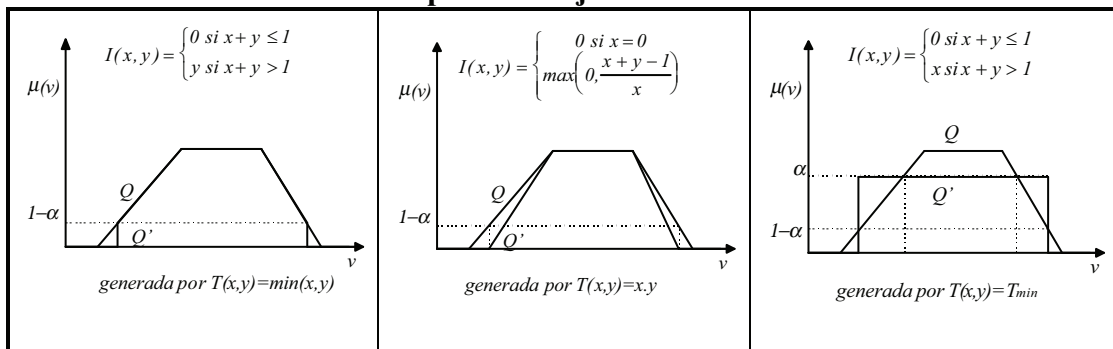
t-norma de Lukasiewicz



En otros casos, cuando p aumenta el soporte de la conclusión se aproxima al del consecuente, mientras que cuando disminuye se hace cada vez más pequeño, aproximándose al núcleo del consecuente.

4.5.2.2 pseudo-conjunciones:

pseudo-conjunciones



Cuando la implicación es una pseudo-conjunción generada a partir de una t-norma continua la modificación del consecuente se traduce siempre en una reducción de su soporte. En efecto $M_S^T(\alpha, y) = 0$ para $y \leq \alpha$ por lo que el soporte viene dado por:

$$]Q'[= \{v / Q(v) \leq 1-\alpha \}$$

Cuando además la t-norma es positiva se tiene $M_S^T(\alpha, y) = 1$ para $\alpha \neq 0$ e $y = 1$ por lo que el núcleo de la conclusión es igual al del consecuente. Si la t-norma es negativa se tiene $M_S^T(\alpha, y) = 1$ para $y = 1$ y $\alpha = 1$ por lo que se produce una reducción de la altura del consecuente en cuanto la hipótesis no se cumple con grado 1 . La nueva altura viene dada por $M_S^T(\alpha, 1) = g^{[-1]}(g(1) - g(1 - \alpha))$:

$$Q'(v) \leq g^{[-1]}(g(1) - g(1 - \alpha))$$

Por ejemplo para la familia paramétrica de Yager la pseudo-conjunción resultante viene dada por:

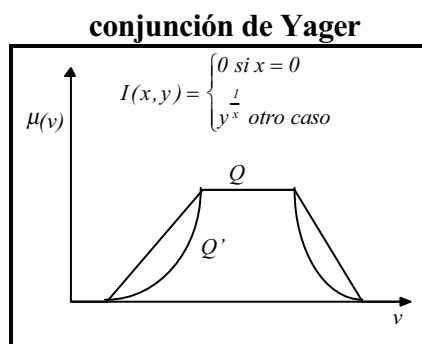
$$C(x, y) = \max(0, (y^p - (1-x)^p)^{\frac{1}{p}})$$

y la altura del consecuente por:

$$C(\alpha, 1) = (1 - (1 - \alpha)^p)^{1/p}$$

Puede comprobarse que si $p = 1$ se recupera la t-norma de Lukasiewicz, para la $C(\alpha, 1) = \alpha$. Por otro lado a medida que p aumenta, la altura del consecuente aumenta y se aproxima a 1, mientras que cuando p disminuye, la altura del consecuente se hace cada vez menor.

4.5.2.3 pseudo-conjunción de Yager:



De nuevo un caso peculiar lo constituye la conjunción M_Y obtenida como función generadora de modus ponens de la implicación de Yager. En este caso tanto el soporte

como el núcleo permanecen inalterados y sólo los grados intermedios de pertenencia al consecuente se ven reducidos en la conclusión.

4.5.3 Resumen

Los resultados quedan resumidos en la siguiente tabla. En la primera columna están los distintos modelos analizados. En la segunda la expresión obtenida para el núcleo de la conclusión o bien la altura de la misma cuando el núcleo está vacío. En la tercera columna figura el soporte de la conclusión, o bien el nivel de indeterminación de la misma cuando el soporte vale V . Por último la última columna resume cualitativamente la modificación que sufre el consecuente para dar lugar a la conclusión.

Modelo	$[Q']$	$]Q[$	$Q(v) \in]0,1[$
(M_R^{T+}, I_R^{T+})	$\{v / Q(v) > \alpha\}$	$]Q[$	igual o aumentan
(M_R^T, I_R^T)	$\{v / Q(v) > \alpha\}$	$Q'(v) \geq f^{-1}(f(0)-f(\alpha))$	aumentan
(M_S^T, I_S^T)	$\{v / Q(v) \geq g^{-1}(g(1)-g(1-\alpha))\}$	$Q'(v) \geq 1-\alpha$	aumentan
(M_S^{T+}, I_S^{T+})	$]Q[$	$Q'(v) \geq 1-\alpha$	igual o aumentan
(M_Y, I_Y)	$]Q[$	$]Q[$	aumentan
(M_R^{T+}, M_R^{T+})	$Q'(v) \leq \alpha$	$]Q[$	igual o disminuyen
(M_R^T, M_R^T)	$Q'(v) \leq \alpha$	$\{v / Q(v) \leq f^{-1}(f(0)-f(\alpha))\}$	disminuyen
(M_S^T, M_S^T)	$Q'(v) \leq g^{-1}(g(1)-g(1-\alpha))$	$\{v / Q(v) \leq 1-\alpha\}$	disminuyen
(M_S^{T+}, M_S^{T+})	$]Q[$	$\{v / Q(v) \leq 1-\alpha\}$	igual o disminuyen
(M_Y, I_Y)	$]Q[$	$]Q[$	disminuyen

Las implicaciones de Yager resultan ser casos particulares ya que tanto el núcleo como el soporte de la conclusión permanecen constantes siempre que la conclusión no sea la conclusión trivial.

4.6 Condiciones para la equivalencia entre la RCI y la MS

El estudio anterior sugiere la posibilidad de analizar si la inferencia mediante la RCI se reduce a la MS solamente en un caso tan sencillo como el de observación puntual, o si existen otras situaciones para las que la equivalencia entre ambos mecanismos se sigue manteniendo, y en su caso, cuál es el índice de similitud empleado en la formulación de la MS equivalente. En el apartado anterior veíamos que cuando la observación es puntual, $P'(u) = \{u_0\}$, la RCI se reduce a:

$$Q'(v) = M(1, I(P(u_0), Q(v))) = I(\alpha, Q(v))$$

siendo $\alpha = P(u_0)$. En este caso la RCI es equivalente a la MS con una medida de similitud dada por $S(P', P) = P(P') = P(u_0)$. En lo que sigue vamos a analizar qué condiciones deben cumplirse para que un modelo (M, I) para la RCI se reduzca a un modelo (S, I) para la MS. El interés de esta equivalencia reside más que en las condiciones que deben cumplirse, en el hecho de que cuando esas condiciones se dan resulta más fácil interpretar el comportamiento de un modelo (M, I) a través de su equivalente (S, I) .

Mientras que el análisis anterior permite conocer cómo se ve modificado el consecuente para la obtención de la conclusión, en este caso el índice de similitud entre P y P' nos da una información valiosa de cómo se hace la comparación entre la observación y la hipótesis.

En lo que sigue utilizaremos por comodidad la formulación de la RCI con valores de verdad borrosos. Supondremos, a menos que se especifique lo contrario, que los conjuntos son normales, convexos y con soporte acotado.

4.6.1 n-implicaciones borrosas

4.6.1.1 El modelo (M_R^T, I_R^T) :

La desigualdad del modus ponens permite plantear una primera condición:

$$\sup_x M_R^T(\tau(x), I_R^T(x, y)) \leq I_R^T(\alpha, y)$$

de donde se puede despejar $\tau(x)$. Para el caso en que la t-norma generadora sea arquímedea se tiene:

$$\begin{aligned}
M_R^T(\tau(x), I_R^T(x, y)) \leq I_R^T(\alpha, y) &\Rightarrow f^{[-1]}(f(\tau(x)) + f(I_R^T(x, y))) \leq I_R^T(\alpha, y) \Rightarrow \\
f(\tau(x)) + f(I_R^T(x, y)) &\geq f(I_R^T(\alpha, y)) \Rightarrow f(\tau(x)) \geq f(I_R^T(\alpha, y)) - f(I_R^T(x, y)) \Rightarrow \\
\tau(x) &\leq f^{[-1]}(f(I_R^T(\alpha, y)) - f(I_R^T(x, y))) = f^{[-1]}(f(y) - f(\alpha) - f(y) + f(x)) \Rightarrow \\
\tau(x) &\leq I_R^T(\alpha, x)
\end{aligned}$$

condición que no es más que la ya conocida T -transitividad de las implicaciones residuadas ([Trillas, Valverde, 85b]), que también verifica la implicación residuada de la t -norma min . De la desigualdad anterior se puede despejar el valor de α que la verifica:

$$\begin{aligned}
\tau(x) \leq I_R^T(\alpha, x) &\Rightarrow \tau(x) \leq f^{[-1]}(f(x) - f(\alpha)) \Rightarrow \\
f(\tau(x)) &\geq f(x) - f(\alpha) \Rightarrow f(\alpha) \geq f(x) - f(\tau(x)) \Rightarrow \\
\alpha &\leq f^{[-1]}(f(x) - f(\tau(x))) = I_R^T(\tau(x), x)
\end{aligned}$$

resultado final también válido para la t -norma min . Por otro lado como α tiene que ser independiente de x se tiene:

$$\alpha = \inf_x I_R^T(\tau(x), x) = \inf_u I_R^T(P'(u), P(u))$$

El valor de α así obtenido solo garantiza que la conclusión $\tau_{QQ}(y)$ está acotada superiormente por $I_R^T(\alpha, y)$, pero no garantiza que la desigualdad del modus ponens se convierta en igualdad. Para garantizar esto último se puede imponer la condición adicional $\tau(\alpha) = I$, en cuyo caso se alcanza el supremo para $x = \alpha$ y por tanto se verifica:

$$\tau_{QQ}(y) = I_R^T(\alpha, y)$$

En definitiva si $\tau(x) \leq I(\alpha, x)$ y $\tau(\alpha) = I$ entonces el modelo (M_R^T, I_R^T) es equivalente al modelo (S, I_R^T) con índice de compatibilidad dado por:

$$S(P, P') = \inf_{u/P'(u)=I} P(u) = N(P/[P'])$$

que se deduce de las dos condiciones $\tau(x) \leq I(\alpha, x)$ y $\tau(\alpha) = I$.

4.6.1.2 El modelo (M_S^T, I_S^T) :

Si la implicación I del modelo es una s -implicación generada a partir de una t -conorma S dual de T , y la FGMP M es su máxima FGMP, cuando T es arquímedea se cumple:

$$M_S^T(x, y) = \begin{cases} g^{[-1]}(g(y) - g(I - x)) & \text{si } x + y > I \\ 0 & \text{otro caso} \end{cases}$$

$$I_S^T(x, y) = g^{[-1]}(g(I - x) + g(y))$$

donde como se vio en el capítulo 3 la función $g(x): [0, +\infty[\rightarrow [0, 1]$ es una función creciente con $g(0)=1$. La primera condición para la equivalencia puede expresarse como:

$$\sup_x M_S^T(\tau(x), I_S^T(x, y)) \leq I(\alpha, y)$$

y en función del generador aditivo g se tiene:

$$\begin{aligned} g^{[-1]}(g(1-x)+g(y)-g(1-\tau(x))) &\leq g^{[-1]}(g(1-\alpha)+g(y)) \Rightarrow \\ g(1-x)+g(y)-g(1-\tau(x)) &\leq g(1-\alpha)+g(y) \Rightarrow \\ g(1-\tau(x)) \geq g(1-x)-g(1-\alpha) &\geq \Rightarrow 1-\tau(x) \geq g^{[-1]}(g(1-x)-g(1-\alpha)) \Rightarrow \\ \tau(x) &\leq 1-g^{[-1]}(g(1-x)-g(1-\alpha))=1-M_S^T(\alpha, 1-x)=I_R^T(\alpha, x) \end{aligned}$$

si además $\tau(\alpha)=1$ entonces se alcanza el supremo para $x=\alpha$ y se verifica la equivalencia, siendo el valor de α el mismo que el obtenido en el apartado anterior.

En definitiva, y al igual que ocurría con el modelo anterior, si $\tau(x) \leq I(\alpha, x)$ y $\tau(\alpha)=1$ entonces el modelo (M_S^T, I_S^T) es equivalente al modelo (S, I_S^T) con índice de similitud dado por:

$$S(P, P') = N(P/[P']).$$

Si la t-norma generadora es la t-norma *min* puede comprobarse que se verifica el mismo resultado. En efecto se tiene:

$$\min(\tau(x), \max(1-x, y)) \leq \max(1-\alpha, y)$$

y la desigualdad se cumple nuevamente para $\tau(x) \leq I_R^{\min}(\alpha, x)$ con:

$$I_R^{\min}(\alpha, x) = \begin{cases} 1 & \text{si } \alpha \leq x \\ x & \end{cases}$$

Si además $\tau(\alpha)=1$, se alcanza el supremo para $x=\alpha$, y por tanto estamos en el mismo caso que los anteriores.

4.6.1.3 El modelo de Yager (M_Y, I_Y) :

Recordemos que este modelo viene dado por:

$$I_Y(x, y) = y^x$$

$$M_Y(x, y) = \begin{cases} y^{\frac{1}{x}} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

La primera condición para que se verifique la equivalencia es la desigualdad del modus ponens:

$$M_Y(\tau(x), I_Y(x, y)) \leq I_Y(\alpha, y)$$

y despejando $\tau(x)$ de la ecuación anterior se tiene

$$I_Y(x, y)^{1/\tau(x)} \leq I(\alpha, y) \Rightarrow y^{x/\tau(x)} \leq y^\alpha \Rightarrow \tau(x) \leq x/\alpha$$

condición que garantiza la DMP. Podemos comprobar en la condición obtenida que $x/\alpha = I_R^P(\alpha, x)$, siendo I_R^P la implicación residuada de la t-norma producto. Si además $\tau(\alpha) = 1$ entonces para $x = \alpha$ se alcanza el supremo en la desigualdad del modus ponens.

Por tanto también para este caso, si $\tau(x) \leq x/\alpha$ y $\tau(\alpha) = 1$, entonces el modelo (M_Y, I_Y) es equivalente al modelo (S, I_Y) con índice de compatibilidad:

$$S(P, P') = N(P/[P'])$$

4.6.2 p-implicaciones borrosas

4.6.2.1 El modelo (M_R^T, M_R^T) :

Si la implicación y la FGMP son ambas iguales a una t-norma T , dada una compatibilidad $\tau(x)$ entre observación e hipótesis, la conclusión dada por la RCI con el modelo propuesto es:

$$\tau_{QQ'}(y) = \sup_x T(\tau(x), T(x, y))$$

La desigualdad del modus ponens permite establecer una primera condición, que supuesta T arquímedea, se escribe como:

$$f^{f^{-1}}(f(\tau(x)) + f(f^{f^{-1}}(f(x) + f(y))) \leq f^{f^{-1}}(f(\alpha) + f(y))$$

y despejando $\tau(x)$ se obtiene:

$$\tau(x) \leq f^{f^{-1}}(f(\alpha) - f(x)) = I(x, \alpha)$$

De esta misma desigualdad se puede despejar el valor de α que la verifica:

$$\begin{aligned} \tau(x) \leq f^{[-1]}(f(\alpha)-f(x)) &\Rightarrow f(\tau(x)) \geq f(\alpha)-f(x) \Rightarrow \\ f(\tau(x))+f(x) \geq f(\alpha) &\Rightarrow \alpha \leq f^{[-1]}(f(\tau(x))+f(x)) = T(\tau(x),x) \end{aligned}$$

y como debe ser independiente de x , la mínima cota superior de $\tau(x)$ se obtiene para:

$$\alpha = \sup_x T(\tau(x),x)$$

Como ya hemos visto, el valor de α así obtenido solo garantiza el cumplimiento de la desigualdad del modus ponens, pero no la igualdad. Esta puede garantizarse imponiendo la condición adicional $\tau(\alpha)=1$. Sin embargo para este modelo puede verse que otras condiciones como por ejemplo $\tau(1)=\alpha$ también son suficientes para garantizar la igualdad. Esto es debido al hecho de que puede obtenerse un resultado de carácter más general, teniendo en cuenta que la asociatividad de las t-normas permite escribir:

$$\tau_{QQ}(y) = T(\sup_x T(\tau(x),x), y)$$

para cualquier valor borroso $\tau(x)$. Esto significa que el modelo (M_R^T, M_R^T) es siempre equivalente al modelo (S, M_R^T) con índice de similitud dado por:

$$S(P, P') = \sup_u T(P'(u), P(u))$$

4.6.2.2 El modelo (M_S^T, M_S^T) :

Cuando la implicación es una pseudo-conjunción M_S^T igual a su FGMP, dado un valor de verdad borroso $\tau(x)$ de la hipótesis dada la observación, la conclusión obtenida con la RCI y el modelo propuesto es:

$$\tau_{QQ}(y) = \sup_x M_S^T(\tau(x), M_S^T(x,y))$$

Si M_S^T procede de una t-norma S arquímedea con generador g la desigualdad del modus ponens permite plantear la condición:

$$g^{[-1]}(g(g^{[-1]}(g(y)-g(1-x))) - g(1-\tau(x))) \leq g^{[-1]}(g(y)-g(1-\alpha))$$

y despejando $\tau(x)$ se obtiene:

$$\tau(x) \leq 1 - g^{[-1]}(g(1-\alpha) - g(1-x)) = 1 - M_S^T(x, 1-\alpha) = I_R^T(x, \alpha)$$

Nuevamente parecería necesario imponer una condición adicional para conseguir la igualdad como se ha visto para otros modelos, como por ejemplo $\tau(\alpha)=1$. Sin embargo, como también se vio en el capítulo 3 apartado 3.4.2.3, las pseudo-conjunciones verifican una propiedad similar a la asociatividad, dada por $C(a,C(b,c))=C(T(a,b),c)$ donde T es una t-norma continua, y C es la pseudo-conjunción generada por dicha t-norma. Esta propiedad permite escribir:

$$\begin{aligned}\tau_{QQ'}(y) &= \sup_x M_S^T(\tau(x), M_S^T(x,y)) = \sup_x M_S^T(M_R^T(\tau(x),x),y) \Rightarrow \\ \tau_{QQ'}(y) &= M_S^T(\sup_x M_R^T(\tau(x),x),y) = M_S^T(\sup_x T(\tau(x),x),y)\end{aligned}$$

lo que significa que el modelo (M_S^T, M_S^T) es siempre equivalente al modelo (S, M_S^T) con índice de similitud dado por:

$$S(P, P') = \sup_u T(P'(u), P(u))$$

Cuando la t-norma generadora es la t-norma *min* entonces se puede comprobar que si se verifica $\tau(x) \leq I_R^{min}(x, \alpha)$ entonces se verifica la desigualdad del modus ponens. En efecto dado que se tiene:

$$M_S^{min}(x, y) = \begin{cases} 0 & \text{si } x + y \leq 1 \\ y & \text{otro caso} \end{cases}$$

y que:

$$I_R^{min}(x, y) = \begin{cases} 1 & \text{si } x \leq y \\ y & \text{otro caso} \end{cases}$$

escogiendo la mínima $I_R^{min}(x_0, \alpha)$ posible superior a $\tau(x)$, en algún punto $x_0 \in [\alpha, 1]$ se tiene $\tau(x_0) = I_R^{min}(x_0, \alpha)$. Si $x_0 = \alpha$ estamos en el caso $x \leq \alpha$ y como $\tau(x_0) = I_R^{min}(x_0, \alpha) = 1$ se tiene $M_S^{min}(\tau(x_0), M_S^{min}(x_0, y)) = M_S^{min}(1, M_S^{min}(\alpha, y)) = M_S^{min}(\alpha, y)$. Si por el contrario $x_0 > \alpha$ entonces $\tau(x_0) = I_R^{min}(x_0, \alpha) = \alpha$ y por tanto $M_S^{min}(\tau(x), M_S^{min}(x, y)) = M_S^{min}(\alpha, M_S^{min}(x, y))$ y para $x = 1$ se alcanza el supremo. Por tanto también en este caso el modelo (M_S^{min}, M_S^{min}) es siempre equivalente al modelo (S, M_S^{min}) con índice de similitud dado por $S(P, P') = \sup_x T(\tau(x), x)$.

4.6.2.3 El modelo (M_R^T, M_S^T) :

Para este modelo la desigualdad del modus ponens permite plantear una primera condición dada por:

$$M_R^T(\tau(x), M_S^T(x, y)) \leq M_S^T(\alpha, y)$$

Supuesto que T es arquímedea, la desigualdad anterior se puede formular en función del generador aditivo:

$$f^{-1}(f(\tau(x)) + f(M_S^T(x,y))) \leq M_S^T(\alpha,y)$$

y despejando $\tau(x)$ se tiene:

$$f(\tau(x)) \geq f(M_S^T(\alpha,y)) - f(M_S^T(x,y))$$

$$\begin{aligned} \tau(x) &\leq f^{-1}(f(M_S^T(\alpha,y)) - f(M_S^T(x,y))) = I_R^T(M_S^T(x,y), M_S^T(\alpha,y)) \\ &= 1 - M_S^T(M_S^T(x,y), 1 - M_S^T(\alpha,y)) \end{aligned}$$

Por otro lado vimos en el capítulo 3 (apartado 3.4.2.3) que $C(C(a,b),c) = C(a,C(c,b))$, siendo C una pseudo-conjunción obtenida a partir de una t-norma continua. Tomando $C(x,y) = M_S^T(x,y)$ la expresión anterior puede reescribirse como:

$$\tau(x) \leq 1 - M_S^T(x, M_S^T(1 - M_S^T(\alpha,y), y))$$

y la expresión $M_S^T(1 - M_S^T(\alpha,y), y)$ puede simplificarse:

$$\begin{aligned} 1 - f^{-1}(f(1-y) - f(1 - M_S^T(\alpha,y))) &= 1 - f^{-1}(f(1-y) - f(1 - (1 - f^{-1}(f(1-y) - f(\alpha))))) = \\ 1 - f^{-1}(f(1-y) - f(f^{-1}(f(1-y) - f(\alpha)))) &= 1 - f^{-1}(f(1-y) - f(1-y) + f(\alpha)) = 1 - \alpha \end{aligned}$$

con lo que finalmente se tiene:

$$\tau(x) \leq 1 - M_S^T(x, 1 - \alpha) = I_R^T(x, \alpha)$$

Esta condición solo asegura el cumplimiento de la desigualdad pero no de la igualdad del modus ponens. Imponiendo la condición adicional de $\tau(\alpha) = 1$ se garantiza la igualdad ya que se tiene:

$$M_R^T(1, M_S^T(\alpha,y)) = M_S^T(\alpha,y)$$

Por tanto si $\tau(x) \leq I_R^T(x, \alpha)$ y $\tau(\alpha) = 1$ el modelo (M_R^T, M_S^T) es equivalente al modelo (S, M_S^T) con índice de similitud dado por:

$$S(P, P') = \sup_{x/\tau(x)=1} x = \sup_{u/P'(u)=1} P(u)$$

deducido de las condiciones $\tau(x) \leq I_R^T(x, \alpha)$ y $\tau(\alpha) = 1$. Puede comprobarse que se obtiene el mismo resultado cuando la t-norma generadora es la t-norma *min*.

4.6.2.4 El modelo (M_S^T, M_R^T) :

La desigualdad del modus ponens permite plantear una primera condición:

$$M_S^T(\tau(x), M_R^T(x, y)) \leq M_R^T(\alpha, y)$$

Supuesto que T es arquímedea, la desigualdad anterior se puede formular en función del generador aditivo:

$$1 - f^{t-1}(f(1 - M_R^T(x, y)) - f(\tau(x))) \leq M_R^T(\alpha, y)$$

y despejando $\tau(x)$ se tiene:

$$\begin{aligned} f^{t-1}(f(1 - M_R^T(x, y)) - f(\tau(x))) &\geq 1 - M_R^T(\alpha, y) \\ f(1 - M_R^T(x, y)) - f(\tau(x)) &\leq f(1 - M_R^T(\alpha, y)) \\ f(\tau(x)) &\geq f(1 - M_R^T(x, y)) - f(1 - M_R^T(\alpha, y)) \\ \tau(x) &\leq I_R^T(1 - M_R^T(\alpha, y), 1 - M_R^T(x, y)) \end{aligned}$$

Como la expresión anterior debe ser cierta para cualquier valor de y podemos escribir de forma equivalente:

$$\tau(x) \leq \inf_y I_R^T(1 - M_R^T(\alpha, y), 1 - M_R^T(x, y))$$

Para $x \leq \alpha$ se tiene $1 - M_R^T(\alpha, y) \leq 1 - M_R^T(x, y)$ por lo que la restricción anterior se convierte en $\tau(x) \leq 1$. Para $x > \alpha$ puede comprobarse que el ínfimo se obtiene para $y=1$, ya que la expresión $I_R^T(1 - M_R^T(\alpha, y), 1 - M_R^T(x, y))$ es decreciente con y . Esto conduce entonces a:

$$\inf_y I_R^T(1 - M_R^T(\alpha, y), 1 - M_R^T(x, y)) = I_R^T(1 - \alpha, 1 - x)$$

y por tanto se obtiene $\tau(x) \leq I_R^T(1 - \alpha, 1 - x)$ para $x > \alpha$ (recordemos que con los otros p-modelos la condición era $\tau(x) \leq I_R^T(x, \alpha)$). Esto permite también despejar el valor de α para el que se obtiene una cota superior de la conclusión en la forma $M_R^T(\alpha, y)$.

$$\begin{aligned} \tau(x) \leq I_R^T(1 - \alpha, 1 - x) &\Rightarrow f(\tau(x)) \geq f(1 - x) - f(1 - \alpha) \Rightarrow \\ f(1 - \alpha) \geq f(1 - x) - f(\tau(x)) &\Rightarrow 1 - \alpha \geq f^{t-1}(f(1 - x) - f(\tau(x))) \Rightarrow \\ \alpha \geq 1 - f^{t-1}(f(1 - x) - f(\tau(x))) &= M_S^T(\tau(x), x) \end{aligned}$$

y en el límite se tiene:

$$\alpha = S(P, P') = \sup_x M_S^T(\tau(x), x) = \sup_u M_S^T(P'(u), P(u))$$

Para garantizar además que la desigualdad se transforma en igualdad podemos nuevamente plantear la condición adicional $\tau(\alpha)=1$ de forma que se tiene:

$$M_S^T(1, M_R^T(\alpha, y)) = M_R^T(\alpha, y)$$

Por tanto si $\tau(x) \leq I_R^T(1-\alpha, 1-x)$ y $\tau(\alpha)=1$ el modelo (M_S^T, M_R^T) es equivalente al modelo (S, M_R^T) con índice de similitud dado por:

$$S(P, P') = \sup_{u/P'(u)=1} P(u)$$

y puede comprobarse que se obtiene el mismo resultado cuando la t-norma generadora es la t-norma *min*.

4.6.2.5 El modelo de Yager (M_Y, M_Y) :

En este caso la condición $Q'(y) = M_Y(\alpha, Q(y))$ y la DMP llevan a:

$$M_Y(\tau(x), M_Y(x, y)) \leq M_Y(\alpha, y)$$

reemplazando por sus expresiones y despejando $\tau(x)$ se tiene:

$$\begin{aligned} M_Y(x, y)^{1/\tau(x)} &\leq M_Y(\alpha, y) \Rightarrow \\ y^{1/(x \cdot \tau(x))} &\leq y^{1/\alpha} \Rightarrow \\ \tau(x) &\leq \alpha/x \end{aligned}$$

condición que garantiza la DMP. Por otro lado y como se demuestra en el apartado 4.6.5, la pseudo-conjunción de Yager es asociativa, esto es $C(x, C(y, z)) = C(C(x, y), z)$, lo que permite directamente obtener:

$$\begin{aligned} \sup_x M_Y(\tau(x), M_Y(x, y)) &= \sup_x M_Y(M_Y(\tau(x), x), y) = \\ &M_Y(\sup_x M_Y(\tau(x), x), y) = M_Y(\alpha, y) \end{aligned}$$

Por lo que el modelo (M_Y, M_Y) es equivalente al modelo (S, M_Y) con índice de similitud dado por:

$$S(P, P') = \sup_x M_Y(\tau(x), x) = \sup_u P(u)^{1/P'(u)}$$

También podríamos haber impuesto la condición $\tau(\alpha)=1$, con lo que el índice de similitud obtenido habría sido $\sup\{x / \tau(x)=1\}$ que efectivamente se deduce del índice anterior más general, haciendo $P'(u)=1$, valor para el que se alcanza el supremo.

4.6.3 Influencia de la FGMP

Para los modelos (M_R^T, I_R^T) y (M_S^T, I_S^T) , si la n-implicación (que llamaremos I_n) que en el modelo se combina con M_{In} , se combina con otra FGMP M , dado que esta será necesariamente menor, se tiene:

$$M(\tau(x), I_n(x, y)) \leq M_{In}(\tau(x), I_n(x, y)) \leq I_n(\alpha, y)$$

siempre que $\tau(x) \leq I_R(\alpha, x)$. Si además $\tau(x)$ sigue cumpliendo $\mu_{\alpha} \leq \tau(x)$ se sigue alcanzando el supremo en la RCI para $x = \alpha$, quedando:

$$\sup_x M(\tau(x), I_n(x, y)) = M(\tau(\alpha), I_n(\alpha, y)) = M(I, I_n(\alpha, y)) = I_n(\alpha, y)$$

por lo que se sigue cumpliendo la equivalencia RCI-MS bajo las mismas condiciones y con el mismo índice de similitud, siendo nula la influencia de la FGMP.

Supongamos ahora que analizamos el comportamiento de los modelos (M_R^T, M_S^T) y (M_R^T, M_S^T) , cuya p-implicación llamaremos I_p y cuya FGMP llamaremos M_{Ip} . Cualquier FGMP M válida para una n-implicación lo es también para una p-implicación. Sin embargo a menos que se verifique que $M(x, y) \leq M_{Ip}(x, y)$ no se puede garantizar el cumplimiento de la desigualdad del modus ponens dada por:

$$M(\tau(x), I_p(x, y)) \leq I_p(\alpha, y)$$

lo que significa que el valor de verdad borroso $\tau(x)$ estará limitado por otras cotas superiores distintas de las ya obtenidas. Si además $\mu_{\alpha} \leq \tau(x)$ el supremo (en el caso en que se verifique la DMP) se seguirá alcanzando para $x = \alpha$ y por tanto se tendrá:

$$\sup_x M(\tau(x), I_p(x, y)) = M(\tau(\alpha), I_p(\alpha, y)) = M(I, I_p(\alpha, y)) = I_p(\alpha, y)$$

por lo que se puede seguir cumpliendo la equivalencia RCI-MS aunque las condiciones pueden verse modificadas.

Por último, para los modelos (M_R^T, M_R^T) y (M_S^T, M_S^T) , si se escoge una FGMP M menor y se sigue verificando $\tau(x) \leq I_R^T(x, \alpha)$ se tiene:

$$M(\tau(x), I_p(x, y)) \leq I_p(\tau(x), I_p(x, y)) \leq I_p(\alpha, y)$$

Sin embargo es necesaria alguna condición adicional para garantizar la igualdad, como por ejemplo $\mu_{(\alpha)} \leq \tau(x)$. En el caso en que la FGMP sea mayor se verá modificado el límite superior de $\tau(x)$, que tenderá a ser menor.

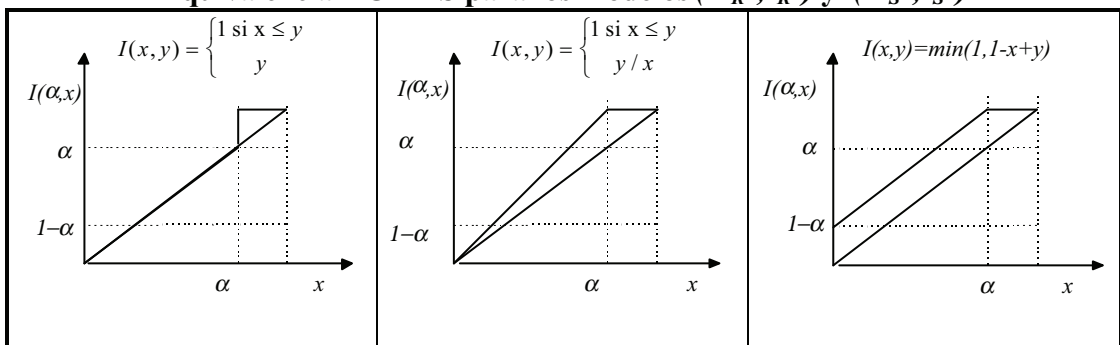
Por tanto aunque la FGMP no influye en el índice de compatibilidad sí puede influir en las condiciones exigidas a P y P' para que la equivalencia RCI-MS se cumpla.

4.6.4 Resumen

4.6.4.1 n-modelos

Las siguientes figuras representan gráficamente la cota superior de $\tau(x)$ para que pueda darse la equivalencia RCI-MS para los modelos analizados basados en algunas de las n-implicaciones más comunes:

Equivalencia RCI-MS para los modelos (M_R^T, I_R^T) y (M_S^T, I_S^T)



En primer lugar es interesante señalar que para los modelos analizados el índice de similitud que resulta de la equivalencia viene dado, para todos ellos, por:

$$S(P, P') = \inf_{u/P'(u)=1} P(u) = N(P/[P'])$$

lo que significa que la modificación del consecuente se basa únicamente en el grado de inclusión de $[P']$ en P .

Para los modelos (M_R^T, I_R^T) y (M_S^T, I_S^T) una condición necesaria para que pueda darse la equivalencia RCI-MS viene dada por $\tau_{PP'}(0) \leq I_R^T(\alpha, 0)$ y es especialmente interesante ya que impone restricciones de distinto tipo según la t-norma generadora sea positiva o negativa.

En efecto cuando la t-norma generadora es positiva se tiene $I_R^{T^+}(\alpha, 0) = 0$ y por tanto $\tau_{PP}(0) = \sup_{0/P(u)=0} P'(u) = 0$ lo que significa que debe verificarse $]P'[\subseteq]P[$. Sin embargo cuando la t-norma generadora es negativa $\tau_{PP}(0) > 0$ y por tanto deja de ser necesaria la condición anterior. La inclusión del núcleo de P' en $]P[$ sí es necesaria ya que en otro caso se tendría $\tau_{PP}(0) = 1$ y la conclusión obtenida sería la conclusión trivial.

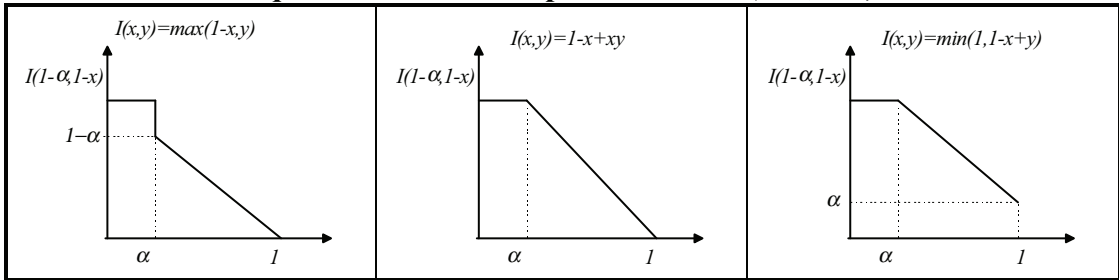
Esta diferencia según el carácter de la t-norma era previsible. En efecto una t-implicación basada en una t-norma positiva no puede introducir un nivel de indeterminación en la conclusión cuando se utiliza en un modelo (S, I) , como se vio en el apartado anterior, cuando se estudiaba la modificación del consecuente ante observación puntual. Sin embargo la RCI sí introduce un nivel de indeterminación en la conclusión cuando éste existe en la comprobación de la hipótesis, esto es cuando $\sup_{u/P(u)=0} P'(u) \neq 0$ (como se analizará detenidamente cuando se estudie la condición de activación de los modelos) por lo que en estos casos no puede darse la equivalencia. En cuanto a las implicaciones basadas en t-normas negativas, sí pueden introducir un nivel de indeterminación, como ocurre por ejemplo con la t-norma de Lukasiewicz, por lo que sí puede darse la equivalencia aun cuando exista cierta incertidumbre en la comprobación de la hipótesis.

Por otro lado las condiciones conjuntas $\tau(\alpha) = 1$ y $\tau(x) \leq I_R^T(\alpha, x)$ imponen una restricción a las pendientes relativas de P y P' , o dicho de otro modo a su contenido de vaguedad. En efecto en el caso de la t-norma *min* observando las condiciones de equivalencia se ve que la pendiente de $\tau_{PP}(x)$ en un entorno de α debe ser infinita lo que se traduce en que el contenido de vaguedad de la observación debe ser nulo, como por ejemplo el caso de una observación puntual. Para otras implicaciones como por la de Lukasiewicz la condición resultante es mucho menos restrictiva. Por ejemplo si los valores absolutos de las pendientes de P' son mayores que los de las pendientes de P , es decir si P' es menos vago que P , entonces siempre es posible encontrar un valor de α no nulo tal que $\tau(x)$ cumpla las condiciones de equivalencia sin ser idénticamente igual a 1.

4.6.4.2 p-modelos

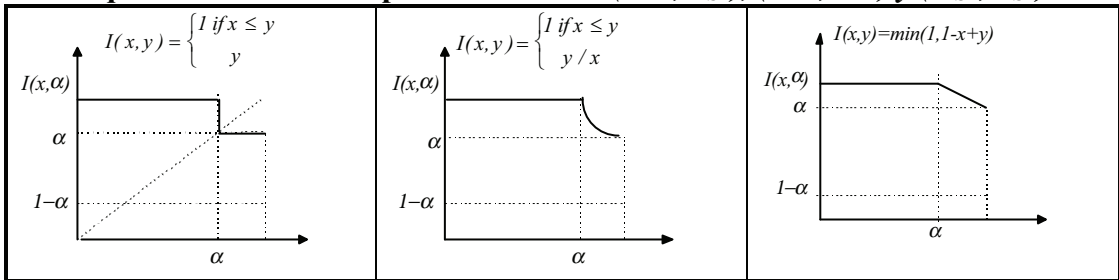
Las siguientes figuras representan gráficamente la cota superior de $\tau(x)$ para que pueda darse la equivalencia RCI-MS para el modelo (M_S^T, M_R^T) :

Equivalencia RCI-MS para el modelo (M_S^T, M_R^T)



Por el contrario para el modelo (M_R^T, M_S^T) la cota superior aparece reflejada en las siguientes figuras (nótese que también es la cota superior para el modelo (M_Y, M_Y) , y los modelos (M_R^T, M_R^T) y (M_S^T, M_S^T) , modelos estos últimos que se comentarán más adelante):

Equivalencia RCI-MS para los modelos (M_R^T, M_S^T) , (M_R^T, M_R^T) y (M_S^T, M_S^T)



En primer lugar es interesante señalar que para los dos p-modelos (M_S^T, M_R^T) y (M_R^T, M_S^T) el índice de similitud que resulta de la equivalencia viene dado por:

$$S(P, P') = \sup_{u/P'(u)=1} P(u) = \Pi(P/[P'])$$

lo que significa que la modificación del consecuente se basa únicamente en el grado de intersección de $[P']$ y P , es decir en la versión optimista de la medida $N(P/[P'])$.

Las figuras anteriores junto con la restricción $\tau_{PP'}(\alpha) = 1$ imponen límites al contenido de vaguedad relativo entre P y P' para grados de pertenencia a P altos. Por ejemplo para los modelos (M_S^T, M_R^T) la restricción es mucho más severa que para los modelos (M_R^T, M_S^T) , y lo es más todavía para los modelos (M_S^{T+}, M_R^{T+}) ya que para estos debe verificarse:

$$\tau_{PP'}(1) = \sup_{u/P(u)=1} P'(u) = 0$$

que se traduce en que la intersección entre $[P]$ y P' tiene que ser nula. Lógicamente la restricción obtenida para (M_S^T, M_R^T) y para (M_S^T, M_R^T) es la misma cuando la tnorma

generadora es la t-norma de Lukasiewicz, ya que en este caso ambos modelos coinciden. Los modelos (M_R^T, M_S^T) no requieren que $[P]$ y P' tengan una intersección nula para que se pueda verificar la equivalencia.

Para los p-modelos (M_R^T, M_R^T) y (M_S^T, M_S^T) la cota superior obtenida para $\tau(x)$ es la misma que la del modelo (M_R^T, M_S^T) aunque no se requiere la condición adicional $\tau(\alpha)=1$, por lo que la equivalencia se verifica siempre, y esto hace que el índice de similitud venga dado por:

$$\Pi_R^T(P/P') = \sup_x M_R^T(\tau(x), x) = \sup_u M_R^T(P'(u), P(u))$$

que no es más que la típica medida de posibilidad de P dado P' basada en la t-norma T . Esta medida es mayor que la medida utilizada por los otros modelos, esto es $S(P, P') = \Pi(P/[P'])$

Como ya hemos visto, para estos modelos la equivalencia se verifica siempre, y en cierto modo podemos decir que únicamente deja de verificarse cuando $\alpha=0$ que se corresponde con la obtención de la conclusión trivial. Si la t-norma generadora es positiva entonces se tiene:

$$\Pi_R^{T^+}(P/P') > 0 \Leftrightarrow]P'[\cap]P[\neq 0$$

mientras que si es negativa se tiene:

$$\Pi_R^{T^-}(P/P') > 0 \Leftrightarrow f(P'(u)) + f(P(u)) < f(0)$$

que es una condición más restrictiva que la obtenida para t-normas positivas, ya que no es suficiente tener $]P'[\cap]P[\neq 0$. Nótese que sí es condición suficiente que se tenga $[P']\cap]P[\neq 0$, o bien $]P'[\cap[P]\neq 0$ para cualquier generador f .

En la siguiente tabla están resumidas las condiciones para las cuales se ha mostrado la equivalencia entre la RCI y la MS para los distintos modelos analizados. En la primera columna figura el modelo, en la segunda las citadas condiciones (y en algunos casos una condición necesaria suficientemente significativa) y en la tercera el índice de similitud resultante para la MS equivalente.

Modelo	Cond. eq. RCI-MS	$S(P',P)$
(M_R^{T+}, I_R^{T+})	$\mu_{f(\alpha)} \leq \tau(x) \leq I_R^{T+}(\alpha, x)$ C.N. $]P'[\subset]P[$	$\alpha = \inf_{u/P'(u)=1} P(u)$
(M_R^T, I_R^T)	$\mu_{f(\alpha)} \leq \tau(x) \leq I_R^T(\alpha, x)$	$\alpha = \inf_{u/P'(u)=1} P(u)$
(M_S^T, I_S^T)	$\mu_{f(\alpha)} \leq \tau(x) \leq I_R^T(\alpha, x)$	$\alpha = \inf_{u/P'(u)=1} P(u)$
(M_S^{T+}, I_S^{T+})	$\mu_{f(\alpha)} \leq \tau(x) \leq I_R^{T+}(\alpha, x)$ C.N. $]P'[\subset]P[$	$\alpha = \inf_{u/P'(u)=1} P(u)$
(M_Y, I_Y)	$\mu_{f(\alpha)} \leq \tau(x) \leq I_R^{prod}(\alpha, x)$ C.N. $]P'[\subset]P[$	$\alpha = \inf_{u/P'(u)=1} P(u)$
(M_R^{T+}, M_S^{T+})	$\mu_{f(\alpha)} \leq \tau(x) \leq I_R^{T+}(x, \alpha)$	$\alpha = \sup_{u/P'(u)=1} P(u)$
(M_R^T, M_S^T)	$\mu_{f(\alpha)} \leq \tau(x) \leq I_R^T(x, \alpha)$	$\alpha = \sup_{u/P'(u)=1} P(u)$
(M_S^{T+}, M_R^{T+})	$\mu_{f(\alpha)} \leq \tau(x) \leq I_R^{T+}(1-\alpha, 1-x)$ C.N. $]P'[\cap]P[=0$	$\alpha = \sup_{u/P'(u)=1} P(u)$
(M_S^T, M_R^T)	$\mu_{f(\alpha)} \leq \tau(x) \leq I_R^T(1-\alpha, 1-x)$	$\alpha = \sup_{u/P'(u)=1} P(u)$
(M_R^{T+}, M_R^{T+})	$\tau(x) \leq I_R^{T+}(x, \alpha) \Leftrightarrow]P'[\cap]P[\neq 0$	$\alpha = \sup_u M_R^{T+}(P'(u), P(u))$
(M_R^T, M_R^T)	$\tau(x) \leq I_R^T(x, \alpha) \Leftrightarrow f(P'(u)) + f(P(u)) < f(0)$	$\alpha = \sup_u M_R^T(P'(u), P(u))$
(M_S^T, M_S^T)	$\tau(x) \leq I_R^T(x, \alpha) \Leftrightarrow f(P'(u)) + f(P(u)) < f(0)$	$\alpha = \sup_u M_R^T(P'(u), P(u))$
(M_S^{T+}, M_S^{T+})	$\tau(x) \leq I_R^{T+}(x, \alpha) \Leftrightarrow]P'[\cap]P[\neq 0$	$\alpha = \sup_u M_R^{T+}(P'(u), P(u))$
(M_Y, M_Y)	$\tau(x) \leq I_r^{prod}(x, \alpha) \Leftrightarrow]P'[\cap]P[\neq 0$	$\alpha = \sup_u P(u)^{1/P'(u)}$

4.6.5 Nueva familia de modelos de regla

Nos queda por último señalar algo que consideramos puede resultar de interés de cara a la generación de nuevas implicaciones y modelos de regla. Como se ha podido comprobar la implicación de Yager aparece íntimamente relacionada con la t-norma producto. Parece como si una t-norma además de generar las implicaciones residuales y fuertes pudiese utilizarse para generar otro tipo de implicaciones, que en el caso del producto sería la implicación de Yager. Recordemos que el modelo basado en la n-implicación de Yager viene dado por:

$$I_Y(x, y) = y^x$$

$$M_Y(x, y) = \begin{cases} y^{\frac{1}{x}} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

y que el generador aditivo de la t-norma producto, con la que dicho modelo parece estar estrechamente relacionado, es $f(x) = -\ln(x)$ siendo $f^{[-1]}(x) = e^{-x}$ su pseudo-inversa. Por simple inspección se puede comprobar que existe una relación directa entre el modelo de Yager y el generador aditivo de la t-norma producto, ya que verifican:

$$M_Y(x, y) = f^{[-1]}(f(y)/x)$$

$$I_Y(x,y) = f^{[-1]}(x.f(y))$$

Generalizando el resultado anterior, planteamos la obtención de una pseudo-conjunción a partir del generador aditivo de una t-norma mediante la expresión:

$$H(x,y) = \begin{cases} f^{[-1]} \left(\frac{f(y)}{x} \right) & \text{para } x \neq 0 \\ 0 & \text{en otro caso} \end{cases}$$

que es un operador conjuntivo y asociativo, aunque no conmutativo.

En primer lugar es un operador conjuntivo ya que se tiene $H(1,1) = f^{[-1]}(f(1)/1) = f^{[-1]}(0) = 1$, $H(0,y) = 0$ por definición y $H(x,0) = f^{[-1]}(f(0)/x)$ y como $f(0)/x > f(0)$ se tiene $H(x,0) = f^{[-1]}(f(0)/x) = 0$. Por último $f(y)/x$ decrece cuando x crece. Como además $f^{[-1]}$ es decreciente, entonces $H(x,y)$ es creciente con x . Del mismo modo, si y crece $f(y)$ decrece y por tanto $x.f(y)$ decrece con lo que $H(x,y)$ es creciente con y . Por tanto queda demostrado que $H(x,y)$ es un operador conjuntivo.

Por otro lado una propiedad de importancia en el ámbito de la inferencia viene dada por $H(1,y) = y$, y en nuestro caso puede comprobarse que en efecto se verifica, ya que $H(1,y) = f^{[-1]}(1.f(y)) = y$.

Es un operador asociativo ya que se verifica:

$$H(x, H(y,z)) = f^{[-1]}(f(H(y,z)) / x) = f^{[-1]}((f(y)/z) / x) = f^{[-1]}((f(y)/x) / z) = H(H(x,y), z)$$

aunque es claramente no conmutativo.

Por tanto esta familia de conjunciones parece apropiada para a partir de ella residuar una implicación, obteniéndose un modelo de regla que verifica el MPE, siendo un caso particular el modelo basado en la n-implicación de Yager (M_Y, I_Y). Sería interesante comprobar si efectivamente se mantiene la estrecha relación entre estos nuevos modelos y la t-norma que los genera (a través del generador aditivo), y si t-normas no arquímedeas como la t-norma *min* también puede dar lugar a modelos de este tipo.

4.7 Núcleo y soporte de las conclusiones según el tipo de modelo

La conclusión obtenida a partir de los modelos (M_R^T, I_R^T) y (M_S^T, I_S^T) puede expresarse de forma general como:

$$Q'(v) = \sup_u M(P'(u), I(P(u), Q(v)))$$

por lo que el núcleo se determina obteniendo aquellos valores $v \in V$ para los cuales $Q'(v) = 1$:

$$[Q'] = \{ v \in V / \sup_u M(P'(u), I(P(u), Q(v))) = 1 \}$$

Por otro lado la conclusión dada por los modelos (M_R^T, M_R^T) y (M_S^T, M_S^T) viene dada por:

$$Q'(v) = \sup_u M(P'(u), M(P(u), Q(v)))$$

y el soporte se obtiene el conjunto de aquellos valores $v \in V$ tales $Q'(v) = 0$, esto es:

$$]Q'[= \{ v \in V / \sup_u M(P'(u), M(P(u), Q(v))) = 0 \}$$

Por último la conclusión dada por los modelos (M_R^T, M_S^T) y (M_S^T, M_R^T) viene dada por:

$$Q'(v) = \sup_u M_R^T(P'(u), M_S^T(P(u), Q(v)))$$

y

$$Q'(v) = \sup_u M_S^T(P'(u), M_R^T(P(u), Q(v)))$$

y el soporte se obtiene el conjunto de aquellos valores $v \in V$ tales $Q'(v) = 0$, y llamando M_1 y M_2 a las dos conjunciones involucradas, se tiene:

$$]Q'[= \{ v \in V / \sup_u M_1(P'(u), M_2(P(u), Q(v))) = 0 \}$$

4.7.1 Núcleos de los modelos basados en n-implicaciones

4.7.1.1 El modelo (M_R^T, I_R^T) :

El núcleo vendrá dado por:

$$[Q'] = \{ v \in V / \sup_u M_R^T(P'(u), I_R^T(P(u), Q(v))) = 1 \}$$

En este caso $M_R^T(x,y)=T(x,y)$ y por tanto $M_R^T(x,y)=1$ ssi $x=y=1$, por lo que debe darse simultáneamente $P'(u)=1$ e $I_R^T(P(u),Q(v))=1$. Como además $I_R^T(x,y)=1$ ssi $x \leq y$, en este caso el núcleo de la conclusión viene dado por:

$$[Q'] = \{ v \in V / Q(v) \geq \inf_{u/P'(u)=1} P(u) \} = \\ \{ v \in V / Q(v) \geq N(P/[P']) \}$$

El tomar el ínfimo se debe a que basta con que exista un valor de u tal que $P'(u)=1$ y $Q(v) \geq P(u)$ para que v pertenezca al núcleo. Nótese que el núcleo viene determinado por la medida $N(P/[P'])$ que era el índice de similitud que se obtuvo al analizar la equivalencia RCI-MS para este modelo.

4.7.1.2 El modelo (M_S^T, I_S^T) :

Si el modelo está generado por una t-norma positiva se tiene $M_S^{T+}(x,y)=1$ ssi $y=1$ y $x \neq 0$, por lo que debe tenerse $I(P(u),Q(v))=1$ y $P'(u) \neq 0$. Como además $I_S^{T+}(x,y)=1$ ssi $x=0$ o bien $y=1$, el núcleo de $Q'(v)$ resulta ser igual al de $Q(v)$, siempre que se tenga $\inf_{u/P'(u)>0} P'(u) > 0$ (o lo que es lo mismo $]P'[\subseteq]P[$ que se corresponde, como veremos en el siguiente apartado, con la activación del modelo):

$$[Q'] = [Q]$$

Si la t-norma es negativa, entonces $M_S^{T-}(x,y)=1$ para $x=y=1$, es decir debe darse $P'(u)=1$ y $I(P(u),Q(v))=1$. Como además $I_S^{T-}(x,y)=1$ ssi $g(1-x)+g(y) \geq g(1)$, condición que depende de g , debe tenerse $g(1-P(u))+g(Q(v)) \geq g(1)$, el núcleo de Q' es por tanto variable con g , y viene dado por:

$$[Q'] = \{ v / Q(v) \geq \inf_{u/P'(u)=1} g^{[1]}(g(1)-g(1-P(u))) \} = \\ \{ v / Q(v) \geq \inf_{u/P'(u)=1} 1-f^{[1]}(f(0)-f(P(u))) \} = \\ \{ v / Q(v) \geq 1-f^{[1]}(f(0)-f(N(P/[P']))) \}$$

Puede comprobarse que $Q' \neq V$ siempre que $[P'] \subseteq]P[$, ya que sólo para $x=0$ se cumple $g(1-x)+g(y) \geq g(1)$ independientemente de y .

4.7.2 Soportes de los modelos basados en p-implicaciones

4.7.2.1 El modelo (M_R^T, M_R^T) :

Si T es positiva se tiene $M_R^{T+}(x,y)=0$ si $x=0$ o bien $y=0$, por lo que, supuesto que existe algún valor de u tal que $P(u) \neq 0$ y $P'(u) \neq 0$ (de forma que la conclusión no sea $Q'(v)=0$) debe cumplirse $M_R^{T+}(P(u), Q(v))=0$. Esto significa que $Q'(v)=0$ para $Q(v)=0$, por lo que el soporte de la conclusión no se ve modificado, y se tiene:

$$]Q' [=]Q[$$

Si T es negativa entonces $M_R^{T-}(x,y)=0$ para $f(x)+f(y) \geq f(0)$, por lo que debe tenerse $f(P(u)+f(Q(v)) \geq f(0)$. Utilizando la asociatividad de las t-normas tenemos:

$$\sup_u T(T(P'(u), P(u)), Q(v)) = 0$$

por tanto se obtiene:

$$]Q' [= \{ v / f(Q(v)) + f(\Pi_R^T(P/P')) > f(0) \} = \\ \{ v / Q(v) \leq f^{t-1}(f(0) - f(\Pi_R^T(P/P'))) \}$$

condición que en general depende de f .

4.7.2.2 El modelo (M_S^T, M_S^T) :

Denotando M_S^T por C , y dado que $C(a, C(b, c)) = C(T(a, b), c)$ podemos escribir:

$$\sup_u C(P'(u), C(P(u), Q(v))) = \sup_u C(T(P'(u), P(u)), Q(v)) = \\ C(\sup_u T(P'(u), P(u)), Q(v)) = 0$$

y como $C(x,y)=0$ para $x+y \leq 1$ se tiene:

$$]Q' [= \{ v / g(Q(v)) - g(1 - \Pi_R^T(P/P')) > g(0) \} = \\ \{ v / Q(v) \leq 1 - \Pi_R^T(P/P') \}$$

4.7.2.3 El modelo (M_R^T, M_S^T) :

La condición para obtener el soporte de la conclusión viene dada por:

$$]Q'[= \{ v \in V / \sup_u M_R^T(P'(u), M_S^T(P(u), Q(v))) = 0 \}$$

Si la t-norma es positiva, entonces para que $M_R^{T+}(P'(u), M_S^{T+}(P(u), Q(v)))$ se anule deberá cumplirse $M_S^{T+}(P(u), Q(v))=0$ o bien $P'(u)=0$. Suponiendo que existe algún valor de u tal que $P(u) \neq 0$ y $P'(u) \neq 0$ la condición anterior se traduce en:

$$]Q'[= \{ v \in V / \sup_{u/P'(u)>0} (P(u)+Q(v)) \leq 1 \}$$

que es equivalente a:

$$]Q'[= \{ v \in V / Q(v) \leq 1 - \sup_{u/P'(u)>0} P(u) \} = \\ \{ v \in V / Q(v) \leq 1 - \Pi(P/]P'[) \}$$

En el caso en que la t-norma sea negativa $T(x,y)=0$ se traduce en $f(x)+f(y) \geq f(0)$, lo que aplicado al caso que nos ocupa:

$$f(P'(u)) + f(M_S^T(P(u), Q(v))) \geq f(0) \Rightarrow M_S^T(P(u), Q(v)) \leq f^{t-1}(f(0) - f(P'(u))) \Rightarrow \\ 1 - f^{t-1}(f(1 - Q(v)) - f(P(u))) \leq f^{t-1}(f(0) - f(P'(u))) \Rightarrow \\ f^{t-1}(f(1 - Q(v)) - f(P(u))) \geq 1 - f^{t-1}(f(0) - f(P'(u))) = M_S^T(P'(u), 1) \Rightarrow \\ f(1 - Q(v)) \leq f(M_S^T(P'(u), 1)) + f(P(u)) \Rightarrow \\ Q(v) \leq 1 - f^{t-1}(f(M_S^T(P'(u), 1)) + f(P(u))) = 1 - T(M_S^T(P'(u), 1), P(u))$$

con lo que finalmente queda:

$$]Q'[= \{ v / Q(v) \leq 1 - \sup_u M_R^T(M_S^T(P'(u), 1), P(u)) \} = \\ \{ v / Q(v) \leq 1 - \sup_u T(M_S^T(P'(u), 1), P(u)) \} =$$

expresión compleja que depende del generador f de T . En cualquier caso se puede observar que dado que $M_S^{T+}(x,y) \geq M_S^{T-}(x,y)$ y que $M_R^{T+}(x,y) \geq M_R^{T-}(x,y)$ se verifica:

$$1 - \sup_u M_R^{T-}(M_S^{T-}(P'(u), 1), P(u)) \geq 1 - \sup_u M_R^{T+}(M_S^{T+}(P'(u), 1), P(u))$$

y por tanto el soporte que se obtiene es mayor que cuando la t-norma generadora es positiva, lo que significa que estará más próximo a $]Q[$, y mayor cuanto mayor sea la t-norma en cuestión.

4.7.2.4 El modelo (M_S^T, M_R^T) :

La condición para obtener el soporte de la conclusión viene dada por:

$$]Q' [= \{ v \in V / \sup_u M_S^T(P'(u), M_R^T(P(u), Q(v))) = 0 \}$$

Para que $M_S^T(P'(u), M_R^T(P(u), Q(v)))$ se anule debe verificarse que $P'(u) + M_R^T(P(u), Q(v)) \leq 1$, independientemente del carácter positivo o negativo de la t-norma generadora. Podemos despejar $Q(v)$ de la expresión anterior con lo que se obtiene:

$$\begin{aligned} P'(u) + T(P(u), Q(v)) &\leq 1 \Rightarrow \\ f(Q(v)) &\geq f(1 - P'(u)) - f(P(u)) \Rightarrow \\ Q(v) &\leq f^{t-1}(f(1 - P'(u)) - f(P(u))) = 1 - M_S^T(P(u), P'(u)) \end{aligned}$$

con lo que finalmente se obtiene:

$$\begin{aligned}]Q' [= \{ v \in V / Q'(v) \leq 1 - \sup_u M_S^T(P(u), P'(u)) \} = \\ \{ v \in V / Q'(v) \leq 1 - \Pi_S^T(P'/P) \} \end{aligned}$$

donde hemos definido la medida de posibilidad Π_S^T como:

$$\Pi_S^T(P'/P) = \sup_u M_S^T(P'(u), P(u))$$

que es la medida de posibilidad generada por la pseudo-conjunción M_S^T y que en general no es simétrica.

4.7.3 Resumen

La siguiente tabla resume los resultados de este apartado. En la primera columna está el modelo analizado. En la segunda el núcleo o el soporte de la conclusión dependiendo del tipo de implicación del modelo. Para los basados en n-implicaciones figura el núcleo de la conclusión, mientras que para los basados en p-implicaciones el soporte de la misma.

Modelo	Núcleo / soporte
(M_R^{T+}, I_R^{T+})	$[Q'] = \{ v \in V / Q(v) \geq N(P/[P']) \}$
(M_R^T, I_R^T)	$[Q'] = \{ v \in V / Q(v) \geq N(P/[P']) \}$
(M_S^T, I_S^T)	$[Q'] = \{ v / Q(v) \geq 1 - f^{T+}(f(0) - f(N(P/[P']))) \}$
(M_S^{T+}, I_S^{T+})	$[Q] = [Q] \text{ o } \emptyset \text{ si }]P'[\cap]P[= 0$
(M_R^{T+}, M_S^{T+})	$]Q'[= \{ v \in V / Q(v) \leq 1 - \Pi(P/[P']) \}$
(M_R^T, M_S^T)	$]Q'[= \{ v / Q(v) \leq 1 - \sup_u M_R^T(M_S^T(P'(u), 1), P(u)) \}$
(M_S^T, M_R^T)	$]Q'[= \{ v \in V / Q'(v) \leq 1 - \Pi_S^T(P'/P) \}$
(M_S^{T+}, M_R^{T+})	$]Q'[= \{ v \in V / Q'(v) \leq 1 - \Pi_S^{T+}(P'/P) \}$
(M_R^{T+}, M_R^{T+})	$]Q'[=]Q[$
(M_R^T, M_R^T)	$]Q'[= \{ v / Q(v) \leq f^{T+}(f(0) - f(\Pi_R^T(P/P'))) \}$
(M_S^T, M_S^T)	$\{ v / Q(v) \leq 1 - \Pi_R^T(P/P') \}$
(M_S^{T+}, M_S^{T+})	$\{ v / Q(v) \leq 1 - \Pi_R^{T+}(P/P') \}$

Como se puede comprobar los resultados obtenidos son consistentes con los obtenidos para observación puntual que es un caso particular de lo que aquí se ha presentado.

4.8 Condiciones de activación de modelos de reglas bajo la RCI

Otro punto de vista diferente para el estudio del comportamiento de modelos de reglas consiste en determinar cuando un modelo se activa, entendiendo por activación el hecho de que se comience a inferir una conclusión mínimamente informativa, o lo que es lo mismo una conclusión no trivial. Definimos la **condición de activación** de un modelo (M, I) para una regla $P(u) \rightarrow Q(v)$ como la condición que la observación $P'(u)$ y la hipótesis de la regla $P(u)$ deben cumplir para obtener como resultado de la inferencia una conclusión no trivial.

4.8.1 La implicación material clásica y la implicación material clásica restringida o producto cartesiano

Para la implicación material clásica es sencillo comprobar que cuando se verifica $P' \subseteq P$ se puede deducir la conclusión $Q' = Q$, mientras que en caso contrario sólo se deduce la conclusión trivial $Q' = V$. La condición de activación es por tanto $P' \subseteq P$. Cuando se utiliza la implicación material, si los conjuntos se interpretan como distribuciones de posibilidad en $\{0, 1\}$ puede observarse que cuando se cumple la condición de activación se tiene la certeza de que $u \in P$, y por tanto la hipótesis “ u es P ” (o lo que es lo mismo “ u cumple P ” o “ u está en P ”) es verdadera. Esto quiere decir que es necesario que v tome su valor en Q . Cuando la condición de activación no se cumple, entonces $\exists u / u \in P' \text{ y } u \notin P$, se concluye $Q' = V$, y sólo se tiene la certeza de que v está en V , que es obviamente una afirmación cierta pero carente de información. En ambos casos Q' es un conjunto en el que necesariamente tiene que estar el valor de v , y que sólo excluye a aquellos valores de V que es imposible que la variable v pueda tomar. Según lo hasta aquí expuesto podemos decir que la conclusión Q' obtenida del condicional material es una distribución de necesaria posibilidad, tal y como se definió en el apartado 3.6.2. Por ello las conclusiones de reglas paralelas modeladas con el condicional material deben agregarse conjuntivamente (lo que se tratará en el capítulo 5). Además la regla de inferencia RCI tal y como fue enunciada en el capítulo 3 exige que todas las distribuciones involucradas tengan el mismo carácter, lo que significa que el condicional material puede interpretarse como una distribución de necesaria posibilidad. Esto, como veremos en el capítulo 5, define el tipo de agregación que debe usarse para reglas modeladas con este tipo de condicional.

Con el producto cartesiano, también llamado condicional material restringido, cuando la observación P' verifica $P' \cap P \neq 0$, el resultado de la inferencia es $Q' = Q$, mientras que en caso contrario se obtiene la conclusión trivial $Q' = 0$. La condición de activación es por tanto $P' \cap P \neq 0$. Utilizando esta implicación, y cuando los conjuntos se interpretan como

distribuciones de posibilidad en $\{0,1\}$, si además la condición de activación se cumple, se puede afirmar que es posible que u tome su valor en P (aunque podría tomarlo fuera) por lo que la proposición “ u es P ” (o “ u está en P ”) es posible pero no cierta. La conclusión que se obtiene en este caso, $Q'=Q$, es un conjunto en el que es posible que v tome su valor. Existe una cierta evidencia, al no existir contradicción entre observación e hipótesis, de que v pueda estar en Q . Cuando la condición de activación no se cumple la conclusión es el conjunto vacío. En este caso no hay ninguna evidencia a favor de ningún subconjunto Q' de V por lo que sólo se concluye la conclusión trivial $Q'=0$. Podemos decir que Q' sólo incluye a aquellos elementos de v para los que hay alguna evidencia en su favor. La conjunción puede considerarse como una implicación válida siempre que sea debidamente interpretada, ya que permite razonar en términos posibilísticos y sus conclusiones son lógicas, siempre que se sea consciente de que no son ciertas o necesarias, sino tan solo posibles. La conclusión que se obtiene de una regla modelada con la conjunción es una distribución de posible posibilidad según se definió en el capítulo 3. Así queda también determinado que la agregación de conclusiones procedentes de reglas paralelas modeladas con una conjunción deben agregarse disyuntivamente. Nuevamente la uniformidad de tipos de distribuciones en la aplicación de la RCI permite interpretar el condicional basado en el producto cartesiano como una distribución de posible posibilidad, lo que determina el tipo de agregación lógica compatible con este tipo de condicionales, como veremos en el capítulo 5.

Puede pensarse desde un punto de vista teórico que $P' \cap P \neq 0$ y $P' \not\subset P$ corresponde al desconocimiento total sobre la veracidad de la afirmación “ u es P ” y que por tanto carece totalmente de utilidad. Sin embargo en la práctica la posibilidad de concluir, a partir de un conjunto suficientemente exhaustivo de posibles afirmaciones que describen el estado de un sistema, aquellas que al menos no son contradictorias (o totalmente contradictorias) con la información real que del sistema se tiene, puede ser de mucha utilidad cuando no se dispone de mejor información. Este puede ser el caso de un sistema de diagnóstico para el que se disponga de un conjunto de diagnósticos suficientemente exhaustivo, capaz de describir las posibles anomalías del proceso diagnosticado. Cuando en el sistema diagnosticado se detecte una anomalía que no se pueda diagnosticar con certeza, queda la opción de determinar todos aquellos diagnósticos que no son contradictorios con el estado actual del proceso, y que por tanto son posibles causas de la citada anomalía.

4.8.2 La conclusión trivial de p y n-implicaciones

Cuando se infiere utilizando la RCI la conclusión trivial se obtiene cuando $P(u)$ y $P'(u)$ son suficientemente distintos, situación que se corresponde con el caso en que los soportes de

ambos conjuntos no intersecan. Dado que $M(0,y)=0$ el supremo de la RCI se obtiene para $P'(u)=1$ que se corresponde con $P(u)=0$, teniéndose:

$$Q'_{trivial}(v)=M(1,I(0,Q(v)))$$

Además $M(1,y)=y$ por lo que la conclusión trivial de un modelo (M,I) puede expresarse como $I(0,y)$, independientemente de M . No ocurre así por ejemplo con el par (M_T, I_T) que no es estrictamente un modelo según nuestra definición, y para el que la conclusión trivial debe expresarse como $M(1,I(0,y))$.

Por tanto para una n-implicación la condición de activación puede expresarse como:

$$\sup_u M(P'(u), I(P(u), Q(v))) < I(0, Q(y))=1$$

Como además I es no decreciente con su segundo argumento, el caso más desfavorable se da para $Q(y)=0$, y puede decirse que el modelo (M,I) se activa cuando:

$$\forall u \in U / M(P'(u), I(P(u), 0)) < 1$$

Análogamente para una p-implicación se tiene:

$$\sup_u M(P'(u), I(P(u), Q(v))) > I(0, Q(y))=0$$

y al ser I no decreciente con su segundo argumento la condición de activación puede expresarse como:

$$\exists u \in U / M(P'(u), I(P(u), 1)) > 0$$

4.8.3 Condición suficiente para la no activación de un modelo

Dado un modelo (M,I) basado en una n-implicación I , podemos establecer una condición suficiente para la no activación del mismo, es decir para que se cumpla:

$$\exists u \in U / M(P'(u), I(P(u), 0)) = 1$$

Dado que $I(0,y)=1$ y que $M(1,y)=y$, es suficiente para que el modelo no se active que se cumpla $P'(u)=1$ y $P(u)=0$ para un mismo valor de u , o lo que es lo mismo:

$$[P'] \cap \neg P \neq \emptyset \Leftrightarrow [P'] \not\subseteq P$$

Por el contrario si I es una p-implicación la no activación del modelo requiere:

$$\forall u \in U / M(P'(u), I(P(u), 1)) = 0$$

y dado que $I(0, 1) = 0$ y $M(0, y) = y$ es suficiente para que el modelo no se active que si $P'(u) \neq 0$ entonces $P(u) = 0$ o bien si $P'(u) = 0$ entonces $P(u) \neq 0$, o lo que es lo mismo:

$$]P'[\cap]P[= 0$$

4.8.4 n-implicaciones borrosas

Dadas las funciones:

$$I_{n-max}(x, y) = \begin{cases} 0 & \text{si } x = 1, y = 0 \\ 1 & \text{en otro caso} \end{cases}$$

$$I_{n-min}(x, y) = \begin{cases} 1 & \text{si } x = y = 0 \text{ o } x = 0 \\ 0 & \text{en otro caso} \end{cases}$$

y dada una n-implicación $I(x, y)$ cualquiera, se tiene:

$$I_{n-min}(x, y) \leq I(x, y) \leq I_{n-max}(x, y)$$

Aunque ninguna de estas dos funciones puede considerarse estrictamente una n-implicación al no cumplir $I(1, y) = y$, nos servirán ambas para probar que cualquier n-implicación basa su condición de activación en algún tipo de inclusión de P' en P .

En efecto utilizando como implicación la función I_{n-max} la condición de activación se traduce en:

$$\begin{aligned} & \forall u \in U / M(P'(u), I_{n-max}(P(u), 0)) < 1 \\ \Leftrightarrow & \forall u \in U / P(u) < 1 \text{ y } M(P'(u), 0) < 1 \end{aligned}$$

mientras que para la función I_{n-min} la condición de activación se traduce en:

$$\begin{aligned} & \forall u \in U / M(P'(u), I_{n-min}(P(u), 0)) < 1 \\ \Leftrightarrow & \forall u \in U / P(u) = 0 \text{ y } M(P'(u), 0) < 1 \end{aligned}$$

Vemos por tanto que la función I_{n-max} permite obtener la condición de activación menos restrictiva, ya que sólo exige $P(u) \neq 1$ mientras que I_{n-min} exige $P(u) = 0$. Escogeremos por tanto la función de agregación M de forma distinta en ambos casos, para así obtener condiciones de activación los más extremas posibles.

Escogiendo M para I_{n-max} tal que $M(x, 1) = 1$ si $x \neq 0$ se obtiene la condición de activación:

$$]P'[\subseteq]P[$$

mientras que escogiendo M para I_{n-min} tal que $M(x, 1) = 0$ si $x \neq 1$, se tiene la condición de activación:

$$[P'] \subseteq]P[$$

con lo que queda probado que cualquier n-implicación basa su condición de activación en algún tipo de inclusión de P' en P , siendo los dos casos extremos $]P'[\subseteq]P[$ y $[P'] \subseteq]P[$.

4.8.4.1 El modelo (M_R^T, I_R^T)

Si la t-norma T es arquímedea la condición de activación puede expresarse en función de su generador aditivo f :

$$\forall u \in U, f^{-1}(f(P'(u)) + f(0) - f(P(u))) < 1$$

Al ser $f^{-1}(x)$ una función estrictamente decreciente para $x \in [0, f(0)]$, con $f^{-1}(0) = 1$, el modelo se activa si:

$$f(P'(u)) + f(0) - f(P(u)) > 0$$

La expresión anterior es mínima para $P(u)$ mínimo y $P'(u)$ máximo. Además no se produce la activación del modelo si $P(u) = 0$ y $P'(u) = 1$ para un mismo valor de u . Dado que f es estrictamente creciente el miembro izquierdo de la desigualdad es positivo si $\forall u \in U, P(u) = 0$ entonces $P'(u) \neq 1$ o bien si $\forall u \in U, P'(u) = 1$ entonces $P(u) \neq 0$. La condición de activación es por tanto:

$$[P'] \subseteq]P[$$

Si T es la t-norma min, entonces la condición de activación queda:

$$\forall u \in U, \min(P'(u), I_R^{\min}(P(u), 0)) < 1$$

El supremo vuelve a obtenerse para $P(u)$ mínimo y $P'(u)$ máximo y puede comprobarse que $P'(u)=1$ y $P(u)=0$ no pueden darse para un mismo valor de u . Por otro lado si $P'(u) < 1$ y $P(u)=0$ entonces se tiene $\min(P'(u), 1)=P'(u) < 1$ que es cierto. Del mismo modo si $P'(u)=1$ y $P(u) \neq 0$ entonces $I_R^{\min}(P(u), 0)=0$ que es obviamente menor que 1 . Por tanto la condición de activación de este modelo es también:

$$[P'] \subseteq]P[$$

El nivel de indeterminación cuando T es arquímedea viene dado por:

$$Q'(v) \geq \sup_u f^{f-1}(f(P'(u)) + f(0) - f(P(u)))$$

Su resolución es en general compleja y depende tanto de f como de P y P' . El supremo se obtiene para $f(P'(u))+f(0)-f(P(u))$ mínimo, al ser f^{f-1} decreciente, lo que equivale a minimizar $f(P'(u))-f(P(u))$.

Podemos particularizar la expresión anterior a ciertos casos particulares notables. Para la t-norma producto T_P , el nivel de indeterminación de la conclusión se obtiene para un valor de u que minimice $-\ln(P'(u))-\ln(P(u))$, es decir:

$$\frac{P'(u)}{P(u)} \text{ máximo}$$

que es una forma de medir la distancia entre $P(u)$ y $P'(u)$. Para la t-norma de Lukasiewicz T_L el nivel de indeterminación se obtiene para un valor de u que minimice $(1-P'(u))-(1-P(u))$, es decir:

$$P'(u)-P(u) \text{ máximo}$$

Otro posible ejemplo es la t-norma de Hamacher, cuya expresión recordemos es $T(x,y)=(1-x)/x$, para la que se obtiene:

$$\frac{P'(u)-P(u)}{P(u)P'(u)} \text{ máximo}$$

Los tres casos anteriores son formas alternativas de, en cierto modo, medir la distancia máxima entre P' y P (aunque estrictamente hablando la medida no sea simétrica).

Volviendo a la expresión general obtenida, podemos particularizarla a los dos casos extremos dados por máximo $P'(u)$ y $P(u)=0$, o bien mínimo $P(u)$ y $P'(u)=1$, siendo ambos posibles candidatos a nivel de indeterminación. Por ejemplo si T es positiva

entonces $f(0)$ es infinito y a menos que $P(u)=0$ se tiene $f(P'(u))+f(0)-f(P(u))=\infty$, y por tanto $f^{f^{-1}}(\infty)=0$. En este caso sólo el primer caso de los dos mencionados es candidato a supremo, y viene dado por:

$$Q'(v) \geq \sup_u f^{f^{-1}}(f(P'(u)) + f(0) - f(P(u))) = \sup_{u/P(u)=0} f^{f^{-1}}(f(P'(u))) = \sup_{u/P(u)=0} P'(u)$$

La medida anterior puede expresarse como una medida de posibilidad entre $\neg]P[$ y P' como sigue:

$$\Pi(\neg]P[/P') = \sup_{u/P(u)=0} P'(u)$$

o también en función de la medida de necesidad:

$$N(]P[/P') = 1 - \Pi(\neg]P[/P') = 1 - \sup_{u/P(u)=0} P'(u) = \inf_{u/P(u)=0} (1 - P'(u))$$

con lo que el nivel de indeterminación queda:

$$Q'(v) \geq \Pi(\neg]P[/P') = 1 - N(]P[/P')$$

Para el caso de la t-norma min se puede fácilmente evaluar ambas expresiones y comprobar que se obtiene la misma aproximación, $\Pi(\neg]P[/P')$.

Si por el contrario T es negativa entonces ambos casos son candidatos a supremo, y para el segundo de ellos se obtiene:

$$Q'(v) \geq \sup_u f^{f^{-1}}(f(P'(u)) + f(0) - f(P(u))) = \sup_{u/P'(u)=1} f^{f^{-1}}(f(0) - f(P(u))) = f^{f^{-1}}(f(0) - f(\inf_{u/P'(u)=1} P(u))) = f^{f^{-1}}(f(0) - f(N(P/[P'])))$$

De esta forma tomando el máximo de ambos casos podemos obtener una cota inferior para el nivel de indeterminación. Dado que el nivel de indeterminación se obtiene cuando la expresión $f(P'(u))-f(P(u))$ es mínima, la simplificación propuesta consiste en suponer que este nivel de indeterminación se puede evaluar como la máxima diferencia entre la observación P' y la hipótesis P (despreciando en cierto modo la no linealidad de f) y que además ésta se da en los extremos mencionados, es decir para máximo $P'(u)$ y $P(u)=0$, o bien mínimo $P(u)$ y $P'(u)=1$. A menos que se particularicen tanto la función f como las funciones de pertenencia P y P' no parece que se pueda obtener una solución exacta suficientemente general.

Es importante tener en cuenta que no hay garantía de que la aproximación propuesta sea suficientemente parecida al valor real. Únicamente se basa en que f es decreciente, y que por tanto es de esperar que el mínimo de $f(P'(u))-f(P(u))$ se dé para un valor u tal que $P'(u)$ sea pequeño respecto a $P(u)$. Cuanto más lineal sean las funciones f y lineales a trozos las funciones $P'(u)$ y $P(u)$ más exacta es la aproximación (ya que en el caso lineal se reduce a un problema lineal con restricciones y el supremo tiene que darse en alguno de los vértices de la región definida por las restricciones). Sin embargo desde un punto de vista semántico no parece descabellado considerar que la incertidumbre de la conclusión pueda estimarse mediante la máxima diferencia entre $P'(u)$ y $P(u)$ medida como $P'(u)-P(u)$, que para distribuciones convexas sencillas (como por ejemplo triangulares o trapezoidales) se da en los extremos mencionados.

4.8.4.2 El modelo (M_S^T, I_S^T)

Si I_S^T es una implicación generada por una t-conorma arquímedea, la condición de activación puede expresarse en función de su generador aditivo:

$$\forall u \in U, g^{f^{-1}}(g(1-P(u))-g(1-P'(u))) < 1$$

Si T es una t-norma positiva, $g^{f^{-1}}(x)$ es una función no decreciente con x , y $g^{f^{-1}}(x) < 1$ para $x \neq \infty$. Por tanto el argumento de $g^{f^{-1}}$ no puede ser ∞ :

$$g(1-P(u))-g(1-P'(u)) \neq \infty$$

es decir que si $1-P(u)=1$, caso en el que $g(1-P(u))$ se hace infinito, entonces debe tenerse $g(1-P(u))=g(1-P'(u))$ y al ser g estrictamente creciente se traduce en que si $\forall u \in U, P(u)=0$ entonces $P'(u)=0$, o bien si $\forall u \in U, P'(u)=0$ entonces $P(u)=0$. La condición de activación es pues:

$$]P'[\subseteq]P[$$

Si T es negativa entonces $g^{f^{-1}}(x)$ es una función no decreciente con x , y $g^{f^{-1}}(x) < 1$ para $x < g(1)$, que en este caso es finito. La condición es entonces equivalente a:

$$\forall u \in U, g(1-P(u))-g(1-P'(u)) < g(1)$$

o bien en función del generador de la t-norma:

$$\forall u \in U, f(P(u))-f(P'(u)) < f(0)$$

Dado el carácter de f , el caso más desfavorable se da para P mínimo y P' máximo, por tanto si $P(u)=0$ entonces debe tenerse $f(P'(u))>0$ para que se cumpla la condición, es decir si $P(u)=0$ entonces $P'(u)\neq 1$, y del mismo modo si $P'(u)=1$ entonces $P(u)\neq 0$. La condición de activación es por tanto:

$$]P'] \subseteq]P[$$

recuperándose la condición obtenida para las r -implicaciones. Nótese que la implicación de Lukasiewicz es tanto una r -implicación como una s -implicación por lo que son aplicables tanto las condiciones de activación de los modelos (M_R^T, I_R^T) como las de los modelos (M_S^T, I_S^T) , que como era de esperar coinciden.

Si T es la t -norma *min* la condición de activación es:

$$\forall u \in U, M_S^{\min}(P'(u), 1-P(u)) < 1$$

con:

$$M_S^{\min}(x, y) = \begin{cases} 0 & \text{si } x + y \leq 1 \\ y & \text{si } x + y > 1 \end{cases}$$

El caso más desfavorable se da para mínimo $P(u)$ y máximo $P'(u)$. Puede comprobarse que si $P'(u)=1$ y $P(u)=0$ se dan para un mismo valor de u entonces el modelo no se activa. De hecho si $P(u)=0$ debe cumplirse $\forall u \in U, M_S^{\min}(P'(u), 1) < 1$, es decir $P'(u)\neq 1$, y del mismo modo si $P'(u)=1$ entonces $P(u)\neq 0$, por lo que la condición de activación queda:

$$]P'[\subseteq]P[$$

El nivel de indeterminación que aparece en la conclusión de estos modelos viene dado en general por:

$$Q'(v) \geq \sup_u M_S^T(P'(u), 1-P(u))$$

Al igual que en el caso anterior la resolución general parece compleja, pero puede aproximarse analizando los dos casos extremos planteados en la sección anterior, esto es maximizando $P'(u)$ para $P(u)=0$, o bien minimizando $P(u)$ para $P'(u)=1$. Para t -normas positivas el modelo se activa sólo si para $P(u)=0$ se tiene $P'(u)\neq 1$. Por tanto suponiendo que el modelo está activado, el nivel de indeterminación puede estimarse minimizando $P(u)$ para $P'(u)=1$:

$$Q'(v) \geq \sup_{u/P'(u)=1} M_S^T(1, 1-P(u)) = \\ \sup_{u/P'(u)=1} (1-P(u)) = \\ 1 - \inf_{u/P'(u)=1} P(u) = 1 - N(P/[P'])$$

Para normas negativas ambos casos son posibles candidatos. La expresión que se obtiene maximizando $P'(u)$ para $P(u)=0$ es:

$$Q'(v) \geq \sup_{u/P(u)=0} M(P'(u), 1) = \sup_{u/P(u)=0} P'(u) = 1 - N([P]/P')$$

Y por tanto la aproximación propuesta deberá tomar el mayor de los valores de las expresiones anteriores.

4.8.5 p-implicaciones borrosas

Dadas las funciones:

$$I_{p-max}(x,y) = \begin{cases} 0 & \text{si } x=y=0 \\ 1 & \text{en otro caso} \end{cases}$$

$$I_{p-min}(x,y) = \begin{cases} 1 & \text{si } x=y=1 \\ 0 & \text{en otro caso} \end{cases}$$

y dada una p-implicación $I(x,y)$ cualquiera, se tiene:

$$I_{p-min}(x,y) \leq I(x,y) \leq I_{p-max}(x,y)$$

Ninguna de estas dos funciones es una p-implicación al no cumplir $I(1,y)=y$, pero ambas servirán para probar que cualquier p-implicación basa su condición de activación en algún tipo de intersección de P' en P .

En efecto tomando I_{p-max} la condición de activación se traduce en:

$$\exists u \in U / M(P'(u), I_{p-max}(P(u), 1)) > 0 \\ \Leftrightarrow \exists u \in U / P(u) > 0 \text{ y } M(P'(u), 1) > 0$$

mientras que tomando la función I_{p-min} la condición de activación que se obtiene es:

$$\exists u \in U / M(P'(u), I_{p-min}(P(u), 1)) > 0 \\ \Leftrightarrow \exists u \in U / P(u) = 1 \text{ y } M(P'(u), 1) > 0$$

Vemos por tanto que la función I_{p-max} permite obtener la condición de activación menos restrictiva, ya que sólo exige $P(u) \neq 0$ mientras que I_{p-min} exige $P(u)=1$ dando lugar a la condición más restrictiva. Al igual que hicimos con las n -implicaciones escogeremos por tanto la función de agregación M de forma distinta en ambos casos, para así obtener condiciones de activación los más extremas posibles.

Escogiendo M para I_{p-max} tal que $M(x,1)=1$ si $x \neq 0$ y $M(x,1)=0$ en otro caso, la activación se da si existe algún valor de u tal que $P(u) \neq 0$ y $P'(u) \neq 0$ simultáneamente, y se obtiene la condición de activación:

$$]P'[\cap]P[\neq 0$$

Escogiendo M para I_{p-min} tal que $M(x,1)=0$ si $x \neq 1$, el modelo se activa si $P(u)$ y $P'(u)$ valen 1 simultáneamente, por lo que la condición de activación queda:

$$[P']\cap[P] \neq 0$$

Queda por tanto probado que cualquier p -implicación basa su condición de activación en algún tipo de intersección de P' con P , siendo los dos casos extremos $]P'[\cap]P[\neq 0$ y $[P']\cap[P] \neq 0$.

4.8.5.1 El modelo (M_R^T, M_R^T)

Si la t -norma es arquímedea la condición de activación puede expresarse en función del generador aditivo:

$$\exists u \in U / f^{t-1}(f(P'(u)+f(P(u))) > 0$$

Si T es una t -norma positiva entonces $f^{t-1}(\infty)=0$. Por tanto la condición de activación exige que para algún valor de u se tenga:

$$f(P'(u)+f(P(u))) \neq \infty$$

y como $f(x)$ se hace infinito para x nulo, el modelo se activa si $\exists u \in U / P(u) \neq 0$ y $P'(u) \neq 0$ simultáneamente, es decir:

$$]P'[\cap]P[\neq 0$$

Si T es negativa, el modelo se activa si:

$$\exists u \in U / f(P'(u)) + f(P(u)) < f(0)$$

cuya solución depende en general de f . Por ejemplo para la familia de Yager la condición de activación queda:

$$(1-P'(u))^k + (1-P(u))^k < 1$$

En este ejemplo la condición de activación es variable con k de modo que cuanto mayor es k menos restrictiva es. Para $k=1$ se recupera la t-norma de Lukasiewicz cuya condición de activación es:

$$P'(u) + P(u) > 1$$

Podemos señalar que es condición suficiente aunque no necesaria, para que cualquiera de estos modelos se active, que se cumpla:

$$[P'] \cap P \neq 0 \text{ o bien }]P' \cap [P \neq 0$$

En efecto $f([P](u)) = f(1) = 0$ y $f(]P](u)) < f(0)$ al ser $]P](u) > 0$. Sin embargo la condición $]P' \cap [P \neq 0$ no es en ningún caso suficiente para t-normas negativas.

Si T es la t-norma \min entonces $\min(P(u), P'(u)) > 0$ que es equivalente a:

$$]P' \cap [P \neq 0$$

La altura de la conclusión viene dada por $\sup_u M_R^T(P'(u), M_R^T(P(u), 1))$:

$$Q'(v) \leq \Pi_R^T(P/P') = \sup_u T(P'(u), P(u)).$$

4.8.5.2 El modelo (M_S^T, M_S^T)

Si la t-norma de la que proviene la pseudo-conjunción es arquimedea la condición de activación puede expresarse en función de su generador aditivo:

$$\sup_u M_S^T(P'(u), M_S^T(P(u), 1)) > 0$$

es decir:

$$\exists u \in U / g^{[-1]}(g(1) - g(1 - P(u)) - g(1 - P'(u))) > 0$$

y dado el $g^{f-1}(x)$ sólo se anula para $x=0$, puede escribirse:

$$\exists u \in U / g(1-P(u)) + g(1-P'(u)) < g(1)$$

condición que en general depende de g .

Si g es el generador de una t-conorma positiva, $g(1)=\infty$, basta con que se tenga $1-P(u) \neq 1$ y $1-P'(u) \neq 1$ simultáneamente de forma que el miembro de la derecha de la desigualdad sea finito. En este caso la condición de activación es $P(u) \neq 0$ y $P'(u) \neq 0$ simultáneamente, o lo que es lo mismo:

$$]P'[\cap]P[\neq 0$$

Si g es el generador de una t-conorma negativa, pongamos por ejemplo la familia de Yager, la condición de activación, dependiente de g , es de la forma:

$$(1-P'(u))^k + (1-P(u))^k > 1$$

que es la misma que teníamos para los modelos basados en una p-implicación utilizando una t-norma. De hecho para cualquier modelo (M_S^T, M_S^T) generado a partir de una t-norma negativa la condición de activación es la misma que la del modelo (M_R^T, M_R^T) generado a partir de la misma t-norma, ya que tomando $f(x)=g(1-x)$ se tiene:

$$g(1-P(u)) + g(1-P'(u)) < g(1)$$

o bien expresándolo en función del generador de la t-norma:

$$\exists u / f(P(u)) + f(P'(u)) < f(0)$$

De nuevo puede comprobarse que en el ejemplo de la familia de Yager, cuando $k=1$ se recupera la t-norma de Lukasiewicz y la condición de activación queda:

$$\exists u / P'(u) + P(u) > 1$$

mientras que a medida que k aumenta la condición se hace cada vez más restrictiva exigiéndose que $P'(u)$ y $P(u)$ sean cada vez mayores para un mismo valor de u .

Si T es la t-norma min, la pseudo-conjunción resultante es:

$$M_S^{min}(x,y) = \begin{cases} 0 & \text{si } x + y \leq 1 \\ y & \text{si } x + y > 1 \end{cases}$$

y la condición de activación queda $P'(u) \neq 0$ y $P(u) \neq 0$, es decir:

$$]P'[\cap]P[\neq 0$$

La altura de la conclusión viene dada por:

$$\begin{aligned} \sup_u M_S^T(P'(u), M_S^T(P(u), 1)) &= \sup_u M_S^T(T(P'(u), P(u)), 1) = \\ M_S^T(\sup_u T(P'(u), P(u)), 1) &= M_S^T(\Pi_R^T(P/P'), 1) \end{aligned}$$

que en general depende de si T es positiva o negativa. Si T es positiva se tiene $M_S^{T+}(x, 1) = 1$ por lo que la altura vale 1. En el caso de que sea negativa $M_S^{T+}(x, 1) = g^{[-1]}(g(1) - g(1-x)) = 1 - f^{[-1]}(f(0) - f(x))$, que en general depende de f , con lo que queda:

$$Q'(v) \leq 1 - f^{[-1]}(f(0) - \Pi_R^T(P/P'))$$

4.8.5.3 El modelo (M_R^T, M_S^T)

La condición de activación viene dada por:

$$\sup_u M_R^T(P'(u), M_S^T(P(u), 1)) > 0$$

Si la t-norma es positiva sabemos que $M_S^{T+}(P(u), 1) = 1$ salvo si $P(u) = 0$, en cuyo caso $M_S^{T+}(P(u), 1) = 0$, y como $M_R^{T+}(P'(u), 0) = 0$ es necesario que se tenga $P(u) \neq 0$. Por otro lado si $P(u) > 0$ la condición $M_R^{T+}(P'(u), 0) = 0$ exige $P'(u) > 0$ por lo que finalmente se obtiene la condición de activación:

$$]P[\cap]P'[\neq 0$$

En el caso en que la t-norma generadora sea negativa, entonces se tiene:

$$\begin{aligned} \sup_u M_R^{T-}(P'(u), M_S^T(P(u), 1)) > 0 &\Rightarrow \\ \sup_u f^{[-1]}(f(P'(u)) + f(1 - f^{[-1]}(f(0) - f(P(u)))) > 0 \end{aligned}$$

condición relativamente compleja que depende del generador aditivo de la t-norma generadora, pero que en cualquier caso es más restrictiva que la anterior. Si puede

comprobarse que es suficiente que $[P] \cap]P'[\neq 0$ o bien $]P[\cap [P'] \neq 0$ para que se produzca la activación del modelo, aunque a priori no es condición necesaria.

Cuando la t-norma es positiva la altura de la conclusión viene dada por:

$$Q'(v) \leq \sup_u M_R^T(P'(u), M_S^T(P(u), 1)) = \sup_{u/P(u)>0} M_R^T(P'(u), 1) \Rightarrow Q'(v) \leq \sup_{u/P(u)>0} P'(u) = \Pi(JP[/P')$$

Cuando la t-norma es negativa la expresión es más compleja, depende de f , y viene dada por:

$$Q'(v) \leq \sup_u M_R^T(P'(u), M_S^T(P(u), 1)) = \sup_u f^{f^{-1}}(f(P'(u)) + f(1 - f^{f^{-1}}(f(0) - f(P(u)))))$$

Sí es importante señalar que al igual que la condición de activación es más restrictiva cuando la t-norma es negativa, también es menor la altura de la conclusión.

4.8.5.4 El modelo (M_S^T, M_R^T)

La condición de activación viene dada por:

$$\sup_u M_S^T(P'(u), M_R^T(P(u), 1)) = \sup_u M_S^T(P'(u), P(u)) = \Pi_S^T(P'/P) > 0$$

y por tanto puede sencillamente expresarse en función de la medida de posibilidad generada por la pseudo-conjunción M_S^T . Por otro lado sabemos que $M_S^T(P'(u), P(u)) = 0$ para $P(u) + P'(u) \leq 1$ por lo que la condición de activación se traduce en:

$$P(u) + P'(u) > 1$$

La altura de la conclusión viene dada por $\sup_u M_S^T(P'(u), M_R^T(P(u), 1))$ y por tanto se tiene:

$$Q'(v) \leq \sup_u M_S^T(P'(u), P(u)) = \Pi_S^T(P/P')$$

4.8.6 Resumen

Las condiciones de activación para los cuatro modelos (M_R^T, I_R^T) , (M_S^T, I_S^T) , (M_R^T, M_R^T) y (M_S^T, M_S^T) , así como los niveles de indeterminación de las conclusiones para los basados en n-implicaciones, o bien las alturas para los basados en p-implicaciones, quedan resumidos en la siguiente tabla. La primera columna, como en tablas anteriores, es el

modelo analizado. La segunda columna es la condición de activación obtenida para dicho modelo, y la tercera depende del tipo de modelo. Para modelos basados en n-implicaciones es una cota mínima del nivel de indeterminación, mientras que para modelos basados en p-implicaciones es directamente la altura de la conclusión obtenida.

Modelo	Condición de activación	Q(v) mínimo para I _n Q(v) máximo para I _p
(M_R^{T+}, I_R^{T+})	$[P'] \subseteq]P[$	$Q'(v) \geq 1 - N(]P[/P')$
(M_R^T, I_R^T)	$[P'] \subseteq]P[$	$Q'(v) \geq f^{-1}(f(0) - f(N(]P[/P')))$ $Q'(v) \geq 1 - N(]P[/P')$
(M_S^T, I_S^T)	$[P'] \subseteq]P[$	$Q'(v) \geq 1 - N(P/[P'])$ $Q'(v) \geq 1 - N(]P[/P')$
(M_S^{T+}, I_S^{T+})	$]P'[\subseteq]P[$	$Q'(v) \geq 1 - N(P/[P'])$
(M_R^{T+}, M_S^{T+})	$]P'[\cap]P[\neq \emptyset$	$Q'(v) \leq \Pi(]P[/P')$
(M_R^T, M_S^T)	$\exists u / \sup_u M_R^T(P'(u), M_S^T(P(u), 1)) > 0$	$Q'(v) \leq \sup_u M_R^T(P'(u), M_S^T(P(u), 1))$
(M_S^T, M_R^T)	$\exists u / P(u) + P'(u) > 1$	$Q'(v) \leq \sup_u \Pi_S^T(P/P')$
(M_S^{T+}, M_R^{T+})	$\exists u / P(u) + P'(u) > 1$	$Q'(v) \leq \sup_u \Pi_S^{T+}(P/P')$
(M_R^{T+}, M_R^{T+})	$]P'[\cap]P[\neq \emptyset$	$Q'(v) \leq \Pi_R^{T+}(P/P')$
(M_R^T, M_R^T)	$\exists u / f(P'(u)) + f(P(u)) < f(0)$	$Q'(v) \leq \Pi_R^T(P/P')$
(M_S^T, M_S^T)	$\exists u / f(P'(u)) + f(P(u)) < f(0)$	$Q'(v) \leq 1 - f^{-1}(f(0) - \Pi_R^T(P/P'))$
(M_S^{T+}, M_S^{T+})	$]P'[\cap]P[\neq \emptyset$	$Q'(v) \leq 1$

La condición de activación de cualquier modelo basado en una n-implicación se corresponde con algún tipo de inclusión de la observación P' en la hipótesis P . Del mismo modo, la de cualquier modelo basado en una p-implicación se corresponde con algún tipo de intersección de P' con P . Podemos interpretarlo del siguiente modo. Dada una hipótesis “ u es P ” y una observación “ u es P' ”, una medida de inclusión de P' en P es una medida de la necesidad o certeza que se tiene de que la hipótesis P sea cierta, o lo que es lo mismo de cómo de necesario es que “ u es P ” sea verdadero dado P' . Dependiendo de unos u otros modelos veremos que esta medida puede interpretarse como un grado de verdad o un grado de certidumbre, pero que siempre se traduce en un grado de certidumbre en la conclusión, puesto que determina su nivel de indeterminación. Por tanto las reglas cuyo modelo se basa en una n-implicación pueden considerarse como reglas de necesidad, ya que concluyen distribuciones de necesaria posibilidad y se activan cuando la necesidad de que la hipótesis sea verdadera es mayor que 0. Por ello hemos denominado a este tipo de modelos modelos de necesidad o n-modelos.

Así mismo una medida de la intersección de P' y P es una medida de la posibilidad que se tiene de que la hipótesis P sea verdadera dada la observación P' . La posibilidad es en cierto modo una medida de la contradicción existente entre el significado de P y el de P' .

Las reglas cuyo modelo se basa en una p-implicación pueden considerarse como reglas de posibilidad ya que se activan cuando alguna medida de posibilidad de que la hipótesis sea verdadera es mayor que cero, y diremos que su modelo es un modelo de posibilidad. Por eso hemos denominado a este tipo de modelos modelos de posibilidad o p-modelos.

4.9 Interpretación del comportamiento global de los modelos analizados

Recordamos por comodidad los resultados obtenidos en los apartados anteriores (4.5, 4.6, 4.7 y 4.8), resumidos en forma de tablas (recordar que en la primera de las tablas $P'(u)=\{u_0\}$ y $\alpha=P(u_0)$):

Observación puntual

Modelo	$[Q']$	$]Q'$	$Q(v) \in]0,1[$
(M_R^{T+}, I_R^{T+})	$\{v / Q(v) > \alpha\}$	$]Q[$	igual o aumentan
(M_R^T, I_R^T)	$\{v / Q(v) > \alpha\}$	$Q'(v) \geq f^{-1}(f(0)-f(\alpha))$	aumentan
(M_S^T, I_S^T)	$\{v / Q(v) \geq g^{-1}(g(1)-g(1-\alpha))\}$	$Q'(v) \geq 1-\alpha$	aumentan
(M_S^{T+}, I_S^{T+})	$]Q[$	$Q'(v) \geq 1-\alpha$	igual o aumentan
(M_Y, I_Y)	$]Q[$	$]Q[$	aumentan
(M_R^{T+}, M_R^{T+})	$Q'(v) \leq \alpha$	$]Q[$	igual o disminuyen
(M_R^T, M_R^T)	$Q'(v) \leq \alpha$	$\{v / Q(v) \leq f^{-1}(f(0)-f(\alpha))\}$	disminuyen
(M_S^T, M_S^T)	$Q'(v) \leq g^{-1}(g(1)-g(1-\alpha))$	$\{v / Q(v) \leq 1-\alpha\}$	disminuyen
(M_S^{T+}, M_S^{T+})	$]Q[$	$\{v / Q(v) \leq 1-\alpha\}$	igual o disminuyen
(M_Y, I_Y)	$]Q[$	$]Q[$	disminuyen

Equivalencia RCI-MS

Modelo	Cond. eq. RCI-MS	$S(P',P)$
(M_R^{T+}, I_R^{T+})	$\mu_{ \alpha } \leq \varpi(x) \leq I_R^{T+}(\alpha, x) \text{ C.N. }]P'[\subset]P[$	$\alpha = \inf_{u/P'(u)=1} P(u)$
(M_R^T, I_R^T)	$\mu_{ \alpha } \leq \varpi(x) \leq I_R^T(\alpha, x)$	$\alpha = \inf_{u/P'(u)=1} P(u)$
(M_S^T, I_S^T)	$\mu_{ \alpha } \leq \varpi(x) \leq I_R^T(\alpha, x)$	$\alpha = \inf_{u/P'(u)=1} P(u)$
(M_S^{T+}, I_S^{T+})	$\mu_{ \alpha } \leq \varpi(x) \leq I_R^{T+}(\alpha, x) \text{ C.N. }]P'[\subset]P[$	$\alpha = \inf_{u/P'(u)=1} P(u)$
(M_Y, I_Y)	$\mu_{ \alpha } \leq \varpi(x) \leq I_R^{prod}(\alpha, x) \text{ C.N. }]P'[\subset]P[$	$\alpha = \inf_{u/P'(u)=1} P(u)$
(M_R^{T+}, M_S^{T+})	$\mu_{ \alpha } \leq \varpi(x) \leq I_R^{T+}(x, \alpha)$	$\alpha = \sup_{u/P'(u)=1} P(u)$
(M_R^T, M_S^T)	$\mu_{ \alpha } \leq \varpi(x) \leq I_R^T(x, \alpha)$	$\alpha = \sup_{u/P'(u)=1} P(u)$
(M_S^{T+}, M_R^{T+})	$\mu_{ \alpha } \leq \varpi(x) \leq I_R^{T+}(1-\alpha, 1-x) \text{ C.N. }]P'[\cap]P[=0$	$\alpha = \sup_{u/P'(u)=1} P(u)$
(M_S^T, M_R^T)	$\mu_{ \alpha } \leq \varpi(x) \leq I_R^T(1-\alpha, 1-x)$	$\alpha = \sup_{u/P'(u)=1} P(u)$
(M_R^{T+}, M_R^{T+})	$\varpi(x) \leq I_R^{T+}(x, \alpha) \Leftrightarrow]P'[\cap]P[\neq 0$	$\alpha = \sup_u M_R^{T+}(P'(u), P(u))$
(M_R^T, M_R^T)	$\varpi(x) \leq I_R^T(x, \alpha) \Leftrightarrow f(P'(u)) + f(P(u)) < f(0)$	$\alpha = \sup_u M_R^T(P'(u), P(u))$
(M_S^T, M_S^T)	$\varpi(x) \leq I_R^T(x, \alpha) \Leftrightarrow f(P'(u)) + f(P(u)) < f(0)$	$\alpha = \sup_u M_R^T(P'(u), P(u))$
(M_S^{T+}, M_S^{T+})	$\varpi(x) \leq I_R^{T+}(x, \alpha) \Leftrightarrow]P'[\cap]P[\neq 0$	$\alpha = \sup_u M_R^{T+}(P'(u), P(u))$
(M_Y, M_Y)	$\varpi(x) \leq I_r^{prod}(x, \alpha) \Leftrightarrow]P'[\cap]P[\neq 0$	$\alpha = \sup_u P(u)^{1/P'(u)}$

Núcleos y soportes

Modelo	Núcleo / soporte
(M_R^{T+}, I_R^{T+})	$[Q'] = \{ v \in V / Q(v) \geq N(P/[P']) \}$
(M_R^T, I_R^T)	$[Q'] = \{ v \in V / Q(v) \geq N(P/[P']) \}$
(M_S^T, I_S^T)	$[Q'] = \{ v / Q(v) \geq 1 - f^{T+}(f(0) - f(N(P/[P']))) \}$
(M_S^{T+}, I_S^{T+})	$[Q] = [Q] \circ \emptyset$ si $]P'[\cap]P[= 0$
(M_R^{T+}, M_S^{T+})	$]Q' [= \{ v \in V / Q(v) \leq 1 - \Pi(P/[P']) \}$
(M_R^T, M_S^T)	$]Q' [= \{ v / Q(v) \leq 1 - \sup_u M_R^T(M_S^T(P'(u), 1), P(u)) \}$
(M_S^T, M_R^T)	$]Q' [= \{ v \in V / Q'(v) \leq 1 - \Pi_S^T(P'/P) \}$
(M_S^{T+}, M_R^{T+})	$]Q' [= \{ v \in V / Q'(v) \leq 1 - \Pi_S^{T+}(P'/P) \}$
(M_R^{T+}, M_R^{T+})	$]Q' [=]Q[$
(M_R^T, M_R^T)	$]Q' [= \{ v / Q(v) \leq f^{T+}(f(0) - f(\Pi_R^T(P/P'))) \}$
(M_S^T, M_S^T)	$\{ v / Q(v) \leq 1 - \Pi_R^T(P/P') \}$
(M_S^{T+}, M_S^{T+})	$\{ v / Q(v) \leq 1 - \Pi_R^{T+}(P/P') \}$

Condiciones de activación

Modelo	Condición de activación	Q(v) mínimo para I _n Q(v) máximo para I _p
(M_R^{T+}, I_R^{T+})	$]P'[\subseteq]P[$	$Q'(v) \geq 1 - N(P/[P'])$
(M_R^T, I_R^T)	$]P'[\subseteq]P[$	$Q'(v) \geq f^{T+}(f(0) - f(N(P/[P'])))$ $Q'(v) \geq 1 - N(P/[P'])$
(M_S^T, I_S^T)	$]P'[\subseteq]P[$	$Q'(v) \geq 1 - N(P/[P'])$ $Q'(v) \geq 1 - N(P/[P'])$
(M_S^{T+}, I_S^{T+})	$]P'[\subseteq]P[$	$Q'(v) \geq 1 - N(P/[P'])$
(M_R^{T+}, M_S^{T+})	$]P'[\cap]P[\neq 0$	$Q'(v) \leq \Pi(P/[P'])$
(M_R^T, M_S^T)	$\exists u / \sup_u M_R^T(P'(u), M_S^T(P(u), 1)) > 0$	$Q'(v) \leq \sup_u M_R^T(P'(u), M_S^T(P(u), 1))$
(M_S^T, M_R^T)	$\exists u / P(u) + P'(u) > 1$	$Q'(v) \leq \sup_u \Pi_S^T(P/P')$
(M_S^{T+}, M_R^{T+})	$\exists u / P(u) + P'(u) > 1$	$Q'(v) \leq \sup_u \Pi_S^{T+}(P/P')$
(M_R^{T+}, M_R^{T+})	$]P'[\cap]P[\neq 0$	$Q'(v) \leq \Pi_R^{T+}(P/P')$
(M_R^T, M_R^T)	$\exists u / f(P'(u)) + f(P(u)) < f(0)$	$Q'(v) \leq \Pi_R^T(P/P')$
(M_S^T, M_S^T)	$\exists u / f(P'(u)) + f(P(u)) < f(0)$	$Q'(v) \leq 1 - f^{T+}(f(0) - \Pi_R^T(P/P'))$
(M_S^{T+}, M_S^{T+})	$]P'[\cap]P[\neq 0$	$Q'(v) \leq 1$

El análisis de los resultados obtenidos permite deducir la existencia de dos tipos de modelos fundamentalmente distintos: los modelos de necesidad y los modelos de posibilidad. Diremos que un modelo para una regla es un **modelo de necesidad** cuando su condición de activación sea una medida de la inclusión de P' en P o lo que es lo mismo, cuando esté generado a partir de una π -implicación. Del mismo modo diremos que un modelo para una regla es un **modelo de posibilidad** cuando su condición de activación sea una medida de la intersección de P' y P , o lo que es lo mismo cuando esté generado a partir de una p -implicación. La condición de activación del modelo permite interpretar las conclusiones de los modelos de necesidad como distribuciones de necesaria posibilidad, y

las de los modelos de posibilidad como distribuciones de posible posibilidad. Nótese que la interpretación propuesta para las distribuciones resultantes determina el modo en que conclusiones procedentes de reglas modeladas con modelos del mismo tipo deben ser agregadas, ya sea con operadores conjuntivos o disyuntivos, y que este resultado es perfectamente coherente con el modo en que habitualmente se agregan tanto en trabajos teóricos como en aplicaciones numéricas.

4.9.1 Los modelos de necesidad

Las siguientes tablas resumen los resultados obtenidos sobre los modelos de necesidad estudiados, reorganizados de distinto modo. La primera de ellas refleja el comportamiento general de los modelos de necesidad, de forma que en la primera columna está el modelo analizado, en la segunda y tercera el núcleo y las cotas inferiores aproximadas del nivel de indeterminación de la conclusión, y en la última la condición de activación. En la segunda tabla está el comportamiento de los mismos modelos cuando se verifican las condiciones de equivalencia entre la RCI y la MS, de forma que en la primera columna está el modelo analizado, en la segunda y tercera el núcleo y el nivel de indeterminación de la conclusión, y en la cuarta la condición para que se dé la equivalencia entre ambos mecanismos de inferencia.

Comportamiento general del los modelos de necesidad

Modelo	núcleo	nivel de indeterminación	activación
(M_R^{T+}, I_R^{T+})	$\{v \in V / Q(v) \geq N(P/[P'])\}$	$Q'(v) \geq 1 - N(P/[P])$	$[P'] \subseteq]P[$
(M_R^T, I_R^T)	$\{v \in V / Q(v) \geq N(P/[P'])\}$	$Q'(v) \geq f^{-1}(f(0) - f(N(P/[P'])))$ $Q'(v) \geq 1 - N(P/[P])$	$[P'] \subseteq]P[$
(M_S^T, I_S^T)	$\{v / Q(v) \geq 1 - f^{-1}(f(0) - f(N(P/[P'])))\}$	$Q'(v) \geq 1 - N(P/[P'])$ $Q'(v) \geq 1 - N(P/[P])$	$[P'] \subseteq]P[$
(M_S^{T+}, I_S^{T+})	$[Q]$	$Q'(v) \geq 1 - N(P/[P'])$	$]P' [\subseteq]P [$

Comportamiento de los modelos de necesidad con equivalencia RCI-MS

Modelo	núcleo	nivel de indeterminación	cond. equiv.
(M_R^{T+}, I_R^{T+})	$\{ v \in V / Q(v) \geq N(P/[P']) \}$	$]Q[$	$\mu_{\alpha_j} \leq \tau(x) \leq I_R^{T+}(\alpha, x)$ $\Rightarrow]P'[\subseteq]P[$
(M_R^T, I_R^T)	$\{ v \in V / Q(v) \geq N(P/[P']) \}$	$Q'(v) \geq f^{-1}(f(0)-f(N(P/[P'])))$	$\mu_{\alpha_j} \leq \tau(x) \leq I_R^T(\alpha, x)$
(M_S^T, I_S^T)	$\{ v / Q(v) \geq 1-f^{-1}(f(0)-f(N(P/[P']))) \}$	$Q'(v) \geq 1-N(P/[P'])$	$\mu_{\alpha_j} \leq \tau(x) \leq I_R^T(\alpha, x)$
(M_S^{T+}, I_S^{T+})	$]Q[$	$Q'(v) \geq 1-N(P/[P'])$	$\mu_{\alpha_j} \leq \tau(x) \leq I_R^{T+}(\alpha, x)$ $\Rightarrow]P'[\subseteq]P[$

4.9.1.1 Grado de verdad y grado de certidumbre

Aceptando las simplificaciones propuestas durante el análisis de las condiciones de activación, podemos considerar que los modelos de necesidad analizados se basan principalmente en dos medidas de inclusión.

La primera de ellas puede interpretarse como una generalización del grado de verdad de una hipótesis dada una observación puntual, al caso en que la observación es un conjunto, en el sentido en que es el grado de compatibilidad más pesimista de la hipótesis con el conjunto de elementos totalmente observados (pertenecientes al núcleo de la observación). Esta medida puede interpretarse como una estimación del grado de verdad de la hipótesis, que llamaremos **grado de verdad** (o compatibilidad) de la hipótesis dada la observación, y se obtiene como el grado de inclusión de $[P']$ en P :

$$N(P/[P']) = \inf_{u/P'(u)=1} P(u)$$

Por otro lado la segunda medida que juega un papel importante en la inferencia viene dada por:

$$N(]P[/P') = \inf_{u/P(u)=0} (1-P'(u))$$

o bien su complemento:

$$\Pi(-]P[/P') = 1 - N(]P[/P') = \sup_{u/P(u)=0} P'(u)$$

La medida $N(]P[/P')$ es el grado de inclusión de P' en $]P[$ y puede interpretarse como una medida de la certidumbre o confianza que se tiene en que el grado de verdad de P sea no nulo, puesto que $\Pi(-]P[/P')$ mide el grado máximo en que han sido observados aquellos elementos del universo de discurso no tenidos en cuenta en modo alguno por la hipótesis P . En el momento en que existen elementos u tales que $P'(u)=1$ y $P(u)=0$ la

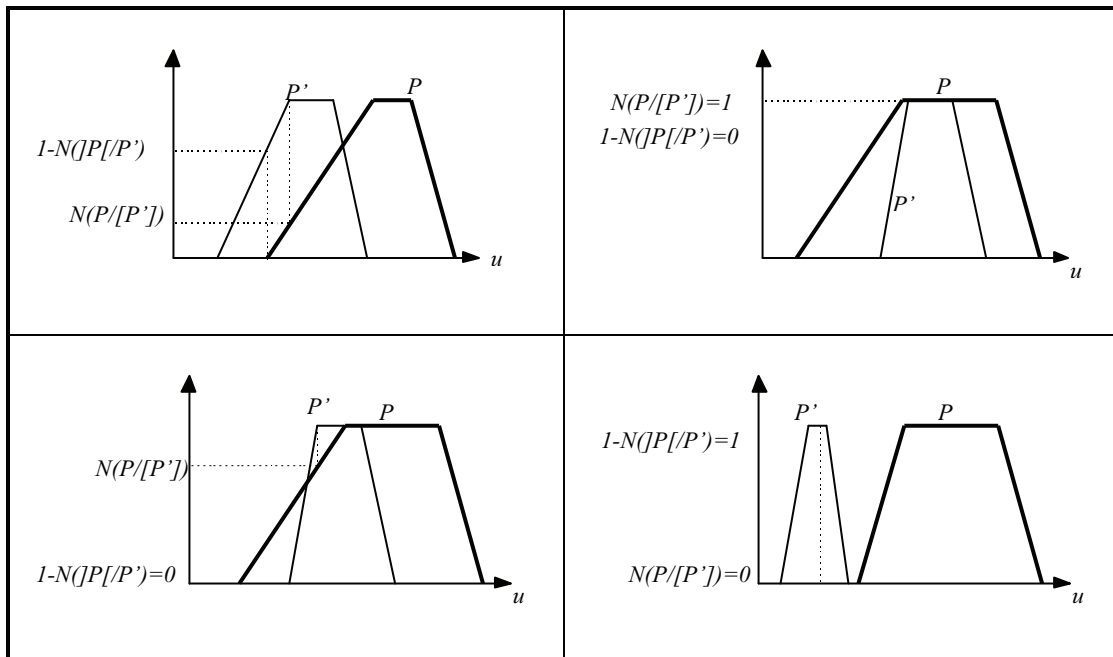
certidumbre se hace lógicamente nula. $\Pi(\neg]P[/P')$ es la incertidumbre que en el dominio de lo necesario se tiene en que la hipótesis sea al menos parcialmente verdadera, o lo que es lo mismo, cómo de necesario es que la hipótesis sea al menos parcialmente verdadera.

Cuando el grado de verdad estimado por $N(P/[P'])$ es nulo, esto es, cuando el núcleo de P' no está incluido en el soporte de P , aunque el grado de verdad de P podría ser (en el ámbito de lo posible) mayor que 0, la hipótesis en el ámbito de lo necesario debe considerarse falsa. En este caso $N(]P[/P')$ da como resultado una certidumbre nula, que indica que no puede asegurarse que la hipótesis tenga un grado de verdad distinto de 0.

Podría calcularse la certidumbre de que la hipótesis fuese falsa o parcialmente falsa, lo que se haría contrastando $\neg P$ con P' (en lugar de P con P') dando lugar a la medida de certidumbre $\Pi(] \neg P[/P') = \Pi(]P[/P')$.

La medida $N(]P[/P')$ será denominada a partir de ahora **grado de certidumbre** de la hipótesis P , o simplemente **certidumbre** de P y como hemos dicho puede interpretarse como la certidumbre de que P sea al menos parcialmente verdadero. Su grado de verdad lo estimaremos mediante la medida que hemos denominado **grado de verdad**, es decir $N(P/[P'])$.

Nótese que mientras que el grado de verdad compara (en términos de inclusión) $]P']$ con P , el grado de certidumbre compara P' con $]P[$. Algunas de las distintas situaciones que pueden presentarse entre P y P' quedan reflejadas en la siguientes figuras:



La comparación entre los distintos modelos de necesidad estudiados permite distinguir tres grandes modos de comportamiento, que se basan en tres formas distintas de modificar el consecuente para obtener la conclusión. Por un lado están los modelos que transforman grado de verdad en grado de verdad y grado de certidumbre en grado de certidumbre, por otro están aquellos que transforman grado de verdad en grado de certidumbre, y por último aquellos que presentan un comportamiento mixto entre los dos anteriores. Veremos que estos comportamientos pueden expresarse en función de las dos medidas aquí propuestas.

Estas diferencias semánticas entre unos y otros modelos se corresponden con las ya detectadas en [Dubois, Prade, 93] de las llamadas implicaciones graduales e implicaciones de certidumbre. En dicho trabajo se analiza el tipo de modificación del consecuente llevada a cabo por una implicación, asimilando diferentes comportamientos con diferentes semánticas. Sin embargo no utilizan el concepto de modelo de regla ni analizan las condiciones de activación de los mismos, a nuestro entender claves para conocer en profundidad el comportamiento de un modelo. Las condiciones de activación permiten por un lado conocer qué condiciones mínimas deben verificar la hipótesis y la observación para inferir una conclusión no trivial, de donde se deduce un índice de similitud que caracteriza parte del comportamiento del modelo, o y por otro, permiten interpretar las distribuciones de posibilidad resultantes en términos de distribuciones de posible posibilidad o necesaria posibilidad, que como hemos visto presentan propiedades substancialmente distintas, y deben ser interpretadas de forma también diferente. Tampoco se tiene en cuenta la equivalencia RCI-MS que permite determinar en qué índice de similitud se basa un modelo para llevar a cabo la modificación del consecuente, lo que permite profundizar en su interpretación semántica.

4.9.1.2 Modelos de necesidad graduales

Son aquellos que transforman grado de verdad en grado de verdad, de forma que cuanto más verdadera es la hipótesis en relación con la observación, más verdadera es la conclusión en relación con el consecuente de la regla. Esto no quiere decir que la conclusión esté carente de incertidumbre, ya que estos modelos también transforman incertidumbre en incertidumbre, de forma que la incertidumbre de la hipótesis aparece también en la conclusión de la regla.

Su comportamiento cualitativo puede resumirse del siguiente modo:

$$[Q'] \geq [Q]$$

$$]Q'[\geq]Q[$$

$N(Q/[Q'])$ función de $N(P/[P'])$

$N(JQ/[Q'])$ función de $N(JP/[P'])$

Estos modelos están generados por r-implicaciones residuadas de t-normas positivas, $(M_R^{T^+}, I_R^{T^+})$, y modifican el núcleo del consecuente en base a al grado de verdad de la hipótesis dada la observación. Además a menos que exista incertidumbre en la comprobación de la hipótesis, no modifican su soporte y se activan únicamente en base al grado de verdad de P dado P' .

Los modelos de necesidad graduales se caracterizan por no excluir del conjunto de valores totalmente posibles a ningún elemento del universo de discurso del consecuente cuyo grado de pertenencia a éste sea mayor que un determinado valor, función del grado de verdad de la hipótesis de la regla. Aún cuando el valor de verdad de la hipótesis sea igual a la unidad, tampoco excluyen a ningún elemento del universo de discurso del consecuente, en el caso en que la certidumbre de la hipótesis sea menor que 1.

Supuesto que deseamos comparar el comportamiento de Q y Q' como posibles observaciones para contrastar una hipótesis cualquiera G , un aumento del núcleo de Q' respecto del de Q significa que los grados de verdad de G dados Q y Q' estarán en la relación:

$$N(G/[Q]) \geq N(G/[Q']), \text{ función de } N(P/[P'])$$

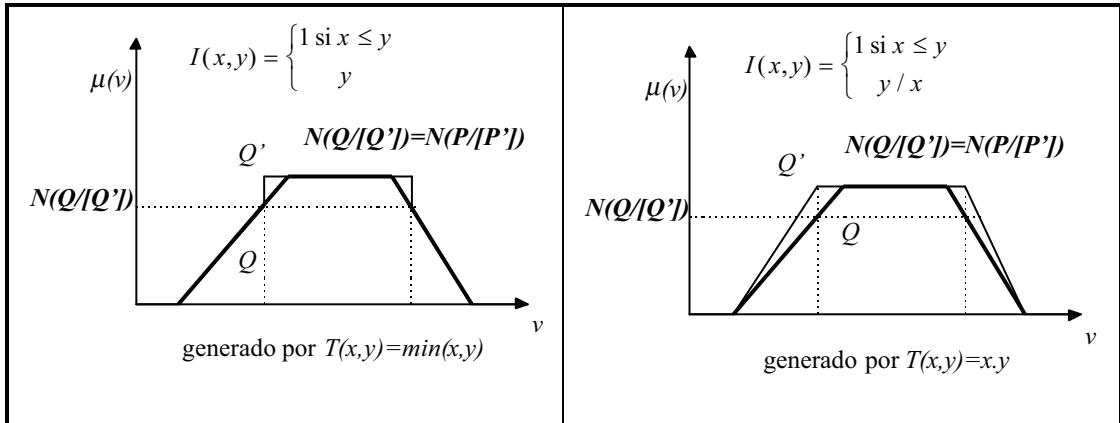
siendo $N(G/[Q'])$ menor cuanto menor sea $N(P/[P'])$. Esto significa que cuando el grado de verdad de la hipótesis dada la observación se reduce, se reduce también el grado de verdad de una hipótesis cualquiera contrastada con la conclusión de la regla.

Del mismo modo para las certidumbres de dicha hipótesis G se tiene:

$$N(JG/[Q]) \geq N(JG/[Q']), \text{ función de } N(JP/[P'])$$

siendo $N(JG/[Q'])$ menor cuanto menor sea $N(JP/[P'])$.

La siguiente figura muestra el comportamiento de los modelos graduales generados a partir de las t-normas *min* y *prod*, cuando se verifican las condiciones de equivalencia entre la RCI y la MS (lo que significa que el nivel de incertidumbre de la conclusión es igual al del consecuente).



4.9.1.3 Modelos de necesidad de certidumbre

Los modelos de necesidad de certidumbre son aquellos que transforman grado de verdad en grado de certidumbre, de forma que cuanto más verdadera es la hipótesis dada la observación, mayor es la certidumbre del consecuente en comparación con la de la conclusión de la regla. Su comportamiento puede resumirse cualitativamente como sigue:

$$[Q'] = [Q]$$

$$]Q'[\gg]Q[\text{ (aunque en rigor }]Q'[= \emptyset)$$

$$N([Q]/Q') \text{ función de } N(P/[P'])$$

Estos modelos son los generados a partir de implicaciones fuertes provenientes de t-normas positivas, es decir (M_S^{T+}, I_S^{T+}) . Modifican el soporte del consecuente introduciendo un nivel de indeterminación en la conclusión, mientras que no modifican su núcleo. No transmiten la certidumbre puesto que si la certidumbre de la hipótesis no es la unidad los modelos no se activan (debido a su restrictiva condición de activación, que veremos puede relajarse utilizando una FGMP menor).

Los modelos de necesidad de certidumbre se caracterizan por no excluir del conjunto de valores posibles a ningún elemento del universo de discurso del consecuente cuando el grado de verdad de la hipótesis es inferior a la unidad, y el grado mínimo de inclusión es función del grado de verdad de la hipótesis de la regla.

En el caso en que se utilicen Q y Q' para contrastar una hipótesis G cualquiera, al ser los núcleos de Q y Q' iguales, los grados de verdad de G dados Q y Q' serán iguales, esto es:

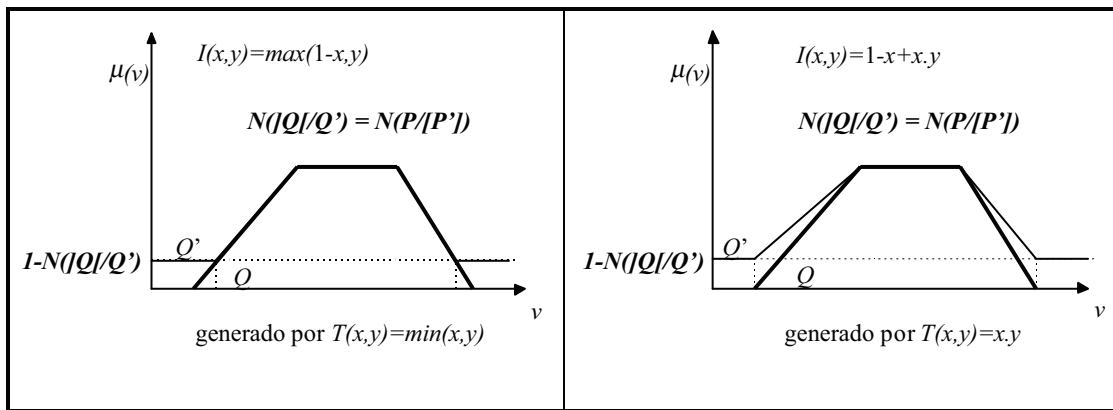
$$N(G/[Q]) = N(G/[Q'])$$

mientras que las certidumbres de dicha hipótesis guardan la relación:

$$N(JG[/Q]) \geq N(JG[/Q']), \text{ función de } N(P/[P'])$$

siendo $N(JG[/Q'])$ menor cuanto menor sea $N(P/[P'])$.

La siguiente figura muestra el comportamiento de los modelos de certidumbre generados por las t-normas *min* y *prod*, cuando se verifican las condiciones de equivalencia entre la RCI y la MS.



4.9.1.4 Modelos de necesidad mixtos

Llamamos modelos mixtos a aquellos que presentan un comportamiento que comparte características de los modelos graduales y de certidumbre. Estos modelos se caracterizan porque el grado de certidumbre de la conclusión depende tanto del grado de verdad como del grado de incertidumbre de la hipótesis, mientras que el grado de verdad de la conclusión depende únicamente del grado de verdad de la hipótesis. Su comportamiento cualitativo puede resumirse como sigue:

$$[Q'] \geq [Q]$$

$$]Q'[\gg]Q[\text{ (en rigor }]Q'[= \emptyset)$$

$$N(Q/[Q']) \text{ función de } N(P/[P'])$$

$$N(Q/Q') \text{ función de } N(P/[P']) \text{ y de } N(P/[P'])$$

Estos modelos, en concreto (M_R^{T-}, I_R^{T-}) y (M_S^{T-}, I_S^{T-}) , tienen todos ellos la misma condición de activación que se corresponde con la de los modelos graduales, menos restrictiva que la de los modelos de certidumbre. Por otro lado modifican el núcleo del consecuente en función del grado de verdad de la hipótesis, aunque el modelo (M_R^{T-}, I_R^{T-}) lo hace igual que

el modelo (M_R^{T+}, I_R^{T+}) , mientras que el modelo (M_S^{T-}, I_S^{T-}) genera un grado de verdad de la conclusión dado el consecuente igual o mayor que el obtenido para la hipótesis dada la observación. En cuanto a los soportes de las conclusiones, cuando la certidumbre de la hipótesis es I el modelo (M_R^{T-}, I_R^{T-}) introduce en la conclusión un nivel de indeterminación que es función del grado de verdad de la hipótesis, y menor o igual que su complemento a I . Por el contrario (M_R^{T+}, I_R^{T+}) introduce un nivel de indeterminación siempre igual al complemento a I de dicho grado de verdad. En caso de que la certidumbre sea menor que I , el nivel de indeterminación de la conclusión viene determinado tanto por el grado de certidumbre como por el grado de verdad de la hipótesis, dependiendo de sus magnitudes relativas. Esto hace que este tipo de modelos puedan considerarse mixtos en el sentido en que presentan comportamientos graduales y de certidumbre.

Los modelos de necesidad mixtos se caracterizan por no excluir del conjunto de valores posibles a ningún elemento del universo de discurso del consecuente cuando los grados de verdad o certidumbre de la hipótesis son inferiores a la unidad, y el grado mínimo de inclusión es función de ambos. Por otro lado también incluyen en el conjunto de valores totalmente posibles a todos aquellos elementos cuyo grado de pertenencia al consecuente sea mayor que un determinado valor, función del grado de verdad de la hipótesis. Nótese que al igual que los modelos graduales, aún cuando el valor de verdad de la hipótesis sea I , tampoco excluyen a ningún elemento del universo de discurso del consecuente (siempre que la certidumbre de la hipótesis sea menor que la unidad).

Si utilizamos Q y Q' para contrastar una hipótesis G cualquiera, entonces, dado que el núcleo y el soporte de Q' son siempre mayores que los de Q , se tiene:

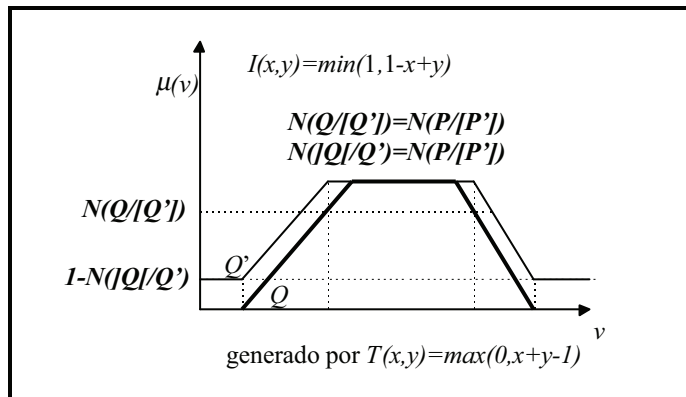
$$N(G/[Q]) \geq N(G/[Q']), \text{ función de } N(P/[P'])$$

siendo $N(G/[Q'])$ menor cuanto menor sea $N(P/[P'])$, y para las certidumbres se tiene también:

$$N(JG/[Q]) \geq N(JG/[Q']), \text{ función de } N(JP/[P']) \text{ y de } N(P/[P'])$$

siendo $N(JG/[Q'])$ menor cuanto menores sean $N(JP/[P'])$ y $N(P/[P'])$.

Un ejemplo de este tipo de modelos es el generado por la t-norma de Lukasiewicz, que se ilustra gráficamente en la siguiente figura, cuando se verifica la equivalencia RCI-MS:



4.9.1.5 Modificación de las condiciones de activación:

A partir de un modelo (M,I) formado a partir de una n -implicación I es posible obtener un nuevo modelo con distinta condición de activación pero con mismo tipo de modificación del consecuente cuando se verifica la equivalencia RCI-MS. Para ello basta con utilizar una FGMP distinta. En efecto eligiendo una función generadora de modus ponens M' menor, que sabemos que es una FGMP válida, (dado un modelo (M,I) cualquier función generadora de modus ponens $M' \leq M$ es función generadora de modus ponens para I) se obtiene un nuevo modelo con una condición de activación menos restrictiva. De este modo para obtener un modelo de necesidad con condición de activación dada por $[P'] \subseteq [P]$, podemos utilizar la implicación fuerte de Reichenbach generada a partir de la t -norma producto:

$$I(x,y) = I_S^P(x,y) = 1 - x + xy$$

y como función generadora de modus ponens la t -norma de Lukasiewicz:

$$M(x,y) = M_R^L(x,y) = \max(0, x+y-1)$$

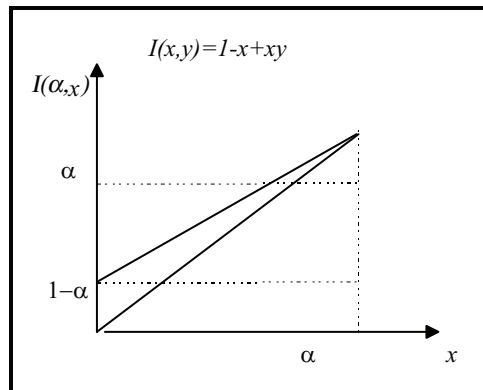
con lo que la semántica es de certidumbre, pero la condición de activación es menos exigente que la de los modelos de certidumbre vistos con anterioridad, y puede comprobarse que viene dada por $\sup_u P'(u) - P(u) > 0$. En cuanto al índice de similitud utilizado para la modificación del consecuente, recordando la forma de obtener la equivalencia entre la RCI y la MS se tiene:

$$\sup_x M(\tau(x), I(x,y)) \leq I(\alpha, y)$$

que aplicada al caso que nos ocupa conduce a:

$$\sup_x \tau(x) - x + xy = \sup_x \tau(x)(1 - x/\tau(x) + x/\tau(x) \cdot y) \leq 1 - \alpha + \alpha y$$

por tanto sí se verifica $x/\tau(x) \leq \alpha$, que equivale a $\tau(x) \leq x/\alpha$, que resulta ser la implicación residual generada por la t-norma producto. Esto da lugar al siguiente valor de verdad borroso como límite para $\tau(x)$:



y significa que el índice de similitud empleado sigue siendo:

$$\alpha = N(P/[P'])$$

y que grados de verdad parciales se transforman en grados de certidumbre, cuando la certidumbre de la hipótesis es 1, pero con la diferencia de que el modelo se activa aunque esta certidumbre no sea la unidad.

Para obtener este tipo de modelos es necesario encontrar FGMP suficientemente pequeñas, y una buena elección inicial suele ser la t-norma de Lukasiewicz que es una de las t-normas continuas más pequeñas.

Por otro lado es de esperar que una certidumbre parcial del antecedente se traduzca también en una certidumbre parcial del consecuente, por lo que en este sentido, cuando no se verifique la equivalencia RCI-MS, la certidumbre del consecuente se vea modificada igual que ocurre con los modelos mixtos, pero sin que haya modificación alguna de su grado de verdad.

4.9.2 Los modelos de posibilidad

Las siguientes tablas resumen los resultados obtenidos para los modelos de posibilidad. Para la primera tabla, en la primera columna está el modelo analizado, en la segunda y tercera el núcleo (o estrictamente hablando la altura) y el soporte de la conclusión, y en la última la condición de activación, que se corresponde con la condición de equivalencia RCI-MS para los modelos (M_R^T, M_R^T) y (M_S^T, M_S^T) . En la segunda tabla está el comportamiento de los modelos (M_R^T, M_S^T) y (M_S^T, M_R^T) cuando se da la equivalencia RCI-MS:

Comportamiento general del los modelos de posibilidad

Modelo	altura	soporte	activación \equiv c. eq. RCI-MS
(M_R^{T+}, M_S^{T+})	$Q'(v) \leq \Pi(P/P')$	$\{v / Q(v) \leq 1 - \Pi(P/P')\}$	$]P'[\cap]P[\neq 0$
(M_R^T, M_S^T)	$Q'(v) \leq \sup_u M_R^T(P'(u), M_S^T(P(u), 1))$	$\{v / Q(v) \leq 1 - \sup_u M_R^T(M_S^T(P'(u), 1), P(u))\}$	$\sup_u M_R^T(P'(u), M_S^T(P(u), 1)) > 0$
(M_S^T, M_R^T)	$Q'(v) \leq \Pi_S^T(P/P')$	$\{v / Q(v) \leq 1 - \Pi_S^T(P/P')\}$	$P'(u) + P(u) > 0$
(M_S^{T+}, M_R^{T+})	$Q'(v) \leq \Pi_S^{T+}(P/P')$	$\{v / Q(v) \leq 1 - \Pi_S^{T+}(P/P')\}$	$P'(u) + P(u) > 0$
(M_R^{T+}, M_R^{T+})	$Q'(v) \leq \Pi_R^{T+}(P/P')$	$]Q[$	$]P'[\cap]P[\neq 0$
(M_R^T, M_R^T)	$Q'(v) \leq \Pi_R^T(P/P')$	$\{v / Q(v) \leq f^{-1}(f(0) - f(\Pi_R^{T+}(P/P')))\}$	$f(P'(u)) + f(P(u)) < f(0)$
(M_S^T, M_S^T)	$Q'(v) \leq 1 - f^{-1}(f(0) - \Pi_R^T(P/P'))$	$\{v / Q(v) \leq 1 - \Pi_R^T(P/P')\}$	$f(P'(u)) + f(P(u)) < f(0)$
(M_S^{T+}, M_S^{T+})	$]Q[\text{ o } \emptyset \text{ si }]P'[\cap]P[= 0$	$\{v / Q(v) \leq 1 - \Pi_R^{T+}(P/P')\}$	$]P'[\cap]P[\neq 0$

Comportamiento de los modelos (M_R^T, M_S^T) y (M_S^T, M_R^T) con equivalencia RCI-MS

Modelo	altura	soporte	cond. equiv.
(M_R^{T+}, M_S^{T+})	$]Q[$	$\{v / Q(v) \leq 1 - \Pi(P/P')\}$	$\mu_{(\alpha)} \leq \alpha(x) \leq I_R^{T+}(x, \alpha)$
(M_R^T, M_S^T)	$Q'(v) \leq g^{[-1]}(g(1) - g(1 - \Pi(P/P')))$	$\{v / Q(v) \leq 1 - \Pi(P/P')\}$	$\mu_{(\alpha)} \leq \alpha(x) \leq I_R^T(x, \alpha)$
(M_S^T, M_R^T)	$Q'(v) \leq \Pi(P/P')$	$\{v / Q(v) \leq f^{-1}(f(0) - f(\Pi(P/P')))\}$	$\mu_{(\alpha)} \leq \alpha(x) \leq I_R^{T+}(1 - \alpha, 1 - x)$ C.N. $]P'[\cap]P[= 0$
(M_S^{T+}, M_R^{T+})	$Q'(v) \leq \Pi(P/P')$	$]Q[$	$\mu_{(\alpha)} \leq \alpha(x) \leq I_R^T(1 - \alpha, 1 - x)$

A la vista de los resultados anteriores podemos destacar la existencia de dos familias distintas de modelos de posibilidad. La primera familia constituida por los modelos (M_R^T, M_S^T) y (M_S^T, M_R^T) basa su comportamiento en varios índices de similitud entre P y P' , que se corresponden con medidas de posibilidad no simétricas.

La segunda familia de modelos, constituida por (M_R^T, M_R^T) y (M_S^T, M_S^T) basa la modificación del consecuente en un único índice de similitud entre P y P' , la medida de posibilidad habitual Π_R^T que es una medida simétrica. Además estos modelos siempre verifican la equivalencia RCI-MS, a diferencia de los anteriores.

4.9.2.1 Posible grado de verdad y posible grado de certidumbre

Dadas dos distribuciones de posibilidad parece razonable suponer que si en el ámbito de lo necesario habíamos definido el grado de verdad de P dado P' como $N(P/[P'])$ entonces en el ámbito de lo posible se pueda definir el posible grado de verdad como una versión optimista de $N(P/[P'])$. Si a partir de:

$$N(P/[P']) = \inf_{u/P'(u)=1} P(u)$$

obtenemos su medida de posibilidad dual se tiene:

$$\Pi(P/[P']) = 1 - N(\neg P/[P']) = 1 - \inf_{u/P'(u)=1} (1 - P(u)) = \sup_{u/P'(u)=1} P(u)$$

Esta medida es el índice de similitud empleado por los modelos (M_R^T, M_S^T) y (M_S^T, M_R^T) para modificar el consecuente cuando se verifican las condiciones de equivalencia RCI-MS. Como ya se dijo es una versión optimista de $N(P/[P'])$, y puede interpretarse como el posible grado de verdad que podría tener P dado P' . La expresión anterior supone que en cuanto se verifica $[P'] \cap P = \emptyset$ entonces su valor es 0. Esto significa que, dado que en general las distribuciones de posible posibilidad no tienen por qué ser normales, en cuanto la altura de P' no sea 1 el valor de la medida propuesta será nulo, por lo que en principio parece una medida demasiado restrictiva.

Por otro lado aplicando nuevamente la dualidad a la medida de incertidumbre en el dominio de lo necesario se recupera una de las medidas que aparece en los modelos de posibilidad. En efecto, a partir de la medida:

$$N(JP/[P']) = \inf_{u/P(u)=0} 1 - P'(u)$$

su medida dual viene dada por:

$$\Pi(JP/[P']) = 1 - N(\neg JP/[P']) = 1 - \inf_{u/P(u) > 0} 1 - P'(u) = \sup_{u/P(u) > 0} P'(u)$$

que resulta ser una (o la) versión optimista de $N(JP/[P'])$.

Si analizamos con más detalle las medidas en las que de forma aproximada se basan los modelos de posibilidad podemos señalar lo siguiente. Cuando las condiciones de equivalencia no se verifican, el soporte de la conclusión para los modelos (M_R^{T+}, M_S^{T+}) , (M_S^{T-}, M_R^{T-}) y (M_S^{T+}, M_R^{T+}) viene determinado respectivamente por las medidas de posibilidad $\Pi(P/]P'[)$, $\Pi_S^{T-}(P'/P)$, y $\Pi_S^{T+}(P'/P)$, mientras que la altura viene determinada respectivamente por las medidas $\Pi(]P[/P')$, $\Pi_S^{T-}(P/P')$, y $\Pi_S^{T+}(P/P')$.

En el caso del primer modelo la medida $\Pi(P/]P'[)$ es una generalización más optimista de la medida $\Pi(P/[P'])$ y puede considerarse como el posible grado de verdad que podría tener P dado P' . La medida $\Pi(]P[/P')$ puede a su vez interpretarse como una medida de la certidumbre de P , que viene principalmente condicionada por la certidumbre de la observación P' , al menos en la parte que interseca con P . Esta medida se obtiene como la altura de P' en la parte que interseca con P .

En el segundo caso (y algo similar ocurre con el modelo (M_R^{T-}, M_S^{T-})) las dos medidas empleadas son muy parecidas, y en el caso particular en que $T=T_L$ (es decir para la t-norma de Lukasiewicz) ambas coinciden y se tiene $\Pi_S^L(P'/P) = \Pi_S^L(P/P')$.

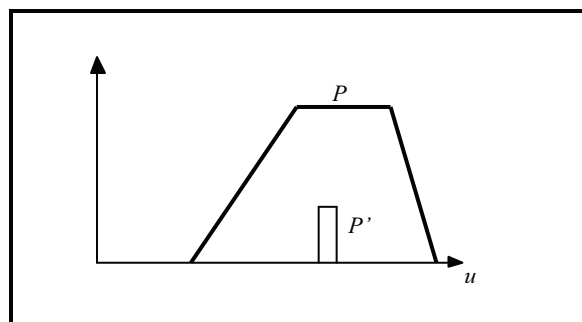
En el tercero de los casos, para la t-norma *min* las medidas $\Pi_S^{T+}(P'/P)$ y $\Pi_S^{T+}(P'/P)$ vienen dadas por:

$$\begin{aligned}\Pi_S^{\min}(P'/P) &= P'(u) \text{ si } P(u) + P'(u) \leq 1, \text{ 0 en otro caso} \\ \Pi_S^{\min}(P/P') &= P(u) \text{ si } P(u) + P'(u) \leq 1, \text{ 0 en otro caso}\end{aligned}$$

La primera de ellas, que es la que determina el soporte de la conclusión, es una medida de la altura de P' para aquellos valores de u para los que no es nula, esto es, tales que $P(u) + P'(u) > 0$. Por tanto puede interpretarse como una medida restrictiva de la certidumbre de P , y de hecho puede considerarse que está entre las medidas $\Pi(]P[/P')$ y $\Pi([P]/P')$. Por contra la segunda medida, $\Pi_S^{T+}(P/P')$, es, para aquellos valores de u para los que no es nula, una medida de la altura de P y por tanto puede considerarse como una medida restrictiva del grado de verdad de P , situada aproximadamente entre $\Pi(P/]P'[)$ y $\Pi(P/]P'[)$. Por tanto los modelos (M_R^{T+}, M_S^{T+}) y (M_S^{T+}, M_R^{T+}) modifican el consecuente para la obtención de la conclusión de forma opuesta.

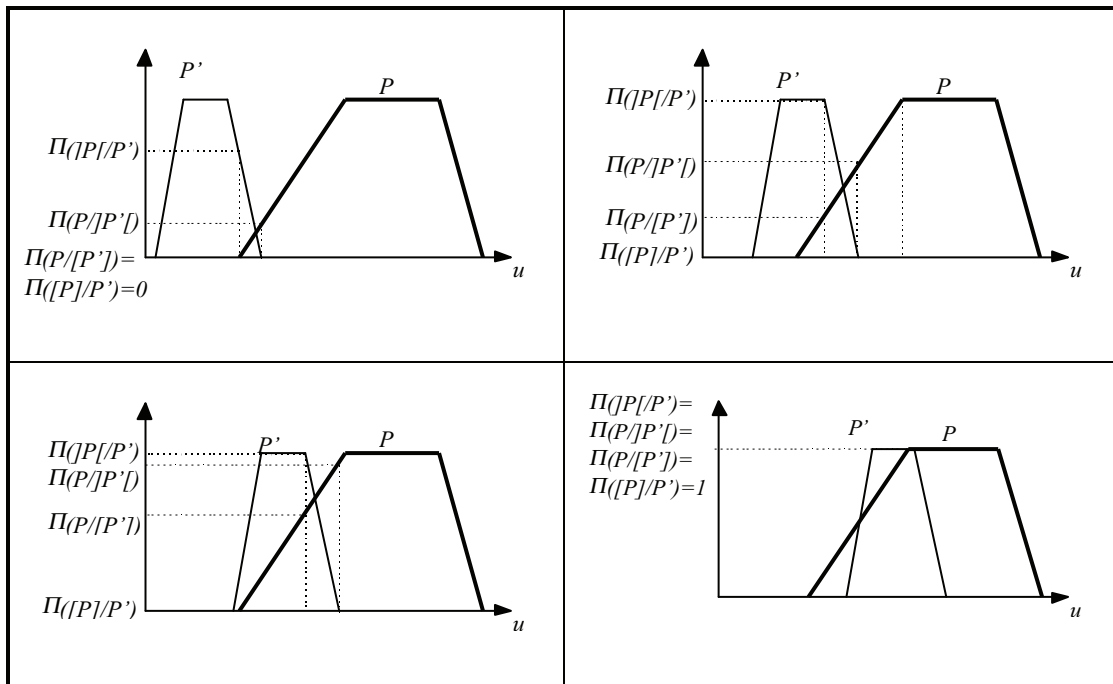
En definitiva el comportamiento de estos cuatro modelos puede considerarse basado en una medida del grado de verdad de P y en una medida de su grado de certidumbre (ambas en el dominio de lo posible), o bien en una medida conjunta como también ocurre para los otros cuatro modelos de posibilidad analizados. Las medidas del grado de verdad de P son medidas del grado de intersección de P con distintas formas (no borrosas) de considerar P' , desde $[P']$ hasta $]P[$. Las medidas de incertidumbre son por el contrario medidas de la intersección de P' con distintas formas (no borrosas) de considerar la

hipótesis P , desde $]P[$ hasta $[P]$. Nótese sin embargo que la medida $\Pi(P'/[P])$ no aparece, probablemente debido a que es demasiado restrictiva, pudiendo ser 0 a pesar de que su equivalente en el dominio de lo necesario (dado por $N(]P[/math>/ $P')$) no lo sea, lo que resulta ser una clara contradicción. La aparición de una posible certidumbre parcial de la hipótesis se debe a la incertidumbre propia de P' , al menos en la parte que interseca con $]P[$. Un ejemplo interesante consiste en comprobar que si la altura de P' es muy pequeña, lo que como se vio en el capítulo 3 indica que su posible certidumbre es muy baja, es lógico pensar que el resultado de comparar cualquier hipótesis P con P' deberá llevar asociada también una gran incertidumbre aunque el grado de verdad de la comparación pueda ser alto. Esta situación extrema queda reflejada en la siguiente figura:$



donde en efecto el posible grado de verdad de P sería 1 (ya que se calcula suponiendo posible la observación) pero su certidumbre en el dominio de lo posible es baja debido a que la observación es muy poco posible.

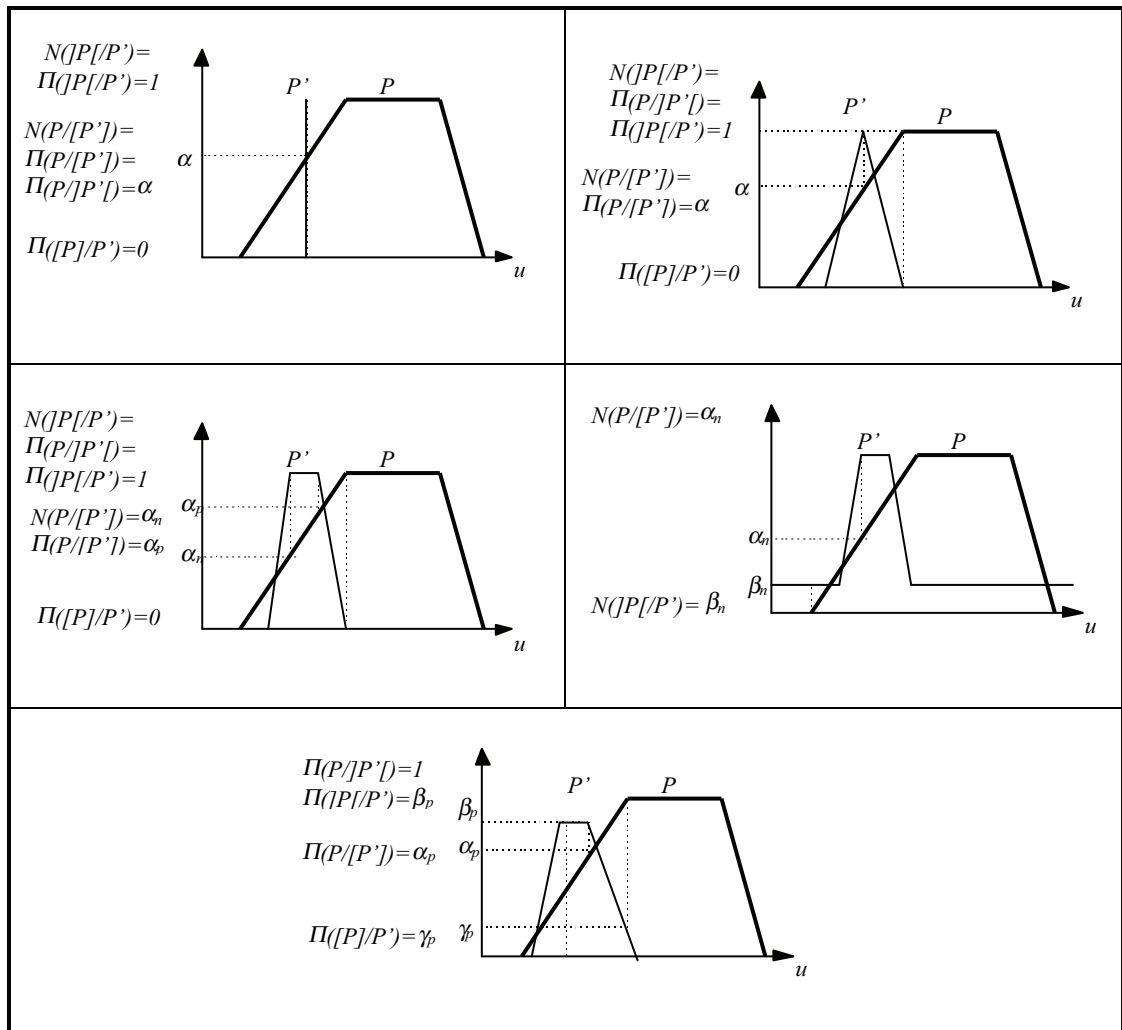
Algunas de las distintas situaciones que pueden darse entre P y P' para algunas de las medidas mencionadas quedan reflejadas en las siguientes figuras:



Las figuras anteriores evidencian las diferencias que pueden existir entre formas extremas de estimar el posible grado de verdad y el posible grado de certidumbre de una hipótesis P dada una observación P' mediante las medidas anteriormente definidas.

Una situación paradójica que puede presentarse es que mientras la estimación del grado de verdad dada por $\Pi(P/[P'])$ puede ser 0 la estimación de la certidumbre dada por $\Pi(P'/P)$ puede ser distinta de 0. Esto es debido a que son estimaciones con distintos niveles de optimismo. Parece por tanto razonable utilizar conjuntamente estimaciones con igual nivel de optimismo, si se quieren evitar estas aparentes contradicciones. A este respecto cabe considerar la pareja de medidas dada por $\Pi(P'/P)$ y $\Pi([P']/P)$. Por otro lado resulta interesante comprobar que la medida de posibilidad habitual, $\Pi_R^T(P/P')$ es una medida intermedia entre el grado de verdad y el grado de incertidumbre, de forma que utilizada para estimar simultáneamente ambos grados, parece como si el incremento en el grado de verdad se compensase por un decremento de la certidumbre. Algo similar ocurre con las medidas $\Pi_S^{T-}(P/P')$ y $\Pi_S^{T-}(P'/P)$ próximas entre sí, y que por tanto producen grados de verdad y de certidumbre similares.

Las siguientes figuras permiten comparar algunos de los grados de verdad y certidumbre definidos, tanto en el dominio de lo necesario como en el de lo posible, dependiendo de los tipos de las distribuciones involucradas.



Las figuras anteriores evidencian el hecho de que la medida de posible certidumbre dada por $\Pi([P]/P')$ es en general demasiado pesimista, pudiendo incluso ser menor que las medidas de necesidad utilizadas para estimar el grado de verdad y el grado de certidumbre, por lo que su uso parece claramente desaconsejable.

En lo que sigue, al igual que se ha hecho con los modelos de necesidad, identificaremos las principales semánticas de los modelos de posibilidad analizados hasta ahora, destacando sus principales características.

4.9.2.2 Modelos de posibilidad graduales

Estos modelos son aquellos que reducen el soporte del consecuente en base al grado de posibilidad entre la hipótesis y la observación o en base a una estimación del grado de

verdad de la hipótesis. Son graduales en el sentido en que son los menores grados de pertenencia al consecuente los que primero se ven afectados, viendo su valor reducido, por lo que no afectan significativamente a la certidumbre de la conclusión. Estos modelos son los pares (M_S^{T+}, M_S^{T+}) , y (M_R^{T+}, M_S^{T+}) y se caracterizan por la tendencia que tienen a no modificar la altura del consecuente y modificar principalmente su soporte, en función de, o bien la posibilidad habitual de P y P' , o bien el grado de verdad (optimista) de P dado P' . Las implicaciones de estos modelos son las llamadas implicaciones antigrauales en [Dubois, Prade, 93]).

En el caso del modelo (M_R^{T+}, M_S^{T+}) un grado de verdad parcial de la hipótesis medido con $\Pi(P/]P'[)$ (que es su versión más optimista) supone una reducción del soporte de Q , y la existencia de una certidumbre inferior a 1, calculada como $\Pi(]P[/P')$ (que es su versión más optimista) hace que esa certidumbre parcial se refleje también en la conclusión. De forma cualitativa se tiene:

$$\begin{aligned} Alt(Q') &\leq Alt(Q) \\]Q'[&\leq]Q[\end{aligned}$$

$$\begin{aligned}]Q'[&\text{ función de } 1-\Pi(P/]P'[) \\ Alt(Q') &\text{ función de } \Pi(]P[/P') \end{aligned}$$

Para el modelo (M_S^{T+}, M_S^{T+}) se tiene:

$$\begin{aligned} [Q'] &= [Q] \\]Q'[&<]Q[\end{aligned}$$

$$]Q'[\text{ función de } 1-\Pi_R^{T+}(P/P')$$

Como ya se ha dicho, los modelos de posibilidad graduales se caracterizan por su tendencia a excluir en el conjunto de valores posibles a aquellos elementos del universo de discurso del consecuente cuyo grado de pertenencia a este último sea menor que un valor que es función, o bien del grado de posibilidad de la hipótesis de la regla dada la observación, o bien de una estimación de su posible grado de verdad.

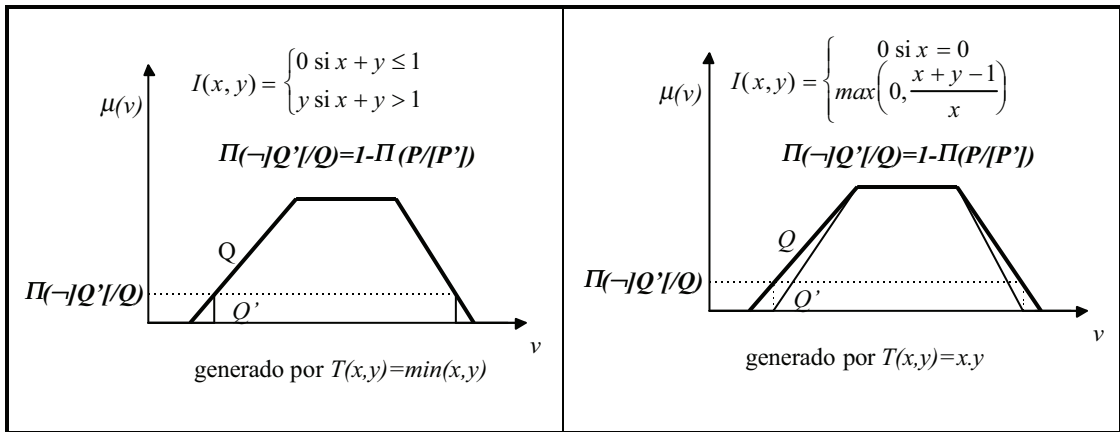
Para el modelo (M_R^{T+}, M_S^{T+}) , supuesto que deseamos comparar el comportamiento de Q y Q' como posibles observaciones para contrastar una hipótesis cualquiera G , la disminución de la altura y del soporte de Q' respecto de Q supone, en lo que respecta a la estimación del posible grado de verdad:

$$\Pi(G/]Q[) \leq \Pi(G/]Q'[) \text{ función de } \Pi(P/]P'[)$$

y a la estimación del posible grado de certidumbre:

$$\Pi(JG[Q] \leq \Pi(JG[Q']) \text{ función de } \Pi(JP[P'])$$

La siguiente figura muestra el comportamiento de los modelos de posibilidad graduales generados por las t-normas *min* y *prod*, cuando se verifica la equivalencia RCI-MS:



4.9.2.3 Modelos de posibilidad de certidumbre

Estos modelos son aquellos que reducen la altura del consecuente en función del grado de posibilidad entre la hipótesis y la observación, o bien de una medida pesimista del grado de verdad de P dado P' , esto es $\Pi_S^{T+}(P/P')$. Este tipo de modelos está constituido por las familias (M_R^{T+}, M_R^{T+}) y (M_S^{T+}, M_R^{T+}) . Las implicaciones de estos modelos se corresponden con las implicaciones con semántica de posibilidad definidas en [Dubois, Prade, 93].

El comportamiento cualitativo del modelo (M_S^{T+}, M_R^{T+}) puede resumirse en:

$$\begin{aligned} Alt(Q') &\leq Alt(Q) \\]Q'[&\leq]Q[\text{ (aunque en general muy similares)} \end{aligned}$$

$$\begin{aligned}]Q'[&\text{ función de } 1 - \Pi_S^{T+}(P'/P) \\ Alt(Q') &\text{ función de } \Pi_S^{T+}(P/P') \end{aligned}$$

y presenta un comportamiento opuesto al del modelo gradual (M_S^{T+}, M_S^{T+}) puesto que modifica el soporte en base a un grado de certidumbre $(\Pi_S^{T+}(P'/P))$ y la altura en base a un grado de verdad $(\Pi_S^{T+}(P/P'))$.

Para el modelo (M_R^{T+}, M_R^{T+}) se tiene:

$$\begin{aligned} Alt(Q') &\leq Alt(Q) \\]Q'[&=]Q[\end{aligned}$$

$$\Pi(Q'/Q) \text{ (altura de } Q' \text{ supuesto } Q \text{ normal) función de } \Pi_R^{T+}(P/P')$$

Los modelos de posibilidad de certidumbre se caracterizan por no incluir en el conjunto de valores totalmente posibles a ningún elemento del universo de discurso del consecuente cuando el grado de posibilidad de P es inferior a 1, o cuando el posible grado de verdad de P calculado de la forma más pesimista es inferior a 1. Sin embargo el modelo (M_S^{T+}, M_R^{T+}) también presenta una componente gradual, debido a que el soporte de Q' viene dado por $\{v / Q(v) \leq 1 - \Pi_S^{T+}(P'/P)\}$ siendo $\Pi_S^{T+}(P'/P)$ como ya hemos dicho una medida del grado de certidumbre de P .

Los dos modelos de certidumbre presentan ambas condiciones de activación opuestas en el sentido en que el modelo (M_R^{T+}, M_R^{T+}) tiene la condición de activación menos restrictiva mientras que (M_S^{T+}, M_R^{T+}) tiene la más restrictiva de todas.

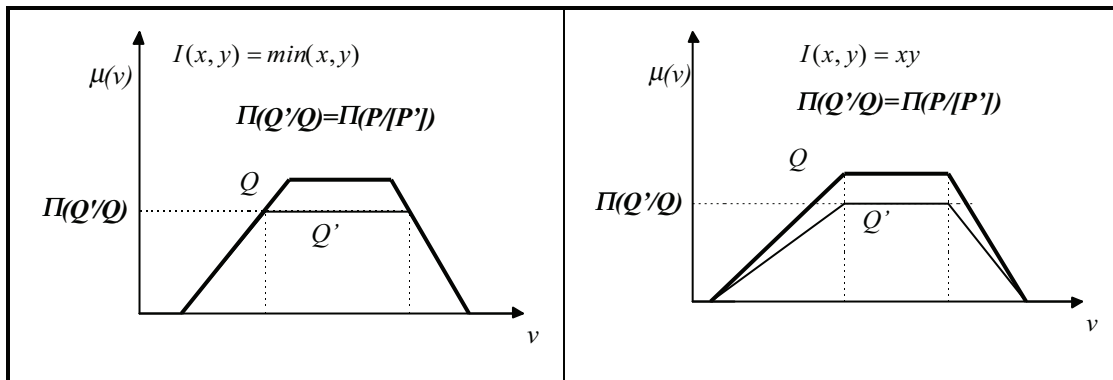
Para el modelo (M_S^{T+}, M_R^{T+}) , cuando se contrasta una hipótesis G cualquiera con Q y Q' se tiene para la estimación más optimista del grado de verdad:

$$\Pi(G/]Q[) \leq \Pi(G/]Q'[) \text{ función de } \Pi_R^{T+}(P'/P)$$

y para la estimación más optimista del grado de certidumbre:

$$\Pi(]G[/Q) \leq \Pi(]G[/Q') \text{ función de } \Pi_R^{T+}(P'/P)$$

La siguiente figura muestra el comportamiento de los modelos generados por las t-normas *min* y *prod*, cuando se verifica la equivalencia RCI-MS:



4.9.2.4 Modelos de posibilidad mixtos

Un tercer grupo de modelos de posibilidad son los que hemos denominado mixtos, que presentan un comportamiento que comparte simultáneamente características de los dos tipos de modelos de posibilidad anteriores.

Estos modelos son los pares (M_R^{T-}, M_R^{T-}) , (M_S^{T-}, M_S^{T-}) , (M_R^{T-}, M_S^{T-}) y (M_S^{T-}, M_R^{T-}) . Tienen tendencia a reducir casi por igual tanto la altura como el soporte del consecuente. Sus condiciones de activación varían según el modelo, pero son en general más restrictivas que las de los modelos graduales.

El comportamiento de estos modelos, sin entrar en detalles, queda resumido cualitativamente como sigue:

$$\begin{aligned} altura(Q') &< altura(Q) \\]Q'[&<]Q[\end{aligned}$$

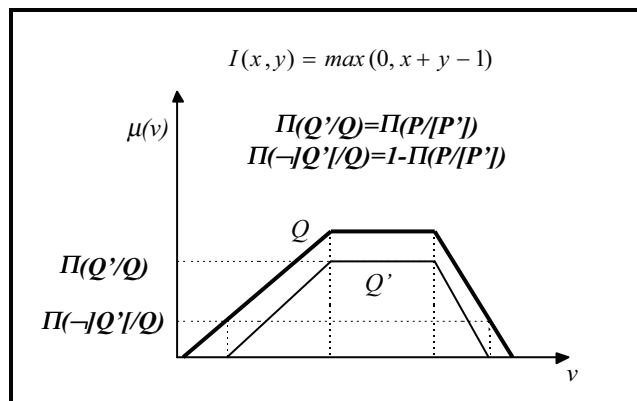
Como ya hemos dicho los modelos mixtos analizados se caracterizan por generar una conclusión cuya altura es, siempre que la activación no sea total, inferior a la altura del consecuente, y cuyo soporte es también inferior al del consecuente, y lo mismo ocurre con el núcleo. Ambas modificaciones del consecuente se basan en medidas iguales o similares, por lo que tienen lugar casi siempre de forma simultánea.

Cuando para los modelos (M_R^{T-}, M_S^{T-}) y (M_S^{T-}, M_R^{T-}) los conjuntos Q y Q' se utilizan para contrastar una hipótesis cualquiera G , se tiene:

$$\begin{aligned} \Pi(G/]Q[) &\leq \Pi(G/]Q'[) \text{ función de } \Pi_1 \\ \Pi(]G[/Q) &\leq \Pi(]G[/Q') \text{ función de } \Pi_2 \end{aligned}$$

donde Π_1 y Π_2 son medidas próximas entre sí, e intermedias entre el grado de verdad y el grado de certidumbre de la hipótesis.

Un típico ejemplo de este tipo de modelos lo constituyen los generados a partir de la t-norma de Lukasiewicz, para los que los cuatro modelos analizados se hacen iguales. La siguiente figura ilustra el comportamiento de este modelo cuando se verifica la equivalencia RCI-MS:



4.9.2.5 Modificación de las condiciones de activación

Es también posible en el caso de los modelos de posibilidad modificar las condiciones de activación para hacerlas menos restrictivas escogiendo una FGMP menor, aunque entonces no se puede en principio garantizar que la equivalencia RCI-MS siga manteniéndose sin condiciones, como sí ocurría con los n-modelos estudiados.

4.10 Medidas generalizadas de posibilidad y necesidad

4.10.1 Definiciones

Los resultados de los apartados anteriores, así como el análisis del comportamiento de las medidas de necesidad y posibilidad llevado a cabo en el capítulo 3 sugiere la conveniencia de definir unas nuevas medidas de posibilidad y necesidad más acordes con el modelo de regla utilizado. Parece razonable que las nuevas medidas sean tales que dado un determinado modelo de necesidad, cuando la medida de necesidad (escogida) de P dado P' sea no nula signifique que de P' se puede inferir algo no trivial, y análogamente debe ocurrir para los modelos de posibilidad con la medida de posibilidad que se escoja. Esto hace pensar en definir estas medidas a partir de las condiciones de activación del modelo en cuestión, de forma que se puede plantear:

$$\begin{aligned} N_{(M,I_n)}(P/P') &= 1 - \sup_u M(P'(u), I_n(P(u), 0)) \\ \Pi_{(M,I_p)}(P/P') &= \sup_u M(P'(u), I_p(P(u), 1)) \end{aligned}$$

donde I_n e I_p son respectivamente una implicación de necesidad y una implicación de posibilidad, y M una FGMP válida. Esta definición generalizada de medidas de necesidad y posibilidad permite expresar de forma compacta las condiciones de activación de los modelos (M, I_n) y (M, I_p) que las generan:

$$\begin{aligned} (M, I_n) \text{ se activa ssi } N_{(M,I_n)}(P/P') &> 0 \\ (M, I_p) \text{ se activa ssi } \Pi_{(M,I_p)}(P/P') &> 0 \end{aligned}$$

ya que en efecto $N_{(M,I_n)}(P/P')$ mide el complemento a 1 del nivel de indeterminación de la conclusión del modelo de necesidad (M, I_n) mientras $\Pi_{(M,I_p)}(P/P')$ que mide la altura de la conclusión del modelo de posibilidad (M, I_p) .

Puede comprobarse que las medidas de necesidad y posibilidad definidas hasta ahora son medidas de incertidumbre en el universo del consecuente, ya que miden el nivel de indeterminación de la conclusión de los modelos. Puede también comprobarse que verifican los axiomas de las medidas de necesidad y posibilidad. Por ejemplo para la medida de necesidad se tiene:

$$N((P_1 \cap P_2)/P') = 1 - \sup_u M(P'(u), I_n(\min(P_1(u), P_2(u)), 0))$$

y como I_n es no creciente con su primer argumento queda:

$$N((P_1 \cap P_2)/P') = 1 - \sup_u M(P'(u), \max(I_n(P_1(u), 0), I_n(P_2(u), 0)))$$

y al ser M no decreciente con ninguno de sus argumentos:

$$\begin{aligned} & 1 - \sup_u \max(M(P'(u), I_n(P_1(u), 0)), M(P'(u), I_n(P_2(u), 0))) = \\ & \min(1 - \sup_u M(P'(u), I_n(P_1(u), 0)), 1 - \sup_u M(P'(u), I_n(P_2(u), 0))) = \\ & \min(N(P_1/P'), N(P_2/P')) \end{aligned}$$

De forma análoga puede mostrarse que para la medida de posibilidad se tiene:

$$\Pi((P_1 \cup P_2)/P') = \max(\Pi(P_1/P'), \Pi(P_2/P'))$$

Nótese que en los dos casos anteriores la intersección y unión de conjuntos borrosos se han modelado con los operadores \min y \max , y que en general la elección de otros operadores hace que dejen de verificarse las propiedades anteriores.

Para la medida de necesidad $N_{(M, I_n)}$ puede comprobarse que $N_{(M, I_n)}(P/P') = 1$ no garantiza $Q' = Q$, únicamente garantiza que el nivel de indeterminación de Q' vale 0. Por ejemplo para los modelos de necesidad graduales aún cuando el núcleo de la conclusión sea mayor que el del consecuente puede tenerse $N_{(M, I_n)}(P/P') = 1$. En cierto modo una conclusión con una incertidumbre elevada es una conclusión muy próxima de la distribución de necesidad trivial y no resulta útil para ser utilizada como nueva observación para contrastar una nueva hipótesis cualquiera, mientras que una conclusión con un grado de verdad parcial puede dar lugar a un conjunto suficientemente informativo y útil de cara a nuevas inferencias. La activación total del modelo, entendida como $Q' = Q$, exige ciertamente condiciones más restrictivas, como son $N_{(M, I_n)}(P/P') = 1$ y $N_{(M, I_n)}(P/[P']) = 1$.

Por otro lado, aunque el comportamiento de algunos de los modelos de posibilidad se base principalmente en una única medida de posibilidad dada por $\Pi_R^T(P/P') = \sup_u T(P'(u), P(u))$, no debe confundirse ésta con la aquí propuesta, $\Pi_{(M, I_p)}(P/P')$ aun cuando I_p sea una t-norma T :

$$\Pi_{(M, I_p)}(P/P') = \sup_u M(T(P'(u), P(u)), 1) = M(\sup_u T(P'(u), P(u)), 1)$$

La expresión $\Pi_{(M, I_p)}(P/P')$ es nuevamente una medida de la certidumbre de la conclusión, en el ámbito de lo posible, ya que mide directamente la altura de Q' . Del mismo modo que para los modelos de necesidad, una conclusión con una elevada incertidumbre (en el dominio de lo posible) carece de valor informativo y es de escasa utilidad de cara a nuevas inferencias, independientemente de la hipótesis que se quiera contrastar, mientras que no ocurre lo mismo con los grados de verdad parciales. Tampoco es correcto suponer que $\Pi_{(M, I_p)}(P/P') = 1$ garantice $Q' = Q$ puesto que puede haberse producido una reducción del

soporte de Q sin afectar a su altura, que es lo que realmente mide $\Pi_{(M,I_p)}(P/P')=1$. En definitiva el grado de activación de una regla está siendo medido por el contenido de incertidumbre en la conclusión, y no por el grado de verdad o de certidumbre de la hipótesis, aunque dependiendo del modelo empleado, la incertidumbre de la conclusión vendrá determinada por uno, por otro, o por ambos. En este sentido las medidas de necesidad y posibilidad aquí definidas son medidas de incertidumbre en el universo del consecuente, pero pueden corresponder con medidas del grado de verdad o del grado de certidumbre en el universo del antecedente.

Puede comprobarse también que $N_{(M,I_n)}(P/P)=1$ y $\Pi_{(M,I_p)}(P/P)=1$. De hecho para los modelos basados en n-implicaciones se tiene:

$$N_{(M,I_n)}(P/P)=1-\sup_u M(P(u), I_n(P(u), 0))$$

y dado que $M(x, I_n(x, y)) \leq y$, por definición se deduce automáticamente que $N_{(M,I_n)}(P/P)=1$. Para modelos generados a partir de p-implicaciones se tiene:

$$\Pi_{(M,I_p)}(P/P)=\sup_u M(P(u), I_p(P(u), 1)) \leq 1$$

y para u tal que $P(u)=1$ se alcanza el supremo por lo que $\Pi_{(M,I_p)}(P/P)=1$. Esto resuelve uno de los problemas que plantean por ejemplo las medidas de posibilidad y necesidad que hemos denominado habituales en el capítulo 3, que no cumplen estas propiedades aparentemente intuitivas, como se señala por ejemplo en [Magrez, Smets, 89b].

4.10.2 Dualidad

Una propiedad interesante entre medidas de necesidad y posibilidad es que en general una medida de posibilidad puede ser definida a partir de una medida de necesidad (y viceversa) mediante la dualidad, esto es:

$$\Pi(P/P')=1-N(\neg P/P')$$

En general como veremos a continuación las medidas propuestas no son duales, salvo algunos casos particulares.

Supongamos en primer lugar que generamos las medidas a partir de dos modelos (M, I_n) y (M, I_p) , siendo I_n e I_p una n-implicación y una p-implicación duales, que por tanto verifican:

$$I_p(x, y)=1-I_n(x, 1-y)$$

(recordemos que son duales las implicaciones I_S^T y M_R^T y las implicaciones I_R^T y M_S^T) entonces, aplicando la dualidad tanto a I_p como a M , podemos escribir:

$$\begin{aligned}\Pi_{(M,I_p)}(P/P') &= \sup_u M(P'(u), I_p(P(u), 1)) = \\ \sup_u M(P'(u), 1 - I_n(P(u), 0)) &= \sup_u 1 - I(P'(u), I_n(P(u), 0)) = \\ &= 1 - \inf_u I(P'(u), I_n(P(u), 0))\end{aligned}$$

y por tanto a partir de la medida de posibilidad se obtiene una nueva medida de necesidad, que ya ha aparecido a lo largo de este capítulo, y que denotaremos por $N_I(P/P')$ para resaltar la implicación de la que proviene:

$$N_I(P/P') = \inf_u I(P'(u), P(u))$$

donde I es la implicación dual de M . La medida $N_I(P/P')$ es dual de la medida de posibilidad $\Pi_{(M,I_p)}(P/P')$ pero respecto a la negación definida por $\Gamma_{I_n}(x) = I_n(x, 0)$.

Del mismo modo si partimos de la medida de necesidad $N_{(M,I_n)}(P/P')$ y siguiendo un procedimiento análogo al anterior tenemos:

$$N_{(M,I_n)}(P/P') = 1 - \sup_u M(P'(u), I_n(P(u), 0)) = 1 - \Pi_M(P/P')$$

es decir:

$$\Pi_M(P/P') = \sup_u M(P'(u), P(u))$$

y recuperamos también una medida de posibilidad ya conocida, $\Pi_M(P/P')$, que es dual de $N_{(M,I_n)}(P/P')$ pero también respecto a la negación $\Gamma_{I_n}(x) = I_n(x, 0)$.

Tenemos por tanto dos parejas de medidas de necesidad y posibilidad duales respecto a la negación obtenida a partir de la n-implicación I , dual de la FGMP M , dada por $\Gamma_{I_n}(x) = I_n(x, 0)$. Para el primer caso las medidas son:

$$\begin{aligned}\Pi_M(P/P') &= \sup_u M(P'(u), P(u)) \\ N_{(M,I_n)}(P/P') &= 1 - \sup_u M(P'(u), I_n(P(u), 0))\end{aligned}$$

y para el segundo:

$$\begin{aligned}\Pi_{(M,I_p)}(P/P') &= \sup_u M(P'(u), I_p(P(u), 1)) \\ N_I(P/P') &= \inf_u I(P'(u), P(u))\end{aligned}$$

- **Modelo $(M, I_n) = (I_p, I_n)$ compuesto por una t-norma y una s-implicación**

Cuando la implicación I_n es una implicación fuerte generada a partir de la negación habitual se tiene $I_n(x, 0) = 1 - x$, y por tanto:

$$N_{(M, I_n)}(P/P') = 1 - \sup_u M(P'(u), I_n(P(u), 0)) = 1 - \sup_u M(P'(u), 1 - P(u)) = \Pi_M(\neg P/P')$$

y lógicamente las medidas $\Pi_M(P/P')$ y $N_{(M, I_n)}(P/P')$ son duales respecto a $\Gamma(x) = I_n(x, 0) = 1 - x$. Por otro lado la p-implicación I_p dual de I_n es una t-norma, por lo que $I_p(P(u), 1) = P(u)$, con lo que para la otra pareja se tiene:

$$\Pi_{(M, I_p)}(P/P') = \sup_u M(P'(u), I_p(P(u), 1)) = \sup_u M(P'(u), P(u)) = \Pi_M(P/P')$$

$$\begin{aligned} N_{(M, I_n)}(P/P') &= 1 - \sup_u M(P'(u), I_n(P(u), 0)) = \\ &= 1 - \sup_u M(P'(u), 1 - P(u)) = 1 - \inf_u (1 - I(P'(u), P(u))) = \\ &= \inf_u I(P'(u), P(u)) = N_i(P/P') \end{aligned}$$

y por tanto las dos medidas de necesidad son iguales entre sí, y lo mismo ocurre para las medidas de posibilidad:

$$\Pi_{(M, I_p)}(P/P') = \Pi_{(M)}(P/P') = 1 - N_{(M, I_n)}(\neg P/P') = 1 - N_i(\neg P/P')$$

Las cuatro medidas anteriores se reducen a dos, una de posibilidad y otra de necesidad, duales respecto a la negación habitual.

- **Modelo $(M, I_p) = (I_p, I_n)$ compuesto por una pseudo-conjunción y una r-implicación**

En este caso se tiene el valor $I_n(x, 0)$ depende de si la t-norma de la que proviene la implicación es positiva o no.

En efecto, si la **t-norma generadora del modelo es positiva**, se obtiene una negación que no es una negación fuerte, dada por:

$$I_n(P(u), 0) = \begin{cases} 1 & \text{si } P(u) = 0 \\ 0 & \text{en otro caso} \end{cases}$$

ya que al ser $I_n(x, 0) = f^{t-1}(f(0) - f(x))$ y ser $f(0)$ infinito, la expresión anterior vale $f^{t-1}(\infty) = 0$ para todo x no nulo, y $f^{t-1}(0) = 1$ cuando $x = 0$, con lo que se tiene:

$$I_n(P(u), 0) = (\neg]P[\](u)$$

y la primera pareja de medidas queda como sigue:

$$\begin{aligned}\Pi_M(P/P') &= \sup_u M(P'(u), P(u)) \\ N_{(M,In)}(P/P') &= 1 - \sup_u M(P'(u), (\neg]P[)(u)) = \inf_u I(P'(u),]P[(u)) = N_I(]P[/P') \\ \Pi_M(P/P') &= \sup_u M(P'(u), P(u)) = 1 - \inf_u I(P'(u), 1 - P(u)) = 1 - N_I(\neg P/P')\end{aligned}$$

y para la segunda:

$$\begin{aligned}\Pi_{(M,Ip)}(P/P') &= \sup_u M(P'(u),]P[(u)) = \Pi_M(]P[/P') \\ N_I(P/P') &= \inf_u I(P'(u), P(u)) \\ \Pi_{(M,Ip)}(P/P') &= \sup_u M(P'(u),]P[(u)) = 1 - \inf_u I(P'(u), 1 -]P[(u)) = 1 - N_I(\neg]P[/P')\end{aligned}$$

y las dos parejas de medidas no coinciden entre sí, ya que en efecto:

$$\begin{aligned}\Pi_M(P/P') &= 1 - N_I(\neg P/P') \\ \Pi_{(M,Ip)}(P/P') &= 1 - N_I(\neg]P[/P') \\ \Pi_{(M,Ip)}(]P[/P') &= \Pi_M(P/P') = 1 - N_{(M,In)}(\neg P/P') = 1 - N_I(\neg P/P')\end{aligned}$$

Cuando por el contrario cuando la **t-norma generadora es negativa**, la negación puede expresarse en función del generador aditivo de la t-norma, esto es:

$$\Gamma_{In}(x) = I_n(P(u), 0) = f^{f-1}(f(0) - f(P(u))) = 1 - g^{f-1}(g(1) - g(1-x))$$

que sí es una negación fuerte (en efecto $\Gamma^*(x) = g^{f-1}(g(1) - g(x))$ es una negación fuerte al ser $g(1)$ finito y por tanto $\Gamma_{In}(x) = 1 - \Gamma^*(1-x)$ también lo es). En este caso, escribiendo $I(x, 0) = \Gamma_I(x) = f^{f-1}(f(0) - f(P(u)))$, para la primera pareja de medidas se tiene:

$$\begin{aligned}\Pi_M(P/P') &= \sup_u M(P'(u), P(u)) \\ N_{(M,In)}(P/P') &= 1 - \sup_u M(P'(u), \Gamma_I(P(u))) = 1 - \Pi_M(\Gamma_I(P)/P') \\ \Pi_M(P/P') &= \sup_u M(P'(u), P(u)) = 1 - \inf_u I(P'(u), 1 - P(u)) = 1 - N_I(\neg P/P')\end{aligned}$$

y para la segunda:

$$\begin{aligned}\Pi_{(M,Ip)}(P/P') &= \sup_u M(P'(u), I_p(P(u), 1)) = \sup_u M(P'(u), 1 - I_n(P(u), 0)) = \\ &= \Pi_M(\neg \Gamma_I(P(u)) / P') \\ N_I(P/P') &= \inf_u I(P'(u), P(u))\end{aligned}$$

$$\Pi_{(M,I_p)}(P/P') = \sup_u M(P'(u), J P[(u)]) = 1 - \inf_u I(P'(u), 1 - J P[(u)]) = 1 - N_I(\neg J P[/P'])$$

y las dos parejas de medidas tampoco coinciden entre sí, ya que en efecto:

$$\begin{aligned} \Pi_M(P/P') &= 1 - N_I(\neg P/P') \\ \Pi_{(M,I_p)}(P/P') &= 1 - N_I(\neg J P[/P']) \end{aligned}$$

$$\Pi_{(M,I_p)}(J P[/P']) = \Pi_{(M)}(P/P') = 1 - N_{(M,I_n)}(\neg P/P') = 1 - N_I(\neg P/P')$$

4.11 Simplificación de la RCI para su aplicación práctica a sistemas de reglas borrosas: la MS de dos índices

Una de las aportaciones de esta tesis es la propuesta de un nuevo mecanismo de inferencia simplificado, basado en la MS pero con dos índices de similitud, y con suficiente riqueza semántica para poder reproducir los principales comportamientos semánticos detectados en los modelos analizados a lo largo de este capítulo, cuando se infiere con la RCI. Para ello se proponen dos funciones de implicación generales, una n-implicación capaz de modificar el núcleo y el nivel de indeterminación del consecuente de forma independiente, y una p-implicación, capaz de modificar la altura y el soporte del consecuente, también de forma independiente.

Se propone también una forma de escoger los índices de similitud involucrados para reproducir los principales comportamientos semánticos, tanto de necesidad como de posibilidad, detectados en los modelos analizados hasta ahora.

Por último se propone un conjunto de modelos simplificados para ser utilizados como representantes de cada una de las semánticas detectadas, de cara a simplificar la elección de un modelo dada una determinada semántica. Estos modelos serán los que se empleen en el capítulo 7 para el sistema de diagnóstico propuesto.

4.11.1 Modelo general de necesidad

Hasta ahora hemos analizado dos de los principales mecanismos de inferencia borrosa, detectando sus diferencias y similitudes de comportamiento. Cuando los modelos considerados son modelos de necesidad, la RCI es un mecanismo mucho más rico que la MS tal y como ha sido definida hasta ahora (que es como normalmente se define en los trabajos consultados), ya que el comportamiento aproximado de la RCI se basa en dos medidas simultáneas que pueden interpretarse, de forma aproximada, como el grado de verdad y el grado de certidumbre de la hipótesis.

La RCI introduce siempre en la conclusión una mínima incertidumbre dada por $\Pi(JP/P')$. Por otro lado la MS sólo es capaz de introducir incertidumbre cuando la implicación está generada por una t-norma negativa. Sin embargo la MS presenta ventajas computacionales claras, ya que no es necesario para cada valor v del universo V donde está definido Q , recorrer todos los valores u del universo U donde está definido P . Una vez evaluada la medida de similitud, que puede normalmente hacerse recorriendo una única vez el universo U , basta con modificar el consecuente de la regla, lo que se consigue también recorriendo

una única vez el universo V (en lugar de recorrer el espacio producto $U \times V$). Por ello resulta de especial interés la regla de inferencia MS frente a la RCI.

Admitiendo las simplificaciones propuestas para el cálculo del nivel de incertidumbre de la conclusión de los modelos basados en n-implicaciones, la inferencia se basa en las siguientes medidas de similitud. El grado de verdad de P dado P' ha sido definido como:

$$N(P/[P']) = \inf_{u/P'(u)=1} P(u)$$

y el grado de certidumbre de P dado P' como:

$$N(JP/[P']) = \inf_{u/P(u)=0} (1 - P'(u))$$

o bien de forma equivalente, el grado de incertidumbre:

$$\Pi(\neg JP/[P']) = \sup_{u/P(u)=0} P'(u)$$

La modificación del consecuente así como la condición de activación quedan determinadas de forma aproximada por las dos medidas anteriores, con lo que la propuesta de esta tesis consiste en diseñar una función de implicación que permita reproducir de forma aproximada los comportamientos detectados para los modelos estudiados, basándose en las citadas medidas $\Pi(\neg JP/[P'])$ y $N(P/[P'])$, y una única n-implicación general, combinada con los índices apropiados será capaz de reproducir los principales comportamientos semánticos identificados.

Definimos la nueva n-implicación como sigue:

$$I_n(\alpha_n, \beta_n, y) = \min\left(\frac{\beta_n}{\alpha_n}(y - \alpha_n) + 1, 1\right)$$

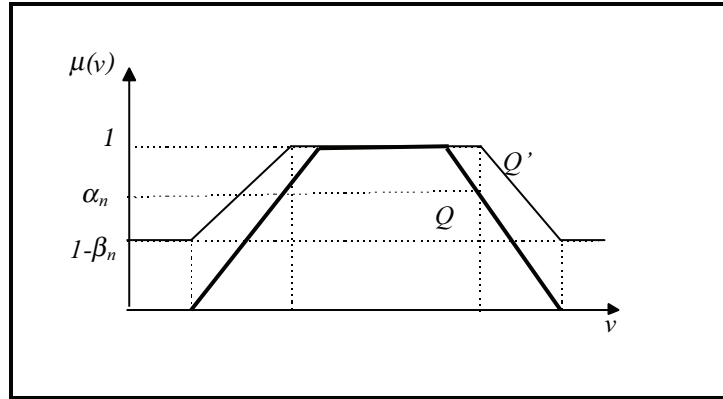
La conclusión Q' vendrá entonces dada por:

$$Q'(v) = I_n(\alpha_n, \beta_n, Q(v)) = \min\left(\frac{\beta_n}{\alpha_n}(Q(v) - \alpha_n) + 1, 1\right)$$

La modificación que se produce sobre el consecuente para la obtención de la conclusión es la siguiente. Por un lado la incertidumbre de la conclusión se obtiene para $Q(v)=0$, y resulta ser igual a $1 - \beta_n$. Por otro el núcleo viene dado por el conjunto $\{v \in V / Q'(v)=1\}$, y como puede comprobarse es igual a $\{v \in V / Q(v) > \alpha_n\}$. Nótese que la implicación de

Lukasiewicz, una de las más utilizadas, es un caso particular de la implicación propuesta correspondiente con $\alpha_n = \beta_n$.

Su comportamiento queda representado gráficamente en la siguiente figura:



Esta implicación es suficientemente general puesto que modifica de forma independiente el núcleo y el nivel de indeterminación de la conclusión, permitiendo modelar diferentes comportamientos. Además la transición del núcleo a la zona con menores grados de pertenencia es gradual, y con la misma forma que las transiciones que presenta Q desde su núcleo hasta su soporte, lo que parece intuitivamente razonable o al menos aceptable, y en general creemos que matemáticamente deseable.

Retomando los resultados simplificados que se han obtenido hasta el momento, podemos proponer qué valores deben tomar α y β de cara a poder reproducir de forma aproximada los comportamientos de los modelos de necesidad analizados, es decir (M_R^{T+}, I_R^{T+}) , (M_R^{T-}, I_R^{T-}) , (M_S^{T-}, I_S^{T-}) y (M_S^{T+}, I_S^{T+}) :

Modelo	α_n	$\beta_n = N(P/P')$	Semántica modelo
(M_R^{T+}, I_R^{T+})	$N(P/[P'])$	$N([P]/P')$	gradual
(M_R^{T-}, I_R^{T-})	$N(P/[P'])$	$1 - \max(1 - N([P]/P'), I_R^{T-}(N(P/[P']), 0))$	mixta
(M_S^{T-}, I_S^{T-})	$1 - I_R^{T-}(N(P/[P']), 0)$	$\min(N([P]/P'), N(P/[P']))$	mixta
(M_S^{T+}, I_S^{T+})	1	$N(P/[P'])$ o 0 si $N([P]/P') < 1$	certidumbre.

El grado de activación de la regla, que antes medíamos con la medida de necesidad generalizada generada por el modelo puede ser estimado con β_n . Por tanto para determinar si un modelo está activado, basta con determinar si se cumple $\beta_n > 0$, lo que garantiza que el nivel de indeterminación de la conclusión será inferior a la unidad. Resulta

también inmediato obtener un modelo de certidumbre con condición de activación similar a la del modelo gradual. Para ello basta con plantear $\beta_n = N(P/[P'])$.

El modelo general obtenido como simplificación de la RCI es un modelo del tipo MS pero basado en dos índices de similitud, uno de los cuales, β_n , determina el nivel de indeterminación de la conclusión, y el otro, α_n , determina su núcleo. Dicho de otro modo β_n es el grado de certidumbre del consecuente y α_n su grado de verdad.

Este tipo de implicación puede resultar de especial interés cuando se manejan distribuciones trapezoidales o λ -trapezoidales, ya que resulta relativamente sencillo determinar los parámetros de la distribución resultante, a partir los parámetros del resto de distribuciones involucradas, de forma que no resulte necesario discretizar los universos de discurso del antecedente o del consecuente.

4.11.2 Modelo general de posibilidad

De forma similar a como ocurre con los modelos de necesidad, las modificaciones sobre el consecuente de los principales modelos de posibilidad se pueden reducir a la modificación de su altura y de su soporte.

Los modelos (M_R^T, M_R^T) y (M_S^T, M_S^T) se basan en una única medida de similitud entre P y P' , dada por:

$$\Pi_R^T(P/P') = \sup_u T(P'(u), P(u))$$

mientras que los modelos (M_S^T, M_R^T) y (M_R^T, M_S^T) se basan en varias medidas distintas de posibilidad, que puede considerarse que estiman, o bien el posible grado de verdad de P , o bien la certidumbre de P , en función de la observación P' en ambos casos.

Del mismo modo que antes es posible definir una función de implicación general, para llevar a cabo la inferencia utilizando el mecanismo MS pero basado en dos índices de similitud distintos, de forma que escogidos apropiadamente sea posible reproducir de forma aproximada los principales comportamientos detectados.

Para ello definimos la una nueva p-implicación propuesta como sigue:

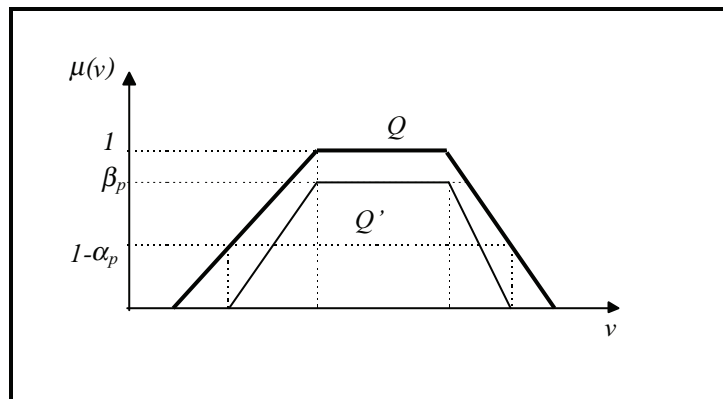
$$I_p(\alpha_p, \beta_p, y) = \max\left(\frac{\beta_p}{\alpha_p}(y-1) + \beta_p, 0\right)$$

La conclusión Q' entonces dada por:

$$Q'(v) = I_p(\alpha_p, \beta_p, Q(v)) = \max\left(\frac{\beta_p}{\alpha_p}(Q(v) - 1) + \beta_p, 0\right)$$

La modificación que se produce sobre el consecuente para la obtención de la conclusión es la siguiente. Por un lado la altura de la conclusión se obtiene para $Q(v)=1$, y resulta ser igual a $1-\beta_p$. Por otro el soporte viene dado por el conjunto $\{v \in V / Q'(v)=0\}$, y puede comprobarse que resulta ser igual a $\{v \in V / Q(v) < \alpha_p\}$. La ϵ -norma de Lukasiewicz resulta ser un caso particular de la implicación propuesta cuando se tiene $\alpha_p = \beta_p$. Es interesante notar que aunque estamos utilizando la función I_p propuesta como una implicación, estamos también definiendo una familia de conjunciones (o pseudo-conjunciones, según se mire).

Su comportamiento queda representado gráficamente en la siguiente figura:



Podemos ahora proponer valores para α_p y β_p de cara a poder reproducir de forma aproximada los comportamientos de los modelos de posibilidad analizados a lo largo de este capítulo, concretamente (M_S^{T+}, M_S^{T+}) , (M_R^{T-}, M_R^{T-}) , (M_S^{T-}, M_S^{T-}) y (M_R^{T+}, M_R^{T+}) :

Modelo	α_p	$\beta_p = \Pi(P/P')$	Semántica modelo
(M_S^{T+}, M_S^{T+})	$\Pi_R^{T+}(P/P')$	1 o 0 si $]P'[\cap]P[=0$	gradual
(M_S^{T-}, M_S^{T-})	$\Pi_R^{T-}(P/P')$	$1 - I_R^{T-}(\Pi_R^{T-}(P/P'), 0)$	mixta
(M_R^{T-}, M_R^{T-})	$1 - I_R^{T-}(\Pi_R^{T-}(P/P'), 0)$	$\Pi_R^{T-}(P/P')$	mixta
(M_R^{T+}, M_R^{T+})	1	$\Pi_R^{T+}(P/P')$	certidumbre

El grado de activación de la regla, que medíamos antes con la medida de posibilidad generalizada generada por el modelo escogido, viene ahora dado por β_p , de forma que el modelo se activa si se verifica $\beta_p > 0$. Al igual que con los modelos de necesidad, β_p puede interpretarse como el posible grado de certidumbre del consecuente, mientras que α_p como su posible grado de verdad.

Por otro lado los modelos (M_R^T, M_S^T) y (M_R^T, M_S^T) basan su comportamiento en varias medidas de posibilidad. Simplificando podemos destacar en primer lugar, dentro de aquellas que estiman el posible grado de verdad de la hipótesis, a las siguientes, enumeradas de menos a más restrictiva:

$$\begin{aligned} \Pi(P/[P']) &= \sup_{u/P'(u)>0} P(u) \\ \Pi_{(M_R, M_S)}(P/P') &= \sup M_R^T(P', M_S^T(P, 1)) \\ \Pi_S^T(P/P') &= \sup_u M_S^T(P'(u), P(u)) \end{aligned}$$

y en el caso en que se cumpla la equivalencia RCI-MS se tiene también:

$$\Pi(P/[P']) = \sup_{u/P'(u)=1} P(u)$$

y dentro de aquellas que estiman el posible grado de certidumbre de la hipótesis, de menos a más restrictiva:

$$\begin{aligned} \Pi(P'/[P]) &= \sup_{u/P(u)>0} P'(u) \\ \Pi_{(M_R, M_S)}(P'/P) &= \sup M_R^T(M_S^T(P', 1), P) \\ \Pi_S^T(P'/P) &= \sup_u M_S^T(P(u), P'(u)) \end{aligned}$$

Esto permite también simplificar los modelos de posibilidad (M_R^T, M_S^T) y (M_R^T, M_S^T) obteniéndose la siguiente tabla:

Modelo	α_p	$\beta_p = \Pi(P/P')$	Semántica modelo
(M_R^{T+}, M_S^{T+})	$\Pi(P'/[P])$	$\Pi(P/[P'])$	gradual
(M_R^T, M_S^T)	$\sup M_R^T(P, M_S^T(P', 1))$	$\sup M_R^T(P', M_S^T(P, 1))$	mixta
(M_S^T, M_R^T)	$\Pi_S^T(P'/P)$	$\Pi_S^T(P/P')$	mixta
(M_S^{T+}, M_R^{T+})	$\Pi_S^{T+}(P'/P)$	$\Pi_S^{T+}(P/P')$	certidumbre

La condición de activación viene nuevamente dada por $\beta_p > 0$. Los dos últimos modelos de la tabla no son iguales ya que las medidas $\Pi_S^{T+}(P/P)$ y $\Pi_S^{T-}(P/P)$ no son iguales al estar una generada por una t-norma positiva y la otra por una t-norma negativa.

Es interesante también resaltar que las dos implicaciones generales propuestas pueden considerarse duales, ya que verifican:

$$I_p(\alpha, \beta, y) = 1 - I_n(\alpha, \beta, 1 - y)$$

4.11.3 Representantes semánticos

4.11.3.1 Representantes semánticos de necesidad

En los modelos anteriores llama la atención el hecho de que mientras que los modelos de necesidad basan su comportamiento en dos medidas de similitud entre hipótesis y observación, los de posibilidad lo hacen, o bien en base a una única medida (aunque diferente de un modelo a otro), o bien en base a dos medidas diferentes, como ocurre con los modelos de necesidad. Dado que los modelos de necesidad presentan un comportamiento más homogéneo, y que de entre los modelos de posibilidad también algunos presentan características similares, parece sugerente la idea de simplificar aún más los modelos presentados en el apartado anterior, pero preservando sus principales características semánticas.

La simplificación de los modelos de necesidad con vistas a obtener un único representante para cada tipo semántico parece razonable, puesto que aunque las tres principales semánticas antes señaladas presentan claras diferencias, resulta complejo, dentro de un mismo tipo, diferenciar los pequeños matices que puedan presentar unos u otros modelos.

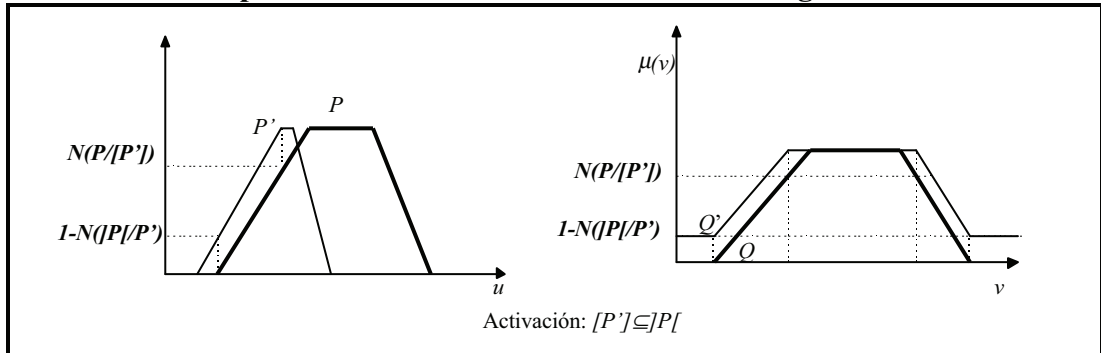
Para los modelos de necesidad proponemos un representante para cada uno de los tipos semánticos identificados, basado en la inferencia MS de dos índices, resumidos en la siguiente tabla:

α_n	$\beta_n = N(P/P')$	Semántica modelo
$N(P/[P'])$	$N(JP/[P'])$	<i>gradual</i>
$N(P/[P'])$	$\min(N(JP/[P']), N(P/[P']))$	<i>mixta</i>
1	$N(P/[P'])$ o 0 si $N(JP/[P']) < 1$	<i>certidumbre estricto</i>
1	$\min(N(JP/[P']), N(P/[P']))$	<i>certidumbre</i>

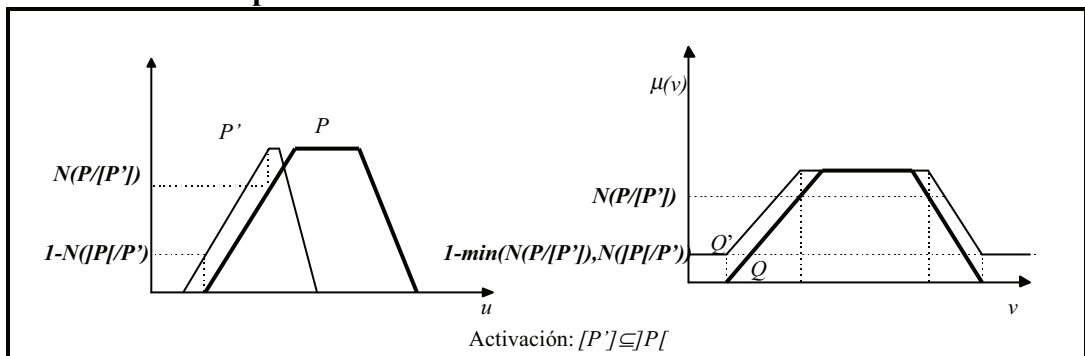
Como puede verse el valor de β_n determina las condiciones de activación del modelo. Al modelo que hasta ahora hemos denominado modelo de necesidad de certidumbre le hemos añadido el calificativo de estricto debido a sus condiciones de activación, y hemos añadido también un representante de certidumbre pero con condiciones de activación

como las de los modelos graduales. El comportamiento de cada uno de los modelos puede verse en las siguientes figuras:

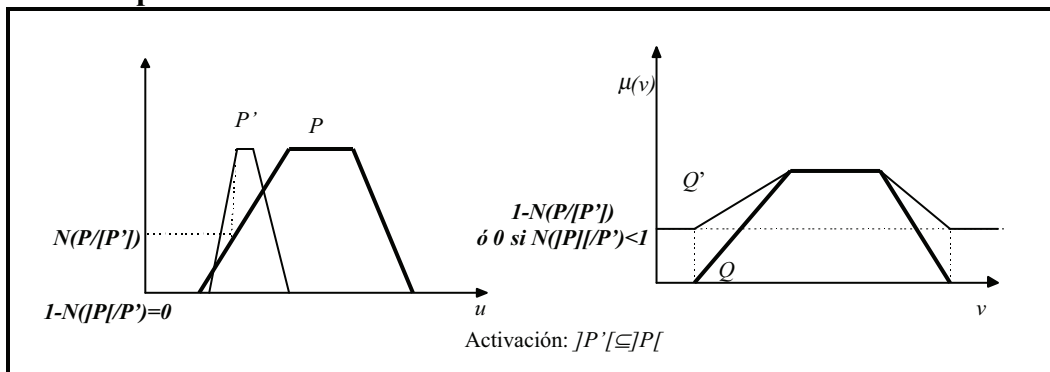
Representante de la semántica de necesidad gradual



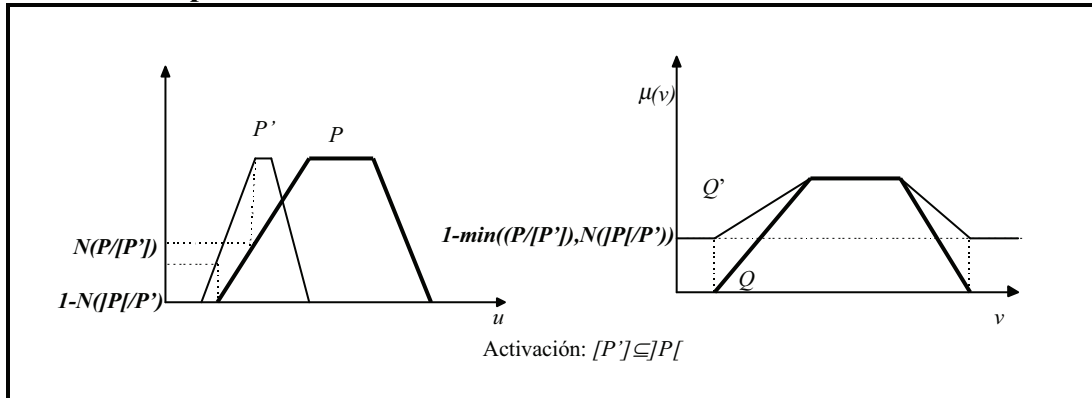
Representante de la semántica de necesidad mixta



Representante de la semántica de necesidad de certidumbre estricta



Representante de la semántica de necesidad de certidumbre



Puede comprobarse por ejemplo que en el caso de los modelos mixtos todos ellos han sido condensados en uno único, aproximadamente equivalente al modelo basado en la implicación de Lukasiewicz utilizando la RCI.

4.11.3.2 Representantes semánticos de posibilidad

Para los modelos de posibilidad, una primera alternativa que cabe plantear consiste sencillamente en simplificar el primer grupo de modelos analizados, lo que da lugar a la siguiente tabla:

α_p	$\beta_p = \Pi(P/P')$	Semántica modelo
$\Pi(P/P')$	1 o 0 si $[P' \cap] P] = 0$	gradual
$\Pi(P/P')$	$\Pi(P/P')$	mixta
1	$\Pi(P/P')$	certidumbre

donde $\Pi(P/P')$ podría ser la medida de posibilidad habitual basada en la t-norma *min*. Sin embargo y como ya hemos señalado estos modelos no presentan la misma riqueza semántica que los otros modelos de posibilidad analizados, o que los modelos de necesidad anteriores.

Por otro lado, teniendo en cuenta las consideraciones hechas en el apartado 4.9.2.1, podemos, de cara a simplificar la formulación final, aproximar las medidas $\Pi_S^{T+}(P'/P)$ y $\Pi_S^{T+}(P/P')$ del siguiente modo:

$$\Pi_S^{T+}(P'/P) \approx \Pi([P' \cap] P] \text{ si } \exists u / P(u) + P'(u) > 1, 0 \text{ en otro caso}$$

y para la otra medida podemos plantear:

$$\Pi_S^{T+}(P/P') \approx \Pi(JP[/P') \text{ si } \exists u/ P(u)+P'(u) > 1, 0 \text{ en otro caso}$$

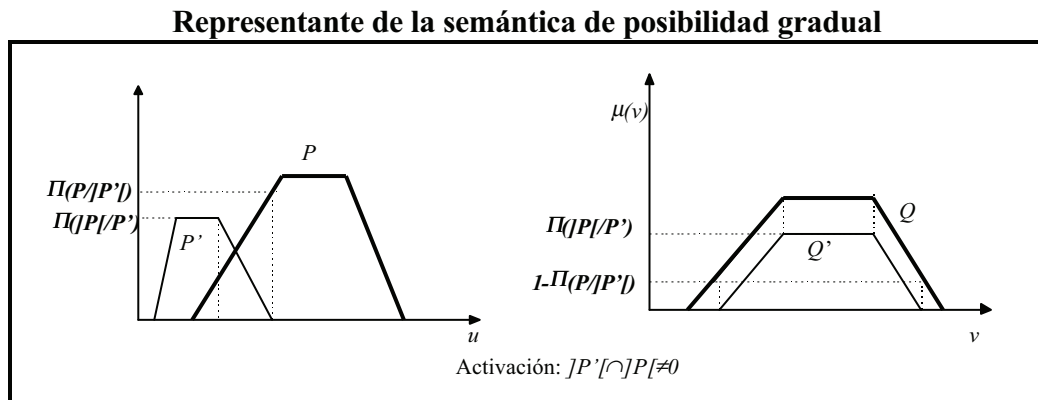
medidas que denominaremos respectivamente $\Pi_r(JP'[/P)$ y $\Pi_r(JP[/P')$ por su carácter más restrictivo, dado por la condición $P(u)+P'(u) > 1$.

Además y dado que las medidas $\sup_u M_R^{T-}(P'(u), M_S^{T-}(P(u), 1))$ y $\sup_u M_R^{T-}(P(u), M_S^{T-}(P'(u), 1))$ son muy próximas entre sí, y lo mismo ocurre con las medidas $\Pi_S^{T-}(P'/P)$ y $\Pi_S^{T-}(P/P')$, parece razonable suponerlas aproximadamente iguales e iguales a la medida de posibilidad basada en la t-norma de Lukasiewicz, que es intermedia a todas ellas, de cara a obtener un único modelo.

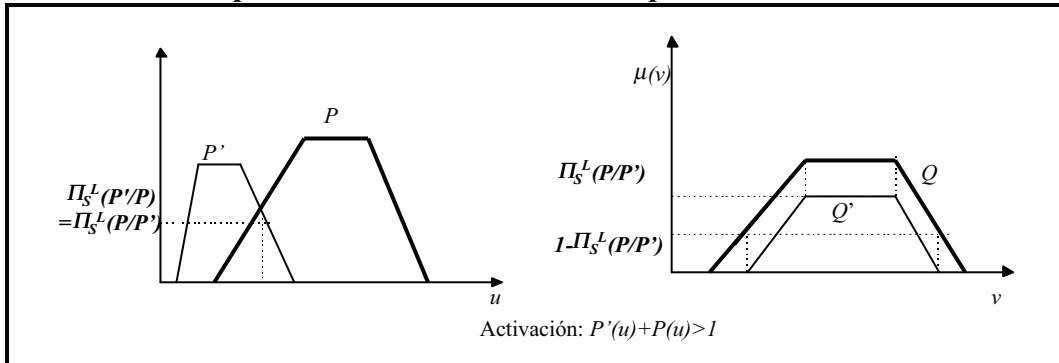
Con estas simplificaciones llegamos a la siguiente tabla de representantes semánticos.

α_p	$\beta_p = \Pi(P/P')$	Semántica modelo
$\Pi(JP'[/P)$	$\Pi(JP[/P')$	gradual
$\Pi_S^L(P'/P) = \Pi_S^L(P/P')$	$\Pi_S^L(P/P') = \Pi_S^L(P'/P)$	mixta
$\Pi_r(JP[/P')$	$\Pi_r(JP'[/P)$	certidumbre

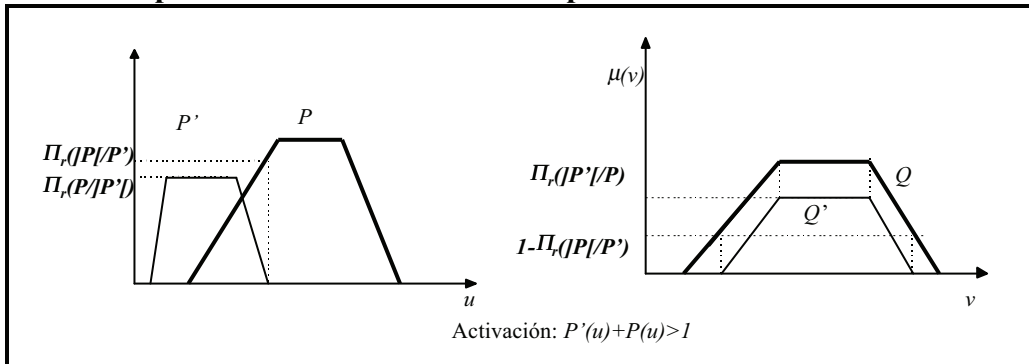
El comportamiento de estos modelos queda reflejado de forma aproximada en la siguientes figuras:



Representante de la semántica de posibilidad mixta



Representante de la semántica de posibilidad de certidumbre



Cabe también la posibilidad de obtener representantes semánticos para los modelos mixtos y de certidumbre con condiciones de activación no restrictivas. En efecto podemos proponer por ejemplo:

α_p	$\beta_p = \Pi(P/P')$	Semántica modelo
$\Pi_R^{min}(P'/P) = \Pi_R^{min}(P/P')$	$\Pi_R^{min}(P'/P) = \Pi_R^{min}(P/P')$	<i>mixta no estricto</i>
$\Pi(JP'/P')$	$\Pi(JP'/P)$	<i>certidumbre no estricto</i>

Una última forma de obtener representantes semánticos consiste en obtener modelos de posibilidad con la misma estructura que los representantes semánticos de necesidad obtenidos en el apartado anterior, a partir de medidas de grado de verdad y grado de certidumbre duales. Estos serán los modelos que finalmente se utilicen en el capítulo 7. Quedan resumidos en la siguiente tabla:

α_p	$\beta_p = \Pi(P/P')$	Semántica modelo
$\Pi(JP'/P)$	$\Pi(JP'/P')$	<i>gradual</i>
$\Pi(JP'/P)$	$\min(\Pi(JP'/P'), \Pi(JP'/P))$	<i>mixta</i>
1	$\Pi(JP'/P)$ o 0 si $P(u)+P'(u) < 1 \forall u$	<i>certidumbre estricto</i>
1	$\min(\Pi(JP'/P'), \Pi(JP'/P))$	<i>certidumbre</i>

La condición de activación más restrictiva, obtenida para el modelo de necesidad de certidumbre, se ha transformado en la condición:

$$P(u)+P'(u) < 1 \forall u$$

que se traduce en la exigencia de un mínimo nivel de intersección para la activación del modelo.

Por último, creemos importante señalar que en reglas donde la interpretación semántica sea importante, como por ejemplo en sistemas de diagnóstico basados en sistemas expertos borrosos, como se mostrará en el capítulo 7 de esta tesis, resulta preferible utilizar algunos de los representantes semánticos propuesto, más ricos en matices de cara a modelar de forma fiel el conocimiento disponible, y de cara a una mejor generación e interpretación de las conclusiones que de cada regla puedan derivarse. Sin embargo en aquellas aplicaciones de marcado carácter numérico se hace a priori necesario analizar con cuidado el comportamiento de unos y otros para poder determinar el modelo numéricamente más apropiado, y es de esperar que en aplicaciones donde la completitud de las reglas disponibles sea determinante (véase el apartado 5.6.3 del capítulo 5) los modelos de posibilidad con condiciones de activación menos restrictivas resulten especialmente atractivos.

5. Sistemas de reglas borrosas

En este capítulo se estudia el funcionamiento de los sistemas de reglas borrosas (sistemas en los que coexisten varias reglas simultáneamente) y se plantean y proponen soluciones a algunos de los principales problemas que surgen durante su diseño. Así por ejemplo, es necesario determinar cómo deben dispararse reglas con modelos de necesidad y reglas con modelos de posibilidad cuando se dispone tanto de distribuciones de posible como de necesaria posibilidad. Por otro lado la existencia de varias fuentes de información para una misma variable requiere la determinación de cómo éstas deben combinarse. Otros problemas surgen de la existencia de reglas paralelas (mismas variables antecedentes y consecuentes) donde se estudia cómo deben ser las particiones de los universos de discurso de las variables, y cómo debe realizarse la inferencia, y de la existencia de reglas susceptibles de encadenarse, donde debe estudiarse el orden de activación de las reglas y su posible encadenamiento. Este capítulo es por tanto la continuación natural del capítulo dedicado a la inferencia, y en él se aborda la problemática de los sistemas de reglas borrosas desde una perspectiva más general.

5.1 Organización del capítulo

Uno de los primeros temas que se abordan en este capítulo es la determinación de cómo deben coexistir los distintos tipos de modelos, de necesidad y posibilidad, y los distintos tipos de distribuciones de posibilidad, de posible y necesaria posibilidad, en un único entorno común, siendo éste uno de los primeros aspectos que inicialmente se plantea en un sistema de reglas modeladas con los modelos estudiados y propuestos en el capítulo anterior. La distinción entre estos dos tipos de distribuciones y de razonamientos (generados por cada tipo de modelo empleado), aunque implícita en algunos trabajos ([Dubois, Prade, 93], [Yager, 93], [Yager, 96]) no se ha llegado a hacer formalmente. En esta tesis se han formalizado y explicitado sus reglas de inferencia (a largo de los capítulos 3 y 4) y en este capítulo se estudia su integración en un único motor de inferencias, siendo una de las aportaciones de esta tesis.

En el siguiente punto se aborda la determinación de cómo deben agregarse varias estimaciones para el valor de una variable, estimaciones procedentes de distintas fuentes de información, cuando se modela simultáneamente la vaguedad, la imprecisión y la incertidumbre. Se estudia el problema de la interactividad entre variables, y se determina, al menos de forma cualitativa, cuándo debe escogerse una u otra t -norma (o t -conorma), según la información adicional de que se disponga y en base a la interpretación propuesta para los distintos tipos de distribuciones de posibilidad.

En el siguiente apartado se trata la multiplicidad de antecedentes en reglas borrosas, con vistas a determinar el operador de agregación a emplear. A este respecto, al igual que para la agregación de observaciones, existen pocos trabajos que profundicen en las diferencias semánticas entre unos y otros operadores, aunque sí coinciden en señalar la conveniencia de elegir el operador apropiado. Utilizando los resultados obtenidos para la agregación de observaciones, propondremos algunas reglas cualitativas que orienten su elección.

Otro problema más extensamente analizado, principalmente en el ámbito del control, es la coexistencia de reglas paralelas (con mismas variables antecedentes y consecuentes) especialmente en lo que respecta a su completitud de cara a garantizar como mínimo el disparo de una de las reglas del sistema y a la verificación del modus ponens. A este respecto puede consultarse por ejemplo [Turksen, Tian, 93b], [Dubois, Prade, 91], [Martin-Clouaire, 89], etc. El enfoque aquí propuesto se basa en el uso de las condiciones de activación de los modelos de reglas, obtenidas en el capítulo anterior, para determinar como debe ser una partición para un sistema de reglas. Veremos que la combinación de propiedades como la completitud y la verificación del modus ponens

estricto permiten determinar configuraciones de particiones óptimas según el modelo de regla escogido.

A continuación se analiza la posible automatización del razonamiento posibilístico, cuando las reglas con modelos de necesidad no son capaces de deducir conclusiones distintas de las triviales, lo que exige la determinación de modelos de posibilidad alternativos a los de necesidad, pero con similares características semánticas. En trabajos como [*Dubois, Prade, 93*] o [*Dubois, Prade, 96*] se estudian y se clasifican los principales tipos de implicaciones analizados en esta tesis, pero no se establece el paralelismo que aquí se propone entre modelos de posibilidad y necesidad. Basándonos en algunas propiedades intuitivamente deseables y analizando el comportamiento de reglas paralelas con distintos modelos determinaremos las parejas de modelos que en el dominio de lo necesario y en el de lo posible presentan comportamientos semánticamente semejantes.

Por último señalaremos brevemente, con un objetivo principalmente práctico, el orden en que deben activarse las reglas de un sistema de reglas para no incurrir en contradicciones, y se propondrá un sencillo algoritmo que permita su ordenación.

5.2 Tipos de modelos de reglas y tipos de distribuciones de posibilidad

Un punto importante en los sistemas borrosos donde coexisten reglas con modelos de necesidad y posibilidad, conjuntamente con distribuciones de posible y necesaria posibilidad, radica en la determinación de cómo todos estos elementos interactúan entre sí para poder especificar con qué tipo de distribuciones debe ser disparado cada tipo de modelo de regla.

5.2.1 Razonamiento con una regla

Nuestro punto de partida es la regla:

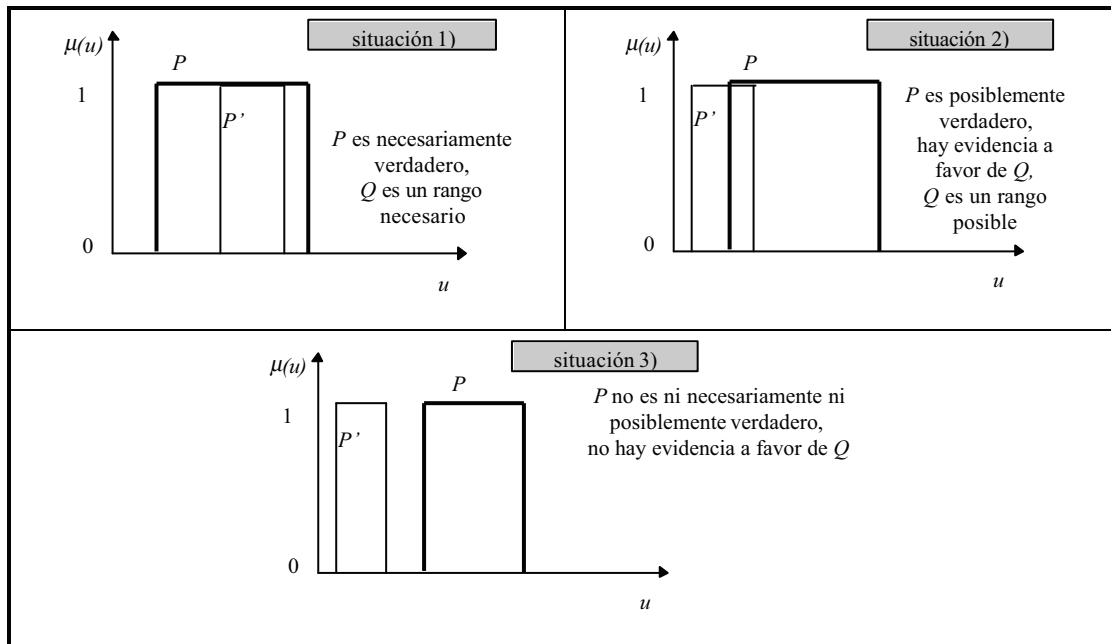
$$P \rightarrow Q$$

donde P es una distribución de posibilidad para la variable u definida en U y Q es una distribución de posibilidad para la variable v definida en V . Por el momento supondremos que tanto P como Q son conjuntos clásicos. Diremos además que una regla es una regla de necesidad cuando el modelo empleado para inferir sea un modelo de necesidad, y una regla de posibilidad cuando el modelo sea un modelo de posibilidad.

Dada una observación P' , también modelada con un conjunto clásico, y en base a lo expuesto en el capítulo 4, la inferencia basada en modelos de necesidad y posibilidad permite inferir las siguientes conclusiones (donde el símbolo \Rightarrow significa “de donde se infiere”):

- situación 1) $P' \subset P \Rightarrow Q$ rango necesario, Q rango posible
- situación 2) $P' \cap P \neq \emptyset \Rightarrow Q$ rango posible (V rango necesario)
(aunque $P' \not\subset P$)
- situación 3) $P' \cap P = \emptyset \Rightarrow (\emptyset$ rango posible, V rango necesario)

Gráficamente las tres situaciones quedan representadas en la siguiente figura:



Recordemos brevemente las diferencias entre distribuciones de necesaria posibilidad (o rangos necesarios) y distribuciones de posible posibilidad (o rangos posibles).

Una distribución de necesaria posibilidad para una variable se caracteriza por excluir sólo a valores que es imposible que la variable pueda tomar, en base a la existencia de evidencia segura en su contra, y puede interpretarse como una cota superior de posibilidad para cada uno de los elementos del universo sobre el que esté definida, en el sentido en que proporciona grados de posibilidad máximos. Así en 2) y en 3) no es posible obtener un rango necesario para Q distinto del trivial, debido a la ausencia de evidencia segura en su contra. Los modelos de necesidad concluyen distribuciones de necesaria posibilidad basándose, como era de esperar, en la inclusión de la observación en la hipótesis.

Una distribución de posible posibilidad se caracteriza porque sólo incluye a valores considerados posibles para la variable, en base a la existencia de algún tipo de evidencia (en general no segura) a su favor. Los grados de posibilidad se interpretan como posibles grados de posibilidad, y por tanto no pueden considerarse como grados máximos o mínimos, ya que son susceptibles de aumentar ante un nuevo aporte de información (nuevas distribuciones de posible posibilidad), o susceptibles de disminuir si hay algún aporte de información fiable (distribuciones de necesaria posibilidad que limiten los valores máximos de posibilidad). Así en 1) y 2) la intersección entre P' y P aporta evidencia a favor de Q y lo convierte en rango posible. Los modelos de posibilidad permiten concluir distribuciones de posible posibilidad cuando, como era de esperar, la observación y la

hipótesis tienen una intersección no nula. Por eso en 2) se infiere Q mientras que en 3) no se infiere nada.

Nótese que en 2) la existencia de otra regla de necesidad podría hacer que Q fuese un rango necesario, mientras que en 3) podría hacer que fuese un rango tanto posible como necesario, dependiendo de la regla y de P y P' . Dado que las conclusiones de las reglas con modelos de necesidad se agregan con operadores conjuntivos, la acción de una de ellas cuando su modelo no se activa no es en ningún caso concluyente. Análogamente ocurre con reglas con modelos de posibilidad. Sin embargo la existencia de evidencia en contra de ciertos valores de la variable consecuente en distribuciones de necesaria posibilidad, o la existencia de evidencia a favor de ciertos valores de la misma en distribuciones de posible posibilidad, sí son concluyentes, ya que permanecen a pesar de la activación de otras reglas de su mismo tipo.

El razonamiento basado en modelos de posibilidad es un razonamiento esencialmente inexacto, en comparación con el basado en modelos de necesidad, por la forma misma de interpretar las conclusiones que de ellos se derivan. Este tipo de razonamiento expresa claramente la existencia de preferencias que podríamos caracterizar de intuitivas, razonables o incluso subjetivas entre conclusiones formalmente posibles, pero con distintos grados de evidencia a favor de unas y otras. Así son conclusiones preferibles aquellas que se derivan de reglas para las que las hipótesis son posibles, frente a aquellas para las que las hipótesis no lo son. Esto hace que sea inexacto pero acorde con la intuición. Es habitual oír expresiones del tipo “parece que Q es posible aunque no estoy seguro”, lo que se puede modelar convenientemente con modelos de posibilidad. El utilizar conjuntamente modelos de necesidad y posibilidad permite flexibilizar el razonamiento, aun cuando se manejen conjuntos clásicos. Así por ejemplo la inclusión de P' en P es un situación más próxima a una intersección no nula entre P y P' que a una intersección nula. Del mismo modo la conclusión pasa de ser “ Q es un rango necesario” a ser “ Q es un rango posible” y finalmente a la conclusión vacía o trivial. De este modo parece intuitivamente razonable e incluso deseable de cara a flexibilizar el razonamiento, diferenciar, como se está haciendo, las situaciones 2) y 3), que en los sistemas tradicionales han sido normalmente confundidas.

Cuando las distribuciones son en general conjuntos borrosos, la aplicación de un modelo de necesidad y otro de posibilidad permite obtener entre otras las siguientes conclusiones:

- situación 1) $P' \subset P \neq 0 \quad \Rightarrow \quad Q'_n$ rango necesario, $Q'_p = Q$ rango posible
- situación 2) $P' \cap P \neq 0 \quad \Rightarrow \quad Q'_p$ rango posible (V rango necesario)
- situación 3) $P' \cap P = 0 \quad \Rightarrow \quad (\emptyset$ rango posible, V rango necesario)

donde las distribuciones Q'_p y Q'_n son las distribuciones de posible y necesaria posibilidad concluidas por el modelo de posibilidad y el de necesidad respectivamente. Nótese también que en 1) la condición $P' \subset P \neq \emptyset$ debe interpretarse como que alguna medida de la inclusión de P' en P es no nula, medida que en general depende del tipo de modelo de necesidad empleado.

Es necesario además tener en cuenta un caso adicional, debido a que en general para proposiciones borrosas $N(P/P') > 0$ no implica $\Pi(P/P')=1$. Esto da lugar por tanto a una situación intermedia entre 1) y 2) la situación 1-2) dada por:

- situación 1-2) $P' \subset P \neq \emptyset, P' \cap P \neq \emptyset \Rightarrow Q'_n$ rango necesario, Q'_p rango posible

Por otro lado cuanto mayor sea el grado de inclusión de P' en P mayor será lógicamente el grado de intersección, por lo que más parecidas serán las distribuciones Q'_n y Q'_p . Nótese que en cualquier caso siempre se cumple:

$$Q'_n \geq Q \geq Q'_p$$

5.2.2 Razonamiento con reglas encadenadas

Supongamos que tenemos dos reglas, de forma que la conclusión de la primera puede ser utilizada como observación o hecho para la segunda:

$$\begin{array}{l} A \rightarrow B \\ B^* \rightarrow C \end{array}$$

donde A es una distribución de posibilidad para la variable u definida en U , B y B^* son dos distribuciones de posibilidad para la variable v definida en V , y C es una distribución de posibilidad para la variable w definida en W . (Hemos cambiado la notación para evitar el uso de la letra R que hemos preferido destinar a reglas o relaciones borrosas). Supongamos además que todas las distribuciones son conjuntos clásicos y que disponemos de una observación A' . La aplicación de modelos de posibilidad y necesidad para ambas reglas permite inferir lo siguiente:

- situación 1) $A' \subset A \Rightarrow B$ rango necesario
 - $B \subset B^* \Rightarrow C$ rango necesario
 - $B \cap B^* \neq \emptyset \Rightarrow C$ rango posible, (W rango necesario)
 - $B \cap B^* = \emptyset \Rightarrow (\emptyset)$ rango posible, (W rango necesario)

- situación 2) $A' \cap A \neq 0 \Rightarrow B$ rango posible, (V rango necesario)
 $B \cap B^* \neq 0 \Rightarrow C$ rango posible, (W rango necesario)
 $B \cap B^* = 0 \Rightarrow (\emptyset)$ rango posible, (W rango necesario)
- situación 3) $A' \cap A = 0 \Rightarrow (\emptyset)$ rango posible, (V rango necesario)
independientemente
de B y $B' \Rightarrow (\emptyset)$ rango posible, (W rango necesario)

Una de las preguntas que cabe plantearse en este caso, a diferencia de lo que ocurría con una única regla, es si de alguna forma es válido utilizar distribuciones de necesaria posibilidad para activar modelos de posibilidad, o distribuciones de posible posibilidad para activar modelos de necesidad. Así por ejemplo, ¿sería válido utilizar como hecho “ V rango necesario” en 2) y 3), para activar el modelo de posibilidad? Nótese que esto permitiría inferir la conclusión “ C rango posible” en 2) y en 3).

El razonamiento con modelos de necesidad es restrictivo y requiere condiciones de activación severas para garantizar la fiabilidad de las conclusiones (o lo que es lo mismo, la obtención de conclusiones con grados de certidumbre cualitativamente altos), condiciones que se basan en la inclusión de la observación en la hipótesis. Sería incurrir en una contradicción el pretender concluir de una distribución de posible posibilidad, que es una información poco fiable, una distribución de necesaria posibilidad, que es una información fiable, y que podría estar excluyendo definitivamente valores entre los que podría estar el valor real de la variable consecuyente. No es por tanto correcto activar modelos de necesidad con distribuciones de posible posibilidad. Como ejemplo véase que su uso en 2) permitiría concluir “ C rango necesario”, que equivale a afirmar que todos los valores de $\neg C$ son imposibles para la variable w , cuando en realidad sólo se podría concluir que los valores de $\neg C$ podrían ser imposibles. Hacer afirmaciones rotundas en base a hipótesis únicamente posibles, y por tanto con escasa fiabilidad o certidumbre, es una evidente contradicción.

Sin embargo afirmar que un determinado conjunto es un rango posible para una variable es siempre poco comprometido puesto que la variable puede seguir tomando su valor real fuera del mismo, ya que ningún valor está siendo definitivamente excluido (o incluso incluido). La mayor contradicción en la que se puede incurrir es considerar posibles a valores para los que se tiene la certeza de su imposibilidad. Esto no es más que la restricción fundamental que las distribuciones de posible posibilidad deben cumplir (véase el apartado 3.6.2), y que se traduce en que una distribución de posible posibilidad no debe ser nunca mayor que una de necesaria posibilidad para la misma variable.

Para mantener la coherencia con el criterio escogido, basado en la intersección no nula entre observación e hipótesis para medir la evidencia en favor de la conclusión de una

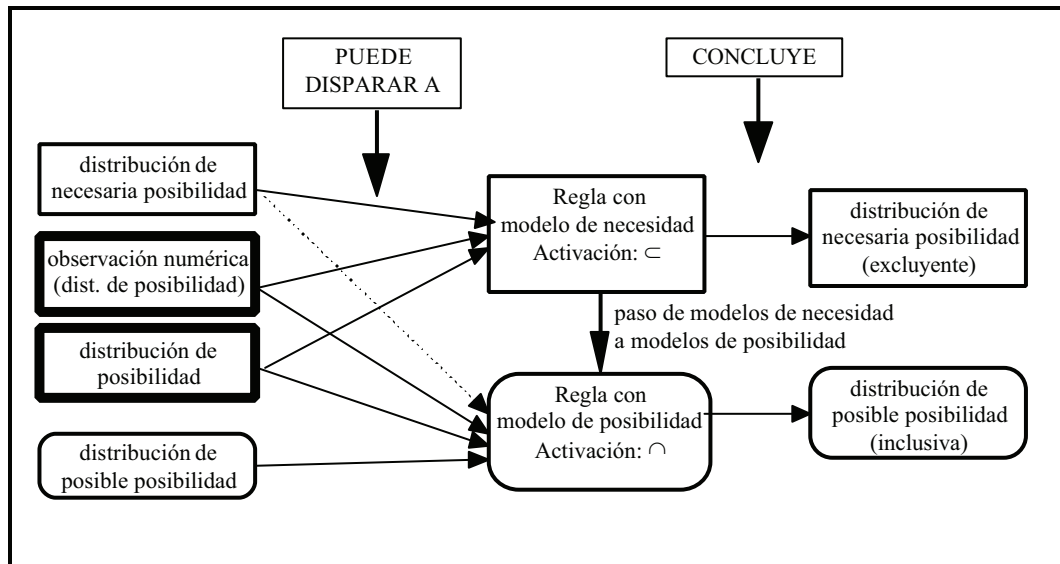
regla, parece que se impone la necesidad de distinguir entre el primer caso de 2) por un lado, y el segundo caso de 2) y la situación 3) por otro. Así por ejemplo el uso de la conclusión “ V rango necesario” para activar el modelo de posibilidad haría que tanto en 2) como en 3) se concluyese “ B rango posible”, que en efecto es una conclusión posible, pero para la que no existe ninguna evidencia a favor tanto en el segundo caso de 2) como en 3). Por eso el criterio escogido en esta tesis es activar los modelos de posibilidad con distribuciones de posible posibilidad. Sin embargo la no existencia de contradicciones en el uso de distribuciones de necesaria posibilidad con modelos de posibilidad puede en algún caso excepcional sugerir el uso de estas últimas, con vistas a obtener conclusiones allí donde no las haya, a costa de reducir aún más su fiabilidad.

Nótese la interesante alternativa de utilizar otras medidas para cuantificar la evidencia a favor de una conclusión, en base al parecido entre hipótesis y observación. Una fiabilidad intermedia entre la ofrecida por una distribución de necesaria posibilidad y una de posible posibilidad vendría generada por un condicional que basase su activación en una medida intermedia entre la inclusión y la intersección, lo que daría lugar a nuevos modelos de reglas, distintos de los ya estudiados de posibilidad y necesidad, y cuyas conclusiones deberían también interpretarse de distinto modo. En este sentido, y como ya se ha comentado con anterioridad, en el apéndice de la tesis se propone como una posible alternativa las medidas que se derivan de la teoría de los ϕ -condicionales, intermedias entre las medidas de inclusión y las de intersección vistas hasta ahora. Otras alternativas muy sugerentes son las basadas en criterios probabilísticos como por ejemplo la proporción del área de intersección entre observación e hipótesis, respecto al área total de la observación. En este caso las distribuciones podrían interpretarse como distribuciones de probable posibilidad. Nótese que en cualquier caso todos estos modelos alternativos son formas de modelar razonamiento de sentido común, en general no monótono (*Trillas, 92b*), que puede caracterizarse por su capacidad de inferir conclusiones inexactas pero razonables (posibles/probables/verosímiles/etc.) en presencia de información incompleta o de un conocimiento parcialmente contradictorio.

Cuando las distribuciones son en general borrosas aparecen situaciones adicionales a las ya planteadas, como ocurría en el caso de una única regla. Sin embargo las mismas razones de coherencia en los criterios de activación de los modelos de posibilidad, exigidas para el caso de que las distribuciones eran conjuntos clásicos, son también aplicables ahora.

Por tanto nuestra elección consistirá en activar siempre modelos de necesidad con distribuciones de necesaria posibilidad, y modelos de posibilidad con distribuciones de posible posibilidad, teniendo en cuenta que en esta última situación puede ser en algún caso interesante, tal y como se comenta en la siguiente sección, disparar modelos de posibilidad con distribuciones de necesidad (que también es una distribución de posible posibilidad válida). Nótese que cuando se dice que la variable u toma el valor $\{u_0\}$ se está

especificando la distribución de posibilidad más informativa (de posible y necesaria posibilidad), ya que el único valor incluido es el valor real de la variable, y cualquier otro valor de su universo queda definitivamente excluido. La siguiente figura resume lo expuesto en este apartado.



5.2.3 Razonamiento con información incompleta

En ciertas situaciones puede resultar imposible o muy complicado conocer el valor de una variable, o poder precisar una distribución de posibilidad que restrinja sus posibles valores. Esto puede ocurrir por ejemplo cuando la medición de la variable sea costosa o compleja y no se lleve a cabo en continuo, cuando existan anomalías en su correspondiente sensor, etc.

En estas situaciones pueden adoptarse varias alternativas para llevar a cabo el proceso de inferencia. Una de ellas consiste en intentar estimar, o bien el valor de la variable, o bien y de forma más general, una distribución de posibilidad que restrinja sus posibles valores. Ahora bien, esto sólo puede hacerse si se conoce su relación con otras variables conocidas, ya sea numérica o en términos más cualitativos, pudiéndose en el primer caso optar por un modelo matemático que estime su valor, y en el segundo por un modelo borroso que permita obtener una distribución de posible o necesaria posibilidad. Sobre modelado borroso hay una extensa bibliografía entre la que destacamos por ejemplo [Pedrycz, 93] donde se estudia los modelos relacionales borrosos y su identificación, [Pedrycz, 96] donde se hace una recopilación de distintas técnicas para la construcción

de modelos borrosos, [Wang, 92b] donde se propone un método para la creación de un modelo borroso a partir de datos numéricos, y por último [Nakanishi, Turksen, Sugeno, 93] donde se propone un interesante enfoque para la identificación de estos modelos. Queremos hacer notar sin embargo que en el ámbito donde habitualmente se estudian y se aplican los modelos borrosos suele ser necesario desborrosificar la distribución de posibilidad resultante, con vistas a obtener resultados numéricos. El problema de la desborrosificación es delicado y la justificación de unas u otras técnicas se basan principalmente en un conocimiento empírico (véase por ejemplo [Villar, Sanz, 94] donde se estudia el comportamiento de distintos métodos de desborrosificación en función de la funciones de implicación empleada). Sin embargo en nuestro caso no existe la necesidad de obtención de una solución numérica, puesto que la distribución de posibilidad resultante puede utilizarse para activar el modelo de una regla. A este respecto hay que hacer notar que la distinción aquí propuesta entre distribuciones de necesaria y posible posibilidad así como modelos de regla de necesidad y posibilidad es fundamental para que el razonamiento sea coherente. Por ello, dependiendo del tipo de modelo borroso empleado, debe determinarse el tipo de distribución de posibilidad obtenida como estimación de la variable incógnita, para así disparar con ella los modelos de regla adecuados.

Otra opción menos costosa pero también menos informativa consiste en modelar el valor de la variable mediante su distribución de necesidad, que siempre existe, y que en el peor de los casos será igual a su universo de discurso. En este último caso todos los modelos de necesidad entre cuyos antecedentes esté alguno relacionado con la variable en cuestión no podrán ser activados, por lo que un amplio grupo de reglas puede quedar totalmente inutilizado. Sin embargo, si para dichas reglas se dispone de un modelo de posibilidad alternativo al de necesidad, puede razonarse con los modelos de posibilidad y obtener conclusiones no triviales, aunque menos fiables que las que se hubiesen podido obtener si se hubiese conocido una distribución de posibilidad en base a criterios menos subjetivos o arbitrarios. A su vez estas conclusiones pueden utilizarse para reforzar o descartar la conveniencia de medir las variables desconocidas, dependiendo de su trascendencia. Este es un caso interesante en el que los modelos de posibilidad se muestran especialmente útiles. Queda sin embargo por determinar qué modelos de posibilidad deben utilizarse conocido el modelo de necesidad de una regla dada. Trataremos este punto más en detalle cuando analicemos la dualidad semántica de modelos en el apartado 5.7.

Existen por último situaciones en las que los modelos de posibilidad no permiten tampoco obtener conclusiones no triviales. En situaciones así, donde sea determinante obtener algún tipo de conclusión por carecer de explicaciones para un determinado fenómeno, interesa encontrar algún mecanismo que lo permita. Para ello pueden relajarse algunas de las restricciones impuestas en el encadenamiento de las reglas. De esta forma, si dado un modelo de posibilidad para una regla, no es posible obtener a partir de él y de una determinada observación una conclusión no trivial, puede ser interesante utilizar para su

activación la distribución de necesaria posibilidad de la variable, que al ser mayor que la de posibilidad facilitará la activación del modelo. Esto es equivalente a establecer hipótesis del tipo ‘ B es un rango posible’ para una variable v cuando no se dispone de información alguna que permita hacer una hipótesis mejor. Sin embargo esta forma de proceder reduce la fiabilidad de las conclusiones, y además debe tenerse en cuenta que la decisión de qué hipótesis deben hacerse puede resultar en general muy compleja, y fuertemente dependiente del contexto.

5.3 Agregación de varias observaciones relativas a una misma variable

Dependiendo de cómo se modele la incertidumbre, ya sea mediante distribuciones de probabilidad clásicas, ya sea mediante la teoría de Dempster-Shafer, ya sea mediante distribuciones de posibilidad (opción esta última escogida en esta tesis), los mecanismos disponibles para la agregación de información son diferentes unos de otros. A este respecto existen varios trabajos interesantes, como por ejemplo [Dubois, Prade, 88b] donde se comparan distintas técnicas de representación de incertidumbre y sus mecanismos de agregación asociados, proporcionando una interesante visión de conjunto. Se proponen las agregaciones conjuntivas para informaciones fiables, las disyuntivas para las poco fiables, y las medias como un compromiso entre ambas. En [Aczel, Alsina, 87] se estudian los operadores para la agregación de juicios sobre una medida de un objeto (peso, volumen, importancia, etc.) analizando aquellas propiedades deseables según la semántica de la agregación, dando lugar a operadores basados en medias. En [Dubois, Prade, 81] se analiza la aplicación del principio de extensión para extender la suma de variables cuando éstas son interactivas (bajo distintos tipos de interactividad) pero aunque se proponen soluciones computacionales, no se proponen interpretaciones al uso de t-normas distintas de *min*, tipo de interactividad que denominan débil.

Sin embargo en el ámbito borroso no existen muchos trabajos que profundicen en las diferencias semánticas que resultan por ejemplo de utilizar una u otra t-norma para una agregación conjuntiva y su relación con la interactividad entre variables, y se limitan normalmente a comprobaciones empíricas en aplicaciones numéricas, o a señalar sencillamente la fuerte dependencia de la agregación con el contexto. El enfoque propuesto en este punto, otra de las aportaciones originales de la tesis, se basa en modelar la información fiable con distribuciones de necesaria posibilidad y la no fiable con distribuciones de posible posibilidad, con lo que los operadores de agregación vienen determinados por la interpretación misma de cada tipo de distribución. Además se aborda el problema de la interactividad entre variables, para determinar, al menos de forma cualitativa, cuándo debe escogerse una u otra t-norma (o t-conorma), según el enfoque propuesto. El uso de otros operadores de agregación como por ejemplo las medias quedaría relegado a otros tipos de distribuciones de posibilidad, como se apunta en el apéndice de la tesis.

5.3.1 Agregación de varias observaciones

Cuando se dispone de varias fuentes de información sobre el valor de una variable se plantea el problema de cómo utilizar o combinar dichas fuentes para su uso directo por un sistema de reglas. Situaciones así pueden presentarse cuando se dispone por ejemplo de varios sensores para medir una misma variable, o bien de varios juicios emitidos por distintos expertos sobre los posibles valores que una variable puede tomar. Supondremos que en cualquier caso la información se expresa en forma de distribuciones de posibilidad.

- Fuentes de información fiables:

Si suponemos que todas las fuentes son igualmente fiables y con un grado de fiabilidad (o certidumbre) alto, entonces podemos interpretar las distribuciones de posibilidad como distribuciones de necesaria posibilidad (como se vio en el capítulo 3), y por tanto la forma lógica de agregarlas consiste en utilizar algún operador de tipo conjuntivo. El uso de una conjunción expresa que al ser las dos estimaciones fiables el valor real de la variable estimada debe pertenecer a ambas, o lo que es lo mismo ambas estimaciones son rangos necesarios. En virtud del principio de mínima especificidad, en ausencia de información adicional la agregación conjuntiva debe basarse en el uso del operador *min*, y se dice entonces que las variables son no interactivas. Esta agregación permite obtener la distribución menos específica que cumple las restricciones impuestas por las distribuciones que se agregan (véase el capítulo 3, [Dubois, Prade, 88b] para la agregación de información según distintos esquemas de representación de incertidumbre y [Dubois, Prade, 91] para el principio de mínima especificidad). En este caso el resultado final es una distribución de necesaria posibilidad y puede entonces usarse para disparar modelos de necesidad y también, como se ha visto, modelos de posibilidad en algunos casos excepcionales.

Cuando entre las estimaciones que proporcionan cada una de las fuentes se conoce la existencia de una relación funcional, entonces las estimaciones de las fuentes son a priori interactivas y el operador *min* no genera la agregación apropiada.

Supongamos que dos fuentes proporcionan estimaciones para una variable u . A la variable correspondiente con la estimación de la primera fuente la llamaremos u_1 , y a la de la segunda u_2 , y a los valores concretos que tomen (ya sean numéricos o distribuciones de posibilidad) los llamaremos realizaciones.

Si ambas fuentes no son independientes podemos suponer que existe una relación funcional entre las estimaciones (aunque caben otros tipos de relaciones más complejas) que expresaremos como:

$$f(u_1, u_2) = 0$$

Si ambas fuentes son totalmente fiables y precisas las realizaciones de las estimaciones son numéricas e iguales al valor real de la variable estimada. En este caso la función $f(u_1, u_2) = 0$ sólo puede ser $u_1 - u_2 = 0$, o lo que es lo mismo $u_1 = u_2$. Al ser las estimaciones numéricas y necesariamente iguales no cabe ambigüedad a la hora de agregarlas. En este caso el resultado final es igual al obtenido utilizando la agregación *min*.

En general si las estimaciones son fiables, y sus realizaciones se modelan con distribuciones de necesaria posibilidad P_1 y P_2 , se tiene la garantía de que el valor real de la variable estimada pertenece a ambas realizaciones. Si suponemos que son distribuciones clásicas, y ambas son fiables deberá entonces tenerse $P_1 \cap P_2 \neq 0$ de forma que el verdadero valor de la variable esté incluido en ambas distribuciones. Si además existe una cierta dependencia entre las estimaciones de ambas fuentes, expresada mediante la relación funcional $f(u_1, u_2) = 0$, existirán elementos u_1 de P_1 que no verifiquen $f(u_1, u_2) = 0$ con ningún elemento u_2 de P_2 y que por tanto podrán ser descartados como valores posibles para la variable u , y viceversa. La distribución final es por tanto la intersección de ambas distribuciones de la que además se deben haber eliminado todos los pares de elementos (u_1, u_2) pertenecientes a $P_1 \times P_2$ que no verifican $f(u_1, u_2) = 0$. La dependencia entre las fuentes hace que la distribución resultante de la agregación sea siempre menor o igual que su intersección.

En el caso más general en que las distribuciones de necesaria posibilidad sean conjuntos borrosos, la aplicación del principio de extensión permite evaluar la influencia de la dependencia funcional entre las estimaciones de las distintas fuentes de información. En efecto su aplicación permite restringir los valores de una de las variables conocida su relación funcional con la otra, y la distribución de posibilidad de esta última. Así para la variable u_2 la aplicación del principio de extensión permite obtener:

$$P_{2-1}(u_2) = \sup_{u_1 / f(u_1, u_2) = 0} P_1(u_1)$$

donde $P_{1-2}(u_2)$ es la distribución de posibilidad que restringe los valores de la variable u_2 debido la restricción sobre la variable u_1 y debido a la relación funcional entre u_1 y u_2 . Como además u_2 viene restringida también por $P_2(u_2)$ la distribución final para la variable u_2 es:

$$P_{F2}(u_2) = \min(P_{2-1}(u_2), P_2(u_2)) = \min(\sup_{u_1 / f(u_1, u_2) = 0} P_1(u_1), P_2(u_2))$$

Del mismo modo para la variable u_1 se tiene:

$$P_{1-2}(u_1) = \sup_{u_2 / f(u_1, u_2) = 0} P_2(u_2)$$

$$P_{F1}(u_1) = \min(P_{1-2}(u_1), P_1(u_1)) = \min(\sup_{u_2 / f(u_1, u_2)=0} P_2(u_2), P_1(u_1))$$

Ahora bien, dado que las dos variables u_1 y u_2 son estimaciones para la misma variable u , también lo son las distribuciones de posibilidad disponibles tanto para u_1 como para u_2 . Por tanto para obtener una estimación final para la variable u podemos agregar conjuntamente las dos distribuciones de posibilidad obtenidas, y en virtud del principio de mínima especificidad escoger la t-norma *min*, ya que las relaciones de dependencia han sido previamente tenidas en cuenta:

$$\begin{aligned} P_F(u) &= \min(P_{F1}(u), P_{F2}(u)) = \\ & \min(\sup_{u_2 / f(u_1, u_2)=0} P_2(u_2), P_1(u_1), \sup_{u_1 / f(u_1, u_2)=0} P_1(u_1), P_2(u_2)) \\ & \leq \min(P_1(u), P_2(u)) \end{aligned}$$

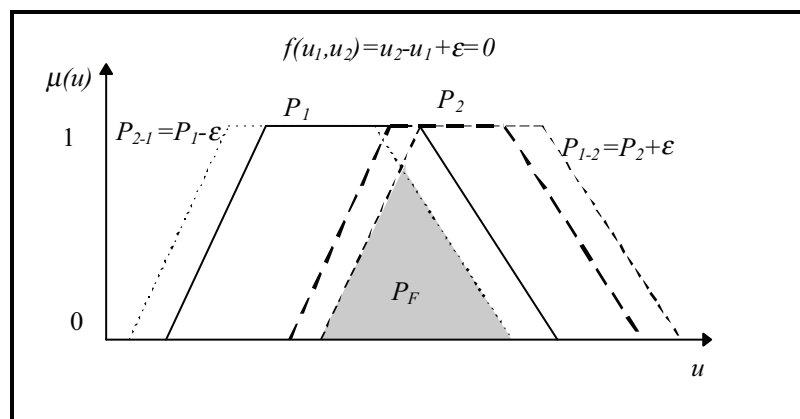
La distribución así obtenida es menor o igual que la que resulta de una agregación con la t-norma *min* y por tanto más informativa, lo que es lógico ya que resulta de tener en cuenta una información adicional dada por $f(u_1, u_2) = 0$. Si por ejemplo f es de la forma:

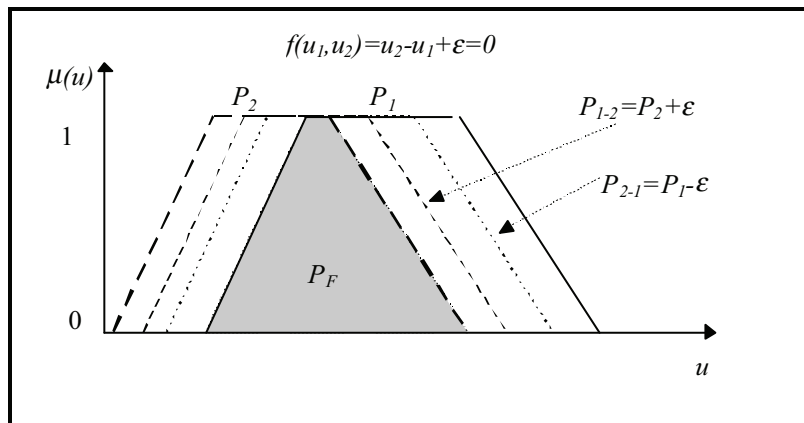
$$f(u_1, u_2) = 0 \text{ si } u_1 = u_2 + \epsilon$$

se tiene, aplicando aritmética borrosa (véase el capítulo 3):

$$\begin{aligned} P_{1-2}(u_1) &= P_2(u_2) + \epsilon \\ P_{2-1}(u_2) &= P_1(u_1) - \epsilon \end{aligned}$$

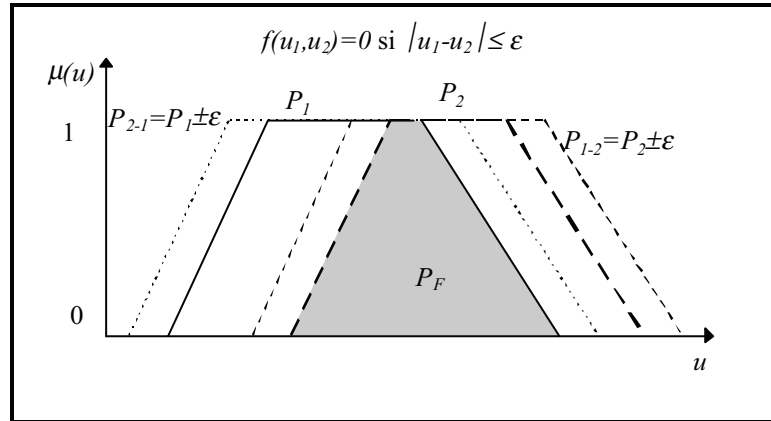
donde $P_2(u_2) + \epsilon$ simboliza en este caso la suma borrosa entre los números borrosos $P_2(u_2)$ y ϵ . Podemos entonces representar gráficamente el resultado de la agregación en las siguientes figuras, para dos situaciones relativas de P_1 y P_2 :





En las figuras se aprecia que la agregación final es en general inferior o igual a la obtenida con el operador *min*. Nótese que la agregación no es simétrica y depende de si P_1 está a la derecha o a la izquierda de P_2 , de forma que en algunos casos la restricción $f(u_1, u_2)$ no tiene efecto sobre el resultado final. De hecho, y con carácter general, podemos señalar que la distribución final resultante de la agregación conjunta es menor que la obtenida con el operador *min* cuando las realizaciones de las estimaciones presenta una relación contradictoria con la que supuestamente deberían mantener las estimaciones. Así en este caso dado que la relación entre las estimaciones viene dada por $u_1 = u_2 + \varepsilon$, la realización P_1 de u_1 debería ser mayor que la realización P_2 de u_2 lo que se traduce en que P_1 debería situarse a la derecha de P_2 . En las figuras vemos que cuando esto se cumple la distribución final es igual a la obtenida con el operador *min* (primera figura), mientras que cuando no se cumple (segunda figura) se obtiene una distribución final menor. Todos aquellos valores pertenecientes a las realizaciones pero contradictorios con la relación funcional entre las estimaciones son eliminados.

Otro ejemplo interesante viene dado por la relación $f(u_1, u_2) = 0$ si $|u_1 - u_2| \leq \varepsilon$. En este caso podemos escribir que $P_{1,2} = P_2 \pm \varepsilon$ y $P_{2,1} = P_1 \pm \varepsilon$ por lo que la representación gráfica conduce a:



y por tanto el suponer que existe una cierta tolerancia o error inferior a ϵ entre las dos estimaciones no supone modificación alguna sobre la agregación *min* y la distribución final es la misma. En efecto, dado que cualquier valor no incluido en P_1 y P_2 queda excluido como posible (al ser distribuciones de necesaria posibilidad), solo serán también posibles los valores de u_2 posibles según P_2 y que verifiquen $u_1 - \epsilon$ y $u_1 + \epsilon$ para cualquier valor de u_1 posible según P_1 . Como en este caso no existe contradicción alguna entre P_1 y P_2 y la relación de dependencia entre las fuentes, la distribución final es la misma que la que se obtiene agregando P_1 y P_2 con el operador *min*.

Cuando la relación funcional es tal que cualquier par de valores u_1 y u_2 la verifica, es decir f es idénticamente nula, entonces se tiene:

$$P_{2-1}(u_2) = \sup_{u_1 / f(u_1, u_2) = 0} P_1(u_1) = 1$$

$$P_{1-2}(u_1) = \sup_{u_2 / f(u_1, u_2) = 0} P_2(u_2) = 1$$

y se recupera la agregación con la t-norma *min*, de forma que:

$$P_F(u) = \min(P_1(u), P_2(u))$$

Este resultado también se obtiene para la función f tal que $f(u_1, u_2) = 0$ para $u_1 = u_2$. En efecto puede comprobarse que entonces se tiene:

$$P_{2-1}(u_2) = \sup_{u_1 / f(u_1, u_2) = 0} P_1(u_1) = \sup_{u_1 / u_1 = u_2} P_1(u_1) = P_1(u_2)$$

$$P_{1-2}(u_1) = \sup_{u_2 / f(u_1, u_2) = 0} P_2(u_2) = \sup_{u_2 / u_2 = u_1} P_2(u_2) = P_2(u_1)$$

La no interactividad entre estimaciones para una misma variable es compatible con la existencia de una relación funcional $f(u_1, u_2) = 0$ que las liga, pero de tal modo que existe un cierto valor $\epsilon > 0$ para el que el conjunto de pares (u_1, u_2) definido por $|u_1 - u_2| \leq \epsilon$ está

incluido en el conjunto de pares definido por $f(u_1, u_2) = 0$, o lo que es lo mismo cuando $P_{2-1}(u) \geq P_1(u)$ y $P_{1-2}(u) \geq P_2(u)$. La interactividad por el contrario se manifiesta cuando existe una relación funcional entre las estimaciones que refleja una contradicción entre realizaciones concretas de las estimaciones.

Podemos encontrar un tipo de interactividad a nivel lógico, y que por tanto no puede expresarse únicamente como una relación funcional entre las estimaciones, sino que tiene en cuenta los grados de pertenencia de u_1 y u_2 a sus respectivas distribuciones de posibilidad. Esta interactividad refleja la existencia de una relación funcional más compleja, dependiente de los grados de pertenencia y por tanto implica una relación a nivel lógico entre las estimaciones dadas por ambas fuentes.

Un ejemplo sencillo de este tipo de interactividad consiste en plantear la siguiente función f :

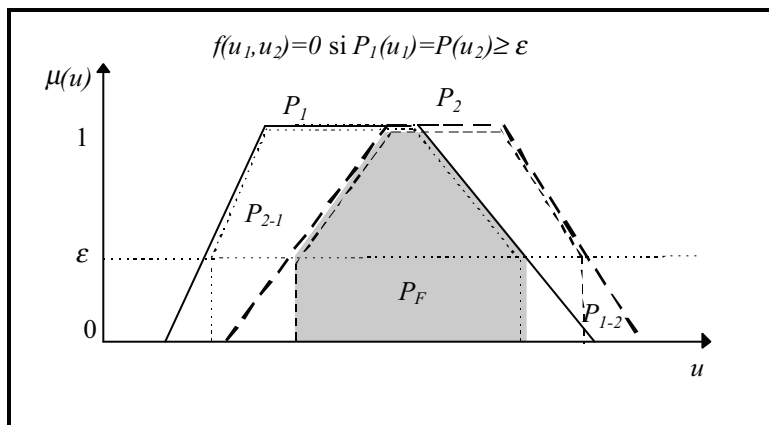
$$f(u_1, u_2) = 0 \text{ para } P_1(u_1) = P_2(u_2) \geq \epsilon$$

La aplicación del principio de extensión permite obtener las distribuciones siguientes:

$$P_{2-1}(u_2) = \sup_{u_1 / P_1(u_1) = P_2(u_2) \geq \epsilon} P_1(u_1) = \begin{cases} P_1(u_2) & \text{si } P_1(u_2) \geq \epsilon \\ 0 & \text{en otro caso} \end{cases}$$

$$P_{1-2}(u_1) = \sup_{u_2 / P_1(u_1) = P_2(u_2) \geq \epsilon} P_2(u_2) = \begin{cases} P_2(u_1) & \text{si } P_2(u_1) \geq \epsilon \\ 0 & \text{en otro caso} \end{cases}$$

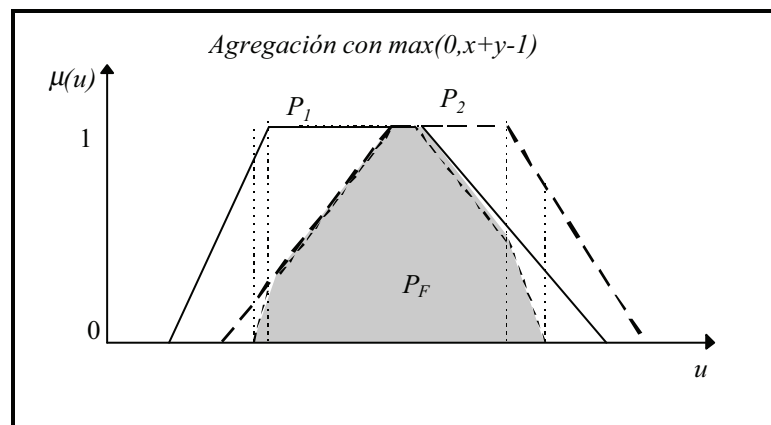
de forma que las distribuciones que se obtienen para cada una de las variables u_1 y u_2 son el resultado de una cierta operación de concentración (véase en el capítulo 3 el apartado 3.4.3.) de la distribución que se obtendría con la agregación *min*. Gráficamente puede representarse como:



Esta agregación sí es simétrica y guarda un cierto parecido con las agregaciones basadas en t-normas distintas (y por tanto menores) que la t-norma *min*. A efectos interpretativos

nótese que la función f definida en este ejemplo muestra la existencia de una contradicción permanente entre las realizaciones de las estimaciones, independiente de sus posiciones relativas sobre su universo de discurso. Esto implica que en el resultado final se rechazan siempre valores con grados de pertenencia bajos a ambas realizaciones.

El resultado obtenido en el ejemplo anterior es cualitativamente similar al que podría obtenerse con una t-norma menor que min . En particular la agregación mediante la t-norma de Lukasiewicz produce una distribución igual a la que aparece en la siguiente figura:



Al igual que en el ejemplo anterior, podemos decir que el uso de una t-norma menor que min refleja la existencia de una contradicción permanente entre las realizaciones de las estimaciones. Si los grados de posibilidad de las distribuciones P_1 y P_2 para una misma posible estimación u_0 de la variable u han sido obtenidos a partir de informaciones contradictorias o procedimientos distintos, entonces el uso de una t-norma pequeña para su agregación, como por ejemplo la t-norma de Lukasiewicz, está especialmente indicado. Si por el contrario han sido obtenidos mediante informaciones no contradictorias o procedimientos idénticos, entonces la t-norma min es la más indicada para su agregación. Si por último han sido obtenidos a partir de informaciones (o procedimientos) que pueden ser o no contradictorias, con un cierto grado de aleatoriedad, entonces una t-norma intermedia, como por ejemplo el producto, parece lo más indicado. Esta relación entre el tamaño de la t-norma y la contradicción entre las observaciones que se agregan ha sido también puesta de manifiesto en otros trabajos desde ópticas diferentes, como por ejemplo [Bonissone, Decker, 86].

Supongamos por ejemplo que para estimar el valor de una variable se dispone de dos sensores que cometen un cierto error en la medida, y se decide utilizar una distribución de posibilidad para modelar la medida que cada uno de ellos proporciona. Si los sensores son

fiables, pero utilizan procedimientos de medida distintos, entonces el uso de una t-norma pequeña como la de Lukasiewicz está especialmente indicado. Si por el contrario ambos procedimientos de medida son semejantes entonces el uso de la t-norma *min* es el más indicado. Por último, si existe una cierta aleatoriedad en el procedimiento de cada uno de los sensores entonces el uso de la t-norma *prod* es el más indicado por ser intermedia entre la de Lukasiewicz y *min*.

Aunque no es directamente comparable con lo que se ha analizado en esta sección, en [Kruse, Gebhardt, Klawonn, 94] puede verse un ejemplo ilustrativo en el que se muestra la aplicación de las tres principales t-normas y sus diferencias semánticas en un caso concreto. En dicho ejemplo se puede comprobar como efectivamente el uso de la t-norma *min* está asociado con la agregación de elementos semejantes, el uso de la t-norma de Lukasiewicz (la menor t-norma continua) está asociado con la agregación de elementos contradictorios, y el uso de la t-norma *prod* esta asociado con la agregación de elementos que pueden ser con la misma probabilidad semejantes y contradictorios.

- Fuentes de información no fiables:

Cuando la fiabilidad de las fuentes es pobre, entonces la información que proporcionan puede interpretarse como distribuciones de posible posibilidad, y deben por tanto agregarse mediante un operador disyuntivo. El uso de un operador disyuntivo expresa que al ser poco fiables las estimaciones, el valor real de la variable estimada podría pertenecer a la primera estimación, pero también podría pertenecer a la otra, o lo que es lo mismo ambas estimaciones son rangos posibles pero no necesarios. Nótese que se está asumiendo implícitamente que otros valores no pertenecientes a las distribuciones de posibilidad agregadas podrían también ser posibles. Como se vio en el capítulo 3, este tipo de distribuciones, en ausencia de información adicional, deben agregarse disyuntivamente con el operador *max*. En este caso la distribución resultante es una distribución de posible posibilidad y sólo puede utilizarse para activar modelos de posibilidad si no se quiere incurrir en contradicciones lógicas.

Si existe una cierta dependencia funcional entre las estimaciones de dos fuentes distintas, retomando la notación utilizada en el apartado anterior, la aplicación del principio de extensión permite restringir los valores de una de las variables u_2 conocida su relación funcional $f(u_1, u_2) = 0$ y la distribución de posible posibilidad $P_1(u_1)$:

$$P_{2-1}(u_2) = \sup_{u_1 / f(u_1, u_2) = 0} P_1(u_1)$$

Sin embargo las distribuciones que se obtienen en este caso son distribuciones de posibilidad y deben por tanto agregarse disyuntivamente mediante el operador *max* ya que las dependencias ya están siendo tenidas en cuenta explícitamente:

$$P_{F2}(u_2) = \max(P_{2-1}(u_2), P_2(u_2)) = \max(\sup_{u_1 / f(u_1, u_2)=0} P_1(u_1), P_2(u_2))$$

Del mismo modo para la variable u_1 se tiene:

$$P_{1-2}(u_1) = \sup_{u_2 / f(u_1, u_2)=0} P_2(u_2)$$

$$P_{F1}(u_1) = \max(P_{1-2}(u_1), P_1(u_1)) = \max(\sup_{u_2 / f(u_1, u_2)=0} P_2(u_2), P_1(u_1))$$

Agregando toda la información de que se dispone, la estimación final para la variable u viene dada por:

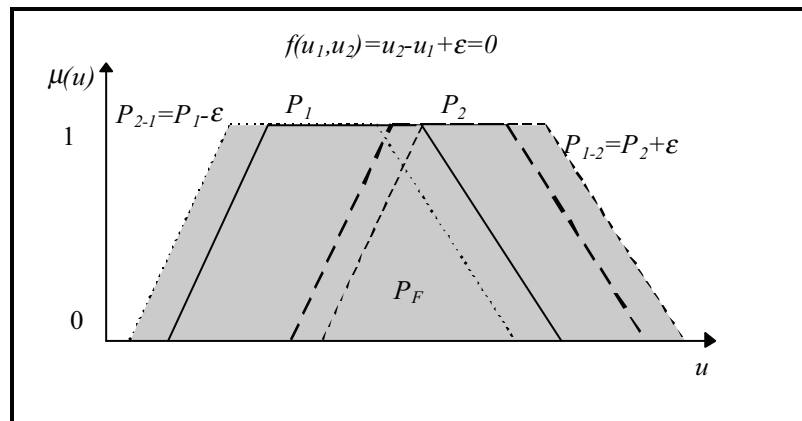
$$P_F(u) = \max(P_{F1}(u), P_{F2}(u)) =$$

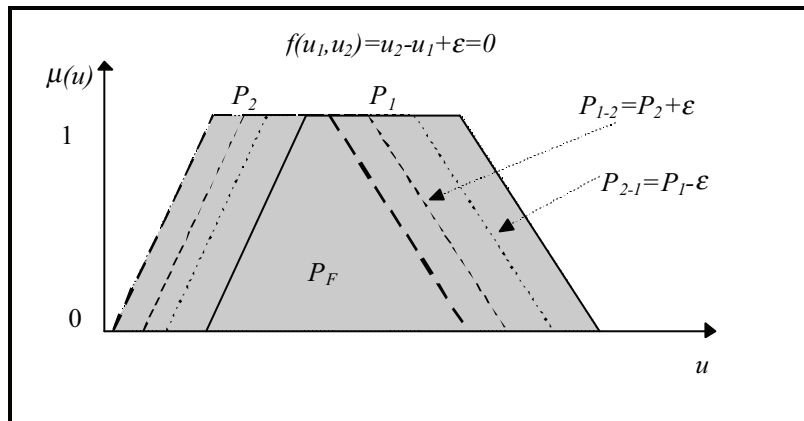
$$\max(\sup_{u_2 / f(u_1, u_2)=0} P_2(u_2), P_1(u_1), \sup_{u_1 / f(u_1, u_2)=0} P_1(u_1), P_2(u_2))$$

$$\geq \max(P_1(u), P_2(u))$$

La función f refleja la existencia de valores, que no tenidos en cuenta por alguna de las fuentes de información, deben ser incluidos como posibles en la distribución final. En efecto, debido a su interpretación, las distribuciones de posible posibilidad no excluyen como posible a ningún elemento del universo de discurso. En este caso la distribución así obtenida es mayor o igual que la que resulta de la agregación con la t-norma max .

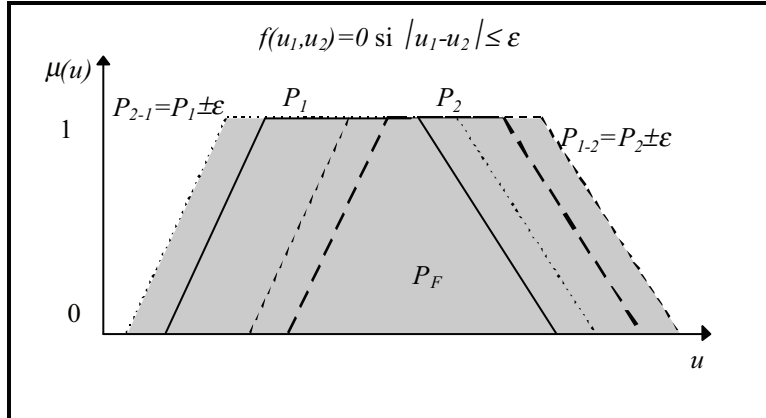
Retomando los ejemplos del apartado anterior, para el primero de ellos se tiene:





Puede también notarse cómo nuevamente la agregación tampoco es simétrica y depende de las posiciones relativas de P_1 y P_2 . En este caso la función f también expresa la existencia de una cierta contradicción entre las realizaciones de las estimaciones. Pero como en este caso las realizaciones no son fiables, la alternativa más sensata es incluir todos los valores necesarios para eliminar la contradicción existente, en lugar de excluir a aquellos valores que provocan la contradicción.

Para el segundo ejemplo, esto es cuando $f(u_1, u_2) = 0$ si $|u_1 - u_2| \leq \varepsilon$, se tiene:



y por tanto el suponer que existe una cierta tolerancia o error inferior a ε entre las dos estimaciones se traduce en que un aumento de la distribución final depende de ε . Este resultado era esperable en el sentido en que si u_1 es un valor posible en cierto grado, entonces dado que tanto P_1 como P_2 no permiten excluir a ningún valor, también son igualmente posibles los valores $u_1 - \varepsilon$ y $u_1 + \varepsilon$.

Cuando la relación funcional es tal que ningún par de valores u_1 y u_2 la verifica, es decir f es distinta de 0 para todo par (u_1, u_2) entonces se tiene:

$$P_{2-1}(u_2) = \sup_{u_1 / f(u_1, u_2) = 0} P_1(u_1) = 0$$

$$P_{1-2}(u_1) = \sup_{u_2 / f(u_1, u_2) = 0} P_2(u_2) = 0$$

y entonces se recupera la agregación con la t-conorma *max*, de forma que:

$$P_F(u) = \max(P_1(u), P_2(u))$$

En efecto en este caso *f* no relaciona a ningún par de elementos (u_1, u_2) . Este resultado se obtiene también si *f* es tal que $f(u_1, u_2) = 0$ para $u_1 = u_2$, en cuyo caso se tiene:

$$P_{2-1}(u_2) = \sup_{u_1 / f(u_1, u_2) = 0} P_1(u_1) = \sup_{u_1 / u_1 = u_2} P_1(u_1) = P_1(u_2)$$

$$P_{1-2}(u_1) = \sup_{u_2 / f(u_1, u_2) = 0} P_2(u_2) = \sup_{u_2 / u_2 = u_1} P_2(u_2) = P_2(u_1)$$

Puede manejarse nuevamente el concepto de interactividad entre estimaciones, correspondiente a la existencia de una relación *f* para la que la agregación resultante es mayor que la obtenida con el operador *max*. Las estimaciones son interactivas si dada una relación funcional $f(u_1, u_2) = 0$ que las liga, existe un cierto valor $\varepsilon > 0$ tal que el conjunto de pares (u_1, u_2) definido por $|u_1 - u_2| \leq \varepsilon$ está incluido en el conjunto de pares definido por $f(u_1, u_2) = 0$, o lo que es lo mismo cuando $P_{2-1}(u) \geq P_1(u)$ y $P_{1-2}(u) \geq P_2(u)$.

Para que las estimaciones proporcionadas por dos fuentes no sean interactivas, independientemente de su fiabilidad, retomando los resultados obtenidos para distribuciones fiables, podemos concluir que entre las estimaciones sólo puede existir una relación de igualdad, es decir:

$$f(u_1, u_2) = 0 \text{ si } u_1 = u_2$$

Nótese que este resultado es lógico ya que la función *f* así definida no aporta ninguna información extra, ni desde un punto de vista de necesidad ni de posibilidad. En efecto se corresponde con la suposición por defecto de que las fuentes de información proporcionan la misma estimación para la variable estimada, sean o no fiables.

Al igual que antes puede darse interactividad a nivel lógico. Con un ejemplo similar al del apartado anterior, escogiendo la relación entre las estimaciones como:

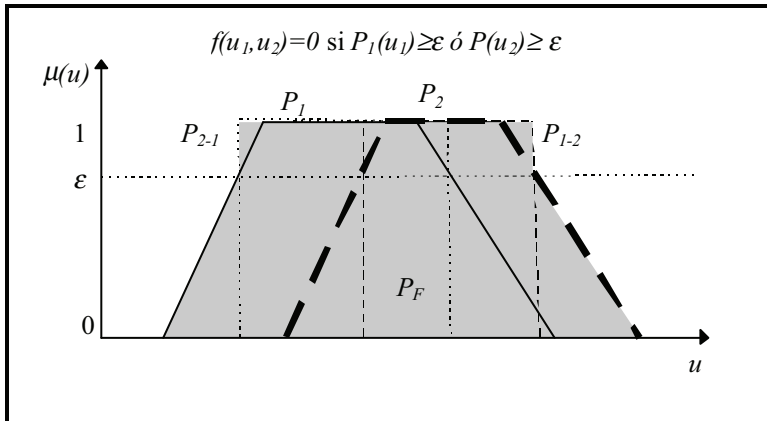
$$f(u_1, u_2) = 0 \text{ para } P_2(u_2) \geq \varepsilon \text{ ó } P_1(u_1) \geq \varepsilon$$

las distribuciones obtenidas mediante el principio de extensión son:

$$P_{2-1}(u_2) = \sup_{u_1 / P_2(u_2) \geq \epsilon} P_1(u_1) = \begin{cases} 1 & \text{si } P_1(u_2) \geq \epsilon \\ 0 & \text{en otro caso} \end{cases}$$

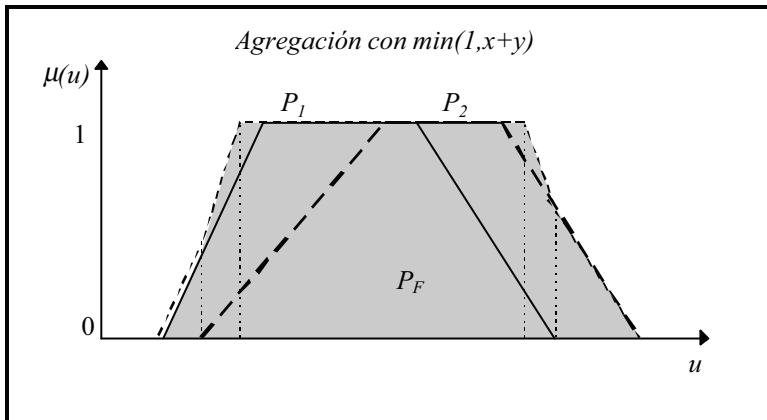
$$P_{1-2}(u_1) = \sup_{u_2 / P_1(u_1) \geq \epsilon} P_2(u_2) = \begin{cases} 1 & \text{si } P_2(u_1) \geq \epsilon \\ 0 & \text{en otro caso} \end{cases}$$

y en este caso el resultado final que se obtiene es:



Como ocurría en el caso de informaciones fiables, en este caso la función de ligadura entre u_1 y u_2 permite recuperar un comportamiento simétrico y similar al que se tendría utilizando una t-conorma mayor que *max*. A efectos interpretativos la función f escogida permite incluir como estimaciones totalmente posibles a estimaciones que sólo lo eran parcialmente en alguna de las distribuciones de partida. Esto significa un reforzamiento de todos los valores con grados de pertenencia altos a alguna de las estimaciones disponibles. Se produce por tanto un aporte adicional de evidencia en favor de los elementos poco dudosos.

Este resultado es cualitativamente similar al que podría obtenerse con la t-conorma Lukasiewicz, que produce una distribución igual a la que aparece en la siguiente figura:



El uso de unas u otras t-normas es, como en el caso anterior, reflejo de las contradicciones existentes entre las informaciones que se agregan. Así cuando los grados de posibilidad de las distribuciones P_1 y P_2 para una misma posible estimación u_0 de la variable u han sido obtenidos a partir de informaciones contradictorias o procedimientos diferentes, entonces el uso de una t-conorma grande como la de Lukasiewicz es el más indicado. Cuando se hayan obtenido a partir de informaciones no contradictorias o procedimientos similares, entonces la t-conorma más apropiada es la t-conorma *max*. Por último cuando se hayan obtenido a partir de informaciones (o procedimientos) que pueden con cierta aleatoriedad ser o no contradictorios entonces el uso de una t-conorma intermedia como la t-conorma dual del producto es el más indicado.

Así por ejemplo en el caso de que se disponga de dos sensores para estimar el valor de una variable a partir de sus medidas, que los sensores sean poco fiables, y que se modelen dichas medidas mediante distribuciones de posibilidad, cuando ambos sensores utilicen procedimientos de medida distintos, podrá utilizarse la t-conorma de Lukasiewicz. Cuando por el contrario se basen en el mismo principio de medida, entonces el uso de la t-conorma *max* es el más apropiado. Por último, si existe una cierta aleatoriedad en el procedimiento utilizado por cada uno de los sensores, entonces el uso de la t-norma *prod* es el más indicado, por ser intermedia entre la de Lukasiewicz y *min*.

- Fuentes con distinta fiabilidad:

Cuando se dispone de fuentes con distinta fiabilidad una forma de proceder consiste en determinar cuales de ellas tienen una fiabilidad alta, cuales baja, y obtener para la misma variable una distribución de necesaria posibilidad y otra de posible posibilidad mediante los procedimientos anteriores. Entonces los modelos de necesidad deberán activarse con la distribución de necesaria posibilidad obtenida y los modelos de posibilidad con la de posible posibilidad. La ventaja de manejar conjuntamente información fiable y no fiable reside en que en general la información fiable es menos específica y por tanto menos informativa que la información no fiable, lo que la convierte a menudo en poco útil en el proceso de razonamiento. Lógicamente será preferible la información no fiable, siempre y cuando las conclusiones que permita obtener sean suficientemente específicas, o lo que es lo mismo suficientemente alejadas de las conclusiones triviales.

Un argumento que puede esgrimirse en contra de esta forma de proceder consiste en señalar que cuando se agregan distribuciones poco fiables mediante operadores disyuntivos, la fiabilidad de la distribución resultante aumenta en detrimento de su precisión, que obviamente disminuye, por lo que podría tal vez utilizarse en algunos casos como distribución de necesaria posibilidad, siempre que se tuviese la certeza de que la

distribución así obtenida no excluye al valor real de la variable. Esto es obviamente cierto cuanto mayor sea la distribución resultante, y en efecto, el caso límite es una distribución de posible posibilidad igual al universo de discurso de la variable en cuestión, en cuyo caso la distribución final es directamente una distribución de posibilidad (o necesaria posibilidad):

$$\pi_n(u) \geq \pi_p(u) = U \Rightarrow \pi_n(u) = \pi_p(u) = U$$

Sin embargo nótese que a medida que el tamaño de la distribución aumenta, aunque su fiabilidad aumenta, también disminuye su capacidad de activar modelos de necesidad, y por tanto el posible incremento de su fiabilidad es cada vez menos útil de cara a la inferencia. El caso límite se da cuando la distribución resultante es el universo de discurso de la variable en cuestión en cuyo caso la fiabilidad es máxima, la información carece de incertidumbre, pero no puede activar ningún modelo de necesidad.

Cabe por último la posibilidad de que dos estimaciones sean poco fiables, pero se tenga la certeza de que la estimación final debe estar en alguna de las dos. En este caso la agregación disyuntiva produce una distribución de posibilidad (de posible y necesaria posibilidad). Sin embargo hay que hacer notar que esto se debe a la existencia implícita de una información adicional. En efecto, es contradictorio pensar que dos fuentes no fiables no puedan proporcionar una información fiable (aunque esta aumente con el tamaño de la distribución). Por tanto suponer que la agregación disyuntiva es fiable, es equivalente a tener una tercera fuente de información fiable, que proporciona una tercera distribución de posibilidad, de necesaria posibilidad en este caso, coincidente con la agregación disyuntiva de ambas, pero no resultante de ella. La información conjunta permite en este caso obtener una distribución de posibilidad, al coincidir las de posible con la de necesaria posibilidad, y por tanto utilizable para disparar los dos tipos de modelos.

5.3.2 Resolución de incoherencias

Cuando se agregan distribuciones de necesaria posibilidad pueden obtenerse en algunos casos distribuciones no normalizadas. Sin embargo esto es claramente un contrasentido ya que si ambas fuentes son fiables la intersección debe tener elementos totalmente posibles, y cuando los conjuntos son clásicos significa que no puede estar vacía (véase por ejemplo [Dubois, Prade, 88b]). Para resolver este tipo de inconsistencia proponemos las siguientes alternativas:

- revisar las fuentes de información para determinar si algunas pueden considerarse más fiables que otras. En caso afirmativo se obtendrían como resultado final dos distribuciones, una de posible posibilidad y otra de necesaria posibilidad. Nótese que

las fuentes de información pueden ser reglas de la base de conocimiento, y que por tanto la revisión puede dar lugar a cambiar modelos de necesidad en modelos de posibilidad para algunas de las reglas, por estar mal formuladas.

- si las fuentes no pueden ser catalogadas con distintas fiabilidades, suponer que todas ellas tienen poca fiabilidad y en ese caso obtener una distribución de posible posibilidad. En este caso sólo se podrán activar los modelos de posibilidad y la certeza de las conclusiones quedará mermada.
- si el grado de contradicción existente entre las fuentes es pequeño, lo que se traduce por una altura próxima a 1 en la distribución resultante de la agregación, se puede optar por normalizarla, por ejemplo dividiéndola por su altura. Esta alternativa es similar a la regla de combinación de Dempster, donde se realiza una agregación conjuntiva modificada por un factor de normalización que tiene en cuenta el grado de contradicción de las informaciones que se agregan ([Dubois, Prade, 88b]). Nótese que esto es en cierta medida incompatible con el hecho de agregar las distribuciones con una t-norma menor que *min* ya que el efecto que produce es justamente el contrario.

Cuando se agregan distribuciones de posible posibilidad no aparecen inconsistencias lógicas, a menos que la distribución de posible posibilidad resultante sea mayor que la distribución de necesaria posibilidad para la misma variable. En este caso puede interpretarse que la agregación de distribuciones de posible posibilidad aumenta suficientemente la fiabilidad de la distribución resultante, en detrimento de su precisión, de forma que se hace igual o mayor que la distribución de necesaria posibilidad. La distribución de necesaria posibilidad debe entonces interpretarse como distribución de posibilidad y ser utilizada para activar tanto modelos de posibilidad como modelos de necesidad, ya que utilizar una distribución más grande carecería de sentido. Para garantizar que la restricción fundamental de las distribuciones de posible posibilidad se sigue verificando se puede optar por corregir la distribución de posibilidad obtenida con la siguiente expresión:

$$\pi_p'(u) = \min(\pi_n(u), \pi_p(u))$$

siendo $\pi_n(u)$ y $\pi_p(u)$ las distribuciones de necesaria y posible posibilidad obtenidas a partir de una agregación de información procedente de distintas fuentes, y $\pi_p'(u)$ la distribución de posible posibilidad que debe finalmente ser utilizada.

5.4 Agregación de varios antecedentes en una misma regla

El problema de la agregación de múltiples antecedentes en una regla no reside en determinar si ésta debe ser conjuntiva o disyuntiva. En efecto, es el significado mismo de la regla que se enuncia y se quiere modelar, el que determina el tipo de agregación lógica que debe ser utilizado, al igual que ocurre con las agregaciones de distribuciones sobre distintos universos, como ya se comentó en el capítulo 3. Sin embargo una vez escogido el tipo de agregación, es necesario determinar una vez más el operador más apropiado que la modele de forma fiel, de entre la gran familia de operadores disponibles.

Al igual que en el estudio de una regla con un único antecedente uno de los requisitos básicos que razonablemente puede exigirse es el cumplimiento del modus ponens estricto. Sin embargo consideraciones semánticas pueden ser tenidas en cuenta para delimitar aún más las opciones disponibles.

Nos limitaremos al estudio de reglas con antecedentes agregados conjuntivamente, ya que es así como suele expresarse tradicionalmente el conocimiento cuando se modela mediante reglas de producción.

5.4.1 Reglas con múltiples antecedentes

Dadas dos variables u_1 y u_2 definidas en U_1 y U_2 respectivamente, dada una regla

$$\text{si } u_1 \text{ es } P_1 \text{ y } u_2 \text{ es } P_2 \text{ entonces } v \text{ es } Q$$

que denotaremos de forma abreviada como

$$P_1, P_2 \rightarrow Q$$

y dados dos conjuntos borrosos P'_1 y P'_2 definidos en U_1 y U_2 respectivamente, observaciones de u_1 y u_2 en forma de distribuciones de posibilidad, que limitan los valores que las variables u_1 y u_2 pueden tomar, la aplicación del MP utilizando la RCI puede formularse como:

$$Q'(v) = \sup_{u_1, u_2} M(T_1(P'_1(u_1), P'_2(u_2)), I(T_2(P_1(u_1), P_2(u_2)), Q(v))) \\ \sup_{w=(u_1, u_2)} M(P'(w), I(P(w), Q(v)))$$

donde $P'(w) = T_1(P'_1(u_1), P'_2(u_2))$ y $P(w) = T_2(P_1(u_1), P_2(u_2))$.

Uno de los primeros requisitos, al igual que con reglas de un solo antecedente, es la verificación del MPE, esto es de $P_1'(u_1)=P_1(u_1)$ y $P_2'(u_2)=P_2(u_2)$ debe deducirse la conclusión $Q'(v)=Q(v)$. El MPE se verifica si se toman las dos t-normas iguales, $T_1=T_2=T$, ya que se tiene:

$$Q'(v) = \sup_{u_1, u_2} M(T(P_1(u_1), P_2(u_2)) , I(T(P_1(u_1), P_2(u_2)), Q(v))) \\ \sup_{w=(u_1, u_2)} M(P(w), I(P(w), Q(v))) = Q(v)$$

alcanzándose el supremo para $P(w)=1$ siempre que P_1 y P_2 estén normalizados. De hecho para cualquier t-norma T_1 tal que $T_1(x,y) \leq T_2(x,y)$ el MPE sigue verificándose, puesto que se tiene:

$$\sup_{u_1, u_2} M(T_1(P_1(u_1), P_2(u_2)) , I(T_2(P_1(u_1), P_2(u_2)), Q(v))) \leq \\ \sup_{u_1, u_2} M(T_2(P_1(u_1), P_2(u_2)) , I(T_2(P_1(u_1), P_2(u_2)), Q(v))) = Q(v)$$

y el supremo para la primera de las expresiones vuelve a alcanzarse para $P(w)=1$ (puesto que el núcleo de la agregación de $P_1(u_1)$ y $P_2(u_2)$ con una t-norma menor sigue siendo el mismo conjunto).

Como se vio en la sección anterior, el uso de una t-norma u otra tiene una interpretación semántica, relacionada con el grado de contradicción de lo que se está agregando. Se puede repetir un razonamiento paralelo al llevado a cabo cuando se agregaban estimaciones para una misma variable, con la diferencia de que si antes se obtenía como resultado de la agregación un conjunto borroso, ahora se obtiene una relación borrosa en el universo producto $U_1 \times U_2$.

Así por ejemplo si los conceptos expresados en la regla por P_1 y P_2 son contradictorios, como los son por ejemplo “*precio coche barato*” y “*tamaño coche grande*” entonces parece razonable utilizar una t-norma pequeña para su agregación conjuntiva. Como la regla se dispara cuando las observaciones tienen un cierto parecido con los antecedentes de la regla (evaluado de forma distinta dependiendo del modelo de regla empleado), significa que las observaciones serán también ligeramente contradictorias, al serlo los antecedentes, y por tanto la agregación con la misma t-norma utilizada para los antecedentes parece una elección semánticamente coherente, y que como hemos visto, preserva el MPE. Nótese que el uso de una t-norma menor que *min* para la agregación de dos conjuntos borrosos no viene justificado únicamente por la existencia de una dependencia funcional entre las variables, sino por la existencia de una contradicción entre el significado de los conjuntos que se agregan y la relación funcional entre las variables a las que dichos conjuntos hacen referencia. Así por ejemplo la agregación de “*precio coche alto*” y “*tamaño coche grande*” no incurre en contradicciones con la dependencia directa

existente entre precio y tamaño, por lo que en este caso la agregación debería de hacerse con la t-norma *min*.

Nótese por último que la interactividad entre dos variables puede venir directamente expresada mediante una distribución de posibilidad bidimensional $\pi(u_1, u_2)$ que delimite los valores que puedan conjuntamente tomar. En este caso, la agregación conjuntiva de u_1 es P_1 y u_2 es P_2 debe corregirse con la distribución conjunta, quedando pues:

$$\min(P_1(u_1), P_2(u_2), \pi(u_1, u_2))$$

Es también conveniente recordar que el carácter conjuntivo de la agregación de distribuciones sobre distintos universos, para la obtención de una distribución de posibilidad conjunta exige, como ya se vio en el capítulo 3, una agregación de tipo conjuntivo (y no debe ser confundido con la agregación de distribuciones sobre un mismo universo, cuyo carácter conjuntivo o disyuntivo está relacionado con la fiabilidad de las distribuciones). Sin embargo, en el caso en que la regla fuese del tipo:

$$\text{si } u_1 \text{ es } P_1 \text{ o bien } u_2 \text{ es } P_2 \text{ entonces } v \text{ es } Q$$

entonces los antecedentes deberían agregarse mediante alguna t-conorma según consideraciones similares a las hechas en esta sección, y la distribución de posibilidad asociada a la observación no sería ya la distribución de posibilidad conjunta de la variable (u, v) , que se entiende como “ u y v ”, sino la distribución de otro tipo de variable disyuntiva “ u o v ”.

5.4.2 Agregación de antecedentes y condiciones de activación

En esta sección vamos a analizar de forma cualitativa la influencia que tiene la t-norma utilizada para la agregación de los antecedentes de una regla con su condición de activación, en función del tipo de modelo empleado.

- Modelos de necesidad:

Las condiciones de activación de los modelos de necesidad estudiados en el capítulo 4 pueden resumirse en los siguientes puntos:

- si $[P'] \subset [P]$ entonces:
 - los modelos graduales (M_R^{T+}, I_R^{T+}) se activan
 - los modelos mixtos (M_R^{T-}, I_R^{T-}) y (M_S^{T-}, I_S^{T-}) se activan

- si $]P'[\subset]P[$ entonces:
 - los modelos de certidumbre (M_S^{T+}, I_S^{T+}) se activan

El uso de una t-norma menor que *min* significa que la distribución $P'(w)=T(P_1'(u_1), P_2'(u_2))$ es menor que la que se habría obtenido con una t-norma mayor, pero aunque su soporte puede en general ser menor (dependiendo de si *T* es positiva o negativa) no así su núcleo que se mantiene igual, ya que para cualquier t-norma *T* se tiene $T(1, 1)=1$. Por la misma razón el soporte $]P[$ de $P=T(P_1(u_1), P_2(u_2))$ también será en general menor.

Para los modelos con condición de activación $]P'[\subset]P[$ una reducción de $]P[$ supone un endurecimiento de la condición de activación. Esto significa que todos los modelos con semántica gradual, es decir (M_R^{T+}, I_R^{T+}) , (M_R^{T-}, I_R^{T-}) y (M_S^{T-}, I_S^{T-}) ven endurecidas sus condiciones de activación. Para estos modelos la existencia de un cierto grado de contradicción entre antecedentes se traduce en un endurecimiento de las condiciones de activación de la regla.

Para los modelos con condición de activación $]P'[\subset]P[$, se produce una reducción tanto de $]P'[/math> como de $]P[$, con lo que la condición de activación queda cualitativamente inalterada. Esto es lo que ocurre con los modelos de certidumbre, es decir (M_S^{T+}, I_S^{T+}) . Es conveniente recordar que cuando en el modelo de certidumbre se usa una t-norma o FGMP menor que la máxima puede recuperarse la condición de activación de los modelos graduales.$

- Modelos de posibilidad:

Las condiciones de activación de estos modelos pueden resumirse en los siguientes puntos:

- si $]P'[\cap]P[\neq 0$ entonces:
 - los modelos de certidumbre (M_R^{T+}, M_R^{T+}) se activan
 - los modelos graduales (M_S^{T+}, M_S^{T+}) y (M_R^{T+}, M_S^{T+}) se activan
- si $f(P'(u))+f(P(u)) < f(0)$ entonces:
 - los modelos mixtos (M_R^{T-}, M_R^{T-}) y (M_S^{T-}, M_S^{T-}) se activan
- si $P(u)+P'(u) > 1$ entonces:
 - los modelos mixtos (M_S^{T-}, M_R^{T-}) se activan
 - los modelos de certidumbre (M_S^{T+}, M_R^{T+}) se activan

- si $\sup_u M_R^{T^-}(P'(u), M_S^{T^-}(P'(u), I)) > 0$ entonces:
 - los modelos mixtos $(M_R^{T^-}, M_S^{T^-})$ se activan

En este caso una reducción de $]P'[$ y $]P[$ supone un endurecimiento de todas las condiciones de activación aunque en general mayor para aquellas que son menos restrictivas.

5.4.3 La MS con múltiples antecedentes

Cuando la inferencia se realiza mediante la MS, entonces la conclusión puede formularse como:

$$Q'(v) = I(S(T(P'_1(u_1), P'_2(u_2)), T(P_1(u_1), P_2(u_2))) , Q(v))$$

donde u_1 y u_2 son dos variables definidas sobre dos universos diferentes U_1 y U_2 , y se ha supuesto además que la variable conjunta (u_1, u_2) no está sujeta a ninguna restricción. Si la medida de similitud S presenta la propiedad de separabilidad respecto a la agregación T entonces el grado de similitud global viene dado por:

$$S(T(P'_1(u_1), P'_2(u_2)) , T(P_1(u_1), P_2(u_2))) = H(S(P'_1(u_1), P_1(u_1)) , S(P'_2(u_2), P_2(u_2)))$$

donde el operador H puede ser en general distinto de T . Esto significa que en lugar de tener que hacer $N_{U_1} \cdot N_{U_2}$ operaciones basta con hacer $N_{U_1} + N_{U_2}$, siendo N_{U_1} y N_{U_2} los números de muestras tomadas para discretizar los universos U_1 y U_2 , para poder calcular numéricamente el resultado de la inferencia. De esta forma es posible calcular el grado de similitud de cada antecedente con su observación, en lugar de manejar distribuciones conjuntas.

Este problema se aborda por ejemplo en [Cross, Sudkamp, 93] donde al primero de los enfoques se le denomina enfoque basado en *compatibilidad-agregación*, mientras que el segundo es el basado en *agregación-compatibilidad*. Sin embargo no parece correcto suponer directamente, como se sugiere en [Cross, Sudkamp, 93], que la inferencia basada en la MS pueda formularse agregando después de obtener los índices S , sin analizar qué tipo de implicaciones lógicas puede acarrear, teniendo además en cuenta que la extensión de uno a varios antecedentes lleva precisamente a la formulación contraria. Sí se señala sin embargo que ambos métodos pueden presentar comportamientos significativamente diferentes.

Por otro lado, dada la menor riqueza semántica de la MS frente a la RCI, en el capítulo 4 se propuso un mecanismo alternativo basado en dos índices de similitud, de forma que podía expresarse de una manera más compacta tanto la modificación del consecuente según distintos criterios semánticos, como la condición de activación, emulándose aproximadamente el comportamiento de los principales modelos analizados bajo la RCI. Por eso nos limitaremos al análisis de la MS propuesta en esta tesis para los modelos representantes semánticos de los distintos comportamientos identificados.

Las medidas de la MS propuesta para los modelos de necesidad, que recordemos hemos denominada grado de verdad y certidumbre, son:

$$N(P/[P']) = \inf_{u/P'(u)=1} P(u)$$

$$N(JP/[P']) = \inf_{u/P(u)=0} 1-P'(u)$$

teniendo en cuenta que la segunda medida se puede reemplazar por la incertidumbre, cuya expresión es:

$$\Pi(\neg]P/[P']) = \sup_{u/P(u)=0} P'(u)$$

Para el primer grupo de modelos de posibilidad analizados basta con el cálculo de la medida de posibilidad habitual dada por:

$$\Pi_R^T(P/P') = \sup_u T(P'(u), P(u))$$

mientras que para los representantes semánticos propuestos las medidas a calcular son:

$$\Pi(JP'/[P]) = \sup_{u/P'(u)>0} P(u)$$

$$\Pi(JP/[P']) = \sup_{u/P(u)>0} P'(u)$$

$$\Pi_S^L(P'/P) = \Pi_R^L(P'/P) = \sup_u \max(0, P(u) + P'(u) - 1)$$

$$\Pi_r(JP'/[P]) = \sup_{u/P(u)+P'(u)>1} P(u)$$

$$\Pi_r(JP/[P']) = \sup_{u/P(u)+P'(u)>1} P'(u)$$

Recuérdese sin embargo que las medidas de similitud, aunque se basen en las expresiones anteriores, pueden en algunos casos ser más complejas ya que por ejemplo pueden encontrarse expresiones del tipo:

$$1 - I_R^{T^-}(N(P/[P']), 0)$$

$$1 - f^{f^{-1}}(f(0) - \Pi_R^{T^-}(P/P'))$$

Sin embargo de cara a obtener simplificaciones computacionales es suficiente con encontrar algún tipo de separabilidad de las medidas básicas señaladas.

Pasamos a analizar en primer lugar la separabilidad de la medida de incertidumbre $\Pi(\neg]P[/P')$. Partiendo de las siguientes expresiones:

$$\begin{aligned}\Pi(\neg]P_1[/P'_1) &= \sup_{u_1 / P_1(u_1)=0} P'_1(u_1) \\ \Pi(\neg]P_2[/P'_2) &= \sup_{u_2 / P_2(u_2)=0} P'_2(u_2)\end{aligned}$$

dada una regla con dos antecedentes, la medida de similitud viene dada por:

$$\begin{aligned}\Pi(\neg] T(P_1, P_2) [/ T(P'_1, P'_2)) = \\ \sup_{u_1, u_2 / T(P_1(u_1), P_2(u_2))=0} T(P'_1(u_1), P'_2(u_2))\end{aligned}$$

En el caso de t-normas positivas se verifica que $T(P_1(u_1), P_2(u_2))=0$ para $P_1(u)=0$ o bien $P_2(u)=0$, por lo que hay que obtener el supremo de la agregación $T(P'_1(u_1), P'_2(u_2))$ para aquellos pares (u_1, u_2) que anulen o bien a $P_1(u_1)$ o bien a $P_2(u_2)$. Esto significa que en este caso particular se verifica un cierto tipo de separabilidad, planteando el cálculo de la medida de similitud global como sigue:

$$\begin{aligned}\Pi(\neg] T(P_1, P_2) [/ T(P'_1, P'_2)) = \\ \sup(T(\sup_{u_1 / P_1(u_1)=0} P'_1(u_1), \sup_{u_2} P'_2(u_2)) , \\ T(\sup_{u_2 / P_2(u_2)=0} P'_2(u_2), \sup_{u_1} P'_1(u_1)))\end{aligned}$$

separabilidad que no se verifica para t-normas negativas.

En lo que respecta al grado de verdad se tiene:

$$\begin{aligned}N(T(P_1, P_2) [/ T(P'_1, P'_2)]) = \\ \inf_{u_1, u_2 / T(P_1(u_1), P_2(u_2))=1} T(P'_1(u_1), P'_2(u_2)) = \\ \inf_{u_1, u_2 / P_1(u_1)=1 \text{ y } P_2(u_2)=1} T(P'_1(u_1), P'_2(u_2)) = \\ T(\inf_{u_1 / P_1(u_1)=1} P'_1(u_1), \inf_{u_2 / P_2(u_2)=1} P'_2(u_2)) = \\ T(N(P_1/[P'_1]) , N(P_2/[P'_2]))\end{aligned}$$

por lo que $N(P/[P'])$ es una medida separable.

La separabilidad de la medida $\Pi_R^T(P/P') = \sup_u T(P'(u), P(u))$ cuando los antecedentes se agregan mediante la misma t-norma T queda como sigue ([Cross, Sudkamp, 93]):

$$\begin{aligned}\Pi_R^T(T(P_1, P_2) [/ T(P'_1, P'_2)]) = \\ \sup_{u_1, u_2} T(T(P'_1(u_1), P'_2(u_2)) , T(P_1(u_1), P_2(u_2))) = \\ \sup_{u_1, u_2} T(T(P'_1(u_1), P_1(u_1)) , T(P'_2(u_2), P_2(u_2))) = \\ T(\sup_{u_1} T(P'_1(u_1), P_1(u_1)) , \sup_{u_2} T(P'_2(u_2), P_2(u_2))) =\end{aligned}$$

$$T(\Pi_R^T(P_1/P'_1) , \Pi_R^T(P_2/P'_2))$$

por lo que en este caso la medida de posibilidad es separable. Sin embargo la separabilidad obtenida se apoya en la asociatividad de la t-norma que aparece en la medida, que es la t-norma generadora del modelo, que por tanto determina su semántica, y que en general no tiene por qué coincidir con la t-norma utilizada para agregar los antecedentes. En otros casos la separabilidad no puede garantizarse. En cualquier caso sí parece razonable, como se ha visto en la sección anterior, y como puede también verse en [Cross, Sudkamp, 93], que la t-norma utilizada para agregar observación y antecedentes sea la misma, aunque esto no es condición suficiente para la separabilidad.

Para la medida de certidumbre en el dominio de lo posible $\Pi(JP[/P')$ (simétrica respecto a sus argumentos) se tiene:

$$\Pi(]T(P_1, P_2)[/ T(P'_1, P'_2)) = \sup_{u_1, u_2 / T(P_1(u_1), P_2(u_2)) \neq 0} T(P'_1(u_1), P'_2(u_2))$$

para el caso en que la t-norma sea positiva $T(P_1(u_1), P_2(u_2)) \neq 0$ se da cuando tanto $P_1(u_1)$ como $P_2(u_2)$ son ambos distintos de 0, y puede entonces escribirse:

$$\begin{aligned} \Pi(]T(P_1, P_2)[/ T(P'_1, P'_2)) &= \\ \sup_{u_1, u_2 / P_1(u_1) \neq 0 \text{ y } P_2(u_2) \neq 0} T(P'_1(u_1), P'_2(u_2)) &= \\ T(\sup_{u_1 / P_1(u_1) \neq 0} P'_1(u_1) , \sup_{u_2 / P_2(u_2) \neq 0} P'_2(u_2)) &= \\ T(\Pi(JP_1[/P'_1) , \Pi(JP_2[/P'_2)) & \end{aligned}$$

por lo que para t-normas positivas la medida es separable. Sin embargo tampoco se verifica la separabilidad para t-normas negativas.

La medida $\Pi(JP'[/P)$ tiene un análisis idéntico al de la medida $\Pi(JP[/P)$ por lo que solamente se verifica su separabilidad para t-normas positivas.

La medida $\Pi_S^L(P'/P)$ no requiere un análisis propio puesto que puede también ponerse de la forma $\Pi_R^T(P'/P)$ que ya ha sido analizada, y que como hemos visto no es separable. Esto hace sugerente la idea de utilizar $\Pi_R^{T+}(P'/P)$ en su lugar par los modelos en los que interviene, aunque debe tenerse presente que esto modifica también las condiciones de activación del modelo en cuestión.

Nos quedan por último dos medidas dadas por:

$$\begin{aligned} \Pi_r(JP'[/P) &= \sup_{u/P(u)+P'(u) > 1} P(u) \\ \Pi_r(JP[/P') &= \sup_{u/P(u)+P'(u) > 1} P'(u) \end{aligned}$$

Para la primera de ellas se tiene:

$$\Pi_r(T(P_1, P_2) /]T(P'_1, P'_2)[) = \sup_{u_1, u_2 / T(P'_1(u_1), P'_2(u_2)) + T(P'_1(u_1), P'_2(u_2)) > 1} T(P_1(u_1), P_2(u_2))$$

mientras que para la agregación de cada una de las medidas:

$$T(\Pi_r(T(P_1) /](P'_1)[), \Pi_r(T(P_2) /](P'_2)[)) = T(\sup_{u_2 / P'_1(u_1) + P'_1(u_1) > 1} P_1(u_1), \sup_{u_2 / P'_2(u_2) + P'_2(u_2) > 1} P_2(u_2))$$

por lo que no parece evidente ninguna separabilidad.

Por lo tanto, si la rapidez computacional es un requisito restrictivo, pueden encontrarse casos no demasiado particulares para los que la separabilidad se verifica, y en los que por tanto la pérdida de riqueza semántica se vea fuertemente compensada por la eficiencia en los cálculos.

A modo de resumen podemos señalar que se han obtenido las siguientes condiciones de separabilidad:

$$\begin{aligned} & \Pi(-]T^+(P_1, P_2)[/ T^+(P'_1, P'_2)) = \\ & \quad \sup_{u_1 / P_1(u_1)=0} T^+(P'_1(u_1), \sup_{u_2} P'_2(u_2)), \\ & \quad T^+(\sup_{u_2 / P_2(u_2)=0} P'_2(u_2), \sup_{u_1} P'_1(u_1)) \\ & \Pi(-]T^-(P_1, P_2)[/ T^-(P'_1, P'_2)) \neq \\ & \quad \sup_{u_1 / P_1(u_1)=0} T^-(P'_1(u_1), \sup_{u_2} P'_2(u_2)), \\ & \quad T^-(\sup_{u_2 / P_2(u_2)=0} P'_2(u_2), \sup_{u_1} P'_1(u_1)) \\ & N(T(P_1, P_2) / [T(P'_1, P'_2)]) = T(N(P_1/[P'_1]) , N(P_2/[P'_2])) \\ & \Pi_R^T(T(P_1, P_2) / T(P'_1, P'_2)) = T(\Pi_R^T(P_1/P'_1) , \Pi_R^T(P_2/P'_2)) \\ & \Pi_R^{T1}(T_2(P_1, P_2) / T_2(P'_1, P'_2)) \neq T_2(\Pi_R^{T1}(P_1/P'_1) , \Pi_R^{T1}(P_2/P'_2)) \\ & \Pi(]T^+(P_1, P_2)[/ T^+(P'_1, P'_2)) = T^+(\Pi(]P_1/[P'_1] , \Pi(]P_2/[P'_2)) \\ & \Pi(]T^-(P_1, P_2)[/ T^-(P'_1, P'_2)) \neq T^-(\Pi(]P_1/[P'_1] , \Pi(]P_2/[P'_2)) \\ & \Pi(]T^+(P'_1, P'_2)[/ T^+(P_1, P_2)) = T^+(\Pi(]P'_1/[P_1] , \Pi(]P'_2/[P_2)) \\ & \Pi(]T^-(P'_1, P'_2)[/ T^-(P_1, P_2)) \neq T^-(\Pi(]P'_1/[P_1] , \Pi(]P'_2/[P_2)) \\ & \Pi_r(T(P_1, P_2) /]T(P'_1, P'_2)[) \neq T(\Pi_r(T(P_1) /](P'_1)[), \Pi_r(T(P_2) /](P'_2)[)) \\ & \Pi_r(T(P'_1, P'_2) /]T(P_1, P_2)[) \neq T(\Pi_r(T(P'_1) /](P_1)[), \Pi_r(T(P'_2) /](P_2)[)) \end{aligned}$$

Nótese por último que el problema de la separabilidad es similar al problema que se plantea cuando a partir de una regla $A, B \rightarrow C$ se buscan dos reglas $A \rightarrow C$ y $B \rightarrow C$ equivalentes, como se señala en [Lee, 90], [Lee, 90b] y [Cross, Sudkamp, 93] y no parece que exista una solución general.

5.5 La agregación conjuntiva de la observación y el condicional de una regla borrosa

Una aparente paradoja en el ámbito de la lógica borrosa es la compaginación del principio de mínima especificidad con la diversidad de modelos de reglas vistos en el capítulo anterior.

El principio de mínima especificidad establece que una agregación conjuntiva debe hacerse, en ausencia de información adicional sobre las distribuciones que se agregan, utilizando como operador conjuntivo el operador *min*, que es el que genera la distribución de posibilidad menos específica, y se aplica, como hemos visto a lo largo del capítulo 3, a la agregación de distribuciones de posible posibilidad, o a agregaciones semánticamente conjuntivas.

Por otro lado el estudio de los modelos de reglas, conjuntamente con la verificación del modus ponens estricto sugiere el uso de otras t-normas y operadores conjuntivos distintos de *min* para la agregación del condicional $I(P(u), Q(v))$ que representa a la regla, y la observación $P'(u)$ disponible para su activación. Esta paradoja se señala por ejemplo en [Dubois, Prade, 91].

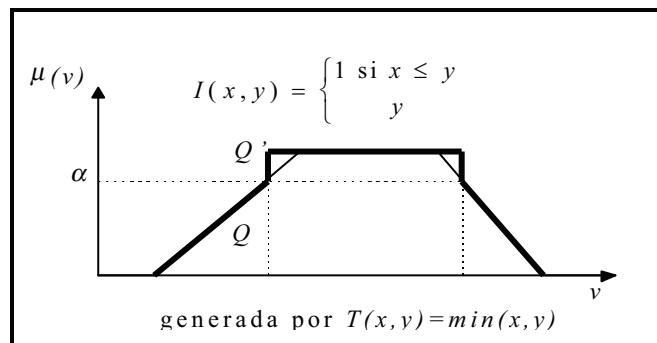
La aplicación rigurosa del principio de mínima especificidad exige el uso de la implicación residuada de la t-norma *min*, implicación de Godel, cuya expresión recordamos es:

$$I_R^{min}(x, y) = \begin{cases} 1 & \text{si } x \leq y \\ y & \text{en otro caso} \end{cases}$$

y que se obtiene residuando la t-norma *min* mediante la expresión:

$$I(x, y) = \{ \sup c \in [0, 1] / T(x, c) \leq y \} \text{ con } T = \text{min}$$

Nótese que en la expresión anterior, cuanto menor sea la t-norma *T* mayor puede ser el valor de *c* que la verifique, y por tanto mayor será la implicación resultante. La modificación lingüística llevada a cabo sobre el consecuente por la implicación de Godel, cuando la observación es puntual, queda representada en la siguiente figura:



El modelo generado por la t-norma *min* es, de entre los modelos (graduales) que transforman grado de verdad en grado de verdad e incertidumbre en incertidumbre ($N(Q/[Q'])=N(P/[P'])$ y $N(JQ/[Q'])=N(JP/[P'])$), el que genera la distribución condicional mayor, y la conclusión más específica. En efecto, el uso de cualquier t-norma continua menor que *min* genera implicaciones residuales mayores. Estas a su vez producen una modificación lingüística sobre el consecuente que da como resultado distribuciones también mayores que la obtenida con I_R^{min} , ante una misma observación. Así, cuanto más débil es la relación de implicación de u a v entonces menor debe ser la t-norma utilizada para obtener la implicación apropiada, y a la inversa. Un claro ejemplo es el uso de la implicación de Lukasiewicz, que produce como conclusiones distribuciones que incluyen a todos los elementos del universo de discurso en un cierto grado no nulo, debido a su nivel de indeterminación, en cuanto la hipótesis de la regla no se verifica con total exactitud. La implicación de Lukasiewicz permite modelar una relación de implicación entre las variables u y v bastante menos fuerte que la que se modelaría con la implicación I_R^{min} .

Por tanto el modelo por defecto es el modelo gradual más puro, en el sentido en que no tiene ningún matiz de certidumbre (a menos que la hipótesis sea incierta), como por ejemplo ocurre con el modelo basado en la implicación residuada de la t-norma *prod*, que sí que introduce una cierta incertidumbre en comparación con I_R^{min} , aunque ésta no se manifieste con un nivel de indeterminación en la conclusión y sí con grados de posibilidad en la conclusión mayores que con I_R^{min} .

Si se obtiene el condicional a partir de la implicación de Godel y se combina con la observación mediante una t-norma menor que *min*, se obtiene una distribución final también menor. Este efecto puede interpretarse como la existencia de una mayor relación entre las variables u y v ligadas por la regla. En general el uso de una FGMP menor permite modelar la existencia de una implicación de u a v más fuerte que la que por defecto vendría expresada por $I(P(u), Q(v))$ para una implicación I dada. De hecho una FGMP puede variar sensiblemente las condiciones de activación que son un reflejo del grado de implicación de u a v cuando las variables pertenecen a los conjuntos P y Q .

En general si se admite la existencia de diferentes semánticas para una regla $P(u) \rightarrow Q(v)$, entonces se asume la existencia de distintas relaciones entre las variables u y v compatibles con el concepto intuitivo de implicación. Los casos extremos son el uso de relaciones de necesidad y de relaciones de posibilidad. Dentro de las relaciones de necesidad, el uso de una t-norma T inferior a min en la ecuación de residuación puede interpretarse como la existencia de una relación de implicación más débil que la que por defecto vendría dada por la implicación de Godel. Una relación de implicación más débil hace que se concluyan distribuciones de necesaria posibilidad mayores (y por tanto menos específicas) que las obtenidas con la implicación de Godel.

El uso de FGMP no conmutativas como las pseudo-conjunciones, para residuar las implicaciones, permite obtener otras relaciones de implicación diferentes entre las variables u y v . Las distribuciones resultantes que se obtienen con este tipo de implicaciones vienen caracterizadas, cuando el modelo está generado por una t-norma positiva, por una estrecha coincidencia con el consecuente de la regla para valores de pertenencia altos y un distanciamiento para valores de pertenencia bajos, debido a la aparición de un nivel de indeterminación en la conclusión en cuanto la hipótesis de la regla no se verifica con exactitud. Estas implicaciones son características de un gran desconocimiento de la relación de implicación entre u e v en cuanto la hipótesis de la regla no se verifica, y como ya hemos visto son especialmente útiles para modelar la incertidumbre sin introducir ninguna componente de gradualidad en la relación entre u y v .

En el dominio de lo necesario la relación condicional, obtenida a partir de la implicación de Godel se calcula de forma que, para todos los posibles pares (u, v) , si $P(u) \leq Q(v)$ entonces $R(u, v) = 1$ y en cualquier otro caso $R(u, v) = Q(v)$, y es la máxima n-implicación compatible con la FGMP min .

Por otro lado la implicación $P \rightarrow Q$ puede considerarse posible si $P \rightarrow \neg Q$ no es necesaria, es decir si la implicación $P \rightarrow \neg Q$ no se verifica. La relación condicional que lo exprese expresará también que $P \rightarrow Q$ es posible. La implicación $P \rightarrow \neg Q$ da lugar a la desigualdad del modus ponens:

$$min(P(u), R(u, v)) \leq 1 - Q(v)$$

cuya solución es:

$$R(u, v) = \begin{cases} 1 - Q(v) & \text{si } P(u) > 1 - Q(v) \\ 1 & \text{en otro caso} \end{cases}$$

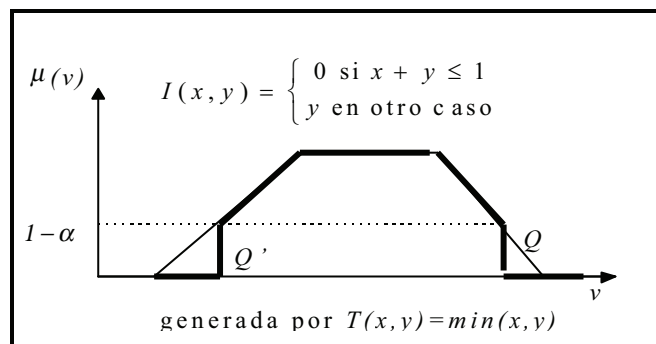
siendo la relación buscada la negación de la relación anterior, es decir:

$$R(u, v) = \begin{cases} Q(v) & \text{si } P(u) > 1 - Q(v) \\ 0 & \text{en otro caso} \end{cases}$$

que como puede verse se obtiene como $I(P(u), Q(v))$ siendo I la pseudo-conjunción que se obtiene de la t-norma *min*:

$$I(x, y) = \begin{cases} 0 & \text{si } x + y \leq 1 \\ y & \text{en otro caso} \end{cases}$$

que es la implicación de posibilidad gradual generada por la t-norma *min* ya analizada en capítulos anteriores. La modificación que produce en el consecuente viene reflejada en la siguiente figura:



Por tanto la opción por defecto en el dominio de lo posible es la implicación de posibilidad M_S^T y el modelo por defecto es el modelo (M_R^{\min}, M_S^{\min}) , analizado en el capítulo anterior, que es el modelo de posibilidad gradual más puro, en el sentido en que modifica el soporte del consecuente a partir de una medida del grado de verdad de la hipótesis, y su altura a partir de una medida de la certidumbre de la hipótesis. Esto sugiere la proximidad semántica de los modelos (M_R^T, I_R^T) con los modelos (M_R^T, M_S^T) que como se puede ver comparten misma FGMP, proximidad ya vista a lo largo del capítulo 4. En el apéndice de la tesis se propone una familia más general de funciones de implicación que engloba a las γ y p -implicaciones, y un procedimiento para la obtención de modelos intermedios a los de necesidad y posibilidad, a partir de un modelo de posibilidad y uno necesidad ambos con misma FGMP, como en este caso ocurre para los modelos (M_R^T, M_S^T) y (M_R^T, I_R^T) , lo que contribuye a resaltar su proximidad semántica.

5.6 Sistemas de reglas paralelas y particiones borrosas

Diremos que dos o más reglas son paralelas cuando la estructura de sus antecedentes y consecuentes sean semejantes, es decir tengan mismo número de antecedentes y hagan referencia a las mismas variables, y de igual modo para los consecuentes. Nos limitaremos sin pérdida de generalidad al estudio de dos reglas R_1 y R_2 con un antecedente y un consecuente, que denotaremos de forma abreviada por:

$$\begin{aligned} R_1: P_1 \rightarrow Q_1 \\ R_2: P_2 \rightarrow Q_2 \end{aligned}$$

donde P_1 y P_2 son conjuntos borrosos definidos para una variable u sobre un universo de discurso U , y Q_1 y Q_2 son conjuntos borrosos definidos para una variable v sobre un universo de discurso V . Hemos visto que para que las conclusiones de dos o más reglas puedan combinarse, las reglas deben estar modeladas con modelos del mismo tipo, es decir o bien modelos de necesidad, en cuyo caso las conclusiones son distribuciones de necesaria posibilidad y se agregan con operadores conjuntivos, o modelos de posibilidad en cuyo caso las conclusiones se interpretan como distribuciones de posible posibilidad y se agregan con operadores disyuntivos. Esto significa que el estudio de reglas paralelas debe restringirse a reglas con mismo tipo de modelo.

5.6.1 Inferencia con reglas paralelas y la RCI: inferencias FATI y FITA

Ya sea en el ámbito clásico como en el borroso, cuando hay reglas paralelas existen dos formas alternativas de llevar a cabo la inferencia. Dadas las dos reglas:

$$\begin{aligned} R_1: P_1 \rightarrow Q_1 \\ R_2: P_2 \rightarrow Q_2 \end{aligned}$$

con condicionales respectivos $R_1(u,v)$ y $R_2(u,v)$, la inferencia puede llevarse a cabo mediante el llamado método FATI ([Turksen, Tian, 93b]), “*First Aggregate Then Infer*”, que consiste en agregar primero los condicionales para obtener un único condicional del que después se infiere la conclusión, y el llamado FITA por otro, “*First Infer Then Aggregate*”, que consiste en agregar las conclusiones inferidas por cada uno de los condicionales por separado.

Si las reglas R_1 y R_2 tienen como modelos respectivos (M, I_1) y (M, I_2) , la conclusión obtenida por el método FITA viene dada por:

$$Q'(v) = H(\sup_u M(P'(u), I_1(P_1(u), Q_1(v))), \sup_u M(P'(u), I_2(P_2(u), Q_2(v)))) = H(Q'_1(v), Q'_2(v))$$

$$\begin{aligned} \text{con } Q'_1(v) &= \sup_u M(P'(u), I_1(P_1(u), Q_1(v))) \\ Q'_2(v) &= \sup_u M(P'(u), I_2(P_2(u), Q_2(v))) \end{aligned}$$

donde H es un operador conjuntivo cuando (M, I_1) y (M, I_2) son modelos de necesidad, y disyuntivo cuando son modelos de posibilidad. Para preservar el anonimato de las reglas el operador H debe ser conmutativo. Si además se asume la inexistencia de interactividad entre las reglas R_1 y R_2 los operadores a escoger son entonces la t-norma min y la t-conorma max . Si se conoce de la existencia de contradicciones entre conclusiones de reglas paralelas, susceptibles de dispararse simultáneamente, entonces puede ser interesante utilizar una t-norma menor o t-conorma mayor. Esto sin embargo obliga a tener identificadas en todo momento a las reglas y las conclusiones que producen, lo que puede complicar significativamente el mecanismo de inferencia, por lo que supondremos que los operadores empleados son min y max .

Cuando se utiliza el método FATI la conclusión viene dada por:

$$Q'(v) = \sup_u M(P'(u), H(I_1(P_1, Q_1(v)), I_2(P_2, Q_2(v)))) = \sup_u M(P'(u), R(u, v))$$

$$\text{con } R(u, v) = H(I_1(P_1, Q_1(v)), I_2(P_2, Q_2(v)))$$

donde H es un operador conjuntivo (preferentemente el operador min) cuando (M, I_1) y (M, I_2) son modelos de necesidad, y disyuntivo (preferentemente max) cuando son modelos de posibilidad. En [Dubois, Prade, 84] se justifica el porqué del uso de operadores conjuntivos para modelos de necesidad y disyuntivos para modelos de posibilidad desde una perspectiva axiomática, mientras que en [Dubois, Prade, 96] se hace desde un enfoque semántico, más próximo al adoptado en esta tesis.

Con el enfoque propuesto en esta tesis las únicas agregaciones posibles para condicionales de necesidad, son los operadores conjuntivos, y para los condicionales de posibilidad los disyuntivos. Según la regla de inferencia RCI tal y como se propone en el capítulo 3 todas las distribuciones involucradas deben ser del mismo tipo. Por otro lado en el capítulo 4 se muestra que las distribuciones de posibilidad concluidas mediante la RCI a partir de condicionales de necesidad pueden ser convenientemente interpretadas como distribuciones de necesaria posibilidad. Esto hace que los condicionales de necesidad se

puedan interpretarse como distribuciones condicionales de necesaria posibilidad, por lo que deben agregarse conjuntivamente. Del mismo modo puede decirse que los condicionales de posibilidad son distribuciones condicionales de posible posibilidad, y por tanto deben agregarse con operadores disyuntivos. En el caso clásico esta interpretación es todavía más directa. En efecto, como se muestra en [Dubois, Prade, 96] o también en [Trillas, 93], cualquier condicional clásico $R(u,v)$ para una regla $P(u) \rightarrow Q(v)$ verifica:

$$P(u) \cap Q(v) \leq R(u,v) \leq \neg P(u) \cup Q(v)$$

por lo que puede escribirse, en términos de distribuciones de posibilidad condicionadas:

$$\pi_p(v/u) = P(u) \cap Q(v) \leq \pi_n(v/u) = \neg P(u) \cup Q(v)$$

Esta desigualdad es la misma que existe entre una distribución de posible posibilidad y una de necesaria posibilidad en el caso general. Sin embargo en este caso $\pi_p(v/u) = P(u) \cap Q(v)$ es una cota inferior para el condicional R , ya que cualquier relación menor no verifica el modus ponens (de hecho en el ámbito clásico se puede considerar que $P(u) \cap Q(v)$ y $\neg P(u) \cup Q(v)$ son los únicos condicionales existentes). Su extensión al ámbito borroso es sencillamente una generalización, y se traduce en la elección de los condicionales apropiados dentro de la amplia familia existente, para poder modelar la semántica real de la regla, que en el ámbito clásico es única y sin matices.

El hecho de que pueda considerarse que un condicional tiene una cota inferior y otra superior ha dado lugar a enfoques basados en el uso de condicionales intervalo valorados, de forma que el condicional supuestamente desconocido se modela mediante una pareja de condicionales que lo acotan. A este respecto puede consultarse por [Turksen, Jiang, 93] donde se utiliza este tipo de técnicas. Sin embargo nos parece más apropiado relegar el uso de condicionales de posibilidad para aquellos casos en los que los de necesidad no permitan inferir conclusiones distintas de las triviales o suficientemente informativas, o no sea posible encontrar modelos de necesidad apropiados (debido por ejemplo a un conocimiento incompleto).

- Modelos de necesidad

Cuando en los modelos (M, I_1) y (M, I_2) las funciones de implicación I_1 e I_2 son n-implicaciones, entonces se verifica ([Dubois, Prade, 84]):

$$\begin{aligned} & \sup_u M(P'(u) , \min(I_1(P_1, Q_1(v)), I_2(P_2, Q_2(v)))) = \\ & \sup_u \min(M(P'(u), I_1(P_1, Q_1(v))) , M(P'(u), I_2(P_2, Q_2(v)))) \leq \\ & \min(\sup_u M(P'(u) , I_1(P_1(u), Q_1(v))) , \sup_u M(P'(u) , I_2(P_2(u), Q_2(v)))) \end{aligned}$$

lo que significa que en general la conclusión que se obtiene agregando las conclusiones de cada una de las reglas es menos informativa. Un caso especialmente llamativo es el que se da cuando la observación viene dada por $P' = P_1 \cup P_2$ ya que en este caso el método FATI permite concluir $Q_1 \cup Q_2$ mientras que el método FITA sólo proporciona en general la conclusión trivial $Q' = V$ (por ejemplo cuando se tiene $P_1 \cap P_2 = \emptyset$). Esto no es un problema específico del ámbito borroso, ya que se da igualmente cuando las distribuciones de posibilidad son clásicas, con la ventaja de que en este caso los cálculos resultan más sencillos y por tanto más ilustrativos del origen del problema.

En efecto, si modelamos las reglas mediante condicionales materiales y nos restringimos al ámbito clásico, para el método FATI, se tiene:

$$\begin{aligned} R_1 &= (P_1 \cap Q_1) \cup (\neg P_1 \cap V) \\ R_2 &= (P_2 \cap Q_2) \cup (\neg P_2 \cap V) \end{aligned}$$

y agregándolos con una disyunción tenemos (la negación de P la hemos denotado mediante \overline{P}):

$$\begin{aligned} R_1 \cap R_2 &= [(P_1 \cap Q_1) \cup (\overline{P_1} \cap V)] \cap [(P_2 \cap Q_2) \cup (\overline{P_2} \cap V)] = \\ & [(P_1 \cap P_2) \cap (Q_1 \cap Q_2)] \cup [(P_1 \cap \overline{P_2}) \cap Q_1] \cup [(\overline{P_1} \cap P_2) \cap Q_2] \cup [(\overline{P_1} \cap \overline{P_2}) \cap V] = \\ & [(P_1 \cap P_2) \cap (Q_1 \cap Q_2)] \cup [(P_1 \cap \overline{P_2}) \cap Q_1] \cup [(\overline{P_1} \cap P_2) \cap Q_2] \cup [(\overline{P_1} \cup \overline{P_2}) \cap V] \end{aligned}$$

En este tipo de expresiones de unión de intersecciones (suma de productos), cada intersección es una regla con condición de activación basada en la intersección del primer término con la observación, y la condición de activación global viene dada por términos del tipo (*expresión* $\cap V$). Por tanto la expresión obtenida permite deducir:

- Debido al término $(\overline{P_1 \cup P_2} \cap V)$ la condición de activación es $P' \cap (\overline{P_1 \cup P_2}) = 0$, que equivale a $P' \subset (P_1 \cup P_2)$.
- Si $P_1 \cap P_2 = \emptyset$ el término $(P_1 \cap P_2) \cap (Q_1 \cap Q_2)$ no se activa nunca y puede ser ignorado. En este caso $P' = P_1$ activa el segundo término y la conclusión es Q_1 . Del mismo modo de $P' = P_2$ se deduce Q_2 .
- Si $P_1 \cap P_2 \neq \emptyset$, entonces $P' = P_1$ activa los primer y segundo términos pero la conclusión sigue siendo Q_1 , y del mismo modo de $P' = P_2$ se deduce también Q_1 .
- Si $P_1 \cap P_2 \neq \emptyset$, entonces de $P' = P_1 \cap P_2$ se deduce $Q_1 \cap Q_2$.

- La observación $P' \subset (P_1 \cup P_2)$ produce la activación de los términos segundo y tercero, siempre que $(P_1 \cap \overline{P_2}) \neq 0$ y $(\overline{P_1} \cap P_2) \neq 0$, lo que exige $P_1 \neq P_2$, en cuyo caso se concluye $Q_1 \cup Q_2$. Esto se corresponde con un comportamiento intuitivamente razonable. Por ejemplo de las reglas “si el coche es caro entonces su velocidad es grande” y “si el coche es barato entonces su velocidad es pequeña” y de la observación “el coche es caro o barato” se deduce “la velocidad es grande o pequeña”.

Cuando por el contrario se utiliza el método FITA las conclusiones del condicional material se agregan mediante un operador conjuntivo, con lo que se tiene:

- La condición de activación es o bien $P' \subset P_1$ o bien $P' \subset P_2$, sin embargo $P' \subset (P_1 \cup P_2)$ no activa ninguna de las reglas, ya que cada regla se tiene que activar por separado. Esta condición es más restrictiva que la obtenida para la conjunción de los condicionales.
- Si $P_1 \cap P_2 = 0$ entonces las dos reglas no pueden activarse simultáneamente y por tanto de $P' = P_1$ se deduce Q_1 y de $P' = P_2$ se deduce Q_2 .
- Si $P_1 \cap P_2 \neq 0$ entonces las dos reglas pueden activarse simultáneamente. Si además existe una relación de inclusión $P_1 \subset P_2$ entonces de $P' = P_1$ se deduce $Q_1 \cap Q_2$. Sin embargo en general de $P' = P_1$ se deduce Q_1 , de $P' = P_2$ se deduce Q_2 y de $P_1 \cap P_2$ se deduce $Q_1 \cap Q_2$.
- En ningún caso puede obtenerse la conclusión $Q_1 \cup Q_2$ de la observación $P_1 \cup P_2$ como ocurriría con la conjunción de condicionales, a menos que exista una regla explícita que tenga como antecedente $P_1 \cup P_2$ y como consecuente $Q_1 \cup Q_2$. Esto significa por otro lado que es posible mantener la estructura de reglas independientes y tener un comportamiento similar al obtenido con la agregación de condicionales, si se añaden ciertas reglas adicionales (como la mencionada) a la base de conocimiento disponible.

Podemos por último comprobar que efectivamente la agregación disyuntiva de condicionales de necesidad presenta propiedades poco deseables. En efecto se tiene:

$$\begin{aligned} R_1 \cup R_2 &= [(P_1 \cap Q_1) \cup (\overline{P_1} \cap V)] \cup [(P_2 \cap Q_2) \cup (\overline{P_2} \cap V)] = \\ &= (P_1 \cap Q_1) \cup (P_2 \cap Q_2) \cup (\overline{P_1} \cap V) \cup (\overline{P_2} \cap V) = \\ &= (P_1 \cap Q_1) \cup (P_2 \cap Q_2) \cup (\overline{P_1 \cap P_2} \cap V) \end{aligned}$$

de donde se deduce:

- Debido al término $(\overline{P_1 \cap P_2} \cap V)$ la condición de activación del modelo es $P' \cap (\overline{P_1 \cap P_2}) = 0$, o lo que es lo mismo $P' \subset (P_1 \cap P_2)$. Esto significa que ni la observación $P'=P_1$ ni la observación $P'=P_2$ activan el modelo y por tanto sólo permiten obtener la conclusión trivial.
- Los términos primero y segundo podrían proporcionar un comportamiento razonable si no fuese por la condición de activación tan restrictiva que impone el tercer término.

Algunos autores ([Dubois, Prade, 91], [Turksen, Tian, 93a], [Martin-Clouaire, 89]) analizan el comportamiento de ambos mecanismos de inferencia y proponen algunas características razonablemente exigibles a los sistemas de reglas paralelas:

- generalización del modus ponens: De P_1 debe deducirse Q_1 y de P_2 debe deducirse Q_2 (véase [Turksen, Tian, 93a] y [Turksen, Tian, 93b]). Esto como hemos visto impone restricciones de inclusión o intersección entre los antecedentes P_1 y P_2 .
- modus ponens estricto frente a la unión de antecedentes: de $P_1 \cup P_2$ debe poder deducirse $Q_1 \cup Q_2$. El método FATI preserva este MPE, mientras que el método FITA no.
- modus ponens estricto frente a la intersección de antecedentes: de $P_1 \cap P_2$ debe poder deducirse $Q_1 \cap Q_2$. Este MPE es el menos problemático ya que si $P_1 \cap P_2 \neq 0$ ambos métodos de inferencia lo preservan.

El método de inferencia FATI plantea a nuestro entender dos problemas de difícil solución. En primer lugar la facilidad de enunciar el conocimiento en forma de reglas así como la posibilidad de utilizar las reglas de forma independiente unas de otras han sido siempre características importantes de los sistemas expertos basados en reglas. Esto se debe a la facilidad de disparar las reglas independientemente unas de otras, y a la posibilidad de justificar adecuadamente las conclusiones obtenidas en base a las reglas disparadas. Por otro lado la posible coexistencia de reglas paralelas con distintos modelos, y por tanto con distintas funciones generadoras de modus ponens, y distintas condiciones de activación, complica sensiblemente el análisis de como debería ser el método FATI para este caso, supuesto que fuese aplicable.

Si se opta por el método FITA se puede diseñar un algoritmo de inferencia más complejo que tenga también en cuenta observaciones del tipo $P' \subset P_1 \cup P_2$. A este respecto puede consultarse por ejemplo [Martin-Clouaire, 89], donde se propone un algoritmo aproximado de cálculo de conclusiones tal y como se obtendrían según el método FATI. Dado que los modelos de necesidad se disparan con distribuciones de necesidad, y que éstas habitualmente tienden a ser conjuntos convexos (por ejemplo como resultado de

agregaciones conjuntivas) la casuística se puede reducir al análisis de las relaciones de inclusión entre la observación P' y uniones de conjuntos adyacentes de la partición. Otra forma más directa podría consistir en generar automáticamente reglas del tipo:

$$P_1 \cup P_2 \rightarrow Q_1 \cup Q_2$$

para cada par de reglas paralelas R_1 y R_2 , con antecedentes P_1 y P_2 adyacentes. Sin embargo si el número de reglas paralelas es elevado, el número total de reglas resultante puede dispararse. Por otro lado nótese que este tipo de inferencia genera conclusiones que son unión de consecuentes y que por tanto tienden a ser conjuntos de gran tamaño, poco informativos. En el ejemplo anterior la conclusión “*la velocidad es grande o pequeña*” es escasamente informativa. Por último la pérdida de información que se tiene con observaciones del tipo $P' \subset P_1 \cup P_2$ cuando se utilizan modelos de necesidad se resuelve parcialmente si se usan modelos de posibilidad alternativos, para los que a costa de reducir la certeza de la conclusión se obtienen distribuciones más informativas. Esto vuelve a plantear una vez más la necesidad de encontrar modelos de necesidad y posibilidad semánticamente duales.

- Modelos de posibilidad

Cuando los modelos que se emplean son modelos de posibilidad, la regla de inferencia puede formularse como:

$$Q'(v) = \sup_u M(P'(u) , \max(I_1(P_1, Q_1(v)), I_1(P_1, Q_1(v)))) = \\ \sup_u \max(M(P'(u), I_1(P_1, Q_1(v))) , M(P'(u), I_1(P_1, Q_1(v)))) = \\ \max(\sup_u M(P'(u) , I_1(P_1(u), Q_1(v))) , \sup_u M(P'(u) , I_2(P_2(u), Q_2(v))))$$

Esto significa, a diferencia de lo que ocurre con los modelos de necesidad, que la conclusión que se obtiene agregando las conclusiones de cada una de las reglas es idéntica a la que se obtiene agregando las reglas primero y concluyendo después.

En efecto si modelamos las reglas mediante el producto cartesiano, para el método FATI, se tiene:

$$R_1 = (P_1 \cap Q_1) \\ R_2 = (P_2 \cap Q_2)$$

Su unión viene entonces dada por:

$$R_1 \cup R_2 = (P_1 \cap Q_1) \cup (P_2 \cap Q_2)$$

- en este caso la condición de activación viene dada por $P' \cap (P_1 \cup P_2) \neq 0$.
- de P_1 se deduce Q_1 y de P_2 se deduce Q_2 y si ambos intersecan se deduce $Q_1 \cup Q_2$.
- de $P_1 \cup P_2$ se deduce obviamente $Q_1 \cup Q_2$.

La intersección de ambas relaciones da como resultado:

$$R_1 \cap R_2 = (P_1 \cap Q_1) \cap (P_2 \cap Q_2)$$

- La condición de activación viene dada por $P' \cap (P_1 \cup P_2) \neq 0$.
- De $P' = P_1$ se concluye Q_1 y de $P' = P_2$ se concluye Q_2 siempre que $P_1 \cap P_2 = 0$, ya que en caso contrario (cuando la intersección no es nula) de $P' = P_1$ o de $P' = P_2$ se concluye $Q_1 \cap Q_2$, que es lo mismo que se concluye para $P' = P_1 \cup P_2$. Este tipo de conclusión no parece acorde con lo intuitivamente razonable.

5.6.2 Inferencia con reglas paralelas y la MS

Si las reglas R_1 y R_2 tienen como modelos respectivos (S_1, I_1) y (S_2, I_2) , la conclusión obtenida final obtenida mediante la MS viene dada por:

$$Q'(v) = H(I_1(S_1(P/P'), Q_1(v)), I_2(S_2(P/P'), Q_2(v))) = H(Q'_1(v), Q'_2(v))$$

$$\begin{aligned} \text{con } Q'_1(v) &= I_1(S_1(P/P'), Q_1(v)) \\ Q'_2(v) &= I_2(S_2(P/P'), Q_2(v)) \end{aligned}$$

donde nuevamente H es un operador conjuntivo cuando (S_1, I_1) y (S_2, I_2) son modelos de necesidad, y disyuntivo cuando son modelos de posibilidad. Supondremos que los operadores empleados son *min* y *max*.

En este caso no parece obvia a priori que posible alternativa podría existir puesto la distribución condicional no está explícita, como sí que ocurría con la RCI.

5.6.3 Particiones, completitud y coherencia de reglas paralelas

5.6.3.1 Particiones borrosas

Una partición borrosa de un universo de discurso U es un conjunto de n conjuntos borrosos definidos en U , $\mathcal{P}_U = \{P_1, P_2, P_3, \dots, P_n\}$ que dividen el universo de discurso en n categorías.

El grado de recubrimiento de una partición se define habitualmente como:

$$\inf_u \sum_i P_i(u)$$

Algunos autores exigen ciertas propiedades a \mathcal{P}_U para que sea una partición válida, entre las cuales destaca por ejemplo:

$$\forall u \in U, \exists P_i / P_i(u) > 0$$

propiedad relacionada con el grado de recubrimiento, de forma que cualquier elemento u de U puede ser catalogado como perteneciente a alguna de las categorías de \mathcal{P}_U . Una propiedad más restrictiva es:

$$\sum_i P_i(u) = 1$$

Se dice que dos conjuntos P_i y P_j se solapan cuando se verifica:

$$s = \sup_u \min(P_i(u), P_j(u)) > 0$$

y a s lo denominaremos grado de solape entre los conjuntos P_i y P_j .

En este apartado trataremos de analizar algunos de los requisitos más importantes en el diseño de un sistema de reglas paralelas, para determinar cómo deben ser las particiones borrosas utilizadas para su construcción. Nos limitaremos al método de inferencia FITA por las razones presentadas en el apartado anterior. Los principales requisitos que serán considerados en el diseño de una partición borrosa, para su uso en un sistema de reglas, son:

1. verificación del modus ponens estricto: cuando alguna de las observaciones coincida exactamente con los antecedentes de alguna regla, la conclusión obtenida debe ser el consecuente de dicha regla.
2. paso gradual de la activación de una regla a otra adyacente, de forma que cuando el grado de activación de una de las reglas va disminuyendo el de alguna regla adyacente debe ir aumentando. Esto confiere a los sistemas de reglas un comportamiento caracterizado por una variación suave de su salida ante variaciones suaves de sus entradas, con un cierto carácter interpolativo ([Dubois, Prade, 92]), aunque depende del tipo de modelo empleado.
3. completitud del sistema de reglas, de forma que no existan situaciones en las que no se produzca la activación de ninguna regla. Este requisito está estrechamente relacionado con el anterior.

Supondremos sin pérdida de generalidad que analizamos un sistema de reglas paralelas R_1 , R_2 y R_3 , con antecedentes respectivos P_1 , P_2 y P_3 .

5.6.3.2 Particiones borrosas, verificación del MPE y completitud del sistema de reglas

La verificación del MPE se traduce, como ya hemos comentado, en que cuando la observación es igual al antecedente de alguna regla R_i entonces la conclusión debe ser igual al consecuente de dicha regla R_i .

Una forma natural de garantizar este requisito consiste en garantizar, que bajo las condiciones señaladas, ninguna otra regla R_j se activa. Para ello basta tener en cuenta las distintas condiciones de activación obtenidas para los modelos, y diseñar la partición acorde con ellas.

Por otro lado se dice que un sistema de reglas es completo cuando es capaz de producir una conclusión no trivial ante cualquier entrada ([Pedrycz, 93]). Esta característica es especialmente importante en los sistemas de control (o en el modelado borroso), donde el modelo debe ser capaz de producir una acción de control (o una estimación) para cada posible estado de la planta (o de la entrada). Nótese que no se exige la activación total de una regla para cualquier entrada, basta con una activación parcial que permita obtener una solución no arbitraria. Nótese también que la completitud depende estrechamente de con qué distribuciones de posibilidad se modelan las entradas del sistema de reglas.

La completitud puede garantizarse si la partición y la observación son tales que siempre se produce la activación de al menos alguna de las reglas del sistema borroso. Esto se traduce para modelos de necesidad y posibilidad en restricciones en cuanto al solape y tamaño de los conjuntos de la partición. Para los modelos de necesidad interesa que la observación sea lo más precisa posible, mientras que los modelos de posibilidad presentan la característica contraria.

- Modelos de necesidad:

Para los modelos de necesidad estudiados, las condiciones de activación pueden resumirse como:

- si $[P'] \subset]P[$ entonces los:
 - modelos graduales (M_R^{T+}, I_R^{T+}) se activan
 - modelos mixtos (M_R^{T-}, I_R^{T-}) y (M_S^{T-}, I_S^{T-}) se activan
- si $]P' [\subset]P[$ entonces los:
 - modelos de certidumbre (M_S^{T+}, I_S^{T+}) se activan

La verificación del MPE exige que dos modelos no se activen simultáneamente cuando la observación es del tipo $P' = P_i$ siendo P_i alguno de los conjuntos de la partición. Esta condición puede endurecerse exigiendo la verificación del MP para observaciones más específicas, del tipo $P' \subset P_i$, siempre que estén normalizadas.

Modelos de necesidad graduales:

Para los modelos de necesidad graduales, es decir (M_R^{T+}, I_R^{T+}) , (M_R^{T-}, I_R^{T-}) y (M_S^{T-}, I_S^{T-}) , las particiones deben ser tales que se garantice que el núcleo de un conjunto de la partición no esté incluido en el soporte de ningún otro:

$$[P_i] \not\subset]P_j[\text{ para } i \neq j$$

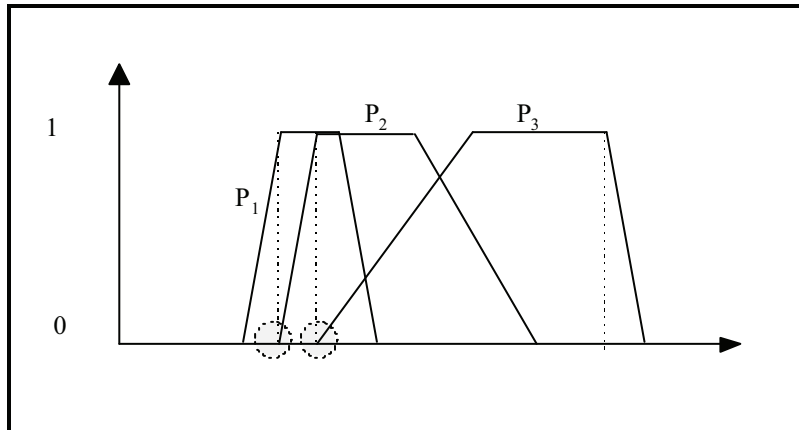
de forma que si $P' = P_i$ la regla R_i se active totalmente, mientras que ninguna regla R_j con $i \neq j$ se active siquiera parcialmente.

Nótese que si se quiere verificar el MP cuando la observación es puntual tiene que cumplirse:

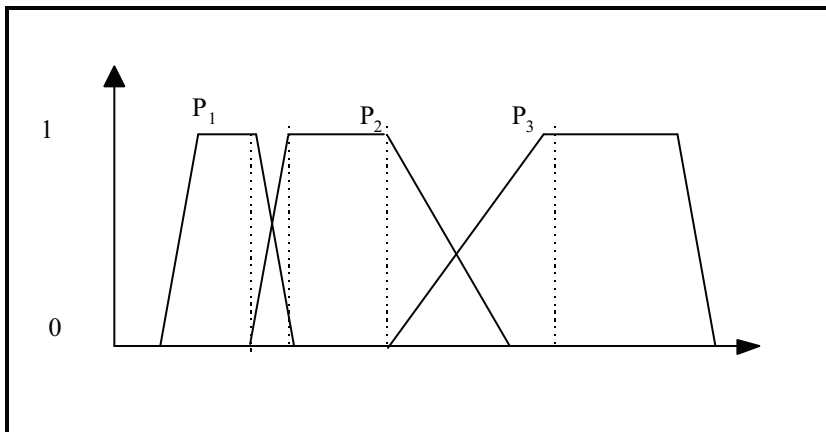
$$[P_i] \cap]P_j[= \emptyset \text{ para } i \neq j$$

De este modo si la observación puntual activa totalmente a una regla R_i , en cuyo caso pertenece al núcleo de P_i , entonces no debe activar ni siquiera parcialmente a ninguna otra, lo que exige que no pertenezca a ningún soporte $]P_j[$ con $i \neq j$.

Por tanto dados tres conjuntos P_1 , P_2 y P_3 fijos, la partición con solape máximo que verifique el MP para observaciones del tipo $P' = P_i$ es la que aparece en la siguiente figura:



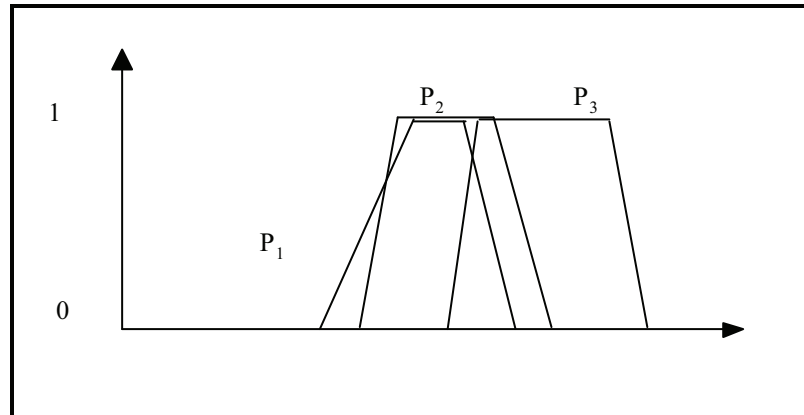
Mientras la partición con solape máximo que verifica el MP con observaciones puntuales, $P' = \{u_0\}$, daría lugar a la siguiente partición:



donde puede comprobarse que existen zonas de transición donde la observación no activa completamente a ninguno de los modelos.

Dado que en general interesa tener un solape alto para facilitar la completitud, el grado de solape puede variar entre los dos señalados, dependiendo del tamaño mínimo del núcleo de la distribución de posibilidad con el que se modele la observación P' . Para la verificación del MP dada una observación del tipo $P' \subset P_i$ el solape puede ser mayor

cuanto mayor sea el núcleo de P' , siendo los dos extremos las dos particiones propuestas en las figuras anteriores. En cualquier caso no sería válida la siguiente partición:

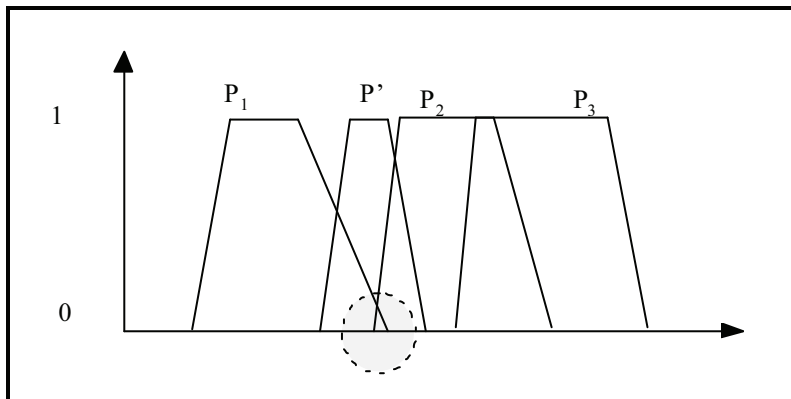


puesto que la inclusión del núcleo de P_1 en P_2 hace que una observación igual a P_1 active totalmente a R_1 y parcialmente a R_2 , obteniéndose como conclusión la intersección de las consecuencias de ambas reglas, que en general no producirá en resultado deseado.

La completitud exige que para cualquier posible observación alguna de las reglas del sistema se active al menos parcialmente. Cuanto mayor sea el solape entre conjuntos de la partición, mejor será la completitud del sistema, ya que esto contribuye a la activación de un mayor número de reglas, pero en general peor será la verificación del MP. Cuanto mayores puedan ser las distribuciones de posibilidad utilizadas para modelar la observación, mayores deberán ser los conjuntos de la partición y mayor el solape entre conjuntos adyacentes, para garantizar la completitud y la transición gradual de una regla a otra. Una primera condición necesaria para la completitud del sistema viene dada por:

$$\forall P', \exists i / [P'] \subset P_i$$

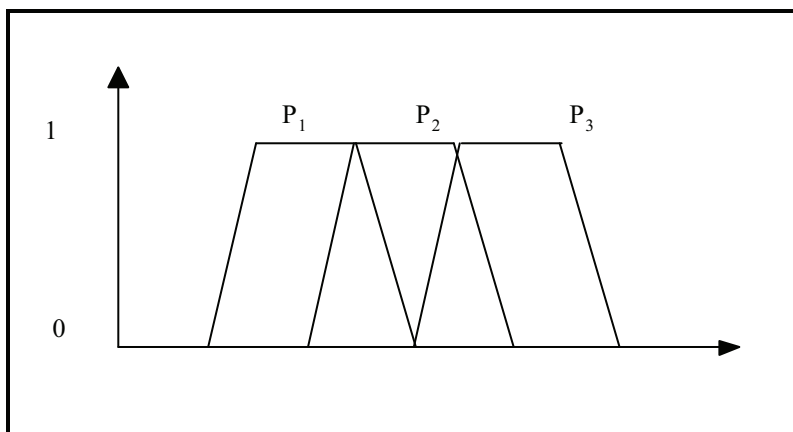
de forma que alguna regla del sistema se pueda activar, al menos parcialmente, para cualquier observación. Por otro lado es necesario exigir condiciones adicionales al solape de conjuntos, como ilustra la siguiente figura, en la que puede verse un ejemplo de cuándo la completitud no queda garantizada:



En efecto $[P'] \not\subset [P_1]$ y $[P'] \not\subset [P_2]$ hace que ni R_1 ni R_2 se activen, ni siquiera parcialmente. De hecho para que dadas dos reglas adyacentes, cuando una deje de activarse la otra comience a hacerlo debe verificarse la siguiente condición:

$$[P'] \subset ([P_i] \cap [P_j]) \text{ para } P_i \text{ y } P_j \text{ adyacentes}$$

Por tanto una buena partición obtenida a partir de conjuntos de igual tamaño y forma, y suponiendo que las observaciones vienen modeladas por conjuntos menores o iguales que los conjuntos de la partición, es la que aparece en la siguiente figura, ya que tanto el MPE como la completitud viene garantizadas:



La verificación del MPE no asegura sin embargo la verificación del MP para observaciones estrictamente menores que los conjuntos de la partición. En el caso en que los conjuntos de la partición tengan tamaños diferentes el requisito de completitud recomienda que el solape sea el máximo posible, pero de tal modo que se verifique el MPE.

Modelos de necesidad de certidumbre:

Para los modelos de necesidad de certidumbre, dado que tienen una condición de activación más restrictiva, la condición de verificación del MPE viene dada por:

$$]P_i[\not\subset]P_j[\text{ para } i \neq j$$

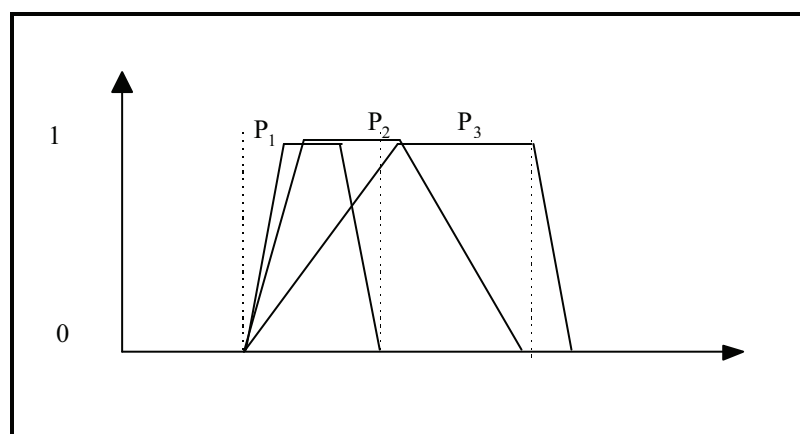
de forma que si $P' = P_i$ la regla R_i se activa totalmente, mientras que ninguna regla R_j con $i \neq j$ se activa siquiera parcialmente. En este caso los núcleos de los conjuntos de la partición no están sujetos a ningún requisito.

Si se quiere verificar el MP cuando la observación es puntual tiene que cumplirse, al igual que ocurría con los modelos de necesidad graduales:

$$[P_i] \cap]P_j[= \emptyset \text{ para } i \neq j$$

para que si la observación puntual activa totalmente a una regla R_i no active ni siquiera parcialmente a ninguna otra, lo que exige, al igual que con los modelos anteriores, que no pertenezca a ningún soporte $]P_j[$ con $i \neq j$. La similitud con el caso anterior se debe a que en el caso de informaciones puntuales su soporte y núcleo coinciden, y por tanto las condiciones $[P'] \subset]P[$ y $]P' [\subset]P[$ son equivalentes.

Por tanto dados tres conjuntos P_1, P_2 y P_3 fijos, la partición con solape máximo, esto es para observaciones del tipo $P' = P_i$, y que verifica el MP, es la que aparece en la siguiente figura (nótese que es el límite máximo de solape que se corresponde con la situación en la que los soportes de unos conjuntos están justamente incluidos en los de otros, y que por tanto, con todo rigor se estaría garantizando el MPE debido a la activación de dos reglas):



Para observaciones puntuales, la partición con máximo solape que verifica el MP es la misma que la que se obtenía para los modelos graduales.

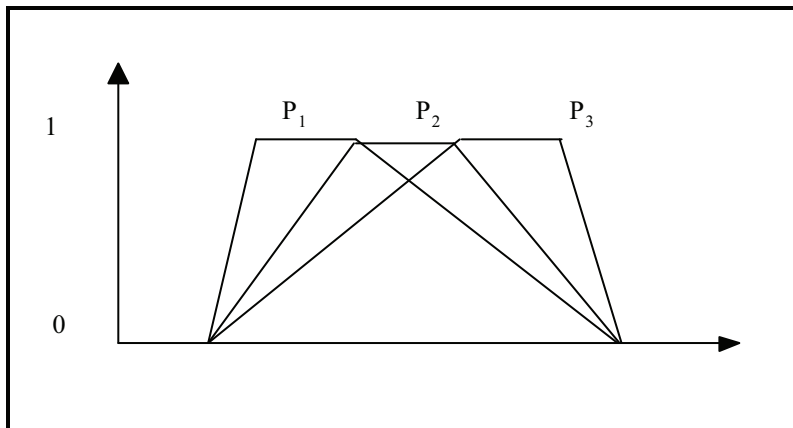
En lo que respecta a la completitud, al igual que ocurría en el caso anterior, cuanto mayor sea el solape entre conjuntos de la partición, mejor será la completitud del sistema. En este caso una primera condición indispensable viene dada por:

$$\forall P', \exists i /]P'[\subset]P_i[$$

Esta condición es ya muy restrictiva puesto que limita considerablemente más que antes el tamaño máximo que pueda tener la observación. Además, para que cuando una regla deje de activarse y una adyacente comience a hacerlo debe también cumplirse:

$$]P'[\subset (]P_i[\cap]P_j[) \text{ para } P_i \text{ y } P_j \text{ adyacentes}$$

que es también una condición mucho más restrictiva que antes. Se deduce por tanto que para modelos de certeza la completitud y la verificación del MPE, esto es para observaciones con mismo tamaño que los conjuntos de la partición, conducen a una curiosa partición en la que todos los conjuntos tienen el mismo soporte, y que además no pueden tener misma forma y tamaño. En efecto la condición $]P'[\subset]P_i[$ conduce a $]P_i[\subset (]P_i[\cap]P_j[)$ lo que obliga a $]P_i[\subset]P_j[$, con lo que todos los conjuntos de la partición tienen en mismo soporte. Esto conduce a una partición como la de la siguiente figura (que es nuevamente un caso límite):



Nótese que la completitud del sistema depende fuertemente del modelo que se utilice para la observación, pero también del tamaño que pueda llegar a tener la observación. En general para los modelos de necesidad es interesante modelar la observación con conjuntos borrosos menores que los conjuntos de la partición de forma que cualquiera de los modelos pueda activarse. De hecho cuanto menor sea mejor es la completitud del sistema ya que más fácil es garantizar las condiciones de activación de estos modelos, basados en la inclusión de la observación en la hipótesis. Así pues en sistemas de control o

aproximación funcional basados en modelos de necesidad es recomendable modelar las observaciones con distribuciones de posibilidad pequeñas (es decir más precisas), a diferencia de lo que ocurre como veremos a continuación para los modelos de posibilidad.

- Modelos de posibilidad:

Para los modelos de posibilidad las condiciones de activación quedan resumidas como sigue:

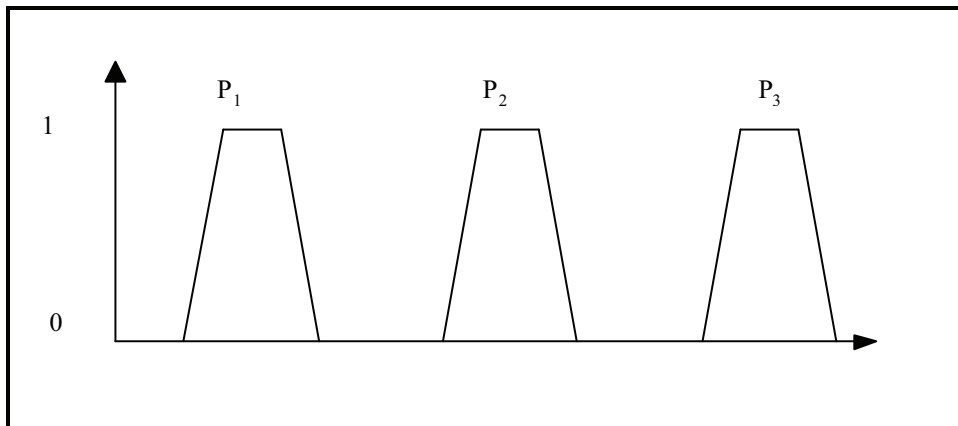
- si $]P'[\cap]P[\neq 0$ entonces:
 - los modelos de certidumbre (M_R^{T+}, M_R^{T+}) se activan
 - los modelos graduales (M_S^{T+}, M_S^{T+}) y (M_R^{T+}, M_S^{T+}) se activan
- si $f(P'(u)) + f(P(u)) < f(0)$ entonces:
 - los modelos mixtos (M_R^{T-}, M_R^{T-}) y (M_S^{T-}, M_S^{T-}) se activan
- si $P(u) + P'(u) > 1$ entonces:
 - los modelos mixtos (M_S^{T-}, M_R^{T-}) se activan
 - los modelos de certidumbre (M_S^{T+}, M_R^{T+}) se activan
- si $\sup_u M_R^{T-}(P'(u), M_S^{T-}(P'(u), 1)) > 0$ entonces:
 - los modelos mixtos (M_R^{T-}, M_S^{T-}) se activan

Modelos graduales y de certidumbre:

La condición $]P'[\cap]P[\neq 0$ correspondiente a la de los modelos de posibilidad graduales y uno de los de certidumbre es la menos restrictiva. En base a esta condición de activación, se garantiza el cumplimiento del MPE si la partición es tal que la intersección de dos conjuntos de la partición es nula:

$$]P_i[\cap]P_j[= \emptyset \text{ para } i \neq j$$

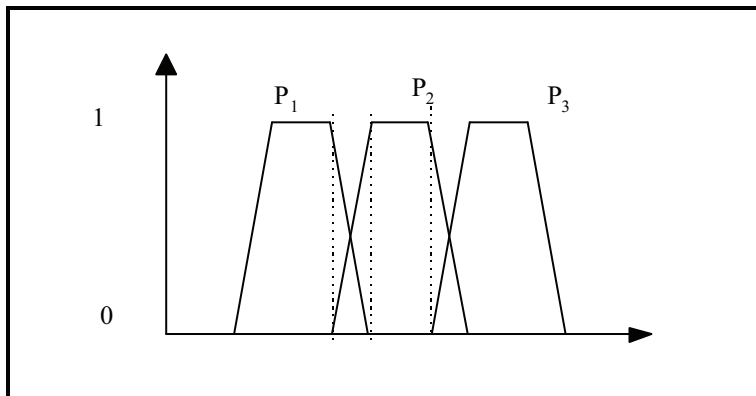
En efecto en caso contrario, si la observación coincide con alguno de los conjuntos de las particiones, entonces siempre se produce la activación total de una de las reglas y parcial de alguna otra, con lo que no se puede garantizar el cumplimiento del MPE. Con particiones generadas con idénticos conjuntos e iguales a la observación se obtiene entonces:



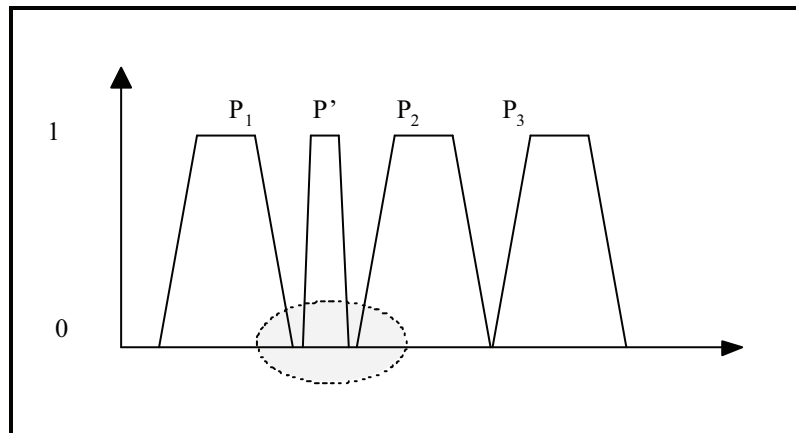
La verificación del MP para observaciones puntuales permite aumentar el solape de la partición, ya que la condición es ahora:

$$[P_i] \cap]P_j[= \emptyset \text{ para } i \neq j$$

de forma que la activación total de una regla no pueda darse simultáneamente con la activación parcial de alguna otra. En este caso la partición con máximo solape posible aparece en la siguiente figura:



La completitud nuevamente exige condiciones adicionales al solape. Así por ejemplo en el caso de la siguiente figura ésta no se puede garantizar:



Una condición necesaria para la completitud viene dada por:

$$\forall P', \exists i /]P'[\cap]P_i[\neq 0$$

de forma que al menos alguna regla del sistema se active. En general la completitud está garantizada si se verifica:

$$]P'[\subset (]\neg P_i[\cap]\neg P_j[)$$

cuando se produce la transición de una regla a otra.

Modelos de posibilidad mixtos:

Para estos modelos se tiene varias condiciones de activación relativamente similares, que vienen dadas por:

$$\begin{aligned} f(P'(u))+f(P(u)) &< f(0) \\ P'(u)+P(u) &> 1 \\ \sup_u M_R^{T-}(P'(u), M_S^{T-}(P'(u), 1)) &> 0 \end{aligned}$$

y se pueden hacer consideraciones similares a las del apartado anterior, aunque particularizadas para cada una de las condiciones anteriores.

Por ejemplo para la primera condición la verificación del MPE exige:

$$\forall u / f(P_i(u))+f(P_j(u)) > f(0)$$

de forma que cuando la observación es $P'=P_i$ no se produzca la activación de R_j .

Si se quiere garantizar el MP para observaciones puntuales entonces la condición se convierte en:

$$\forall u / f([P_i](u)) + f(P_j(u)) > f(0)$$

Dependiendo de la función f pueden obtenerse las particiones con conjuntos de misma forma y tamaño y con máximo solape, que verifique MPE con observaciones modeladas con conjuntos de misma forma y tamaño que los de la partición, o el MP con observaciones puntuales. Por último señalar que la completitud quedaría garantizada si para cualquier observación $P'(u)$ se verifica:

$$\exists i / f(P'(u)) + f(P_i(u)) < f(0) \text{ para algún valor de } u.$$

5.6.3.3 Coherencia de un sistema de reglas paralelas

Para los modelos de necesidad, donde las distribuciones resultantes son distribuciones de necesaria posibilidad y que por tanto deben ser normales, la coherencia de un sistema de reglas se traduce en que para cualquier entrada la conclusión final debe ser normal. La pérdida de normalidad se debe a que reglas con conclusiones contradictorias se activen simultáneamente y la intersección de sus conclusiones tenga un núcleo vacío. Sobre la coherencia de reglas paralelas puede consultarse algunos donde esta se estudia en profundidad, por ejemplo [Dubois, Prade, Ughetto, 96] y [Dubois, Prade, 96].

La coherencia puede estudiarse teniendo en cuenta las condiciones de activación y el tipo de observación que pueda darse. Para reglas con modelos de necesidad, la observación que más reglas dispara es la observación puntual. Así dos reglas $P_i \rightarrow Q_i$ y $P_j \rightarrow Q_j$ no serán coherentes si, susceptibles de dispararse simultáneamente, sus conclusiones son contradictorias. Si se modelan con modelos de necesidad graduales y mixtos, entonces la condición de activación de una regla $P \rightarrow Q$ ante una observación P' viene dada por $[P'] \subset [P]$ por lo que serán coherentes si:

$$\text{si }]P_i[\cap]P_j[\neq \emptyset \text{ entonces } [Q_i] \cap [Q_j]$$

para cualquier pareja de reglas de la base de conocimiento. Nótese que dado que en general Q'_i y Q'_j tienen respectivamente un núcleo mayor que Q_i y Q_j la condición impuesta puede ser excesivamente restrictiva, aunque sí es suficiente. Dado que el núcleo de los consecuentes se ve agrandado en base al grado de verdad de la hipótesis dada la observación, $N(P/[P'])$, dependiendo de las pendientes relativas de Q_i y Q_j , se pueden establecer condiciones necesarias y suficientes, y por tanto menos restrictivas, en términos de los α -cortes de Q_i y Q_j (véase [Dubois, Prade, 96]).

Cuando los modelos empleados son modelos de certidumbre, la condición de activación viene dada por $]P'[\subset]P[$ lo que significa que si se sigue asumiendo que la observación puede ser puntual, entonces se sigue verificando la misma condición para la coherencia de las reglas. Además en este caso el núcleo de los consecuentes de las reglas no se ve modificado, por lo que para garantizar la coherencia es condición necesaria y suficiente la que antes era suficiente para los modelos graduales y mixtos.

La coherencia de un sistema de reglas modeladas con modelos de posibilidad plantea problemas desde su definición. Por ejemplo en [Dubois, Prade, Ughetto, 96] y [Dubois, Prade, 96] se dice que dado que las conclusiones de estos modelos son en general no normales, no puede exigirse la condición de normalidad para demostrar su coherencia, y que su estudio no permite determinar si las reglas con modelos de posibilidad son o no coherentes. Por otro lado en el ámbito del control y de la aproximación numérica suele asumirse que dos reglas son coherentes si cuando se disparan simultáneamente, la distribución obtenida es unimodal, de forma que no existen conflictos aparentes para llevar a cabo la desborrosificación. Este criterio es arbitrario en el sentido en que es perfectamente admisible obtener distribuciones multimodales, en cuyo caso todos aquellos elementos con grados de pertenencia altos son posibles soluciones finales. Es el método de desborrosificación el que en este caso es poco adecuado al tipo de distribución obtenida.

Las reglas con modelos de posibilidad se disparan con mayor facilidad que con modelos de necesidad, existiendo distribuciones capaces de disparar a todas las reglas del sistema, por lo que los análisis relacionados con los tipos de distribuciones resultantes se complican tanto computacional como semánticamente. Estas y otras razones que serán presentadas en el apartado de este mismo capítulo donde se estudia la dualidad de condicionales, sugieren que los sistemas de reglas deben basarse siempre (en la medida de lo posible) en modelos de necesidad, y que las propiedades de completitud y coherencia así como la bondad de las particiones deben estudiarse con este tipo de modelos. De hecho veremos por ejemplo que intentar garantizar el MPE con modelos de posibilidad conduce a particiones con comportamientos paradójicos que enmascaran otras propiedades interesantes de los modelos de posibilidad, y que sólo quedan patentes cuando se verifica el MPE con las mismas reglas pero con modelos de necesidad.

5.7 Modelos de necesidad y posibilidad semánticamente duales

Un problema que ha venido apareciendo repetidamente a lo largo del capítulo, y que consideramos de gran importancia para la automatización de un sistema de razonamiento flexible, consiste en determinar para una regla modelada con un modelo de necesidad determinado, cuál debería ser el modelo de posibilidad a emplear, si se quisiesen relajar las condiciones de activación de la regla en detrimento de la fiabilidad de la conclusión. Diremos que estos modelos son semánticamente duales. En este apartado intentaremos determinar qué modelos de posibilidad pueden ser considerados semánticamente duales de los modelos de necesidad.

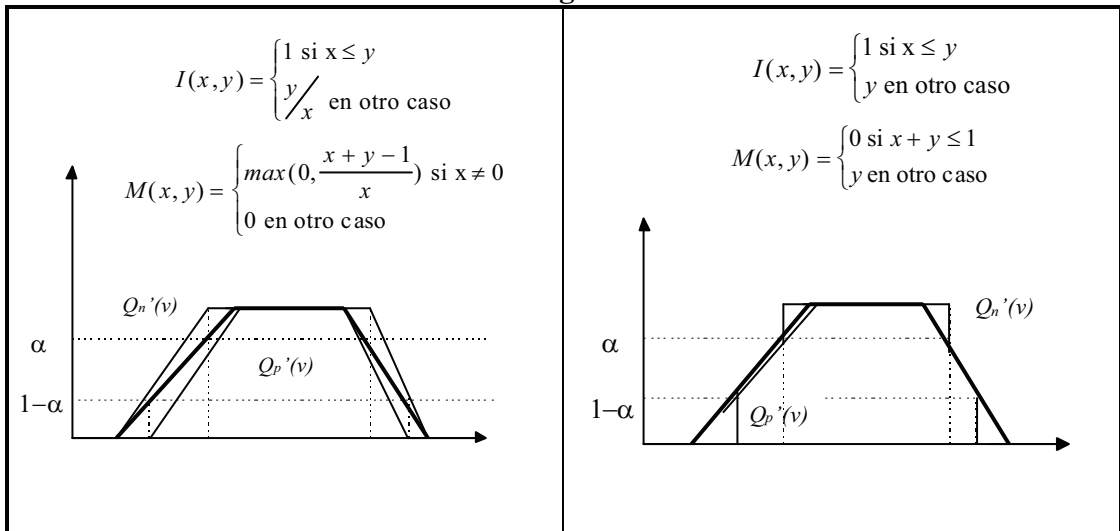
Conviene recordar que en todo momento se ha asumido que las variables son disyuntivas, en el sentido en que sólo pueden tomar un valor. El uso de variables conjuntivas (por ejemplo el idioma que una persona puede hablar, que es una variable conjuntiva ya que puede tomar simultáneamente varios valores como son *{francés, español}*) lleva a un análisis semántico diferente. Así ocurre que los modelos de posibilidad graduales pueden tener una interpretación más directa para este tipo de variables, como puede consultarse en [Dubois, Prade, 96]. Además las variables conjuntivas han muy sido poco tratadas en la literatura ([Dubois, Prade, 92b]).

La primera consideración que cabe hacer es la existencia de una gran proximidad semántica entre las implicaciones que hemos denominado duales, en la que nos basaremos como punto de partida. Creemos que esta proximidad semántica es la base para determinar qué modelos pueden considerarse semánticamente duales.

5.7.1 Modelos graduales

Las siguientes figuras reflejan el comportamiento de las principales implicaciones duales con semántica gradual cuando la observación es puntual. En este caso la FGMP no interviene en la inferencia, y la RCI es equivalente a la MS, por lo que el comportamiento de las implicaciones es directamente el comportamiento de los modelos:

Modelos graduales



Los modelos de necesidad graduales basan su comportamiento en la exclusión de valores del universo como posibles para la variable consecuyente. Análogamente los modelos de posibilidad graduales basan su comportamiento en la inclusión de valores del universo como posibles valores para la variable consecuyente.

Supongamos que un elemento v_0 del universo V es considerado posible en grado α por el consecuyente de una regla. Si se usa un modelo de necesidad gradual, a medida que el grado de verdad de la hipótesis disminuye, el grado de posibilidad del elemento aumenta, ya que no puede ser excluido con un grado tan elevado como α , mientras que para el modelo de posibilidad gradual, a medida que el posible grado de verdad disminuye, el grado de posibilidad del elemento disminuye, ya que no puede ser incluido con un grado tan elevado como α .

Por otro lado, recordando lo que ocurre para el encadenamiento de reglas cuando se emplea el representante semántico de necesidad gradual:

$$N(G/[Q]) \geq N(G/[Q']) \text{ función de } N(P/[P'])$$

$$N(JG/[Q]) \geq N(JG/[Q']) \text{ función de } N(JP/[P'])$$

y para el representante de posibilidad gradual se tiene:

$$\Pi(G/[Q]) \leq \Pi(G/[Q']) \text{ función de } \Pi(P/[P'])$$

$$\Pi(JG/[Q]) \leq \Pi(JG/[Q']) \text{ función de } \Pi(JP/[P'])$$

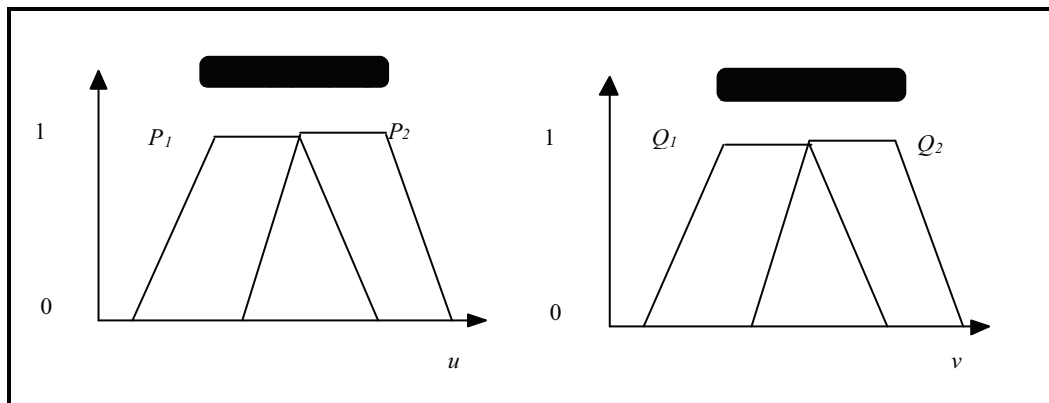
con lo que ambos modelos propagan de manera similar grados de verdad y de certidumbre parciales.

Por otro lado el análisis del comportamiento semántico se suele analizar a menudo examinando las reglas por separado, lo que conduce a interpretaciones que pueden resultar en algunos casos paradójicas. Esto lleva por ejemplo a menospreciar la utilidad de las pseudo-conjunciones como modelos de posibilidad para variables disyuntivas, como puede verse en [Dubois, Prade, Ughetto, 96]. Sin embargo su estudio desde la perspectiva global que confiere el estudio de un sistema de varias reglas paralelas conduce a resultados más coherentes e interesantes. Supongamos que tenemos dos reglas paralelas:

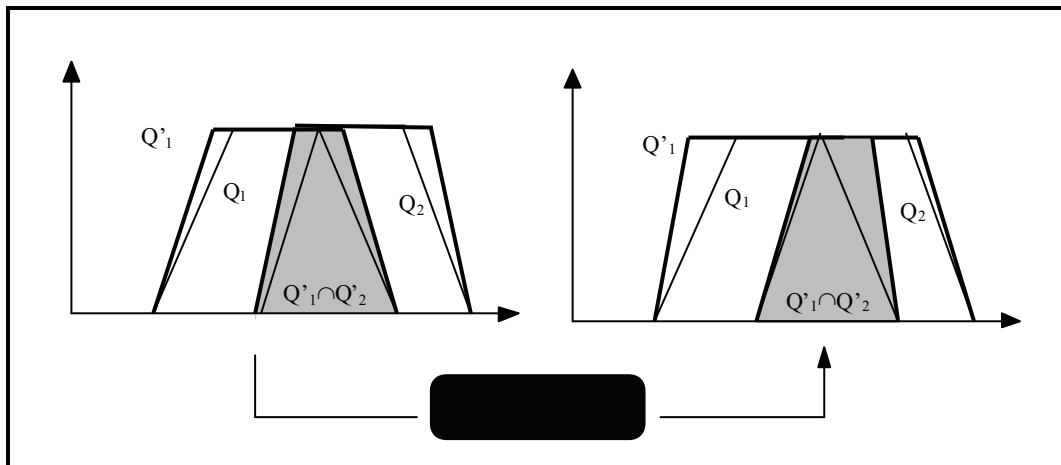
$$P_1 \rightarrow Q_1$$

$$P_2 \rightarrow Q_1$$

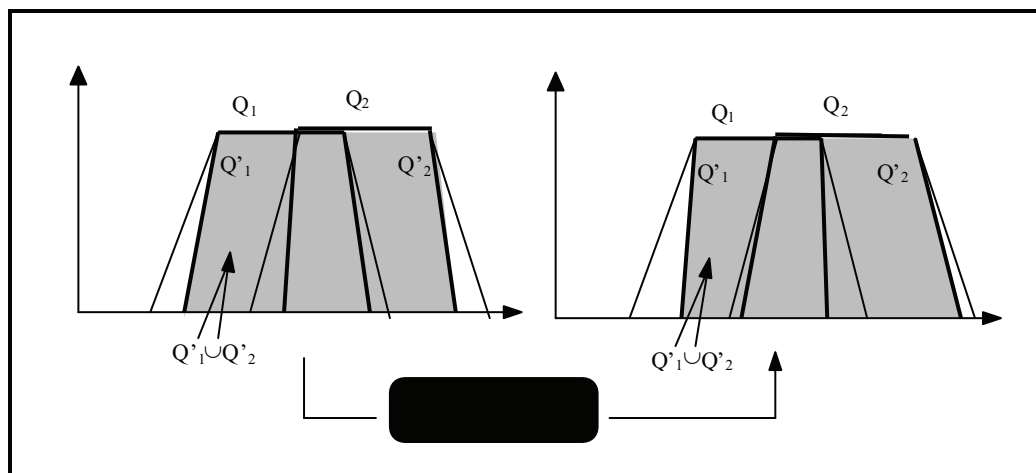
generadas a partir de las particiones que aparecen en las siguientes figuras:



Cuando utilizamos modelos de necesidad graduales, a medida que el grado de activación de una regla se reduce, el grado de activación de la regla adyacente aumenta, y como conclusión se obtiene una distribución normalizada que se desplaza sobre el universo de la variable consecuyente. La siguiente figura ilustra gráficamente la distribución obtenida cuando el grado de activación de una de las reglas decrece y el de la otra crece, para una observación puntual:

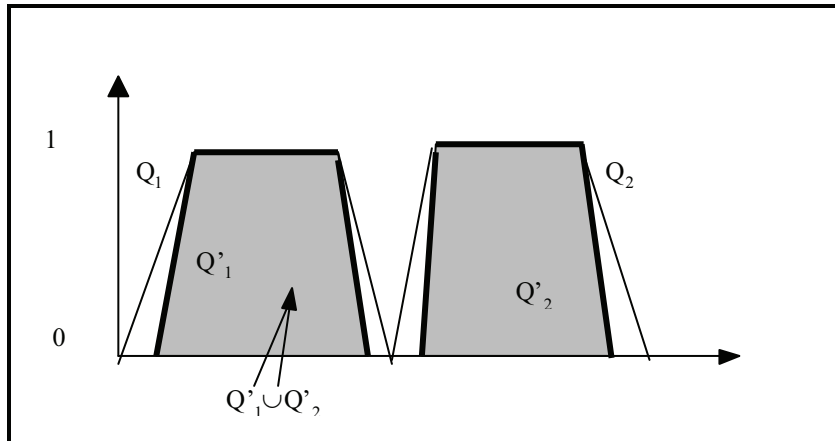


Para los modelos de posibilidad graduales el resultado es también similar (en contra de lo que a primera vista podría parecer), obteniéndose también una distribución normalizada que se desplaza sobre el universo de discurso de la variable consecuyente. Si utilizamos las mismas particiones anteriores la siguiente figura ilustra el resultado obtenido cuando el grado de activación de una de las reglas disminuye y aumenta el de la otra, ante observación es puntual:



Utilizando una misma partición para ambos tipos de modelos, el resultado obtenido para ambos ha sido también similar. La mayor diferencia viene motivada por el hecho de que el núcleo de la conclusión, sobre el universo de discurso del consecuyente, se desplaza de forma más brusca para los modelos de posibilidad que para los de necesidad, lo que puede resolverse parcialmente con particiones con conjuntos con núcleos pequeños (aunque es conveniente que exista un cierto solape entre núcleos de conjuntos adyacentes). Para los modelos de necesidad graduales, la gradualidad se pierde a medida que el solape entre los soportes de los conjuntos de la partición disminuye.

Cuando la partición se diseña para garantizar completitud y MPE para los modelos de posibilidad graduales, es cuando su comportamiento resulta particularmente insatisfactorio, como se aprecia en la siguiente figura.



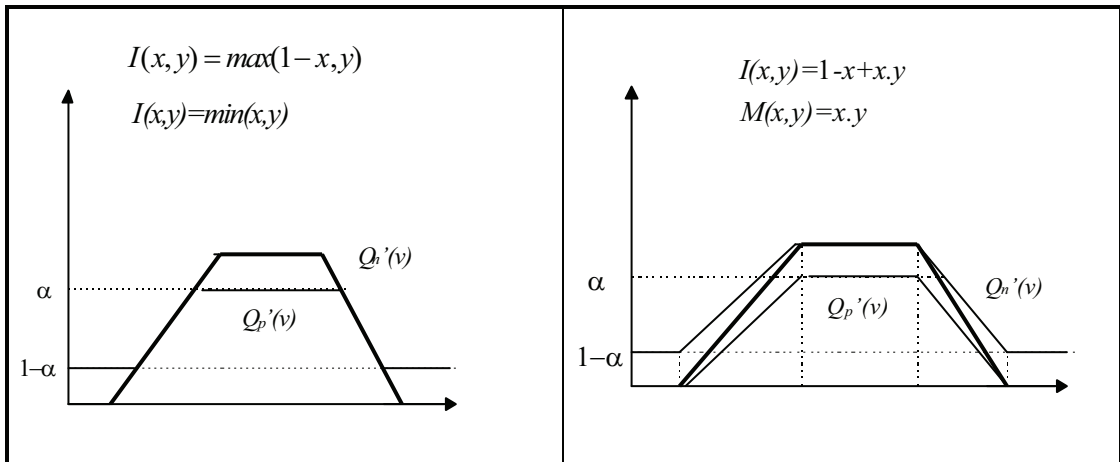
Este resultado sugiere la conveniencia de manejar los modelos de posibilidad graduales con las particiones diseñadas para modelos graduales para maximizar el solape.

Supuesto que en un sistema de reglas con modelos de necesidad graduales el paso a modelos de posibilidad graduales se hace cuando los primeros son incapaces de inferir una conclusión distinta de la trivial (o suficientemente informativa), cuando además la partición empleada y las reglas del sistema son las mismas en ambos casos, el resultado obtenido parece ser coherente, lo que corrobora la propuesta que se hace en esta tesis de sustituir de forma automática los modelos de necesidad por modelos de posibilidad cuando los anteriores no permiten obtener una conclusión satisfactoria, sin necesidad alguna de modificar las reglas de la base de conocimiento. Desde este punto de vista los modelos de necesidad y posibilidad graduales parecen ser modelos semánticamente duales, y su comportamiento es en cierto sentido óptimo cuando las particiones utilizadas para enunciar las reglas se diseñan para los modelos de necesidad, especialmente se persigue un buen comportamiento numérico.

5.7.2 Modelos de certidumbre

Las siguientes figuras muestran la modificación lingüística llevada a cabo por las implicaciones con semántica de certidumbre:

Modelos de certidumbre



El comportamiento de los modelos de certidumbre se basa en introducir un nivel de incertidumbre que afecta por igual a todos los elementos del universo de discurso, para ambos tipos de modelos, comportamientos que parecen a nuestro entender semánticamente coherentes. Así ocurre que en ambos casos la verificación parcial de la condición de activación del modelo se traduce en la aparición de un cierto grado de incertidumbre en la distribución resultante, ya sea en el dominio de lo necesario o en el de lo posible.

En cuanto al encadenamiento podemos recordar que para los modelos de necesidad de certidumbre se tiene:

$$N(G/[Q]) = N(G/[Q'])$$

$$N(JG[/Q]) \geq N(JG[/Q']) \text{ función de } N(P/[P'])$$

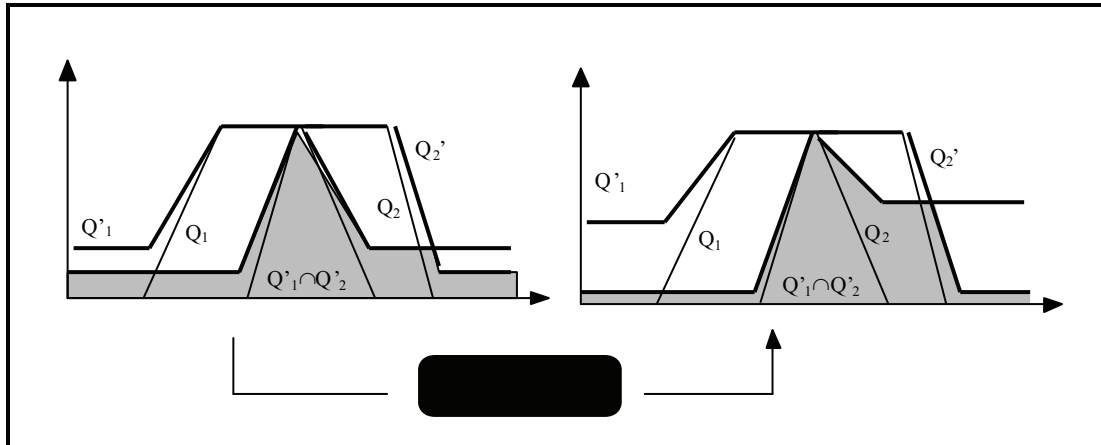
y para los modelos de posibilidad de certidumbre:

$$\Pi(G/[Q]) \leq \Pi(G/[Q']) \text{ función de } \Pi_r(JP'/[P])$$

$$\Pi(JG[/Q]) \leq \Pi(JG[/Q']) \text{ función de } \Pi_r(JP/[P'])$$

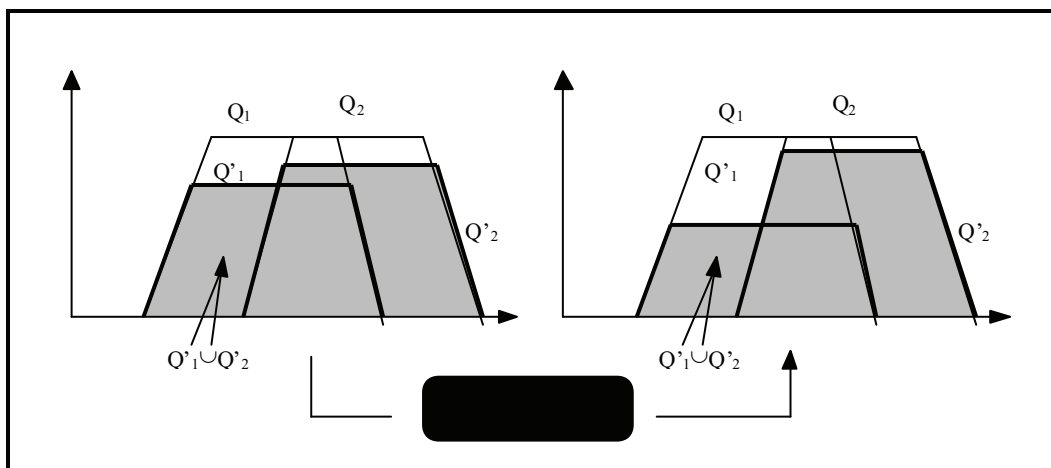
de forma que nuevamente ambos propagan de forma similar grados de verdad y certidumbre parciales. Para ambos modelos la reducción de certidumbre de la conclusión se debe a un grado de verdad parcial de la hipótesis, aunque el modelo de posibilidad propaga la certidumbre parcial dando lugar a una reducción del soporte de la conclusión. En cualquier caso cuanto menor es el grado de verdad de la hipótesis, mayor es la incertidumbre en la conclusión.

Por otro lado, siguiendo el mismo razonamiento que el llevado a cabo para los modelos graduales, la conclusión obtenida para modelos de necesidad de certidumbre, para observación puntual, viene representada en la siguiente figura:



En este caso la obtención de una conclusión normalizada exige un solape en la partición mayor que el de una partición para modelos de necesidad graduales. La conclusión que se obtiene presenta, siempre que las reglas no se activen totalmente, un cierto nivel de incertidumbre que se reduce a medida que alguna de las reglas aumenta su grado de activación. La distribución resultante se desplaza de forma más brusca sobre su universo que como lo hacía la obtenida con modelos graduales.

Cuando para las mismas reglas y particiones empleamos los modelos de posibilidad de certidumbre el resultado obtenido queda reflejado en la siguiente figura:



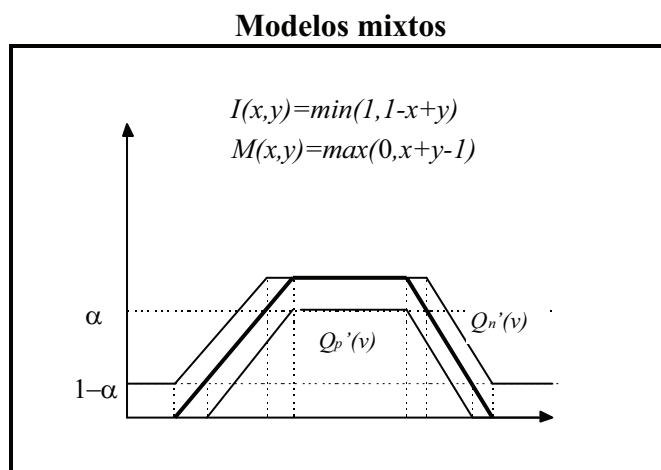
La distribución que se obtiene es una distribución que generalmente no está normalizada y que presenta un comportamiento muy similar al obtenido con modelos de necesidad pero

en el ámbito de lo posible. El uso de los modelos de posibilidad certidumbre con particiones diseñadas para los modelos de necesidad de certidumbre hace que las distribuciones resultantes sean unimodales, lo que en circunstancias particulares puede considerarse como un síntoma de coherencia de las reglas que forman la base de conocimiento. Como hemos visto estas inconsistencias pueden atribuirse en algunos casos a un diseño inapropiado de la partición utilizada. En vista de la similitud entre ambos, parece razonable suponer que los modelos de posibilidad y necesidad de certidumbre pueden considerarse semánticamente duales.

Para los dos tipos de modelos, tanto graduales como de certidumbre, hemos visto que podemos considerar que modelos basados en implicaciones duales pueden considerarse semánticamente duales, especialmente cuando la partición se diseña para el modelo de necesidad. Esto hace que para los modelos de posibilidad se pierda la verificación del MP, pero esta condición parece mucho menos interesante que la coherencia de las distribuciones resultantes, más si se tiene en cuenta que las reglas de la base de conocimiento deberían formularse en términos de necesidad, para los que se debería cumplirse el MPE, mientras que el uso de los modelos de posibilidad es en cierto modo un mal menor que la ausencia total de conclusiones, peor no un objetivo en sí mismo.

5.7.3 Modelos mixtos

Para los modelos mixtos son aplicables los resultados y comentarios de los modelos graduales y de certidumbre. La siguiente figura muestra la modificación lingüística llevada a cabo sobre el consecuente por este tipo de modelos:



En este caso la distribución resultante presenta características comunes a las obtenidas con modelos graduales y de certidumbre. Además de desplazarse sobre el universo de discurso de la variable consecuente, presenta en general un nivel de indeterminación que varía de forma gradual a medida que una regla reduce su grado de activación y otra adyacente lo aumenta.

En cuanto al encadenamiento de este tipo de reglas se tiene para los modelos de necesidad mixtos:

$$N(G/[Q]) \geq N(G/[Q']) \text{ función de } N(P/[P'])$$

$$N(JG[/Q]) \geq N(JG[/Q') \text{ función de } N(JP[/P') \text{ y de } N(P/[P'])$$

y para los modelos de posibilidad mixtos se puede de forma aproximada escribir:

$$\Pi(G/[Q]) \leq \Pi(G/[Q']) \text{ función de } \Pi_S^T(P/P')$$

$$\Pi(JG[/Q]) \leq \Pi(JG[/Q') \text{ función de } \Pi_S^T(P/P')$$

por lo que en este caso tanto grados de verdad parciales como certidumbres parciales se propagan como incertidumbres parciales en el universo del consecuente.

El problema de la dualidad semántica, a nuestro entender complejo, no queda del todo resuelto con las consideraciones hasta ahora hechas. Por una lado parece claro que implicaciones duales presentan comportamientos semánticamente muy parejos, lo que contribuye a pensar que modelos semánticamente duales deben basarse en implicaciones duales. Por otro lado mientras que el comportamiento de los modelos de necesidad parece más claro (en gran parte debido a las simplificaciones realizadas para llevar a cabo su estudio), el de los de posibilidad, más abundantes en número, es un comportamiento que, aunque interpretable en los mismos términos, parece presentar más diversidad de medidas o al menos una mayor complejidad a la hora de proceder a su simplificación. Veremos en el capítulo 6 que existen ciertas diferencias significativas entre ambos tipos de modelos, que hacen más difícilmente interpretables los modelos de posibilidad. Esto hace que no sea obvio el detectar comportamientos totalmente paralelos en el dominio de lo necesario y lo posible lo que complica su análisis. En el apéndice de la tesis se propone una visión unificada de los condicionales, los φ -condicionales, que podría ser una buena vía de investigación para profundizar más en los distintos tipos de condicionales y sus posibles similitudes semánticas.

5.8 Sistemas de reglas encadenadas

Existen dos aspectos importantes en torno al encadenamiento de reglas. Por un lado, con un interés principalmente teórico, está el estudio de la equivalencia de dos reglas encadenadas, con una única regla más sencilla. El interés es teórico puesto que como se ha señalado en otros apartados es preferible mantener las reglas individualmente en la base de conocimiento, incluso a pesar de la pérdida de información que puede acarrear (como en el caso de reglas paralelas, véase el punto 5.6) ya que esto facilita su disparo y el seguimiento de la activación de cada una de ellas, para poder entender la línea de razonamiento llevada a cabo por el sistema.

Otro de los problemas asociados con reglas encadenadas es la determinación del orden de activación de las mismas. En efecto un orden incorrecto puede conducir a utilizar distribuciones de posibilidad todavía indeterminadas para activar modelos de reglas a destiempo. Propondremos para ello un sencillo procedimiento que permita determinar el orden de activación correcto de un grupo de reglas, clasificándolas en distintos niveles y generando un árbol de reglas.

5.8.1 Encadenamiento de reglas y transitividad de condicionales

Dadas dos reglas encadenadas R_1 y R_2 :

$$R_1: A \rightarrow B$$

$$R_2: B \rightarrow C$$

una propiedad aparentemente deseable es que su equivalencia con una tercera regla $R: A \rightarrow C$. Esta propiedad está estrechamente relacionada con la M -transitividad de la implicación I del modelo (M, I) empleado para las reglas respecto a su FGMP, como se muestra en [Dubois, Prade, 84]. Sin embargo su análisis se complica considerablemente si combinamos reglas con modelos distintos.

El problema del encadenamiento se ha estudiado tradicionalmente para reglas modeladas con el mismo modelo a través de la M -transitividad de la implicación I , por lo que pierde validez según el enfoque propuesto en esta tesis, en el que se potencia la coexistencia de distintos modelos para modelar distintas semánticas.

Por otro lado el conocimiento de las condiciones de activación de modelos permite abordar el problema del encadenamiento desde una perspectiva más práctica. Así por

ejemplo es fácil comprobar que si las dos reglas encadenadas están modeladas con modelos de certidumbre con condición de activación $]P'[\subset]P[$, la activación parcial de la primera regla garantiza la no activación de la segunda debido al nivel de indeterminación de la conclusión de la primera. Por otro lado conocidas las condiciones de activación de unos y otros modelos, el diseñador puede elegirlos de forma consciente, por lo que la no equivalencia de dos reglas encadenadas con una tercera se convierte en una elección voluntaria. Por todo ello no dedicaremos más tiempo a este punto.

5.8.2 Orden de activación de reglas en un sistema borroso

Una de las características de los sistemas de reglas borrosos, a diferencia de los compuestos de reglas clásicas, consiste en que el orden de activación de las reglas debe ser cuidadosamente establecido, ya que pueden en caso contrario obtenerse conclusiones erróneas.

Un ejemplo sencillo puede ilustrar este problema. Supongamos que se dispone de las reglas:

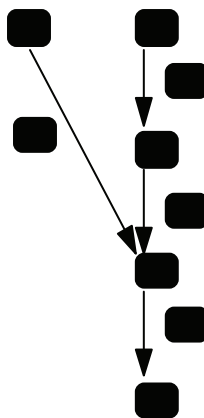
$$R_1: A \rightarrow B$$

$$R_2: C \rightarrow D$$

$$R_3: D \rightarrow B$$

$$R_4: B \rightarrow E$$

donde los conjuntos borrosos A , B , C , D y E están definidos sobre distintos universos de discurso y en los que están definidas las variables a , b , c , d , y e . Gráficamente estas reglas se puede representar como sigue:



Si las reglas se disparan en el orden R_1, R_2, \dots la conclusión obtenida será errónea, ya que la información utilizada para disparar la regla R_4 es incompleta al no haberse tenido en cuenta la información procedente de la regla R_3 : en efecto la distribución de posibilidad asociada a la variable b está todavía por determinar.

El orden correcto de activación hubiese sido R_1, R_2, R_3 y R_4 , de forma que la regla R_4 no se activa hasta que haya sido agregada toda la información disponible para la variable b , esto es las conclusiones de las reglas R_1 y R_3 . Para obtener el orden de activación correcto para un grupo cualquiera de reglas proponemos el siguiente algoritmo de creación de un árbol de reglas.

La determinación del orden correcto de activación de reglas se basa en la clasificación de las variables del sistema en niveles. Así por ejemplo las variables a y c pertenecerían al primer nivel, la variable d al segundo, la variable b al tercero y la variable e al cuarto. Esto permite también clasificar las reglas por niveles, con las siguientes características:

- El orden de activación de las reglas de un mismo nivel es indiferente.
- Las reglas de un nivel no deben activarse hasta que no se hayan activado todas las reglas de los niveles anteriores.
- Conocidos los niveles de reglas, es posible determinar en cada instante qué reglas pueden activarse en cada instante.

La clasificación de las variables y reglas del sistema por niveles puede hacerse con el siguiente procedimiento:

- Obtención de los niveles de variables:
 - Obtención del primer nivel de variables: consiste en la obtención de todas las variables que aparecen en antecedentes de reglas y no aparecen en ningún consecuente. Diremos que estas variables pertenecen al nivel 0 de variables.
 - Obtención del nivel de variables $k+1$ a partir de las variables del nivel k : se trata de obtener todas las reglas que se activan con las variables del nivel k o anteriores, de forma que todas aquellas variables que aparezcan en los consecuentes de dichas reglas son las variables del nivel $k+1$.

- Obtención del último nivel: El último nivel N de variables se obtiene cuando a partir del mismo no somos capaces de encontrar ninguna regla que tenga a dichas variables por antecedentes.
- Purga de los niveles de variables: la purga de los niveles provisionales de variables consiste en eliminar del nivel k , todas aquellas variables que también estén en cualquier nivel $k-i$ anterior (i entero tal que $0 \leq i \leq k$). De esta forma una variable solo pertenece al último nivel en el que ha aparecido durante la formación de los niveles provisionales.
- Obtención de los niveles de reglas: diremos ahora que una regla pertenece al nivel k si se activa con todas las variables de los niveles k y anteriores, pero debe ser necesaria al menos una variable del nivel k para su activación. De esta forma garantizamos que una regla solo pertenece a un único nivel. Podemos decir por tanto que el nivel de una regla es el último nivel de variables con el cual la regla se activa parcialmente. Por tanto, la regla se activará totalmente con las variables de su nivel y niveles anteriores, y parcialmente con las de su nivel.

A modo de ejemplo supongamos que tenemos que crear el árbol de reglas a partir de las siguientes reglas:

$$\begin{aligned}
 R_1: A \rightarrow B \\
 R_2: A \rightarrow C \\
 R_3: A, B \rightarrow D \\
 R_4: A, C \rightarrow E \\
 R_5: C \rightarrow D \\
 R_6: D \rightarrow E \\
 R_7: B \rightarrow F \\
 R_8: A \rightarrow E \\
 R_9: E \rightarrow H
 \end{aligned}$$

Los distintos niveles obtenidos para variables y reglas son, aplicando el procedimiento anterior:

	niveles provisionales de variables	niveles definitivos de variables	niveles definitivos de reglas
nivel 0	a	a	R_1, R_2, R_8
nivel 1	b, c, e	b, c	R_3, R_4, R_5, R_7
nivel 2	f, d, e, h	f, d	R_6
nivel 3	e	e	R_9
nivel 4	h	h	

6. Valores de verdad borrosos

Los valores de verdad borrosos son una generalización al ámbito de la lógica borrosa de los valores de verdad binarios de la lógica clásica, y permiten modelar simultáneamente la existencia de grados de verdad parciales, así como la incertidumbre en el conocimiento exacto de los mismos. El valor de verdad borroso de una hipótesis dada una observación se puede calcular mediante la llamada compatibilidad borrosa introducida por Zadeh, que resulta ser una medida borrosa de la inclusión de la observación en la hipótesis. Dependiendo del tipo de distribución con el que esté modelada la observación, la compatibilidad borrosa da lugar a distribuciones de posible posibilidad o necesaria posibilidad, distribuciones que hemos denominado respectivamente valores de posibilidad y valores de necesidad borrosos, y que deben ser interpretados de distinto modo. Los valores de verdad borrosos también pueden ser utilizados como modificadores lingüísticos, en términos de posibilidad y necesidad, lo que a su vez da lugar a dos tipos de modificadores que hemos denominado modificadores de necesidad y modificadores de posibilidad. En este capítulo repasaremos algunas definiciones y propiedades básicas de los valores de verdad borrosos y los modificadores, y los interpretaremos en términos de grados de verdad y grados de certidumbre. Finalmente propondremos una familia general de valores de verdad borrosos basada en las funciones de implicación generales introducidas en el capítulo 4 para su uso en sistemas borrosos y estudiaremos su agregación.

6.1 Organización del capítulo

En primer lugar se definen los valores de verdad borrosos y se introducen algunos de los problemas que plantea su uso, y más concretamente el uso de los que tradicionalmente se han propuesto y utilizado en la literatura, debido principalmente a que se basan en una elección arbitraria de su geometría, lo que hace más compleja su interpretación.

En el siguiente apartado se repasa la obtención de valores de verdad borrosos a partir de la compatibilidad borrosa entre predicados, y se introducen los valores de posibilidad y necesidad borrosos, basados en los conceptos de distribuciones de posible y necesaria posibilidad, así como los modificadores de posibilidad y de necesidad, que como se verá en posteriores apartados, permiten cualificar proposiciones borrosas tanto en el dominio de lo posible como en el dominio de lo necesario.

En los dos puntos siguientes se relacionan los valores de verdad borrosos con las principales medidas de grados de verdad y certidumbre obtenidas a lo largo del capítulo 4 de la tesis, y con las medidas generalizadas de posibilidad y necesidad también introducidas en el capítulo 4.

El siguiente apartado está dedicado al análisis de la inferencia con valores de verdad borrosos, en donde se analiza cómo los distintos modelos transforman el valor de verdad borroso de la hipótesis para generar un modificador, que aplicado al consecuente de una regla, genera una nueva distribución con un comportamiento coherente con la semántica del modelo.

A continuación se propone una familia general de valores de verdad, suficientemente rica desde un punto de vista semántico, para ser utilizados como modificadores lingüísticos, y que permita cualificar proposiciones borrosas en términos grados de verdad y grados de certidumbre parciales, en los dominios de lo posible y de lo necesario.

En último lugar se analiza el resultado que produce la agregación de valores de verdad borrosos en los grados de verdad y grados de certidumbre parciales, de cara a profundizar en su comportamiento y a proponer una posible simplificación semánticamente coherente que facilite su agregación.

6.2 Definición

Los valores de verdad borrosos son conjuntos borrosos definidos sobre el universo de discurso $X=[0,1]$ donde cada número real $x \in X$ se interpreta como un grado de verdad parcial. Si a una proposición se le asocia un grado de verdad x , ésta será más verdadera cuanto más próximo sea x de 1, y más falsa cuanto más próximo sea de 0. Asociar a una proposición un valor de verdad en $[0,1]$ permite modelar la existencia de grados de verdad parciales. Asociar a una proposición un valor de verdad borroso permite, además de modelar la existencia de grados parciales de verdad, modelar el desconocimiento que en general puede tenerse sobre el valor exacto del grado de verdad (en general parcial) de la proposición, es decir su incertidumbre.

Un valor de verdad borroso es por tanto una distribución de posibilidad definida sobre el universo de discurso $[0,1]$ de valores de verdad, que delimita el conjunto de grados de verdad que una proposición puede tomar. Esto significa que en general, al igual que otra variable cualquiera, el valor del grado de verdad de una proposición puede quedar descrito mediante distribuciones de necesaria y posible posibilidad definidas en $[0,1]$, tal y como se vio en el capítulo 3.

A menudo se decide trabajar con un conjunto cerrado de valores de verdad, de forma que si por algún procedimiento se obtiene un valor de verdad cualquiera, este se reduce mediante algún mecanismo de comparación a alguno de los valores de la familia escogida. En la literatura la familia tradicional de valores de verdad ha sido casi siempre la propuesta por Zadeh en sus primeros trabajos ([Zadeh, 73], [Zadeh, 78b]). Esta familia está compuesta por los siguientes valores de verdad:

<i>verdadero:</i>	$\tau_v(x)=x$
<i>muy verdadero:</i>	$\tau_{mv}(x)=x^2$
<i>más o menos verdadero:</i>	$\tau_v(x)=x^{1/2}$
<i>falso:</i>	$\tau_f(x)=1-x$
<i>muy falso:</i>	$\tau_f(x)=1-x^2$
<i>más o menos falso:</i>	$\tau_f(x)=1-x^{1/2}$
<i>desconocido:</i>	$\tau_d(x)=1$

Lógicamente pueden generarse nuevos valores de verdad utilizando algunos de ellos como modificadores de otros. La elección propuesta por Zadeh, aunque es coherente con el sentido común no deja de ser arbitraria, y además limita sensiblemente las posibilidades semánticas disponibles, como veremos a lo largo de este capítulo.

Por otro lado suele ser conveniente mantener un conjunto cerrado de valores de verdad borrosos, de forma que cualquier nuevo valor que aparezca se pueda reducir a alguno de los del conjunto inicial. Sin embargo el parecido entre los valores de verdad inferidos con los modelos analizados en el capítulo 4 y los valores de verdad de la familia anterior es muy escaso, por lo que su reducción a alguno de los constituyentes de la familia de Zadeh solo puede hacerse en base a criterios de comparación confusos, que pueden a su vez dar lugar a simplificaciones en muchos casos inadmisibles.

Una forma habitual de obtener valores de verdad borrosos es mediante la compatibilidad borrosa definida por Zadeh (véase el apartado 6.3) que se obtiene como una distribución de posibilidad que describe el grado de inclusión de una observación en una hipótesis, para los distintos grados de pertenencia a la hipótesis. Sin embargo dependiendo del tipo de distribución con el que se modele la observación, así será el tipo de distribución que describa la compatibilidad borrosa obtenida. La consideración de los dos tipos de valores de verdad borrosos nos permitirá diseñar una familia general de valores de verdad borrosos, con una semántica más rica que la familia propuesta por Zadeh, y permitirá cualificar de forma homogénea una proposición, tanto en el dominio de lo posible como en el de lo necesario, en términos de grados de verdad parciales y grados de certidumbre parciales.

6.3 Compatibilidad borrosa, valores borrosos de necesidad y posibilidad

Como ya se vio en el capítulo 3, el valor de verdad o compatibilidad de una proposición P dado un hecho u observación P' se obtiene mediante la llamada compatibilidad borrosa:

$$\tau_{PP'}(x) = \begin{cases} \sup\{P'(u) / u \in P^{-1}(\{x\})\} & \text{if } P^{-1}(\{x\}) \neq \emptyset \\ 0 & \text{en otro caso} \end{cases}$$

$$\tau_{PP'}(x) = \begin{cases} \sup_{u/P(u)=x} P'(u) & \\ 0 & \text{if } P^{-1}(x) = \emptyset \end{cases} = P(P')$$

obteniéndose un valor de verdad borroso, que proporciona el máximo grado de pertenencia a P' para cada grado de pertenencia a P , y que es equivalente a evaluar, mediante el principio de extensión, la expresión borrosa $P(P')$ (véase por ejemplo [Dubois, Prade, 91]). El valor de verdad borroso $\tau_{PP'}(x)$ se puede interpretar como la distribución de posibilidad del grado de verdad de P conocido P' , siendo $\tau_{PP'}(x)$ el grado de posibilidad de que el grado de verdad de P dado P' sea x . Por tanto $\tau_{PP'}(x)$ es una distribución de posibilidad que restringe los grados de verdad de la proposición ' P es verdadero dado P' '.

La distribución $\tau_{PP'}(x)$ proporciona los máximos grados de pertenencia a P' para cada grado x de pertenencia a P por lo que permite obtener información precisa sobre la inclusión de P' en P (en cierto modo la distancia $\tau_{PP'}(x)-x$ es una medida de la inclusión de P' en P) así como sobre su intersección. Esto hace que a partir de $\tau_{PP'}(x)$ se pueda medir cómo de necesario y cómo de posible es que P sea verdadero conocido P' .

Por ello las distribuciones $\tau_{PP'}(P(u))$ y $P'(u)$ son equivalentes desde un punto de vista lógico, ya que ambas contienen la misma información (referente tanto a la inclusión como a la intersección de P' y P), lo que en definitiva permite que la RCI se pueda formular mediante valores de verdad borrosos, tanto con modelos de necesidad como con modelos de posibilidad.

De la expresión de la compatibilidad borrosa se deduce:

$$P'(u) \leq \tau_{PP'}(P(u)), \text{ o lo que es lo mismo: } P' \subset \tau_{PP'}(P)$$

lo que significa que $\tau_{PP'}(P(u))$ es siempre mayor que $P'(u)$ y por tanto puede interpretarse como una distribución de necesaria posibilidad compatible con $P'(u)$, pero equivalente a $P'(u)$ en las zonas que son de interés de cara a la inferencia. La distribución $\tau_{PP'}(P(u))$ es una observación equivalente a $P'(u)$ a efectos de comprobación de la hipótesis P , aunque es evidente que solamente es válida para contrastar dicha hipótesis. Veremos que esta observación es especialmente importante cuando se infiere con valores de verdad borrosos, de cara a interpretar correctamente las conclusiones que se obtienen.

Si la observación con la que se contrasta P es una distribución de necesaria posibilidad P_n' , entonces $\tau_{PP_n'}(P(u))$ también es una distribución de necesaria posibilidad respecto a P' . En efecto se tiene:

$$P_n'(u) \geq P'(u) \Rightarrow \tau_{PP_n'}(x) \geq \tau_{PP'}(x)$$

De modo similar, si P_p' es una distribución de posible posibilidad, entonces $\tau_{PP_p'}(P(u))$ también es una distribución de posible posibilidad para el grado de verdad de P , ya que se tiene:

$$P_p'(u) \leq P_n'(u) \Rightarrow \tau_{PP_p'}(x) \leq \tau_{PP_n'}(x)$$

El valor borroso $\tau_{PP'}(x)$ es por tanto, dependiendo del carácter de P' , una distribución de posible o necesaria posibilidad del grado de pertenencia a P' para cada grado de pertenencia a P .

En el caso en que P' sea una distribución de necesaria posibilidad $\tau_{PP'}(x)$ debe interpretarse como el máximo grado de posibilidad con el que han sido observados los elementos del universo U cuyo grado de pertenencia a P es x , y lo denominaremos **valor de necesidad borroso**. Si por el contrario P' es una distribución de posible posibilidad, entonces se interpreta como un posible grado de posibilidad con el que han sido observados los elementos del universo U cuyo grado de pertenencia a P es x (debido a que P' proporciona solamente posibles grados de observación de los elementos de U), y lo denominaremos **valor de posibilidad borroso**.

Cuando se utiliza un valor de verdad borroso como modificador lingüístico su comportamiento depende también del tipo de distribuciones involucradas. En concreto cuando $\tau(x)$ es la distribución de necesaria posibilidad dada por $\tau(x)=x$, $\tau(x)$ puede interpretarse (y así lo denominaremos a partir de ahora) como el **valor borroso necesario**, y se caracteriza porque no modifica geoméricamente la distribución a la que se aplica, ni tampoco su tipo. En efecto la proposición “ $(u \text{ es } P(u)) \text{ es necesario}$ ” siendo $P(u)$ una distribución de necesaria posibilidad es equivalente a “ $u \text{ es } \tau(P(u))=P(u)$ ” donde $P(u)$ debe interpretarse también como una distribución de necesaria posibilidad. Si por el contrario $P(u)$ es una distribución de posible posibilidad, afirmar que “ $(u \text{ es } P(u))$

es necesario” es equivalente a afirmar que es necesario que $P(u)$ sea un rango posible para u que es a su vez equivalente a afirmar que $P(u)$ es un rango posible para u .

Cuando por el contrario $\tau(x)$ es la distribución de posible posibilidad dada por $\tau(x)=x$ entonces $\tau(x)$ puede interpretarse (y así lo denominaremos a partir de ahora) como el **valor borroso posible**, y se caracteriza porque cuando se aplica a una proposición cualquiera siempre genera como resultado una distribución de posible posibilidad. En efecto si P es una distribución de necesaria posibilidad para la variable u , la proposición “*u es $P(u)$ es posible*” es equivalente a decir que $P(u)$ es un rango posible para u , o lo que es lo mismo “*u es $\tau(P(u))=P(u)$ ” donde ahora $P(u)$ debe interpretarse como una distribución de posible posibilidad. Si por el contrario P es una distribución de posible posibilidad, afirmar que es posible que P sea un rango posible para u es equivalente a afirmar directamente que P es un rango posible para u .*

A un valor de verdad borroso que verifique $\tau(x) \geq x$ lo denominaremos **modificador de necesidad**, y se caracteriza porque utilizado como modificador lingüístico aplicado a una distribución P siempre genera una distribución P' mayor, que por tanto puede interpretarse como una distribución de necesaria posibilidad respecto de P . Un modificador de necesidad aplicado a una distribución de necesaria posibilidad produce una nueva distribución de necesaria posibilidad. Sin embargo cuando se aplica a una distribución de posible posibilidad produce una distribución de tipo indeterminado, por lo que este último uso carece de sentido.

A un valor de verdad posibilidad borroso que verifique $\tau(x) \leq x$ lo denominaremos **modificador de posibilidad**, y se caracteriza porque utilizado como modificador lingüístico aplicado a una distribución P siempre genera una distribución P' menor que puede interpretarse como una distribución de posible posibilidad respecto de P . Un modificador de posibilidad aplicado a una distribución de necesaria posibilidad o a una de posible posibilidad produce una distribución siempre menor que por tanto puede interpretarse como una distribución de posible posibilidad.

Utilizaremos (como lo hemos venido haciendo) el término **valor de verdad borroso** cuando nos refiramos a cualquier distribución de posibilidad definida en el universo $[0,1]$ de valores de verdad.

Como veremos en el apartado 6.5 los modificadores de necesidad y posibilidad juegan un papel muy importante en la inferencia cuando ésta se formula en función de valores de verdad borrosos.

6.4 Interpretación de la compatibilidad borrosa a través de su relación con las medidas de grado de verdad y de certidumbre

Como se vio a lo largo del capítulo 4 existen algunas medidas de posibilidad y necesidad que, bajo ciertas simplificaciones, son especialmente útiles para describir el comportamiento en la inferencia de los distintos modelos analizados, además de proporcionar interesantes interpretaciones semánticas en términos de grados de verdad y grados de certidumbre.

Es posible relacionar las distintas formas de estimar el grado de verdad y el grado de certidumbre de una hipótesis P dada una observación P' , y la compatibilidad borrosa de P dado P' . Recordemos que el grado de verdad puede evaluarse en el dominio de lo necesario y en el dominio de lo posible mediante las estimaciones proporcionadas por las medidas:

$$\begin{aligned} N(P/[P']) &= \inf_{u/P'(u)=1} P(u) \\ \Pi(P/[P']) &= \sup_{u/P'(u)>0} P(u) \end{aligned}$$

y para el grado de certidumbre se tiene:

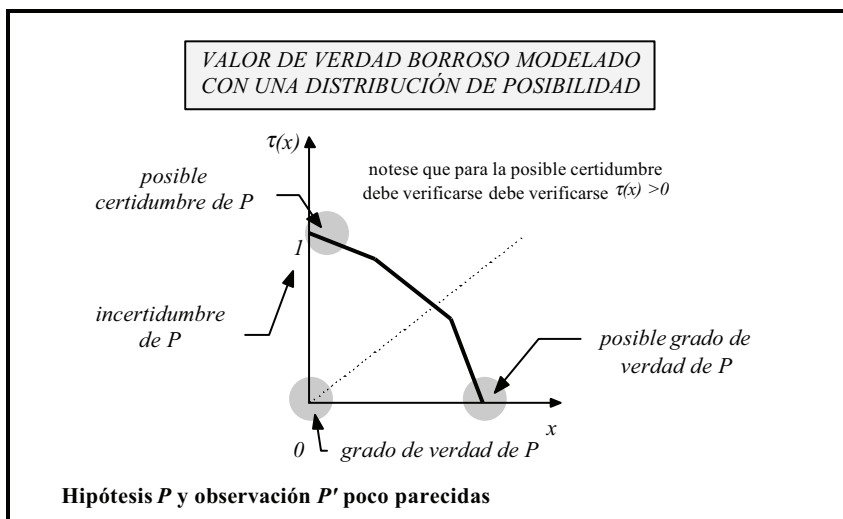
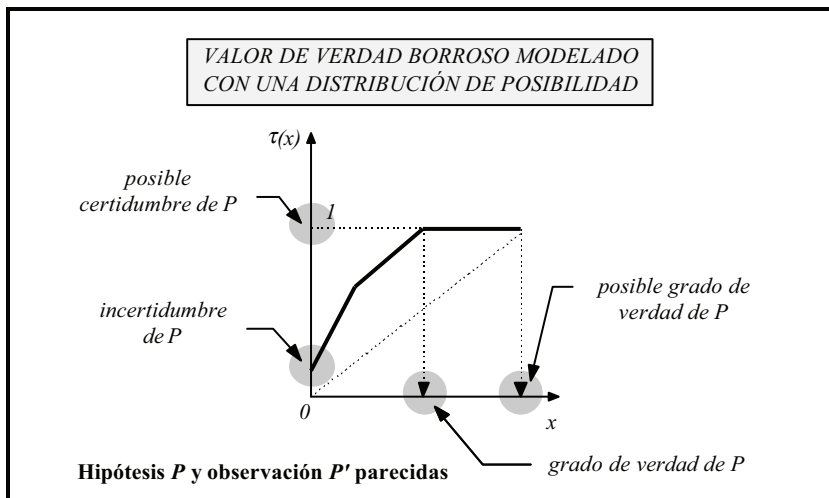
$$\begin{aligned} N(JP[/P']) &= \inf_{u/P(u)=0} (1-P'(u)) \\ \Pi(JP[/P']) &= \sup_{u/P(u)>0} P'(u) \end{aligned}$$

La definición de estas medidas permite obtener su valor a partir de la compatibilidad borrosa de P dado P' facilitando así la interpretación de esta última. En efecto se tiene, para cada una de ellas:

$$\begin{aligned} N(P/[P']) &= \inf_{u/P'(u)=1} P(u) = \inf \{ x / \tau_{PP'}(x) = 1 \} \\ \Pi(P/[P']) &= \sup_{u/P'(u)>0} P(u) = \sup \{ x / \tau_{PP'}(x) > 0 \} \\ N(JP[/P']) &= \inf_{u/P(u)=0} (1-P'(u)) = \inf \{ 1 - \tau_{PP'}(x) / x = 0 \} = 1 - \tau_{PP'}(0) \\ \Pi(JP[/P']) &= \sup_{u/P(u)>0} P'(u) = \sup \{ \tau_{PP'}(x) / x > 0 \} \end{aligned}$$

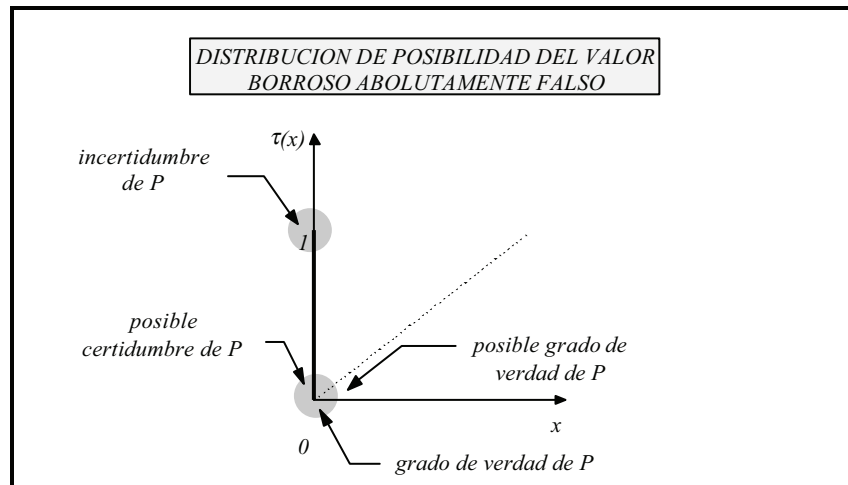
lo que permite interpretar numérica y gráficamente los valores de verdad borrosos, simplificando la información que contienen, de forma que puede fácilmente extraérseles una estimación del posible o necesario grado de verdad de la hipótesis, así como una estimación de cómo de posible o de necesario es que la hipótesis sea al menos parcialmente verdadera (esto es, de su certidumbre).

Con las relaciones anteriores podemos interpretar cualitativamente por inspección visual el significado de las distintas formas que la compatibilidad borrosa puede adoptar. Las siguientes figuras ilustran gráficamente su interpretación cualitativa a través de las estimaciones propuestas para los grados de verdad y de certidumbre, tanto en el dominio de lo posible como en el de lo necesario, cuando la compatibilidad borrosa es una distribución de posibilidad, esto es cuando se contrasta una hipótesis P con una distribución de posibilidad P' :



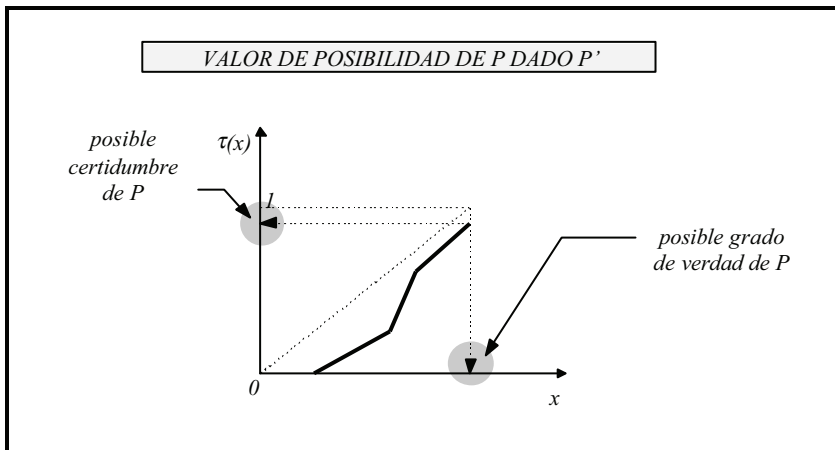
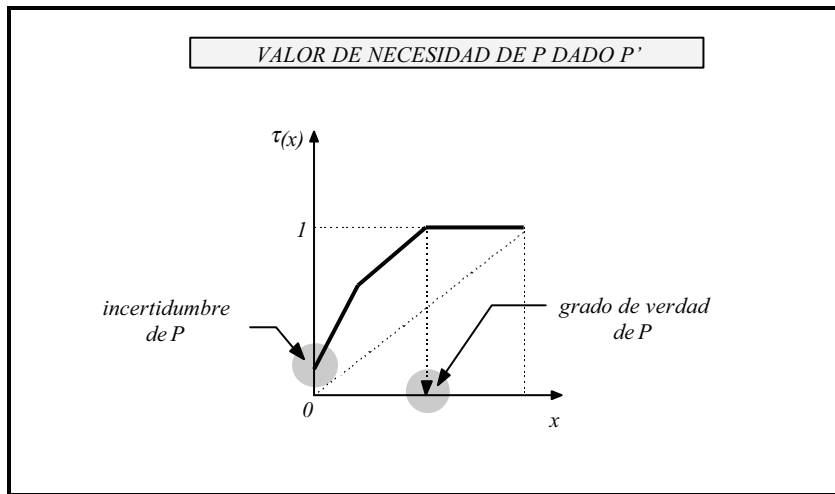
Las medidas de certidumbre y grado de verdad en el dominio de lo posible solamente se hacen 0 cuando la intersección entre P y P' es el conjunto vacío, en cuyo caso la

compatibilidad borrosa que resulta es el conjunto clásico $\{0\}$, que se conoce como el valor de verdad borroso *absolutamente falso* (siempre que P' sea una distribución de necesaria posibilidad), y que se muestra en la siguiente figura:



Como ya hemos señalado cuando la observación viene modelada por una distribución de necesaria posibilidad P_n' la distribución que resulta para la compatibilidad borrosa de P es también una distribución de necesaria posibilidad. Esta distribución contiene igualmente información tanto sobre el grado de inclusión de P_n' en P como sobre el grado de intersección, y en general ambos pueden ser de interés de cara a la inferencia. Sin embargo si se conoce una distribución de posible posibilidad P_p' para la variable u (tal y como se vio en el capítulo 5 sobre la coexistencia en un entorno único de los dos tipos de distribuciones) la información de interés de $\tau_{PP_n'}(x)$ de cara a la inferencia será solo aquella relativa a la inclusión de P' en P , ya que la información relativa a lo posible que pueda ser P vendrá mejor representada por $\tau_{PP_p'}(x)$. Del mismo modo la distribución $\tau_{PP_p'}(x)$ solo será interesante de cara a obtener información sobre cómo de posible es la hipótesis (y no sobre cómo de necesaria). Se podría evaluar el grado de inclusión de P_p' en P pero no tiene utilidad puesto que al ser la observación P_p' poco fiable no se puede por ejemplo afirmar nada sobre el (necesario) grado de verdad de P .

Las siguientes figuras ilustran la información que puede extraerse de los valores de necesidad y posibilidad borrosos:



Como puede verse en general el valor de verdad se estima viendo el mínimo valor de x para el que $\tau(x)$ es 1 (o suficientemente elevado, si se quiere ser menos estricto). El posible valor de verdad se puede estimar como el máximo valor de x para el que $\tau(x)$ es superior a 0 (o bien suficientemente elevado, si se quiere ser más estricto).

Del mismo modo el grado de incertidumbre se evalúa viendo directamente el valor $\tau(0)$ mientras que la posible certidumbre se evalúa como el máximo valor de $\tau(x)$ para x no nulo.

6.5 Relación de la compatibilidad borrosa con las medidas de posibilidad y necesidad generalizadas

La relación entre las medidas de posibilidad y necesidad habituales y la compatibilidad borrosa viene dada por las siguientes expresiones:

$$\begin{aligned}\Pi(P / P') &= \sup_x \min(P(u), P'(u)) \\ &= \sup_x \min(x, \tau_{PP'}(x)) = \Pi(\tau_v(x), \tau_{PP'}(x))\end{aligned}$$

$$\begin{aligned}N(P / P') &= 1 - \Pi(\neg P, P') = \inf_u \max(P(u), 1 - P'(u)) \\ &= \inf_x \max(x, 1 - \tau_{PP'}(x)) = N(\tau_v(x), \tau_{PP'}(x))\end{aligned}$$

Nótese que el cálculo de $N(P/P')$ solo tiene sentido si $\tau_{PP'}(x)$ es una distribución de necesaria posibilidad, mientras que el cálculo de $\Pi(P/P')$ debe en general hacerse a partir de una distribución de posible posibilidad. Las relaciones anteriores permiten también establecer algunas propiedades interesantes de $\tau_{PP'}(x)$ en relación con las medidas anteriores:

$$\begin{aligned}\text{si } \tau_{PP'}(1) &= 1 \text{ entonces } [P'] \cap [P] \neq \emptyset \Rightarrow \Pi(P/P') = 1 \\ \text{si } \tau_{PP'}(1) &= 0 \text{ entonces } [P'] \cap [P] = \emptyset \Rightarrow \Pi(P/P') < 1\end{aligned}$$

$$\begin{aligned}\text{si } \tau_{PP'}(0) &= 1 \text{ entonces } [P'] \cap \neg P \neq \emptyset \Rightarrow \Pi(\neg P/P') = 1 \text{ y } N(P/P') = 0 \\ \text{si } \tau_{PP'}(0) &= 0 \text{ entonces } [P'] \cap \neg P = \emptyset \Rightarrow \Pi(\neg P/P') < 1 \text{ y } N(P/P') > 0\end{aligned}$$

La medida de compatibilidad borrosa también puede relacionarse con las medidas de posibilidad y necesidad generalizadas propuestas en el apartado 4.7 del capítulo 4. Recordemos que a partir de un modelo de necesidad (M, I_n) y un modelo de posibilidad (M, I_p) es posible definir una medida de necesidad y otra de posibilidad dadas por:

$$\begin{aligned}\Pi_{(M, I_p)}(P/P') &= \sup_u M(P'(u), I_p(P(u), 1)) \\ N_{(M, I_n)}(P/P') &= 1 - \sup_u M(P'(u), I_n(P(u), 0))\end{aligned}$$

Para la primera de ellas se tiene:

$$\begin{aligned}\Pi_{(M, I_p)}(P/P') &= \sup_u M(P'(u), I_p(P(u), 1)) = \\ &= \sup_x M(\tau_{PP'}(x), I_p(x, 1))\end{aligned}$$

mientras que para la segunda:

$$N_{(M,I_n)}(P/P') = 1 - \sup_u M(P'(u), I_n(P(u), 0)) = \\ 1 - \sup_x M(\tau_{PP'}(x), I_n(x, 0))$$

La simplificación de las expresiones anteriores solo puede hacerse particularizando las implicaciones. Es un problema similar al que se daba cuando se analizó la dualidad de estas medidas en el apartado 4.7.2. En efecto dependiendo de la implicación las expresiones $I_p(x, 1)$ e $I_n(x, 0)$ toman valores diferentes. Recordando los resultados del apartado 4.3.3.2. del capítulo 4 se tiene:

$$\begin{array}{ll} I_p(x, 1) = x & \text{si } I_p(x, y) = M_R^T(x, y) \\ I_p(x, 1) = 1 & \text{si } I_p(x, y) = M_S^{T+}(x, y) \text{ y } x \neq 0 \\ I_p(x, 1) = g^{[1]}(g(1) - g(1-x)) & \text{si } I_p(x, y) = M_S^T(x, y) \\ \\ I_n(x, 0) = 1-x & \text{si } I_n(x, y) = I_S^T(x, y) \\ I_n(x, 0) = 0 & \text{si } I_n(x, y) = I_R^{T+}(x, y) \text{ y } x \neq 0 \\ I_n(x, 0) = f^{[-1]}(f(0) - f(x)) & \text{si } I_n(x, y) = I_R^T(x, y) \end{array}$$

expresiones que permiten simplificar, según los distintos casos particulares que se plantean, la relación entre el valor de verdad borroso y las medidas de necesidad y posibilidad generalizadas. Por ejemplo cuando se verifica $I_p(x, y) = M_R^T(x, y)$ y $I_n(x, y) = I_S^T(x, y)$ se tiene:

$$\Pi(P / P') = \sup_u M(P'(u), M(P(u), 1)) = \sup_x M(x, \tau_{PP'}(x)) = \Pi(\tau_v(x), \tau_{PP'}(x))$$

$$N(P / P') = 1 - \sup_u M(P'(u), I(P(u), 0)) = \\ 1 - \sup_x M(\tau_{PP'}(x), I(x, 0)) = N(\tau_v(x), \tau_{PP'}(x))$$

obteniéndose en este caso el mismo tipo de relación que con las medidas estándar.

6.6 Valores de posibilidad y necesidad borrosos e inferencia

Cuando la RCI se formula en función de valores de verdad borrosos las conclusiones que se obtienen son también valores de verdad borrosos, que permiten modificar el consecuente de la regla para obtener la conclusión, de acuerdo con la semántica del modelo empleado para la regla. Como se vio en el capítulo 4 se tiene:

$$Q'(v) = \tau_{QQ'}(Q(v))$$

$$\tau_{QQ'}(y) = \sup_x M(\tau_{PP'}(x), I(x, y))$$

y para los modelos de necesidad se verifica:

$$\tau_{QQ'}(y) \geq y$$

mientras que para los de posibilidad:

$$\tau_{QQ'}(y) \leq y$$

Además cuando los modelos se activan por completo, las desigualdades se convierten en igualdades, pero la conclusión que se obtiene de unos y otros modelos se interpreta de modo distinto, lo que hace que el valor borroso $\tau_{QQ'}(y)=y$ concluido para ambos deba también interpretarse de forma diferente.

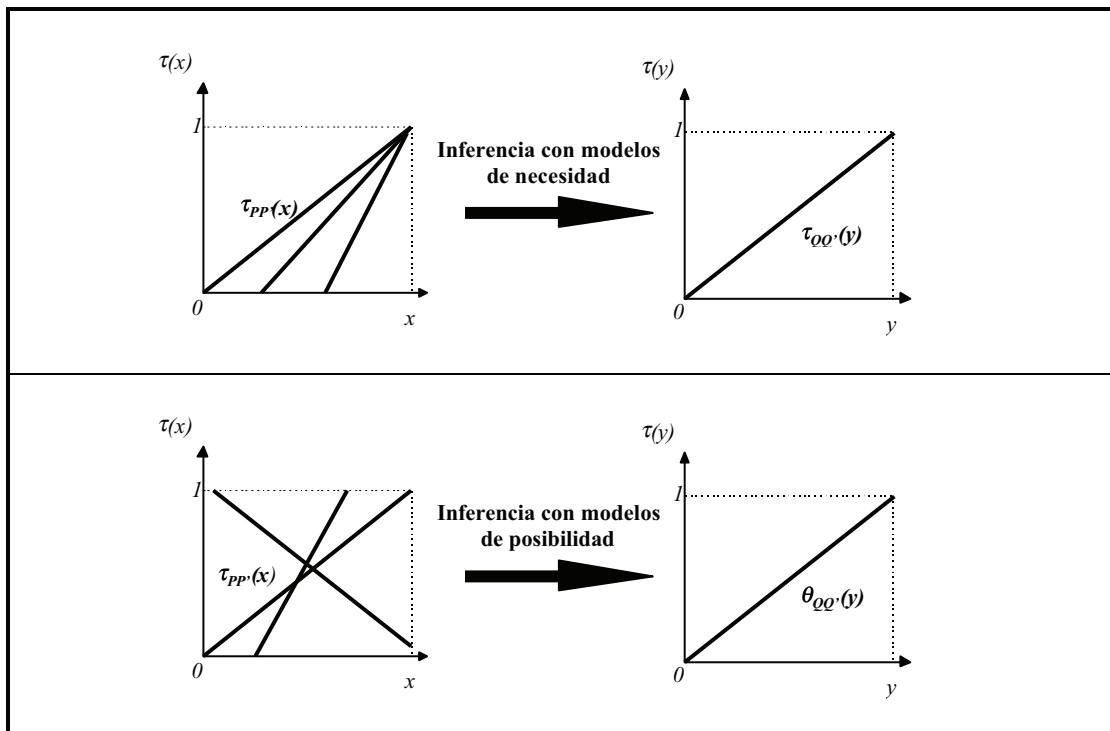
En efecto cuando los modelos se activan por completo no se produce modificación del consecuente para la obtención de la conclusión. Los modelos de necesidad concluyen “*Q es necesario*” lo que equivale a decir que Q es un rango necesario para la variable v . Como la conclusión $Q'=Q$ equivale a “*Q' es $\tau_{QQ'}(Q)=Q$* ” y Q' se interpreta como una distribución de necesaria posibilidad, el valor de verdad borroso $\tau_{QQ'}(y)=y$ que en este caso se concluye es el **valor de verdad borroso necesario**, introducido en el apartado 6.3.

Para los modelos de posibilidad la activación total da lugar a la conclusión “*Q es posible*”, lo que equivale a decir que Q es un rango posible para la variable v . Como la conclusión $Q'=Q$ equivale a “*Q' es $\tau_{QQ'}(Q)=Q$* ”, y Q' se interpreta como una distribución de posible posibilidad, en este caso $\tau_{QQ'}(y)=y$ debe interpretarse como el **valor de verdad borroso posible**.

Además como para los modelos de necesidad se tiene $\tau_{QQ'}(y) \geq y$, el valor de necesidad borrosos $\tau_{QQ'}(y)$ es para estos modelos un modificador de necesidad. Del mismo modo

para los modelos de posibilidad se verifica $\tau_{QQ'}(y) \leq y$ por lo que $\tau_{QQ'}(y)$ es un modificador de posibilidad.

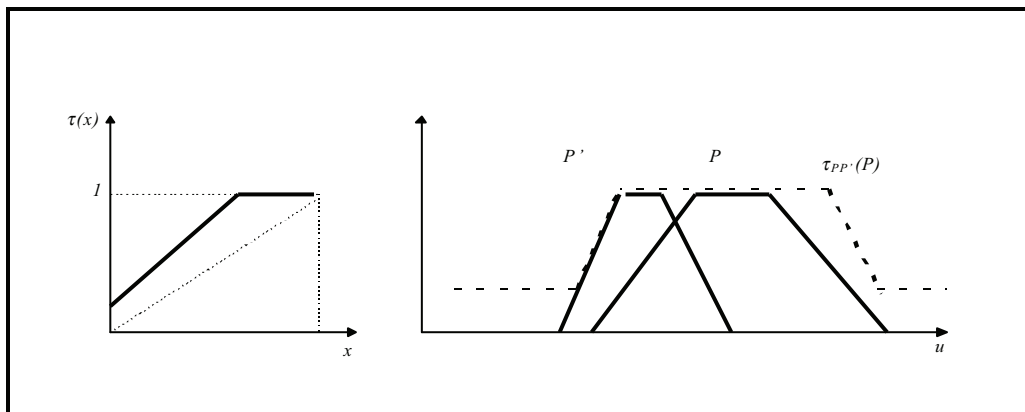
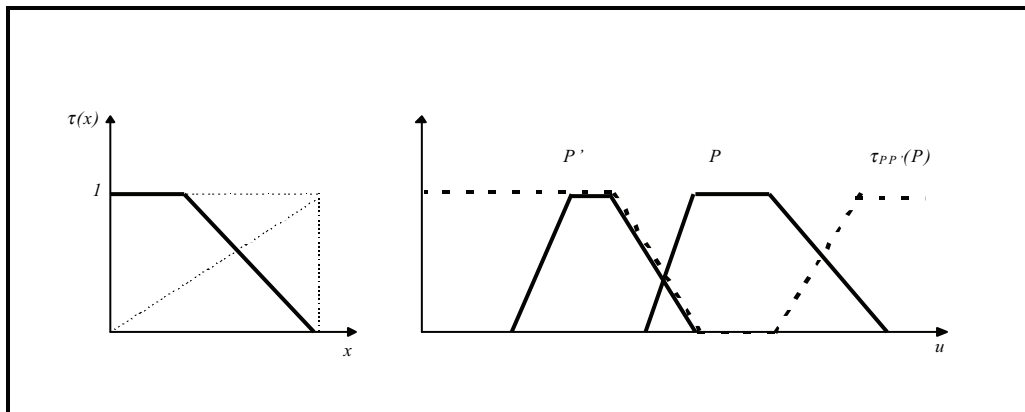
Existen diferencias significativas entre el comportamiento de los modelos de posibilidad y los necesidad, en la forma en que transforman los valores de verdad borrosos, como puede apreciarse en las siguientes figuras. En ellas se muestran algunos de los posibles valores de compatibilidad de P dado P' que dan lugar a la misma conclusión $\tau_{QQ'}(y)=y$:



Los modelos de necesidad transforman valores de verdad borrosos en valores de necesidad borrosos similares, ya que solo pueden activarse (aunque depende de su condición de activación) cuando la compatibilidad borrosa de P dado P' es un modificador de necesidad, en cuyo caso generan siempre como resultado otro modificador de necesidad.

Si la compatibilidad borrosa de P es $\tau_{PP}(x)$ entonces resulta equivalente contrastar P con P' que contrastarlo con la observación $\tau_{PP}(P)$. Si por ejemplo se infiere con el modelo de necesidad gradual resulta equivalente tomar como observación $\tau_{PP}(x)$ que tomar $\tau_{QQ'}(x)$. En efecto la modificación lingüística que $\tau_{QQ'}(x)$ produce sobre P es equivalente, de cara a contrastar P , a la que produce $\tau_{PP}(x)$. Sin embargo se produce una pérdida importante de

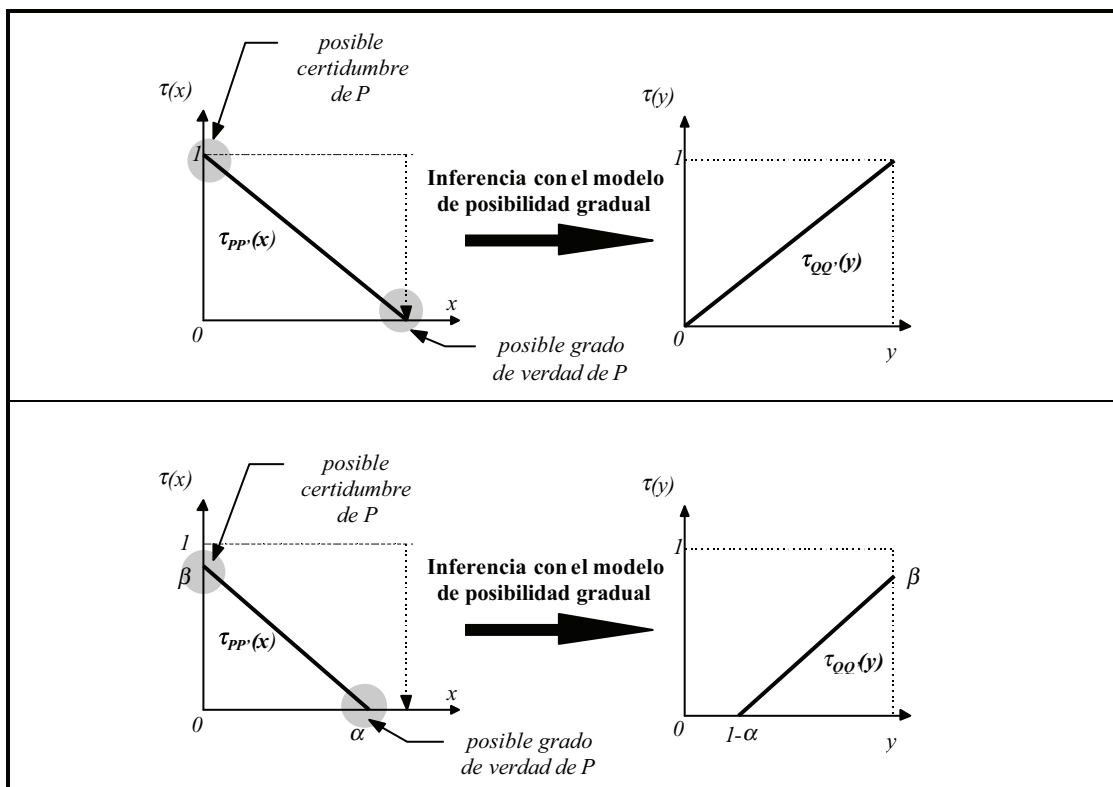
información si para contrastar otra hipótesis cualquiera se utiliza $\tau_{PP'}(P)$ en lugar de P' . Esta pérdida de información es tanto mayor cuanto más diferente es $\tau_{PP'}(x)$ de x , ya es entonces cuando más diferentes son $\tau_{PP'}(P)$ y P' , como queda reflejado en las siguientes figuras:



En la situación representada en la primera de las figuras ningún modelo de necesidad se activaría, con lo que la conclusión vendría dada por $\tau_{QQ}(y)=1$. En la segunda sin embargo los modelos de necesidad graduales y mixtos se activarían dando lugar a valores $\tau_{QQ}(y)$ similares a $\tau_{PP'}(x)$. En cierto sentido los modelos de necesidad se activan cuando $\tau_{PP'}(P)$ es suficientemente parecido a P' , en cuyo caso concluyen valores $\tau_{QQ}(y)$ similares a $\tau_{PP'}(x)$. Esto significa que los modelos de necesidad se activan cuando P' y $\tau_{PP'}(P)$ son prácticamente reemplazables el uno por el otro, no solo para contrastar P sino también para contrastar otras posibles hipótesis.

Por el contrario los modelos de posibilidad transforman valores de verdad borrosos en otros valores de verdad borrosos que en muchos casos pueden ser completamente diferentes, ya que muchos de ellos pueden activarse (aunque depende de su condición de activación) en cuanto la compatibilidad de P dado P' es diferente del valor de verdad borroso *absolutamente falso*, en el caso en que P' sea una distribución de necesaria posibilidad, o del valor de verdad borroso *posiblemente absolutamente falso* (que es una distribución de posible posibilidad igual a la del valor de necesidad borroso *absolutamente falso*).

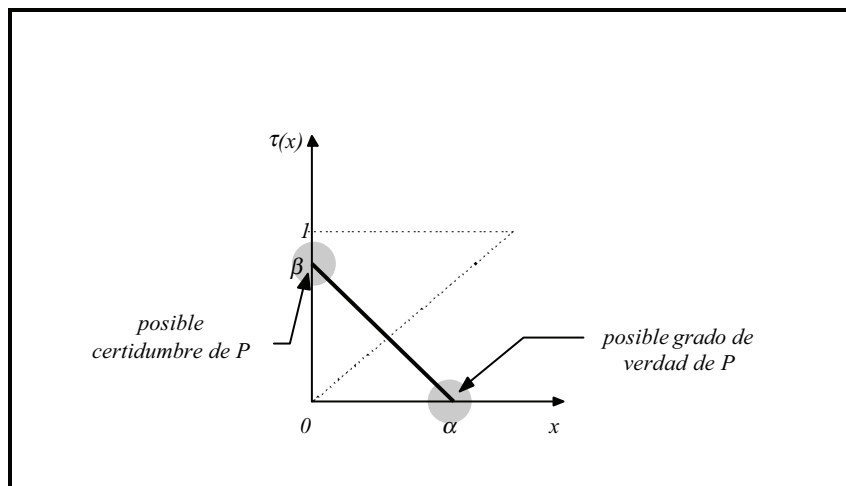
Por ejemplo para el modelo de posibilidad gradual, a partir de una observación $P' = \neg P$, se verifica $\Pi(P'/]P])=1$ y $\Pi(P/]P')=1$ por lo que concluye el valor de verdad borroso $\tau_{QQ'}(y)=y$ que se interpreta como el valor borroso posible, mientras que Q se interpreta como un rango posible para v . La siguiente figura muestra el comportamiento del modelo de posibilidad gradual en función de valores de verdad borrosos:



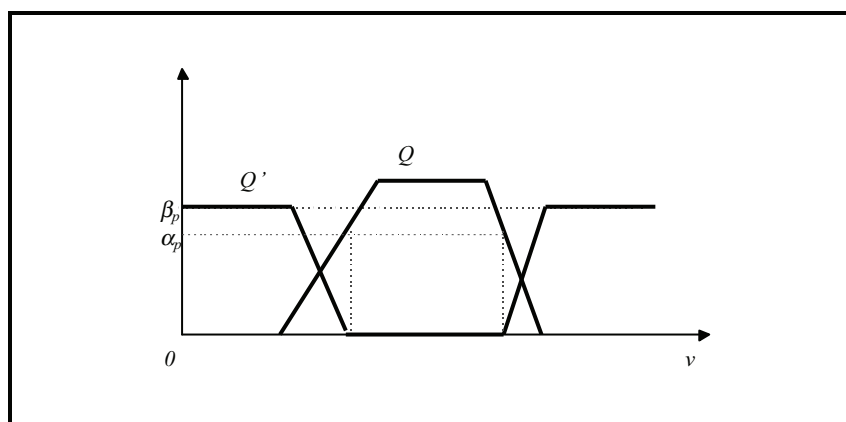
Aunque $\tau_{PP'}(P)$ sea muy diferente de P' , siempre que exista un mínimo grado de intersección entre P y P' los modelos de posibilidad pueden activarse. Por tanto aunque resulte equivalente contrastar P con P' que contrastarlo con $\tau_{PP'}(P)$, dado que el parecido

entre $\tau_{PP'}(P)$ y P' puede ser muy pobre, no puede decirse que $\tau_{PP'}(P)$ y P' sean reemplazables el uno por el otro a efectos de comprobación de otras posibles hipótesis. Por eso no es lógico que en este caso los modelos de posibilidad concluyan valores borrosos $\tau_{QQ'}(y)$ similares a $\tau_{PP'}(x)$.

En efecto si el modelo de posibilidad transformase $\tau_{PP'}(x)$ en un valor de posibilidad borroso $\tau_{QQ'}(y)$ similar, al igual que hacen los modelos de necesidad, el comportamiento de la conclusión así obtenida sería completamente contrario al deseado. Si para simplificar suponemos que el comportamiento del modelo es tal que $\tau_{QQ'}(x) \approx \tau_{PP'}(x)$ cuando la distribución $\tau_{PP'}(x)$ es tal y como aparece en la siguiente figura:

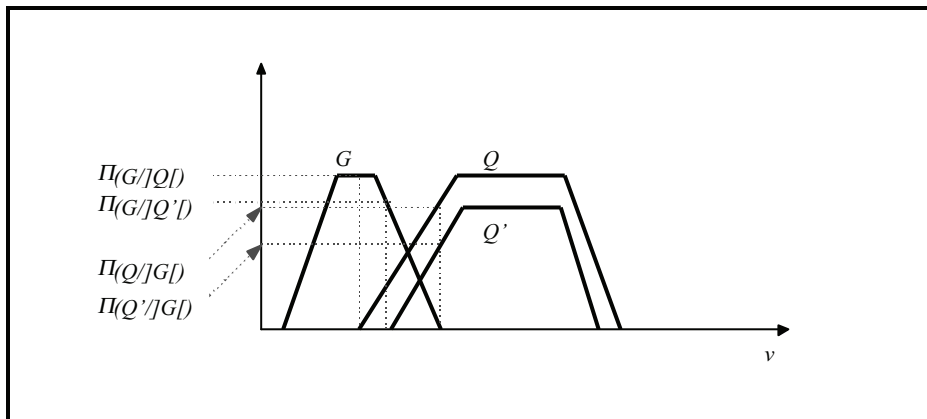


entonces la modificación llevada a cabo en el consecuente es la que queda reflejada en la siguiente figura:



y si bien es cierto que Q dado Q' daría lugar al mismo posible valor de verdad que el de P dado P' , una observación como Q' haría que otras hipótesis tuviesen en general grados de verdad mucho mayores que si se contrastasen con Q , lo cual es claramente poco deseable. La forma única forma razonable de obtener una distribución de posible posibilidad Q' a partir de Q compatible con la semántica de los modelos de posibilidad es reduciendo Q , y por tanto utilizando como valor de posibilidad borroso un modificador de posibilidad.

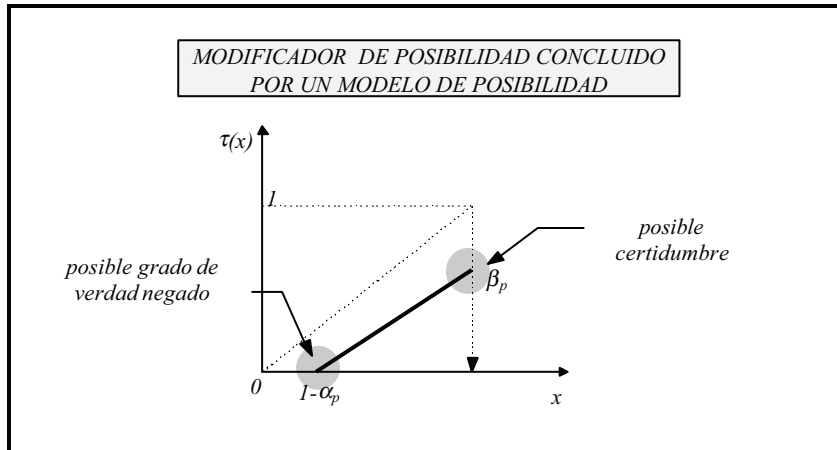
Aunque para los modelos de posibilidad el grado de verdad de Q respecto de Q' es en general igual a 1, la obtención de Q' mediante una reducción de Q da lugar a comportamientos razonables. Así por ejemplo en la siguiente figura puede verse como en efecto tanto el posible grado de verdad como el posible grado de certidumbre de una hipótesis G contrastada con Q' son en general menores que si se obtuviesen contrastando G con Q , o cualquier otra distribución mayor que Q .



Por tanto cualificar una distribución Q en términos de posibilidad, para que la distribución resultante utilizada como observación conlleve en general una reducción del posible grado de verdad y del posible grado de certidumbre de otra hipótesis cualquiera, exige una reducción de Q . Esto significa que aunque de $\tau_{PP}(x)$ se deduzca que el posible grado de verdad de P es α y su posible grado de certidumbre β , la mejor forma de obtener la distribución que mejor represente a “ u es P es posiblemente verdadero en grado α y su posible certidumbre es β ” no es mediante “ u es $\tau_{PP}(P(u))$ ” sino mediante “ u es $\tau(P(u))$ ” donde $\tau(x)$ debe ser un modificador de posibilidad con una semántica apropiada. La modificación “ u es $\tau_{PP}(P(u))$ ” da resultados exactos de cara a contrastar P pero da resultados aberrantes de cara a contrastar otras hipótesis. Por el contrario “ u es $\tau(P(u))$ ” no da resultados exactos de cara a contrastar P , pero da resultados razonables de cara a

la proposición inicial. Del mismo modo, dado que la certidumbre se mide como $N(JP[/P')$, decir que una proposición P es parcialmente cierta deberá ser visto por una hipótesis cualquiera como un incremento de su soporte, lo que se traduce en la aparición de un cierto nivel de indeterminación.

La siguiente figura muestra una interpretación análoga de los modificadores de posibilidad:



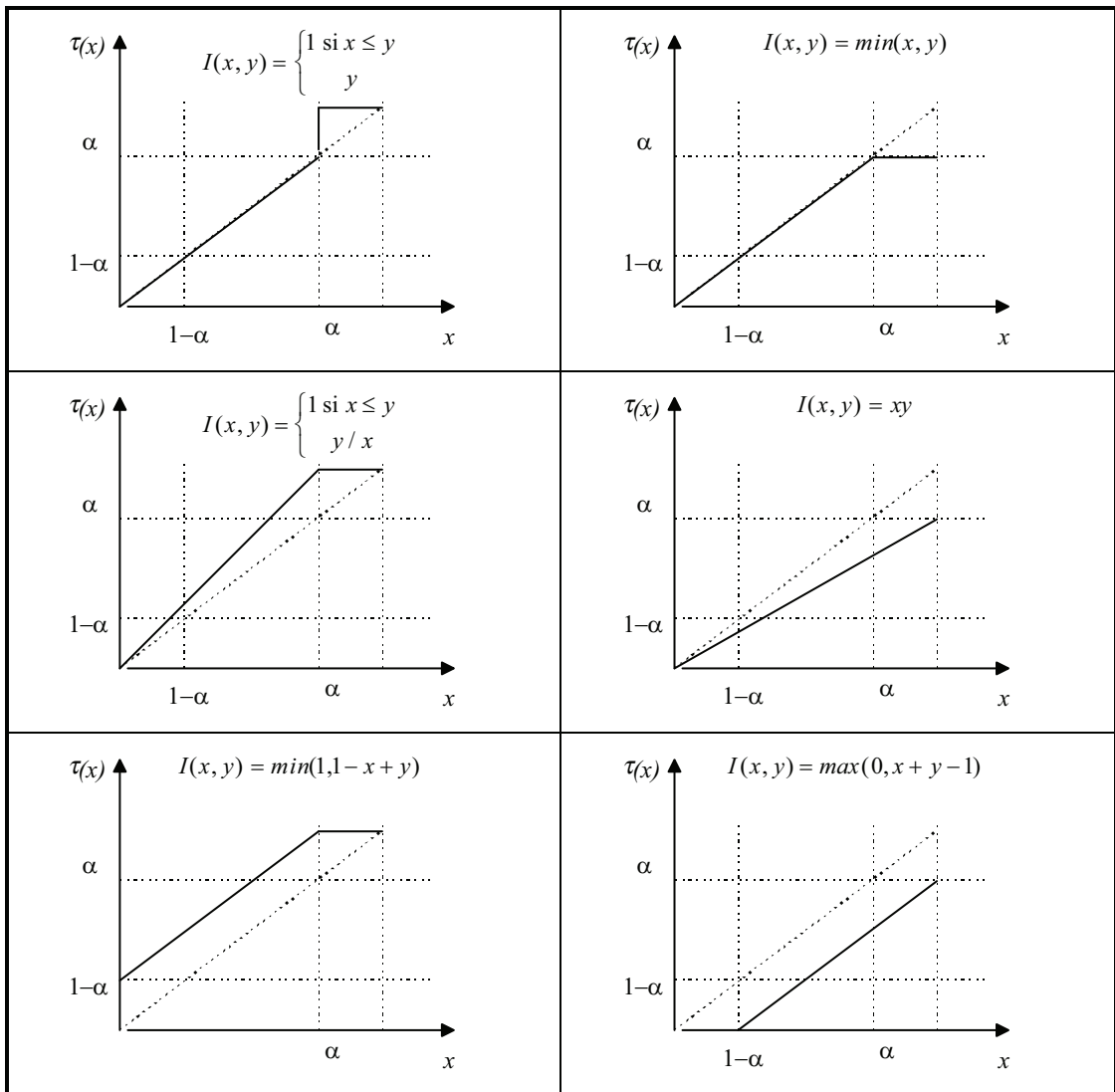
En este caso para conseguir un comportamiento satisfactorio de la conclusión del modelo, es necesario modificar el consecuente de modo muy diferente al que produciría los valores de $\tau_{pp}(x)$ que hacen que los modelos de posibilidad se activen. Esto hace que el modificador de posibilidad deba interpretarse de distinto modo a como se interpretan los valores de posibilidad borrosos que se obtienen evaluando la compatibilidad borrosa de una hipótesis dada una observación.

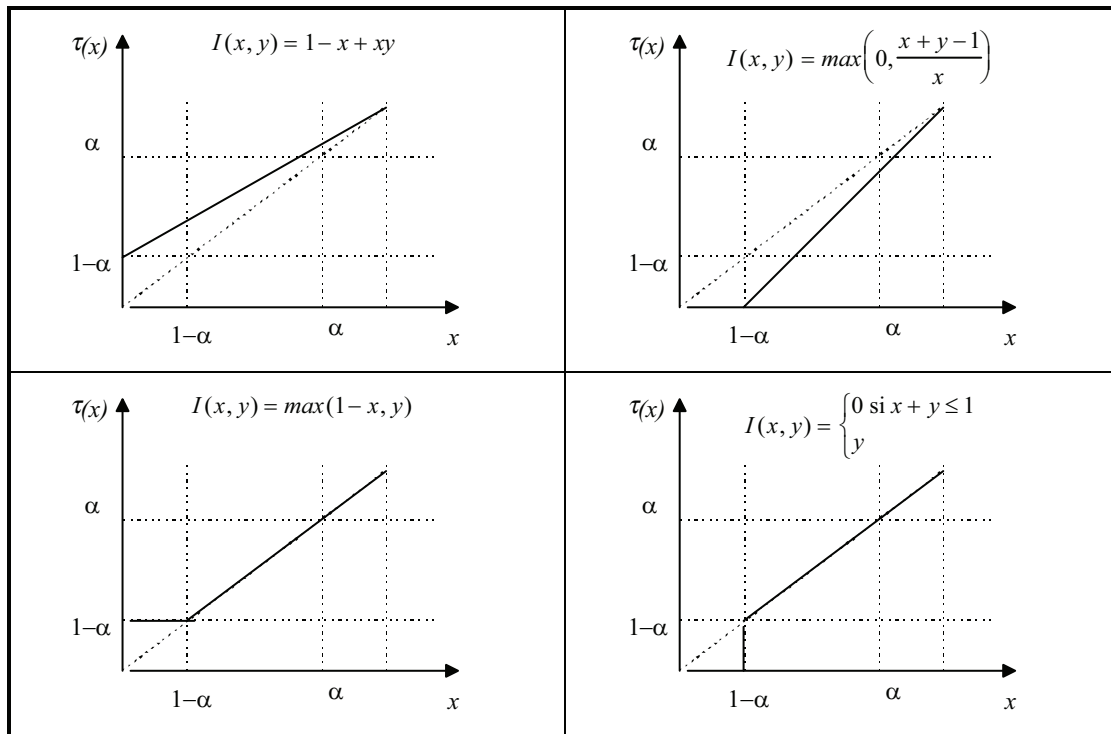
Así se tiene que en el dominio de lo posible el grado de verdad de una hipótesis se mide como $\Pi(P[/P')$, por lo que cualificar una proposición P como parcialmente verdadera tiene que implicar una decremento de su soporte de forma que la comprobación de otra hipótesis cualquiera garantice que su grado de verdad se verá también decrementado. Del mismo modo, dado que la certidumbre se mide como $\Pi(JP[/P')$, decir que una proposición P es parcialmente cierta deberá ser visto por una hipótesis cualquiera como un decremento de su núcleo, lo que se traduce en una reducción de su altura.

La interpretación propuesta permite comprender mejor el significado de cualificar una proposición con alguno de los modificadores de necesidad o posibilidad que se obtiene según los distintos tipos de modelos empleados. Por ejemplo, cuando se verifican las condiciones de equivalencia entre la RCI y la MS, para un modelo basado en una implicación I , los modificadores concluidos tienen todos la forma:

$$\tau(y) = I(\alpha, y)$$

Algunos de los modificadores obtenidos con las principales implicaciones aparecen representados en las siguientes figuras:





Nótese que las simetrías que pueden apreciarse entre unos y otros no son casuales, y en [Dubois, Prade, 84b] pueden encontrarse las transformaciones matemáticas que permiten pasar de unos valores de verdad a otros. En concreto puede mostrarse de forma sencilla que valores de verdad borrosos generados por condicionales duales son simétricos respecto a la curva $\tau(x)=x$, como puede apreciarse en las figuras anteriores.

A modo de ejemplo los valores de verdad borrosos obtenidos a partir de algunas de las n-implicaciones más comunes pueden interpretarse como sigue:

$$\begin{aligned} \tau(x) &= I_R^{\min}(\alpha, x) \text{ grado de verdad } \alpha, \text{ grado de certidumbre } 1 \\ \tau(x) &= I_S^{\min}(\alpha, x) \text{ grado de verdad } 1, \text{ grado de certidumbre } 1-\alpha \\ \tau(x) &= I_S^L(\alpha, x) \text{ grado de verdad } \alpha, \text{ grado de certidumbre } 1-\alpha \end{aligned}$$

Del mismo modo una interpretación razonable para los valores de posibilidad borrosos inferidos por las principales p-implicaciones analizadas es:

$$\begin{aligned} \tau(x) &= M_R^{\min}(\alpha, x) \text{ posible grado de verdad } \alpha, \text{ posible certidumbre } 1 \\ \tau(x) &= M_S^{\min}(\alpha, x) \text{ posible grado de verdad } 1 \text{ posible certidumbre } \alpha \\ \tau(x) &= M_S^L(\alpha, x) \text{ posible grado de verdad } \alpha \text{ posible certidumbre } \alpha \end{aligned}$$

Estamos finalmente en condiciones de proponer una familia general de valores de verdad borrosos que puedan ser utilizados como modificadores de necesidad y posibilidad e interpretados en términos de grados de verdad y grados de certidumbre parciales, y que permitan cualificar proposiciones tanto en el dominio de lo posible como en el dominio de lo necesario, diferenciando grados de certidumbre y grados de verdad.

6.7 Familia general de valores de verdad borrosos

El análisis del comportamiento de los principales modelos de reglas condujo en el capítulo 4 a formular un mecanismo de inferencia simplificado pero capaz de reproducir los principales comportamientos semánticos detectados. El mecanismo de inferencia propuesto es del tipo MS pero con dos índices de similitud, y se basa en el uso de dos implicaciones generalizadas, una para los modelos de necesidad y otra para los modelos de posibilidad.

La implicación propuesta para los modelos de necesidad viene dada por:

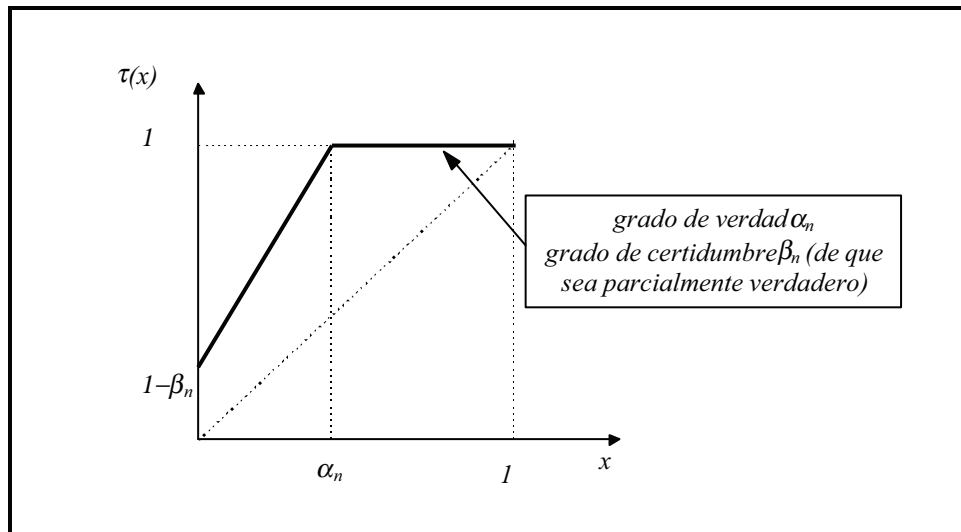
$$I_n(\alpha, \beta, y) = \min\left(\frac{\beta}{\alpha}(y - \alpha) + 1, 1\right)$$

Esta implicación permite generar la familia de modificadores de necesidad $I_n(\alpha_n, \beta_n, x)$ cuya interpretación según se ha visto en apartados anteriores viene dada por:

$$\tau(x) = I_n(\alpha_n, \beta_n, x) \text{ grado de verdad } \alpha_n, \text{ grado de certidumbre } \beta_n$$

de forma que escogiendo adecuadamente α_n y β_n se tiene una familia de valores de necesidad borrosos con una riqueza semántica mucho mayor que por ejemplo la de la familia que habitualmente se utiliza. Así la afirmación “ P es verdadero en grado α , y su certidumbre es β ” puede considerarse equivalente a la afirmación “ $(u \text{ es } P(u)) \text{ es } I_n(\alpha_n, \beta_n, x)$ ”, o lo que es lo mismo a “ $u \text{ es } I_n(\alpha_n, \beta_n, P(u))$ ”. En este caso la contrastación de P con $I_n(\alpha_n, \beta_n, P(u))$ permite recuperar los valores α y β . A menudo utilizaremos la notación (N, α_n, β_n) para designar al valor de necesidad borroso $I_n(\alpha_n, \beta_n, x)$.

La siguiente figura representa la familia de valores de necesidad borrosos cuando modelan grados de verdad:

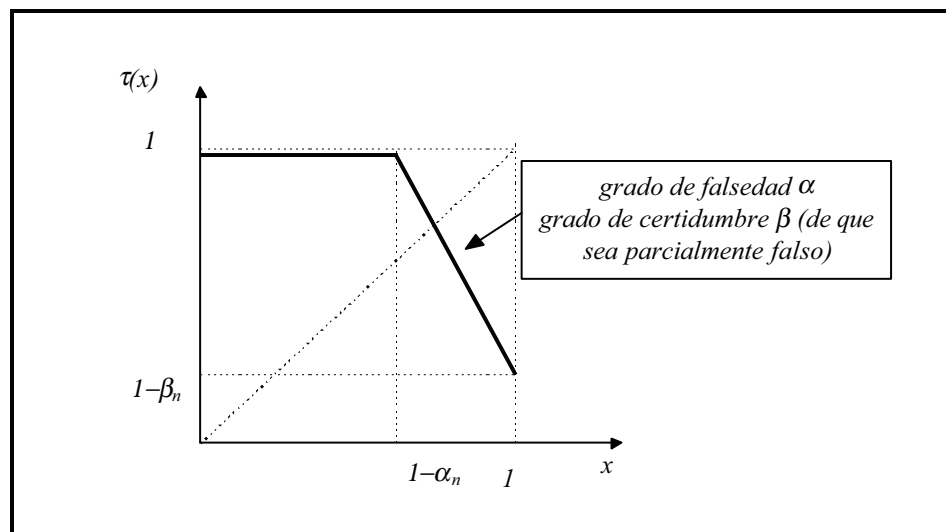


Para modelar cómo de falso es un predicado mediante un modificador basta con obtener el antónimo de los modificadores de necesidad anteriores, que en este caso coincide con su complemento, de forma que se tiene:

$$\tau(x) = 1 - I_n(\alpha_n, \beta_n, x)$$

grado de falsedad α_n ,
grado de certidumbre (de que P sea al menos parcialmente falso) β_n

La siguiente figura representa los valores de necesidad borrosos cuando modelan grados de falsedad:



Podemos utilizar el mismo proceder con los valores de posibilidad borrosos. Así utilizando la implicación de posibilidad obtenida en el capítulo 4, que recordamos venía dada por:

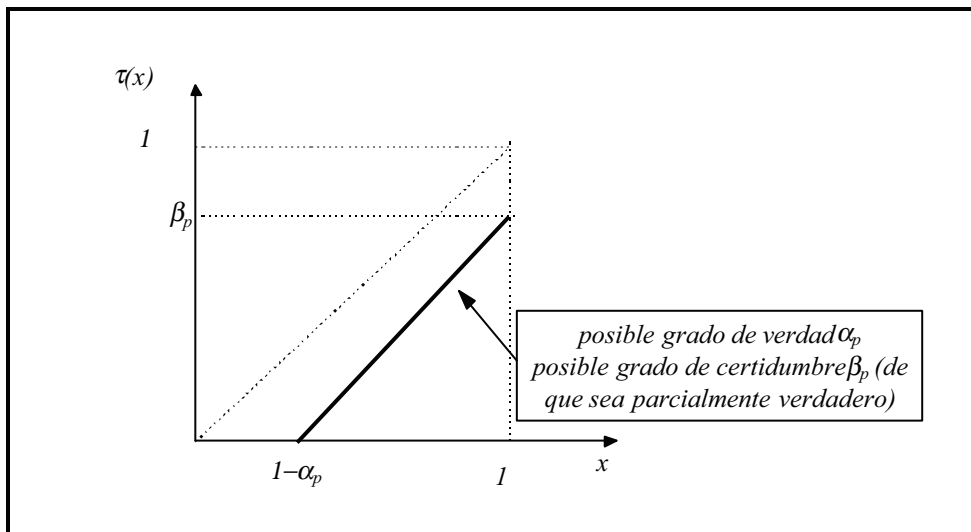
$$I_p(\alpha, \beta, y) = \max\left(\frac{\beta}{\alpha}(y-1) + \beta, 0\right)$$

podemos generar la familia de modificadores de posibilidad $I_p(\alpha_p, \beta_p, x)$:

$$\tau(x) = I_p(\alpha_p, \beta_p, x) \text{ posible grado de verdad } \alpha_p \\ \text{ posible grado de certidumbre } \beta_p$$

La afirmación “ P es posiblemente verdadero en grado α , y su posible certidumbre es β ” puede considerarse aproximadamente equivalente a la afirmación “(u es $P(u)$) es $I_p(\alpha_p, \beta_p, x)$ ”, o lo que es lo mismo a “u es $I_p(\alpha_p, \beta_p, P(u))$ ”. En este caso la contrastación de P con $I_n(\alpha_n, \beta_n, P(u))$ solo permite recuperar el valor β ya que el grado de verdad que se obtiene es 1, por lo que la equivalencia es solo aproximada. Sin embargo como ya hemos visto $I_n(\alpha_n, \beta_n, P(u))$ parece ser la distribución de posibilidad que, de forma global y no solo en relación con P , mejor representa a la afirmación anterior. A menudo utilizaremos la notación (Π, α_p, β_p) para designar al valor de necesidad borroso $I_p(\alpha_p, \beta_p, x)$.

La siguiente figura muestra la representación gráfica de los valores de posibilidad borrosos cuando modelan grados de verdad:

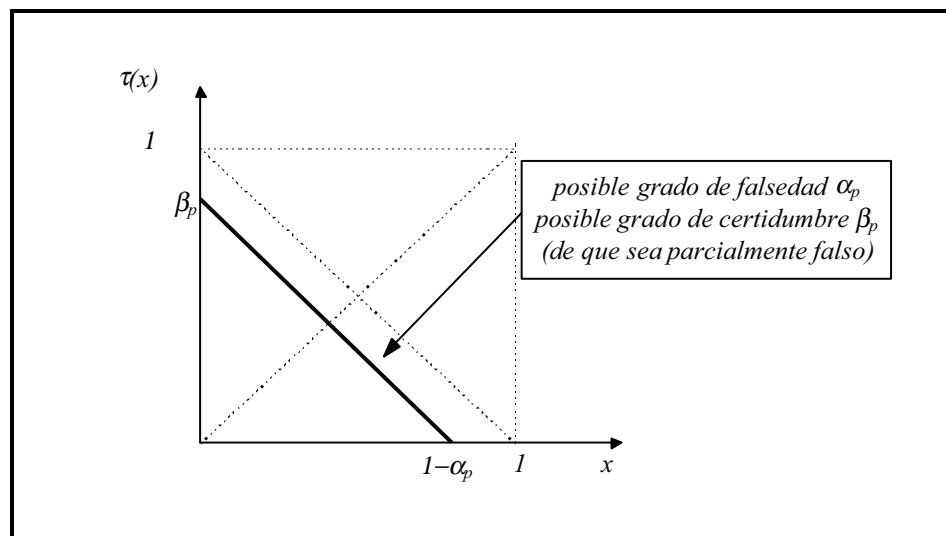


Podemos también generar valores de posibilidad borrosos para modelar grados de falsedad:

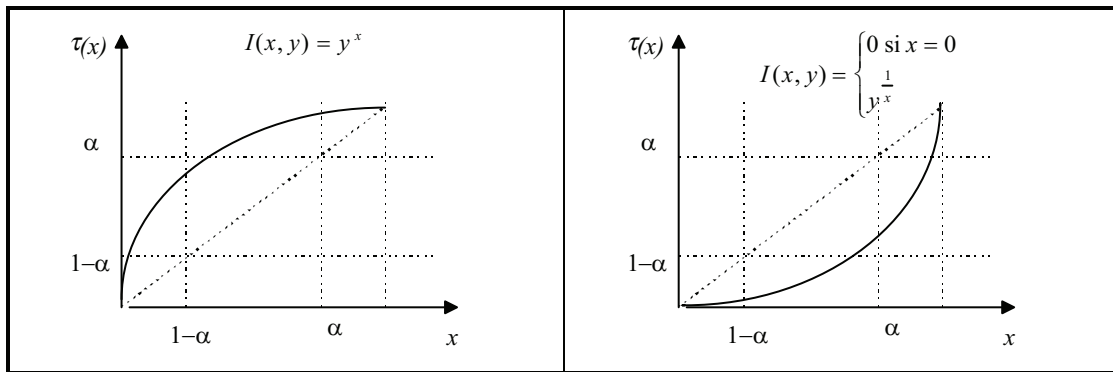
$$\tau(x) = 1 - I_p(\alpha_p, \beta_p, x) \text{ posible grado de falsedad } \alpha$$

posible grado de certidumbre β (de que P sea al menos parcialmente falso)

La siguiente figura muestra la representación gráfica de los valores de posibilidad borrosos cuando modelan grados de falsedad:



Los valores de verdad borrosos de la familia aquí propuesta no están representados en la familia propuesta por Zadeh, pero su riqueza semántica es mayor ya que permite expresar conceptos más complejos en términos de grados de verdad parciales y grados de certidumbre parciales, tanto en el dominio de lo necesario como en el dominio de lo posible. Por otro lado es interesante señalar que la familia propuesta por Zadeh puede generarse también a partir de una implicación, la implicación de Yager (que fue parcialmente analizada a lo largo del capítulo 4), y puede considerarse como una familia de valores de necesidad borrosos. Su FGMP máxima permitiría del mismo modo que se ha hecho aquí obtener la correspondiente familia de valores de posibilidad borrosos. Esta familia, como se vio en el capítulo 4, resulta más difícil de interpretar que la que aquí ha sido propuesta, debido a que la modificación semántica que produce en el consecuente no modifica ni su núcleo ni su soporte, que son precisamente sus zonas más significativas a la hora de ser interpretado. La siguiente figura representa los valores de verdad borrosos obtenidos a partir de las implicaciones de Yager:



6.8 Agregación de valores de verdad borrosos

Cuando se dispone de más de un valor de verdad borroso que describa el conocimiento que sobre una determinada proposición P se tiene, estos pueden ser agregados entre sí para obtener un único valor de verdad borroso, y el tipo de agregación será función del tipo de las distribuciones involucradas. Los valores de necesidad borrosos deberán agregarse mediante el operador \min mientras que los valores de posibilidad borrosos deberán ser agregados mediante el operador \max .

En efecto dadas dos proposiciones “ $(u \text{ es } P(u)) \text{ es } \tau_1(x)$ ” y “ $(u \text{ es } P(u)) \text{ es } \tau_2(x)$ ” donde tanto P como τ_1 y τ_2 son distribuciones de necesaria posibilidad, la distribución que resulta viene dada por:

$$u \text{ es } \min(\tau_1(P(u)), \tau_2(P(u)))$$

que es equivalente a:

$$u \text{ es } (\min(\tau_1(x), \tau_2(x)))(P(u))$$

Del mismo modo si las distribuciones son distribuciones de posible posibilidad entonces la distribución resultante viene dada por:

$$u \text{ es } \max(\tau_1(P(u)), \tau_2(P(u)))$$

que es equivalente a:

$$u \text{ es } (\max(\tau_1(x), \tau_2(x)))(P(u))$$

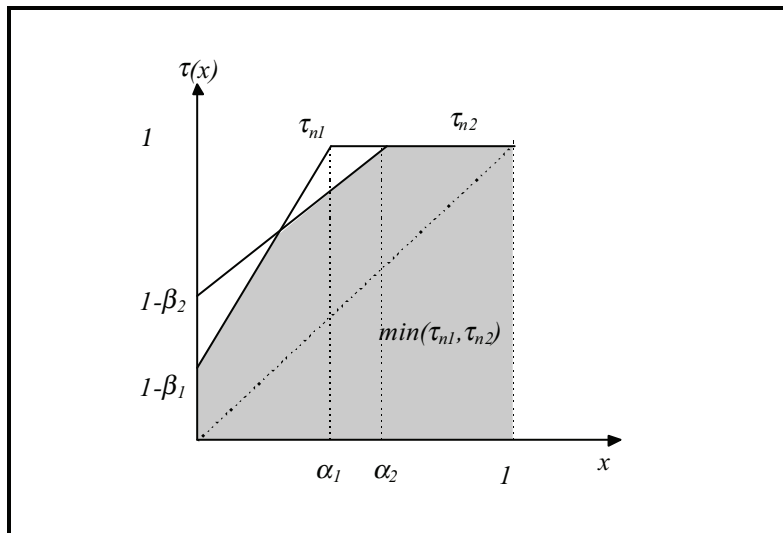
El interés de analizar el comportamiento de la agregación de valores de verdad borrosos reside en la posibilidad de analizar como se agregan entre sí los grados de verdad y los grados de certidumbre parciales, según la interpretación propuesta en esta tesis.

6.8.1 Agregación de valores de necesidad borrosos

Dados dos valores de necesidad borrosos para una misma proposición, el objetivo es determinar el valor de necesidad final de dicha proposición. Sean $\tau_{n1}(x) = I_n(\alpha_1, \beta_1, x)$ y $\tau_{n2}(x) = I_n(\alpha_2, \beta_2, x)$ dos valores de necesidad borrosos para una misma proposición p . El valor de necesidad final se obtiene como:

$$\tau_n(x) = \min(\tau_{n1}(x), \tau_{n2}(x)) = \min(I_n(\alpha_1, \beta_1, x), I_n(\alpha_2, \beta_2, x))$$

La siguiente figura ilustra de forma gráfica el resultado de la agregación:



Aunque el resultado final no es del tipo general $I_n(\alpha, \beta, t)$ si puede señalarse que los coeficientes de la distribución final α y β pueden obtenerse a partir de las distribuciones de partida, de forma que se tiene:

$$\alpha = \max(\alpha_1, \alpha_2)$$

$$1 - \beta = \min(1 - \beta_1, 1 - \beta_2) \Rightarrow \beta = \max(\beta_1, \beta_2)$$

Lo que significa que el grado de verdad de la proposición p es el máximo de los grados de verdad y que el grado de certidumbre es también el máximo de los grados de certidumbre.

Teniendo en cuenta cómo se agregan los grados de verdad y de certidumbre, y aunque la distribución final no pertenezca a la familia de valores propuestos, puede obtenerse una aproximación razonable suponiendo que viene dada por:

$$\tau_n(x) = I_n(\max(\alpha_1, \alpha_2), \max(\beta_1, \beta_2), x)$$

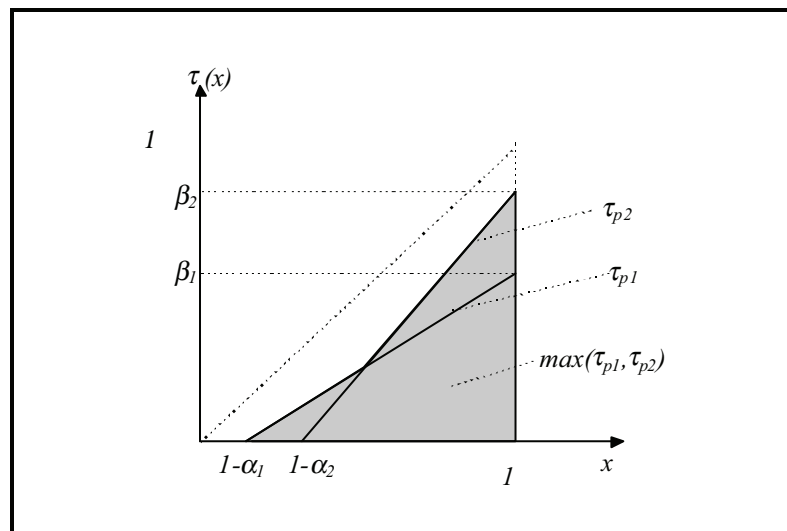
con lo que se simplifica notablemente su procedimiento de cálculo.

6.8.2 Agregación de valores de posibilidad borrosos

Cuando la información sobre el grado de verdad de una proposición p viene expresado mediante dos valores de posibilidad borrosos, el valor de posibilidad borroso final debe obtenerse agregándolos mediante el operador max . Así dados $\tau_{p1}(x) = I_p(\alpha_1, \beta_1, x)$ y $\tau_{p2}(x) = I_p(\alpha_2, \beta_2, x)$ dos valores de posibilidad borrosos para una misma proposición p , el valor de posibilidad final se obtiene como:

$$\tau_p(x) = \max(\tau_{p1}(x), \tau_{p2}(x)) = \max(I_p(\alpha_1, \beta_1, x), I_p(\alpha_2, \beta_2, x))$$

La siguiente figura ilustra de forma gráfica el resultado de la agregación:



Aunque al igual que antes, el resultado final no es del tipo general $I_p(\alpha, \beta, x)$ pero sí puede señalarse que los coeficientes de la distribución final α y β pueden obtenerse a partir de las distribuciones de partida:

$$1 - \alpha = \min(1 - \alpha_1, 1 - \alpha_2) \Rightarrow \alpha = \max(\alpha_1, \alpha_2)$$

$$\beta = \max(\beta_1, \beta_2)$$

El posible grado de verdad de la proposición p viene dado por el máximo de los posibles grados de verdad, mientras que el posible grado de certidumbre (o grado de posibilidad) viene también dado por el máximo de los posibles grados de certidumbre.

El resultado final puede aproximarse mediante un valor de posibilidad borroso perteneciente a la familia general propuesta, de forma que se mantiene cerrado el conjunto de valores de posibilidad y se simplifica notablemente su cálculo:

$$\tau_p(x) = I_p(\max(\alpha_1, \alpha_2), \max(\beta_1, \beta_2), x)$$

Se concluye por tanto que los grados de verdad y de certidumbre se agregan siempre del mismo modo, de forma disyuntiva, independientemente de si hacen referencia al dominio de lo necesario o al dominio de lo posible.

7. Metodología de diseño de un sistema de diagnóstico basado en la teoría de conjuntos borrosos

En este capítulo se propone una metodología para el diseño de sistemas de diagnóstico, basada en las herramientas analizadas en los capítulos anteriores de esta tesis.

El sistema de diagnóstico que se describe utiliza como sistema de detección de anomalías el que se propone y detalla en [Muñoz, 96] y que se repasa brevemente en este capítulo, al que se ha incorporado un mecanismo que permite integrar de forma eficiente la información que genera para ser posteriormente usada por el sistema de identificación de anomalías. Este a su vez es un sistema experto de diagnóstico que utiliza reglas de conocimiento con modelos borrosos, y su estructura, la de las reglas de conocimiento y el motor de inferencias son descritos en detalle.

Por último se presenta un ejemplo de aplicación del sistema de identificación de anomalías propuesto al diagnóstico de transformadores de potencia, mediante el análisis de los gases disueltos en el aceite que se utiliza como aislante y refrigerante dentro de la cuba en la que están sumergidos.

7.1 Organización del capítulo

En primer lugar y a modo de introducción se enumeran algunos de los principales pasos para el diseño de un sistema de diagnóstico según la metodología propuesta en esta tesis.

En segundo lugar se analiza brevemente la descomposición en componentes y elección de variables significativas del proceso, y se proponen estructuras de representación de conocimiento para el conocimiento estructural y su uso posterior por los distintos módulos del sistema de diagnóstico. También se analizan los distintos modos de describir el valor de una variable, y se proponen algunas recomendaciones para la elección de sus estados.

En el siguiente punto se describe el mecanismo de integración del módulo de detección de anomalías propuesto en [Muñoz, 96], de forma que sus salidas sean directamente representables en forma de hechos interpretables por el sistema de diagnóstico, y en concreto por el módulo de identificación de anomalías.

A continuación se describe en detalle la estructura del módulo de identificación de anomalías propuesto en la tesis, y los esquemas de representación de conocimiento para enunciar, almacenar y utilizar las reglas de conocimiento, basados en el uso de técnicas borrosas. Se describe la concepción general del sistema de diagnóstico, las estrategias de diagnóstico escogidas, y el motor de inferencias utilizado para razonar.

Por último se estudia como caso ejemplo el diagnóstico de un transformador de potencia, basado en el análisis de los gases disueltos en el aceite, aislante y refrigerante utilizado dentro de la cuba en la que está sumergido el transformador. El objetivo de este ejemplo es ilustrar algunas de las principales aportaciones de esta tesis, en concreto el uso de reglas con modelos borrosos para modelar adecuadamente las relaciones semánticas existentes entre consecuentes y antecedentes, la determinación de estos modelos, y el mecanismo de inferencia simplificado basado en el cómputo de grados de verdad y de certidumbre, ya sea en el dominio de lo necesario como en el dominio de lo posible.

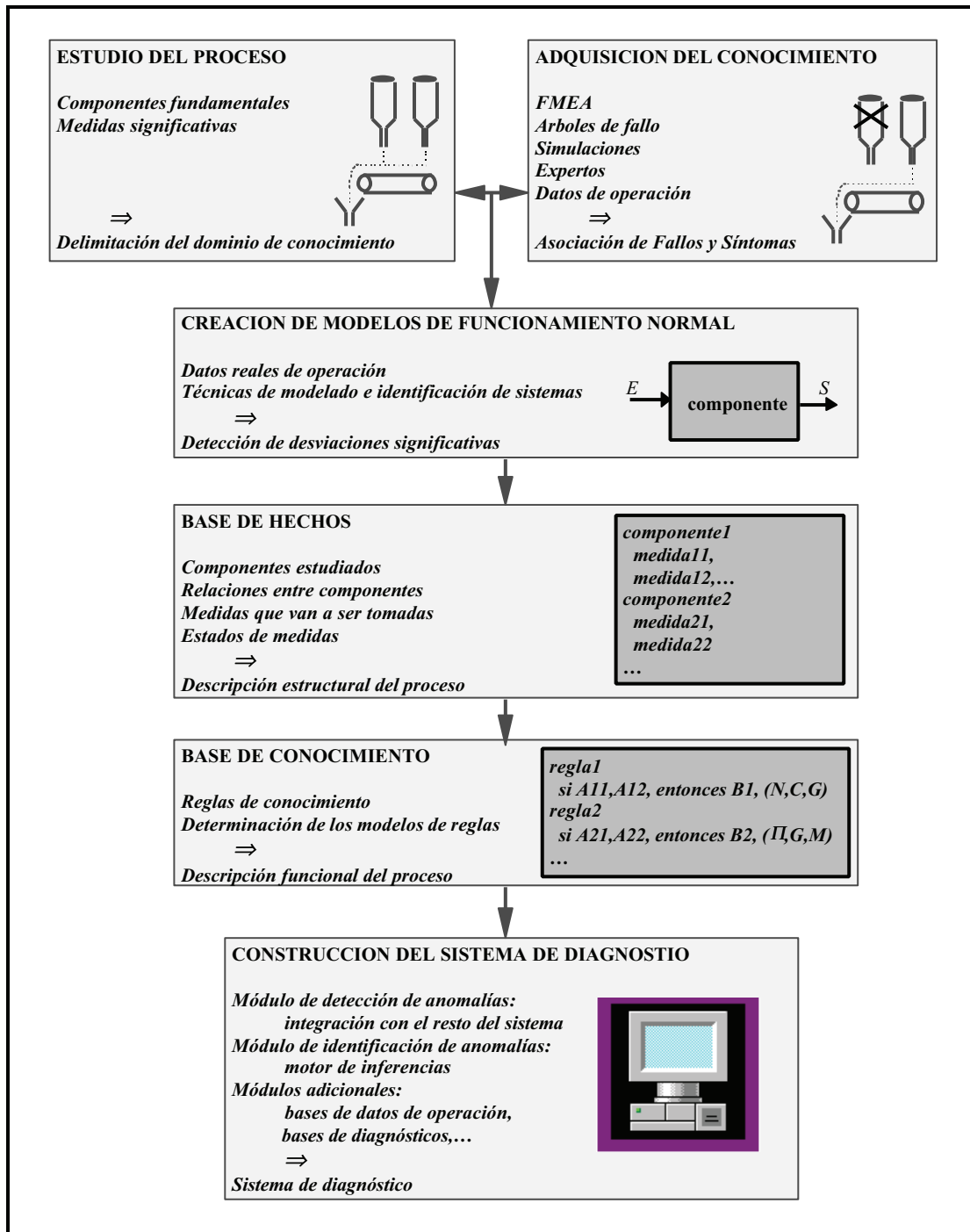
7.2 Pasos necesarios para la construcción de un sistema de diagnóstico

La construcción de un sistema experto de diagnóstico, según la estructura general que se comentó en el capítulo 2 y la metodología que se propone en esta tesis, pasa por los siguientes puntos:

- Análisis de la estructura del proceso para delimitar el dominio de conocimiento e identificar los componentes y sus medidas más relevantes de cara al diagnóstico.
- Adquisición del conocimiento. Para la adquisición del conocimiento es necesario llevar a cabo un análisis de los fallos que el proceso puede presentar, así como detección de los síntomas que puedan ser utilizados para su identificación. Para enumerar exhaustivamente los posibles fallos que puedan presentarse, sus síntomas asociados, su gravedad en el conjunto del proceso, etc., es recomendable hacer uso de técnicas clásicas de fiabilidad como la elaboración de un FMEA ([Pau, 81], [Sanz, 92], [Hopgood, 93], [Hunt, Price, Lee, 93]) o de árboles de fallo ([Pau, 81], [Sanz, 92]), o sus extensiones para representar la incertidumbre, como por ejemplo mediante el uso de la teoría de conjuntos borrosos (véase por ejemplo [Cayrac, Dubois, 96] para la elaboración de FMEA borrosos, o [Gmytrasiewicz, Hassberger, Lee, 90], [Ulieru, 94] para árboles de fallo borrosos). Otra fuente importante de obtención de conocimiento es la simulación del proceso en funcionamiento normal o en presencia de alguna anomalía mediante el uso de modelos, ya sean numéricos ([Ulieru, Isermann, 93], [Goode, Chow, 95]) o cualitativos ([Kitamura et al., 89], [Villar, 91], [Muñoz, Villar, Alba, 91]). Sin embargo la principal fuente de conocimiento deben ser las entrevistas con expertos, ya sea de cara a formular el conocimiento de diagnóstico, ya sea de cara a refinar el conocimiento inicialmente adquirido con alguno de los métodos anteriores. No hay que olvidar además, que dado que en general no se dispone de un conjunto de medidas suficientemente exhaustivo para la identificación de los posibles fallos, éstos no quedarán unívocamente determinados, y ya desde la fase de adquisición del conocimiento deberá escogerse algún mecanismo de representación de incertidumbre, como por ejemplo la teoría de conjuntos borrosos.
- Creación de modelos de funcionamiento normal de los componentes, con objeto de predecir medidas representativas de su comportamiento, de forma que puedan detectarse desviaciones entre el comportamiento medido y predicho, que serán indicio de la existencia de posibles anomalías. Este paso se llevará a cabo según se especifica en [Muñoz, 96], y se describe brevemente en el apartado 7.4.

- Creación de la base de hechos que contiene la descripción estructural del proceso (componentes, relaciones entre componentes, medidas) mediante su codificación con algún mecanismo de representación apropiado (véase el apartado 2.3.1 del capítulo 2). El detalle alcanzado en la descripción estructural debe ser acorde con las necesidades que se detecten para la formulación del conocimiento funcional. Es fundamental una correcta descripción de los estados de las medidas, pieza clave para la formulación del conocimiento funcional del proceso. La descomposición en componentes se aborda en el apartado 7.3 y en el apartado 7.3.2. se analiza algunos aspectos relacionados con las medidas del sistema, con el modelado del valor de una medida y con la descripción de sus estados.
- Utilizando el conocimiento funcional relativo a los fallos y sus síntomas asociados, creación de la base de conocimiento a partir del conocimiento adquirido en los puntos anteriores, mediante el enunciado de reglas borrosas y su representación con algún esquema apropiado. Para cada regla que se enuncie deberá identificarse su tipo semántico para poder asignarle un modelo borroso apropiado. Los modelos de reglas y su identificación se tratan en el punto 7.5.4 y en el ejemplo del apartado 7.6.
- Construcción del sistema de detección de anomalías y del módulo que integra sus salidas con la base de hechos del sistema de identificación de anomalías, que se aborda en el apartado 7.4.
- Construcción del sistema de identificación de anomalías basado en la utilización de reglas de conocimiento con modelos borrosos. La estructura de este sistema se describe en detalle a lo largo del apartado 7.5.
- Construcción de los gestores de las bases de datos de operación y diagnósticos para un posterior refinamiento del sistema de diagnóstico.

El proceso anterior queda resumido en la siguiente figura:



7.3 Identificación de componentes y medidas del proceso

El primer paso para llevar a cabo las tareas del diagnóstico consiste normalmente en hacer una descomposición jerárquica del proceso en componentes. Para poder representar este conocimiento es recomendable utilizar técnicas de representación estructuradas, apropiadas para almacenar información en forma de árbol. Los atributos y estructuras de representación del conocimiento estructural que aquí se proponen son extensiones de las ya utilizadas en el sistema experto TRAFES descrito en [Sanz, Villar, Palacios, 95].

7.3.1 Componentes del proceso

Para obtener la descripción de la estructura del proceso organizada en componentes son recomendables, entre otros, los siguientes pasos:

- Identificar y describir los componentes del proceso especificando su relación jerárquica. No es necesario considerarlos todos de forma exhaustiva, sino solo aquellos que se usen para formular posteriormente el conocimiento funcional. Para ello es necesario nombrar cada componente del proceso, y especificar en qué unidad superior se integra (a qué componente pertenece). Además muchos de los componentes del proceso pueden normalmente identificarse como pertenecientes a un mismo tipo, y por tanto compartir un conjunto de propiedades, ya sea un mismo conocimiento funcional, ya sea un mismo conjunto de medidas estáticas (aunque probablemente con distintos valores). Estas características hacen que el uso de un mecanismo de representación estructurada (como los frames o los objetos, véase el apartado 2.5.2) sea especialmente adecuado para almacenar este tipo de información. Además ambos mecanismos admiten la herencia entre tipos de datos, lo que permite definir tipos y subtipos de componentes. En el caso concreto de los objetos, cuando se utilizan lenguajes compilados, el uso de la herencia exige que los tipos de componentes estén codificados directamente en el lenguaje, lo que resta mucha transportabilidad al sistema en su conjunto, por lo que a menudo suele simularse internamente, con un cierto coste adicional. No ocurre así en los lenguajes interpretados, que aunque más lentos, son en general más flexibles. El uso de herramientas específicas de desarrollo de sistemas expertos puede facilitar la tarea pero puede también llevar a cuellos de botella difícilmente salvables, ya que hay que adaptarse a las herramientas y limitaciones que tienen.
- Especificar aquellas relaciones entre componentes necesarias para poder formular el conocimiento funcional. Así por ejemplo deberán especificarse relaciones del tipo *conectado_con* cuando el conocimiento de este tipo de relación sea necesario para

formular alguna de las reglas de conocimiento, caso típico de las plantas químicas, circuitos hidráulicos, circuitos de refrigeración o de lubricación, etc.

- Especificar para cada uno de los componentes definidos en los pasos anteriores las medidas estáticas que los describen, necesarias para formular el conocimiento funcional. Las medidas estáticas de los componentes se corresponden básicamente con aquellas características físicas que permanecen constantes independientemente del punto de operación del componente (véase el apartado 2.3.1 del capítulo 2). Las medidas estáticas pueden estar a menudo asociadas a tipos de componentes más que a componentes concretos. Esto hace que sea conveniente desarrollar algún tipo de herencia que permita que las características definidas para un tipo de componente determinado sean comunes a todos los componentes de dicho tipo.
- Especificar para cada uno de los componentes todas aquellas medidas dinámicas que los describen, necesarias para formular el conocimiento funcional. Las medidas dinámicas de los componentes se corresponden básicamente con aquellas características físicas que dependen del punto de operación del componente, y cuyo valor es adquirido o bien a través de un sistema de adquisición de datos desde los sensores instalados en los componentes, o bien a partir de análisis de laboratorio o inspecciones visuales (véase el capítulo 2).
- Especificar para cada componente sus medidas calculadas. Este tipo de medida no mencionado anteriormente puede en general ser el resultado de una expresión aritmética entre otras medidas dinámicas y estáticas del sistema. Su tratamiento es similar al de las medidas dinámicas puesto que su valor puede variar cada vez que el sistema adquiere un nuevo dato. Sin embargo para flexibilizar su uso es conveniente tener previsto algún mecanismo que permita definir expresiones aritméticas y evaluarlas en continuo, para calcular los valores de estas medidas, lo que suele requerir el uso de un interprete de expresiones aritméticas.

Como hemos mencionado es conveniente diferenciar entre tipo de componente y componente. Mientras que tipo de componente se corresponde con el concepto abstracto de clase en la terminología de la orientación a objetos, componente se corresponde con lo que comúnmente se denomina instancia de la clase. Los tipos de componentes permiten definir una serie de atributos comunes a una gama de componentes, mientras que el componente permite definir componentes concretos del proceso. Para definir un tipo de componente se puede utilizar una estructura como la siguiente:

<i>TIPO_COMPONENTE</i>	<i>nombre_tipo</i>
<i>TIPO</i>	<i>nombre_tipo</i>
<i>CONTIENE_A</i>	<i>nombre_tipo1, nombre_tipo2, ...</i>

<i>RELACION_1</i>	<i>nombre_tipo11, nombre_tipo12</i>
<i>RELACION_2</i>	<i>nombre_tipo21, nombre_tipo22</i>
...	
<i>MEDIDAS</i>	<i>medida1, medida2, ...</i>
<i>MODELOS</i>	<i>modelo1, modelo2, ...</i>

El atributo *tipo_componente* comienza la declaración de un nuevo tipo, mientras que el atributo *tipo* permite definir varios niveles de herencia, de forma que un tipo de componente puede heredar los atributos, relaciones, etc. de otros tipos de componentes previamente definidos.

El atributo *contiene_a* permite definir qué tipo de componentes contiene el tipo de componente definido originalmente. Se puede por ejemplo definir el tipo *coche*, y enunciar como tipos que incluye las *ruedas*, el *motor*, la *carrocería*, etc, así como que relaciones guardan entre sí. Por ejemplo el motor está conectado con la transmisión y la transmisión con las ruedas, etc.

Por otro lado el tipo de componente puede tener asociada unas determinadas medidas características. En el caso del coche parece razonable definir las medidas *velocidad*, *revoluciones*, *peso*, etc., comunes a cualquier coche estándar.

Definir un componente del tipo *coche* permite que automáticamente queden definidos los tipos de componentes y las medidas que contiene. Los tipos de componentes solo suponen una ventaja cuando sea necesario definir más de un componente de un mismo tipo.

Dado que la detección de anomalías del sistema propuesto se basa en el uso de modelos de funcionamiento normal también parece necesario definir una estructura para la representación de modelos, y proporcionar la flexibilidad suficiente para definir tipos de modelos al igual que se ha hecho con los componentes. Tipos de modelos pueden ser modelos muy abstractos, diferenciados por la técnica matemática empleada (por ejemplo modelos de regresión, series temporales, modelos basados en redes neuronales, etc.) y modelos más concretos asociados a tipos de componentes para los que ya se hacen explícitas las variables de entrada y salida del modelo.

Para describir un componente se puede utilizar una estructura similar a la anterior:

<i>COMPONENTE</i>	<i>nombre_componente</i>
<i>TIPO</i>	<i>tipo_componente</i>
<i>PERTENCE_A</i>	<i>nombre_componente</i>
<i>RELACION_1</i>	<i>nombre_componente11, nombre_componente12, ...</i>

<i>RELACION_2</i>	<i>nombre_componente21, nombre_componente22, ...</i>
<i>MEDIDAS</i>	<i>medidad1, medidad2, ...</i>
<i>MODELOS</i>	<i>modelo1, modelo2, ...</i>

El atributo *tipo* especifica a qué tipo de componente pertenece el componente en cuestión, y puede en algunos casos estar vacío. El atributo *pertenece_a* permite especificar la relación jerárquica entre componentes, y el resto son atributos que ya han sido definidos.

Puede optarse por no enumerar las medidas de un componente o tipo de componente en su definición, siempre y cuando se especifique a qué componente o tipo de componente pertenecen cuando sean posteriormente definidas. Sin embargo sí parece conveniente de cara a una mayor legibilidad de la descripción estructural del proceso en cuestión.

En la especificación de los modelos asociados a componentes concretos deberán incluirse los parámetros del modelo, obtenidos a partir de datos de funcionamiento normal, y/o posiblemente los mecanismos de reajuste automático en el caso en que se considere necesario, temas estos complejos que quedan fuera del alcance de esta tesis.

7.3.2 Medidas del proceso

Al igual que se ha hecho con los componentes y podría hacerse con los modelos, parece interesante considerar la posibilidad de definir tipos de medidas asociados o no a tipos de componentes. Algunos de los atributos podrán ir asociados a un tipo de medida determinado mientras que otros serán propios de medidas concretas. Algunos de los atributos más importantes son:

- Nombre del tipo de medida o de medida concreta, para poder identificar y diferenciar las medidas del sistema. En el caso en que se utilicen nombres de tipos de medidas, una medida puede quedar definida directamente por su tipo y por el componente concreto al que pertenezca. Esto puede facilitar el uso de los llamados *comodines*, de forma que el conocimiento se exprese de forma genérica para tipos de medidas y tipos de componentes, y se particularice en tiempo de ejecución o durante la lectura de la descripción estructural a cada una de las correspondientes medidas del proceso.
- Componente o tipo de componente al que pertenece la medida o tipo de medida.
- Tipo de medida, estática, manual, automática o calculada (no confundir con un tipo general de medida). Debido a sus distintas propiedades es necesario utilizar un atributo para diferenciar las medidas que hemos denominado manuales de las medidas que

hemos denominado automáticas (véase el apartado 2.3.1). En general unas y otras provienen de sitios distintos, y por tanto son adquiridas por el sistema de diagnóstico desde módulos diferentes, y posiblemente mediante procedimientos diferentes. Además su tratamiento matemático suele también ser diferente, ya que no es lo mismo, por ejemplo, calcular la tendencia de una medida que se toma con una periodicidad elevada, que la de una medida cuyos datos en general escasos, llegan además de forma aperiódica. Otro tipo de medida que ya ha sido mencionado son las medidas calculadas, obtenidas a partir de una expresión matemática función de otras medidas estáticas, dinámicas o manuales del sistema. El valor de estas medidas deberá actualizarse cada vez que se obtenga un nuevo valor para algunas de las medidas de las que dependa.

- Tipo de valor. Como ya hemos visto las medidas también pueden distinguirse por el tipo de valores que toman. Así podemos diferenciar las medidas digitales (o binarias) que sólo pueden tomar dos valores, las medidas simbólicas cuyo valor estará en general dentro de un conjunto limitado de cadenas de caracteres, y las medidas numéricas cuyo valor puede pertenecer a un subconjunto de la recta real, subconjunto que puede estar limitado por los rangos de sensibilidad de los sensores utilizados para su medición o por características físicas de la propia medida. En general, debido al tipo de razonamiento que se propone más adelante para la estrategia de diagnóstico, el valor de una medida vendrá caracterizado por dos tipos de valores, el valor necesario y el valor posible. En el caso particular de medidas numéricas, el valor necesario vendrá modelado por una distribución de necesaria posibilidad, mientras que el valor posible por una distribución de posible posibilidad. En el apartado 7.3.2.1 se describen distintas formas de caracterizar el valor de una medida.
- Otros atributos de interés. Para facilitar la interacción con el usuario es conveniente añadir ciertos atributos a las medidas, que aunque irrelevantes para el proceso de inferencia, pueden facilitar la legibilidad de los resultados al usuario. Ejemplos típicos son las unidades, la escala de cara a representaciones gráficas de la evolución de los valores de la medida, etc.
- Estados de la medida. Para una representación eficiente y compacta del conocimiento funcional es necesario definir de la mejor manera los posibles estados de la medida. Los valores de las medidas y la definición de sus estados se tratan en los puntos 7.3.2.1 y 7.3.2.2.
- Propiedades. Una forma de agrupar las medidas en torno a características comunes, de cara a la formulación del conocimiento funcional (apartado 2.3.1) es el uso de lo que hemos denominado *propiedades*. Así a aquellas medidas que se desean agrupar se les asocia una propiedad común. La agrupación de medidas es útil de cara a formular de

forma más compacta el conocimiento funcional y recuerda a lo que comúnmente se denomina *comodines* ([Sanz, Villar, Palacios, 95]). Así el conocimiento puede hacer referencia a todas aquellas medidas con una determinada propiedad, evitándose de este modo la enumeración exhaustiva.

Una posible estructura para definir un tipo de medida puede ser:

<i>TIPO_MEDIDA</i>	<i>nombre_medida</i>
<i>TIPO_COMPONENTE</i>	<i>nombre_componente</i>
<i>TIPO</i>	<i>tipo</i>
<i>SENSOR</i>	<i>sensor</i>
<i>EXPRESION</i>	<i>expresión</i>
<i>TIPO_DINAMICA</i>	<i>estática, automática o manual</i>
<i>TIPO_VALOR</i>	<i>digital, simbólico o numérico</i>
<i>UNIDADES</i>	<i>unidades</i>
<i>RANGO</i>	<i>rango</i>
<i>ESTADOS</i>	<i>estado1, estado2, ...</i>
<i>PROPIEDADES</i>	<i>propiedad1, propiedad2, ...</i>

Y la misma estructura ligeramente modificada puede ser utilizada para definir medidas concretas (a diferencia de tipos de medidas) del sistema:

<i>MEDIDA</i>	<i>nombre_medida</i>
<i>COMPONENTE</i>	<i>nombre_componente</i>
<i>TIPO</i>	<i>tipo_medida</i>
<i>SENSOR</i>	<i>sensor</i>
<i>TIPO_DINAMICA</i>	<i>estática, automática, manual o calculada</i>
<i>EXPRESION</i>	<i>expresión</i>
<i>TIPO_VALOR</i>	<i>digital, simbólico o numérico</i>
<i>UNIDADES</i>	<i>unidades</i>
<i>RANGO</i>	<i>rango</i>
<i>ESTADOS</i>	<i>estado1, estado2, ...</i>
<i>PROPIEDADES</i>	<i>propiedad1, propiedad2, ...</i>
<i>VALOR</i>	

La mayor parte de los atributos enunciados han sido ya comentados. El atributo *valor* es propio de las medidas estáticas ya que son las únicas para las que éste no se obtiene en continuo a través de los mecanismos de adquisición de datos. El atributo *expresión* es propio de las medidas calculadas y se utiliza para definir la expresión matemática que permitirá calcular su valor en continuo cada vez que lleguen nuevos valores de las medidas dinámicas o manuales involucradas. Para evitar errores es necesario determinar un

pequeño árbol de evaluación de medidas calculadas, de forma que si alguna de ellas depende de otras medidas también calculadas su valor se haya actualizado previamente. En cierto modo la creación de este árbol es un proceso similar a la creación del árbol de reglas propuesto en el apartado 5.8.2.

7.3.2.1 Descripción del valor de una medida

La descripción que el valor de una medida puede tomar depende en general del tipo de valor de la medida, ya sea digital, simbólico o numérico, y de si se modelan o no características como la imprecisión o la incertidumbre. Por otro lado, y para el correcto funcionamiento del sistema de identificación de anomalías descrito a lo largo del apartado 7.5, es necesario distinguir entre valores posibles (modelados mediante distribuciones de posible posibilidad) y valores necesarios (modelados mediante distribuciones de necesaria posibilidad). Esto hace que haya que mantener internamente y de forma permanente estos dos tipos de valores, vigilando su consistencia como se describe en el apartado 5.3.2 del capítulo 5.

- Medidas digitales:

Para las digitales sus únicos valores posibles son los valores 0 ó 1 . Su valor puede ser descrito escogiendo alguno de los valores del conjunto $\{0,1\}$ (que no es más que una distribución de posibilidad clásica), en cuyo caso el valor de la medida se conoce con total certidumbre. Si por el contrario existe cierta incertidumbre que impide conocerlo con precisión, este puede describirse mediante una distribución de posibilidad en el universo $\{0,1\}$.

- Medidas simbólicas:

Para las medidas simbólicas existen varias formas más o menos completas de describir su valor, en función del conjunto de posibles valores que pueda tomar.

Supongamos que una medida s puede tomar sus valores en el conjunto $S=\{s_1, s_2, s_3, \dots\}$. Una primera opción consiste en describir el valor de la medida directamente mediante alguno de sus posibles valores s_i . En este caso no es posible modelar ningún tipo de incertidumbre, a menos que ésta esté directamente codificada en el significado de s_i , como por ejemplo sería el caso del conjunto:

$$S=\{alto, posiblemente_alto, \dots\}$$

Otra opción consiste en describir su valor mediante una distribución de posibilidad sobre el universo S , de forma que a cada s_i se le asocia un grado de posibilidad. En este caso se puede modelar la incertidumbre tal y como se modela en universos continuos, como se describió en el apartado 3.6.4.

Una tercera alternativa es asociar a cada valor simbólico s_i un modificador que describa su grado de verdad y su grado de certidumbre. En este caso se obtienen dos distribuciones numéricas definidas en S , una distribución $\alpha(s_i)$ de grados de verdad y una distribución $\beta(s_i)$ de grados de certidumbre. La primera distribución expresa en qué medida el valor de s puede no parecerse a s_i y por tanto solo tiene sentido cuando puedan darse parecidos parciales, o lo que es lo mismo cuando el concepto expresado por s pueda verificarse de forma gradual. La segunda distribución expresa la certidumbre que se tiene de que el concepto expresado por s_i pueda verificarse al menos parcialmente, y por tanto tiene también sentido para conceptos que no sean graduales, en cuyo caso un valor positivo $\beta(s_i)$ expresa que la certidumbre que se tiene en que el valor de s sea exactamente igual a s_i . A este respecto supongamos que la variable s es el estado en que se encuentra un determinado componente, descrito mediante un conjunto de diagnósticos. Así por ejemplo para el estado de una tubería podríamos tener el conjunto de valores posibles dado por:

$$S = \{ \text{obstrucción, rotura} \}$$

en cuyo caso los conceptos expresados por cada uno de los diagnósticos son graduales, ya que una obstrucción puede ser mayor o menor dependiendo de la sección de paso que le quede al fluido que circula por la tubería, y de igual modo una rotura puede también considerarse mayor o menor dependiendo de la cantidad de fluido que escape de la tubería.

El valor de la medida s queda entonces descrito por dos distribuciones que se complementan, una de grados de verdad y otra de grados de certidumbre, pudiendo ambas estar definidas en el dominio de lo posible, o en el dominio de lo necesario.

En el ejemplo de la tubería, la pareja de coeficientes (α_n, β_n) (que son el grado de verdad y el grado de certidumbre en el dominio de lo necesario) para el diagnóstico *rotura* tienen la siguiente interpretación:

- a) La certidumbre que se tiene en que exista una rotura, independientemente de su magnitud (grande o pequeña), viene dada por β_n . Si β_n es 0 no se tiene ninguna certidumbre de que haya rotura. El coeficiente β_n es el grado en que se considera necesario que el diagnóstico sea *rotura* (en mayor o menor grado).

- b) Si por $\alpha_n=1$ se entiende que la rotura es media entonces un valor menor de α_n expresa de forma cualitativa que la rotura es media, pudiendo tender a grande o a pequeña. Cuanto menor sea el valor de α_n mayor o menor puede ser la magnitud de la rotura. En el dominio de lo necesario la cualificación con un grado de verdad parcial de un diagnóstico permite incluir como posibles a diagnósticos conceptualmente próximos. Si por $\alpha_n=1$ se entiende que la rotura es grande, y no se consideran roturas mayores, entonces un valor menor de α_n expresa de forma cualitativa que la rotura es grande o ligeramente media. Cuanto menor sea α_n menor podría ser la magnitud de la rotura.

Por otro lado la pareja de coeficientes (α_p, β_p) (que son el grado de verdad y el grado de certidumbre en el dominio de lo posible) para el diagnóstico *rotura* tienen la siguiente interpretación:

- a) El grado de contradicción que existe entre el diagnóstico *rotura* y los hechos disponibles viene dado por $1-\beta_p$. Dicho de otra forma β_p se interpreta como el grado en que es posible que el diagnóstico sea *rotura*. Este grado permite determinar de entre todos los diagnósticos cuales son los más posibles, o lo que es lo mismo cuales son los menos contradictorios con la situación real.
- b) Si nuevamente suponemos que la magnitud implícita del diagnóstico *rotura* es media, entonces un valor $\alpha_p=1$ indica que la rotura podría ser media, sin afectar a otras posibles magnitudes. Valores menores expresan que la rotura sólo puede ser media en el grado expresado por α_p , aunque sin afectar nuevamente a otras posibles magnitudes. En el dominio de lo posible un grado de verdad parcial excluye al diagnóstico al que cualifica, sin afectar a otros diagnósticos próximos o menos próximos conceptualmente.

Las parejas de coeficientes (α_n, β_n) o (α_p, β_p) son en realidad una notación reducida de los valores de verdad borrosos de necesidad y posibilidad, descritos en el capítulo 6, apartado 6.7. Recordemos que la expresión general de los valores de necesidad borrosos es:

$$(N, \alpha_n, \beta_n) = I_n(\alpha_n, \beta_n, x) \text{ grado de verdad } \alpha_n, \text{ grado de certidumbre } \beta_n$$

donde I_n es la implicación general de necesidad dada por:

$$I_n(\alpha, \beta, y) = \min\left(\frac{\beta}{\alpha}(y - \alpha) + 1, 1\right)$$

y la de los de posibilidad:

$$(\Pi, \alpha_p, \beta_p) = I_p(\alpha_p, \beta_p, x) \text{ posible grado de verdad } \alpha_p \\ \text{ posible grado de certidumbre } \beta_p$$

donde I_p es la implicación general de posibilidad dada por:

$$I_p(\alpha, \beta, y) = \max\left(\frac{\beta}{\alpha}(y-1) + \beta, 0\right)$$

Por tanto las distribuciones de grados de verdad y de grados de certidumbre sobre el universo de los posibles valores se obtienen asociando a cada elemento del universo un valor de verdad borroso.

- Medidas numéricas:

Cuando las medidas son numéricas la forma más general de describir su valor es mediante una distribución de posibilidad (o distribuciones de posible y necesaria posibilidad). En este caso no tiene sentido asociar valores de necesidad o posibilidad borrosos a cada posible valor de la variable ya que el parecido del valor real de la medida con un determinado valor viene implícitamente determinado por los grados de posibilidad asociados a valores cercanos, debido a la relación de orden existente en los conjuntos numéricos ordinarios, mientras que la certidumbre viene expresada directamente por el grado de posibilidad asociado a cada posible valor.

Para representar las distribuciones de posibilidad de forma compacta proponemos el uso de los conjuntos borrosos λ_{np} -trapezoidales, definidos en el capítulo 3, que recordemos vienen descritos como se indica en la siguiente tabla:

λ_{np} -Trapezoidal	$P(u) = \begin{cases} \lambda_n & \text{si } u < a \\ \lambda_n + \frac{\lambda_p - \lambda_n}{a-b}(u-a) & \text{si } u \in [a, b] \\ \lambda_p & \text{si } u \in [b, c] \\ \lambda_n + \frac{\lambda_n - \lambda_p}{d-c}(u-d) & \text{si } u \in [c, d] \\ \lambda_n & \text{si } u > d \end{cases}$	
-----------------------------	--	--

En general cuando utilizamos un conjunto λ_{np} -trapezoidal para representar una distribución de necesaria posibilidad, el parámetro λ_p tomará siempre el valor 1. Cuando utilizemos un

conjunto λ_{np} -trapezoidal para representar a una distribución de necesaria posibilidad será el parámetro λ_n el que tomará el valor 0.

A una distribución de posibilidad numérica se le puede asociar valores de verdad borrosos para matizar su significado, dando lugar a una nueva distribución de posibilidad.

7.3.2.2 Estados de una medida

Los estados de una medida son los posibles valores (numéricos o cualitativos) que una medida pueda tomar, siempre que sean significativamente diferentes de cara a formular el conocimiento funcional del proceso. Por tanto son susceptibles de adoptar, en principio, cualesquiera de las formas propuestas en el apartado anterior para describir el valor de una medida.

En general el número de estados para cada medida depende del nivel de detalle del conocimiento experto, y de la medida en sí misma. Un número reducido de estados puede llevar a confundir situaciones distintas, de forma que el sistema no sea capaz de diferenciar fallos significativamente diferentes. Por otro lado un número excesivo puede conducir a enunciar un número excesivo de reglas todas ellas muy similares y que por tanto podrían haberse condensado en una sola, complicándose de este modo la formulación del conocimiento. En general los estados deben determinarse después de haber formulado, al menos de forma esquemática, el conocimiento funcional, puesto que es éste el que determina los estados necesarios para el correcto diagnóstico del sistema. La formulación de los estados por un lado, y el conocimiento funcional o más concretamente las reglas por otro, permite independizar el conocimiento del proceso del componente concreto al que se aplique. En la mayoría de los casos las reglas de conocimiento expresan conceptos de carácter general aplicables a muchos tipos de componentes similares, y son los estados de las medidas de cada uno de esos componentes los que permiten particularizar dicho conocimiento a sus características propias (variables de unos a otros, como son por ejemplo las dimensiones de un tanque de almacenamiento, o la potencia de un motor, etc.).

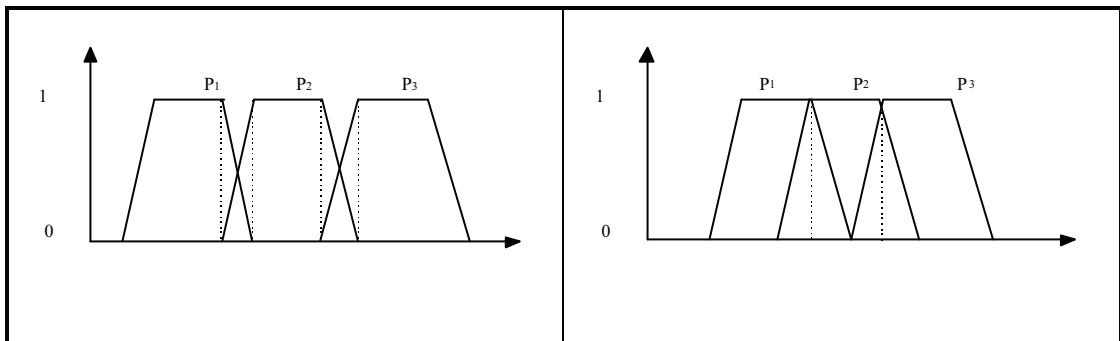
Por otro lado no es normal enunciar reglas en cuya parte antecedente deba modelarse algún tipo de incertidumbre, sino por el contrario suele ser suficiente modelar imprecisión y vaguedad que son las características principales de los conceptos utilizados por los expertos en el lenguaje hablado (véase el apartado 2.3.2.1). Por ejemplo en la regla:

si
temperatura del motor es alta
entonces
falta agua en el circuito de refrigeración
termostato averiado

podemos comprobar que en la parte antecedente no hay incertidumbre asociada con el concepto *alta*. El concepto *alta* será en general borroso para expresar el hecho de que unas temperaturas pueden considerarse más altas que otras, pero no existe incertidumbre alguna en su definición. La incertidumbre o la imprecisión aparecen normalmente asociadas a las observaciones. Si modelamos la temperatura medida mediante un conjunto borroso estaremos modelando principalmente la imprecisión del sensor, y será la comprobación de la hipótesis la que dará lugar a un grado de verdad y a un grado de incertidumbre, tal y como se definen en el apartado 4.9. La incertidumbre puede también aparecer en la parte consecuente debido a que en diagnóstico es necesario, en la mayoría de los casos, razonar por abducción. Así un conjunto de antecedentes puede estar asociado a varias conclusiones que pueden ser excluyentes, lo que las hace de por sí inciertas, aunque los antecedentes se verifiquen totalmente. En la regla ejemplo aunque se tenga la absoluta certeza de que la temperatura es alta, no se puede afirmar con total certidumbre cual es la causa que lo ha provocado. Por tanto, en base a estas consideraciones, los estados de las medidas podrán describirse mediante conjuntos borrosos trapezoidales.

El enunciar el conjunto de estados que una determinada medida puede tomar equivale al diseño de una partición borrosa. Por tanto es conveniente tener en cuenta criterios de completitud y de verificación del modus ponens para las reglas que posteriormente sean enunciadas. En este sentido es importante consultar las consideraciones realizadas en el apartado 5.6.3.2 sobre el diseño de particiones borrosas.

En general un buen punto de partida es utilizar una partición de las que mejor se adaptan a los sistemas de reglas con modelos de necesidad graduales. Por un lado y como se vio en dicho capítulo la partición debe orientarse preferentemente al uso de reglas con modelos de necesidad, ya que el objetivo prioritario es obtener conclusiones necesarias (es decir con un alto grado de fiabilidad). Por otro lado estas particiones proporcionan un recubrimiento que garantiza una buena completitud del sistema de reglas que posteriormente puedan ser enunciadas. Recordemos las dos particiones extremas obtenidas para modelos de necesidad fueron:



La primera partición garantiza la completitud y el MPE ante observaciones puntuales, pero no se garantiza la completitud cuando las observaciones son conjuntos con más de un elemento, ya que puede darse el caso de que el núcleo de la observación no esté incluido en ninguno de los soportes de los conjuntos de la partición. La segunda de las particiones resuelve este problema siempre que las observaciones vengan modeladas por conjuntos borrosos menores o iguales que los de la partición, pero tiene el inconveniente de que una observación puntual, aunque esté incluida en el soporte de alguno de los conjuntos de la partición, puede estar activando, en grado máximo, a más de una regla simultáneamente.

De cara a la completitud de la base de conocimiento es conveniente diseñar la partición de una medida de acuerdo al modelo que se utilice para los valores medidos de la misma. En efecto, si se comete un cierto error de medida es de esperar que los estados no pretendan distinguir más allá de lo que los medidores permiten. Si el error se considera constante a lo largo de todo el universo entonces las zonas de transiciones de los conjuntos de la partición tenderán a tener anchos también constantes, pero si el error es proporcional a la magnitud entonces parece razonable que también los anchos de las zonas de transiciones de los conjuntos dependan también de la magnitud de los elementos que contienen, correspondiendo este último caso a una visión logarítmica del universo de discurso de la medida, como ocurre por ejemplo con la variable *frecuencia* en el estudio de sistemas lineales. En este punto creemos que la completitud es un aspecto importante que debe primar sobre el MPE ya que la situación opuesta reflejaría la imposibilidad de catalogar un determinado valor de la variable, o lo que es lo mismo una cierta laguna del conocimiento disponible. Dado que se ha optado por disparar las reglas paralelas por separado (que se corresponde con el método FITA descrito en el apartado 5.6.1) deberá tenerse especial cuidado con la completitud, dado que este método se comporta mal cuando las observaciones son conjuntos anormalmente grandes, como por ejemplo uniones de conjuntos de la partición (véase el punto 5.6.1).

La segunda de las particiones además de mejorar la completitud tiene la ventaja adicional de permitir generar nuevos conceptos como por ejemplo $(P_1 \text{ y } P_2)$ o $(P_1 \text{ o } P_2)$ mediante cualquier t-norma o t-conorma, obteniéndose conjuntos normalizados, y unimodales en el

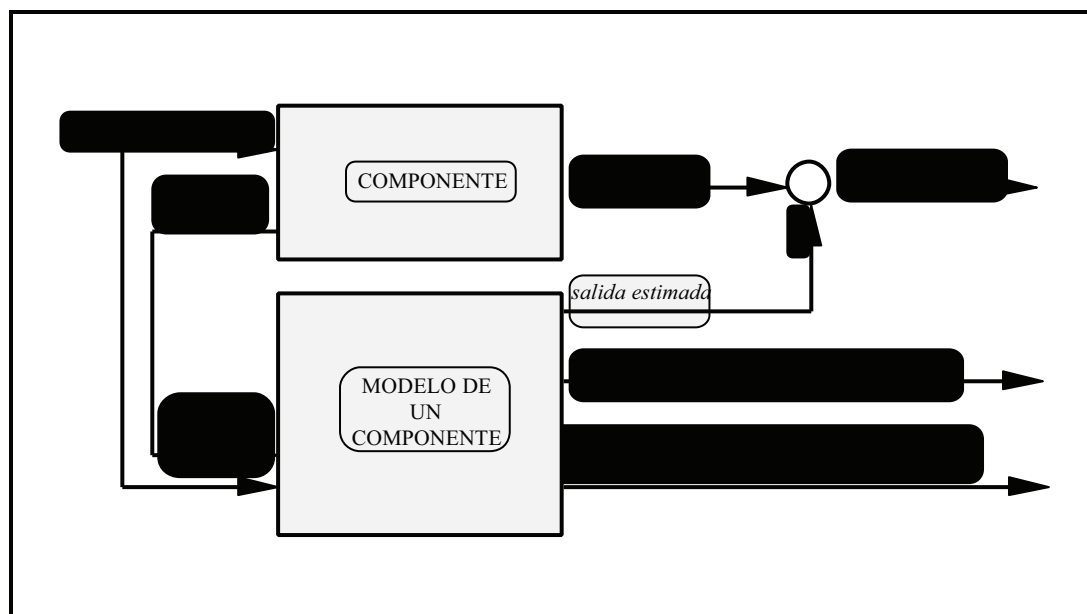
caso de que P_1 y P_2 sean adyacentes. Para la primera de las particiones solo sería unimodales para t-conormas grandes (pero no para t-normas). Este tipo de consideraciones deben también ser tenidas en cuenta de cara a formular el conocimiento.

Por otro lado, el hecho de que la partición se diseñe para modelos de necesidad graduales y no para modelos de necesidad de certidumbre viene motivado por el hecho de que en estos últimos modelos la condición de activación es más restrictiva y en general la completitud no es un requisito importante. En el caso en que se utilice un modelo de certidumbre con condición de activación modificada, entonces la partición apropiada es también la de los modelos graduales, al compartir misma condición de activación (véanse los apartados 4.9 y 4.11).

7.4 Módulo de detección de anomalías y su integración en el sistema de diagnóstico

El módulo de detección de anomalías es el encargado de detectar la existencia de anomalías en los componentes del proceso monitorizado. La detección de anomalías que proponemos se basa en el sistema que se detalla en [Muñoz, 96], y cuyas principales características recordaremos aquí brevemente.

Este sistema utiliza como punto de partida el modelado matemático de los componentes, con el objeto de conseguir modelos capaces de proporcionar estimaciones para medidas características del funcionamiento normal de los mismos, así como la confianza/fiabilidad que en la estimación proporcionada se tiene. La estructura general de este tipo de modelos queda esquematizada en la siguiente figura:



Sus principales características son:

- Cada modelo matemático está asociado con un componente cuyo comportamiento modela, y cada componente puede tener más de un modelo o no tener ninguno. Los parámetros del modelo se ajustan con datos reales de operación representativos del funcionamiento normal del componente.
- Sus entradas son medidas dinámicas del proceso.

- La primera salida es una estimación de alguna de las medidas dinámicas (o calculadas) del componente. Esta estimación puede ser comparada con los valores reales de la medida, obteniéndose un residuo $R(k)$, de forma que cuanto mayor es $R(k)$ puede deducirse que en general mayor es el grado de anomalía del componente. La estimación de la salida va acompañada de una estimación de la varianza del residuo $\sigma(k)$, lo que permite establecer una cota máxima y determinar cuándo la desviación entre la estimación y el valor real puede considerarse significativa.
- La otra salida del modelo es una medida de la confianza que se tiene en cada momento en las estimaciones que suministra el modelo, y que en concreto, para el tipo de modelos propuesto en [Muñoz, 96], se basa en la estimación de la función de densidad de la probabilidad de las entradas, $f_{dp}(e(k))$. Así cuando $f_{dp}(e(k))$ es alto las entradas pueden considerarse conocidas y la estimación fiable, mientras que cuando es bajo es debido a entradas poco conocidas (por estar escasamente representadas en el conjunto de entrenamiento del modelo) y la confianza en la estimación nunca podrá ser elevada, aunque no se disponga de criterios para invalidarla.

Con estas hipótesis de partida vamos a proponer un mecanismo de obtención de hechos que describen el conocimiento que sobre las anomalías de los componentes se tiene. Su formato es el siguiente:

anomalía en componente es τ

donde τ es en general un valor de verdad necesidad o posibilidad borroso, dependiendo de la fiabilidad de la estimación proporcionada por el modelo, y que indica tanto el grado de verdad de la anomalía como la certidumbre en la existencia de una anomalía (véase el apartado 6.7). Por tanto cada componente debe llevar asociada una medida llamada *anomalía* cuyo valor es un valor de verdad borroso. La generación de hechos con este formato permitirá utilizar toda la información suministrada por el módulo de detección de anomalías e integrarla de forma eficiente en el sistema de diagnóstico.

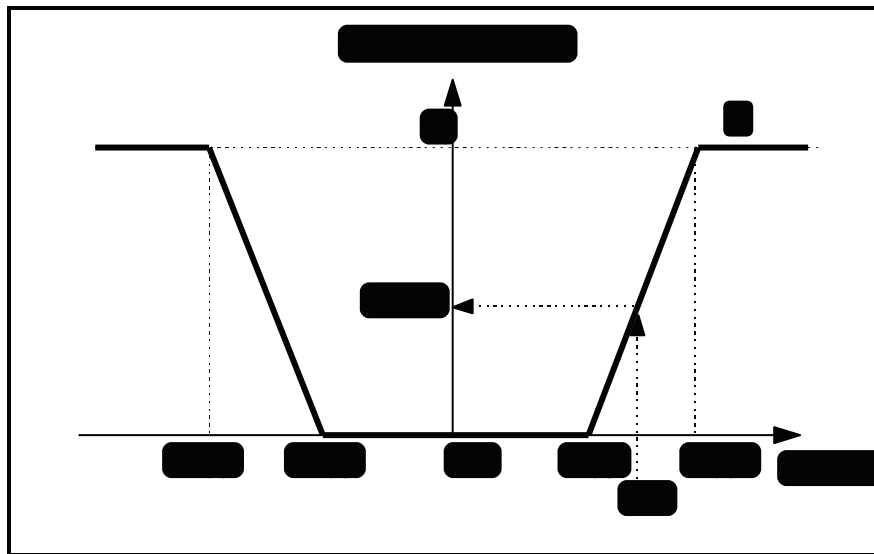
7.4.1 Obtención del grado de anomalía de un componente

A partir de sus entradas cada modelo es capaz de generar una estimación para el valor de su salida, así como una estimación de la varianza del error cometido en la estimación para ese mismo valor de las entradas, lo que le confiere una gran precisión en la tarea de detección de anomalías, puesto que la magnitud del residuo se mide respecto a su varianza

que se estima de forma local ([Muñoz, 96]), y no global, como típicamente hacen los modelos estadísticos tradicionales.

La detección de anomalías se basa en comparar el residuo con su estimación, utilizando un múltiplo de la varianza para determinar el error máximo admitido en la comparación. Admitiendo que el error de estimación se distribuye según una normal de media 0, entonces entre la estimación más dos veces la varianza y la estimación menos dos veces la varianza deberían encontrarse un 95% de los errores obtenidos. Por otro lado suponer que el error de estimación se distribuye según una normal de media 0 es precisamente uno de los contrastes utilizados habitualmente para validar el modelo utilizado para generar la estimación, por lo que la suposición es razonable.

Para obtener el grado de anomalía se propone cuantificarlo comparando el residuo con su varianza. Para ello definimos un conjunto borroso asociado al concepto *grado de anomalía*, que aparece representado en la siguiente figura:



El grado de anomalía será obtenido comparando el valor del residuo con su varianza $\sigma(k)$. De esta forma cuando el módulo del residuo sea inferior a $2\sigma(k)$, se considerará que es suficientemente pequeño y que por tanto el grado de anomalía es 0. A medida que el residuo se haga mayor que $2\sigma(k)$ el grado de anomalía irá aumentando hasta que cuando se haga mayor o igual que $3\sigma(k)$ se considerará que el grado de anomalía del componente es máximo, es decir 1.

Esta forma de obtener el grado de anomalía en función del residuo y de su varianza, puede parecer artificiosa, pero está en concordancia con la práctica habitual ([Muñoz, 96]). Aunque desde un punto de vista probabilístico comparar el residuo con su varianza permite estimar la certidumbre de que el residuo pueda considerarse 0 (y no su grado de anomalía como aquí se pretende), esto se debe en parte a que este enfoque solamente considera los estados "residuo 0" y "residuo no 0". Sin embargo parece obvio que la magnitud del residuo está claramente relacionada con la magnitud de la anomalía.

El enfoque propuesto elimina la necesidad de escoger un límite brusco para determinar si un residuo es o no síntoma de anomalía. El conjunto borroso propuesto es sólo orientativo y pueden escogerse otros parámetros diferentes para el conjunto borroso propuesto, y en el caso en que se puedan encontrar criterios físicos más precisos que permitan cuantificar el grado de anomalía en función de la magnitud del residuo, estos serán siempre preferibles al criterio aquí propuesto. La ventaja de utilizar un conjunto borroso reside precisamente en su robustez, ya que pequeñas variaciones en los límites del conjunto propuesto, en su geometría o en el valor del residuo, provocan también pequeñas variaciones en el grado de anomalía, y por tanto también en las consecuencias que del mismo puedan derivarse. Esto hace que la determinación de los parámetros del conjunto borroso propuesto no resulte crítica.

Definimos por tanto el grado de anomalía $A(k)$ en el instante k como:

$$A(k) = A(R(k))$$

estando A definido del siguiente modo:

$$A(R) = \begin{cases} 1 & \text{si } R \leq -3\sigma \text{ o } R \geq 3\sigma \\ 0 & \text{si } R \in [-2\sigma, 2\sigma] \\ \frac{-1}{2\sigma} R - \frac{1}{2} & \text{si } R \in [-3\sigma, -2\sigma] \\ \frac{1}{2\sigma} R - \frac{1}{2} & \text{si } R \in [2\sigma, 3\sigma] \end{cases}$$

La utilización del último valor del residuo para determinar el grado de anomalía es especialmente sensible a ruidos ya que éstos pueden provocar el sobrepaso puntual, pero no significativo, de sus límites máximo o mínimo. Así una ligera variación del residuo puede provocar variaciones bruscas en las conclusiones del sistema, cuando por ejemplo oscile en torno a su valor máximo (o mínimo) provocando alternativamente diagnósticos de normalidad y fallo. Para solventar este problema se utilizan habitualmente técnicas heredadas del control de calidad, de forma que sólo se considera que hay anomalía cuando

dos de tres residuos consecutivos sobrepasan su valor máximo, o cuando los últimos ocho valores están todos por encima o por debajo del valor esperado.

Para los sistemas tradicionales el grado de anomalía en el instante k se define como:

$$A(k) = \begin{cases} 1 & \text{si } |R(k)| \geq R_{max} \\ 0 & \text{en otro caso} \end{cases}$$

pudiendo tomarse por ejemplo $R_{max}=2\sigma$ o $R_{max}=3\sigma$. Se considera que el sistema debe detectar una anomalía en el instante k cuando se verifica:

$$A(k-2)+A(k-1)+A(k) \geq 2$$

Este mismo filtro puede también ser aplicado cuando el grado de anomalía pertenece al conjunto $[0, 1]$ en lugar de a $\{0, 1\}$, como en el sistema propuesto en esta tesis en el que el grado de anomalía viene dado por $A(k)=A(R(k))$. Para ello basta con redefinir un nuevo grado de anomalía efectivo como:

$$A^*(k)=(A(k-2)+A(k-1)+A(k))/3.$$

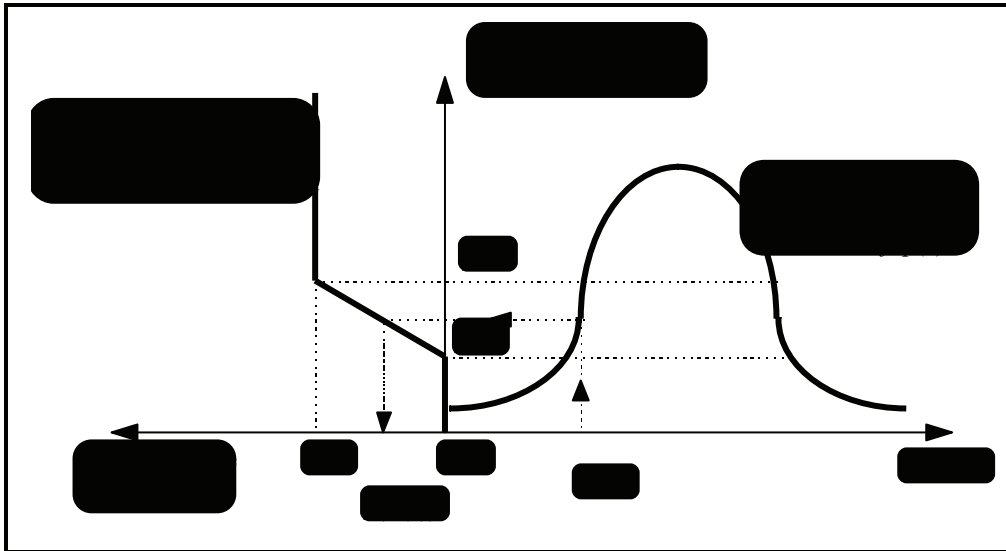
de forma que si $A^*(k)>0$ existe una anomalía en el grado indicado por $A^*(k)$, filtrándose de este modo posibles oscilaciones poco significativas.

7.4.2 Obtención de la certidumbre sobre el grado de anomalía de un componente

En este apartado se pretende convertir los grados de significación probabilística a grados de certidumbre obtenidos como grados de pertenencia a un conjunto borroso. Según las hipótesis de partida, los modelos utilizados deben ser capaces de dar una medida de la certidumbre de su estimación, o lo que es lo mismo de la certidumbre que se tiene sobre el grado de anomalía que generan. La razón de que esta certidumbre sea variable reside en el hecho de que en general los parámetros de los modelos de funcionamiento normal se ajustan utilizando datos reales de operación del proceso monitorizado. Sin embargo es muy difícil disponer de datos para cada uno de los modos de operación posibles, lo que hace que en muchos casos no se pueda garantizar la precisión de la estimación del modelo ante determinadas entradas. La estimación de la certidumbre de la predicción se basa en comprobar si las entradas actuales son conocidas por el modelo, o lo que es lo mismo, aparecen en los datos de entrenamiento utilizados para el ajuste de sus parámetros. Para

ello tal y como se propone en [Muñoz, 96] puede estimarse la función de densidad de probabilidad (f_{dp}) de las entradas del modelo utilizadas para su entrenamiento, y comparar el valor obtenido con las entradas en el instante k con un límite p_{lim} preestablecido a partir de un determinado nivel de significación. En general la certidumbre de la estimación para unas entradas $e(k)$ es grande si el valor $f_{dp}(e(k))$ es grande en comparación con p_{lim} .

Para obtener una medida borrosa de la fiabilidad de la estimación proponemos el uso de un conjunto borroso que permita realizar un contraste de hipótesis también borroso, de forma que la hipótesis no se acepte o se rechace únicamente a partir de la comparación de $f_{dp}(e(k))$ con un único umbral p_{lim} . La siguiente figura ilustra el mecanismo propuesto:



Como puede verse el procedimiento propuesto es el siguiente. En lugar de utilizar un único valor p_{lim} para decidir si la estimación es o no fiable, como se haría en un test de hipótesis tradicional, utilizamos dos valores p_{max} y p_{min} que permiten establecer una transición suave de la hipótesis no fiable a la hipótesis fiable. Estos valores pueden por ejemplo obtenerse como los valores de $f_{dp}(e)$ para niveles de significación respectivos de 90% y 99%. Si estuviésemos trabajando con distribuciones de probabilidad p_{min} sería el valor de probabilidad para el cual el 99% de los datos del conjunto de entrenamiento tienen una probabilidad mayor que p_{min} .

De este modo a partir de la función de densidad de probabilidad $f_{dp}(e)$ estimada mediante los datos de entrenamiento del modelo, dada una entrada cualquiera $e(k)$, definimos la certidumbre de la estimación proporcionada en el instante k mediante la expresión:

$$\beta_n(A(k)) = \begin{cases} 1 & \text{si } fdp(e(k)) \geq p_{max} \\ 0 & \text{si } fdp(e(k)) \leq p_{min} \\ \frac{fdp(e(k)) - p_{min}}{p_{max} - p_{min}} & \text{si } p_{min} < fdp(e(k)) < p_{max} \end{cases}$$

Así cuando $fdp(e(k))$ es menor que p_{min} la certidumbre de la estimación se toma como nula, cuando $fdp(e(k))$ está entre p_{max} y p_{min} la certidumbre viene dada por un valor entre 0 y 1, y cuando $fdp(e(k))$ es mayor que p_{max} la estimación del modelo se supone totalmente cierta y se le asigna la certidumbre 1. De este modo se obtiene una estimación razonable que nos indica cómo de necesario es que la estimación del modelo sea correcta.

Cuando $\beta_n(A(k))$ es mayor que 0 se puede utilizar la información del sistema de detección de anomalías para disparar modelos de necesidad. De este modo, a partir de un grado de anomalía $A(k)$ o $A^*(k)$ con grado de certidumbre dado por $\beta_n(A(k))$ puede generarse el hecho:

$$\text{anomalía en componente es } (N, \alpha, \beta) = I_n(A(k), \beta_n(A(k)), x)$$

o bien alternativamente:

$$\text{anomalía en componente es } (N, \alpha, \beta) = I_n(A^*(k), \beta_n(A(k)), x)$$

dependiendo de si se aplica o no el filtrado antes mencionado, y siendo I_n la implicación general de necesidad definida en el capítulo 4. Un hecho así formulado puede entonces ser utilizado para disparar reglas con el siguiente formato:

*si hay anomalía en motor
y ...
entonces ...*

Cuando $\beta_n(A(k))$ vale 0, pero el sistema de detección de anomalías predice la existencia de una anomalía, por ejemplo cuando $A(k) > 0$ o bien $A(k-2) + A(k-1) + A(k) \geq 2$, no existe forma a priori de determinar si hay o no una anomalía. Esto significa que puede racionalmente suponerse que la posibilidad de la anomalía es 1 al existir evidencia a su favor, ya que la estimación puede ser correcta a pesar de que no se disponga de información suficiente para probarlo. Puede incluso pensarse en matizar su posibilidad, al menos cualitativamente, dependiendo de la capacidad de generalización del modelo, aunque esto pueda ser muy complejo de cuantificar. En cualquier caso cuando $\beta_n(A(k))$

vale 0 pero el sistema de detección detecta la presencia de una anomalía puede generarse un hecho alternativo al anterior dado por:

$$\text{anomalía en componente es } (\Pi, \alpha, \beta) = I_p(A(k), I, x)$$

o bien:

$$\text{anomalía en componente es } (\Pi, \alpha, \beta) = I_p(A^*(k), I, x)$$

donde I_p es la implicación general de posibilidad general propuesta en el capítulo 4.

Las reglas con modelos de necesidad no pueden dispararse con hechos cuyo grado de certidumbre es 0. Sin embargo si mediante algún mecanismo se transforman las reglas de necesidad en reglas de posibilidad, tal y como se describe en el apartado 7.5, estas últimas sí pueden ser disparadas y obtener conclusiones que aunque menos fiables sean más informativas. Estas conclusiones pueden ser utilizadas posteriormente para determinar por ejemplo, en el caso de que el sistema de diagnóstico sea incapaz de identificar ningún fallo, la necesidad de reajustar los parámetros del modelo con los datos que han provocado la supuesta anomalía.

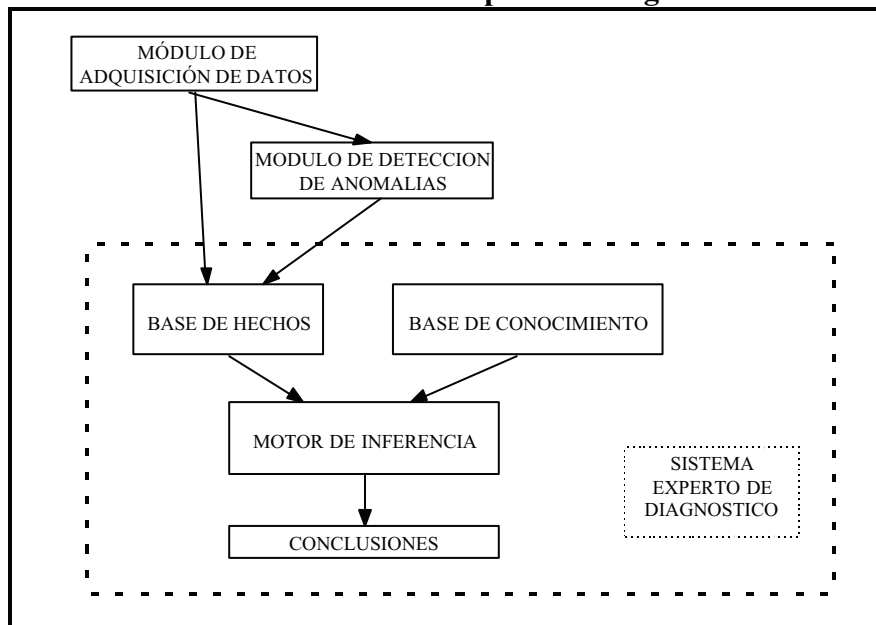
7.5 El módulo de identificación de anomalías: sistema experto de diagnóstico y reglas de conocimiento

7.5.1 El sistema experto de diagnóstico

El módulo de identificación de anomalías es el encargado de, a partir de la información procedente del módulo de detección de anomalías y de la información sobre la operación del proceso (conocida a través de las medidas dinámicas de los componentes), identificar las anomalías detectadas por el módulo de detección, proporcionando diagnósticos que expliquen o contradigan su existencia.

El módulo de identificación de anomalías propuesto se basa en un sistema experto de diagnóstico con una base de conocimiento compuesta de reglas con modelos borrosos. Como todo sistema experto, se compone de una base de hechos (también denominada memoria de trabajo), que contiene la descripción estructural del proceso así como los hechos que describen el conocimiento de operación del mismo, incluidos los generados por el módulo de detección de anomalías. Por otro lado está la base de conocimiento compuesta de reglas de conocimiento borrosas, y por último un motor de inferencias combina la información de la base de hechos y de la base de conocimiento para disparar las reglas con los hechos observados, y generar las conclusiones pertinentes. Esta estructura queda representada en la siguiente figura:

Estructura del sistema experto de diagnóstico



7.5.2 Estrategias de diagnóstico

7.5.2.1 Estrategia de diagnóstico en continuo

Siempre que sea posible, es preferible adoptar una estrategia de diagnóstico en continuo, como ya ha sido señalado con anterioridad a lo largo de la tesis, aunque la estructura de sistema de diagnóstico que se propone es aplicable también a sistemas que funcionen de forma discreta, siempre que sean sistemas de razonamiento hacia adelante (véase el apartado 2.4.2). Ya sea continuo o discreto, el funcionamiento es tal que la llegada de nuevas medidas (manuales o automáticas) pone en funcionamiento el módulo de detección de anomalías que procede a su análisis. Cuando detecta una anomalía genera los hechos que la describen según el procedimiento indicado en los apartados 7.4.1 y 7.4.2 de este mismo capítulo, hechos que son introducidos en la base de hechos para su uso posterior por el sistema de identificación de anomalías, es decir por el sistema experto de diagnóstico, que se activa ante la detección de nuevas anomalías.

7.5.2.2 Conclusiones necesarias y conclusiones posibles

Para flexibilizar al máximo las capacidades de razonamiento del sistema, se propone aplicar los resultados obtenidos a lo largo de los capítulos anteriores en lo que concierne al razonamiento en el dominio de lo necesario y en el dominio de lo posible. De esta forma el

objetivo que se persigue es obtener primero conclusiones necesarias (basadas en el uso de modelos de regla de necesidad) más fiables pero en general más difíciles de obtener que las conclusiones posibles. Cuando el sistema sea incapaz de generar conclusiones necesarias satisfactorias, por su escasez o por inadecuación a los síntomas observados, el sistema optará por inferir conclusiones posibles (basadas en el uso de modelos de regla de posibilidad), menos fiables, pero en general más abundantes e informativas, debido a una menor exigencia en el cumplimiento de las condiciones de las reglas de posibilidad para activarse.

7.5.2.3 Relaciones semánticas entre consecuentes y antecedentes

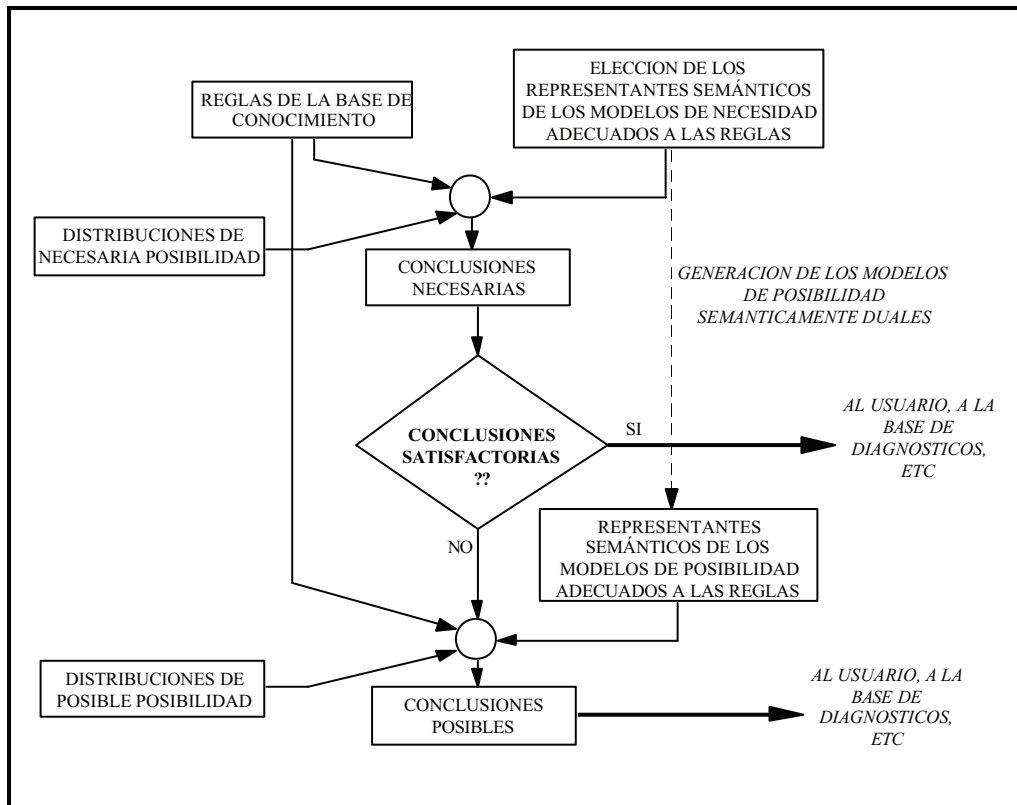
Por otro lado se pretende aprovechar al máximo las capacidades semánticas detectadas en los modelos analizados a lo largo del capítulo 4 de esta tesis, y materializadas en los modelos que al final de dicho capítulo se proponen. Para ello será necesario identificar la relación semántica que mantiene cada consecuente de una regla con sus antecedentes para así proceder a la elección del modelo adecuado.

Los modelos podrán ser escogidos de entre los representantes semánticos propuestos al final del capítulo 4, que ofrecen de forma simplificada los principales tipos de comportamientos semánticos detectados. Así se dispone de modelos graduales, de modelos de certidumbre y de modelos mixtos, tanto en el dominio de lo posible como en el dominio de lo necesario. Es necesario señalar que para una regla con varios antecedentes y un consecuente la relación semántica con el consecuente puede variar de antecedente a antecedente.

7.5.3 El motor de inferencias

7.5.3.1 Obtención de conclusiones necesarias y conclusiones posibles

La siguiente figura materializa la estrategia de diagnóstico escogida para la combinación en un entorno único de modelos de necesidad y posibilidad con distintas semánticas y distribuciones de necesaria y posible posibilidad. La integración de reglas de necesidad y posibilidad así como de distribuciones de necesaria y posible posibilidad fue descrita en el capítulo 5, y aquí solo se presenta un visión gráfica de conjunto de la estrategia propuesta.



Cada regla de la base de conocimiento lleva asociada una descripción semántica que especifica si la relación entre las variables que intervienen en la parte consecuente y la parte antecedente es una relación gradual, una relación de certidumbre o una relación mixta, para cada pareja antecedente-consecuente. Por defecto se considera que las reglas son reglas de necesidad, en el sentido en que si los antecedentes se cumplen, entonces las conclusiones que se obtienen son necesarias, o lo que es lo mismo se tiene la certidumbre de que son válidas. Las conclusiones de este tipo de reglas pueden interpretarse como distribuciones de necesaria posibilidad. La base de conocimiento debe formularse en la medida de lo posible utilizando reglas de necesidad, de forma que el conocimiento tenga la mayor certidumbre posible, y así las conclusiones que de ellas se deriven sean lo más fiables posible.

Sin embargo cuando el conocimiento no sea suficientemente completo puede ser necesario enunciar reglas de posibilidad. Para las reglas de posibilidad, además de especificar el tipo de relación entre su parte antecedente y su parte consecuente, al igual que para las de necesidad, debe especificarse que la relación descrita por la regla solo se considera válida en el dominio de lo posible, por lo que las conclusiones de estas reglas, cuando los antecedentes se cumplan, solo podrán considerarse posibles (es decir no contradictorias con los hechos) pero nunca necesarias. Sin embargo en la mayoría de los casos las reglas

de posibilidad serán obtenidas por el sistema de forma automática a partir de las reglas de necesidad.

A partir de los tipos de reglas anteriores, la estrategia de diagnóstico propuesta es la siguiente. En primer lugar el principal objetivo del sistema es obtener conclusiones fiables, es decir válidas en el dominio de lo necesario. Para ello a cada regla de necesidad se le asigna un modelo de necesidad borroso, cuya semántica debe ser la especificada en la regla. Para simplificar el proceso se utilizará alguno de los modelos representantes semánticos propuestos al final del capítulo 4, cuyo mecanismo de inferencia asociado es la MS de dos índices.

Combinando los modelos de necesidad con las distribuciones de necesaria posibilidad o hechos necesarios, el motor de inferencias obtiene todas las conclusiones necesarias que se puedan derivar de las reglas y hechos del sistema mediante la aplicación de la MS de dos índices. En general la mayor parte de las observaciones realizadas vendrán modeladas mediante distribuciones de posibilidad, y por tanto también de necesaria posibilidad.

Si las conclusiones son satisfactorias, entonces la información generada sigue su curso habitual (presentación al usuario, incorporación en la base de diagnósticos, etc.) y se termina el proceso de inferencia hasta que no se disponga de nuevas anomalías que activen de nuevo al sistema de diagnóstico.

Las conclusiones pueden no ser suficientemente satisfactorias. Por un lado, dado que las condiciones de activación de los modelos de necesidad son restrictivas, es posible que ante una nueva anomalía ninguno de los modelos de necesidad se active y el sistema no sea capaz de emitir ninguna conclusión. Por otro lado, aunque puedan deducirse conclusiones necesarias, es posible que ninguna de ellas consiga explicar satisfactoriamente la anomalía detectada, por lo que sería deseable generar nuevas conclusiones para tener un mayor rango de posibles explicaciones. En ambos casos puede ser interesante generar todas aquellas conclusiones que aunque no pueda asegurarse que son válidas, sí pueda afirmarse que no son contradictorias (o del todo contradictorias) con los hechos observados.

En este caso a cada regla de necesidad se le asocia un modelo de posibilidad acorde con la semántica de la regla, semánticamente dual del modelo de necesidad empleado en el paso anterior, que escogeremos de entre los representantes semánticos propuestos en el capítulo 4. Además todas las reglas de posibilidad de la base de conocimiento podrán también ser incorporadas en este nuevo paso del proceso deductivo con su correspondiente modelo. La combinación de los modelos de posibilidad de las reglas con las distribuciones de posible posibilidad del sistema permiten deducir todas aquellas conclusiones posibles en mayor o menor grado. La información así generada debe entonces seguir su curso habitual (presentación al usuario, incorporación en la base de

diagnósticos, etc.) y el proceso de inferencia se da por finalizado hasta la llegada de nuevas observaciones al sistema. En general la mayor parte de las observaciones realizadas vendrán modeladas mediante distribuciones de posibilidad, y por tanto también de posible posibilidad. Hay hechos que sin embargo vienen modelados directamente mediante distribuciones de posible posibilidad, como son por ejemplo, bajo ciertas combinaciones de sus entradas, las salidas del sistema de detección de anomalías, o también las conclusiones obtenidas a partir de otras reglas con modelos de posibilidad. Estos hechos, solamente válidos en el dominio de lo posible, sólo pueden ser utilizados durante este segundo paso del proceso de inferencia, donde lo que se deducen son conclusiones posibles y no necesarias.

7.5.3.2 Orden de disparo de las reglas del sistema experto

Según la estrategia descrita en el punto anterior, primero deben dispararse las reglas de necesidad y luego, en función de las conclusiones obtenidas en este primer paso de inferencia, las reglas de posibilidad.

Por otro lado, tal y como se vio en el apartado 5.8.2, las reglas no pueden dispararse en un orden arbitrario, puesto que se corre el riesgo de activar los modelos con observaciones incompletas, que pueden conducir a resultados erróneos.

Por tanto el primer paso, durante el arranque del sistema, debe consistir en la ordenación de todas las reglas según los criterios propuestos en dicho apartado, sin atender a sus características semánticas o tipo.

En ausencia de metarreglas el sistema debe disparar las reglas según el orden anterior para garantizar que todas ellas utilizan información completa sobre cada una de las variables a las que hacen referencia. En el primer paso de inferencia deben activarse solamente las reglas de necesidad, y en función de los resultados obtenidos, las reglas de posibilidad.

Si las reglas se organizan por ejemplo por grupos temáticos, pueden utilizarse metarreglas para decidir cuándo un determinado grupo de reglas debe ser comprobado o no, y así eliminar la necesidad de comprobar sistemáticamente todas las reglas de la base de conocimiento. Las metarreglas también deben estar ordenadas en el conjunto global de reglas. En este caso proponemos seguir el siguiente orden de activación:

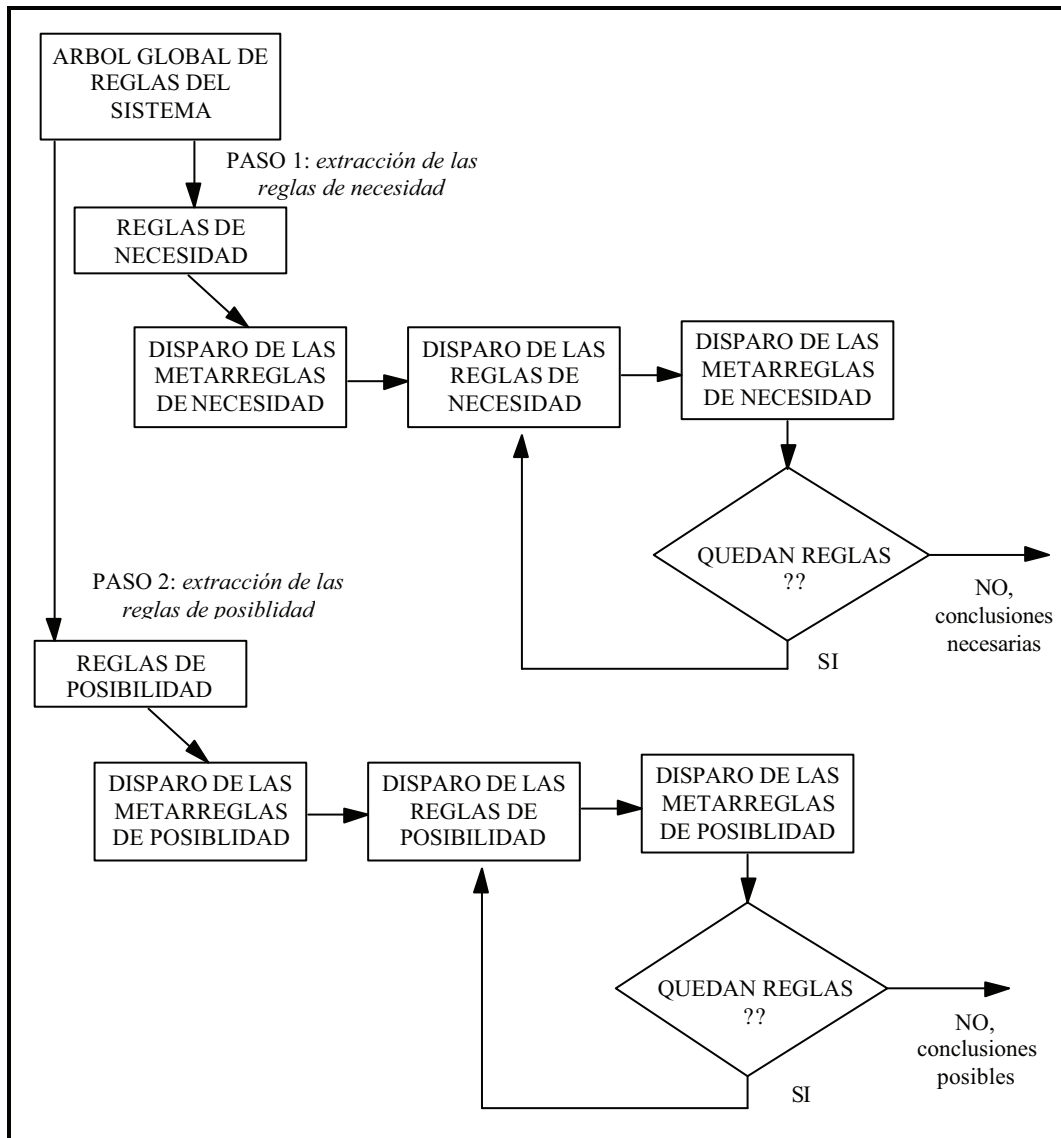
- 1) En primer lugar deben activarse todas las metarreglas del sistema para determinar si con los hechos disponibles se activa algún grupo de reglas.

- 2) En segundo lugar deben dispararse todas aquellas reglas que no aparecen en ninguna metarregla de activación, y todas aquellas reglas que pertenezcan a algún grupo de los activados en el paso anterior.
- 3) En tercer lugar se procederá a disparar nuevamente todas las metarreglas del sistema, por si existen nuevos hechos obtenidos en el paso que las activen. El disparo de nuevas metarreglas puede activar nuevos grupos de reglas.
- 4) En este punto se repite el proceso desde el segundo punto, y la inferencia termina cuando no existen reglas por disparar. Hay que tener especial cuidado en verificar que no aparecen procesos cíclicos que conduzcan a bucles infinitos.

Una buena forma de analizar las inconsistencias de las reglas que puedan dar lugar a procesos cíclicos consiste en dibujar el árbol global de reglas y utilizarlo para seguir de forma gráfica el proceso de inferencia llevado a cabo, durante la fase de puesta a punto. Tanto los procesos cíclicos como el transcurrir del razonamiento se pueden seguir fácilmente en un gráfico de este tipo, por lo que es muy recomendable basar la traza del razonamiento, consistente en la determinación de la activación y disparo de las reglas y el orden en que se hace, en un análisis gráfico sobre el árbol de reglas del sistema.

Nótese que cuando se utilizan metarreglas para activar unos u otros grupos de reglas, el orden del árbol de reglas sigue siendo válido pero particularizado para el conjunto de reglas susceptibles de ser disparadas. Recuérdese también que el proceso de razonamiento se lleva a cabo en dos pasos, en primer lugar la obtención de conclusiones necesarias y en segundo lugar la obtención de conclusiones posibles.

La siguiente figura ilustra gráficamente el orden de disparo de reglas propuesto:



7.5.4 Estructura, formato y disparo de las reglas del sistema experto

En este apartado se propone una estructura general para las reglas de conocimiento para poder codificar de forma fiel y flexible el conocimiento experto, y se determina cómo deben obtenerse cada una de sus conclusiones.

7.5.4.1 Estructura general de una regla de conocimiento

La estructura general que se propone para una regla de conocimiento está basada en la que se ha utilizado en el proyecto TRAFES desarrollado para UEFSA por el equipo investigador al que pertenezco, en colaboración con otros socios, y que se describe brevemente en [Sanz *et al*, 93]. A esta estructura se le ha añadido la posibilidad de incorporar modelos para expresar las distintas relaciones semánticas entre antecedentes y consecuentes. El formato propuesto es el siguiente:

```

REGLA      código
NOMBRE     nombre
GRUPO     grupo1, grupo2, ...
si
           antecedente1, antecedente2, ...
entonces
           consecuente1, modc1
           consecuente2, modc2
...
diagnósticos
           diagnóstico1, modd1
           diagnóstico2, modd2...
acciones
           acción1, modal
           acción2, modal ...
COMENTARIOS: .....

```

El código de la regla es un campo que permite identificar la regla dentro del sistema, mientras que el nombre es una pequeña frase explicativa del contenido de la regla. Los grupos son los grupos temáticos a los que se supone pertenece la regla, lo que permite activar la regla desde las metarreglas del sistema. A continuación se enuncian los antecedentes, los consecuentes, los diagnósticos y las acciones a tomar. Hemos distinguido entre consecuentes, diagnósticos y acciones ya que, aunque formalmente todos ellos son consecuentes, interesa distinguirlos por la diferente estructura que en general tendrán para nosotros, lo que permite considerarlos como categorías propias.

Para cada consecuente, diagnóstico y acción se debe (o se puede) especificar un modelo (simbolizado con *mod*) que haga explícito el tipo de relación semántica que mantiene con los antecedentes, ya sea ésta gradual, de certidumbre o mixta, pudiendo además darse en el dominio de lo necesario o de lo posible. Además la relación semántica de un

consecuente y los antecedentes de la regla puede variar de uno o otro, lo que da lugar a tantas relaciones semánticas como antecedentes tenga la regla. Por otro lado, una regla como la enunciada da internamente lugar a tantas reglas como consecuentes, diagnósticos y acciones tenga, pudiendo cada una de ellas tener asociado modelos diferentes.

Para poder controlar la activación de grupos de reglas que puedan por ejemplo relacionarse con determinados síntomas, de forma que solo se proceda a su comparación cuando estos sean observados, proponemos una sencilla estructura de metarregla, cuyo formato es:

```

METARREGLA   código
NOMBRE       nombre
si
    antecedente1, antecedente2, ...
entonces
    activar grupo1, grupo2, ...

```

La principal diferencia con las reglas reside en que para estas reglas su parte consecuente consiste en la acción de poner en el grupo de reglas que el motor de inferencias debe comprobar las reglas pertenecientes a los grupos *grupo1*, *grupo2*, Para estas reglas bastará que la comprobación de la parte antecedente tenga una certidumbre no nula (o superior a un pequeño umbral) para que la regla se dispare. En este caso el cumplimiento parcial o total de la parte antecedente no influye en el comportamiento de la regla. Supondremos además que las metarreglas son todas reglas de necesidad y que por tanto son aplicables tanto cuando se infieran conclusiones necesarias como cuando se infieran conclusiones posibles, aunque se podría también ofrecer la flexibilidad de incorporar metarreglas de posibilidad. Sin embargo el modo de comprobación de los antecedentes dependerá del paso de inferencia en el que nos encontremos, ya sea en el dominio de lo necesario o en el de lo posible.

En general tanto los antecedentes como los consecuentes de las reglas pueden ser de diferentes tipos, que serán descritos a lo largo de esta sección.

7.5.4.2 Regla de inferencia, modelos de reglas y relaciones semánticas

A cada consecuente o diagnóstico de una regla hemos asociado un modelo que especifica qué tipo de relación mantiene con la parte antecedente de la regla, ya sea gradual, de certidumbre o mixta. Como hemos señalado, cada consecuente o diagnóstico da lugar internamente a una regla diferente con el modelo especificado. Además el tipo de influencia de cada uno de los antecedentes en un determinado consecuente puede variar de un

antecedente a otro por lo que será necesario especificar, para cada consecuente, tantas relaciones semánticas como antecedentes tenga la regla.

La regla de inferencia escogida es la MS basada en dos índices y se describe en detalle en el apartado 4.11. Como ya se dijo esta regla ofrece capacidades semánticas comparables a las de la RCI pero es más rápida de calcular y tiene una mayor transparencia de cara al análisis e interpretación de su comportamiento. De hecho surge del estudio de la RCI y de su posterior simplificación.

La MS de dos índices se basa en el uso de las implicaciones de necesidad y posibilidad generalizadas. La primera de las implicaciones permite inferir en el dominio de lo necesario dando como resultado distribuciones de necesaria posibilidad y recordemos viene dada por:

$$I_n(\alpha_n, \beta_n, y) = \min\left(\frac{\beta_n}{\alpha_n}(y - \alpha_n) + 1, 1\right)$$

donde α_n y β_n son medidas de similitud entre los antecedentes y las observaciones, basadas en alguna medida de inclusión. Las segunda de las implicaciones permite inferir en el dominio de lo posible dando como resultado distribuciones de posible posibilidad y viene dada por:

$$I_p(\alpha_p, \beta_p, y) = \max\left(\frac{\beta_p}{\alpha_p}(y - 1) + \beta_p, 0\right)$$

donde α_p y β_p son medidas de similitud entre los antecedentes y las observaciones, basadas en alguna medida de intersección.

Los modelos a emplear pueden escogerse entre los representantes semánticos propuestos al final del capítulo 4. Para convertir un modelo de necesidad en un modelo de posibilidad, cuando el sistema pasa de la fase de razonamiento en el dominio de lo necesario a la fase de razonamiento en el dominio de lo posible, debe utilizarse un modelo de posibilidad semánticamente dual del modelo de necesidad empleado durante la primera fase.

Los modelos para la MS de dos índices determinan los índices de similitud que deben ser empleados para la obtención de α y β en el cálculo de la conclusión. Los modelos que utilizaremos para razonar en el dominio de lo necesario son los que aparecen en la siguiente tabla:

α_n	$\beta_n = N(P/P')$	Semántica modelo
$N(P/[P'])$	$N(JP/[P'])$	<i>gradual</i>
$N(P/[P'])$	$\min(N(JP/[P']), N(P/[P']))$	<i>mixta</i>
1	$N(P/[P'])$ o 0 si $N(JP/[P']) < 1$	<i>certidumbre estricto</i>
1	$\min(N(JP/[P']), N(P/[P']))$	<i>certidumbre</i>

En el dominio de lo posible utilizaremos como modelos semánticamente duales los siguientes (véanse los apartados 4.11.3.2 y 5.7):

α_p	$\beta_p = \Pi(P/P')$	Semántica modelo
$\Pi(JP'/[P])$	$\Pi(JP/[P'])$	<i>gradual</i>
$\Pi(JP'/[P])$	$\min(\Pi(JP/[P']), \Pi(JP'/[P]))$	<i>mixta</i>
1	$\Pi(JP'/[P])$ o 0 si $P(u)+P'(u) < 1 \forall u$	<i>certidumbre estricto</i>
1	$\min(\Pi(JP/[P']), \Pi(JP'/[P]))$	<i>certidumbre</i>

Para especificar el modelo de un consecuente, diagnóstico o acción se propone el siguiente formato:

$$(tipo, semántica, grado_verdad, certidumbre) = (tipo, sem1, sem2, \dots, \alpha, \beta)$$

El tipo *tipo* que puede ser *N* o *Π* permite especificar si la relación semántica entre la parte antecedente y el consecuente o diagnóstico en cuestión es válida en el dominio de lo posible o en el de lo necesario.

La relación semántica deberá especificarse para cada uno de los antecedentes de la regla, mediante los parámetros *sem1*, *sem2*, De este modo el parámetro *sem1* especifica la relación semántica entre el consecuente considerado y el primer antecedente, *sem2* su relación con el segundo antecedente y así sucesivamente. Los valores considerados para los parámetros *sem* son *G* para relaciones graduales, *C* para relaciones de certidumbre y *M* para relaciones mixtas. Puede considerarse también la posibilidad de distinguir la semántica de certidumbre estricta, *CE*, cuando la condición de activación es la más restrictiva.

Los parámetros α y β son respectivamente el grado de verdad y la certidumbre de la conclusión cuando la parte antecedente de la regla se cumple. Como veremos más adelante para los consecuentes la expresión anterior suele reducirse a $(tipo, sem)$ ya que normalmente el grado de verdad y la certidumbre viene implícita en las distribuciones de posibilidad involucradas, o es posible obtenerla utilizando el valor de verdad borroso como modificador. Así por ejemplo el consecuente:

$$temperatura \text{ en motor es alta } (tipo, sem, \alpha, \beta)$$

es equivalente al consecuente:

$$\text{temperatura en motor es } I_{\text{tipo}}(\alpha, \beta, (\text{alta})) \text{ (tipo, sem)}$$

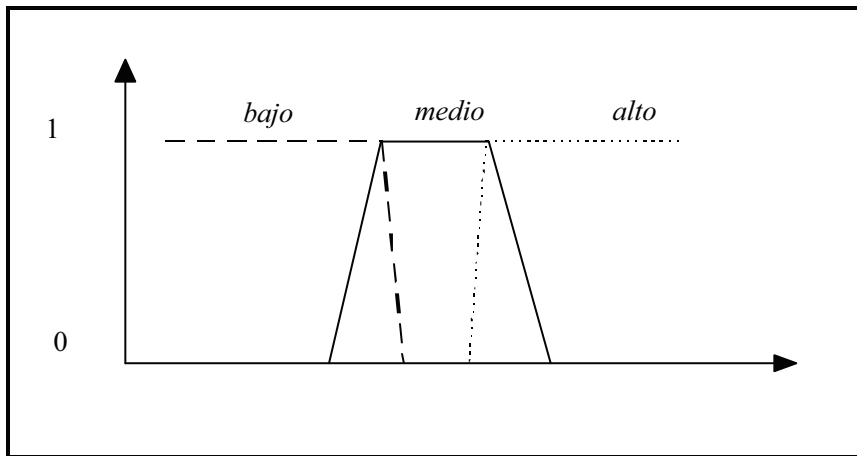
siempre que *alta* sea una distribución de posibilidad sobre un universo numérico.

- Determinación de los modelos de una regla

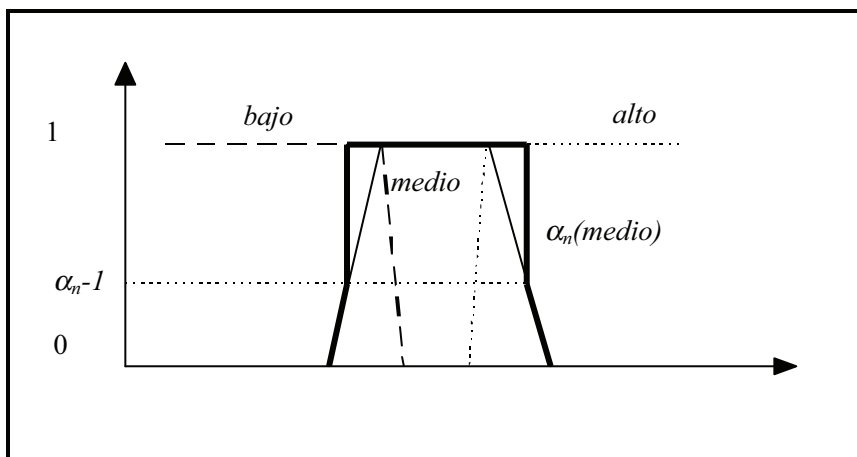
La semántica de las reglas es en general compleja y debe ser cuidadosamente analizada para conseguir modelarla correctamente. En cierto modo puede considerarse similar a la identificación de sistemas (en un ámbito principalmente cualitativo) donde no existen procedimientos definitivos y suficientemente sistemáticos que permitan obtener los modelos buscados, aunque sí es posible encontrar recomendaciones y metodologías racionales que pueden servir de guía, facilitando dicho proceso.

En este sentido el método que se propone consiste en determinar cómo varía el grado de verdad o el grado de certidumbre de cada uno de los diagnósticos de la regla cuando varía el grado de verdad de cada uno de sus antecedentes. Por simplicidad supondremos que trabajamos en universos ordenados.

Un decrecimiento del grado de verdad del antecedente se traduce en un aumento del núcleo del conjunto borroso que modela a su correspondiente estado. Si observamos la partición que aparece en la siguiente figura podemos comprobar que cuando el grado de verdad del estado *medio* disminuye, puede considerarse que el valor de la medida se hace cada vez más compatible con el concepto *bajo* o con el concepto *alto*, aunque sin dejar de ser compatible con el concepto *medio*.



Considerando por tanto el estado como un hipotético valor cualitativo de la medida, la reducción de su grado de verdad puede ser vista como la existencia de un posible incremento o decremento de dicho valor cualitativo, como se aprecia en la siguiente figura:



Una forma de amplificar este efecto consiste en sustituir, en el antecedente analizado, el estado original por un valor cualitativo mayor o menor, pero en cualquier caso adyacente. Si por ejemplo el antecedente analizado fuese:

u es medio

lo reemplazaríamos alternativamente por los antecedentes *u es bajo* y *u es alto* dando lugar a dos nuevas reglas cuyo nuevo consecuente debe ser determinado utilizando el conocimiento experto disponible. En el caso en se pueda considerar que de las reglas así obtenidas solo varía el grado de verdad del consecuente, o que su estado ha tomado un valor cualitativo adyacente (próximo), entonces la relación podrá considerarse una relación gradual. En el caso en que su estado tome otro valor significativamente diferente, o que sea

necesario formular otro consecuente diferente, entonces la relación podrá considerarse de certidumbre. En el caso en que se den ambos fenómenos la relación podrá considerarse mixta. Esto puede ocurrir por ejemplo cuando en las nuevas reglas se mantenga un consecuente similar al inicial pero sea necesario formular otro adicional.

En el caso de los antecedentes que involucran a conjuntos situados en los extremos de la partición, la variación del grado de verdad significa necesariamente que el valor cualitativo del estado puede haber aumentado o disminuido, pero no ambos simultáneamente. Por tanto sólo se obtendrá una nueva regla.

Hay que tener en cuenta que las variaciones de estado así provocadas son una amplificación de la variación del grado de verdad del antecedente original, y pueden no ser todo lo sutiles que debieran para una determinación correcta del modelo. Por tanto las conclusiones que así se obtengan no deben considerarse como definitivas sino solo como indicativas.

Cuando se disponga de un sistema de reglas suficientemente completo se puede variar el valor del estado de un determinado antecedente al valor inmediatamente inferior o superior, y buscar, dentro de la base de conocimiento, la o las regla del sistema de reglas que más se asemejen a las nuevas reglas obtenidas. Comparando los consecuentes de las reglas encontradas con los de la regla original se podrá establecer el tipo de relación semántica existente. Este es el procedimiento que se ha seguido en el ejemplo del apartado 7.6.

- Relaciones graduales:

Las relaciones graduales están normalmente indicadas cuando las variables que intervienen en el antecedente y el consecuente están semánticamente relacionadas mediante una *razonable* continuidad cualitativa. Por ejemplo:

si
revoluciones del motor son elevadas
entonces
temperatura en motor es alta

A medida que el número de revoluciones del motor se hace más compatible con el concepto *medio*, también la temperatura se hace más compatible con el concepto *media*. Dicho de otro modo, si el número de revoluciones tiende a disminuir también puede considerarse que la temperatura tiende a disminuir y viceversa (siempre que no tengamos en consideración otros fenómenos adicionales).

Formalmente la regla anterior puede formularse como:

REGLA *REVOLUCIONES*
NOMBRE *relación entre revoluciones y temperatura*
GRUPO *general*

Si
 revoluciones en motor elevadas
entonces
diagnósticos

temperatura en motor alta (N,G,1,1)

COMENTARIOS:

El descriptor (N,G,1,1) se interpreta como la existencia de una relación de necesidad gradual. Además cuando las hipótesis se verifican por completo se considera que el diagnóstico es verdadero en grado 1 y con una certidumbre también de valor 1.

Cuando el conocimiento que se tiene para una relación de este tipo no verifica las condiciones de certeza anteriores se puede recurrir a un modelo de posibilidad gradual. Si por ejemplo se pretende expresar:

si
 revoluciones del motor son elevadas
entonces
 temperatura en motor es posiblemente alta

debe usarse un modelo de posibilidad gradual. La regla expresa que los hechos *revoluciones del motor son elevadas* y *temperatura en motor es posiblemente alta* no son contradictorios, y que por tanto pueden darse simultáneamente. En cierto modo expresa también la existencia de una cierta evidencia de que tiendan a darse simultáneamente, aunque no se considera en ningún grado necesario que deban hacerlo. Formalmente la regla se puede enunciar como:

REGLA **REVOLUCIONES**
NOMBRE *relación entre revoluciones y temperatura*
GRUPO *general*
Si
 revoluciones en motor elevadas
entonces
diagnósticos
 temperatura en motor alta (II,G,1,1)

COMENTARIOS:

El descriptor (II,G,1,1) se interpreta como la existencia de una relación de posibilidad gradual. Además cuando las hipótesis se verifican por completo se considera que el diagnóstico es posiblemente verdadero en grado 1 y con una posibilidad (certidumbre en el dominio de lo posible) también de valor 1.

La relación gradual puede también aparecer entre los antecedentes y determinado tipo de acciones. Así por ejemplo, podría fácilmente concebirse una regla del tipo:

si
 freno de mano del coche frena poco
entonces
 tensar freno de mano

En este caso puede verse fácilmente que (supuesto que tensar sea suficiente para mejorar el funcionamiento del freno de mano) cuanto menos frene el freno de mano más deberá tensarse, lo que pone de manifiesto una clara semántica gradual.

Las reglas graduales son especialmente útiles cuando forman sistemas completos coherentes, en el sentido en que como se vio en el capítulo 5, en el apartado 5.7.1, tienden a generar distribuciones unimodales que se desplazan de forma suave en el universo sobre el que concluyen, y pueden ser utilizadas como interpoladores ([Dubois, Prade, 96]).

- Relaciones de certidumbre

Las relaciones de certidumbre se traducen en una reducción de la certidumbre de los consecuentes cuando se produce una reducción del grado de verdad de los antecedentes, y expresan la imposibilidad de graduar el grado de verdad del consecuente en función del grado de verdad del antecedente.

Un ejemplo sencillo de este tipo de relaciones se da cuando la variable consecuente es una variable que solo admite valores discretos sin que pueda establecerse una relación de

orden entre los mismos, y que al no admitir situaciones intermedias no admite grados de verdad parciales. Este es el caso de la siguiente regla:

si
 temperatura en motor es muy_ alta
entonces
 motor gripado

En efecto supuesto que el estado del motor no admite situaciones intermedias, y que o bien funciona correctamente o bien está gripado y deja de funcionar, la relación semántica entre el antecedente y el consecuente de la regla anterior es una relación de certidumbre. En el caso en que el grado de verdad del antecedente decrezca, que se traduce en este caso en que la temperatura del motor tiende a ser menos alta, la certidumbre del consecuente se reduce. Estas relaciones no son sólo aplicables a conclusiones que no admitan grados de verdad parciales, y simplemente expresan la imposibilidad de relacionar el grado de verdad del consecuente con el del antecedente. Como se verá en el ejemplo del apartado 7.6 la existencia de una relación de certidumbre entre un antecedente y un consecuente no se debe al tipo de valor de este último, sino a cómo éste varía ante pequeñas variaciones del antecedente. Si variaciones leves de este último producen variaciones bruscas del consecuente (y en cierto modo impredecibles) la relación semántica será una relación de certidumbre.

Formalmente la regla anterior puede expresarse como:

REGLA *MOTOR_GRIPADO*
NOMBRE *motor gripado por elevación de temperatura*
GRUPO *general*

Si
 temperatura en motor muy_ alta
entonces
diagnósticos
 motor gripado (N,C,1,1)

COMENTARIOS:
El descriptor (N,C,1,1) se interpreta como la existencia de una relación de necesidad de certidumbre. Además cuando las hipótesis se verifican por completo el diagnóstico es verdadero en grado 1 y con una certidumbre también de valor 1.

En el caso de esta regla no parece necesario utilizar un modelo de certidumbre estricto. Los modelos de certidumbre estrictos, es decir con condición de activación restrictiva, pueden ser especialmente interesantes cuando en su parte consecuente puedan existir

acciones que impliquen cierto riesgo, y por tanto no se sea admisible que se disparen cuando la hipótesis tenga una cierta componente de incertidumbre, solo cuando tenga un grado de verdad parcial no nulo. Este es por ejemplo el caso de la siguiente regla, que podría recomendar una de las acciones a llevar a cabo durante una revisión rutinaria de un vehículo:

si
 temperatura del motor es baja
entonces
 abrir el radiador y comprobar nivel de agua

En este caso la acción de abrir el radiador no admite situaciones intermedias, dando lugar a una relación de certidumbre. Por otro lado, el riesgo que esta acción entraña exige que se tenga la certidumbre de que la temperatura es baja, en el grado de verdad que sea (por supuesto no nulo), lo que sugiere el uso de la condición de activación más restrictiva. Formalmente la regla puede expresarse como:

REGLA **COMPROBAR_AGUA**
NOMBRE *comprobar agua del circuito de refrigeración*
GRUPO *revisión_del_vehículo*
Si
 temperatura en motor baja
entonces
diagnósticos
 comprobar agua al circuito de refrigeración (N,CE,1,1)

COMENTARIOS:

El descriptor (N,CE,1,1) se interpreta como el último descriptor visto pero la clave CE indica que el modelo de certidumbre debe ser un modelo estricto. En el caso de que exista la más mínima certidumbre de que la temperatura pueda no ser baja la regla no se activa por el peligro que la acción entraña.

- Relaciones mixtas

Estas relaciones presentan una semántica que combina características de los dos tipos anteriores. Son especialmente indicadas por ejemplo cuando un determinado síntoma puede ser debido a varias causas o diagnósticos, y simultáneamente existe una relación gradual entre alguna de las causas y el citado síntoma.

Así por ejemplo en la regla:

si
temperatura del motor es alta
entonces
falta agua en el circuito de refrigeración
correa ventilador rota

podemos ver que entre el antecedente y el primer diagnóstico existe una relación gradual, pero existe también una relación de incertidumbre debido a que la causa del síntoma podría ser el segundo de los diagnósticos. En efecto cuanto menor sea el grado de verdad del antecedente, es decir cuanto más tienda la temperatura del motor a ser media, menor tenderá a ser la falta de agua del motor (semántica gradual), pero sólo bajo el supuesto de que ésta sea la verdadera causa del incremento de temperatura. Para modelar el desconocimiento que se tiene sobre ésta verdadera causa se introduce un nivel de incertidumbre en la conclusión, variable también con el grado de verdad del antecedente. Para el segundo diagnóstico sólo existe una relación de certidumbre, ya que el estado de la correa no admite situaciones intermedias, dando lugar a una semántica de certidumbre. Formalmente esta regla puede expresarse como:

REGLA *TEMPERATURA_MOTOR_ALTA*
NOMBRE *causas para explicar la elevación de temperatura del motor*
GRUPO *general*
Si
temperatura en motor es alta
entonces
falta agua en el circuito de refrigeración (N,M,1,0.7)
correa ventilador rota (N,C,1,0.4)

COMENTARIOS:

El descriptor (N,M,1,1) se interpreta como el último descriptor visto pero la clave M indica que el modelo es un modelo mixto. Además se está asumiendo la creencia en que la falta de agua es una causa que en general explica mejor el síntoma detectado, que la rotura de la correa del ventilador. Esta regla ilustra las ventajas de combinar distintos modelos según los consecuentes, diagnósticos o acciones, dentro de una misma regla.

Las relaciones de necesidad mixtas se caracterizan por generar las distribuciones menos específicas, ya que simultáneamente aumentan el núcleo del consecuente y le añaden un nivel de indeterminación. Esto puede justificar el uso de este tipo de relaciones cuando se desconozca con precisión la relación semántica entre la parte antecedente y la parte consecuente de la regla, por lo que puede considerarse como una opción por defecto: producen un decremento tanto del grado de verdad como del grado de certidumbre.

- Reglas heurísticas de modificación de certidumbre

Existen a menudo situaciones en las que el experto necesita expresar un conocimiento del tipo:

si
 se cumple que u es P
entonces
 me creo más que v sea Q

pero sin que sea capaz de establecer una relación directa entre el síntoma P y el diagnóstico Q . Por tanto esta regla debe interpretarse como *supuesto que Q es un diagnóstico con certidumbre no nula, la existencia de P incrementa la certidumbre en Q .*

Siguiendo con el ejemplo de la regla anterior podría plantearse la siguiente regla:

si
 tapón depósito agua deteriorado
entonces
 incrementar certidumbre en falta agua en el circuito de refrigeración

La semántica de este tipo de reglas no puede expresarse con los modelos vistos hasta ahora por lo que necesitan un tratamiento especial. Supongamos que tras haber activado y disparado todas las reglas del primer paso de inferencia (reglas de necesidad) se concluye que:

falta agua en el circuito refrigeración (N,0.8,0.5)

con lo que en el caso de que fuese cierto, la falta de agua sería notable (al tener un grado de verdad de 0.8), pero la certidumbre de este diagnóstico no sería muy elevada, en concreto de 0.5. Supongamos que la regla de modificación de certidumbre se formula del siguiente modo:

si
 tapón depósito agua deteriorado
entonces
 incrementar certidumbre en falta agua en
 el circuito de refrigeración 0.7

Para modificar la certidumbre de 0.5 anteriormente obtenida proponemos utilizar la siguiente expresión:

$$0.5^{1-0.7} \approx 0.8$$

con lo que efectivamente se produce un cierto incremento en la certidumbre final del diagnóstico. Nótese que incrementos consecutivos de certeza son conmutativos, por lo que el orden de disparo de estas reglas es indiferente:

$$((\beta)^a)^b = ((\beta)^b)^a$$

Además la certidumbre solamente puede incrementarse cuando es distinta de 0. En efecto cuando no se tiene ninguna certidumbre de que el grado de verdad de un diagnóstico pueda ser mayor que 0, no tiene sentido incrementarla, por lo que el comportamiento obtenido es intuitivamente razonable: si $\beta=0$ entonces $\beta^a=0^a=0$.

Del mismo modo cuando la certidumbre es 1 tampoco se produce un incremento de la misma puesto que ya está en su valor máximo ($1^a=1$).

Las reglas de modificación de certidumbre pueden ser utilizadas para matizar las conclusiones de otras reglas mediante la modificación relativa de sus certidumbres. Pueden también ser utilizadas para modelar excepciones al comportamiento general, de forma que cuando se producen determinadas observaciones, se modifiquen las certidumbres que hasta entonces hayan sido concluidas.

Estas reglas pueden formularse tanto en el dominio de lo necesario como en el de lo posible, y su comportamiento es en ambos casos idéntico, salvo que el argumento al que se aplican es en un caso una certidumbre en el dominio de lo necesario o grado de necesidad, y en el segundo una certidumbre en el dominio de lo posible o grado de posibilidad. Estas reglas están sujetas al mismo orden de disparo que el resto con alguna salvedad. Deben intervenir igual que otra regla cualquiera en el árbol de reglas. Sin embargo para una variable determinada, estas reglas son las últimas que deben ser activadas, es decir después de las reglas graduales, de certidumbre o mixtas, ya que están destinadas a incrementar o decrementar una certidumbre final obtenida mediante los otros tipos de reglas.

En el caso de que se pretenda decrementar la certidumbre se puede plantear el modelo de regla alternativo, descrito mediante el descriptor $(N,D,0.7)$. Utilizando el mismo ejemplo anterior, la modificación de la certidumbre se puede llevar a cabo mediante la siguiente expresión:

$$0.5^{1/(1-0.7)} \approx 0.8$$

de esta forma si la cuantía del decremento es 1 el valor final de la certidumbre es 0 , mientras que si la cuantía del decremento es 0 entonces el valor final de certidumbre permanece inalterado.

Formalmente este tipo de reglas pueden enunciarse como sigue:

REGLA *DETERIORO_TAPON_AGUA*
NOMBRE *incremento de certidumbre en falta de agua por deterioro del tapón*
GRUPO *general*
Si

tapón en depósito_agua deteriorado
 entonces
 diagnóstico
 falta_agua en el circuito_refrigeración (N,I,0.7)

COMENTARIOS:

El descriptor (N,I,0.7) se interpreta como que el incremento de certidumbre sobre el diagnóstico especificado debe hacerse en el ámbito de lo necesario (y por tanto también en el de lo posible) y en grado 0.7 cuando la hipótesis se cumpla.

Cuando el valor de necesidad borroso del diagnóstico antes del disparo de esta regla viene dado por $(N, \alpha d, \beta d)$, el valor de necesidad borroso de la hipótesis es $(N, \alpha h, \beta h)$ y el incremento propuesto por la regla viene dado por $(N, I, \beta m)$ es necesario tener en cuenta la certidumbre de la hipótesis para llevar a cabo el incremento (o decremento de certidumbre del consecuente). Para ello proponemos en primer lugar recalcular el incremento que debe ser llevado a cabo, teniendo en cuenta la certidumbre de los antecedentes:

$$\beta = \beta m^{1-\beta h}$$

Así cuando la certidumbre de la hipótesis es nula no se produce ningún incremento de la certidumbre del diagnóstico, mientras que cuando es 1 se produce el incremento que indica la regla.

A continuación se procede a incrementar la certidumbre como ya se ha indicado, de forma que la certidumbre final del diagnóstico viene dada por:

$$\beta d_{final} = (\beta d)^{1-\beta}$$

Al igual que con los modelos ordinarios podría establecerse toda una casuística de si β debe obtenerse a partir de αh o bien a partir de βh , pero en la que por simplicidad no entraremos.

7.5.4.3 Determinación de las conclusiones de una regla

Dada una regla:

$$P_1, P_2, \dots \rightarrow Q_1, Q_2, \dots$$

de varios antecedentes y consecuentes, y teniendo en cuenta que para cada consecuente puede existir un modelo de regla diferente que determine sus relaciones semánticas con los antecedentes, el motor de inferencias genera tantas reglas como consecuentes, obteniéndose por tanto:

$$P_1, P_2, \dots \rightarrow Q_1$$

$$P_1, P_2, \dots \rightarrow Q_2$$

...

...

A continuación analizaremos varias formas de obtener cada una de las conclusiones Q_1' , Q_2' , ..., a partir de la comprobación de los antecedentes de cada una de las reglas.

- Conclusión de una regla con una única relación semántica

Supondremos en este apartado que partimos de una regla del tipo:

$$P_1, P_2, \dots \rightarrow Q$$

con un modelo tal que la relación semántica entre Q y cada uno de los antecedentes P_i es la misma. En este caso la obtención de la conclusión puede hacerse del siguiente modo:

- se obtienen las distribuciones conjuntas que modelan a la parte antecedente de la regla y a la observación. Este tipo de distribuciones se calculan mediante una agregación conjunta con lo que se obtienen las distribuciones:

$$P(u_1, u_2, \dots) = T(P_1(u_1), P_2(u_2), \dots)$$

$$P'(u_1, u_2, \dots) = T(P_1'(u_1), P_2'(u_2), \dots)$$

- se determinan los cuatro índices globales de similitud entre antecedentes y observaciones necesarios para la inferencia (que recordemos dependen del modelo de

regla escogido). Los dos primeros índices son válidos para inferir en el dominio de lo necesario, los dos segundos en el dominio de lo posible:

$$\alpha_n = S_{n1}(T(P_1'(u_1), P_2'(u_2), \dots), T(P_1(u_1), P_2(u_2), \dots))$$

$$\beta_n = S_{n2}(T(P_1'(u_1), P_2'(u_2), \dots), T(P_1(u_1), P_2(u_2), \dots))$$

$$\alpha_p = S_{p1}(T(P_1'(u_1), P_2'(u_2), \dots), T(P_1(u_1), P_2(u_2), \dots))$$

$$\beta_p = S_{p2}(T(P_1'(u_1), P_2'(u_2), \dots), T(P_1(u_1), P_2(u_2), \dots))$$

- se obtiene la conclusión necesaria de la regla mediante la implicación general de necesidad, y la conclusión posible mediante la implicación general de posibilidad, utilizando en cada caso sus correspondientes índices:

$$Q_n'(v) = I_n(\alpha_n, \beta_n, Q(v))$$

$$Q_p'(v) = I_p(\alpha_p, \beta_p, Q(v))$$

- Conclusión de una regla con una única relación semántica, aplicando la separabilidad de los índices de similitud

El mecanismo de inferencia propuesto en el apartado anterior no tiene en cuenta las ventajas que se pueden derivar de la separabilidad que bajo ciertas condiciones pueden verificar las medidas de similitud. Por otro lado los modelos propuestos al final del capítulo 4 se basan en cuatro medidas de similitud, cuya separabilidad ha sido analizada en el capítulo 5, en el apartado 5.4.3. Recordemos que en lo que concierne a los modelos de reglas que se propone utilizar (véase el apartado 7.5.4.2) los principales resultados eran:

$$a) \Pi(-]T^+(P_1, P_2)[/ T^+(P'_1, P'_2)) =$$

$$\sup(T^+(\sup_{u1 / P1(u1)=0} P'_1(u_1), \sup_{u2} P'_2(u_2)),$$

$$T^+(\sup_{u2 / P2(u2)=0} P'_2(u_2), \sup_{u1} P'_1(u_1)))$$

$$b) N(T(P_1, P_2) / [T(P'_1, P'_2)]) = T(N(P_1/[P'_1]), N(P_2/[P'_2]))$$

$$c) \Pi(]T^+(P_1, P_2)[/ T^+(P'_1, P'_2)) = T^+(\Pi(]P_1/[P'_1]), \Pi(]P_2/[P'_2]))$$

$$d) \Pi(]T^+(P'_1, P'_2)[/ T^+(P_1, P_2)) = T^+(\Pi(]P'_1/[P_1]), \Pi(]P'_2/[P_2]))$$

Si además se tiene en cuenta que para los modelos de necesidad las distribuciones de necesidad deben estar normalizadas la propiedad a) puede reescribirse como:

$$a) \Pi(-]T^+(P_1, P_2)[/ T^+(P'_1, P'_2)) =$$

$$\sup(\sup_{u1 / P1(u1)=0} P'_1(u_1), \sup_{u2 / P2(u2)=0} P'_2(u_2))$$

Según estas expresiones se tiene que, para los modelos de necesidad:

- $N(P/[P'])$: es separable respecto a T^+
- $N(JP/[P'])=1-\Pi(-JP/[P'])$: también verifica una cierta separabilidad respecto a T^+
- $\min(N(JP/[P']),N(P/[P']))$: como ambas medidas son separables se pueden calcular independientemente y luego obtener su mínimo.
- $N(P/[P']) = 0$ si $N(JP/[P']) < 1$: ambas medidas son separables respecto a T^+ .

y para los modelos de posibilidad:

- $\Pi(JP'/[P])$: separable respecto a T^+ .
- $\Pi(JP/[P'])$: separable respecto a T^+
- $\min(\Pi(JP/[P']),\Pi(JP'/[P]))$: ambas son separables por lo que se pueden calcular por separado y luego obtener el mínimo.
- $\Pi(JP'/[P]) = 0$ si $P(u)+P'(u) < 1 \quad \forall u$: en este caso aunque $\Pi(JP'/[P])$ sí es separable respecto a una t-norma la suma no lo es. En efecto $T(P_1(u_1),P_2(u_2)) + T(P_1'(u_1),P_2'(u_2)) \neq T(P_1(u_1)+P_1'(u_1),P_2(u_2)+P_2'(u_2))$ y de hecho la segunda expresión puede no estar definida. Si por ejemplo utilizamos la t-norma \min se puede escribir:

$$\begin{aligned} & \min(P_1(u_1),P_2(u_2))+\min(P_1'(u_1),P_2'(u_2))= \\ & \min(P_1(u_1)+P_1'(u_1),P_2(u_2)+P_2'(u_2), P_1(u_1)+P_2'(u_2), P_2(u_2)+P_1'(u_1)) \leq \\ & \min(P_1(u_1)+P_1'(u_1),P_2(u_2)+P_2'(u_2)). \end{aligned}$$

Esto significa que la condición $\min(P_1(u_1)+P_1'(u_1),P_2(u_2)+P_2'(u_2)) < 1$ garantiza que el modelo no se activa pero es una condición menos restrictiva que la condición real del modelo. En efecto garantiza una intersección mínima en cada universo pero no a nivel global. Sin embargo conviene tener en cuenta que este modelo es apropiado solamente para una semántica muy particular y que por tanto no es de esperar que una base de conocimiento contenga reglas de este tipo. Además los modelos de certidumbre estricta son útiles principalmente cuando son modelos de necesidad. Recordando el ejemplo de regla de certidumbre estricta dada en el apartado 7.5.4.2, es prácticamente inútil afirmar que *es bastante posible que la temperatura del motor sea alta*, ya que el riesgo que entraña la

apertura del depósito de agua exige una seguridad total de que el antecedente se verifica con total certidumbre (y por tanto en el dominio de lo necesario).

Por tanto si se agregan los antecedentes de las reglas mediante una t-norma positiva, las medidas de similitud de los modelos propuestos son separables lo que permite simplificar considerablemente el proceso de inferencia de regla con múltiples antecedentes, dando lugar al siguiente algoritmo de inferencia:

- se obtienen en primer lugar los índices parciales de similitud para cada uno de los antecedentes de la regla, tanto en el dominio de lo necesario como en el dominio de lo posible:

$$\alpha_{n1} = S_{n1}(P_1'(u_1), P_1(u_1))$$

$$\beta_{n1} = S_{n2}(P_1'(u_1), P_1(u_1))$$

$$\alpha_{n2} = S_{n1}(P_2'(u_2), P_2(u_2))$$

$$\beta_{n2} = S_{n2}(P_2'(u_2), P_2(u_2))$$

...

...

$$\alpha_{p1} = S_{p1}(P_1'(u_1), P_1(u_1))$$

$$\beta_{p1} = S_{p2}(P_1'(u_1), P_1(u_1))$$

$$\alpha_{p2} = S_{p1}(P_2'(u_2), P_2(u_2))$$

$$\beta_{p2} = S_{p2}(P_2'(u_2), P_2(u_2))$$

...

...

$$P(u_1, u_2, \dots) = T(P_1(u_1), P_2(u_2), \dots)$$

$$P'(u_1, u_2, \dots) = T(P_1'(u_1), P_2'(u_2), \dots)$$

- aplicando la separabilidad se obtienen los índices de similitud globales.

$$\alpha_n = H_1(\alpha_{n1}, \alpha_{n2}, \dots)$$

$$\beta_n = H_2(\beta_{n1}, \beta_{n2}, \dots)$$

$$\alpha_p = H_3(\alpha_{p1}, \alpha_{p2}, \dots)$$

$$\beta_p = H_4(\beta_{p1}, \beta_{p2}, \dots)$$

- se obtiene la conclusión necesaria de la regla mediante la implicación general de necesidad, y la conclusión posible mediante la implicación general de posibilidad, utilizando en cada caso sus correspondientes índices:

$$Q_n'(v) = I_n(\alpha_n, \beta_n, Q(v))$$

$$Q_p'(v) = I_p(\alpha_p, \beta_p, Q(v))$$

7.5.4.4 Mecanismo propuesto para la determinación de los valores de verdad y certidumbre de un consecuente: esquema general de inferencia simplificado

A continuación describimos el mecanismo general de inferencia que se propone utilizar en esta tesis.

- Determinación de la conclusión de una regla con distintas relaciones semánticas entre consecuente y antecedentes.

Supongamos que partimos de una regla general del tipo:

$$G_1, G_2, \dots, C_1, C_2, \dots, M_1, M_2, \dots \rightarrow Q$$

donde G_1, G_2, \dots , son antecedentes que mantienen una relación gradual con D , C_1, C_2, \dots , son antecedentes que mantienen una relación de certidumbre con D y M_1, M_2, \dots , son antecedentes que mantienen una relación mixta con el consecuente Q . Vamos a tratar de determinar cuánto deben valer el grado de verdad y el grado de certidumbre de Q en función de los antecedentes de la regla.

En primer lugar para no manejar distribuciones conjuntas es necesario escoger un esquema de inferencia que verifique propiedades de separabilidad. Por ello optaremos por utilizar la modificación por similitud basada en los índices $N(P/[P'])$, $N(JP/[P'])$, $\min(N(JP/[P']), N(P/[P']))$ y $N(P/[P'])$ o 0 si $N(JP/[P']) < 1$ para los modelos de necesidad y $\Pi(JP'/[P])$, $\Pi(JP/[P'])$ y $\min(\Pi(JP/[P']), \Pi(JP'/[P]))$ para los modelos de posibilidad.

Además haremos la agregación de los antecedentes mediante una t-norma positiva T , en concreto la t-norma \min , de forma que sabemos que se verifican las propiedades de separabilidad vistas en el apartado anterior. De estas propiedades se deduce que cuando se agregan antecedentes, el grado de verdad de la conclusión se obtiene mediante una conjunción de los grados de verdad de los antecedentes que mantengan una relación gradual con el consecuente, mientras que el grado de incertidumbre se obtiene como la disyunción de todos aquellos grados que en el dominio de los antecedentes se traduzcan,

dependiendo del modelo, en incertidumbre en el dominio del consecuente. Dicho de otro modo podemos considerar que los grados de certidumbre también se obtienen como una conjunción de aquellas medidas que en el dominio de los antecedentes se traduzcan, dependiendo del modelo, en certidumbre en la conclusión.

Con estas consideraciones se puede proponer un esquema de inferencia simplificado para una regla de un único consecuente, de forma que el grado de verdad del consecuente venga dado por:

$$\alpha(Q) = \min(\alpha(G_1), \alpha(G_2), \dots, \alpha(M_1), \alpha(M_2), \dots)$$

y el grado de certidumbre por:

$$\beta(Q) = \min(\beta(G_1), \beta(G_2), \dots, \alpha(C_1), \beta(C_1), \alpha(C_2), \beta(C_2), \dots, \alpha(M_1), \beta(M_1), \alpha(M_2), \beta(M_2), \dots)$$

y se obtiene de este modo un mecanismo de inferencia intuitivamente razonable, rico en matices y sencillo de implementar, que es el que se propone en esta tesis para la aplicación del modus ponens generalizado.

- Agregación de las conclusiones procedentes de diferentes reglas

Para las reglas que concluyan una distribución de posibilidad para una determinada variable del sistema, como ya se vio a lo largo del capítulo 3, la agregación depende del tipo de distribución concluido, y por tanto del tipo de modelo empleado para la regla. Para las reglas con modelos de necesidad se escogerá una agregación conjuntiva, y para las reglas con modelos de posibilidad se escogerá una agregación disyuntiva. En ausencia de información adicional, que probablemente se corresponda con el caso más habitual, estos operadores deberán ser *min* y *max*.

Para aquellas reglas que concluyen valores de verdad borrosos que cualifiquen una determinada propiedad o concepto, es obvio que éstos están sujetos a las mismas reglas que cualquier otra distribución de posibilidad, y por tanto puede uno remitirse al caso anterior. Sin embargo utilizando la familia general propuesta en el capítulo 6 apartado 6.8.2, caracterizada por dos únicos parámetros, el primero correspondiente con un grado de verdad parcial y el segundo con un grado de posibilidad parcial, se vio en el capítulo 6 que es posible agregarlos de forma simplificada, independientemente del tipo de distribución de posibilidad que representan, mediante el operador *max*, con lo que recordando el resultado obtenido se tiene, para los valores de necesidad borrosos:

$$\begin{aligned}\tau_n(x) &= \min(\tau_{n1}(x), \tau_{n2}(x)) = \\ & \min(I_n(\alpha_1, \beta_1, x), I_n(\alpha_2, \beta_2, x)) \approx \\ & I_n(\max(\alpha_1, \alpha_2), \max(\beta_1, \beta_2), x)\end{aligned}$$

y del mismo modo para los valores de posibilidad borrosos:

$$\begin{aligned}\tau_P(x) &= I_p(\max(\alpha_1, \alpha_2), \max(\beta_1, \beta_2), x) = \\ & \max(\tau_{p1}(x), \tau_{p2}(x)) \approx \\ & \max(I_p(\alpha_1, \beta_1, x), I_p(\alpha_2, \beta_2, x))\end{aligned}$$

Esta simplificación es similar a suponer que el resultado de agregar varias distribuciones de posibilidad trapezoidales es otra distribución de posibilidad trapezoidales. Esto aunque supone una simplificación (cuya conveniencia convendría evaluar detalladamente) puede aportar grandes beneficios de cara a facilitar el cómputo de conclusiones, y de cara también a la interpretación final. En el caso de los valores de verdad borrosos resulta obvio que es más fácil interpretarlos a través de sus parámetros que a través de las distintas formas geométricas que puedan adoptar.

- Resumen

El mecanismo de inferencia que se propone aplicar, y que será ilustrado en el ejemplo apartado 7.6, queda resumido como sigue. A partir de una regla del tipo:

$$G_1, G_2, \dots, C_1, C_2, \dots, M_1, M_2, \dots \rightarrow Q$$

donde G_1, G_2, \dots , son antecedentes que mantienen una relación gradual con D , C_1, C_2, \dots , son antecedentes que mantienen una relación de certidumbre con D y M_1, M_2, \dots , son antecedentes que mantienen una relación mixta con el consecuente Q , proponemos obtener el grado de verdad del consecuente mediante la expresión:

$$\alpha(Q) = \min(\alpha(G_1), \alpha(G_2), \dots, \alpha(M_1), \alpha(M_2), \dots)$$

y el grado de certidumbre mediante:

$$\beta(Q) = \min(\beta(G_1), \beta(G_2), \dots, \alpha(C_1), \beta(C_1), \alpha(C_2), \beta(C_2), \dots, \alpha(M_1), \beta(M_1), \alpha(M_2), \beta(M_2), \dots)$$

Por otro lado, obtenidos para un mismo consecuente Q los grados de verdad $\alpha_1, \alpha_2, \dots$, y los grados de certidumbre β_1, β_2, \dots , procedentes de diferentes reglas, proponemos obtener los grados de verdad y de certidumbre finales mediante el operador max :

$$\begin{aligned}\alpha_F(Q) &= \max(\alpha_1(Q), \alpha_2(Q), \dots) \\ \beta_F(Q) &= \max(\beta_1(Q), \beta_2(Q), \dots)\end{aligned}$$

7.5.4.5 Tipos de antecedentes

- Antecedente básico

La inferencia tal y como se ha estudiado a lo largo de los capítulos 4 y 5 presuponía que un antecedente venía dado por una expresión del tipo:

si u es P

y que el valor real de u venía modelado por una distribución de posibilidad $P'(u)$ con la que debía contrastarse la hipótesis u es P . En lo que sigue utilizaremos la notación $\pi(u)$ para hacer referencia a la distribución de posibilidad que describe el valor de la variable u , con lo que para el caso anterior se tiene $\pi(u)=P'(u)$. Con esta notación, ante un hecho $\pi(u)=P'(u)$ la comprobación del antecedente anterior equivale a calcular los grados de verdad y de certidumbre de la proposición:

si u con distribución $\pi(u)=P'(u)$ es $P(u)$

El cálculo de las medidas de similitud ha sido analizado para antecedentes de tipo básico a lo largo del capítulo 4. Sin embargo es posible concebir múltiples tipos de antecedentes cuya formulación inicial puede ser más compleja que la de los antecedentes de tipo básico. Para poder contrastarlos con los hechos, es decir para calcular sus grados de verdad y de certidumbre los reduciremos a algún antecedente de tipo básico equivalente. Llamaremos evaluación de un antecedente a su reducción a un antecedente básico, de forma que el resultado de la evaluación sea por un lado una distribución de posibilidad observada $\pi(u)$, y por otro un conjunto borroso $P(u)$ que es la hipótesis que se contrasta. En algunas situaciones particularmente sencillas, la evaluación del antecedente puede dar lugar directamente a un valor de verdad borroso o clásico. Sin embargo y en general, en un antecedente básico la variable u puede ser una variable multidimensional, que puede expresarse de forma más detallada como:

si $u=(u_1, u_2, \dots, u_n)$ con distribución $\pi(u_1, u_2, \dots, u_n)$ es $P(u_1, u_2, \dots, u_n)$

En el caso en que se tenga:

$$\text{si } u_1 \text{ es } P_1(u_1) \text{ y } u_2 \text{ es } P_2(u_2)$$

se puede obtener el antecedente básico dado por:

$$\text{si } (u_1, u_2) \text{ con distribución } \pi(u_1, u_2) = \min(\pi(u_1), \pi(u_2)) \text{ es } \min(P_1(u_1), P_2(u_2))$$

en el que interviene una variable bidimensional. Por otro lado y teniendo en cuenta la separabilidad de los índices de similitud propuestos, este antecedente es equivalente a tantos antecedentes como operandos aparezcan en la conjunción, y por tanto se puede tratar como la conjunción de varios antecedentes básicos.

Es conveniente señalar que aunque estemos utilizando únicamente una letra para nombrar de forma abreviada a una medida, por ejemplo u , en general cuando se formule el conocimiento (en el fichero en el que se enuncian las reglas de conocimiento) deberá especificarse con un formato más completo, que por ejemplo incluya el nombre de la medida y el componente al que pertenece:

$$\text{nombre_medida en nombre_componente}$$

para que quede unívocamente determinada.

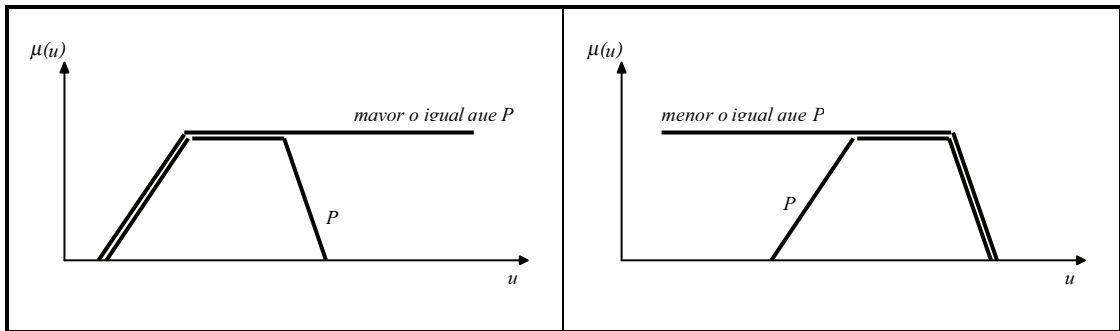
En lo que sigue describiremos algunos antecedentes que se consideran importantes de cara a la formulación del conocimiento funcional, e iremos proponiendo uno a uno un mecanismo de evaluación que permita reducirlo a un antecedente básico.

- Antecedentes ordinales

Llamamos antecedentes ordinales a aquellos en los que se tiene que llevar a cabo una comparación ordinal entre la observación de una variable y la hipótesis que se plantea. Su estructura es:

$$\begin{aligned} &u \text{ es mayor o igual que } P \\ &u \text{ es menor o igual que } P \end{aligned}$$

Estos antecedentes pueden reducirse a antecedentes básicos obteniendo los conjuntos borrosos *mayor_o_igual_que_P* y *menor_o_igual_que_P*, a partir del conjunto P , tal y como se ilustra en las siguientes figuras:

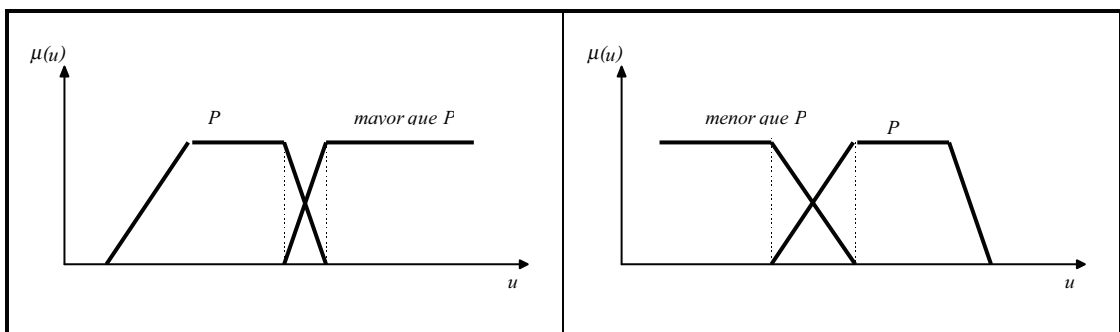


En el caso en que los antecedentes expresen relaciones de orden estrictas:

u es mayor que P

u es menor que P

pueden también ser reducidos a antecedentes básicos del siguiente modo:



- Antecedentes funcionales

Diremos que un antecedente es funcional cuando tenga la siguiente estructura:

$$si v=f(u_1, u_2, \dots, u_3) es P$$

donde P es un posible estado de la función f .

De entre los antecedentes funcionales cabe destacar algunos de los más comunes, como son:

si tendencia(u) es alta
si correlación(u₁,u₂) es baja

Si los valores de las medidas involucradas son numéricos entonces el valor de la función es también un resultado numérico. Si por el contrario los valores de las medidas vienen descritos mediante distribuciones de posibilidad, entonces la aplicación del principio de extensión permite obtener la distribución de posibilidad que describe el valor de la función f :

$$\pi(v) = \sup_{u_1, u_2, \dots, u_n / v=f(u_1, u_2, \dots, u_n)} \min(\pi(u_1), \pi(u_2), \dots, \pi(u_n))$$

con lo que obtenida la distribución $\pi(v)$ el antecedente funcional se reduce a un antecedente básico:

si v es P

donde la observación de v viene dada por $\pi(v)$.

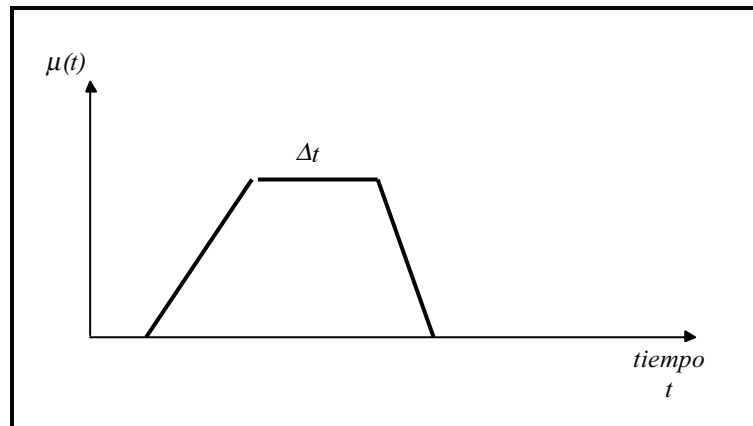
- Antecedentes temporales

Diremos que un antecedente es temporal cuando exprese una relación temporal entre valores de medidas del sistema, lo que puede ser de gran utilidad para aquellos casos en que las secuencias temporales de los valores de algunas medidas sean decisivas en el proceso de diagnóstico.

El antecedente temporal más sencillo es:

si u₁ es P₁ Δt antes que u₂ es P₂

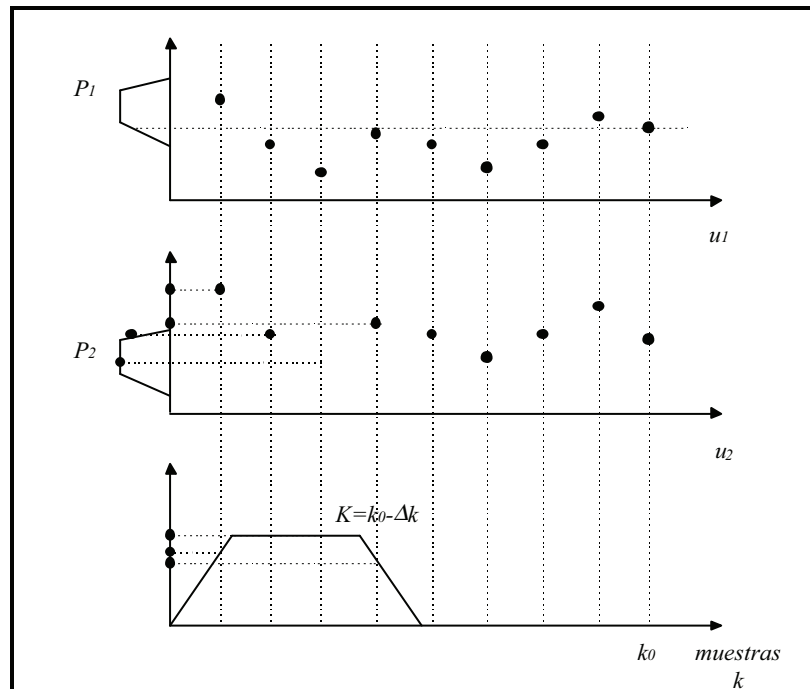
donde tanto P_1 como P_2 son borrosos, y Δt expresa un intervalo temporal borroso, como se ilustra en la siguiente figura:



Llamando Δk la traducción a muestras del intervalo temporal borroso Δt , el antecedente propuesto puede plantearse de forma equivalente como:

$$\text{si } u_1(k_0) \text{ es } P_1 \text{ y } u_2(k_0 - \Delta k) \text{ es } P_2$$

Cuando las medidas u_1 y u_2 tienen valores numéricos entonces podemos representar gráficamente los distintos elementos que intervienen en el antecedente mediante la siguiente figura:



En este caso podemos decir que la variable u_1 en el instante k_0 se distribuye $\{u_1(k_0)\}$ siendo $u_1(k_0)$ un valor numérico. A la variable u_2 por el contrario le corresponde una distribución $\{u_2(k_i)\}$ para cada instante k_i . Como además debe cumplirse que k_i debe pertenecer al conjunto borroso $K=k_0-\Delta k$ podemos expresar la distribución de la variable $u_2(k_0-\Delta k)$ mediante un conjunto borroso obtenido según la expresión:

$$\pi_K(u_2) = \max_i (\min(\{u_2(k_i)\}, K(k_i)))$$

con lo que el antecedente temporal se reduce al antecedente básico:

$$\text{si } (u_1, u_2) \text{ con distribución } \pi(u_1, u_2) \text{ es } P(u_1, u_2)$$

donde $\pi(u_1, u_2) = \min(\{u_1(k_0)\}, \pi_K(u_2))$ y $P(u_1, u_2) = \min(P_1(u_1), P_2(u_2))$.

En el caso en que las medidas u_1 y u_2 tomen valores borrosos, expresados mediante distribuciones de posibilidad, siendo $\pi_i(u_j)$ la distribución de posibilidad de la variable u_j para la muestra k_i , podemos generalizar el resultado anterior. Para estimar la distribución de posibilidad de la variable $u_2(k_0-\Delta k) = u_2(K) = \pi_K(u_2)$ utilizamos la siguiente expresión:

$$\pi_K(u_2) = \max_i (\min(\pi_i(u_2), K(k_i)))$$

con lo que la distribución conjunta de las variables $u_1(k_0)$ y $u_2(k_0-\Delta k)$ puede obtenerse como:

$$\pi(u_1, u_2) = \min(\pi_0(u_1), \pi_K(u_2))$$

y nuevamente el antecedente temporal se reduce al antecedente básico:

$$\text{si } (u_1, u_2) \text{ con distribución } \pi(u_1, u_2) \text{ es } P(u_1, u_2)$$

con $P(u_1, u_2) = \min(P_1(u_1), P_2(u_2))$.

- Antecedentes de existencia y de relación

A menudo es necesario expresar conocimiento funcional condicionado a la existencia de algún tipo de componente especial en el proceso, o bien a la existencia de alguna relación determinada entre componentes. Su estructura es del tipo:

si existe componente
si componente1 relación componente2

Estos antecedentes en principio no contienen ningún tipo de vaguedad ni de incertidumbre, y por tanto producen un valor de verdad que es directamente 0 o 1. Por ello no es necesario reducirlos al tipo básico, ya que basta con comprobar si el componente en cuestión o la relación entre los componentes han sido o no definidas en la descripción estructural del proceso. En caso afirmativo el antecedente se ignora y se procede a la reducción al tipo básico de los demás antecedentes de la regla. En caso negativo la regla no puede dispararse por no verificarse sus antecedentes y es descartada. Esto significa que estos antecedentes deberán ser los que primero se comprueben.

7.5.4.6 Tipos de consecuentes

Hemos denominado consecuentes únicamente a aquellos consecuentes que concluyen sobre el valor de una variable del sistema, para distinguirlos de los diagnósticos y de las acciones:

$$v \text{ es } Q(v)$$

El consecuente anterior puede cualificarse con un valor de verdad borroso que matice el concepto expresado por $Q(v)$:

$$v \text{ es } Q(v) (N, \alpha, \beta)$$

siendo entonces equivalente al consecuente de tipo básico:

$$v \text{ es } I_n(\alpha, \beta, Q(v))$$

7.5.4.7 Diagnósticos

Los diagnósticos pueden considerarse como medidas especiales asociadas a componentes del proceso, que describen su estado de salud, y cuyos valores pueden ser, o bien un valor de verdad borroso (concluido por el sistema experto de diagnóstico) y que no necesitan ser definidos, o bien un descriptor simbólico o borroso que cualifique por ejemplo la gravedad o magnitud del fallo al que el diagnóstico hace referencia. En el caso en que para un diagnóstico puedan definirse estados (que no sean valores de verdad borrosos) el diagnóstico debe describirse como otra medida más en la base de hechos del sistema. Una medida se interpreta como un diagnóstico (es decir que su valor se presenta al usuario al final de cada proceso de inferencia, y pasa a formar parte de la base de diagnósticos) en el momento en que aparece en alguno de los diagnósticos de una regla, que recordemos venían encabezados por la palabra clave *diagnósticos*.

En el caso los posibles valores de un diagnóstico sean valores de verdad borrosos el diagnóstico queda descrito por su nombre, el componente al que está asociado, y su valor de verdad borroso:

diagnóstico en componente es (tipo, α, β)

donde $(\text{tipo}, \alpha, \beta)$ es un valor de verdad borroso tal y como se definió en el capítulo 6. Por otro lado la nomenclatura utilizada ya ha sido definida en el apartado 7.3.2.1. El valor de este tipo de diagnósticos queda descrito directamente por su nombre y su valor borroso inferido. Recordemos que cuando varias reglas concluyen sobre un mismo diagnóstico de este tipo, de cada una de ellas se infiere un valor de verdad borroso y como vimos en el capítulo 6, independientemente del tipo de valor de verdad concluido (valor de necesidad o valor de posibilidad borroso) los grados de verdad y certidumbre finales se obtienen tomando el máximo de todos los grados de verdad y todos los grados de certidumbre provenientes de cada una de las reglas disparadas.

Para los diagnósticos con estados simbólicos su valor después de la inferencia queda descrito por el valor de verdad borroso asociado a cada uno de sus posibles estados siempre que la certidumbre sea no nula, por lo que todos ellos deben ser enumerados, y sólo es necesario llevar a cabo agregaciones cuando varias reglas concluyan sobre el mismo valor de un determinado diagnóstico.

Para los diagnósticos con estados modelados con conjuntos borrosos sobre universos numéricos, su valor queda descrito por la distribución de posibilidad obtenida después del proceso de inferencia. La distribución se obtiene agregando las distribuciones inferidas por cada una de las reglas involucradas. Recordemos que las distribuciones de necesaria posibilidad pueden agregarse utilizando el operador conjuntivo *min* y que para las distribuciones de posible posibilidad debe utilizarse el operador disyuntivo *max*.

Como hemos visto en la descripción de los modelos de las reglas hay un modelo que permite incrementar o decrementar la certidumbre que sobre un diagnóstico se pueda tener previamente, y su formato es:

diagnóstico en componente es (tipo, tipo_incremento, β)

donde tipo es *N* o *Π*, tipo_incremento es *I* (incrementar) o *D* (decrementar) y β indica en qué medida se debe incrementar o decrementar dicha certidumbre.

7.5.4.8 Tipos de acciones

Puede considerarse que en general las acciones son específicas del proceso considerado, por lo que no resulta obvio clasificarlas a priori. Sí que parece al menos razonable considerar dos tipos de acciones de carácter muy general, como pueden ser:

- Acciones de interacción con el sistema de control del proceso para cambiar su punto de operación, lo que a menudo se conoce como “*failure acomodation*” ([Polycarpou, Vemuri, 95], [Frank, 90]).
- Acciones de interacción con el sistema de mantenimiento para replanificar las acciones de mantenimiento en base a la información de diagnóstico disponible.

Las acciones son desde el punto de vista de la inferencia similares a los diagnósticos, por lo que también admiten un formato similar. Aunque no parece necesario asociarlas a un componente determinado, siempre puede recurrirse al componente *proceso* cuando no puedan localizarse en componentes de jerarquía inferior. Por tanto su formato puede ser:

acción en componente (tipo, α, β)

o bien, en el caso en que se asocien al proceso omitirlo en su formulación:

acción (tipo, α, β)

La certidumbre de las acciones es un concepto delicado, ya que una acción se lleva a cabo o no. Sin embargo es inevitable tener en cuenta que desde el momento en que la comprobación de las hipótesis de una regla conlleva una cierta incertidumbre, esa incertidumbre debe manifestarse de uno u otro modo en las acciones que dicha regla recomienda. Esta certidumbre podrían tal vez utilizarse para establecer, conjuntamente con otros criterios (como por ejemplo nivel de complicación, nivel de peligrosidad de no llevar a cabo la acción, etc.) un conjunto de prioridades entre las acciones que el sistema recomiende que sean llevadas a cabo. En concreto a acciones críticas para el buen funcionamiento del sistema se les puede asociar certidumbres muy elevadas, de forma que se conviertan en prioritarias frente a acciones con certidumbres menores, para igual cumplimiento de los antecedentes. También se podría graduar la prioridad de la acción en términos de alta, media y baja, y establecer relaciones graduales con los antecedentes apropiados.

7.6 Caso ejemplo: diagnóstico de un transformador de potencia mediante el análisis de los gases disueltos en el aceite

Nos basaremos, para este ejemplo, en algunos de los resultados obtenidos durante el desarrollo, por el equipo investigador al que pertenezco, del sistema experto de diagnóstico TRAFES (en el que yo mismo participé activamente en la construcción del sistema experto de diagnóstico). Este sistema fue desarrollado para UEFSA, financiado por el PIE (Plan de Investigación Electrotécnica) y con la colaboración de otras entidades como la Escuela Técnica Superior de Ingenieros Industriales de Madrid, ABB Trafosur, etc.

TRAFES ([Sanz et al., 93] y [Sanz, Villar, Palacios, 95]) es un sistema de experto para el diagnóstico en continuo de transformadores de potencia, cuyas principales fuentes de información son, por un lado, el conocimiento experto almacenado en forma de reglas, y por otro, las medidas tanto manuales como automáticas que llegan al sistema. La conjunción de ambas fuentes permite al sistema, mediante reglas de inferencia, deducir nuevas conclusiones que explican las anomalías que el sistema de detección de anomalías haya podido detectar.

Un transformador de potencia es una máquina eléctrica que con un rendimiento elevado transforma una señal eléctrica en su entrada, en otra de igual potencia, y mayor o menor tensión, dependiendo de la relación de transformación. Su comportamiento eléctrico es similar al de otros tipos de transformadores, pero sus características físicas y estratégicas son diferentes.

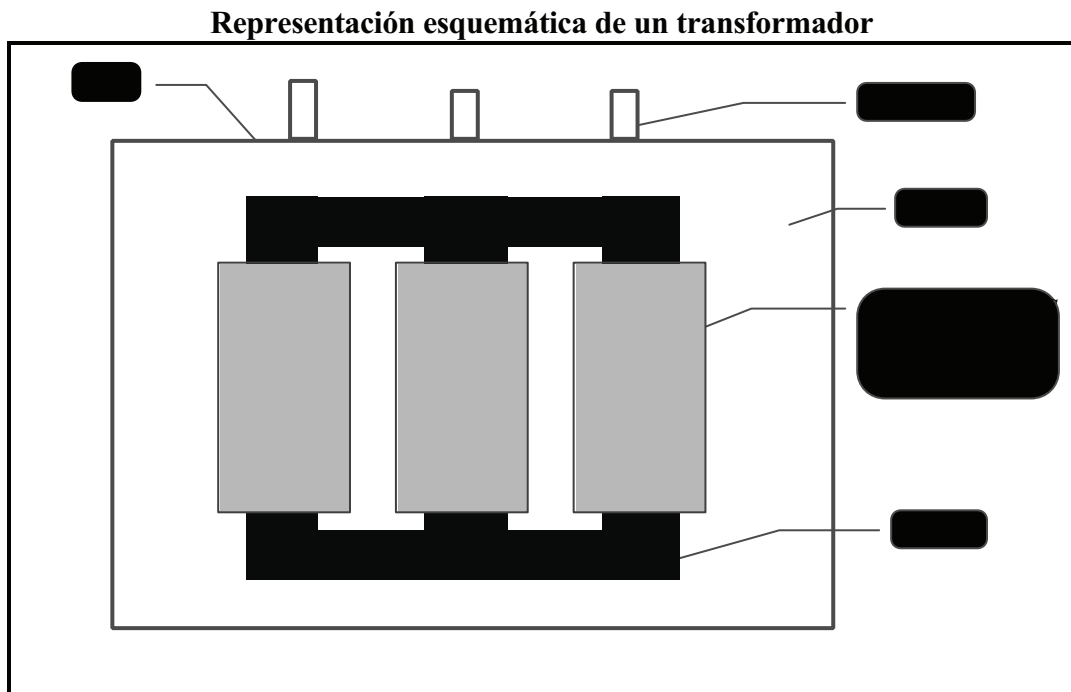
Al transformar señales eléctricas de elevada potencia, fallos internos en este tipo de máquinas pueden dar lugar a explosiones violentas que deterioren considerablemente a la máquina, con elevados costes de reparación. Por otro lado, de su buen funcionamiento suele depender el suministro eléctrico de zonas geográficas importantes, por lo que su puesta fuera de servicio puede acarrear importantes deterioros de la calidad del servicio eléctrico. Esto hace que sea importante conocer en cada momento el estado de salud del transformador de cara a poder determinar su disponibilidad y planificar adecuadamente su mantenimiento.

Una característica común a muchos transformadores de cierto tamaño es la utilización de aceite como aislante eléctrico y refrigerante. El transformador se construye de forma que se introduce en una cuba llena de aceite, que es mejor aislante eléctrico que el aire y contribuye además a refrigerar el transformador. Sus mejores propiedades dieléctricas

permiten reducir las dimensiones del transformador frente a las que se obtendrían utilizando aire en su lugar.

Muchos de los posibles fallos de un transformador dejan algún tipo de huella en su aceite en forma de gases disueltos, debido a su propia descomposición o a la de otros aislantes sólidos también presentes (como por ejemplo materiales de celulosa, etc.). La medida de las concentraciones de estos gases permite orientar la identificación del origen de los posibles fallos o defectos que se puedan presentar.

Una representación muy esquemática de un transformador trifásico aparece en la siguiente figura:



7.6.1 Principales fallos detectables mediante el análisis de los gases disueltos en el aceite

Una de las formas más extendidas para el seguimiento y vigilancia del estado de salud de los transformadores de potencia es el análisis periódico de los gases disueltos en el aceite de la cuba.

Como ya hemos comentado, debido a la descomposición del aceite o de los demás aislantes sólidos que se utilizan en los transformadores, las concentraciones de determinados gases característicos pueden verse significativamente afectadas. La medida de estas concentraciones permite facilitar el seguimiento y diagnóstico de los fallos que se puedan originar.

La descomposición del aceite puede ser debida a dos fenómenos diferentes, descomposición por bombardeo iónico y descomposición térmica. La descomposición por bombardeo iónico se debe a la aparición de descargas parciales (por ruptura dieléctrica del aceite) y da principalmente lugar a la creación de hidrógeno (H_2). La descomposición térmica se debe a un calentamiento excesivo del aceite, por ejemplo por la existencia de zonas mal refrigeradas, y dependiendo de la temperatura que llegue a alcanzar puede dar lugar a la aparición de distintos tipos de hidrocarburos. Así cuando la temperatura es similar a la temperatura de funcionamiento normal (por debajo de 150 °C) se produce principalmente gas metano (CH_4) pero en pocas cantidades. Cuando la temperatura es ligeramente superior a la anterior la producción de metano aumenta. Para temperaturas mayores comienzan a aparecer etano (C_2H_6) y etileno (C_2H_4), y para altas temperaturas la cantidad de acetileno (C_2H_2) puede llegar a ser significativa. En definitiva cuanto mayor es la temperatura del fallo, menor es la proporción de hidrógeno en los hidrocarburos producidos.

Las descargas parciales (eléctricas) son siempre fenómenos de baja energía, por lo que además de producción de hidrógeno pueden dar lugar a producción de metano, pero no a otros hidrocarburos más saturados.

Las descargas eléctricas, al igual que las parciales, dan lugar a hidrógeno por rotura dieléctrica, pero al ser fenómenos de mayor energía pueden dar lugar a la aparición de hidrocarburos más saturados.

Por último, los defectos térmicos o puntos calientes son calentamientos locales del aceite y solo dan lugar a la producción de hidrocarburos, más o menos saturados dependiendo de que la falta térmica sea de mayor o menor temperatura.

El análisis de los gases disueltos, que como decimos es una práctica común, viene recomendado por ejemplo en una norma de la Comisión Electrotécnica Internacional, (CEI o IEC, véase [IEC, 78]) donde se proponen criterios concretos para su utilización en diagnóstico. En este ejemplo nos hemos basado en esta norma para formular las reglas de conocimiento utilizadas para ilustrar la existencia de diferentes relaciones semánticas entre antecedentes y consecuentes, y para la determinación de las mismas de cara a formular los modelos de reglas más apropiados. En dicha norma se propone analizar el valor de los cocientes de concentraciones siguientes:

$$\frac{CH_4}{H_2} \quad \frac{C_2H_4}{C_2H_6} \quad \frac{C_2H_2}{C_2H_4}$$

El análisis de los mismos permite establecer criterios de diagnóstico. En lo que respecta a la medida CH_4/H_2 se tiene:

si CH_4/H_2 bajo (H_2 alto y/o CH_4 bajo)

*descargas parciales de baja energía
descargas parciales de alta energía*

si CH_4/H_2 medio (H_2 y/o CH_4 similares)

*sin fallo
descargas eléctricas de baja energía
descargas eléctricas de alta energía
defecto térmico de baja temperatura*

si CH_4/H_2 alto (H_2 bajo y/o CH_4 alto)

*defecto térmico de baja-media temperatura
defecto térmico de media temperatura
defecto térmico de alta temperatura*

Para la medida C_2H_4/C_2H_6 se tiene:

si C_2H_4/C_2H_6 bajo (C_2H_6 alto y/o C_2H_4 bajo)

*sin fallo
descargas parciales de baja energía
descargas parciales de alta energía
defecto térmico de baja-media temperatura*

si C_2H_4/C_2H_6 medio (C_2H_6 y C_2H_4 similares)

*descargas eléctricas de baja energía
defecto térmico de baja temperatura
defecto térmico de media temperatura*

si C_2H_4/C_2H_6 alto (C_2H_6 bajo y/o C_2H_4 alto)

*descargas eléctricas de alta energía
defecto térmico de alta temperatura*

Por último para la medida C_2H_2/C_2H_4 se tiene:

si C_2H_2/C_2H_4 bajo (C_2H_2 alto y/o C_2H_4 bajo)

sin fallo

descargas parciales de baja energía

defecto térmico de baja-media temperatura

defecto térmico de baja temperatura

defecto térmico de media temperatura

defecto térmico de alta temperatura

si C_2H_2/C_2H_4 medio (C_2H_2 y C_2H_4 similares)

descargas parciales de alta energía

descargas eléctricas de baja energía

descargas eléctricas de alta energía

si C_2H_2/C_2H_4 alto (C_2H_2 bajo y/o C_2H_4 alto)

no se contempla nada más que como límite, cuando

se producen descargas eléctricas de baja energía

Otros parámetros importantes relacionados con las concentraciones de los gases anteriores, que también deben ser vigilados, son las tendencias o velocidades de crecimiento de sus concentraciones. Aunque la norma CEI no es especialmente clara a este respecto, si establece ciertos criterios que pueden dar lugar a algunas reglas de conocimiento.

Si llamamos Δ_{gas} a la variación de la concentración del gas "gas", el estudio de su dependencia con distintos parámetros de funcionamiento puede aportar información sobre el tipo de falta presente. Por un lado la velocidad de crecimiento es indicativa de la gravedad (o proximidad) de la falta, y por otro lado la dependencia con unos u otros parámetros de funcionamiento del transformador permite facilitar la identificación de la falta.

Una dependencia lineal de Δ_{gas} con el tiempo durante el cual el transformador está en servicio, pero independiente de la carga, permite apuntar a la existencia de descargas parciales debidas a fallos de aislamiento (dependientes de las tensiones) o a cortocircuitos en las chapas del núcleo magnético (sometidas a tensión siempre que el transformador está en servicio).

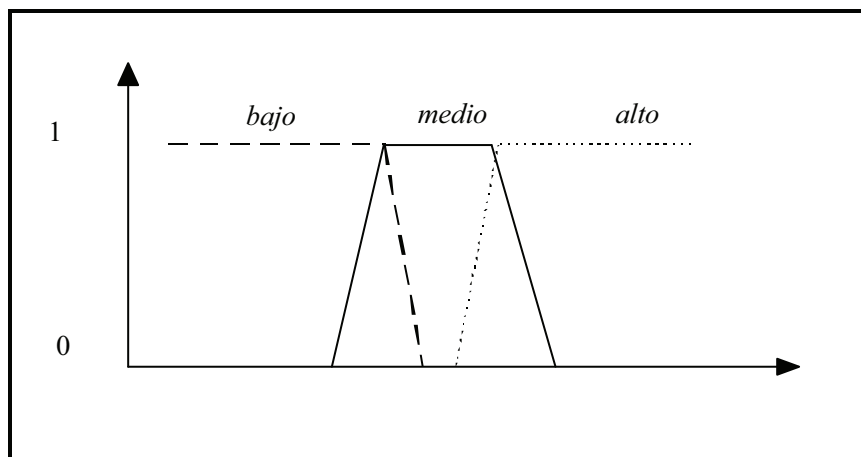
Una dependencia lineal de Δ_{gas} con la integral de la intensidad al cuadrado a lo largo del tiempo, es decir con las pérdidas óhmicas del transformador, ponen de manifiesto la posible existencia de defectos térmicos relacionados con sobrecalentamiento de conductores (lo que se produce para intensidades elevadas).

Señalar por último que aunque la norma hace otras consideraciones adicionales, respecto a los niveles mínimos que deben tener las concentraciones para que el resultado de los cocientes se pueda interpretar con los criterios que propone, éstas serán por simplicidad ignoradas a lo largo de todo este ejemplo. Sin embargo sería posible matizar la certidumbre de las observaciones atendiendo precisamente a la certidumbre de las medidas y a la operación matemática, en este caso el cociente, que permite obtener las medidas que realmente son utilizadas en las reglas de conocimiento que se formulen.

7.6.2 Las medidas utilizadas y sus estados

Según las recomendaciones de la norma CEI se proponen para cada una de las variables CH_4/H_2 , C_2H_4/C_2H_6 y C_2H_2/C_2H_4 límites rígidos para determinar si están en uno u otro estado.

Como ya se ha dicho la utilización de límites rígidos puede producir cambios bruscos de unos diagnósticos a otros, sin que existan transiciones de los grados de verdad o de certidumbre, haciendo más errático el funcionamiento del sistema de diagnóstico, y demasiado sensible a variaciones posiblemente no significativas. Además como veremos a lo largo de este ejemplo el uso de estados borrosos junto con la elección apropiada de los modelos de las reglas permite obtener una información más precisa que la suministrada por un sistema clásico con mismo número de reglas. Para borrosificar los límites propuestos en la norma hemos escogido particiones del tipo que aparece en la siguiente figura:

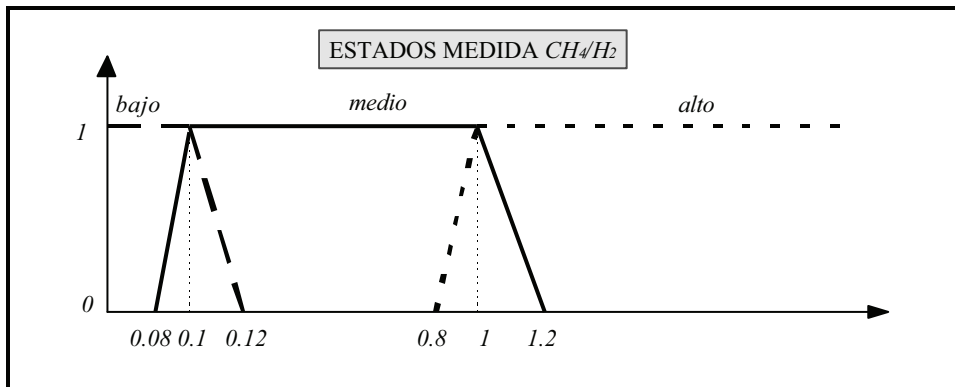


Recordemos que este tipo de particiones presenta buenas propiedades, que ya han sido comentadas en apartados anteriores (5.6.3.2 y 7.3.2.2).

Para la primera de las variables los estados implícitamente propuestos en la norma son:

$CH_4/H_2 < 0.1$	CH_4/H_2 <i>baja</i>
$0.1 < CH_4/H_2 < 1$	CH_4/H_2 <i>media</i>
$1 < CH_4/H_2$	CH_4/H_2 <i>alta</i>

Para borrosificarlos se puede determinar para qué valores comienza a ser dudoso el concepto expresado por el estado, y para qué valores dicho concepto deja de verificarse. Es obvio que la concentración de $CH_4/H_2=0.099$ no pasa bruscamente de ser baja a alta cuando aumenta por ejemplo en 0.001 . El límite superior del núcleo del estado *baja* puede determinarse como aquel valor de concentración de CH_4/H_2 a partir del cual comienza a ser dudoso o discutible que la concentración de CH_4/H_2 sea baja. El límite superior del soporte puede determinarse como aquel valor a partir del cual se tiene la certeza de que la concentración no puede considerarse baja. En el caso en que distintos expertos o distintas fuentes de conocimiento proporcionen valores distintos, se puede recurrir a una media que pueda tener más o menos en consideración todos los criterios disponibles. Conviene tener en cuenta que es más importante disponer de una transición gradual razonable, de la pertenencia a la no pertenencia, que el valor exacto de las mismas. Por último el ancho de las zonas de transición puede ser constante a lo largo de todo el universo o variar con la magnitud de los valores del mismo. Este último caso se corresponde con una visión logarítmica del universo de discurso de la medida, donde lo que prima es el orden de magnitud de sus valores frente al valor concreto, y es el que aquí se ha elegido. La siguiente figura representa los estados y la partición escogida:



Con estas consideraciones la medida CH_4/H_2 podría quedar definida como sigue:

medida

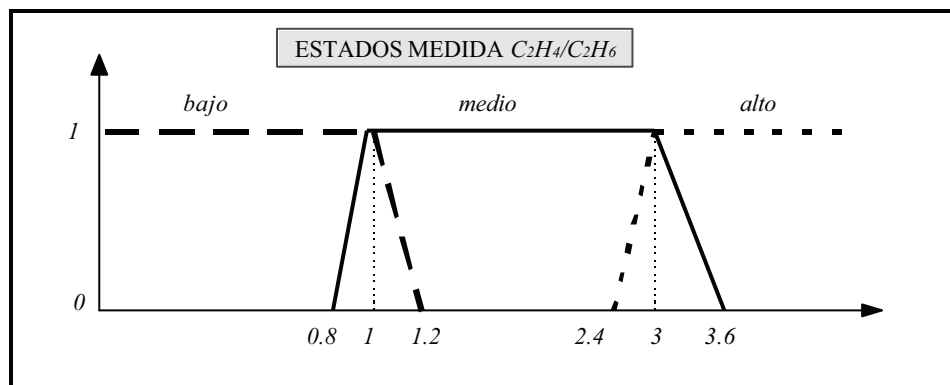
<i>nombre</i>	CH_4/H_2
<i>tipo</i>	<i>manual</i>
<i>propiedades</i>	<i>cociente_gas</i>

pertenece_a *aceite*
escala *0 10*
estados *baja(0,0,0.1,0.12)*
 media(0.08,0.1,1,1.2)
 alta(0.8,1,10,10)

Del mismo modo para la medida C_2H_4/C_2H_6 se pueden obtener los siguientes estados clásicos:

$C_2H_4/C_2H_6 < 1$	C_2H_4/C_2H_6 <i>baja</i>
$1 < C_2H_4/C_2H_6 < 3$	C_2H_4/C_2H_6 <i>media</i>
$3 < C_2H_4/C_2H_6$	C_2H_4/C_2H_6 <i>alta</i>

Los nuevos estados borrosos y la partición propuesta quedan reflejados en la siguiente figura:



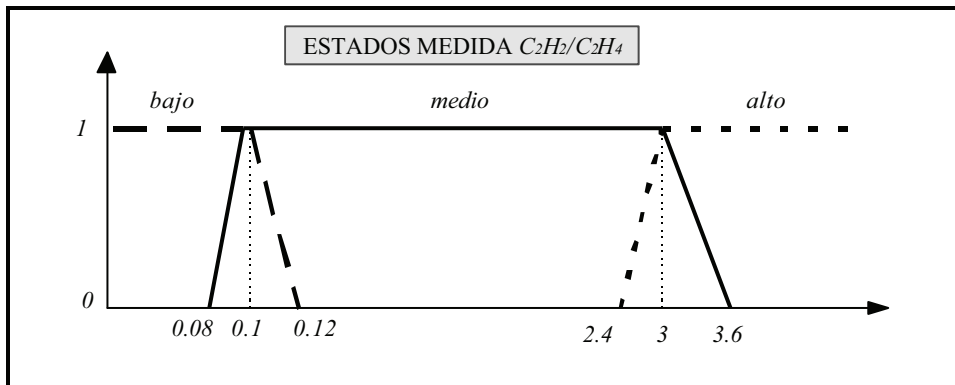
La definición de la medida podría ser:

medida
nombre C_2H_4/C_2H_6
tipo *manual*
propiedades *cociente_gas*
pertenece_a *aceite*
escala *0 10*
estados *baja(0,0,1,1.2)*
 media(0.8,1,3,3.6)
 alta(2.4,3,10,10)

Por último los estados clásicos propuestos en la norma para la medida C_2H_2/C_2H_4 son:

$$\begin{array}{ll}
 C_2H_2/C_2H_4 < 0.1 & C_2H_2/C_2H_4 \text{ baja} \\
 0.1 < C_2H_2/C_2H_4 < 3 & C_2H_2/C_2H_4 \text{ media} \\
 3 < C_2H_2/C_2H_4 & C_2H_2/C_2H_4 \text{ alta}
 \end{array}$$

Los estados y la partición aquí propuestos quedan reflejados en la siguiente figura:



y una posible definición de la medida podría ser:

medida

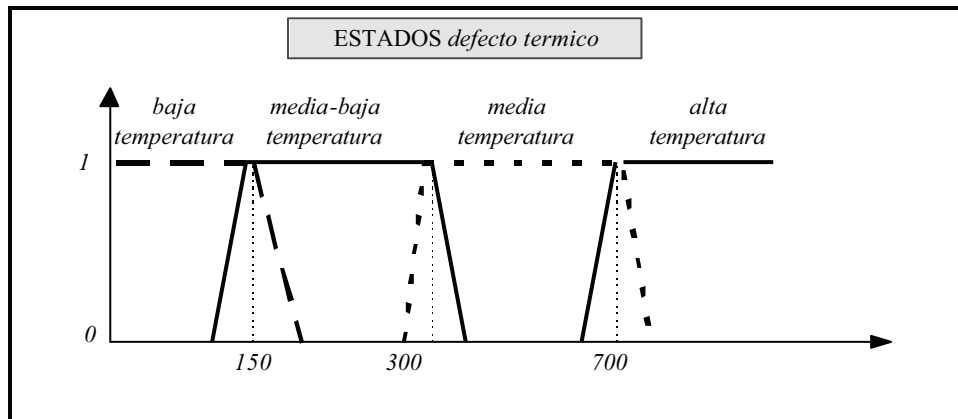
<i>nombre</i>	C_2H_2/C_2H_4
<i>tipo</i>	manual
<i>propiedades</i>	cociente_gas
<i>pertenece_a</i>	aceite
<i>escala</i>	0 10
<i>estados</i>	baja(0,0,0.1,0.12)
	media(0.08,0.1,3,3.6)
	alta(2.4,3,10,10)

Por otro lado como ya se ha comentado los diagnósticos pueden tratarse como medidas especiales de los componentes. En el ejemplo que nos ocupa las medidas que se derivan de los diagnósticos son:

descargas parciales
descargas eléctricas
defecto térmico

Además cada una de ellas lleva asociado un conjunto de posibles estados cualitativos. En concreto la medida *descargas parciales* puede tomar los valores *alta energía* y *baja energía*, la medida *descargas eléctricas* puede tomar los valores *alta energía* y *baja energía* y por último la medida *defecto térmico* puede tomar los valores *baja*

temperatura, media-baja temperatura, media temperatura y alta temperatura. Estos estados, en tanto en cuanto son cualitativos, no es necesario definirlos en la base de hechos del sistema, ya que pueden ser automáticamente creados durante la lectura de la base de conocimiento. Sin embargo en los tres casos, y especialmente en el último de ellos resulta bastante directo definir sobre un universo numérico los citados estados. Así por ejemplo, para el defecto térmico se puede definir la partición:



donde los límites de los conjuntos borrosos empleados estarían por determinar. Se podría entonces inferir directamente una distribución de posibilidad que describiese de qué temperatura el sistema estima que es el defecto térmico. Esta descripción se incluiría en la base de hechos del sistema.

7.6.3 Reglas de conocimiento

Relacionando cada diagnóstico por separado con los síntomas que permiten detectarlo se pueden formular de forma esquemática, al igual que hace la norma, las siguientes reglas de conocimiento:

regla G1

si

CH_4/H_2 medio

C_2H_4/C_2H_6 bajo

C_2H_2/C_2H_4 bajo

entonces

sin fallo

regla G2

si

CH_4/H_2 bajo

C_2H_4/C_2H_6 bajo

C_2H_2/C_2H_4 bajo

entonces

descargas parciales de baja energía

regla G3

si

CH_4/H_2 bajo

C_2H_4/C_2H_6 bajo

C_2H_2/C_2H_4 medio

entonces

descargas parciales de alta energía

regla G4

si

CH_4/H_2 medio

C_2H_4/C_2H_6 medio

C_2H_2/C_2H_4 medio

entonces

descargas eléctricas de baja energía

regla G5

si

CH_4/H_2 medio

C_2H_4/C_2H_6 alto

C_2H_2/C_2H_4 medio

entonces

descargas eléctricas de alta energía

regla G6

si

CH_4/H_2 medio

C_2H_4/C_2H_6 medio

C_2H_2/C_2H_4 bajo

entonces

defecto térmico de baja temperatura

regla G7

si

CH_4/H_2	<i>alto</i>
C_2H_4/C_2H_6	<i>bajo</i>
C_2H_2/C_2H_4	<i>bajo</i>

entonces

defecto térmico de baja-media temperatura

regla G8

si

CH_4/H_2	<i>alto</i>
C_2H_4/C_2H_6	<i>medio</i>
C_2H_2/C_2H_4	<i>bajo</i>

entonces

defecto térmico de media temperatura

regla G9

si

CH_4/H_2	<i>alto</i>
C_2H_4/C_2H_6	<i>alto</i>
C_2H_2/C_2H_4	<i>bajo</i>

entonces

defecto térmico de alta temperatura

Teniendo en cuenta las consideraciones relacionadas con la velocidad de producción de gases y su dependencia con parámetros de funcionamiento del transformador es posible enunciar dos reglas adicionales. Sin embargo, y en base a la norma, el conocimiento relativo a la producción de gases es escaso, y no parece que pueda utilizarse directamente para identificar faltas, aunque sí se desprende su posible uso para confirmar diagnósticos identificados mediante otros criterios más seguros como los de las reglas anteriores. Por ello este conocimiento se utilizará en reglas de incremento de certidumbre para posibles defectos identificados, pero no para su identificación.

regla TGPI

si

tendencia(gas) alta
dependencia(gas, tiempo_servicio) alta

entonces

incrementar certeza en descargas parciales

regla TGP2*si**tendencia(gas) alta**dependencia(gas,perdidas_ohmicas) alta**entonces**incrementar certeza en defecto térmico*

En estas dos reglas se está utilizando la propiedad *gas*, que supuestamente deberían compartir todas las medidas de concentraciones de gases. Aunque no han sido definidas en este ejemplo, obviamente deberían estarlo en el sistema real. Además, de cara a estas dos reglas, no sería necesario siquiera definir estados para estas medidas, sólo sería necesario definirlos para la función *tendencia*. Nótese que si el resultado de esta función se normaliza adecuadamente pueden tratarse de forma uniforme las tendencias de cada uno de los gases, dando lugar a un único conjunto de estados para una única función *tendencia*. Cada una de las reglas *TGP1* y *TGP2* se expandirían en tantas reglas como medidas con la propiedad *gas* hubiesen sido definidas en la base de hechos o fichero de configuración del sistema.

7.6.4 Determinación de los modelos de las reglas de conocimiento

El método que se ha seguido ha sido el propuesto en el apartado 7.5.4.2, que ilustraremos con algunos ejemplos concretos, determinando el modelo de algunas de las reglas anteriores. Posteriormente utilizaremos estas reglas para ilustrar el mecanismo de inferencia propuesto en el apartado 7.5.3.1.

- Ejemplo regla *G1*:

Para determinar la semántica de la regla analizamos el efecto que producen variaciones del grado de verdad parcial de los antecedentes en el grado de verdad o grado de certidumbre del consecuente. La regla en cuestión es:

regla *G1**si**CH₄/H₂ medio**C₂H₄/C₂H₆ bajo**C₂H₂/C₂H₄ bajo**entonces**sin fallo*

Podría pensarse que cualquier cambio en los estados de los antecedentes implica un cambio drástico de diagnóstico (cierto al menos a nivel lingüístico), pasando de la situación normal a la existencia de algún tipo de fallo. Con este supuesto la relación entre cada uno de los antecedentes y el diagnóstico debería ser una relación de certidumbre. Sin embargo esta suposición puede llevar a pensar que pequeños cambios de cualquiera de los estados de los antecedentes pueden dar lugar a cualquier tipo de diagnóstico, lo que no parece del todo acertado.

Si se analiza la influencia de cada uno de los antecedentes puede comprobarse cómo para algunos de ellos se produce una transición gradual de la situación de normalidad a las situaciones de fallo, pasando primero por situaciones en las que estos fallos pueden considerarse leves y muy próximos a la condición de normalidad. Por otro lado puede considerarse que cualificar el diagnóstico *sin fallo* mediante un grado de verdad parcial equivale a admitir la posibilidad de que existan fallos leves, lo que incita a considerar como primera opción el empleo de modelos graduales.

Para el primer antecedente el paso de *medio* a *bajo* o a *alto* da lugar a las reglas *G2* o *G7*:

<p>regla G2</p> <p><i>si</i></p> <p style="padding-left: 40px;">CH_4/H_2 <i>bajo</i></p> <p style="padding-left: 40px;">C_2H_4/C_2H_6 <i>bajo</i></p> <p style="padding-left: 40px;">C_2H_2/C_2H_4 <i>bajo</i></p> <p><i>entonces</i></p> <p><i>descargas parciales de baja energía</i></p>	<p>regla G7</p> <p><i>si</i></p> <p style="padding-left: 40px;">CH_4/H_2 <i>alto</i></p> <p style="padding-left: 40px;">C_2H_4/C_2H_6 <i>bajo</i></p> <p style="padding-left: 40px;">C_2H_2/C_2H_4 <i>bajo</i></p> <p><i>entonces</i></p> <p><i>defecto térmico de baja-media temperatura</i></p>
--	--

Ambas reglas concluyen diagnósticos de poca gravedad (*descargas parciales de baja energía* o *defecto térmico de baja-media temperatura*), y por tanto próximos a la situación de normalidad, por lo que podría admitirse la existencia de una relación gradual. Podría tal vez pensarse en una relación semántica mixta si se tiene en cuenta que el diagnóstico de la regla *G7* no es el defecto térmico de menor gravedad, ya que existe también el defecto térmico de baja temperatura, aunque este matiz parece poco significativo.

El disparo simultáneo de la regla *G1* con cualquiera de las reglas *G2* o *G7* debe precisar el significado del diagnóstico *sin fallo* cualificado con un grado de verdad parcial. Esto se debe a que en el sistema de identificación de anomalías propuesto, el uso combinado de varias reglas de conocimiento con modelos apropiadamente escogidos, es de especial importancia de cara a determinar con precisión el diagnóstico más acertado, especialmente

en las zonas de transición de unos estados a otros, que se corresponden con transiciones de unas reglas a otras.

Para el segundo antecedente el paso de *bajo* a *medio* da lugar a la regla *G6*:

regla G6

si

CH_4/H_2 *medio*

C_2H_4/C_2H_6 *medio*

C_2H_2/C_2H_4 *bajo*

entonces

defecto térmico de baja temperatura

El diagnóstico de esta regla es también de escasa gravedad, incluso menor que el que proporcionaba la regla *G7*. Por tanto parece clara la existencia de una relación gradual entre el segundo antecedente y el diagnóstico en la regla *G1*.

Para el tercer antecedente el paso de *bajo* a *medio* no da lugar a ninguna regla conocida. Esto puede significar o bien un agujero en el conocimiento disponible, y por tanto la necesidad de considerar la existencia de una nueva regla, cuyos antecedentes vendrían dados por

Nueva regla

si

CH_4/H_2 *medio*

C_2H_4/C_2H_6 *bajo*

C_2H_2/C_2H_4 *medio*

entonces

diagnóstico

o bien la imposibilidad de que dicha situación se dé en la práctica. Esto último puede deberse a consideraciones físicas que conlleven dependencias entre las medidas involucradas en la regla. Este último caso se corresponde con la suposición de que el sistema de reglas es completo, y por tanto contempla todas las posibles combinaciones de valores de las variables.

En el supuesto de que el sistema sea completo, se propone buscar las reglas más próximas dentro de la base de conocimiento disponible. En este caso se obtienen las reglas *G3* y *G4*, pero permitiendo que varíe otro de los antecedentes de la regla, en concreto el primero para obtener *G3* y el segundo para obtener *G4*:

<p>regla G3</p> <p>si</p> <table> <tr> <td>CH_4/H_2</td> <td>bajo</td> </tr> <tr> <td>C_2H_4/C_2H_6</td> <td>bajo</td> </tr> <tr> <td>C_2H_2/C_2H_4</td> <td>medio</td> </tr> </table> <p>entonces</p> <p>descargas parciales de alta energía</p>	CH_4/H_2	bajo	C_2H_4/C_2H_6	bajo	C_2H_2/C_2H_4	medio	<p>regla G4</p> <p>si</p> <table> <tr> <td>CH_4/H_2</td> <td>medio</td> </tr> <tr> <td>C_2H_4/C_2H_6</td> <td>medio</td> </tr> <tr> <td>C_2H_2/C_2H_4</td> <td>medio</td> </tr> </table> <p>entonces</p> <p>descargas eléctricas de baja energía</p>	CH_4/H_2	medio	C_2H_4/C_2H_6	medio	C_2H_2/C_2H_4	medio
CH_4/H_2	bajo												
C_2H_4/C_2H_6	bajo												
C_2H_2/C_2H_4	medio												
CH_4/H_2	medio												
C_2H_4/C_2H_6	medio												
C_2H_2/C_2H_4	medio												

Los diagnósticos obtenidos son significativamente diferentes de la situación de normalidad lo que apunta a la existencia de una relación de certidumbre entre el diagnóstico y el tercero de los antecedentes de la regla.

La conclusión de la regla puede, de acuerdo con las relaciones semánticas identificadas, obtenerse en el dominio de lo necesario a partir de la implicación general de necesidad $I_n(\alpha_n, \beta_n, \gamma)$ de forma que el valor α_n se determina agregando conjuntamente los grados de verdad de los antecedentes con relaciones graduales o mixtas, mientras que β_n se determina agregando conjuntamente grados de certidumbre de los modelos graduales, mixtos y de certidumbre, junto con los grados de verdad de los modelos mixtos y de certidumbre, según se describe en el apartado 7.5.4.2, y como se ilustra también al final de este ejemplo.

Formalmente la regla *G1* se puede formular del siguiente modo:

REGLA *G1*
NOMBRE *sin fallo*
GRUPO *norma_cei*
Si

CH_4/H_2	<i>en aceite es</i>	<i>medio</i>
C_2H_4/C_2H_6	<i>en aceite es</i>	<i>bajo</i>
C_2H_2/C_2H_4	<i>en aceite es</i>	<i>bajo</i>

entonces

diagnósticos:

sin fallo en transformador (N,G,G,C,1,1)

COMENTARIOS:
El descriptor (N,G,G,C,1,1) indica que la relación semántica del diagnóstico con el primer y segundo antecedente es gradual, y con el tercero es de certidumbre. Además cuando los tres antecedentes tengan grados de verdad y de certidumbre unitarios, también los tendrá el diagnóstico.

El proceso llevado a cabo para determinar el modelo correcto para la regla evidencia la complejidad que puede llegar a suponer, y sugiere la conveniencia de desarrollar alguna herramienta automática para la adquisición del conocimiento (uno de los grandes cuellos de botella de los sistemas basados en conocimiento) y para el establecimiento de modelos de regla apropiados, detección de inconsistencias, etc. Esta herramienta permitiría por un lado determinar los modelos de las reglas, y por otro detectar posibles inconsistencias o lagunas en el conocimiento. Nótese a este respecto que en general, dado que los procesos físicos suelen verificar ciertas propiedades de continuidad, un conocimiento suficiente amplio y detallado debería dar lugar principalmente a modelos con semántica gradual que además son los más informativos.

- Ejemplo regla G2:

La regla G2 viene dada por:

regla G2

si

CH_4/H_2 *bajo*

C_2H_4/C_2H_6 *bajo*

C_2H_2/C_2H_4 *bajo*

entonces

descargas parciales de baja energía

Analizando la relación del diagnóstico con el primero de los antecedentes vemos que cuando su estado pasa de *bajo* a *medio* obtenemos la regla G1 ya analizada en el apartado anterior, pudiendo establecerse una relación gradual entre este antecedente y el diagnóstico de la regla.

Cuando variamos el estado del segundo antecedente no se dispone de ninguna regla que permita determinar de forma directa la relación existente. La regla más próxima es la regla G6 que viene dada por:

regla G6

si

CH_4/H_2 *medio*

C_2H_4/C_2H_6 *medio*

C_2H_2/C_2H_4 *bajo*

entonces

defecto térmico de baja temperatura

Aunque el diagnóstico de $G2$ y el de $G6$ pueden considerarse defectos leves, se deben a causas diferentes por lo que parece razonable suponer la existencia de una relación de certidumbre entre el segundo de los antecedentes y el diagnóstico. Esto significa que a medida que el grado de verdad parcial del segundo antecedente se reduce cualquier otro diagnóstico aparece como posible explicación a la supuesta anomalía.

Variando el estado del tercer antecedente obtenemos la regla $G3$:

regla G3

si

CH_4/H_2	bajo
C_2H_4/C_2H_6	bajo
C_2H_2/C_2H_4	medio

entonces

descargas parciales de alta energía

donde el diagnóstico es similar pero correspondiente a un defecto de alta en lugar de baja energía, lo que apunta a la existencia de una relación gradual. Con estos supuestos la regla $G2$ queda finalmente:

REGLA G2

NOMBRE *descargas parciales de baja energía*

GRUPO *norma_cei*

Si

CH_4/H_2	en aceite es	bajo
C_2H_4/C_2H_6	en aceite es	bajo
C_2H_2/C_2H_4	en aceite es	bajo

entonces

diagnósticos:

descargas parciales de baja energía en transformador (N,G,C,G,I,I)

COMENTARIOS:

El descriptor (N,G,C,G,I,I) indica que la relación semántica del diagnóstico con el primer y tercer antecedente es gradual, mientras que con el segundo es de certidumbre. Además cuando los tres antecedentes tengan grados de verdad y de certidumbre unitarios, también los tendrá el diagnóstico.

- Ejemplo regla G4:

La regla $G4$ viene dada por:

regla G4*si* CH_4/H_2 *medio* C_2H_4/C_2H_6 *medio* C_2H_2/C_2H_4 *medio**entonces**descargas eléctricas de baja energía*

Esta regla es especialmente delicada en el sentido en que los estados de los tres antecedentes pueden variar tanto para aumentar hacia el estado *alto* como para disminuir hacia el estado *bajo*, dando lugar a un mayor número de combinaciones.

La variación del estado del primer antecedente a *bajo* o *medio* da como reglas más próximas G8 y G3 (aunque variando también los estados de otros antecedentes):

regla G3 <i>si</i> CH_4/H_2 <i>bajo</i> C_2H_4/C_2H_6 <i>bajo</i> C_2H_2/C_2H_4 <i>medio</i> <i>entonces</i> <i>descargas parciales de alta energía</i>	regla G8 <i>si</i> CH_4/H_2 <i>alto</i> C_2H_4/C_2H_6 <i>medio</i> C_2H_2/C_2H_4 <i>bajo</i> <i>entonces</i> <i>defecto térmico de media temperatura</i>
--	---

Los diagnósticos que se obtienen son significativamente distintos por lo que la relación semántica parece ser de certidumbre. Nótese que éste es otro de los casos en los que cabría plantearse, o bien la existencia de una clara dependencia entre las medidas de los antecedentes, o bien la existencia de una posible laguna en el conocimiento, ya que no están contempladas en la base de conocimiento ninguna de las siguientes situaciones:

CH_4/H_2 <i>bajo</i> C_2H_4/C_2H_6 <i>medio</i> C_2H_2/C_2H_4 <i>medio</i>	CH_4/H_2 <i>alto</i> C_2H_4/C_2H_6 <i>medio</i> C_2H_2/C_2H_4 <i>medio</i>
--	--

En cuanto al segundo antecedente su variación de *medio* a *bajo* y de *medio* a *alto* da lugar a las reglas G1 y G3 por un lado (aunque variando el estado de otro de los antecedentes) y a la regla G5 por otro:

<p>regla G1 <i>si</i></p> <table style="margin-left: 40px;"> <tr><td>CH_4/H_2</td><td><i>medio</i></td></tr> <tr><td>C_2H_4/C_2H_6</td><td><i>bajo</i></td></tr> <tr><td>C_2H_2/C_2H_4</td><td><i>bajo</i></td></tr> </table> <p><i>entonces</i> <i>sin fallo</i></p>	CH_4/H_2	<i>medio</i>	C_2H_4/C_2H_6	<i>bajo</i>	C_2H_2/C_2H_4	<i>bajo</i>	<p>regla G3 <i>si</i></p> <table style="margin-left: 40px;"> <tr><td>CH_4/H_2</td><td><i>bajo</i></td></tr> <tr><td>C_2H_4/C_2H_6</td><td><i>bajo</i></td></tr> <tr><td>C_2H_2/C_2H_4</td><td><i>medio</i></td></tr> </table> <p><i>entonces</i> <i>descargas parciales de alta energía</i></p>	CH_4/H_2	<i>bajo</i>	C_2H_4/C_2H_6	<i>bajo</i>	C_2H_2/C_2H_4	<i>medio</i>
CH_4/H_2	<i>medio</i>												
C_2H_4/C_2H_6	<i>bajo</i>												
C_2H_2/C_2H_4	<i>bajo</i>												
CH_4/H_2	<i>bajo</i>												
C_2H_4/C_2H_6	<i>bajo</i>												
C_2H_2/C_2H_4	<i>medio</i>												
<p>regla G5 <i>si</i></p> <table style="margin-left: 40px;"> <tr><td>CH_4/H_2</td><td><i>medio</i></td></tr> <tr><td>C_2H_4/C_2H_6</td><td><i>alto</i></td></tr> <tr><td>C_2H_2/C_2H_4</td><td><i>medio</i></td></tr> </table> <p><i>entonces</i> <i>descargas eléctricas de alta energía</i></p>		CH_4/H_2	<i>medio</i>	C_2H_4/C_2H_6	<i>alto</i>	C_2H_2/C_2H_4	<i>medio</i>						
CH_4/H_2	<i>medio</i>												
C_2H_4/C_2H_6	<i>alto</i>												
C_2H_2/C_2H_4	<i>medio</i>												

En este caso pasamos de *descargas eléctricas de baja energía* a *descargas eléctricas de alta energía* por un lado, y a situación de normalidad o *descargas parciales de alta energía* por otro. Se detecta una cierta semántica gradual aunque desvirtuada por la posibilidad de concluir *descargas parciales* en uno de los casos, por lo que una relación mixta podría ser la apropiada.

Por último para el tercero de los antecedentes sólo se obtiene la regla *G6*:

regla G6
si

CH_4/H_2	<i>medio</i>
C_2H_4/C_2H_6	<i>medio</i>
C_2H_2/C_2H_4	<i>bajo</i>

entonces
defecto térmico de baja temperatura

con un diagnóstico diferente por lo que la relación más apropiada es una relación de certidumbre. Con estos supuestos la regla *G4* queda como sigue:

REGLA G4
NOMBRE descargas eléctricas de baja energía
GRUPO norma_cei

Si

CH_4/H_2 en aceite es medio

C_2H_4/C_2H_6 en aceite es medio

C_2H_2/C_2H_4 en aceite es medio

entonces

diagnósticos:

descargas eléctricas de baja energía en transformador (N,C,M,C,1,1)

COMENTARIOS:

El descriptor (N,C,M,C,1,1) indica que la relación semántica del diagnóstico con el primer y tercer antecedente es de certidumbre mientras que con el segundo puede considerarse mixta. Además cuando los tres antecedentes tengan grados de verdad y de certidumbre unitarios, también los tendrá el diagnóstico.

- Ejemplo regla G5:

La regla G5 viene dada por:

regla G5

si

CH_4/H_2 medio

C_2H_4/C_2H_6 alto

C_2H_2/C_2H_4 medio

entonces

descargas eléctricas de alta energía

Si el valor del estado de la primera medida varía de *medio* a *bajo* no es posible encontrar, dentro del conjunto de reglas enunciadas, una regla suficientemente próxima, aún permitiendo que los estados de los otros antecedentes a su vez varíen. Esto pone de manifiesto la existencia de una relación de certidumbre puesto que una variación del grado de verdad conduce a una situación de total indeterminación. Cuando el valor del estado aumenta de *medio* a *alto* la regla más próxima que puede hallarse es la regla G9, aunque modificando también el estado del tercer antecedente:

regla G9*si* CH_4/H_2 *alto* C_2H_4/C_2H_6 *alto* C_2H_2/C_2H_4 *bajo**entonces**defecto térmico de alta temperatura*

Dado que se obtiene un diagnóstico significativamente diferente del de la regla G5 se confirma la relación de certidumbre.

Para el segundo de los antecedentes solo cabe la variación de *alto* a *medio*. En este caso la regla más próxima es la regla G4 que concluye sobre el mismo tipo de defecto, aunque de baja en lugar de alta energía:

regla G4*si* CH_4/H_2 *medio* C_2H_4/C_2H_6 *medio* C_2H_2/C_2H_4 *medio**entonces**descargas eléctricas de baja energía*

Por tanto en este caso la relación puede considerarse gradual.

Para el tercer antecedente se puede analizar el comportamiento de la conclusión cuando varía su estado de *medio* a *bajo* o a *alto*. Las reglas más próximas son la G6 y la G9 y ambas concluyen sobre un diagnóstico diferente, por lo que para este tercer antecedente la relación con el diagnóstico es una relación de certidumbre.

Por tanto la regla G5 queda formalmente del siguiente modo:

REGLA G5
NOMBRE descargas electricas de alta energía
GRUPO norma_cei

Si

CH_4/H_2 en aceite es medio

C_2H_4/C_2H_6 en aceite es alto

C_2H_2/C_2H_4 en aceite es medio

entonces

diagnósticos:

descargas eléctricas de alta energía en transformador (N,C,G,C,1,1)

COMENTARIOS:

El descriptor (N,C,G,C,1,1) indica que la relación semántica del diagnóstico con el primer y tercer antecedente es de certidumbre mientras que con el segundo es gradual. Además cuando los tres antecedentes tengan grados de verdad y de certidumbre unitarios, también los tendrá el diagnóstico.

- Ejemplo regla G6:

La regla G6 viene dada por:

regla G6

si

CH_4/H_2 medio

C_2H_4/C_2H_6 medio

C_2H_2/C_2H_4 bajo

entonces

defecto térmico de baja temperatura

La variación del estado del primer antecedente de *medio* a *alto* da lugar a la regla G8:

regla G8

si

CH_4/H_2 alto

C_2H_4/C_2H_6 medio

C_2H_2/C_2H_4 bajo

entonces

defecto térmico de media temperatura

La proximidad del diagnóstico de esta regla con la regla *G6* apunta a la existencia de una relación gradual. Sin embargo si el estado del antecedente varía de *medio* a *bajo* se obtiene como regla más próxima la regla *G2*:

regla G2

si

CH_4/H_2 *bajo*

C_2H_4/C_2H_6 *bajo*

C_2H_2/C_2H_4 *bajo*

entonces

descargas parciales de baja energía

siendo su diagnóstico significativamente distinto, aunque también de escasa gravedad. Por tanto la relación sería en este caso de certidumbre o mixta.

Para el segundo de los antecedentes su variación a *bajo* da lugar a la regla *G1* que aunque tiene un diagnóstico diferente, *sin fallo*, es próximo al de la regla *G6*, lo que apuntaría a una relación gradual. Si por el contrario pasa a *alto* las reglas más próximas son la *G5* y la *G9*.

<p>regla G5</p> <p>si</p> <p>CH_4/H_2 <i>medio</i></p> <p>C_2H_4/C_2H_6 <i>alto</i></p> <p>C_2H_2/C_2H_4 <i>medio</i></p> <p>entonces</p> <p><i>descargas eléctricas de alta energía</i></p>	<p>regla G9</p> <p>si</p> <p>CH_4/H_2 <i>alto</i></p> <p>C_2H_4/C_2H_6 <i>alto</i></p> <p>C_2H_2/C_2H_4 <i>bajo</i></p> <p>entonces</p> <p><i>defecto térmico de alta temperatura</i></p>
---	--

Mientras que la regla *G9* propone un diagnóstico similar, la regla *G5* propone un diagnóstico diferente. Parece razonable en este caso optar por una relación mixta.

Por último, para el tercero de los antecedentes la variación de *bajo* a *medio* lleva a la regla *G4*:

regla G4*si* CH_4/H_2 *medio* C_2H_4/C_2H_6 *medio* C_2H_2/C_2H_4 *medio**entonces**descargas eléctricas de baja energía*

y dado que el diagnóstico es diferentes podemos considerar que la relación es en este caso de certidumbre.

Por tanto la regla G6 queda finalmente como:

REGLA *G6***NOMBRE** *defecto térmico de baja temperatura***GRUPO** *norma_cei**Si* CH_4/H_2 *en aceite es* *medio* C_2H_4/C_2H_6 *en aceite es* *medio* C_2H_2/C_2H_4 *en aceite es* *bajo**entonces**diagnósticos:**defecto térmico de baja temperatura en transformador(N,M,M,C,1,1)***COMENTARIOS:**

El descriptor (N,M,G,C,1,1) indica que la relación semántica del diagnóstico con dos primeros antecedentes es mixta, mientras que con el tercero es de certidumbre. Además cuando los tres antecedentes tengan grados de verdad y de certidumbre unitarios, también los tendrá el diagnóstico.

- Ejemplo regla G8:

La regla G8 viene dada por:

regla G8

si

CH_4/H_2	alto
C_2H_4/C_2H_6	medio
C_2H_2/C_2H_4	bajo

entonces

defecto térmico de media temperatura

La regla obtenida al variar el estado del primer antecedente es la regla G6 dada por:

regla G6

si

CH_4/H_2	medio
C_2H_4/C_2H_6	medio
C_2H_2/C_2H_4	bajo

entonces

defecto térmico de baja temperatura

El nuevo diagnóstico, *defecto térmico de baja temperatura*, es compatible con el de la regla original, aunque existe otro diagnóstico intermedio, correspondiente con el defecto de *baja-media temperatura*. Por ello parece razonable optar por una relación o bien gradual o bien mixta, como ya se ha visto en ejemplos anteriores.

Si analizamos el comportamiento con el segundo antecedente, tenemos que considerar o bien un decremento del valor de su estado o bien un incremento del mismo, Para el primero de los casos la regla más próxima es:

regla G7

si

CH_4/H_2	alto
C_2H_4/C_2H_6	bajo
C_2H_2/C_2H_4	bajo

entonces

defecto térmico de baja-media temperatura

con lo que el diagnóstico ha variado de forma gradual. Cuando el estado de la medida C_2H_4/C_2H_6 aumenta, la regla más próxima de la base de conocimiento es la regla G9:

regla G9*si*

CH_4/H_2	<i>alto</i>
C_2H_4/C_2H_6	<i>alto</i>
C_2H_2/C_2H_4	<i>bajo</i>

*entonces**defecto térmico de alta temperatura*

y podemos verificar que se mantiene claramente la relación gradual.

Para el tercer antecedente solo cabe la posibilidad de que aumente, por lo que la regla más próxima debería tener como antecedentes:

si

CH_4/H_2	<i>alto</i>
C_2H_4/C_2H_6	<i>medio</i>
C_2H_2/C_2H_4	<i>medio</i>

Sin embargo no existe en la base de conocimiento ninguna regla con los mismos antecedentes. Buscando nuevamente la regla más próxima se obtiene la regla G4:

regla G4*si*

CH_4/H_2	<i>medio</i>
C_2H_4/C_2H_6	<i>medio</i>
C_2H_2/C_2H_4	<i>medio</i>

*entonces**descargas eléctricas de baja energía*

donde como en otros ejemplos ha sido necesario modificar el estado de otro de los antecedentes, en este caso el primero. El diagnóstico de esta nueva regla ha variado por completo con respecto a la regla original, pasando de *defecto térmico* a *descargas eléctricas*. Por ello parece clara la existencia de una relación de certidumbre.

Formalmente la regla G8 puede expresarse como sigue:

REGLA G8
NOMBRE defecto térmico de media temperatura
GRUPO norma_cei

Si
 CH_4/H_2 en aceite es alto
 C_2H_4/C_2H_6 en aceite es medio
 C_2H_2/C_2H_4 en aceite es bajo

entonces
 diagnósticos:
 defecto térmico de media temperatura en
 transformador(N,G,G,C,1,1)

COMENTARIOS:
 El descriptor (N,M,G,C,1,1) indica que la relación semántica del diagnóstico con los dos primeros antecedentes es gradual, mientras que con el tercero es de certidumbre. Además cuando los tres antecedentes tengan grados de verdad y de certidumbre unitarios, también los tendrá el diagnóstico.

- Ejemplo regla G9:

La regla G9 viene dada por:

regla G9
 si
 CH_4/H_2 alto
 C_2H_4/C_2H_6 alto
 C_2H_2/C_2H_4 bajo
 entonces
 defecto térmico de alta temperatura

En este caso una variación del estado del primer antecedente conduce a las reglas G5 y G6:

<p>regla G5 si CH_4/H_2 medio C_2H_4/C_2H_6 alto C_2H_2/C_2H_4 medio entonces descargas eléctricas de alta energía</p>	<p>regla G6 si CH_4/H_2 medio C_2H_4/C_2H_6 medio C_2H_2/C_2H_4 bajo entonces defecto térmico de baja temperatura</p>
---	--

Aunque uno de los diagnósticos es similar al de la regla original el otro es significativamente distinto por lo que la relación sería en este caso una relación mixta.

Parece especialmente notable la ausencia de una regla que cubriese el hueco existente entre las reglas *G9* y *G6*, como podría ser por ejemplo:

regla G9-6

si

CH_4/H_2 *medio*

C_2H_4/C_2H_6 *alto*

C_2H_2/C_2H_4 *bajo*

entonces

defecto térmico de media temperatura

o bien, utilizando el estado *media-baja* en lugar de *media*:

regla G9-6

si

CH_4/H_2 *medio*

C_2H_4/C_2H_6 *alto*

C_2H_2/C_2H_4 *bajo*

entonces

defecto térmico de media-baja temperatura

Este es un caso claro en el que se puede sospechar la existencia de una laguna en el conocimiento, y donde sería interesante confirmarla o rechazarla mediante una revisión de la posible documentación disponible y a través de entrevistas con expertos.

Para el segundo de los antecedentes la variación de su estado conduce a la regla *G8*:

regla G8

si

CH_4/H_2 *alto*

C_2H_4/C_2H_6 *medio*

C_2H_2/C_2H_4 *bajo*

entonces

defecto térmico de media temperatura

por lo que la relación entre el segundo antecedente y el diagnóstico puede considerarse gradual.

Para el tercer antecedente la regla más próxima es la regla G5:

regla G5

si

CH_4/H_2 medio

C_2H_4/C_2H_6 alto

C_2H_2/C_2H_4 medio

entonces

descargas eléctricas de alta energía

en cuyo caso la relación es una relación de certidumbre.

Finalmente la regla G9 queda de la forma:

REGLA G9

NOMBRE defecto térmico de alta temperatura

GRUPO norma_cei

Si

CH_4/H_2 en aceite es alto

C_2H_4/C_2H_6 en aceite es alto

C_2H_2/C_2H_4 en aceite es bajo

entonces

diagnósticos:

defecto térmico de media temperatura en transformador(N,M,G,C,1,1)

COMENTARIOS:

El descriptor (N,G,G,C,1,1) indica que la relación semántica del diagnóstico con el primer antecedente es mixta, con el segundo gradual y con el tercero de certidumbre. Además cuando los tres antecedentes tengan grados de verdad y de certidumbre unitarios, también los tendrá el diagnóstico.

- Ejemplo reglas G2 y G3:

Estas reglas vienen dadas por:

<p>regla G2</p> <p>si</p> <p>CH_4/H_2 bajo</p> <p>C_2H_4/C_2H_6 bajo</p> <p>C_2H_2/C_2H_4 bajo</p>	<p>regla G3</p> <p>si</p> <p>CH_4/H_2 bajo</p> <p>C_2H_4/C_2H_6 bajo</p> <p>C_2H_2/C_2H_4 medio</p>
--	---

<i>entonces</i> <i>descargas parciales de baja energía</i>	<i>entonces</i> <i>descargas parciales de alta energía</i>
---	---

Supongamos que fuese imposible disponer de una medida de la concentración C_2H_2 . Esto significaría que los grados de verdad y de certidumbre del tercer antecedente, calculados en el dominio de lo necesario, valdrían siempre 0, y no sería posible activar estas reglas con los modelos de necesidad antes obtenidos. En efecto la concentración desconocida de C_2H_2 hace que el cociente C_2H_2/C_2H_4 deba modelarse con una distribución igual al universo de discurso. Sin embargo recordemos que el razonamiento en el dominio de lo posible sí permitiría activar estas reglas, en cuyo caso sí podría concluirse que, cuando CH_4/H_2 y C_2H_4/C_2H_6 sean ambos *bajo*, es posible que existan descargas parciales de baja o alta energía.

Otra posible alternativa, aunque supondría una modificación de la base de conocimiento, sería utilizar una única regla de necesidad con dos diagnósticos, puesto que sabemos que con los valores de medidas señalados, se tienen descargas parciales, ya sean de baja o alta energía. Así se podría formular la regla de necesidad dada por:

regla G2-3

si

CH_4/H_2 *bajo*

C_2H_4/C_2H_6 *bajo*

entonces

descargas parciales de baja energía

descargas parciales de alta energía

Esta regla permite ilustrar que la determinación del modelo asociado a cada diagnóstico requiere no solo la relación semántica con cada uno de los antecedentes, sino también, y por comparación con el otro diagnóstico, la determinación de su certidumbre para un total cumplimiento de los antecedentes.

Una variación del estado del primer antecedente daría lugar a la regla *G1*, supuesto que prescindimos del tercero de los antecedentes. Dado que el primero de los diagnósticos puede considerarse un defecto leve próximo a la situación normal, parece razonable suponer la existencia de una relación gradual. Por otro lado, la existencia de varios diagnósticos simultáneos evidencia un cierto grado de incertidumbre que puede ser modelado con un modelo mixto. Para el segundo de los diagnósticos la relación con el primer antecedente es de certidumbre al ser significativamente distinto del diagnóstico de *G1*.

Una variación del estado del segundo antecedente da lugar a una combinación de antecedentes no presente en ninguna de las reglas de la base de conocimiento. Las reglas más próximas se obtienen variando también el primero son las reglas *G4* y *G6*:

<p>regla G4</p> <p>si</p> <p style="padding-left: 40px;">CH_4/H_2 medio</p> <p style="padding-left: 40px;">C_2H_4/C_2H_6 medio</p> <p style="padding-left: 40px;">C_2H_2/C_2H_4 medio</p> <p>entonces</p> <p>descargas eléctricas de baja energía</p>	<p>regla G6</p> <p>si</p> <p style="padding-left: 40px;">CH_4/H_2 medio</p> <p style="padding-left: 40px;">C_2H_4/C_2H_6 medio</p> <p style="padding-left: 40px;">C_2H_2/C_2H_4 bajo</p> <p>entonces</p> <p>defecto térmico de baja temperatura</p>
--	--

En este caso la diversidad de diagnósticos obtenidos parece apuntar claramente a relaciones de certidumbre entre los dos diagnósticos y el segundo antecedente.

Por último es necesario determinar la certidumbre de cada uno de los diagnósticos cuando los grados de verdad y de certidumbre de todos los antecedentes son 1. Los dos diagnósticos no pueden tener certidumbre unitaria puesto que sería como suponer que ambos se dan simultáneamente, cuando son en cierto modo excluyentes. En este caso es necesario recurrir al juicio subjetivo de uno o más expertos que expresen la credibilidad que cada diagnóstico les merece cuando se dan los antecedentes de la regla. A falta de información más precisa supondremos que el defecto de menor energía es más creíble que el de mayor energía. Nótese que en estas situaciones, a falta de mejores datos para poder discriminar entre los diagnósticos disponibles, el grado de certidumbre que se asigne se basará en general en consideraciones probabilísticas subjetivas posiblemente influenciadas por casos particulares, en tanto en cuanto será más creíble aquel diagnóstico que se haya dado más veces.

Con estos supuestos la regla *G2-3* quedaría como sigue:

REGLA G2-3
NOMBRE descargas parciales
GRUPO norma_cei
 Si
 CH_4/H_2 en aceite es alto
 C_2H_4/C_2H_6 en aceite es alto
 entonces
 diagnósticos:
 descargas parciales de baja energía en transformador(N,M,C,1,0.8)
 descargas parciales de alta energía en transformador(N,C,C,1,0.6)

- Ejemplo reglas PG1 y PG2:

Las reglas PG1 y PG2 son reglas de incremento de certeza para las que solamente hemos considerado un único tipo de modelo. Formalmente ambas reglas pueden expresarse como sigue:

REGLA PG1
NOMBRE producción de gases y tiempo de servicio
GRUPO norma_cei
 Si
 tendencia(gas) alta
 correlación(gas,tiempo_servicio) alta
 entonces
 diagnósticos:
 descargas parciales (N,I,0.6)

REGLA PG2
NOMBRE producción de gases y tiempo de servicio
GRUPO norma_cei
 Si
 tendencia(gas) alta
 correlación(gas,perdidas_ohmicas) alta
 entonces
 diagnósticos:
 defecto térmico (N,I,0.6)

El funcionamiento de estas reglas debe ser el siguiente. Una vez se hayan disparado todas las reglas que hacen referencia al diagnóstico *descargas parciales* se dispara la regla PG1, de forma que se incrementa el grado de certidumbre de los diagnósticos *baja*

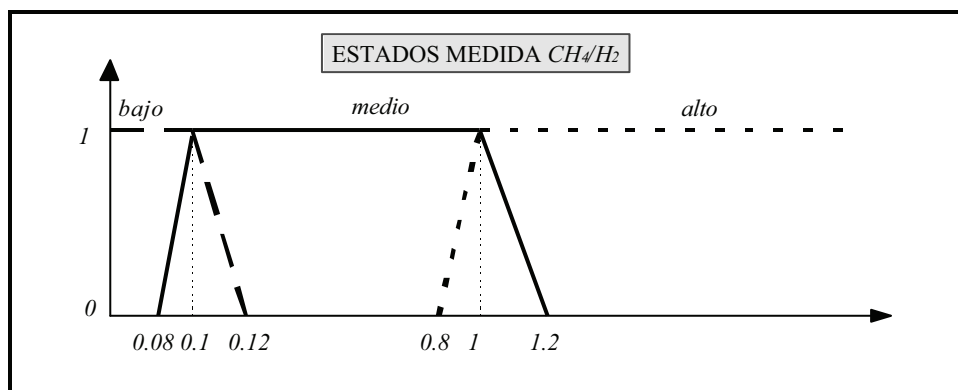
energía y *alta energía*, siempre que esta no sea nula. El funcionamiento de la segunda regla es similar, incrementando el grado de certidumbre de todos los tipos de defectos térmicos que hayan sido identificados con certidumbre no nula. En el caso en que los estados de este fallo se modelen con conjuntos borrosos sobre el universo de temperaturas, el incremento de certidumbre se aplica al conjunto borroso final resultante de la agregación las conclusiones de todas las reglas paralelas que concluyan sobre este tipo de fallo.

Es evidente que para que estas reglas puedan dispararse es necesario haber definido previamente lo que se entiende por tendencia y correlación, así como las medidas *tiempo_servicio* y *perdidas-ohmicas*.

7.6.5 Inferencia

- Modelo de las observaciones:

Aunque las medidas que realmente se toman son las concentraciones de los gases, y a partir de ellas se calculan los cocientes utilizados para el diagnóstico, por simplicidad vamos a suponer que las medidas que llegan al sistema son directamente dichos cocientes (que son también suministrados directamente por los laboratorios que normalmente realizan este tipo de análisis). Como ya se ha comentado las particiones escogidas para las medidas del sistema se corresponden con una visión logarítmica del universo de discurso, en el sentido en que la vaguedad del concepto aumenta con la magnitud de los valores que considera. En efecto recordemos por ejemplo que la partición de la medida CH_4/H_2 es:

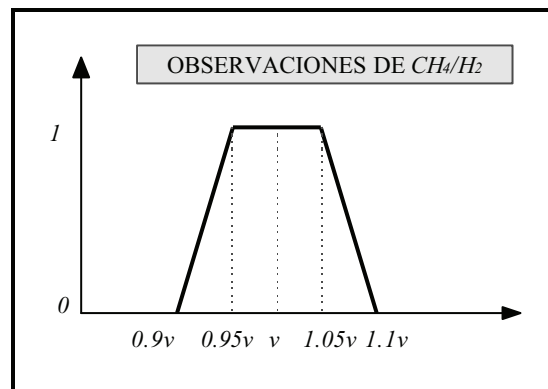


donde la zona de transición de la pertenencia a la no pertenencia depende del orden de magnitud de los elementos involucrados.

La completitud recomienda en este caso que el núcleo de la observación no sea mayor que $0.12-0.08$ en un caso y que $1.2-0.8$ en otro (véase el apartado 5.6.3.2). Podemos ver que esto es equivalente a decir que el núcleo de la observación debe ser menor que el 40% de su valor central, o lo que es lo mismo, que con una imprecisión en la medida menor del $\pm 20\%$ el sistema puede funcionar correctamente en cuanto a completitud se refiere. Esto significa que el error se puede también modelar como un error relativo a la magnitud medida, lo que es acorde con la visión logarítmica que se desprende de la partición empleada. Dado que el valor del $\pm 20\%$ es un valor límite posiblemente elevado respecto a la precisión real de las medidas, escogeremos un valor menor de por ejemplo $\pm 5\%$.

En cuanto a las transiciones de la pertenencia a la no pertenencia de las observaciones, parece también razonable utilizar un ancho proporcional a la magnitud de la observación. La determinación de este ancho es un tanto arbitraria y dependiente del procedimiento empleado para medir, del tipo de sensor, etc. Nosotros utilizaremos un ancho del 5%. Esto significa que si la observación toma el valor v estamos seguros de que cualquier valor que no pertenezca al intervalo $v \pm 10\%$ es un valor totalmente imposible para la medida en cuestión.

Con estas consideraciones el modelo de las observaciones para la primera medida es tal y como aparece en la siguiente figura:



Para la segunda y tercera medidas y en base a las mismas consideraciones utilizaremos también el mismo modelo.

A continuación ilustraremos el procedimiento propuesto para inferir en el apartado 7.5.4.3 con algunos ejemplos concretos.

- La inferencia con la MS de dos índices

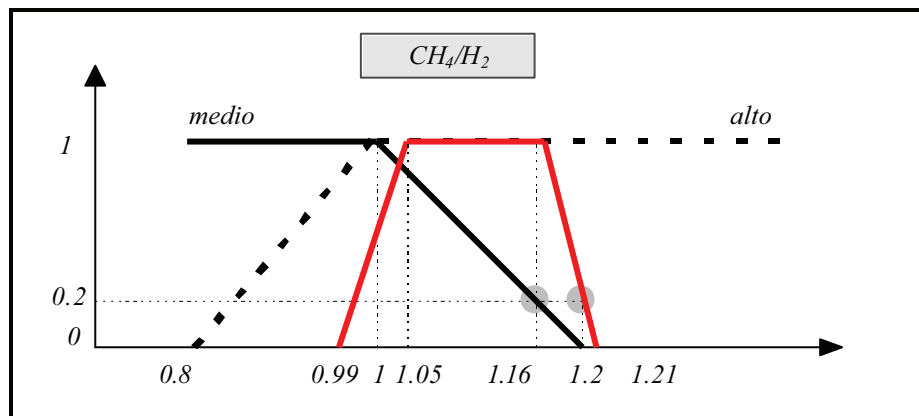
Supongamos que disponemos de los siguientes valores medidos:

$$\begin{array}{ll} CH_4/H_2 & 1.1 \\ C_2H_4/C_2H_6 & 3.2 \\ C_2H_2/C_2H_4 & 0.085 \end{array}$$

Con el modelo de observación propuesto dado por $(0.9v, 0.95v, 1.05v, 1.1v)$ donde v es el valor medido, los conjuntos borrosos que resultan, para cada una de las medidas anteriores, son:

$$\begin{array}{ll} CH_4/H_2 & (0.99, 1.05, 1.16, 1.21) \\ C_2H_4/C_2H_6 & (2.88, 3.04, 3.36, 3.52) \\ C_2H_2/C_2H_4 & (0.077, 0.081, 0.089, 0.094) \end{array}$$

A partir de estos conjuntos podemos calcular el grado de verdad y el de certidumbre en el dominio de lo necesario de cada uno de los estados de las medidas. Para la primera de ellas se tiene:



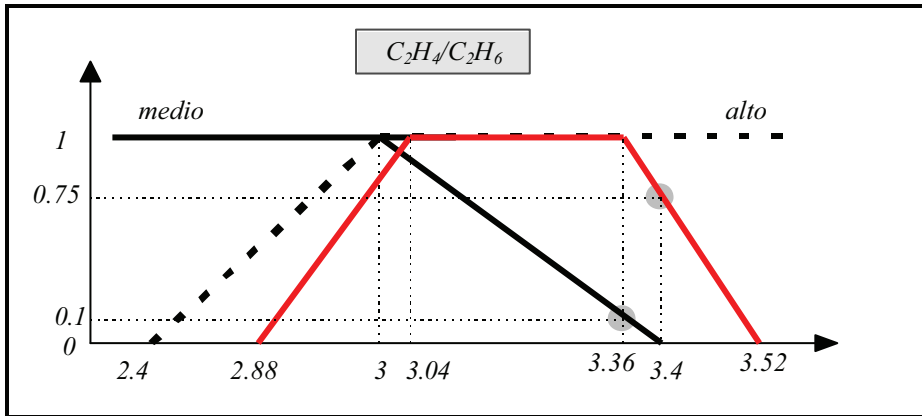
Recordemos que los grados de verdad y certidumbre van a ser calculados como:

$$\begin{array}{l} \alpha_n = N(P/[P']) \\ \beta_n = N(JP/[P']) \end{array}$$

Llamando a la observación m , esto es $m=(0.99, 1.05, 1.16, 1.21)$, la siguiente tabla refleja los grados de verdad y de certidumbre de los distintos estados de la medida, calculados en el dominio de lo necesario:

$\alpha_n(CH_4/H_2 \text{ bajo} / m)=0$	$\beta_n(CH_4/H_2 \text{ bajo} / m)=0$
$\alpha_n(CH_4/H_2 \text{ medio} / m)=\text{medio}(1.16)=0.2$	$\beta_n(CH_4/H_2 \text{ medio} / m)=1-m(1.2)=0.8$
$\alpha_n(CH_4/H_2 \text{ alto} / m)=1$	$\beta_n(CH_4/H_2 \text{ alto} / m)=1$

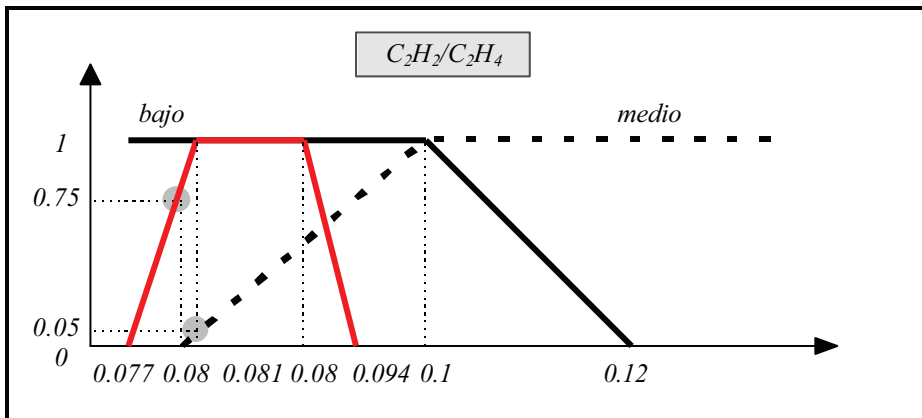
Para la segunda medida se tiene:



Del mismo modo que antes, con $m=(2.88,3.04,3.36,3.52)$, se tiene:

$\alpha_n(C_2H_4/C_2H_6 \text{ bajo}/m)=0$	$\beta_n(C_2H_4/C_2H_6 \text{ bajo}/m)=0$
$\alpha_n(C_2H_4/C_2H_6 \text{ medio}/m)=\text{medio}(3.36)=0.1$	$\beta_n(C_2H_4/C_2H_6 \text{ medio}/m)=1-m(3.4)=0.25$
$\alpha_n(C_2H_4/C_2H_6 \text{ alto}/m)=1$	$\beta_n(C_2H_4/C_2H_6 \text{ alto}/m)=1$

Por último para la tercera medida:



y tomando $m=(0.077,0.081,0.089,0.094)$:

$\alpha_n(C_2H_2/C_2H_4 \text{ bajo}/m)=1$	$\beta_n(C_2H_2/C_2H_4 \text{ bajo}/m)=1$
$\alpha_n(C_2H_2/C_2H_4 \text{ medio}/m)=\text{medio}(0.081)=0.05$	$\beta_n(C_2H_2/C_2H_4 \text{ medio}/m)=1-m(0.089)=0.25$
$\alpha_n(C_2H_2/C_2H_4 \text{ alto}/m)=0$	$\beta_n(C_2H_2/C_2H_4 \text{ alto}/m)=0$

La fotografía del estado del transformador en función de los valores de necesidad borrosos de sus medidas queda reflejado en la siguiente tabla:

	<i>bajo</i>	<i>medio</i>	<i>alto</i>
CH_4/H_2	$(N, 0, 0)$	$(N, 0.2, 0.8)$	$(N, 1, 1)$
C_2H_4/C_2H_6	$(N, 0, 0)$	$(N, 0.1, 0.25)$	$(N, 1, 1)$
C_2H_2/C_2H_4	$(N, 1, 1)$	$(N, 0.05, 0.25)$	$(N, 0, 0)$

Con estos datos se está en condiciones de calcular la conclusión, en el dominio de lo necesario, de cualquiera de las reglas de la base de conocimiento propuesta. Las reglas que se activan son las reglas $G4$, $G5$, $G6$, $G8$ y $G9$. Sin embargo el grado de activación de muchas de ellas es significativamente bajo. La comprobación de antecedentes de cada una de las reglas queda reflejada en la siguiente tabla (donde en cada una de las filas salvo la primera, que indica la regla y su modelo, está el valor de necesidad borroso de cada uno de los antecedentes):

$G4$ $(N, C, M, C, 1, 1)$	$G5$ $(N, C, G, C, 1, 1)$	$G6$ $(N, M, M, C, 1, 1)$	$G8$ $(N, M, G, C, 1, 1)$	$G9$ $(N, M, G, C, 1, 1)$
<i>medio</i> $(N, 0.2, 0.8)$	<i>medio</i> $(N, 0.2, 0.8)$	<i>medio</i> $(N, 0.2, 0.8)$	<i>alto</i> $(N, 1, 1)$	<i>alto</i> $(N, 1, 1)$
<i>medio</i> $(N, 0.1, 0.25)$	<i>alto</i> $(N, 1, 1)$	<i>medio</i> $(N, 0.1, 0.25)$	<i>medio</i> $(N, 0.1, 0.25)$	<i>alto</i> $(N, 1, 1)$
<i>medio</i> $(N, 0.05, 0.25)$	<i>medio</i> $(N, 0.05, 0.25)$	<i>medio</i> $(N, 0.05, 0.25)$	<i>bajo</i> $(N, 1, 1)$	<i>bajo</i> $(N, 1, 1)$

Aplicando el mecanismo de inferencia propuesto en 7.5.4.3 a cada una de las reglas se obtiene la siguiente tabla de resultados, donde α_n y β_n son respectivamente el grado de verdad y el grado de certidumbre, calculados en el dominio de lo necesario, del diagnóstico de cada una de las reglas:

	$G4$ $(N, C, M, C, 1, 1)$	$G5$ $(N, C, G, C, 1, 1)$	$G6$ $(N, M, M, C, 1, 1)$	$G8$ $(N, M, G, C, 1, 1)$	$G9$ $(N, M, G, C, 1, 1)$
α_n	0.1	1	$\min(0.2, 0.1) = 0.1$	$\min(1, 0.1) = 0.1$	$\min(1, 1) = 1$
β_n	$\min(0.2, 0.8, 0.1, 0.25, 0.05, 0.25) \approx 0$	$\min(0.2, 0.8, 1, 0.05, 0.25) \approx 0$	$\min(0.2, 0.8, 0.1, 0.25, 0.05, 0.25) \approx 0$	$\min(1, 1, 0.25, 1, 1) = 0.25$	$\min(1, 1, 1, 1, 1) = 1$

De la tabla anterior se deducen los diagnósticos:

$G8$: defecto térmico de media temperatura $(N, 0.1, 0.25)$

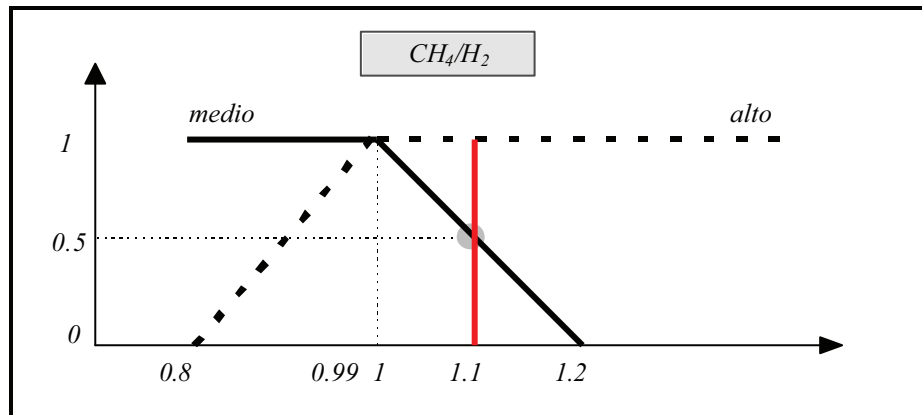
$G9$: defecto térmico de alta temperatura $(N, 1, 1)$

Por un lado, y debido a la regla *G8* el diagnóstico *defecto térmico de media temperatura* tiene un grado de verdad de 0.1 con una certidumbre de 0.25. Por sí sola esta regla solo apunta a un defecto de tipo térmico (aunque con una certidumbre escasa) y a que dicho defecto sería presumiblemente de media, baja o alta temperatura. Una certidumbre baja significa que otros diagnósticos podrían también ser candidatos para explicar la supuesta anomalía, o que incluso podría no existir fallo alguno. Sin embargo la regla *G9* confirma la información suministrada por la regla *G8*, ya que se confirma que efectivamente el defecto es de tipo térmico, y que está en el rango de las altas temperaturas. Esta acción combinada de ambas reglas es de especial interés en el ámbito borroso y especialmente para la metodología propuesta ya que en las zonas de transición de unos estados a otros, donde las reglas suelen dispararse parcialmente, reglas adyacentes con modelos apropiados permiten obtener conclusiones que se apoyan mutuamente, de forma que la combinación de varias conclusiones similares permite delimitar con precisión los diagnósticos finales. Si por ejemplo la regla *G9* se hubiese activado también con un grado de verdad parcial se podría deducir que el defecto habría sido de tipo térmico y en el rango de temperaturas situadas entre las medias y las altas, dando más información que cualquiera de las dos reglas por separado. Por otro lado la gran incertidumbre que se deduce de la regla *G8* es debido a que las observaciones se han modelado con un error significativamente grande en comparación con el solape de la partición, aunque respetando los criterios de completitud del sistema de reglas.

Si modelamos las observaciones con menor error, se mejora la completitud del sistema, tal y como se apuntaba en el apartado 5.6.3.2, de forma que se facilita la activación de mayor número de reglas. En efecto supongamos que modelemos las observaciones directamente con conjuntos unitarios:

CH_4/H_2	1.1
C_2H_4/C_2H_6	3.2
C_2H_2/C_2H_4	0.085

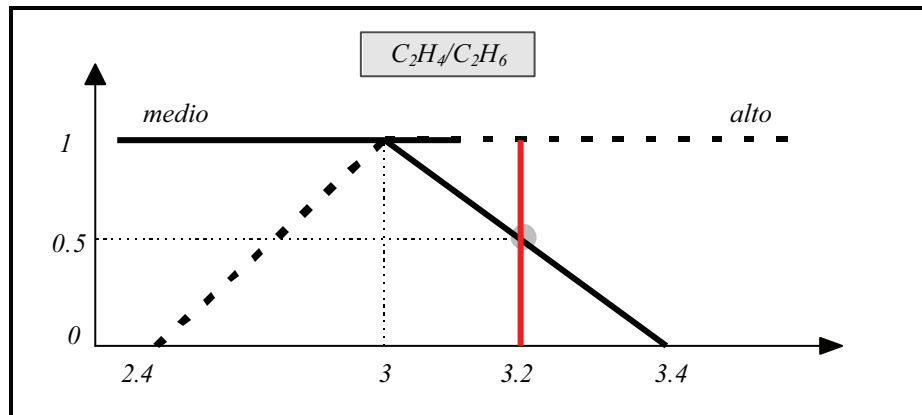
La comprobación de los estados de la primera medida da lugar a:



Y con $m=\{1.1\}$ se tiene:

$\alpha_{np}(CH_4/H_2 \text{ bajo}/m)=0$	$\beta_{np}(CH_4/H_2 \text{ bajo}/m)=0$
$\alpha_{np}(CH_4/H_2 \text{ medio}/m)=\text{medio}(1.1)=0.5$	$\beta_{np}(CH_4/H_2 \text{ es medio}/m)=1$
$\alpha_{np}(CH_4/H_2 \text{ alto}/m)=1$	$\beta_{np}(CH_4/H_2 \text{ alto}/m)=1$

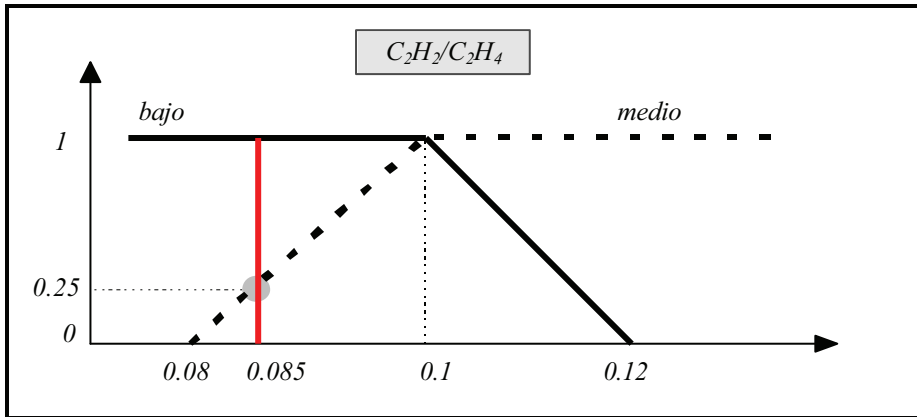
La notación con los subíndices np refleja el hecho de que para observaciones puntuales los grados de verdad y de certidumbre son iguales en el dominio de lo necesario y lo posible, por lo que en este caso solo tiene sentido razonar en el dominio de lo necesario. Para la segunda medida se tiene:



y tomando $m=\{3.2\}$

$\alpha_{np}(C_2H_4/C_2H_6 \text{ bajo}/m)=0$	$\beta_{np}(C_2H_4/C_2H_6 \text{ bajo}/m)=0$
$\alpha_{np}(C_2H_4/C_2H_6 \text{ medio}/m)=\text{medio}(3.2)=0.5$	$\beta_{np}(C_2H_4/C_2H_6 \text{ medio}/m)=1$
$\alpha_{np}(C_2H_4/C_2H_6 \text{ alto}/m)=1$	$\beta_{np}(C_2H_4/C_2H_6 \text{ alto}/m)=1$

Por último para la tercera medida se tiene:



y con $m=\{0.0085\}$ se tiene:

$\alpha_{np}(C_2H_4/C_2H_6 \text{ bajo}/m)=1$	$\beta_{np}(C_2H_2/C_2H_4 \text{ bajo}/m)=1$
$\alpha_{np}(C_2H_2/C_2H_4 \text{ medio}/m)=\text{medio}(0.085)=0.25$	$\beta_{np}(C_2H_2/C_2H_4 \text{ medio}/m)=1$
$\alpha_{np}(C_2H_2/C_2H_4 \text{ alto}/m)=0$	$\beta_{np}(C_2H_2/C_2H_4 \text{ alto}/m)=0$

Con estos datos la foto del estado del transformador es ahora:

	<i>bajo</i>	<i>medio</i>	<i>alto</i>
CH_4/H_2	$(N, 0, 0)$	$(N, 0.5, 1)$	$(N, 1, 1)$
C_2H_4/C_2H_6	$(N, 0, 0)$	$(N, 0.5, 1)$	$(N, 1, 1)$
C_2H_2/C_2H_4	$(N, 1, 1)$	$(N, 0.25, 1)$	$(N, 0, 0)$

Como puede comprobarse los grados de verdad han variado algo, pero se ha producido un incremento substancial de la certidumbre (lo que depende obviamente del núcleo de la observación por un lado, y del soporte por otro). Una alternativa intermedia a ambos modelos de observación sería el uso de conjuntos triangulares, que aun permitiendo modelar un cierto error en la observación favorecen la completitud del sistema de reglas.

La comprobación de los antecedentes de las reglas queda reflejada en la siguiente tabla:

G4 <i>(N,C,M,C,1,1)</i>	G5 <i>(N,C,G,C,1,1)</i>	G6 <i>(N,M,M,C,1,1)</i>	G8 <i>(N,M,G,C,1,1)</i>	G9 <i>(N,M,G,C,1,1)</i>
<i>medio</i> <i>(N,0.5,1)</i>	<i>medio</i> <i>(N,0.5,1)</i>	<i>medio</i> <i>(N,0.5,1)</i>	<i>alto</i> <i>(N,1,1)</i>	<i>alto</i> <i>(N,1,1)</i>
<i>medio</i> <i>(N,0.5,1)</i>	<i>alto</i> <i>(N,1,1)</i>	<i>medio</i> <i>(N,0.5,1)</i>	<i>medio</i> <i>(N,0.5,1)</i>	<i>alto</i> <i>(N,1,1)</i>
<i>medio</i> <i>(N,0.25,1)</i>	<i>medio</i> <i>(N,0.25,1)</i>	<i>medio</i> <i>(N,0.25,1)</i>	<i>bajo</i> <i>(N,1,1)</i>	<i>bajo</i> <i>(N,1,1)</i>

Con estos datos la comprobación de las reglas que se activan da lugar a la siguiente tabla de resultados:

	G4 <i>(N,C,M,C,1,1)</i>	G5 <i>(N,C,G,C,1,1)</i>	G6 <i>(N,M,M,C,1,1)</i>	G8 <i>(N,M,G,C,1,1)</i>	G9 <i>(N,M,G,C,1,1)</i>
α_n	0.5	1	$\min(0.5,0.5)=0.5$	$\min(1,0.5)=0.5$	$\min(1,1)=1$
β_n	$\min(0.5,1,0.5,1,0.25,1)=0.25$	$\min(0.5,1,1,0.25,1)=0.25$	$\min(0.5,1,0.5,1,0.25,1)=0.25$	$\min(1,1,1,1,1)=1$	$\min(1,1,1,1,1)=1$

De los resultados anteriores se deducen esta vez los diagnósticos:

- G4: descargas eléctricas de baja energía (N,0.5,0.25)*
- G5: descargas eléctricas de alta energía (N,1,0.25)*
- G6: defecto térmico de baja temperatura (N,0.5,0.25)*
- G8: defecto térmico de media temperatura (N,0.5,1)*
- G9: defecto térmico de alta temperatura (N,1,1)*

Es importante señalar algunos puntos de interés. Por un lado si los modelos se hubiesen tomado todos por defecto graduales los resultados obtenidos hubiesen sido radicalmente distintos, ya que por ejemplo para el primer diagnóstico se hubiese obtenido:

descargas eléctricas de baja energía (N,0.25,1)

lo que apuntaría claramente a un defecto de descargas eléctricas, aunque de baja o alta energía. Por tanto la determinación correcta de los modelos es importante para el correcto diagnóstico por parte del sistema.

En segundo lugar aunque se han obtenido muchos diagnósticos, los correspondientes a descargas eléctricas no se refuerzan unos a otros debido a su baja certidumbre, mientras

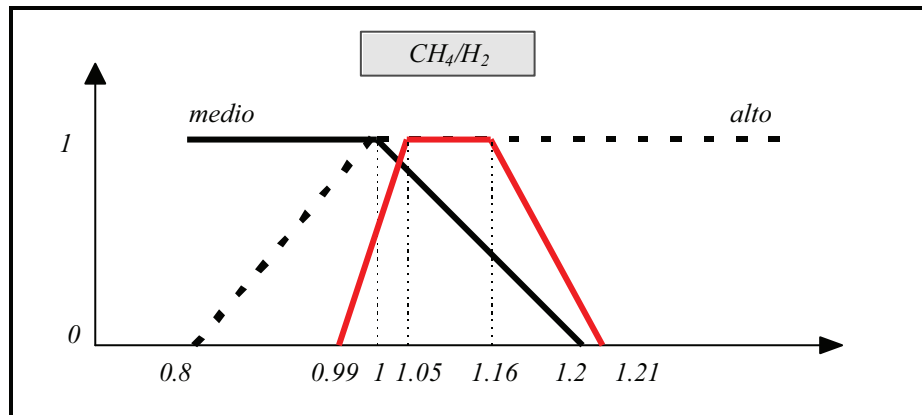
que los de defecto térmico se apoyan unos a otros, pudiendo claramente deducirse como diagnóstico final un defecto térmico de media-alta temperatura. Por tanto las conclusiones son cualitativamente las mismas que anteriormente. Sin embargo este ejemplo muestra de forma más clara como varias reglas que se han disparado simultáneamente pueden reforzarse mutuamente, y que los diagnósticos que proponen pueden afinarse mutuamente, y reafirmarse como la mejor explicación, y cómo otras reglas permiten casi descartar otros posibles defectos por su escasa certidumbre.

Supongamos por último que las observaciones se modelan como inicialmente se ha propuesto pero que no se dispone del valor de la medida C_2H_2/C_2H_4 . En este caso no se puede inferir en el dominio de lo necesario puesto que el valor de esta medida puede ser cualquiera, lo que como ya se ha visto se modela mediante una distribución de posibilidad igual a su universo de discurso. Para obtener conclusiones con la misma base de conocimiento se puede recurrir a los modelos de posibilidad semánticamente duales de los modelos de necesidad obtenidos para cada una de las reglas, para lo que es necesario calcular los grados de verdad y de certidumbre en el dominio de lo posible. Recordemos que el grado de verdad y el de certidumbre, calculados en el dominio de lo posible, han sido definidos como:

$$\alpha_p = \Pi(P/J/P')$$

$$\beta_p = \Pi(J/P/P')$$

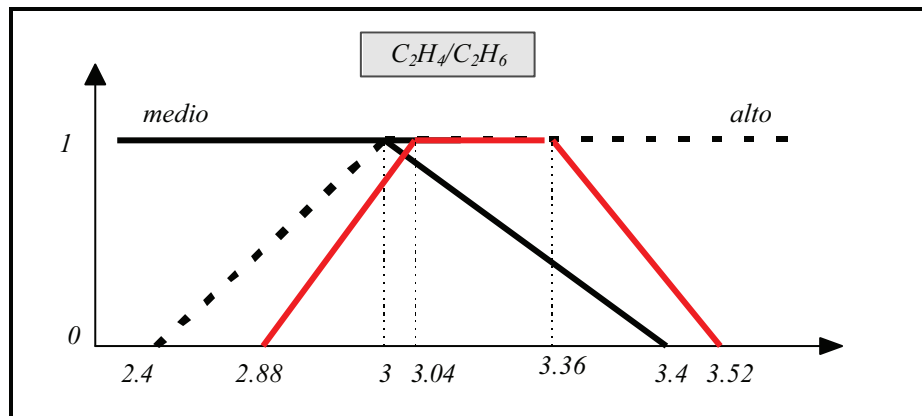
A este respecto para la primera medida se tiene:



De donde se deduce, para $m=(0.99,1.05,1.16,1.21)$, los grados de verdad y de certidumbre de los distintos estados de la medida, calculados en el dominio de lo posible:

$\alpha_p(\text{CH}_4/\text{H}_2 \text{ bajo} / m)=0$	$\beta_p(\text{CH}_4/\text{H}_2 \text{ bajo} / m)=0$
$\alpha_p(\text{CH}_4/\text{H}_2 \text{ medio} / m)=1$	$\beta_p(\text{CH}_4/\text{H}_2 \text{ medio} / m)=1$
$\alpha_p(\text{CH}_4/\text{H}_2 \text{ alto} / m)=1$	$\beta_p(\text{CH}_4/\text{H}_2 \text{ alto} / m)=1$

Para la segunda medida se tiene:



con $m=(2.88,3.04,3.36,3.52)$, se tiene:

$\alpha_p(\text{C}_2\text{H}_4/\text{C}_2\text{H}_6 \text{ bajo}/m)=0$	$\beta_p(\text{C}_2\text{H}_4/\text{C}_2\text{H}_6 \text{ bajo}/m)=0$
$\alpha_p(\text{C}_2\text{H}_4/\text{C}_2\text{H}_6 \text{ medio}/m)=1$	$\beta_p(\text{C}_2\text{H}_4/\text{C}_2\text{H}_6 \text{ medio}/m)=1$
$\alpha_p(\text{C}_2\text{H}_4/\text{C}_2\text{H}_6 \text{ alto}/m)=1$	$\beta_p(\text{C}_2\text{H}_4/\text{C}_2\text{H}_6 \text{ alto}/m)=1$

Y para la tercera medida se tiene, con $m=(0,0,10,10)$:

$\alpha_n(C_2H_4/C_2H_6 \text{ bajo}/m)=1$	$\beta_n(C_2H_2/C_2H_4 \text{ bajo}/m)=1$
$\alpha_n(C_2H_2/C_2H_4 \text{ medio}/m)=1$	$\beta_n(C_2H_2/C_2H_4 \text{ medio}/m)=1$
$\alpha_n(C_2H_2/C_2H_4 \text{ alto}/m)=1$	$\beta_n(C_2H_2/C_2H_4 \text{ alto}/m)=1$

La fotografía del estado del transformador en función de los valores de posibilidad borrosos de sus medidas queda reflejado en la siguiente tabla:

	<i>bajo</i>	<i>medio</i>	<i>alto</i>
CH_4/H_2	$(\Pi, 0, 0)$	$(\Pi, 1, 1)$	$(\Pi, 1, 1)$
C_2H_4/C_2H_6	$(\Pi, 0, 0)$	$(\Pi, 1, 1)$	$(\Pi, 1, 1)$
C_2H_2/C_2H_4	$(\Pi, 1, 1)$	$(\Pi, 1, 1)$	$(\Pi, 1, 1)$

Las conclusiones se pueden ahora calcular en el dominio de lo posible. Las reglas que se activan son las mismas que antes, esto es $G4$, $G5$, $G6$, $G8$ y $G9$, aunque en general el número de reglas activadas en el dominio de lo posible será casi siempre mayor que en el dominio de lo necesario. La comprobación de los antecedentes de las reglas da lugar a la siguiente tabla:

$G4$ $(\Pi, C, M, C, 1, 1)$	$G5$ $(\Pi, C, G, C, 1, 1)$	$G6$ $(\Pi, M, M, C, 1, 1)$	$G8$ $(\Pi, M, G, C, 1, 1)$	$G9$ $(\Pi, M, G, C, 1, 1)$
<i>medio</i> $(\Pi, 1, 1)$	<i>medio</i> $(\Pi, 1, 1)$	<i>medio</i> $(\Pi, 1, 1)$	<i>alto</i> $(\Pi, 1, 1)$	<i>alto</i> $(\Pi, 1, 1)$
<i>medio</i> $(\Pi, 1, 1)$	<i>alto</i> $(\Pi, 1, 1)$	<i>medio</i> $(\Pi, 1, 1)$	<i>medio</i> $(\Pi, 1, 1)$	<i>alto</i> $(\Pi, 1, 1)$
<i>medio</i> $(\Pi, 1, 1)$	<i>medio</i> $(\Pi, 1, 1)$	<i>medio</i> $(\Pi, 1, 1)$	<i>bajo</i> $(\Pi, 1, 1)$	<i>bajo</i> $(\Pi, 1, 1)$

Y el disparo de las reglas anteriores (aplicando el mecanismo de inferencia propuesto en 7.5.4.3) da lugar a la siguiente tabla (donde α_p y β_p son respectivamente el grado de verdad y el grado de certidumbre, calculados en el dominio de lo posible, del diagnóstico de cada una de las reglas):

	$G4$ $(N, C, M, C, 1, 1)$	$G5$ $(N, C, G, C, 1, 1)$	$G6$ $(N, M, M, C, 1, 1)$	$G8$ $(N, M, G, C, 1, 1)$	$G9$ $(N, M, G, C, 1, 1)$
α_n	1	1	1	1	1
β_n	1	1	1	1	1

De la tabla anterior se deducen los diagnósticos:

$G4$: descargas eléctricas de baja energía $(\Pi, 1, 1)$

- G5: descargas eléctricas de alta energía (II,1,1)
- G6: defecto térmico de baja temperatura (II,1,1)
- G8: defecto térmico de media temperatura (II,1,1)
- G9: defecto térmico de alta temperatura (II,1,1)

En este caso, y como era de esperar, la información disponible no permite discernir entre defectos térmico y debido a descargas eléctricas. Sin embargo sí es posible afirmar que estos dos defectos son los únicos que no son contradictorios con las medidas del transformador, lo que permite descartar otros posibles defectos. Además debe optarse por una interpretación disyuntiva en el sentido en que la verdadera causa podría ser o bien las descargas eléctricas, o bien el defecto térmico o incluso ambos (esto contrasta con el razonamiento con modelos de necesidad donde las conclusiones se interpretan conjuntivamente). Cabe señalar que la escasez de matices (todos los valores de α_p y β_p obtenidos son 1) se debe a que en este caso las observaciones se han modelado con un error significativamente grande. Nótese además que para observaciones precisas (modeladas con conjuntos unitarios), las medidas de similitud coinciden en este caso tanto en el dominio de lo necesario como en el de lo posible, por lo que resulta de escasa utilidad razonar en este último, a menos que, como en este caso, se desconozca el valor de alguna una medida.

Si en efecto suponemos que las primeras medidas son precisas pero que se sigue desconociendo en el valor de la última, para la primera y segunda medidas los grados de verdad y certidumbre son los calculados con anterioridad, mientras que para la tercera todos valen 1, con lo que en resumen se tiene:

$\alpha_{np}(CH_4/H_2 \text{ bajo}/m)=0$	$\beta_{np}(CH_4/H_2 \text{ bajo}/m)=0$
$\alpha_{np}(CH_4/H_2 \text{ medio}/m)=\text{medio}(1.1)=0.5$	$\beta_{np}(CH_4/H_2 \text{ es medio}/m)=1$
$\alpha_{np}(CH_4/H_2 \text{ alto}/m)=1$	$\beta_{np}(CH_4/H_2 \text{ alto}/m)=1$

$\alpha_{np}(C_2H_4/C_2H_6 \text{ bajo}/m)=0$	$\beta_{np}(C_2H_4/C_2H_6 \text{ bajo}/m)=0$
$\alpha_{np}(C_2H_4/C_2H_6 \text{ medio}/m)=\text{medio}(3.2)=0.5$	$\beta_{np}(C_2H_4/C_2H_6 \text{ medio}/m)=1$
$\alpha_{np}(C_2H_4/C_2H_6 \text{ alto}/m)=1$	$\beta_{np}(C_2H_4/C_2H_6 \text{ alto}/m)=1$

$\alpha_n(C_2H_4/C_2H_6 \text{ bajo}/m)=1$	$\beta_n(C_2H_2/C_2H_4 \text{ bajo}/m)=1$
$\alpha_n(C_2H_2/C_2H_4 \text{ medio}/m)=1$	$\beta_n(C_2H_2/C_2H_4 \text{ medio}/m)=1$
$\alpha_n(C_2H_2/C_2H_4 \text{ alto}/m)=1$	$\beta_n(C_2H_2/C_2H_4 \text{ alto}/m)=1$

Con estos datos la foto del estado del transformador es ahora:

	<i>bajo</i>	<i>medio</i>	<i>alto</i>
CH_4/H_2	$(\Pi, 0, 0)$	$(\Pi, 0.5, 1)$	$(\Pi, 1, 1)$
C_2H_4/C_2H_6	$(\Pi, 0, 0)$	$(\Pi, 0.5, 1)$	$(\Pi, 1, 1)$
C_2H_2/C_2H_4	$(\Pi, 1, 1)$	$(\Pi, 1, 1)$	$(\Pi, 1, 1)$

La comprobación de los antecedentes de las reglas queda reflejada en la siguiente tabla:

G4 $(\Pi, C, M, C, 1, 1)$	G5 $(\Pi, C, G, C, 1, 1)$	G6 $(\Pi, M, M, C, 1, 1)$	G8 $(\Pi, M, G, C, 1, 1)$	G9 $(\Pi, M, G, C, 1, 1)$
<i>medio</i> $(\Pi, 0.5, 1)$	<i>medio</i> $(\Pi, 0.5, 1)$	<i>medio</i> $(\Pi, 0.5, 1)$	<i>alto</i> $(\Pi, 1, 1)$	<i>alto</i> $(\Pi, 1, 1)$
<i>medio</i> $(\Pi, 0.5, 1)$	<i>alto</i> $(\Pi, 0.5, 1)$	<i>medio</i> $(\Pi, 0.5, 1)$	<i>medio</i> $(\Pi, 0.5, 1)$	<i>alto</i> $(\Pi, 1, 1)$
<i>medio</i> $(\Pi, 1, 1)$	<i>medio</i> $(\Pi, 1, 1)$	<i>medio</i> $(\Pi, 1, 1)$	<i>bajo</i> $(\Pi, 1, 1)$	<i>bajo</i> $(\Pi, 1, 1)$

Con estos datos la comprobación de las reglas que se activan da lugar a la siguiente tabla de resultados:

	G4 $(\Pi, C, M, C, 1, 1)$	G5 $(\Pi, C, G, C, 1, 1)$	G6 $(\Pi, M, M, C, 1, 1)$	G8 $(\Pi, M, G, C, 1, 1)$	G9 $(\Pi, M, G, C, 1, 1)$
α_n	0.5	1	$\min(0.5, 0.5)=0.5$	$\min(1, 0.5)=0.5$	$\min(1, 1)=1$
β_n	$\min(0.5, 1, 0.5, 1, 0.25, 1)=0.25$	$\min(0.5, 1, 1, 1, 1)=0.5$	$\min(0.5, 1, 0.5, 1, 1)=0.5$	$\min(1, 1, 1, 1, 1)=1$	$\min(1, 1, 1, 1, 1)=1$

de donde se deducen los diagnósticos:

G4: descargas eléctricas de baja energía $(\Pi, 0.5, 0.5)$

G5: descargas eléctricas de alta energía $(\Pi, 1, 0.5)$

G6: defecto térmico de baja temperatura $(\Pi, 0.5, 0.5)$

G8: defecto térmico de media temperatura $(\Pi, 0.5, 1)$

G9: defecto térmico de alta temperatura $(\Pi, 1, 1)$

En este caso las conclusiones obtenidas son más específicas que en el caso anterior. Por un lado los diagnósticos posibles son nuevamente las descargas eléctricas y el defecto térmico pero este último es más posible (o menos contradictorio con los hechos observados) que el anterior. Por otro lado en el caso de que el defecto fuese debido a descargas eléctricas estas serían más bien de alta energía, mientras que si fuese un defecto térmico este se situaría próximo a las altas temperaturas. Estos resultados, aunque obtenidos en el dominio

de lo posible son como era de esperar similares a los obtenidos en el dominio de lo necesario cuando se conocía el valor de la tercera medida, y vienen a corroborar la validez de estos resultados.

Finalmente señalar que toda la gama de matices obtenidos mediante el mecanismo de inferencia propuesto puede perderse si se recurre a otros mecanismos de inferencia más simples. Por un lado la aplicación directa de la RCI resulta ser semánticamente la opción semántica más completa. Sin embargo la riqueza semántica del mecanismo MS de dos índices propuesto y su facilidad de cálculo hacen a este última una alternativa a la RCI con características similares pero más fácil de interpretar y de calcular.

Si se recurre a la aplicación de la MS de un solo índice, las relaciones semánticas que se obtienen son menos completas. Así por ejemplo, en un modelo gradual, que se obtendría utilizando una implicación gradual, no es posible propagar la incertidumbre que pueda aparecer en la comprobación de la hipótesis de la regla. Para propagar la incertidumbre es necesario recurrir a una implicación de certidumbre, en cuyo caso se pierde la gradualidad que pueda existir entre antecedente y consecuente, o bien a un modelo mixto, que propaga de igual modo grados de verdad y de incertidumbre, con lo que la conclusión siempre tendrá una cierta incertidumbre de origen dudoso.

Si se recurre a la utilización de conjuntos clásicos junto con la regla del modus ponens clásica, la exactitud de las conclusiones obtenidas es bastante menor. En efecto, dado que solo se admiten grados de verdad y certidumbre unitarios, la distinción que se ha hecho entre ambos desaparece. No es posible representar la incertidumbre parcial, por lo que cualquier observación con una incertidumbre no nula da lugar a conclusiones triviales. Las condiciones de activación de las reglas en el ámbito clásico son más restrictivas que las obtenidas para los modelos borrosos estudiados. Esto hace que se tenga que optar por ciertas alternativas, todas ellas a nuestro entender poco satisfactorias. Una primera consiste en no concluir nada en cuanto existe una cierta incertidumbre o alguna diferencia, aún cuando sea poco significativa, entre la hipótesis y la observación, con lo que el sistema solo es capaz de diagnosticar cuando está absolutamente seguro de lo que concluye, dejando al usuario la tarea de analizar e interpretar cualquier situación dudosa, aunque sea potencialmente peligrosa. Una segunda alternativa puede consistir en sobredimensionar los conjuntos que modelan a las hipótesis y reducir los que modelan a las observaciones, de cara a obtener conclusiones para un mayor número de casos. Esto hace que ante situaciones significativamente diferentes se obtengan idénticas conclusiones, dejando nuevamente al usuario la tarea de distinguir entre conclusiones más o menos seguras en base a su experiencia y a los datos disponibles. Una tercera opción consiste en diseñar particiones muy pobladas de estados, lo que aunque da lugar a bases de conocimiento con un gran número de reglas, permite sin embargo distinguir un mayor número de situaciones que puedan ser consideradas significativamente diferentes. La obtención de conclusiones

más exactas tiene como contrapartida una mayor complejidad a la hora de enunciar el conocimiento al haber un mayor número de reglas, con los problemas asociados de consistencia y completitud.

8. Conclusiones, aportaciones y líneas de futuros desarrollos

En éste último capítulo se presentarán las conclusiones más significativas que se deducen del trabajo realizado, se señalarán las principales aportaciones de estas tesis al campo de la teoría de conjuntos borrosos y el diagnóstico, y por último se presentarán algunas líneas de futuros desarrollos que han ido surgiendo a lo largo de dicho trabajo.

8.1 Conclusiones

Dado que los resultados obtenidos a lo largo de la tesis se apoyan en un estudio de algunas técnicas básicas de la teoría de conjuntos borrosos para su aplicación posterior a los sistemas de diagnóstico, las conclusiones se pueden articular entorno a ambos puntos.

- **Conclusiones referentes a la teoría de conjuntos borrosos**

La teoría de conjuntos borrosos, a través de las distribuciones de posibilidad, es una herramienta flexible de representación del conocimiento, ya que permite modelar tanto grados de verdad parciales como grados de incertidumbre parciales, por lo que pueden modelarse con una misma herramienta conceptos vagos y/o inciertos. Sin embargo es conveniente distinguir con claridad, por un lado, lo que son grados de verdad y certidumbre parciales, y por otro, dos tipos de incertidumbre que dan lugar a otros dos tipos fundamentales de distribuciones de posibilidad, que deben interpretarse de distinto modo: las distribuciones de necesaria posibilidad y las de posible posibilidad. Mientras que las primeras proporcionan un grado máximo en que se considera posible un determinado elemento de un universo de discurso, las segundas dan una estimación de dicho grado de posibilidad en base a algún tipo de evidencia a su favor.

La distinción entre los dos tipos de distribuciones de posibilidad es especialmente importante de cara a entender e interpretar correctamente los resultados de la inferencia tanto en lógica borrosa como clásica. Existen dos modos complementarios de inferir conclusiones, el primero de ellos basado en lo que se ha llamado modelos de reglas de necesidad, que permite razonar en el dominio de lo necesario, y el otro basado en los modelos de posibilidad, que permite razonar en el dominio de lo posible. En el primero se deducen distribuciones de necesaria posibilidad y por tanto conclusiones necesarias o con un grado de certidumbre cualitativamente alto, mientras que el segundo da lugar a distribuciones de posible posibilidad, que se corresponden con conclusiones posibles o lo que es lo mismo no del todo contradictorias con los hechos observados, y por tanto con un grado de certidumbre cualitativamente bajo (aunque intuitivamente aceptable). Dada la diferente interpretación de ambos tipos de conclusiones es importante escoger correctamente los modelos a emplear.

La siempre deseable completitud y coherencia de las bases de conocimiento puede analizarse utilizando las condiciones de activación de los modelos de reglas obtenidas en el capítulo 4. Se ha visto cómo los modelos de necesidad tienen condiciones de activación mucho más restrictivas que los modelos de posibilidad, a costa de generar conclusiones con una certeza cualitativamente mayor. Esto explica en gran medida por qué en

aplicaciones numéricas, donde a menudo la completitud es uno de los requisitos fundamentales frente a la estimación de la certeza de las conclusiones, se tiende a emplear los modelos de posibilidad frente a modelos de necesidad. Ejemplos típicos de este tipo de aplicaciones son el control borroso y el modelado borroso donde el mecanismo de inferencia empleado se suele reducir a la llamada regla *max-min*. Esta aplicación masiva de los modelos de posibilidad, y en concreto de la regla *max-min*, ha hecho que en muchos casos se olvidase la existencia de otros modelos alternativos o complementarios que se han descartado de muchas aplicaciones reales donde correctamente utilizados hubiesen podido dar mejores resultados, o al menos similares. También se ha visto que desde el punto de vista de obtención de conclusiones lógicas las particiones deben diseñarse para el correcto funcionamiento de los modelos de necesidad, pero que esas mismas particiones pueden emplearse con los modelos de posibilidad cuando sea necesario, dando lugar a resultados intuitivamente satisfactorios.

Una de las principales ventajas de la lógica borrosa frente a la lógica clásica es su mejor capacidad de representación del conocimiento. Esta capacidad se ve incrementada mediante la posibilidad de diferenciar entre distintos tipos de relaciones condicionales dentro de un mismo modo de razonamiento, que permite establecer relaciones graduales, de certidumbre y mixtas, entre los consecuentes y cada uno de los antecedentes de una regla de conocimiento. La elección correcta de estos modelos es una pieza clave para la validez de las conclusiones obtenidas.

Una crítica que a menudo se hace a la teoría de conjuntos borrosos y a la teoría de la posibilidad es, por un lado la dificultad de diseñar sistemas borrosos debido a la gran cantidad de operadores y opciones disponibles, y por otro, la dificultad de interpretación de los resultados, debido en gran medida a los mismos inconvenientes que se plantean durante el diseño. El trabajo llevado a cabo en esta tesis ha puesto de manifiesto que el comportamiento de la regla composicional de inferencia puede ser analizado en detalle e interpretado en base a criterios de sentido común. Como ya se ha comentado, las relaciones condicionales no son solamente funciones matemáticas opacas, sino que reflejan distintos tipos de relaciones semánticas entre los antecedentes y los consecuentes que relacionan. Además esas relaciones existen realmente en la práctica y es posible determinarlas analizando el comportamiento de los consecuentes de reglas de conocimiento cuando se producen variaciones leves en sus antecedentes, lo que ha permitido ofrecer criterios claros y concisos para el diseño de sistemas borrosos, para la determinación correcta de los modelos a emplear y para la interpretación de sus conclusiones, desde un punto de vista teórico y no sólo empírico. Del mismo modo uno de los elogios que se hace a la teoría de conjuntos borrosos es su robustez frente a cambios de operadores, funciones de pertenencia, etc. Sin embargo y como aquí se ha visto, algunos de estos cambios pueden dar lugar a diferencias en el comportamiento final que pueden llegar a ser importantes, y que por tanto no pueden ni deben de ser ignoradas.

Otro de los inconvenientes que a menudo se señala es la gran cantidad de cálculos que es necesario hacer para operar con conjuntos borrosos, especialmente en el terreno de la inferencia lógica, debido a la formulación algo tediosa de la regla composicional de inferencia. Sin embargo y como se ha visto a lo largo del capítulo 4 es posible simplificar considerablemente esta regla manteniendo su riqueza semántica, dando lugar a un mecanismo de inferencia mucho más simple y rápido de calcular. Esta simplificación conduce a un tipo de modificación por similitud, pero donde es necesario utilizar dos índices en lugar de uno, como tradicionalmente viene haciéndose, para medir la similitud entre la hipótesis de una regla y la observación disponible. Además estos índices pueden ser interpretados, uno, como el grado de verdad o grado en que se parece la hipótesis a la observación, y otro, como el grado de incertidumbre de la hipótesis, que viene a cuantificar en qué medida la hipótesis tiene al menos un parecido no nulo con la observación. Estos índices pueden definirse tanto en el dominio de lo necesario como en el de lo posible, por lo que la simplificación propuesta sigue manteniendo los dos modos de razonamiento complementarios.

Por último se ha visto que es posible encontrar una familia de valores de verdad borrosos a partir de las implicaciones generalizadas propuestas, que permiten cualificar afirmaciones en términos de grados de verdad y en términos de grados de certidumbre, tanto en el dominio de lo necesario como en el dominio de lo posible. Estos valores de verdad borrosos permiten enriquecer la expresión del conocimiento, e integrar en un entorno único resultados procedentes de otros módulos de razonamiento diferentes. Se ha visto como la familia tradicional de valores de verdad borrosos propuesta por Zadeh puede resultar un tanto arbitraria, difícil de interpretar y por tanto no siempre adecuada a los modelos de reglas utilizados habitualmente, y en concreto a la mayor parte de modelos analizados a lo largo del capítulo 4 de esta tesis.

- **Conclusiones referentes al diagnóstico**

En primer lugar se ha mostrado cómo es posible mejorar y aprovechar, utilizando técnicas borrosas, los mecanismos de detección de anomalías basados en modelos, y en concreto el propuesto en la tesis [Muñoz, 96], gracias a los valores de verdad borrosos que se describen en el capítulo 6. Se ha presentado una forma de integración eficiente que permite que el módulo de identificación de anomalías utilice de forma homogénea la información procedente del módulo de detección de anomalías y del módulo de adquisición de datos, aún cuando dicha información tenga una fuerte componente de incertidumbre o vaguedad.

La base de conocimiento de un sistema de diagnóstico debe construirse a partir del conjunto de fallos que puede presentar el proceso y de sus síntomas asociados, y deben

emplearse principalmente reglas con modelos de necesidad, a diferencia de como se hace tradicionalmente en los sistemas de control borroso. Sin embargo y en caso de que esto no sea posible se deja abierta la posibilidad de emplear modelos de posibilidad. A partir de la base de conocimiento es posible utilizar un procedimiento bastante sistemático para la determinación precisa de los modelos de reglas, esto es de las relaciones existentes entre cada uno de los consecuentes y antecedentes de las reglas, procedimiento que además puede permitir evaluar parcialmente la completitud de la base de conocimiento disponible. Como ya se ha comentado, la determinación correcta de estos modelos es clave para la validez de las conclusiones obtenidas por el sistema.

El uso de ambos tipos de modelos de regla da lugar a una estrategia de diagnóstico que se estructura en dos niveles de razonamiento. En el primer nivel se hace uso de toda la información con un grado de certidumbre cualitativamente elevado (distribuciones de necesaria posibilidad y modelos de necesidad), información que hemos denominado fiable, y permite deducir conclusiones también fiables, pero que en general son difíciles de obtener. Cuando la información disponible no es fiable, o cuando en el primer nivel no se pueden obtener conclusiones que expliquen adecuadamente las anomalías detectadas, se procede a razonar en el segundo nivel, capaz de conjuntar distribuciones de posible posibilidad y modelos de posibilidad. En éste se determinan aquellas conclusiones que sin ser del todo fiables, al menos no son del todo contradictorias con los hechos disponibles que describen el estado del proceso. Se consigue así un aprovechamiento máximo del conocimiento disponible, aún cuando éste sea incompleto o parcialmente inconsistente.

Los dos modos de razonamiento señalados, que habitualmente se presentan como contradictorios o alternativos, son en realidad complementarios, permitiendo llevar a cabo procesos deductivos en dos niveles, y flexibilizando de este modelo el motor de inferencias que es el corazón del sistema de identificación de anomalías. La elección no es por tanto escoger entre uno y otro, sino optar por un mecanismo que los integre a ambos, y dentro de cada uno de ellos escoger apropiadamente los modelos que reflejen con precisión la semántica del conocimiento que se quiere representar.

8.2 Aportaciones de la tesis

La principal aportación de esta tesis ha sido la de realizar un estudio exhaustivo de los principales mecanismos de inferencia en lógica borrosa, para determinar con claridad las diferencias existentes entre unos y otros, no sólo a nivel numérico o matemático sino también a nivel semántico, y la de aplicarlo al razonamiento en el ámbito del diagnóstico. Se han catalogado distintas formas de inferir y se han identificado los operadores que deben ser utilizados en cada una de ellas. Esto ha permitido proponer un sistema de identificación de anomalías basado en reglas con modelos borrosos, con mayor riqueza semántica y por tanto con mayor capacidad de representación del conocimiento que otros sistemas similares propuestos con anterioridad. A continuación se detallan las aportaciones más significativas.

- **Distribuciones de necesaria y posible posibilidad**

Una de las primeras aportaciones ha consistido en distinguir claramente dos tipos de distribuciones de posibilidad que a menudo se confunden, y de clarificar las distintas reglas de inferencia que pueden aplicarse con unas y con otras. La distinción entre ambos tipos de distribuciones ha permitido diferenciar el razonamiento en dos dominios de certidumbre diferentes. El dominio de lo necesario por un lado, donde las afirmaciones tienen un grado de certidumbre cualitativamente alto, y donde el razonamiento produce conclusiones con grados de certidumbre también altos. Por otro lado el dominio de lo posible, donde las afirmaciones se basan en estimaciones razonables o plausibles pero con poca certidumbre, y donde el razonamiento produce como conclusiones aquellas que no están en total contradicción con los hechos disponibles, pero que no pueden probarse con seguridad. La forma de definir ambos tipos de distribuciones ha permitido en gran medida justificar las reglas de inferencia a aplicar, sin necesidad de recurrir a consideraciones empíricas. Aunque todavía queda mucho trabajo por realizar en este camino, en esta tesis se han dado algunos de los primeros pasos.

- **Condicionales borrosos y funciones de implicación**

Se ha hecho un estudio exhaustivo del comportamiento de la regla composicional de inferencia cuando se utilizan distintos tipos de funciones de implicación, desde una óptica diferente a la que tradicionalmente se ha empleado en otros trabajos. Se ha mostrado cómo pueden diferenciarse dos grandes categorías de funciones de implicación, aquellas que se obtienen como generalizaciones de la implicación material clásica, las *n*-implicaciones, que dan lugar a condicionales que pueden interpretarse como

distribuciones de necesaria posibilidad, y aquellas que se obtienen como generalizaciones del producto cartesiano clásico, las p -implicaciones, que dan lugar a condicionales interpretables como distribuciones de posible posibilidad. El análisis realizado ha dado lugar a una metodología destinada a estudiar e interpretar el comportamiento de cualquier modelo de regla. Para ello los principales pasos que deben seguirse son el análisis de la modificación llevada a cabo sobre el consecuente cuando la observación es puntual, el estudio de las condiciones de equivalencia entre la RCI y la MS que permiten determinar el índice de similitud empleado por la RCI para modificar el consecuente ante la ausencia de incertidumbre, y el estudio de la condición de activación del modelo que permite identificar en qué dominio de razonamiento trabaja. Además, por un lado, se ha propuesto una n -implicación y una p -implicación generales que permiten, aunque de forma aproximada, reproducir los comportamientos semánticos de las principales funciones de implicación utilizadas en la regla composicional, y que hemos llamado implicaciones generalizadas. Por otro lado se ha ido más lejos generalizado estos dos tipos de implicaciones a un tipo de implicación que los engloba y que hemos llamado las φ -implicaciones, que permiten generar los φ -modelos de reglas que se introducen en el apéndice de la tesis, y cuya interpretación, todavía por explorar pero susceptible de hacerse con la misma metodología propuesta, ha sido brevemente señalada para algunos casos particulares.

- **Simplificación numérica y semántica de la regla composicional de inferencia**

Las implicaciones generalizadas propuestas en ésta tesis han permitido simplificar la regla composicional de inferencia a un tipo de modificación por similitud basado en dos índices. Se ha visto cómo escogiendo adecuadamente los índices de similitud se pueden reproducir los principales comportamientos semánticos de la regla composicional de inferencia. Se ha visto también que la simplificación se puede llevar más lejos y obtener un mecanismo de inferencia para la aplicación del modus ponens, sencillo de interpretar y de implementar, basado en el cálculo de los grados de verdad y de certidumbre de los consecuentes, de directa aplicación práctica al diseño de sistemas borrosos, y en concreto de sistemas de diagnóstico, a los que se ha aplicado desde un punto de vista teórico, e ilustrado en un caso práctico mediante el estudio de un ejemplo concreto. De hecho es posible comprobar que si se utilizan únicamente conjuntos borrosos λ_{np} -trapezoidales se pueden encontrar expresiones sencillas que proporcionen directamente los citados índices de similitud, simplificando aún más la aplicación de la regla de inferencia.

- **Combinación de dos modos de razonamiento en un mismo entorno**

Además de detectar e interpretar dos modos de razonamiento significativamente diferentes, uno en el dominio de lo necesario y otro en el dominio de lo posible, en esta tesis se han

propuesto mecanismos destinados a combinarlos en un entorno único. En efecto se ha mostrado cómo ambos mecanismos no son alternativos, como a menudo se presentan, sino complementarios. Así se ha propuesto, en primer lugar, obtener todas aquellas conclusiones que hemos calificado de necesarias (o con alto grado de certidumbre), mediante el uso de distribuciones de necesaria posibilidad y modelos de reglas de necesidad. Sin embargo en muchas ocasiones las exigentes condiciones de activación de estos modelos pueden dar lugar a la obtención de conclusiones triviales y carentes de interés informativo. En esos casos es posible, mediante el uso de los modelos de posibilidad semánticamente duales, razonar en el dominio de lo posible, y obtener al menos aquellas conclusiones menos contradictorias con los hechos observados. Esto permite razonar en dos niveles, de forma que en entornos con información fiable y precisa las conclusiones sean acordes a la misma, pero que en entornos inciertos, donde se pueden desconocer valores de variables clave, y donde pueden existir contradicciones profundas en la base de conocimiento disponible, se obtengan a pesar de todo conclusiones razonables, flexibilizando aún más las capacidades deductivas de los sistemas borrosos. El paso de un nivel a otro lo determinan la cuantía y certidumbre de las conclusiones obtenidas en el primero.

- **Modelos de reglas**

Los distintos modos de razonamiento señalados se basan en el uso de los dos tipos de distribuciones y de dos tipos de modelos de reglas, basados en n-implicaciones y p-implicaciones. Además en cada modo de razonamiento se han identificado diferentes relaciones semánticas entre los antecedentes y los consecuente de las reglas. Para modelar con la mayor precisión posible estas relaciones semánticas se ha propuesto el uso de modelos de reglas que especifican para cada consecuente el tipo de relación que mantiene con cada uno de los antecedentes, lo que permite una gran riqueza de expresión de conocimiento ausente en otros sistemas similares, donde a menudo se confunden grados de verdad y grados de certidumbre. Además se ha propuesto un procedimiento semiautomático de obtención de estos modelos a partir de una base de conocimiento, y se ha mostrado su aplicación práctica en un ejemplo concreto. Con estas características se obtienen reglas de conocimiento con mucha más información que las reglas clásicas o reglas borrosas con un único modelo implícito, de forma que para un mismo número de reglas las conclusiones que se obtienen pueden llegar a ser más precisas y acordes a las observaciones disponibles. Se ha visto también cómo los valores de verdad borrosos pueden obtenerse a partir de funciones de implicación, de forma que comparten o heredan su mismo significado semántico. A este respecto puede señalarse que la familia propuesta por Zadeh puede generarse a partir de la implicación de Yager, y que ambas pecan de una cierta opacidad a la hora de ser interpretadas. Por otro lado la metodología propuesta ha permitido detectar una estrecha relación entre la implicación de Yager y la t-norma

producto, y ha dado lugar a un nuevo mecanismo de generación de funciones de implicación a partir de la familia de las t-normas.

- **Sistema experto de diagnóstico basado en conocimiento vago e incierto. Métodos de formalización de reglas.**

Las aportaciones señaladas en los puntos anteriores han sido aplicadas al establecimiento de una metodología de diseño de sistemas expertos de diagnóstico basados en conocimiento vago e incierto, conocimiento modelado mediante el uso de conjuntos borrosos, distribuciones de posibilidad y reglas de conocimiento con modelos borrosos. Se han propuesto estructuras de representación del conocimiento estructural (componentes y medidas) y funcional (reglas de conocimiento), mecanismos para la formalización del conocimiento en forma de reglas, un procedimiento para la determinación de los modelos de regla más adecuados, y se han descrito con precisión las estrategias de diagnóstico y el motor de inferencias encargado de llevarlas a cabo. Se ha hecho especial hincapié en describir en detalle varios de los aspectos fundamentales, por ejemplo en lo que respecta a las estructuras de representación del conocimiento propuestas, al orden de activación de las reglas, a la determinación de las conclusiones de una regla, a la agregación de conclusiones procedentes de diferentes reglas, etc., con vistas a presentar una metodología suficientemente completa de cara a la realización práctica de sistemas de diagnósticos para uso industrial. Además la metodología propuesta se ha aplicado a un ejemplo real que ha permitido ilustrar sus características más señaladas, en particular en lo referente a la determinación de los modelos de reglas y al mecanismo de inferencia propuesto.

8.3 Líneas de futuros desarrollos

- **Análisis detallado del comportamiento e interpretación de los φ -modelos**

A menudo el desarrollo de teorías más generales permiten obtener conclusiones aplicables a casos particulares, que no habían sido detectadas durante el estudio de éstos últimos, así como un mejor entendimiento de los mismos. En ese sentido nos parece que un estudio del comportamiento e interpretación de los φ -modelos propuestos en el apéndice de la tesis contribuiría notablemente a una mejor comprensión de los n y p -modelos analizados, que son los que habitualmente se utilizan en los sistemas borrosos, así como a encontrar nuevas y fructíferas aplicaciones de los mismos. Este estudio podría llevarse a cabo con una metodología similar a la empleada en el capítulo 4, de forma que también se pudiesen identificar distintos tipos de relaciones semánticas, como las ya comentadas graduales, de certidumbre y mixtas.

- **Análisis de la completitud y la coherencia de una base de conocimiento**

Una posible aplicación de los φ -modelos podría consistir en la obtención de medidas de la calidad de una base de conocimiento. Por un lado se ha visto que cuando las reglas están caracterizadas por n -modelos es posible detectar inconsistencias debidas a conclusiones contradictorias de reglas que puedan dispararse simultáneamente. Por otro lado, y supuesto que la base de conocimiento es completa, debería ser posible, ante cualquier situación, obtener al menos conclusiones posibles que permitan explicar los hechos observados (las conclusiones necesarias pueden no existir debido a hechos incompletos o poco precisos). Por tanto, en el caso en que el uso de los p -modelos semánticamente duales de los n -modelos originales no permitan obtener alguna conclusión que no esté en contradicción con los hechos, esto es reflejo de lagunas en la base de conocimiento. Estas consideraciones evidencian por tanto que cada tipo de modelo permite detectar un tipo de defecto de la base de conocimiento, y ambos son casos extremos dentro de la familia de los φ -modelos. Esto a su vez hace sugerente la idea de encontrar medidas de la calidad de una base de conocimiento, posiblemente basadas en el uso de los φ -modelos que mejores conclusiones permitan inferir, en cuanto a certidumbre y especificidad de las mismas.

- **Análisis de los modelos del tipo Yager generados por t -normas**

Como se ha visto en el capítulo 4 es posible encontrar una expresión matemática que permite generar la implicación de Yager a partir de la t -norma producto. Esta expresión permite a su vez generar nuevas implicaciones, y por tanto nuevos modelos de reglas, a

partir de otras t-normas diferentes. Una línea de investigación común a muchos trabajos en el ámbito de la inferencia borrosa, consiste precisamente en encontrar expresiones generales para la obtención de funciones de implicación con distintas propiedades de las habitualmente utilizadas, y en esta misma dinámica un posible desarrollo consistiría en profundizar en este mecanismo de generación, y explorar los nuevos modelos de reglas que de él se deducen, posiblemente de cara a encontrar nuevos comportamientos semánticos.

- **Análisis de la evolución temporal de los grados de verdad y de certidumbre de un diagnóstico**

El funcionamiento continuado de un sistema de diagnóstico desarrollado según la metodología propuesta en la tesis da lugar a la obtención de series temporales de la evolución de los grados de verdad y de certidumbre de cada uno de los posibles diagnósticos del sistema. Estas series constituyen una fuente de información valiosa desde dos puntos de vista. Por un lado el análisis de sus características, como por ejemplo la tendencia, son una forma de identificación incipiente de anomalías, ya que el diagnóstico puede detectarse inicialmente con grados de verdad y de certidumbre bajos, pero que cuya evolución futura al alza o a la baja, debe permitir confirmar o rechazar su validez antes que su propio valor absoluto. Por otro lado esta fuente de información puede ser utilizada para refinar el conocimiento disponible en la base de diagnóstico. Por ejemplo la evolución temporal progresiva de un diagnóstico a otro puede dar información sobre posibles interdependencias entre ambos, sobre tiempos de degradación de componentes a partir de posibles sobrecargas o defectos en otros componentes cercanos, etc. También parece viable la creación de modelos lógicos que, teniendo como salidas los grados de verdad y de certidumbre de los diagnósticos del sistema, utilicen como entradas sus fuentes de información en forma de grados de pertenencia de las medidas a sus estados más representativos.

- **Modelado borroso**

Aunque en muchas aplicaciones numéricas se prescinde de la certidumbre que puedan tener las conclusiones obtenidas, en otras, como ocurre por ejemplo con el módulo de detección de anomalías propuesto en [Muñoz, 96], esta certeza es una pieza clave de la información final suministrada. Una posible aplicación de los mecanismos de inferencia propuestos en esta tesis podría ser el campo del modelado borroso, donde además de una estimación numérica, el modelo fuese capaz de cuantificar la certidumbre de la estimación, mediante un grado de certidumbre asociado. La distinción que se ha hecho entre grados de verdad y grados de certidumbre, y que ha permitido la integración eficiente del citado módulo, apunta a la posibilidad de aplicar las citadas técnicas directamente al terreno del

modelado borroso, y en concreto al módulo de detección de anomalías. De esta forma las medidas de incertidumbre obtenidas serían directamente medidas de necesidad y posibilidad, y por tanto directamente interpretables por el sistema de identificación.

- **Aplicación de las técnicas propuestas a otros campos de actividad**

Esta tesis se ha dedicado en su mayor parte al estudio y análisis de las herramientas que ofrece la teoría de conjuntos borrosos en lo que respecta a la inferencia, para su aplicación a la tarea de identificación de anomalías en un sistema de diagnóstico. Aunque en el último capítulo se propone y se ilustra una metodología completa para el desarrollo de un sistema de este tipo, dada su gran amplitud en aplicaciones reales, sería conveniente, de cara a validar los resultados propuestos, y a detectar y corregir los inconvenientes que una aplicación real puede presentar, proceder al desarrollo de una aplicación completa para el diagnóstico de un proceso complejo. Esto permitiría contemplar de forma más sistemática cada uno de estos inconvenientes y las posibles soluciones que pudiesen aportarse, con vistas a crear una metodología de diagnóstico lo más completa y sistemática posible. Es importante también abordar algunos problemas que en esta tesis no han sido analizados, como por ejemplo la evaluación del impacto de algunas de las simplificaciones que la metodología propuesta, principalmente en lo que concierne a la inferencia y al tipo de conjuntos considerados, exige para su aplicación. Sería también interesante trasladar los resultados obtenidos a otros campos de actividad como por ejemplo el control borroso, y evaluar las implicaciones prácticas derivadas del uso de unos u otros tipos de modelos, partiendo siempre de particiones y reglas bien diseñadas, para poder obtener criterios de comparación válidos.

A. Los φ -condicionales

La definición de función de implicación propuesta en el capítulo 4, en el apartado 4.2.2.2, contiene como casos particulares a las implicaciones que hemos denominado n-implicaciones y p-implicaciones, que se caracterizan respectivamente porque verifican $I_n(0,y)=1$ e $I_p(0,y)=0$. La observación de estas dos propiedades sugiere la posibilidad de estudiar un tipo más general de implicaciones, que hemos llamado φ -implicaciones, caracterizadas por $I_\varphi(0,y)=\varphi$, entre las cuales se cuentan como casos particulares las n y p-implicaciones. Estas implicaciones dan lugar a los φ -condicionales, que permiten generar nuevos modelos de reglas con nuevas semánticas por analizar. En este apéndice se formaliza este tipo de implicaciones y condicionales, se analizan algunas de sus propiedades, se introduce el concepto de distribución de φ -posibilidad y se sugieren algunas posibles interpretaciones.

A.1 φ -implicaciones y φ -condicionales

A.1.1 Definiciones

La definición de condicional propuesta en el capítulo 4 es de carácter muy general y da lugar a muchas posibles particularizaciones, dos de las cuales, extensamente analizadas a lo largo de dicho capítulo, son las que hemos llamado n-implicaciones y p-implicaciones.

Recordemos que una función $I(x,y)$ es una función de implicación borrosa, o simplemente una **función de implicación**, o implicación, cuando verifique:

$$\begin{cases} I:[0,1] \times [0,1] \rightarrow [0,1] \\ I(1,y) = y \\ I(x,y) \text{ no decreciente con } y \end{cases}$$

Llamaremos **φ -implicación** a una función de implicación I_φ que verifique $I_\varphi(0,y) = \varphi$. Así puede comprobarse que una **n-implicación** es una **1-implicación** mientras que una **p-implicación** es una **0-implicación**.

Así como a una n-implicación le exigíamos que fuese no creciente con x , y a una p-implicación que fuese no decreciente con x , a una φ -implicación le exigiremos que sea no creciente con x para $y \geq \varphi$, y no decreciente con x para $y \leq \varphi$.

Diremos que un φ -condicional para una regla $P \rightarrow Q$ es el condicional generado por una φ -implicación I_φ como sigue:

$$R_\varphi(u,v) = I_\varphi(P(u), Q(v))$$

A.1.2 Generación de φ -implicaciones

Dada la naturaleza intermedia de las φ -implicaciones, situadas entre las n y p-implicaciones, parece razonable utilizar estas últimas para mediante algún tipo de agregación generar φ -implicaciones. Dado un par de implicaciones (I_n, I_p) usado para generar una φ -implicación I_φ diremos que I_n y I_p son las implicaciones generadoras de I_φ .

Una posible forma de generar φ -implicaciones a partir de una n-implicación y de una p-implicación es por ejemplo mediante el uso de una media generalizada:

$$I_{\varphi}(x, y) = \left(\frac{(I_n(x, y))^{\alpha} + (I_p(x, y))^{\alpha}}{2} \right)^{\frac{1}{\alpha}} = m_{\alpha}(I_n, I_p)$$

donde:

1. Para $\alpha = \infty$ se tiene $m_{\alpha}(x, y) = \max(x, y)$ por lo que se recupera la n-implicación.
2. Para $\alpha = -\infty$ se tiene $m_{\alpha}(x, y) = \min(x, y)$ por lo que se recupera la p-implicación.
3. Para $\alpha \in [-\infty, 0]$ se obtienen distintos tipos de p-implicaciones ya que $I_{\varphi}(0, y) = 0$.
Si por ejemplo $\alpha = -2$, $I_{\varphi}(0, y) = m_{\alpha}(1, 0) = 0$.
4. Para $\alpha \in]0, 1]$ se obtiene φ -implicaciones con $\varphi \in [0, 0.5]$
5. Para $\alpha \in [1, +\infty]$ se tienen φ -implicaciones con $\varphi \in [0.5, 1]$

Puede calcularse el valor de α necesario para obtener un φ determinado, según:

$$\alpha = \frac{\ln(0.5)}{\ln(\varphi)}$$

Otra posible alternativa para obtener φ -implicaciones viene dada por la siguiente expresión:

$$I_{\varphi}(x, y) = \varphi \cdot I_n(x, y) + (1 - \varphi) \cdot I_p(x, y)$$

donde:

1. Para $\varphi = 1$ se recupera la n-implicación generadora.
2. Para $\varphi = 0$ se recupera la p-implicación generadora.

Una última alternativa que proponemos consiste en utilizar la mediana definida por ejemplo en [Yager, 91], que es un operador idempotente, asociativo y conmutativo:

$$I_{\varphi}(x, y) = \text{Med}_{\varphi}(I_n(x, y), I_p(x, y)) = \begin{cases} \max(I_n(x, y), I_p(x, y)) = I_n(x, y) & \text{si } \max(I_n(x, y), I_p(x, y)) = I_n(x, y) \leq \varphi \\ \min(I_n(x, y), I_p(x, y)) = I_n(x, y) & \text{si } \min(I_n(x, y), I_p(x, y)) = I_n(x, y) \geq \varphi \\ \varphi & \text{en otro caso} \end{cases}$$

A.2 Función generadora de modus ponens de un φ -condicional y φ -modelos de reglas

Una φ -implicación generada por alguno de los procedimientos anteriores es siempre inferior a la n -implicación generadora, y mayor que la p -implicación generadora. En efecto, para el primero de los procedimientos, dado que se basa en hacer una medida de las implicaciones generadoras, se tiene:

$$I_p(x, y) \leq I_\varphi(x, y) = \left(\frac{(I_n(x, y))^\alpha + (I_p(x, y))^\alpha}{2} \right)^{\frac{1}{\alpha}} \leq I_n(x, y)$$

y en el segundo, al ser una combinación lineal, con $\varphi \in [0, 1]$, también se verifica.

$$I_p(x, y) \leq I_\varphi(x, y) = \varphi \cdot I_n(x, y) + (1 - \varphi) \cdot I_p(x, y) \leq I_n(x, y)$$

Y del mismo modo puede comprobarse que para la tercera forma de generar φ -implicaciones también se verifica $I_p(x, y) \leq I_\varphi(x, y) \leq I_n(x, y)$.

Por otro lado dada una función M generadora de modus ponens para la n -implicación se tiene:

$$\begin{aligned} M(x, I_n(x, y)) &\leq y \\ \sup_x M(x, I_n(x, y)) &= y \end{aligned}$$

y como M es no decreciente con ninguno de sus argumentos, y como además se ha visto que $I_\varphi(x, y) \leq I_n(x, y)$ se cumple que:

$$M(x, I_\varphi(x, y)) \leq y$$

por lo que si M es FGMP de I_n también es FGMP para la nueva φ -implicación I_φ .

Definimos un **φ -modelo** para la RCI como el par (M, I_φ) donde I_φ es una φ -implicación y M es una FGMP para I_φ .

A.3 Condiciones de activación de φ -condicionales

Como se vio en el capítulo 4 la condición de activación de un modelo para una regla $P(u) \rightarrow Q(v)$ es la relación que deben verificar la observación $P'(u)$ y la hipótesis $P(u)$ para obtener una conclusión $Q'(v)$ no trivial. Recordemos que $P'(u)$ y $P(u)$ son conjuntos borrosos definidos en U , y $Q'(v)$ y $Q(v)$ son conjuntos borrosos definidos en V .

Dado un modelo (M, I) la expresión general que permite obtener la conclusión Q' utilizando la regla composicional de inferencia es:

$$Q'(v) = \sup_u M(P'(u), I(P(u), Q(v)))$$

Por otro lado se vio en el capítulo 4 que el concepto de conclusión trivial depende del tipo de modelo empleado. Así por ejemplo mientras la conclusión trivial de los modelos de necesidad (basados en n -implicaciones) viene dada por $Q'(v) = I$, la de los modelos de posibilidad viene dada por $Q'(v) = 0$. Esto parece sugerir el que la conclusión trivial de un φ -modelo venga dada por $Q'(v) = \varphi$, y puede verificarse que en efecto, cuando la hipótesis P y la observación P' no presentan ningún parecido, esto es cuando verifican $P \cap P' = 0$, el supremo en la expresión de $Q'(v)$ se obtiene para $P'(u) = I$ y $P(u) = 0$, y resulta ser igual a φ para todo valor de v : la conclusión se corresponde con $M(I, I_\varphi(0, y)) = I_\varphi(0, y) = \varphi$.

Por tanto la condición de activación de un φ -modelo puede plantearse de forma general como:

$$Q'(v) = \sup_u M(P'(u), I_\varphi(P(u), Q(v))) \neq \varphi$$

Por otro lado, como $I_\varphi(0, \varphi) = \varphi$ y además I_φ es no decreciente con su segundo argumento, se tiene $I_\varphi(0, \varphi^+) \geq \varphi$ y $I_\varphi(0, \varphi^-) \leq \varphi$, y por tanto en la zona definida por los valores $v / Q(v) > \varphi$ se tiene $I_\varphi(0, y) \geq \varphi$ mientras que en la zona definida por $v / Q(v) = 0$ se tiene $I_\varphi(0, y) \leq \varphi$.

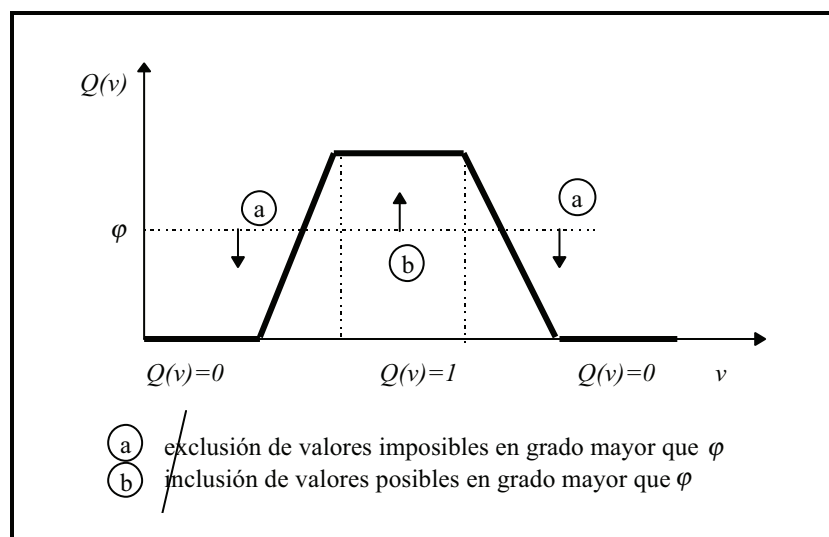
Esto significa que pueden plantearse dos condiciones de activación distintas para los dos casos extremos posibles, esto es para $Q(v) = I$ en cuyo caso la activación se da cuando $\exists v / Q'(v)_{v/Q(v)=I} > \varphi$, o bien para $Q(v) = 0$ en cuyo caso la activación se da si $\forall v / Q'(v)_{v/Q(v)=0} < \varphi$. Obviamente para un modelo de necesidad solo la segunda de las condiciones tiene sentido puesto que $\varphi = I$, mientras que para un modelo de posibilidad solo la primera de ellas lo tiene al ser $\varphi = 0$.

Nótese que estrictamente hablando la condición de activación es única, y será, de las dos propuestas, la que menos restricciones imponga a P' con respecto a P . Sin embargo tiene sentido a efectos de interpretación estudiar las dos situaciones que se plantean, esto es:

- a) $\sup_u M(P'(u), I_\varphi(P(u), 0)) < \varphi$
- b) $\sup_u M(P'(u), I_\varphi(P(u), 1)) > \varphi$

Retomando la interpretación propuesta para las distribuciones de posible y necesaria posibilidad parece que estaríamos en un caso intermedio en el que en para $Q'(v) \geq \varphi$ la distribución se interpreta como de posible posibilidad, mientras que para la zona en que $Q'(v) \leq \varphi$ la distribución se interpreta como de necesaria posibilidad. Esta interpretación tan directa solamente es válida cuando el operador de agregación empleado para combinar las implicaciones generadoras es la mediana anteriormente definida. En otros casos la interpretación es más compleja aunque no por ello deja de ser interesante. Analizaremos muy brevemente dos casos particulares.

Supongamos que generamos la φ -implicación a partir del operador mediana. En este caso I_φ vale o bien I_n o bien I_p . En este caso $Q'(v)$ puede interpretarse, para valores de v tales que $Q'(v) \geq \varphi$, como posibles grados de posibilidad de v , y para valores de v tales que $Q'(v) \leq \varphi$, $Q'(v)$ puede interpretarse como un valor de posibilidad máximo para v . En efecto cuando la φ -implicación es igual a la n-implicación generadora, el condicional es una distribución condicional de necesidad, y el resultado puede interpretarse también como una distribución de necesaria posibilidad. Análogamente ocurre cuando la φ -implicación es igual a la p-implicación generadora. Esto queda ilustrado en la siguiente figura:



Además en este caso la condición b) se da siempre antes que la a), puesto que cuando $I_\varphi=I_n$ se obtiene una condición de inclusión de P' en P , mientras que cuando $I_\varphi=I_p$ se obtiene una condición de intersección que es siempre menos restrictiva. Nótese que se

exige tanto un grado de intersección no nulo superior o igual a φ como un grado de inclusión no nulo también superior o igual a φ .

Si por el contrario suponemos que la φ -implicación se obtiene utilizando la media generalizada con $\alpha=1$, entonces se tiene:

$$I_{\varphi}(x, y) = \frac{I_n(x, y) + I_p(x, y)}{2} = m(I_n, I_p)$$

En este caso se está haciendo una medida aritmética del condicional de necesidad y del condicional de posibilidad. En cierto sentido para cada par (x, y) se está obteniendo su grado de posibilidad medio, ni su máximo grado de posibilidad, como se haría con la n-implicación, ni una estimación de cómo de posible es, como se haría con la p-implicación. Es por tanto un enfoque con ciertos matices probabilísticos, en el que la distribución de posibilidad condicional es una distribución de posibilidad cuyos valores podrían interpretarse como grados de posibilidad medios, y que podría calificarse como distribución de probable posibilidad. En este caso el grado de posibilidad correspondiente con la indeterminación es el valor 0.5 (como era de esperar) ya que $\varphi=0.5$.

La condición de activación es ahora más compleja que en el caso anterior, y es una condición media entre la intersección y la inclusión. En efecto ni la condición a) ni la condición b) se traducen directamente en medidas de intersección o inclusión, y puede suponerse a priori que la condición a) será una condición de inclusión relajada (es decir menos restrictiva) al ser $I_{\varphi}(P(u), 0) \leq I_n(P(u), 0)$. Del mismo modo la condición b) será una condición de intersección endurecida, puesto que $I_{\varphi}(P(u), 0) \geq I_n(P(u), 0)$. Para el caso concreto en que $\alpha=1$ se tiene:

$$a) \sup_u M(P'(u), (I_n(P(u), 0) + I_p(P(u), 0))/2) = \sup_u M(P'(u), I_n(P(u), 0)/2) < 0.5$$

$$b) \sup_u M(P'(u), (I_n(P(u), 1) + I_p(P(u), 1))/2) = \sup_u M(P'(u), I_p(P(u), 1)/2) > 0.5$$

Esto sugiere la riqueza de interpretaciones que pueden darse a este nuevo tipo de condicionales, interpretación que depende tanto del operador de agregación que se utilice para generar la φ -implicación, como del valor de φ escogido. Además esta interpretación será también función del tipo de n y p-implicaciones generadoras, de forma que a priori parece previsible que existan al menos las semánticas graduales, de certidumbre y mixtas.

A.4 Agregación de φ -condicionales

Un problema que lógicamente se plantea, si se desea construir un sistema de reglas con φ -modelos, consiste en determinar cómo deben agregarse las distribuciones de posibilidad que intervienen, que denominaremos φ -distribuciones.

Como ya vimos para los modelos de necesidad la agregación debía ser de tipo conjuntivo, cuyas propiedades fundamentales son:

$$\begin{aligned} T(0,0) &= T(1,0) = T(0,1) = 0 \\ T(1,1) &= 1 \\ T(x,y) & \text{ comutativo y asociativo} \end{aligned}$$

o bien para los de posibilidad operadores de tipo disyuntivo:

$$\begin{aligned} S(0,0) &= 0 \\ S(0,1) &= S(1,0) = S(1,1) \\ S(x,y) & \text{ comutativo y asociativo} \end{aligned}$$

En el caso de las φ -distribuciones, la comutatividad sigue pareciendo una propiedad esencial ya que el orden de agregación no debería influir en el resultado final, pero la asociatividad puede en principio ser una propiedad más discutible, en concreto cuando los grados de posibilidad se interpretan como grados medios puesto que proceden del cálculo de medias que son operadores no asociativos. Parece por tanto que las propiedades que deberían exigirse a un operador de agregación de φ -distribuciones podrían ser:

$$\begin{aligned} A(0,0) &= 0 \\ A(1,1) &= 1 \\ A(0,1) &= A(1,0) = A(\varphi, \varphi) = \varphi \\ A(x,y) & \text{ comutativo} \end{aligned}$$

Estas propiedades las verifican algunos de los operadores de combinación n y p-implicaciones propuestos para la generación de φ -implicaciones. En efecto para la media generalizada se tiene:

$$\begin{aligned} A(x,y) &= \left(\frac{x^\alpha + y^\alpha}{2} \right)^{\frac{1}{\alpha}} \\ \alpha &= \frac{\ln(0.5)}{\ln(\varphi)} \end{aligned}$$

y puede comprobarse que para $\alpha = \infty$ se tiene $A(x,y) = \max(x,y)$ y para $\alpha = -\infty$ $A(x,y) = \min(x,y)$, recuperándose los dos operadores de agregación extremos de las distribuciones de posible y necesaria posibilidad, cuando la φ -implicación es una implicación de posibilidad o de necesidad. Este operador no es asociativo, salvo en sus límites extremos. Si se desea obtener un operador asociativo se puede recurrir a la mediana, que además es idempotente y conmutativo:

$$Med_{\alpha}(x,y) = \begin{cases} \max(x,y) & \text{si } \max(x,y) \leq \alpha \\ \min(x,y) & \text{si } \min(x,y) \geq \alpha \\ \alpha & \text{otro caso} \end{cases}$$

Nuevamente se pueden recuperar como casos límite los operadores \max y \min . Este operador es de hecho el único operador de agregación que situado entre \min y \max es asociativo e idempotente.

Aunque no profundizaremos en ello, parece razonable suponer que deba existir un cierto paralelismo semántico entre el tipo de condicional empleado y el operador de agregación de las distribuciones involucradas. Recordemos que según la interpretación propuesta en esta tesis, las distribuciones involucradas en una misma expresión deben ser interpretables del mismo modo. Esto sugiere la posible conveniencia de utilizar la mediana para las distribuciones resultantes del φ -modelo generado mediante una p-implicación y una n-implicación combinadas también con la mediana, o la media aritmética cuando el φ -modelo se haya obtenido también mediante la misma media aritmética.

Podría plantearse la posibilidad de utilizar una combinación lineal para la agregación de distribuciones, puesto que también permite obtener φ -implicaciones:

$$A(x,y) = \varphi x + (1-\varphi)y$$

sin embargo este operador no es conmutativo ni cumple el requisito básico $A(\varphi, \varphi) = \varphi$.

Cabe plantearse también si es posible obtener un operador de agregación agregando una t-norma y una t-conorma del mismo modo que agregábamos las implicaciones de necesidad y posibilidad. Por ejemplo utilizando la combinación lineal:

$$A(x,y) = \beta S(x,y) + (1-\beta)T(x,y)$$

si por ejemplo $S=\max$ y $T=\min$ entonces el operador de agregación así definido cumple las propiedades antes propuestas. Si las t-normas y t-conormas empleadas son la máxima y mínima respectivamente, esto es:

$$S_{\max}(x,y) = \begin{cases} 0 & \text{si } x = y = 0 \\ 1 & \end{cases}$$

$$T_{\min}(x,y) = \begin{cases} 1 & \text{si } x = y = 1 \\ 0 & \end{cases}$$

entonces también se obtiene un operador de agregación válido. Los dos casos señalados son casos extremos. Para el primero de los casos se tiene:

$$\beta \cdot \beta + (1-\beta) \cdot \beta = \beta$$

y para el segundo:

$$\beta \cdot 1 + (1-\beta) \cdot 0 = \beta$$

Para otras t-normas y t-conormas aunque se cumplen el resto de las propiedades no se verifica $A(\beta, \beta) = \beta$, y lo mismo ocurre con el operado definido como:

$$A(x,y) = \left(\frac{S^\alpha(x,y) + T^\alpha(x,y)}{2} \right)^{\frac{1}{\alpha}} = m_\alpha(S(x,y), T(x,y))$$

A.5 Relajación de hipótesis en sistemas de reglas

Supongamos que se tiene un sistema de reglas paralelas:

$$R_i: P_i \rightarrow Q_i$$

con P_i definido en U , Q_i definido en V .

La tendencia habitual en el mundo de la ingeniería y más concretamente en el mundo del control, cuando se utilizan herramientas basadas en la teoría de conjuntos borrosos, consiste en hacer uso de alguna p-implicación, generalmente alguna de las t-normas *min* o *prod*, lo que conduce a modelos de posibilidad. Sin embargo la justificación de esta elección suele limitarse a constataciones empíricas de su buen comportamiento, sin que existan bases teóricas sólidas para su uso. A lo largo de esta tesis hemos visto algunos de los argumentos a favor de estos modelos, principalmente sus condiciones de activación menos restrictivas que los hacen más atractivos en aplicaciones numéricas. Además estos modelos se comportan mejor cuando se utiliza como método de desborrosificación el centro de gravedad (véase por ejemplo [Villar, Sanz, 94]), aunque en principio depende de los criterios que se empleen para diseñar las particiones borrosas.

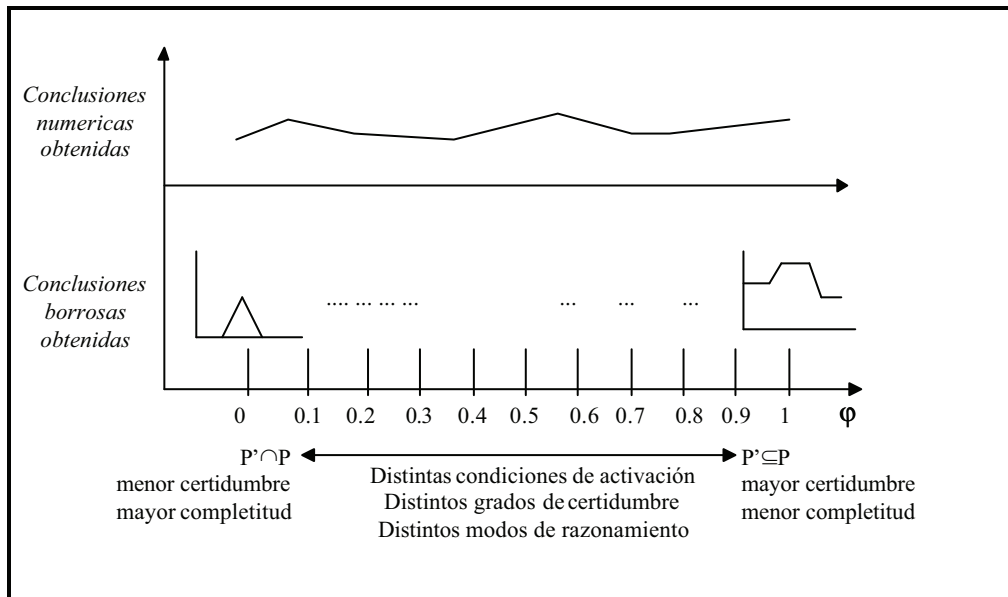
En el ámbito de control y modelado borroso lo que prima es la obtención de un resultado numérico no indeterminado, aunque su certidumbre no sea todo lo grande que se podría desear. Por ello interesa utilizar reglas con condiciones de activación poco restrictivas, lo que apunta directamente a los modelos de posibilidad, para los que además las inconsistencias de la base de conocimiento no se manifiestan de forma traumática (véase el apartado 5.6.3.3 del capítulo 5). La certidumbre de la conclusión no puede anteponerse a la obtención de la misma (por incierta que sea) ya que un sistema de control debe suministrar permanentemente un mando ante cualquier entrada, al igual que un modelo una estimación.

Por otro lado los resultados aquí presentados permiten señalar la existencia de otros tipos de modelos situados a medio camino entre los modelos de necesidad y los modelos de posibilidad, para los que se puede considerar que tienen condiciones de activación, para grados de pertenencia altos, semejantes a medidas de intersección entre la observación P' y una hipótesis P^* obtenida relajando P (es decir modelada mediante un conjunto mayor que P), y para grados de pertenencia bajos, semejantes a medidas de inclusión entre la observación P' y una hipótesis P^{**} obtenida restringiendo P (es decir modelado mediante un conjunto menor), lo que significa que ambas condiciones tienden a acercarse.

Este tipo de modelos ofrece la posibilidad de disminuir o aumentar la certidumbre que se tiene en la conclusión, modificando también la precisión de la información que suministran (cuanto menor es la certidumbre mayor es la precisión que se puede obtener). El valor φ de un φ -condicional mide cuánto de lejos o de cerca se está de que la certidumbre de la conclusión pueda medirse el dominio de lo necesario. En cualquier caso para un valor fijo de φ , dependiendo del grado de cumplimiento de la hipótesis se obtendrán conclusiones más o menos específicas.

Supongamos que se tiene un grupo de reglas con modelos de necesidad que representan el conocimiento sobre comportamiento de un sistema, en términos de entradas y salidas al mismo. Ante unas entradas concretas al sistema, una posible estrategia, para a partir de las citadas reglas obtener el valor de las salidas, podría ser la siguiente:

- Utilizar como punto de partida los modelos de necesidad originales para intentar obtener una conclusión no trivial y suficientemente informativa, que permita, en caso de que fuese necesario, llevar a cabo una desborrosificación razonable.
- Si la conclusión obtenida contiene demasiada incertidumbre, se puede obtener un φ -modelo semánticamente dual del modelo de necesidad, y modificar el valor de φ desde 1 hasta 0, hasta que las conclusiones obtenidas resulten ser suficientemente satisfactorias. El valor final de φ obtenido puede considerarse como una medida de cómo de lejos están las reglas enunciadas de ser válidas en el dominio de lo necesario, y por tanto valores bajos pueden ser indicativos de la conveniencia de refinar el conocimiento disponible. Se obtiene de este modo un beneficio doble. Por un lado se pueden obtener conclusiones tan específicas como el sistema de reglas lo permita, sin necesidad de recurrir sistemáticamente a modelos de posibilidad. Por otro lado el valor de φ obtenido resulta ser una medida de la certidumbre del modo de razonamiento que ha sido necesario utilizar para obtener la conclusión deseada. La siguiente figura ilustra el paso progresivo de un modelo de necesidad a un modelo de posibilidad variando el parámetro φ de un φ -modelo.



Por otro lado, como ya se ha señalado, la coherencia de un sistema de reglas puede analizarse basándose en modelos de necesidad. En efecto un sistema no será coherente cuando varias reglas susceptibles de activarse simultáneamente tengan conclusiones contradictorias. Del mismo modo su completitud puede analizarse utilizando modelos de posibilidad. Cuando utilizando los modelos de posibilidad no sea posible obtener una conclusión distinta de la conclusión trivial se interpreta como que todas las conclusiones resultan estar en contradicción con los hechos observados. Si asumimos que el conjunto de conclusiones disponible es suficientemente exhaustivo, esta situación pone de manifiesto la necesidad de añadir nuevas reglas o modificar las existentes. La transición de uno a otro tipo de modelo variando el parámetro ϕ permite no sólo aumentar el grado de especificidad de las conclusiones, sino salvar posibles inconsistencias de la base de conocimiento, a costa de reducir la certidumbre de las conclusiones obtenidas.

Queda por último señalar el hecho de que, en el supuesto de que se desee desborrosificar la conclusión obtenida, puede ser necesario investigar cual es el método de desborrosificación más adecuado al tipo de modelo empleado. Además, es de esperar que exista una cierta dependencia del método de desborrosificación con el parámetro ϕ que caracterice al modelo empleado. En [Villar, Sanz, 94] se puso de manifiesto que el método que calcula la media de los valores con máximo grado de pertenencia (conocido normalmente como MOM) parecía más adecuado para los modelos, mientras que el centro de gravedad de la distribución parecía más adecuado para los modelos de posibilidad. Sin embargo los criterios utilizados en dicho trabajo para diseñar las particiones no coinciden con los que se ha propuesto utilizar en esta tesis (véase el capítulo 5). Esto pone de manifiesto la necesidad de analizar en profundidad la conveniencia de

unos u otros métodos según el tipo de modelo empleado. En cualquier caso nótese que los dos métodos mencionados son casos particulares del siguiente:

$$v_0 = \frac{\int_{v/Q'(v) \geq \varphi} Q'(v) \cdot v \cdot dv}{\int_{v/Q'(v) \geq \varphi} Q'(v) \cdot dv}$$

que efectivamente depende del valor de φ que caracteriza al modelo empleado.