

# Mining Temporal Causal Relations in Medical Texts

A. Sobrino<sup>1</sup>, C. Puente<sup>2</sup>, J. A. Olivas<sup>3</sup>

<sup>1</sup> Facultad de Filosofía, Plaza de Mazarelos, s.n., Universidad de Santiago de Compostela,  
Spain  
alejandro.sobrino@usc.es

<sup>2</sup> Advanced Technical Faculty of Engineering ICAI, Pontifical Comillas University,  
Madrid, Spain  
crisrina.puente@comillas.edu

<sup>3</sup> J. A. Olivas, Department of Information Technologies and Systems,  
University of Castilla-La Mancha, Ciudad Real, Spain  
joseangel.olivas@uclm.es

**Abstract.** Causal sentences are a main part of the medical explanations, providing the causes of diseases or showing the effects of medical treatments. In medicine, causal association is frequently related to time restrictions. So, some drugs must be taken before or after meals, being ‘after’ and ‘before’ temporary constraints. Thus, we conjecture that medical papers include a lot of time causal sentences. Causality involves a transfer of qualities from the cause to the effect, denoted by a directed arrow. An arrow connecting the node cause with the node effect is a causal graph. Causal graphs are an imagery way to show the causal dependencies that a sentence shows using plain text. In this paper, we will provide several programs to extract time causal sentences from medical Internet resources and to convert the obtained sentences in their equivalent causal graphs, providing an enlightening image of the relations that a text describes, showing the cause-effect links and the temporary constraints affecting their interpretation.

**Keywords:** Causality, time, mining causal sentences, causal graphs, time constrained causal graphs.

## 1. Introduction: Time and Causation

Causality and time are largely related [1]. Causal effects can dramatically change depending on time restrictions. This is a matter of fact that time is very relevant in the area of medicine [2]. Thus, the Institute for Safe Medication Practices published in 2011 the Guidelines for Timely Administration of Scheduled Medications, distinguishing time-critical from non-time scheduled medications. So, some drugs require administration within a specified period of time before, after, or with meals. For example, the medication for curing hypothyroidism must be taken at least one hour before breakfast.

Time and causality are key factors in medical practice, as we can note in many medical reports. Doctors usually record the incidences of a disease including chronological information in clinical histories writing or making databases with time-stamps [3]. So, the expected data to change medication after an operation or the estimated time to wait for medical discharge are frequently in medical reports. Medicine is, also, a causal science. Etiology or diagnoses are frequently based on

causal schemas or mechanisms. as the following examples show: ‘Sinusitis is usually caused by a virus’, ‘Common side effects of antihistamines are drowsiness’, etc..

Time is a fundamental ingredient of causality, as shows the traditional characterization of causal relations. In effect, a causal relation is defined by three general postulates [4]:

1. Irreflexivity: Nihil is causa sui.
2. Anti-symmetry: Provided that X and Y are distinct, if X is a cause of Y, then Y is not a cause of X.
3. Precedence: If X causes Y, then X precedes Y.

Principles 1, 2 and 3 are related, as the anti-symmetry is derived from the temporal precedence, and the temporal precedence excludes the reflexivity. Precedence means that the cause starts before the effect does, or that the effect ends later the cause does. The possibility that both events occur instantaneously is discarded. In any interpretation of causal relations, the cause precedes the effect and no event precedes itself: thus, postulates 1 and 2 are a corollary of the third one. Even being time a determining factor in the characterization of causality, it is not explicitly represented in the most frequent graphical methods used for illustrating it.

Traditionally, causality is presented as a crisp categorical relation. So, many causal examples are matched with the schema: A causes B. But an empirical inspection overthrows this conviction. Causes not always necessarily provoke the effect, but simply contribute to it in a certain degree. Cause favors the effect in some extent.

Causes, effects and the cause-effect link are often qualified by degrees (percentages) or words denoting different levels of strength. Puente et al. [5] evidenced that, even in scientist texts, causality is not expressed categorically, but in a qualified manner. Short papers extracted from the Hawking’s Physics Colloquium were processed and the mined phrases showed that, even in the realm of Physics, causal relations are gradual, in so far as the causal link is frequently tinged with semantic hedges or vague adjectives, as ‘too’, ‘definitely’, ‘nearly’, and so on (in bold in table 1).

**Table 1.** Causal sentences from Hawking’s *Physics Colloquium*.

<ul style="list-style-type: none"> <li>• If the field is <b>nearly</b> constant in a region, the gradient terms will be <b>small</b>, and the energy momentum tensor, will be minus half V, times the metric.</li> <li>• However, Lindeh pointed out that if the potential is not <b>too</b> steep, the expansion of the universe, will slow down the rate at which the field rolls down the potential, to the minimum.</li> <li>• The observations do not yet indicate that the universe is <b>definitely</b> open, or that lambda is non zero, but it is beginning to look like one or the other, if not both</li> <li>• A <b>very</b> small perturbation if you are with one of the big banks.</li> <li>• If the dilation had a <b>low</b> value, the effective coupling would be <b>weak</b>, and string theory would be a <b>good</b> quantum theory.</li> </ul>
---

Time relations can be also crisp or fuzzy. So, claiming that a medicine has to be taken 1 hour before the meal denotes a crisp temporal restriction (1 hour before), but if we say that a test can help detect early cancer, a fuzzy time restriction is used (early). Allen in [6] defined an ontology of precise temporal relationships based on three types of relationships: (a) Interval-interval basic relationships, (b) point-point basic relationships and c) point-interval basic relationships. Fuzzyfication of Allen’s ontology can be reached using fuzzy values or fuzzy intervals. Later on, we will discuss this issue.

Medical texts frequently include timed and fuzzy causal sentences. In this paper, we will mine some medical Internet sources in order to check this fact. Causal sentences affected by time restrictions are extracted. A program depicts the causal graph corresponding to a causal sentence and the time restriction is annotated. Time restrictions can be associated to the cause or to the effect. In both cases, an arrow is

directed to the time line, pointing to the time restriction affecting the causal relation: the cause provokes the effect if the time condition is satisfied or the cause induces the effect with a given cadence. Thus, we intend to achieve an illustrative and intuitive representation of the causal relationship between two entities conditioned by time in crisp and fuzzy scenarios.

In order to get that objective, the paper is structured as follows: In section 2, we present the Allen time ontology for crisp time relationships and its fuzzyfication. In point 3, the methodology for transforming a causal sentence in a causal graph is presented. In point 4 time annotations to the causal graph are provided. In Section 5 we present the main achievements and future work and point 6 concludes with References and Agreements.

## 2 Allen’s Temporal Relations and its fuzzyfication

According to Allen, [7], the crisp ontology of time is made from points and intervals. A time point is a single element denoting an instantaneous event. The interval is an ordered pair of time points, being the first point less than the second (in some time scale). The more typical relations are: i) between time intervals, ii) time points-time interval, and iii) between time points as can be seen in the following tables that illustrate these relations:

**Table 2.** Basic temporal relations between intervals.

Relation	Symbol	Inverse	Illustration
A before B	<	>	AAA BBB
A occurs at the same time as B	=	=	AAA BBB
A meets B	m	mr	AAA BBB
A overlaps B	o	or	AAA BBB
A during B	d	dr	AAA BBBBBB
A starts B	s	sr	AAA BBBBBB
A finishes B	f	fr	AAA BBBBBB

**Table 3.** Relation between time points-time interval.

Relation	Symbol	Inverse	Illustration
tp before B	<	>	■ BBB
tp during B	=	There isn’t	BBB■BBB
tp starts B	s	There isn’t	■ BBB
<b>A overlaps B</b>	o	There isn’t	■ BBB

**Table 4.** Relations between time points.

Relation	Symbol	Illustration
$p_i$ occurs at the same time as $p_j$	=	■ ●
$p_i$ before $p_j$	<	■ ●
$p_i$ inverse_before $p_j$ (after)	>	● ■

Instants and intervals are taken in the Allen ontology as crisp time items. But causal sentences are frequently stained with fuzzy temporal restrictions. That is especially apparent in medical reports. So, a drug causes a therapeutic effect if it is administrated a little before the meal, being ‘a little’ a fuzzy instant, or to lose weight you can not eat protein at night, being ‘at night’ a fuzzy interval, because the night period varies depending on different seasons and cultures. Sentences as those suggest that we should soften the crisp representations of the primitives of the Allen ontology defining a fuzzy time point and a fuzzy interval.

According to Dubois et al. [8], a fuzzy time point  $p$  is a point that has no a precise situation in a clock. So, he arrived at about 5 o'clock or he is almost 15 years old are examples of fuzzy time points. Fuzzy time points contrasts with crisp time ones, as ‘Five o'clock’. Fig 1 represents both times: The crisp case is depicted by a one-step function and the fuzzy one by a soft fall function (a triangular function).

A fuzzy time point is modeled by a possibility distribution function,  $\pi_p : T \rightarrow [0,1]$ . Then,  $\forall t \in T, \pi_p(t) \in [0,1]$  is the numerical estimation of the possibility that  $p$  precisely be  $t$ .



Figure 1. Representation of the fuzzy time ‘about 5 o'clock’

If an action concerns more than an instant  $t$ , then an interval must be supplied. Dubois et al. described a fuzzy interval as a pair of fuzzy sets  $]A, B[$  and  $[A, B]$ . The set  $]A, B[$  is a set of time points that are more or less certainly between  $A$  and  $B$ , and the set  $[A, B]$  is a set of time points that are possibly between  $A$  and  $B$ . These proximity measures can be calculated by:

$$\mu_{]A, B[}(t) = \sup_{s \leq t \leq s'} \min(\pi_A(s), \pi_B(s')) = [A, +\infty) \cap (-\infty B)$$

$$\mu_{[A, B]}(t) = \min(\mu_{]A, +\infty[}(t), \mu_{(-\infty, B]}(t)) = (-\infty, A)^C \cap [A, +\infty)^C$$

Figure 2 illustrates a case of fuzzy time interval (with no-overlapping distributions) as that corresponding to the sentence, ‘between about 3 and about 6’.

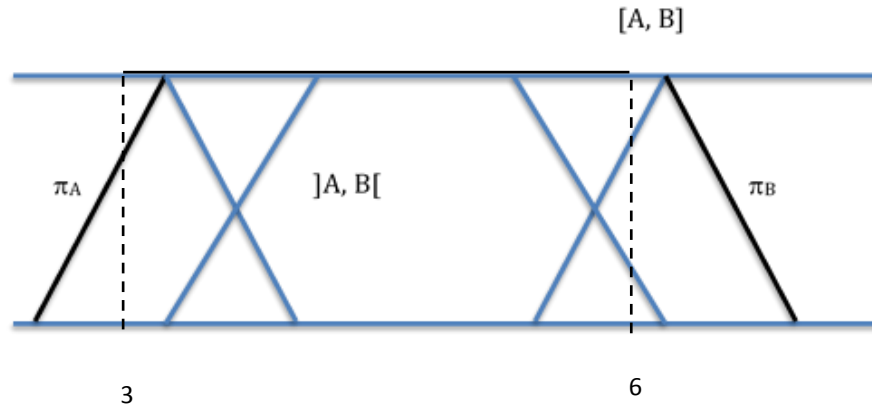


Figure 2. Representation of the fuzzy interval ‘between about 3 and about 6’.

### 3 Mining Causal Relations

In [5] we showed 48 different patterns of causal and conditional sentences in English. We select the 20 ones with less ambiguity problems. This selection provides the basis for our experiments. Table 5 shows some noted examples.

**Table 5.** Set of some conditional and causal structures implemented.

Structure 1: if + present simple + future simple.
Structure 2 : if + present simple + may/might.
Structure 3 : if + present simple + must/should.
Structure 4 : if + past simple + would + infinitive.
Structure 5 : if + past simple + might/could.
Structure 6 : if + past continuous + would + infinitive.
Structure 7 : if + past perfect + would + infinitive.
Structure 8 : if + past perfect + would have+ past participle.

With these patterns, we built a four-state automaton with the GNU *Flex* morphological analyzer, capable of extracting this type of sentences from a given text, so to obtain a causal base of knowledge about a theme. The results were quite good, as we realized that in scientific texts were much better than in general texts, such as novels, news or even movie-scripts. For this reason, we decided to move our experiments to the medical field, in particular to documents related to cancer.

#### 3.1 Creating a Causal Graph

Once the knowledge base about a scope was been created, the user has to introduce a question to run the whole process. In [9], we created another program, using *Flex*, *C* programming language and the *Stuttgart tree tagger POST* [10], to select the main concept from an input query and know whether the user is asking for causes or consequences. For example, if the user asks *What provokes lung cancer?*, the *POST* tagger would return as shown in figure 3.

What	WP
provokes	VVZ
lung	NN
cancer	NN
?	SENT

Figure 3. POST tagger output.

*POST* output shows that the nominal clause is *lung cancer*. Processing this clause with the morphological analyzer, the program, that detects the word *provokes* plus the interrogative pronoun *what*, would assume that the user is asking for the cause of lung cancer.

Once the nominal clause has been selected and isolated, another program has been developed to extract those sentences in which these concepts are contained. The searching set is the file created with the conditional and causal sentences. The retrieved set of sentences will serve as the input for the sentence summary process.

With the concept ‘Lung Cancer’ over the previously filtered database, we obtained a set of sentences to be represented into the causal graph related to this concept (figure 5). A schema of this process is presented in figure 4:

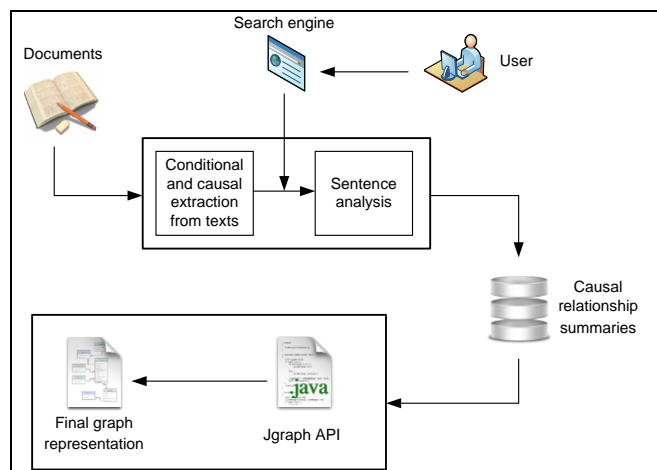


Figure 4. Process to obtain a causal graph from plain text.

This schema is stored in a database which is accessed by the graphing process. Once all the phrases have been processed and schematized, the graph algorithm, by reading the concepts associated to the cause or the consequence, creates then the causal graph, as seen in figure 5.

A relationship is specified by a modifier if the arrow is inside of the modifier cell, otherwise the arrow will be pointing at the border of the node. If the relationship has more than one modifier, the arrow will point to an intermediate node, aiming the arrows to the corresponding modifiers. The intensity of the relationship’s arrow (quantifiers) if so, is marked in red, and the type of causal connective in black. As we can see in the graph, there are four nodes with the word *smoking*, or related, and another four with the words *lung cancer*. The rest of the nodes have been retrieved in the process. So with a graph like this, we could establish new relationships hidden at first sight, for instance between smoking and other nodes. Fuzzy probabilities and fuzzy quantifiers are labeled in the lower side of the arrow.

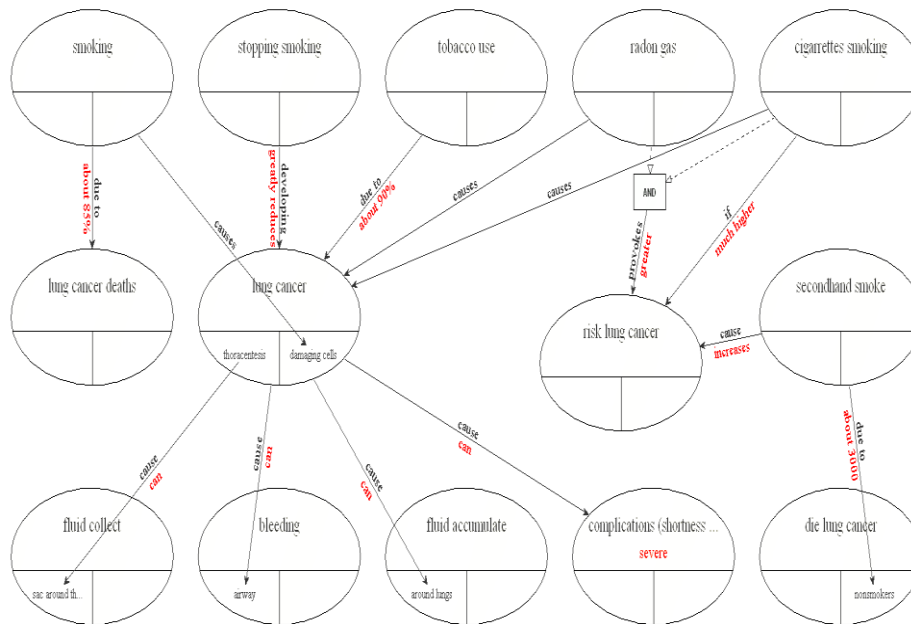


Figure 5. Automatically retrieved causal graph.

#### 4 Causation and Time

To introduce time restrictions in the graph, we performed another search with each one of the nodes to select those sentences that had time modifiers. So, we included in our Flex morphological analyzer the most common words indicating time, like *after that, time, hour, minute, day, before, now, today, tomorrow, yesterday, always, ever, etc.* From a set of 1214 causal sentences connected to lung cancer, this program obtained 370 involving time markers.

Once isolated, the next step was to identify if the time was associated to the antecedent node or to the consequent. For that we detected the modifier and used the same program as in the previous step where we formulated a question and the system had to know if we were asking for a cause or an effect. Using the Stanford POST, we are able to establish the dependencies of each word of the sentence with the others, knowing if the modifier is affecting to the cause or to the effect. For example, in the sentence “*If, for some reason, surgery is not an option with early stage 1 and 2, treatment is usually radiation therapy*”, the time constraint works on the cause. In this case, the output of the parser would return the results of table 7:

**Table 7.** Output of the Stanford Parser for the sentence “*If for some reason, surgery is not an option with early stage 1 and 2, treatment is usually radiation therapy*”.

root ( ROOT-0 , therapy-23 )	case ( stage-14 , with-12 )
mark ( 1-15 , If-1 )	amod ( stage-14 , early-13 )
case ( reason-5 , for-3 )	nmod ( not-9 , stage-14 )
det ( reason-5 , some-4 )	advcl ( therapy-23 , 1-15 )
nmod ( 1-15 , reason-5 )	cc ( 1-15 , and-16 )
nsubj ( 1-15 , surgery-7 )	conj ( 1-15 , 2-17 )
cop ( 1-15 , is-8 )	nsubj ( therapy-23 , treatment-19 )
neg ( 1-15 , not-9 )	cop ( therapy-23 , is-20 )
det ( option-11 , an-10 )	advmod ( therapy-23 , usually-21 )
nmod:npm ( not-9 , option-11 )	compound ( therapy-23 , radiation-22 )

So, in this example we can appreciate that the time modifier ‘early’ is associated to the word *stage*. Introducing via a *C* program the word or words that compose the node, the program with this method is able to locate if the modifier is affecting to the antecedent or to the consequence-as shown below:

```

>Line number:17914 If , for some reason , surgery is not an option with early
stage 1 and 2 , treatment is usually radiation therapy .
Modificator found: early , contex: with early stage 1 and
Modificador found: usually , contex: is usually radiation therapy

```

Figure 6. Detection of the time modifier in a sentence.

Once the node affected by the time modifier is located the constraint is drawn in a time line. Thus, in the sentence “If you stop smoking before a cancer develops, your damaged lung tissue gradually starts to repair itself.” the cause must happen ‘before’ the cancer begins to develop. So, the time constraint is associated to the cause and the restriction is stop smoking ‘before’ if the effect wants to be achieved. Figure 7 illustrates that:

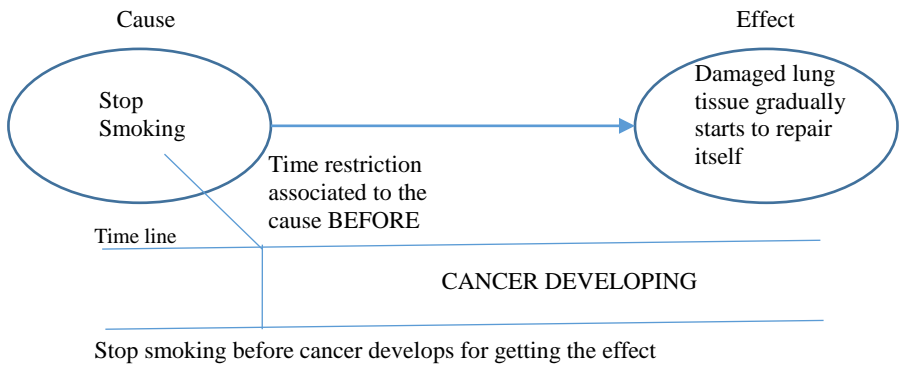


Figure 7. Cause depending on the ‘before’ crisp time constraint in the time line.

Other type of time modifiers exist: they are fuzzy time constraints, as fuzzy intervals. As previously quoted, fuzzy intervals are represented by fuzzy membership functions, as triangular or trapezoidal ones. For instance, in the sentence “*Some lung cancers are found early by accident as a result of tests for other medical conditions.*” the modifier *early* points to the effect and denotes a fuzzy time, drawn by a decreasing line illustrating that ‘early’ reaches poor values as it moves away from the beginning, as figure 8 pictures:

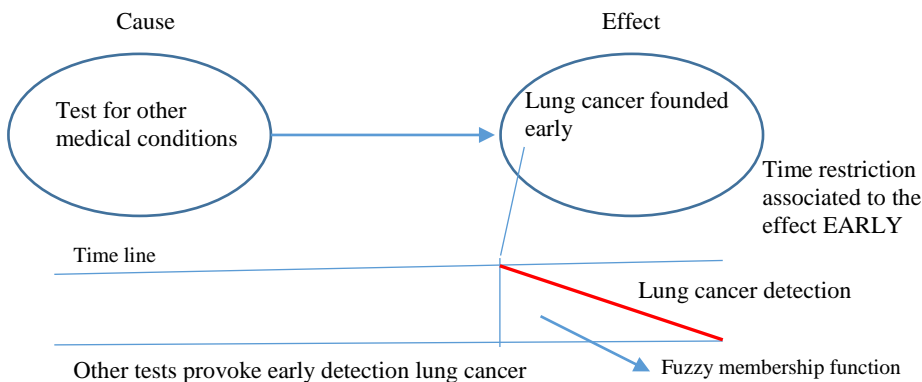


Figure 8. Effect depending on the ‘early’ fuzzy time constraint in the time line.



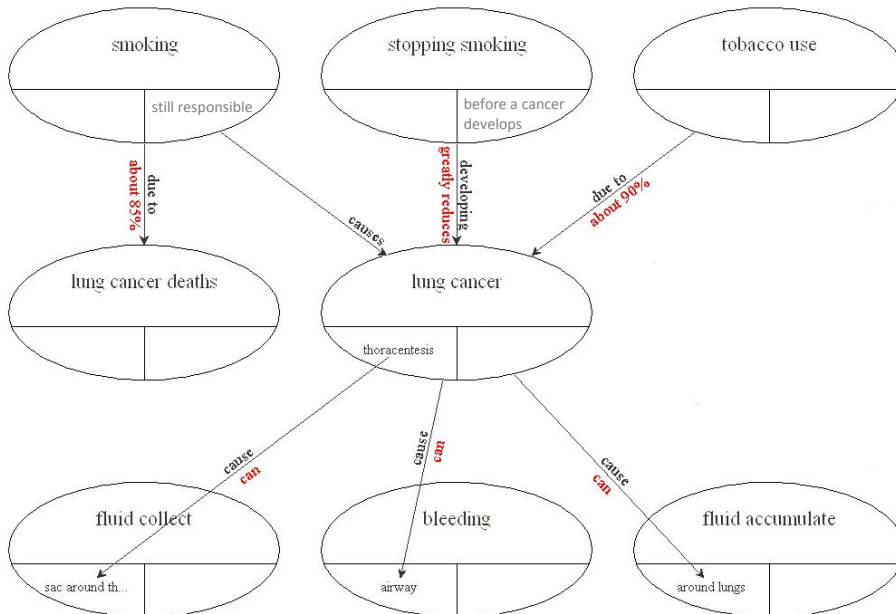
So, with these schemas we analyze each time modifier of the sentence to evaluate which of the classifications is correct.

The next step is to modify the original causal graph with this representation of the time. To do so, we performed a new search of causal sentences by means of another C program to get all the sentences related to a node with time modifiers. For instance, the sentences related to smoking, are shown in table 8.

**Table 8.** Sentences related to the node 'Smoking'.

<p>SMOKING</p> <p>If you stop smoking <b>before</b> a cancer develops your damaged lung tissue gradually starts to repair itself</p> <p>Although decades have passed since the link between smoking and lung cancers became clear smoking is <b>still</b> responsible for most lung cancer deaths</p> <p>Even if you have already been diagnosed with lung cancer there are <b>still</b> benefits to quitting smoking</p>
---

The encountered time modifiers related to smoking are *still* and *before*, but it seems that *before* has a stronger perception than *still*, that is why in the node, will appear the representation of *before* as a time line restriction. To obtain the time related causal graph, we have followed the same process, giving as result the graph presented in figure 9, where we can observe that some nodes have a temporal representation according with the sentences found related to a concept that had a time restriction. In the case that a node had sentences related with different time modifiers, we have prevailed the representation with the strongest one as in the case described. To complete the graph, we have used the bottom-right space to indicate, if so, the time constraint, or the word/words denoting it:



**Figure 9.** Causal graph obtained with temporal restrictions.

## 5 Conclusions and Future Work

In this paper we showed the relevance of causal sentences influenced for time constraints in the area of medicine. We approached several programs in order to automatically mine those sentences in medical texts. We showed how to get a causal graph from causal sentences, as causality is usually depicted as nodes related by a directed arrow. Finally, temporal restrictions are added to the graph with a time line, representing crisp time restrictions (points or intervals) and fuzzy time restrictions (fuzzy points or fuzzy intervals).

Future work should address the improvement of temporal lexicon classification in order to better tune the restrictions endorsed to the cause or the effect and to represent the temporary constraints in a more compact way, perhaps as a part of the causal graph.

**Acknowledgments.** Work supported by the Spanish Ministry for Economy and Innovation and by the European Regional Development Fund (ERDF/FEDER) under grant TIN2014-56633-C3-1-R.

## References

1. Frisch, M., ‘Time and causation’, in *The Blackwell Companion to the Philosophy of Time*, H. Dyke and A. Bardom (eds.), Wiley-Blackwell, 2013, 282-300.
2. Zhou, L. & G, Hripcsak, Temporal reasoning with medical data—A review with emphasis on medical natural language processing. *Journal of Biomedical Informatics* 40 (2007) 183–202.
3. Shahar Y. & Combi C. Timing is everything. Time-oriented clinical information systems. *West Journal Medicine* 1998;168(2):105–13.
4. Taylor, C., A formal logical analysis of causal relations. DPhil. Thesis. University of Sussex, Brighton, UK, (1993).
5. Puente, C.; A. Sobrino; J. A. Olivas & R. Merlo (2010), ‘Extraction, Analysis and Representation of Imperfect Conditional and Causal Sentences by means of a Semi-Automatic Process’, *Proceedings FUZZ-IEEE 2010*, pp. 1423-1430, 19-23 July, Barcelona.
6. Allen, J. F., ‘Maintaining knowledge about temporal intervals’, *Communications of the ACM*, 26, 1983, 832-843.
7. Allen, J., Towards a general theory of action and time. *Artificial Intelligence*, 23(2), 123–154, (1984).
8. Dubois, D. et al., ‘Fuzziness and Uncertainty in Temporal Reasoning’, *Journal of Universal Computer Science* 9(9), 1168-1194, 2003.
9. Sobrino A., Puente, C., Olivas J. A., “Extracting Answers from causal mechanisms in a medical document”. *Neurocomputing* 135 (2014) 53–60. ISSN: 0925-2312.
10. Schmid H., Probabilistic Part-of-Speech Tagging Using Decision Trees. *Proceedings of International Conference on New Methods in Language Processing*, 1994.