XRCE-T: XIP temporal module for TempEval campaign

Caroline Hagège
XEROX Research Centre Europe
6, chemin de Maupertuis
38240 MEYLAN, FRANCE
Caroline.Hagege@xrce.xerox.com

Xavier Tannier
XEROX Research Centre Europe
6, chemin de Maupertuis
38240 MEYLAN, FRANCE
Xavier.Tannier@xrce.xerox.com

Abstract

We present the system we used for the TempEval competition. This system relies on a deep syntactic analyzer that has been extended for the treatment of temporal expressions, thus making temporal processing a complement to a better general purpose text understanding system.

1 General presentation and system overview

Although interest in temporal and aspectual phenomena is not new in NLP and AI, temporal processing of real texts is a topic that has been of growing interest in the last years (Mani et al. 2005).

The work we have done concerning temporal processing of texts is part of a more general process in text understanding, integrated into a more generic tool.

In this article, we present briefly our general purpose analyzer XIP and explain how we perform our three-level temporal processing. TempEval experiments of our system are finally described and results we obtained are discussed.

1.1 XIP – a general purpose deep syntactic analyzer

Our temporal processor, called XTM, is an extension of XIP (Xerox Incremental Parser (Aït Mokhtar et al., 2002). XIP extracts basic grammatical relations and also thematic roles in the form of dependency links. See (Brun and Hagège 2003) for details on deep linguistic processing using XIP.
XIP is rule-based and its architecture can roughly be divided into the three following parts:

- A pre-processing stage handling tokenization, morphological analysis and POS tagging.
- A surface syntactic analysis stage consisting in chunking the input and dealing with Named Entity Recognition (NER).
- A deep syntactic analysis

1.2 Intertwining temporal processing and linguistic processing

The underlying idea is that temporal processing is one of the necessary steps in a more general task of text understanding. All temporal processing at the sentence level is performed together with other tasks of linguistic analysis. Association between temporal expressions and events is considered as a particular case of the more general task of attaching thematic roles to predicates (the TIME and DURATION roles). We will detail in sections 3.1 and 3.2 how low-level temporal processing is combined with the rest of the linguistic processing.

2 Three levels of temporal processing

Temporal processing has the following purposes:
1) Recognizing and interpreting temporal expressions, 2) Attaching these expressions to the corresponding events[1] they modify, 3) Ordering these events using a set of temporal expressions we present above.

We deliberately decided not to change our system's output in order to match TempEval gold-standard EVENTs and TIMEX3s. This would have

[1] We consider as events: verbs, deverbal nouns or any kind of non-deverbal nouns from a pre-defined list (e.g.: "sunrise" or "war").

Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval-2007), pages 492–495,
Prague, June 2007. ©2007 Association for Computational Linguistics

implied to change our parser's behavior. As linking events and temporal expressions is only a part of a general syntactico-semantic process, changing this part would have had bad consequences for the other aspects of the parsing.

2.1 Local level

Recognition of temporal expressions is performed by local rules that can make use of left and/or right context. Together with contextual rules, some actions are associated. These actions are meant to attribute a value to the resulting temporal expression. Figure 1 illustrates this stage for a simple anchor date. An ADV (adverbial) node with associated Boolean features is built from linguistic expressions such as "4 years ago". Note that there is a call to a Python function (Roux, 2006) "merge_anchor_and_dur" whose parameters are three linguistic nodes (`#0` represents the resulting left-hand expression). The representation of the values is close to TimeML format (Saurí et al, 2006).

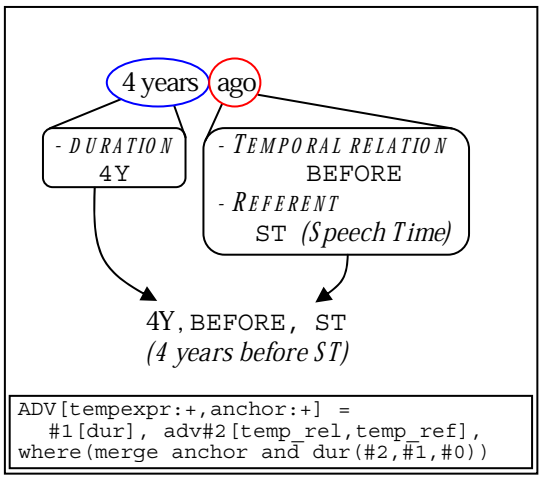

Figure 1: Local level processing, anchor date.

2.2 Sentence level

The sentence level is the place where some links between temporal expressions and the events they modify are established, as well as temporal relations between events in a same sentence.

Attaching temporal expressions to events

As a XIP grammar is developed in an incremental way, at a first stage, any prepositional phrase (PP, included temporal PP) is attached to the predicate it modifies through a very general MOD (modifier) dependency link. Then, in a later stage, these dependency links are refined considering the nature and the linguistic properties of the linked constituents.

In the case of temporal expressions, a specific relation TEMP links the temporal expression and the predicate it is attached to.

For instance, in the following sentence (extracted from trial data):

```
People began gathering in Abuja
Tuesday for the two day rally.
```

The following dependencies are extracted

```
TEMP(began, Tuesday)
TEMP(rally, two day)
```

"Tuesday" is recognized as a date and "two day" as a duration.

Temporal relations between events in the same sentence

Using the temporal relations presented above, the system can detect in certain syntactic configurations if predicates in the sentence are temporally related and what kind of relations exist between them. When it is explicit in the text, a temporal distance between the two events is also calculated.

The following example illustrates these temporal dependencies:

```
This move comes a month after
Qantas suspended a number of
services.
```

In this sentence, the clause containing the verb "suspended" is embedded into the main clause headed by "comes". These two events have a temporal distance of one month, which is expressed by the expression "a month after". We obtain the following dependencies:

```
ORDER[before](suspended, comes)
DELTA(suspended, comes, a month)
```

Verbal tenses and aspect

Morphological analysis gives some information about tenses. But the final tense of a complex verbal chain is calculated considering not only morphological clues, but also aspectual information. Tenses of complex verbal chains may be underspecified when there is insufficient context.

For instance, for the chain "has been taken", we extract "take" as the semantic head of the verbal chain. The aspect is perfective and the tense of the auxiliary "has" is present.
From this information, we deduce that this form is either in present or in past. This is expressed the following way:

```
PRES-OR-PAST(taken).
```

2.3 Document level

Beyond sentence-level, the system is at the first stage of development. We are only able to complete relative dates when it refers to the document creation time, and to infer new relations with the help of composition rules, by saturating the graph of temporal relations (Muller and Tannier, 2004).

3 Adapting XTM to TempEval specifications

The TempEval track consists of three different tasks described in (Verhagen et al. 2007). TempEval guidelines present several differences with respect to our own methodology. These differences concern definitions of relations and events, as well as choices about linking.

3.1 TIMEX3 definition

TimeML definition of a temporal expression (TIMEX3) is slightly different from what we consider to be a temporal expression in XTM:

- First, we incorporate signals (in, at…) into temporal expressions boundaries. But, as TIMEX3s are provided in the test collection, a simple mapping is quite easy to perform.
- We also have a different tokenization for complex temporal expressions. This tokenization is based on syntactic and semantic properties of the whole expression.

For example, our criteria make that we consider "ten days ago yesterday" as a single temporal expression, while "during 10 days in December" should be split into "during 10 days" and "in December".

3.2 TIMEX3 linking

XTM does not handle temporal relations between events and durations. In our temporal model, an event can have duration. However, this is not represented by a temporal relation, but by an attribute of the event. Durations included in a larger temporal expression (like in "two days later") introduce an interval for the temporal relation: AFTER(A, B, interval: two days). Here again no temporal relation is attributed with respect to the duration.

Therefore, we had to adapt our system so that it is able to infer at least some relations between events and durations. We used two ways to do so:

- An event having an explicit duration attributed by XTM gets the relation OVERLAP with this duration.
- An event occurring, for example, "two days after another one" (resp. "two days before") gets the relation AFTER (resp. BEFORE) with this duration.

Other relations are found (or not) by composition rules.

3.3 TIMEX3 values

TempEval test collection provides a "value" attribute for each TIMEX3. However we did not use this value, because we wanted to obtain an evaluation as close as possible to a real world application. The only value we used was the given Document Creation Time.

3.4 EVENTs mapping

Event lists do not match either between TempEval corpus and our system analysis. Unfortunately, when a TempEval EVENT is not considered as an event by XTM, we did not find any successful way to map this EVENT to another event of the sentence.

3.5 Temporal relation mapping

The set of temporal relations we use is the following: AFTER, BEFORE, DURING, INCLUDES, OVERLAPS, IS_OVERLAPPED AND EQUALS.

This choice is explained in more details in (Muller and Tannier, 2004).

Obtaining TempEval relations from our own relations is straightforward: AFTER and BEFORE are kept just as they are. The other relations or disjunctions of these relations are turned into OVERLAP. Disjunctions of relations containing AFTER (resp. BEFORE) and OVERLAP-like relations are turned into OVERLAP-OR-AFTER (resp. BEFORE-OR-OVERLAP).

4 Results

The trial, training and test sets of document provided were all subsets of the annotated TimeBank corpus. For each task, two metrics are used, the strict measure and the relaxed measure (see also (Muller and Tannier, 2004)).

Our rule-based analyzer is designed to favor precision. As our system is intended for use in information extraction, finding correct relations is more important than finding a large number of relations. That is why, at least for tasks A and B, we do not assign a temporal relation when the parser does not find any link. For the same reason, in our opinion, the strict measure is not as valuable as the relaxed one. We would argue that it does not really make sense to use a strict metric in combination with disjunctive relations.

Tasks A and B were evaluated together. We obtained the best precision for relaxed matching (0.79), but with a low recall (respectively 0.50). Strict matching is not very different. Another interesting figure is that less than 10% of the relations are totally incorrect (e.g.: BEFORE instead of AFTER). As we said, this was our main aim.

Note that if we choose a default behavior (OVERLAP for task A, BEFORE for task B, which are respectively the most frequent relations) for every undefined relation, we obtain precision and recall of 0.69, which is lower than but not far from the best team results.

Task C was more exploratory. Even more than for task AB, the fact that we chose not to use the provided TIMEX3 values makes the problem harder. Our gross results are quite low. We used a default OVERLAP for each unfound relation[2] and finally got equal precision and recall of 0.57.

However, assigning OVERLAP to all 258 links led to precision and recall of 0.508; no team managed to bring a satisfying trade-off in this task.

5 Conclusion

We described in this paper the system that we adapted in order to participate to TempEval 2007 evaluation campaign. We obtained a good precision score and a very low rate of incorrect relations, which makes the tool robust enough for information extraction applications. Errors and low recall are mostly due to parsing errors or underspecification and to the fact that we gave priority to our own theoretical choices concerning event and temporal expression definitions and event-temporal expression linking.

References

James Allen, 1984. Toward a general theory of action and time. *Artificial Intelligence*, 23:123-154.

Salah Aït-Mokhtar, Jean-Pierre Chanod and Claude Roux. 2002. *Robustness beyond Shallowness: Incremental Deep Parsing*. Natural Language Engineering, 8:121-144

Caroline Brun and Caroline Hagege, 2003. *Normalization and Paraphrasing using Symbolic Methods*, 2nd Workshop on Paraphrasing, ACL 2003.

Inderjeet Mani, James Pustejovsky and Robert Gaizauskas (ed.) 2005. *The Language of Time A reader.*

Philippe Muller and Xavier Tannier 2004. *Annotating and measuring temporal relations in texts*. In Proceedings of COLING 2004.

James Pustejovsky, Patrick Hanks, Roser Saurí, Andrew See, Robert Gaizauskas, Andrea Setzer and Beth Sundheim. 2003. The TIMEBANK Corpus. *Corpus Linguistics*. Lancaster, U.K.

Claude Roux. 2006. *Coupling a linguistic formalism and a script language*. CSLP-06, Coling-ACL.

Roser Saurí, Jessica Littman, Bob Knippen, Robert Gaizauskas, Andrea Setzer and James Pustejovsky. TimeML Annotation Guidelines. 2006.

Marc Verhagen, Robert Gaizauskas, Frank Schilder, Mark Hepple, Graham Katz and James Pustejovsky. 2007. *SemEval-2007 - Task 15: TempEval Temporal Relation Identification*. SemEval workshop in ACL 2007.

[2] The OVERLAP relation is the most frequent for task C training data.