Robust Temporal Processing: from Model to System

Tommaso Caselli and Irina Prodanof

ILC-CNR, Pisa
firstName.secondName@ilc.cnr.it

Abstract. This paper shows the functioning and the general architecture of an empirically-based model for robust temporal processing of text/discourse. The starting point for this work has been the understanding of how humans process and recognize temporal relations. The empirical results show that the different salience of the linguistic and commonsense knowledge sources of information calls for specific computational components and procedures to deal with them.

1 Introduction

Temporal processing of text/discourse has recently become one of the most active areas in NLP, boosted by the presence of specific markup languages (ISO-TimeML, SemAF/Time Project) and by a growing number of initiatives (CLEF, TERN, SemEval-TempEval2).

Natural languages have a variety of devices to communicate information about events and their temporal organization and the identification of the temporal relations in a text/discourse is not a trivial task. Previous research has explored and analyzed what sources of information are at play when inferring the temporal orders of eventualities such as tense, temporal adverbs, signals, viewpoint aspect, lexical aspect, discourse relations, commonsense and pragmatic knowledge. Most sources of information for inferring a temporal relation very rarely code in an explict and clear-cut way the specific temporal relation holding between two entities (i.e. eventuality - eventuality, eventuality - temporal expression) and this may lead to biases and incorrect tagging. Substantial linguistic processing is required for a system to perform temporal inferences and commonsense knowledge can hardly be encoded in domain independent programs. One of the main issues which has not been answered so far is how the linguistic devices which languages have at disposal to codify temporal relations

© A. Gelbukh (Ed.)
Special issue: Natural Language Processing and its Applications.
Research in Computing Science 46, 2010, pp. 29-40

Received 23/11/09
Accepted 16/01/10
Final version 09/03/10

interact both with each other and under which conditions they are autonomous, i.e. able to codify a temporal relations between eventualities without the support of non-purely linguistic elements, like discourse structure or world-knowledge based inferences. This calls for the development of procedures and techniques which maximize the role of the sources of information and the conditions under which they are necessary and sufficient to determine the current temporal relation.

This paper presents a general empirically-based model for robust temporal processing of text/discourse. Though the experiments have been conducted on Italian, the model is language independent. The lack of a complete system is mainly due the absence of temporally annotated resources for Italian over which systems could be developed and evaluated. The remaining of the paper is organized as follows: in sect. 2 we illustrate the methodology and the experimental results on the basis of which the model has been developed. Section 3 reports the overall architecture of the model and the functioning of its core components. Finally, sect. 4 presents the conclusion and observations for future work.

2 Linguistic Information and Pragmatic Mechanisms: Defining an Order of Application

Recent psychological studies ([1], [2]) have established correspondences between the formal aspect of the temporal structure of discourse and the mental representations interpreters built. The order in which eventualities are presented in a text/discourse *vs.* their real chronological order, the use of particular tenses, the presence of elements which explicitly mark a temporal relations are all features used when constructing a mental model of a text/discourse. As [3] pointed out, these features are of the kind which can be constructed automatically by information extraction systems. Knowing how these features interact with respect to their different nature, i.e. linguistic *vs.* world-knowledge based, is a necessary step to have robust automatic extraction systems.

To develop a model for temporal processing we have decided to investigate through an experimental study if it is possible to determine a hierarchical order of application of the linguistic and non-linguistic sources of information and under which conditions purely linguistic information is necessary and sufficient to determine the temporal relations between the entities in analysis. The aim of this study is that of identifying how deep must the computation of information go, that is how many modules must be activated in order to obtain a reliable temporal representation of the text/discourse.

2.1 Methodology

In order to verify the existence of a salience order of the sources of information and to obtain cues on the way the model should be implemented we have elaborated a test which was submitted to two groups of subjects: a first group of 29 subjects, none of them having knowledge in linguistics (Group 1), and a second group of 6 subjects, all MA students in Linguistics. The two groups were submitted with comparable, though not identical, test data, provided their different backgrounds and the level of metalinguistic analysis required.

In both experiments the subjects were presented with a set of 52 discourse excerpts, automatically extracted from the Italian Syntactic Semantic Treebank (ISST), and were asked to temporally order two highlighted eventualities in the discourse segments. To improve the reliability and avoid inconsistency, the subjects were asked to choose the temporal relations among a restricted set of 5 predetermined values, namely BEFORE, AFTER, SIMULTANEOUS, OVERLAP, and NO TEMPORAL RELATION. No binary interpretation of the temporal relations was allowed. In order to discover the existence of a salience order of application of the sources of information and to determine in a reliable way under which conditions linguistic (grammatical and lexical) information is autonomous (i.e. necessary and sufficient) for the identification of the temporal relations with respect to non-purely linguistic i.e. con-textual one, the subjects were asked to state what source of information had *mostly* helped them in the identification of the temporal relation. Similarly to the first task and to keep the experiments under control, we provided the subjects with a predetermined set of possible answers according to their background. Group 1 had at disposal the values TENSE, TEMPORAL EXPRESSIONS, and NOT SPECIFIED, while Group 2 had a larger set, i.e. TENSE, TEMPORAL EXPRESSIONS, SIGNAL, ASPECT, SEMANTICS and NOT SPECIFIED.

2.2 Data Analysis and Results

In Table 1 we report the results obatined for the identification of temporal relations. The agreement among the subjects have been computed by means of the K statistics.

Table 1. Agreement of the subjects on temporal relation identification

Agreement of temporal relations	K value
Overall agreement	0.58
Agreement in presence of temporal expressions	0.64
Agreement in presence of signals	0.73
Agreement in presence of shifts in tense	0.70

As the results illustrate, the identification of temporal relations is a challenging task. Only in presence of specific *markers* of temporal relations, such as shifts in tense, temporal expressions and signals, the agreement raises up to reliable values. The data have also shown that the temporal representations humans construct are varied: in absence of specific information they are mainly coarse grained, while unique and clear-cut values can be obtained only in presence of explicit information, i.e. of markers which guides the interpretation process.

One of the main results is the identification of a set of *constraints* and *preferences*. The constraints apply both to the role and to the relationships among the sources of information, while the preferences deal with tense patterns and associated temporal relations. Different constraints are activated for each source of information. The constraints can be conceived as representing the condition under which each source of information is a necessary and sufficient element for recovering the correct temporal relation. In particular, we claim that:

- **Constraint 1 (Tense)**: sequences of adjacent eventualities must have different tense meaning, otherwise other sources of information are responsible for their ordering;

- **Constraint 2 (Temporal Expressions)**: temporal expressions may represent explicit information for the ordering of eventualities provided that: a.) they are related to the moment of the event, E, or to a secondary deictic moment, R; b.) if more than one temporal expressions is present, they must stand in an anchoring relation one with each other to signal a temporal relation; c.) in case there is just one temporal expression related to an eventuality, its contribution is relevant for the anchoring of eventuality on the time line but it is null for the determing the temporal relation between two eventualities;
- **Constraint 3 (Signals)**: signals represent salient information for temporal relations only when their semantics is explicit, like for *dopo* [after], *intanto* [simultaneously]. When they are implicit, like *quando* [when], *per* [for/since], they offer ancillary information which reinforces the contribution of other sources, like tense, viewpoint aspect, lexical aspect, temporal expression, and commonsense knowledge;
- **Constraint 4 (Viewpoint and Lexical Aspect)**: these two types of information have a constraint similar to that for tense: when adjacent eventualities have the same values, either for viewpoint or lexical aspect, their knowledge is necessary but not sufficient to determine the temporal relation.

When all other sources of information fail to provide distinguishing cues, commonsense knowledge is used. We claim that commonsense knowledge is the most salient source of information for recovering temporal relations but also the less affordable, since it may introduce biases and disagreement. The identification that a sequence of sentences forms a text/discourse is the pre-condition for the existence of any kind of relations between the discourse entities. It is only in this sense that temporal relations are a by-product of the computation of the general discourse structure, and, as the data have shown, discourse structure cannot be considered the primary source of information for the identification of temporal relations. Of course, knowledge of discourse structure can improve the automatic recognition of temporal relations, as [4] have demonstrated. To illustrate how the constraints work, consider this example:

(1) Marco è *caduto*$_{e1}$. Giovanni l'*ha spinto*$_{e2}$.
Marco fell$_{e1}$. *Giovanni pushed him*$_{e2}$.

In 1 the two eventualities cannot be ordered by exploiting only linguistic information. They are not compliant neither with **Constraint 1** (different tense meaning) nor with **Constraint 4** (different viewpoint and lexical aspect). **Constraints 2** and **3** do not apply since no temporal expression and signal is present. The only available source of information is the commonsense knowledge on the basis of which can infer that $_{e2}$ stands in a precedence relation with $_{e1}$. In case we had more specific information (about the context of occurrence), the temporal relation could be overriden.

The analysis of the correlation between tense patterns and temporal relations has suggested that it is possible to associate a preference order for temporal relations according to the combination of the tense forms. In particular, it appears from the data that certain tense forms when appearing in particular tense patterns, like the *trapassato I* [past perfect], tend to grammaticalize particular temporal relations, while others are more prone to code a larger set of relations, like the *passato composto* [present perfect or simple past]. The preferences have been introduced to reduce the possible temporal relations which may be computed. To clarify how preference rules work consider the following example which is an adapatation from example 1. In this reduced and simplified formalization, t represent the beginning or ending point of the eventutalities, S the moment of utterance and R, as already stated a possible secondary deictic moment/point necessary to describe the semantics of some tense forms, like the *trapassato I.*

(2) Marco è *caduto*$_{e1}$. Giovanni l'*aveva spinto*$_{e2}$.
Marco fell$_{e1}$. *Giovanni had pushed him*$_{e2}$.
- discourse sequence: *passato composto*$_{e1}$ - *trapassato I*$_{e2}$
- $e_1 = ((E_1 \prec S) \wedge (t_1 \leq E_1 \leq t_2))$ [tense analysis for $_{e1}$]
- $e_2 = ((E_2 \prec R_2) \wedge (R_2 \prec S) \wedge (t_3 \leq R_2 \leq t_4))$ [tense analysis for $_{e2}$]
- $(t_1 \prec t_3) \wedge (t_2 \prec t_4) \wedge (t_1 \prec t_4) \wedge (t_2??t_3))$
- possible temporal relations: $((E_2 \prec E_1) \vee (E_2\ m\ E_1) \vee (E_2\ o\ E_1))$

The final output does not provide a unique temporal relations due to the missing information between the ending point of e_1 and the beginning point of e_2, i.e. $(t_2??t_3)$. The application of the preference rule for sequences of *passato composto - trapassato I* states that the reliable temporal relation is that of precedence:

- possible temporal relations: $((E_2 \prec E_1) \vee (E_2\ m\ E_1) \vee (E_2\ o\ E_1))$
- Preference Rule: `if the sequence is` *passato composto - trapassato I*
  `then reduce the output to` $E_2 \prec E_1$
- final output: $(E_2 \prec E_1)$

It is important to point out that the preference rules do not apply for all tense patterns. For instance, with the *futuro composto* [future pefect] and the *futuro nel passato* [future-in-the-past], where the relationship between E and S cannot be reliably stated, no preference rules apply and the output of the component is obtained by disjunctive finely grained relations which can be rearranged in terms of coarse grained temporal relations.

Finally, it has been possible to formulate a saliency-based hierarchical order of application of the sources of information as reported in **Formula 1**. The symbol $\lessapprox$ stands for "in absence of more specific linguistic information, X is the most salient source of information" and $\lesssim$ for "in absence of more specific information, X is the most salient source of information". Notice that when stating "in absence of more specific (linguistic) information", we are referring to the constraints we have identified for the saliency of the sources of information:

Formula 1 (Hierarchical order of information) : COMMONSENSE KNOWLEDGE $\lessapprox$ (IMPLICIT SIGNALS $\lesssim$ TENSE $\lesssim$ VIEWPOINT ASPECT $\lesssim$ LEXICAL ASPECT $\lesssim$ TEMPORAL EXPRESSIONS $\lesssim$ EXPLICIT SIGNALS)

The saliency based hierarchy is an abstraction. A human interpreter always has at disposal all the sources of information. On the basis of the experimental data, we have deducted that the most probable order for processing the information is the one illustrated in the hierarchy since as soon as the subjects have found a reliable solution they should have blocked their inferencing processes. The behavior of the pragmatic, i.e. commonsense, knowledge seems to offer further support to this observation. In fact, this source was selected as the most salient only when all the others were "absent", i.e. when the constraints we have illustrated were not respected.

3 The Model Architecture

This section is devoted to the illustration of the general architecture and mechanisms of a computational model for automatically resolve temporal relations in a text/discourse. The model is based on the empirical data and results illustrated above. Its modular organization is proposed as a strategy to improve the reliability of the output and avoid failure. Each module has some specialized functions and components which are conceived to deal with a source of information at time. The modules are organized on a pipeline according to which the output of one module represents the input for the following. On the basis of the hierarchy we have illustrated by means of the **Formula 1** and as a general strategy, the specialized components of each module should be activated only when necessary. Figure 1 illustrates the overall workflow of the model, from raw input text to the final output.

The first module is responsible for two primary tasks: the identification, normalization and assignment of temporal relations between the temporal expressions by means of a Timex Grammar and Normalizer and the identification of the eventualities and the assignment of their default lexical aspect through an Event Detector component. The two components take in input shallow parsed text since the chunks' extent approximates the extent of temporal expressions and eventualities. Moreover, chunks can be easily combined together for items whose extent corresponds to more than one token.

The second module has three main components which compute three different types of information strictly connected with each other, namely tense, viewpoint aspect and lexical aspect. The main result of the analysis of this internal submodule is the formalization of each eventuality in its corresponding interval representation. The output of these three components is necessary for two elements:

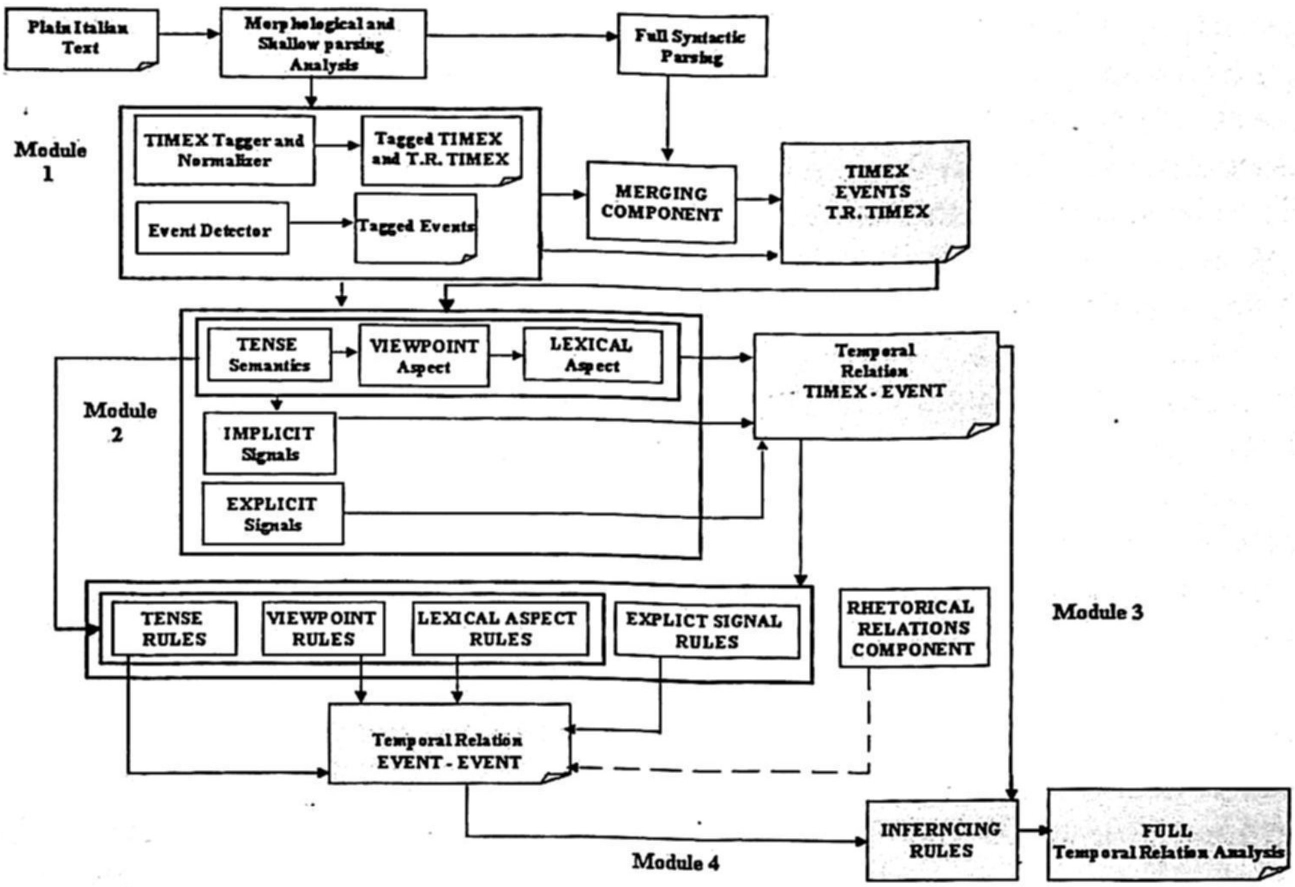

Fig. 1. Workflow of the Model.

firstly, it is used to to determine the temporal relations between temporal expressions and eventualities when associated to the output of the temporal expression component of module 1, and secondly, it is used to activate the components of module 3 only when they can provide a reliable output. The temporal relations which are assumed to be valid are all [5]'s 13 interval relations and [6]'s 8 instant-interval relations. The two signal components are activated when the connection between the eventuality and the temporal expression is "mediated" by a signal.

The third module is responsible for the identification of the temporal relations between eventualities. Each component has a set of heuristics based on the empirical data and provides as output the temporal relation value(s). The heuristics are divided into two main groups: one for complex sentence contexts and the others for adjacent eventualities in discourse segments. The four internal components are mutually exclusive one with respect to the other. A general principle, which results from the constraints illustrated in sect. 2, guides their functioning. Each component is activated if and only if a set of preconditions is respected, otherwise the temporal ordering is completely inferred by means of the external component (module 4). The module assumes as basic temporal relations [5]'s 13 interval relations and [6]'s 8 instant-interval relations. A special predicate, *hold*, is postulated to account for the measure of the duration of the eventualities. As a general principle, the temporal relations between two adjacent eventualities are computed by considering the relations between the beginning and ending points of their interval representations. However, much of this information is missing or only vaguely present in the text/discourse. In order to deal with this issue, the final outputs of the internal components can differ in terms of the preciseness of the temporal knowledge expressed so that we can have (i.) precise temporal knowledge, when a single temporal relation can be stated or (ii.) coarse grained temporal knowledge, when more than one temporal relation can be inferred. In presence of coarse grained knowledge, the multiple temporal relations do not represent contradicting temporal representations, but related or conceptually adjacent temporal relations. Instead of expressing these types of temporal relations by means of disjunctive finely grained relations, we have decided to use of coarse grained knowledge based on [7]'s notion of conceptual neighbors. The main advantage of such a representation is two-folded: on the one hand, the model is somehow cognitively similar to the temporal representations that humans may have, and, on the other hand, it avoids that the inferencing module could fail to complete the whole set of temporal relations. According to our analysis, we have instances of precise temporal relations only when the output is obtained by three components, namely (i.) tense, (ii.) explicit signals and (iii.) discourse relations. The viewpoint and lexical aspect components will produce coarse grained temporal knowledge.

Module 4 is responsible for the inferencing process of temporal relations. This module takes in input both the output of module 2 and that from module 3 and it activates two different types of inferencing mechanisms according to which the module provides its output. When Module 2 provides the input, the eventualities are already connected by means of a temporal relation to a temporal expression. In this case, module 4 activates a set of inferencing rules according to which the relations between the temporal expressions are transferred to their connected eventualities. Things are more complicate when the input comes from module 3. In this case, module 4 looks for couples of adjacent eventualities with one of them in common. Once identified, it will activate inferencing rules based on a transitivity table which preserves the insights of Allen's table and the coarse grained knowledge, as proposed by [7]. As for eventualities realized by nouns or other parts of speech different than verbs, our model implements a strategy based on [3]. Event nouns do not present information on their temporal location, thus the identification of the temporal relations requires an extended use of commonsense knowledge. Following [3]'s proposal, an abstract device, a Chronoscope, apply. This device allows temporal representation abstraction. The Chronoscope requires that we index temporal relations to a certain level of time granularity g. In our account, event nouns (and all parts-of-speech other than verbs) will be considered as simultaneous with the verbal eventualities with which they co-occur. This means that, the finely grained distinction between tensed eventualities will be maintained and preserved, and, at the same time, there is no need to make reference to commonsense knowledge in order to extend the model to event nouns. The only operation is to abstract their temporal representation on a level of same temporal granularity as that of the verbal eventualities.

4 Conclusion and Future Work

This paper illustrates a general architecture for robust temporal processing of text/discourse. The workflow of the model has been elaborated on the basis of experimetal data which have suggested a saliency-based order of application of the linguistic and non-linguistic sources of information involved in this task. With respect to previous research, this work has presented a first systematization of the linguistic devices involved in temporal processing and how they interact with each other. The development of a system compliant to the model has also advantages for the interpretation of errors and may facilitate their solution.

The actual relization of the model is ongoing. Some components (temporal expression tagger [8], an intrasentential tagger for temporal relations between eventualities and temporal expressions in presence of signals [9]) and procedures (use of lexical resources for the identification of eventualities) have been realized and evaluated on specifically annotated documents by using different techiniques (rule-based systems *vs.* machine learning techniques). As a side effect of this work is the realization of an annotated corpus for Italian for events, temporal expressions and temporal relations, comparable to the English TimeBank[10].

As future work, we plan to experiment the validity of the experimental results by developing classifiers which apply different orders of application of the linguistic information with respect to those illustrated in sect. 2. A further element of analysis will be the evaluation of the flow of information from one module to the other and how errors can influence their functioning.

References

1. van der Meer, E., Beyer, R., Heinze, B., Badel, I.: Temporal order relations in language comprehension. Journal of Experimental Psychology: Learning, Memory and Cognition **28** (2002) 770–79
2. Kelter, S., Kaup, B., Claus, B.: Representing a described sequence of events: A dynamic view of narrative comprehension. Journal of Experimental Psychology: Learning, Memory and Cognition **30** (2004) 451–64
3. Mani, I.: *Chronoscopes*: A theory of underspecified temporal representation. In Schilder, F., Katz, G., Pustejovsky, J., eds.: Annotating, Extracting and Reasoning about Time and Events. LNAI. Springer-Verlag, Berlin Heidelberg (2007) 127–39
4. Forascu, C., Pistol, I., Cristea, D.: Temporality in relation with discourse structure. In: Proceedings of the Fifth International conference on Language Resources and Evaluation (LREC-06). (2006) 65–70
5. Allen, J.: Maintainig knowledge about temporal intervals. Communications of ACM **26**(11) (1983) 832–43
6. Allen, J., Hayes, P.: Moments and points in an interval-based temporal logic. Computational Intelligence **5**(3) (1989) 225–38
7. Freska, C.: Temporal reasoning based on semi-intervals. Artificial Intelligence **54** (1992) 199–227
8. Caselli, T., dell'Orletta, F., Prodanof, I.: A timeML compliant timex tagger for italian. In: Proceedings of the International Multiconference on Computer Science and Information Technology – IMCSIT 2009. Volume 4. (2009) 185 – 192
9. Caselli, T., dell'Orletta, F., Prodanof, I.: Temporal relations with signals: the case of italian temporal prepositions. In Lutz, C., Raskin, J.F., eds.: 16th International Symposium on Temporal Representation and Reasoning, 2009. TIME 2009. (2009) 125 – 132
10. Pustejovsky, J., Hanks, P., Saurí, R., See, A., Gaizauskas, R., Setzer, A., Radev, D., Sundheim, B., Day, D., Ferro, L., Lazo, M.: The TIMEBANK corpus. In: Corpus Linguistics 2003. (2003)