Temporality in relation with discourse structure

Corina Forăscu[1,2], Ionuț Cristian Pistol[1], Dan Cristea[1,3]

[1] Faculty of Computer Science, University "Al. I. Cuza" of Iaşi, Romania
16, Berthelot St., 700483 – Iaşi, Romania
[2] Institute for Artificial Intelligence, Romanian Academy, Bucharest - Romania
[3] Institute for Computer Science, Romanian Academy, Iaşi - Romania
{corinfor,ipistol,dcristea}@info.uaic.ro

Abstract

Temporal relations between events and times are often difficult to discover, time-consuming and expensive. In this paper a corpus study is performed to derive a strong relation between discourse structure, as revealed by Veins theory, and the temporal links between entities, as addressed in the TimeML annotation standard. The data interpretation helps us gain insight on how Veins theory can improve the manual and even (semi-) automatic detection of temporal relations.

1. Introduction

Temporal relations are binary relations that connect temporal events or temporal expressions in texts. On the other hand, texts display discourse structures, as those expressed by Rhetorical Structure Theory - RST (Mann & Thompson, 1988). Different authors suggested that temporality should manifest an intimate correlation with the discourse structure (Asher & Lascarides, 2003).

1.1. Main objectives

This paper indicates a method that leads to a quantitative investigation of the connection between RST and the temporal relations in news texts and shows to what extent the link could be used for two purposes: reducing the human effort in manual annotation of temporal relations and improving the automatic annotation of temporal relations.

1.2. Related work

One of the first discourse models that show how different knowledge sources constrain the interpretation of temporal information in the utterances of a whole discourse is the Discourse Representation Theory. In (Kamp & Reyle, 1993) inferences about the temporal structure of discourse are correlated with the semantic information given by tense and aspect.

In a formal account of the pragmatic influences in ordering events in discourse (Lascarides & Asher, 1993), the interactions between temporal structure and discourse structure are investigated using both linguistic and world knowledge. Five discourse relations (*Narration, Explanation, Elaboration, Background* and *Result*) are used to interpret temporal aspects of discourse.

The relations between tense and temporal relations, on one hand, and a theory of anaphoric reference and discourse modelling, on the other, are used by Webber (1988) to treat the tensed clauses as anaphors. This framework is further extended by Song & Cohen (1991) by using tense interpretation to determine the temporal ordering between states and events in simple narratives. Their algorithm is evaluated on stories without indirect speech.

Hitzeman et al. (1995) describe a detailed representation of the temporal structure of discourse. Different event ordering possibilities are obtained by assigning weights to the various knowledge sources used in the system: temporal adverbials, discourse cues to rhetorical relations, tense and aspect are treated as mutually constraining, while the remaining ambiguities are handled through the semantic closeness of the utterances, instead of background world knowledge.

More recent work uses annotated corpora to investigate temporal phenomena in texts. Mani & Wilson (2000) describe a temporal annotation scheme for representing dates and times in temporal expressions, in an automatic tagger based on hand-crafted and machine-discovered rules, relying on lexical features. Using the tagged expressions, the verb-denoted events are aligned on a calendrical time.

In (Schilder & Habel, 2001) the temporal expressions (time-denoting and event-denoting) and the temporal relations between them are located in German newswire articles, by using a semantic tagging system for temporal expressions. The implemented system extracts temporal relation information from sentences where the seven temporal relations are explicitly signaled by prepositions, adverbials, nominal modifiers and elliptical relative clauses.

Another approach (Filatova & Hovy, 2001) assigns a time stamp (either time-points or time-intervals) to every verb-denoted event. The rule-based system takes as input a set of news stories broken into separate sentences and produces as output a text that combines all the events from all the articles, organized in chronological order.

In (Katz & Arosio, 2001), through an intrasentential annotation, the temporal relations between events (only verbs) are detected within each sentence from a multilingual corpus.

Setzer (2001) introduces STAG, a fine-grained annotation scheme that captures anchorable events, temporal expressions (points or intervals when events happen), and explicit and implicit temporal relations between events or between events and times. The annotation scheme has been evaluated through the construction of a trial news corpus.

The TIMEX2[1] scheme, used to annotate time expressions according to a canonical representation, is applied to the Automatic Content Extraction (ACE[2]) evaluation, Relation Detection and Characterization [RDC] task, whose goal is to detect and characterize Entity-Entity and Event-Entity relations according to specified types.

Even if no consensus is yet reached with regard to the modalities to annotate temporal information in text, a collaborative effort leaded to TimeML[3], a metadata standard for marking events, their temporal anchoring and links in news articles. The TERQAS[4] workshop, which leaded to TimeML, has as another major deliverable the TimeBank 1.1 corpus[5] (Pustejovky et al., 2003), one of the corpora used also for the investigations in this paper.

The TARSQI project[6], started in 2005, intends to develop a technology for annotating temporal information in natural language text, extracting temporal information from text, and reasoning about temporal information. The 5 components of TARSQI, a modular system for automatic temporal annotation that adds time expressions, events and temporal relations to news texts is already created and partially evaluated. The temporal links are added through GUTenLINK tagger that uses hand-developed syntactic and lexical rules. SputLink is a temporal closure component that takes known temporal relations in a text and derives news implied relations from them.

1.3. Motivation

Many NLP applications, such as topic detection and tracking, question answering (questions like "when", "how often" or "how long"), information extraction or information retrieval, machine translation, summarization, would benefit significantly from the ability to accurately position detected events in time, either relatively with respect to other events or on an absolute time axis.

As the detection of temporal relations is known to be very time consuming and expensive (Pustejovky et al., 2002), and, in manual work, even far from being complete, i.e. about 4% in an initial manual annotation phase (Setzer et al., 2005), the use of discourse theories and their adequate annotations can facilitate the uncovering of these temporal links.

The Veins Theory - VT (Cristea et al., 1998; Cristea et al., 2000), a model of global discourse interpretation that takes from RST its notions of nuclearity and relations, reveals a "hidden" structure in the discourse tree, called *vein*, which permits to determine, for each discourse unit, a sub-span of the whole discourse in which all anaphors belonging to that discourse unit are allowed to find an antecedent. The vein expression of a discourse unit gives the minimal span necessary to understand that particular unit in the context of the whole discourse. Veins can be used to trace antecedents of referential expressions or to obtain summaries of the text focussed on particular discourse entities (Cristea et al., 2005).

In a recent work, Mani et al. (2005) show that a discourse parser built using, for example, the rhetorical structure (Marcu, 1997) can improve the events ordering rules by reasoning based on RST.

The obvious claim investigated in this paper is that a similar relationship as the one between discourse structure and referentiality should also exists between discourse structure and temporality.

To our knowledge, no systematic study of the relation between RST or VT and temporal links in discourse has yet been carried out.

1.4. Paper presentation

Section 2 gives an overview of Veins Theory and of the temporal information in natural language, as encoded through the TimeML standard. Section 3 describes the method used to investigate the relationship between discourse structure and temporality. An evaluation of the initial claim is performed in section 4, while in the last two sections the results are discussed, and conclusions and limitations are formulated.

2. TIME along with Veins

2.1. Veins

The fundamental intuition underlying the unified account on discourse structure and accessibility in VT is that the RST-specific distinction between nuclei and satellites constrains the range of referents to which anaphors can be resolved; in other words, the nucleus-satellite distinction, superimposed over a tree-like structure of discourse, induces for each anaphor a *domain of referential accessibility* (DRA). More precisely, for each anaphor *x* in a discourse unit *u*, VT hypothesizes that *x* can be resolved by examining discourse entities from a subset of the discourse units that precede *u*.

The discourse structure assumptions in VT are, to a great extent, the same as in RST:

- the basic units of a discourse are non-overlapping spans of text: usually a dot-to-dot sentence, but they may be sometimes reduced to clauses;
- discourse structures are represented as binary trees in which internal nodes are rhetorical relations and leaf nodes are *elementary discourse units* (*edus*);
- unlike RST, VT ignores relation names, but retains the notion of nuclearity: among the daughters of a relation, at least one node, called *nucleus,* is considered essential for the writer's purpose, while non-nuclear nodes, called *satellites*, include spans of text that increase understanding but are not essential to the writer's purpose.

Vein expressions defined over a discourse tree are sub-sequences of the sequence of units making up the discourse. The vein expression of a node intends to reflect the sequence of units that are significant to understand the span of text covered by the node, in the context of the whole discourse.

The part of the theory relating referentiality to the structure of discourse formulates two claims, empirically

1 http://timex2.mitre.org/
2 http://www.nist.gov/speech/tests/ace/index.htm
3 http://www.timeml.org/site/index.html
4 http://www.timeml.org/site/terqas/index.html
5 Since February 2006 the 183 news articles form TimeBank 1.2 corpus can be browsed at http://www.timeml.org/site/timebank/browser_1.2/; it encodes the last version of TimeML, v. 1.2.1, October 2005
6 http://www.timeml.org/site/tarsqi/index.html

proved on annotated corpora (Cristea et al., 2000): that the span of text given by the vein expression of a discourse unit is sufficient to recuperate the antecedents anchored in the referential expressions on that unit, and that the effort to recuperate these antecedents is always less when going on veins. Till now, there has been made no tentative to verify whether the relationship between discourse structure and referentiality does not extend also in the temporality domain.

2.2. Annotated temporal information

The TimeML 1.1. specification language is intended to use for automatically extracting information about the event-structure of narrative texts, especially of English news. The language consists of a collection of tags inserted into a text, intended to make explicit information about the events reported in the text and their temporal relations.

The metadata standard marks:

- Events through the tags:
  - EVENT - indicates situations that happen or occur, states or circumstances in which something obtains or holds true: *We are waiting for him.* There are 7 possible classes of EVENTs: OCCURRENCE, PERCEPTION, REPORTING, ASPECTUAL, STATE, I_STATE, I_ACTION.
  - MAKEINSTANCE: it marks how many different instances or realizations have a given event; the tag also carries the tense and aspect of the verb-denoted event: *John learns **twice** on Monday.*
- temporal anchoring of events through:
  - TIMEX3 - for explicit temporal expressions: times of a day, dates – calendar dates or ranges, durations: *Tuesday*; *15 November 2005*; *two days*.
  - SIGNAL - function words that indicate how temporal objects are to be related to each other.
- links between events and/or timexes:
  - TLINK – indicates 13 types of temporal relations (analogous to Allen's relations) between two temporal elements (event-event, event-timex): BEFORE, AFTER, INCLUDES, IS_INCLUDED, DURING, SIMULTANEOUS, IAFTER, IBEFORE, IDENTITY, BEGINS, ENDS, BEGUN_BY, ENDED_BY.
  - ALINK, Aspectual Link - the relationship (Initiation, Culmination, Termination, Continuation) between an aspectual event and its argument event.
  - SLINK - Subordination Link (of type Modal, Factive, Evidential, Negative) for contexts introducing relations between two events.

The TimeBank 1.1 corpus[7] is a set of 186 English news report documents annotated with the 1.1 version of the TimeML standard for temporal annotation. It is considered to be preliminary, as it still needs improvements and reviews.

[7] Available for academic and research purposes at http://nrrc.mitre.org/NRRC/Docs_Data/MPQA_04/approval_time.htm

3. Method

The experimental data used in this study consists of ten newspaper texts, drawn from the Wall Street Journal – WSJ, from the Rhetorical Structure Discourse Treebank (Carlson et al., 2003) which have been found to be also annotated for temporal expressions, events and relations, in the TimeBank 1.1. corpus. Some basic statistics performed on the corpora are illustrated in the next figures.

	TimeBank 1.1	RST Discoourse Treebank
# words	2548	2333
# sentences / *EDUs*	120	271
# EVENTs	334	-
# TIMEX3	65	-
# TLINKs	256	-

Figure 1. Basic statistics on corpora

The difference between the total number of words in the two sub-corpora used in the study appears because the RST Treebank does not include the header capturing document format, structure information and the date of document's creation. In the rest of this section, the processing steps are described.

First, the original Lisp-like notation of the RST structure in the RST Discourse Treebank is automatically translated onto an XML notation. Following (Cristea et al., 1998), to each *edu*, its vein expression is computed and added (Pistol, 2005).

The original RST+veins annotation is then fused with the TimeML annotation. The resulted information includes the RST structure, vein structure, and temporal structure. These merged files are used to determine quantitatively the relation between discourse structure and time, and to extract the examples (positive and negative) which will be further discussed.

4. Evaluation

Four types of evaluations are performed:

a) relations-on-the-veins evaluation: for each temporal relation, we count whether or not the second term (a time expression or event) belongs to a unit included on the vein expression of the unit the first term belongs to.

This has produced an accuracy of 45.31% computed as the report between the number of temporal relations found on the vein expression over the total number of temporal relations. One explanation of this rather low score is that in the TimeML files there has been added a "creation date" at the beginning of each file, as a reference for temporal expressions in that file. This time anchor does not exist in the RST annotation, and, as such, it is not included in any computed vein. If we disregard these errors, an accuracy of 62.03% is obtained. Previously, Song & Cohen (1991) and Cristea et al. (2000) have shown that indirect speech acts do not correlate with the discourse structure, hence neither with the computed veins. If they are disregarded, the accuracy goes up to 77.85%. The difference of more than 15% shows that indirect speech acts do have a significant impact on the performance, and a different approach is required in order

to improve the identification of temporal relations involving indirect speech.

b) annotation effort evaluation: for each time expression we compute how many units back the second term is located, if computed on the vein versus computed linearly. This way we have an indication of the effort saved in manual annotation of temporal links if the annotator would use the vein of the unit the temporal expression is located on, as an abridged version of the whole original text.

In the evaluation corpora we have found that 67.18% from the temporal relations link events in the same or adjacent elementary discourse units. From the rest of 32.82%, which are more relevant as providing a more difficult challenge to the human annotator, 27.38% can be detected quicker following the computed veins than following the text order.

c) To what extend different types of temporal relations behave in connection with the discourse structure? We have performed the evaluation reported at (a) for each type of temporal relation and have sorted the resulted values in descendent order.

Tlink type	total	found
SIMULTANEOUS	28	17
BEFORE	71	23
AFTER	26	11
IS_INCLUDED	51	28
ENDED_BY	19	5
INCLUDES	19	8
IDENTITY	27	14
ENDS	1	1
BEGUN BY	5	4
IAFTER	1	1
IBEFORE	0	0
BEGINS	3	3
DURING	5	1

Figure 2. TLINKs on veins found on corpora

The numbers in Figure 2. show that the best results are obtained for the SIMULTANEOUS, IS_INCLUDED and IDENTITY relations.

d) how temporal relations behave in connection with different types of rhetorical relations? Among the temporal relations that fail to find antecedents along veins we investigate the types of rhetorical relations involved.

The most frequent rhetorical relations that do not connect with the temporal relations are ATTRIBUTION (30%) and SPAN (32%). ATTRIBUTION involves elementary discourse units with verbs denoting an indirect speech act, so this comes as further proof that indirect speech is not correlated with the discourse structure. SPAN is a rhetorical relation found between two adjacent discourse units, and it is the most common rhetorical relation appearing in the RST corpus.

5. Discussions

The evaluation shows that the expectations are fulfilled to a great extend, in the sense that temporal links are found to be mainly located on the veins of the discourse structure. It also shows that there is a subset of temporal relations that behaves better than the rest in correlation with the discourse structure. Mainly the bad performance characterises relations involving events indicating indirect speech. It seems that speech act events do not actually correlate with the discourse structure. This is proved reciprocally: the class of verbs involving speech acts anchors temporal relations addressing co-terms outside veins, and the failings to find co-terms in time relations on the veins happen mainly around the ATTRIBUTION rhetorical relation, which is the one involving speech acts.

The study evidenced a method to quicken the manual annotation of temporal relations in texts which include already a discourse structure annotation, known to be very time consuming and expensive (Pustejovky et al., 2002): an interface could be build to hide to the annotator the whole text excepting the units belonging to the vein of the unit containing the time expression s/he is concentrated upon. Automatic annotation of temporal relations could also benefit provided it is performed in correlation with the discourse parsing.

By mimicking the use of such an interface, some TLINKs not included in TimeBank were discovered. The following example is from the document wsj_0610. Suppose the current *edu* is the one with ID 55. Its vein expression, illustrated below, is `8, 42, 49, 54, 55`:

> 8>>> The oat-bran craze[e190] has cost[e189] the world's largest cereal maker market share.
> 42>>> much of Kellogg's erosion[e204] has been in such core brands as Corn Flakes, Rice Krispies and Frosted Flakes,
> 49>>> that it soon[t207] will begin[e33] selling[e34] boxes for as little as 99 cents,
> 54>>> "Cheerios and Honey Nut Cheerios have eaten[e36] away sales
> 55>>> normally going[e38] to Kellogg's corn-based lines

All temporal entities are marked in text and the (automatic) annotator is supposed to look through these 5 *edus*, in order to discover TLINKs. To keep the notation as simple as possible, a TLINK indicating that e190 is BEFORE e189 is depicted as: e190 < e189. A careful check of the temporal links involving the set of the above temporal entities shows that the following temporal ordering exists between four events, which are in a window of 10 sentences (or 46 *edus*) before the current *edu*:

e38 < e190 < e36 < e189.

In TimeBank the only relations already captured are e190 < e189 and e38 < e36. The correct ordering of these four events would have to include TLINKS for three relations e38 < e190, e190 < e36 and e36 < e189. The

initial intrasentential TLINKS can be inferred from these three.

An attentive examination of the example shows that, even after using the time event closure algorithm, a context of three sentences, as used by the Event Diagram interface (Pustejovsky et al., 2002), is not enough for the manual recuperation of the time relations among entities from large texts (the text from which the example has been extracted, wsj_0610, contains 44 sentences with 53 EVENTs, 9 TIMEXs and 54 TLINKS).

6. Conclusions

This study may be relevant to several HLT areas which are sensitive to manual and (semi-) automatic temporal annotation. It gives corpus-evidence on how temporal annotation may be improved by making use of a previous discourse structure annotation. The discourse structure could be either given, as in the case of a gold standard corpus, or automatically obtained using an adequate discourse parser.

The method used to obtain summaries focussed on chosen entities of a free text (Cristea et al., 2005) could be combined with the temporal structure of the text in order to obtain chronological focussed summaries. In multi-document summarization (Mani, 2001), the combination between discourse and temporal information may also improve the performance of the system.

The relationship between discourse structure and temporality, as proven by the present study, may be exploited bidirectionally: discourse structure could be improved if the temporal links are known, and vice versa, more temporal links could be found if the discourse structure is known.

Acknowledgements. We are grateful to Daniel Marcu for making available the corpus of RST trees used in this study. This work has been partially supported by the CEEX research grant ROTEL-29 of AMCSIT (MEC code PC-D03-PT00-205).

7. References

Asher, N. and Lascarides, A. (2003): Logics of Conversation. Cambridge University Press.

Carlson, L., Marcu, D. and Okurowski, M. (2003): Building a Discourse-Tagged Corpus in the Framework of Rhetorical Structure Theory. In Jan van Kuppevelt and Ronnie Smith (Eds.), *Current Directions in Discourse and Dialogue*, Kluwer Academic Publishers, pp. 85-112.

Cristea, D., Ide, N. and Romary, L. (1998). Veins Theory. An Approach to Global Cohesion and Coherence. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and the 17th International Conference on Computational Linguistics (ACL-98/COLING-98)*, pp. 281-285.

Cristea, D., Ide, N., Marcu, D., Tablan, M.-V. (2000): Discourse Structure and Co-Reference: An Empirical Study, In *Proceedings of The 18th International Conference on Computational Linguistics COLING'2000*, Luxembourg, pp. 208-214.

Cristea, D. (2003): The relationship between discourse structure and referentiality in Veins Theory, in W. Menzel and C. Vertan (Eds.): *Natural Language Processing between Linguistic Inquiry and System Engineering*, "Al.I.Cuza" University Publishing House, Iasi, pp. 9-22.

Cristea, D., Postolache, O., Pistol, I. (2005). Summarisation through Discourse Structure. In Alexander Gelbukh (Ed.): *Computational Linguistics and Intelligent Text Processing, 6th International Conference CICLing 2005*, Springer LNSC, vol. 3406, pp. 632-644.

Filatova, E. and Hovy, E. (2001). Assigning Time-Stamps to Event-Clauses. In *Proceedings of the ACL-2001 Workshop on Temporal and Spatial Information Processing, ACL-2001*, pp. 88-95.

Hitzeman, J., Moens, M. and Grover. C. (1995) Algorithms for analysing the temporal structure of discourse. In *Proceedings of the Annual Meeting of the European Chapter of the Association for Computational Linguistics*, Dublin, Ireland, pp. 253-260.

Kamp, H. and Reyle, U. (1993). *From Discourse to Logic*. Introduction to Model theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory, Dordrecht ,Kluwer Academic Publishers.

Katz, G. and Arosio, F. (2001). The Annotation of Temporal Information in Natural Language Sentences. In *Proceedings of the ACL-2001 Workshop on Temporal and Spatial Information Processing, ACL-2001*, pp. 104-111.

Lascarides, A. and N. Asher (1993) Temporal interpretation, discourse relations and common sense entailment. In *Linguistics and Philosophy*, 16.5, pp. 437-493.

Mani, I. and Wilson, G. (2000). Robust Temporal Processing of News. In *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics (ACL'2000)*, pp. 69-76.

Mani, I., (2001). *Automatic Summarization*. John Benjamins.

Mani, I., Pustejovsky, J., Gaizauskas, R. (eds.) (2005). *The Language of Time - a Reader*, Oxford University Press.

Mann, W.C., and Thompson, S.A. (1988): Rhetorical Structure Theory: Toward a Functional Theory of Text Organization. In *Text* 8(3), pp. 243–281.

Marcu, D. (1997). The Rhetorical Parsing of Natural Language Texts. In *Proceedings of 35th Annual Meeting of Association for Computational Linguistics (ACL 97)*, pp. 96-103.

Pistol, I. (2005). *The automated parsing of the linguistic discourse*. Master dissertation, "Al.I.Cuza" University of Iasi.

Pustejovky, J., Belanger, L., Castaño, J., Gaizauskas, R., Hanks, P., Ingria, B., Katz, G., Radev D., Rumshisky, A., Sanfilippo, A., Sauri, R., Setzer, A., Sundheim, B. and Verhagen, M. (2002): NRRC Summer Workshop on Temporal and Event Recognition for QA Systems.

Pustejovsky, J., Hanks, P., Saurí, R., See, A., Gaizauskas, R., Setzer, A., Radev, D., Sundheim, B., Day, D., Ferro L., and Lazo, M. (2003). The TIMEBANK Corpus. *Proceedings of Corpus Linguistics 2003*, pp. 647-656.

Schilder, F., Habel, C. (2001). From Temporal Expressions to Temporal Information: Semantic Tagging of News Messages. In *Proceedings of the ACL-2001 Workshop on Temporal and Spatial Information Processing, ACL-2001*, pp. 65-72.

Setzer, A. (2001) Temporal Information in Newswire Articles: an Annotation Scheme and Corpus Study, PhD dissertation, University of Sheffield.

Setzer, A.; Gaizauskas, R.; Hepple M. (2005) Using Semantic Inference for Temporal Annotation Comparison. In I. Mani, J. Pustejovsky, R. Gaizauskas (Eds.), *The Language of Time - a Reader*, Oxford University Press, pp. 575-584.

Song, F. and Cohen, R. (1991). Tense Interpretation in the Context of Narratives. In *Proceedings of the Ninth National Conference on Artificial Intelligence (AAAI-91)*, pp. 131-136.

Webber, B.L. (1988). Tense as discourse anaphor. In *Computational Linguistics*, 14(2), pp. 61-73.