

Annotating Events and Temporal Information in Newswire Texts

Andrea Setzer and Robert Gaizauskas

Department of Computer Science
University of Sheffield
Regent Court
211 Portobello Street
Sheffield S1 4DP, UK
{A.Setzer, R.Gaizauskas}@dcs.shef.ac.uk

Abstract

If one is concerned with natural language processing applications such as information extraction (IE), which typically involve extracting information about temporally situated scenarios, the ability to accurately position key events in time is of great importance. To date only minimal work has been done in the IE community concerning the extraction of temporal information from text, and the importance, together with the difficulty of the task, suggest that a concerted effort be made to analyse how temporal information is actually conveyed in real texts. To this end we have devised an annotation scheme for annotating those features and relations in texts which enable us to determine the relative order and, if possible, the absolute time, of the events reported in them. Such a scheme could be used to construct an annotated corpus which would yield the benefits normally associated with the construction of such resources: a better understanding of the phenomena of concern, and a resource for the training and evaluation of adaptive algorithms to automatically identify features and relations of interest. We also describe a framework for evaluating the annotation and compute precision and recall for different responses.

1. Introduction

Most facts are not true eternally but for a limited period of time. To know that something happened may be worthless without knowing when it happened. For example, if one is interested in knowing the impact that a management change has had on the share price of a company, knowing that Jones succeeded Smith as president of MegaCorp may be useless without knowing when this event occurred with respect to, say, the record profits posted by MegaCorp in the third quarter of 1998.

Thus, if one is concerned with natural language processing applications such as information extraction (IE) (Cowie and Lehnert, 1996; Gaizauskas and Wilks, 1998), which typically involve extracting information about temporally situated scenarios (terrorist attacks, management successions events, rocket launches), the ability to accurately position key events in time is of great importance.

To date only minimal work has been done in the IE community concerning the extraction of temporal information from text, as evidenced by tasks set in recent Message Understanding Conferences (MUC). The MUC-6 named entity subtask required the identification of absolute time expressions in text (MUC, 1995), and the MUC-7 named entity subtask extended this requirement to include relative time expressions (MUC, 1998), but none of these tasks required placing events in time, or temporally relating events to each other. The MUC-5 and -7 scenario tasks did require participants to assign a calendrical time to certain specified event types (joint venture announcements and rocket launchings, respectively). However, this task is quite limited and the scores were low, indicating its difficulty.

The importance of extracting temporal information from texts, together with the difficulty of the task, suggest that a concerted effort be made to analyse how temporal information is actually conveyed in real texts. To this end we have devised an annotation scheme for annotating those

features and relations in texts which enable us to determine the relative order and, if possible, the absolute time, of the events reported in them. Such a scheme could be used to construct an annotated corpus which would lead to: a better understanding of the phenomena of concern (both their variety and distribution); a resource for training adaptive algorithms to automatically identify features and relations of interest; and, a resource for evaluating algorithms which purport to identify the features and relations of interest.

This paper presents a proposal for annotating temporal information in newswire texts. The paper addresses the conceptualisation of time underlying the annotation scheme, describes the annotation scheme and discusses the application of the scheme to a small corpus in an experiment to determine to what extent events in text can be ordered using only explicit temporal information and to what extent their ordering requires lexical semantic or world knowledge.

2. Conceptualising Time

Before an annotation scheme for temporal information can be proposed we must make clear the sorts of temporal entities and relations we suppose exist. Much has been written concerning the appropriate temporal ontology and set of temporal relations for analysing temporal phenomena in natural language (Allen, 1984; Galton, 1990; Steedman, 1997). However, our goal here is not to arrive at some indisputably 'true' description of temporal reality. Rather we wish to provide a framework that can be used in classifying expressions in real texts in a fashion that enables us to gain useful insights into how temporal information is conveyed in written language. Ultimately our aim is to develop an algorithm which can identify events in newswire texts and determine their temporal order or position in calendrical time well enough to answer questions about the ordering or times of events at the level we would expect

of an average human reader. Thus, our temporal ontology and the annotation scheme derived from it must be judged by the pragmatic criterion of the extent to which they contribute to this goal, rather than by an abstract criterion of philosophical truth.

Given this perspective, we may summarise our descriptive framework quite simply. It presumes the world contains the following primitive types: events, states, times, temporal relations and subevent relations – each of these is discussed in detail below. Of course this framework is not complete, but we believe it provides a useful starting point. Further event-event relations which are not temporal *per se*, but have temporal implications (for example causation) we have ignored for now as being difficult to define and annotate, and being of uncertain value for the project of building algorithms to extract temporal information from texts. Should they prove essential, our approach could be expanded to include them later.

Our proposals have been shaped by the genre of texts on which we are concentrating. Newspaper reporting is a distinct genre unto itself and several of its characteristics are significant when analysing how temporal information is conveyed (Bell, 1991). Firstly, newspaper articles typically exhibit a non-chronological time structure, which results from obeying news values rather than ordinary narrative norms. Secondly, newspaper articles frequently adopt the so-called ‘instalment method’ by which an event is introduced and returned to in more detail two or more times later in the text. Thirdly, and strongly related to this second point, news articles tend to follow the ‘inverted pyramid style’ which means that all main points are made in the beginning and then the article processes through decreasingly important information.

From this it is clear that different linguistic mechanisms may be used to convey temporal information in news-wire articles than in other types of narrative. In particular, event coreference plays an important role whereas aspect, in the sense of Aktionsarten, or lexical aspect (see Vendler (1967)), does not play a large role, mainly because the short paragraphs are not necessarily connected.

2.1. Events

Intuitively an event is something that happens, something that one can imagine putting on a time map. Events can be ongoing or conceptually instantaneous, we do not distinguish between these. What defines an event is very much dependent on the application and domain, but generally events have to be anchorable in time and they are usually conveyed by finite verbs and some nominalisations. Examples of events are:

A small single-engine plane crashed into the Atlantic Ocean.

By midafternoon, several vessels and a helicopter were combing the area.

Searchers found the plane’s landing gear, seat cushions and other debris.

The 1996 crash of the TWA 747 remains unexplained.

The view of those analysing language in news media, for example Bell (1991), is that an event is characterised by the following features: WHO, WHAT, WHERE, WHEN, ATTRIBUTION.

From our perspective we are obviously interested in the *when* part, because we want to analyse the (explicit) temporal information separately from the event. However, we are not interested in the internal structure of events at this stage and thus the *who*, *where* and *what* merge into one. We propose to treat events as black boxes without distinguishing the action (*what*) from the participants (*who*) or location (*where*) because we want to concentrate on the temporal information. It will, of course, be possible to include more detailed information about these other aspects of events at a later stage should it become clear that such information would lead to better results.

This leaves the *attribution* which tells us who reported the event, who or what is the source of information. This is usually conveyed by an event itself with the reported event being conveyed by a subordinate clause as in

The Coast Guard said the craft had taken off from Allaire Airport in Monmouth County, N.J.

where the source of information is the Coast Guard and the reported event is the plane taking off. This brings us to a discussion of event classes.

2.2. Event Classes

During our analysis of newswire articles it became clear that events can be classified into groups. This is not a new idea – for example, see Halliday (1985). The main classes in his classification system are *material processes* (or processes of doing like *the lion caught the tourist*), *mental processes* (or processes of sensing like *Tim realised that he was in a big city*), *relational processes* (like *Peter has a piano*), and *behavioural processes* (like *breathing* or *dreaming*). His classification, however, is not suitable for our purposes. For example, a *relational process* can also be *The fair is on Tuesday*, which we would want to place on a time map whereas we would regard *Peter has a piano* as stative and not put it on a time map. What we need is a classification which helps us in the project of locating events in time, and so we propose the following classification scheme.

Most events are what we call *occurrence events* – these are the events we want to place on a time map. Examples are:

A small single-engine plane crashed into the Atlantic Ocean about eight miles off New Jersey on Wednesday

By midafternoon, several vessels and a helicopter were combing the area about eight miles east of Sea Bright, N.J.

Revenue in Avco and Textron Financial rose 19 percent last year.

As mentioned above, there are also *reporting events*, whose main function is to associate the source of information with an (occurrence) event. Examples are:

The Coast Guard reported finding aircraft debris and fuel slick.

Searchers found the plane's landing gear, seat cushions and other debris, Petty Officer Fenn said.

The Coast Guard said the craft had taken off from Allaire Airport.

The most common case is that the reported event happened in the past and thus before the reporting event. In some cases this helps to temporally locate an event more accurately than just before the date of the article, as the following example shows.

Aeroflot general manager for Hong Kong Vassili Tkatchenko said on Tuesday he was unaware the writ had been filed.

The following example illustrates another type of event:

The plane was seen hitting the water shortly after 11 a.m. by a fisherman.

We call these events *perception events* and although they are relatively rare, the benefits of annotating them justify their being included into the scheme. Perceived events (like *the plane hitting the water* in the example) happen roughly at the same time as the perception event (like *was seen by a fisherman* above) and although their exact temporal relationship might not be known, we are able to locate the perceived event more accurately in time than we would have been able to using only verb tense of the perceived event.

Attitude events are similar to *reporting* and *perception* events in that they all take another event as an argument. Examples are *John hopes to go to New York on Friday* or *Mary believes that the plane hit the the Atlantic Ocean*. Like reporting events and unlike perception events, attitude events do not guarantee the reality of the participant event (X may report or believe that Y without Y's being the case; in contrast, if X sees that Y then Y must be true). However, while both reporting and perception events take place after the event reported or perceived, attitude events stand in no such clear temporal relation to their participant event. Thus, it is useful to distinguish this event class from the others when taxonomising events with respect to their temporal properties.

The final class of events we distinguish is *aspectual events*, such as

Afterwards, the statement said, there were two apparent interruptions of power to the cockpit voice recorder. The first interruption began 1 minute and 39 seconds after the sound and lasted 1 minute and 17 seconds; the second was just before the end of the tape.

which basically follow the structure mentioned above and involve *aspectual verbs* like *start, stop, finish etc.* Their temporal consequence is that the aspectual event indicates the start or ending of the related event.

2.3. Times

Like events, times can be viewed as having extent (intervals) or as being punctual (points). Rather than trying to reduce one perspective to the other, as has happened in much philosophical discussion on time, we shall simply treat both as *time objects*. A time object must, however, be capable of being placed on a time line (fictional or real).

Following general convention, and the approach taken in MUC, we distinguish between two classes of time objects, DATES and TIMES, times which are larger or smaller than a day, respectively.

2.4. States

A state is a relation between entities or the holding of an attribute of an entity which, while capable of change, is ongoing over a time span, usually longer than the time span covered by the article. Examples are:

The plane, which can carry four people, ...

The water is about 125 feet deep in that area.

Typically, a change of state constitutes an event. At this point we are less interested in states, and we have not taken them into account in our annotation scheme.

2.5. Temporal Relations

Events stand in certain temporal relations to other events and to times. Times may be temporally related to other times as well, although this does not happen very often in the articles we analysed so far. A few examples will illustrate how events are related to events and times.

*The plane crashed **after** the pilot and his crew ejected.*

*...**before** the craft fell, its three rotor blades shot off.*

Here events are related to other events by using temporal subordinate conjunctions, which is the most common case. Temporal prepositional phrases are usually employed when relating events to times, as shown here:

A small single engine plane crashed into the Atlantic Ocean on Wednesday.

Neither relating events to events nor relating events to times need necessarily be explicit, as this example shows:

Sunday, an F-14D crashed into the Pacific Ocean off southern California, killing its two crew members.

The full set of temporal relations we suppose at present is:

included *The plane crashed on Wednesday*

includes *By midafternoon, several vessels were combing the area*

after *The plane crashed after the pilot and his crew ejected.*

before *...before the craft fell, its three rotor blades shot off*

simultaneous *All 75 people on board the Aeroflot Airbus died when it ploughed into a Siberian mountain in March 1994*

Note that this set consists of two pairs of converse relations (A includes/is after B iff B is included in/is before A) and the **simultaneous** relation. Thus, there really need be just three primitive relations and the remaining two can be viewed as being defined in terms of the others. This is a minimal set we defined after analysing a number of news-wire articles and it can easily be expanded should it prove necessary and beneficial.

2.6. Subevent Relations

As mentioned earlier, the *instalment method* used widely in news-wire articles means events introduced early in the text are returned to at a later point in the same text, where further information about the event is provided. To exploit this method, we suppose a further kind of event relation – the *subevent* and *identity* relation, (identity can be viewed as a special case of subeventness).

This relation is exhibited in cases like the following:

*The National Transportation Safety Board is borrowing a Boeing 737 from Seattle's Museum of Flight as part of its investigation into why a similar jetliner **crashed** near Pittsburgh in 1994.*

[. . .]

*The museum's aircraft, ironically enough, was donated by USAir, which operated the airplane that **crashed, killing** 132 people on board.*

where the first sentence introduces the event and the second sentence refers to the same event - event identity for *crashed* - and a subevent (*killing 132 people on board*). The next example shows how event coreference can help locating events in time more accurately.

*A small single-engine plane **crashed** into the Atlantic Ocean about eight miles off New Jersey on Wednesday.*

[. . .]

*The plane, which can carry four people, was seen **hitting** the water shortly after 11 a.m. by a fisherman, ...*

Again, the first sentence introduces the event and the second event refers to a subevent (*hitting the water*), with the subevent providing more precise information, i.e. that the plane crashed happened around 11 a.m. on the Wednesday in question.

3. Annotating Temporal Information in Text

Given this conceptual framework, we can now turn to proposing an annotation scheme, which enables events, times and temporal and subevent relations to be marked up in texts. As with most modern annotation schemes, our scheme is defined in SGML. We first describe the scheme, then discuss the vexed issue of annotating implicit temporal relations, a discussion which leads us to advance the notion

of a minimal temporal model of a text which can serve as a basis for comparing the results of different annotators annotating the same text. Finally, we discuss the actual process of annotation.

3.1. The Annotation Scheme

3.1.1. Annotating Events

We annotate events by simply marking a representative of the event in the text, as the following example illustrates (please note that we do not necessarily include all attributes in our examples that would be included 'in real life', to keep the examples as short and legible as possible).

```
A small single-engine plane
<event eid=16 class=OCCURRENCE tense=past>
  crashed
</event>
into the Atlantic Ocean about eight miles off New
Jersey
```

The first candidate for event representative is the head of the finite verb group. If the event is conveyed by a nominalisation then we chose the head of the nominalisation as the representative. Events can also be represented by non-finite clauses, as in

```
The plane crashed, <event> killing </event> all
passengers on board
```

in which case we annotate the non-finite verb as the representative.

The set of attributes assigned to an event comprises:

event ID: The event ID uniquely identifies an event in the text.

class: The event class, as described in section 2.1.

tense of the verb: Only simple tenses (past, present and future) are indicated.

aspect: The aspect of the verb can be indicated here, only progressive or perfective aspect is marked.

relatedToEvent: If an event is related to one or more events, be it via a temporal or a subevent relation, then the IDs of the related events are stored in this attribute.

relatedToTime: Same as above, this time an event is related to one or more time objects and the IDs of the time objects are stored here.

eventRelType: The types of relationship, either the temporal relations described in section 2.5. or the coreference relations (*subevent*, *identity*) are possible values for this attribute.

timeRelType: Similar to *eventRelType*, but only the temporal relations from section 2.5. are valid.

signal: The word(s) or textspan(s) signalling the temporal relation holding between two entities can be realised in the text or not. In the former case, we keep track of the 'signal' by keeping the ID in this attribute.

argEvent: *Reporting, perception, attitude and aspectual* events have other events (mostly *occurrence* events) as arguments and the IDs of these argument events are stored here.

The following example shows the annotation of a reporting event and the argument occurrence event with all attributes.

```
The Coast Guard
<event eid=2 class=REPORTING tense=past
  argEvent=3>
  reported </event>
<event eid=3 class=OCCURRENCE tense=past >
  finding </event>
aircraft debris and a fuel slick, but no bodies or
survivors.
```

3.1.2. Annotating Times

To be annotated as times, time expressions must be anchorable on a time line and like events, time expressions are uniquely identified by an ID to enable events to be associated with them.

```
<timex tid=5 type=DATE> Tuesday </timex>
<timex tid=5 type=TIME> 11 a.m. </timex>
```

MUC also distinguishes between ABSOLUTE time expressions (indicating a specific segment of time as in *20 minutes after 10, midnight, 10th of October*) and RELATIVE time expressions (indicating a date relative to the date of the document as in *yesterday or last month*). In our opinion, most time expressions, whether they are what MUC calls specific segments or not, are relative to the date of the document. If we consider, for example, the expression *October 10th* then it depends on the context whether the following or past October is referred to, as the following examples illustrate.

The plane crashed on October 10th.
The CEO will give a talk on October 10th.

An alternative would be to define an absolute time expression as one where the time object represented by the expression can be placed unambiguously on a calendrical timeline or timemap without any additional information from any other part of the text it occurs in. And a relative time expression, accordingly, as one where one does need additional information.

The latter distinction is the one we would favour, but since it is not clear at this point whether it would lead to better results if it would be incorporated into the annotation scheme, we have not yet done so.

3.1.3. Annotating Temporal Relations

Events may be related to times or to other events. If two events are related then one of them carries the ID of the other as an attribute to link them and also the type of the relation. Which event has those attributes associated with it depends on the the type of relation (apart from *simultaneous* which is reflexive). If the word or textspan signalling two events being related is realised, as in the example below, then the ID of the signal is stored as an attribute of

the event, so the link between the signalling word and the events is not lost. If an event is related to a time object then the event carries the ID of the time object and the type of relation in its attributes. Should the signalling word be realised, again as in the example, then the ID of the signal becomes an attribute of the event. The following examples illustrate the approach.

```
All 75 people on board the Aeroflot Airbus
<event eid=4 class=OCCURRENCE tense=past
  relatedToEvent=5 eventRelType=simultaneous
  signal=7>
  died </event>
<tr_signal sid=7> when </tr_signal>
it
<event eid=5 class=OCCURRENCE tense=past >
  ploughed </event>
into a Siberian mountain.
```

```
A small single-engine plane
<event eid=9 class=OCCURRENCE tense=past
  relatedToTime=5 timeRelType=included
  signal=9>
  crashed </event>
into the Atlantic Ocean about eight miles off
New Jersey
<tr_signal sid=9> on </tr_signal>
<timex tid=5> Wednesday </timex>.
```

3.2. Annotating Implicit Temporal Relations

In many cases temporal relations are not explicitly signalled in text. For example, an event can be related to a time object by simple juxtaposition:

A senior investigator who looked at the cockpit wreckage Tuesday ...,

Event coreference (as described above) is hardly ever expressed explicitly; rather, the reader is presumed to be able to draw upon script-like knowledge of the temporal structure of stereotypical scenarios. Narrative sequence is another example of implicit temporal relation but is rarely used in newswire articles.

Implicit temporal information is easily annotated with our annotation scheme (section 3.1.). However, equivalent implicit relations can be marked up in a variety of ways. For instance, if A and B are *simultaneous* and C is implicitly *after* A and B, then C might be annotated as *after* A or *after* B or both. Hence it is a very difficult task to produce an annotation guideline which would result in the same annotation independent of the annotator. Since, as the preceding example shows, much of the temporal information that could be marked is redundant, we would expect high inter-annotator disagreement.

With this problem in mind, we have to ask ourselves how to compare and score different annotations of the same text produced by different annotators (human or machine) where implicit, as well as explicit, temporal relations are marked. This problem of implicit temporal information is similar to the problem of coreference annotation in MUC-6 and -7 (1995; 1998). There too the relation being annotated (coreference) could be marked in superficially different, but semantically identical ways (equivalent coreference

chains can be specified by annotating different links). The solution adopted in MUC was to define a model-theoretic scoring scheme which relied upon comparing the equivalence classes of the set of linked entities produced by the annotator 1 with that produced by annotator 2 (Vilain et al., 1995). Our approach is similar, but the problem more complex because there are more relations and not all of the relations are equivalence relations.

3.3. Comparing Temporal Annotations

The events and times annotated in a text, or rather their respective IDs, form two sets, E and T , respectively. Since all of our temporal annotations are binary relations relating events or times to other events or times, the denotation of each relation as specified in the text can be viewed as a subset of $(E \cup T) \times (E \cup T)$. For each temporal relation certain formal properties pertain. For example *simultaneous* is an equivalence relation, while *before*, *includes* and *subevent* are transitive, but asymmetric and irreflexive. Therefore, given a partially specified model of the temporal relations in a text (specified as a set of pairs comprising a part of the denotation of the relation) the deductive closure of each relation can be computed to arrive at a total model. If the deductive closures of two partially specified models are identical, then these two models are equivalent, though not themselves identical. Further if any (partial) model \mathcal{M} is such that no proper subset of \mathcal{M} has an equivalent deductive closure to \mathcal{M} , then \mathcal{M} is a *minimal* model of the temporal relations in the text.

Let us denote sets of pairs from $(E \cup T) \times (E \cup T)$ which constitute the denotations of *simultaneous*, *before*, *includes* and *subevent* by S , B , I , and E respectively. The set of inference rules we need to compute the deductive closure is not yet complete, but will at least contain the following rules:

$\forall x, y, z \in (E \cup T) \times (E \cup T) :$

- $(x, y) \in S \Rightarrow (y, x) \in S$
- $(x, y) \in S \wedge (y, z) \in S \Rightarrow (x, z) \in S$
- $(x, y) \in B \wedge (y, z) \in S \Rightarrow (x, z) \in B$
- $(x, y) \in I \wedge (y, z) \in S \Rightarrow (x, z) \in I$
- $(x, y) \in E \wedge (y, z) \in E \Rightarrow (x, z) \in E$
- $(x, y) \in S \wedge (y, z) \in I \Rightarrow (x, z) \in S$
- $(x, y) \in S \wedge (x, z) \in B \Rightarrow (y, z) \in B$
- $(x, y) \in I \wedge (y, z) \in B \Rightarrow (x, z) \in B$
- $(x, y) \in I \wedge (y, z) \in I \Rightarrow (x, z) \in I$
- $(x, y) \in B \wedge (y, z) \in B \Rightarrow (x, z) \in B$
- $(x, y) \in E \Rightarrow (y, x) \in I$

We then can denote the deductive closure of S , B , I , and E by S^{\models} , B^{\models} , I^{\models} , and E^{\models} respectively. Given these definitions we are now in a position to specify what precision and recall mean in this framework. Letting S_k and S_r

denote the annotated *simultaneous* relations in the answer key and system response respectively and S_k^{\models} and S_r^{\models} their deductive closures, respectively (and similarly for B , I and E). The recall and precision for the *simultaneous* relation is given by:

$$R = \frac{|S_k^{\models} \cap S_r^{\models}|}{|S_k^{\models}|}$$

$$P = \frac{|S_k^{\models} \cap S_r^{\models}|}{|S_r^{\models}|}$$

Recall and precision measures can be defined in a parallel fashion for the other relations. An overall recall and precision measure for all temporal relation could be defined as follows (note that subeventness is factored out in this equation because, as the final inference rule above implies, the members of E^{\models} are also included in I^{\models} and they must not be counted twice):

$$R = \frac{|S_k^{\models} \cap S_r^{\models}| + |B_k^{\models} \cap B_r^{\models}| + |I_k^{\models} \cap I_r^{\models}|}{|S_k^{\models}| + |B_k^{\models}| + |I_k^{\models}|}$$

$$P = \frac{|S_k^{\models} \cap S_r^{\models}| + |B_k^{\models} \cap B_r^{\models}| + |I_k^{\models} \cap I_r^{\models}|}{|S_r^{\models}| + |B_r^{\models}| + |I_r^{\models}|}$$

3.4. The Process of Annotation

The annotation takes place in several stages. In the first stage, the annotator marks up all the explicit information contained in the text. As a result, all events, times and explicit temporal information are annotated. In a second phase, the annotator will mark the implicit information, as described in section 3.2. This means subevents are annotated, or rather linked to the events containing them, and the cases where time objects are related to events but the signalling word is missing. The second phase could be supported by a program that would prompt the annotator about missing relations or this could be introduced as a third phase. Our general framework supports both approaches.

In order to build a complete model of the temporal information contained in a text (as described in section 3.3.), we first analyse the annotated information and draw all inferences possible and assign as many pairs as possible from $(E \cup T) \times (E \cup T)$ to one of the relations in S , B , I and E . The annotator is then prompted for unknown pairs. To minimise the number of questions that need to be asked, we will try and develop a scheme which asks those questions first which allow as many inferences as possible. This model then will allow us to evaluate inter-annotator agreement, as well as computing precision and recall (as described earlier).

To enable simple and user friendly annotation, we developed a graphical annotation tool in Perl/Tk which allows the annotator to simply mark up a portion of the text and click on the appropriate type with which he or she wants to annotate. At this point the annotator will be prompted to enter all necessary attributes, some of which he or she can choose from a selection.

4. Working with the Annotation Scheme

An interesting question is to what extent events in text can be ordered using only explicit temporal information and to what extent their ordering requires lexical semantic and world knowledge. To this end we have conducted an experiment in which the annotation scheme was applied to three New York Times newswire texts. The annotation was then separated from the text and analysed independently to determine how well events could be located in time without actual textual content, i.e. solely based on information like `<event eid=16 class=OCCURRENCE tense=past>`. Located in time means here either marked with a calendrical date or related temporally to another event in the text. We then manually created the deductive closure, as described in section 3.3., and compared the results to the complete model of the temporal information contained in the text. This complete model was also manually built, using all available knowledge (annotation, textual content and general world knowledge).

This method resulted in 37% Recall (and since no false temporal relations were annotated, 100% Precision), but more articles need to be examined to confirm these results. We recently revised the evaluation scheme and thus only three texts could be evaluated, but the experiment shows how the annotation scheme may provide insights into the mechanisms by which temporal information is conveyed in text. More information about working with the annotation scheme and the corpus built with it can be found in (Setzer and Gaizauskas, 2000).

5. References

- J.F. Allen. 1984. Towards a General Theory of Action and Time. *Artificial Intelligence*, 23:123–154.
- A. Bell. 1991. *The Language of News Media*. Blackwell.
- J. Cowie and W. Lehnert. 1996. Information extraction. *Communications of the ACM*, 39(1):80–91.
- R. Gaizauskas and Y. Wilks. 1998. Information Extraction: Beyond Document Retrieval. *Journal of Documentation*, 54(1):70–105.
- A. Galton. 1990. A Critical Examination of Allen’s Theory of Action and Time. *Artificial Intelligence*, 42.
- M.A.K. Halliday. 1985. *An Introduction to Functional Grammar*. Edward Arnold.
1995. *Proceedings of the Sixth Message Understanding Conference (MUC-6)*. Morgan Kaufman.
1998. *Proceedings of the Seventh Message Understanding Conference (MUC-7)*. Morgan Kaufman. available at <http://www.saic.com>.
- A. Setzer and R. Gaizauskas. 2000. Building a Temporally Annotated Corpus for Information Extraction. In *Proceedings of the Workshop on Information Extraction Meets Corpus Linguistics held in conjunction with the Second International Conference on Language Resources and Evaluation (LREC2000)*.
- M. Steedman. 1997. Temporality. In J. van Bentham and A. ter Meulen, editors, *Handbook of Logic and Language*, pages 895–938. Elsevier.
- Z. Vendler. 1967. *Linguistics in Philosophy*. Cornell University Press, Ithake, New York.
- M. Vilain, J. Burger, J. Aberdeen, D. Connolly, and L. Hirschman. 1995. A Model-Theoretic Coreference Scoring Scheme. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*. Morgan Kaufman.