Preface to the Workshop Proceedings

The ability to identify and analyse temporal information is important for a variety of natural language applications, such as information extraction, question answering, and multi-document summarisation. Nevertheless, this area of research has been relatively unexplored. It became evident during an ACL 2001 workshop on temporal and spatial information processing that that some research institutions have started to work on different aspects of temporal information, but no consensus has been achieved yet on what and how temporal information should be identified in text.

This workshop aims to provide a forum for researchers to present their work in this field and to discuss future developments such as building shared resources e.g. temporally annotated corpora. It is timely to coordinate the effort being undertaken in the community at this stage of research into temporal information.

We received 11 submissions and accepted 8 of them. The quality of the papers was very high and the topics diverse within the boundaries of the workshop theme. The selection process was therefore difficult. We would like to thank the members of the programme committe for their time and effort during the evaluation stage and for their valuable comments to the authors of the submitted papers.

The workshop will comprise presentation of each author and two talks by invited speakers. This will be followed by a panel discussion session at the end of the workshop. We would like to thank the invited speakers and the authors in advance for their participatation and we are looking forward to a hopefully enjoyable and enriching workshop which will contribute to this interesting and important emerging field within the natural language processing community.

Andrea Setzer, University of Sheffield
Robert Gaizauskas, University of Sheffield
(proceedings editors)

The Workshop Programme

09:15 - 09:30	Introduction
09:30 - 10:00	Andrew Salway *Temporal Information in Collateral Texts for Indexing Movies*
10:00 - 10:30	David Bree *Features required to distinguish between temporal uses of the preposition FOR*
10:30 - 11:00	Eleni Galiotou *A Representational Scheme for temporal and causal Information Processing*
11:00 - 11:30	coffee break
11:30 - 12:00	Joel Tetreoult *Tense and Implicit Role Reference*
12:00 - 12:30	Estele Saquete *Recognising and Tagging Temporal Expressions in Spanish*
12:30 - 13:00	Invited speaker: Stephanie Strassel *Temporal Annotation and Relation Tagging for Automatic Content Extraction*
13:00 - 14:30	lunch
14:30 - 15:00	Rich Campbell *A Language-Neutral Representation of Temporal Information*
15:00 - 15:30	Andrea Setzer *On the Importance of Annotating Temporal Event-Event Relations in Text*
15:30 - 16:00	Jerry Hobbs *Towards an Ontology for Time for the Semantic Web*
16:00 - 16:30	Invited speaker: James Pustejovsky *TimeML: Time and Event Recognition for Question Answering*
16:30 - 17:00	coffee
17:00 - 18:00	panel discussion

Workshop Organisers

Bill Black, UMIST, UK
Rob Gaizauskas, University of Sheffield, UK
Graham Katz, University of Osnabrueck, Germany
Andrea Setzer, University of Sheffield, UK
George Wilson, the MITRE Corporation, USA

Workshop Programmme Committee

Bill Black, UMIST, UK
Rob Gaizauskas, University of Sheffield, UK
Graham Katz, University of Osnabrueck, Germany
Inderjeet Mani, the MITRE Corporation and Georgetown University, USA
Allan Ramsay, UMIST, UK
James Pustejovsky, Brandeis University, USA
Frank Schilder, University of Hamburg, Germany
Andrea Setzer, University of Sheffield, UK
Beth Sundheim, SPAWAR Systems Center, USA
Janyce Wiebe, University of Pittsburgh, UK
George Wilson, the MITRE Corporation, USA

Table of Contents

Papers

Author Index

Features required to distinguish between temporal uses of the preposition FOR

David S. Brée

Computer Science Department, Manchester University
Oxford Road, Manchester M13 9PL, UK
bree@cs.man.ac.uk

Abstract

What are the features of a sentence that enable a reader to distinguish between the various different temporal uses of prepositions? To answer this question, we analyzed all the sentences in the million word Brown corpus containing a temporal use of FOR. The preposition FOR has been taken as a case study as it is both highly frequent and has several different uses. The results show that many different aspects of the FOR PP itself and of its matrix play a role. The most important are: the temporal nature of the noun in the PP as this indicates the temporal use, the definite v. indefinite nature of the determiner in the PP (in particular if there is no determiner, then the distinction between absence due to a proper noun v. a plural), the nature and aspect of the verb in the matrix of the PP, the position of the PP relative to other components of the matrix and finally the presence of certain qualifiers and post-modifiers in the PP which attach durations to the time axis. It is recommended that attention be paid to these features when tagging a text for temporal analysis.

1. Introduction

We use a dozen or so prepositions to convey temporal information. A few of do so in several different ways. How do listeners determine which of these ways the speaker intended? They use other features in the sentence, such as the aspect of the verb. Determining what these features are is the central concern of this talk.

We focus on just one of these prepositions: FOR. FOR is the most prolific of the temporal prepositions after IN, which I have already dealt with elsewhere (Brée, in press). And, like IN, but unlike the other high frequency temporal preposition AT, it is used in several different ways. The main distinction is between temporal durations (lengths of time not directly attached to the time line, e.g. FOR *five minutes*) and named temporal intervals, which are attached to the time line, e.g. FOR *1961*.

The most frequent temporal use of FOR is to indicate that a state given by the matrix in which the FOR Prepositional Phrase (PP) is embedded, holds for a length of time, i.e. FOR acts as a universal quantifier over a duration:

f32:084 She helped with teaching as well as office work FOR a few years ...

During any sub duration of the duration *a few years*, the claim is being made that *she helped with teaching*. Indeed, the possibility or otherwise of adding a durational FOR phrase to a verb phrase is a well-known test of a state versus an event verb.

The duration may be attached to the time axis to indicate an interval, with one end or the other being the time of reference (`tor`) (Reichenbach, 1947):

n09:100b Tom had been laying for Aaron McBride FOR a long time ...

p14:189 I won't be in town FOR a couple of days ...

One of our tasks is to discover how to determine whether the duration indicated by the FOR prepositional phrase (PP) is attached to the `tor` or not.

FOR, unlike IN, cannot in general be used with a temporal interval, c.f.:

g14:001 There were fences IN/*FOR the old days when we were children.

So how is universal quantification over a temporal interval indicated? By using IN, even though with durations it is an existential quantifier, or, occasionally and for emphasis, *throughout*.

g32:061 Little enough joy was afforded Wright *in/*for/throughout* the spring of 1925 ...

However, in certain circumstances FOR can be used with a temporal interval:

j66:011 "Evil, man, evil", he said, and that's all he said FOR *the rest of the night.*

b01:038 DeKalb's budget FOR *1961* is a record one ...

h29:018b I look for TV sales and production to be approximately equal at 5.7 million sets FOR *the year* ...

There are obviously restrictions on the use of FOR with an interval rather than a duration, and another of our tasks will be to determine what these restrictions are.

As with other temporal prepositions, FOR can be used not only to indicate a duration or an interval, but it also has special and idiomatic uses which will need to be recognized as such:

k19:060 Rector asked him to move it FOR the time being; ... [*special*]

n23:015 And FOR the hundredth time that week, he was startled at her beauty. [*idiomatic*]

We will now look in more detail at these different ways in which FOR is used temporally before searching for the features of the context that may discriminate between these different uses. Then, by examining sentences with a temporal FOR PP in a corpus, we will develop a set of heuristic rules for actually making such discriminations using these features, and others if necessary.

The corpus that will be used is the Brown University corpus of a million words taken from 499 sample texts of American English (Kučera & Frances, 1967) selected randomly from 15 genres, labelled *a* ... *r*. Sentences from the corpus are indicated by the use this letter, followed by a number to indicate the text, then a semi-colon and finally the sequential number of the sentence within its text. For example, n19:100b indicates the 100th sentence in the 19th text within the genre *n*. The final *b* indicates that the FOR being analyzed is its second occurrence in that sentence.

There is, in the Penn Tree Bank, a tagged version of the Brown corpus which was originally used. However, for the level of analysis that is to be undertaken here, the tagging proved to be insufficient. Since this analysis was made, a newly tagged and parsed version of the Brown corpus has been made available by the Language Data Consortium (Penn Tree version 3), but only 8 of the 15 genre have been parsed. This tagging includes an identification of the temporal use of prepositions. However, a check of the temporal uses of IN against a hand analysis showed that there were about 80% more instances tagged as the temporal use of IN (821) compared to those found by hand (483) in these same genres (Brée, in press). Consequently this new version of the Brown corpus is not suitable as a basis for the level of analysis that we are about to conduct.

2. Uses

2.1. Duration

We begin with cases in which FOR specifies a duration. It is usually indicated by an indefinite determiner in the FOR PP. In its simplest form the duration is unattached in any way to the time axis. Otherwise the duration is attached to the time axis at the `tor`, to indicate an interval either beginning or ending at the `tor`. Taken together these are the Duration uses.

The simplest Duration use is to specify a pure duration during which the matrix proposition holds, either just once or on a regular basis:

a24:035 One house was without power FOR about *half an hour* ...

k14:074 He worked FOR *two hours a day* with each model sent by the rabbi.

However, the duration may also be attached to the time axis just after or before the `tor`. The matrix state then lasts for the whole of the interval with the `tor` at one end and lasting the duration specified in the FOR phrase:

p14:189 I won't be in town FOR *a couple of days*, ...[`tor+`]

f34:092 Cereal grains have been used FOR *centuries* to prepare fermented beverages. [`tor−`]

Combinations with PPs introduced by other temporal prepositions give interesting possibilities here:

After: updates the `tor`, and the duration follows the `tor` (`tor+`):

n13:034 ... FOR a good minute *after* they rounded the bend ... Matilda could not speak at all.

Before: updates the `tor`, and the duration precedes the `tor` (`tor−`):

f13:080 ... usually you would do better to rent a place FOR a year or two *before* you buy.

Since: marks the beginning of the duration ending at the `tor` (`tor−`):

f20:073 FOR three years, *since* the liquor territorial conference, Torrio had ... tolerated O'Banion's impudent double-crossing.

Until: marks the end of the duration beginning at the `tor` (`tor+`):

d15:075 They went along the pass ... FOR a short while *until* they came to a river ...

Occasionally an indefinite duration will be found with a special use of FOR (see Section 2.3.):

p20:162 We went to the Louvre FOR *a* few hours ...

Note that the *going to the Louvre* did not last a few hours; it was the subsequent stay in the Louvre that lasted that time.

2.2. Interval

FOR phrases with intervals, usually indicated by a definite determiner, can be used in the same three ways as for IN phrases: attached to the VP, attached to an NP and as an interval over which a count in the matrix is to be taken. Taken together these are the Interval uses. We will look at each in turn.

2.2.1. Attached to the VP

As we have seen already, only certain intervals can be used in a FOR PP whose matrix is a VP. From the Brown corpus we find that there are just three types of interval that may be used.

The specification of the interval includes its duration. The interval in the FOR phrase is a duration but the duration is qualified or modified so that it becomes unique and definite:

- The duration is attached to the `tor` and made definite by using a suitable qualifier such as *past/last/next*:

  g08:010 But the South is, and has been FOR *the past* century, engaged in ...

  n19:053 FOR *the last* half hour Mary Jane had criss-crossed half the length of the Gardens ...

  l07:116 FOR *the next* hour he scrambled happily up and down the ladder, ...

- The duration is attached to some other time interval using a suitable post-modifier such as *following/ending/up to ...*:

  g54:071 ... FOR *the* weeks *following* it Tom did not know whether his return to Harvard could be arranged.

  a43:036 Operating revenues were ... up FOR *the* 12 months *ending* in March.

  g46:040 The plan is admirably fulfilled FOR *the* period *up to* 1832.

The interval is part of a larger definite interval. The part may be the beginning, the end or somewhere in the middle of the interval and is given either by an adjective or a complex NP:

- The beginning of a definite interval, using *the first*:

  f16:061 FOR *the first* three weeks, the ship skirted up the east coast of Great Britain, then turned westward.

- The end of a definite interval, using the qualifiers *the rest of/remainder of/remaining*:

  r08:022 FOR *the remainder of* the movie, Chancellor Neitzbohr proceeds to lash the piano stool ...

j79:009a Moreover, by ...always using the optimal Af-stage policy FOR *the remaining* stages, ...

- Some other part of a definite interval, using, e.g. *most/much/whole (duration)/all of*:

a18:027 FOR *most of* the 25 years the operation was under feminine direction.

a18:039 Col. Clifton Lisle ...headed the Troop Committee FOR *much of* its second and third decades ...

f15:021b ...it is possible to be exempt from the normal obligation of parenthood for a long time and even FOR *the whole duration of* married life ...

The tor is included in the interval using a noun phrase such as *now, the present, the moment*:

m03:016 "FOR *now*, it is clear that we were in the wrong.

g15:025 FOR *the present* it is enough to note that ...

l14:068 Try to forget motive FOR *the moment*.

In summary, the interval in a FOR phrase whose matrix is a VP must be either:

- a duration made definite by a qualifier or modifier linking it to the tor or another interval; or
- part of a larger interval; or
- an interval that includes the tor specifically, e.g. *now, the present, the moment*.

Note, in particular, that proper nouns indicating clock times and calendar dates, such as *7 p.m., May, 1996*, are not found.

2.2.2. FOR modifying a noun phrase

The matrix of FOR phrase may itself be a noun phrase referring to a regularly recurring event or object. The FOR phrase indicates a unique interval which serves to pick out a specific occurrence of this event or object. Here there is no restriction on the type of interval specified in the FOR PP:

- Regularly recurring events, e.g. plays

k05:140 The *play* FOR Saturday night was to be a benefit performance of The Octoroon.

- Regularly recurring financial objects, e.g. tax returns, budgets, costs, revenues etc.:

h24:009 Since ...your *return* FOR the calendar year 1961 will be timely filed.

b01:038 DeKalb's *budget* FOR 1961 is a record one ...

a28:013 FOR the year to date, *sales* ...still lag about 5 % behind 1960.

- Other regularly recurring objects, e.g. newspapers, journals, appointments, horoscopes, rates and even obligations:

g38:040 The final issue of the *Englishman*, No. 57 FOR February 15, ran to some length ...

j56:055 Eber L. Taylor of Manchester Depot recorded ...in his *diary* FOR 1906.

g49:003 ...having made a formal *engagement* by letter FOR the next week, ...

p21:082 I may settle on some makeshift *arrangements* FOR the summer.

r09:078 ...I'd say that your *horoscope* FOR this autumn is the reverse of rosy.

j07:030b ...to give reliable impact *rates* FOR the periods of exposure.

b17:009a Our only *obligation* FOR this day is to vote, ...

It could be argued that the use of a FOR phrase to qualify a noun is not essentially temporal. In many of the above examples the time in the FOR phrase could be replaced by an appropriate proper noun, e.g. FOR *IBM*, FOR *John*. The FOR phrase would then indicate possession or purpose, e.g. *the budget of IBM, the arrangements for John* (i.e. the special use of FOR, see Section 2.3.). We have included it here as we have chosen to include as temporal all uses of FOR that have a temporal noun in the PP.

2.2.3. There is a count over an interval given by the FOR PP

This is closely related to the above, but there is no NP which the FOR PP is modifying. Rather there is a count over an interval given by the FOR PP:

a43:038 FOR the year, the road earned 133 per cent of its interest costs ...

Normally IN is used for indicating the interval over which a count is taken, but when the interval is given by a duration made definite by the determiner *the*, then FOR is used.

The remaining examples of the use of a definite noun phrase with FOR are non-standard. These have been categorized as a special use (Section 2.3.) and an idiomatic use (Section 2.4.).

2.3. Special use

The special use is one in which the FOR phrase gives an interval but not one in which the matrix proposition holds. Rather it is an interval, later than the tor, in which some event, that is related to the matrix proposition in some way, will take place. The relationship may be a simple prediction, planning or purpose:

h29:008 I believe a further gain *is in prospect* FOR 1961. [*prediction*]

a10:045 The dinner is sponsored by organized labor and *is scheduled* FOR 7 p.m.. [*plan*]

l24:118 We *decided to* leave the third one intact FOR tomorrow. [*purpose*]

Note how in all these examples, the matrix proposition holds at the tor rather than in the interval given by the FOR PP. In:

h29:008 the *gain* is already *in prospect* at the tor; the actual gain will be realised later, in *1961*.

a10:045 the *dinner* is already *scheduled* by the tor; the actual dinner will take place later, at *7 p.m.*.

l24:118 the *third one* is *left intact* at the tor; the reason for this is so that it can serve some purpose *tomorrow*.

It is the matrix that is indicative of this special usage. It needs to indicate that an action has been undertaken that will lead to some event to take place or state to be the case at a future interval that is given by the FOR PP. How this is indicated is not simple. The matrix verb, plus particle if present, gives the indication: *be in prospect, be scheduled,* *decide to X*. A list of verbs that were found in the corpus to indicate special use will be given in Section 4.2.

While the FOR PPs in the purposeful sentences such as l24:118 clearly have a temporal function, they are not only temporal. There is also a sense of purpose which is another use of FOR:

a01:060 Vandiver opened his race FOR governor in 1958 ...

Should they be included as a temporal use of FOR? With prepositions with both temporal and spatial uses, such as AT or IN, the temporal use can be detected from the temporal nature of the noun in the PP (Brée & Pratt, 1997). If we wish to use this general heuristic to identify when a PP is being used temporally rather than spatially, we need to keep these as a special temporal use of FOR.

Note also that, as with NP attachment of the FOR interval, proper clock and calendar nouns can be used in the FOR PP when there is a special use.

2.4. Idiomatic

For the Nth time, where *Nth* is any ordinal, generally *first*, but including *last*, was classified as idiomatic use:

p22:041 It was weeks before we even kissed FOR *the first time*.

19:008 And FOR *the thousandth time*, I answered myself.

The reason for considering this an idiomatic use is that the noun *time* is here not being used in its temporal meaning but as an indication of position in a sequence. Confirmation of this difference is given by the different translations given in other languages. In Dutch, for instance, the normal word for time is *tijd*; but in the expression *for the Nth time*, it is *keer*.

This concludes the list of uses for temporal FOR phrases. We turn now to an examination of the components of the phrase in the expectation that they will distinguish between these different uses.

3. Components of the prepositional phrase

Our task now is to code features of the sentences which may indicate which of the temporal uses of the FOR PP was intended by the writer. The following information was coded for each sentence in the Brown corpus in which the preposition FOR was being used temporally. The features of the PP, being those specified in a standard grammar (Quirk et al., 1985), arranged in order of occurrence in the PP as in *Except for most of the first glorious year of his Presidency*, were:

- Any pre-modifier of FOR (16):[1] *not, except, only, save*.
- Any pre-determiner of the NP in the FOR PP (68): a fraction (*of*), a cardinal number *of, all, only, even, most of, much of, longer than, more than, about, approximately, nearly, almost, up to, at (the) least, over*.
- Any determiner of the NP: definite, indefinite, demonstrative, possessive, quantifier, none or zero. When there was no determiner a choice was made, by hand, whether to give a 'none' or a 'zero' entry. A 'none' entry was chosen if the noun was a noun, usually proper, which could not take any determiner, i.e. a particular year, month or date. Otherwise a 'zero' entry was chosen. The reason for this distinction is that the 'none' entries behave in the same way as definite determiners, whereas the 'zero' entries behave in the same way as indefinite determiners.[2]

  The distinction between 'definite' and 'indefinite' was important in distinguishing between Duration and Interval uses of FOR. The 'demonstrative' and 'possessive' entries all had the same effect as the 'definite' and 'none' determiners. We will refer to them as Definite determiners (302). The 'quantifiers' were the existentials *some* and *any*; they occurred in FOR phrases which were given a duration meaning. Therefore, they were grouped with 'indefinite' and 'zero'; this group will be referred to as Indefinite determiners (704).
- Any post-determiner of the NP (541): (*total of*) *N* [a cardinal number], *Ns of*, an ordinal number, *single, couple of, dozen, number of, few, several, (so) many, matter of, another, more, past, preceding, last, next, only*.
- Any qualifier of the head noun in the NP, e.g. *glorious*.
- The head noun in the NP, sub divided by type (Brée & Pratt, 1997):

  Measure of a temporal duration (865), e.g. *year*;

  Calendar interval (47), e.g. *1966, future*;

  Cyclical (22), e.g. *morning, spring*;

  Part of a longer event (10), e.g. *act, inning, movement, session, spell, stage*;

  Adverbials (25): *ever, long, now, once, tomorrow, tonight*;

  Life (2): *birthday, lifetime*;

  Complex, always with a post-modifier (2): *length, remainder*
- If there was a phrase post-modifying the NP, then the preposition or conjunction introducing this post-modifying expression (126): *of, when, in (advance), during*, any demonstrative, *ended, ending, up to, after(ward), since, following, commencing, starting, or, and, at, each, every, a(n), now, being, to (come), longer, forward, as (a whole), than.*

Certain properties of the verb in the matrix were also noted as these were required to make the distinction between different uses:

Tense (only coded when required): past, present, past/present participle, infinitive;

Aspect (174): perfect (146), progressive (2), perfect progressive (26);

[1]The numbers in parentheses give the number of occurrences in the total of all 1006 examples of temporal use of FOR in the Brown corpus.

[2]Two 'nouns', actually adverbials being used as nouns, never take a determiner, but the determiner was coded as 'Zero' rather than 'None' as they always occurred with Durations rather than Intervals: *long* (17), *ever* (1).

Modal (92): *will* (46), *shall* (5), *can* (30), *may* (6), *have to* (3).

Also with Interval uses, the position of the FOR PP relative to other components of the matrix, in particular its verb.

The categories that turned out to be of most help in discriminating between different uses of FOR were the determiner, the noun itself and the post-modifiers, which we now need to examine in some detail.

3.1. Post-modifiers

As we have seen, the noun in the FOR phrase is frequently post-modified (126). There are of several kinds of post-modifiers: adverbials, conjunctions, demonstratives and other temporal prepositions followed by a noun phrase. Each can be used for different functions depending on whether the determiner is Definite or Indefinite. We now will look in detail at which post-modifiers indicate which function. This section can be skipped without loss of continuity; a summary is given in Section 3.1.3..

3.1.1. With an Indefinite determiner

When the determiner was Indefinite (82), usually associated with one of the Duration uses (see Section 4.3.), the post-modifying phrase has one of the following functions:

- It can simply add information about the duration (22), usually just to say that a *period* is temporal (*period of time*, 7), but also to indicate more precisely its length (using *period of*, 4; *as long as*, 4) or even its nature (*of*, 5):

  j11:081 ... many have been made FOR a very short *period of time*.
  j08:054a All samplers were operated FOR a *period of* two hours ...
  j16:012b ... until used they were stored FOR *as long as* 2 weeks.
  f11:076 An average national figure FOR two to three years *of* treatment would be $ 650 to $ 1,000 .

- It can signal that the duration is generic (18), either directly (*at a time*, 2) or by giving the cyclical frequency (*a(n)*, 13; *every*, 3; *each*, 2):

  k14:074 He worked FOR two hours *a* day ...
  g37:010 ... the missionaries skipped FOR hours *at a time*.
  f04:117 Rugged outdoor exercise FOR an hour and a half *every* day

- It can link the duration to the `tor` (14), either as preceding it (*now*, 10) or following it (*to come*, 6; *longer*, 3; *forward*, 1)

  b08:066 She's been in and out of my house FOR a dozen years *now* ...
  b15:005b The demand for these lots can be met FOR some time *to come*.
  p07:005 Spencer was quiet FOR a moment *longer* ...
  h02:081a ... they must plan their own complex investment programs FOR at least 5 years *forward* ...

- It can link the duration to an interval other than the `tor` (20), either being in the interval (*in*, 6; *during*, 2; a demonstrative, 2) or beginning at the interval (*after(ward)*, 3; *since*, 1; *following*, 3; *starting*, 1; *commencing*, 1):

  a28:031 Except FOR a few months *in* late 1960 and early 1961 ...
  b23:087a A truth-revealing crisis erupted in Katanga FOR a couple of days *this* month ...
  k29:016 FOR many nights *afterward* ...
  j56:088a FOR a time *following* the abandonment of the local plant ...

- It can even alter the extent of the duration given by the head noun in the FOR PP (11), either by giving an alternative duration (using *or*, 5), by including another type of duration (*and*, 1) or by setting the given duration as a maximum (*as long as*, 4) or a minimum (*more ... than*, 1):

  f43:021 FOR a moment *or* two, both scenes are present simultaneously ,
  r06:063 ... her husband hasn't been home FOR two days *and* nights."
  j16:012b ... until used they were stored FOR *as long as* 2 weeks.

3.1.2. With a Definite determiner

When the determiner was Definite (44) the post-modifying phrase has one of the following functions:

- To specify the interval over which the count FOR *the N^{th} time* is to be taken (21), using *in* (14), a demonstrative (4) or *since* (3):

  l12:154 FOR the first time *in* his life ...
  n23:015 And FOR the hundredth time *that* week, he was startled at her beauty.
  r01:035 Then, FOR the first time *since* his arrest ...

- To specify the larger interval from which a part is selected by the head noun, using *of* (7):

  p03:078 FOR the first few months *of* their marriage ...
  r08:022 FOR the remainder *of* the movie ...

- When the head noun is a Measure noun, to specify the end of the duration thus selecting an interval and making the FOR NP Definite (5), using *ending/ed* (3), *up to* (1), *following* (1):

  h27:001 Sales and net income FOR the year *ended* December 31, 1960 ...
  h29:011 FOR the year as a whole, retail sales of TV sets ...

- To describe a further attribute of the interval which enables it to be identified uniquely (6), using *of* possessively (5), or *when* with a clause (1):

  a05:025b ... come up with recommendations for possible changes in time FOR the next session *of* the General Assembly.
  h15:081 A flashlight or electric lantern also should be available FOR those periods *when* a brighter light is needed.

- Other (5): relating an imprecise interval to the `tor` (*to date*, 1; *time being*, 1), giving a range (*to*, 1) or giving emphasis (*as a whole*, 2):

  a28:013 FOR the year *to date* ...

h13:008 ... total annual rail commutation dropped 124 million FOR 1947 *to* 1957.
h29:011 FOR the year *as a whole*, retail sales of

3.1.3. In summary

Most of the post-modifiers can be used with both a Definite and an Indefinite determiner:

- *of*:
  Indefinite: property of the duration,
  Definite: possessive, after a temporal part noun, to indicate which larger interval it is a part of;
- *in, during*: and the demonstratives *this, that*:
  Indefinite: linking the duration to an interval,
  Definite: to give the interval over which the count in *the* N^{th} *time* is to be taken;
- *ending, up to, after, since, following, commencing, starting*: after a Measure noun, indicating that the duration ends or begins at the interval in the post-modifier PP (the FOR PP may be either Definite or Indefinite);
- *or, and*: coordination of either a duration or an interval.

Certain post-modifiers only appear after an Indefinite determiner:

- *at a time, each, every, a(n)* indicate that the duration is generic (and give its frequency);
- *now, to come, longer, forward* attach a duration to the `tor`;
- *as long as, more ... than* affect the extent of a duration.

This concludes the description of the various features in a temporal FOR PP, and its matrix that are likely to indicate which of the different possible uses of temporal FOR was intended by a writer. Most attention has been given to the post-modifying phrase as this is the least well understood part of the FOR PP. While the post-modifier is important, it is by no means the only feature of significance. We will see that the determiner, the qualifier and the verbal aspect also influence the use.

4. Distinguishing uses

We turn now to the task of distinguishing between the different uses of temporal FOR phrases, using the features introduced in the previous section.

There are 9482 instances of FOR in 7710 sentences (out of a total of 52355 sentences) in the Brown corpus. If the noun following the FOR was one that is known to indicate temporal use (Brée & Pratt, 1997), then it was included in the sample. Correspondingly, if the noun clearly indicated another use, e.g. purpose, then it was discarded. For all the remaining instances, the complete sentence was inspected to determine whether or not there was a temporal use. All the temporal instances were classified by use. Finally, the features, as given in Section 3., were coded. The numbers of instances of each temporal use of FOR in the Brown corpus are shown in Table 4.

We begin with the two most distinctive uses, idiomatic and special. Then we will see how to distinguish Durations from Intervals. Finally we will look at the different types of Duration and Interval use.

Use	Instances	Sub-totals
pure duration	212	
`tor−`	168	
`tor+`	265	
All Durations		677
interval: Count	8	
interval: Identify	94	
interval: Universal	73	
All Intervals		175
Special		97
Idiomatic		57
Total		1006

Table 1: Temporal uses of FOR PPs in the Brown corpus

4.1. The idiomatic use

The idiomatic use of FOR (57) in which the noun *time* indicates a position in a sequence rather than the passage of time is readily detected. All occurrences of the singular noun *time* with a post-determiner were idiomatic use. The post-determiner was usually an ordinal number (54) but could also be *last* (2) or *only* (1):

j64:033 FOR *the only time* in the opera, words are not set according to their natural inflection ...

4.2. The special use

Detecting the special use of FOR (97) to give the interval in which a consequence of the matrix will take place, is not simple. Simple features turn out not to be useful. For example, usually (78/97) the determiner is Definite, but Definite determiners are more frequently used (167) for simple intervals.

It is the matrix verb that indicates the special use of FOR. These verbs all have some element of purpose, planning or prediction. Here are examples of such uses grouped by these features. Note that there is no hard and fast way to separate these verbs into these three categories.

planning i.e. carrying out an activity at the `tor` that sets up a later event or state in the interval given by the FOR PP (46), e.g. *be available, have, issue* something, *invite* somebody, *plan, schedule, set*:

a18:081 Their Majesties ... have jointly *issued* invitations FOR Shrove Tuesday evening at midnight.
a42:006 ... Mr. Kennedy *invited* Stevenson to Cape Cod FOR the weekend.
a43:051a Higher tolls are *planned* FOR July 1, 1961, ...
l12:101 Andy's performance was *scheduled* FOR eleven o'clock.
a18:048c A preview party for sponsors of the event and for the artists is *set* FOR April 8.

predicting from what is known at the `tor` that some event or state will occur at a later time given by the FOR PP (8): *auger well/badly, be in (prospect), be out, continue, leave, realize, suggest* something (once each):

b20:076 He has, moreover, another qualification which *augurs* well FOR the future.

b10:048 To me, Brandt looks as though he could *be in* FOR a fine year.

h29:008 I believe a further gain *is in prospect* FOR 1961.

a23:004 ... Multnomah, as of Aug. 22, had spent $ 58,918 out of its budgeted $ 66,000 in the category, *leaving* only $ 7,082 FOR the rest of the month.

j39:091b ... and to appraise the forecast that its interpretation *suggests* FOR the future of farm prices over the years immediately ahead.

purpose in which some activity is conducted at the `tor` in order to achieve some state in a time period given in the FOR PP (33), e.g. *be at, be out, bed down, develop* something, *go (back/to), live, prescribe, protect, shut up, stash (away)*:

p11:030 I made a lemon sponge, ..., so there *would be* something nice in the icebox FOR the weekend.

k12:002 Once, they *were at* Easthampton FOR the summer.

l02:107 ..., but when you ... know that you can't eat until he's *bedded down* FOR the night, ...

e28:104b Does your company have a program for selecting and *developing* sales and marketing management personnel FOR the longer term?

p20:162 We *went* to the Louvre FOR a few hours.

l11:082-4 "You *live* in the present?" "In the present', Felix proclaimed. "FOR the future. ..."

f34:027 The artist who paints in oil uses drying oils to carry the pigments and to *protect* his finished work FOR the ages.

k26:027 The women ... sounded like chickens *shut up* in a coop FOR the night.

l24:144 I've got a little *stashed* FOR a rainy day, and I guess this is rainy enough.

Clearly this list of verbs is incomplete. Nor is it obvious what test should be applied to verbs to decide whether or not they can indicate a special use of a temporal FOR phrase. They should admit of prediction, purpose or planning, but how to develop this insight into a test is not clear.

4.3. Duration versus Interval

Having distinguished idiomatic and special uses of temporal FOR, we turn to Duration and Interval uses (852). Both uses can be one-off or generic. Since distinguishing between one-off and generic sentences is a general problem whose solution does not depend on the FOR PP, we will not attempt to make the distinction here. Instead we now examine how to make the main distinction, that between Duration uses on the one hand and Interval uses on the other. In general this is easy: an Indefinite determiner indicates a Duration use (8/685), a Definite determiner indicates an Interval use (4/167). However, there are exceptions.

Very occasionally (8/685) an Indefinite determiner does not indicate a Duration use. This is when:

- There is the universal quantifier *all* as a pre-determiner (4):

  j19:044 The managers stay the same, so that A[fj] is the same FOR *all* weeks.

  f30:044 Probably a lawyer once said it best FOR *all* time in the Supreme Court of the United States.

  These four sentences are the only ones in which *all* occurs as a pre-determiner. Hence it is a reliable if infrequent indicator of an interval use.

- A duration is made a particular interval by being directly attached to the `tor`, using the post-determiners *last, next* (2):

  b06:083 South Viet Nam's rice surplus FOR *next* year ... may have been destroyed.

  a27:024 A substantial rise in new orders and sales of durable goods was reported FOR *last* month.

  This is similar to, but different from, the attachment of a duration to the `tor`, indicated by a perfect aspect or modal (see Sections 4.4. and 4.5.). The difference arises as the noun phrase itself is sufficient to detect that there is attachment to the `tor`. In fact, while there is no determiner in these examples, a definite determiner is acceptable: FOR *the next/last month*.

- A duration is made a particular interval by being directly attached to a point of time, using certain postmodifying expressions, such as *commencing, following* (4):

  h07:095 ... they could levy taxes FOR an interim period of nine months, *commencing* with September 30 and ending with June 30.

  a03:054 Full payment of nursing home bills FOR up to 180 days *following* discharge from a hospital.

  d08:007 ... which succeeded in reuniting China and keeping it together FOR a longer period (*from* 202 B.C. *to* A.D. 220).

  a23:002 FOR a second month *in a row*, Multnomah County may be short of general assistance money ...

  This is somewhat different from the previous two cases, as here the determiner can't be definite. Moreover, several other gerundive post-modifiers would have this same effect, e.g. *beginning, starting, ending*.

- There is no determiner, but a definite determiner could have been used (2):

  a26:086b ... estimated sales of domestic cars in the U.S. FOR first three months of 1961 were ...

  j37:022a ... the rumors of election dates appeared once again, first FOR spring of 1958 and later for the summer.

  In a26:086b it could be argued that a definite determiner should have been included. But j37:022a is quite acceptable as it stands. In fact, a definite determiner may be dropped with a season, and the season then acts as a proper noun.

Only 4 of the FOR PPs with a Definite determiner turned out to have a Duration rather than an Interval use, all but one with a superlative in the FOR PP.[3] Superlatives can be definite but still indicate a Duration use, here always `tor+`:

[3] The one exception was:

p25:068 Richard's dark eyes came up and seemed FOR *the tiniest moment* to reflect sharp light.

All the superlative qualifiers indicated `tor+` use.

In summary, our simple rule for distinguishing between Durations and Intervals needs to be elaborated to:

- If the determiner is Indefinite, then there is a Duration use, unless:
  - there is the pre-determiner *all*;
  - *last* or *next* are qualifiers, indicating that the duration is attached to the `tor`;
  - there is a post-modifier that attaches the duration to some other time point, e.g.: *beginning, commencing, following, starting, ending*;
  - there is no determiner and the noun is a season;

  when there is an Interval use.
- If the determiner is Definite, then there is an Interval use, unless there is a superlative qualifier when there is a Duration use.

We now turn to distinguishing between the different Interval and Duration uses.

4.4. Distinguishing `tor−` from other Duration uses

The duration can be attached to the time axis to indicate an interval ending at the `tor` (`tor−`) (168). Can this `tor−` use be distinguished from other Durations? A likely indicator is the perfect aspect as it is used to place an event or state before the `tor`. While perfect aspect was an indicator of `tor−` use with IN PPs other features of the matrix, sometimes in conjunction with the perfect aspect, were better indicators: a negative, superlative or ordinal in the matrix (and sometimes if there was a cardinal number and perfect aspect), or if the IN PP was topicalized (Brée, in press). Turning to the data for FOR PPs we find, by contrast:

- The perfect aspect frequently (150/165) did indicate a `tor−` use:

  f31:080 They *have* not been friendly FOR years".
  n09:100b Tom *had* been laying for Aaron McBride FOR a long time ...

  However, 15 FOR PPs occurring in a matrix with a perfect aspect weren't `tor−` use. One reason is that a perfect aspect with a past tense is used to indicate a time earlier than the `tor` rather than the time just prior to the `tor` (10/15):[4]

  k28:165 After that they *had* sat FOR five minutes without saying a word.

  If the past perfect aspect is being used consistently, e.g. as it is in the text preceding k28:165:

  k28:161 "I'll give you ...", Miss Ada *had* said.

---

j04:095 The temperature was maintained to within about A[fj] FOR *the* period of time required to make the measurement (usually about one hour).

[4]In all the remaining 5 examples it is a present tense perfect aspect that fails to indicate a `tor−` use. It seems unlikely that these examples could be excluded without resorting to domain knowledge.

Rule	When use is: `tor−`	all	Cumulative: OK	miss
Matrix is an *after* phrase	9	9	9	0
Perfect aspect in matrix	150	165	159	15
Matrix verb is past participle	5	7	164	17
Out of a maximum of:			168	509

Table 2: Evaluation of the heuristic rules for `tor−` use

  then it is likely that the past perfect in the matrix is also setting the time of the event to before a past `tor`, rather than to the time immediately preceding the `tor`. The past perfect aspect occurs much more frequently with the `tor−` use than with the other Duration uses (65/10), so does not in itself provide an exclusion heuristic.

- The FOR PP was embedded in an *after* phrase (9), generally with a present participle (7):

  a17:001 *After being* closed FOR seven months, the Garden of the Gods Club will have its gala summer opening Saturday, June 3.
  f16:118 Once, *after* the Discovery lay FOR a week in rough weather ...

  Inserting a perfect aspect in these after phrases is quite possible, but clearly not necessary. All occurrences of an Indefinite FOR PP in a *after* phrase were `tor−` use.

- If the verb in the matrix was a past participle then there was usually (5/7) a `tor−` meaning:

  a12:074 Halfback Bud Priddy, *slowed* FOR almost a month by a slowly-mending sprained ankle, ...

  It is not possible to have a perfect aspect with a past participle:

  a12:074′ Halfback Bud Priddy, *had* slowed FOR almost a month by a slowly-mending sprained ankle, ...

  The 2 instances of a past participle in the matrix which were not `tor−` meaning were both generic durations. Distinguishing generic from episodic sentences is non trivial!

These rules explain 164 of the 168 instances of `tor−` use, but without any means for detecting that the past perfect had been established as the tense of story, they would also include 17 of the other 509 Duration uses (see Table 4.4.). So there is no hard and fast rule for distinguishing `tor−` use from other Duration uses, but the following heuristics would be most useful. The duration of FOR PP is likely to end at the `tor` if:

- there is a perfect aspect in the matrix (unless the perfect past has been established as the tense of use); or
- the matrix is an *after* phrase (generally one with a present participle rather than a full tense); or
- the tense verb in the matrix is a past participle.

4.5. Distinguishing `tor+` from other Duration uses

Attachment of the duration to after the `tor`, the `tor+` use, is also possible and indeed very frequent, occurring in almost half the Duration cases (297). With temporal IN PPs, the rules for distinguishing `tor+` use from pure durations were heuristic, so the same is to be expected with FOR PPs. The indicators of the `tor+` use with FOR PPs that were found are:

- Topicalization of the temporal PP, without a perfect aspect, always signalled `tor+` use with IN (Hitzeman, 1997; Brée, in press). The `tor+` use of IN is unusual in that the time of the matrix event is set to be at a time interval which is at the end of a duration, given by the IN PP, after the `tor`, rather than within an interval starting at the `tor` and lasting this duration, as one might expect. But even so topicalization without a perfect aspect also usually indicated `tor+` use with FOR PPs (62/80):[5]

  l10:172 FOR exactly one week, she was able to continue in this manner.

  but not always (18):

  a19:034 FOR a number of years the board used a machine to keep a permanent record but abandoned the practice about two years ago.

- *Then* is a marker for the `tor`. When *then* is in the matrix (15) or begins the clause that follows the matrix (which may be the next sentence) (53) then `tor+` is almost always (68/73) indicated:

  g67:080 *Then* he kept Blackman awake FOR more than an hour . . .
  f31:098a She worked as a domestic, first in Newport FOR a year, and *then* in . . .

- The vague noun *time*, without a qualifier or post-modifier, was an indicator of `tor+` use with IN, as in the expression *in time*. With FOR, several nouns were found more frequently with a `tor+` use than with a pure duration (177/203), e.g.: *instant* (10/11), *moment* (56/63), *second* (14/16), *minute* (50/63), *(a)while* (28/29), and *time* when there was no qualifier (19/21):

  n29:174 The kid showed FOR an *instant*, and his arm was cocked back.
  n07:110 Barton waited FOR a long *moment* . . .
  n12:049 He studied the problem FOR a few *seconds* and thought of a means by which it might be solved.
  p20:079 She was silent FOR a *while* . . .
  l14:048 Detective Pearson, Eighteenth Precinct, thought FOR *a time* he might be on to something.

  Why these nouns? *While*, like *time*, refers to only a vague duration, so is not useful for specifying the duration of an event unless the duration is attached to the `tor`. The other nouns (*instant, second, moment, minute*) refer to short or very short durations, almost too short for any ordinary event to take place within; so they are not often used for that purpose. When they are, it is usually in a generic context, where there is no explicit `tor`:

  k25:040 He looked at her out of himself, she thought, as he did only FOR *an instant at a time* . . .
  k25:048a *Every* few minutes she would awaken FOR a *moment* to review things . . .

  or there is a pre-modifier:

  g35:004 *Not* FOR a *moment* do we forget that our own fate is firmly fastened to that of these countries . . .

- The post-determiner *another* (6) and the qualifier *additional* (1), always indicated `tor+`:

  f28:005b . . . it would be best for all if the plantation were operated FOR *another* year.
  j27:054b . . . the group felt a topic under study should not be dropped FOR an *additional* week . . .

- Some post-modifiers, such as *to come* (6), *longer* (3), *afterward* (1), also always indicate `tor+`:

  h03:027 May the Divine Speaker in Heaven bless this country with Sam Rayburn's continued service here FOR years *to come*.
  l12:102 He stalled FOR a half-hour *longer* . . .
  k29:016 FOR many nights *afterward*, the idea of . . . would return

Rule	When use is: `tor+`	all	Cumulative: OK	miss
another, additional	7	7	7	0
to come, longer, afterward	10	10	17	0
Noun: *time* (not qualified)	19	21	35	2
Nouns: *instant, second*	52	56	85	5
then in matrix or following	68	73	135	10
Noun: *moment*	56	63	176	15
Noun: *minute*	50	63	210	25
FOR PP topicalized	62	80	219	42
Out of a maximum of:			297	212

Table 3: Evaluation of the heuristic rules for `tor+` use

The affect of applying these rules together, rather than individually, is shown in Table 4.5.. The rules are applied in order of most effective first, starting with the post-determiners, qualifiers and post-modifiers that always indicate a `tor+` use, then adding in those nouns that are most indicative, then the presence of *then*, then the remaining nouns and finally topicalization. The additional effect of each rule can be seen in the two rightmost columns. Note in particular that, although topicalization was a good indicator of `tor+` use with IN PPs and that it frequently occurs with this use of FOR PPs, it detracts from the effectiveness of the rule set, adding only 9 `tor+` uses but 17 pure duration uses; so topicalization, while a good heuristic in itself, is not a good heuristic in conjunction with the other rules in the set.

[5] The numbers in parentheses indicate the number of `tor+` occurrences out of the total of `tor+` plus pure duration occurrences. Occurrences with `tor−` are not included as `tor−` usage is almost always indicated by a perfect aspect or a matrix which is an NP.

To explain the remaining instances of `tor+` use, further heuristics are necessary, for example the nature of the verb in the matrix. One of the heuristics useful for detecting `tor+` use with IN PPs was the presence of one of the modals *will*, *shall* or *may*. However, this heuristic is not specific enough (25/ 51) to be useful with FOR.

We have selected the following heuristics (in addition to there being no perfect aspect, the matrix not being an *after* phrase and the adverbial *now* not being in the matrix, all of which indicate `tor−` use) as indicating `tor+` rather than other Durative uses:

- there is a post-determiner or qualifier in the class of words such as *another*, or
- there is a post-modifier in the category of words such as *to come*.
- the noun in the FOR phrase is very short (*instant*, *second*) or vague (*while*, or
- the noun in the FOR phrase is *time*, unqualified, or
- *then* appears in the matrix or begins the following clause (which may be the next sentence), or
- the noun in the FOR phrase is shortish: *moment*, *minute*.

This concludes the rules for distinguishing between different Duration uses. The curious reader can check them for himself by taking any simple sentence in which there is a temporal FOR PP with a pure duration use and altering it in each of the above ways. They will find that the pure duration use switches to a `tor+` use. The skeptical might like to put this hypothesis to a psychological test.

4.6. Distinguishing between different interval uses

We now turn to the distinguishing between the three different Interval uses. The main distinction to be made is whether the FOR PP is attached to a VP (73) or an NP (94). There are only a few (8) count uses.

Detecting NP attachment is a classic problem. In this context the noun to which a FOR PP is attached is usually a regularly recurring event or object, such as financial figures (see Section 2.2.2.). However, there are many such types of noun and it is not clear how to use this information to detect that a FOR PP is attached to an NP. It can sometimes (63/94) be detected by the following syntactic heuristics:

- Always whenever the FOR PP occurs after a noun but before the verb (41):

  h24:024 His return FOR the period January 1 to June 20, 1961, is due April 16, 1962.

- The FOR PP occurs not only after the verb but also after another PP (18):[6]

  b27:081 Sir Robert Watson-Watt wrote, *on* page 50 of SR Research FOR 4 March 1961: ...

- Whenever the FOR PP follows one of a sequence of nouns after the verb (4):

  h30:035 An alert dean will confer all through the year on personnel needs, plans FOR the future, qualifications of those on the job , and ...

There are various other clues, but there will remain about a third of the sentences, in which the FOR PP follows a single NP after the verb but which should be classified as NP attachment:

e28:098 Have you estimated your sales manpower needs FOR the future ...

Of course, if the verb in the matrix is an event, as in e28:098, then the FOR PP can't be given VP attachment. This would require classifying the verbs by Vendler's types (Vendler, 1957), a difficult but not impossible task. However, recall that attachment to the VP can only occur under rather restricted circumstances (Section 2.2.1.):

- The specification of the interval, signalled by a Definite determiner, includes its duration, attaching the duration to the `tor` using the post-determiners *past*, *last*, *next* (25/35):

  b16:038 My husband's hours away from home FOR the *past* years have been ...

- The FOR NP picks out a particular duration from part of a longer interval (24/31),[7] using:
  - the N^{th} interval of that duration, usually the *first* (11/16);
  - the end, using *rest of*, *remainder of* (6/ 7);
  - (almost) all, using *all*, *most*, *much*, *whole of* (8/ 9)

- The temporal noun is one that specifically includes the `tor` (13/13): *moment, present, now*:

  b02:067 ...this approach might be expected to head off Mr. Khrushchev FOR the *moment*.

These three heuristics pick out 70 of the 81 VP attachments, but also include 16 cases of NP attachment. This is more restrictive than the conditions for attachment to an NP, so it would be simpler to look for VP attachment rather than NP attachment. Fortunately, the heuristics for detecting NP attachment detect all but 6 of these 16 cases. So once it has been signalled that VP attachment is likely, a check should be made that after all it isn't NP attachment.

Finally, the occasional use (8) to indicate an interval over which a count is taken is easily spotted by there being a cardinal in the matrix giving the count. Clearly there can also be a cardinal in the matrix with VP attachment, but this happened not to be present in this FOR extract from the brown corpus.

An overview of the effect of each rule, starting with the rule for the count use, is shown in Table 4.6.. Note how adding the simple rule that a FOR PP before the verb must be NP attachment reduces considerably the misclassification as VP attachment. Adding the two additional rules for identifying miss classifications of NP attachment as VP

[6] There are 2 exceptions which are VP attachment:

n26:046 Blue Throat, who had ruled the town *with* his six-shooter FOR the last six months ...

a43:036 Operating revenues were off in the first three months of 1961, but up FOR the 12 months ending in March.

[7] The number after the forward slash gives the number of occurrences in all the interval readings taken together, here 31.

Use	Rule	When use is: this use	all	Cumulative: OK	miss
Count	Cardinal in matrix	8	8	8	0
VP attachment	Noun: *moment, present, now*	13	13	21	0
	Qualifier: *rest/remainder of*	6	7	27	0
	Pre-determiner: *all, much/most/whole of*	8	9	35	1
	Post-determiner: Nth	11	16	45	6
	Post-determiner: *past, last, next*:	25	35	70	16
Count and VP attachment		In total:		81	94
NP attachment	FOR PP before verb	41	41	111	9
	Sequence of nouns	4	4	115	8
	FOR PP after another PP	18	20	133	6
	NP attachment as default			164	6
All Interval uses		In total:		175	175

Table 4: Evaluation of the heuristic rules for Interval Uses

attachment helps only marginally. So the only rule that I propose for identifying NP attachment is the obvious one that the FOR PP is before the matrix verb. Then anything already identified as some Interval use but that isn't count use or VP attachment is, by default, NP attachment.

4.7. Summary

Bringing this all together we see that some uses of temporal FOR phrase can be readily distinguished:

- The expression *for the Nth time* indicates the idiomatic use;
- The special use is identified by the presence of predictive, planning or purposeful verbs in the matrix.
- *A Definite determiner indicates an Interval use:
  - provided that the adverbial 'nouns', such as *long, ever* are not classified as proper nouns, even though they can never take a determiner;
  - *unless there is a superlative qualifier.
- *An Indefinite determiner indicates one of the Duration uses, unless:
  - there is the pre-determiner *all*; or
  - *there is a qualifier that attaches the duration to the `tor`, e.g.: *last, next*; or
  - *there is a gerundive post-modifier that attaches the duration to some time point other than the `tor`, e.g.: *beginning, commencing, following, starting, ending*; or
  - there is no determiner and the noun is a season;

  in which cases there is an Interval use.

Other uses are difficult to distinguish, but the following heuristics would be of help. For distinguishing between the different Duration uses:

- Attachment of a duration to before the `tor` rather than pure duration is indicated by:
  - *a perfect aspect in the matrix, *unless the perfect aspect has been established as the tense of use; or
  - the matrix is an *after* phrase (generally with its verb in present participle form); or
  - the verb tense in the matrix is a past participle.
- Attachment of a duration to after the `tor` rather than pure duration is indicated by the following heuristics:
  - there is a post-determiner or qualifier in the class of words such as *another*; or
  - there is a post-modifier in the category of words such as *to come*; or
  - the noun in the FOR phrase is very short (*instant, second*) or vague (*while*); or
  - the noun in the FOR phrase is *time*, unqualified, or
  - *then* appears in the matrix or begins the following clause (which may be the next sentence); or
  - the noun in the FOR phrase is shortish: *moment, minute*.

To distinguish between the different Interval uses:

- Use of the FOR PP to specify an interval over which a count is to be made is indicated by there being a cardinal in the matrix.
- Attachment of an interval to the VP rather than an NP is indicated by:
  - the specification of the interval includes its duration which is attached to the `tor` using the post-determiners *past, last, next*; or
  - the FOR NP picks out a particular duration from part of a longer interval; or
  - the temporal noun is one that specifically includes the `tor`: *moment, present, now*;

  *unless the FOR PP occurs after a noun but before the verb, when NP attachment is signalled.
- *Otherwise there is attachment of the interval to an NP.

Those rules above that are marked with an asterisk (*) are also those used for distinguishing between the different uses of IN in (Brée, in press). The rules for detecting the idiomatic and special uses are obviously going to be different between these two prepositions. The rules for discriminating Duration from Interval uses show a large overlap. There are some minor differences, most probably due to the size of the samples, that I believe could be avoided by the inclusion of exceptions under both prepositions. Discriminating the `tor−` use from other Duration uses is again largely the same. However, the heuristics for detecting the `tor+` use are very different. (With IN topicalization and modals were the main clues.) The `tor+` use with IN indicates that the matrix holds at a time after the `tor` by the duration indicated by the NP, rather than between the `tor` and this time, as is the case with FOR, so it is unlikely that a common set of heuristic rules will serve to detect the `tor+` use for both prepositions. For the Interval uses, the heuristics for detecting NP attachment and count use are very similar and could be generalised. (No heuristics were provided for detecting VP attachment for IN PPs.)

Clearly providing a common set of (heuristic) rules for discriminating between the different uses (except for `tor+`) for both FOR and IN is the next order of the day. There are also two further prepositions, both infrequent, that have both Duration and Interval uses and so should be examined with the same heuristics in mind: WITHIN, OVER.

5. Conclusions

For such a fine grained analysis as we have made here, a million word corpus is about the maximum size that can be managed by hand; but it is clearly not large enough to give testable results. A corpus of the order of 10 million words, such as the British National Corpus, is needed. To perform this type of analysis such a corpus would have first to be suitably coded. What features would be need to be coded to make the distinctions by the method we have developed?

Some features are easily coded:

- First and foremost the temporal use of FOR, and other prepositions, can be readily detected from the noun in their PP. A list of about 100 such nouns is already available (Brée & Pratt, 1997). Some nouns are marginally temporal, e.g. *in history*, but these are mercifully few. This should overcome the problem, in the latest Penn Tree version of the Brown corpus, of tagging as temporal many non-temporal uses of prepositions .
- The distinction between different types of temporal noun, as given in Section 3., would be useful, in particular, recognising seasons, short time periods, such as *now, present, instant, second, moment, minute* and that some nouns, such as *a while*, give vague times, would be useful.
- The idiomatic use of temporal prepositions needs to be recognised, here simply *for the* N^{th} *time*.
- The contrast between Definite and Indefinite determiners needs to be recognised. Mostly this is easy, but the distinction between a 'Zero' and a 'None' entry needs to be made on the basis of the nature of the noun. Perhaps this is best left to the user, but then a provision has to be made for recognising proper nouns.
- The presence of a superlative in the PP as, while this is Definite, it is not a signal for an Interval use, but for a Duration use.
- The aspect of the verb in the matrix.
- The distinction between pre-determiners, determiners and post-determiners needs to be made along the lines in (Quirk et al., 1985).
- The position of the FOR PP relative to other components in its matrix, in particular the verb.

Other features are more difficult to code:

- To recognise the special use of FOR, the verb in the matrix must be coded as one of prediction, planning or purpose.
- Recognising those qualifiers, such as *past, last, next*, which attach a duration to the `tor`.
- Recognising those post-modifiers, such as *beginning, commencing, following*, which attach a duration to the `tor`.
- Distinguishing the temporal use of *then* to signal the `tor` from other uses, such as inferential.
- It would also assist the analysis if the entire matrix could be categorised as a state, process or event, or something similar to Vendler's types, although this feature has not been explicitly used here.

This is some wish list. But most features would be useful not only for distinguishing between the different temporal uses of FOR but for all the other temporal prepositions.

6. References

D. S. Brée. in press. The semantics of temporal IN. In C. Zelinsky, editor, .

D. S. Brée and I. E. Pratt. 1997. Using prepositions to tell the time. In M. G. Shafto and P. Langley, editors, *Proceedings of the Nineteenth Conference of the Cognitive Science Society*, page 873, New Jersey. Erlbaum.

J. Hitzeman. 1997. Semantic partition and the ambiguity of sentences containing temporal adverbials. *J. Natural Language Semantics*, 5:87–100.

H. Kučera and W. N. Frances. 1967. *Computational analysis of present-day American English*. Brown University Press.

R. Quirk, S. Greenbaum, G. Leech and J. Svartvik. 1985. *A comprehensive grammar of the English language*. Longmans.

H. Reichenbach. 1947. *Elements of symbolic logic*. Macmillan.

Z. Vendler. 1957. Verbs and times. *The Philosphical Rev.*, 66:143–160.

A Language-Neutral Representation of Temporal Information

Richard Campbell*, Takako Aikawa, Zixin Jiang, Carmen Lozano, Maite Melero and Andi Wu

Microsoft Research
One Microsoft Way, Redmond, WA 98052 USA
{richcamp, takakoa, jiangz, clozano, maitem, andiwu}@microsoft.com
*to whom inquiries should be addressed

Abstract

We propose a framework for representing semantic tense that is language-neutral, in the sense that it represents what is expressed by different tenses in different languages in a shared formal vocabulary. The proposed framework allows the representation to retain surface distinctions for particular languages, while allowing fully semantic representations, such as a representation of event sequence, to be derived from it. The proposed framework also supports the incorporation of semantic tense information that does not derive from grammatical tense, but derives instead from other expressions such as time adverbials. The framework is currently implemented in NLPWin, a multi-lingual, multi-application natural language understanding system currently under development at Microsoft Research, but the representational framework is in principle independent of any particular system.

1. Introduction[1]

Multilingual applications face (at least) two problems in the domain of semantic tense: First, there is the problem that grammatical, or morphological, tenses in different languages do not mean the same thing. In English, for example, grammatical past tense situates an event prior to the utterance ("speech time" in Reichenbach's (1947) terminology), and grammatical present tense situates an event simultaneous with the utterance. In contrast Japanese past tense situates an event prior to some reference time, and present tense situates an event simultaneous with some reference time, where the reference time may or may not be the utterance time. Neither language has a tense that expresses exactly what is expressed by past or present in the other. This poses a problem for applications such as machine translation (MT), since a given grammatical tense in one language does not automatically translate into the same surface tense in another language:

(1) 彼女は病気だと言った。
kanozyo-wa [byooki **da**] to itta
she -Top sick be**-Pres** that say-Past
'she said [she **was** sick]'

In (1), for example, the grammatical present tense in the embedded clause (indicated by brackets) translates into English as grammatical past tense, both of which allow the interpretation that the event described in the embedded clause is simultaneous with that described in the main clause.

Another problem is that what is expressed as grammatical tense in one language is sometimes only expressible as an adverbial construction in another language. For example, Chinese has no grammatical tense per se (see Section 3.3 for more details); consequently a single form can in principle express past, present or future; this is illustrated in the following examples:

(2) 昨天他来看我
zuotian ta lai kan wo
yesterday he come see me
'Yesterday he came to see me.'

(3) 明天他来看我
mingtian ta lai kan wo
tomorrow he come see me
'Tomorrow he will come to see me.'

In (2) and (3), the adverbials *zuotian* 'yesterday' and *mingtian* 'tomorrow' are all that indicate that these sentences are set in the past and future, respectively.

In this paper, we propose a framework for representing semantic tense, by which we mean information about the sequence of events. Our framework is *language-neutral*, in the sense that it represents surface tense marking of various languages using a shared formal vocabulary. Our framework also allows the incorporation of semantic tense information that is not expressed as grammatical tense, for example, that (2) is about a past time. Also, since a large part of what is expressed by tenses concerns the sequence of events and states, one aspect of our framework is enabling an explicit representation of temporal sequence. The analyses reported here are currently implemented in the NLPWin system under development at Microsoft Research (Heidorn, 2000).

Most (if not all) other proposals for a language-neutral representation of tense, such as Van Eynde (1997), are explicit attempts to represent the semantics of tense directly. However, the kind of semantic representation of tense may vary considerably depending on application. For example, some applications may require tense to be represented in first-order predicate calculus, perhaps incorporating Davisonian event arguments (Davidson, 1980), while others might require only an explicit sequence of events, as in Filatova and Hovy (2001).

The novelty of our approach lies in the fact that it does not attempt to be a particular semantic representation. Our goal is to preserve syntactic information about semantic tense so that various semantic representations of

[1] We would like to thank three anonymous reviewers and our colleagues in the Natural Language Processing group at MSR for their helpful comments and discussion, especially Michael Gamon, Marisa Jimenez, Jessie Pinkham and Hisami Suzuki.

tense can be constructed if necessary for a particular application. For example, our representation is compatible with both the referential theory of tense (e.g. Enç, 1987) and the quantificational theory of tense (e.g. Ogihara, 1995). Also, although it does not express sequence of events directly, a representation of such a sequence can be derived from our representation.

Our framework owes much to Reichenbach (1947); but while a strictly Reichenbachian approach to tense may work well for European languages, such an approach becomes unwieldy when faced with a set of languages with more typologically diverse tense systems, including Japanese and Chinese, aspects of which are discussed below. We therefore do not rely on the Reichenbachian notions of reference and event times, as does e.g. Van Eynde (1997), but adapt what we take to be Reichenbach's essential insights to a wider range of tense systems.[2]

Before proceeding, it is necessary to say something about the terms *tense* and *aspect*, and to lay out what the scope of the paper is. By *semantic tense*, we mean information about how events or situations are sequenced; this includes some of what in some traditions is called aspect, such as the interpretation of the English perfect, etc. It also includes information that may not be recorded by grammatical tense, as shown in (2) and (3). By *aspect*, we mean temporal information that goes beyond temporal sequence, such as (im)perfectivity, progressive, stative, habitual, and the like. In this paper, we are concerned with semantic tense, not primarily with aspect, though some aspectual features are considered in Section 3.3.2, below.

The paper is organized as follows: In Section 2 we outline the general framework of Language-Neutral Syntax (LNS) (Campbell, 2002; Campbell & Suzuki, 2002), within which we situate the current proposal; in Section 3, we lay out our proposal for the representation of semantic tense; in Section 4, we compare our system to other proposals for representing semantic tense; Section 5 offers a conclusion.

2. Language-neutral syntax

In this section we describe the basic properties and motivation for LNS. For more detailed descriptions, the reader is referred to Campbell (2002) and Campbell & Suzuki (2002).

LNS is a level of representation that is more abstract than a surface-syntactic analysis, yet not as abstract as a fully-articulated semantic analysis; rather, it is intermediate between the two. The basic design principle of LNS is that it be close enough to the surface syntax of individual languages to allow reconstruction of the surface structure of a given sentence (i.e., LNS can serve as the input to a language-particular generation function), yet abstract and language-independent enough to allow derivation of deeper semantic representations, where necessary, by a language-independent function. The role of LNS is illustrated schematically in Figure 1.

[2] However, we do use the terms "reference time" and "event time" informally below.

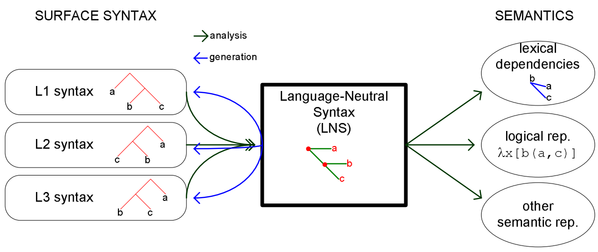


Figure 1: Language-Neutral Syntax

The primary motivation for such an intermediate representation is to mediate between languages in multi-lingual applications, given that fully articulated semantic representations are typically not needed in most such applications. For example, the Adjective + Noun combinations *black cat* and *legal problem* have identical surface structures, but very different semantics: the first is interpreted as λx[black(x) & cat(x)], i.e., as describing anything that is both a cat and black; the second, however, does not have the parallel interpretation as a description of something which is both a problem and legal: rather, it typically describes a problem having to do with the law. To accurately analyze this distinction would require extensive and detailed lexical annotation for adjective senses and, most likely, for lexicalized meanings of particular Adj + Noun combinations; such extensive annotation, if it is even possible, would make a system that depends on it very brittle. For most applications, however, this semantic difference is immaterial, and the extensive and brittle annotation unnecessary: for example, all that we need to know to translate these phrases into French *chat noir* lit. 'cat black' and *probléme legal* lit. 'problem legal' is that the adjective modifies the noun in some way. LNS is a representation in which *black cat* and *legal problem* have the same structure, despite their deep semantic difference, and in which *black cat* and *chat noir* have the same structure, despite their superficial syntactic difference.

An LNS representation is an annotated tree, in which constituents are unordered, and linked to their parent by labeled arcs, the labels corresponding to semantically motivated grammatical functions such as semantic head, logical subject, time, etc. The LNS tree is annotated with semantically motivated features and relations expressing long-distance dependencies (such as binding and control) and discourse-oriented functions (such as topic and focus). An example (somewhat simplified, and with tense not represented for the time being) is given below; this figure represents the LNS for this noun phrase before the implementation of the framework for tense representation presented below.

```
(4)  the cat that was seen yesterday
NOMINAL1 (+Def +Sing)
|_SemHeads--cat1
|_L_Attrib--FORMULA1 (+Pass +Proposition)
                |_SemHeads--see1
                |_L_Sub---_X1
                |_L_Obj---NOMINAL2
                          |_SemHeads--that1
                          |_Cntrlr:  cat1
                |_L_Time-- yesterday1
```

The root node (NOMINAL1) is in the upper left; the daughters of a given node are indicated by labeled arcs such as *SemHeads* (semantic head), *L_Attrib* (logical attributive modifier), *L_Obj* (logical object), and the like. In addition to these attributes indicating deep grammatical relatons, there are other attributes which express additional relations among nodes in the tree. For example, the relative pronoun NOMINAL2 has a *Cntrlr* attribute, whose value is *cat1*, and indicates that *cat1* is the antecedent of the relative pronoun. The Cntrlr attribute is not part of the LNS tree per se; that is, the value of Cntrlr must be part of the LNS tree independently of the Cntrlr relation (in this case, as the semantic head of NOMINAL1). We refer to attributes such as Cntrlr as *non-tree* attributes. For display purposes only in this paper, we display non-tree attributes as labeled arcs, even though they are not part of the LNS tree per se; they will be displayed slightly differently, however, in that the value of the attribute is introduced by a colon, instead of by a dashed line.

In this example we see also that passives are normalized in terms of their argument structure, but the fact that the relative clause is passive is recorded in the feature +Pass on FORMULA1. This reflects a basic design principle of LNS: The basic structure is normalized for variation both within and among languages, but surface distinctions (such as the active/passive distinction) are retained as much as possible.

Thus an LNS representation needs to be close enough to the surface syntax to indicate meaningful distinctions, yet abstract enough to normalize meaningless cross-linguistic variation.

3. Framework for semantic tense

The LNS representation of semantic tense must therefore satisfy the following design criteria:

(5) *Design criteria for LNS representation of tense*:
a. Each individual grammatical tense in each language is recoverable from LNS.
b. The explicit sequence of events entailed by a sentence is recoverable from LNS by a language-independent function.

Criterion (5)(a) says that we must be able to reconstruct, by a distinct generation function for each language, how the semantic tense was expressed in the surface form of that language; this criterion will be satisfied if the LNS representation is different for each tense in a particular language. Criterion (5)(b) says that we must be able to derive an explicit representation of the sequence of events from an LNS representation by means of a language-independent function. This criterion will be satisfied if the representation of each tense in each language is truly language-neutral. In this section we detail a framework for semantic tense that meets the design criteria in (5). We begin by giving the details of the basic formalism (which we will add to in subsequent subsections), followed by a discussion of the motivation and function of its various aspects.

3.1. Basic framework: simple tenses

3.1.1. Tense features and relations

In our proposal each tensed clause contains a distinct Tense node, which is in the *L_Tense* ("logical tense") relation with the clause, and which is specified with semantic tense features, representing the meaning of each particular tense, and attributes indicating its relation to other nodes (including other Tense nodes) in the LNS tree. Semantic tense features can be either *global* or *anchorable*.[3]

The basic tense features, along with their interpretations, are given in the following tables; Table I shows the global features, and Table II the anchorable ones ('U' stands for the utterance time: 'speech time' in Reichenbachian terms):[4]

Feature	Meaning
G_Past	before U[5]
G_NonPast	not before U
G_Future	after U

Table I: Global tense features

Feature	Meaning
Befor	before Anchr if there is one; otherwise before U
NonBefor	not before Anchr if there is one; otherwise not before U
Aftr	after Anchr if there is one; otherwise after U
NonAftr	not after Anchr if there is one; otherwise not after U

Table II: Anchorable tense features

The tense features of a given Tense node are determined on a language-particular basis according to the interpretation of individual grammatical tenses. For example, the simple past tense in English is [+G_Past], the simple present is [+G_NonPast +NonBefor], etc.

Additional features may turn out, on further analysis, to be necessary; for example, many languages make a grammatical distinction between immediate future and general future, or between recent past and remote or

[3] The distinction between global and anchorable tense features is very similar to Comrie's (1985) distinction between 'absolute' and 'relative' tenses. We have adopted the different terminology to emphasize that the global/anchorable distinction is for features, not for tenses per se, as in Comrie's taxonomy.

[4] Note that, given their meanings, some pairs of Tense features are semantically incompatible with each other, and cannot occur on the same node. For example, a given Tense cannot be [+G_Past +G_NonPast].

[5] Strictly speaking the meaning of the global tense features is to express a relation between a given time t and a globally specified reference time, G. Conceivably, the value of G could vary, depending on various factors including genre, discourse context, etc. However, we currently have no theory as to how G might be set to any value other than U, so we will assume throughout that the global referene time is always the same as the utterance time.

general past. We have nothing to say about these specific contrasts, however, other than to note that the framework we propose is flexible enough to accommodate new tense features, if necessary.

A Tense node T will also under certain conditions have a non-tree attribute called *Anchr*, which indicates a relation that T bears to some other Tense node (the value of the Anchr attribute must be another Tense node). Like other non-tree attributes such as Cntrlr, Anchr should be thought of as an annotation on the basic tree, not as part of the tree itself; that is, the value of the Anchr attribute must fit into the LNS tree in some independent way. A Tense node has an Anchr attribute if (a) it has anchorable tense features; and (b) meets certain structural conditions. For simple tenses, the structural condition that it must meet to have an Anchr is that the clause containing it is an argument (i.e., logical subject or object) of another clause; in this case the value of Anchr is the Tense node in the governing clause. In the discussion of compound tenses below we will augment the set of sufficient structural conditions for having an Anchr.[6]

3.1.2. Past tense in English and Japanese

As indicated in Table II, if a Tense node with anchorable features has no Anchr, then it is interpreted as if anchored to the utterance time U. This means that, for example, a [+G_Past] Tense and an unanchored [+Befor] Tense have the same interpretation, all else being equal. Consider the following English and Japanese sentences, with the relevant parts of their LNS structure shown:[7]

(6) She was sick.
FORMULA1
|_SemHeads----sick1
|_L_Tense----_Tense1 (**+G_Past**)

(7) 彼女は病気だった。
kanozyo-wa byooki datta
she -Top sick be-Past
'She was sick.'
FORMULA1
|_SemHeads----病気1 (sick)
|_L_Tense--_Tense1 (**+Befor**)

The English and Japanese past tenses are represented differently because they are semantically different, though in these simple examples that difference is neutralized. The English simple past tense is [+G_Past], indicating that it denotes a time that is before U. The Japanese simple past tense on the other hand is [+Befor], indicating that it denotes a time that is before its Anchr. However, in this simple root sentence, there is no Anchr, so it is interpreted as if anchored to U; hence the interpretation is before U. Thus the design criterion (5)(b) is met, at least for these simple cases: a simple language-independent function would yield the correct sequence *be_sick* < *U* for both these examples.

The semantic difference between the English and Japanese past tenses comes into play when the Anchr attribute is present, which for simple tenses is in clauses that are arguments of a higher clause. Consider the following English and Japanese examples, in which the tense in question (in boldface) is in an embedded sentence (indirect speech), represented in LNS as the logical object (L_Obj) of the matrix clause:

(8) She said she was sick.
FORMULA1
|_SemHeads--say1
|_L_Tense--_Tense1 (+G_Past)
|_L_Obj--FORMULA2
    |_SemHeads--sick1
    |_L_Tense--**_Tense2** (+G_Past)

(9) 彼女は病気だったと言った。
kanozyo-wa byooki datta to itta
she -Top sick be-Past that say-Past
'she said she was sick'
FORMULA1
|_SemHeads--言う1 (say)
|_L_Tense--_Tense1 (+Befor)
|_L_Obj--FORMULA2
    |_SemHeads--病気1 (sick)
    |_L_Tense--**_Tense2** (+Befor)
        |_Anchr: _Tense1

Since the embedded tense in (8) is +G_Past, its interpretation is before U; left unspecified is whether the situation described by the embedded clause (FORMULA2) is reported to have occurred before, or simultaneous with, the situation described by the matrix clause. In fact, both interpretations are possible in this case: her reported sickness may be simultaneous with her saying that she was sick (i.e., she said "I am sick"), or it may have preceded it (i.e., she said "I was sick").[8] The structure we assign to it captures that underspecification succinctly.

In (9), on the other hand, the embedded tense, *_Tense2*, is +Befor; since it has an anchorable feature, and its clause is the logical object of another clause, it must be anchored to the tense of that matrix clause, i.e., to *_Tense1*. Consequently, it denotes a time that is before the time denoted by *_Tense1* (which, like *_Tense1* in (7), denotes a time before U). So the only interpretation (9) has is that her reported sickness is prior to her saying that she was sick; i.e., it can only mean 'she said "I was sick"'; it cannot mean 'she said "I am sick"'. This construction illustrates the essential difference between the English and Japanese past tense forms: the former directly expresses a

[6] We have not ruled out the possibility of language-particular anchoring conditions, but so far have not encountered any need for them.

[7] In this paper we show only the parts of the LNS necessary to illustrate the treatment of tense; for example, we leave out logical subject, etc., unless otherwise necessary. Note also that the copula is regularly omitted from the LNS (see Campbell, 2002).

[8] A third logical possibility, consistent with the interpretation of G_Past, is that her sickness was in the past (i.e., before U), but after her saying that she was sick; i.e., she said "I will be sick". But this kind of interpretation seems to be universally disallowed without some kind of irrealis marking on the clause (such as a modal), and therefore does not need to be separately indicated.

relation to U, while the latter directly expresses a relation to some "reference" time, which may or may not be U.

Examples such as (8) and (9) illustrate precisely why the English and Japanese grammatical past tenses have different representations in the current framework. Suppose for example that the Japanese past tense were [+G_Past] (like the English past), instead of [+Befor]; then Japanese (9) should have the same range of interpretations as English (8), in particular it should be able to serve as a description of an event in which she said "I am sick"—i.e., where the time of her being sick coincides with the time that she said she was sick. As noted, however, this interpretation is not available for (9), as it is for (8).

Our analysis of the English and Japanese past tenses differs from the approach taken by e.g. Ogihara (1995), who claims that English and Japanese past tenses mean the same thing, and that differences such as that between (8) and (9) below are due to the optional application in English of a rule that deletes the embedded past tense from the logical form component. Our analysis gives a uniform description to both the English and Japanese grammatical past tenses.

It is important to note that there is only one sense of the feature Befor (the same holds true for all the anchorable features in Table II), and hence only one meaning for Japanese past tense, in our system. This is a crucial point which is easily overlooked: phrased in strictly Reichenbachian terms, we may appear to be saying that the Japanese past tense means *either* E<R (if it is anchored) *or* E<S (if not anchored). But this appearance of bi-vocalism is due, we believe, to an overly rigid adherence to Reichenbach's notation; our own notation is more flexible, allowing us to characterize the Japanese past tense as univocal, while still retaining what we regard as Reichenbach's essential insights, namely that some tenses relate to U and others to a structurally determined "reference" time.

3.1.3. Present tense in English and Japanese

Another good illustration of the differences between global and anchorable tense features is provided by the English and Japanese present tenses. As in the case of past tense, the two tenses receive the same interpretation in simple sentences:

(10) She is sick.
FORMULA1
|_SemHeads—sick1
|_L_Tense--_Tense1 (**+G_NonPast** +NonBefor)

(11) 彼女は病気だ。
kanozyo-wa byooki da
she-Top sick be-Pres
'She is sick'
FORMULA1
|_SemHeads—病気1 (sick)
|_L_Tense--_Tense1 (**+NonBefor**)

Since the English present tense in (10) is [+G_NonPast] (as well as [+NonBefor]; see below), it must denote a time that is not before U. The Japanese present tense is just [+NonBefor], so it denotes a time that is not before its Anchr; since it lacks an Anchr, in this case, it must denote a time that is not before U. Consequently (10) and (11) receive the same interpretation.

Note that nothing in these representations directly expresses anything about the "present": G_NonPast is interpreted as "not before" U, but does not have to be simultaneous with U. This is by design: the English grammatical present tense allows a future interpretation as well as a "present" one, as in *We speak tomorrow* (see Section 4, below). Our assumption is that present-time reference is the default denotation for any Tense whose features and relations to other time expressions are consistent with that interpretation. Similar comments hold for the Japanese present tense, which is [+NonBefor] in our analysis. As in English, the Japanese present tense also allows a future-time construal (see Section 3.3.3, below).

As in the case of the past tenses, the difference between the English and Japanese present tenses shows up when there is an Anchr:

(12) She said she is sick.
FORMULA1
|_SemHeads--say1
|_L_Tense--_Tense1 (+G_Past)
|_L_Obj--FORMULA2
    |_SemHeads--sick1
    |_L_Tense--**_Tense2 (+G_NonPast +NonBefor)**
        |_ Anchr: _Tense1

(13) 彼女は病気だと言った。
kanozyo-wa byooki da to itta
she -Top sick be-Pres that say-Past
'she said she was sick'
FORMULA1
|_SemHeads--言う1 (say)
|_L_Tense--_Tense1 (+Befor)
|_L_Obj--FORMULA2
    |_SemHeads--病気1 (sick)
    |_L_Tense--**_Tense2 (+NonBefor)**
        |_ Anchr: _Tense1

In this case, both embedded tenses are anchored, since both have the anchorable feature [+NonBefor]. The English present tense is [+G_NonPast], however, so *_Tense2* denotes a time that is not before U; it is also [+NonBefor], so it also denotes a time that is not before the (past) time denoted by *_Tense1*. Consequently, the period of her sickness must overlap both the time of her saying that she was sick and the utterance time U (see also Note 8); in fact, as Enç (1987) notes, this construction has exactly that meaning. This example also illustrates the fact that a given tense may have any collection of mutually-compatible tense features, including both global and anchorable ones.

In contrast, the Japanese example (13) (the same as (1)) does not imply that the period of her sickness includes the utterance time; instead, the possibility that she is still sick at the present moment is left open, unlike (12). In our framework, this is because the Japanese present lacks a global tense feature. *_Tense2* is [+NonBefor] and not [+G_NonPast] like (12), so its only requirement is that it denote a time that is not before the time denoted by its Anchr, *_Tense1*. As indicated in the gloss, the best English translation of (13) is with the past tense. Examples like (12) and (13) illustrate precisely why the

English and Japanese present tenses are to be represented differently.

3.2. Compound tenses

One of the great insights of Reichenbach's (1947) analysis of tense is his treatment of compound tenses, such as the English present- and past-perfect. In this subsection, we outline our representation of compound tenses, which, despite notational differences, is essentially Reichenbachian.

We begin by making a formal distinction between *primary* and *secondary* tenses, the latter being tenses, such as English *have* + past participle, which require an Anchr within the same clause, the former being all others. Thus each language-particular tense must be specified as to its features, and whether it is primary or secondary. Consider the following example of the past perfect in English:

(14) He had arrived.
FORMULA1
|_SemHeads—arrive1
|_L_Tense--_Tense1 (+G_Past)
--_Tense2 (**+Befor**)
|_ **Anchr: _Tense1**

We treat English perfect constructions as consisting of two tenses: a secondary tense that is [+Befor], anchored to a primary tense, in this case simple past (hence [+G_Past]). There is no principled upper limit to the number of Tense nodes in a given clause (though particular grammars presumably impose de facto limits), though the following conditions must be met for well-formedness: (1) each clause has one and only one Tense that is not anchored within the clause (though it may be anchored outside the clause); this is the Tense that designates the "reference" time; and (2) each clause has one and only one Tense which is not the Anchr of another Tense in the same clause (though it may be the Anchr of another Tense in another clause); this is the Tense that designates the "event" time. In (14), the first condition is satisfied by *_Tense1*, and the second condition is satisfied by *_Tense2*. In the simple tense examples discussed in Section 3.1, both conditions are satisfied by the same Tense node.

The advantages of treating the perfect construction as a compound tense, instead of as a simple tense, are two-fold: (1) it allows us to distinguish English present perfect and simple past without additional features (thus helping to satisfy the design criterion (5)(a)); and (2) it captures the fact that the perfect construction co-occurs with every simple tense in English, with the same interpretation.

3.3. Survey of tenses across languages

The framework described above is not a theory of tense, in that it does not uniquely determine a representation for each grammatical tense in each language, but provides a language-neutral vocabulary for expressing differences among grammatical tenses across typologically diverse languages. To implement the framework in an NLP system, then, we need to have actual analyses of specific tenses. In this section we provide such analyses for several tenses in several languages.

3.3.1. English

The discussion above gives examples of the past, present and perfect tenses in English and their combinations. Here we give two more examples of English grammtical tenses: the future with *will*[9] and the past with *used to*.

Future: Though an argument might be made that the future with *will* is actually a compound tense, we take the simpler route here and analyze it as a distinct primary tense with the feature [+G_Future], as in the following example:

(15) You will be sick.
FORMULA1
|_SemHeads—sick1
|_L_Tense--_Tense1 (+G_Future)

Past with used to: The past tense formed with *used to*, as in *he used to work here*, like the simple past tense is [+G_Past], but differs from the simple past not only in aspectual properties (not treated here), but also in that it has the anchorable feature [+Befor]. Consider the following example:

(16) He said he used to work here.
FORMULA1
|_SemHeads—say1
|_L_Tense--_Tense1 (+G_Past)
|_L_Obj—FORMULA2
|_SemHeads—work1
|_L_Tense--_Tense2 (**+G_Past +Befor**)
|_ **Anchr: _Tense1**

Since the embedded *_Tense2* is [+Befor], it denotes a time that is not only before U, but also before the (past) time denoted by *_Tense1*. This reflects the fact that in (16), the time that he worked here must be before the time that he said he used to work here (compare to (8), above); that is, it can only mean that he said "I used to work here", and cannot mean that he said "I work here".

3.3.2. Other European languages

Apart from aspectual differences, the tense systems of Western European languages such as French, German and Spanish are very similar to that of English. The aspectual differences are of course important, and must be represented in LNS. Although a complete discussion of aspect goes beyond the scope of the present paper, we include a brief discussion of some diferences between English and Spanish here.

One notable difference between Spanish and English is that Spanish has two distinct grammatical past tenses, the perfective, or preterite, and the imperfective. The

[9] Needless to say, this is not the only way to express future-time reference in English. The simple present can sometimes be used, and there are at least two other constructions that are future only: *be going to* + infinitive, and *be about to* + infinitive. The latter construction has a different meaning from the others (immediate future), and should be distinguished, perhaps with a feature. The difference between *will* and *be going to* is hard to detect, if it exists at all, but in keeping with design criterion (5)(a) they should be distinguished in some way.

difference is entirely aspectual, and does not appear to affect the interpretation of sequence of events per se. Another notable difference between English and these other languages is that most of them use the simple grammatical present tense to refer to an event ongoing at the utterance time, as in the following Spanish example:

(17) Llueve.
rain-Pres
'It's raining.'

The simple present in English, however, cannot be used this way; English *it rains* has only a generic or habitual sense.

For both of these distinctions, a feature indicating the aspectual difference is used; in our system, the relevant features are *Discrete* and *NonDiscrete*; the former indicating that events are viewed in their entirety, the latter that events are subdivided into arbitrarily small subintervals. Thus the Spanish preterite is [+Discrete], while the imperfect is unmarked for either of these features. Also, the simple present in English is [+Discrete], while the simple present in e.g. Spanish is umarked for this feature.

Aside from such aspectual differences, the most notable tense difference between Spanish and English is that the Spanish present progressive, in contrast to the simple present, is incompatible with future time reference:

(18) Vuelvo mañana.
return-1sg tomorrow
'I return/am returning tomorrow'

(19) Estoy volviendo (*mañana).
be-1sg returning tomorrow
'I am returning tomorrow.'

This is handled by assigning the present progressive the features [+G_NonPast +NonBefor +NonAftr] (in addition to aspectual features), which differs from the simple present in being [+NonAftr]. In (19) there is no Anchr, so the [+NonAftr] feature dictates that the time referred to is not after U; i.e, is not in the future; this accounts for this tense's incompatibility with a future time adverbial.

3.3.3. Japanese

The discussion above gives some examples of the simple past and present in Japanese, analyzed in our framework as [+Befor] and [+NonBefor], respectively. Since there is no separate future tense in Japanese, future time reference is normally achieved with the simple present tense, as in the following example:

(20) 明日雨が降る。
ashita ame-ga furu
tomorrow rain-Nom fall-Pres
'Tomorrow, it will rain.'

FORMULA1
|_SemHeads—降る1 (fall)
|_L_Time—明日1 (tomorrow) (+G_Future)
|_L_Tense--_Tense1 (+NonBefor)

The feature [+NonBefor] on *_Tense1* is compatible with future time reference, as discussed in Section 3.1.3, above. The future, as opposed to present, reading of (20) comes from the presence of the adverbial *ashita* 'tomorrow'. In Section 4, we discuss how semantic tense information from adverbials is incorporated into our framework.

3.3.4. Chinese

Unlike the other languages discussed above, Chinese has no grammatical tense. As noted in the introduction vis-a-vis examples (2) and (3), semantic tense, when expressed, is often expressed via adverbials, and not with grammatical tense; this is discussed in more detail in Section 4, below. However, Chinese does have a limited number of particles, traditionally referred to as aspect markers, which, besides indicating aspect, also indicate semantic tense information. The aspectual meaning of these particles is beyond the scope of this paper, but we will discuss a few examples to show how they express semantic tense, and how that information is represented in our framework.

We discuss here the particles *le*, *guo* and *jiang*, as in the following examples:

(21) 他说他买**了**书
ta shuo ta mai **le** shu
he say he buy Aspect book
'He says/said that he has/had bought books.'

FORMULA1
|_SemHeads--说1 (say)
|_L_Tense--_Tense1
|_L_Obj—FORMULA2
|_SemHeads--买1 (buy)
|_L_Tense--**_Tense2 (+Befor)**
|_Anchr: _Tense1

(22) 他说他买 **过**书
ta shuo ta mai **guo** shu
he say he buy Aspect book
'He says/said that he has/had (once) bought books.'

FORMULA1
|_SemHeads--说1 (say)
|_L_Tense--_Tense1
|_L_Obj—FORMULA2
|_SemHeads--买1 (buy)
|_L_Tense--**_Tense2 (+Befor)**
|_Anchr: _Tense1

(23) 他说他**将**到美国去
ta shuo ta **jiang** dao meiguo qu
he say heAspect to US go
'He says/said that he will/would go to the US.'

FORMULA1
|_SemHeads--说1 (say)
|_L_Tense--_Tense1
|_L_Obj—FORMULA2
|_SemHeads--买1 (buy)
|_L_Tense--**_Tense2 (+Aftr)**
|_Anchr: _Tense1

In all these examples, the tense of the main clause (*_Tense1*) has no features; we take this to be the default case in Chinese, in which an unmarked clause can be interpreted as past, present or future (see the discussion of examples (2) and (3) in the Introduction, and Section 4, below). However, aspectual particles such as *le*, *guo* and *jiang* can also contribute semantic tense information, which we represent as if it were grammatical tense.

The particles *le* and *guo* are both [+Befor] (their difference is aspectual, not represented here); in (21) and (22), the embedded clause Tense is anchored to the matrix, indicating that the buying of books took place before his saying. In contrast, *jiang* in (23) is [+Aftr], so this example means that the going to the US takes place after his saying.

4. Deriving semantic tense from syntactic context

It is often the case that semantic tense information is not represented as grammatical tense per se, but can come, at least in part, from adverbials or other features of the syntactic environment. We have seen that this is one of the main sources of semantic tense information in Chinese; an example from English is *We speak tomorrow*, which is grammatically present tense (hence [+G_NonPast +NonBefor], but semantically is unambiguously about the future. To deal with this situation, we propose to augment the framework outlined in Section 3 with an additional non-tree attribute *Spcfrs*, which indicates, for a given Tense node, any other temporal expressions in the clause that contributes to the semantic tense of that clause. Like Anchr, Spcfrs is not part of the LNS tree per se, but is an annotation on the tree. The representation is given below:

(24) We speak tomorrow.
FORMULA1
|_SemHeads—speak1
|_L_Time—tomorrow1 (+G_Future)
|_L_Tense--_Tense1 (+G_NonPast +NonBefor)
                |_ Spcfrs: tomorrow1

_Tense1 has only the features of any present tense, so the representation satisfies the first design criterion (5)(a); but its Spcfrs is the adverb *tomorrow1*, which itself has the feature [+G_Future], since tomorrow is unambiguously in the future. This relation indicates to the language-independent function that derives the explicit sequential representation that the temporal reference of the clause is to a time that is after U, thus satisfying the second design criterion (5)(b).

Note that design criterion (5)(a) is satisfied in another way, as well: the structure of (24) is different from the structure of a sentence with a future tense, which presumably makes use of the feature [+G_Future] (see below); thus the distinction between the “scheduled” future (Comrie, 1985) in (24) and the more basic future of *We will speak tomorrow* is preserved.

The need for the Spcfrs relation is much more prevalent in languages that make little or no use of grammatical tense, such as Chinese. Consider the following examples:

(25) 昨天他来看我
zuotian ta lai kan wo
yesterday he come see me
‘Yesterday he came to see me.’
FORMULA1
|_SemHeads--来1 (come)
|_L_Time--昨天1 (yesterday) (+G_Past)
|_L_Tense--_Tense1
                |_Spcfrs: 昨天1

(26) 明天他来看我
mingtian ta lai kan wo
tomorrow he come see me
‘Tomorrow he will come to see me.’
FORMULA1
|_SemHeads--来1 (come)
|_L_Time--明天1 (tomorrow) (+G_Future)
|_L_Tense--_Tense1
                |_Spcfrs: 明天1

The Spcfrs relation thus permits specification of semantic tense features that are not expressed as grammatical tense.

5. Comparison to other frameworks

Our proposal is for a system of representation of semantic tense that is language-neutral; i.e., that represents the tense distinctions of different languages in a formal vocabulary that has the same meaning in all languages. As such, our proposal is very different from proposals to represent the semantics of tense in a particular language such as English, both in the obvious respect that we consider other languages, and in the less obvious respect that our proposal is not a semantic one in any deep sense, but rather a syntactic representation that is language-neutral, as sketched in Section 2 (Campbell & Suzuki, 2002).

As such, the nearest thing to a comparable proposal that we have encountered in the computational literature is Van Eynde (1997), which explicitly provides a Reichenbachian semantic framework for multiple languages, and incorporates information from temporal adverbs in addition to grammatical tense. Unlike our proposal, however, Van Eynde’s framework is explicitly Reichenbachian, characterizing tenses in terms of three possible values for sTENSE, expressing the relation between the reference and speech times, and six values for sASPECT, expressing the relation between the event and reference times. Although our framework encodes the same essential insight, it does so without rigidly adhering to the reference time/event time distinction, which leads to a simpler representation in our view.

6. Application

Having a language-neutral representation of semantic tense has clear implications for multi-lingual applications such as MT. Consider again the Japanese example (13), in which an embedded present tense is to be translated into past tense in English. A simple transfer of the language-particular present tense yields the wrong result, since *She said she is sick* (=(12)) means something very different from (13). Instead, what needs to be transferred is the whole temporal structure of *_Tense2*, including its features and its Anchr, since this is the information that determines that it denotes a time that is before U. Such context-sensitive transfers are possible in an MT system such as that described by Richardson, *et al*. (2001).

Similarly, consider the Chinese example (25), in which there is no grammatical tense specified. A Chinese-English MT system must transfer not the grammatical tense (which yields no information whatsoever), but rather the whole temporal structure, which in this case includes its Spcfrs, in order to give the English generation system the information it needs to generate past tense.

7. Conclusion

We have presented and exemplified a framework for representing semantic tense in a language-neutral fashion, which meets the competing design criteria in (5): that each language-particular tense be reconstructible by a generation function, and that an explicit representation of temporal sequence be derivable by means of a language-independent function.

The framework we have proposed allows us to get semantic tense information from grammatical tense, or from adverbial modifiers, and represents this information in a semantically motivated, language-neutral fashion.

8. References

Campbell, R., 2002. *Language-neutral syntax*. MSR Tech Report (in preparation).

Campbell, R. & H. Suzuki. 2002. A language-neutral representation of syntactic structure. SCANALU-2002.

Comrie, B., 1985. *Tense*. Cambridge University Press.

Davidson, D. 1980. The logical form of action sentences. In D. Davidson, ed., *Essays on actions and events*, 105-122. Clarendon Press, Oxford.

Enç, M., 1987. Anchoring conditions on Tense. *Linguistic Inquiry* 18, 633-657.

Filatova, E. & E. Hovy. 2001. Assigning time-stamps to event-clauses. In Proceedings of ACL-EACL.

Heidorn, G.E. 2000. Intelligent writing assistance. In R. Dale, H. Moisl and H. Somers, eds., *Handbook of natural language processing*. Marcel Dekker, New York.

Ogihara, T. 1995. The Semantics of Tense in Embedded Clauses. *Linguistic inquiry* 26, 663-679.

Reichenbach, H. 1947. *Elements of symbolic logic*. The Free Press, New York, and Collier-Macmillan, London.

Richardson S., W. Dolan, A. Menezes and J. Pinkham. 2001. Achieving commercial-quality translation with example-based methods. In *Proceedings of the VIIIth MT summit*, Santiago de Compostela, Spain. 293-298.

Van Eynde, F. 1997. Mood, Tense & Aspect. In F. Van Eynde and P. Schmidt (eds.), *Linguistic specifications for typed feature structure frameworks*. EU Commission, Luxembourg.

A Representational Scheme for Temporal and Causal Information Processing

Eleni Galiotou[1,2], Gérard Ligozat[3]

[1]Department of Informatics, TEI of Athens
Ag. Spyridona, GR-122 10 Egaleo, Greece,
[2]Dept. of Informatics and Telecommunications, University of Athens
Panepistimiopolis, GR-157 84 Athens, Greece
egali@di.uoa.gr

[3]LIMSI-CNRS, Paris-Sud University
Bldg. 508, P.O.Box 133, F-91403, Orsay, France
ligozat@limsi.fr

Abstract

In this paper, we propose a representational scheme for temporal and causal information processing which is based on the exploitation of linguistic elements of narrative texts having a temporal and/or causal interest. We show that the framework of generalized intervals provides an adequate representational device for temporal information as it is conveyed from descriptions in natural language. The linguistic elements which are used in order to determine the temporal structure of the text are: temporal adverbials, verbs, derived nouns, temporal and/or causal connectives. Reasoning over the resulting temporal constraint network allows the disambiguation of ambiguous temporal relations and provides answers to questions concerning the temporal order of events in the text, their relative positioning and their causal interpretation.

1. Introduction

The extraction and processing of temporal information from natural language texts requires an effective representational scheme which facilitates the answer to questions such as: «What happened»?, «Why?», «Did event A take place before event B?», etc. In this paper, we propose such a scheme and we focus on its theoretical justification.

We show that the framework of generalized intervals (Ligozat 1991; Ligozat 1997; Bestougeff and Ligozat 1992) which extends Allen's interval framework (Allen 1983; Allen 1984) provides an adequate representational device for temporal information as it is conveyed from descriptions in natural language. The formalism of generalized intervals allows a direct representation of qualitative temporal relations possessing an internal structure. Therefore, it is suitable for the representation of events which are classified according to the tripartite ontology of events (Moens 1987; Moens and Steedman 1988). This representation allows reference to their internal structure and proves to be well adapted to the representation of temporal information

In our approach, the linguistic elements which are used in order to determine the temporal structure of the text are: temporal adverbials, verbs, derived nouns, temporal and/or causal connectives. The above mentioned elements contribute to determining the order of events in the text, their relative position, their degree of completion and their duration.

Reasoning over the temporal constraint network which reflects the temporal structure of the text allows the disambiguation of ambiguous temporal relations and provides answers to questions concerning the temporal order of events in the text, their relative positioning and their causal interpretation.

The structure of the paper is as follows: In section 2 the theoretical assumptions of the approach are presented. Section 3 is concerned with the representational scheme while in section 4 the final conclusions are presented.

2. Theoretical assumptions

Our choice of a representation language is motivated by the complexity and richness of the linguistic information required for the processing of temporal and causal phenomena. Indeed, most of the well known formalisms such as the interval-based framework proposed by Allen (1983; 1984) prove to be inadequate for the representation of complex linguistic information. In the following we briefly present the formalism of generalized intervals as proposed by Ligozat (1991; 1996; 1997) and Bestougeff and Ligozat (1992).

2.1. The generalized interval framework

In this approach, the main temporal object is a a poly-typed string (henceforth PTS) which consists of a generalized interval, i.e. an increasing sequence of typed boundaries (time points) corresponding to an event, and its associated linguistic information. The typing of the boundaries (opening, closing, undefined) allows to reason over the degree of completion of an action and therefore proves to be adequate for the representation of the grammatical aspect. A sequence of n typed boundaries is called an n-string and it defines a partitioning of time into 2n+1 zones numbered from 0 to 2n. Odd numbers correspond to boundaries and even numbers correspond to open intervals defined by these boundaries. Using this numbering we are able to define all the combinatorial possibilities of relations between two n-strings: Consider a p-string X and a q-string Y. The relations between X and Y is determined by specifying for each boundary of X which zone of Y it belongs to. We thus obtain a non-decreasing sequence of p integers between 0 and 2q.

A sequence of polytyped strings reflecting a consistent piece of discourse is called a temporal site. This temporal site is represented by a temporal constraint network, that is, a graph where the edges correspond to typed strings and the arcs to relations between them.

2.2. Events in the generalized interval framework

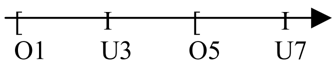
The processing of temporal and causal information as it is conveyed by natural language requires an ontology which is constituted on the cognitive basis of relations such as causation and enablement (grouped under the more general notion of consequentiality) between events and states rather than on purely temporal primitives. An ontology which satisfies the above requirements and is also able to capture most of the relevant linguistic phenomena is the tripartite ontology of events (Moens 1987; Moens and Steedman 1988). The basic element of this ontology is an «event nucleus» which is composed by a preparatory process, leading to a culmination point followed by a consequent state. In this way, we can explain the fact that the same event (for example, the construction of a bridge) can be used by referring to its preparatory process (drawing the plans), to its culmination point (the bridge was inaugurated) or to its consequent state (traffic problems were solved). Temporal and aspectual categories as conveyed by natural language utterances are classified into states and events. Following (Parsons 1990) we will use the term *eventuality* in order to describe an event or a state. Events are further classified into culminated processes, processes, culminations and points (punctual events which are not associated to a preparatory process or consequent state). Reference to the appropriate part of the nucleus determines the exact nature of the event or state at hand.

3.2.3. Representing temporal knowledge

In the generalized interval framework, different types of processes are associated to typical schemata in terms of sequences of typed boundaries.

In the following we sketch the representation of temporal knowledge in the generalized interval framework (Galiotou 1999; Galiotou and Ligozat 1997).

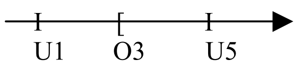
A culminated process (the nucleus) situated in the past will be represented by a sequence of typed boundaries numbered by odd numbers from 1 to 7.



Preparatory process: (1 3), Culmination point: 3, Consequent state: (3 5), Speech time: 7, Aspect: Perfect, Tense: Past

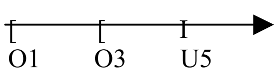
The boundaries take their values among the following: **[** (**O**pening boundary), **]** (**C**losing boundary), **U** (**U**ndefined boundary)

A culmination situated in the past will be represented by a sequence of three boundaries as in the following:



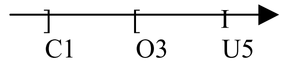
Culmination point: 1, Consequent state: (1 3), Speech time: 5, Aspect: Perfective, Tense:Past

A process situated in the past will represented as:



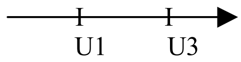
Process: (1 3), Speech time: 5 , Aspect: Imperfective, Tense: Past

A state will represented as:



State: (1 3), Speech time: 5, Aspect: Imperfective, Tense: Past

A punctual event will be represented as:

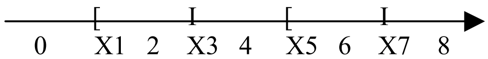


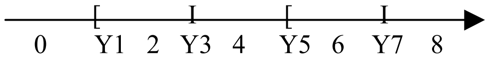
Point: 1, Speech time: 3, Aspect: Perfective, Tense: Past

3.2.3. Building the temporal site

Using this schematic representation we try to express the corresponding temporal relations (see also Galiotou 1999, Galiotou and Ligozat 1997). In the following, we give an example of a temporal relation expressed by means of relations between typed boundaries:

Suppose that we consider two 4-intervals X and Y each referring to an event nucleus. X is defined by the sequence of boundaries x_1, x_2, x_3, x_4 and Y is defined by the the sequence of boundaries y_1, y_2, y_3, y_4. The variables x_i and y_i , $i \in \{1,3,5,7\}$ have a type among the following: {O(pening), C(losing), U(ndefined)}. The open intervals defined by the boundaries will be numbered by even numbers. So, the representation of the two 4-intervals will be:





The relation of temporal precedence will be represented by the following:

([0,2], [0,2], [0,6], 7)

The culmination point of X (x_3) is related to the culmination point of Y (y_3) by a relation of temporal precedence while the position of x_5 is not completely determined.

Based on this representation, we build a temporal constraint network reflecting the temporal structure of a text. Actually, the arcs of the network correspond to two types of relations: on the one hand, purely temporal relations, and on the other hand discourse relations reflecting the hierarchical structure of the text and having a temporal or causal significance (Asher 1993): narration, elaboration, explanation, background, and result. We also take into account the discourse relation of contrast which has a rather non-temporal character but proves to be useful in establishing contextual constraints for the processing of causal information.

At this point of the representation we can anticipate different ways of processing (Bestougeff and Ligozat 1992):

- Restrict *a priori* the number of possible relations between two n-intervals in a natural language utterance.
- In case of an incomplete graph, calculate the possible relations between two generalized intervals using the formula of composition of relations between two generalized intervals (Ligozat 1991). An application of this procedure can be found in (Galiotou 1999).
- Verify that the set of temporal relations in the temporal constraint network is consistent.
- Verify that updating the network does not introduce any inconsistency.

Every time a new set of constraints between two generalized intervals is introduced in the temporal constraint network, the set of relations is updated. For the propagation of constraints in the graph we use a variant of Allen's constraint propagation algorithm (Allen 1983; Allen 1984) suitably adapted for the generalized interval framework (Ligozat 1991; Galiotou 1999).

3. A proposal for a representational scheme

Before proceeding with the definition of the scheme, we have to specify the linguistic elements of the text which contribute to determining the temporal relations.

3.1. Linguistic elements in the text

In our experimentation we have used a corpus of short newspaper articles in Modern Greek describing car accidents. The study of temporal and causal phenomena in such a corpus had two goals:

- extract the events which have led to an accident as well as its consequences;
- provide all plausible explanations of the causes of the accident.

In general, the causes of the accident are explicitly mentioned in the text using causal connectives and/or causal expressions. This explicit causal information contributes to the determination of the temporal order and the relative positioning of events in conjunction with the «Causes precede Effects» rule.

Our corpus consists of narrative texts composed of simple clauses containing temporal adverbials, subordinates introduced by temporal and/or causal connectives, verb phrases, derived nouns and adjectives. In this type of narration, events are situated in the past so, verbal forms of the future do not appear in the text. The most frequent verbal forms are those of the preterite, while the present tense is used in order to describe consequent states. As far as grammatical aspect is concerned, all possible values for Greek are present, i.e. imperfective, perfective and perfect (Mackridge 1985). The lexical aspect follows the tripartite ontology of events as already stated in section 2.

Temporal adverbials are classified (Tzevelekou 1995) into:

- those which require a different starting point from the starting point of the narration: kat' ar`xin[1] (at first), sti si`nexia (afterwards) etc.
- those which establish their reference either from the starting point of the narration or form another starting point: ta ksime`romata tis kiriakis (on Sunday morning), `liγo prin ta me`sanixta tis paraske`vis (short before Friday midnight) etc.
- deictic adverbs which establish their reference uniquely from the starting point of the narration: xθes (yesterday), pro`xθes (the day before yesterday), etc.

Temporal clauses are introduced by temporal connectives which on the one hand provide information on the temporal order of events and on the other hand help determine implicit causal relations (i.e. causal relations which are not expressed by causal markers in the text). Consider the case of the temporal connective `otan (when). Its use in the corpus is in accordance with the observation that its basic meaning is not temporal in the first place (Moens and Steedman 1988). Indeed, `otan (when) establishes a consequentiality relation between the main clause event and the subordinate clause event provided that the eventualities in the text follow the organizational scheme of the tripartite ontology of events. Therefore, it is a candidate for expressing causality in the text. Take for example, the case of the construction

Q(Past-perfective) `otan P(Pat-Perfective)

In this case the processes Q and P are related with a consequentiality relation as in the following example:

tris `nei `exasan ti zo`i tus otan to afto`kinito tus ana`trapike
Three young men lost their lives when their car overturned.

Explicit causal information in the text is expressed by causal connectives such as *epi`δi (because)* of causal expressions such as *me apo`telesma (with result*) or *`loγo (because of).*

It is also expressed by noun phrases which are introduced in simple sentences such as:

e`pisimi e`tia tu δisti`ximatos ine i olisθi`rotita tu `δromu
The official cause of the accident is the slipperiness of the road.

or, verbal phrases with a causal sense such as:

to δi`stixima o`filete stin ipervoli`ki ta`xitita tu aftoki`nitu
The accident is due to the excessive speed of the car.

The combination of a causal expression with a derived noun phrases such as *i olisθi`rotita tu `δromu (the slipperiness of the road), i api`ria tu oδi`γu (the inexperience of the driver), I aprose`ksia ton pe`zon (the inattention of the pedestrians)* etc is used in order to

[1] Greek words are transcribed according to the characters of the International Phonetic Alphabet. When necessary, stress is indicated with symbol « ` » before any stressed syllable.

establish discourse relations and to build the temporal structure of the text. For example:

to δi`stixima o`filete stin apì`ria tu oδi`γu a`la ke stin aprose`ksia ton pe`zon

The accident is due to the inexperience of the driver but also, to the inattention of the pedestrians.

Here, a relation of explanation holds between the propositions and therefore, a relation of consequentiality between the corresponding eventualities. The derived nominal represents a state. So, the consequentiality relation is specified as an enablement relation between the state represented by the derived nominal and the main clause event which in this case is the accident. Thus we obtain an overlap relation between the related temporal intervals.

3.2. The representational scheme

The theoretical assumptions described in section 2 lead to the definition of a representational scheme taking into account the linguistic elements of the text which are described in the previous subsection. As already stated, our treatment of temporal information in the text is based solely on the exploitation of linguistic knowledge.

Therefore, we propose a single-dimensional scheme consisting of a list of labels relative to the temporal and/or causal knowledge as it is expressed by the linguistic elements of the text.

Note that, since the starting point of our research was the study of temporal and causal phenomena in Modern Greek, we had to take into account the particularities of the temporal and aspectual system of the language in elaborating a suitable tag set. Yet, in our opinion, this does not constitute a major problem since the set of values is easily extendable in order to capture particularities of other languages provided that the general guidelines are kept.

The text segments to be annotated are taken at the clause level.

3.2.3. The tag set

The proposed tag set takes into account the following classes:

BOUNDARY_LIST : The generalized interval expressed in terms of a list of typed boundaries for example,

<U1, 03, U5>

TENSE: Values are limited to

<past>, <non_past>

since in Modern Greek future is considered as modal.

ASPECT: <imperfective>, <perfective>, <perfect>

LEXASP: The values follow the tripartite ontology of event and they are limited to:

<state>, <culm_process>, <process>, <culmination>, <point>.

The number of categories could easily be modified provided that the principles of the ontology are respected.

CONNECTIVE: We take into account causal and/ or temporal connectives therefore, the proposed tags are:

<cn_causal>, <cn_temp>, <cn_ct>

TEMPADV: Temporal adverbials play a crucial role in our processing of temporal and causal information but there was no need for a more fine-grained characterization. So, the only possible tag is:

<tempadv>

In the current state of the scheme we have decided to encode dates under the <tempadv> tag.

CUE_PHRASE: Cue_phrases contribute to the determination of temporal and explicitly causal information, so they enter our scheme with the tag:

<cue_phrase>

3.2.3. The decision rules

Following (Klein, 1999) we provide a decision tree aiming at giving an overview of all possible tags and how they are related to each other. We also provide rules that help to mutually constraint the tags between categories and ensure the coherence of the annotation:

In figure 1, we give an example of the decision rules :

if the segment is tagged with <culmination> and <past >
then label with <U1 O3 U5>
….
if the segment is tagged with <past> and <perfect>
then label with <culm_process>
…
if the segment is tagged with <past> and <imperfective>
then label with <state>
…

Figure 1. Example of decision rules

3.2.3. Examples and evaluation

In table 1 we give an example of the application of the scheme to a fragment of a text describing a car accidents (in Modern Greek).

A quick overview this application has shown that as far as the temporal entities are concerned there was an approximately even distribution of culminated processes and culminations. and to a lesser extent states and processes. The fifth category of temporal entities, namely that of punctual events, did not appear in any text.

As for the verbal forms, there was no occurrence of the future tense. The most frequent forms were those of the preterite, while the present tense was used in order to describe the consequent states of events and conclusions.

The distribution of temporal and causal connectives indicated that the most frequent appearance is that of the *`otan (when)* connective used both as a temporal and as a causal one.

We must also report an extensive use of cue_phrases for the expression of causal knowledge in the text. Contrary to what one may have expected in a corpus of short newspaper articles describing car accidents, causal connectives were not the most frequent elements used in order to express causality. Explicit causal information was mostly put forward using cue phrases in simple sentences.

The scheme proved to be quite easy to apply by the human annotator. Yet, the tagging procedure, as far the boundary list was concerned proved to be more difficult to implement. This has led us to the conclusion that the annotation procedure should be automated at least as far as the boundary_list category is concerned.

ORIGINAL TEXT	ANNOTATED TEXT
tris \`nei \`exasan ti zo\`i tus otan to afto\`kinito tus ana\`trapike (Three young men lost their lives when their car overturned).	<U1 O3 U5> <culm> tris \`nei <past><imperfective>\`exasan </past></imperfective> ti zo\`i tus </culm> </U1 O3 U5> <O1 U3 O5 U7> <culm_process> <cn_temp>otan </cn_temp> to afto\`kinito tus <past><perfective>ana\`trapike</past></perfective> </culm_process></O1 U3 O5 U7>

Table 1: Application of the scheme to a fragment of a text

4. Conclusions

In this paper, we proposed a representational scheme which provides the background for the processing of temporal and causal information.

We have insisted on the theoretical assumptions of the scheme concerning:

- the temporal positioning of events;
- the appropriate temporal ontology;
- the interaction of temporal and causal information in texts;
- the causal interpretation of events.

We also discussed the nature of the linguistic elements of a text which can be used to determine the temporal structure of the text.

We then briefly described a representational scheme based on the theoretical assumptions mentioned above.

This representational scheme was applied to a corpus of short newspaper articles describing car accidents in Modern Greek. These articles were selected with respect to their informational interest and the temporal and causal relations they contained. Therefore, the particularities of the temporal and aspectual system of the Greek language were taken into account in the proposed tag set. Nevertheless, this tag set could easily be modified to take into account the particularities of other languages provided that the general approach is followed.

The scheme has proved to be quite reliable in its application to the particular corpus. Yet, the task of the human annotator would be greatly facilitated if at least certain procedures such as the tagging with the boundaries_list notation were automated.

5. References

Allen, J.F., 1983. Maintaining Knowledge about Temporal Intervals. *Communications of the ACM 26,* 11:832-843.

Allen, J.F., 1984. Towards a General Theory of Action and Time. *Artificial Intelligence, 23,* 123-154.

Asher, N., 1993. *Reference to Abstract Objects of Discourse*, Dordrecht:Kluwer.

Bestougeff, H. and G. Ligozat, 1992. *Tools for Temporal Knowledge Representation*, Academic Press.

Galiotou, E., 1997. Temporal and Causal relations in Greek Texts. In A. Ralli, M. Grigoriadou, G. Philokyprou, D. Christodoulakis and E. Galiotou (eds.) *Working Papers in Natural Language Processing*, Athens: Diavlos, 109-122.

Galiotou, E., 1999. *Modélisation de la Causalité à partir de l'Analyse des Phénomènes Temporels dans les Langues Naturelles: Application au Grec Moderne,* Ph.D. Thesis, Université Paris-Sud, (=Publication Notes et Documents LIMSI 2000-11, LIMSI-CNRS).

Galiotou, E. and G. Ligozat, 1997. Représentation du Temps et de l'Aspect en vue de la Génération d'Explications Causales. *Actes GAT'97*, Université Stendhal, Grenoble, 85-96.

Kitis, E., 1995. Connectives and Ideology. *Proceedings of the 4th International Symposium on Critical Discourse Analysis: Language and Social Thought*, University of Athens.

Klein, M., 1999. Standardisation Efforts on the Level of Dialogue Act in the MATE Project*.* *Proceedings of the ACL Workshop: Towards Standards and Tools for Discourse Tagging,* University of Maryland, 85-41.

Lascarides, A., and N. Asher, 1993. *Temporal Interpretation, Discourse Relations and Common Sense Entailment*, Research Paper HCRC/RP-16, University of Edinburgh.

Ligozat, G., 1991. On Generalized Interval Calculi. *Proceedings of AAAI'91*, 234-240.

Ligozat, G., 1996. A New Proof of Tractability for ORD-Horn Relations. *Proceedings of AAAI'96,* 395-401.

Ligozat, G., 1997. Time, Space and Natural Language Processing. In Ralli et al. (eds.), Athens: Diavlos, 155-174.

Mackridge, P., 1985. *The Modern Greek Language*, Oxford University Press.

Moens, M., 1987. *Tense, Aspect and Temporal Reference*, Ph.D. Thesis, University of Edinburgh.

Moens, M. and M. Steedman, 1988. Temporal Ontology and Temporal Reference. *Computational Linguistics, 14*, 2: 15-28.

Parsons, T., 1990. *Events in the Semantics of English*, MIT Press.

Soria, C. and V. Pirrelli, 1999. A Recognition-based Meta-scheme for Dialogue Acts Annotation, *Proceedings of the ACL Workshop: Towards Standards and Tools for Discourse Tagging,* University of Maryland, 75-83.

Teufel S., and M. Moens, 1999. Discourse-level argumentation in scientific articles: human and automatic annotation, *Proceedings of the ACL Workshop: Towards Standards and Tools for Discourse Tagging,* University of Maryland, 84-93.

Tzevelekou, M., 1995. *Catégorisation Lexicale et Aspect: Le Système Aspectuel du Grec Moderne*, Ph.D. Thesis, Université Paris 7.

Toward an Ontology of Time for the Semantic Web

Jerry R. Hobbs

Artificial Intelligence Center
SRI International
Menlo Park, California 94025
USA
hobbs@ai.sri.com

Abstract

In connection with the DAML project for bringing about the Semantic Web, an ontology of time is being developed for describing the temporal content of Web pages and the temporal properties of Web services. The bulk of information on the Web is in natural language, and this information will be easier to encode for the Semantic Web insofar as community-wide annotation and automatic tagging schemes and the DAML time ontology are compatible with each other.

1. Introduction

The DARPA Agent Markup Language (DAML) project is an effort aimed at bringing into reality the Semantic Web, in which Web users and automatic agents will be able to access information on the Web via descriptions of the content and capabilities of Web resources rather than via key words. An important part of this effort is the development of representative ontologies of the most commonly used domains. We are beginning to develop such an ontology of temporal concepts, for describing the temporal content of Web pages and the temporal properties of Web services. This effort is being informed by temporal ontologies developed at a number of sites and is intended to capture the essential features of all of them and make them easily available to a large group of Web developers and users.

The bulk of information on the Web is in natural language, and this information will be easier to encode for the Semantic Web insofar as community-wide annotation and automatic tagging schemes and the DAML time ontology are compatible with each other.

In this paper I outline the temporal ontology as it has been developed so far, in order to initiate a dialog between the two communities. Five categories of temporal concepts are considered, and for each the principal predicates and their associated properties are described.

A note on notation: Conjunction ($\wedge$) takes precedence over implication($\supset$) and equivalence ($\equiv$). Formulas are assumed to be universally quantified on the variables appearing in the antecedent of the highest-level implication. Thus,

$$p_1(x) \wedge p_2(y) \supset q_1(x,y) \wedge q_2(y)$$

is to be interpreted as

$$(\forall\, x,y)[[p_1(x) \wedge p_2(y)] \supset [q_1(x,y) \wedge q_2(y)]]$$

2. Topological Temporal Relations

2.1. Instants and Intervals

There are two subclasses of temporal-entity: $instant$ and $interval$.

$$instant(t) \supset temporal\text{-}entity(t)$$
$$interval(T) \supset temporal\text{-}entity(T)$$

(In what follows, lower case t is used for instants, upper case T for intervals and for temporal-entities unspecified as to subtype. This is strictly for the reader's convenience, and has no formal significance.)

$start\text{-}of$ and $end\text{-}of$ are functions from temporal entities to instants.

$$temporal\text{-}entity(T) \supset instant(start\text{-}of(T))$$
$$temporal\text{-}entity(T) \supset instant(end\text{-}of(T))$$

For convenience, we can say that the start and end of an instant is itself.

$$instant(t) \supset start\text{-}of(t) = t$$
$$instant(t) \supset end\text{-}of(t) = t$$

$inside$ is a relation between an instant and an interval.

$$inside(t,T) \supset instant(t) \wedge interval(T)$$

This concept of $inside$ is not intended to include starts and ends of intervals, as will be seen below.

Infinite and half-infinite intervals can be handled by positing time instants at positive and negative infinity, and using them as start and end points.

It will be useful in characterizing clock and calendar terms to have a relation between instants and intervals that says that the instant is inside or the start of the interval.

$$in\text{-}interval(t,T)$$
$$\equiv [start\text{-}of(T) = t \vee inside(t,T)]$$

$interval\text{-}between$ is a relation among a temporal entity and two instants.

$$interval\text{-}between(T,t_1,t_2)$$
$$\supset temporal\text{-}entity(T) \wedge instant(t_1)$$
$$\wedge\, instant(t_2)$$

The two instants are the start and end points of the temporal entity.

$$interval\text{-}between(T,t_1,t_2)$$
$$\equiv start\text{-}of(T) = t_1 \wedge end\text{-}of(T) = t_2$$

The ontology is silent about whether the interval from t to t, if it exists, is identical to the instant t.

The ontology is silent about whether intervals *consist of* instants.

The ontology is silent about whether intervals are uniquely determined by their starts and ends.

We can define a proper interval as one whose start and end are not identical.

$$proper\text{-}interval(t) \equiv \\ interval(t) \wedge start\text{-}of(t) \neq end\text{-}of(t)$$

The ontology is silent about whether there are any intervals that are not proper-intervals.

2.2. Before

There is a $before$ relation on temporal entities, which gives directionality to time. If temporal entity T_1 is before temporal entity T_2, then the end of T_1 is before the start of T_2. Thus, before can be considered to be basic to instants and derived for intervals.

$$before(T_1, T_2) \\ \equiv before(end\text{-}of(T_1), start\text{-}of(T_2))$$

The end of an interval is not before the start of the interval.

$$interval(T) \\ \supset before(end\text{-}of(T), start\text{-}of(T))$$

The start of a proper interval is before the end of the interval.

$$proper\text{-}interval(T) \\ \supset before(start\text{-}of(T), end\text{-}of(T))$$

If one instant is before another, there is an interval between them.

$$instant(t_1) \wedge instant(t_2) \wedge before(t_1, t_2) \\ \supset (\exists T) interval\text{-}between(T, t_1, t_2)$$

The ontology is silent about whether there is an interval from t to t.

If an instant is inside a proper interval, then the start of the interval is before the instant, which is before the end of the interval. The converse is true as well.

$$instant(t) \wedge proper\text{-}interval(T) \\ \supset [inside(t, T) \\ \equiv before(start\text{-}of(T), t) \\ \wedge before(t, end\text{-}of(T))]$$

Intervals are contiguous with respect to the $before$ relation, in that an instant between two other instants inside an interval is inside the interval.

$$before(t_1, t_2) \wedge before(t_2, t_3) \\ \wedge inside(t_1, T) \wedge inside(t_3, T) \\ \supset inside(t_2, T)$$

The $before$ relation is anti-symmetric and transitive.

$$before(T_1, T_2) \supset \neg before(T_2, T_1)$$

$$before(T_1, T_2) \wedge before(T_2, T_3) \\ \supset before(T_1, T_3)$$

The relation $after$ is defined in terms of $before$.

$$after(T_1, T_2) \equiv before(T_2, T_1)$$

The ontology is silent about whether time is linearly ordered.

2.3. Interval Relations

The relations between intervals defined in Allen's temporal interval calculus (Allen and Kautz, 1985) can be defined in a straightforward fashion in terms of $before$ and identity on the start and end points.

$$interval(T_1) \wedge interval(T_2) \\ \supset [int\text{-}equals(T_1, T_2) \\ \equiv start\text{-}of(T_1) = start\text{-}of(T_2) \\ \wedge end\text{-}of(T_1) = end\text{-}of(T_2)]$$

$$interval(T_1) \wedge interval(T_2) \\ \supset [int\text{-}before(T_1, T_2) \equiv before(T_1, T_2)$$

$$interval(T_1) \wedge interval(T_2) \\ \supset [int\text{-}after(T_1, T_2) \equiv after(T_1, T_2)$$

$$interval(T_1) \wedge interval(T_2) \\ \supset [int\text{-}meets(T_1, T_2) \\ \equiv end\text{-}of(T_1) = start\text{-}of(T_2)$$

$$interval(T_1) \wedge interval(T_2) \\ \supset [int\text{-}met\text{-}by(T_1, T_2) \\ \equiv int\text{-}meets(T_2, T_1)]$$

$$interval(T_1) \wedge interval(T_2) \\ \supset [int\text{-}overlaps(T_1, T_2) \\ \equiv before(start\text{-}of(T_1), start\text{-}of(T_2)) \\ \wedge before(start\text{-}of(T_2), end\text{-}of(T_1)) \\ \wedge before(end\text{-}of(T_1), end\text{-}of(T_2))]$$

$$interval(T_1) \wedge interval(T_2) \\ \supset [int\text{-}overlapped\text{-}by(T_1, T_2) \\ \equiv int\text{-}overlaps(T_2, T_1)]$$

$$interval(T_1) \wedge interval(T_2) \\ \supset [int\text{-}starts(T_1, T_2) \\ \equiv start\text{-}of(T_1) = start\text{-}of(T_2) \\ \wedge before(end\text{-}of(T_1), end\text{-}of(T_2)]$$

$$interval(T_1) \wedge interval(T_2) \\ \supset [int\text{-}started\text{-}by(T_1, T_2) \\ \equiv int\text{-}starts(T_2, T_1)]$$

$$interval(T_1) \wedge interval(T_2) \\ \supset [int\text{-}during(T_1, T_2) \\ \equiv (before(start\text{-}of(T_2), start\text{-}of(T_1)) \\ \wedge before(end\text{-}of(T_1), end\text{-}of(T_2))]$$

$$interval(T_1) \wedge interval(T_2) \\ \supset [int\text{-}contains(T_1, T_2) \\ \equiv int\text{-}during(T_2, T_1)]$$

$$interval(T_1) \wedge interval(T_2) \\ \supset [int\text{-}finishes(T_1, T_2) \\ \equiv before(start\text{-}of(T_2), start\text{-}of(T_1)) \\ \wedge end\text{-}of(T_1) = end\text{-}of(T_2)]$$

$$interval(T_1) \wedge interval(T_2) \\ \supset [int\text{-}finished\text{-}by(T_1, T_2) \\ \equiv int\text{-}finishes(T_2, T_1)]$$

In addition, it will be useful below to have a single predicate for "starts or is during". This is called $int\text{-}in$.

$$int\text{-}in(T_1, T_2) \\ \equiv [int\text{-}starts(T_1, T_2) \vee int\text{-}during(T_1, T_2)]$$

It will also be useful to have a single predicate for intervals intersecting in at most an instant.

$$
\begin{aligned}
&\mathit{int\text{-}disjoint}(T_1, T_2) \\
&\quad \equiv [\mathit{int\text{-}before}(T_1, T_2) \lor \mathit{int\text{-}after}(T_1, T_2) \\
&\qquad \lor \mathit{int\text{-}meets}(T_1, T_2) \\
&\qquad \lor \mathit{int\text{-}met\text{-}by}(T_1, T_2)]
\end{aligned}
$$

So far, the concepts and axioms in the ontology of time would be appropriate for scalar phenomena in general.

2.4. Linking Time and Events

The time ontology links to other things in the world through four predicates—$\mathit{at\text{-}time}$, during, holds, and $\mathit{time\text{-}span\text{-}of}$. We assume that another ontology provides for the description of events—either a general ontology of event structure abstractly conceived, or specific, domain-dependent ontologies for specific domains.

The term "eventuality" will be used to cover events, states, processes, propositions, states of affairs, and anything else that can be located with respect to time. The possible natures of eventualities would be spelled out in the event ontologies.

The predicate $\mathit{at\text{-}time}$ relates an eventuality to an instant, and is intended to say that the eventuality holds, obtains, or is taking place at that time.

$$\mathit{at\text{-}time}(e, t) \supset \mathit{eventuality}(e) \land \mathit{instant}(t)$$

The predicate during relates an eventuality to an interval, and is intended to say that the eventuality holds, obtains, or is taking place during that interval.

$$\mathit{during}(e, T) \supset \mathit{eventuality}(e) \land \mathit{interval}(T)$$

If an eventuality obtains during an interval, it obtains at every instant inside the interval.

$$\mathit{during}(e, T) \land \mathit{inside}(t, T) \supset \mathit{at\text{-}time}(e, t)$$

Whether a particular process is viewed as instantaneous or as occuring over an interval is a granularity decision that may vary according to the context of use, and is assumed to be provided by the event ontology.

Often the eventualities in the event ontology are best thought of as propositions, and the relation between these and times is most naturally called holds. holds can be defined in terms of $\mathit{at\text{-}time}$ and during:

$$
\begin{aligned}
&\mathit{holds}(e, t) \land \mathit{instant}(t) \equiv \mathit{at\text{-}time}(e, t) \\
&\mathit{holds}(e, T) \land \mathit{interval}(T) \equiv \mathit{during}(e, T)
\end{aligned}
$$

The event ontology may provide other ways of linking events with times, for example, by including a time parameter in predications.

$$p(x, t)$$

This time ontology provides ways of reasoning about the t's; their use as arguments of predicates from another domain would be a feature of the ontology of the other domain.

The predicate $\mathit{time\text{-}span\text{-}of}$ relates eventualities to instants or intervals. For contiguous states and processes, it tells the entire instant or interval for which the state or process obtains or takes place.

$$
\begin{aligned}
&\mathit{time\text{-}span\text{-}of}(T, e) \\
&\quad \supset \mathit{temporal{-}entity}(T) \land \mathit{eventuality}(e) \\
&\mathit{time\text{-}span\text{-}of}(T, e) \land \mathit{interval}(T) \\
&\quad \supset \mathit{during}(e, T) \\
&\mathit{time\text{-}span\text{-}of}(t, e) \land \mathit{instant}(t) \\
&\quad \supset \mathit{at\text{-}time}(e, t)
\end{aligned}
$$

$$
\begin{aligned}
&\mathit{time\text{-}span\text{-}of}(T, e) \land \mathit{interval}(T) \\
&\qquad \land \neg \mathit{inside}(t_1, T) \land \neg \mathit{start\text{-}of}(t_1, T) \\
&\qquad \land \neg \mathit{end\text{-}of}(t_1, T) \\
&\quad \supset \neg \mathit{at\text{-}time}(e, t_1) \\
&\mathit{time\text{-}span\text{-}of}(t, e) \land \mathit{instant}(t) \land t_1 \neq t \\
&\quad \supset \neg \mathit{at\text{-}time}(e, t_1)
\end{aligned}
$$

$\mathit{time\text{-}span\text{-}of}$ is a predicate rather than a function because until the time ontology is extended to aggregates of temporal entities, the function would not be defined for non-contiguous eventualities. Whether the eventuality obtains at the start and end points of its time span is a matter for the event ontology to specify. The silence here on this issue is the reason $\mathit{time\text{-}span\text{-}of}$ is not defined in terms of necessary and sufficient conditions.

The event ontology could extend temporal functions and predicates to apply to events in the obvious way, e.g.,

$$
\begin{aligned}
&\mathit{ev\text{-}start\text{-}of}(e) = t \\
&\quad \equiv \mathit{time\text{-}span\text{-}of}(T, e) \land \mathit{start\text{-}of}(T) = t
\end{aligned}
$$

This would not be part of the time ontology, but would be consistent with it.

Different communities have different ways of representing the times and durations of states and events (processes). In one approach, states and events can both have durations, and at least events can be instantaneous. In another approach, events can only be instantaneous and only states can have durations. In the latter approach, events that one might consider as having duration (e.g., heating water) are modeled as a state of the system that is initiated and terminated by instantaneous events. That is, there is the instantaneous event of the start of the heating at the start of an interval, that transitions the system into a state in which the water is heating. The state continues until another instantaneous event occurs—the stopping of the heating at the end of the interval. These two perspectives on events are straightforwardly interdefinable in terms of the ontology we have provided. This is a matter for the event ontology to specify. This time ontology is neutral with respect to the choice.

3. Measuring Durations

3.1. Temporal Units

This development assumes ordinary arithmetic is available.

There are at least two approaches that can be taken toward measuring intervals. The first is to consider units of time as functions from Intervals to Reals, e.g.,

$\mathit{minutes}$: Intervals $\rightarrow$ Reals
$\mathit{minutes}([5:14, 5:17]) = 3$

The other approach is to consider temporal units to constitute a set of entities—call it TemporalUnits—and have a single function $duration$ mapping Intervals × TemporalUnits into the Reals.

$$duration([5:14, 5:17], *Minute*) = 3$$

The two approaches are interdefinable:

$$\begin{array}{l} seconds(T) = duration(T, *Second*) \\ minutes(T) = duration(T, *Minute*) \\ hours(T) = duration(T, *Hour*) \\ days(T) = duration(T, *Day*) \\ weeks(T) = duration(T, *Week*) \\ months(T) = duration(T, *Month*) \\ years(T) = duration(T, *Year*) \end{array}$$

Ordinarily, the first is more convenient for stating specific facts about particular units. The second is more convenient for stating general facts about all units.

The aritmetic relations among the various units are as follows:

$$\begin{array}{l} seconds(T) = 60 * minutes(T) \\ minutes(T) = 60 * hours(T) \\ hours(T) = 24 * days(T) \\ days(T) = 7 * weeks(T) \\ months(T) = 12 * years(T) \end{array}$$

The relation between days and months (and, to a lesser extent, years) will be specified as part of the ontology of clock and calendar below. On their own, however, month and year are legitimate temporal units.

In this development durations are treated as functions on intervals and units, and not as first class entities on their own, as in some approaches. In the latter approach, durations are essentially equivalence classes of intervals of the same length, and the length of the duration is the length of the members of the class. The relation between an approach of this sort (indicated by prefix D-) and the one presented here is straightforward.

$$\begin{array}{l} (\forall\, T, u, n)[duration(T, u) = n \\ \quad \equiv\ (\exists\, d)[D\text{-}duration\text{-}of(T) = d \\ \quad\quad \wedge\, D\text{-}duration(d, u) = n]] \end{array}$$

At the present level of development of the temporal ontology, this extra layer of representation seems superfluous. It may be more compelling, however, when the ontology is extended to deal with the combined durations of noncontiguous aggregates of intervals.

3.2. $Hath$

The multiplicative relations above don't tell the whole story of the relations among temporal units. Temporal units are *composed of* smaller temporal units. The basic predicate used here for expressing the composition of larger intervals out of smaller clock and calendar intervals is $Hath$, from statements like "30 days hath September" and "60 minutes hath an hour." Its structure is

$$Hath(S, N, u, x)$$

meaning "A set S of N calendar intervals of type u hath the calendar interval x." That is, if $Hath(S, N, u, x)$ holds, then x is composed of the disjoint union of N intervals of type u; S is the set of those intervals. For example, if x is some month of September and S is the set of the successive days of that September, then $Hath(S, 30, *Day*, x)$ would be true.

The principal properties of $Hath$ are as follows:

The type constraints on its arguments: S is a set, N is an integer, u is a temporal unit, and x is an interval:

$$\begin{array}{l} Hath(S, N, u, x) \\ \quad \supset\ set(S)\ \wedge\ integer(N) \\ \quad\quad \wedge\, temporal\text{-}unit(u)\ \wedge\ interval(x) \end{array}$$

The elements of S are intervals of duration u:

$$\begin{array}{l} Hath(S, N, u, x) \\ \quad \supset\ (\forall\, y)[member(y, S) \\ \quad\quad \supset\ interval(y)\ \wedge\ duration(y, u) = 1] \end{array}$$

S has N elements:

$$Hath(S, N, u, x)\ \supset\ card(S) = N$$

The elements of S are disjoint:

$$\begin{array}{l} Hath(S, N, u, x) \\ \quad \supset\ (\forall\, y_1, y_2)[member(y_1, S) \\ \quad\quad\quad \wedge\, member(y_2, S)\ \wedge\ y_1 \neq y_2 \\ \quad\quad \supset\ int\text{-}disjoint(y_1, y_2)] \end{array}$$

There are elements in S that start and finish x:

$$\begin{array}{l} Hath(S, N, u, x) \\ \quad \supset\ (\exists\, y_1)[member(y_1, S) \\ \quad\quad \wedge\, int\text{-}starts(y_1, x)] \end{array}$$

$$\begin{array}{l} Hath(S, N, u, x) \\ \quad \supset\ (\exists\, y_2)[member(y_2, S) \\ \quad\quad \wedge\, int\text{-}finishes(y_2, x)] \end{array}$$

Except for the first and last elements of S, every element of S has an element that precedes and follows it:

$$\begin{array}{l} Hath(S, N, u, x) \\ \quad \supset\ (\forall\, y_1)[member(y_1, S) \\ \quad\quad \supset\ [int\text{-}finishes(y_1, x) \\ \quad\quad\quad \vee\, (\exists\, y_2)[member(y_2, x) \\ \quad\quad\quad\quad \wedge\, int\text{-}meets(y_1, y_2)]]] \end{array}$$

$$\begin{array}{l} Hath(S, N, u, x) \\ \quad \supset\ (\forall\, y_2)[member(y_2, S) \\ \quad\quad \supset\ [int\text{-}starts(y_2, x) \\ \quad\quad\quad \vee\, (\exists\, y_1)[member(y_1, x) \\ \quad\quad\quad\quad \wedge\, int\text{-}meets(y_1, y_2)]]] \end{array}$$

If time is linearly ordered, the existential quantifier $\exists$ in the last four axioms can be replaced by $\exists\,!$.

Finally, we would like to say that the set S covers x. A simple way to say this is as follows:

$$\begin{array}{l} Hath(S, N, u, x) \\ \quad \supset\ (\forall\, t)[inside(t, x) \\ \quad\quad \supset\ (\exists\, y)[member(y, S) \\ \quad\quad\quad \wedge\, in\text{-}interval(t, y)]] \end{array}$$

That is, if an instant t is inside x, there is a smaller unit y that t is inside or the start of.

However, this is a good place to introduce notions of granularity. In describing the temporal properties of some class of events, it may make sense to specify their time with respect to some temporal unit but not with respect to a smaller temporal unit. For example, one might want to talk about an election as a point-like event being at some instant, and specifying the day that instant is in, but not specifying the hour or minute.

To accomodate this, the above axiom can be loosened by applying it only when the instant t is located in *some interval* of size u. The axiom above would be modified as follows:

$$
\begin{aligned}
&Hath(S, N, u, x) \\
&\quad \supset (\forall\, t, y_1)[inside(t, x) \wedge inside(t, y_1) \\
&\qquad \wedge\, duration(y_1, u) \\
&\quad\; \supset (\exists\, y)[member(y, S) \\
&\qquad \wedge\, in\text{-}interval(t, y)]]
\end{aligned}
$$

Essentially, the conjuncts $inside(t, y_1) \wedge duration(y_1, u)$ specify that t can be viewed at a granularity of u.

This treatment of $Hath$ could be extended to measurable quantities in general.

3.3. The Structure of Temporal Units

We now define predicates true of intervals that are one temporal unit long. For example, $week$ is a predicate true of intervals whose duration is one week.

$$
\begin{aligned}
&second(T) \equiv seconds(T) = 1 \\
&minute(T) \equiv minutes(T) = 1 \\
&hour(T) \equiv hours(T) = 1 \\
&day(T) \equiv days(T) = 1 \\
&week(T) \equiv weeks(T) = 1 \\
&month(T) \equiv months(T) = 1 \\
&year(T) \equiv years(T) = 1
\end{aligned}
$$

We are now in a position to state the relations between successive temporal units.

$$
\begin{aligned}
&minute(T) \supset (\exists\, S) Hath(S, 60, {*Second*}, T) \\
&hour(T) \supset (\exists\, S) Hath(S, 60, {*Minute*}, T) \\
&day(T) \supset (\exists\, S) Hath(S, 24, {*Hour*}, T) \\
&week(T) \supset (\exists\, S) Hath(S, 7, {*Day*}, T) \\
&year(T) \supset (\exists\, S) Hath(S, 12, {*Month*}, T)
\end{aligned}
$$

The relations between months and days are dealt with in Section 4.4.

4. Clock and Calendar

4.1. Time Zones

What hour of the day an instant is in is relative to the time zone. This is also true of minutes, since there are regions in the world, e.g., central Australia, where the hours are not aligned with GMT hours, but are, e.g., offset half an hour. Probably seconds are not relative to the time zone.

Days, weeks, months and years are also relative to the time zone, since, e.g., 2002 began in the Eastern Standard time zone three hours before it began in the Pacific Standard time zone. Thus, predications about all clock and calendar intervals except seconds are relative to a time zone.

This can be carried to what seems like a ridiculous extreme, but turns out to yield a very concise treatment. The Common Era (C.E. or A.D.) is also relative to a time zone, since 2002 years ago, it began three hours earlier in what is now the Eastern Standard time zone than in what is now the Pacific Standard time zone. What we think of as the Common Era is in fact 24 (or more) slightly displaced half-infinite intervals. (We leave B.C.E. to specialized ontologies.)

The principal functions and predicates will specify a clock or calendar unit interval to be the nth such unit in a larger interval. The time zone need not be specified in this predication if it is already built into the nature of the larger interval. That means that the time zone only needs to be specified in the largest interval, that is, the Common Era; that time zone will be inherited by all smaller intervals. Thus, the Common Era can be considered as a function from time zones to intervals.

$$CE(z) = T$$

Fortunately, this counterintuitive conceptualization will usually be invisible and, for example, will not be evident in the most useful expressions for time, in Section 4.5 below. In fact, the CE predication functions as a good place to hide considerations of time zone when they are not relevant.

Time zones should not be thought of as geographical regions. Most places change their time zone twice a year, and a state or county might decide to change its time zone, e.g., from Central Standard to Eastern Standard. Rather it is better to have a separate ontology articulate the relation between geographical regions X times and time zones. For example, it would state that on a certain day and time a particular region changes its time zone from Eastern Standard to Eastern Daylight.

Moreover, time zones that seem equivalent, like Eastern Standard and Central Daylight, should be thought of as separate entities. Whereas they function the same in the time ontology, they do not function the same in the ontology that articulates time and geography. For example, parts of Indiana are always on Eastern Standard Time, and it would be false to say that they shift in April from that to Central Daylight time.

In this treatment it will be assumed there is a set of entities called time zones. Some relations among time zones are discussed in Section 4.5.

4.2. Clock and Calendar Units

The aim of this section is to explicate the various standard clock and calendar intervals. A day as a calender interval begins at and includes midnight and goes until but does not include the next midnight. By contrast, a day as a duration is any interval that is 24 hours in length. The day as a duration was dealt with in Section 3. This section deals with the day as a calendar interval.

It is useful to have three ways of saying the same thing: the clock or calendar interval y is the nth clock or calendar interval of type u in a larger interval x in time zone z. This can be expressed as follows for minutes:

$$min(y, n, x)$$

Because y is uniquely determined by n and x, it can also be expressed as follows:

$$minFn(n, x) = y$$

For stating general properties about clock intervals, it is useful also to have the following way to express the same thing:

$$clock\text{-}int(y, n, u, x)$$

This expression says that y is the nth clock interval of type u in x. For example, the proposition $clock\text{-}int(10 : 03, 3, *Minute*, [10 : 00, 11 : 00])$ holds.

Here u is a member of the set of clock units, that is, one of *Second*, *Minute*, or *Hour*.

In addition, there is a calendar unit function with similar structure:

$$cal\text{-}int(y, n, u, x)$$

This says that y is the nth calendar interval of type u in x. For example, the proposition $cal\text{-}int(12Mar2002, 12, *Day*, Mar2002)$ holds. Here u is one of the calendar units *Day*, *Week*, *Month*, and *Year*.

The unit *DayOfWeek* will be introduced below in Section 4.3.

The relations among these modes of expression are as follows:

$$\begin{aligned}
& sec(y, n, x) \equiv secFn(n, x) = y \\
& \quad \equiv clock\text{-}int(y, n, *sec*, x) \\
& min(y, n, x) \equiv minFn(n, x) = y \\
& \quad \equiv clock\text{-}int(y, n, *min*, x) \\
& hr(y, n, x) \equiv hrFn(n, x) = y \\
& \quad \equiv clock\text{-}int(y, n, *hr*, x) \\
& da(y, n, x) \equiv daFn(n, x) = y \\
& \quad \equiv cal\text{-}int(y, n, *da*, x) \\
& mon(y, n, x) \equiv monFn(n, x) = y \\
& \quad \equiv cal\text{-}int(y, n, *mon*, x) \\
& yr(y, n, x) \equiv yrFn(n, x) = y \\
& \quad \equiv cal\text{-}int(y, n, *yr*, x)
\end{aligned}$$

Weeks and months are dealt with separately below.

The am/pm designation of hours is represented by the function $hr12$.

$$\begin{aligned}
& hr12(y, n, *am*, x) \equiv hr(y, n, x) \\
& hr12(y, n, *pm*, x) \equiv hr(y, n + 12, x)
\end{aligned}$$

Each of the calendar intervals is that unit long; a calendar year is a year long.

$$\begin{aligned}
& sec(y, n, x) \supset second(y) \\
& min(y, n, x) \supset minute(y) \\
& hr(y, n, x) \supset hour(y) \\
& da(y, n, x) \supset day(y) \\
& mon(y, n, x) \supset month(y) \\
& yr(y, n, x) \supset year(y)
\end{aligned}$$

A distinction is made above between clocks and calendars because they differ in how they number their unit intervals. The first minute of an hour is labelled with 0; for example, the first minute of the hour [10:00,11:00] is 10:00. The first day of a month is labelled with 1; the first day of March is March 1. We number minutes for the number just completed; we number days for the day we are working on. Thus, if the larger unit has N smaller units, the argument n in clock$-$int runs from 0 to $N - 1$, whereas in $cal\text{-}int$ n runs from 1 to N. To state properties true of both clock and calendar intervals, we can use the predicate $cal\text{-}int$ and relate the two notions with the axiom

$$cal\text{-}int(y, n, u, x) \equiv clock\text{-}int(y, n - 1, u, x)$$

The type constraints on the arguments of $cal\text{-}int$ are as follows:

$$\begin{aligned}
& cal\text{-}int(y, n, u, x) \\
& \quad \supset interval(y) \wedge integer(n) \\
& \qquad \wedge temporal\text{-}unit(u) \wedge interval(x)
\end{aligned}$$

There are properties relating to the labelling of clock and calendar intervals. If N u's hath x and y is the nth u in x, then n is between 1 and N.

$$\begin{aligned}
& cal\text{-}int(y, n, u, x) \wedge Hath(S, N, u, x) \\
& \qquad \wedge member(y, S) \\
& \quad \supset 0 < n <= N
\end{aligned}$$

There is a 1st small interval, and it starts the large interval.

$$\begin{aligned}
& Hath(S, N, u, x) \\
& \quad \supset (\exists\, y)[member(y, S) \wedge cal\text{-}int(y, 1, u, x)] \\
& Hath(S, N, u, x) \wedge cal\text{-}int(y, 1, u, x) \\
& \quad \supset int\text{-}starts(y, x)
\end{aligned}$$

There is an nth small interval, and it finishes the large interval.

$$\begin{aligned}
& Hath(S, N, u, x) \\
& \quad \supset (\exists\, y)[member(y, S) \\
& \qquad \wedge cal\text{-}int(y, N, u, x)] \\
& Hath(S, N, u, x) \wedge cal\text{-}int(y, N, u, x) \\
& \quad \supset int\text{-}finishes(y, x)
\end{aligned}$$

All but the last small interval have a small interval that succeeds and is met by it.

$$\begin{aligned}
& cal\text{-}int(y1, n, u, x) \wedge Hath(S, N, u, x) \\
& \qquad \wedge member(y_1, S) \wedge n < N \\
& \quad \supset (\exists\, y_2)[cal\text{-}int(y_2, n + 1, u, x) \\
& \qquad \wedge int\text{-}meets(y_1, y_2)]
\end{aligned}$$

All but the first small interval have a small interval that precedes and meets it.

$$\begin{aligned}
& cal\text{-}int(y_2, n, u, x \wedge Hath(S, N, u, x) \\
& \qquad \wedge member(y_2, S) \wedge 1 < n \\
& \quad \supset (\exists\, y_1)[cal\text{-}int(y_1, n - 1, u, x) \\
& \qquad \wedge int\text{-}meets(y_1, y_2)]
\end{aligned}$$

If time is linearly ordered, the existential quantifier $\exists$ can be replaced by $\exists\,!$ in the above axioms.

4.3. Weeks

A calendar week starts at midnight, Saturday night, and goes to the next midnight, Saturday night. It is independent of months and years. However, we can still talk about the nth week in some larger period of time, e.g., the third week of the month or the fifth week of the semester. So the same three modes of representation are appropriate for weeks as well.

$$
\begin{aligned}
& wk(y,n,x) \equiv wkFn(n,x) = y \\
& \quad \equiv \mathit{cal\text{-}int}(y,n,{*}Week{*},x)
\end{aligned}
$$

As it happens, the n and x arguments will often be irrelevant.

A calendar week is one week long.

$$wk(y,n,x) \supset week(y)$$

The day of the week is a temporal unit (*DayOfWeek*) in a larger interval, so the three modes of representation are appropriate here as well.

$$
\begin{aligned}
& dayofweek(y,n,x) \\
& \quad \equiv dayofweekFn(n,x) = y \\
& \quad \equiv \mathit{cal\text{-}int}(y,n,{*}DayOfWeek{*},x)
\end{aligned}
$$

Whereas it makes sense to talk about the nth day in a year or the nth minute in a day or the nth day in a week, it does not really make sense to talk about the nth day-of-the-week in anything other than a week. Thus we can restrict the x argument to be a calendar week.

$$dayofweek(y,n,x) \supset (\exists\, n_1,x_1) wk(x,n_1,x_1)$$

The days of the week have special names in English.

$$
\begin{aligned}
& dayofweek(y,1,x) \equiv Sunday(y,x) \\
& dayofweek(y,2,x) \equiv Monday(y,x) \\
& dayofweek(y,3,x) \equiv Tuesday(y,x) \\
& dayofweek(y,4,x) \equiv Wednesday(y,x) \\
& dayofweek(y,5,x) \equiv Thursday(y,x) \\
& dayofweek(y,6,x) \equiv Friday(y,x) \\
& dayofweek(y,7,x) \equiv Saturday(y,x)
\end{aligned}
$$

For example, $Sunday(y,x)$ says that y is the Sunday of week x.

A day of the week is also a day of the month (and vice versa), and thus a day long.

$$
\begin{aligned}
& (\forall\, y)[[(\exists\, n,x) dayofweek(y,n,x)] \\
& \quad \equiv [(\exists\, n_1,x_1) da(y,n_1,x_1)]]
\end{aligned}
$$

One correspondance will anchor the cycle of weeks to the rest of the calendar, for example, saying that January 1, 2002 was the Tuesday of some week x.

$$
\begin{aligned}
& (\forall\, z)(\exists\, x) Tuesday(dayFn(1, monFn(1, \\
& \qquad yrFn(2002, CE(z)))), x)
\end{aligned}
$$

We can define weekdays and weekend days as follows:

$$
\begin{aligned}
& weekday(y,x) \\
& \quad \equiv [Monday(y,x) \vee Tuesday(y,x) \\
& \qquad \vee Wednesday(y,x) \vee Thursday(y,x) \\
& \qquad \vee Friday(y,x)] \\
& weekendday(y,x) \\
& \quad \equiv [Saturday(y,x) \vee Sunday(y,x)]
\end{aligned}
$$

4.4. Months and Years

The months have special names in English.

$$
\begin{aligned}
& mon(y,1,x) \equiv January(y,x) \\
& mon(y,2,x) \equiv February(y,x) \\
& mon(y,3,x) \equiv March(y,x) \\
& mon(y,4,x) \equiv April(y,x) \\
& mon(y,5,x) \equiv May(y,x) \\
& mon(y,6,x) \equiv June(y,x) \\
& mon(y,7,x) \equiv July(y,x) \\
& mon(y,8,x) \equiv August(y,x) \\
& mon(y,9,x) \equiv September(y,x) \\
& mon(y,10,x) \equiv October(y,x) \\
& mon(y,11,x) \equiv November(y,x) \\
& mon(y,12,x) \equiv December(y,x)
\end{aligned}
$$

The number of days in a month have to be spelled out for individual months.

$$
\begin{aligned}
& January(m,y) \\
& \quad \supset (\exists\, S) Hath(S,31,{*}Day{*},m) \\
& March(m,y) \supset (\exists\, S) Hath(S,31,{*}Day{*},m) \\
& April(m,y) \supset (\exists\, S) Hath(S,30,{*}Day{*},m) \\
& May(m,y) \supset (\exists\, S) Hath(S,31,{*}Day{*},m) \\
& June(m,y) \supset (\exists\, S) Hath(S,30,{*}Day{*},m) \\
& July(m,y) \supset (\exists\, S) Hath(S,31,{*}Day{*},m) \\
& August(m,y) \\
& \quad \supset (\exists\, S) Hath(S,31,{*}Day{*},m) \\
& September(m,y) \\
& \quad \supset (\exists\, S) Hath(S,30,{*}Day{*},m) \\
& October(m,y) \\
& \quad \supset (\exists\, S) Hath(S,31,{*}Day{*},m) \\
& November(m,y) \\
& \quad \supset (\exists\, S) Hath(S,30,{*}Day{*},m) \\
& December(m,y) \\
& \quad \supset (\exists\, S) Hath(S,31,{*}Day{*},m)
\end{aligned}
$$

The definition of a leap year is as follows:

$$
\begin{aligned}
& (\forall\, z)[\mathit{leap\text{-}year}(y) \\
& \quad \equiv (\exists\, n,x)[year(y,n,(CE(z)) \\
& \qquad \wedge [divides(400,n) \\
& \qquad\quad \vee [divides(4,n) \wedge \neg divides(100,n)]]]
\end{aligned}
$$

We leave leap seconds to specialized ontologies.

Now the number of days in February can be specified.

$$
\begin{aligned}
& February(m,y) \wedge \mathit{leap\text{-}year}(y) \\
& \quad \supset (\exists\, S) Hath(S,29,{*}Day{*},m) \\
& February(m,y) \wedge \neg \mathit{leap\text{-}year}(y) \\
& \quad \supset (\exists\, S) Hath(S,28,{*}Day{*},m)
\end{aligned}
$$

A reasonable approach to defining month as a unit of temporal measure would be to specify that the start and end points have to be on the same days of successive months.

$$
\begin{aligned}
& month(T) \\
& \quad \equiv (\exists\, d_1,d_2,n,x,m) \\
& \qquad [\mathit{in\text{-}interval}(\mathit{start\text{-}of}(T),d_1) \\
& \qquad \wedge \mathit{in\text{-}interval}(\mathit{end\text{-}of}(T),d_2) \\
& \qquad \wedge da(d_1,n,monFn(m,x)) \\
& \qquad \wedge da(d_2,n,monFn(mod{+}(m,1,12),x))]
\end{aligned}
$$

Here $mod+$ is modulo addition to take care of months spaning December and January. So the month as a measure of duration would be related to days as a measure of duration only indirectly, mediated by the calendar.

To say that July 4 is a holiday in the United States one could write

$$(\forall\, d, m, y)[da(d, 4, m) \wedge July(m, y) \\ \supset holiday(d, USA)]$$

4.5. Time Stamps

Standard notation for times list the year, month, day, hour, minute, and second. It is useful to define a predication for this.

$$time\text{-}of(t, y, m, d, h, n, s, z) \\ \equiv in\text{-}interval(t, secFn(s, minFn(n, hrFn(h, \\ daFn(d, monFn(m, yrFn(y, CE(z))))))))$$

For example, an instant t has the time

5:14:35pm PST, Wednesday, February 6, 2002

if the following properties hold for t:

$$time\text{-}of(t, 2002, 2, 6, 17, 14, 35, *PST*) \\ (\exists\, w, x)[in\text{-}interval(t, w) \\ \wedge Wednesday(w, x)]$$

The second line says that t is in the Wednesday w of some week x.

The relations among time zones can be expressed in terms of the time−of predicate. Two examples are as follows:

$$h < 8 \supset [time\text{-}of(t, y, m, d, h, n, s, *GMT*) \\ \equiv time\text{-}of(t, y, m, d-1, h+16, n, s, *PST*)] \\ h \geq 8 \\ \supset [time\text{-}of(t, y, m, d, h, n, s, *GMT*) \\ \equiv time\text{-}of(t, y, m, d, h-8, n, s, *PST*)]$$

$$time\text{-}of(t, y, m, d, h, n, s, *EST*) \\ \equiv time\text{-}of(t, y, m, d, h, n, s, *CDT*)$$

5. Deictic Time

Deictic temporal concepts, such as "now", "today", "tomorrow night", and "last year", are more common in natural language texts than they will be in descriptions of Web resources, and for that reason we are postponing a development of this domain until the first three are in place. But since most of the content on the Web is in natural language, ultimately it will be necessary for this ontology to be developed. It should, as well, mesh well with the annotation standards used in automatic tagging of text.

We expect that the key concept in this area will be a relation now between an instant and an utterance or document.

$$now(t, d)$$

The concept of "today" would also be relative to a document, and would be defined as follows:

$$today(T, d) \\ \equiv (\exists\, t, n, x)[now(t, d) \wedge in\text{-}interval(t, T) \\ \wedge da(T, n, x)]$$

That is, T is today with respect to document d if and only if there is an instant t in T that is now with respect to the document and T is a calendar day (and thus the nth calendar day in some interval x).

Present, past and future can be defined in the obvious way in terms of now and before.

Another feature of a treatment of deictic time would be an axiomatization of the concepts of "last", "this", and "next" on anchored sequences of temporal entities.

6. Aggregates of Temporal Entities

A number of common expressions and commonly used properties are properties of sequences of temporal entities. These properties may be properties of all the elements in the sequence, as in "every Wednesday", or they may be properties of parts of the sequence, as in "three times a week" or "an average of once a year". We are also postponing development of this domain until the first three domains are well in hand.

This may be the proper locus of a duration arithmetic, since we may want to know the total time an intermittant process is in operation.

7. Vague Temporal Concepts

In natural language a very important class of temporal expressions are inherently vague. Included in this category are such terms as "soon", "recently", "late", and "a little while". These require an underlying theory of vagueness, and in any case are probably not immediately critical for the Semantic Web. This area will be postponed for a little while.

Acknowledgments

I have profited from discussions with James Allen, George Ferguson, Pat Hayes, Adam Pease, and Stephen Reed, among others, none of whom however would necessarily agree entirely with the way I have characterized the effort. The research was funded by the Defense Advanced Research Projects Agency under Air Force Research Laboratory contract F30602-00-C-0168 and by the Advanced Research and Development Agency.

8. References

Allen, James F. and Henry A. Kautz. 1985. "A model of naive temporal reasoning." *Formal Theories of the Commonsense World*, ed. by Jerry R. Hobbs and Robert C. Moore, Ablex Publishing Corp., pp. 251-268.

Temporal Information in Collateral Texts for Indexing Movies

Andrew Salway and Eleftheria Tomadaki

Department of Computing
University of Surrey
Guildford, Surrey
GU2 7XH
United Kingdom
a.salway@surrey.ac.uk

Abstract

This paper suggests that video indexing is an interesting and important natural language application for which it is crucial to identify temporal information in collateral text that articulates the semantic content of moving images. Recently a rich source of information about the content of films and television programmes has become available in the form of audio description scripts. The analysis of the expression of temporal information in a corpus of audio description scripts leads to a discussion of some consequences for schemes to annotate such information in a video indexing application.

1. Introduction

The further development of digital libraries to retrieve, manipulate, browse and generate complex multimedia artefacts depends upon the machine-based representation of those artefacts, and in particular their 'semantic content'. An image can be understood at different levels of meaning: an image sequence, like a movie, can also tell a story by depicting a sequence of events. A crucial part of a film's semantic content is the narrative that it relates. As the story unfolds, the viewer constructs their understanding of the story guided by the director's careful sequencing of scenes and editing of shots. A machine-level representation of a film should maintain its rich structure and detail the entities, events and themes depicted; but how can a representation be instantiated for a given film?

One general approach to video indexing is based on the association of moving images and *collateral text* so that keywords, and potentially richer representations, are extracted from text fragments. Consider, for example, the speech of news and documentary presenters, sports commentaries and even newspaper film reviews. The challenge is to explicate the relationship between the moving image and the text. This involves dealing with temporal information in various ways; for example it is necessary to associate text fragments with the video intervals for which they are true; temporal relationships between the events depicted in the moving image must be extracted from the text; and, the time at which the action takes place must be ascertained.

A newspaper film review gives an incomplete and temporally re-ordered account of the events in a film. The speech of a newsreader is temporally aligned with the moving image but does not always refer to the visual information directly – much of a news broadcast comprises head and shoulders shots of newsreaders or stock video footage used to illustrate a story, thus keywords are more likely to be indicative of general story content than refer directly to what can be seen. By contrast, an audio description is a kind of 'narrative monologue' that gives a detailed account of what can be seen on screen in which the text order tends to follow the order of events in a programme or film. Audio description enhances the enjoyment of most kinds of films and television programmes for visually impaired viewers. In the gaps between existing speech the audio description gives key information about scenes, people's appearances and on-screen actions so that in effect the story conveyed by the moving image is retold in words.

We are interested in applying information extraction technology to generate machine-level representations of video content from audio description. It is hoped that the enhanced representation of video content could facilitate more complex querying ("find all clips showing X happening" – where X is a detailed description of events) and perhaps also contribute further to systems for video generation and maybe even question answering about what happened in a movie and why. As a first step towards information extraction we are considering the annotation of a corpus of audio description scripts to explicate what and how information is conveyed. At the moment priority is being given to temporal information because it seems to be crucial for the proper integration of moving images and collateral text.

Audio description is scripted before it is recorded. An audio description script is thus a text which is 'written to be spoken' and includes time-codes to indicate when each utterance is to be spoken. The task of processing audio description scripts is constrained because audio describers follow guidelines that restrict the language they use, i.e. normally the present tense, simple sentences and few pronominal references. This restricted language, the presence of time-codes and the relatively straightforward chronological order of the texts make audio description scripts a good starting point for extracting information for video indexing.

Though it is straightforward to associate a time-coded text fragment approximately with a video interval, a more precise association requires consideration of tense and aspect. For example, consider how the following fragments relate to intervals in the moving image: they are from audio description for *The English Patient* – time codes are in the format [minute:second][1].

[1] This sample is reproduced from *The English Patient*. Please note that further examples in the paper are fictitious but based closely on actual audio description and maintain grammatical structure (i.e. only names and events have been changed).

[11:43] Hanna passes Jan some banknotes – a near instantaneous event in the present tense, so the fragment relates to a short video interval at the time of speaking;

[11:55] Laughing, Jan falls back into her seat – the present participle indicates that 'laughing' is ongoing and so relates to a longer video interval that includes the instantaneous 'falls back';

[12:01] An explosion on the road ahead – use of nominalisation to refer to an event;

[12:08] The jeep has hit a mine – the present perfect indicates that the event is completed and the video interval that the text relates to must have start and end points before the time-code of the text fragment (general knowledge tells us that this event occurred before the explosion and was its cause).

Once text fragments have been associated with video intervals the events depicted in the steady flow of the moving image must be related to each other according to a different time-line – that of the diegetic world depicted by the movie. For example the 'hit mine' event happens immediately before the 'explosion' described above and it might be appropriate to label the relationship with causality. There are also examples of simultaneous and included events, such as – *he prevents her from leaving, holding her firmly*. Events in a movie are grouped in scenes where each scene has a (normally) unique combination of time and location. In audio description an explicit time reference might be used to introduce a new scene, e.g. *October 1944*; *later* is also used to introduce scenes and indicate story progression.

This paper suggests that video indexing is an interesting and important natural language application for which it is crucial to identify and analyse temporal information in collateral texts that articulate the semantic content of moving images. A review of video retrieval systems shows that the use of collateral text is important, but in order to extend the approach to more kinds of video material and collateral text it will be essential to process temporal information. The conceptualisation of time and events with respect to semantic content in digital video systems is outlined, particularly for films (Section 2).

We attempt to formalise the challenge of integrating video data and collateral text by describing three tasks that would contribute to the use of collateral text for video indexing. These tasks guided the analysis of an audio description script corpus (70,856 words): prominent expressions of temporal information are quantified and exemplified (Section 3). The results begin to give a basis for discussing what would be required of a scheme to annotate temporal information in this scenario: existing annotation schemes are reviewed and some tentative extensions are proposed (Section 4). The paper closes by considering further directions for this work (Section 5).

2. Digital Video Systems

Video data can be indexed with visual features based on the distribution of pixels, e.g. colour, texture, shape and motion: however a 'semantic gap' appears between video databases and users who often conceive their information needs in terms of the relationships between entities, events and themes to be depicted in the video sequence of interest. Indexing could be achieved by attaching keywords and other descriptors manually to either whole video data files or intervals and regions within them. A cheaper alternative is to use language technology to process 'collateral text'; Srihari introduced this term to refer to textual information associated with visual information, specifically photo captions (1995). Video data sometimes includes an *integral* textual component in the form of speech and closed-captions. Other *external* textual information arises in the production and distribution of video material, e.g. scripts and production notes, and now audio description (legislation in a number of countries makes the provision of audio description mandatory for an increasing amount of digital TV and film output).

In news broadcasts and documentary programmes much of the information content is carried by the spoken words of the presenters, and the subjects on which they are speaking will reflect, albeit to varying degrees, the entities, events and themes shown in the accompanying moving images. The *Informedia* system indexes news broadcasts and documentary programs by keywords that are extracted from speech and closed captions: since the speech is time-aligned with the moving images the keywords can be associated with specific video intervals (Wactlar et al., 1999). This approach has been extended into a multi-lingual context in the Pop-Eye and Olive projects, and to deal with sports footage in the current MUMIS project (de Jong et al., 2000). Other researchers have applied text segmentation techniques to the speech stream of video data in order to segment video sequences (Mani et al., 1997; Takeshita, Inoue and Tanaka, 1997). The transcribed speech of news presenters has been exploited in a system for browsing through news broadcasts by following hypertext links between terms and viewing associated video sequences (Shahraray, 1999). Research and systems focused on accessing broadcast news, including further tasks like multi-stream segmentation, combined name/face recognition and multimedia summarisation are collected in Maybury (2000).

There are moving images that do not contain 'integral' text, but that can be indexed with text that was produced specifically to elucidate the video's content. The *WebSEEK* system, which has indexed hundreds of thousands of images and videos on the WWW, selects keywords from the text of hyperlinks to images and videos on WWW-pages (Smith and Chang, 1997); note that this system only indexes whole videos rather than intervals. Another system, developed at the Japan Broadcasting Corporation, parses the notes kept in the production of wildlife documentary programs that describe the entities and events in the recorded footage and are time-coded. Queries for video intervals can be made in terms of the relationships between entities and actions (Kim and Shibata, 1996).

More recently there have been developments to combine visual and textual features for the classification of video sequences. For example, visual features may indicate the location of a scene (indoors/outdoors) and whether there is one or many people in the shot, and textual features may indicate the nature of the spoken words (a news report / a political speech): taken together these features suggest whether a video sequence depicts a political rally, an outside news broadcast, or a political party's conference (Satoh, Nakamura and Kanade, 1999).

Textual information from TV sit-com scripts has been combined with visual features, through a process involving user interaction, so a system can locate scenes containing a particular character (Wachman and Picard, 2001).

Collateral text could potentially be used for extracting information about other kinds of video, including those with rich semantic content like films and dance sequences. In specialist domains, like dance, there is an extensive range of collateral texts available (dance programmes, newspaper reviews, textbooks, choreographer's notes, biographies, etc.) and spoken commentaries can be elicited from experts asked to 'describe' and to 'interpret' sequences. The KAB system was developed to index fixed-length intervals of dance videos with keywords from such commentaries: this work also specified requirements for a general system to process diverse collateral texts (Salway and Ahmad, 1998; Salway, 1999). A key requirement is a video data model and representation scheme that captures semantic video content, including temporal organisation, at an appropriate level of detail to facilitate complex queries, browsing and even video generation; the link with collateral text also needs to be modelled.

In the video database literature semantic content is usually treated as comprising the objects and events depicted by a moving image and the spatio-temporal relationships that hold between them; for a survey see Chen, Kashyap and Ghafoor (2000). Descriptions of objects and events (as keywords and propositions, for example) are associated with intervals of video data which can be modelled either as a hierarchy of discrete intervals (Weiss, Duda and Gifford, 1995) or as multi-layered overlapping intervals (Davenport, Aguierre Smith and Pincever, 1991).

Allen's (1983) temporal logic has been applied widely in video data models to facilitate reasoning about video content and more complex queries: the number of the 13 possible interval relationships that are used varies between applications. A hierarchical model is appropriate for dealing with film in terms of scenes and shots (see Corridoni et al., 1996). However, to capture more detail about the events within a shot it might be necessary to allow for overlapping video intervals and more description of the relationships between events.

Knowledge representation schemes aim to provide unambiguous representations of meaning and to facilitate inferencing: a number of proposals have been made to use such schemes for semantic video content. Regarding the composition of events/sub-events in moving images, particularly in stereotypical situations, a framework was developed based on Schank and Abelson's scripts (1977), see (Parkes, 1989; Nack and Parkes, 1997). Semantic networks have been used in a video browsing system to elaborate the description of events, for example to specify participants and causal relationships between events (Roth, 1999). The use of conceptual dependency graphs and story grammars has also been discussed (Tanaka, Ariki and Uehara, 1999). Independent to this, but sharing some similar goals, researchers in computer vision have proposed levels to deal with complex visual information at stages from raw visual input to final representation, e.g. 'change – event – verb – history' (Nagel, 1988), and specifically for human motion 'movement – activity – action' (Bobick, 1997).

3. Temporal Information in Collateral Text

This section characterises the expression of temporal information in a corpus of audio description scripts with respect to three tasks we consider important for video indexing with collateral text. First though, in order to extend the use of collateral text to index films it is necessary to explicate how a linear text relates to a film with multi-faceted content. The discussion here is limited to film content that is conveyed visually, and hence accessible through audio description - we are not currently considering dialogue and sound effects. The focus is on films and accompanying audio description but much of what is discussed could be relevant to other kinds of video and other collateral text types.

3.1. Integrating Moving Images and Text

In order to integrate audio description text with film at a semantic level it is necessary to deal with film in terms of the shots and scenes by which it is structured. It is also important to recognise two timelines: (i) film time, i.e. the time it takes to watch the film; and, (ii) story time, i.e. the time in which the events depicted take place. Figure 1 shows how a film (stored as a video data file) can be modelled in terms of shots which are defined as continuous pieces of filming, and scenes which are characterised by each having a unique combination of location and time. The story timeline is shown in parallel with layers of events taking place. Of course the relative position of events may differ between the two timelines, e.g. the film may depict events in a different order than they happen in the story, and events that are happening at the same time but in different locations will be depicted in different scenes. For video retrieval purposes it is important to maintain temporal relationships between events; different sub-sets of Allen's relationships will be required for different applications.

The structure of film provides some useful constraints when dealing with temporal information. It is reasonable to assume that all events depicted within a scene take place close together in the story timeline, and are likely to form larger events (information about scene boundaries may be available from sources like film scripts and automatic video analysis). When considering how events are depicted at the shot level it is important to note film-making techniques that are used to convey that an event is taking place, or has taken place, without showing it in its entirety; a director may choose to portray only the end result of an event and allow the viewer to infer that the event took place.

The collateral text is shown as a series of time points that indicate the time at which the speaker starts the utterance (assuming a temporally aligned collateral text, like an audio description). The three tasks outlined next relate to the extraction of temporal information from collateral text to: (i) associate an utterance with the video interval for which it is true, be it a shot, scene or some other interval; (ii) specify event-event relationships – here we only consider relationships holding within a scene (in film time); (iii) establish the time at which scene is set (in story time).

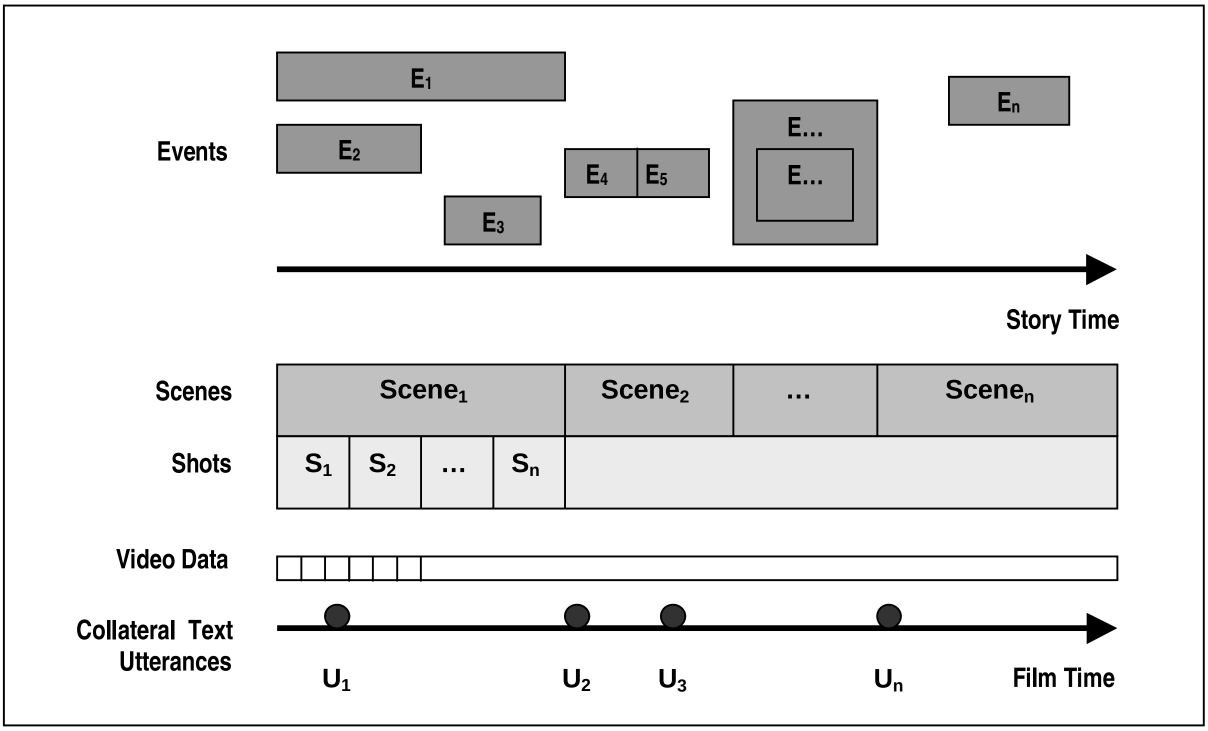


Figure 1. The organisation of a film's content in terms of shots and scenes (which relate to film time) and the events that comprise the semantic video content (which relate to story time). Collateral text such as audio description is temporally aligned with the video data in film time.

3.1.1. Task 1: Associate an audio description fragment with the interval in film time for which it is true.

Given a time coded text fragment it is relatively straightforward to associate it approximately with the video interval for which it is true, i.e. the interval in which the event it refers to is taking place; the time-code plus and minus an arbitrary number of seconds works as a crude approximation of start and end times. However, it is desirable to be more precise about at least one of: start time, end time or duration. (A greater challenge, not addressed here, is to ascertain whether the event is depicted on-screen throughout the duration). As well as events, it is also appropriate to deal with states if they change significantly during the movie, e.g. to indicate scenes in which a character is a child or grown-up.

The problem can be gauged to some extent by considering an earlier feasibility study for indexing moving images with audio description (Turner, 1998). A small sample of video material with accompanying audio description was analysed (including a film and various kinds of television programme). Results showed that overall about 50% of shots were described but only about 40% of the audio description referred directly to the shot on-screen at the time of speaking.

To ascertain appropriate video intervals it may help to consider some of the aspectual features of events classified by Comrie (1976). Whether an event has internal structure (*punctual* / *durative*) gives some information about its duration; this may be an inherent characteristic of a verb but may be modified grammatically, e.g. with the progressive. Knowing about an event's end result, if it has one (*telic* / *atelic*), gives information about its completion (and in audio description may be all that is referred to).

3.1.2. Task 2: Event-event relationships in story time (within the same scene).

Moving images can depict many events at the same time, and in the case of film the temporal organisation of events and relationships such as event / sub-event and causality are crucial to a viewer's understanding. As discussed previously some or all of Allen's 13 temporal relationships might be needed, though whether they can all be extracted from collateral text remains to be seen. In a narrative dialogue by default events are mentioned in the order in which they occur – however, events may occur simultaneously, or there may be stylistic reasons to mention them out of order.

Ascertaining basic temporal relationships, like before / after / overlapping, may be possible just from the collateral text. However, to construct rich representations of composite events within scenes perhaps relies more on prior 'world knowledge' than it does on information immediately available in a narrative monologue (cf. the use of Schank's scripts to deal with semantic video content); the problem becomes harder still when event-event relationships across different scenes are considered. A lexical resource, like WordNet, might help as a first step to relate events, in light of the entailment relations for verbs described by Fellbaum (1998). When considering temporal inclusion some sets of verbs are *co-extensive*, e.g. 'march and walk', 'whisper and talk'; whereas other share a relationship of *proper inclusion*, e.g. 'sleep and snore'. Having access to these relationships may help to associate descriptions of the same event and to establish sub-event relationships. Other relationships allow events to be associated according to *backward presupposition*, e.g. 'forget and know', and on the basis of *causality*, e.g. 'show and see'.

3.1.3. Task 3: Establish the time a scene takes place (in story time).

A viewer's appreciation of a film requires knowing when it is set, and if it is set over a long time period then the time of each scene must be known – thus information needs to be extracted to give each scene a time. Unless a film is based upon true-life events then it is normally set within a time period without specific dates being implied. Similarly, within the course of a day in story time exact times are usually less important than whether the viewer knows it is morning, afternoon, evening or night (of course exact times will be crucial for some plots). Unless otherwise indicated the assumption is that scenes are ordered sequentially according to the story timeline, but for some movies the use of flashback will have to be dealt with.

3.2. Temporal Information in Collateral Text: a case study with audio description scripts

The intention of this case study is to quantify and exemplify prominent expressions of temporal information in audio description scripts: the analysis is organised around the three tasks for video indexing described in the previous section. The corpus comprises audio description scripts for 12 movies, covering a range of movie genres, and written by six different describers. It currently totals 70,856 words – this will be expanded to around 500,000 words in coming months.

When carrying out the analysis we considered the variety of ways temporal information can be expressed in English as outlined by Quirk et al. (1985), i.e. by using tense, aspect, adverbials, prepositional phrases, subordinate clauses, nouns and proper nouns. Of course in narrative monologues text order is highly informative about the order in which events take place. Our 'conceptualisation of time' is guided by approaches to video data modelling, i.e. the association of event and state descriptions with video intervals, the specification of interval relationships following Allen (1983), and the organisation of complex events using knowledge representation schemes. Theoretical perspectives on events, such as Comrie's classification of aspect (1976) and Fellbaum's entailment relations (1998) were also considered.

Based on the 50 most frequent verbs in the corpus it appears that the majority of events are material processes (84%), with some mental processes (10%), a few relational processes (4%) and a few behavioural processes (2%), following Halliday (1994).

3.2.1. Information to Associate Text Fragments with Video Intervals

The present tense proliferates in the audio description corpus. It is even used to describe events that are about to happen, for example to describe speech acts which cannot be described at the time they occur – *the doctor questions Tom*. The occasional use of the present perfect is important to describe events after they happened (possibly because there was not an opportunity to describe them at the time they occurred, or because only the end result is depicted on screen) – *the cake has been eaten*. Past events are also sometimes referred to in relative clauses used to identify unnamed characters – *the woman who visited Paul is walking down the street*. In order to be more precise about the start, end or durations of events it seems that a variety of aspectual information is important, especially aspectual verbs and the inherent aspect of verbs.

In the audio description corpus the verbs *start*, *stop*, *begin* and *finish* occur relatively far more frequently than they do in the general language British National Corpus (BNC), Table 1 shows just 3rd person singular forms. These verbs almost always appeared in the present tense to refer to another event so it would be straightforward to use them to compute their arguments' start and end times.

Verb	Abs. Freq	Ratio with BNC
stops	105	65.79
starts	60	25.13
begins	19	4.67
finishes	3	25.65

Table 1: Showing prominent aspectual verbs in Surrey's audio description corpus (only for 3rd person singular). The third column is calculated by dividing relative frequency in the audio description corpus with relative frequency in the British National Corpus (BNC)

Regarding the duration of events the adverb *still* is frequently used with durative events that have not finished at the time of speaking (85 occurrences of *still* in the corpus; 62 of these are in the time sense). There was little sign of time expressions giving information about exact durations but relatively short periods were frequently indicated with *for a moment* (29 occurrences). Other frequent, adverbs like *slowly* (111 occurrences) and *quickly* (20) might make a small contribution to understanding the duration of an event.

The grammatical marking of progressive aspect does not appear to be significant for the task of associating text fragments with video intervals. In a narrative monologue we learn nothing about the duration of the event from the distinction between 'he runs' and 'he is running'; the simple present and the progressive are used interchangeably in the corpus. In fact it is probably an event's inherent aspect that is most important to determine, at least approximately, its duration in film time. In general language this will be problematic given that the many senses of common verbs often have different aspectual characteristics, however in specialist domains it may be possible to store default durations for events, like dance movements.

3.2.2. Information to Specify Event-Event Relationships in Film Time

The most frequent conjunction to indicate events happening at the same time in the audio description corpus was *as* (350 occurrences in the time sense) – *the children* *play as the crowd moves away*; sometimes *as* indicates more of a connection between events – *she continues to hide as the monster approaches*. Used only to indicate simultaneity (without implying further connection between events) *while* occurred 37 times. Both these conjunctions indicate some degree of overlap between events but further information is required to know whether the events are strictly simultaneous, whether one is included within the other, or if they simply overlap. Non-finite verbs with sub-ordinate clauses tend to indicate that the second event is included in the first – *Coughing, Mary gives the medicine to Tom*. When linking events *and* was 'ambiguous' as to whether the events occurred serially or in parallel.

Occurring only in its time sense, *then* (173 occurrences) was still relatively more frequent in the audio description corpus than the BNC. Though it is redundant as far as indicating sequence is concerned (that is already conveyed by text order) it does seem to imply the completion of the first event before the start of the next one – *Sarah chops the tomatoes then fries an egg*. Furthermore in many examples it suggests that the events meet in time, i.e. the endpoint of one equals the start of the other. (This kind of information could be useful in relation to our Task 1). The less frequent *when* (29 occurrences) and *until* (20) were used respectively to indicate the start and end of events in relation to other events or states, often suggesting that the first event led to the second – *she walks through the forest until she finds the house*.

Like *then*, *now* (40 occurrences) adds little or nothing by way of basic temporal information in these narrative monologues, however it does seem to indicate a change or contrast between two events across a passage of audio description – *Jane is dancing with George … Now she is dancing with her cousin*. Perhaps surprisingly, *after* and *before* occur relatively infrequently in the corpus (compared with the BNC) and when they are used they only serve to emphasise the sequence of events already conveyed by text order, i.e. we find 'after E1, E2' but not 'E2 after E1', and 'E1 before E2' but not 'before E2, E1'.

The adverb *again* is prominent in the corpus (141 occurrences – 2.5 times relatively more often than in the BNC). It generally indicates that an event is happening for a second time within a scene – for video retrieval purposes it might be useful to relate the two instances.

3.2.3. Information to Specify When Scenes Take Place in Story Time

The most frequent time expressions used to locate a scene on the story timeline relate to non-specific times of day: *night* (37 occurrences), *morning* (19), *evening* (11), *dusk* (6), *dawn* (3). Less frequent were expressions for non-specific times of year, i.e. months (without years), seasons and festival days (17 occurrences in total). This probably reflects the fact that the progression of time during a film is more often at the granularity of days. The relative paucity of specific times and dates (there were only a few examples) is explained in part by the fact that for many films the viewer need only understand a general time period. This may be conveyed by costumes, props and, for times of day, lighting: these will all be referred to by audio description.

Scenes are sometimes introduced with one of the time expressions mentioned above – indicating a change of time is a shortcut to indicate a new scene. Quite often *later* (32 occurrences) is used for this purpose and as such may be a useful cue for scene segmentation.

4. An Annotation Scheme for Temporal Information in Collateral Text?

Based on the preceding analysis this section discusses some tentative requirements of an annotation scheme for temporal information in collateral text: such a scheme would be a step in applying information extraction technology to the task of video indexing. The extent to which existing schemes would cater for these requirements is reviewed. A number of factors suggest that some extensions to existing schemes will be required: (i) there seems to be a need to maintain two timelines; (ii) if practical in terms of time and inter-annotator agreement, it would be desirable to record aspectual information regarding the internal structure of events and end-states; (iii) also subject to practicality, it is important for film to specify sub-event and causal relationships.

A canonicalized representation of times was proposed as part of a set of guidelines for annotating temporal expressions by Mani et al. (2001), who targeted a variety of text genres such as both print and broadcast news, and meeting scheduling dialogues. The emphasis of their approach was on detailing different classes of time expressions like points in time (when), durations (how long) and frequencies (how often) and handling context-dependent expressions. It also addressed fuzzy temporal boundaries that arise from the use of phrases that refer to times of year and times of the day, e.g. *summer* and *morning*, and addressed non-specific times, such as *a sunny day in April* (not a specific day, nor a specific year).

Of the time expressions dealt with it is points in time that seem to apply most directly in our scenario in order to locate events (at the granularity of scenes) on the story timeline (our Task 3). Though duration and frequency relate to the kinds of aspectual characteristics that we would like to describe for events, they annotate only words and phrases that express this information directly; though there were some frequent phrases in the audio description corpus for which it might be applicable – *for a moment*.

Another scheme that has been proposed is more concerned with associating temporal information with events, and annotating the temporal relationships between events (Setzer and Gaizauskas, 2000; Setzer, 2001); this scheme was developed initially for newswire texts but is extensible. Annotations are attached to the heads of finite verb groups as representatives of events, as well as to temporal expressions. It is possible to specify the type of event (Occurrence / Perception / Reporting / Aspectual) as well as the tense and grammatical aspect of the verb. The annotations have attributes to specify five event-event and event-time relationships: 'before', 'after', 'includes', 'included' and 'simultaneous'. The features of the scheme that have been summarised here are exemplified in Appendix A which shows the annotation of 9 utterances of audio description.

The annotation of event-event relationships within a scene (our Task 2) would be dealt with quite comprehensively by Setzer's and Gaizauskas' scheme: although as many as 13 temporal relationships (from Allen) are discussed in the video retrieval literature the five used in this scheme would probably serve most purposes. Being able to annotate aspectual events, i.e. to indicate the start and end time of occurrence events, is certainly important given their frequency in the audio description corpus - cf. our Task 1. For other parts of Task 1 it might be necessary to extend the scheme to specify the start and end of events when there is no explicit time expression, or to do it relative to the time-code in the text; a further minor extension would be to allow for the annotation of states as well as events. It certainly would be desirable to be able to specify causal and sub-event relationships between events as these are crucial to the narrative structure of movies, however this would depend upon annotators' ability to apply them consistently.

5. Closing Remarks

Dealing with temporal information is an important first step towards generating machine-level representations of video content from collateral text, especially when dealing with a complex multimedia artefact, like film, and richly informative collateral text, like audio description. This work is in its early stages but the three tasks outlined here begin to give us a handle on some of the challenges involved in integrating moving images and narrative monologues. The corpus analysis showed an extensive range of temporal information that needs to be dealt with in respect to these tasks. Progress will be made by more extensive application of existing annotation schemes leading to decisions about exactly what is required by way of extensions. Such decisions need to be informed by considerations of any new scheme's practicality (is it simple enough to be applied consistently and quickly) and the extent to which it captures important information (the criteria for which will vary between video applications and users). The final test would perhaps be a comparison of video retrieval using: (i) unannotated audio description (i.e. relying on time codes and text order alone); (ii) annotated audio description (with no further processing); and, (iii) machine-based representations generated from annotated audio description.

6. Acknowledgements

This research was carried out as part of the Television in Words project (TIWO) supported by an Engineering and Physical Sciences Research Council (EPSRC) grant, GR/R67194/01. The authors would like to thank the members of the TIWO Round Table for sharing their knowledge of audio description and providing samples for our corpus. Finally, we are very grateful for the comments of two anonymous reviewers and have tried to take heed.

7. References

Allen, J.F., 1983. Maintaining Knowledge About Temporal Intervals. *Communications of the ACM* 26 (11):832-843.

Bobick, Aaron F., 1997. Movement, Activity, and Action: The Role of Knowledge in the Perception of Motion. *Philosophical Transactions of the Royal Society of London Series B – Biological Sciences* 352 (1358):1257-1265.

Chen, S.-C., R.L. Kashyap, and A. Ghafoor, 2000. *Semantic models for Multimedia Database Searching and Browsing*. Kluwer Academic Publishers.

Comrie, B., 1976. *Aspect: an introduction to the study of verbal aspect and related problems*. Cambridge University Press.

Corridoni, J.M., A. Del Bimbo, D. Lucarella, and H. Wenxue, 1996. Multi-perspective Navigation of Movies. *Journal of Visual Languages and Computing,* 7:445-466.

Davenport, G., T. Aguierre Smith, and N. Pincever, 1991. Cinematic Primitives for Multimedia. *IEEE Computer Graphics and Applications* July:67-74.

de Jong, F., J.-L. Gauvain, D. Hiemstra, and K. Netter, 2000. Language-Based Multimedia Information Retrieval. *Proceedings RIAO 2000 Content-Based Multimedia Information Access, Paris, April 2000*, 713-722.

Fellbaum, C., 1998. A Semantic Network of English Verbs. In Fellbaum (eds.), *WordNet: an electronic lexical database*. Cambridge MA: The MIT Press.

Halliday, M. A. K., 1994. *An Introduction to Functional Grammar*. London: Edward Arnold, 2nd edition.

Kim Y.-B., and M. Shibata, 1996. Content-Based Video Indexing and Retrieval – A Natural Language Approach. *IEICE Transactions on Information and Systems* E79-D(6):695-705.

Mani I., D. House, M.T. Maybury, and M. Green, 1997. Towards Content-Based Browsing of Broadcast News Video. In M. Maybury (ed.), *Intelligent Multimedia Information Retrieval*. Menlo Park CA / Cambridge MA: AAAI Press / MIT Press.

Mani, I., G. Wilson, L. Ferro, and B. Sundheim, 2001. Guidelines for Annotating Temporal Information. *Procs. HLT 2001, First International Conference on Human Language Technology Research*, San Francisco: Morgan Kaufmann.

Maybury, Mark, 2000 (ed.). Special Issue – 'News on Demand'. *Communications of the ACM,* 43(2).

Nack, F., and A. Parkes, 1997. Toward the Automated Editing of Theme-Oriented Video Sequences. *Applied Artificial Intelligence*, 11:331-366.

Nagel, H.-H., 1988. From image sequences towards conceptual descriptions. *Image and Vision Computing,* 6(2):59-74.

Parkes, A. P., 1989. The Prototype CLORIS System: Describing, Retrieving and Discussing Videodisc Stills and Sequences. *Info. Proc. and Management,* 25(2):171-186.

Quirk, R., S. Greenbaum, G. Leech, and J. Svartvik, 1985. *A Comprehensive Grammar of the English Language.* London and New York: Longman.

Roth, V., 1999. Content-based retrieval from digital video. *Image and Vision Computing,* 17:531-540.

Salway, A., 1999. *Video Annotation: the role of specialist text*. Ph.D. thesis, University of Surrey.

Salway, A., and K. Ahmad, 1998. Talking Pictures: Indexing and Representing Video with Collateral Texts. *Procs. 14th Workshop on Language Technology -*

Language Technology for Multimedia Information Retrieval, 85-94.

Satoh, S., Y. Nakamura, and T. Kanade, 1999. Name-it: Naming and detecting faces in news videos. *IEEE Multimedia*, 6 (1):22-35.

Schank, R. C. and R. P. Abelson, 1977. *Scripts, Plans, Goals and Understanding: an inquiry into human knowledge structures*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Setzer, A., 2001. *Temporal Information in Newswire Articles: An Annotation Scheme and Corpus Study*. Ph.D. thesis, University of Sheffield.

Setzer, A., and R. Gaizauskas, 2000. Annotating Events and Temporal Information in Newswire Texts. *Procs. LREC 2000, 2nd Int. Conf. On Language Resources and Evaluation*, 1287-1293.

Shahraray, B., 1999. Multimedia Information Retrieval Using Pictorial Transcripts. In B. Furht (ed.), *Handbook of Multimedia Computing*. Florida: CRC Press.

Smith J.R., and S.-F. Chang, 1997. Visually Searching the Web for Content. *IEEE Multimedia* July-Sept:12-20.

Srihari, R.K., 1995. Computational Models for Integrating Linguistic and Visual Information: A Survey. *Artificial Intelligence Review*, 8(5-6):349-369.

Takeshita, A., T. Inoue, and K. Tanaka, 1997. Topic-based Multimedia Structuring. In M. Maybury (ed.), *Intelligent Multimedia Information Retrieval*. Menlo Park CA / Cambridge MA: AAAI Press / MIT Press.

Tanaka, K., Y. Ariki, and K. Uehara, 1999. Organization and Retrieval of Video Data. *IEICE Trans. on Information and Systems*, E82-D(1):34-44.

Turner, J.M., 1998. Some Characteristics of Audio Description and the Corresponding Moving Image. *ASIS Annual Meeting*, 35:108-117.

Wachman, J. S., and R.W. Picard, 2001. Tools for Browsing a TV Situation Comedy Based on Content Specific Attributes. *Multimedia Tools and Applications*, 13(3):255-284.

Wactlar, H.D., M.G. Christel, and Y. Gong, and A.G. Hauptmann, 1999. Lessons Learned from Building a Terabyte Digital Video Library. *Computer*, Feb:66-73.

Weiss, R., A. Duda, and D.K. Gifford, 1995. Composition and Search with a Video Algebra. *IEEE Multimedia*, Spring 1995:12-25.

Appendix A

Annotation of an audio description script following Setzer's scheme

The following passage of audio description (from *The English Patient*) has been annotated following the scheme and guidelines given by Setzer (2001). The sample here exemplifies how: (i) tense and aspect features can be associated with an event's verb; (ii) how the class of an event can be noted; (iii) how relationships between events can be specified. The sequence of events inherent in the text order has not been annotated, though it could have been – only exceptions to the 'default' have been annotated, e.g. simultaneous events, and events that are mentioned in a different order to which they occur.

[11.43] Hanna <event eid=1 tense=present class=occurrence> passes </event> Jan some banknotes.
[11.55] <event eid=2 tense=present class=occurrence aspect=progressive relatedToEvent=3 eventRelType=includes> Laughing </event>, Jan <event eid=3 tense=present class=occurrence relatedToEvent=4 eventRelType=simultaneous signalID=1> falls </event> back into her seat <signal sid=1> as </signal> the jeep <event eid=4 tense=present class=occurrence> overtakes </event> the line of the lorries.
[12.01] An <event eid=5 tense=present class=occurrence relatedToEvent=6 eventRelType=after> explosion </event> on the road ahead.
[12.08] The jeep has <event eid=6 tense=present class=occurrence aspect=perfective > hit </event> a mine.
[12.09] Hanna <event eid=7 tense=present class=occurrence> jumps </event> from the lorry.
[12.20] Desperately she <event eid=8 tense=present class=occurrence> runs </event> towards the mangled jeep.
[12.27] Soldiers <event eid=9 tense=present class=occurrence> try </event> to stop her.
[12.31] She <event eid=10 tense=present class=occurrence> struggles </event> with the soldier who <event eid=11 tense=present class=occurrence> grabs </event> hold of her firmly.
[12.35] He <event eid=12 tense=present class=occurrence> lifts </event> her bodily from the ground <event eid=13 tense=present class=occurrence aspect=progressive relatedToEvent=12 eventRelType=simultaneous> holding </event> her tightly in his arms.

Recognizing and Tagging Temporal Expressions in Spanish

Estela Saquete, Patricio Martínez-Barco and Rafael Muñoz
Departamento de Lenguajes y Sistemas Informáticos. Universidad de Alicante
Estela.Saquete@ua.es
{patricio, rafael}@dlsi.ua.es

Abstract

This paper shows a system about the recognition of temporal expressions in Spanish and the resolution of their temporal reference. For the identification and recognition of temporal expressions we have based on a temporal expression grammar and for the resolution on an inference engine, where we have the information necessary to do the date operation based on the recognized expressions. For further information treatment, the output is proposed by means of XML tags in order to add standard information of the resolution obtained. Different kinds of annotation of temporal expressions are explained in another articles [WILSON2001][KATZ2001]. In the evaluation of our proposal we have obtained successful results.

1. Introduction

The study of anaphora phenomena has been carried out for a lot of researches. Most of these researches have focused on pronominal anaphora and a few of them on definite descriptions. Most of temporal expressions could be considered as a type of definite description, but a few of them are temporal adverbs like "mañana" (*tomorrow*). The research work developed in definite description is focused on establishing a relationship between anaphoric expressions and their antecedents. In these work, if the definite description is a temporal expression it has been only solved establishing the relationship but not inferring the new date. The resolution of temporal expressions involves the recognition of them and the inference of the new date.

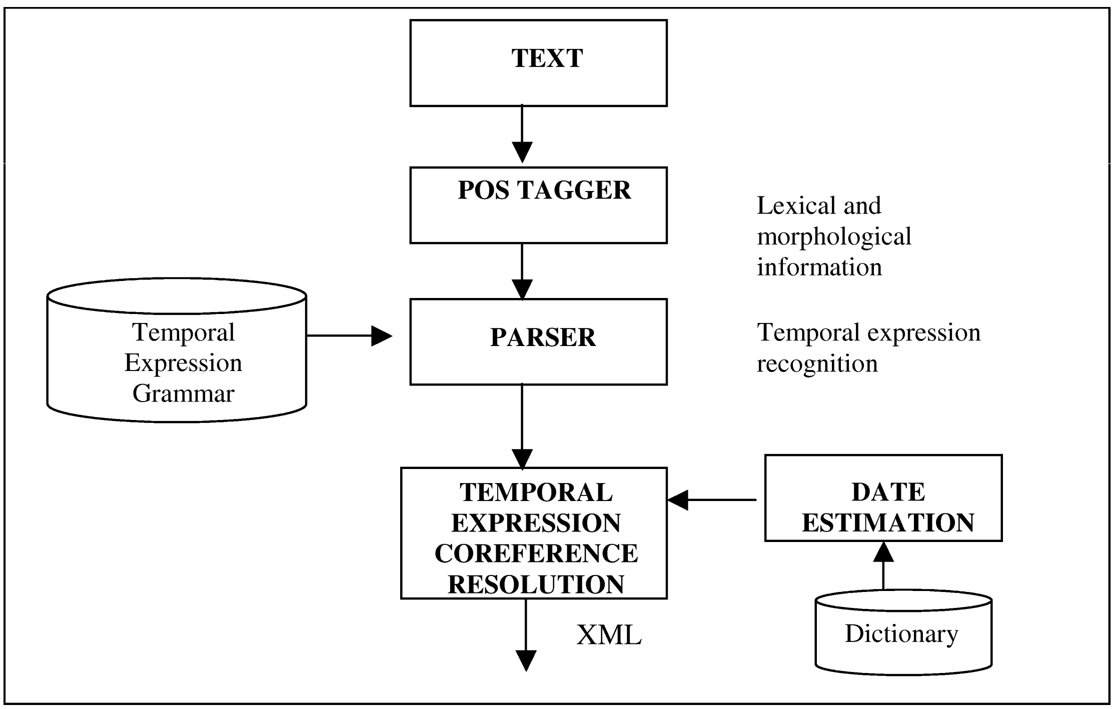


Fig. 1. Graphic representation of the system proposed

Work focused on temporal expression should to solve both tasks. In the literature we can found several studies focused on temporal expressions (Guillen et *al*. 1995) (Wiebe et *al.* 1998). These studies are based on the use of a temporal model that is able to interpret different formats for date and time expressions. Some of them are based on the application of empirical methods using the focus theory proposed in (Grosz et al. 1986).

In this paper a proposal of a grammar for the recognition of temporal expressions in Spanish is presented, as well as an approximation to the resolution of the coreference introduced by them, as is explained in (Saquete and Martínez-Barco 2000). Moreover, in this paper a set of tags is used to annotate the temporal expressions in the corpus.

In a text, there are dates with typical representations like, for example: “23/01/2000” o “23 de enero del 2000” (*23rd of January of 2000*), but we can find references to dates named previously too, for example: “dos días antes” (*two days before*), “la semana anterior” (*the previous week*), etc. This kind of coreference should be solved and mapped to dates with a standard format for a more efficient analysis of the text. For that, we use a grammar for the recognition of temporal expressions with their correspondent temporal parser, and an inference engine to solve and to map these expressions in a standard format: mm/dd/yyyy (hh:mm, for time expressions). Once the kind of reference and the interpretation of the expression are solved, the text is tagged with XML tags.

2. System structure

The system proposed is shown in Figure 1. The system has the plain text as input. These texts are tagged with lexical and morphological information using the POS-tagger developed by Pla (Pla 2000) and this information is the input of the temporal parser. The temporal parser is implemented using an ascending technique and it is based in a temporal grammar shown below. Once the temporal expressions are recognized, these are introduced into the resolution unit called Temporal Expression Coreference Resolution, which will update the value of the reference according to the date that it is referring to, and then it will generate the XML tags for each expression.

There are two different kinds of rules that are used for the grammar because there are two different kinds of temporal expressions too:

1. There are anaphoric and not anaphoric expressions. That is why there are rules for the date and time recognition (non-anaphoric expressions like “12/06/1975”).

2. There are rules for the temporal reference recognition (anaphoric temporal expressions that need another complete temporal expression to be understood “*two days before*”).

Temporal references could be divided in two groups: time adverbs (i.e. *yesterday, today*) and nominal phrases that refer to temporal relationships (i.e. *two years before*).

Tables 1 and 2 show some rules used for the recognition of dates and the detection of anaphoric temporal expressions, respectively.

3. Coreference resolution based on a temporal model

Previous to coreference resolution, the parser identifies temporal expressions. Temporal expressions can be anaphoric or non-anaphoric. For this reason, we have split the rules for the identification or recognition of temporal expression in two different sets. The first set is made up by the rules for the identification of non-anaphoric temporal expressions (table 1) and the second one is made up by the rule to recognize the anaphoric temporal expressions (table 2). The coreference module should be applied to the temporal expressions recognized by the rules of the second set.

For the coreference resolution we use an inference engine that contains the interpretation for every reference named before. If we compare this system with traditional anaphoric systems, the algorithm for the treatment of temporal expressions needs to carry out an additional step. In our algorithm to solve the coreference of anaphoric temporal expressions, two different tasks can be distinguished:

1. Looking for the antecedent. This task is similar to the traditional approach to anaphora resolution. The algorithm chooses the antecedent from a list of candidates. Two main candidates are usually chosen. The first candidate is related to the date of the text, i.e. the date when the newspaper was written. This date is considered as *default date,* called in this paper *FechaP*. The second candidate is the previous non-anaphoric temporal expressions, called in this paper *FechaAnt* (*two days before* has as antecedent the previous date in the text). If none is found the default date is considered as antecedent. Sometimes, the temporal expression includes prepositional phrase with information about and event or process (*the day of the final match*), in this case the algorithm should look for the date associated to the event or process from the list of candidates. The following process is carried out:

   3. By default, the newspaper’s date is used as a base referent (temporal expression) if it exits. If not, the system date is used. (“ayer” *(yesterday)* Day(FechaP) –1 / Month(FechaP) / Year(FechaP).

   4. In case of finding a non-anaphoric temporal expression, it is stored as FechaAnt storing the old FechaAnt in a list of candidates. This value is updated every time that a non-anaphoric temporal expression appears in the text.

2. Providing the new date. Once the antecedent is selected, the new date should be inferred. This new step is related to provide a new date or time. The references are estimated using the antecedent selected in the previous step. This model is based on the two rules below and it is only applicable to these dates that are not *FechaP*, since for *FechaP* there is nothing to resolve

In Table 3 some of the entries of the dictionary used in the inference engine are shown. Moreover, the inference engine has the correspondence between numeric and string expressions of days and months, that is, *one* have the value *1* and *July* is *07*.

The module that makes the estimation of the dates will accede to the right entry in the inference engine in each case and it will apply the function specified obtaining a date in the format *mm/dd/yyyy or* a range of dates. So, at that point the coreference will have been resolved.

date	dd + ” /” + mm + “/” + (yy)yy (12/06/1975) *(06/12/1975)*
date	dd + “-“ + month + “ -“+ (yy)yy (12-junio-1975) *(12th-June-1975)*
date	dd + “de” + mm + “de” + (yy)yy (12 de junio de 1975) *(12th of June of 1975)*
date	(“El”) + day_of_week+ dd + ”de”+ month + “de” + (yy)yy (El domingo 12 de junio de 1975) *(Sunday, 12th of June of 1975)*
date	month+ ”de”+ yy(yy) (Febrero de 1975) *(February of 1975)*
date	dd + ”de”+ month + “de” + (yy)yy + “a las”+ time (12 de junio de 1975 a las 6 y media) *(12th of June of 1975 at half past six)*
dd	[“01”|”02”|”03”|….|”31”]
day	[“uno”| ”dos” |… |”treinta y uno”] *[one|two|…|thirty one]*
mm	[“01”|”02”|”03”|…|”12”]
month	[“enero”| “febrero”| “marzo” | “abril” | “mayo” | “junio” | “julio” | “agosto” | “septiembre” | “octubre” | “noviembre” | “diciembre”] *(January|February|March|April|May|June|July|August|September|October|November|December)*
a	[“1”| “2” | “3”|…”9”| “0”]
day_of_week	[“lunes”| “martes”| “miércoles” | “jueves” | “viernes” | “sábado” | “domingo”] *(Monday|Tuesday|Wednesday|Thursday|Friday|Saturday|Sunday)*
time	[hh:mm(:ss) | hh (y|menos) cuarto | hh y media |...]

Table 1. Sample of rules for the date recognition

Time Adverbs	reference “ayer” *(yesterday)*
	reference “mañana” *(tomorrow)*
	reference “anteayer” *(the day before yesterday)*
	reference “anoche” *(last night)*
Temporal Nominal Phrases	reference ”el”+ “próximo” + [“día” | “mes” | “año”] *(the next day|month|year)*
	reference “un” + [“día” | “mes” | “año”] + “después” *(a day|month|year later)*
	reference num + [“días” | “meses” | “años”] + “después” *(num days|months|years later)*
	reference “un” + [“día” | “mes” | “año”] + “antes” *(a day|month|year before)*
	reference num + [“días” | “meses” | “años”] + “antes” *(num days|months|years before)*
	reference “dentro” + “de” + “un” + [“día” | “mes” | año”] *(within a day|month|year)*
	reference “dentro” + “de” + num +[“días”|“meses”| “años”] *(within num days|months|years)*
	reference “el” + “pasado” + [“día” | “mes” | “año”] *(the last day|month|year)*
	reference “el” + [“día” | “mes” | “año”] + “siguiente” *(the next day|month|year)*
	reference “los” + num + [“días” | “meses” | “años”] + “siguientes” *(the num next days|months|years)*
	reference “el” + [“día” | “mes” | “año”] + “pasado” *(the last day|month|year)*
	reference “los” + num + [“días” | “meses” | “años”] +“pasados” *(the last num days|months|years)*
	num [“dos” | “tres” | “cuatro” | “cinco” | ….] *(two|three|four|five|…)*

Table 2. Sample of rules for the reference recognition

REFERENCE	DICCIONARY ENTRY
"ayer" *(yesterday)*	Day(FechaP) -1 / Month(FechaP) / Year(FechaP)
"mañana" *(tomorrow)*	Day(FechaP) +1 / Month(FechaP) / Year(FechaP)
"anteayer" *(the day before yesterdary)*	Day(FechaP) -2 / Month(FechaP) / Year(FechaP)
"anoche" *(last night)*	Day(FechaP) -1 / Month(FechaP) / Year(FechaP) [09:00-05:00]
"el"+ "próximo"+"día" *(the next day)*	Day(FechaP)+1 / Month(FechaP) / Year(FechaP)
"un"+"mes"+"después" *(a month later)*	[DayI/Month(fechaAnterior)+1/Year(fechaAnterior)-- DayF/Month(fechaAnterior) +1/ Year(fechaAnterior)]
num+"años"+ "después" *(num years later)*	[01/01/ Year(fechaAnterior) + num -- 31/12/ Year(fechaAnterior) + num]
"un" + "día" + "antes" *(a day before)*	Day(fechaAnterior)-1/Month(fechaAnterior)/Year(fechaAnterior)
num+"meses"+ "antes" *(num months before)*	[DayI/Month(fechaAnterior) -num / Year(fechaAnterior) - DayF/ Month(fechaAnterior) - num / Year(fechaAnterior)]
"dentro"+"de"+"un"+ "año" *(within a year)*	[01/01/ Year(fechaAnterior) +1 - 31/12/ Year(fechaAnterior) +1]
"dentro"+"de"+num+ "días" *(within num days)*	Day(fechaAnterior)+num / Month(fechaAnterior) / Year(fechaAnterior)
"el" + "pasado" + "día" *(the last day)*	Day(fechaAnterior)-1/Month(fechaAnterior) / Year(fechaAnterior)
"el"+"mes"+"siguiente" *(the next month)*	[DayI / Month(fechaAnterior) +1 / Year(fechaAnterior) -- DayF / Month(fechaAnterior) +1 / Year(fechaAnterior)]
"los"+num+"años"+ "siguientes" *(the num years later)*	[01/01/Year(fechaAnterior) -- 31/12 / Year(fechaAnterior) +num]
"el" + "día" + "pasado" *(the last day)*	Day(fechaAnterior)-1/Month(fechaAnterior) / Year(fechaAnterior)
"los"+num+"meses"+ "pasados" *(the num last months)*	[DayI/Month(fechaAnterior) - num / Year(fechaAnterior) - DayF/Month(fechaAnterior) - 1 / Year(fechaAnterior)]

Table 3. Sample of some of the entries of the dictionary

4. Tagging of temporal expressions

Several proposals for the annotation of temporal expressions have been arisen in the last few years (Wilson et *al.* 2001) (Katz and Arosio 2001) since this kind of research has started. In this section, we proposed doing this annotation using XML tags, in order to standardize anaphoric and non-anaphoric temporal expressions.

4.1. XML

In our proposal we have chosen XML to define the set of tags we are going to use. XML stands for *eXtensible Markup Language* and it provides a subset of the SGML (*Standard Generalized Markup Language*). XML offers a non-ambiguous text-based method to develop data structures. XML documents represent data by means of tags.

XML was developed by a Generic SGML Editorial Review Board formed under the auspices of the W3 Consortium in 1996 and chaired by Jon Bosak of Sun Microsystems, with the participation of a Generic SGML Working Group also organized by the W3C (W3C 2002). Since then, the use of XML has been generalizing until becoming the universal standard for data electronic exchange.

XML has specific rules that must be strictly followed in order to make a new document. The XML document must fulfill the set of constraints being established in the *Document Type Declaration* (DTD). This DTD contains the structure of the document, and the validity of a XML document could be tested through this DTD: a well-formed document is valid only if it contains a proper document type declaration and if the document obeys the constraints of that declaration (element sequence and nesting is valid, required attributes are provided, attribute values are of the correct type, etc.).

As a result, XML provides us several advantages to our proposal:

- Both persons and machines easily interpret it. As a consequence that makes easy both the manual and the automatic extraction of dates from a text.
- XML documents are easily tagged from an automatic process, but also a manual annotator could make use of commercial XML editors to develop this task.
- XML is standard.
- A DTD has been built to check the validity of each XML document. So manual and automatic annotations can be automatically tested looking for possible mistakes.

4.2. Annotation schema

An appropriate annotation schema has been defined to mark every temporal expression found. This schema is base on the following ideas:

First, every temporal expression is going to be marked. That includes the markup of dates and times that could be expressed in whatever format, including anaphoric and not anaphoric expression. In this way, the following rules are going to be applied:

a) Full date expressions, that is, non-anaphoric temporal expressions are going to be marked using the standard format for dates: mm/dd/yyyy.

10 de abril de 2002 04/01/2002
(*10th of April of 2002*)

b) Full date and time expressions (again, non-anaphoric expressions) are going to be marked in the same way, including in this case the time parameter, in the standard 24-hour format: hh:mm

10 de abril de 2002, a las nueve
04/10/2002, 09:00
(*10th of April of 2002, at nine o'clock*)

c) Time expressions (without explicit date, but referring to an omitted date) are anaphoric expressions. Then, the coreference resolution module is applied before tagging. Once the date of the time is calculated, date and time are tagged.

A las nueve 04/10/2002, 09:00
(*At nine o'clock*)

d) Some kind of time expressions without explicit date are not anaphoric expressions because they do not refer to an omitted date. In this case only the time parameter is need, so the coreference resolution module is not used.

Todos los días a las nueve 09:00
(*Everyday at nine o'clock*)

e) Anaphoric date expressions need the coreference resolution module to define the absolute expressions to which they refer to. After that, the appropriate tag will be marked.

El próximo miércoles 04/17/2002
(*Next Wednesday*)

f) Anaphoric date and time expressions follow the previous rule, calling the coreference resolution previous to be marked.

El próximo miércoles, a las nueve
04/17/2002, 9:00
(*Next Wednesday, at nine o'clock*)

g) Non-anaphoric ranges of dates and/or time are directly tagged by means of the initial and the final date.

Del 10 de abril al 20 de abril de 2002
04/10/2002 — 04/20/2002
(*From 10 of April to 20 of April of 2002*)[1]

El 10 de abril, de 9 a 11 y media
04/10/2002, 9:00 — 11:30
(*10th of April from 9 to half past 11*)

Del 10 de abril de 2002 a las nueve al 20 de abril de 2002 a las doce
04/10/2002, 9:00 — 04/20/2002, 12:00
(*From 10 of April of 2002 at nine o'clock to 20 of April of 2002 at twelve o'clock)*

h) Anaphoric ranges of dates and/or time are previously solved using the coreference resolution module. Then the full date is marked.

[1] Direct translation from Spanish

i) What we have called *fuzzy temporal expressions* have not a concrete date or time related to. For this reason the coreference resolution module is useless. However, in order to be identified as temporal expressions we decided to mark them using a special parameter with the “FUZZY” value.

The grammar used to identify anaphoric and not anaphoric expressions acts as a trigger launching the appropriate rule in each case.

4.3. Tag definitions

The structure of tags used to define temporal expression data is the following:

```
< DATE_TIME TYPE=”value”
      VALDATE1=”value”
      VALTIME1=”value”
      VALDATE2=”value”
      VALTIME2=”value” >
      Expression
</DATE_TIME>
```

In this structure the next elements are used:

- DATE_TIME is the name of the tag for non-anaphoric temporal expressions.
- VALDATE# store the range of dates obtained from the inference engine.
- VALTIME# store the range of times obtained from the inference engine.
- TYPE attribute could have the following values: CONCRETE, PERIOD and FUZZY:
  - § CONCRETE is referring to only a date.
  - § PERIOD is referring to a period of time.
  - § FUZZY attribute is used when we really do not know the date or period of time when a temporal expression is referring to.

Moreover, VALDATE1, VALDATE2, VALTIME1 and VALTIME2 are optional attributes:

- VALDATE2 and VALTIME2 are used to establish ranges. So, if we try to tag a concrete date (TYPE adopt the value CONCRETE) then these attributes are omitted.
- VALTIME1 could be omitted if only a date is specified.
- VALDATE1 could be omitted if only a time must be specified. This is the case in which the date does not mind. For example, *todos los días a las nueve* (*everyday at nine o’clock*). However, when only a time expression is specified, such as *a las nueve* (*at nine o’clock*), the VALDATE1 of this time must be computed.

The use of XML allows us to take advantage of the XML schema in which the tag language is defined. This schema let an application know if the XML file is valid and well-formed. The schema defines the different kind of elements, attributes and entities that are allowed, and can express some limitations to combine them. Moreover, use the same syntax as XML and the schemas are extensible. Once the XML file has been generated, a parser of our XML needs to be defined to make the information useful.

Tables 4a and 4b show several examples for tagging temporal expressions (non-anaphoric and anaphoric). Table 5 shows an example of an annotated text in which the features of the used tags are shown. In this example we assume that the newspaper’s date is 04/25/2000. The system, for the reference “el próximo año”*(the next year),* will return “01/01/2001-12/31/2001”. For the reference “mañana” *(tomorrow)* it will return 04/26/2000.

ABSOLUTE TAGS (non-anaphoric temporal expressions)
<**DATE_TIME VALDATE1**=”06/12/2001”>12 de junio de 2001</**DATE_TIME**>
<**DATE_TIME VALDATE1**=”06/12/2001” **VALTIME1**=”20:30”>12 de junio de 2001 a las ocho y media de la tarde</**DATE_TIME**>
<**DATE_TIME VALDATE1=**”06/12/2001” **VALTIME1**=”20:00” **VALDATE2**=”06/12/2001” **VALTIME2**=”21:00” >12 de junio de 2001 entre las ocho y las nueve de la tarde</**DATE_TIME**>

Table 4a. Sample of the tags generated by this system

REFERENCE TAGS (anaphoric temporal expressions)
<**DATE_TIME_REF VALDATE1**="06/11/2002">ayer</**DATE_TIME_REF**>
<**DATE_TIME_REF VALDATE1**="01/01/2002" **VALDATE2**="12/31/2007">los 5 años siguientes</**DATE_TIME_REF**>

Table 4b. Sample of the tags generated by this system

```
"La oficina de Congresos de la Universidad ha propuesto 5 congresos
para <DATE_TIME_REF TYPE= "PERIOD" VALDATE1="01/01/2000" VALDATE2=
"12/31/2000> este año</DATE_TIME_REF>, sin embargo, el crecimiento para
<DATE_TIME_REF  TYPE="PERIOD" VALDATE1= "01/01/2001"
VALDATE2="12/31/2001> el próximo año</DATE_TIME_REF> será superior a los
15. Por otro lado, el Director de la oficina ofrece <DATE_TIME_REF TYPE=
"CONCRETE" VALDATE1= "04/26/2000" > mañana</DATE_TIME_REF>  una
conferencia."

( The University Conference Office has proposed 5 conferences for
<DATE_TIME_REF TYPE="PERIOD" VALDATE1="01/01/2000" VALDATE2="12/31/2000>
this year </DATE_TIME_REF>, however, the increase for <DATE_TIME_REF
TYPE="PERIOD" VALDATE1= "01/01/2001" VALDATE2="12/31/2001> the next year
</DATE_TIME_REF>  will be over 15. On the other hand, the Office Manager
offers <DATE_TIME_REF TYPE= "CONCRETE" VALDATE1= " 04/26/2000" > tomorrow
</DATE_TIME_REF>  a lecture ).
```

Table 5. Example of annotated text

5. System evaluation

For implementing the system we need two different units, it is necessary to implement a parser and for that we have used LPA Prolog, because this language is based on rules as the parser is. The other unit is implemented in Visual Basic because this language has several time functions. This unit generated the XML tags too. The implementation of a XML parser is an optional possibility. The evaluation of the system has been done with a sample extracted from 16 articles that belong to the digital edition on the Internet of two Spanish newspapers describing different topics. The results obtained for the articles showed a precision and a recall of 95.59 % and 82.28% respectively.

The total has been calculated according to the number of successes being 195, the number of treated references is 204 and the number of total references is 237.

5.1. Error analysis

However, some fails in the system have been detected and we show their possible improvements below:

- The unit that resolves temporal references is not able to resolve undetermined temporal references like "hace unos cuantos días" (*some days before*) accurately. Here, one possible solution is the use of the semantic information. For example, if the sentence is "*some days before*", the system will suppose that is less time than a week, because we usually use the word "*week*" referring to seven days.

- It is possible that we have non-anaphoric expressions that make reference to an event or a fact and, despite they are not temporal expressions themselves, they mean a date or period of time too. For example: "ganó el mundial y al día siguiente se lesionó" (*he won the World Champion and the next day he hurt himself*).

- In newspaper articles, sometimes we find expressions like "el sábado hubo un accidente" (*Saturday there was an accident*). To resolve this expression we should know some extra information of the context where the reference is. This extra information could be the sentence verb. If it is a past verb that means that the sentence is referring to *the last Saturday*. However, if the verb is future, it is referring to *the next Saturday*. In our system, we are not using this kind of information, so we assume that this kind of reference if referring to the last day, not the next, because the news usually tells us facts occurred previously.

6. Conclusions

In this paper a system for temporal expressions recognition in Spanish and their reference resolution has been presented, based on a temporal model proposed. The system has two different units: the parser based on a temporal expression grammar, which allows to identify these kind of expressions and a coreference resolution unit which is based in a inference engine and make a transformation of the expressions to dates, resolving their reference in this way. The evaluation of the system shows successful results of precision and recall for our proposal.

For future works, it is pretended to extend the system with the temporal references that are not treated in this paper. Moreover, the study of the verbal forms in the sentences where the references are found will improve the efficiency of the system solving some kind of expressions.

7. Acknowledgements

This paper has been supported by the Spanish Government (MCYT) under grant TIC2000-0664-C02-01/02.

8. References

Katz, G. and Arosio, F. (2001). The Annotation of Temporal Information in Natural Language Sentences. In *Proceedings of the Workshop On Temporal And Spatial Information Processing (ACL'2001).*

Pla, F. (2000). Etiquetado Léxico y Análisis Sintáctico Superficial basado en Modelos Estadísticos. Ph D. Thesis. Departamento de Sistemas Informáticos y Computación. Universidad de Politécnica de Valencia.

Saquete, E. & Martinez-Barco, P (2000). Grammar specification for the recognition of temporal expressions. In *Proceedings of the*

Wilson, G., Sundheim, B., & Ferro, L. (2001). A Multilingual Approach to Annotating and Extracting Temporal Information. In *Proceedings of the Workshop On Temporal And Spatial Information Processing (ACL'2001).*

W3C (2002). Extensive Markup Language web page. W3C World Wide Web Consortium. http:// www.w3.org/XML.

On the Importance of Annotating Event-Event Temporal Relations in Text

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

Many natural language processing applications, such as information extraction, question answering, topic detection and tracking, and multi-document summarisation, would benefit significantly from the ability to accurately position reported events in time, either relatively with respect to other events or absolutely with respect to calendrical time. However, only recently has concerted work started on the automatic extraction of temporal information from text. The overall aim of our work is to automatically establish the temporal relations holding between events as well as between events and calendrical times in newspaper articles. This information makes it possible to create a 'time-event graph' to represent the temporal information contained in a text, and would in turn support the applications mentioned above. In this paper we first argue for the superiority of the time-event graph over a time-stamped event sequence as a target representation for extracted temporal information and discuss the importance of annotating temporal relations. We then give a brief account of the annotation scheme we have devised which allows us to annotate relational information as well as temporal referring expressions. We also discuss a pilot study in which we assessed the utility and feasibility of the scheme and the annotation tool we have developed to aid the annotation process. Finally, we discuss potential improvements in the annotation tool which are aimed at making the annotation of larger scale corpora possible.

1. Introduction

Many natural language processing applications, such as information extraction, question answering, topic detection and tracking, and multi-document summarisation, would benefit significantly from the ability to accurately position reported events in time, either relatively with respect to other events or absolutely with respect to calendrical time. However, only recently has concerted work been started on the automatic extraction of temporal information from text.

In addressing the goal of extracting temporal information from text, it is necessary to:

1. specify the target temporal representation which we wish to obtain for a text;
2. identify ancillary information which we **may** want to extract because of its utility in arriving at the target temporal representation (by analogy with, e.g. part-of-speech tagging or parsing as intermediate goals towards semantic interpretation).

For example, one candidate for target representation is an association of a calendrical time point or interval with each event in a text, i.e. a list of pairs of calendrical times and events. Arriving at this representation might require extracting additional information, such as temporal relational information, about events. For example, assigning "before 1984" to an event A might only be possible by recognising that event B occurs in 1984 and that A occurs before B. Thus, the capability to determine temporal relations between events might be a useful component capability in a temporal information extraction system, even if the information identified by such a component is not directly included in the target representation.

Our view is that target representation should be a **time-event graph** where the nodes in the graph are either times or events and the arcs are temporal relations. This is somewhat different from the "time-stamping" representation introduced in the preceding paragraph and one of our major goals in this paper is to argue that it is a superior representation.

With respect to ancillary temporal information to be extracted, our view is that time-referring expressions, event representatives, and temporal relations as signalled by, e.g. prepositions and temporal adverbials, all convey important temporal information and should be extracted. This information is necessary to derive a time-event graph for a text; but of course it is useful for creating a time-stamp representation as well – arguably both necessary and sufficient.

In this paper, we first give an overview over existing approaches to temporal annotation and information extraction in Section 2. Then in Section 3. we discuss the importance of a target representation that captures temporal relations and describe the annotation scheme we have developed to do so. Section 4. presents some results of a pilot study we have conducted based on the scheme. Further improvements to the process of annotation, to support the creation of larger annotated resources, are discussed in section 5.

2. Overview of Existing Approaches

Existing approaches to capturing temporal information in text can be divided broadly into the following three groups: (1) approaches that concentrate on an accurate and detailed annotation of temporal referring expressions, (2) time-stamping approaches that aim to associate a calendrical time with some or all events in the text, and (3) approaches that focus on the temporal relations between events and times, between events and events or both. We give a brief overview of existing work on each approach in this section.

2.1. Annotating Temporal Referring Expressions

The most extensive work on annotating temporal referring expressions so far has been done as part of the MUC language technology evaluations or the subsequent TIDES[1] and ACE [2] programmes.

2.1.1. MUC Named Entity Task

Between 1987 and 1998 the DARPA-sponsored Message Understanding Conferences (MUCs) developed a quantitative evaluation regime for message understanding (MU) systems, now generally called information extraction (IE) systems. The last MUC, MUC-7, was held in 1998, but related work continues within the ACE workshops. For more information about the message understanding conferences see MUC (1998).

While MUC evaluations typically defined several evaluation tasks, the relevant task here is the *named entity (NE) recognition* task, introduced in MUC-5. The NE task required the recognition and classification of specified named entities such as persons, locations, organisations, monetary amounts and, most importantly in the current context, time expressions (timex). The aim of the timex task was to mark up time expressions in text using SGML tags and to classify these expressions using a TYPE attribute. Type DATE referred to complete or partial date expressions and type TIME referred to complete or partial expressions of time of day. Both absolute and relative time expressions had to be marked up, although these two types were not distinguished in the annotation.

In the MUC-7 evaluation, the best systems were able to obtain F-measure scores approaching 94% on this task.

2.1.2. An Annotation Scheme for Temporal Expressions

Wilson et al. (2001) describe a set of guidelines[3] being developed within the TIDES programme for annotating time expressions and associating with them a canonical representation of the times to which they refer. A method for extracting such time expressions from multiple languages is also introduced. The main novel features as compared to the MUC temporal annotation task are:

1. In MUC the task called merely for surface time expressions to be annotated and crudely classified, whereas the Wilson et al. (2001) guidelines also call for each expression to be *evaluated*, i.e., to have associated with it a normalised representation of the time referred to.

2. The range of expressions flagged is much wider.

3. Context-dependent time expressions like *today* are handled in addition to fully specified time expressions like *September 3rd, 1997*. Context can be local (within the same sentence) or global (outside the sentence). Indexical time expressions, that require knowledge about the time of speech, like *now* are also included. A corpus study (Wilson and Mani, 2000) showed that two-thirds of time expressions in print and broadcast news are context dependent, so this feature is significant.

Wilson et al. (2001) have developed a tagger to do time expression tagging as described in the TIDES guidelines, and report F-measure scores of 96.2% on expression identification and 83.2% on evaluating these expressions.

2.2. Time-Stamping of Events

Annotating temporal referring expressions is only a first step towards extracting rich temporal information from text. The approaches introduced in this section aim at 'stamping' some or all events in a text with a calendar time – possibly the time value of an associated temporal referring expression.

2.2.1. MUC-5 and MUC-7 Time Slots

In addition to the Named Entity time expression tagging task, MUC-5 and MUC-7 also required relations between times and events to be established as part of the scenario template task. Participants were required to assign a calendrical time to certain specified event types (joint venture announcements and rocket launchings, respectively).

Scenario template filling requires the identification of specific relations holding between template elements. For example, the MUC-7 scenario template filling task concerned rocket launch events. The scenario template contains information about vehicles, pay load, launch site, mission function etc. It also contained a slot called LAUNCH_DATE, which was to be filled with a link to a time entity which in turn contained slots for a normalised representation of the start and end times of the temporal interval containing the launch event, if the interval could be determined from the text.

Temporal relations between events and other events were not explicitly addressed, though insofar as they were necessary to infer correct slot fills, systems needed to take them into account. Scores were quite low on this slot reflecting the difficulty of correctly assigning to it.

2.2.2. Assigning Time-Stamps to Event Clauses

In the MUC task, times were only to be determined for the events of interest, the scenario events. A more ambitious goal is to attempt to associate calendrical times or time intervals with *every* event in a text.

Filatova and Hovy (2001) describe a method for breaking news stories into their constituent events and assigning time-stamps to them. The time-stamps assigned are either full specified calendrical dates, sets of dates, closed date ranges (both end points specified), or date ranges open at one end or the other, indicating some time before or after the specified date.

The syntactic units conveying events are assumed to be simple clauses and they are identified using a parser which produces semantically labelled syntactic parse trees. Some problems are ignored in this approach, for example multiple verbs with different tenses in one sentence.

The time-stamper uses two time-points for anchoring. One time-point is the time of the article (at the moment only the date is used and the time of day is not taken into

[1] See http://www.darpa.mil/ipto/research/tides/.

[2] See http://www.itl.nist.gov/iaui/894.01/tests/ace/.

[3] The full set of guidelines are available as Ferro et al. (2000)

account) and the other time-point is the last time-point assigned within the same sentence. The procedure of time-stamping is as follows:

1. The text is divided into event clauses
2. All date phrases in the text are extracted
3. A date is assigned to each event clause based on either
   (a) the most recent date phrase in the same sentence, or
   (b) if this is not defined, then the date of the article.

In assigning dates various time assignment rules are used. When a date phrase is present in the sentence these rules both take into account nearby prepositions, such *on*, *after, before*, and carry out fuller specification. For example if the date phrase is simply a day of the week, then the article date is also used to derive a date-stamp that is fully specified with respect to year and position within the year. If no date phrase is present in the sentence then tense information is used to assign a time interval relative to the date of the article.

After all events have been stamped with a time, the event clauses are arranged in chronological order. The authors report scores of 77.85% correct time-stamp assignment to event clauses which have been manually (i.e. correctly) extracted from sample texts of a small trial corpus.

2.2.3. Temporal Semantic Tagging of Newswire Texts

The ultimate goal for Schilder and Habel (2001), as for ourselves, is to establish the temporal relations between all events in news articles.

In Schilder and Habel's approach temporal expressions are classified into *time-denoting expressions* that refer to a calendar or clock time and *event-denoting expressions* which refer to events. They view their goal as anchoring these temporal expression on the absolute time-line, so as to produce a linearly ordered anchored set of temporal entities; hence a time-stamp representation appears to be their target representation. For time-denoting expressions this may mean resolving indexicals (*now*, *yesterday*) or fleshing out expressions like *Thursday* to fully specified calendar dates. For event-denoting expressions a calendar time which is the time of the event must be associated with the event, possibly by extracting temporal relations which are signalled by prepositional phrases like *on Friday*. The set of temporal relations proposed is *before, after, incl, at, starts, finishes* and *excl* (equivalent to Allen (1983)'s relations).

They have developed a semantic tagging system for temporal expressions in newswire articles. The main part of their system is a Finite State Transducer (FST) based on handwritten rules. Their target language is German. The FST tags all time-denoting expressions, all verbs and an experimental version tags event-signalling nominal expressions. A semantic representation is then proposed, based on which inferences are drawn, especially about temporal relations. In its current state, the FST establishes temporal relations between times and events. The tagger was evaluated with respect to a small corpus (10 news articles) and an overall precision rate of 84.49% was achieved.

2.3. Annotating Temporal Relations

The work described in the preceding section aims at associating a calendar time with some or all events reported in a text, but none of these approaches view the identification of temporal relations as a explicit goal in its own right. These temporal relations are clearly of importance, even for time-stamping approaches. The work described in this section, as well as the approach we develop in the next section, address temporal relations directly.

2.3.1. Annotation of Intrasentential Temporal Information

Katz and Arosio (2001) aim to create a large multilingual corpus, in which intrasentential temporal relations are tagged in addition to standard morphological and syntactic features. To aid this, they have developed a language-neutral and theory-neutral method for annotating sentence internal temporal relations. With this corpus, Katz and Arosio (2001) hope to be able to automatically acquire the lexical knowledge required for determining temporal interpretation in narrative discourse.

A temporal interval is associated with each verb in the sentence; it is the temporal relations between these intervals that are of concern. The temporal interpretation should be closely linked to the syntactic context (which is of importance since it is not known beforehand to what degree the cues used by the speaker are lexical and to what degree they are grammatical). This linking is needed to keep track of both the semantic relations among times as well as the syntactic relations among the words in the sentences that refer to these times.

The authors add a layer of semantic annotation to already syntactically annotated text. The verbs in the sentence are linked via secondary edges labelled with a temporal relation. Precedence and inclusion and their duals are the possible relations. Indexical information is included by introducing the symbol $\circ$ for the speech time, which is automatically prefaced to all sentences prior to annotation.

A searchable multi-language annotated treebank has been created where each sentence is stored in a relational database with both syntactic and temporal annotations. This makes is possible to query the corpus ("Find the sentences containing a relative clause which is interpreted as temporally overlapping the main clause" (Katz and Arosio, 2001)).

This work is valuable, especially for linguists interested in the studying, cross-lingually, the complex interrelationship of lexical and syntactic mechanisms used to convey temporal relations between events in the same sentence. However, if one's goal is extraction of the full temporal content of a text, it is limited in only considering intrasentential temporal relations.

3. Annotating Temporal Information in Text

From the preceding overview of existing work on temporal information extraction it is clear that the bulk of work so far has gone into the identification of temporal referring expressions and the assignment of time-stamps to events. Only Katz and Arosio (2001) focus directly on the problem of identifying temporal relations between events, and in their case only between events in the same sentence.

In this section we start by arguing that a time-event graph, in which not all events are necessarily directly anchored on a time-line, is a superior target representation for a text to a time stamped representation. We then present the conceptual underpinning for the approach we advocate for annotating temporal information in text, followed by the details of the annotation scheme itself.

3.1. Why Annotate Temporal Relations?

Recall that a time-event graph is a graph in which the nodes are either times or events and the arcs are temporal relations. There are two principal arguments for preferring a time-graph representation to a time-stamp representation.

First, in many cases texts position events in time only by relation to other events and any attempt to coerce these events onto a time-line must either lose information, invent information, or rely on a notion of an underspecified time point constrained by temporal relations (i.e. introduce a representation of temporal relations by the back door).

Consider this example:

> After the plane crashed, a search was started. Afterwards the coast guard reported finding debris.

and assume that an earlier sentence specifies the calendrical time of the plane crash.

An approach attempting to map the information presented in this example onto a time-line is faced with the situation depicted in Figure 1.

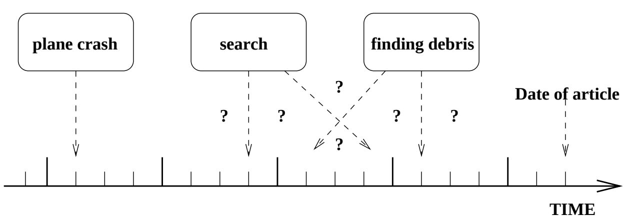


Figure 1: A Time-line Representation

While the crash event can be placed on the time-line the other two events cannot. Either time points must be guessed, or an interval be assigned. The first option is clearly not satisfactory. But if an interval is assigned the only possible interval, for both the searching and finding events is the interval from the crash till the date of the article. But if this is assigned to both events then the information about their ordering with respect to each other is lost.

A simpler representation which while not attempting to be as specific actually carries more information is shown in Figure 2.

This representation preserves the information that the searching event precedes the finding event, without forcing any early commitment to points on a time-line.

The second argument for preferring a time-event graph representation that captures event-event temporal relations as well as time-event relations is that to position events on a time-line accurately requires the extraction of event-event relational information. In the example, the placing of the searching and finding events in the interval between the plane crash and the date of the article requires the recognition that these events occurred after the crash as signalled by the words "after" and "afterwards". Without identifying the relations conveyed by these words the searching and finding events could only be positioned before the time of the article, and not following the plane crash. Thus, even if a time-stamp representation is viewed as the best target representation, achieving it requires the extraction of temporal relational information. In this case adopting a time-event graph as an intermediate representation is still a good idea, which begs the question of why it should not simply be taken as the final target representation.

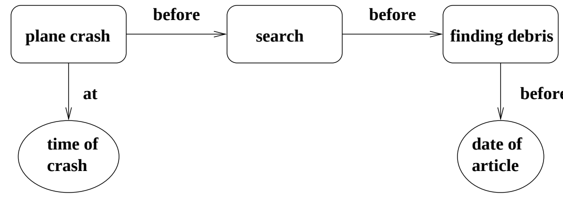


Figure 2: A Time-Event Graph Representation

3.2. Conceptualising Time

Before we describe the annotation scheme we have developed, we will very briefly explain what kind of temporal entities and relations we suppose exist. We presume the world contains the following primitives: events, states, times, temporal relations and event identity. Each primitive is described briefly below.

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 on a time-line and they are usually conveyed in language by finite verbs or by nominalisations. Examples of events are:

> *A small single-engine plane* ***crashed*** *into the Atlantic Ocean.*
>
> *The 1996* ***crash*** *of the TWA 747 remains unexplained.*

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, the focus of much of the philosophical debate 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, time objects which are larger or smaller than a day, respectively.

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 text of interest. 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.

Temporal Relations Events stand in certain temporal relations to other events and to times. Times are temporally related to other times as well, but this phenomenon is not only very rarely explicitly expressed in text, it is also of lesser importance and is not taken into account here.

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

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

The full set of temporal relations we suppose at present is { *included, includes, after, before, simultaneous* } . This is a minimal set, which was defined after analysing a number of newspaper articles, and can easily be expanded.

3.3. The Annotation Scheme

Given this conceptual framework, we can describe the annotation scheme we have defined. For more details see Setzer and Gaizauskas (2000).

Annotating Events Events are marked by annotating a representative in the clause conveying the event. The first choice for a representative is the head of the finite verb group. If a nominalisation conveys the event, then the head of the nominalisation serves as the representative. In the rare case of an event being conveyed by a non-finite clause, the non-finite verb is marked as the representative.

An events carries attributes for some or all of the following properties: unique event ID, event class, verb tense, verb aspect, other event to which it is related and temporal relation by which it is related, time object to which it is related and temporal relation by which it is related, the word(s) by which the temporal relation is signalled, and the ID of events it might have as an argument. For example, ignoring temporal relations for the moment:

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

Annotating Times We distinguish between *simple* and *complex* time referring expressions. Simple time referring expressions refer to times directly, as in example (1). Complex time referring expressions, as in (2), refer to a point in time by relating (*after*) an interval (*17 seconds*) to an event (*hearing the sound*). The point in time referred to is the point at the end of the interval.

(1) *last Thursday*

(2) *17 seconds after hearing the sound ...*

For simple time referring expressions we annotate the whole text span conveying the time-object:

```
<timex tid=5 type=DATE calDate=12041997>
  last Thursday </timex>
```

Each time referring expression has a unique ID, an attribute flagging whether it is a time or a date, and an attribute carrying the calendar date the expression refers to.

Complex time referring expressions, like the one in example (2), include a time interval (*17 seconds*), a preposition (*after*) and an event (*hearing the sound*) or time. The way these are annotated is similar to the way events are annotated. The interval is chosen as the representative for the time referring expression and related to the event expression via the temporal relation, usually signalled by the preposition.

```
<timex tid=5 type=complex eid=3 signalID=7
relType=after> 17 seconds </timex>
<signal sid=7> after <signal>
<event eid=3> hearing</event> the sound...
```

Annotating Temporal Relations Events and times can be related to other events or times. If two events are related then one of the events carries the ID of the other as well as the temporal relation in which they stand to each other. If an event is related to a time then the event carries the ID of the time object and the temporal relation. In either case, if the relation is signalled explicitly in the text, then the ID of this signal is an attribute as well, as the following two examples illustrate.

```
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>.
```

If the temporal relation is implicitly expressed, then the only difference is that the attribute for the signal is simply left out.

One problem with this annotation scheme is that it is not possible to relate one event to two or more other events or times, though by and large we have not found this to be a problem in annotating real text. This problem has been addressed by the TERQAS[4] workshop, which is working towards defining a general time markup language and has adopted many aspects of the current annotation scheme.

[4]See http://www.cs.brandeis.edu/ jamesp/arda/time/.

The solution proposed there is to introduce independent SGML `LINK` entities, which consume no text, to serve as relational objects tying events and times together. One event can then participate in as many links as is necessary.

4. The Pilot Study

To study the feasibility of the annotation scheme and to gain insight into the linguistic mechanisms conveying temporal information in text, we have applied the annotation scheme to a small trial corpus.

4.1. The Corpus

The trial corpus consists of 6 newswire articles taken from the New York Times, 1996, which were part of the MUC7 (MUC, 1998) training data. Basic statistics about the corpus are presented in table 1.

	sentences	words	number of annotators
text1	26	448	3
text2	18	333	2
text3	13	269	3
text4	13	213	2
text5	10	211	3
text6	13	399	3
total	93	1873	3

Table 1: The corpus

Each text was annotated by either two or three annotators, in addition to one of the authors, who produced what in the following is taken to be the 'gold standard' or 'key' annotation.

4.2. The Process of Annotation

The annotation takes place in two stages, both of which are described briefly in this section. To aid the annotator with her or his task, we have developed an annotation tool which not only allows the annotation of the information required by the scheme but which also interactively supports the annotator during the second phase, where additional temporal relations are established.

Stage I During Stage I, all event and time expressions are annotated as well as all signal expressions. Afterwards, those temporal relations that are explicitly expressed, e.g. by temporal prepositional phrases or subordinate clauses, and hold between events or events and times are established and stored as event attributes. Some implicitly expressed temporal relations are also established during this stage, for example, when events are clearly positioned in time but the signal expression has been omitted, as in *The army said Friday [...]*. In addition, *ing*-clauses without a subject can also be used to implicitly express a temporal relation between two events and are annotated during this stage.

Stage II The annotation scheme we have developed is aimed at establishing as many temporal relations in the text as possible. To relieve the burden on the annotator, and to increase the number of temporal relations annotated, we introduced stage II, which is cyclical in nature. Based on the information available, which in the beginning consists of the events, times and the temporal relations annotated in stage I, all inferences possible are drawn, according to an axiomatisation of the temporal relations *included, includes, after, before, simultaneous*. This is conducted automatically by the annotation tool which computes the deductive closure over these temporal relations. If the temporal relation between any pair of events or events and times is still unknown, the annotator is prompted for one of these [5] and, again, all possible inferences are automatically drawn. The process continues until every event-event and event-time pair in the text has been related.

4.3. The Results

In this section, we briefly describe the distribution of temporal phenomena over the trial corpus, as far as this is relevant to the issues discussed in this paper. Note that although this is a trial corpus, the results are indicative. We will not talk here about recall and precision values of the individual annotators with respect to the gold standard here – see section 5. For more information about the pilot study and its outcome see Setzer (2001).

Table 2 shows the number of event expressions, time expressions, and the number of event-event relations annotated in each text of the corpus in Stage I of the annotation process – i.e. these are the temporal relations that are explicitly expressed in the texts.

	# sentences	# words	# event expr.	# event-event relations	# event-time relations
text1	26	448	40	10	12
text2	18	333	30	10	5
text3	13	269	19	7	3
text4	13	213	27	5	0
text5	10	211	16	1	4
text6	13	399	26	13	5
total	93	1873	158	46	30

Table 2: Number of event expressions and explicit temporal relations per text

Table 3 shows for each text the number of event and time expressions in the text, the number of explicit temporal relations annotated in Stage I, the number of relations inferred from these without any further input from the annotator, the number of relations solicited from the annotator (i.e. the implicit temporal relations), and the number of inferred temporal relations overall.

4.4. Discussion

In Section 3.1. we criticised the time-stamped event sequence as a target representation on two grounds:

1. Forcing events to be placed on a time-line may result in the loss of event-event ordering information,

[5] Note that *unknown* is a possible value for a temporal relation here.

	event + time expr.	annotated ev.-ev. and ev.-time relations	inferred relations based on annotated relations	solicited rel.	total inferred relations
text1	32 + 11	10 + 12	222	124	1005
text2	26 + 5	10 + 5	122	93	380
text3	17 + 3	7 + 3	21	49	141
text4	18 + 0	5 + 0	8	45	120
text5	10 + 4	1 + 4	13	18	110
text6	24 + 5	13 + 5	107	52	514
total	127 + 28	46 + 29	493	381	2270

Table 3: Annotated, solicited, and inferred temporal relations

since the time-stamps assigned to distinct events may be identical even though we know the events occurred at separate times and know their order.

2. Event-event relational information must be extracted in order to position events on a time-line. Given this, why not choose a target representation that includes this richer information.

While both of these observations are true in general, ideally we would like to substantiate them empirically and quantitatively with respect to the trial corpus. Unfortunately we have not as yet been able to carry out the analysis for the whole corpus. However, we have chosen one text from the corpus (text6) and investigated it in detail.

To corroborate the first point above, we read text6 and, assuming perfect knowledge of the temporal information contained, then represented this information on a time-line, associating an interval with each event. In other words, without worrying about *how* the temporal information is extracted we time-stamped each event, where each time-stamp contains a start and end time expressed as calendar dates or, for at most one of the times, a symbol indicating the time is unknown.

For example, the sentences *A senior investigator looked at the wreckage Tuesday* and *Flight 800 exploded midair 20 days before Tuesday and then plunged into the ocean*[6] can be represented on a time-line as shown in Figure 3.

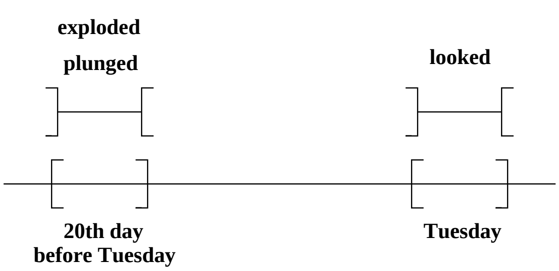


Figure 3: Example of a time-line Representation

Note that the events *exploded* and *plunged* have to be associated with the interval which encompasses the 20th day before Tuesday. We have lost the information that the plane plunged into the ocean **after** the explosion. This information can be easily represented in a time-event graph, as shown in figure 4.

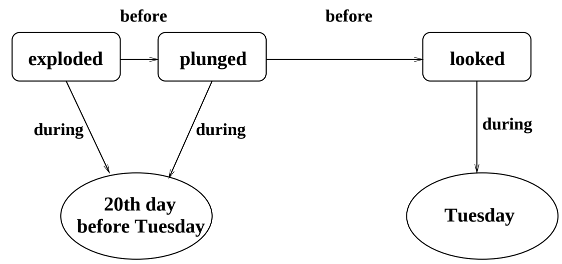


Figure 4: Example of a time-event graph Representation

Overall, 7 event-event relations that were explicitly mentioned in the text were lost in the time-line representation. While we have not performed the detailed analysis to let us say how many of the 514 inferred temporal relations in text6 are dependent on these 7 relations, it seems fair to assume that a significant number are.

To corroborate the second point we investigated how many of the 107 relations inferred for text6 from the explicitly annotated event-event and event-time relations resulted in new event-time relations involving events for which no event-time relation existed already. This corresponds to the intuitive notion of how many events are placeable on the time-line solely due to event-event relational information. For text6 we discovered that 20 of the 107 new relations were time-event relations for events for which no previous time-event relation existed. These 20 relations mentioned 4 distinct events (i.e. these 20 relations involved relating 4 events to different times, perhaps redundantly, but also potentially defining separate start and end points for intervals associated with them). Thus, 4 of the 24 events in text6 can be placed on a time-line using event-event relational information which is explicitly present in the text – positional information that otherwise would either be lost or require knowledge of implicit relations to extract.

Finally, we can make the general observation of the trial corpus that from 127 event-event relations plus 28 event-time relations, a total of 2270 additional temporal relations has been inferred. Even though we do not have the exact figure of how many of these inferred temporal relations are based on annotated event-event relations, it seems likely that the event-event relations contribute significantly to the number of relations inferred. We base this observation on the fact that there are nearly twice as many event-event relations as event-time relations annotated, and that subsequent inferences in the deductive closure calculation build on these initial relations. This observation adds weight to our claim that annotating event-event relations is important for temporal information extraction.

5. Improvements to the Annotation Process

The pilot study has shown that the interannotator agreement and the recall and precision figures need to be improved and that the burden on the annotator needs to be lessened, before the annotation scheme can be used to cre-

[6]The sentences have been slightly altered to make them more comprehensible out of context, but the temporal information they convey is the same as in the original text.

ate a larger corpus. Larger corpora will be necessary to train and evaluate temporal information extraction systems.

In Setzer and Gaizauskas (2001) we identified five main causes of low annotator precision and recall scores (with respect to the gold standard): imprecision/incompleteness of the guidelines; imperfect annotator understanding of the task; intrinsic difficulty of identifying the appropriate temporal relation in some cases; annotator fatigue; and annotator carelessness. In this section we do not address all of these problems, but focus on a number of proposals to enhance the annotation process, thereby lightening the load on annotators and increasing the accuracy of the annotations.

Pre-tagging An automatic first annotation pass could be used to reduce the amount of manual annotation and to raise recall. A part-of-speech tagger or word group parser, could be used to mark up finite verbs and signals and a time expression tagger such as Wilson et al. (2001)'s could be used to tag time referring expressions. Using the corpus as an indication, we know that a large percentage of the finite verbs will indicate events and the annotator can easily add attribute information to those or delete the mark up of mistakenly flagged verbs which do not indicate events. The high accuracy of time expression taggers mean that most of the work of tagging time expressions would be done automatically with the annotator left only to confirm details and scan for missed expressions.

Signals are a slightly different case. These are mostly prepositions and subordinating conjunctions, but a smaller number will have to be marked up. Here we have two options. We can mark up all prepositions and conjunctions and leave it to the annotator to delete inappropriate annotations, which is an easy process. Alternatively we could only automatically annotate those prepositions which are followed by a time referring expression. This approach carries the danger of not pre-annotating all signals, and the annotator, concentrating on the pre-annotated sections, might not catch all signals.

Intelligent Interaction with the Annotation Tool: Question Ordering The second phase of stage II of the annotation process, during which the deductive closure over the temporal relations is calculated and the annotator is prompted for unknown temporal relations, is problematic for the following reasons.

1. It is a long process, during which the annotator was prompted for 62 temporal relations per text on average, even for the short texts in the pilot corpus.

2. There is, for now, only marginal consistency checking and it is not possible to correct errors. Once the annotator notices that she or he made a mistake earlier in the process, then the whole stage II annotation process has to be restarted.

One possible solution for the first problem would be to optimise the order in which unknown temporal relations are prompted for. As we explained in section 4.2., after each temporal relation solicited from the user, all possible inferences are drawn. The larger the number of the inferences, the smaller the number of remaining unknown temporal relations will be. The following simple example illustrates the effect non-optimal soliciting can have. Imagine four events, forming a 'precedence chain':

$$e_1 < e_2 < e_3 < e_4$$

Imagine also that the link between e_2 and e_3 is missing in the response:

$$e_1 < e_2 \qquad e_3 < e_4$$

If the first question establishes the temporal relation holding between e_2 and e_3, then all other temporal relations can be inferred, based on the transitivity of before. The temporal model can be completed with one question. However, the order of questioning could be very different, establishing the temporal relations between e_1 and e_4, then between e_1 and e_3, e_2 and e_4 and then between e_2 and e_3. In this case four questions are asked to establish the relations holding between them.

Thus, question order can be important in determining how many questions the annotator ultimately gets asked. Clearly, one wants to minimise the number of questions asked, but it is not clear (to us) whether there is a question order that is guaranteed to minimise this number, and if so how to determine it. We propose to investigate initially a naive approach in which given two temporally-ordered event chains we first ask questions which attempt to link their end points, simply on the grounds that such questions could lead to maximal gains. However, considerably more empirical and theoretical investigation needs to be carried out here.

Intelligent Interaction with the Annotation Tool: Correcting Mistakes The second point requires a more elaborate solution. Once an incorrect temporal relation has been added an indeterminate number of further incorrect inferences may have been drawn on the basis of it. Two solutions suggest themselves:

1. Provide the possibility of check-pointing, i.e. saving intermediate stages to which the annotator can return when an error has been detected. This could be done automatically after each new user-solicited relation is added. This has the advantage of being easy to implement but the disadvantage of erasing possibly correct temporal relations added after the error, but independently of it, with the consequence that work that will have to be redone unnecessarily.

2. Implement a sort of truth maintenance system (Doyle, 1987; de Kleer, 1987), whereby only the incorrect temporal relation and those temporal relations which were inferred from it are deleted. This has the advantage of minimising the amount of work the annotator needs to redo unnecessarily, but the disadvantage of being more complex to implement.

Clearly the second solution is the better in the long run, as annotator effort is the chief quantity to conserve. We are working on solution whereby all temporal relations added to the temporal fact database record with them a justification which includes a reference to any facts from which they have been derived. Removing a temporal fact f then becomes a recursive procedure which begins with a search for all facts f' whose justification mentions f followed by a recursive call to delete f'. This will ensure that all dependents of f will be removed, while not touching any facts, solicited or derived, that may have been added after f in the annotation process, but which are logically independent of it.

6. Conclusion

We have argued that when extracting temporal information from texts a target representation, such as a time-event graph, which explicitly admits event-event temporal relations as well as time-event relations, is superior to one which does not, such as a time-stamped event sequence representation. In essence the arguments are that a time-stamp representation forces overspecification leading to information loss, and that event-event relations must be extracted even if a time-stamp representation is the target, and hence might as well be retained.

We also described the annotation scheme we have developed, which enables us to annotate temporal relations as well as events and time referring expressions, thus providing the necessary information to build time-event graphs for texts. A trial corpus which we constructed based on this scheme was described and used to corroborate the argument in support of the time-event graph approach.

One potential practical argument against the time-event graph approach is that building annotated resources capturing the required information is costly and error-prone. In the final section of the paper we introduced ideas for improving quality and reducing effort in the annotation process, improvements which we hope will make future larger scale application of the annotation scheme feasible.

7. References

J.F. Allen. 1983. Maintaining Knowledge About Temporal Intervals. *Communications of the ACM*, 26:832–843.

J. de Kleer. 1987. An Assumption-based TMS. In M.L. Ginsberg, editor, *Readings in Nonmonotonic Reasoning*, pages 280–298. Morgan Kaufman Publishers, Los Altos, California.

J. Doyle. 1987. A Truth Maintenance System. In M.L. Ginsberg, editor, *Readings in Nonmonotonis Reasoning*, pages 259–279. Morgan Kaufman Publishers, Los Altos, California.

L. Ferro, I. Mani, B. Sundheim, and G. Wilson. 2000. TIDES Temporal Annotation Guidelines. Technical Report MTR 00W0000094, The MITRE Corporation, October. Draft-Version 1.0.

E. Filatova and E. Hovy. 2001. Assigning Time-Stamps to Event-Clauses. In *Proceedings of ACL-EACL 2001, Workshop for Temporal and Spatial Information Processing*, pages 88–95, Toulouse.

G. Katz and F. Arosio. 2001. The Annotation of Temporal Information in Natural Language Sentences. In *Proceedings of ACL-EACL 2001, Workshop for Temporal and Spatial Information Processing*, pages 104–111, Toulouse. Association for Computational Linguistics.

1998. *Proceedings of the Seventh Message Understanding Conference (MUC-7)*. Morgan Kaufman. http://www.itl.nist.gov/iaui/894.02/related_projects/muc/.

F. Schilder and C. Habel. 2001. From Temporal Expressions to Temporal Information: Semantic Tagging of News Messages. In *Proceedings of ACL-EACL 2001, Workshop for Temporal and Spatial Information Processing*, pages 65–72, Toulouse.

A. Setzer and R. Gaizauskas. 2000. Annotating Events and Temporal Information in Newswire Texts. In *Proceedings of the Second International Conference on Language Resources and Evaluation (LREC2000)*, pages 1287–1293, Athens, Greece.

A. Setzer and R. Gaizauskas. 2001. A Pilot Study On Annotating Temporal Relations In Text. In *Proceedings of ACL-EACL 2001, Workshop for Temporal and Spatial Information Processing*, pages 73–80, Toulouse, France.

A. Setzer. 2001. *Temporal Information in Newswire Articles: An Annotation Scheme and Corpus Study*. Ph.D. thesis, University of Sheffield.

G. Wilson and I. Mani. 2000. Robust Temporal Processing of News. In *Proceedings of the 38th Meeting of the Association of Computational Linguistics (ACL 2000)*, pages 69–76, Hong Kong, October.

G. Wilson, I. Mani, B. Sundheim, and L. Ferro. 2001. A Multilingual Approach to Annotating and Extracting Temporal Information. In *Proceedings of ACL-EACL 2001, Workshop for Temporal and Spatial Information Processing*, pages 81–87, Toulouse.

Tense and Implict Role Reference

Joel Tetreault

University of Rochester
Department of Computer Science
Rochester, NY, 14627
tetreaul@cs.rochester.edu

Abstract

This paper presents a probe study into the use of temporal information in the resolution of implicit roles. It has been shown that temporal information can influence rhetorical relations between utterances which in turn can influence the resolution of referential entities. We describe a focusing-based algorithm for resolving implicit roles using such information and test the algorithm on an annotated corpus. Our results show that temporal information can be useful in resolving implicit roles in some cases, but a much larger and varied corpus is needed to strengthen this claim.

1. Introduction

This paper describes preliminary work relating tense to implicit role reference. Past work has shown that tense can influence the resolution of other reference types such as pronouns as well as discourse structure ((Webber, 1988a) and (Hwang and Schubert, 1992)). We extend this claim to the reference of implicit roles. Verbs have certain required roles, which refer to discourse entities, that are necessary for comprehending the verb phrase, and thus aid in natural language processing. Roles that do not have explicit antecedents in the verb phrase are deemed implicit. We annotated a small corpus for NPs and VPs and tense information and show, in some cases, that one can improve resolution rates of implicit roles by using simple heuristics incorporating tense with focusing. To our knowledge, this is the first time that an automated corpus study has been done analyzing the effects of temporal information in reference.

First, we describe implicit role reference in more detail and how temporal information can be used in resolving implicit roles. In section 3, we describe our annotation scheme, in section 4, our focusing-based algorithm for implicit roles and finally the results of our algorithm on an annotated corpus.

2. Implicit Role Reference

We claim that in addition to canonical reference types such as pronominal reference, VP ellipsis, and discourse deixis, verb phrases have certain required roles that can be viewed as anaphoric. These required roles refer to discourse entities and are necessary for the interpreter to understand the verb phrase, and thus the complete utterance. For example, in order to use the verb "take" one needs to understand that an entity is being moved, that it is being moved to one place from some other place, and that there is some entity that is responsible for moving it.

Implicit role reference has been briefly studied as a side effect of bridging and discourse relations ((Poesio, 1994) and (Asher and Lascarides, 1999)) but no major empirical work has been done in the area.

Resolution of implicit roles occurs frequently in naturally occurring dialog. Consider the following, modified from Asher and Lascarides (p. 90):

```
(1)  Take engine E1 from Avon to Dansville.
(2a) Pick up the box-
car and take it to Broxburn.
(2b) And then take the box-
car from Corning.
(2c) Also take the boxcar.
(3)  Leave E1 there but move the box-
car some more to Evansville.
```

For the sake of simplicity, assume that the verb "take" has these roles: "Theme": the entity being moved; "To-Loc": the location we are taking the "theme"; and "From-Loc" the location we are leaving.

In utterance (2a) one needs to know the At-Loc of the boxcar, in order to send it to Broxburn. This role is implicit and is resolved to Dansville. In order to resolve "there" in utterance (3) after utterance (2b) one must resolve the implicit "From-Loc" in "take" in the previous sentence. But the main point is that in a natural language understanding/planning system, one must keep track of entities and their locations in order for it to plan and carry out the task.

Asher and Lascarides point out that use of rhetorical roles, specifically whether utterances are in a narrative or parallel relationship, can aid in reference resolution. For example, the relationship between (1) and (2a) is a narrative while (1) and (2c) is parallel. While it is hard to annotate rhetorical relations we believe that one can approximate them by calculating the temporal relation between the two utterances. For instance, if we know that there is a narrative relation then we know that the entity that serves as the To-Loc role will probably serve as the From-Loc role in the next utterance, since entities move from the place they were just taken.

In our corpus we found the following distribution (see Figures 1 & 2) for the roles we focus on in this study (From-Loc and To-Loc) and their antecedents. These figures show that for a given role, how many sentences back (depth) its antecedent is found and in what role focus list it is located in. The trend is that antecedents for a From-Loc or To-Loc are predominantly found in the current utterance or the previous two utterances.

We describe our work in implicit roles in more detail in (Tetreault, 2002).

Depth	From-Loc	To-Loc
1	11	9
2	4	1
3	0	0
4	1	0
5+	0	0
%	61.5%	38.5%

Figure 1: From-Loc

Depth	Theme	From-Loc	To-Loc
1	0	1	1
2	0	0	2
3	0	0	0
4	1	0	1
5+	0	1	0
%	14.3%	28.6%	57.1%

Figure 2: To-Loc

3. Annotation

We use a subset of the TRAINS-93 Corpus (Heeman and Allen, 1994) annotated with coreference information for pronouns (Byron and Allen, 1998). The dialogs typically consist of short sentences, usually 10 words or less and are annotated using a sgml-style encoding. Our corpus consists of a 86-utterance dialog in which two human participants are given a task involving moving commodities and trains around a fictional world. We manually annotated each NP with an unique ID and its class (engine, tanker, location, food). Each VP was annotated with an ID, a time ID, and what NP ID(s) each role refers to. If a role is not mentioned explicitly in the text such as the "at-loc" role in (2a), then it is marked as implicit. The roles from each verb are taken from the TRIPS natural language system lexicon (Allen et al., 2000). For all the roles that are marked in this study (instrument, theme, from-loc, to-loc) roughly 30% are implicit.

A time point is associated with each verb event and constraints with previously mentioned time points are included in the time tag. The first element of each time tag is the time point associated with that event and is a string of a character followed by a number such as "t0." There are two types of constraint relations: either time x precedes a time y: "$x < y$" or x follows y: "$x > y$". Multiple constraints for a time point are encoded by linking the individual constraints with an ampersand: "$t1 > t2\&t1 < t0$" which says that t1 comes after t2 and t1 precedes t0. It should be noted that this is a very naive encoding scheme and that complex verb tenses are reduced to their root forms. A sample annotation (modified for readability) is shown in Figure 3.

Annotation of time points was difficult because the goal of each dialogs was to create a plan not to execute a plan in real-time. This means that the two speakers will often talk abstractly about parts of the plan and create hypothetical plans that may be abandoned if the speakers feel that they would not meet the constraints outlined by the experiment. Often utterances such as "We will need to move the boxcar to Avon by midnight" would appear and be followed by statements related to the introduced task. For our purposes, these multiple stand-alone plans complicate annotation because all time points in the discourse are not necessarily related. To deal with this, we give each sub-plan or hypothetical plan its own code, so one sub-plan may have its events labeled with "u": "$u0, u1, u2$..." while another distinct plan would have "v."

4. Algorithm

We have developed a preliminary model for resolving implicit roles that uses a combination of focusing and temporal reasoning. Our algorithm for resolving implicit roles in a discourse is as follows: first, as one progresses through the discourse, each utterance maintains a focus list for each role, such that when a NP is encountered, its discourse entity representation is placed at the top of the appropriate focus stack(s). When a verb is encountered, we check all of its roles and place explicit ones (those found in surface form of the sentence) on the top of the appropriate focus stack. If a role is implicit then it is resolved as determined by its type:

- Instrument: search through current utterance first for an entity that meets the verb's constraints. If one is not found, then search through each past utterance's focus stacks: looking at the instrument and theme stacks in that order.

- Theme: same as above except that the search order of instrument and theme focus stacks is reversed

- From/To-Loc: use temporal reasoning to determine what order to search past To-Loc and From-Loc lists for each utterance.

Take Engine E1 from Avon to Dansville. Pick up the boxcar.

<ve id=ve122 time=t0 theme=ne12 from-loc=ne5 to-loc=ne6> **Take** <ne id=ne12>**engine E1**</ne> **from** <ne id=ne5>**Avon**</ne> **to** <ne id=ne6>**Dansville**</ne></ve>. <ve id=ve123 time=$t1 > t0$ from-loc=ne6 theme=ne13 implicit=from-loc> **Pick up** <ne id=ne13>**the boxcar**</ne></ve>.

Figure 3: Example Annotation

Algorithm	Instrument	Theme	From-Loc	To-Loc	Overall
R-L	78.9%	55.6%	65.4%	22.2%	61.9%
L-R	78.9%	44.4%	88.5%	44.5%	73.0%
Time, L-R	78.9%	55.6%	61.5%	55.6%	65.1%
Time, R-L	78.9%	44.5%	69.3%	55.6%	66.7%
Total	19	9	26	9	

Figure 4: Implicit Role Reference Results

Our temporal reasoning scheme amounts to determining whether the current sentence u_j is in a narrative or parallel relation with a preceding utterance u_i being searched through for an antecedent. Since we annotated event times we can use the following simple algorithm to assign a narrative or parallel relation: If u_j's event time occurs after u_i's event time then we assume that a narrative relation holds between the two and that a From-Loc role in u_j should search through the To-Loc list in u_i. This is because in a narrative, there is a linear movement from place to place. If no such temporal relation is found, then we assume that a parallel relation holds between u_j and u_i and we search the From-Loc of u_i for antecedents first. The same method is used for To-Loc roles.

5. Results

We implemented the implicit role algorithm in a LISP system and and tested it on our dialog. Figure 4 shows the percentage correct for each version of the algorithm on each implicit role. The first two versions of the algorithm do not use temporal reasoning, while the last two do. R-L indicates that each focus list is searched from right to left, or from most recent to least recent. L-R indicates that the focus list is searched in reverse order, meaning that the subject of that utterance would be prominent. The last line is the number of times that role appears implicitly in the corpus.

6. Discussion

The conclusion of this study is that simple temporal reasoning has a mixed effect on the resolution rate of a verb's implicit roles. While there is a moderate improvement over the resolution of To-Loc's (55.6% to 44.5%), the naive method for resolving From-Loc's clearly outperforms its temporal reasoning counterpart (88.5% to 69.3%). Since our corpus is so small it is hard to draw concrete conclusions on whether not temporal reasoning works, especially since a most-recent strategy performs very well. This is not too surprising however since our statistics show that implicit roles typically have antecedents found locally.

It should be noted that this is a work in progress. Our annotation scheme is very basic and our error analysis shows that many of the From-Loc errors using temporal reasoning are due to deficiencies in the annotation (such as reducing complex verb phrases to their one root verb). We believe that a more detailed annotation of tense would make result in a finer temporal ordering which would improve performance. Another area of concern is our very small corpus. Many empirical studies such as (Strube, 1998) and (Tetreault, 2001) have corpora of hundreds or even thousands of annotated sentences. The larger and more varied the corpus, the more reliable the results. We also acknowledge the fact that automating the annotation of temporal relations is complicated task all to itself and that it is an area of future research.

Recent work on this corpus has looked into the effects of breaking up conjoined utterances on reference resolution as suggested by (Kameyama, 1998). and implemented by (Strube, 1998). We found that this simple metric improved scores for all implicit roles (without using temporal reasoning) as well as for pronouns in another corpus (Tetreault, 2001). We tested temporal reasoning with the utterances broken apart and found it did not improve the score any higher.

Currently, we are annotating a much larger corpus of a similar domain (emergency rescue planning for a city). We hope that using this new data will address the problems discussed above.

In short, preliminary results indicate that temporal reasoning could be useful in reference resolution, but a better annotation scheme and a larger corpus are needed to strengthen this claim.

7. Acknowledgments

I would like to thank James Allen for many discussions on implicit role reference and tense. I am also grateful to Mark Core and Donna Byron for their comments. Partial support for the research reported in this paper was provided by the DARPA research grant no. F30602-98-2-0133 to and the ONR grant no. N00014-01-1-1015, both to the University of Rochester.

8. References

Allen, Byron, Dzikovska, Ferguson, Galescu, and Stent. 2000. An architecture for a generic dialogue shell. *NLENG: Natural Language Engineering, Cambridge University Press*, 6.

Asher, Nicholas and Alex Lascarides. 1999. Bridging. *Journal of Semantics*, 15:83–113.

Byron, D. and J. Allen. 1998. Resolving demonstrative pronouns in the TRAINS93 corpus. pages 68 – 81.

Heeman, P. and J. Allen. 1994. The TRAINS93 dialogues. Technical Report TRAINS TN 94-2, University of Rochester.

Hwang, Chung Hee and Lenhart K. Schubert. 1992. Tense trees as the “fine structure” of discourse. In *Proceedings of the 30th Annual Meeting of the Association for Computational Linguistics*.

Kameyama, Megumi. 1998. Intrasentential centering: A case study. In *Centering Theory in Discourse*.

Poesio, M. 1994. Definite descriptions, focus shift and a theory of discourse interpretation. In *In Proceedings of the Conference on Focus in Natural Language Processsing*.

Strube, Michael. 1998. Never look back: An alternative to centering. In *Association for Computational Lingusitics*, pages 1251–1257.

Tetreault, Joel R. 2001. A corpus-based evaluation of centering and pronoun resolution. *Computational Lingusitics*, 27(4).

Tetreault, Joel R. 2002. Implicit role reference. Unpublished.

Webber, B. L. 1988a. Discourse deixis: Reference to discourse segments. In *Proc. of the 26th ACL*, pages 113–122.

Webber, B. L. 1988b. Tense as discourse anaphora. *Computational Linguistics*.