In I. Mani and R. Gaizauskas (eds), Readings on Temporal Information Processing, 2004.

Assigning Time-Stamps to Event-Clauses

Elena Filatova
Information Sciences Institute
University of Southern California
elena@isi.edu

Eduard Hovy
Information Sciences Institute
University of Southern California
hovy@isi.edu

Abstract

We describe a procedure for arranging into a timeline the contents of news stories describing the development of some situation. We describe the parts of the system that deal with breaking sentences into clause-sized event units and resolving both explicit and implicit temporal references for every clause in the text. Evaluations show a performance of 52% compared to humans.

1 Introduction

Linguists who have analyzed news stories (Schokkenbroek, 1999; Bell, 1997; Ohtsuka and Brewer, 1992, etc.) noticed that "narratives[1] are about more than one event and these events are temporally ordered. Though it seems most logical to recapitulate events in the order in which they happened, i.e. in chronological order, the events are often presented in a different sequence". The same paper states that "it is important to reconstruct the underlying event order[2] for narrative analysis to assign meaning to the sequence in which the events are narrated at the level of discourse structure….If the underlying event structure cannot be reconstructed, it may well be impossible to understand the narrative at all, let alone assign meaning to its structure".

Several psycholinguistic experiments show the influence of event arrangement in news stories on the ease of comprehension by readers. Duszak (1991) had readers reconstruct a news story from the randomized sentences. According to his experiments, readers follow a default strategy by which—in the absence of cues to the contrary—they re-impose chronological order on events in the discourse.

The problem of reconstructing the chronological order of events becomes more complicated when we deal with separate news stories, written at different times and describing the development of some situation, as is the case in multidocument summarization.

By judicious definition, one can make this problem easy or hard. Selecting only specific items to assign time-points to, and then measuring correctness on them alone, may give high performance but leave much of the text unassigned. We address the problem of assigning a time-point to *every* clause in the text.

Our approach is to break the news stories into their clause-sized constituent events and to assign time-stamps—either time-points or time-intervals—to these events. When assigning time-stamps we analyze both implicit time references (mainly through the tense system) and explicit ones (temporal adverbials) such as "on Monday", "in 1998", etc. The result of the work is a prototype program which takes as input set of news stories broken into separate sentences and produces as output a text that combines all the events from all the articles, organized in chronological order.

2 Data

As data we used a set of news stories about an earthquake in Afghanistan that occurred at the end of May in 1998. These news stories were taken from CNN, ABC, and APW

[1] Schokkenbroek (1999) uses the term *narrative* for news stories that relate more than one event.

[2] i.e., chronological order.

websites for the DUC-2000 meeting. The stories were all written within one week. Some of the texts were written on the same day. In addition to a description of the May earthquake, these texts contain references to another earthquake that occurred in the same region in February 1998.

3 Identifying Events

To divide sentences into event-clauses we use CONTEX (Hermjakob, 1997), a parser that produces a syntactic parse tree augmented with semantic labels. CONTEX applies machine learning techniques to induce a grammar from a given treebanks.

To divide a sentence into event-clauses the parse tree output by CONTEX is analyzed from left to right (root to leaf). The *::CAT* field for each node provides the necessary information about whether the node under consideration forms a part of its upper level event or whether it introduces a new event. *::CAT* features that indicate new events are: S-CLAUSE, S-SNT, S-SUB-CLAUSE, S-PART-CLAUSE, S-REL-CLAUSE. These features mark clauses that contain both subject (one or several NPs) and predicate (VP containing one or several verbs).

The above procedure classifies a clause containing more than one verb as a simple clause. Such clauses are treated as one event and only one time-point will be assigned to them. This is fine when the second verb is used in the same tense as the first, but may be wrong in some cases, as in *He* ***lives*** *in this house now and* ***will stay*** *here for one more year*. There are no such clauses in the analyzed data, so we ignore this complication for the present.

The parse tree also gives information about the tense of verbs, used later for time assignment.

In order to facilitate subsequent processing, we wish to rephrase relative clauses as full independent sentences. We therefore have to replace pronouns by their antecedents, where possible. Very often the parser gives information about the referential antecedents (in the example below, *Russia*). Therefore we introduce the rule: if it is possible to identify a referent, insert it into the event-clause:

> *1. Russia <2.> said;*
>
> *2.* ***which*** *<Russia> has loaned helicopters in previous disasters,*
>
> *3.* ***it*** *<Russia> would consider sending aid.*

But sometimes the antecedent is identified incorrectly:

> *Qulle charged that the United Nations and non-governmental organizations involved in the relief were poorly coordinated,* ***which*** *was costing lives.*

Here the antecedent for *which* is identified as *the relief*, giving *which <the relief> was costing lives* instead of *which <poor coordination> was costing lives.* Fortunately, in most cases our rule works correctly.

Although the event identifier works reasonably well, breaking text into event-clauses needs further investigation. Table 1 shows the performance of the system. Two kinds of mistakes are made by the event identifier: those caused by CONTEX (it does not identify clauses with omitted predicate, etc.) and those caused by the fact that our clause identifier does too shallow an analysis of the parse tree.

4 The Time-stamper

According to (Bell, 1997) "time is expressed at different levels—in the morphology and syntax of the verb phrase, in time adverbials whether lexical or phrasal, and in the discourse structure of the stories above the sentence".

4.1 Representation of Time-points and -intervals

For the present work we adopt a slight modification of the time representation suggested in (Allen, 1991), using the following formats:

Recall = (# of event-clauses correctly identified by system) / (# of event-clauses identified manually)
Precision = (# of event-clauses correctly identified by system) / (# of event-clauses identified by system)

Text number	# of clauses by human	# of clauses by system	# correct	Recall	Precision
Text 1	*7*	*6*	*5*	*5/7 = 71.42%*	*5/6 = 83.33%*
Text 2	*27*	*31*	*15*	*15/27 = 55.55%*	*15/31 = 48.38%*
Text 3	*5*	*8*	*3*	*3/5 = 60%*	*3/8 = 37.5%*
Text 4	*28*	*28*	*18*	*18/28 = 64.28%*	*18/28 = 64.28%*
Text 5	*33*	*36*	*19*	*19/33 = 57.57%*	*19/36= 52.77%*
Text 6	*58*	*63*	*36*	*36/58=62.07%*	*36/63 = 57.14%*
Total	**158**	**172**	**96**	**96/158 = 60.76%**	**96/172 = 55.81%**

Table 1. Recall and precision scores for event identifier.

- **{YYYY:DDD:W}**[3] Used when it is possible to point out the particular day the event occurred.
- **{YYYY1:DDD1:W1},{YYYY2:DDD2:W2}...** Used when it is possible to point out several concrete days when the events occurred.
- **{YYYY1:DDD1:W1}---{YYYY2:DDD2:W2}** Used when it is possible to point out a range of days when the event occurred.
- **<<<{YYYY:DDD:W}** Used when it is possible to say the event occurred {YYYY:DDD:W} or earlier.
- **>>>{YYYY:DDD:W}** Used when it is possible to say the event occurred {YYYY:DDD:W} or later.

4.2 4.2 Time-points Used for the Time-stamper

We use two anchoring time points:

<u>1. Time of the article</u>

We require that the first sentence for each article contain time information. For example:

> ***T1 (05/30/1998:Saturday 18:35:42.49)*** *PAKISTAN MAY BE PREPARING FOR ANOTHER TEST.*

The date information is in bold. We denote by ***Ti*** the reference time-point for article *i*. The symbol *Ti* is used as a comparative time-point if the time the article was written is unknown. The information in brackets gives the *exact date* the article was written, which is the main anchor point for the time-stamper. The information about hours, minutes, and seconds is ignored for the present.

<u>2. Last time point assigned in the same sentence</u>

While analyzing different event-clauses within the same sentence we keep track of what time-point was most recently assigned within this sentence. If needed, we can refer to this time-point. In case the most recent time information assigned is not a date but an interval we record information about both time boundaries. When the program proceeds to the next sentence, the variable for the most recently assigned date becomes undefined. In most cases this assumption works correctly (examples 5.2–5.3):

> *5.2.1 In the village of Kol, hundreds of people swarmed a United Nations helicopter*

[3] YYYY—year number, DDD—absolute number of the day within the year (1–366), W—umber of the day in a week (1–Monday, … 7–Saturday). If it is impossible to identify the day of the week then W is assigned 0.

5.2.2 that <a United Nations helicopter> touched down three days after Saturday's earthquake

5.2.3 <after Saturday's earthquake> struck a remote mountainous area rocked three months earlier by another massive quake

5.2.4 that <another massive quake> claimed some 2,300 victims.

5.3.1 On Monday and Tuesday, U.N. helicopters evacuated 50 of the most seriously injured to emergency medical centers.

The last time interval assigned for sentence 5.2 is {1998:53:0}---{1998:71:0}, which gives an approximate range of days when the previous earthquake happened (nominally, 3 months earlier). But the information in sentence 5.3 is about the recent earthquake and not about the previous one of 3 months earlier, so it would be a mistake to place Monday and Tuesday within that range. Similarly, Mani and Wilson (2000) point out that "over half of the errors [made by their time-stamper] were due to propagation of spreading of an incorrect event time to neighboring events". The rule of dropping the most recently assigned date as an anchor point when proceeding to the next sentence very often helps us to avoid this problem.

There are however cases where dropping the most recent time as an anchor when proceeding to the next sentence causes errors:

4.8.1 But in February a devastating earthquake in the same region killed 2,300 people and left thousands of people homeless.

4.9.1 At the time international aid workers suffered through a logistical nightmare to reach the snow-bound region with assistance.

It is clear that sentence 4.9 is the continuation of sentence 4.8 and refers to the same time point (February earthquake). In this case our rule assigns the wrong time to 4.9.1. Still, since we don't interpret phrases like "at the time", we retain this rule because it is more frequently correct than incorrect.

4.3 Preprocessing

First, the text divided into event-clauses is run through a program that extracts all explicit date-stamps (made available by Kevin Knight, ISI). In most cases this program does not miss any date-stamps and extracts only the correct ones. The only cases in which it did not work properly for the texts were:

*1.H1 PAKISTAN **MAY** BE PREPARING FOR ANOTHER TEST.*

Here the modal verb *MAY* was assumed to be the month, given that it started with a capital letter.

*6.24 Tuberculosis is already common in the area where people live in close **quarters** and have poor hygiene*

Here the noun *quarters*, which in this case is used in the sense *immediate contact or close range* (Merriam-Webster dictionary), was interpreted in the sense *the fourth part of a measure of time* (Merriam-Webster dictionary).

After extracting all the date-phrases we proceed to time assignment.

4.4 Rules of Time Assignment

When assigning a time to an event, we select the time to be either the most recently updated value or, if it is undefined, to the date of the article. We use a set of rules to perform this selection. These rules can be divided into two main categories: those that work for sentences containing explicit date information, and those that work for sentences that do not.

4.4.1 Assigning Time-Stamps to Clauses with Explicit Date Information

<u>Day of the Week</u>

If the day-of-the-week used in the event-clause is the same as that of the article (or the most recently assigned date, if it is defined), and no words before it could signal that the described event happened earlier or will happen later, then the time-point of the article (or the most recently assigned date, if it is defined) is assigned to this event. If before or after a day-of-the-week there is a word/words signaling that the event happened earlier of will happen later then the time-point is assigned in accordance with this signal-word and the most recently assigned date, if it is defined.

If the day-of-the-week used in the event-clause is not the same as that of the article (or the most recently assigned date, if it is defined), then if there are words pointing out that the event happened before the article was written or the tense used in the clause is past, then the time for the event-clause is assigned in accordance with this word (such words we call signal words), or the most recent day corresponding to the current day-of-the-week is chosen. If the signal word points out that the event will happen after the article was written or the tense used in the clause is future, then the time for the event-clause is assigned in accordance with the signal word or the closest subsequent day corresponding to the current day-of-the-week. For example,

> *5.3.1* ***On Monday and Tuesday****, U.N. helicopters evacuated 50 of the most seriously injured to emergency medical centers.*

Since the time for article 5 is *(06/06/1998:Tuesday 15:17:00)*, the time assigned to event-clause 5.3.1 is *{1998:151:1}*, *{1998:152:2}*.

Name of the Month

The rules are the same as for a day-of-the-week, but in this case a time-range is assigned to the event-clause. The left boundary of the range is the first day of the month, the right boundary is the last day of the month, and though it is possible to figure out the days of weeks for these boundaries, this aspect is ignored for the present.

> *4.8.1 But in* ***February*** *a devastating earthquake in the same region killed 2,300 people and left thousands of people homeless.*

Since the time for article 4 is *(05/30/1998:Saturday 14:41:00)*, the time assigned to this event-clause is *{1998:32:0}*---*{1998:60:0}*.

In the analyzed corpus there is a case where the presence of a name of month leads to a wrong time-stamping:

> *6.3.1 Estimates say*
>
> *6.3.2 up to 5,000 people died from the May 30 quake,*
>
> *6.3.3 more than twice as many fatalities as in the* ***February*** *disaster.*

Because of *February*, a wrong time-interval is assigned to clause 6.3.3, namely *{1998:32:0}*---*{1998:60:0}*. As this event-clause compares latest news to earlier figures it should have the time-point of the article. Such cases present a good possibility for the use of machine learning techniques to disambiguate between the cases where we should take into account date-phrase information and where not.

Weeks, Days, Months, Years

We might have date-stamps where the words *weeks, days, months, years* are used with modifiers. For example

> *5.2.1 In the village of Kol, hundreds of people swarmed a United Nations helicopter*
>
> *5.2.2 that <a United Nations helicopter> touched down three days after Saturday's earthquake*
>
> *5.2.3 after Saturday's earthquake struck a remote mountainous area rocked* ***three months earlier*** *by another massive quake*
>
> *5.2.4 that <another massive quake> claimed some 2,300 victims.*

In event-clause 5.2.3 the expression *three months earlier* is used. It is clear that to obtain the time for this event it is not enough to subtract 3 months from the time of the article because the above expression gives an approximate range within which this event could happen and not a particular date. For such cases we invented the following rule:

Time=multiplier*length[4]; (in this case, 3*30);

Day=DDD-Time; for *years* Year=YYYY-Time;

Left boundary of the range=

Day-round (10%*Day)

(for *years* = Year - round(10%*Year));

Right boundary of the range=

[4] For *days*, length is equal to 1, *weeks*–7, *months*–30.

Day + *round* (10%(Day));

(for *years* = Year + *round* (10%*Year)).

Thus for event 5.2.3 the time range will be *{1998:53:0}---{1998:71:0}* (the exact date of the article is {1998:152:2}).

If the modifier used with *weeks, days, months* or *years* is *several*, then the multiplier used equals 2.

When, Since, After, Before, etc.

If an event-clause does not contain any date-phrase but contains one of the words *when, since, after, before*, etc., it might mean that this clause refers to an event, the time of which can be used as a reference point for the event under analysis. In this case we ask the user to insert the time for this reference event manually.

This rule can cause problems in cases where *after* or *before* are used not as temporal connectors but as spatial ones, though in the analyzed texts we did not face this problem.

4.4.2 Assigning Time-Stamps to Clauses without Explicit Date Information

Present/Past Perfect

If the current event-clause refers to a time-point in Present/Past Perfect tense, then an open-ended time-interval is assigned to this event. The starting point is unknown; the end-point is either the most recently assigned date or (if that is undefined) the time-point of the article.

Future Tense

If the current event-clause contains a verb in future tense (one of the verbs *shall, will, should, would, might* is present in the clause) then the open-ended time-interval assigned to this event-clause has the starting point at either the most recently assigned date or (if that is undefined) the date of the article.

Other Tenses

Other tenses that can be identified with the help of CONTEX are *Present* and *Past Indefinite*. In the analyzed data all the verbs in Present Indefinite are given the most recently assigned date (or the date of the article). The situation with Past Indefinite is much more complicated and requires further investigation of more data. News stories usually describe the events that already took place at some time in the past, which is why even if the day when the event happened is not over, past tense is very often used for the description (this is especially noticeable for US news of European, Asian, African and Australian events). This means that very often an event-clause containing a verb in Past Indefinite Tense can be assigned the most recently assigned date (or, if that is undefined, the date of the article). It might prove useful to use machine learned rules for such cases.

No verb in the event-clause

If there is no verb in the event-clause then the most recently assigned date (or the date of the article) is assigned to the event-clause.

4.5 Sources of Errors for Time-stamper

We ran the time-stamper program on two types of data: a list of event-clauses extracted by the event identifier and a list of event-clauses created manually. Tables 2 and 3 show the results[5]. In the former case we analyzed only the correctly identified clauses. One can see that even on manually created data the performance of the time-stamper is not 100%. Why?

Some errors are caused by assigning the time based on the date-phrase present in the event-clause, when this date-phrase is not an adverbial time modifier but an attribute. For example,

1. Estimates say

2. up to 5,000 people died from the May 30 earthquake,

3. more than twice as many fatalities as in the February disaster.

As described in Section 4.4.1, the third event describes the May 30 earthquake but the time interval given by our rules for this event is {1998:32:0}---{1998:60:0} (i.e., the event happened in February). It might be possible to use machine learned rules to correct such cases.

[5] If an event happened at some time-point but according to the information in the sentence we can assign only a time-interval to this event (for example, *February earthquake*) then we say that the time-interval is assigned correctly if the necessary time-point is within this interval.

Text number	Number of event-clauses identified correctly	Number of time points correctly assigned to correctly identified clauses	Percentage correct assignment
text 1	*5*	*4*	*80.00*
text 2	*15*	*15*	*100*
text 3	*3*	*2*	*66.67*
text 4	*18*	*17*	*94.44*
text 5	*19*	*17*	*89.47*
text 6	*36*	*24*	*66.66*
Total	**96**	**79**	**82.29**

Table 2. Time-stamper performance on automatically claused texts (only correctly identified clauses are analyzed).

Text number	Number of manually created event-clauses	Number of time points correctly assigned to manually created clauses	Percentage correct assignment
target 1	*7*	*6*	*85.71*
target 2	*27*	*20*	*74.07*
target 3	*5*	*4*	*80.00*
target 4	*28*	*26*	*92.85*
target 5	*33*	*30*	*90.91*
target 6	*58*	*37*	*63.79*
Total	**158**	**123**	**77.85**

Table 3. Time-stamper performance on manually (correct) claused texts.

Another significant source of errors is writing style:

1. “When I left early this morning,

2. everything was fine.

3. After the earthquake, I came back,

4. and the house had collapsed.

5. I ***looked for two days*** *and gave up.*

6. Everybody gave up…

When the reader sees *early this morning* he or she tends to assign to this clause the time of the article, but later on seeing *looked for two days* realizes that the time of the clause containing *early this morning* is two days earlier than the time of the article. Such verbatim quotes introduce distinct temporal environments that have to be handled separately. It seems that errors caused by such writing style can hardly be avoided.

The approx. 5% discrepancy between averages in Tables 2 and 3 is interesting. We surmise that the system performs worse on manually delimited (and hence correct) clauses because some of them are simply more difficult to handle. The same reasons that make them hard to delimit—they are brief and syntactically hidden—also make it harder to find time-stamping clues.

5 Time-line for Several News Stories and its Applications

After stamping all the news stories from the analyzed set, we arrange the event-clauses from all the articles into a chronological order. We obtain a new set of event-clauses which can easily be divided into two subsets: the first containing all references to the February earthquake, the second containing event-clauses describing what happened in May, in chronological order. Such rewriting of the text into chronological order may be helpful when creating updates in multidocument summaries, where it is important to include into the final summary not only the most important information but also the most recent events.

6 Related Work

Allen presents a detailed logic-based apparatus for time representation and temporal reasoning. Unfortunately, the problem of what part of text is an event and should be assigned a time stamp is not discussed, which makes the application of this theoretical framework difficult.

A few computational applications have focused on temporal expressions in scheduling dialogues (Busemann et al., 1997; Alexandresson et al., 1997). Since such dialogues contain many explicit temporal references, these systems deliver quite high performance, but the specificity of the application makes it hard to transfer the systems to the general case.

Several other studies describe parts of the time-stamping problem, but interpret temporal information just where it occurs in the text and ignore clauses without any explicit time marking. Of course, the accuracy scores in these studies are much higher than those we obtain here. For example, systems participating in the Message Understanding Conferences (MUC, 1997) also extracted time information from given documents. However, MUC systems were restricted to a predefined template and thus extracted time phrases only for specific events in the template's focus, not for each clause. In fact, they usually extracted only a single phrase per text for a given template.

More recently, interest has been growing in creating a corpus manually annotated with time-stamps (Setzer and Gaizauskas, 2000). In this corpus time phrases are given values according to the time phrase type, and verbs are marked as events and assigned the corresponding time-stamps.

The prior work most relevant to ours is (Mani and Wilson, 2000), who implemented their system on news stories, introduced rules spreading time-stamps obtained with the help of explicit temporal expressions throughout the whole article, and invented machine learning rules for disambiguating between specific and generic use of temporal expressions (for example, whether *Christmas* is used to denote the 25th of December or to denote some period of time around the 25th of December). They also mention the problem of disambiguating between temporal expressions and proper names, as in the newspaper name *USA Today*.

7 Conclusion

We agree with Bell (1997) that "more research is needed on the effects of time structure on news comprehension. The hypothesis that the non-canonical news format does adversely affect understanding is a reasonable one on the basis of comprehension research into other narrative genres, but the degree to which familiarity with news models may mitigate these problems is unclear". Research of this topic may greatly improve the performance of the time-stamper and might lead to a list of machine learning rules for time detection.

In this paper we make an attempt to not just analyze and decode temporal expressions where they occur explicitly, but to apply this analysis throughout the whole text and assign time-stamps to each clause, which may later be used to form new sentences in such applications as multidocument summarization.

References

Alexandresson, J., N. Reithinger, and E. Maier. 1997. Insights into the Dialogue Processing of VERBMOBIL. *Proceedings of the Fifth Conference on Applied Natural Language Processing*, 33–40.

Allen, J. 1991. Time and Time Again: The Many Ways to Represent Time. *International Journal of Intelligent Systems* 4, 341–356.

Allen, J. and G. Ferguson. 1994. Actions and Events in Interval Temporal Logic. The University of Rochester, Technical Report 521.

Bell, A. 1997. The Discourse Structure of News Structure. In A Bell (ed.), *Approaches to Media Discourse*, 64–104.

Busemann, S., T. Declerck, A.K. Diagne, L. Dini, J. Klein, and S. Schmeier. 1997. Natural Language Dialogue Service for Appointment Scheduling Agents. *Proceedings of the Fifth Conference on Applied Natural Language Processing*, 25–32.

Duszak, A. 1991. Schematic and Topical Categories in News Story Reconstruction. *Text, 11(4), 503-522.*

Hermjakob, U., and Mooney, R. 1997 Learning Parse and Translation Decisions from Examples with Rich Context. *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics, 482–489.*

Mani, I. and G. Wilson, 2000. Robust Temporal Processing of News. *Proceedings of the ACL-2000 Conference*.

Message Understanding Conference (MUC). 1997. Proceedings of the Seventh Message Understanding Conference.

Schokkenbroek, S. 1999. News Stories: Structure, Time and Evaluation. *Time and Society* 8(1), 59–98.

Setzer, A. and R. Gaizauskas. 2000. Annotating Events and Temporal Information in Newswire Texts. *Proceedings of the Second International Conference on Language Resources and Evaluation*.