Ines BERRAZEGA

TRIME: An Approach for Temporal Relation Identification between Main Events

Mémoire de Mastère / ISG / Juin 2012

Université de Tunis

Institut Supérieur de Gestion

Ecole Doctorale Sciences de Gestion

TRIME: An Approach for Temporal Relation Identification between Main Events

Mémoire en vue de l'obtention du Mastère en

Informatique de Gestion

Présenté par:

Ines BERRAZEGA

Sous la direction de:

Dr. Rim FAIZ (LARODEC, IHEC)

Année Universitaire 2011/2012

TRIME An Approach for Temporal Relation Identification between Main Events

Ines BERRAZEGA

Acknowledgements

Thanks God, the merciful and mighty, for providing me the perseverance and the patience to achieve this work.

I gratefully acknowledge the jury's members who honored me by judging my dissertation. I am grateful for the time they spent to do this.

I would like to express my deep gratitude and sincere appreciation to my supervisor Dr. Rim Faiz, for her encouragement, guidance and patience she demonstrated during my study.

My heartiest thanks to my father Rafik, my mother Hayet and my sister Imen for their understanding, support and confidence to which I am forever indebted.

Finally, I would like to offer my sincere thanks to the members of my family and my freinds who have supported me during this year.

Contents

List of Figures

List of Tables

Introduction

Information Extraction (IE) field is gaining increased attention by researchers who seek to acquire knowledge from huge amount of Natural Language contents. Many works have been proposed to extract valuable information from text. Existing approaches are classified into three categories, namely knowledge-based approaches (Riloff, 1993); machine-learning approaches (Seymore and al., 1999) and hybrid approaches (Califf and Mooney, 2003). Even though IE approaches have tackled a variety of issues, dynamic facts like *temporal information* have been neglected, although time is a crucial dimension in any information space. This limitation can be explained by the complexity of such task. For example, the classical techniques used to extract named entities and events from textual contents are unable to identify the temporal relations between events or to infer the chronological ordering of these events. Such processes require a greater effort to analyze *how temporal information is conveyed in textual contents*, especially when temporal information is implicitly expressed. In this concern, recent researchers have aimed to expand the capabilities of existing Natural Language Processing (NLP) systems to account for the temporal dimension of language. Thus the Temporal Information Extraction (TIE)tasks first appeared in the scope of the Fifth Message Understanding Conference (MUC-5), when it was asked to assign a calendrical time to a joint venture event (Kufman, 1993) .

The usefulness of Temporal Information interested researchers in different areas, who have approached the problem in different ways. In fact, the scope of re-

search in Temporal Information is broadened, ranging form classical theories of time and language to current computational approaches. The first attempts to understand temporal information in Natural Language was done from philosophical and linguistic perspectives. Several works have been proposed. They discussed a wealth of theories related to time and language. The most outstanding works have studied linguistic mechanism of time namely tense and aspect (Reichenbach, 1947);(Vendler, 1967); Temporal Reasoning (Allen, 1983); and Temporal Structure of Discourse (Lascarides and Asher, 1993). These works have not been of interest to the larger Artificial Intelligence community until very recently.
With the recent emergence of NLP tasks like Question Answering, Summarization and Information Extraction, many computational approaches have been proposed (Faiz, 2006); (Mani and al., 2006); (Chambers and al., 2007); (H.Llorens and al., 2010) . They are handling two main issues namely *automatic recognition of temporal entities in narrative* (namely temporal expressions and events), and *temporal relations identification between these entities*. Such tasks are considered quite complex; tackling them cannot be limited to simple pattern recognition. It requires a deep comprehension and studying of Natural Language contents, especially that temporal information can be expressed in implicit way and some expressions can be ambiguous.

It's worth noting that, even though existing approaches have achieved great performances in recognizing temporal expressions and events, the identification of temporal relations remains a challenging task. In fact, current approaches are limited on *morpho-syntactic analysis* and a minority of them has reached the *semantic level*. Also, they did not take benefit from the wealth of the classical works carried on *the Temporal Structure of Discourse*. However, identifying temporal relations is a complex task. *It requires linguistic knowledge at all language analysis levels, including semantic, pragmatic and discourse.*

These findings motivated us to propose a new approach able to come with the limitations of related works. This is precisely our motivation for applying mechanisms related to the pragmatic influences of discourse, mainly *causality*, in addition to morpho-syntactic and semantics ones to process temporal relations between events. We are interested in the scope of this work in a specific type of temporal relations namely main *Event-Event relations of consecutive sentences.* For this purpose, we will investigate related works and we will define and implement a model for applying all linguistic analysis levels, ranging from morpho-syntactic to pragmatic analysis. We aim with this attempt to contribute in the area by improving the state of art systems' performances.

In the first chapter, we present the general framework of our work, namely the *Temporal Information Extraction field*. We first define its basic concepts. Then we study the classical works on Temporal Information Processing. After that, we study the current main issues on Temporal Information Extraction field, its real world applications and some computational approaches.

In the second chapter, we attempt to summarize related works on *Temporal Relation Identification between main events in two consecutive sentences.* First, we state the basic temporal relations concepts. Then, we highlight the computational Temporal Relation Identification tasks. Later, we discuss the most similar approaches to our work. Finally, we identify their limitations and we introduce our new proposal.

In the third chapter, we detail the different steps of our new approach for Temporal Relation Identification between Main Events: TRIME.

In the fourth chapter, we present the validation of our approach by the system that we have developed as part of this work. An evaluation and comparison of obtained results with state-of-art approaches are also presented.

Finally, we conclude by highlighting areas for improvement in our approach as well as future directions.

Chapter 1

The Literature Review on Temporal Information Extraction

1.1 Introduction

In last decade, with the expansion of communication tools, million of users access to the Web every day to look for different kind of information. Therefore, data provided by the Web has become the most important source of information. Most of these contents describe dynamic facts namely events of all kinds (political, economic, cultural...). Such information is valuable for many Natural Language Processing applications like Question Answering, Summarization, and Information Retrieval. Thus, extracting temporal information from text has become a new challenging field.

In the scope of this work, we are interested in issues of this new field, namely the temporal relation identification between main events. This first chapter is devoted to introduce the Temporal Information Extraction field. In section 1.2, we state a brief survey of Information Extraction. Then, we present the main concepts related to Temporal Information in section 1.3. We review in section 1.4 the classical works on Temporal Information Processing. After that, we focus on the current issues in Temporal Information Processing in section 1.5. Then, we state some real-world applications in section 1.6, followed by some computational approaches in section 1.7.

1.2 From Information Extraction to Temporal Information Extraction

Daily, huge amounts of unstructured data are produced everywhere. These contents need to be exploited and analyzed to acquire relevant information useful in different domains. This issue has been the topic of several fields of research such as Information Retrieval (IR) and Information Extraction (IE).

1.2.1 Information Retrieval vs. Information Extraction

Information Retrieval (IR) involves searching and finding documents that answer the user's requirement. According to (Gaizauskas and Wilks, 1998), Information Retrieval is based on document retrieval with the aim to answer the question "how to find, in a set of documents, those that interest me?" An Information Retrieval system generally proceed by query interpretation, document representation, indexing and ranking retrieved documents. On the other side, Information Extraction (IE) aims to extract interesting information from documents for an automatic analysis by a computer. The extraction techniques have to deal with the understanding of the meaning of natural language. (Sarawagi, 2008) defines Information Extraction as "the automatic extraction of structured information such as entities, relationships between entities, and attributes describing entities from unstructured sources. This enables much richer forms of queries on the abundant unstructured sources than possible with keyword searches alone." Information Extraction dates back to the late 1970s in the early days of NLP. Later, start-ing from 1987, the scope of IE was strongly influenced by two competitions: the Message Understanding Conferences (MUC) (Grishman and B.Sundheim, 1996); (Chinchor, 1998); (Sundheim, 1991) and Automatic Content Extraction (ACE)[1] program. In the scope of these competitions, several tasks have been discussed: named entity recognition, co-reference resolution, event extraction, etc.

1.2.2 Limitations of classical Information Processing techniques

With the growing amount of unstructured texts, there is a need to exploit new Natural Language Processing techniques able to account for *the temporal dimension* and to handle *the ambiguity in natural language contents*. Thus, the techniques used in classical fields of research like Information Retrieval and Information Extraction need to be consolidated with new ones. Thereby, the Temporal Information Extraction field has emerged. Processing temporal information is valuable in many NLP tasks like Question Answering, Summarization and Information Retrieval. Thus, we are interested in temporal information, mainly *the temporal relation between main events*. We present in the next section the main Temporal Information concepts.

1.3 Basic Concepts

When studying temporality, it is important to start with defining basic concepts related to time. In this section we try to provide a clear description of temporal entities and relations in textual contents that enables us to gain useful insights into how temporal information is conveyed in written language.

[1]NIST.Automatic content extraction (ACE) program.http://www.itl.nist.gov/iad/mig/tests/ace/ (accessed 10/03/2012)

1.3.1 Eventualities: Events and States

The first two main temporal entities in texts are *Events* and *States*. These two entities are grouped under the term of *eventualities* (Bach, 1986). Events are things that happen or occur in the world (like weddings, birthdays, and parties...) at a certain time or over a given period of time and in a given place. They are typically dynamic occurrences that have causes and effects, a clear beginning and end, and bring about some perceptible change in the world (Asher, 1993).
States, on the other hand, are considered as the existence of a set of properties over a given period (like happiness, owning a car...) often without a clearly defined beginning and end.
Thus, we can indicate that the main difference between events and states is attributed to the *dynamic/static* distinction, where events are seen as dynamic and resulting in some change, while states, on the other hand, do not involve a perceptible change.
When studying eventualities, (Bittar, 2010) tries to highlight the relation between events and states. He mentions that "the occurrence of an event may bring about the existence of a state. Similarly, an event may also bring about the end of a state. In other words, a change of state in the world typically indicates the occurrence of an event."

1.3.2 Temporal Expression

Temporal expressions are the second device expressing temporality in language. A temporal expression can be defined as being any element in language giving an information about when a given event happened, how long it lasted or how often it occured. Such information is generally expressed in terms of quantifiable temporal units. In accordance with (Sauri and al., 2006), there are four categories of temporal expressions according to the types of instants or intervals they describe:

dates, times, durations and sets.

- **Date:** are expressions that refer to a particular period based on *the Gregorian calendar*. This includes units that are larger than a year, such as centuries and millennia, as well as subintervals of a typical year, such as seasons. The basic unit on which the calendar is based is the day. Dates can be expressed in an absolute form (eg, *Tuesday 18th*) or in a relative form (eg, *last week*).
- **Time:** are expressions which denote a particular subdivision of a day. They are understood to refer to particular moments that subdivide a day. Also, times represent a relatively simple category of temporal expressions and may correspond to the moments we measure on a clock, or express more general parts of a typical day (eg, *in the morning/ at 9 a.m*)
- **Duration:** are expressions which refer to an extended period of time, very often specifying the temporal extent of an eventuality. They are measured using calendrical units (*years, months, days* etc.) or clock units (*hours, minutes, seconds* etc.) of temporal measure (eg, *3 hours last Monday*).
- **Set:** are expressions that refer to the regularity or reoccurrence of an eventuality, either in the absolute, or relative to a period of time (eg, *twice a week*).

1.3.3 Temporal Relations

In natural language, temporal relations hold between temporal entities: between two events, between two temporal expressions or between an event and a temporal expression. (Longacre, 1983) defines a temporal relation as "an inter-propositional relation that communicates the simultaneity or ordering in time of events or states". On this basis, reasoning about time implies the temporal representation of these relations. In this concern, most of computational models for reasoning about time are based on the Allen's Interval Algebra (Allen, 1983) to capture the temporal dimension of a narrative. More details about this algebra are given in Appendix A.

1.4 Classical Works in Temporal Information Processing

The scope of research in Temporal Information Processing is broadened, ranging form classical linguistic theories of tense and aspect to current computational approaches. Prior to describe how Temporal Information Processing is currently addressed, it is necessary to review the background studies on time and language. The following subsections review the most outstanding classical Temporal Information Processing issues from three different perspectives (Mani and al., 2005). The first subsection addresses the core linguistic theories of tense and aspect. The second subsection reviews the temporal reasoning from an AI perspective. The third subsection explores the works on the temporal structure of discourse. Finally, the fourth subsection presents current approaches based on annotations schemes.

1.4.1 Linguistic theories of tense and aspect

Temporal information is conveyed in natural language through grammatical mechanisms to represent time and temporal relations. In this concern, tense and aspect are considered as the most important grammatical categories in language. The most prominent works on linguistic theories are the Reichenbach's theory for tense called *'The Tenses of Verbs'* (Reichenbach, 1947), and the Vendler's work (Vendler, 1967) for aspect which has been the basis for subsequent researchers.

We present these two works with more details in the Appendix A.

1.4.2 Temporal Reasoning

Temporal Reasoning involves the temporal representation of events and their temporal anchoring within natural language text. This topic has attracted great attention due to its potential applications in several NLP tasks like Summarization, Question Answering and so on. This attention was accompanied with the development of efficient representation models and has been a central area of research in AI since the 1960s.

The state-of art of Temporal Reasoning proposes several *constraint based models*. In a constraint based model, the temporal information is represented as a temporal constraint network (TCN) in which the events are denoted by nodes and the ordering constraints between events are denoted by edges. In this case, reasoning about time becomes a Temporal Constraint Satisfaction Problem (TCSP). Different TCNs are defined depending on the representation of the temporal entity as time intervals, durations or points, and the class of constraints namely qualitative, quantitative, metrics or its combination (Sanampudi and Kumari, 2010). Several researches fall into this framework (Allen, 1983); (Filatova and Hovy, 2001); (Setzer and Gaizauskas, 2002). We present more details about these models in Appendix B.

1.4.3 Temporal structure of discourse

Interpreting relations between temporal entities at discourse level is required to correctly understand natural laguage contents. In fact, sentences are usually interpreted in context instead of isolation. For this purpose, several works have been proposed to define a formal representation for the temporal structure of discourse.

In this context,(Kamp, 1981) proposes an influential work on *Discourse Representation Theory* (DRT). This work introduces a new level of mental representation, known as *the discourse representation structures* (DRSs). In fact, Kamp considers that a hearer of a given discourse builds up a mental representation for each sentence of this discourse in a cumulative way. This theory gives the possibility to interpret sentences in the context of a discourse, rather than sentences in isolation using a construction procedure extending a given DRS.
Other remarkable work on the temporal structure of discourse is Dowty's *Temporal Discourse Interpretation Principle* (TDIP)(Dowty, 1986). Another well known work on temporal structure of discourse is Webber's theory of *anaphoric reference* (Webber, 1988). In this work, Webber develops an account for noun-phrase anaphora and shows how tensed clauses can be treated as anaphors like noun phrases. This work is also based on the Reichenbach's theory for tense. We can also mention the Hwang and Schubert's *Tense Trees* (Hwang and Schubert, 1992). Authors try to consider the compositional semantics of complex sentences. For this purpose, they automatically convert the logical form of sentences into a temporal structure as a tense tree. Other works on temporal structure of discourse are interesting namely the approach proposed by (ter Meulen, 1995) about *the Dynamic Aspect Trees* (DATs) ; *the Reference Point* (Rpt) work presented by (Kamp and Reyle, 1993) and *the Temporal Conceptual Graphs* proposed by (Webber, 1988).
The pragmatic influences of discourse, such as *causality*, have been also considered in DRT. (Lascarides and Asher, 1993) pointed out that *common-sense and real world knowledge* - or what they called *defeasible reasoning* - are needed for the correct interpretation of temporal relations.

1.4.4 Discussion

Even though the presented works highlighted a variety of issues and proposed many solutions for understanding and representing temporal information, they were criticized because they were not analyzed and evaluated over real linguistic data. In this concern, it's worth noting that Temporal Information Processing and Natural Language Processing in general have undergone a mutation from the *rationalist strategy* based on such formal theories of language analysis to an *empiricist strategy* based on the analysis of real language use (i.e., textual corpora) (Manning and Schutze, 1999). This mutation was accompanied by the development of *computational semantic models* and the establishment of *evaluation frameworks* to annotate temporal information in a corpus based vision. In the next section, we highlight this new trend.

1.5 Current issues of Temporal Information Processing

As mentioned in the previous section, *Temporal Annotation* has become in the last decade the most prominent issue in the Temporal Information Processing area, especially with the development of annotation schemes and annotated corpora. With this new trend, researches are focusing on two tasks: on one hand *automatically recognizing and extracting temporal entities in narrative* (namely temporal expressions and events), and on the other hand on *discovering temporal relations between these entities and inferring the type of each recognized one.* The following subsections are dedicated to present these two tasks (section 1.5.1) and (section 1.5.2) as well as the most prominent annotation schemes and annotated corpora (section 1.5.3).

1.5.1 Temporal entity recognition

This task consists on identifying and extracting temporal entities form natural language texts. For events, this means finding which textual entity constitutes an event. For example, in the sentence *"she bought% 15 of the shares"*, (Filatova and Hovy, 2001) consider that the entire clause represents an event. While for (Pustejovsky and al., 2003a), the appropriate span is just the verb group or just the head of the verb group (*"bought"* in this example). Furthermore, event recognition involves the identification of some attributes related to each event. Such attributes depend on the topic being studied. In TimeBank corpus, Pustejovsky and al, define five attributes namely tense, aspect, modality, polarity and class(Pustejovsky and al., 2003b).

For temporal expressions, entity recognition amounts to identify the type of each temporal expression (time, date, duration or frequency) and to find its corresponding value.

1.5.2 Temporal relation identification

Besides temporal entities, natural language texts contain other devices expressing logical relations between times and events or between events and events. The temporal relation identification task aims to recognize such relations and to infer the type of each recognized one. Several researches fall into this task. (Allen, 1983); (Filatova and Hovy, 2001) and (Setzer and Gaizauskas, 2002) propose representation models to capture temporal relationships. We present more details about these models in Appendix A.

1.5.3 Annotation schemes and annotated corpora

Handling temporal annotation requires the development of annotation schemes and the construction of annotated corpora. In this concern, a lot of ongoing researches are focusing on the development of annotations schemes to extract, model and interpret temporal information in natural language texts (Manning and Schutze, 1999). The most prominent temporal annotation schemes are: MUC-TIMEX (Grishman and B.Sundheim, 1996), TIDES (Ferro and al., 2001), STAG (Setzer and Gaizauskas, 2000), and TimeML (Pustejovsky and al., 2003a). All of them follow a SGML/XMLbased annotation format. We give more details about these schemes in Appendix B.

TimeML (Time Markup Language) is the latest annotation scheme. It was developed under the sponsorship of ARDA as the natural evolution of STAG. This new scheme is considered to be a combination and an extension of preceding schemes, which makes it the most complete and potent one. TimeML has recently been standardized to an ISO international standard for temporal information markup, ISO-TimeML (ISO-TimeML, 2007). Both the TimeML and the ISO-TimeML annotation standards define the following basic XML tags: <EVENT> for the annotation of events, <TIMEX3> for the annotation of time expressions, <SIGNAL> for locating textual elements that indicate a temporal relation, and the tags <TLINK>, <SLINK> and <ALINK> that capture different types of relations.

Regarding the construction of corpora, a corpus is seen as a collection of natural language data contents organized according to a set of criteria with the aim of conducting specific research. These contents are often enriched with descriptive or analytic annotations, in order to explain a particular linguistic phenomenon. In terms of temporally annotated corpora, several corpora have been developed within the framework of the TimeML annotation scheme. TimeBank 1.1, TimeBank 1.2 and the AQUAINT Corpus are all made up of journalistic texts in English annotated with the TimeML annotation language. TimeBank 1.2 (Pustejovsky and al., 2003b) has become somewhat of a reference for the

study of temporal information within the Computational Linguistics community. It's the last major TimeML annotated corpus to and it has become the main reference for temporal annotation in English. It's also used as the basis for the TempEval evaluation campaigns' corpora.
After getting insights on current issues of Temporal Information Processing, it's worth exposing some real world applications (section 1.6) as well as the most prominent computational approaches proposed to tackle Temporal Information Processing tasks (section 1.7).

1.6 Real-World Applications

Extracting Temporal Information from texts is a key problem for many real-world applications like Web Search, Medical-Records, Legal Reasoning, Accounting, Banking, Reservation Systems, and Accident Reports. Examples from these applications are presented in the next subsections.

1.6.1 Web Search

Query log analysis is currently an active topic of research. In fact, there is significant and growing number of contributions to the understanding of online user behavior. In this concern, temporal information has received little attention by researches who want to understand and characterize users' behavior online. Researchers try to identify and characterize the use of temporal expressions in web search queries. In this concern (Nunes and al., 2008) led a query log analysis to investigate the use of temporal expressions in web search queries. They find that temporal expressions are rarely used (1.5% of queries) and, when used, they are related to current and past events. Also, there are specific topics where the use of temporal expressions is more visible.

1.6.2 Medical Domain

Temporal Information Systems are lately used in several medically-related problems and tasks such as diagnosis, monitoring and longitudinal guideline-based care. Several applications in the medical domain are used to extract information about times of clinical investigations (X-rays, ultrasounds, etc.). In this concern, (Roberts and al., 2008) propose an algorithm to extract temporal relations between temporal expressions and clinical investigation events form clinical narratives.

1.6.3 Legal Reasoning

In legal domain, extracting temporal information has a great utility to facilitate the work of lawyers. Several applications have been proposed namely the automatic temporal ordering of legal documents in a time line, or the automatic retrieval of a specific event mentioned in various legal documents according to temporal constraints that may be associated with this event. (Schilder, 2007) presents a prototype system that extracts events from the United States Code on U.S. immigration nationality and links these events to temporal constraints.

1.6.4 Financial Accounting

Many applications in the accountancy field have used temporal information to solve several issues. (Fisher, 2007) presents a prototype system to support the temporal reconstruction of Financial Accounting Standards (FASs). This prototype enables a user to specify an FAS along with a date, and permits a dynamic and continuing codification of a particular FAS.

1.7 Computational approaches

Motivated by the expected applications of Temporal Information Extraction in various fields, many temporal-aware systems have been developed. In the next subsections, we expose some of the current computational approaches for Temporal Information Extraction.

1.7.1 REES: Relation and Event Extraction System

The REES System, provided by (Aone and Ramos-Santacruz, 2000), is a large-scale events and relation extraction system that recognizes and extracts a total of 100 types of events and relations from various domains like business, politics and crime. The system consists of three main components: a tagging component, a co-reference resolution module and a template generation module. Events and relations are associated with templates which contain slots for various properties to be extracted. The properties vary according to the type of relation or event to be extracted. Slots are filled by lexical entries associated with generic syntactic patterns. For events, this involves searching for place and time adjunct information. Event templates contain a TIME slot to be filled with the time of occurrence of the event. Evaluation of the system reports F-scores of 73.74% and 53.75% for relation and event extraction, respectively. The system does not deal at all with nominal event mentions. Added to that, the co-reference module doesn't resolve pronouns (like "it" and "they") which carry little semantic information. The system also suffers from several errors due to a lack of patterns for certain constructions.

1.7.2 TARSQI Tool Kit: Temporal Awareness and Reasoning Systems for Question Interpretation

TTK or TARSQI Tool Kit developed by (Verhagen and al., 2005), is the most famous state-of-art system. It is a complete annotation system based on the TimeML annotation scheme, and composed from five modules:

- **GUTime** is the module for recognizing TIMEX3 elements. This module is based on TempEx tagger (Mani and Wilson, 2000).
- **EVITA** is the module in charge of event recognition. The rules rely on: PoS tagging, lemmatizing and chunking obtained using Alembic Workbench; lexical lookup and contextual parsing; and WordNet information combined with Bayesian learned disambiguation for identifying noun events.
- **GUTenLINK** is a TLINK tagger based on syntactic and lexical information.
- **Slinket** automatically identifies subordinating relations between pairs of events which are represented by SLINK tag in TimeML.
- **SputLink** is a temporal closure component that extends known temporal relations with new relations.

1.7.3 EXEV system

The EXEV system is developed by (Faiz, 2006) to automatically extract all information about events (paragraphs or sentences) from news articles. Based on a morpho-syntactic approach, EXEV extracts temporal information running five modules:

- **Lexical analysis module:** allowing the chunking of a text into sentences and into words.

- **Morphological analysis module:** identifying words while triggering functions that deal with morphological inflexions and generate a morpho-syntactic code for each word.

- **Syntactic analysis module:** re-establishing the order of the morpho-syntactic codes generated by the morphological analyzer with the aim of building some morpho-syntactic structures.

- **Extraction module:** allowing picking out markers in order to identify distinctive sentences which represent events.

- **Interpretation module:** allowing the interpretation of the extracted sentences to identify "Who did what?", "to whom?" and "where?".

1.7.4 TIPSem (Temporal Information Processing based on Semantic information)

(H.Llorens and al., 2010)propose the TIPSem system to extract temporal information from natural language texts for English and Spanish. In their approach, the authors are based on the hypothesis that the linguistic expression of time is a semantic phenomenon and therefore, to achieve a better extraction performance, temporal information must be processed using semantics. For this purpose, they use features based on lexical semantics, semantic roles, and temporal semantics. The approach is based on Conditional Random Field (CRF) and Support Vector Machines (SVM) to learn an annotation model. They processed on 4 steps: recognition, classification, normalization and link-categorization.

1.8 Conclusion

Temporal Information Processing is a recent area gaining more interest among researchers. Our literature review reveals the existence of two trends in the domain: classical works arising from the intersection of linguistics, philosophy and symbolic AI; and current works based on annotation schemes and annotated corpora. We have also seen that recent works handle several interrelated issues namely the automatic recognition of temporal entities and the temporal relations identification. We devote the next chapter to present an overview of the Temporal Relations Identification tasks, and we focus on a specific task which is the Temporal Relation between Main Events of consecutive sentences.

Chapter 2

Temporal Relation Identification between main events

2.1 Introduction

As seen in chapter 1, the Temporal Information is gaining more interest lately given its importance in many Natural Language Processing applications like Question Answering, Summarization and Information Extraction... In this concern, a lot of ongoing researches are focusing on Temporal Information Processing, especially with recent construction of annotation schemes and annotated corpora mainly the TimeBank corpus (Pustejovsky and al., 2003b). Proposed approaches vary according to the strategies adopted and the temporal entities treated. Temporal Entity Recognition approaches have achieved satisfactory results, while the Temporal Relation Identification is considered very complicated even for a human annotator, which makes it challenging in recent years. This motivation led us to focus on thisissue. We are interested in the scope of this work in a specific type of temporal relation namely *the mains Event-Event relations of consecutive sentences.* We have chosen this task among all due to its high complexity. Thus, we consider its achievement as being an interesting challenge, especially that existent approaches which tackled it didn't obtain high performances.

The remainder of this chapter is organized as follows. In section 2.2 we state basic concepts related to Temporal Relation Identification. Then, we present in section 2.3 the computational Temporal Relation Identification tasks. Section 2.4 will be dedicated to study the most similar approaches to our work, followed by the limitations of these approaches in section 2.5 and our new proposal in section 2.6.

2.2 Basic Concepts

Given an approach that identifies temporal expressions and events in a textual content, the next task is determining the relations that may hold between two temporal entities. In this section we describe the main concepts related to Temporal Relation Identification.

2.2.1 Time-Event relation

The Time-Event relationship is the relation that may hold between an event and a temporal expression in a textual content. The temporal expression can be the document creation time (DCT) or any other temporal expression in the text. A Time-Event relationship can be expressed in two manners:

a. Explicitly (eg, *I woke up at 8 a.m*): the relation is expressed via a prepositional phrase.

b. Implicitly (eg, *I woke up at 8 a.m, I prepared the breakfast and then I ate it*): there is no explicit temporal relation between preparing the breakfast and temporal expressions, or eating and temporal expressions; but there is event-event relation of preparing the breakfast and eating with waking up and waking up has an explicit relation with temporal expression.

2.2.2 Event-Event relation

The Event-Event relationship is the relation that may hold between two events. Two cases are possible:

a. **Main Event-Event relation (inter-sentential)**. This consists of categorizing the temporal relation between two main events in consecutive sentences. The main event of a sentence is considered to be the syntactically dominant verb of that sentence.

b. **Subordinated Event-Event relation (intra-sentential)**. This consists of determining the temporal relation between two events in the same sentence, where one event syntactically dominates the other event.

Several researchers have attempted to model the representation of these temporal relations, to enable better understanding and exploitation of these relations in further applications. In next section, we introduce briefly the best known Temporal Relation representation models.

2.3 Computational Temporal Relation Identification tasks

Based on the presented concepts, several works have attempted to identify temporal relations. Initially, based on TimeBank annotations (Pustejovsky and al., 2003b), most of the published works focused on identifying only temporal relations between events. Later in TempEval 2007, (Verhagen and al., 2007) introduce three particular tasks:

2.3.1 Event-Timex relation

This task consists on determining the temporal relation between an event and a temporal expression the in same sentence. It was called Task A in TempEval-1 then Task C in TempEval-2. The following example illustrates the text input and the expected output. The categorization process indicates that "obtained" overlaps with "2010" (Table 2.1).

Raw data	She obtained her diploma in 2010
Annotated data	She <EVENT eid=1>obtained</EVENT> her diploma in <TIMEX3 tid=1>2010</TIMEX3>
Input	<TLINK lid="l1" reltype="**NONE**" leid="1" ltid="1"/>
Output	<TLINK lid="l1" reltype="**OVERLAP**" leid="1" ltid="1"/>

Table 2.1 – Event-Timex relation in the same sentence

2.3.2 Event-Document Creation Time relation

This task consists on determining the temporal relation between an event and the document creation time (DCT). It was called Task B in TempEval-1 then Task D in TempEval-2. Table 2.2 illustrates the text input and the expected output for this task. The categorization process indicates that "born" (1987) was before the document was written or created (2010-06-15).

Raw data	DCT: 2010-06-15 She was born in 1987
Annotated data in	DCT: 2001-01-15 (tid="0") She <EVENT eid=1> was born</EVENT> <TIMEX3 tid=1>1987</TIMEX3>
Input	<TLINK lid="l1" reltype="**NONE**" leid="1" ltid="0"/>
Output	<TLINK lid="l1" reltype="**BEFORE**" leid="1" ltid="0"/>

Table 2.2 – Event-DCT relation

2.3.3 Main Event-Event relation (inter-sentential)

This task consists on determining the temporal relation between main events of two consecutive sentences. It was called Task C in TempEval-1 then Task E in TempEval-2. Table 2.3 illustrates an example of the text input and the expected output for this task. The categorization output indicates that "born" was before "obtained" and "obtained" was after "studied".

Raw data	She was born in Tunisia. She obtained her master degree in Paris. She had studied previously in Tunis
Annotated data	She was <EVENT eid="1">born</EVENT> in Tunisia. She <EVENT eid="2">obtained</EVENT> her master degree in Paris. She had <EVENT eid="3">studied</EVENT> previously in Tunis
Input	<TLINK lid="l1" reltype="**NONE**" leid="1" leid="2" /> <TLINK lid="l2" reltype="**NONE=**" leid="2" leid="3" />
Output	<TLINK lid="l1" reltype="**BEFORE**" leid="1" leid="2"/> <TLINK lid="l2" reltype="**AFTER**" leid="2" leid="3"/>

Table 2.3 – Main events relation in two consecutive sentences

2.3.4 Subordinated Event-Event relation (intra-sentential)

In TempEval 2010, (Verhagen, 2010) introduced another task called task F: This task consists on determining the temporal relation between two events where one event syntactically dominates the other event. The following example illustrates the text input and the expected output. The categorization indicates that "saw" and "explosion" overlap in time.

In the scope of this work, we are interested in determining *temporal relation between main events of two consecutive sentences.* So, we will describe in the next section how researchers approached this task.

Raw data	She was born in Tunisia. She obtained her master degree in Paris. She had studied previously in Tunis
Annotated data	She <EVENT eid=1>saw</EVENT> an <EVENT eid=2> explosion</EVENT>
Input	<TLINK lid="l1" reltype="**NONE**" leid="1" leid="2" />
Output	<TLINK lid="l1" reltype="**OVERLAP**" leid="1" leid="2" />

Table 2.4 – Subordinated events relation in same sentence

2.4 Approaches for Temporal Relation Identification between Main Events

Since the establishment of evaluation frameworks mainly in the scope of the *TempEval campaigns*, several researches have proposed computational approaches to handle temporal relations identification between events situated in two different sentences.

Due to the complexity of this task, TempEval reduces the problem of identifying temporal relations beyond sentence level to the task of relating the main events of two adjacent sentences. Recall that the main event of a sentence is considered to be the syntactically dominant verb of that sentence. Proposed works for Temporal Relation Identification between main events used the TimeBank corpus and the TimeML annotation scheme. In fact, TimeML captures the temporal relations with the TLINK tag, which has event id (to identify the event), timex id (to identify the temporal expression) and temporal relation. This task consists on recognizing the following subset of Allen's relations: BEFORE, AFTER, BEFORE-OR-OVERLAP, OVERLAP-OR-AFTER, VAGUE. As for general Information Extraction tasks, there are three major ways to identify temporal relations between events: rule-based approaches, machine learning approaches and hybrid approaches. We present in the following paragraphs the most outstanding approaches addressing this task, focusing on the contribution of each one compared to its previous.

(Hagège and al., 2007) present a rule-based system called XRCE-T in scope of TempEval-1. Their approach relies on a deep syntactic analyzer that was extended to treat temporal expressions. Temporal processing is integrated into a more generic tool consisting on a linguistic analyzer. Temporal analysis is intertwined with syntactic-semantic text processing like deep syntactic analysis and determination of thematic roles. TempEval specific treatment is performed in a post-processing stage.

At this stage, it's worth noting that, given the high complexity of Temporal Relation Processing, this task requires more advanced techniques than those used in rule-based approaches. That's why we find only one system which approaches the problem with rule-based techniques.

(Mani and al., 2006) propose a Machine-learning based approach for temporally ordering and anchoring events in natural language texts. They build a maximum entropy classifier that assigns a temporal relation class to each pair of events. Their classifier relies on the gold standard features extracted from their OTC corpus, and the pairs of tense and aspect agreement (tense1-tense2/ aspect1-aspect2).
Mani and al., adopt the assumption that human intuition is required to infer temporal ordering of events. For this purpose, they develop pattern matching rules to tackle the event ordering task. In a first step, they develop a TLINK tagger named **GTag**, in which they incorporate 187 hand-coded lexical and syntactic rules. This tagger relies on syntactic information from part-of-speech tagging and chunking to infer and label TLINKs between tagged events. For each pair of events, the tagger searches for the most-confident rule to apply, then it assigns the suitable class of the TLINK relation. The following example shows one of these rules.

If sameSentence=YES &&

sentenceType=ANY &&

conjBetweenEvents=YES &&

arg1.class=EVENT &&

arg2.class=EVENT &&

arg1.tense=PAST &&

arg2.tense=PAST &&

arg1.aspect=NONE &&

arg2.aspect=NONE &&

arg1.pos=VB &&

arg2.pos=VB &&

arg1.firstVbEvent=ANY &&

arg2.firstVbEvent=ANY &&

then infer relation=BEFORE

Experiments showed that the intuition-based strategy used in this baseline has very low accuracy. Even when heuristic preferences are intuited; those preferences need to be guided by empirical data, whereas hand-coded rules are relatively ignorant of the distributions that are found in data. To overcome this deficiency, (Mani and al., 2006) investigate the use of empirically derived lexical relations between verbs. They incorporate in their baseline the set of *happens-before* relations derived from the **Verbocean database**. However, this second strategy didn't make a considerable improvement. This limitation is explained with the data sparseness: most of the Verbocean entries rarely occur in the OTC corpus.

Even though Mani and al., are the first to try incorporating Common-Sense Knowledge to consider event ordering, their approach presents limits. The development of hand-coded lexical and syntactic rules is not effective; it's time consuming and based on intuition, which induce to incompleteness. Also, this strategy needs to be completed with assertions drawn from empirical data. Added to that, although Mani and al., tried to incorporate lexical rules in their system, they used it in a rule-based baseline.

(Chambers and al., 2007) propose a two-stage machine learning approach to learn temporal relations between pairs of events. In the first stage, they learn for each event the five gold standard attributes (tense, aspect, modality, polarity and class) from TimeBank corpus. To do so, they implement both Naïve Bayes and Maximum Entropy classifiers and use a set of morpho-syntactic features. In the second stage, they split the data on two separate training sets: one for events in the same sentence, and the other for events crossing sentence boundaries. Then, they learn for each event pair the relation class on the training data set. For this, they use the five attributes learned in stage one as well as new morpho-syntactic features namely event string, lemmas, Wordnet synset and event class agreement.

Experimental results show that the introduction of new morpho-syntactic features in the temporal relation learning give a 4.3% gain when compared to Mani's approach. However, the event attributes learned in the first stage lead a decrease in the accuracy (50.19%), compared to the original ones found in the corpus (50.97%). This decrease is due to the introduction of imperfect data. Also the split strategy makes the data sparser and involves a drop in performance (59.43%) when compared to results obtained without this split (60.45%).
To sum up, it's clear that the best strategy is to keep the original gold standard attributes as tagged in the training corpus, and to add new features to learn temporal relations between events. Nevertheless, an approach limited to a morpho-syntactic analysis level could not be enough robust for Temporal Information Processing.

Based on the assumption that semantics are needed to improve the performance of current approaches, (H.Llorens and al., 2010) propose TIPSem: a Temporal Information Processing system based on Semantic information. In this

system, the authors focus on semantic roles and semantic networks to train CRF models for identifying relations between temporal entities. To learn main event relations, TIPSem uses features related to the main event pairs and the syntactically closest timexes they may have. In fact, the authors use morpho-syntactic features to detect the tense-aspect agreement and whether the two events are in the same sentence or not. The novelty in this approach is the use of a new semantic feature indicating if both events (event1 and event2) are syntactically linked with two different temporal expressions (timex1 and timex2 respectively). This feature represents the order between timex1 and timex2 (before, equal or after), which helps inferring main event relations.
Even though TIPSem's approach is the best-reported when addressing the Temporal Relation Identification between main events (F-score= 55%), this system is unable to solve complicated inter-sentential relations like causality as shown in the following example:

a. John *pushed* Marc. Marc *fell*.

b. Marc *fell*. John *pushed him*.

In such cases, TIPSem categorizes both relations as before considering that events with the same tense are normally ordered in narrative forward.

In this regard, we believe that using a common-sense knowledge strategy to consider event ordering is crucial. For this purpose, we would try to give some insight how a common-sense knowledge resource could be used in a fully automatic system.
Upon investigating in the state-of-art, we have found only one machine-learning approach (Ha and al., 2010) integrating such resource. In fact, (Ha and al., 2010) try to incorporate in their NCSU system two lexical relation features learned from Verbocean and Wordnet. They also use some morpho-syntactic features inspired from Bethard and Martin's approach (Bethard and Martin, 2007). Then, they

train 4 Naïve Bayes models to learn temporal relations between events and time expressions as defined in the TempEval-2 evaluation campaign. For main event relations in two consecutive sentences, they use a weighted version of the lexical relation features learned from Verbocean and Wordnet. This version includes a score to measure the degree of confidence for each temporal relation.
Experimental results show that the use of the weighted version gave lower results (48%) than those obtained with the unweighted version used to learn subordinated event relations (66%).

2.5 Limitations of current approaches

By studying related works, we have drawn the followings conclusions. Regarding the linguistic knowledge used, most of the systems are limited to the use of morpho-syntactic derived features. The latest proposals integrated semantic mechanisms. This points out that morpho-syntactic properties of language are central for addressing temporal information processing and that the use of semantics for this task is quite novel. Furthermore, we have noticed that current computational approaches did not take benefit from the wealth of theoretical works on the temporal structure of discourse reviewed in chapter 1. Complicated inter-sentential relations like causality are left unsolved. Nevertheless, the context of a discourse is required to reach the correct language comprehension, including the temporal ordering of events. Only one machine learning approach has integrated a common-sense knowledge mechanism. This trend is fairly new due to the limitation of computational resources. Whereas, it's clear that lexical rules have a role to play in the semantic and pragmatic reasoning from language.

2.6 Our contribution

The limitations of related works prompted us to propose a new and effective approach **TRIME** for **T**emporal **R**elation **I**dentification between **M**ain **E**vents. We would try to give some insight how the combination of syntactic, semantic and pragmatic knowledge (mainly causality) could be used in a fully automatic approach.

Unlike most of the approaches that do not exploit any form of semantic understanding (Hagège and al., 2007); (Bethard and Martin, 2007); (Ha and al., 2010), and those that do not tackle complicated inter-sentential relations like causality (H.Llorens and al., 2010), our approach is able to process a natural language text at all analysis levels: morpho-syntactic, lexical, semantic and pragmatic. At each level we explore a large set of features using a variety of tools and methods. The whole process is illustrated by the diagram of Figure 2.1 which can be detailed as follows:

- **Pre-processing**. In this step, we initially employ several Natural Language Processing (NLP) techniques in text analysis. These techniques include splitting, tokenization, lemmatization, part-of-speech tagging, and parsing. Then, we extract main events and their related temporal expressions.

- **Feature extraction**.The feature extraction consists on finding a set of features from the preprocessed data set at all analysis level. The goal of this step is to find out relevant indicators that may help us capturing the accurate class of the temporal relation between two events situated in different sentences.

- **Temporal relation classification**. This step consists on training and testing classification models on our data set in order to learn for each pair of main events their corresponding temporal relation class.

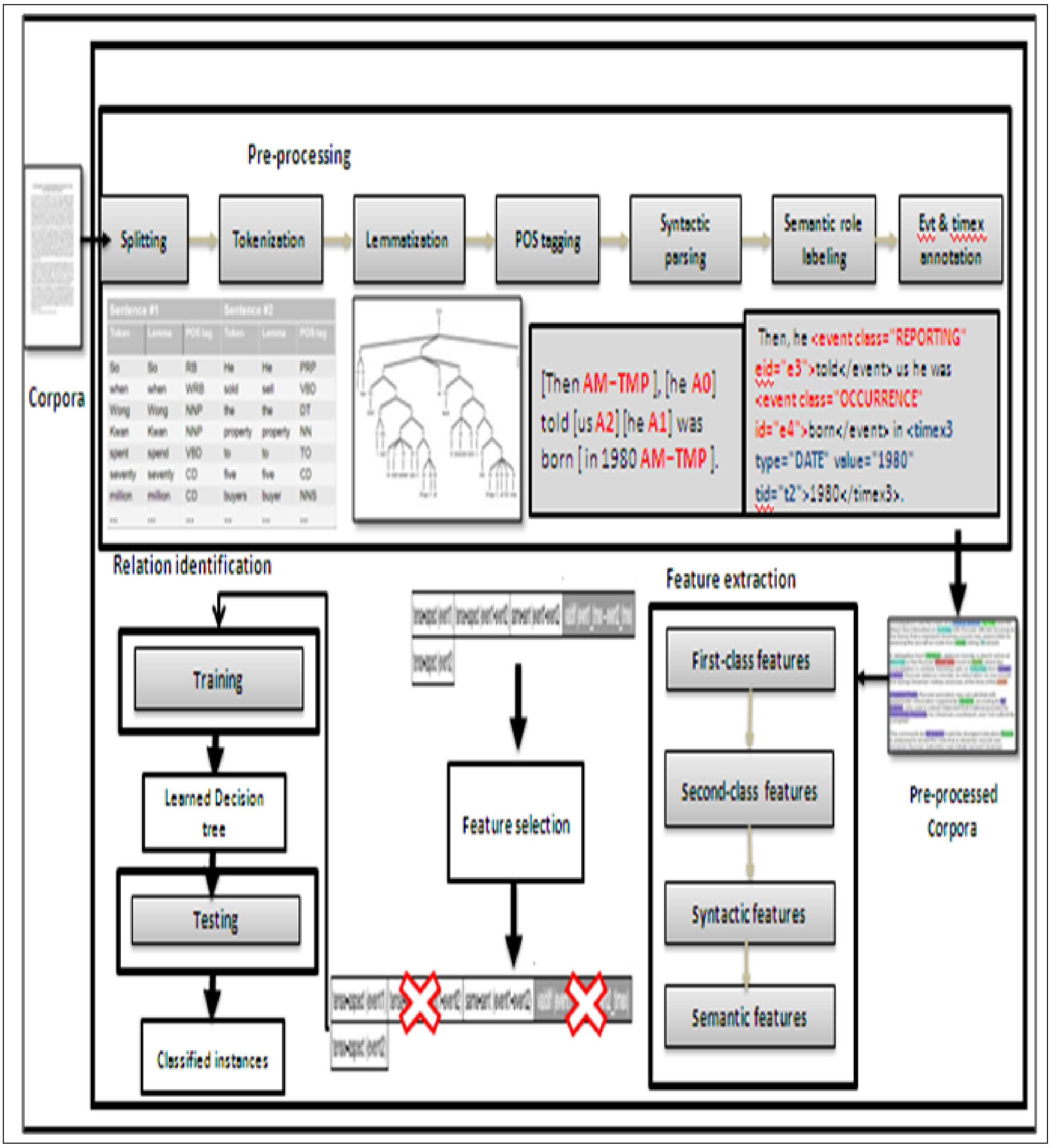


Figure 2.1 – Proposed approach

2.7 Conclusion

In this chapter, we presented an overview on temporal relation identification between main events of two consecutives sentences. This task requires a considerable effort due to its complexity. Most of current approaches are limited to the use morpho-syntactic features, which explain the low results that they have obtained. Some approaches have tackled the task with a semantic analysis level, and only one machine learning based approach has studied causality to infer the temporal relations between main events. These limitations led us to propose a new approach to identify temporal relations between events in textual contents in order to contribute in the area by improving the state of art systems' performance.

Chapter 3

Proposed Temporal Relation Identification approach

3.1 Introduction

Identifying temporal relation between two events situated in different sentences is a very complex task. It requires a deep analysis and the use of a variety of techniques and tools to explore the contribution of different features in the performance of the approach. In the scope of this work, we proceed the Temporal Relation Identification task as a classification problem using a novel machine learning approach. This third chapter is dedicated to describe our proposal: TRIME as Temporal Relation Identification between Main Events. The remainder of this chapter is structured as follows. In Section 3.2, we detail the pre-processing process. In section 3.3, we depict our feature extraction strategy used to find out useful features for the temporal relation identification. In section 3.4, we study several machine-learning classification models to pick out the most suitable one to our work.

3.2 Pre-processing

The pre-processing step consists on employing several Natural Language Processing (NLP) techniques to prepare the data for the Temporal Relation Identification task. Initially, this process consists on splitting each document into sentences, then into words (or tokens); assigning to each word its part-of-speech category and deriving its lemma. Then, this process is followed by a syntactic analysis to find out syntactic relations among the entities of each sentence. Finally, the syntactic analysis is complemented by a semantic analysis, in which we apply a semantic role labeling method to assign semantic roles to each entity. In the following subsections, we describe the pre-cited steps.

3.2.1 Sentence Boundary Detection (Splitting)

Splitting a document consist on identifying its sentences boundaries. This task is hard given that the identification of a point followed by a capital letter is not enough to detect the end or the beginning of a sentence. This task requires taking into consideration more sophisticated indicators, namely the syntactic structure of a sentence and the punctuation markers in a well defined context (Mourad, 2002); (ElKhlifi and Faiz, 2009).

3.2.2 Tokenization

Tokenization is the process of segmenting the text into elementary linguistic units like numbers, words, and punctuation symbols. This process provide a set of elements called tokens used as input for other processing steps like lexical or syntactic parsing. Most of algorithms used for tokenization consider that white spaces and punctuation are clues of token boundaries. But for languages which have no word boundaries, more sophisticated techniques are required. In our approach, we apply a tokenizer to split each sentence into tokens.

3.2.3 POS Tagging

A part of speech is a linguistic category of a word, generally assigned according to its definition as well as its context. In fact, the same word can be a verb in one sentence and a noun in another one. So, part of speech states the way in which a given word is used. Grammatically, a word may have one of this eight parts of speech classes: verb, noun, pronoun, adjective, adverb, preposition, conjunction, and interjection. The part-of-speech tagging process marks up words in a text with their corresponding grammatical classes.

3.2.4 Lemmatization

Lemmatization is the process of determining the lemma for a given word. Since the process may involve complex tasks such as understanding context and determining the part of speech of a word in a sentence, it can be a hard task to implement a lemmatizer for a new language. Lemmatization is closely related to stemming. The difference is that a stemmer operates on a single word without knowledge of the context, and therefore cannot discriminate between words which have different meanings depending on part of speech. However, stemmers are typically easier to implement and run faster, and the reduced accuracy may not matter for some applications.

After splitting the text into sentences, the pre-described steps out-put a set of morphological features at token level. In our approach we use the Stanford CoreNLP , an integrated suite of NLP tools for English for sentence splitting, tokenization, and POS tagging; and Wordnet for lemmatization.

The following example illustrates the output of these four steps. We take the following two consecutive sentences from the TimeBank 1.2 corpus:

"Now with new construction under way, three of his buyers have

backed out. And Wong Kwan will be lucky to break even."

Table 3.1 shows the result of applying the preprocessing techniques on these two sentences.

Sentence#1			Sentence#2		
Token	**Lemma**	**POS**	**Token**	**Lemma**	**POS**
Now	Now	RB	And	and	CC
with	with	IN	Wong	Wong	NNP
new	new	JJ	Kwan	Kwan	NNP
construction	construction	NN	will	will	MD
under	under	IN	be	be	VB
way	way	NN	lucky	lucky	JJ
,	,	,	to	to	TO
three	three	CD	break	break	VB
of	of	IN	even	even	TR
his	his	PRP	.	.	.
buyers	buyer	NNS			
have	have	VBP			
backed	back	VBN			
out	out	PR			

Table 3.1 – Morphological features

3.2.5 Syntactic analysis

Besides the features obtained at token level, it is also crucial to have good dependency features for pairs of entities. We observe from related works that most of the syntactic dependencies strongly indicate temporal relations. For this reason, we apply a syntactic parser to determine the syntactic structure of each sentence. The output from the parser is a syntactic parse tree. In our work, we use the

Stanford parser; a statistical parser providing grammatical relations between the words of a sentence. Figure 3.1 and Figure 3.2 show the syntactic parse trees related to the example given in table 3.1.

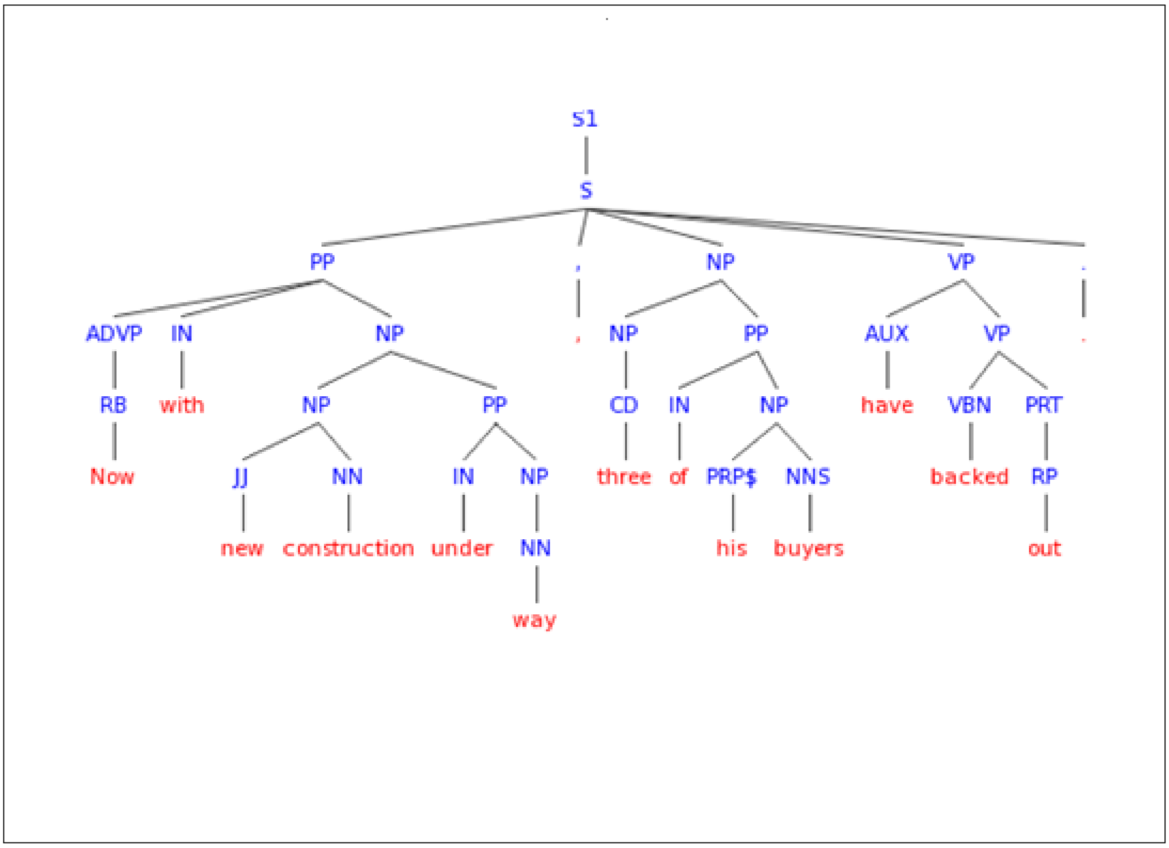


Figure 3.1 – Syntactic parse tree of sentence 1

3.2.6 Semantic analysis

Our semantic analysis consists on studying the semantic relations holding between a syntactic constituent and its predicate. For a predicate, each constituent is an argument (agent, patient, instrument, etc.) or an adjunct (locative, temporal, manner, etc.). Thus, recognizing and labeling semantic arguments is a key task

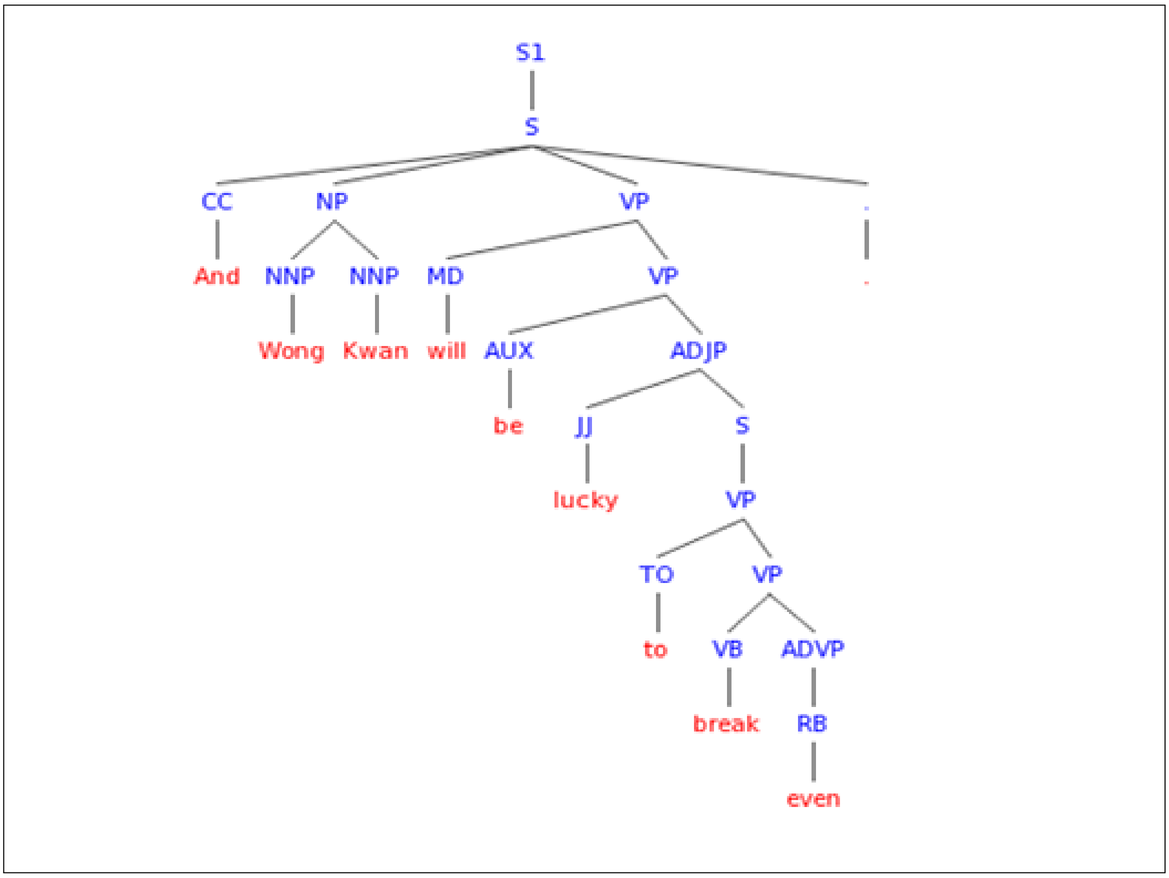


Figure 3.2 – Syntactic parse tree of sentence 2

in our approach. In this concern, we apply a semantic role labeling tool SENNA[1]. Figure 3.3 shows the semantic role labeling output for our example.

3.2.7 Temporal entity recognition

In this step, we first identify and annotate all events and temporal expressions in each sentence as defined in the TimeML markup language. Then we filter out the main event of each sentence and their related temporal expressions (if any). Recognized events and temporal expressions are associated with a set of attributes showing their important aspects namely tense, aspect, modality, polarity and

[1]http://ml.nec-labs.com/senna/ (accessed 24/03/2012)

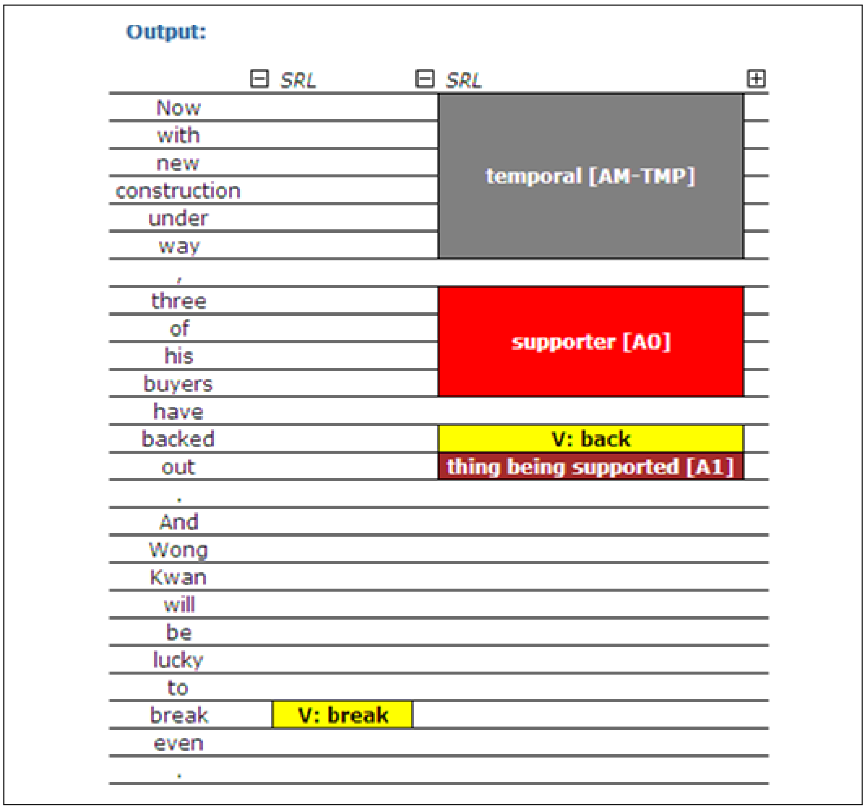


Figure 3.3 – Semantic Role Labeling

class for events; and type and value for temporal expressions (Pustejovsky and al., 2003b). This operation allows us obtaining these basic attributes for each main event and its related temporal expression (if any). To do so, we use the TIPSem-B annotator (H.Llorens and al., 2010).

3.3 Feature extraction

The feature extraction is the task of constructing feature vectors for the Temporal Relation Identification task. In this concern, we use the output of the pre-processing modules and we proceed as follows. First, we extract the gold standard attributes of each event and temporal expression. In a second time, we combine these features to derive other ones. Then we extract new features from the syntactic and semantic analysis. All the extracted features are later described in the next subsections.

3.3.1 First-class features

For each event and temporal expression, we learn a set of features directly obtained from the markups of the training data set. In fact the attributes which are directly obtained from the tags of the training data set have a great impact on performance of machine learning classifiers, compared with effects of other features. For events, tense and grammatical aspects are necessary in any method of temporal relation classification. They are used to find out a clear distinction among the grammatical categories of verbal phrases, and to define temporal location and event structure. Modality and polarity specify non-occurring (or hypothetical) situations. The event class shows the type of event. Tables 3.2 and 3.3 show the range of values for events and temporal expressions attributes based on (Pustejovsky and al., 2003a).

Timex features	Range of values
Type	Date, time, set, duration
Value	ISO 8601 normalized value

Table 3.2 – Temporal expression features and their range of values

3.3.2 Second-class features

Besides the features obtained at temporal expression and event level, it is also crucial to use other features dealing with the dependency between main event pairs. Inspired by (H.Llorens and al., 2010), we use the first-class features to derive three other ones. Table 3.4 gives an explication of these features.

Event features	Range of values
Tense	none, present, past, future
Aspect	none, prog, perfect
Polarity	positive, negative
Modality	none, to, should, would, could can, might
Class	report, aspectual, state, I state I action, perception, occurrence

Table 3.3 – Event features and their range of values

Tense and aspect (event1-event2)	Combination of tense and aspect of the two events. This attribute help us finding the precedence between them.
Same sentence (event1-event2)	This feature indicates whether the two events are in the same sentence or not.Normally, main events relations are identified for events in consecutive sentences. However, sometimes two main events of the same sentence are considered
Time position	If both events (event1 and event2) are syntactically linked with two different temporal expressions (timex1 and timex2 respectively), this feature represents the order between timex1 and timex2 (before, equal or after). Otherwise the value is set to "equal" by default.

Table 3.4 – Temporal expression features and their range of values

3.3.3 Syntactic and semantic features

Our approach aims to overcome the weakness of current systems in Temporal Relation Identification between main events. For this purpose, we have found that exploring contextual indicators may solve many indeterminations in the semantic relationship between these events. We have tried to find out suitable syntactic and semantic features from the analysis done in the preprocessing step. Our feature extraction strategy is done in two steps. First we extract some features similar to those used in related works on Temporal Relation Identification (Bethard and

Martin, 2007); (Ha and al., 2010). Our choice is based on the efficiency of these features and their contribution in the improvement of systems performances. In a second step, we try to find other features able to boost the robustness of our approach. These features are learned from the context of the events.
For each pair of main events, we compute various lexical, syntactic and semantic features. The tuned feature set is shown in table 3.5.

Features	**Explication**
Word features	According to their positions, four categories of words are considered: (1) the words of both events; (2) the words between the two events; (3) the words before event 1; (4) the words after event 2.
Overlap	The number of words separating two main events.
Governing verb	The verb governing each event.
POS-governing verb	The part of speech of the governing verb.
Prepositional phrase	Preposition heads are often indicators of a temporal class, thus we can use a new feature that indicates if an event is part of a prepositional phrase
Connecting indicators	The presence of a subordinated conjunction preceding the event (eg. **Before** the*meeting*, when he was *born*) is a relevant temporal indicator in the determination of the order between events.
Modal indicators	The presence of a modal indicator preceding the event (eg. you **should** *wash* your hands before eating).

Table 3.5 – Syntactic and semantic features

3.3.4 Discourse features

In addition to syntactic and semantic information, pragmatic and background world knowledge play a crucial role to infer the temporal relations between events situated in different sentences. Consider the following examples:

1. John pushed Marc. Marc fell.

2. Marc fell. John pushed him.

3. Naima opened the door. The room was pitch dark.

4. Naima switched off the light. The room was pitch dark.

Examples (1 and 2) or (3 and 4) have the same syntax. In example 1, the order in which the events are described matches their temporal order, whereas in example 2 narrative order mismatches temporal order. If we only consider semantics, we categorize both relations in Examples 1 and 2 as *before*. Again, the event and state in example 3 temporally *overlap*, whereas in example 4 they do not.

To handle such complex inter-sentential temporal relations, we are referring to *the defeasible reasoning theory* proposed by (Lascarides and Asher, 1993). In this work, Lascarides and Asher propose a formal account of the pragmatic influences in event ordering in discourse. They state that a theory is needed to combine common-sense knowledge and pragmatic principles in a formal logic. Lascarides and Asher talk about two issues to be handled with defeasible reasoning: *Explanation* (example 4) and *Result* (example 2). Even though this work interest researchers, it is still not for any practical use because of limitation of common-sense knowledge. By seeking in the literature of the domain, we have found only one computational approach dealing with the defeasible reasoning. In fact, (Chklovski and Pantel, 2004) extract 22 306 semantic relations between 3 477 verbs by finding phrases matching their lexico-syntactic patterns using

Google search. Among five semantic relations (table 3.6), they extract temporal *happens-before relations*, which can be used to handle ambiguous cases like in example 2. Semantic relations are presented as a list of verb pair relations, along with a confidence score. These relations are stored in a lexical database called Verbocean [2].

Relation	Example
SIMILARITY	*produce :: create*
STRENGTH	*wound :: kill*
ANTONYMY	*open :: close*
ENABLEMENT	*fight :: win*
HAPPENS-BEFORE	*buy :: own*

Table 3.6 – Semantic relations between verbs in VERBOCEAN

To resolve the ambiguity of temporal relations between events at discourse level, we need to apply inferences based on such world knowledge bases. That's why we integrate in our approach a new discourse feature. We filter out all *happens-before* semantic relations from Verbocean, and we have stored the obtained relations in a relational database. In a second step, for each main event pair, we compare the lemmatized forms of event against entries of the new database.

At the end of the feature extraction steps, we obtain a set of features which will be used in the Temporal Relation Identification step.

[2]http://demo.patrickpantel.com/demos/verbocean/

3.4 Temporal Relation Identification

The Temporal Relation Identification step consists on assigning to each event pair the corresponding temporal relation class. We address this task as a supervised classification problem. For this purpose, we use the feature vector obtained previously to train and test a classification model.

A variety of supervised Machine Learning techniques can be used as Decision Tree, Naïve Bayes, Neural Network, Logistic Regression, etc. We choose Decision Tree for several reasons: it is a simple graphical model, easy to understand and to interpret by human. Discriminating variables are ranked in an easily readable form of tree. In addition, the construction of this model is less configurable, compared to other techniques. This option reduces the system complexity and speeds up the learning and the implementation process. Added to that, most decision tree algorithms can be implemented in both serial and parallel form while others can only be implemented in either serial or parallel form (Anyanwu and Shiva, 2009).

In order to choose an appropriate decision tree algorithm, we test the IDE3 (Quinlan, 1986) and C4.5 (Quinlan, 1993) under Weka, then we have retained the IDE3 model given that it gives better results over the training data set.

We also test a Naïve Bayes model to compare its performance with the IDE3 model. Details about these results will be presented in the next chapter.

In a second step, we evaluate the obtained Decision Tree and Naïve Bayes classifiers using the test data set provided in the scope of the TempEval2 campaign.

3.5 Conclusion

Identifying Temporal Relation between two events situated in different sentences is a very complex task. It requires a deep analysis and the use of a variety of techniques and tools to explore the contribution of different features in the performance of the approach. In this chapter, we have proposed a new Temporal Relation Identification approach with the aim of improving state-of-art systems' performance. Compared to related works, our proposal tackles all analysis levels, making it more complete and more robust. Next chapter provides the implementation and the evaluation of our approach in the context of the TempEval-2 evaluation campaign.

Chapter 4

TRIME: A system for Temporal Relation Identification between Main Events

4.1 Introduction

In this chapter, we present the evaluation of the new system that we implemented to validate our Temporal Relation Identification approach. We perform tests with the resources provided in the scope of the TempEval-2 evaluation campaign [1]. The remainder of this chapter is organized as follows: in section 4.2, we summarize the different steps and tools used in this implementation. In section 4.3, we describe the steps for the use of the system. We present the experimental data and we discuss our experimental results in section 4.4. Finally, we conclude the chapter in section 4.5.

[1]http://timeml.org/site/timebank/timebank.html

4.2 Implementation

The implementation of our system is performed through several steps, in which we use several tools. Our Temporal Relation Identification pipeline consists of three major parts: preprocessing, feature extracting and a temporal relation classification. In what follows, we summarize the different steps and tools used in this implementation.

4.2.1 Preprocessing Tools

The pre-processing of the data is performed using a pipeline of NLP tools. First, we use the Stanford CoreNLP suite [2] for sentence splitting, tokenization and POS tagging; and Wordnet for the lemmatization of the training and the test documents. Then, we use the Stanford parser for the syntactic parsing and the dependency parsing. The output of the parser for the example taken previsouly in chapter 3 is illustrated as follows:

```
(ROOT
  (S
    (PP
      (ADVP (RB Now))
      (IN with)
      (NP
        (NP (JJ new) (NN construction))
        (PP (IN under)
          (NP (NN way)))))
    (, ,)
    (NP
      (NP (CD three))
      (PP (IN of)
```

[2] http://nlp.stanford.edu/software/lex-parser.shtml

(NP (PRP$ his) (NNS buyers))))

(VP (VBP have)

(VP (VBN backed)

(PRT (RP out))))

(. .)))

(ROOT

(S (CC And)

(NP (NNP Wong) (NNP Kwan))

(VP (MD will)

(VP (VB be)

(ADJP (JJ lucky)

(S

(VP (TO to)

(VP (VB break)

(ADVP (RB even))))))))

(. .)))

Then, the output of the parser is converted to the collapsed form of the Stanford dependency scheme.

advmod(backed-13, Now-1)

amod(construction-4, new-3)

prep_with(backed-13, construction-4)

prep_under(construction-4, way-6)

nsubj(backed-13, three-8)

poss(buyers-11, his-10)

prep_of(three-8, buyers-11)

aux(backed-13, have-12)

root(ROOT-0, backed-13)

prt(backed-13, out-14)

cc(lucky-6, And-1)
nn(Kwan-3, Wong-2)
nsubj(lucky-6, Kwan-3)
aux(lucky-6, will-4)
cop(lucky-6, be-5)
root(ROOT-0, lucky-6)
aux(break-8, to-7)
xcomp(lucky-6, break-8)
advmod(break-8, even-9)

For the semantic parsing, we employ the SENNA[3] semantic role labeler to annotate each predicate with its corresponding role. The output of the semantic role labeling is given below for this sentence: ***"Google announced a new product yesterday."***

Google	S-NP	S-ORG	-	S-A0
announced	S-VP	O	announced	S-V
a	B-NP	O	-	B-A1
new	I-NP	O	-	I-A1
product	E-NP	O	-	E-A1
yesterday	S-NP	O	-	S-AM-TMP
.	O	O	-	O

Finally, we annotate events and temporal expressions with the TimeML EVENT and TIMEX3 tags using the TIPSemB system. The output of the temporal annotation is as follows.

[3] http://ml.nec-labs.com/senna/

<TEXT>

<TIMEX3 type="DATE" value="PRESENT_REF" tid="t1">

Now</TIMEX3> with new <EVENT class="OCCURRENCE" eid="e1">

construction</EVENT> under way , three of his buyers have

<EVENT class="OCCURRENCE" eid="e2">**backed**</EVENT> out.

And Wong Kwan will be lucky to <EVENT class="OCCURRENCE"

eid="e3">**break**</EVENT> even.

</TEXT>

<MAKEINSTANCE eiid="ei1" eventID="e1" pos="NOUN"

tense="PRESENT"aspect="PERFECTIVE" polarity="POS"/>

<MAKEINSTANCE eiid="ei2" eventID="e2" pos="VERB"

tense="PRESENT"aspect="PERFECTIVE" polarity="POS"/>

<MAKEINSTANCE eiid="ei3" eventID="e3" pos="VERB"

tense="FUTURE" aspect="NONE" polarity="POS"/>

4.2.2 Features extraction Tools

After having applied the pre-processing tools on the training data set, we tackle the feature extraction. For this purpose, we implement JAVA methods using the development environment Eclipse to build feature vectors, and we create a relational data base using MySQL Workbench 5.2[4] to save the extracted instances and their corresponding values. The figure 4.1 illustrates our database. Our database includes 5 tables.

a. Document table: for each document of both training and test data sets, we insert an instance including the identifier of the document and its type (training/test)

b. Token table: for each token of both training and test data sets, we insert its identifier, the identifier of its corresponding document, its word, its rank in the document and its rank in the sentence.

[4] http://www.mysql.com/products/workbench/

c. Verbocean table: this table includes the extracted happens-before relations from Verbocean.

d. Timex table: this table includes all temporal expressions with their identifier and their corresponding tokens, types and values.

e. Event table: for each event, this table includes an identifier, the identifier of its token, its tense, aspect, polarity, modality and class.

f. Relation table: this table is used to save event-event instances and their related temporal relations. It includes the identifiers of the two events and the relation class.

After creating the database, we implement a JAVA algorithm to extract form each document morphological, syntactic, dependency, semantic and discourse features; events and temporal expressions tokens; first class-features and second-class features. All these features constitute our feature vector. We store the output of this process in a CSV[5] file to learn a classification model for the temporal relation identification.

4.2.3 Temporal Relation Identification Tools

Temporal relation identification is considered as a classification problem. For this purpose, we have to learn a supervised Machine-Learning model able to correctly classify the type of temporal relations between two events. In this step, we use the obtained feature vector to learn a supervised classifier on the training data set. We employed Weka 3.6 to build Decision Tree and Naive bayes models for train and test.

4.3 Running the system

In this section, we describe the overall use of the *TRIME* system. Once the system is run (cf. Figure 4.2), the user chooses the training option from the setting menu (cf. Figure 4.3). Then he selects a text file (cf. Figure 4.4), and he starts running the

[5]Comma-Separated Values

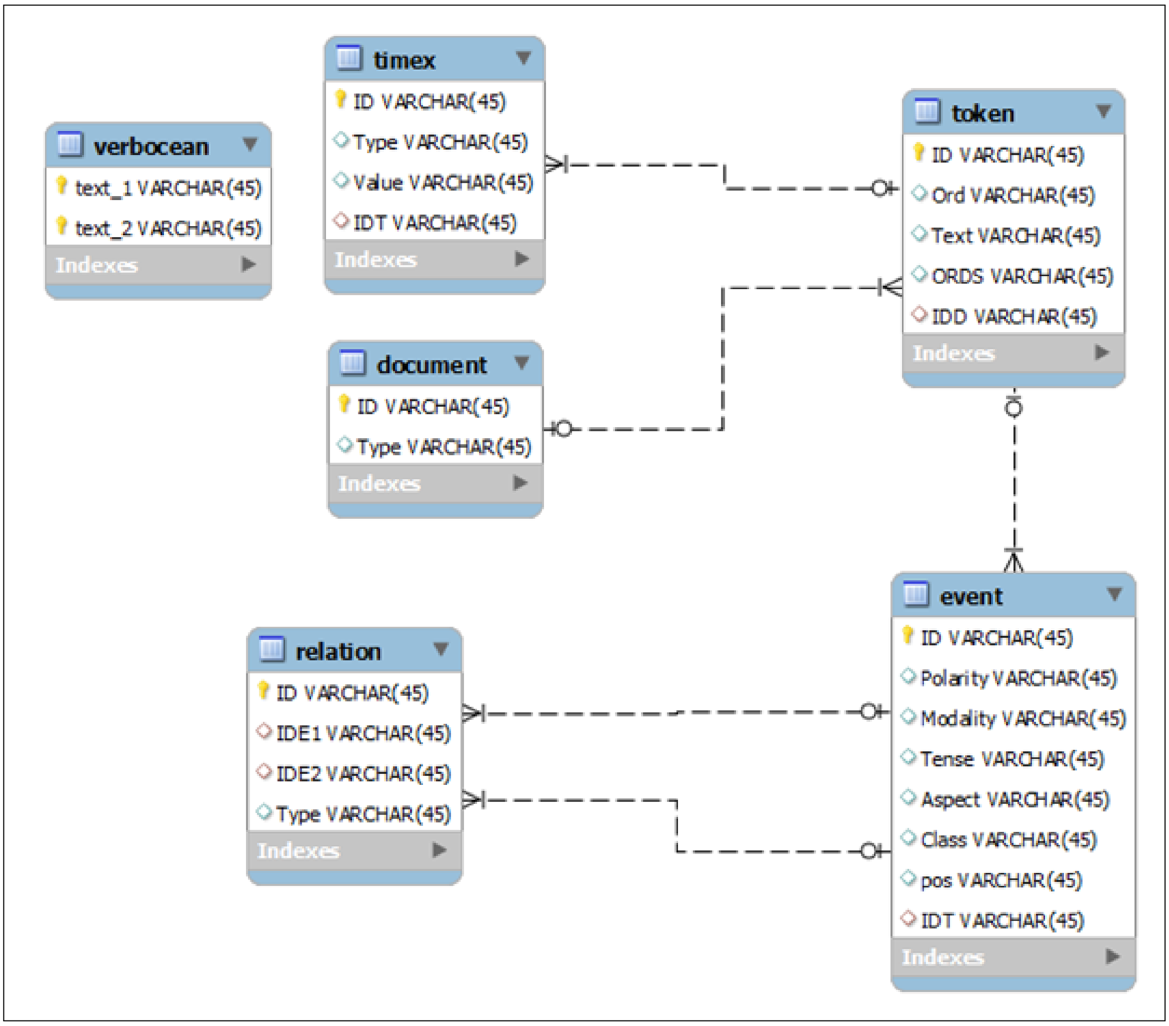


Figure 4.1 – Our database

pre-processing functions: sentences splitting (cf. Figure 4.5), tokenization (cf. Figure 4.6), POS tagging (cf. Figure 4.7), Parsing (cf. Figure 4.8) and Temporal annotation to assign the Event and Timex3 tags (cf. Figure 4.9). The results are stored in the database.

In a second time, the user runs the feature extraction functions (cf. Figure 4.10) to generate the "example_train" file (cf. Figure 4.11) . Each line in the example file represents one training instance of the following format:

<target> <feature>:<value> <feature>:<value> ... <feature>:<value>.

Finally, the training examples are used to build a Decision Tree model able to predict the temporal relation class between two main events.

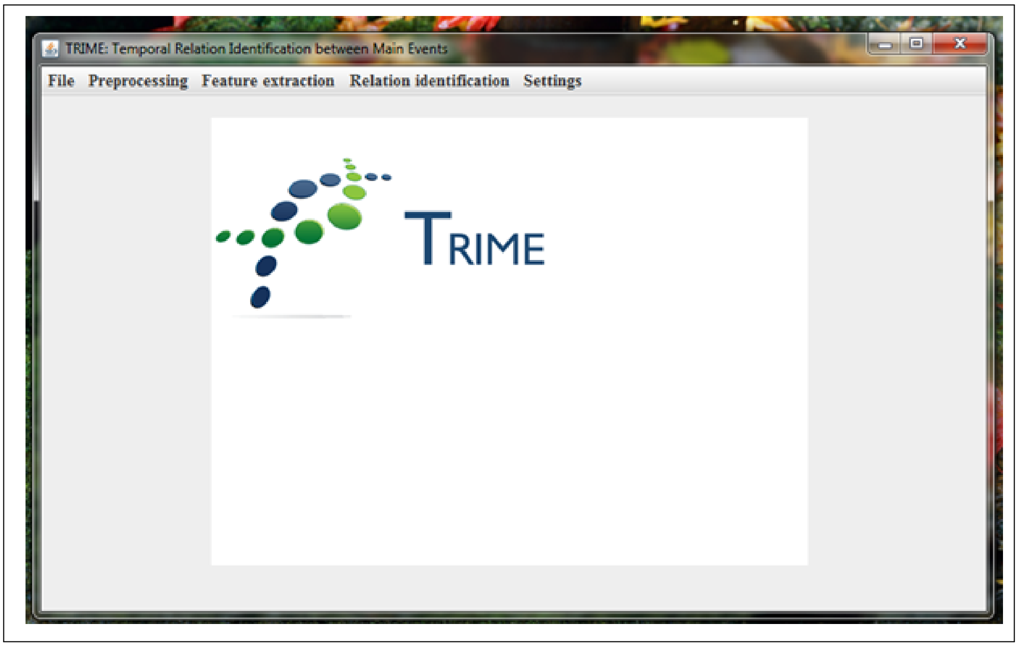


Figure 4.2 – Main Graphical User Interface

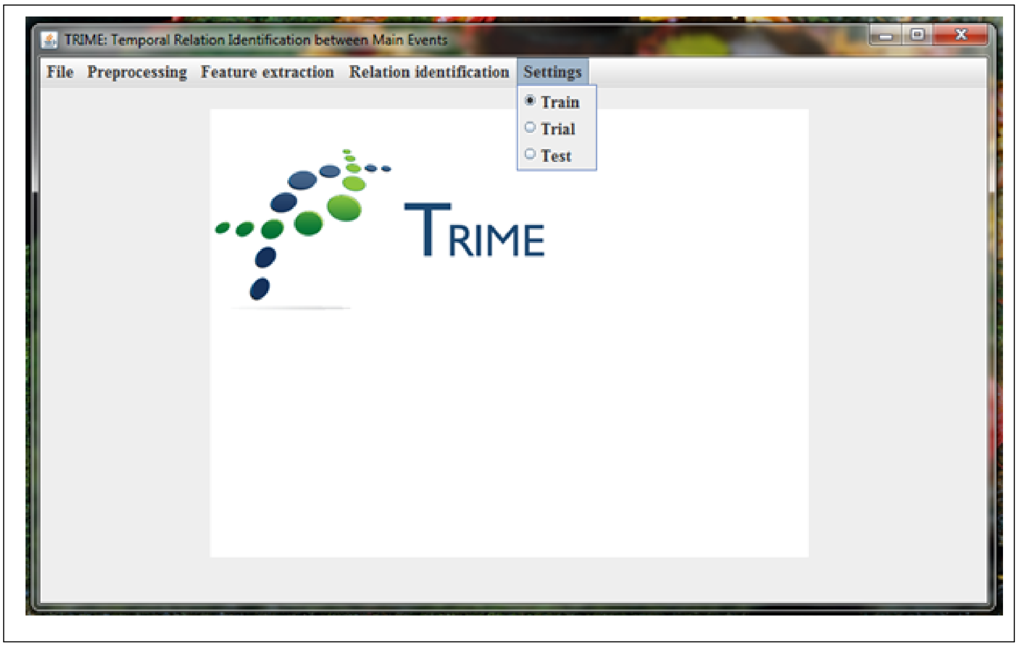


Figure 4.3 – Setting options

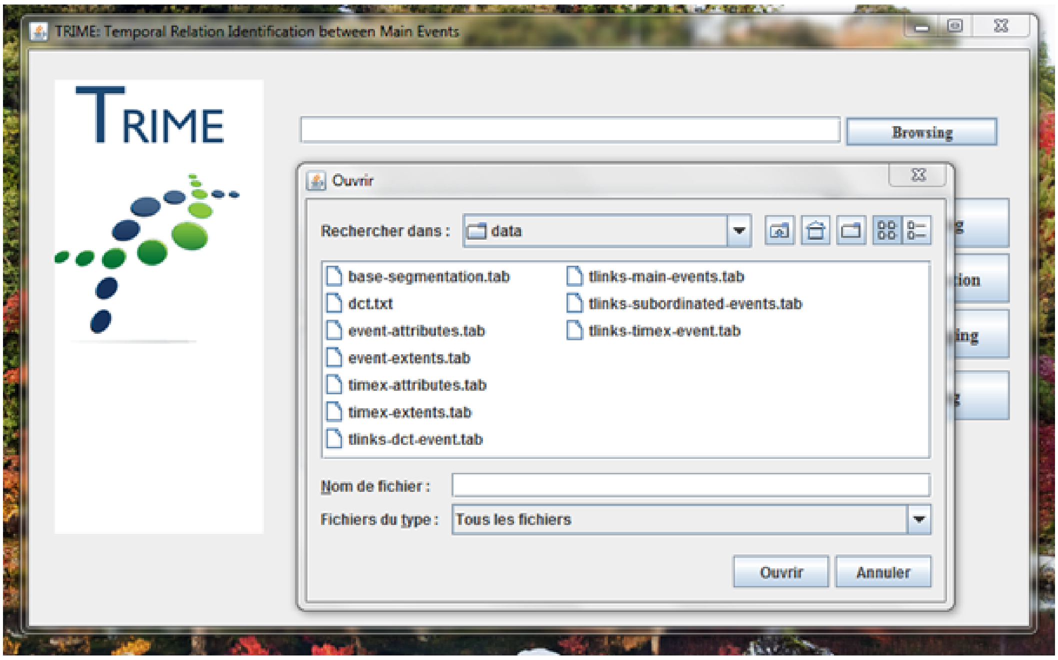


Figure 4.4 – Selection of text file

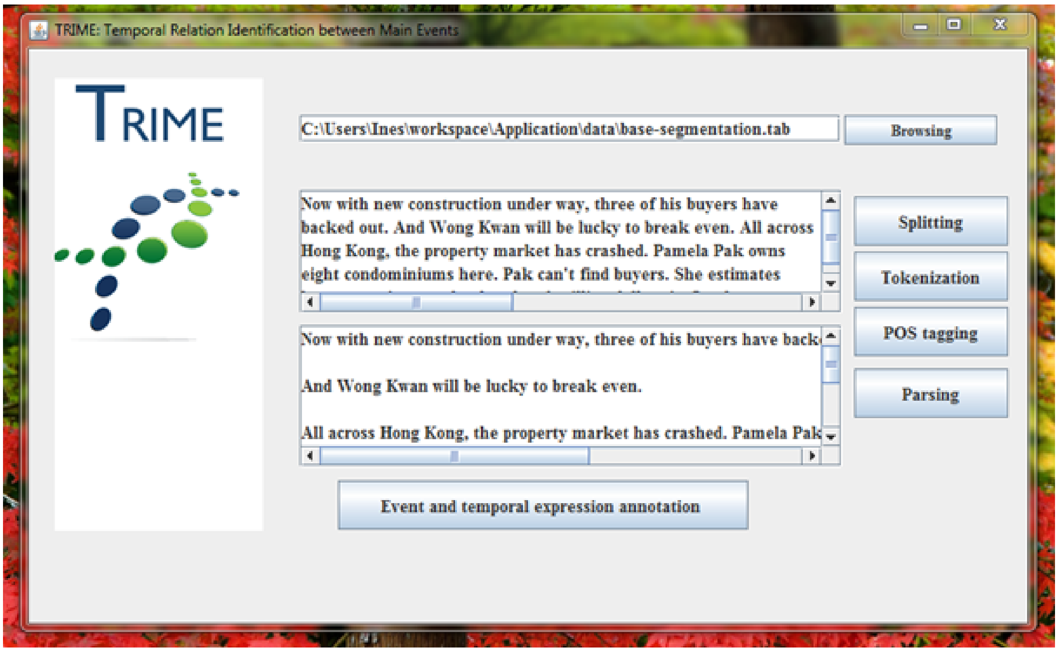


Figure 4.5 – Sentence Splitting

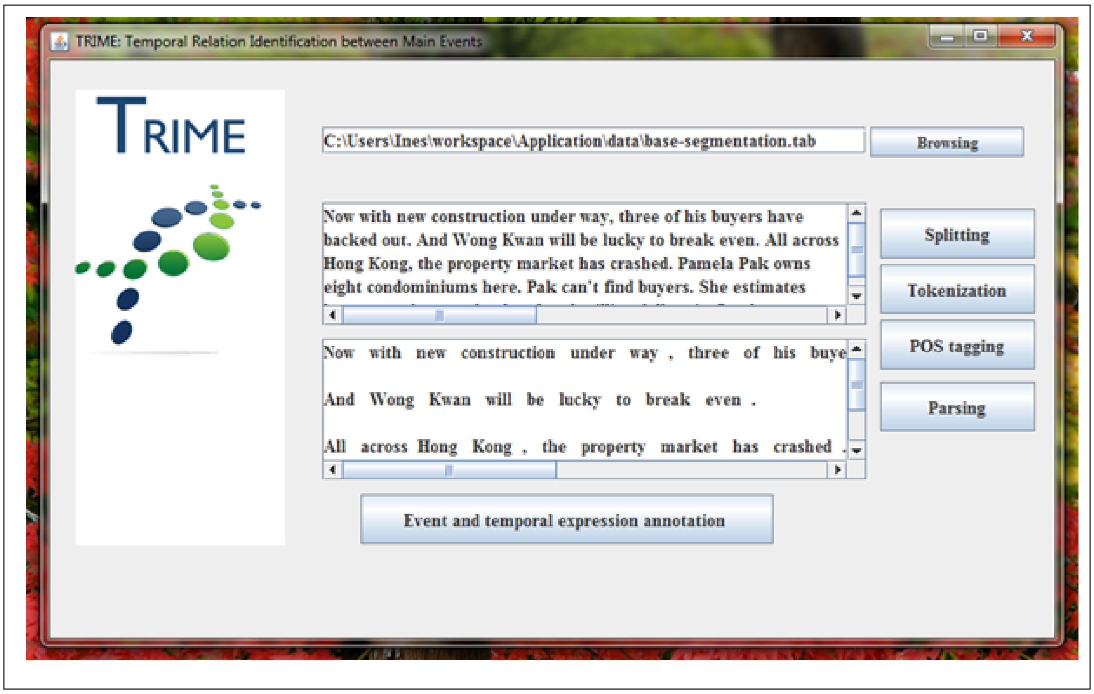


Figure 4.6 – Tokenization

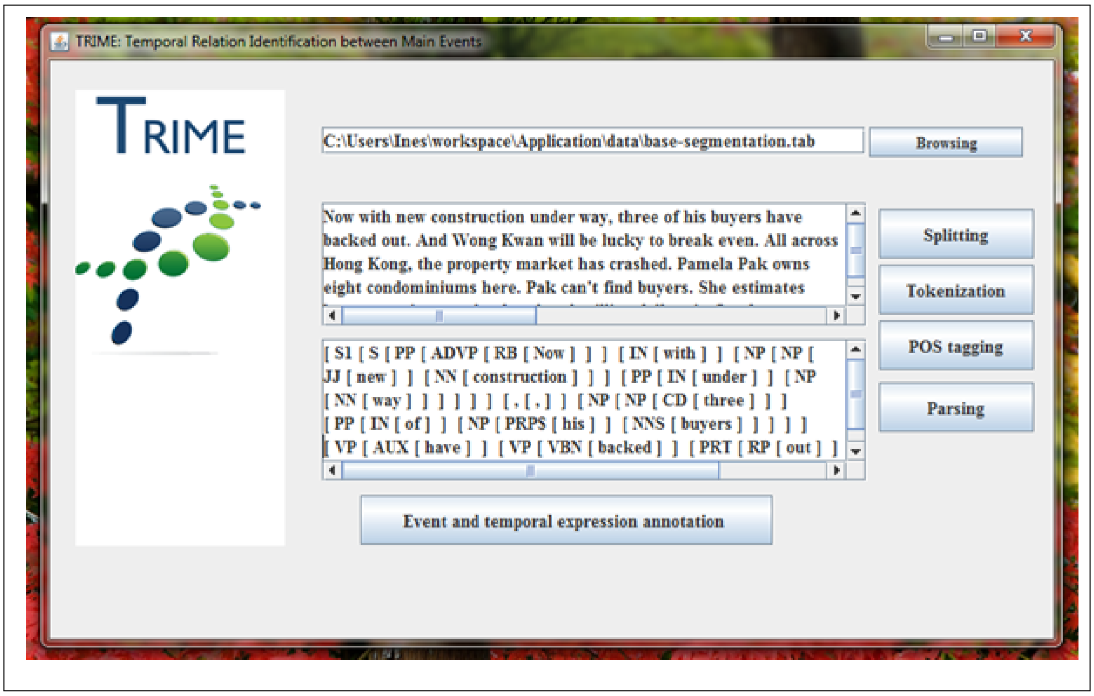


Figure 4.7 – POS tagging

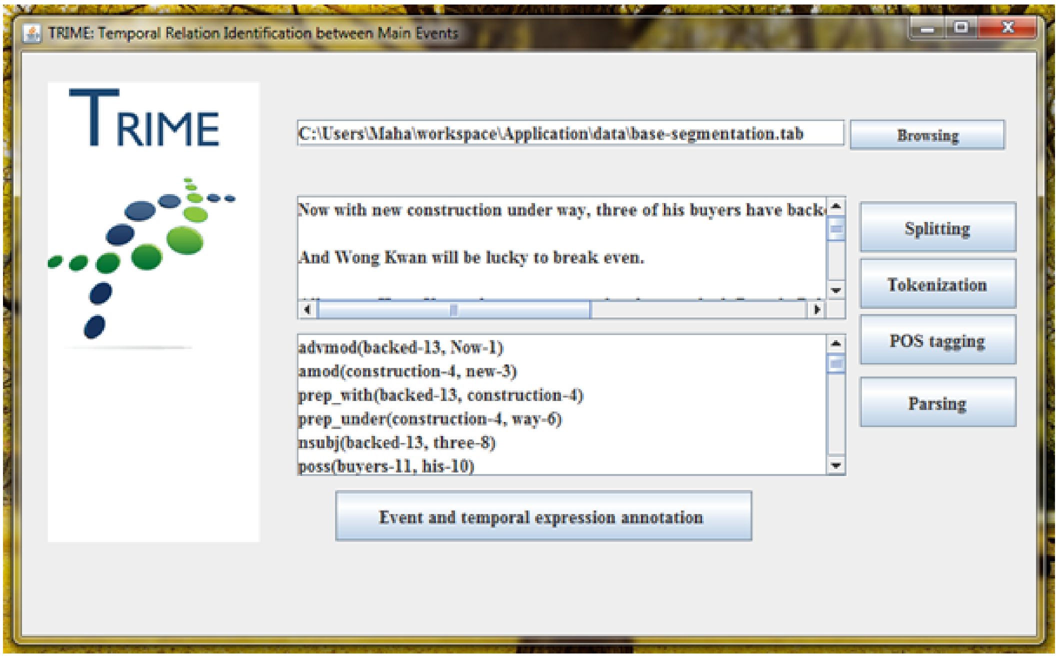


Figure 4.8 – Parsing

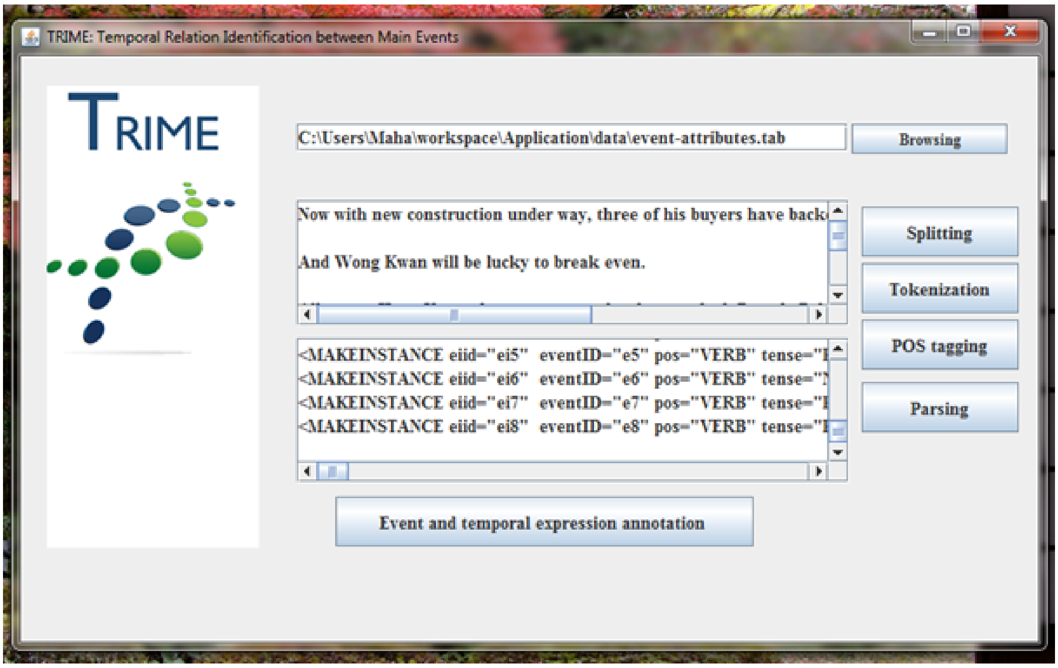


Figure 4.9 – Event and Temporal Expression annotation

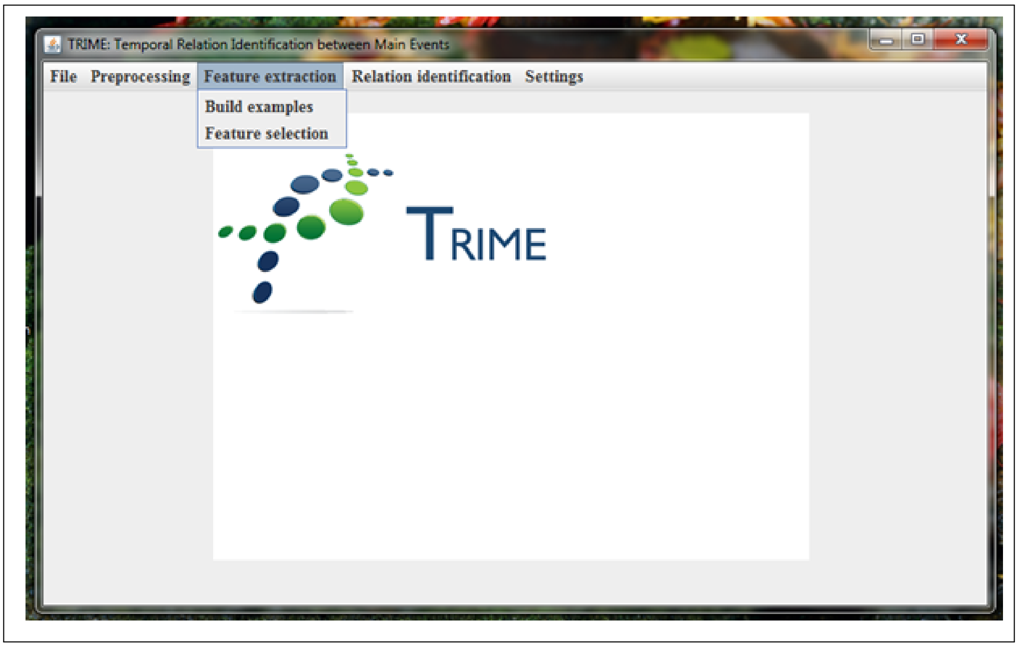
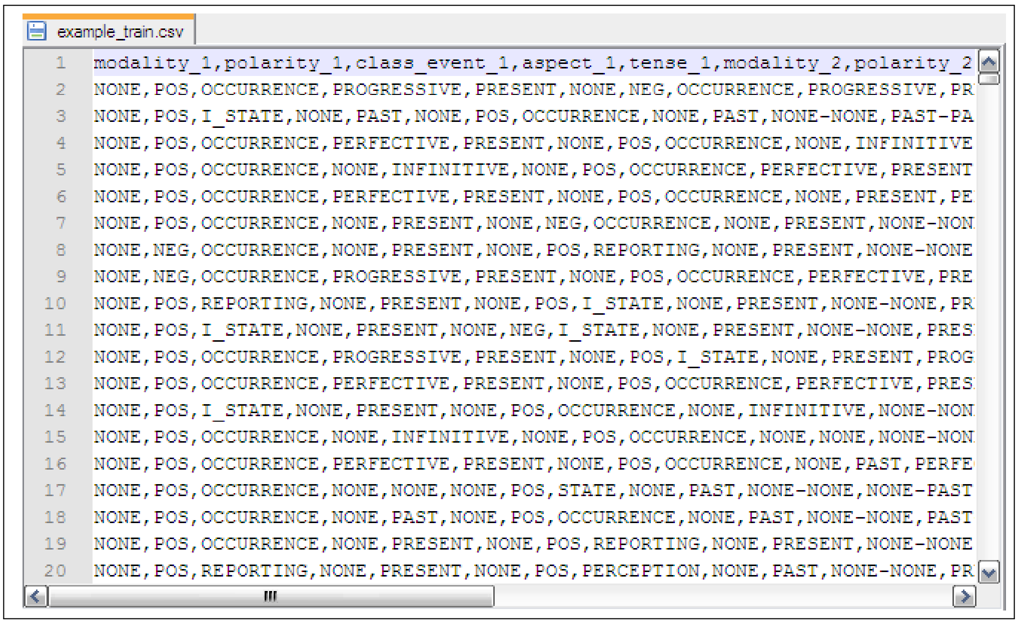


Figure 4.10 – Feature extraction



```
modality_1,polarity_1,class_event_1,aspect_1,tense_1,modality_2,polarity_2
NONE,POS,OCCURRENCE,PROGRESSIVE,PRESENT,NONE,NEG,OCCURRENCE,PROGRESSIVE,PR
NONE,POS,I_STATE,NONE,PAST,NONE,POS,OCCURRENCE,NONE,PAST,NONE-NONE,PAST-PA
NONE,POS,OCCURRENCE,PERFECTIVE,PRESENT,NONE,POS,OCCURRENCE,NONE,INFINITIVE
NONE,POS,OCCURRENCE,NONE,INFINITIVE,NONE,POS,OCCURRENCE,PERFECTIVE,PRESENT
NONE,POS,OCCURRENCE,PERFECTIVE,PRESENT,NONE,POS,OCCURRENCE,NONE,PRESENT,PE
NONE,POS,OCCURRENCE,NONE,PRESENT,NONE,NEG,OCCURRENCE,NONE,PRESENT,NONE-NON
NONE,NEG,OCCURRENCE,NONE,PRESENT,NONE,POS,REPORTING,NONE,PRESENT,NONE-NONE
NONE,NEG,OCCURRENCE,PROGRESSIVE,PRESENT,NONE,POS,OCCURRENCE,PERFECTIVE,PRE
NONE,POS,REPORTING,NONE,PRESENT,NONE,POS,I_STATE,NONE,PRESENT,NONE-NONE,PR
NONE,POS,I_STATE,NONE,PRESENT,NONE,NEG,I_STATE,NONE,PRESENT,NONE-NONE,PRES
NONE,POS,OCCURRENCE,PROGRESSIVE,PRESENT,NONE,POS,I_STATE,NONE,PRESENT,PROG
NONE,POS,OCCURRENCE,PERFECTIVE,PRESENT,NONE,POS,OCCURRENCE,PERFECTIVE,PRES
NONE,POS,I_STATE,NONE,PRESENT,NONE,POS,OCCURRENCE,NONE,INFINITIVE,NONE-NON
NONE,POS,OCCURRENCE,NONE,INFINITIVE,NONE,POS,OCCURRENCE,NONE,NONE,NONE-NON
NONE,POS,OCCURRENCE,PERFECTIVE,PRESENT,NONE,POS,OCCURRENCE,NONE,PAST,PERFE
NONE,POS,OCCURRENCE,NONE,NONE,NONE,POS,STATE,NONE,PAST,NONE-NONE,NONE-PAST
NONE,POS,OCCURRENCE,NONE,PAST,NONE,POS,OCCURRENCE,NONE,PAST,NONE-NONE,PAST
NONE,POS,OCCURRENCE,NONE,PRESENT,NONE,POS,REPORTING,NONE,PRESENT,NONE-NONE
NONE,POS,REPORTING,NONE,PRESENT,NONE,POS,PERCEPTION,NONE,PAST,NONE-NONE,PR
```

Figure 4.11 – Generated examples

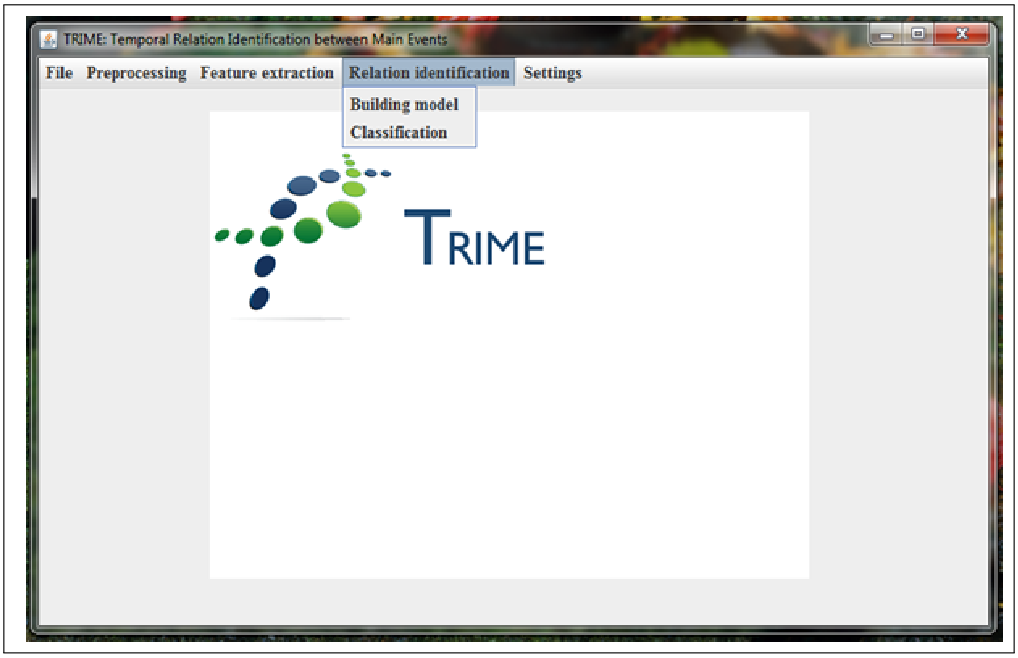


Figure 4.12 – Temporal Relation Identification

4.4 Evaluation

In this section, we describe the experiments we conducted to validate our approach. We present the experimental data, the metrics used to evaluate the performance of our system and we give the obtained results.

4.4.1 Experimental Data

We trained and tested our approach using the data set provided in the context of the TempEval-2 evaluation campaign[6]. This choice enables the comparison of our results to those obtained by other systems in this recent evaluation exercise. Both training and test sets are built from the TimeBank 1.2 corpus (Pustejovsky and al., 2003b). TimeBank 1.2 is a collection of 183 news articles collected from a several sources namely the Automatic Context Extraction (ACE) program (NIST, 2007)[7] , and PropBank

[6] http://timeml.org/site/timebank/timebank.html

[7] NIST (2007).The ACE 2007 (ACE07) Evaluation Plan.National Institute of Standardsand Technology.

Set	Docs	Words	Element
Train	162	53K	TIMEX (1052)
			EVENT (5688)
			TLINK event-timex (959)
			TLINK event-dct (640)
			TLINK main-events (1587)
			TLINK subordinated-events (1721)
Test (entities)	9	5K	TIMEX (81)
			EVENT (498)
Test (relations)	11	5K	TLINK event-timex (65)
			TLINK event-dct (190)
			TLINK main-events (137)
			TLINK subordinated-events (140)

Table 4.1 – TimeML English data sets (TempEval-2)

(Kingsbury and Palmer, 2000). The corpus is annotated according to the TimeML 1.2.1 specification . All annotated documents in the corpus were validated against a Document Type Definition (DTD) and XML schema. The Table 4.1 summarizes the statistics of the experimental data sets.

As shown in table 4.1, the number of main events instances linked with a TLINK tag are respectively 1587 and 137 in the training and testing data sets.

4.4.2 Metrics

The performance of our system is evaluated using the standard Precision, Recall and F1-score metrics. The Precision measures how often the system is correct when it outputs a temporal relation. It is calculated by dividing the number of correct outputs (true positive, TP) by the total number of the outputs. The total number of the outputs is the number of correct outputs plus the number of incorrect outputs (false positive

FP).

$$Precision = \frac{|TP|}{|TP| + |FP|} \tag{4.1}$$

The Recall measures how often the system correctly finds the right classes to output. It is defined as proportion of true positives against potential correct outputs. The total number of potential correct outputs is the number of correct output (true positive, TP) plus the count of objects that should have been outputed but where not (false negative, FN).

$$Recall = \frac{|TP|}{|TP| + |FN|} \tag{4.2}$$

The F1-score attempts to balance the contributions of precision and recall to system performance.

$$F - score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{4.3}$$

4.4.3 Results

In this subsection, we first present the inter anotator agreement results as well as the official results obtained by participating systems in the TempEval-2 evaluation compaign. These results are provided in the scope of the evaluation campaign. Then we present the results that we have obtained.

Table 4.2 and Table 4.3 report respectively the human inter-annotator agreement in the TimeML elements recognition (exact match); and the average precision, recall (P&R) and kappa for the TimeML attributes, namely the temporal expressions types and values, the events classes and the temporal links types.

These values were only calculated over the TimeBank corpus annotation , but since TimeBank and TempEval-2 corpus are mostly the same data, we give these values as indirect assessment.

TimeML tag	agreement
TIMEX	0.83
EVENT	0.78
TLINK	**0.55**

Table 4.2 – TimeML English data inter-annotator agreement

Table 4.2 shows that the inter-annotator agreement score for temporal expressions is the best (0.83) compared to events and temporal links. TLINKs inter-annotator agreement score is quite low (0.55). This fact is due to the large number of event-pairs that can be selected for specifying temporal links. Thus two different annotators can annotate the temporal relations differently.

TimeML tag	P&R	Kappa
TIMEX.type	1.00	1.00
TIMEX.value	0.90	0.89
EVENT.class	0.77	0.67
TLINK.relType	**0.77**	**0.81**

Table 4.3 – P&R and Kappa for TimeML English attributes

The values presented in Table 4.3 indirectly suggest the complexity of the different tasks, in terms of human annotation agreement for temporal attributes. From this data, it can be observed that, for human annotators, event and temporal relation processing is more complex than timex processing.This can be explained by the fact that events and temporal relations depend on various factors, which may contribute to the ambiguity of interpretation.

After having presented the human inter annotator agreement for temporal elements and their corresponding attribures, we present in Table 4.4 the TempEval-2 official F-score=1 scores for the participating systems in the temporal relation identification task between main events.

System	Precision	Recall	F-score=1
TRIOS	0.56	0.42	0.56
TIPSem			0.55
TIPSem-B			0.55
NCSU	0.48	0.48	0.48
JU CSE			0.56
USFD2			0.45

Table 4.4 – TempEval-2 official results for temporal relation identification between main events

As shown in Table 4.4, 6 systems particiated in the TempEval-2 compagian for temporal relation identification between main events. F-score values range from 0.45 to 0.56.

Table 4.5 and table 4.6 show respectively our results from the Naïve Bayes classifier and the ID3 classifier with and without the discourse feature.

TP Rate	**FP Rate**	**Precision**	**Recall**	**F-Measure**	**Class**
0.716	0.44	0.575	0.716	0.638	OVERLAP
0.528	0.159	0.505	0.528	0.516	BEFORE
0.298	0.068	0.477	0.298	0.367	AFTER
0	0.012	0	0	0	OVERLAP-OR-AFTER
0.031	0.009	0.222	0.031	0.055	VAGUE
0.222	0.05	0.13	0.222	0.164	BEFORE-OR-OVERLAP
0	0.002	0	0	0	UNKNOWN
0.51	**0.252**	**0.484**	**0.51**	**0.484**	**Weighted Avg**

Table 4.5 – Our results from the Naive Bayes classifier

TP Rate	FP Rate	Precision	Recall	F-Measure	Class
0.728	0.305	0.669	0.728	0.697	OVERLAP
0.556	0.205	0.463	0.556	0.505	BEFORE
0.59	0.071	0.622	0.59	0.605	AFTER
0	0.019	0	0	0	OVERLAP-OR-AFTER
0.103	0.02	0.316	0.103	0.156	VAGUE
0.053	0.007	0.167	0.053	0.08	BEFORE-OR-OVERLAP
0	0.001	0	0	0	UNKNOWN
0.575	**0.203**	**0.551**	**0.575**	**0.556**	**Weighted Avg**

Table 4.6 – Our results from the ID3 classifier

4.5 Conclusion

In this chapter, we presented the design and implementation of our Temporal Relation Identification system TRIME and we reported the obtained results. For this purpose, we processed the data set with a pipeline of NLP tools. Then, we implemented methods to extract syntactic, semantic and discourse features. After that, we used two supervised classifiers namely Naïve Bayes model and a Decision Tree model to classify the temporal relation between main events. Our experimentation study shows that our event extraction approach gives motivating results compared against the state-of-the-art benchmarks.

Conclusion and Future Works

Processing the temporal dimension of natural language is essential in many Natural Language Processing applications, such as Question Answering, Summarisation and Information Retrieval.

In this work we have proposed a new approach and a system for temporal relation identification between main events in Natural Language texts. In fact, we have considered this task as a classification problem where the aim is to identify the temporal class between main events in two consecutive sentences. Our contribution in this work is automatically applying all linguistic analysis levels on textual contents from morphological analysis to pragmatic analysis.

Along this work, we have achieved the following goals: We have reviewed how temporal information is conveyed in natural language. Then, we have overviewed existing approaches in Temporal Information Processing. For this purpose, we have presented existent resources to handle such issue (namely temporal annotation schemes and annotated corpora), some computational approaches proposed to perform different temporal processing tasks as well as some real-world applications of Temporal Information Processing.

In a second time, we have focused on the Temporal Relation Identification between main events. We have presented the basic concepts and the computational tasks for temporal relation identification. Then, we have discussed related works to our proposal. After that, we have proposed a novel approach for the identification and the classification of main events in two consecutive sentences in textual contents. We have started

by preporcessing texts by appliying several NLP techniques. Then we have extracted a large set of features at all linguistic analysis levels. In the last stage, we have applied Machine-Learning techniques to classify the temporal relations between each main event pair.
In order to evaluate our approach, we have built a system with the aim to be fast, simple and robust. We have also aimed to ensure the applicability of our system to various types of texts in different domains, even though we have trained and tested it on news articles.

Our proposal achieved good results when compared to other systems performing the same task. However, several factors have proved the complexity of the task mainly the low level of human annotators agreement, and the low results obtained by all existant systems, when compared to ohter tasks proposed in the scope of the TempEval campaigns.

Future research directions Our first future work would be to adapt and evaluate our approach on corpora writen in other languages and/or belonging to other genres. Another line of research would be to exploit the wealth of theoretical works on the temporal structure of discourse. At this stage, we would investigate the integration of new mechanims namely the anaphoric resolution at discourse level. For example, we would study the role that temporal adverbials may play to infer the right class of temporal relations between events; Further work stemming from this research involves specific tasks that would improve the functionality of the system described in this thesis, or wider applications that would use this system to address more complex NLP problems.

An essential stage in finding the temporal relation between two temporal entities is detecting the temporal relation that holds between a pair of clauses involved in a syntactic relation. Existing connections between syntax and temporality need to be further investigated at inter-clausal level. For each type of syntactic relation that can hold between two clauses, it would be interesting to extract from a corpus pairs of clauses involved in that relation, and to analyze the correlations that can be identified between the syntactic properties of the two clauses combined with the syntactic relation holding between them and the temporal relation that can be established between the main events of the two clauses. This analysis could suggest improvements to the module that solves the task of inter-clausal temporal ordering.

Appendix A

Linguistic Time mechanisms

Tense and aspect are the most important grammatical categories for representing time and temporal relations in English. In this section we present a brief review on classical works based on linguistic theories that researchers explored for studying and representing tense and aspect (Lyons, 1981).

A.1 Tense

Tense is a temporal linguistic mechanism that expresses the time at which or during which an event takes place. In this purpose, Reichenbach (1947) develops a theory in which tense gives information about the following times: Speech Time (S): the time related to the time point of the speaking act Event Time (E): the time at which the told event happens This two times points let express the basic tense classes with use of the operators of precedence (<) and simultaneity (=).

a. *I played football:* Past tense. (E < S)

b. *I play football:* Present tense. (E= S)

c. *I will play football:* Future tense. (S < E)

However, combinations between these two times points are unable to express all tenses. Consider the following examples:

a. *I played football.*

b. *I had played football.*

Although both refer to events in the past (E < S), representing them in the same way seems incorrect. In example (I had played football) the event of playing seems to refer to another event. To handle this situation, Reichenbach introduce a third temporal point into his model which is the Reference Time (R). According to this extension, the Event Time (E) In example (I had played football) is the time at which I played, and the Reference Time (R) is between the Event Time (E) and the Speech Time (S) "(E <R < S)". With the three points defined by Reichenbach, it's possible to represent all the tenses using a set of relations between these points. Nevertheless, Reichenbach relations still ambiguous, so Song and Cohen (1991) develop an unambiguous set of relations including a new operator (>) and presenting the relations always in S-R-E order. The following table illustrates the original Reichenbach's relations, the unambiguous relations proposed by Song and Cohen, and their mapping to English tenses.

Reichenbach relations	**unambiguous relations**	**English tense**	**example**
E<R<S (Anterior Past)	S>R>E	Past Perfect	*I had played*
E=R<S (Simple Past)	S>R=E	Past Simple	*I played*
R<E<S* (Posterior Past)	S>R<E		*[I expect]I would play*
R<S=E*			
R<S<E*			
E<S=R (Anterior Present)	S=R>E	Present Perfect	*I have palyed*
S=R=E (Simple Present)	S=R=E	Present Simple	*I play*
S=R<E (Posterior Present)	S=R<E	Future Simple	*I will play*
S=E<R*			
E<S<R*			
S<E<R* (Anterior Future)	S<R>E	Future Perfect	*I will have played*
S<R=E (Simple Future)	S<R=E	Future Simple	*I will play*
S<R<E (Posterior Future)	S<R<E		*I sahll be going to play*

Table A.1 – Tense Temporal Relations

* ambiguous relations It's worth noting that tense is a key concept in the identification of temporal relation between two events, especially in cases where these events are situated in different sentences or if there's no time indicator like prepositions (after, before...) in the text.

A.2 Aspect

Aspect is the second device expressing time in natural language. Two types of aspects are expressed in language namely grammatical aspect and lexical aspect. Grammatical aspect expresses the viewpoint from which a particular eventuality is described, it indicates the phase in which an eventuality is to be perceived. While lexical aspect, distinguishes between different subclasses of events based on its following temporal properties: *dynamicity*, *telicity* and *durativity.* In the literature of Aspect, Vendler's work (Vendler 1967) has been the basis for subsequent researchers.Vandler propose an initial distinction between events and states and then classifies the event expressions into three aspectual categories or Aktionsarten: activities, accomplishments, and achievements.

a. **Activities** are events which are durative, or extended in time, but that do not involve an explicit end point. In other words, activities are *durative* and *atelic* events.

b. **Accomplishments** (durative culminated process): Events that imply a duration with a definite end point in which a state changes.

c. **Achievements** (non-durative culminations): Punctual or instantaneous events that do not imply a duration, happening at a defined point in which a state changes.

Even though works addressing theoretical linguistics concepts discussed in this section helped to illuminate complex problems related to temporal relation identification, these works were criticized because they were not analyzed and evaluated over real linguistic data. That's why, in the mid-90's, the field has experienced a considerable mutation from theoretical works on Natural Language Processing to computational ones especially with the availability of digitalized texts (Manning and Schutze, 1999).

Appendix B

Temporal Reasoning Models

B.1 Allen's Interval Algebra

The most influential work to capture the temporal dimension of a narrative is the Allen's Interval Algebra (Allen, 1983). This algebra has inspired many researches to develop temporal reasoning models.
In fact, James Allen developed an Interval Algebra which provides a conceptual model of time that captures the different ways in which eventualities may be related to each other.
Allen considers that every temporal expression or event can be presented as a temporal interval having a start point and an end point on a timeline. In this concern, Allen defines a set of thirteen basic (binary) interval relations, where six are inverses of the other six, excluding equality: **equals** (=), **before** (<), **after** (>), **meets** (m), **met by** (mi), **overlaps** (o), **overlapped by** (oi), **starts** (s), **started by** (si), **finishes** (f), **finished by** (fi),**during** (d), **contains** (di).
The following figure represents those intervals:

Allen represents the temporal structure of his algebra in a network, where the nodes represent individual intervals and the arcs represent the relationship between them. The following figure illustrates this network.

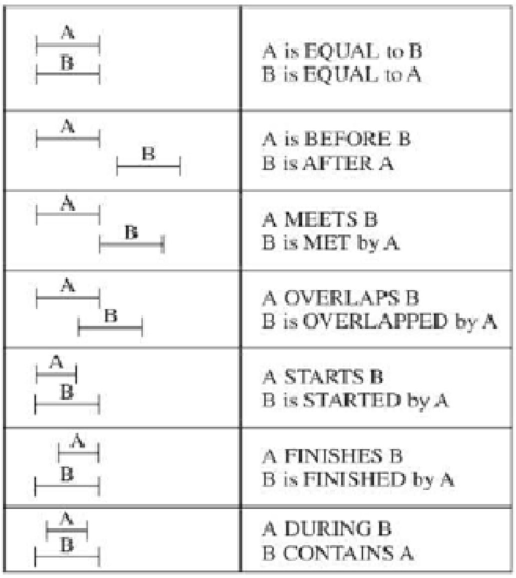


Figure B.1 – Allen's Temporal Interval Relations

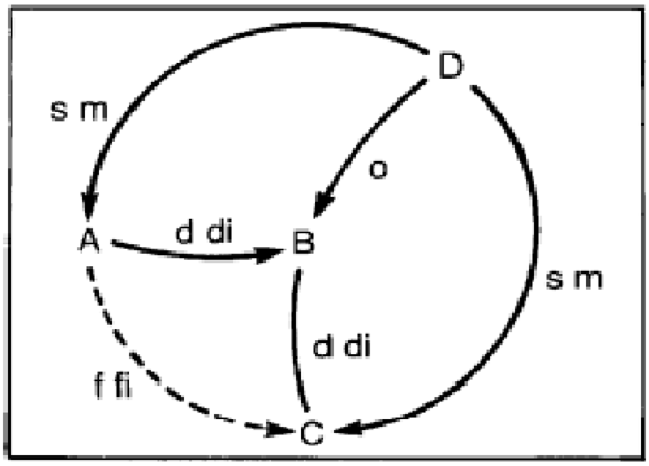


Figure B.2 – Allen's interval-relation network

B.2 Time stamping of events

Filatova and Hovy (2001) also propose a model for temporal relations representation known as *the time stamping of events.* This model serves for arranging the contents of news stories into a time-line. This procedure consists on assigning a calendrical time point or interval to all events in a text.
However, this model doesn't capture information in many cases and sometimes loses information or misinterpret.
The following example and scheme due to Mani and al., (2005) explain this concept and its limitation.
"After the plane crashed, a search was begun. Later the coastguard reported finding debris."

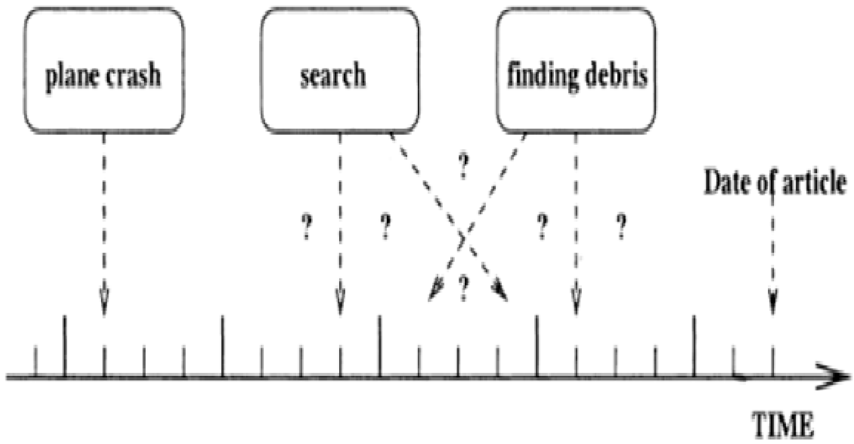


Figure B.3 – A Time-Line Representation

Although we can place the crashing event in a timeline, we can't place the two other events. We either have to guess the time-points, or assign an interval. Guessing the time-point is not an option and if we assign an interval, then for both searching and finding debris the interval is from the crash till the date of the article. In this case, we lose the ordering between searching and finding debris.

B.3 Time-Event Graph Representation

Setzer and Gaizauskas (2002) propose a simpler representation able to capture more information. It's a *Time-Event Graph Representation* used for identifying event-event relationship with event-time relationship. This representation is based on *the Allen algebra.* In fact, after reducing all events and temporal expressions to intervals and after identifying the temporal relations between them, the temporal information in a text can be represented as a graph where events and TEs form the nodes, and the edges are labeled with the temporal relations between them.

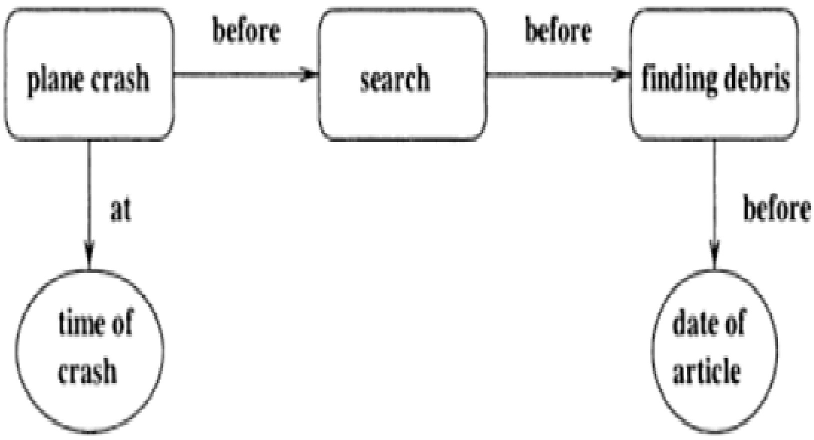


Figure B.4 – A Time-Graph Representation

Appendix C

Annotation Schemes

The most outstanding temporal annotation schemes are: MUC-TIMEX (Grishman and Sundheim, 1996), TIDES (Ferro, Mani, Sundheim, and Wilson, 2000), STAG (Setzer and Gaizauskas, 2000), and TimeML (Pustejovsky, Castano, Ingria, Sauri, Gaizauskas, Setzer, and Katz, 2003). All of them follow a SGML/XMLbased annotation format. These schemes are described below in chronological order, highlighting the novelties each one introduced to its predecessor.

C.1 1995/1997 - MUC-TIMEX

The earliest annotation scheme was created at the 6th DARPA's Message Understanding Conference (MUC-6) (Grishman and Sundheim, 1996). This first temporal scheme focused on the annotation of temporal adverbials and temporal phrases representing explicit dates (e.g., November 2010) and times (e.g., 7 a.m.). The MUC's participants were asked to build systems able to mark these expressions with the TIMEX tag and to indicate their type (DATE or TIME) over a textual corpus. The outputs of participating systems were compared with a manually annotated gold standard in order to evaluate their performance. The following example shows the TIMEX tag: He was born on <TIMEX type="DATE">March 1st, 1980</TIMEX> Afterwards in 1997 at MUC-7 (Gaizauskas and Wilks, 1998), the TIMEX recognition task included the relative temporal expressions5 (e.g., yesterday, two years ago). This first annotation scheme was very simple and it was limited to the identification of temporal expressions in texts.

C.2 2000 - TIDES- TIMEX2

The Translingual Information Detection, Extraction, and Summarization or TIDES scheme (Ferro et al., 2000; Wilson et al., 2001; Ferro et al., 2005) was developed under the support of the DARPA and ACE. This new multilingual annotation scheme replaced the TIMEX tag by TIMEX2 and introduced new types of temporal expressions namely durations (e.g., for two years) and sets (e.g., monthly). Absolute (e.g., October 1st, 1999), relative (e.g., yesterday), period (e.g., two years), and set (e.g., weekly) temporal expressions must be normalized in a new attribute VAL, following the ISO 8601 standard. This schema also added new attributes to capture the semantics of timexes namely MOD (captures temporal modifiers), ANCHOR VAL (contains a normalized form of an anchoring date/time), ANCHOR DIR (captures the relative time direction between VAL and ANCHOR VAL) and SET (identifies expressions denoting sets). The following example shows the TIMEX2 tag:

A rocket was launched <TIMEX2 VAL="1999-10">in October, 1999</TIMEX2>. Three rockets were launched <TIMEX2 VAL="P1Y" ANCHOR VAL="2000">during the next year</TIMEX2>. She visits him <TIMEX2 SET="YES" VAL="XXXX-XX-XX">daily</TIMEX2>. She knows him <TIMEX2 VAL="P1Y" MOD="LESS THAN"> for less than a year</TIMEX2>.
This annotation scheme was used in TERN (2004) evaluation forum and in EVALITA'07. It was the most important scheme (Negri and Marseglia, 2004; Saquete et al., 2006) until the adoption of TimeML as standard.

The problem of this scheme is its limitation to the annotation of temporal expressions. This excludes events and temporal relations which are important entities for temporal reasoning.

C.3 2000/2001 - Sheffield Temporal Annotation Guidelines (STAG)

STAG (Sheffeild Temporal Annotation Guidelines) is the temporal annotation language presented in Andrea Setzer's PhD thesis (Setzer, 2001) with the aim of providing a more complete temporal information annotation scheme to identify events in news, as well as their anchoring to time and their relative ordering. In fact, this scheme was motivated by contemporary works in corpus-based temporal information processing and has suggested extension of temporal information annotation from just timexes inherited from MUC and TIDES to events and temporal relations. With this new scheme, events are annotated with the **EVENT** tag which includes some properties such as event-class. Four classes of events were defined: occurrence, perception, reporting and aspectual. However, stative events were excluded from the scheme due to their complexity. This tag also includes attributes to annotate temporal relations between two entities based on Allen Algebra (Allen, 1983). Added to that, STAG provides a tag named **SIGNAL** to annotate elements that point out temporal relation holding between two temporal entities (timex-event, timex-timex, or event-event) which are denoted by prepositions (e.g., on, after, during) and conjunctions (e.g., while, when). The following example shows the STAG tags:

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 <signal sid="9">on</signal> <timex tid="5" type="DATE" calDate="12031997">Wednesday</timex>
The problem of this scheme is that it does not include stative events and that temporal relations were included in the event tag. This limitation induces a wrong annotation of complex temporal relations in some cases.

C.4 2002/2003 - Time Markup Language (TimeML)

TimeML (Pustejovsky et al., 2002, 2003) was developed under the sponsorship of ARDA as the natural evolution of STAG. It combines and extends features of preceding schemes, which makes it a more powerful annotation scheme. TimeML has recently been standardized to an ISO international standard for temporal information markup, ISO-TimeML (ISO-TimeML, 2007). Both the TimeML and the ISO-TimeML annotation standards define the following basic XML tags: <EVENT> for the annotation of events, <TIMEX3> for the annotation of time expressions, <SIGNAL> for locating textual elements that indicate a temporal relation, and the tags <TLINK>, <SLINK> and <ALINK> that capture different types of relations. TimeML scheme is distinguished from previous attempts with introducing new characteristics as follows:

EVENT tag: The EVENT tag is inherited from STAG and extended with three new classes. Thus, TimeML defined seven classes of events: Reporting, Perception, Aspectual, I_Action, I_State, State and Occurrence.
Also, The STAG attributes used to capture temporal relation are removed in TimeML schema given that temporal relations are represented in separated. The specification of EVENT is shown below:

attributes ::= eid class tense aspect
eid ::= EventID
EventID ::= e<integer>
class ::= 'OCCURRENCE' | 'PERCEPTION' | 'REPORTING' | 'ASPECTUAL' | 'STATE'| 'I_STATE' | 'I_ACTION' | 'MODAL'
tense ::= 'PAST' | 'PRESENT' | 'FUTURE' | 'NONE'
aspect ::= 'PROGRESSIVE' | 'PERFECTIVE' | 'PERFECTIVE_PROGRESSIVE' | 'NONE'

TIMEX3 tag: Timexes were classified by TimeML into four types: Date, Time, Duration and Set. Furthermore, to facilitate the computational interpretation of temporal expressions, a new attribute (value) is added to capture their ISO 8601 normalization. The following table shows an example of the four TIMEX3 types and their normalization values.

TIMEX3	**Type**	**Normalized value**
Jun 2012	Date	2012-06
Tomorrow at 5 a.m	Time	depends on the DCT*
Five days	Duration	P5D
Monthly	Set	XXXX-XX

Table C.1 – Examples of TIMEX3 tags

*DCT: Document Creation Time.

The specification for TIMEX3 is given below:

attributes::=tid type[functionInDocument][temporalFunction]
(value| valueFromFunction)[mod] [anchorTimeID | anchorEventID]
tid ::= TimeID
TimeID ::= t<integer>
type ::= 'DATE' | 'TIME' | 'DURATION'
functionInDocument::='CREATION_TIME'|'EXPIRATION_TIME'|MODIFICATION
_TIME'| 'PUBLICATION_TIME' |'RELEASE_TIME'| 'RECEPTION_TIME' | 'NONE'
temporalFunction ::= 'true' | 'false'
temporalFunction ::= boolean
value ::= CDATA
value ::= duration|dateTime|time|date|gYearMonth|gYear|gMonthDay | gDay | gMonth
valueFromFunction ::= IDREF
valueFromFunction ::= TemporalFunctionID
TemporalFunctionID ::= tf<integer>
mod ::= 'BEFORE'|'AFTER'|'ON_OR_BEFORE'|'ON_OR_AFTER'|'LESS_THAN'|

’MORE_THAN’|’EQUAL_OR_LESS’|’EQUAL_OR_MORE’ | ’START’|’MID’
|’END’| ’APPROX’
anchorTimeID ::= TimeID
anchorEventID ::= EventID

SIGNAL tag: this tag is used to markup textual elements expressing a relation between two temporal entities (timex and event, 2 events or 2 timexes). In fact, a signal can be a preposition (on, in, at, from, to, before, after, during), a conjunction (when, while, before, after) or a special character used in time ranges (- or /). The specification for SIGNAL is given below:
attributes ::= sid
sid ::= ID
sid ::= SignalID
SignalID ::= s<integer>
TLINK tag: represents the temporal relationship holding between events or between an event and a time, and establishes a link between the involved entities, making explicit if they are. The specification for TLINK is given below:
attributes ::= (eventInstanceID | timeID) [signalID] (relatedtoEvent | relatedtoTime) relType [magnitude]
eventInstanceID ::= ei<integer>
timeID ::= t<integer>
signalID ::= s<integer>
relatedToEvent ::= ei<integer>
relatedToTime ::= t<integer>
relType ::= ’BEFORE’ | ’AFTER’ | ’INCLUDES’ | ’IS_INCLUDED’ | ’HOLDS’ ’SIMULTANEOUS’ |
’IAFTER’ | ’IBEFORE’ | ’IDENTITY’ | ’BEGINS’ | ’ENDS’ | ’BEGUN_BY’ | ’ENDED_BY’
magnitude ::= t<integer>

SLINK tag: defines a relation between two events in a syntactic subordination Relation. An SLINK is of one of the following sorts: modal, factive, counter-factive, vidential, negative evidential, or conditional. The specification for the SLINK relation is given below:

attributes ::= [eventInstanceID] (subordinatedEvent |
subordinated EventInstance) [signalID] relType [polarity]
eventInstanceID ::= ei<integer>
subordinatedEvent ::= e<integer>
subordinatedEventInstance ::= ei<integer>
signalID ::= s<integer>
relType ::= 'MODAL'| 'NEGATIVE' | 'EVIDENTIAL' | 'NEG_EVIDENTIAL'| 'FACTIVE'| 'COUNTER_FACTIVE'

ALINK tag: or Aspectual Link represents the relationship between an aspectual event and its argument event. The ALINKs are of one of the following types: initiation, culmination, termination, continuation or reinitiation. The specification for the ALINK relation is given below:
attributes ::= eventInstanceID [signalID] relatedToEvent relType
eventInstanceID ::= ei<integer>
signalID ::= s<integer>
eventID ::= e<integer>
relType ::= 'INITIATES'| 'CULMINATES' | 'TERMINATES'| 'CONTINUES'
TimeML has been widely accepted as the most important markup language for time. This scheme is used in many applications that require access to the temporal information embedded in text namely to annotate the temporal information in large textual corpora.
Currently, TimeML is the standard annotation scheme for temporal information processing because it is the most complete scheme compared to its predecessors.

Bibliography

Allen, J. Maintaining knowledge about temporal intervals. *Commun. ACM*, 26(11): 832–843, ACM, November 1983. ISSN 0001-0782.

Anyanwu, N. and Shiva, S. Comparative analysis of serial decision tree classification algorithms. *International Journal of Computer Science and Security (IJCSS)*, 3, 2009.

Aone, C. and Ramos-Santacruz, M. Rees: A large-scale relation and event extraction system. In *the 6th Applied Natural Language Processing Conference*, pages 76–83. ANLP-2000, 2000.

Asher, N. *Reference to abstract objects in discourse.* Kluwer Academic Publishers, Dordrecht, 1993.

Bach, E. *The Algebra of Events*, chapter 9, pages 5–16. Linguistics and Philosophy, 1986.

Bethard, S. and Martin, J. H. Cu-tmp: Temporal relation classification using syntactic and semantic features. In *Fourth International Workshop on Semantic Evaluations*, pages 129–132, Prague, Czech Republic, 2007. SemEval.

Bittar, A. *Building a TimeBank for French: A Reference Corpus Annotated According to the ISO-TimeML Standard.* PhD thesis, Université Paris Diederot EFR de Linguistique, école doctorale de sciences du langage, laboratoire ALPAGE, 2010.

Califf, M. and Mooney, R. Bottom-up relational learning of pattern matching rules for information extraction. *The Journal of Machine Learning Research*, 4:177–210, JMLR.org, 2003.

Chambers, N., Wang, S., and Jurafsky, D. Classifying temporal relations between events. In *the 45th Annual Meeting on Association for Computational Linguistics*, 2007.

Chinchor, N. A. Overview of muc-7. In *the 7th Message Understanding Conference (1998) Key: citeulike:2209778*, 1998.

Chklovski, T. and Pantel, P. Verbocean: Mining the web for fine-grained semantic verb relations. In *In Dekang Lin and Dekai Wu, editors Proceedings of EMNLP 2004*, pages 33–40, arcelona, Spain, 2004. Association for Computational Linguistics.

Dowty, D. R. *The effects of aspectual class on the temporal structure of discourse: Semantics or pragmatics?*, chapter 9, pages 37–61. Linguistics and Philosophy, 1986.

ElKhlifi, A. and Faiz, R. Automatic annotation approach of events in news articles. *International Journal of Computing and Information Sciences*, 7:40–50, 2009.

Faiz, R. Identifying relevant sentences in news articles for event information extraction. *International Journal of Computer Processing of Oriental Languages*, 19:19–37, 2006.

Ferro, L., Mani, I., Sundheim, B., and Wilson, G. Tides temporal annotation guidelines. Technical Report 1.0.2, The MITRE Corporation, Washington C3 Center McLean, Virginia, 2001.

Filatova, E. and Hovy, E. Assigning time-stamps to event-clauses. In Katz, G., Pustejovsky, J., and Schilder, F., editors, *the workshop on Temporal and spatial information processing*, Dagstuhl Seminar Proceedings, Morristown, NJ, USA, 2001. Association for Computational Linguistics.

Fisher, I. A prototype system for temporal reconstruction of financial accounting standards. *International Journal of Accounting Information Systems*, 8:139–164, JMLR.org, 2007.

Gaizauskas, R. and Wilks, Y. Information extraction: Beyond document retrieval. *Journal of Documentation*, 54(1):70–105, 1998.

Grishman, R. and B.Sundheim. Message understanding conference-6: A brief history. In *the 16th Conference on Computational Linguistics*, pages 466–471, USA, Morristown, NJ, 1996. Association for Computational Linguistics.

Ha, E., Baikadi, A., Licata, C., and Lester, J. Ncsu: Modeling temporal relations with markov logic and lexical ontology. In *5th International Workshop on Semantic Evaluations*, pages 341– 344, Upagessala, Sweden, 2010. SemEval.

Hagège, C., Maupertuis, D., and Tannier, X. Xrce-t: Xip temporal module for tempeval campaign. In *the 4th Workshop on Semantic Evaluations*. SemEval, 2007.

H.Llorens, E.Saquete, and B.Navarro. Tipsem (english and spanish):evaluating crfs and semantic roles in tempeval-2. In *the 5th International Workshop on Semantic Evaluation, ACL 2010*, pages 284–291, Upagessala, Sweden, 2010. Association for Computational Linguistics.

Hwang, C. H. and Schubert, L. K. Tense trees as the “fine structure” of discourse. In *the 30th annual meeting on Association for Computational Linguistics*, Morristown, NJ, USA, 1992. ACL.

Kamp, H. *A theory of truth and semantic representation.* In Formal methods in the study of language, 1981.

Kamp, H. and Reyle, U. *From Discourse to Logic. Introduction to Model-theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory.* Studies in Linguistics and Philosophy, Dordrecht, The Netherlands, 1993.

Kingsbury, P. and Palmer, M. From treebank to propbank. In *Third International Conference on Language Resources and Evaluation*. LREC-02, 2000.

Kufman, M. *Fifth Message Understanding Conference, (MUC-5): proceedings of a conference held in Baltimore, Maryland, August 25-27, 1993*. Morgan Kaufmann Publishers, 1993.

Lascarides, A. and Asher, N. A semantics and pragmatics for the pluperfect. In *Proceedings of the sixth conference on European chapter of the Association for Computational Linguistics*, pages 250–259, Stroudsburg, PA, USA, 1993. Association for Computational Linguistics.

Longacre, R. *The Grammar of Discourse*. New York. Plenum Press, 1983.

Mani, I., Pustejovsky, J., and Gaizauskas, R. *The Language of Time. A Reader*. Oxford University Press, 2005.

Mani, I., Verhagen, M., B.Wellner, Lee., C., and Pustejovsky, J. Machine learning of temporal relations. In *the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*, pages 753–760, Morristown, NJ, USA, 2006. Association for Computational Linguistics.

Mani, I. and Wilson, G. Robust temporal processing of news. In *ACL Annual Meeting*, pages 69–76, NJ, USA, 2000. Association for Computational Linguistics.

Manning, C. D. and Schutze, H. *Foundations of statistical natural language processing*. Mit Press, Cambridge, 1999.

Mourad, G. *Analyse informatique des signes typographiques pour la segmentation de textes et l'extraction automatique de citations. Réalisation des Applications informatiques: SegATex et CitaRE*. PhD thesis, Paris-Sorbonne, 2002.

Nunes, S., Ribeiroand, C., and David, G. Use of temporal expressions in web search. In *ECIR 2008*, pages 580–584. Springer-Verlag Berlin Heidelberg, 2008.

Pustejovsky, J., Castano, J., Ingria, R., Sauri, R., Gaizauskas, R., Setzer, A., and Katz, G. Timeml: Robust specification of event and temporal expressions in text. In *in Fifth International Workshop on Computational Semantics*. IWCS-5, 2003a.

Pustejovsky, J., Hanks, R., Sauri, A., See, R., Gaizauskas, A., Setzer, A., Radev, B., Sundheim, D., Day, L., Ferro, L., and Lazo, M. The timebank corpus. In *Corpus Linguistics*. Lancaster, 2003b.

Quinlan, J. R. Induction of decision trees. *Machine Leaning*, 1(1), Kluwer Academic Publishers, 1986.

Quinlan, J. R. *C45: Programs for Machine Learning*. Morgan Kaufmann, San Mateo, CA, 1993.

Reichenbach, H. *The tenses of verbs*. Elements of Symbolic Logic. New York. Macmillan & Co, 1947.

Riloff, E. Automatically constructing a dictionary for information extraction tasks. In *11th National Conference on Artificial Intelligence*, pages 811– 816, 1993.

Roberts, A., Gaizauskas, R., and Hepple, M. Extracting clinical relationships from patient narratives. In *Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing BioNLP*, pages 10–18, Columbus, Ohio, USA, 2008.

Sanampudi, S. K. and Kumari, G. . Temporal reasoning in natural language processing: A survey. *International Journal of Computer Applications*, 1(4):53–57, February 2010. Published By Foundation of Computer Science.

Sarawagi, S. Information extraction. *Foundations and Trends in Databases*, 1:261–377, 2008.

Sauri, R., Pustejovsky, J., and Littman, J. Argument structure in timeml. In Katz, G., Pustejovsky, J., and Schilder, F., editors, *Annotating, Extracting and Reasoning about Time and Events*, Dagstuhl Seminar Proceedings, Dagstuhl, Germany, 2006.

Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany.

Schilder, F. Event extraction and temporal reasoning in legal documents. In *the proceedings of Annotating, Extracting and Reasoning about Time*, pages 59–71. LNAIVolume 4795/2007, 2007.

Setzer, A. and Gaizauskas, R. Annotating events and temporal information in newswire texts. In *the Second International Conference on Language Resources and Evaluation*, pages 1287–1294. LREC2000, 2000.

Setzer, A. and Gaizauskas, R. On the importance of annotating temporal event-event relations in text. In *the Third International Conference on Language Resources and Evaluation*. LREC2002, 2002.

Seymore, K., McCallum, A., and Rosenfeld, R. Learning hidden markov model structure for information extraction. In *AAAI'99 Workshop on Machine Learning for Information Extraction*, pages 37–42, 1999.

Sundheim, B. M. Overview of the third message understanding evaluation and conference. In *Proceedings of the Third Message Understanding Conference (MUC-3)*, pages 3–16, San Diego, CA, 1991.

ter Meulen, A. *FRepresenting time in natural language: the dynamic interpretation of tense and aspect*. Mit Press, Cambridge, 1995.

Vendler, Z. *Verbs and Times*, chapter 4, pages 97–121. Cornell University Press, Ithaca, New York, 1967.

Verhagen, M. Semeval-2010 task 13: Tempeval 2. In *the International Workshop on Semantic Evaluations*. SemEval, 2010.

Verhagen, M., Gaizauskas, R., Schilder, F., Hepagesle, M., Katz, G., and Pustejovsky, J. Semeval-2007 task 15: Tempeval temporal relation identification. In *the 4th International Workshop on Semantic Evaluations*. SemEval, 2007.

Verhagen, M., Mani, I., Sauri, R., Knipagesen, R., Jang, S. B., Littman, J., .and J. Phillips, A. R., and Pustejovsky, J. Automating temporal annotation with tarsqi. In *Proceedings of the ACL 2005 on Interactive poster and demonstration sessions*, pages 81–84, Stroudsburg, PA, USA, 2005. Association for Computational Linguistics.

Webber, B. L. *Tense as discourse anaphor*, chapter 14, pages 61–73. Linguistics and Philosophy, 1988.