UC3M system: Determining the Extent, Type and Value of Time Expressions in TempEval-2

María Teresa Vicente-Díez, Julián Moreno Schneider, Paloma Martínez
Department of Computer Science
Universidad Carlos III de Madrid
Avda. Universidad, 30
Leganés, 28911, Madrid, Spain.
{tvicente, jmschnei, pmf}@inf.uc3m.es

Abstract

This paper describes the participation of Universidad Carlos III de Madrid in Task A of the TempEval-2 evaluation. The UC3M system was originally developed for the temporal expressions recognition and normalization (TERN task) in Spanish texts, according to the TIDES standard. Current version supposes an almost-total refactoring of the earliest system. Additionally, it has been adapted to the TimeML annotation schema and a considerable effort has been done with the aim of increasing its coverage. It takes a rule-based design both in the identification and the resolution phases. It adopts an inductive approach based on the empirical study of frequency of temporal expressions in Spanish corpora. Detecting the extent of the temporal expressions the system achieved a Precision/Recall of 0.90/0.87 whereas, in determining the TYPE and VALUE of those expressions, system results were 0.91 and 0.83, respectively.

1 Introduction

The study of temporality in NLP is not a new task. However, in the last years it has witnessed a huge interest. Initiatives like TempEval task or the Automatic Context Extraction[1] (ACE) TERN competitions have boosted research on the field and have promoted the development of new resources to the scientific community.

There are two main advantages in participating in these evaluations. On the one hand it is possible to measure the systems' performance under standardized metrics, sharing datasets and other resources. On the other hand, it is possible to make comparative evaluations among distinct participants looking forward the same objectives but using different approaches.

Until recently, most of temporally annotated corpora, as well as temporal taggers, were available in English. Since languages as Spanish start to become prominent in the field it seems interesting the development of specific resources. Tempeval-2 has contributed to this target in a significant way thanks to the release of annotated corpora and the publication of specific guidelines (Sauri et al., 2009), (Saurí et al., 2010).

This paper resumes the participation of the UC3M system in the task of determining the extent and resolving the value of time expressions in texts (Task A). This system was originally developed for the Spanish TERN task proposed in ACE 2007 evaluation (Vicente-Díez et al., 2007), achieving encouraging results although it was in a early stage of development.

The system follows a ruled-based approach whose knowledge base has been inducted from the study of annotated temporal corpora (Vicente-Díez et al., 2008). A machine learning approach was initially discarded due to the limitation of annotated Spanish corpora.

The aims of this work were to improve the coverage of the original system and test its performance against new available datasets with a view to its integration in future domains of application. Main challenges were to move to a new temporal model where interval is considered as the basic time unit as well as the isolation of the internal representation of temporal information from the annotation schema.

[1] Automatic Content Extraction Evaluation. National Institute of Standards and Technology (NIST) http://www.itl.nist.gov/iad/mig//tests/ace/

Proceedings of the 5th International Workshop on Semantic Evaluation, ACL 2010, pages 329–332,
Uppsala, Sweden, 15-16 July 2010.

This paper is organized as follows: Section 2 describes the system operation; Section 3 presents experimentation and results; conclusions and future work are discussed in Section 4.

2 System Description

The UC3M system recognizes and annotates temporal expressions in texts based on a linguistic rules engine for Spanish language.

Our system is divided into three different parts: recognition of temporal expressions, normalization of the detections, and annotation of the temporal expressions according to the TimeML schema.

Following the definition of the Task A, the system is able to determine not only the extent of the temporal expressions but also the value of the features TYPE and VAL. It differentiates among the four TYPE values (dates, durations, sets and times) thanks to the classification of the recognition rules. The system straightforwardly provides a VAL attribute that accomplishes the format defined by TIMEX2 and TIMEX3 standards through its internal model for representing time.

2.1 Recognition

The recognizer detects temporal expressions by means of a set of linguistic rules, focusing on those which are most frequent in Spanish.

We adopted an empirical inductive approach through the analysis of the different types of temporal expressions in news corpora, and we could outline a typology of most common time expressions in the language. The typology together with the patterns that define these expressions form up the knowledge base for a successful automatic identification and resolution of temporal expressions.

The rule engine allows managing different sets of rules independently of the target. In this case, the rules have been created attending to each pattern that is likely to match a temporal expression. Each rule determines the set of tokens that form an expression, the normalization type to be applied and the expression type.

In Table 1 an example of a rule to identify dates is shown. The first line represents the name of the rule. The second line specifies the normalization method that will be used once the expression is recognized. The third line specifies the type of the temporal expression and the annotation pattern. Finally, the fourth line shows the tokens that trigger the rule.

1.	*TEMPORAL_RULE(r1.3)*
2.	*TEMPORAL_ANALYSIS_NORMALIZATION_ TYPE=(abs_dia_mes_anio_3)*
3.	*TEMPORAL_ANALYSIS_TYPE= (date:init:YYYY-MM-DD)*
4.	*RULE= [[el/_] [DIC(DIASEMANA)/_] [dia/_] DIC(DIA) de DIC(MES) DIC(PREP) METHOD(year)]*

Table 1 Rule definition example

The operation of the system is described as follows: first, the text is parsed token by token. Then, for each token, every rule is checked to find out if it triggers through a given token and the following ones.

This operation implies that the higher the number of rules, the slower the text processing. The disadvantage of the processing speed has been accepted as a design criterion for the sake of the simplicity of creating new rules.

2.2 Normalization

The temporal expression normalization is done as an intermediate step between recognition and annotation, isolating the extraction of semantics from the annotation schema while trying to facilitate the second step.

Normalization is important since recognized time expressions are managed and returned in a standard format that avoids semantic ambiguities.

UC3M system applies an interval-based temporal normalization. It means that every temporal expression is represented as an interval with two boundaries: an initial and a final date (including time). This approach is motivated by the belief that the use of intervals as a basic time unit leads to a lower loss of semantics. For instance, when an expression like “en enero” (“in January”) is detected, current task proposes the annotation “2010-01”. However, we think that for many applications that are likely to use this system it would be more useful to have the complete interval that the expression refers (“2010-01-01 - 2010-01-31”). Through a set of procedures (as getting the length of a given month), our system tries to define the interval boundaries as much as possible. Every normalized expression is made up of two dates although it refers to a concrete date or time.

In the internal representation model normalized dates and times adopts the ISO-8601 form, durations are captured as a length related to the unit of measure, and sets are managed in a similar way to durations, adding quantity and frequency modifiers.

The normalization process is dependent on the rule used to recognize each expression. For each new rule added to the engine a new normalization clause is needed.

In Table 2 some temporal expression normalization examples are presented:

Expression	Init Date	Final Date
18 de abril de 2005 *18th of April of 2005*	20050418	20050418
mayo de 1999 *May of 1999*	19990501	19990531
en 1975 *in 1975*	19750101	19751231
el próximo mes *next month*	20100501	20100531

Table 2 Interval-based normalization sample

2.3 Annotation

The annotation process starts from the normalized form of the temporal expression. The system implements a transformation procedure based on patterns. This transformation is dependent on the temporal expression type.

Dates: when dealing with dates, the VAL value is extracted from the initial boundary of the interval in accordance with the annotation pattern defined in the corresponding rule (see Table 1). Some examples are shown in Table 3.

Expression	Norm. Init Date	Pattern	VAL
mayo de 1999 *May of 1999*	19990501	YYYY-MM	1999-05
la semana pasada *last week*	20100405	YYYY-WXX	2010-W14
los años 80 *the 80's*	19800101	YYY	198

Table 3 Annotation patterns for dates

Durations: the model represents durations by capturing the length of action as a quantity. This quantity is stored in the position of the initial boundary whose granularity corresponds with the unit of measure. The annotation patterns indicate the granularity to be considered (Table 4).

Expression	*Norm. Init Date*	*Pattern*	*VAL*
4 años *4 años*	00040000	PXY	P4Y
4 meses, 3 días y 2 horas *4 moths,3 days and 2 hours*	00040003-02:00:00	COMBINED	P4M3DT2H

Table 4 Annotation patterns for durations

Sets are managed similarly to durations. In this case also frequency and quantity modifiers are captured internally together with the interval representation, so that the transformation is immediate.

Expression	Norm. Init Date	Pattern	VAL	FREQ	QUANT
cada 2 años *each 2 years*	00020000 F1QEv	PXY	P2Y	1x	EVERY
2 veces al día *twice a day*	00000001 F2QEv	PXD	P1D	2x	EVERY

Table 5 Annotation patterns for sets

Times: the representation model allows capturing hours, minutes, seconds and milliseconds if they are specified. Similarly to the annotation of dates, VAL value is obtained of the information in the initial boundary in the way the pattern determines (Table 6).

Expression	Norm. Init Date	Pattern	VAL
a las 12:30 PM *at 12:30 PM*	20100405 12:30:00	THXMX	2010-04-05T12H30M
por la tarde *in the evening*	20100405 12:00:00	TDP	2010-04-05TAF

Table 6 Annotation patterns for times

3 Experiments and Results

Precision and recall and f-measure are used as evaluation metrics according to the evaluation methodology (Pustejovsky et al., 2009). To determine the quality of annotation, results are completed with figures concerning to the resolution of TYPE and VAL attributes.

Before evaluation, the system was tested on the training corpus and, once the test datasets were released, it was tested on the corpus for relations detection (tasks C-F) since it contained both files "*timex-extents.tab*" and "*timex-attributes.tab*". The results are shown in Table 7.

Corpus	Timex Extent			Timex Attbs.	
	P	*R*	***F***	*TYPE*	*VAL*
Training	0.93	0.67	**0.78**	0.87	0.82
Relation-Test	0.89	0.63	**0.74**	0.86	0.83

Table 7 Results on training corpus

In Table 8 results of final evaluation are presented and compared with the other participants' figures for the same task and language. Since the test corpora were not aligned, further comparisons for different languages have not been proposed.

Our system achieved a precision rate of 90% and a recall of 87%, being the f-measure of 88%. Thus, it supposes a significant improvement over our earlier work. In more, determining the value of TIMEX3 attributes the system raises good figures, obtaining the best VAL score, what means that normalization is working well.

Team	Timex Extent			Timex Attrbs.	
	P	*R*	***F***	*TYPE*	*VAL*
UC3M	0.90	0.87	**0.88**	0.91	0.83
TIPSem	0.95	0.87	**0.91**	0.91	0.78
TIPSem-B	0.97	0.81	**0.88**	0.99	0.75

Table 8 Results on test corpus

Analyzing the experimental errors several facts can be highlighted:

The percentage of expressions completely and correctly recognized and normalized is good but there are some missing expressions, mainly due to their complexity (or fuzziness) and to the absence of a rule to manage them, i.e.: “durante un largo periodo” (*during a long period*).

Errors in determining the extent of the temporal expressions were mainly due to the inclusion of prepositions or articles that precede to the kernel of the expression, i.e.: “a corto plazo” vs. “corto plazo” (*in short term*).

A number of false positives were due to some inconsistencies in the annotation of the corpus. An example has been observed in fuzzy time expressions that denotes a future reference: “el próximo técnico” (*the next trainer*) (not annotated) vs. “el próximo preparador” (*the next coach*) (FUTURE_REF)

Although normalization figures are good, some annotations are incorrect if their resolution implies context-aware mechanisms.

4 Conclusions and Future Work

In this paper a rule based approach for automatically detecting and annotating temporal expressions according to TimeML TIMEX3 tag has been presented. It is based on an empirical study of temporal expressions frequencies in Spanish that provides the main recognition rules of the knowledge base. At the normalization stage, a representation model based on intervals has been adopted with the aim of capturing most semantics. The annotation process relies on patterns that distinguish among different types and granularities of the expressions to be tagged.

Obtained results suppose a significant improvement over our previous work. Part of this success is due to the specific annotation guidelines for Spanish that have been released with occasion of the TempEval-2. It is a helpful tool to optimize the system performance, since each language has its own peculiarities that should be taken into account. The promotion of a common framework and the development of resources like specific corpora are also very interesting topics to boost research in the field, since both comparative and standardized evaluation of the systems are needed.

Several aspects should be taken into account in future versions of the system. In order to improve the recall new knowledge must be incorporated to the rule engine. That supposes the addition of new rules and annotation patterns. This objective includes the implementation of dictionaries with a broader coverage of translatable temporal expressions, such as holidays, festivities, etc.

We will also explore context extraction techniques that facilitate the resolution of context-aware temporal expressions.

Another pending issue is the enlargement of the system to span the detection of events and the relations among events and time expressions.

Finally, the system will be integrated into a NLP application that benefits from the temporal information management. We want to check the improvement that the extraction of temporal entities supposes on a traditional approach.

Acknowledgments

This work has been partially supported by the Research Network MAVIR (S-0505/TIC-0267), and project BRAVO (TIN2007-67407-C03-01).

References

James Pustejovsky, Marc Verhagen, Xue Nianwen, Robert Gaizauskas, Mark Hepple, Frank Schilder, Graham Katz, Roser Saurí, Estela Saquete, Tommaso Caselli, Nicoletta Calzolari, Kiyong Lee, and Seohyun Im. 2009. TempEval2: Evaluating Events, Time Expressions and Temporal Relations. SemEval Task Proposal.

María Teresa Vicente-Díez, Doaa Samy and Paloma Martínez. 2008. An empirical approach to a preliminary successful identification and resolution of temporal expressions in Spanish news corpora. In Proceedings of the LREC'08.

María Teresa Vicente-Díez, César de Pablo-Sánchez and Paloma Martínez. Evaluación de un Sistema de Reconocimiento y Normalización de Expresiones Temporales en Español. Procesamiento del lenguaje natural. N. 39 pp. 113-120, Sept. 2007.

Roser Saurí, Estela Saquete and James Pustejovsky. 2010. Annotating Time Expressions in Spanish. TimeML Annotation Guidelines. Version TempEval-2010.

Roser Saurí, Olga Batiukova, James Pustejovsky. 2009. Annotating Events in Spanish. TimeML Annotation Guidelines. Version TempEval-2010.