

Evaluation of Complex Temporal Questions in CLEF-QA

E. Saquete, J.L. Vicedo, P. Martínez-Barco, and R. Muñoz

Grupo de investigación en Procesamiento del Lenguaje y Sistemas de Información.

Departamento de Lenguajes y Sistemas Informáticos.

University of Alicante, Spain

{*stela,vicedo,patricio,rafael*}@dlsi.ua.es

Abstract. This paper presents the evaluation of a Temporal QA system for the treatment of temporal complex questions. The system was implemented as a multilayered architecture where complex temporal questions are first decomposed into simple questions, according to the temporal relations expressed in the original question. These simple questions are then processed independently by our standard Question Answering engine and their respective answers are filtered to accomplish each simple question temporal restrictions. Finally, answers to simple decomposed questions are integrated following the temporal relations extracted from the original complex question in order to compose the final answer. This evaluation was performed as a pilot task at the Spanish QA Track from the Cross Language Evaluation Forum 2004.

1 Introduction

Although current operational Question Answering systems deal with simple factual questions, more and more systems dealing with complex questions are needed in order to reach more complex information. One of this kind of questions are temporal questions, that is, questions requiring a date as answer (“When did Bob Marley die?”) or questions that use temporal expressions in their formulation (“Who won the U.S. Open in 1999?”). Processing this kind of questions is usually accomplished by identifying explicit temporal expressions in questions and relevant documents in order to detect temporal expressions that are necessary to answer the queries. Even though, it seems necessary to emphasize the system described in [1] as the only one that also uses implicit temporal expression recognition for Question Answering purposes by applying the temporal tagger developed by Mani and Wilson [2]. However, questions referring to the temporal properties of the entities being questioned and the relative ordering of events mentioned in the questions are beyond the scope of current Question Answering systems:

- “Is Bill Clinton *currently* the President of the United States?”
- “Who was spokesman of the Soviet Embassy in Baghdad *during* the invasion of Kuwait?”

- “Were there any meetings between the terrorist hijackers and Iraq *before* the WTC event?”

This paper shows the participation of the University of Alicante in the Spanish CLEF 2004 Pilot Task, that aims at investigating how Question Answering systems answer complex questions with temporal restriction, with more than a correct answer, or whose answer is a list of items. Our participation is focused in treating these complex temporal questions.

The treatment is based on the decomposition of complex questions into simple questions to be resolved in conventional Question Answering systems as describe above.

2 Answering Temporal Questions

The study of two corpora [3] [7] including information and questions related to time showed a classification of temporal questions presented in [4] [5] about the way of resolving the questions. Then we can distinguish between simple questions, that can be solved directly, and complex questions that must be processed by means of a temporal expression analyzer.

The temporal analyzer used in this task is based on the TERSEO temporal resolution system [6]. The analyzer processes the temporal expressions in the question, and then a question decomposition module is used in order to split the question in simple questions. Finally, the answers to these simple questions have to be combined to answer the complex question.

In order to perform these tasks, the system is able to classify among the following types of question:

- **No temporal Questions:** *Type 0*
- **Simple Temporal Questions**
 - *Type 1: Single event temporal questions without temporal expression.* Resolved by a Question Answering System directly without pre or postprocessing of the question. *When did Jordan close the port of Aqaba to Kuwait?*
 - *Type 2: Single event temporal questions with temporal expression.* There are one or more temporal expressions that need to be recognized, resolved and annotated. *Who won the 1988 New Hampshire republican primary?*
- **Complex Temporal Questions**
 - *Type 3: Multiple events temporal questions with temporal expression.* Questions contain two or more events, related by a temporal signal. The temporal expressions need to be recognized, resolved and annotated. *What did George Bush do after the U.N. Security Council ordered a global embargo on trade with Iraq in August 90?*
 - *Type 4: Multiple events temporal questions without temporal expression.* Questions consist of two or more events, related by a temporal signal. *What happened to world oil prices after the Iraqi “annexation” of Kuwait?*

The processing of these questions will be explained in detail in the following sections.

3 System Description

Our Temporal Question-Answering system is based on a multilayered architecture increasing the functionality of actual Question-Answering systems to solve any type of temporal questions. This architecture superposes an additional processing layer, one by each type of complex question, to the General Purpose Question Answering system.

Our system is focused on the Temporal Q.A. processing module, however other kinds of question could be solved according this architecture (script questions, template questions, ...)

3.1 Temporal Question Answering Module

Figure 1 shows the different parts and interaction between them of the architecture of a Temporal Question Answering System. The main components of the Temporal Question Answering System are:

- Question Decomposition Unit,
- General purpose Question-Answering system, and
- Answer Recomposition Unit.

These components work all together in order to obtain a final answer. The Question Decomposition Unit and the Answer Recomposition Unit are the units that conform the Temporal Question-Answering layer which process the temporal questions, before and after using the General Purpose Question Answering system.

- *Question Decomposition Unit* is a preprocessing unit which has three main tasks. First of all, because we are dealing with questions related to temporality, the recognition and resolution of temporal expressions of the question is needed, if there is any. Regarding the taxonomy of the questions shown before, there are different types of questions and every type has to be treated in a different way from the others. For this reason, type identification needs to be done. After that, complex questions, are split in simple ones. These simple questions are the input of the General Purpose Question-Answering system.
- *General Purpose Question-Answering system*. Any generic Question-Answering system could be used here. In this case, we use the QA System for Spanish developed in the University of Alicante [8].
- *Answer Recomposition Unit* is the last stage in the process. This unit carries out a recomposition of the different answers, using temporal information obtained from the question, like temporal signals (that are explained later) or temporal expressions, and manage to return the correct answer to the original question.

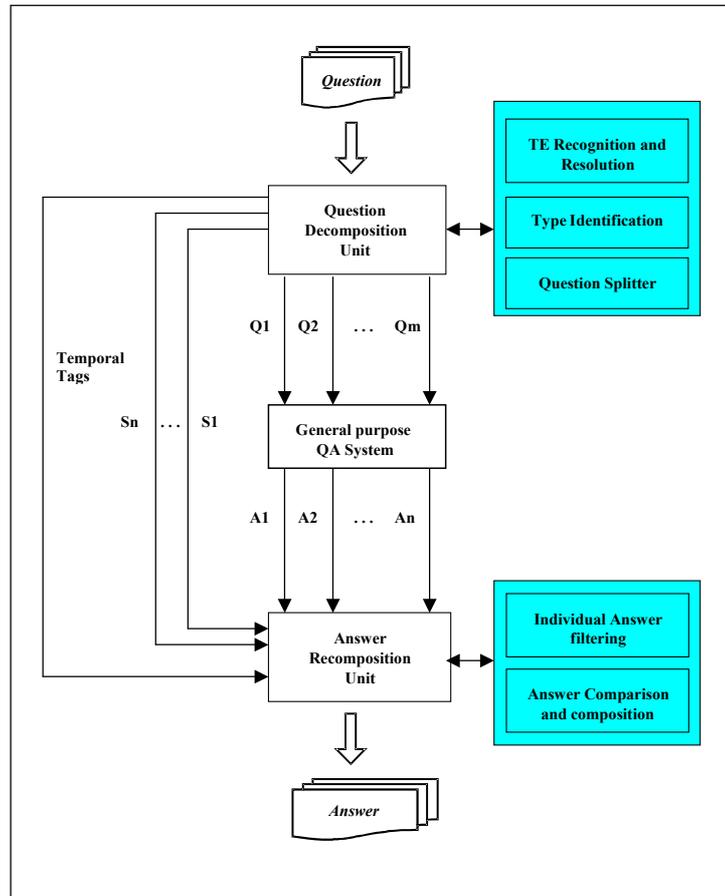


Fig. 1. Graphic representation of the Temporal Question Answering System

4 CLEF Evaluation

Participants in the CLEF Spanish QA Pilot Task have to answer 100 questions, that are equally distributed throughout the following types:

- Factoid questions with temporal restriction
 - Date restriction: A precise date contextualises the object of the question, that can refer either to that particular moment, or before or after that date. “¿Quién gobernaba en Francia en 1988?”. “¿Qué país visitó Berlusconi antes de junio de 1994?”. This kind of questions is recognized by our system as type 2 questions.
 - Period restriction: In this case, questions refer explicitly to a whole period. “¿Quién gobernaba en Irak entre 1985 y 1987?”. “¿Cuántos coches se vendieron en España en la década de los ochenta?”. These questions

also correspond to type 2, although TERSEO solves the temporary expression as a period of time instead of a concrete date.

- Event restriction (embedded question): Temporal restriction refers here to the moment in which a second event occurred. “¿Quién gobernaba en Argentina durante la guerra de las Malvinas?”. “¿Qué le ocurrió al precio del crudo tras la invasión iraquí de Kuwait?”. These questions are recognized by our system as type 3 and 4.
- Lists: Participating groups are given questions whose answer is a list of items, persons, organisations, etc. “Enumere los países que pertenecen a la UE”. “¿Quién ha presidido el gobierno español desde 1992?”. As our system is focused in temporal questions, most of them are not able to be solved. However, taking into account the multilayered architecture, a layer specialized in lists could be integrated to the system to confront this new challenge.

Results will be known during the CLEF 2004 Conference.

5 Conclusions

This paper presents the participation of the University of Alicante in the Spanish CLEF QA Pilot Task for complex questions with a new method for answering complex temporal questions using current factual-based Question Answering systems. The method is based on a multilayered architecture increasing the functionality of actual Question-Answering systems to solve any type of temporal questions. The architecture superposes an additional processing layer to the General Purpose Question Answering system.

The complex temporal QA layer is based on a new proposal for the decomposition of temporal questions where complex questions are divided into simpler ones by means of the detection of temporal signals. The TERSEO system, a temporal information extraction system applied to event ordering, has been used to detect and resolve temporal expressions in questions and answers. This information is needed in the following steps:

- the decomposition of the question into simple events to generate simple questions (sub-questions),
- the ordering of the sub-questions,
- the filtering of the sub-answers,
- and finally, the comparison between sub-answers to build the final complex answer.

Moreover, as TERSEO system is a multilingual system, the layer used to complex temporal questions in Spanish could be easily extended to other languages, obtaining a multilingual Temporal QA system.

Although our participation has been focussed on solving complex temporal questions, the same approach can be applied to process other kinds of complex questions that allow question decomposition such as script questions, or template-like questions.

6 Acknowledgments

This paper has been supported by the Spanish government, project TIC-2003-07158-C04-01.

References

1. E. Breck, J. Burger, L. Ferro, W. Greiff, M. Light, I. Mani, and J. Rennie. Another sys called quanda. In *Ninth Text REtrieval Conference*, volume 500-249 of *NIST Special Publication*, pages 369–378, Gaithersburg, USA, nov 2000. National Institute of Standards and Technology.
2. I. Mani and G. Wilson. Robust temporal processing of news. In ACL, editor, *Proceedings of the 38th Meeting of the Association of Computational Linguistics (ACL 2000)*, Hong Kong, October 2000.
3. D. Radev and B. Sundheim. Using timeml in question answering. <http://www.cs.brandeis.edu/~jamesp/arda/time/documentation/TimeML-use-in-qa-v1.0.pdf>, 2002.
4. E. Saquete, P. Martínez-Barco, and R. Muñoz. Splitting complex temporal questions for question answering systems. In ACL, editor, *Proceedings of the 42nd Meeting of the Association of Computational Linguistics (ACL 2004)*, pages 567–574, Barcelona, July 2004.
5. E. Saquete, P. Martínez-Barco, R. Muñoz, and J.L. Vicedo. Multilayered question answering system applied to temporality evaluation. *Procesamiento del Lenguaje Natural*, (33):25–34, September 2004.
6. E. Saquete, R. Muñoz, and P. Martínez-Barco. Terseo: Temporal expression resolution system applied to event ordering. In TSD, editor, *Proceedings of the 6th International Conference ,TSD 2003, Text, Speech and Dialogue*, pages 220–228, Ceske Budejovice,Czech Republic, September 2003.
7. TREC, editor. *The TREC-8 Question Answering Track Evaluation, Text Retrieval Conference TREC-8*, Gaithersburg,MD, November 1999.
8. J.L. Vicedo, R. Izquierdo, F. Llopis, and R. Muñoz. Question answering in spanish. In CLEF, editor, *Proceedings CLEF-2003 Lecture Notes in Computer Science*, Trondheim, Norway, August 2003.