

Luiz Augusto S. Pizzato and Diego Mollá-Aliod
Centre for Language Technology
Macquarie University, Sydney, Australia
{pizzato, diego} @ics.mq.edu.au

Abstract

This work concerns a question answering tool that uses multiple Web search engines and Web question answering systems to retrieve snippets of text that may contain an exact answer for a natural language question. The method described here treats each Web information retrieval system in a unique manner in order to extract the best results they can provide. The results obtained suggest that our method is comparable with some of today's state-of-the-art systems.

Introduction

Text-based Question Answering (QA) focuses on finding answers for natural language questions by searching collections of textual documents. Our system belongs to the category of Web-based QA that can take advantage of the enormous amount of data available on the World Wide Web and use data-intensive approaches that exploit the inherent redundancy to find answers [1].

An exact answer is defined as a string that does not contain any extraneous information but the answer in it [2].

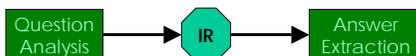
What is the capital of Brazil?

✓ *Brasília*

✗ *Brazilian capital Brasília*

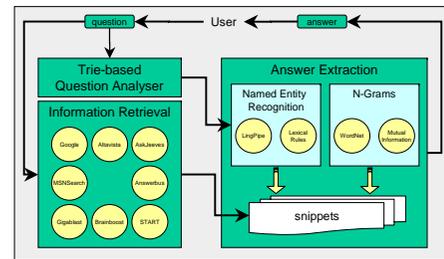
Our approach combines the results of several Web search engines and Web QA systems. Our system works in a similar way of those known as meta-search engines, however we do differentiate between the search engines used in order to extract the best information they may provide.

Our system architecture is similar to the common QA framework. The common framework considers three main phases as shown in the image below.



System Architecture

Our system structure is very similar to the common framework, but the approaches for performing each of the tasks are different. The question analysis is performed using the Trie-based question classifier [3,4] trained over nearly 5500 questions [5]. The information retrieval stage is a combination of several Web search engine results, and the answer extraction combines named-entity, n-grams and lexico-semantic information from WordNet [6].



Web search results combined

Results are combined into four different sets:

- **MSN Search:** Can answer some questions by using encyclopaedia information, as well as providing definitions for words, and a way to make measurement conversion.
- **Start:** Uses predefined Web databases to answer questions on several subjects.
- **Answerbus:** Answers in a snippet-like format.
- **Brainboost:** Similar to Answerbus.
- **AskJeeves:** Useful information about specific questions on celebrities, movies, weather and words definition.
- **Gigablast:** Gigablast presents related term to the search results that are likely to contain the answers to questions.
- **Google:** Considered the best Web search engine. It provides definitions for words and, following MSN Search, Google has recently acquired the ability to answer encyclopaedia questions. We understand that this is a good feature to be used in our system, but the version describe here does not yet consider this feature
- **Altavista:** Well established search engine.

Exact answer extraction

The approach used for pinpointing the exact answer location uses named-entity recognition combined with n-grams extraction and word overlap. We also make use of the semantic classification of terms in WordNet.

Answers are mined using a set of heuristics based on the question type and the sets of results from the IR engines.

Who killed JFK? (HUM:IND)

What is ethology? (DESC:DEF)

Definition questions is answered by asking a definition query to MSN Search and Google. Factoid ones require a search for the correct Named-Entity. Named-Entities receive a score value according to their relative distance from the question words in the extracted passages.

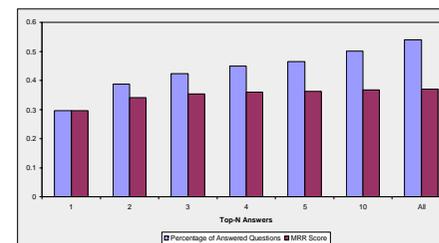
We also compute n-grams and mutual information boosting the score of an n-gram if it exists in the hierarchy of the question words in WordNet. This helps answer questions like:

Which breed of dog has a blue tongue?

We then generate an XML containing the best exact answers, the snippets of text containing the answers' context and their source documents.

Evaluation

The evaluation was performed in a similar way as the main task of the QA track of TREC-2003 [6]. We, however, have just evaluated factoid questions and using web documents instead of Aqaunt ones. We were also obligated to manually judge some answers due to differences regarding updated, modified and previous inexistent answers.



Concluding Remarks

We develop a meta-QA system that combines the results of different Web search/QA systems in order to provide exact answers for natural language questions. By using a trie-based question analysis, named-entity recognition, n-gram computation and lexico-semantic information from WordNet, we are able to achieve results comparable to some of the best state-of-the-art QA systems. The results shows that the exact answer could be found in almost half of the times by considering up to 5 answers for every question, giving a reasonable MRR score of 0.36.

Further work is needed in order to identify the gain in performance by adding, replacing, removing and promoting search engines. There is also a need for the evaluation of the best weights for the features used to pinpoint the location of the answers, and the feasibility of using language independent methods such as n-grams and mutual information to perform a multilingual QA.

There is also the capacity of taking advantage of certain features provided by search engines. For instance, by restricting the search domain by Web site, language, country or even neighbourhoods, it is possible to restrict the QA domain. We already performed some minor tests asking questions in the Macquarie University Website showing promising results.

References

- [1] Eric Brill, Jimmy Lin, Michele Banko, Susan Dumais, and Andrew Ng. 2001. Data intensive question answering. In Ellen M.Voorhees and Donna K. Harman, editors, Proc. TREC 2001, number 500-250 in NIST Special Publication. NIST.
- [2] Ellen M. Voorhees. 2003. Overview of the TREC 2002 question answering track. In Proceedings of TREC-2002.
- [3] Luiz Augusto Sangoi Pizzato. 2004. Using a trie-based structure for question analysis. In Ash Asudeh, Cécile Paris, and Stephen Wan, editors, *Proceedings of the Australasian Language Technology Workshop 2004*, pages 25-31, Macquarie University, Sydney, Australia, December. ASSTA. ISBN: 0 9581946 1 0
- [4] Menno Van Zaanen, Luiz Augusto Pizzato, and Diego Molla. 2005. Question classification by structure induction. In *Proc. of the Nineteenth International Joint Conference on Artificial Intelligence (IJCAI-2005)*, Edinburgh, Scotland, August.
- [5] X. Li and D. Roth. 2002. Learning question classifiers. In Proceedings of the COLING-02, pages 556-562.
- [6] C. Fellbaum, editor. 1998. *WordNet: An electronic Lexical Database*. MIT Press

Acknowledgements

The authors would like to thank Macquarie University and iMurs Scholarship for their support on this work.