

Question Answering Using Web News as Knowledge Base

Zhiping Zheng

Computational Linguistics Department

Saarland University

D-66041 Saarbrücken, Germany

zheng@coli.uni-sb.de

Abstract

AnswerBus News Engine¹ is a question answering system using the contents of CNN Web site² as its knowledge base. Comparing to other question answering systems including its previous versions, it has a totally independent crawling and indexing system and a fully functioning search engine. Because of its dynamic and continuous indexing, it is possible to answer questions on just-happened facts. Again, it reaches high correct answer rate. In this demonstration we will present the living system as well as its new technical features.

Keywords: question answering, QA specific indexing, search engine

1 Introduction

AnswerBus³ ([2,3]) is originally designed as a Web-based open-domain question answering (QA) system. It successfully uses natural language processing and information retrieval techniques and reaches very high correct answer rate. Although it is not designed for TREC, it still correctly answers over 70% of TREC-8 questions with Web resources. Because we use commercial search engines as the search tools for the system, we don't know if special indexing system and other possible techniques will work better for the QA tasks.

In the new experiment, we used the contents of CNN Web site and developed a QA system

called AnswerBus News Engine to automatically answer news related questions. We chose CNN Web site as the knowledge base because it has a good archive of news stories since 1996 and the CNN Web site seems having good reputation on timely updating. The goal of this experiment is to use most techniques used in AnswerBus QA system together with some new techniques, such as QA specific indexing described in [2,3] but not fully implemented in original AnswerBus system, and build a QA system to answer time sensitive questions in the real world.

Before building the AnswerBus News Engine, we did another experiment⁴ ([7]) using part of DUC conference corpus as local archive. The result was exciting. The experimental QA system correctly answered 80% questions designed specially for the local archive.

2 New Features

AnswerBus News Engine has many new features not used in other QA systems including its previous versions.

2.1 Sentence Level Indexing

QA systems usually use some search tools to retrieve documents. These search tools include some commercial search engines like Google, Alta Vista. Some other systems tried local search engines for local data, for example, local Web contents or TREC corpus. We partially deployed the techniques used in Seven Tones Search Engine⁵ ([6,5]) for the search task, since it has a high indexing speed and it is possible to timely update the indexed database part by part.

Comparing to other QA systems, AnswerBus News Engine not only uses a specialized search

¹ <http://www.coli.uni-sb.de/~zheng/answerbus/news/>

² <http://www.cnn.com/>

³ <http://www.answerbus.com/>

⁴ <http://www.coli.uni-sb.de/~zheng/answerbus/local/>

⁵ <http://www.seventones.com/>

engine for QA task, but also crawls and indexes CNN Web site automatically. Also, the special index system is different from other search engine index system in some aspects, for example, sentence level indexing ([4]), temporal indexing and partially updating.

As the results of the new techniques, AnswerBus News Engine is now able to answer some time sensitive questions about the some factual issues just happened half an hour ago.

2.2 Embedded Search Engine

It is normal that some times a QA system cannot find any answer from the working knowledge base for a question. This doesn't mean there is no answer for the question. In this case, AnswerBus redirects the question to the embedded search engine so users will get a bunch of documents instead of answers. Very likely, if there is an answer to the question, the user can dig it out from the documents given by the search engine.

2.3 Scalability

The current size of indexed data has been over 700K Web pages from CNN Web site and some of its sub sites. We believe that it has been the largest size of knowledge base for QA tasks at current time. And the designed size can be much bigger than the size we have already reached. This makes it possible for the future system to index the whole Web and answer questions.

2.4 Speed and System Load

Because of the local indexing, AnswerBus is now able to find the possible answers for a user question in 2-4 seconds. This makes the system fast enough to process more documents to mine the answers.

This also decreases the system load than its previous systems and the system can answer more questions at the same time than its previous versions with same resource.

3 Web Interface

The system has a Web interface as its previous versions. As in Figure 1, the system lists up to ten possible answers to a specific user question. Each of these answers has a dynamic link back to a

specific CNN Web page containing the answer sentence. The navigation bar at the end provides an easy way to try user question with other online systems.

Figure 2 is a screen shot when the system could not find the proper answer and the redirected the user question to embedded search engine. This page only shows 20 items returned by the search engine.

4 Evaluation

Evaluation of question answering techniques has been a very difficult task. It gets more difficult to evaluate this system because we don't have any baseline or comparable systems. And also because of the dynamic content, it is difficult to design a question set to do the evaluation like TREC.

However, the techniques used in this system and in its previous local archive version [7] are almost same. The evaluation data should be able to technically level the performance of the system.

We refer to the milestones described in [1] and designed a set of 50 questions, which covered all 16 Arthur Graesser's questions categories and three other question categories that ranged from easy to very difficult. The test result is very encouraging and the accuracy is 72% in top 1 and 80% in top 5 (Table 1).

We also compared our search engine results with the search result from the LookSmart search engine⁶ used by CNN Web site, and the result from the Google site search⁷. We conclude that our system outperforms these systems.

Question-sentence matching formula used in original AnswerBus system was proved effective in Web-based QA system. However, in the new QA system, it is not working as good as in original AnswerBus QA system. The possible reasons include: 1) The text in CNN Web site is much more formal and the style is much more unique; 2) Fewer redundant information can be found in CNN Web site.

Restricted to the contents of CNN Web site, the system seems working better for news or politics related questions.

⁶ http://cnn.looksmart.com/r_search

⁷ <http://www.google.com/search?q=site:cnn.com>

5 Conclusion

Based on our experiment on our new QA system, we found that QA specific indexing and searching are quite feasible. Most techniques used in original AnswerBus System can be scalable to large size knowledge base and still gets high accuracy.

References

- [1] John Burger et al. Issues, Tasks and Program Structures to Roadmap Research in Question & Answering (Q&A). NIST: http://www-nlpir.nist.gov/projects/duc/papers/qa.Roadmap-paper_v2.doc 2001
- [2] Zhiping Zheng. AnswerBus Question Answering System. *Human Language Technology Conference (HLT 2002)*. San Diego, CA. March 24-27, 2002.
- [3] Zhiping Zheng. Developing a Web-based Question Answering System. *The Eleventh World Wide Conference (WWW 2002)*. Honolulu, HI. May 7-11, 2002.
- [4] Zhiping Zheng. Rule-based Sentence Segmentation for HTML/TEXT Documents. *The Thirteenth meeting of Computational Linguistics in the Netherlands (CLIN 2002)*. Groningen, Netherlands. November 29 2002.
- [5] Zhiping Zheng. Seven Tones: Search for Linguistics and Languages. *The 2nd Meeting of the North American Chapter of Association for Computational Linguistics (NAACL 2001)*. Pittsburgh, PA. June 2-7, 2001.
- [6] Zhiping Zheng and Gregor Erbach. Specialized search in linguistics and languages. *XI International Conference on Computing (CIC 2002)*. Mexico City, Mexico. November 25-29, 2002.
- [7] Zhiping Zheng, Huiyan Huang and Sven Schmeier. Deploying Web-based Question Answering System to Local Archive. *Fifth International Conference on TEXT, SPEECH and DIALOGUE (TSD 2002)*. Brno, Czech Republic. September 9-12, 2002.

Table 1. Evaluation on AnswerBus Local Archive

Question Type	Number	Top1	Top5	Wrong
1. Verification	3	1	1	1
2. Comparison	2	0	1	1
3. Disjunctive	2	2	0	0
4. Concept Completion	6	5	0	1
5. Definition	6	5	0	1
6. Example	3	3	0	0
7. Interpretation	3	2	1	0
8. Feature Specification	5	5	0	0
9. Quantification	6	4	0	2
10. Causal antecedent	3	2	0	1
11. Cause Consequence	0	0	0	0
12. Goal orientation	1	1	0	0
13. Enablement	0	0	0	0
14. Instrumental/Procedural	1	1	0	0
15. Expectational	1	1	0	0
16. Judgmental	3	1	0	0
17. Assertion	3	1	1	1
18. Request/Directive	0	0	0	0
19. Nils question	2	0	0	2

