

# Syntactic Features in Question Answering

Xiaoyan Li

Center for Intelligent Information Retrieval  
Department of Computer Science  
University of Massachusetts Amherst, MA 01003  
[xiaoyan@cs.umass.edu](mailto:xiaoyan@cs.umass.edu)

## ABSTRACT

Syntactic information potentially plays a much more important role in question answering than it does in information retrieval. Although many people have used syntactic evidence in Question Answering, there haven't been many detailed experiments reported in the literature. The aim of the experiment described in this paper is to study the impact of a particular approach for using syntactic information on question answering effectiveness. Our results indicate that a combination of syntactic information with heuristics for ranking potential answers can perform better than the ranking heuristics on their own.

## Categories and Subject Descriptors

H.3.3 Information Search and Retrieval

## General Terms

Algorithm, Experimentation

## KEYWORDS

Syntactic information, question answering

## 1. INTRODUCTION

Question answering (QA) is a task different from information retrieval (IR). Questions submitted to QA systems are full sentences instead of 2-3 keywords typically given to web search engines. Therefore, syntactic information about how a question is phrased and how sentences in documents are structured potentially provides important clues for the matching of the question to answer candidates in the sentences.

In this paper, we present a particular approach to incorporating syntactic information in question answering. In this approach, both questions and selected sentences in top documents are parsed. Syntactic information is extracted from the parser output and used in the answer selection process. There are general syntactic clues that apply to all types of questions, such as matching of phrases in the question and the distance between the main verb and an answer candidate in a sentence. There are also some specific syntactic patterns that apply to different types of

questions.

We have noted that other researchers have used syntactic information in their QA systems [12,3,4,6]. However, we have carried out detailed experiments on the comparison of system performance with and without syntactic information in addition to the differences of how to use syntactic information. In this paper, to study the impact of syntactic evidence on the effectiveness of question answering, a baseline QA system and a new QA system are implemented. The baseline QA system is based on the QA techniques and heuristics that are similar to that used in other QA systems [5,7]. In the new QA system, syntactic information is combined with the heuristics to further improve the accuracy of answer selection. The experimental results show that the combination of heuristics and syntactic information outperform the baseline QA system that used heuristics alone.

## 2. QA WITH HEURISTIC RANKING

In question answering, either an answer or a ranked list of answer candidates is expected. Typically answer candidates are sorted by their belief scores, which are calculated using heuristics or other techniques. Heuristic ranking techniques are common in QA systems. We used heuristic ranking in the baseline QA system, which consists of three main components: query processing module, search engine, and answer extraction module. In the query processing module, each question is classified and the type of answer that this question expects is determined. A query is then generated, and is sent to the INQUERY search engine. The search engine takes the query, searches in its data collection and returns the top 10 documents that the search engine believes they are more likely to have correct answers. In the answer extraction module, answer candidates are extracted and their associated scores are calculated. An answer candidate is a named entity identified by the Identifinder and its type is the same as the question expects. The named entity will not be considered as an answer candidate if it also appears in the question. The heuristic score in the baseline QA System is calculated by the following equation

$$heu\_score = N + 0.5 * Sm + N/W + 0.5/D \quad (1)$$

where four heuristics are considered: the number of matching query words ( $N$ ), whether the matching words are in the same sentence ( $Sm=0/1$ ), the size of the best matching window ( $W$ ), and the distance between an answer candidate and the center of the best matching window ( $D$ ). The answer candidates then are ranked according to their scores and the answer candidate with the highest score appears at the top of the list.

## 3. COMBINING SYNTACTIC INFO WITH HEURISTIC RANKING TECHNIQUES

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference '00, Month 1-2, 2000, City, State.  
Copyright 2000 ACM 1-58113-000-0/00/0000...\$5.00.

Six factors related to syntactic information are considered in the new QA system and the heuristic score is adjusted accordingly, which makes the final belief score for each answer candidate. Table 1 shows the six syntactic features considered in the new QA system, and the syntactic score is given by equation (2).

$$syn\_score = 1.0 * F1 + 0.5 * F2 + 0.5 * F3 + 1.0 * F4 + 1.0 * F5 + 1.0 * F6 \quad (2)$$

where  $F_i$  ( $i=1, \dots, 6$ ) are defined in Table 1. The weights of each factor considered in this process are currently assigned manually, based on our observations of how important the factors are. The weights are assigned 1 if we think their corresponding factors are more important. All the other weights are simply assigned 0.5. The final belief score for each candidate is then calculated using the following equation

$$Final\_score = heu\_score + syn\_score \quad (3)$$

The ranking program ranks candidates for each question by the belief score and the top 5 responses are output.

**Table 1. Six syntactic factors in the new QA system**

F1	Match the sentence with the phrases extracted from the question. If a longer phrase is matched, then the short phrases within it will not be further considered. F1 = the size of total matched phrases/the size of the question.
F2	Consider the distance between the answer candidate and the main verb in token offset. F2 = the distance between the answer candidate and the main verb.
F3	For "PERSON", check the relationship between the answer candidate and the main verb in the sentence to see if it is consistent with the relationship in the question. Some pre-defined syntactic patterns were used here to understand whether it is a "passive" or "active" relationship. F3 = 1 if factor 3 is satisfied, 0 otherwise.
F4	For "LOCATION" questions, check the possessive formats such as, "Venezuela's Orinoco". F4 = 1 if factor 4 is satisfied 0 otherwise.
F5	For "LOCATION" and "DATE" questions, check whether the candidate is inside a prepositional phrase and modifies the main verb. F5 = 1 if factor 5 is satisfied 0 otherwise
F6	For "PERSON" questions, check whether the candidate and all query words are inside a NPA (adjective noun phrase). F6 = 1 if factor 6 is satisfied 0 otherwise.

Experiments are run with TREC-9 questions. The questions are selected according to two criteria. First, their question-types can be determined by the question-classifier that is used in the QA system, and the expected named entities can be recognized by BBN's Identifinder. Second, the correct answer can be found in the top 10 documents returned by INQUERY search engine. 162 questions are left and the experimental results are given in tale 2.

Two evaluation measures are used for comparison. The first evaluation measure is the mean reciprocal answer rank from TREC-9., the new QA system incorporating syntactic information achieves 0.744 over 162 questions, comparing to 0.690 in the baseline QA system. The new QA system outperforms the baseline by 7.8%. The second evaluation measure is the number of questions whose correct answer can be found in the top rank. For 162 questions, there are 94 questions that the correct answer can be found in the top rank using the baseline QA system. There

are 105 questions that the correct answer can be found in the top rank using the new QA system. That indicates the new QA system performs approximately 11.7% better than the baseline QA system in terms of this measure.

**Table 2. Experimental Results**

Question Type	All	Person	Location	Number	Date	Organization
N <sub>questions</sub>	162	57	56	15	25	9
MRR-base	0.690	0.686	0.668	0.650	0.778	0.667
MRR-new	0.743	0.773	0.753	0.724	0.690	0.667
Change	0.054	0089	0.085	0.074	-0.088	0
Change of %	7.8%	13.0%	12.7%	11.4%	-12.7%	0
N <sub>improved</sub>	32	12	14	4	2	0
N <sub>decreased</sub>	14	3	4	1	6	0

A paired t test has been done. At 90% confidence level, the performance of the new system is significantly different from that of the baseline system.

## 4. CONCLUSIONS

Syntactic information potentially plays a much more important role in question answering than it does in information retrieval. Our experimental results indicate that a combination of syntactic information with heuristics for ranking potential answers can outperform the ranking heuristics on their own. The heuristics are also useful for helping filter out passages that are unlikely have correct answers, providing "back off" answers and calculating base belief scores that will be adjusted after considering syntactic information.

## 5. ACKNOWLEDGMENTS

This work was supported in part by the Center for Intelligent Information Retrieval, in part by SPAWARSSYSCEN-SD grant numbers N66001-99-1-8912 and N66001-02-1-8903, and in part by Advanced Research and Development Activity under contract number MDA904-01-C-0984. Any opinions, findings and conclusions or recommendations expressed in this material are the author(s) and do not necessarily reflect those of the sponsor.

## 6. REFERENCES

- [1] C.L.A. Clarke, et al, "Question Answering by Passage Selection", TREC-9, (2000).
- [2] O. Ferret, B. Grau, G. Illouz, C. Jacquwin, and N. Masson, "QALC - the Question-Answering program of the language and Cognition group at LIMSI-CNRS", TREC-8, (1999).
- [3] S. Harabagiu, D. Moldovan et al., "FALCON: Boosting knowledge for answer engines", TREC-9, (2000).
- [4] D.A. Hull, "Xerox TREC-8 Question Answering Track Report", TREC-8, (1999).
- [5] X. Li and W.B. Croft, "Evaluating Question Answering Techniques in Chinese", Proc. HLT 01, 96-101, (2001)
- [6] K.C. Litkowski, "Question-Answering Using Semantic Relation Triples", TREC-8, (1999).
- [7] D. Moldovan et al, "LASSO: A Tool for Surfing the Answer Net," TREC-8, pp 175-183. (1999).