Pattern Optimization and the Application in Question Answering

Yongping Du Ming He
.Institute of Computer Science, Beijing University of Technology, Beijing
{ypdu,heming}@bjut.edu.cn

Abstract

Open Domain Question Answering (QA) represents a challenge of natural language processing, aiming at returning exact answers in response to natural language questions. Recently, it has been found that pattern matching is an effective strategy for QA. The technique of pattern optimization and its applications in QA are demonstrated in the paper. Named entity recognition is used to optimize the answer patterns. It makes the answer patterns have better extending ability. There are two important applications of pattern in QA. Pattern-based query expansion retrieves more exact snippets including the correct answer and the performance is improved. Answer extraction which applies the pattern matching approach is also effective.

Keywords: *Question Answering, Pattern Optimization, Performance Evaluation*

1. Background

Question answering has recently received much attention from the natural language processing communities. The Text REtrieval Conference (TREC) Question Answering track provides a large-scale evaluation for open domain question answering systems. The goal of question answering is to retrieve answers to questions rather than documents as most information retrieval systems currently do.

The traditional approach to Question Answering is to combine the strengths of Information Retrieval (IR), Natural Language Processing (NLP) and Information Extraction (IE) to locate answers to questions within a large text collection[1].

An integrated QA system usually consists of three modules.

Question Analysis Module. This module determines the semantic answer type which will offer help for answer extraction, such as the person name, location name and so on. At the same time, it also translates natural language questions into queries for the search engine.

Search Module. The relevant documents or snippets are retrieved using the queries generated by the first module. These retrieved results can potentially answer the question. This process dramatically reduces the range of the documents used for answer extraction.

Answer Extraction Module. It analyzes the documents or snippets retrieved in the second module, and then extract the answer to the question. This module accomplishes the goal of the QA system finally. For example, question "What country is the holy city of Mecca located in?" is the input of the QA system and the answer "Saudi Arabia" is returned as the output.

The key difference between question answering and usual information retrieval is the answer extraction module, which extracts the precise answer. This module can be realized by various strategies, such as logical inference[2], utilizing WordNet external knowledge resources[3] and pattern matching [4][5].

2. Related Work

Expressions of the answer are flexible in the corpus and it causes great difficulty in identifying the precise answers. Recently, it has been found pattern matching is an effective strategy for QA. The most important task is to build a perfect pattern knowledge base. Among the research groups participated in TREC evaluation, Martin and Sergei uses simple patterns extracted from the corpus and they build the enormous pattern knowledge base manually[6]. They have achieved good performance in TREC but their patterns cannot be learned automatically. In addition, other research institutions accomplish the pattern learning automatically[4][5]. A serious limitation of these patterns is that they contain the morphology without any external knowledge, such as semantic information. We hope to obtain patterns combined with the semantic information for better expansibility and reliability.

The wealth of information on the Web makes it an attractive resource and many systems have make use of the Web knowledge[7][8]. We also take advantage of the variety on Internet for learning different answer patterns in QA. For each type of the question, the corresponding answer patterns can be learned from the Web automatically.

978-0-7695-3382-7/08 $25.00 © 2008 IEEE
DOI 10.1109/ISDA.2008.59

It is noteworthy that our patterns combine the semantic information and have the better robustness. The technique of pattern optimization and its applications in QA are demonstrated in the paper. Named entity recognition is used to generalize the answer patterns. It makes the answer patterns have better extending ability. The constituent elements of answer pattern contain both morphological and semantic information with better robustness. There are two important applications of the pattern in QA. Pattern-based query expansion retrieves more exact snippets including the correct answer and the performance is improved. Another application is answer extraction which applies the pattern matching approach.

This paper first introduces the pattern acquisition and pattern optimization in section 3, and then presents two important applications of the pattern in section 4. Finally, it is the conclusion in section 5.

3. Pattern Acquisition and Optimization

3.1 Pattern Acquisition

Answer patterns describe various context of the answer sentence, and they will be used in the QA system. We acquire corresponding answer patterns according to different question types.

Not all the answer patterns have high credibility and they may extract the wrong answer. Thus, it's necessary for us to evaluate these patterns in order to enhance the overall performance of QA system. We use the Confidence from data mining as the evaluation guideline. The patterns with higher confidence can extract the correct answer in higher probability, while the patterns with lower confidence can not guarantee its accuracy.

More details about answer pattern acquisition and evaluation algorithms are described in [9].

The example of question type and answer patterns with confidence are shown in Table 1.

Here, <A> represents the answer to the question, and the other labels represent different elements of the question as following.

Q_LCN= "Hawaii"

Q_DoVerb= "become"

Q_BNP= "a state"

During the process of pattern matching for answer extraction, information that matched with <A> will be selected as a candidate answer to the question [9].

The same question type may correspond to different questions, and it can also correspond to many answer patterns. In other words, many answer patterns can be used to answer different questions of the same question type.

Question Type/Question	Answer Pattern	Confidence
When did Q_LCN Q_DoVerb Q_BNP ? (Sample question: When did Hawaii become a state?)	1. Q_LCN Q_DoVerb Q_BNP in <A>.	0.91
	2. Q_DoVerb Q_BNP in <A>, Q_LCN	0.82
	3. in <A>, Q_LCN Q_DoVerb Q_BNP	0.73
	4. <A>, Q_LCN Q_DoVerb Q_BNP	0.45
	5. Q_BNP in <A>, Q_LCN	0.39

Table 1 Example of Answer Patterns with Confidence

Our current pattern knowledge base contains 931 different question types and 7852 answer patterns. Each question type has 8 answer patterns on average.

3.2 Pattern optimization

In the process of the answer pattern application, we find out some potential problems that reduces the generality of the pattern as shown below.

Example 1:

Question Type: When be Q_PRN Q_DoVerb? (Q_DoVerb= "born")

Question: When was Abraham Lincoln born?

Snippets: Abraham Lincoln (1809-1865), the sixteenth president of the United States.

Answer: 1809

Answer Pattern: Q_PRN (<A> - 1865)

(Q_PRN= "Abraham Lincoln")

New question: When was Thomas Jefferson born?

As for this new question "When was Thomas Jefferson born?" that has the same question type, answer pattern "Q_PRN (<A> - 1865)" cannot be applied because the existence of the special word "1865".

Example 2:

Question Type: Where be Q_PRN Q_DoVerb ? (Q_DoVerb= "born")

Question：Where was George Washington born?

Snippets：George Washington was born in 1732 in Virginia, he was raised on a farm established by his great-grandfather.

Answer：Virginia

Answer Pattern：Q_PRN be Q_DoVerb in 1732 in <A> ,

(Q_PRN= "George Washington")

New question：Where was Harry Truman born?

The answer patterns in the above two examples have poor generality when they are used to answer new questions of the same question type. The important factor is that these answer patterns contain information closely related to the original question, such as "1865" and "1732", which are meaningless to the new questions. In order to resolve this problem, we need to generalize this kind of pattern to eliminate the impact of the special information.

After observing these special answer patterns, we find out that the special information is usually some proper nouns, such as date, location and organization. This kind of information is exactly in the named entity category, and so we use named entity recognition tool to generalize these answer patterns. We replace the special information by the corresponding named entity label, and it makes the answer patterns have better robustness. This achieves the goal of optimizing patterns. Here, the named entity labels are shown as following.

Date- DATE； Person name-PRN；

Location- LCN ； Organization-ORG ； Number-NUM

The above example answer patterns after optimization becomes the following.

Example 1 Answer Pattern:

Q_PRN (<A> - DATE)

Example 2 Answer Pattern:

Q_PRN be Q_DoVerb in DATE in <A>,

We generalize the answer patterns as stated above and give the experiment results in TREC13 test set shown in Table2. Here, MRR is the Mean Reciprocal Rank (MRR) score[10], a precision-like measure.

Experiment	Precision	MRR
Result_1	40%	0.5
Result_2	42%	0.52
Result_3	36%	0.36
Result_4	40%	0.46
MRR	0.5	0.52

Table 2 Experiment results of pattern optimization

Result_1 denotes the experiment result without pattern optimization when the QA system answers the question type "When be Q_PRN Q_DoVerb?"

Result _2 denotes the experiment result with pattern optimization when the QA system answers the question type "When be Q_PRN Q_DoVerb?"

Result _3 denotes the experiment result without pattern optimization when the QA system answers the question type "Where be Q_PRN Q_DoVerb?"

Result_4 denotes the experiment result with pattern optimization when the QA system answers the question type "Where be Q_PRN Q_DoVerb?"

At present, we accomplish this kind of optimization on some question types, and the overall experiment results show two phenomena. The first is that a certain amount of questions can be answered correctly after pattern optimization. The second is that there is no negative effect on the questions that could be answered correctly before optimization.

4 Application of the Pattern

The pattern knowledge base is an important resource in QA and there are two important applications. They are query expansion and answer extraction.

4.1 Query Expansion

Basic query for the information retrieval module of QA is based on the syntactic elements, and it achieved good results. But there are still some questions which did not return fragments containing the correct answer. The answer patterns describe the different expressions of the answer. This is a very good resource and we can utilize them for query expansion. The retrieval module can achieve better performance after query expansion.

Not all the answer patterns are used for query expansion. We only select the answer patterns with high confidence. After instantiation, these patterns are used as the queries. An example is shown below.

Question：Where is Mount Olympus?

(Q_LCN= "Mount Olympus")

Answer pattern： Q_LCN is located in <A> .（Confidence=0.85）

Answer pattern instantiation： Mount Olympus is located in <A> .

Query Expansion："Mount Olympus is located in"

As the above query is submitted to Google, many fragments containing the correct answer are returned with higher probability, such as "…Mount Olympus is located in Northern Greece ...". This will help us to position the correct answer more accurately.

The performance evaluation of search engine is precision and recall commonly.

$$\Pr ecision = \frac{\text{the retrieved relevant fragments}}{\text{the retrieved fragments}} \quad (Eqn1)$$

$$\mathrm{Re}call = \frac{\text{the retrieved relevant fragments}}{\text{all relevant fragments in the document set}} \quad (Eqn2)$$

We implement the retrieval process using Google and it is impossible to decide all relevant fragments based on the Web information. And so we cannot determine the value of recall. Here, we evaluate the retrieval results in different queries by precision.

We select the *where* questions in TREC13 test set for evaluation, and the evaluation results based on basic query and expansion query are shown in Table 3.

	Basic Query	**Expansion Query**
P@5	42.8%	57.1%
P@10	50%	78.5%
P@20	82.1%	82.1%
P@30	82.1%	82.1%

Table 3. Evaluation results of retrieval module

Here, P@n represents the retrieval precision when the top n results are retrieved by the search engine. Different questions may have different number of patterns used for query expansion. Here, we only check the top 30 results to each query. The user will not look out the results which are listed behind in the ranking commonly.

As shown in Table 3, the precision on P@5 and P@10 to the expansion query get the better performance than that to the basic query. However, the basic query and expansion query have the same precision on P@20 and P@30. The experimental results indicate that the search engine can rank the exact snippets to the forefront of the results list after the query expansion, and the efficiency of retrieval is improved effectively.

4.2 Answer Extraction

The answer patterns can be used to extract the precise answer. This is the most important application which can accomplish the final goal of QA system. The answers of the questions are published by TREC and we have done several experiments to test the pattern matching approach using the TREC question and answer test set.

For each answer pattern and each snippet returned by search module, the words matching tag <A> is selected as the candidate answer. Sort the candidate answers by their answer pattern's confidence and their frequency, and the one with the highest score is selected as the final answer. The answer extraction method using pattern is described in [9].

The system performance according to different interrogatives is shown in Figure 1. Here, the results of (a), (b), (c), (d) denotes the experiment on TREC8, TREC9, TREC10 and TREC11 test set respectively.

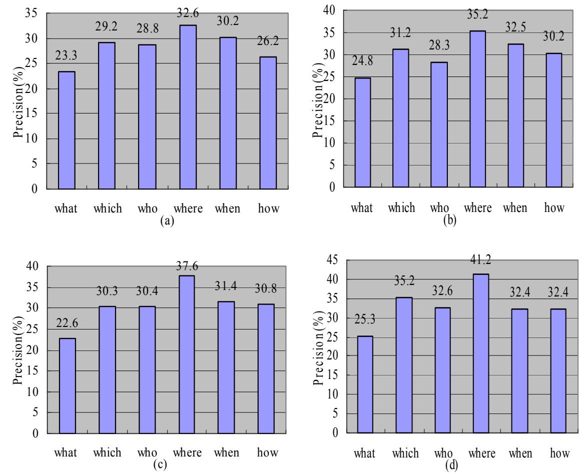

Figure 1. System result according to different interrogatives

We found that the *where* questions have the most stable performance among various interrogatives. It gets the highest precision in different test set. On the contrary, the what questions get the bad precision at all the times. We analyze the reason to this phenomenon.

The answer type of the *where* question is steadfast relatively, and it is the LOCATION name which belongs to the named entities category. The answer context expressions of this kind question are also general and the answer pattern can play an effective role to find the answer.

However, it is difficult to determine the answer type of the *what* question. Most of these questions have different answer types and they are not belonged to the named entity category. And so it cannot be recognized correctly. In addition, the answer context expressions of this kind question are varied and the answer patterns cannot coverage all of them.

5 Conclusions

Pattern matching is an effective method to achieve the answer in QA system. This paper presents the pattern optimization and its applications in QA system.

We take advantage of the variety on Internet for learning different answer patterns in QA. For each type of the question, the corresponding answer patterns can be learned from the Web automatically.

To make the answer patterns have better expansibility, we combine the semantic information to the patterns. The constituent elements of the answer pattern contain both morphological and semantic information with better robustness.

The pattern has two major applications in QA system, and that is query expansion and answer extraction. We instantiate the answer patterns with high confidence for query expansion. The experiment results show that the pattern based query expansion can rank the exact snippets in forefront of the retrieval result list, which makes the search module get better performance. The pattern based answer extraction also achieves the satisfied results.

We know that the expression of language is rich and diverse, and the pattern is not always exhaustive. We will extend our pattern knowledge base and optimize them continually in the future work.

Acknowledgement

This work is supported by the PH.D research foundation of Beijing University of technology under No. 52007012200701.

10. References

[1] Yang, H., Chua, T., Wang, and S. Koh,C. Structured use of external knowledge for event-based open domain question answering. In Proceedings of ACM SIGIR. 2003. New York.

[2] Sanda. Harabagiu, Dan. Moldovan, Marius. Pasca, Mihai. Surdeanu, Rada. Mihalcea, Roxana Girju, Vasile Rus, Finley Lacatusu, Paul Morarescu, Razvan Bunescu. Answering complex, list and context questions with lcc's question answering server. In Proceedings of the TREC-10 Conference. 2001. NIST, Gaithersburg, MD, 355–361.

[3] Moldovan. D, and Mihalcea. R. Using WordNet and lexical operators to improve Internet searches.. IEEE Internet Computing (January-February 2000). 2000: 34-43.

[4] Deepak Ravichandran and Eduard Hovy. Learning Surface Text Patterns for a Question Answering System. In Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL-2002). Philadelphia. 2002: 41-47.

[5] Dell. Zhang, Wee Sun. Lee. Web based Pattern Mining and Matching Approach to Question Answering. Proceedings of the eleventh Text REtrieval Conference. NIST, Gaithersburg, MD, US. 2002: 505-512.

[6] Martin. M. Soubbotin and Sergei. M. Soubbotin. Patterns of Potential Answer Expressions as Clues to the Right Answer. Proceedings of the TREC-10 Conference. NIST, Gaithersburg, MD, US. 2001: 175-182.

[7] Charles L.A. Clarke, Gordon V. Cormack, Thomas R. Lynam. Exploiting redundancy in question answering. Proceedings of the 24th ACM Conference on Research and Development in Information Retrieval (SIGIR-2001) New Orleans, LA. ACM Press. New York. 2001: 358–365.

[8] Cody C.T. KWOK, Oren Etzioni, Daniel S.Weld. Scaling Question Answering to the Web. Proceedings of the 10th World Wide Web Conference (WWW'2001). Hong Kong . 2001:150-161.

[9] Yongping Du, Xuanjing Huang, Lide Wu. Effectively Implementing a Pattern Learning Method in the Question Answering System. Journal of Computer Research and Development. Vol.43, No.3. 2006: 449-455.

[10]Voorhees, E. 2001. Overview of the Question Answering Track. Proceedings of the TREC-10 Conference. NIST,Gathersburg, MD, 157-165.