Towards Interactive QA: suggesting refinement for Questions

Yang Tang
Tsinghua University
tangyang9@gmail.com

Fan Bu
Tsinghua University
bufan000@gmail.com

Zhicheng Zheng
Tsinghua University
zhengzc04@gmail.com

Xiaoyan Zhu
Tsinghua University
zxy-dcs@tsinghua.edu.cn

ABSTRACT
The user's intent can be understood better in a question answering system if there are interactions between the user and the system. Consequently, more accurate answers may be served. In this paper, we propose a method to recommend refinement keywords for an input question to facilitate users to make themselves clear in questioning: an important step for interactive question answering when the question is unclear or unspecific. In this paper, we utilize similar questions to explore the refinements of the input question. We show these refinements in rhetorical questions based on Hownet (Chinese WordNet), which help the user identify the unclearness in the question. Experiments show that the precision of recommendation is about 80%.

Categories and Subject Descriptors
H.3.4 [Systems and Software]: Question Answering

General Terms
Algorithm, Experiment

Keywords
Question Refinement, Question Answering, Hownet

1. INTRODUCTION
Nowadays, the major way of interacting with the search engines is typing several keywords into the search box. In order to help the user express their inquiry intention more clearly, some researchers proposed to use natural language questions instead of keywords queries[1]. Compared with keywords, natural language question is able to express detailed relationship among the inside words. For example, when a user types *Flight Beijing New York*, it's impossible to know whether he wants to find *the flight from Beijing to New York* or *the flight from New York to Beijing*. If he types the exact question *what is the flight number from New York to Beijing*, then his inquiry intention will be much more clearly. Although natural language question is more powerful at expressing user's query intention, it faces a similar problem with the keywords query: the user may forget to supply some constraint context. For example, when a user asks for the question *what is the best restaurant*, he may forget the location constraint for the restaurant. Without the constraint, the system is not able to give a useful answer to him. In such case, question refinements suggestion is quite helpful.

We define the question refinements task as follows: Given an input question, question refinement suggests a list of keywords which provide more specific or restrictive context for the original question.

Suggesting question refinements faces two main challenges: (i) Question style queries are far less than keywords queries. Question refinement is similar with query refinement[2]. However, since the query refinement algorithms make heavy use of query logs, they are unsuitable to be directly applied to question refinement. (ii) There are too many possible constraint context words. These context words need to be well organized so as to be user friendly.

In this paper, we utilize similar questions to explore refinements and show them in rhetorical questions based on Hownet (Chinese WordNet). First, we retrieve a set of similar questions by searching the initial question in a question set. Then, we extract the refinement words (the refinment words reflect the subtopics of the initial question) of these similar questions and map the refinement words to Hownet. Finally, the refinement words are clustered and shown with the form of rhetorical questions. The experiment shows that the algorithm generates accurate question refinements.

The rest of the paper is organized as follows. We first present our approach to suggest question refinements. Then we show some case studies to evaluate our approach. Finally, we conclude the work by summarizing our contributions and outlining future work.

2. Methodology
Question refinement is to suggest some refinement words to the initial question. Different from query refinement, we can't get sufficient refinement information from the query log. In this paper, we propose to find the refinement words in a set of questions which are similar to the initial question.

We denote the initial question as Qr, and it can be represented as a set of words: $Qr = \{w_1, w_2 \ldots w_n\}$. With the vector space model, we retrieve a set of similar questions. We use V to denote the word set consisting of the words that come from the similar questions. Notice that we aim to suggest refinement words from V, so we filter out those words already existing in Qr. However, even after filter, not all of the words in V are suitable to be refinement words. There are two kinds of exceptions: (i) the words which are tightly related to the words in Qr, (ii) the words which are too common to be a refinement words. We illustrate the two exceptions with an example. A user input the initial question *where is the KFC*. Then we retrieve several similar questions, such as *where is the KFC restaurant in the Beijing Airport*, *where is the KFC restaurant near Beijing Zoo*[1], etc. Besides the words in the initial question, V still contains the following words: *restaurant*, *near*, *in*, *Beijing Zoo*, *Beijing Airport*. Among these words, *restaurant* is not a refinement word, since when we mention *KFC*, it implies *KFC restaurant*; *near* and *in are* not refinement words either, since they are both common words. To filter the two kinds of words from V, we use two metrics: PMI[3]

SIGIR Workshop on "entertain me": Supporting Complex Search Tasks, July 28, 2011, Beijing.

[1] The questions are Chinese questions in Baidu Zhidao, see in: http://zhidao.baidu.com/question/94389482.html?an=0&si=1 http://zhidao.baidu.com/question/61613739.html?an=0&si=1

and IDF[4]. PMI(w_1, w_2) reflects the relatedness between w_1 and w_2. The higher the PMI is, the stronger the relatedness between the two words is. In order to filter out the words of first exception, we set an upper bound of PMI (denoted as u). For each word v in V, if there is a word w in Qr satisfies that PMI(v, w) > u, then v is excluded from V. IDF(w) reflects the commonness of a word w. The smaller the IDF is, the more common the word is. In order to filter out those common words, we set a lower bound of IDF (denoted as l). For each word v in V, if it satisfies that IDF(v) < l, then v is excluded from V.

After filtering out the words which are not refinement words to the initial question, we get a set of refinement words (denoted as C). To well organize these refinement words, we use ontology to cluster these words. In practice, we handle with Chinese questions, hence we use Hownet as ontology. Similar with Wordnet, Hownet organize the Chinese words in an ontology form. We cluster the refinement words with following steps. (i) For each constraint word c in C, we map it to a node m in Hownet. Since Hownet is ontology, we can find a path from the root of the ontology to the node m. We denote the path as $\{m_{c1}, m_{c2} \ldots m_{ck}\}$, where m_{c1} is the root and m_{ck} is the node which c is mapped to. (ii) With the paths, we can build a sub tree of the ontology. For example, assume there are three words in C, denoted as $\{w_1, w_2, w_3\}$, and we find three paths: $\{a, b\}$, $\{a, c, d\}$, $\{a, c, e\}$. Then we form a sub tree as shown in Figure 1. Since each node exists at least in one path, each node corresponds to a set of constraint words. In the example, node c is associated with w_1 and w_2. (iii) Scan the nodes in the sub tree in a bottom-up way. If the node contains more than t refinement words, the node is extracted to represent the cluster consisting of the refinement words in it. Then we remove the node and its refinement words from the sub tree. After we remove all the refinement words, we end the clustering process.

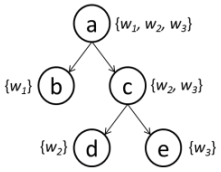

Figure 1. Sub tree of the illustration example.

We can organize these words by clusters instead of simply listing all the refinement words. We design some rhetorical question templates for different clusters. We still take the example *where is the best restaurant*. Different from simply listing all the refinement suggestions, such as *Beijing*, *New York*, *Seafood*, *Steak*, we show the suggestion in a rhetorical question form. With the clustering algorithm, we obtain two clusters from these constraint words. The two clusters are mapped to the nodes *City* and *Food-Type* in Hownet. With manual created templates, we suggest the following text: *Which city it locates? Beijing, New York, etc.* and *Which food type? Seafood, Steak, etc.* In this way, we can help the user identify the missing constraints easier.

3. Experiment

We set up an experiment on a set of Chinese questions. We sampled 100 questions as the initial question set. These questions satisfy the following conditions: (i) The length of the question is less than ten; (ii) We find at least 20 questions which are similar to the question in our question set.

We apply the question refinement suggestion on the initial question set. We get 138 rhetorical question suggestions. Manual judgments show that 111 of them are correct. The precision of question suggestion is 80.4%. We also ask users to manually write question refinements first. Then we compare the system refinement results with manual results. The recall is 56%.

We also show some suggestion examples in Table 1. In the first example, our algorithm generates two rhetorical question suggestions. One implies that the initial question requires constraint of location; the other implies that the initial question requires constraint of the disease being treated in the hospitals. In the second example, the algorithm also generates two rhetorical question suggestions. The first one well implies the requirement for constraint of the place of origin. However, since the cluster is not well explained by Hownet, the second one implies the requirement for constraint of the movie in a weird way.

Table 1. Two Examples of Question Refinement Suggestion

Initial Question	Question Refinement Suggestion
哪家医院比较好 *(Which hospital is better)*	哪个地方(市)?是上海,苏州, …, 还是深圳? (Which city? Shanghai, Suzhou, …, or Shenzhen?) 哪些疾病?是皮肤病,白癜风, ...,还是心脏病? Which kind of disease? Skin disease, vitiligo, ..., or heart disease?
有什么好看的电影推荐 *(Can you recommend some good movies)*	哪个地方?是欧洲,美国, …, 还是香港? (Where? Europe, America or Hongkong?) 什么人?是死人,匪,间谍还是吸血鬼? (Which kind of person? Dead man, robber, spy, or vampire?)

4. Conclusion

In this paper, we utilize similar questions to generate question refinements, and show these refinements in rhetorical questions. We use PMI and IDF to extract refinement words. Then we cluster the refinement words with Hownet. According to the cluster notation given by Hownet, we apply different rhetorical question templates to different clusters of constraint words, and generate different kinds of rhetorical question suggestions. The experiment shows that the algorithm generates accurate suggestion results. However, since Hownet is not able to well explain all the clusters of constraint words, the algorithm sometimes generates some weird suggestions.

The performance of the algorithm is limited by the coverage of Hownet. In the future work, we aim to explore other resources to cluster and explain the constraint words.

Acknowledgement

This work is supported by Canada's IDRC Research Chair in Information Technology program, Project Number: 104519-006.

5. REFERENCES

[1] Voorhees E M. *The TREC-8 Question Answering Track Report*. Proc. of TREC-8, 1999. 77–82.

[2] Eldar Sadikov, Jayant Madhavan, Lu Wang, and Alon Halevy. 2010. *Clustering query refinements by user intent.* In Proc. of the WWW. ACM, New York, NY, USA, 841-850.

[3] Manning C D and Sch¨utze H. *Foundations of Statistical Natural Language Processing.* The MIT PRESS, Cambridge,Massachusets, 1999.

[4] Church K W and Gale W A. *Inverse document frequency: a measure of deviation from Poisson.* In A. et al. (Ed.), NLP using Very Large Corpora. Kluwer Academic Publishers, 1999