A Re-examination of IR Techniques in QA System

Yi Chang, Hongbo Xu, and Shuo Bai

Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China
{changyi,hbxu}@software.ict.ac.cn, bai@ncic.ac.cn
http://www.software.ict.ac.cn

Abstract. The performance of Information Retrieval in the Question Answering system is not satisfactory from our experiences in TREC QA Track. In this article, we take a comparative study to re-examine IR techniques on document retrieval and sentence level retrieval respectively. Our study shows: 1) query reformulation should be a necessary step to achieve a better retrieval performance; 2) The techniques for document retrieval are also effective in sentence level retrieval, and single sentence will be the appropriate retrieval granularity.

1 Introduction

Information Retrieval (IR) and Information Extraction (IE) are generally regarded as the two key techniques to Natural Language Question Answering (QA) System that returns exact answers. IE techniques are incorporated to identify the exact answer while IR techniques are used to narrow the search space that IE will process, that is, the output of the IR is the input of IE.

According to TREC QA overview: [11], most current question answering systems rely on document retrieval to provide documents or passages that are likely to contain the answer to a question. Since document-oriented information retrieval techniques are relative mature while IE techniques are still under developing, most of current researches have focused on answer extraction: [5], [8], [9]. There is little detailed investigation into the IR performance which impacts on overall QA system performance. Clarke et al.: [3] proposed a passage retrieval technique based on passage length and term weights. Tellex et al.: [10] make a quantitative evaluation of various passage retrieval algorithms for QA. Monz: [6] compares the effectiveness of some common document retrieval techniques when they were used in QA. Roberts and Gaizauskas: [7] use coverage and answer redundancy to evaluate a variety of passage retrieval approaches with TREC QA questions.

In most of current researches, the granularity for information retrieval in QA is passage or document. What is the potential of IR in QA and what is the most appropriate granularity for retrieval still need to be explored thoroughly.

We have built our QA system based on the cooperation of IE and IR. According to our score and rank on past several TREC conferences, although we are making progress each year, the results are still far from satisfactory. As our recent study shows, IR results in much more loss comparing with IE. Therefore, we re-examine two important questions that have ever been overlooked: 1) whether a question is a good query for retrieval in QA? 2) Whether the techniques for document retrieval are effective on sentence level retrieval?
In this paper, we compare some alternative IR techniques and all our experiments are based on TREC 2003 QA AQUAINT corpus. To make a thorough analysis, we focus on those questions with short, fact-based answers, called Factoid questions in TREC QA.

In Section 2, we describe our system architecture and evaluate the performance of each module. Then in section 3, according to the comparison of four document retrieval methods, we find the reason to limit our retrieval performance. We then present in Section 4 the results of four sentence level retrieval methods and in Section 5 we research different retrieval granularities. Finally, Section 6 summarizes the conclusions.

2 System Description

Our system to answer Factoid questions contains five major modules, namely Question Analyzing Module, Document Retrieval Engine, Sentence Level Retrieval Module, Entity Recognizing Module and Answer Selecting Module. Figure 1 illustrates the architecture.

In this paper, a Bi-sentence means two consecutive sentences and there is no overlapping between two consecutive Bi-sentences; a phrase means a sequence of keywords or one keyword in a question, where a keyword is a word in the question but not in Stop-word list.

To answer each question, Question Analyzing Module makes use of NLP techniques to identify the right type of information that the question requires. Question Analyzing Module also preprocesses the question and makes a query for further retrieval. Document Retrieval Engine use the question to get relevant documents and