Classification of Multiple-Sentence Questions

Akihiro Tamura, Hiroya Takamura, and Manabu Okumura

Precision and Intelligence Laboratory,
Tokyo Institute of Technology, Japan
aki@lr.pi.titech.ac.jp
{takamura, oku}@pi.titech.ac.jp

Abstract. Conventional QA systems cannot answer to the questions composed of two or more sentences. Therefore, we aim to construct a QA system that can answer such multiple-sentence questions. As the first stage, we propose a method for classifying multiple-sentence questions into question types. Specifically, we first extract the core sentence from a given question text. We use the core sentence and its question focus in question classification. The result of experiments shows that the proposed method improves F-measure by 8.8% and accuracy by 4.4%.

1 Introduction

Question-Answering (QA) systems are useful in that QA systems return the answer itself, while most information retrieval systems return documents that may contain the answer.

QA systems have been evaluated at TREC QA-Track in U.S. and QAC (Question & Answering Challenge) in Japan. In these workshops, the inputs to systems are only single-sentence questions, which are defined as the questions composed of one sentence. On the other hand, on the web there are a lot of multiple-sentence questions (e.g., answer bank, AskAnOwner), which are defined as the questions composed of two or more sentences: For example, “My computer reboots as soon as it gets started. OS is Windows XP. Is there any homepage that tells why it happens?”. For conventional QA systems, these questions are not expected and existing techniques are not applicable or work poorly to these questions. Therefore, constructing QA systems that can handle multiple-sentence questions is desirable.

An usual QA system is composed of three components: question processing, document retrieval, and answer extraction. In question processing, a given question is analyzed, and its question type is determined. This process is called “question classification”. Depending on the question type, the process in the answer extraction component usually changes. Consequently, the accuracy and the efficiency of answer extraction depend on the accuracy of question classification.
Therefore, as a first step towards developing a QA system that can handle multiple-sentence questions, we propose a method for classifying multiple-sentence questions. Specifically, in this work, we treat only questions which require one answer. For example, if the question “The icon to return to desktop has been deleted. Please tell me how to recover it.” is given, we would like “WAY” to be selected as the question type. We thus introduce core sentence extraction component, which extracts the most important sentence for question classification. This is because there are unnecessary sentences for question classification in a multiple-sentence question, and we hope noisy features should be eliminated before question classification with the component. If a multiple-sentence question is given, we first extract the most important sentence for question classification and then classify the question using the only information in the sentence.

In Section 2, we present the related work. In Section 3, we explain our proposed method. In Section 4, we describe our experiments and results, where we can confirm the effectiveness of the proposed method. Finally, in Section 5, we describe the summary of this paper and the future work.

2 Related Work

This section presents some existing methods for question classification. The methods are roughly divided into two groups: the ones based on hand-crafted rules and the ones based on machine learning. The system “SAIQA” [1], Xu et al. [2] used hand-crafted rules for question classification. However, methods based on pattern matching have the following two drawbacks: high cost of making rules or patterns by hand and low coverage.

Machine learning can be considered to solve these problems. Li et al. [3] used SNoW for question classification. The SNoW is a multi-class classifier that is specifically tailored for learning in the presence of a very large number of features. Zukerman et al. [4] used decision tree. Ittycheriah et al. [5] used maximum entropy. Suzuki [6] used Support Vector Machines (SVMs). Suzuki [6] compared question classification using machine learning methods (decision tree, maximum entropy, SVM) with a rule-based method. The result showed that the accuracy of question classification with SVM is the highest of all. According to Suzuki [6], a lot of information is needed to improve the accuracy of question classification and SVM is suitable for question classification, because SVM can classify questions with high accuracy even when the dimension of the feature space is large. Moreover, Zhang et al. [7] compared question classification with five machine learning algorithms and showed that SVM outperforms the other four methods as Suzuki [6] showed. Therefore, we also use SVM in classifying questions, as we will explain later.

However, please note that we treat not only usual single-sentence questions, but also multiple-sentence questions. Furthermore, our work differs from previous work in that we treat real data on the web, not artificial data prepared for the QA task. From these points, the results in this paper cannot be compared with the ones in the previous work.