

## EFFECT OF USING PRAGMATICS INFORMATION ON QUESTION ANSWERING SYSTEM OF ANALECTS OF CONFUCIUS

YE YANG<sup>1</sup>, PEILIN JIANG<sup>1</sup>, SEIJI TSUCHIYA<sup>2</sup> AND FUJI REN<sup>2,3</sup>

<sup>1</sup>Graduate School of Advanced Science Technology Education

<sup>2</sup>Graduate School of Sociotechno Science

The University of Tokushima

Minamijyosanjima 2-1, Tokushima-shi, Tokushima 770-8506, Japan

{ yang; jiang; tsuchiya; ren }@is.tokushima-u.ac.jp

<sup>3</sup>School of Information and Engineering

Beijing University of Post and Telecommunications

Haidian District, Beijing 100876, P. R. China

Received March 2008; revised July 2008

**ABSTRACT.** *In general, the techniques of statistical retrieval and shallow language analysis are chiefly used in question answering(QA) systems in order to improve the accuracy of answers. But these techniques are not contributing effectively to restricted domain QA system of classic Chinese literature such as “Analects of Confucius”. This QA system requires to extract related verses and automatically answer the query in natural Chinese language about Confucius’ thought in the “Analects of Confucius”. Therefore we propose a novel method to integrate the pragmatics information with classical information retrieval technique for QA system so as to improve retrieval efficiency. In this paper, we examined the effect of the pragmatics information on QA system of the “Analects of Confucius”. According to the experiments, the pragmatics information based retrieval results are more accurate than the one without using it.*

**Keywords:** Pragmatics information, Utterance interpretation, Information retrieval, Question answering system, Analects of confucius

**1. Introduction.** Currently in the process of informationization, the technologies to automatically answer question according to user’s requirements by querying in a large amount of documents are attracting attention. Question answering (QA) technology aims to retrieve a desirable answer to a question in natural language from a document set. As the components in a QA system, information retrieval, information extraction, summarization, and conversation interface etc are considerably developed and evaluated with MUC [1] and IREX [2] etc.

Recently, the QA system based on the statistical retrieval model [3] and shallow language analysis [4,5] becomes a main direction of narrowing the answer in the question. The construction of a practical restricted domain QA system using ontology with high accuracy is expected. Up to now several Chinese QA systems have been constructed. For example, [6] proposes a QA system which concerns with sightseeing information and combines the question answering database retrieval and the document retrieval of the web text by integrating the statistical technique with the analytical base. Moreover, there is an applied system for the banking business area by using syntactic-analysis information [7]. And [8] is about human relationships in “Dream of the Red Chamber” using the technology of a shallow inference structure mode. But these methods only address the literal meaning and explicitly communicated meaning of the words and the passages, but