

Generating a Set of Rules to Determine Honorific Expression Using Decision Tree Learning

Kanako Komiya, Yasuhiro Tajima, Nobuo Inui and Yoshiyuki Kotani

Department of Computer, Information and Communication Sciences
Tokyo University of Agriculture and Technology
2-24-16, Nakacho, Koganei, Tokyo, Japan, 184-8588
komiya@fairry.ei.tuat.ac.jp

Abstract. In Japanese language, the speaker must choose suitable honorific expressions depending on many factors. The computer system should imitate this mechanism to make a natural Japanese sentence. We made a system to determine a suitable expression and named it honorific expression determining system (HEDS). It generates a set of rules to determine suitable honorific expression automatically, by decision tree learning. The system HEDS determines one out of the three classes for an input sentence: the respect expression, the modesty expression and the non-honorific expression and determines what expression the verb is. We calculated the accuracy of HEDS using the cross validation method and it was up to 74.88%.

1 Introduction

In Japanese language, one must choose suitable honorific expressions depending on the speaker, the addressees, the subject of the utterance, contents of the dialogue and situations in the conversation. The computer system should imitate this mechanism to make a natural Japanese sentence.

Japanese language has the two types of honorific expression: (1) respect or modesty expression and (2) polite expression. The respect expression is used to display respect to others, or their higher rank, and practically to show second person of the sentential implicit subject, in contrast that the modesty expression shows first person. The modesty expression is an expression that one display modesty to respecting persons. These two honorific expressions cannot be used in a single word at the same time, but the combination of (1) and (2) can be used for one word at the same time. We focus on the type (1) in this paper.

2 Honorific Expression Determining System (HEDS)

The user of HEDS provides honorific expressions and its factors for determining suitable honorific expressions as data. Then HEDS generates a set of selection rules. It