

CHINESE-JAPANESE TRANSLATION OF CAUSATIVE SENTENCES USING SUPER-FUNCTION BASED MACHINE TRANSLATION SYSTEM

LIYING MI¹, XIN LUO¹ AND FUJI REN^{1,2}

¹Faculty of Engineering
University of Tokushima
Tokushima 770-8506, Japan
{ mily; luoxin; ren }@is.tokushima-u.ac.jp

²School of Information Engineering
Beijing University of Posts and Telecommunications
Beijing 100876, P. R. China

Received May 2007; revised October 2007

ABSTRACT. *Causative sentences in Japanese are a matter of affixation. In particular, the causative form takes the shape of a suffix. In Chinese, the causative form constitutes an independent “word”. In our previous studies on Super-Function Based Machine Translation (SFBMT), we have found that causative sentences are very frequently used and difficult to translate correctly, the over use of causative sentences can be dangerous as it may introduce ambiguity in the translation. In this paper, we discuss the challenges in handling Japanese causative sentences in an SFBMT system; we present a shallow method for translating causative sentences by using some fixed rules and Super Functions (SF). In the present research, sufficient Chinese-Japanese causative sentence patterns have been employed as a language-database for experiments, which proves the suggested method can effectively improve translation quality within the range under discussion.*

Keywords: Causative sentence, NLP, Machine translation, SFBMT

1. Introduction. With the rapid development of the internet, there are an increasing number of people who obtain information from it. But among the variety of information they find, some of it is often written in language which is different from their mother tongue and not only greatly hinders their reading speed, but also their understanding. So when assessing highly efficient tools to solve this problem, Machine-translation software is naturally taken into consideration. Machine translation (MT) describes computer-based translation between human languages and is one of the oldest large-scale applications of computer science. There is an increasing demand for MT systems which produce high quality translations and which can be easily adapted to many language pairs, new domains and changing terminology. However, in the field of MT, the lack of logic and the accumulation of various mistakes are still common. It is, therefore, very important to improve the accuracy of MT.

Accordingly, in recent years, some improvements have been achieved concerning the correctness and the agreement with the target language, which can be found in Rule-Based Machine Translation (RBMT); Knowledge-Based Machine Translation (KBMT); Sentence-to-Sentence Machine Translation [1, 2, 3, 4]; Sentence-Pattern Machine Translation and Translation-Memory Machine Translation. However, the scope that these models have covered is quite limited.