COMPUTATIONAL METHODS IN ENGINEERING AND SCIENCE
EPMESC X, *Aug. 21-23, 2006, Sanya, Hainan, China*
©2006 Tsinghua University Press & Springer

Application of Translation Corresponding Tree (TCT) Annotation Schema for Chinese to Portuguese Machine Translation

Chi-Wai Tang *, Fai Wong, Ka-Seng Leong, Ming-Chui Dong, Yi-Ping Li

Faculty of Science and Technology, University of Macau, Macau SAR, China
Email: {kevin, derek, sejohnny, dmc}@inesc-macau.org.mo, ypli@umac.mo

Abstract

In Example Based Machine Translation (EBMT) research, there are three main approaches: Surface Based, Pattern Based and Structure Based approach. In Structure Based EBMT system, such as SSTC approach [1], it has a problem that it relies on two syntax parsers to analyze the translation examples, but robust syntax parsers are not always available. On the other hand, Chinese and Portuguese belong to two different language families and there exist grammatical deviation problem between them. In order to resolve the weakness of the Structure Based EBMT system and linguistic problems between Chinese and Portuguese, Tang and Wong [2] propose a new Portuguese to Chinese Machine Translation method and this method is based on a novel technology called Translation Corresponding Tree (TCT) which is an example based knowledge annotation method for Portuguese to Chinese translation.
In this paper, it adopts the TCT annotation scheme and introduces the knowledge based construction and translation for Chinese to Portuguese MT. In this parper, it presents the additional problems during Chinese analysis and the corresponding solutions. In this research, it also proposes a conversion algorithm to reuse the existing translation knowledge of Portuguese to Chinese MT system, which represented in terms of TCT trees. Based on the transformation algorithm, the knowledge trees of Portuguese to Chinese translation is converted into that of the translation knowledge which can be used to facilitate the Chinese to Portuguese . By this conversion method, existing knowledge can be easily reused without having to re-construct the knowledge from scratch. Based on the research result of this paper, a Chinese to Portuguese prototyping machine translation system is implemented and the empirical results show that the MT system can achieve the translation accuracy of 85% in the domain of Macau Law statements.

REFERENCES

1. AI-Adhaileh MH, Tang EK, Zaharin Y. A synchronization structure of SSTC and its applications in machine translation. The COLING 2002 Post-Conference Workshop on Machine Translation in Asia, 2002.
2. Tang CW, Wong F, Li YP. TCT Schema in EBMT and its application. In Proc. the Symposium on Applied Science and Technology in Macau 2004, 2004, pp. 19-27.