

Exploiting a Mono-Bilingual Dictionary for English-Korean Translation Selection and Sense Disambiguation

Hyun Ah Lee¹, Juntae Yoon², and Gil Chang Kim¹

¹ Dept. of EECS, Korea Advanced Institute of Science and Technology (KAIST),
373-1 GusongDong, YusongGu, Daejeon, 305-701, Republic of Korea

² Daumsoft Inc., 946-12 DaechiDong, GangnamGu, Seoul, 135-280, Republic of Korea
halee@csone.kaist.ac.kr, jtyoon@daumsoft.com, gckim@cs.kaist.ac.kr

Abstract. A mono-bilingual dictionary is a combined dictionary of a monolingual and bilingual dictionary, which contains rich information for translation. In this paper, we propose to exploit a mono-bilingual dictionary for sense disambiguation and translation selection. Based on the ‘word-to-sense and sense-to-word’ relationship, our method selects translation through two steps – sense disambiguation of a source word and selection of a target word. Each step requires various information including clue words for sense disambiguation, sense-to-word mapping and syntactic relation mapping information, and we extract them from a mono-bilingual dictionary. Evaluation results showed that our method selects more appropriate target words.

1 Introduction

Translation selection is a process to select, from a set of target language words corresponding to a source language word, one that conveys the correct sense of a source word and makes more fluent target language sentences. Translation selection is a key problem in machine translation (MT) since the quality of translation varies significantly according to results of translation selection.

The difficulty of translation selection and machine translation is that they link two different languages thus requiring more complex knowledge than other problems concerning only one language. So, the following has become important issues – what kind of knowledge to be used for translation, how to acquire and use the knowledge. Among them, knowledge acquisition is a critical problem for many MT systems and translation selection methods.

A bilingual dictionary is a good knowledge source for MT and translation selection. An entry in a bilingual dictionary generally contains translation definitions, and example sentences in source and target languages. On the other hand, an entry in a common monolingual dictionary includes sense definitions and monolingual example sentences. A mono-bilingual dictionary is a combination of a monolingual and a bilingual dictionary, in which a word entry has sense definitions in a source language, translation definitions, and pairs of example sentences in both languages as shown in Fig. 1.