

論文の内容の要旨

論文題目 PIVOTAL APPROACH FOR LEXICAL TRANSLATION
(中間言語を用いた辞書の翻訳)

氏名 綱 川 隆 司

The bilingual lexicon is an expensive but crucial resource for multilingual applications in natural language processing. This thesis proposes an enrichment of the bilingual lexicons used for machine translation in two aspects: the identification of synonyms and a pivotal approach for acquiring lexical translation.

Construction of synonym lists is one of the important natural language processing (NLP) tasks, because the obtained synonym lists can be used by several different NLP applications such as machine translation (MT) and information retrieval (IR). In the first part of this thesis, a machine learning method for identifying synonyms in a bilingual lexicon is presented. Initially we prepare a bilingual lexicon with synonymous information and generate the pairs of translation equivalents, and attach the presence or absence of synonymous relations to each pair. Then, a classifier is learned using training data by employing features related to spelling variations and so forth. The principal contributions of this work are defining the synonymous relations in a bilingual lexicon, proposing features and algorithms for identifying the synonyms, and verifying the effectiveness of the method for a bilingual lexicon with synonymous information. The experimental results show that our proposed method has an F-score of 91.6% on a pairwise evaluation and significantly outperforms the performances of the baselines and the approach using combinations of monolingual synonyms.

In the second part of this thesis, an integrated framework for building a bilingual lexicon between Chinese and Japanese languages through English as the pivot language is proposed. Since the language pair of Chinese-Japanese does not include English, bilingual resources between these languages are smaller than those with English. One solution to this problem is to build a Chinese-Japanese bilingual lexicon through English as the pivot language. In addition to the pivotal approach, we can make use of the characteristic that Chinese and Japanese languages use Han characters. We incorporate a translation model obtained from a small Chinese-Japanese lexicon and the similarity of hanzi and kanji characters by using the log-linear model. Our experimental results show that the use of the pivotal approach can improve the translation performance over the translation model built from the small Chinese-Japanese lexicon. The results also demonstrate that the similarity of hanzi and kanji characters has a positive effect on the translation of technical terms.