

博士论坛

## 基于多重线性回归模型的翻译等价对获取

张春祥, 赵铁军, 李生

哈尔滨工业大学 计算机科学与技术学院

收稿日期 2005-5-9 修回日期 网络版发布日期 接受日期

**摘要** 翻译等价对在词典编纂、机器翻译和跨语言信息检索中有着广泛的应用。本文从双语句对的译文等价树中抽取翻译等价对。使用译文直译率、短语对齐概率和目标语-源语言短语长度差异等特征对自动获取的等价对进行评价。提出了一种基于多重线性回归模型的等价对评价方法, 并结合N-Best策略对候选翻译等价对进行过滤。实验结果表明: 在开放测试中, 基于多重线性回归模型的等价对评价及过滤方法其性能要优于其它方法。

**关键词** [翻译等价对,多重线性回归模型,N-Best策略](#)

分类号

## Research On Automatic Acquisition Method Of Translation Equivalences Based On Multiple Linear Regression

ChunXiang Zhang,TieJun Zhao,Sheng Li

哈尔滨工业大学 计算机科学与技术学院

### Abstract

Translation equivalence is very useful for bilingual lexicography, machine translation system and cross-lingual information retrieval. In this paper, translation equivalences are extracted from translation corresponding trees of bilingual sentence pairs. Translation literality, phrase alignment probability, and length difference from target language phrase to source language phrase are employed to score for extracted equivalences. An evaluation method based on multiple linear regression is proposed. This new approach is employed to filter equivalences combined with N-Best strategy. Experimental results show that the new method does better than other approaches on evaluation and filtering.

**Key words** [translation equivalence](#) [multiple linear regression](#) [N-Best strategy](#)

DOI:

通讯作者 张春祥 zcx51666 zcx51666 [cxzhang@mtlab.hit.edu.cn](mailto:cxzhang@mtlab.hit.edu.cn)

### 扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF\(OKB\)](#)

▶ [\[HTML全文\]\(OKB\)](#)

▶ [参考文献](#)

服务与反馈

▶ [把本文推荐给朋友](#)

▶ [加入我的书架](#)

▶ [加入引用管理器](#)

▶ [复制索引](#)

▶ [Email Alert](#)

▶ [文章反馈](#)

▶ [浏览反馈信息](#)

相关信息

▶ [本刊中 包含“翻译等价对,多重线性回归模型,N-Best策略”的相关文章](#)

▶ 本文作者相关文章

· [张春祥](#)

· [赵铁军](#)

· [李生](#)