

数据库、信号与信息处理

句法调序的统计机器翻译方法研究

孙广范, 宋金平, 肖健, 袁琦

中国电子信息产业发展研究院, 北京 100044

收稿日期 2008-7-11 修回日期 2008-11-17 网络版发布日期 2009-12-30 接受日期

摘要 为解决基于短语统计机器翻译存在的调序能力不足的问题, 尝试利用句法分析器对基于短语统计机器翻译的输入汉语句子进行句法分析, 然后利用转换器进行调序操作, 并对部分类型短语进行预先翻译, 然后再利用基于短语统计机器翻译的解码器进行翻译。重点测试了汉语中“的”字引导的复杂定语调序、介词短语、特定搭配短语、方位词短语的调序及预翻译产生的效果。实验结果表明, 这些调序及预翻译操作可以显著地提高基于短语的统计机器翻译的英文译文结果的BLEU值。

关键词 [统计机器翻译](#) [句法调序](#) [预先翻译](#)

分类号 [TP391](#)

Research on approaches of syntactic reordering for statistical machine translation

SUN Guang-fan, SONG Jin-ping, XIAO Jian, YUAN Qi

China Center for Information Industry Development, Beijing 100044, China

Abstract

To resolve the problem of the lack of capacity for the reordering in Phrase-Based Statistical Machine Translation (PBSMT), this paper attempts to parse the input Chinese sentences to PBSMT, conduct the reordering of the sentences by a converter, pre-translate some types of phrases, and then translate the output from the above operation by a decoder. This paper focuses on the effects of the reordering for complex attributes linked by “de”, and the effects of the reordering and pre-translation of prepositional phrases, specific collocation phrases and phrases of noun of locality. Experimental results show that the reordering and pre-translation can significantly improve the BLEU values of the English translation by PBSMT.

Key words [statistical machine translation](#) [syntactic reordering](#) [pre-translation](#)

DOI: 10.3778/j.issn.1002-8331.2009.36.042

通讯作者 孙广范 morgan2001_sun@sohu.com

扩展功能	
本文信息	
▶ Supporting info	
▶ PDF(505KB)	
▶ [HTML全文](0KB)	
▶ 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“统计机器翻译”的相关文章	
▶ 本文作者相关文章	
· 孙广范	
· 宋金平	
· 肖健	
· 袁琦	