

## 人工智能

一种基于短语统计机器翻译的高效柱搜索解码器

罗毅<sup>1</sup>; 李淼<sup>2</sup>; 张建<sup>1</sup>

中国科学院 合肥智能机械研究所<sup>1</sup>

中国科学技术大学 信息科学技术学院<sup>2</sup>

收稿日期 2007-2-15 修回日期 网络版发布日期 2007-8-27 接受日期

**摘要** 描述了一种基于短语统计机器翻译的柱搜索解码器。搜索算法的效率是解码的关键, 基于传统的柱搜索解码算法, 提出了提高搜索效率的改进措施: 动态剪枝策略改进了原来固定地剪枝对搜索当前情形反应不足的问题, 提高了剪枝精度; 预剪枝策略限制了较差的扩展, 减少了不必要的扩展, 提高了搜索速度; 在研究了当前主要位置重排限制的基础上, 提出了一种快速位置重排限制策略, 加快了位置重排时的解码速度。此外, 针对领域术语翻译唯一性问题提出了专门处理方法以提高翻译的准确度。分析对比实验结果, 证明了算法的有效性。

**Abstract** An efficient beam search decoder for phrase-based statistical machine translation was described. The efficiency of search algorithm is the key to decoding process. After introducing the conventional beam search decoding algorithm, some efficiency improving measures were proposed. Dynamic pruning strategy enhanced the accuracy of pruning by improving that the original fixed pruning had not enough response to the current situation of search. Pre-pruning strategy was used to limit the poor sprawl, reduce unnecessary expansion and improve search speed. A rapid reordering constrains strategy was presented based on the research of the current major reordering constrains. In addition, the domain term always has the only translation, so a special process approach was put forward to improve the accuracy of the translation. Comparative analysis of the experimental results proves the effectiveness of the algorithm.

**关键词** [统计机器翻译](#) [解码器](#) [柱搜索](#) [动态剪枝](#) [位置重排限制](#)

**Key words** statistical machine translation; decoder; beam search; dynamic pruning; reordering constrains

分类号

**DOI:**

通讯作者:

李淼 [mli@iim.ac.cn](mailto:mli@iim.ac.cn)

作者个人主页: 罗毅 李淼 张建

## 扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF \(642KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [参考文献\[PDF\]](#)

▶ [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](#)

▶ [引用本文](#)

▶ [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

· [罗毅](#)

· [李淼](#)

· [张建](#)