

研究论文

基于分类标注语料库的关键词标引知识自动获取

刘华

暨南大学华文学院<sup>1</sup>

海外华语研究中心<sup>2</sup>

收稿日期 2006-6-11 修回日期 网络版发布日期 2007-7-30 接受日期

摘要 基于大规模层级分类语料库,抽取网页上专家已经标引的关键词形成关键词表;针对关键词的领域不均匀性和邻界域两个特征,提出并模拟计算了关键词表征文本主题特征程度的主题度。以关键词及其主题度为领域知识,结合统计方法,完成了一个知识与统计相结合的关键词自动标引系统。

关键词 [关键词标引](#) [分类语料库](#) [主题度](#)

分类号 [G254.361](#) [TP391.1](#)

DOI:

对应的英文版文章: [2006-1951](#)

通讯作者:

刘华 [liuhua0461@sina.com](mailto:liuhua0461@sina.com)

作者个人主页: 刘华

扩展功能

本文信息

- ▶ [Supporting info](#)
- ▶ [PDF](#) (OKB)
- ▶ [\[HTML全文\]](#) (OKB)
- ▶ [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“关键词标引”的 相关文章](#)
- ▶ 本文作者相关文章
  - [刘华](#)