
基于英汉词典和翻译软件的跨语言信息检索实验

王树锋 张永奎

(山西大学计算机科学系 太原 030006)

摘要: 本文介绍了使用双语词典和翻译软件翻译检索条件, 实现跨语言信息检索的实验. 使用了一种简单高效的汉语分词方法, 比较了两种检索条件翻译方法的性能. 检索结果的评价显示: 用双语词典达到单语言检索性能的 50 - 60%, 使用翻译软件能达到单语言检索性能的 60 - 70%.

关键词: 跨语言信息检索, 双语词典, 机器翻译软件, 分词方法

Experiment on English-Chinese Cross-Language IR Using Bilingual Dictionary and MT Software

Abstract: In this paper, We made use of two techniques of query translation: MT software and bilingual dictionary. We use a simple and effective Chinese word segmentation method and compare CLIR performance. Evaluation shows that English-Chinese CLIR using bilingual dictionary can achieve about 50-60% of monolingual average precision and using MT software can bring this effectiveness to 60-70% of monolingual.

Keyword: CLIR, bilingual dictionary, MT software, word segmentation method

作者简介: 张永奎, 男, 教授, 博士生导师, 主要研究领域为中文信息处理和人工智能。王树锋, 男, 在职硕士研究生, 研究领域为中文信息处理。

通邮: 山西省孝义市孝义铝矿二期 王树锋 032303