

知识组织与知识管理

基于规则的中文时间词和数词的自动识别算法

高霄云<sup>1</sup>;杨建林<sup>1,2</sup>

南京大学信息管理系<sup>1</sup>

收稿日期 2007-1-8 修回日期 2007-2-16 网络版发布日期 2007-3-26 接受日期

**摘要** 对中文时间词和数词在文本中的常见形式进行归纳, 以此为基础构建用于识别时间词、数词的规则集, 提出一个基于规则的时间词和数词自动识别算法, 并对该算法在竞争情报分析领域和机器翻译领域中的应用价值进行论述。

**Abstract** This paper firstly generalizes the formats of Chinese time words and numerals appearing in the text. Based on them, this paper then sets up a rule sets for recognition, proposes a method about Chinese time words and numerals based on rules and discusses its application value in competitive intelligence analysis as well as machine translation field at last.

**关键词** [词语切分](#) [信息抽取](#) [规则](#)

**Key words** Word segmentation; Information extraction; Rule

**分类号** [TP391G252](#)

**DOI:**

通讯作者:

高霄云 [blue-6095@163.com](mailto:blue-6095@163.com)

作者个人主页: 高霄云 杨建林

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“词语切分”的 相关文章](#)
- ▶ 本文作者相关文章
  - [高霄云](#)
  - [杨建林](#)
  -