

论文

基于期待类型的Chart句法分析算法

王焱¹,李中志²

- 1. 成都信息工程学院
- 2. 电子科技大学电子工程学院

摘要:

Chart算法是目前句法分析中应用最广泛的算法之一, 但该算法的计算效率仍有待提高。通过对两种常用Chart算法的分析, 提出一种以由底向上的Chart算法为基础, 结合自顶向下Chart算法的预测能力的算法。算法按严格从左到右、由底向上的方向进行, 根据已有活动边的活动角色类型和句法规则, 产生当前词位置上的期待类型表, 并以此限制后续边的生成。对比实验的结果表明, 分析速度较普通Chart算法提高了约24%, 同时也减少了一半以上因边池溢出而导致的分析失败的语句。

关键词: 自然语言处理 句法分析 Chart natural language processing syntax analysis Chart

Grammatical parsing algorithm based on expected categories

Abstract:

Chart algorithm is one of the most popular algorithms in grammatical parsing. Its computational efficiency, however, still needs to be improved. By analyzing two common used Chart algorithms, a new parsing algorithm was proposed based on the bottom-to-top Chart algorithm, combining with the predicating ability of the top-to-bottom Chart algorithm. The proposed algorithm executed exactly from left to right and bottom to top. At each word position in parsing, an expecting table was formed according to the active roles of existing active arcs and the grammatical rules, and was used to restrict the creating of arcs in the rest process of parsing. Experimental results show that the parsing speed improves about 24% compared with the standard Chart algorithm, meanwhile the number of the failed sentences reduces more than a half.

Keywords:

收稿日期 2008-11-13 修回日期 2009-01-04 网络版发布日期 2009-06-09

DOI:

基金项目:

国家级基金;省部级基金

通讯作者: 王焱

作者简介:

参考文献:

本刊中的类似文章

1. 汤庸 林鹭贤 罗焯敏 潘炎.基于自动问答系统的信息检索技术研究进展[J]. 计算机应用, 2008,28(11): 2745-2748
2. 谈文蓉; 符红光; 刘莉; 杨宪泽.一种基于贝叶斯分类与机读词典的多义词排歧方法[J]. 计算机应用, 2006,26(6): 1389-1391
3. 张秋余;张博.自然语言语义理解在反垃圾邮件中的应用[J]. 计算机应用, 2006,26(6): 1315-1317
4. 张玉艳 杨潇 黄国栋 侯金奎.基于图算法的二元组合文法分析[J]. 计算机应用, 2008,28(7): 1668-1671

扩展功能

本文信息

- ▶ Supporting info
- ▶ PDF(552KB)
- ▶ [HTML全文]
- ▶ 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

本文关键词相关文章

- ▶ 自然语言处理
- ▶ 句法分析
- ▶ Chart
- ▶ natural language processing
- ▶ syntax analysis
- ▶ Chart

本文作者相关文章

- ▶ 王焱
- ▶ 李中志

PubMed

- ▶ Article by Yu,y
- ▶ Article by Li,Z.Z

反馈人	<input type="text"/>	邮箱地址	<input type="text"/>
反馈标题	<input type="text"/>	验证码	<input type="text" value="3241"/>