



## 论文摘要

中南大学学报(自然科学版)

ZHONGNAN DAXUE XUEBAO(ZIRAN KEXUE BAN)

Vol.40 No.6 Dec.2009

[PDF全文下载] [全文在线阅读]

文章编号: 1672-7207(2009)06-1630-06

## 基于依存关系的句法分析统计模型

袁里驰<sup>1, 2</sup>

- (1. 江西财经大学 信息学院数据与知识工程江西省重点实验室, 江西 南昌, 330013;
2. 中南大学 信息科学与工程学院, 湖南 长沙, 410083)

**摘要:** 利用语义、语法等语言知识, 建立一种基于依存关系的句法分析统计模型, 并利用改进的句法分析模型进行句法分析实验。研究表明: 利用依存关系、互信息对词聚类, 能解决模型数据稀疏问题; 模型可同时考虑几种语义依存关系; 该模型是一个词汇化的句法分析模型, 能结合分词、词性标注进行句法分析; 概率上下文无关语法中由概率的上下文无关性假设和祖先结点无关性假设引起的问题在该模型中得到有效解决; 精确率和召回率分别为86.96%和85.25%, 其综合指标 $F$ 与Collins的头驱动句法分析模型的 $F$ 相比提高4.75%。

**关键字:** 自然语言处理; 词聚类; 中心词驱动; 句法分析统计模型

## Statistical language parsing model based on dependency

YUAN Li-chi<sup>1, 2</sup>

- (1. School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330013, China;
2. School of Information Science and Engineering, Central South University, Changsha 410083, China)

**Abstract:** By incorporating linguistic features such as semantic dependency and syntactic relations, a novel statistical Parsing model was proposed. The experiments were conducted for the refined statistical parser. The results show that the model is constructed on word cluster, so the problem of data sparseness is not serious. The model can take advantage of a few semantic dependencies at the same time. The model is a parser based on lexicalized model, it is combined with segmentation and POS tagging model and thus a language parser is built. The questions caused by context-free hypothesis and ancestor-free hypothesis in probability context free grammar are solved well in this model. It achieves 86.96% precision and recall 85.25%,  $F$  value is improved by 4.75% compared with that of the head-driven parsing model introduced by Collins.

**Key words:** natural language processing; word clustering; head-driven parsing model; statistical parsing model

版权所有：《中南大学学报(自然科学版、英文版)》编辑部

地 址：湖南省长沙市中南大学 邮 编： 410083

电 话： 0731-88879765 传 真： 0731-88877727

电子邮箱： zngdxb@mail.csu.edu.cn 湘ICP备09001153号