

<<计算机自然语言处理>>

图书基本信息

书名：<<计算机自然语言处理>>

13位ISBN编号：9787302100898

10位ISBN编号：7302100896

出版时间：2005-4

出版时间：清华大学出版社

作者：王晓龙

页数：170

字数：260000

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

## <<计算机自然语言处理>>

### 内容概要

计算机自然语言处理技术在我国现代化及信息化建设中起着越来越重要的作用，我国政府已经将它列入“国家中长期科学技术发展纲领”。

近年来，语言处理技术，特别是基于国际互联网的中文语言处理技术正在引起我国广大科技工作者的高度重视。

本书既全面阐述了中文语言处理技术的特殊规律，又借鉴了国内外学者在计算语言学领域里的最新成就，还包括了作者的实践经验和体会。

本书可以作为计算机相关专业研究生的专业课教材，也可供相关专业高年级大学生和从事自然语言处理技术研究和应用的科技人员参考。

## &lt;&lt;计算机自然语言处理&gt;&gt;

## 书籍目录

第1章 引言第2章 数学基础 2.1 初等概率理论 2.1.1 基本概念 2.1.2 条件概率与独立  
 2.1.3 全概率公式与贝叶斯公式 2.1.4 随机变量 2.1.5 多维随机变量 2.1.6 数学期望  
 与方差 2.1.7 常用分布 2.2 信息论基础 2.2.1 信息熵 2.2.2 联合熵和条件熵 2.2.3  
 互信息 2.2.4 相关熵 2.2.5 语言与熵 2.2.6 噪声信道模型 2.3 粗糙集 2.3.1 信  
 息系统 2.3.2 不可分辨关系 2.3.3 集合近似 2.3.4 约简 2.3.5 属性依从 2.3.6 决  
 策规则合成 2.4 小结第3章 汉语自动分词技术 3.1 引言 3.2 分词规范 3.3 常用的分词方法  
 3.3.1 正向最大匹配分词 3.3.2 反向最大匹配分词 3.3.3 基于统计的词网格分词 3.4  
 歧义的分类和识别 3.4.1 歧义的分类 3.4.2 歧义的抽取和消歧 3.5 新词的识别 3.5.1  
 统计构词能力 3.5.2 汉字构词模式 3.5.3 未登录词识别算法 3.6 关于分词的若干统计结果  
 3.7 语言单位的统计分布规律(Zipf定律) 3.8 小结第4章 基于数学统计的语言模型 4.1 统计  
 语言模型概述 4.2 现有的主要统计语言模型 4.2.1 上下文无关模型 4.2.2 N元语法模型  
 4.2.3 N\|POS模型 4.2.4 基于决策树的语言模型 4.2.5 动态、自适应、基于缓存的语言模  
 型 4.3 数据平滑技术 4.3.1 数据平滑算法的评价标准 4.3.2 常见平滑方法 4.4 隐马尔科  
 夫模型 4.4.1 随机过程 4.4.2 马尔科夫链和马尔科夫性 4.4.3 马尔科夫模型 4.4.4  
 隐马尔科夫模型 4.5 最大熵模型 4.5.1 模型介绍 4.5.2 模型评价 4.5.3 最大熵语言建  
 模 4.6 小结第5章 基于语言理解的处理方法第6章 音字转换技术第7章 自动文摘技术第8章 信  
 息检索技术第9章 文字识别技术

<<计算机自然语言处理>>

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>