

<<中文信息处理教程>>

图书基本信息

书名：<<中文信息处理教程>>

13位ISBN编号：9787040318968

10位ISBN编号：7040318962

出版时间：2011-6

出版时间：高等教育出版社

作者：宋继华，杨尔弘，王强军 编

页数：283

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>

## <<中文信息处理教程>>

### 内容概要

《普通高等教育“十一五”国家级规划教材：中文信息处理教程》是编者依据自身的教学实践，在学习、吸收、借鉴前辈经验的基础上归纳、提炼而成的中文信息处理教材。书中比较系统地介绍了本学科领域的基本原理、方法和应用技术。

《普通高等教育“十一五”国家级规划教材：中文信息处理教程》共8章，按中文信息处理的语言单位层级——汉字、词语、句子、篇章依次展开。

内容包括：导论、汉字的信息处理、词的信息处理、句子的信息处理、句子语义表达与分析。中文信息处理的基础资源、文本分析与处理、中文信息处理评测。

《普通高等教育“十一五”国家级规划教材：中文信息处理教程》的特色是简明、实用，逻辑性强，可读性好，注重引导学生动手解决实际问题。

每章都附有习题，并给出了大部分习题的参考答案。

《普通高等教育“十一五”国家级规划教材：中文信息处理教程》可作为高等学校计算机、信息管理等专业本科生和研究生的教材，也可供从事中文信息处理研究和应用的科技工作者参考。

## 书籍目录

第1章 导论本章概览知识结构图1.1 基本概念1.1.1 学科由来1.1.2 学科定位1.2 研究内容1.3 难点分析1.3.1 歧义1.3.2 语法1.4 研究路线1.5 习题第2章 汉字的信息处理本章概览知识结构图2.1 基础知识2.1.1 汉字的形、音、义2.1.2 汉字的字频和字量2.1.3 汉字的编码2.2 交换码和内码2.2.1 ASC 2.2.2 编码框架: ISO/IEC20222.2.3 GB23122.2.4 BIG52.2.5 ISO/IEC10646和Unicode2.2.6 GBK2.2.7 GBL80302.2.8 编程务实2.3 汉字的输入2.3.1 键盘输入2.3.2 字形识别2.3.3 语音识别2.4 汉字的输出2.5 中文编码的前沿课题2.5.1 古籍数字化2.5.2 《通用规范汉字表》2.5.3 少数民族文字2.6 习题第3章 词的信息处理本章概览知识结构图3.1 基础知识3.1.1 概率论基础3.1.2 信息论基础3.1.3 n元语法模型3.1.4 语法模型的性能评价3.2 自动分词3.2.1 汉语词的界定3.2.2 自动分词方法3.2.3 未登录词的识别3.3 词性标注3.3.1 词性标注概述3.3.2 基于统计的词性标注方法3.3.3 基于规则的词性标注方法3.4 命名实体识别3.4.1 命名实体识别介绍3.4.2 中文命名实体的定义和标准3.4.3 中文命名实体识别方法3.5 习题第4章 句子的信息处理本章概览知识结构图4.1 形式语言基础4.1.1 形式语言理论4.1.2 自动机理论4.2 短语结构语法4.2.1 汉语短语结构语法4.2.2 分析算法4.2.3 实用策略4.3 依存语法4.3.1 依存句法理论4.3.2 分析算法4.4 句法理论探索4.4.1 Chomsky语法理论4.4.2 广义短语结构语法4.4.3 链语法4.4.4 范畴语法4.5 习题第5章 句子语义表达与分析本章概览知识结构图5.1 格语法5.1.1 基本思想5.1.2 格的分类5.1.3 格的判断5.1.4 用格语法生成句子5.1.5 汉语格语法5.2 概念依存理论5.2.1 基本模型5.2.2 语义推理5.2.3 脚本法5.3 概念层次网络(HNC) 5.4 语义本体5.5 习题第6章 中文信息处理的基础资源本章概览知识结构图6.1 语料库概述6.1.1 语料库发展历史6.1.2 语料库属性6.2 典型中文语料库介绍6.2.1 北京大学计算语言学研究所《人民日报》标注语料库6.2.2 清华大学汉语树库6.2.3 哈尔滨工业大学汉语依存树库6.2.4 中国台湾"中央研究院"语料库6.2.5 国家语言文字工作委员会语料库6.2.6 山东大学语料库6.3 典型中文语言知识库资源介绍6.3.1 北京大学计算语言学研究所综合型语言知识库6.3.2 知网6.3.3 中国台湾"中央研究院"中文词汇网络6.3.4 概念层次网络6.3.5 汉语框架网知识库6.4 中文语言资源联盟6.5 习题第7章 文本分析与处理本章概览知识结构图7.1 文本分类7.1.1 文本表示7.1.2 文本分类算法7.2 信息检索7.2.1 常用的检索模型7.2.2 搜索引擎7.3 问答系统7.3.1 问答系统概述7.3.2 系统构成7.3.3 典型问答系统介绍7.4 信息抽取7.4.1 信息抽取概述7.4.2 系统的基本构成及关键技术7.5 自动文摘7.5.1 自动文摘概述7.5.2 基于统计的机械文摘7.5.3 基于意义的理解文摘7.6 习题第8章 中文信息处理评测本章概览知识结构图8.1 评测概述8.1.1 评测的意义8.1.2 评测的过程8.2 系列评测介绍8.2.1 NIST系列评测8.2.2 ACL-SIGHAN系列评测8.2.3 863技术测评8.2.4 中文信息学会评测8.3 各领域技术评测介绍8.3.1 中文分词和词性标注评测8.3.2 词义消歧评测8.3.3 句法分析评测8.3.4 文本分类器性能评估8.3.5 信息检索系统的评测8.3.6 问答系统评测8.3.7 信息抽取评测8.3.8 自动文摘评测8.4 习题部分习题参考答案参考文献

## <<中文信息处理教程>>

### 编辑推荐

《普通高等教育“十一五”国家级规划教材：中文信息处理教程》共8章，按中文信息处理的语言单位层级——汉字、词语、句子、篇章依次展开。

内容包括：导论、汉字的信息处理、词的信息处理、句子的信息处理、句子语义表达与分析。

中文信息处理的基础资源、文本分析与处理、中文信息处理评测。

本书的特色是简明、实用，逻辑性强，可读性好，注重引导学生动手解决实际问题。

每章都附有习题，并给出了大部分习题的参考答案。

<<中文信息处理教程>>

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>