第一图书网, tushu007.com

<<自然语言处理综论>>

图书基本信息

书名：<<自然语言处理综论>>

13位ISBN编号：9787121007767

10位ISBN编号：7121007762

出版时间：2005

出版时间：电子工业出版社

作者：Daniel Jurafsky,James H. Martin

页数：588

译者：冯志伟,孙乐

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

内容概要

本书是一本全面系统地讲述计算机自然语言处理的优秀教材。
本书英文版出版之后好评如潮，国外许多著名大学纷纷把本书选为自然语言处理和计算语言学课程的主要教材，该书被誉为该领域教材的“黄金标准”。
本书包含的内容十分丰富，分为四个部分，共21章，深入细致地探讨了计算机处理自然语言的词汇、句法、语义、语用等各个方面的问题，介绍了自然语言处理的各种现代技术。
从层次的角度看，本书的论述是按照自然语言的不同层面逐步展开的，首先论述单词的自动形态分析，接着论述自动句法分析，然后论述各种语言单位的自动语义分析，最后论述连贯文本的自动分析、对话与会话的智能代理以及自然语言生成。
从技术的角度看，本书介绍了正则表达式、有限状态自动机、文本-语音转换、发音与拼写的概率模型、词类自动标注、N元语法、隐马尔可夫模型、上下文无关语法、特征与合一、词汇化剖析与概率剖析、一阶谓词演算、词义排歧、修辞结构理论、机器翻译等非常广泛的内容。
本书具有“覆盖全面、注重实用、强调评测、语料为本”四大特色。
在本书的配套网站上，还提供了相关的资源和工具，便于读者在实践中进一步提高。

书籍目录

第1章 导论1.1 语音与语言处理中的知识1.2 歧义1.3 模型和算法1.4 语言、思维和理解1.5 学科现状与近期发展1.6 语音和语言处理简史1.6.1 基础研究：20世纪40年代和20世纪50年代1.6.2 两个阵营：1957年至1970年1.6.3 四个范型：1970年至1983年1.6.4 经验主义和有限状态模型的复苏：1983年至1993年1.6.5 不同领域的合流：1994年至1999年1.6.6 多重发现1.6.7 心理学的简要注记1.7 小结1.8 文献和历史说明第一部分 词汇的计算机处理第2章 正则表达式与自动机2.1 正则表达式2.1.1 基本正则表达式模式2.1.2 析取、组合与优先关系2.1.3 一个简单的例子2.1.4 一个比较复杂的例子2.1.5 高级算符2.1.6 正则表达式中的替换、存储器与eliza2.2 有限状态自动机2.2.1 用fsa来识别羊的语言2.2.2 形式语言2.2.3 另外的例子2.2.4 非确定fsa2.2.5 使用nfsa接收符号串2.2.6 识别就是搜索2.2.7 确定自动机与非确定自动机的关系2.3 正则语言与fsa2.4 小结2.5 文献和历史说明第3章 形态学与有限状态转录机3.1 英语形态学概观3.1.1 屈折形态学3.1.2 派生形态学3.2 有限状态形态剖析3.2.1 词表和形态顺序规则3.2.2 用有限状态转录机进行形态剖析3.2.3 正词法规则和有限状态转录机3.3 把fst词表与规则相结合3.4 与词表无关的fst：porter词干处理器3.5 人是怎样进行形态处理的3.6 小结3.7 文献和历史说明第4章 计算音系学与文本-语音转换4.1 言语语音与语音标音法4.1.1 发音器官4.1.2 辅音：发音部位4.1.3 辅音：发音方法4.1.4 元音4.1.5 音节4.2 音位和音位规则4.3 音位规则和转录机4.4 计算音系学中的一些高级问题4.4.1 元音和谐4.4.2 模板式形态学4.4.3 优选理论4.5 音位规则的机器学习4.6 tts中从文本映射到语音4.6.1 发音词典4.6.2 词典之外的查找：文本分析4.6.3 基于有限状态转录机（fst）的发音词典4.7 文本-语音转换中的韵律4.7.1 韵律的音系学性质4.7.2 韵律的语音和声学性质4.7.3 语音合成中的韵律4.8 人处理音位和形态的过程4.9 小结4.10 文献和历史说明第5章 发音与拼写的概率模型5.1 关于拼写错误5.2 拼写错误模式5.3 非词错误的检查5.4 概率模型5.5 把贝叶斯方法应用于拼写5.6 最小编辑距离5.7 英语的发音变异5.8 发音问题研究中的贝叶斯方法5.8.1 发音变异的决策树模型5.9 加权自动机5.9.1 从加权自动机计算似然度：向前算法5.9.2 解码：viterbi算法5.9.3 加权自动机和切分5.9.4 用切分来进行词表的自动归纳5.10 人类发音研究5.11 小结5.12 文献和历史说明第6章 n元语法6.1 语料库中单词数目的计算6.2 简单的（非平滑的）n元语法6.2.1 n元语法及其对训练语料库的敏感性6.3 平滑6.3.1 加1平滑6.3.2 witten-bell打折法6.3.3 good-turing打折法6.4 回退6.4.1 回退与打折相结合6.5 删除插值法6.6 拼写和发音的n元语法6.6.1 上下文有关的错拼更正6.6.2 发音模型的n元语法6.7 熵6.7.1 用于比较模型的交叉熵6.7.2 英语的熵6.8 小结6.9 文献和历史说明第7章 hmm与语音识别7.1 语音识别的总体结构7.2 隐马尔可夫模型概述7.3 再谈viterbi算法7.4 先进的解码方法7.4.1 a*解码算法7.5 语音的声学处理7.5.1 声波7.5.2 怎样解释波形7.5.3 声谱7.5.4 特征抽取7.6 声学概率的计算7.7 语音识别系统的训练7.8 用于语音合成的波形生成7.8.1 音高和音延的修正7.8.2 单元选择7.9 人的语音识别7.10 小结7.11 文献和历史说明第二部分 句法的计算机处理第8章 词的分类与词类标注8.1 大多数英语词的分类8.2 英语的标记集8.3 词类标注8.4 基于规则的词类标注8.5 随机词类标注8.5.1 说明问题的一个例子8.5.2 实际的hmm标注算法8.6 基于转换的标注8.6.1 怎样应用tbl规则8.6.2 怎样学习tbl规则8.7 其他问题8.7.1 多重标记和多项词8.7.2 未知词8.7.3 基于类的n元语法8.8 小结8.9 文献和历史说明第9章 英语的上下文无关语法9.1 组成性9.2 上下文无关规则和树9.3 句子级的结构9.4 名词短语9.4.1 在中心名词前的成分9.4.2 名词后的成分9.5 并列关系9.6 一致关系9.7 动词短语和次范畴化9.8 助动词9.9 口语的句法9.9.1 不流畅现象9.10 语法等价与范式9.11 有限状态语法和上下文无关语法9.12 语法和人的语言处理9.13 小结9.14 文献和历史说明第10章 基于上下文无关语法的剖析10.1 剖析就是搜索10.1.1 自顶向下剖析10.1.2 自底向上剖析10.1.3 自顶向下剖析与自底向上剖析的对比10.2 基本的自顶向下剖析10.2.1 增加自底向上过滤10.3 基本的自顶向下剖析的问题10.3.1 左递归10.3.2 歧义10.3.3 子树的重复剖析10.4 earley算法10.4.1 预测10.4.2 扫描10.4.3 完成10.4.4 示例10.4.5 从线图中检索剖析树10.5 有限状态剖析方法10.6 小结10.7 文献和历史说明第11章 特征与合一11.1 特征结构11.2 特征结构的合一11.3 语法中的特征结构11.3.1 一致关系11.3.2 中心语特征11.3.3 次范畴化11.3.4 其他词类的次范畴化11.3.5 长距离依存关系11.4 合一的实现11.4.1 合一的数据结构11.4.2 合一算法11.5 带有合一约束的剖析11.5.1 把合一结合到earley剖析器中11.5.2 复制的必要性11.5.3 合一剖析11.6 类型与继承11.6.1 类型的扩充11.6.2 合一的其他扩充11.7 小结11.8 文献和历史说明第12章 词汇化剖析与概率剖析12.1 概率上下文无关语法12.1.1 pcfg的概率cyk剖析12.1.2 pcfg概率的学习12.2 pcfg的问题12.3 概率词汇化的cfg12.4 依存语法12.4.1 范畴语法12.5 人的剖

析12.6 小结12.7 文献和历史说明第13章 语言的复杂性13.1 chomsky层级13.2 怎么判断一种语言不是正则的13.2.1 抽吸引理13.2.2 英语和其他自然语言是正则语言吗13.3 自然语言是上下文无关的吗13.4 计算复杂性和人的语言处理13.5 小结13.6 文献和历史说明第三部分 语义的计算机处理第14章 意义的表示法14.1 意义表示的计算要求14.1.1 可能性验证14.1.2 无歧义表示14.1.3 规范形式14.1.4 推论与变元14.1.5 表达能力14.2 语言的意义结构14.2.1 谓词论元结构14.3 一阶谓词演算14.3.1 fopc基础14.3.2 fopc的语义14.3.3 变量和逻辑量词14.3.4 推论14.4 某些与语言学相关的概念14.4.1 范畴14.4.2 事件14.4.3 时间表示14.4.4 体14.4.5 信念表示14.4.6 缺陷14.5 有关的表示方法14.6 意义的其他表示方法14.6.1 作为行动的意义14.6.2 作为真值的意义14.7 小结14.8 文献和历史说明第15章 语义分析15.1 句法驱动的语义分析15.1.1 给上下文无关语法规则扩充语义15.1.2 量词辖域和复杂项的转译15.2 给英语片断附加语义分析15.2.1 句子15.2.2 名词短语15.2.3 动词短语15.2.4 介词短语15.3 把语义分析结合到earley剖析中15.4 惯用语和组成性15.5 鲁棒的语义分析15.5.1 语义语法15.5.2 信息抽取15.6 小结15.7 文献和历史说明第16章 词汇语义学16.1 词位及其涵义之间的关系16.1.1 同形关系16.1.2 多义关系16.1.3 同义关系16.1.4 上下位关系16.2 wordnet：词汇关系信息库16.3 词的内在结构16.3.1 题元角色16.3.2 选择限制16.3.3 基元分解16.3.4 语义场16.4 语言的创造性与词典16.4.1 隐喻16.4.2 换喻16.4.3 隐喻和换喻的计算方法16.5 小结16.6 文献和历史说明第17章 词义排歧与信息检索17.1 基于选择限制的排歧17.1.1 选择限制的局限性17.2 鲁棒的词义排歧17.2.1 机器学习方法17.2.2 基于词典的方法17.3 信息检索17.3.1 向量空间模型17.3.2 检索词加权17.3.3 检索词的选择和创造17.3.4 同形关系、多义关系和同义关系17.3.5 改进用户的查询条件17.4 信息检索的其他任务17.5 小结17.6 文献和历史说明第四部分 语用的计算机处理第18章 话语18.1 所指判定18.1.1 所指现象18.1.2 同指的句法和语义约束18.1.3 代词解释中的优先关系18.1.4 代词判定算法18.2 文本的连贯18.2.1 现象18.2.2 基于推理的判定算法18.3 话语结构18.4 所指和连贯的心理语言学研究18.5 小结18.6 文献和历史说明第19章 对话与会话智能代理19.1 什么使对话出现差别19.1.1 话轮和话段19.1.2 对话的共同基础19.1.3 会话隐涵19.2 对话行为19.3 对话行为的自动解释19.3.1 对话行为的计划推理解释19.3.2 对话行为的基于提示的解释19.3.3 要点19.4 对话结构与连贯性19.5 会话智能代理中的对话管理19.6 小结19.7 文献和历史说明第20章 自然语言生成20.1 语言生成导引20.2 生成的体系结构20.3 表层实现20.3.1 系统语法20.3.2 功能合一语法20.3.3 要点20.4 话语规划20.4.1 文本说明图20.4.2 修辞关系20.4.3 小结20.5 其他问题20.5.1 微规划20.5.2 词汇选择20.5.3 生成系统评价20.5.4 语音生成20.6 小结20.7 文献和历史说明第21章 机器翻译21.1 语言的相似性和差异性21.2 转换模型21.2.1 句法转换21.2.2 词汇转换21.3 中间语的思想：使用意义21.4 直接转换21.5 使用统计技术21.5.1 流畅性的量化21.5.2 忠实性的量化21.5.3 输出的搜索21.6 可用性与系统开发21.7 小结21.8 文献和历史说明附录a 正则表达式的算符附录b porter stemming算法附录c 标记集c5和c7附录d hmm模型的训练：向前-向后算法参考文献术语表

第一图书网, tushu007.com

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