

目录

总序	王文斌	xii
前言	王璐璐 袁毓林	xvii
第一章 人类的进化和交际方式的演变		1
<hr/>		
1.1 人类的进化		1
1.2 交际方式的演变		7
1.3 语言在人类发展中的作用		11
第二章 语言通信设备的更替和网络的诞生		15
<hr/>		
2.1 语言通信设备的更替		15
2.2 香农的通信模型		21
2.3 网络的诞生		23
2.4 社交网络		26
第三章 网络文本的信息爆炸和搜索引擎的发明		32
<hr/>		
3.1 网络文本的信息爆炸		32
3.2 信息检索与搜索引擎		35
3.2.1 信息检索的概念与类型		35
3.2.2 网络搜索引擎的工作原理		36
3.3 搜索引擎技术		41

3.3.1	布尔检索和倒排索引	41
3.3.2	超链接分析与网页排序	44
3.4	基于关键词匹配方法的局限性	47
3.5	基于语义和概念的搜索引擎技术	50
3.5.1	基于语义的扩充式关键词搜索	51
3.5.2	基于本体知识推理的语义检索	53

第四章 自然语言处理的规则模型和统计模型 57

4.1	自然语言处理技术概说	57
4.2	基于规则的自然语言处理	63
4.2.1	有限状态自动机	64
4.2.2	上下文无关文法	66
4.2.3	基于上下文无关文法的剖析	69
4.3	基于统计的自然语言处理	73
4.3.1	基于概率的语言识别	75
4.3.2	基于统计的机器翻译	78

第五章 语言信息处理和语言知识数据库的建设 82

5.1	语料库和语言知识库概说	82
5.2	语言知识库面面观	86
5.2.1	词汇网络 (WordNet)	86
5.2.2	句法树库 (Treebank)	90
5.2.3	动词网络 (VerbNet)	98
5.2.4	命题库 (PropBank)	101
5.2.5	情境框架网络 (FrameNet)	103

5.3	相关资源的统一和整合	106
5.3.1	词义消歧和义项归组	107
5.3.2	义项与框架对接、框架与实例对勘	109
5.3.3	相关资源的义项映射和框架映射	111
5.3.4	实现相关资源的统一和整合	114
第六章 汉语句法语义知识库的研究与建设		118
<hr/>		
6.1	汉语词类的模糊划分与测试平台	118
6.1.1	词类的范畴性质	119
6.1.2	词类的模糊划分和隶属度分析	122
6.1.3	网络版词类测试平台的设计及实现	125
6.2	北大实词句法语义信息词典与检索系统	131
6.2.1	《形容词信息词典》的知识内容	133
6.2.2	《动词信息词典》的知识内容	137
6.2.3	《名词信息词典》的知识内容	144
6.3	汉语动词蕴涵关系和蕴涵型式库建设	152
6.3.1	语言表达的多样性与文本蕴涵	152
6.3.2	动词蕴涵关系的理论背景与蕴涵型式库 建设的目标	153
6.3.3	蕴涵式的类聚规律与分类体系	159
6.3.4	汉语动词蕴涵型式库的体系结构与功能模块	170
第七章 语言信息处理和语义计算的多层次建模		172
<hr/>		
7.1	理论背景：认知语言学研究的三种范式	173
7.2	基于认知的语义知识的描述和计算	174

7.2.1	领属转移的情境网络	176
7.2.2	领属转移情境的词汇知识库	180
7.2.3	领属转移的情境网络和词汇知识库的连接	182
7.3	亲属关系自动推理系统的研发过程	183
7.3.1	亲属关系的认知建模和情境网络	184
7.3.2	亲属关系的逻辑表示和语义公理	185
7.3.3	亲属关系的词汇—句法知识库	190
7.3.4	亲属关系自动推理的算法设计和程序实现	192
7.4	基于血缘关系结构图的亲属关系推理系统研究与实现	198
7.4.1	亲属关系推理与人工智能机器人	199
7.4.2	亲属关系推理系统研究现状	200
7.4.3	亲属关系推理的复杂性	203
7.4.4	亲属关系推理的关键问题分析	207
7.4.5	亲属关系推理系统	219
第八章 语音的识别与合成和言语信息处理技术		223
8.1	语音识别	223
8.1.1	特征抽取	226
8.1.2	声学解码	228
8.1.3	语言解码	231
8.2	语音合成	235
8.2.1	语言处理模块	236
8.2.2	信号处理模块	240
8.3	人机交互应用	241
8.4	人机对话交流和聊天机器人的研究与开发	244
8.4.1	语言运用与人工智能：从“图灵测试”	

到“中文屋”.....	244
8.4.2 聊天机器人的前辈：从医生 ELIZA 到病人 PARRY.....	247
8.4.3 智能对话问答系统的技术与相关的修辞学 研究课题.....	254
8.4.4 智能对话问答系统研发给语言学的启示.....	261
第九章 语义网替代万维网和本体知识与知识图谱的建构	264
<hr/>	
9.1 万维网的体系结构及其功能缺陷.....	264
9.2 语义网的体系结构.....	267
9.3 基于本体知识的语义网建设.....	270
9.3.1 本体知识的构建.....	270
9.3.2 建议上层共用知识本体.....	272
9.3.3 网络代理.....	274
9.4 本体知识的类型及其建构的困难.....	276
9.5 语义网的杀手级应用：维基百科.....	277
9.6 从语义网到知识图谱.....	280
第十章 深度问答系统和文本分析技术	284
<hr/>	
10.1 深度问答系统.....	284
10.1.1 Watson 系统.....	284
10.1.2 深度问答方法.....	286
10.2 文本分析技术.....	291
10.2.1 文本分类.....	291
10.2.2 文本聚类.....	292

10.2.3	主题模型	292
10.2.4	向量空间模型	293
第十一章 走向文本内容的深度计算和自动理解		295
<hr/>		
11.1	汉语文本内容的自动分析	296
11.1.1	自动分词和词性标注	296
11.1.2	句法分析和语义分析	298
11.2	汉语文本内容的自动理解	303
11.2.1	汉语名名组合的自动释义研究与实现	303
11.2.2	汉语“把”字句的自动释义研究与实现	319
11.2.3	汉语“比”字句的分析模型及计算应用	336
第十二章 语言处理技术的发展趋势与未来议题		356
<hr/>		
12.1	深度学习与自然语言处理	356
12.2	词嵌入与语义的向量表示	359
12.3	基于神经网络的语言模型及其局限性	363
12.3.1	神经网络模型 Transformer	363
12.3.2	双向编码器表示模型 BERT	364
12.3.3	生成式预训练模型 GPT	365
12.3.4	深度学习模型的“不可解释性”及其原因	368
12.4	基于语言知识资源的语义矛盾核查	370
12.5	人工智能呼唤语义理解和常识推理	377
12.6	利用语义资源帮助机器人回答常识性问题	380
12.7	走向未来的自然语言的认知计算	386

后记	390
参考文献	398
推荐文献	416
索引	418