

中文指代消解模型的对比研究

周炫余 刘娟 罗飞 刘洋 颜晗

(武汉大学计算机学院 武汉 430072)

摘要 指代消解的基本任务是把指向现实世界中相同实体的所有实体表达关联起来。与英文指代消解的研究相比,中文指代消解的研究相对较少,至今没有对现存的中文指代消解模型进行公平的测评和比较,主要原因是现存的中文指代消解模型在训练和测评时采用了不同的语料,且所选用的特征属性也不相同。基于上述原因,实现了 5 类基本的中文指代消解模型,并在 ACE2005 中文语料上进行同平台、同语料、同特征的测评。通过测评比较了各类模型的性能,同时分析和探讨了影响中文指代消解模型精度的各种因素。

关键词 人工智能,指代消解模型,指代消解,自然语言处理

中图分类号 TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.2.006

Comparison of Chinese Anaphora Resolution Models

ZHOU Xuan-yu LIU Juan LUO Fei LIU Yang YAN Han

(Computer School, Wuhan University, Wuhan 430072, China)

Abstract The task of anaphora resolution is determining linguistic expressions that refer to the same real-world entity in a text. Compared to the amount of research on English anaphora resolution, relatively little work has been done on Chinese anaphora resolution. To our best knowledge, there is no fair comparison on the Chinese anaphora resolution models, mainly because of the useage of different data and features. We aimed to gain insights into the factors that affect the performance of the models and the advantages of each model via experiment with the same data and features in the ACE2005 Chinese corpus.

Keywords Artificial intelligence, Anaphora resolution models, Anaphora resolution, Natural language processing

1 引言

指代是一种常见的语言学现象,它使得句子变得通顺和简洁,然而篇章中大量的指代现象增大了计算机理解的难度。指代关系中抽象的语言学单元称为照应语,具体的语言学单元称为先行语。指代消解的基本任务就是把指向现实世界中相同实体(Entity)的所有实体表达(Mention)关联起来。指代消解不仅能增进计算机理解文本的程度,也是其余各类自然语言处理技术的关键子技术之一,例如信息抽取^[1]、机器学习^[2]、自动问答^[3]等。

指代消解经过 50 年的发展,现已有大量的模型和方法。早期的中文指代消解方法由于缺少标注语料,主要采用的是基于规则的方法^[4,5]和基于无监督的方法^[6,7]。随着 ACE、OntoNotes 语料的出现,中文指代消解方法逐渐转向基于有监督的方法^[8-15]和基于规则的层次过滤的方法^[16-19]。由于各类模型所采用的语料、特征属性和待消解项不同,因此至今没有对现存的中文指代消解模型进行公平的测评和比较。基于上述原因,对 5 类基本的中文指代消解模型在 ACE2005 中文

语料上进行同平台、同语料、同特征的测评,通过广泛的实验来着重讨论以下几点。

(1)待消解项识别。待消解项识别的主要任务是确定篇章中哪些词语需要指代消解^[20]。本文将从以下 3 个方面来探索待消解项识别:制约中文待消解项识别的因素是什么?待消解项识别对指代消解模型会有哪些影响?如何才能抽取嵌套名词短语?

(2)特征抽取。特征抽取是将文字转换成向量的关键步骤。不同的模型有不同的特征抽取方法,究竟哪种抽取方法更加准确?特征属性抽取不准确的主要原因是什么?

(3)5 类模型的比较。模型划分是中文指代消解的关键步骤。由于各个模型采用的特征属性、待消解项集合都不相同,因此很难通过各自的测评结果衡量模型的消解效果。如何在公平的条件下评判各类模型的性能?究竟哪类模型更适合中文指代消解?

2 中文指代消解模型

根据基于无监督的指代消解模型^[6]、实体表达对模型^[8]、

到稿日期:2015-03-15 返修日期:2015-05-17 本文受国家自然科学基金(61272274),国家自然科学基金青年项目(61402340),湖北省自然科学基金青年项目(2014CFB194)资助。

周炫余(1987—),男,博士生,主要研究方向为自然语言处理,E-mail:zhouxuan-yu@whu.edu.cn;刘娟(1970—),女,教授,博士生导师,主要研究方向为自然语言处理、数据挖掘;罗飞(1980—),男,讲师,主要研究方向为数据挖掘、机器学习;刘洋(1991—),男,硕士,主要研究方向为自然语言处理;颜晗(1990—),男,硕士,主要研究方向为自然语言处理。

实体-实体表达模型^[9]、实体排序模型^[10]、基于规则的层次过滤模型^[16]的思想实现了5类基本的中文指代消解模型。上述5类模型基本涵盖了现有中文指代消解模型。

1) 基于无监督方法的中文指代消解模型

根据 Cardie 等^[6]提出的思想实现无监督的指代消解模型。该思想利用距离函数衡量照应语和先行语之间的指代相似性,当距离函数的值小于人工设定的阈值 r 时,二者存在指代关系,距离函数如式(1)所示。其中 NP_i 表示先行语, NP_j 表示照应语。 $Similar_f(NP_i, NP_j)$ 表示先行语和照应语关于特征 f 的相容函数,返回值为 0 或者 1。 W_f 表示特征 f 的权重,该权重由人工设定。

$$Core(NP_i, NP_j) = \sum_{f \in F} w_f Similar_f(NP_i, NP_j) \quad (1)$$

2) 实体表达对模型(Mention-pair model)

实体表达对模型^[8]是一种使用最为广泛的中文指代消解模型。通过训练语料训练得出一个二元分类模型,并用该模型判断实体表达对之间 $\langle \pi, \delta_i \rangle$ 指向同一实体的概率 P_c , 如式(2)所示。最后通过最近原则或者最优原则的聚类算法把存在指代关系的实体表达对聚成一条指代链。

$$P_c(COREF | \langle \pi, \delta_i \rangle) = \frac{\exp(\sum_{i=1}^n \lambda_i f_i(\langle \pi, \delta_i \rangle, COREF))}{\sum_c \exp(\sum_{i=1}^n \lambda_i f_i(\langle \pi, \delta_i \rangle, c))} \quad (2)$$

实体表达对模型的训练正例:每个照应语 m_i 和与之具有指代关系且距离最近的先行语 m_j 构成正例。训练负例:照应语 m_i 和中间词 $m_{j+1}, m_{j+2}, \dots, m_{i-1}$ 构成负例。通过上述正负例,基于 SVM 训练得出实体表达对模型。测试实例:用训练好的模型判断实体表达对之间是否存在指代关系,选择每一个与 m_i 最近的且存在指代关系的实体表达作为先行语。如果没有先行语,则自己成为一条单独的指代链。

3) 实体-实体表达模型(Entity-Mention model)

实体-实体表达模型^[9]在本质上也是经典的二元分类问题。与实体表达对模型不同的是,实体-实体表达模型估计的是照应语 π 与之前形成的部分聚类体 c_j (部分指向同一实体的不同实体表达集合称为部分聚类体)之间是否存在指代关系,如式(3)所示。如果二者指向同一实体,则把该名词加入部分聚类体。

$$P_c(COREF | \langle c_j, \pi \rangle) = \frac{\exp(\sum_{i=1}^n \lambda_i f_i(\langle c_j, \pi \rangle, COREF))}{\sum_c \exp(\sum_{i=1}^n \lambda_i f_i(\langle c_j, \pi \rangle, c))} \quad (3)$$

实体-实体表达模型的训练正例:每个照应语 m_i 和与之具有指代关系且距离最近的部分聚类体 c_j 构成正例。训练负例:照应语 m_i 和部分聚类体 c_k (c_k 中的实体表达出现在照应语 m_i 和部分聚类体 c_j 的中间)构成负例。通过上述正负例,基于 SVM 训练得出实体-实体表达模型。测试实例:用训练好的模型判断实体与部分聚类体之间是否存在指代关系,选择与 m_i 最近的且存在指代关系的部分聚类体 c_j 作为先行语。

4) 实体排序模型(Mention-ranking model)

实体排序模型^[10]首先针对每个照应语,创建候选先行语集合 A , 然后计算集合 A 中候选先行词 δ_i 是照应语 π 的指代概率,如式(4)所示。根据指代概率进行排序,选择概率值最

高的作为最终的先行语。

$$P_r(\delta_i | \pi) = \frac{\exp(\sum_{i=1}^n \lambda_i f_i(\pi, \delta_i))}{\sum_k \exp(\sum_{i=1}^n \lambda_i f_i(\pi, \delta_k))} \quad (4)$$

实体排序模型的训练正例:每个照应语 m_i 和与之具有指代关系且指代概率最高的实体表达 m_j 构成正例。训练负例:照应语 m_i 和中间词 $m_{j+1}, m_{j+2}, \dots, m_{i-1}$ 构成负例。通过上述正负例,基于 SVM 训练得出实体排序模型。测试实例:照应语 m_i 选取指代概率值最高的候选先行语 m_j 作为最终的先行语。

5) 基于规则的中文指代消解模型

Raghuathan 等^[16]首次提出一种基于规则的层次过滤模型。Lee 等^[17]扩展上述模型,主要增加了语义层、待消解项识别层、后处理层。Lee 的语义层主要是通过 WordNet、Wikipedia infoboxes、Freebase records 获取语义知识。然而中文语义知识库较小,很难包含所有的待消解项。Zhang 等^[19]在 Lee 的基础上删除了语义层,并针对中文的特点进行了一定的修改使之更加适合中文指代消解。由于同一条指代链中的实体表达在语义上存在着很大的语义相关性,本文在 Zhang 等提出模型的基础上增加了语义匹配层,该层通过引入 Web 语义知识很好地克服了中文语义知识库较小的缺陷。同时,针对中文中存在大量的嵌套名词短语,对待消解项识别层进行了一系列的修改,使之更加适合中文指代消解。基于规则的层次过滤模型的主要思想是人工设定一系列的指代消解过滤规则,模型中各个层次的排列顺序按照每层的消解精度从高到低进行排序,每一层的输入以上一层输出的实体聚类体为基础,结合共享属性(性别、单复数等)得到最终的指代链。该模型中主要包含字符串匹配层、宽松字符串匹配、中心词匹配、准确的结构匹配层、语义匹配层、代词层,具体层次顺序如表 1 所列。

表 1 层次过滤模型中的顺序

层次	处理数据类型	特征
1	名词短语	字符串匹配
2	名词短语	实体表达宽松字符串匹配 实体宽松字符串匹配
3	名词短语	严格中心词匹配 宽松中心词匹配
4	名词短语	同位语 缩略语
5	名词短语	HowNet 语义相似度计算 Web 语义相似度计算
6	代词	单复数一致 & 性别一致 & 生命度 & NER 标签

3 实验结果与分析

本节首先介绍实验所采用的数据以及5类基本模型所采用的特征属性,在此基础上对5类基本模型进行测评和比较,并通过测评结果来说明引言中的问题。表4和表5中的P、R、F分别表示系统的正确率、召回率和综合评估值。图1和表5中的MP、EM、RM、Unsupervised、Rules分别表示实体表达对模型、实体-实体表达模型、实体排序模型、基于无监督的模型、基于规则的层次过滤模型;MUC^[21]、BCUBED^[22]、CEAFE^[23]、MentionWise^[24]、BLANG^[25]分别表示5种不同角度的测评方法,Avg_F为上述5种测评方法F值的平均值,该值反映了模型的总体性能。

3.1 实验数据

Automatic Content Extraction(ACE)^[26]是2000年12月由美国国家标准和技术学会(NIST)资助和主管的一个国际公认的信息抽取平台。ACE语料从ACE2003开始就加入中文语料,ACE2005已经达拥有30万字的训练语料和5万字的测试语料,因此采用ACE2005中文语料作为实验数据具有较强的说服力。ACE2005中文语料存在3种不同类型的子语料,分别是BNEWS(网络对话)、NWIRE(新闻广播)、Weblogs(博客)。从上述子语料中分别随机抽取100篇、80篇、30篇作为测评语料。表2为测评语料中实体以及实体表达的统计。

表2 测试语料统计

数据集	语料数	实体	实体表达	实体表达/实体
BNEWS	100	2283	4905	2.15
NWIRE	80	2317	5026	2.17
Weblogs	30	764	1907	2.50
SUM	210	5364	11838	2.21

3.2 特征属性

特征属性抽取是将文本变成特征向量的关键步骤,该步骤很大程度上决定着模型最终的消解精度。从词法、语法、语义、位置来描述5类模型所采用的特征属性,具体抽取方法如表3所列。

表3 特征属性集合

类型	特征	照应语 m_i 与先行语 m_j 之间的特征描述
词汇特征	字符串匹配	如果照应语 m_i 和先行语 m_j 之间存在字符串匹配,则聚成一条指代链(为1),否则不能聚为一条链(为0)。
	宽松字符串匹配	如果照应语 m_i 和先行语 m_j 之间的字符串存在词包含关系,则二者可以聚成一条指代链(为1),否则不能聚为一条链(为0)。
	中心词匹配	如果照应语 m_i 和先行语 m_j 之间存在中心词匹配,则二者可以聚成一条指代链(为1),否则不能聚为一条链(为0)。
语法特征	同位语属性	如果照应语 m_i 和先行语 m_j 之间的是同位语,则为1,否则为0。
	生物属性	如果照应语 m_i 和先行语 m_j 之间的生物属性一致,则为1,否则为0。
	性别属性	如果照应语 m_i 和先行语 m_j 之间的性别一致,则为1,否则为0。
	单复数属性	如果照应语 m_i 和先行语 m_j 之间的单复数一致,则为1,否则为0。
语义特征	Web语义属性	如果照应语 m_i 和先行语 m_j 之间的Web语义相似度小于0.5,则二者可以聚成一条指代链(为1),否则不能聚为一条链(为0)。
	HowNet属性	如果照应语 m_i 和先行语 m_j 之间的HowNet语义相似度大于0.5,则二者可以聚成一条指代链(为1),否则不能聚为一条链(为0)。
位置特征	句子距离属性	如果照应语 m_i 和先行语 m_j 在一个句子中,则为1,在第二个句子为0.8,以此类推。

3.3 待消解项识别

本节将着重讨论引言中待消解项识别的3个问题。实验过程中发现各类中文自然语言处理技术处理文本时会出现各种错误,例如:在分词中会把“文化局长”、“龙映台”分成了“文化局/n 长龙/n 映/v 台/q”。由于后续的词性标注、命名实体识别、句法分析是基于分词的结果进行处理,且上述技术本身也存在一定的误差,因此容易引起错误的逐层传递和不断扩大,例如:在命名实体识别中上述两个词被识别成“文化局/ORG 长龙/O 映/O 台/GPE”。从表4可知,待消解项

识别模块的正确率和召回率分别仅为75.9%和81.1%。

结论1 中文基础自然语言处理技术的不完备,严重影响了待消解项识别的准确率和召回率。

为了更好地评估待消解项对中文指代消解的影响,把待消解项分为自动获取的实体表达(AutoMention)和标准的实体表达(GoldMention)。利用各种基础自然语言处理技术获取的待消解项集合称为AutoMention,直接从标注语料中读取待消解项集合称为GoldMention。对比表5中AutoMention的Avg_F值与GoldMention的Avg_F值,可以清楚地得出各类模型的GoldMention分别比AutoMention高出10.5%、9.8%、9.2%、8.2%、12.1%。由此可知,好的待消解项集合能在很大程度上提高中文指代消解模型的消解结果。

结论2 待消解项识别的精度能显著提高指代消解模型的精度。

ACE2005中文语料的待消解项集合不仅有名词、代词,还有一定数量的嵌套名词短语,例如“负责策划这次艺术节活动的台北市文化局”、“南京西路建成圆环”等。简单地通过词性标注和命名实体识别抽取待消解项很难准确地抽取嵌套名词短语,因此本文用以下两个步骤抽取待消解项。1)抽取阶段:根据分词结果,对每一句进行句法分析,在句法分析树上抽取出所有包含NP、NN、NR的名词、代词以及名词短语作为待消解项。2)过滤阶段:由于上述阶段只考虑了待消解项识别的召回率,抽取结果中存在一定数量的重复名词短语。通过停用词过滤、数词过滤、重复中心词过滤删除非待消解项。在3个子语料中,待消解项识别F值能分别达到79.4%、81.3%、74.5%,具体消解结果如表4所列。

表4 待消解项识别模块性能(%)

待消解项识别模块	语料	P	R	F
Our MD	BNEWS	76.4	82.7	79.4
Our MD	NWIRE	79.3	83.6	81.3
Our MD	Weblogs	72.1	77.2	74.5
Our MD	All_avg	75.9	81.1	78.5
Baseline MD	All_avg	70.3	71.7	70.9

用Zhang等^[19]提出的待消解项模块作为的Baseline系统来验证上述方法的有效性,该系统只是简单地通过词性标注和命名实体识别抽取待消解项,在表4中用Baseline MD表示。表4中的Our MD表示本文所给出的待消解项识别模块。测评结果表明所给出的待消解项识别模块性能比Baseline系统好7.6%。

结论3 在句法分析树上抽取所有的待消解项更有效。

3.4 特征的抽取

从图1和表5可以清楚地看出,实体-实体表达模型的5种测评的F平均值分别高出实体表达对模型和实体排序模型2.3%和2.8%。相对于实体表达对模型和实体排序模型,实体-实体表达模型抽取的是部分聚类体和实体表达之间的特征属性。部分聚类体包含的各类信息会更加准确和丰富,很好地弥补了实体表达对信息不足的缺陷。实体排序模型虽然从整个篇章的角度考虑哪一个先行语最适合照应语,但该模型在抽取特征向量时和实体表达对模型一致,抽取的仍旧是实体表达对之间的特征属性,同样存在信息不足的缺陷。通过图1和表5可以看出实体排序模型的消解精度和实体表达对模型的Avg_F值基本类似,实体排序模型略低于实体-实体表达模型。

结论 4 部分聚类体所包含的属性信息远大于实体表达对之间的信息属性。

3.5 5类模型的比较

为了公平地比较 5 类模型的消解精度, 给定所有模型相同的特征属性以及相同的待消解项集合。通过图 1 和表 5 可以清楚地看出, 基于无监督的指代消解模型在 5 种测评中都低于其余 4 类模型, 主要原因是该模型只是简单通过距离函数来衡量照应语和先行语之间的指代关系。该距离公式中的阈值和所有的特征权重都是人为设定的, 因此存在一定的误差。3 种基于有监督学习的方法在 MentionWise 的测评结果较差, 而在 BLANG 的测评结果远好于 MentionWise 测评, 主要原因是训练语料中的正负例比例失衡, 使得分类器对负例的判断好于对正例的判断。从图 1 和表 5 可以清楚看出基于规则的层次过滤模型无论在 AutoMention 还是在 GoldMention 的 Avg_F 值都好于其余 4 类系统, 主要原因是实体表达对模型、实体-实体表达模型、实体排序模型、基于无监督的指代消解模型都是通过特征向量来描述待消解项, 并通过一个单独的模型判断待消解项之间是否存在指代关系。然而特征

向量中存在一些消解正确率较低的特征, 该类特征会覆盖消解正确率较高的特征, 从而引起模型的错误划分。基于规则的层次过滤模型不需要标注语料, 很好地弥补了中文标注语料不足的缺陷。该模型中各个层次的排列顺序按照每层的消解精度从高到低进行排序, 每一层的输入以上一层输出的实体聚类体为基础, 结合共享属性(性别、单复数等)得到最终的指代链。上述过程使得消解正确率高的特征先于消解正确率低的特征使用, 很好地弥补了单一模型中消解正确率低的特征覆盖消解正确率高的特征的不足。

结论 5 基于规则的层次过滤模型比较适合中文指代消解。

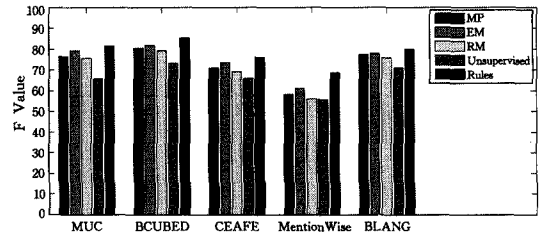


图 1 各类模型的 F 值对比

表 5 各类模型的性能比较(%)

System	ACE2005 中文语料																
	MUC			BCUBED			CEAFE			MentionWise			BLANG			Avg_F	
	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F		
AutoMention	MP	63.1	65.1	64.1	72.4	67.6	69.9	66.2	59.0	62.4	48.6	47.2	47.9	68.3	66.8	67.5	62.4
	EM	65.9	68.2	67.0	75.6	68.0	71.6	69.4	61.2	65.1	51.5	50.0	50.7	69.7	68.1	68.9	64.7
	RM	63.7	65.5	64.6	71.9	68.3	70.1	64.9	58.8	61.7	47.8	46.9	47.3	66.0	65.2	65.6	61.9
	Unsupervised	57.2	53.1	55.1	75.0	63.4	68.7	51.9	63.2	57.0	50.4	46.1	48.1	68.7	55.8	61.5	58.1
	Rules	64.3	71.5	67.7	78.7	73.6	76.0	69.7	61.1	65.1	54.9	50.3	52.5	67.9	72.1	69.9	66.1
GoldMention	MP	78.1	74.6	76.3	83.5	77.8	80.5	76.7	65.9	70.9	59.1	57.5	58.3	78.6	75.0	77.2	72.9
	EM	79.6	78.5	79.1	84.2	79.5	81.8	77.8	69.3	73.4	62.4	60.2	61.2	79.6	76.3	77.9	74.5
	RM	77.2	73.9	75.5	81.3	77.1	79.1	75.4	63.6	69.0	57.5	54.9	56.0	77.8	73.9	75.8	71.1
	Unsupervised	68.9	63.1	65.9	77.6	69.3	73.2	67.4	64.7	66.0	54.2	57.3	55.7	70.1	71.8	70.9	66.3
	Rules	79.3	83.7	81.4	82.9	87.6	85.1	78.6	73.9	76.1	68.2	69.0	68.6	78.5	80.9	79.7	78.2

结束语 本文以 ACE2005 中文语料为实验数据, 实现了 5 类经典的中文指代消解模型。通过大量的实验, 主要从待消解项识别、特征抽取和消解模型 3 个方面分析和探讨了影响中文指代消解模型精度的各种因素。通过比较 5 类模型可知, 基于规则的层次过滤模型不需要标注语料训练模型, 且该模型无论在 AutoMention 还是在 GoldenMention 的 Avg_F 值都好于其余 4 类模型, 因此该模型比较适合中文指代消解。

参考文献

[1] Hardmeier C, Federico M. Modelling pronominal anaphora in statistical machine translation[C]// Proceedings of the International Workshop on Spoken Language Translation. Paris: Konferensbidrag, 2010: 283-289

[2] Doddington G, Mitchell A, Przybocki M. The automatic content extraction (ACE) program-tasks, data, and evaluation [DB/OL]. 2012-05-11. <http://www.comp.nus.edu.sg/rpnlpir/proceedings/lrec-2004/pdf/.pdf>, 2012

[3] Witte R, Krestel R, Bergler S. Context based multidocument summarization using fuzzy coreference cluster graphs [DB/OL]. 2012-05-06. <http://www.nlp.nist.gov/projects/duc/pubs/2006.papers/20.final.pdf>, 2012

[4] Wang Hou-feng, He Ting-ting. Research on Chinese Pronominal

Anaphora Resolution[J]. Chinese Journal of Computers, 2001, 24(2): 136-143 (in Chinese)

王厚峰, 何婷婷. 汉语中人称代词的消解研究[J]. 计算机学报, 2001, 24(2): 136-143

[5] Wang Hou-feng, Mei Zheng. Robust Pronominal Resolution within Chinese Text[J]. Journal of Software, 2005, 16(5): 700-707 (in Chinese)

王厚峰, 梅铮. 鲁棒性的汉语人称代词消解[J]. 软件学报, 2005, 16(5): 700-707

[6] Cardie C, Wagstaff K. Noun phrase coreference as clustering [C]// Proceedings of the 1999 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora. College park: Kluwer Academic Publishers, 1999: 281-289

[7] Zhou Jun-sheng, Huang Shu-jian, Chen Jia-jun, et al. A New Graph Clustering Algorithm for Chinese Noun Phrase Coreference Resolution[J]. Journal of Chinese Information Processing, 2007, 21(2): 77-82 (in Chinese)

周俊生, 黄书剑, 陈家骏, 等. 一种基于图划分的无监督汉语指代消解算法[J]. 中文信息学报, 2007, 21(2): 77-82

[8] Soon W M, Ng H T, Lim D C Y. A machine learning approach to coreference resolution of noun phrases[J]. Computational Linguistics, 2001, 27(4): 521-544

- [12] Yang Chen-yu, Zhu Li-xin, Ling Zhen-hua, et al. Automatic Phrase boundary labeling for a Mandarin TTS corpus using the Viterbi decoding algorithm[J]. Journal of Tsinghua University (Science and Technology), 2011, 51(9): 1267-1281 (in Chinese)
杨辰雨, 朱立新, 凌震华, 等. 基于 Viterb 解码的中文合成音库韵律短语边界自动标注[J]. 清华大学学报(自然科学版), 2011, 51(9): 1276-1281
- [13] Li Jian-feng, Hu Guo-ping, Wang Ren-hua. New Prosody Phrase Prediction Model Based on Whole Sentence Similarity Computing[J]. Journal of Chinese Computer Systems, 2006, 27(10): 1935-1938 (in Chinese)
李剑锋, 胡国平, 王仁华. 基于整句相似性计算的韵律短语预测模型[J]. 小型微型计算机系统, 2006, 27(10): 1935-1938
- [14] Dong Hong-hui, Tao Jian-hua, Xu Bo. Chinese Prosodic Phrasing with a Constraint based Approach[J]. Journal of Chinese Information Processing, 2007, 21(1): 54-59 (in Chinese)
董宏辉, 陶建华, 徐波. 基于约束模型的韵律短语预测[J]. 中文信息学报, 2007, 21(1): 54-59
- [15] Shao Yan-qi, Hui Zhi-fang, Han Ji-qing, et al. A Study on Chinese Prosodic Hierarchy Prediction Based on Dependency Grammar Analysis[J]. Journal of Chinese Information Processing, 2008, 22(2): 116-123 (in Chinese)
邵艳秋, 穗志方, 韩纪庆, 等. 基于依存句法分析的汉语韵律层级自动预测技术研究[J]. 中文信息学报, 2008, 22(2): 116-123
- [16] Yang Hong-wu, Zhu Ling. Predicting Chinese Prosodic boundary based on syntactic features[J]. Journal of Northwest Normal University (Natural Science), 2013, 49(1): 41-45 (in Chinese)
杨鸿武, 朱玲. 基于句法特征的汉语韵律边界预测[J]. 西北师范大学学报(自然科学版), 2013, 49(1): 41-45
- [17] Zhang Yuan-ping, Ling Zhen-hua, Dai Li-rong, et al. Improved decision tree based method for English prosodic phrase boundary Prediction[J]. Application Research of Computers, 2012, 29(8): 2921-2925 (in Chinese)
张元平, 凌震华, 戴礼荣, 等. 一种改进的基于决策树的英文韵律短语边界预测方法[J]. 计算机应用研究, 2012, 29(8): 2921-2925

(上接第 34 页)

- [9] Yang X Q, Su J, Zhou G D, et al. An NP-cluster based approach to coreference resolution[C]//Proceedings of the 20th international conference on Computational Linguistics. Morristown: ACL, 2004: 226-232
- [10] Denis P, Baldridge J. A Ranking Approach to Pronoun Resolution[C]//International Joint Conferences on Artificial Intelligence. Hyderabad: ACL, 2007: 1588-1593
- [11] Rahman A, Ng V. Narrowing the modeling gap: a cluster-ranking approach to coreference resolution [J]. Journal of Artificial Intelligence Research, 2011, 40(1): 469-521
- [12] Kong Fang, Zhu Qiao-ming, Zhou Guo-dong, et al. Coreference Resolution Based on Center Theory [J]. Computer Science, 2009, 36(6): 219-222 (in Chinese)
孔芳, 朱巧明, 周国栋, 等. 基于中心理论的指代消解研究[J]. 计算机科学, 2009, 36(6): 219-222
- [13] Gao Jun-wei, Kong Fang, Zhu Qiao-ming, et al. Research of Chinese Noun Phrase Anaphora Resolution: A SVM-based Approach[J]. Computer Science, 2012, 39(10): 231-234 (in Chinese)
高俊伟, 孔芳, 朱巧明, 等. 基于 SVM 的中文名词短语指代消解研究[J]. 计算机科学, 2012, 39(10): 231-234
- [14] Kong Fang, Zhou Guo-dong. Pronoun Resolution in English and Chinese Languages Based on Tree Kernel[J]. Journal of Software, 2012, 23(5): 1085-1099 (in Chinese)
孔芳, 周国栋. 基于树核函数的中英文代词消解[J]. 软件学报, 2012, 23(5): 1085-1099
- [15] Xi Xue-feng, Zhou Guo-dong. Pronoun Resolution Based on Deep Learning [J]. Acta Scientiarum Naturalium Universitatis Pekinensis, 2014, 50(1): 100-110 (in Chinese)
奚雪峰, 周国栋. 基于 Deep Learning 的代词指代消解[J]. 北京大学学报(自然科学版), 2014, 50(1): 100-110
- [16] Raghunathan K, Lee H, Rangarajan S. A multipass sieve for coreference resolution [M]. MIT, Massachusetts: ACL, 2010: 492-501
- [17] Lee H, Peirsman Y, Chang A, et al. Stanford's multi-pass sieve coreference resolution system at the CoNLL-2011 shared task [C]//Proceedings of the Fifteenth Conference on Computational Natural Language Learning, Shared Task. Oregon: ACL, 2011: 28-34
- [18] Lee H, Chang A, Peirsman Y, et al. Deterministic coreference resolution based on entity-centric, precision-ranked rules [J]. Computational Linguistics, 2013, 39(4): 885-916
- [19] Zhang X C, Wu H. Chinese coreference resolution via ordered filtering[C]//Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning: Shared Task. Jeju: ACL, 2012: 95-99
- [20] Kong Fang, Zhu Qiao-ming, Zhou Guo-dong, et al. Anaphoricity Determination for Coreference Resolution in English and Chinese[J]. Journal of Computer Research and Development, 2012, 49(5): 1072-1085 (in Chinese)
孔芳, 朱巧明, 周国栋, 等. 中英文指代消解中待消解项识别的研究[J]. 计算机研究与发展, 2012, 49(5): 1072-1085
- [21] Marc V, John B, John A, et al. A model theoretic coreference scoring scheme[C]//Proceedings of the 6th Message Understanding Conference. Stroudsburg: ACL, 1995: 45-52
- [22] Amit B, Breck B. Algorithms for scoring coreference chains [C]//Proceedings of LREC. Granada: ACL, 1998: 563-566
- [23] Luo X Q. On coreference resolution performance metrics[C]//Proceedings of HLT-EMNLP Stroudsburg. ACL, 2005: 25-32
- [24] Nong Y. The Handbook of Data Mining [M]. Cleveland CRC Press, 2001: 247- 277
- [25] Marta R, Eduard H. BLANC: Implementing the Rand Index for coreference evaluation[J]. Natural Language Engineering, 2011, 17(4): 485-510
- [26] Doddington G, Mitchell A, Przybocki M. The automatic content extraction (ACE) program-tasks, data, and evaluation [DB/OL]. 2012-5-11. [http:// www. comp. nus. edu. sg/~rpnlp/ proceedings/ lrec-2004 /pdf/ 5. pdf](http://www.comp.nus.edu.sg/~rpnlp/ proceedings/ lrec-2004 /pdf/ 5. pdf)