计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.210100226

基于规则的有标复句关系的自动识别

杨进才[1]　胡巧玲[1]　胡　泉[2]
1 华中师范大学计算机学院　武汉 430079
2 华中师范大学人工智能教育学部　武汉 430079

摘　要　汉语复句的语义表达复杂,复句关系分类问题作为汉语篇章研究与应用的重要内容,一直是自然语言处理领域关注的热点。文中总结与挖掘出复句类别自动识别的十几类字面、句法特征,将特征形式化为规则,用关系词触发规则的机制,对有标复句进行十二类关系类别的识别。实验结果表明该方法取得了较高的准确率,优于现有的方法。

关键词:有标复句;复句关系分类;规则;自动识别
中图法分类号　TP391

Rule-based Automatic Recognition of Relations for Marked Complex Sentences

YANG Jin-cai[1],HU Qiao-ling[1] and HU Quan[2]
1 School of Computer,Central China Normal University,Wuhan 430079,China
2 Faculty of Artificial Intelligence in Education,Central China Normal University,Wuhan 430079,China

Abstract　Semantic expression of Chinese complex sentences is complicated. As an important content of Chinese discourse studies,complex sentences classification has always been a hot spot in the field of natural language processing. This paper summarizes and excavates more than ten types of literal and syntactic features for automatic identification of complex sentence categories,formalizes the features and constitutes rules,and uses the mechanism of relational words to trigger the rules to identify twelve types of relationship categories for marked complex sentences. Experimental results show that this method has achieved a higher accuracy rate,which is better than the existing methods.

Keywords　Marked complex sentences,Complex sentence classification,Rules,Automatic recognition

1　引言

自然语言处理中的句子级分析技术分为词法分析、句法分析、语义分析3个层面。语义分析是自然语言处理的核心技术,决定自然语言处理的发展进程。复句是包含两个或两个以上分句的句子,它连接小句和篇章,并在二者之间起了衔接作用,同时兼有语法、语义等多个属性。复句关系的识别研究为篇章语义分析奠定了基础。

复句关系识别具有广泛的应用,包括篇章分析[1]、信息抽取[2]、自动问答[3]以及机器翻译[4]等自然语言处理相关领域。例如,转折关系可以消除句子内极性的歧义,而因果关系检测可以改善问答系统和事件关系的提取能力。

中文复句中,根据有无关系标志(关系词),可分为有标复句和无标复句;著名语言学家邢福义提出的汉语复句分类三分系统将分句间的逻辑语义关系分为因果、转折、并列三大类[5],因果类复句又分为因果、推断、假设、条件、目的五小类,转折类复句分为转折、让步、假转三小类,并列类复句分为并列、连贯、递进 、选择四小类。复句关系类别识别就是对复句分句间逻辑语义关系的识别。

2　相关工作

从中文信息处理的角度来看,对复句进行的研究主要集中在关系词的自动识别、复句层次的划分以及复句关系类别的自动识别3个方面。

在复句关系词自动识别方面,Hu等[6-7]采用了基于规则的方法对关系词的标识进行了探讨,总结了关系词自动识别中12种规则的约束类型[7],并结合词性标记和关系词搭配理论,提出"正向选择算法"来标识关系词,以及"包含匹配算法"[8]进行规则解析;Jia[9]根据语料库建立规则,利用规则的结论来标识准关系词;Zhou等[10]根据清华汉语树库的标注方法,利用规则从中提取复句关系词并标注其类别;Hu和Yang等[11-13]建立了较完善的关系词识别规则库。Yang等[14]基于统计的方法,使用贝叶斯模型来对复句关系词进行识别。

对于复句关系类别的识别,大多是通过关系词的识别来进行,即先标识出句中的关系词,再对复句的语义、语法、层次结构进行分析,从而判断出复句的关系类别。目前,在关系词识别成果的基础上,主要是利用机器学习、深度学习以及基于统计的方法来进行复句关系的类别识别。基于机器学习的方法中,Huang等[15]提出了一种半监督的学习方法,从小标签数据集和大标签数据集中学习关系标记的概率分布,发现成对出现的关系词为话语关系解析提供了强有力的线索。Chen[16]结合词性、关系词、极性等特征,利用决策树算法来识别汉语句子间的因果与并列关系。基于统计的方法中,Yang等提出语义相关度计算方法,并结合关系标记对非充盈态的

基金项目:国家社科基金(19BYY092)
This work was supported by the National Social Science Fundation of China(19BYY092).
通信作者:杨进才(jcyang@mail.ccnu.edu.cn)

二句式复句进行类别识别[17]。基于深度学习的方法中，Wang 等采用在卷积网络中融合关系词特征的 FCNN 模型，通过学习自动分析两个分句之间的语法语义等特征，从而识别出复句的关系类别[18]。Li 等通过在复句语义关系识别任务上使用局部特征与全局特征相结合的方法，提出了一种基于句内注意力机制的多路 CNN 网络结构 Inatt-MCNN[19]。

上述方法中，基于深度学习的方法是一种黑盒模型，无法解释做出的决策，推理过程不够透明，无法解释关系类别识别使用到的词性、语法特征的作用。深度学习的方法根据已有的数据来学习，需要大量的计算性能来构建，对于小数据集问题，在计算开销和时间相同的情况下，并没有产生比其他方法更好的效果。基于规则的方法可解释性更好，能方便使用语法特征，决策的过程也更加透明，不仅适用于大数据集，也适用于小数据集。考虑到现有的复句语料库中有标复句的数量不够大以及各个类别的数量分布不均匀，本文提出用基于规则的方法识别有标复句的关系类别。

3 有标复句关系识别的特征

基于规则的有标复句类别识别方法的基本思路是：先制定一系列规则，然后构造成规则库，利用关系标志触发规则的机制，对有标复句中的关系类别进行判别。

本文的研究对象为分句中含有关系标记的有标复句。关系标记对复句的关系类别起重要的标示作用，可以借助关系标记来对复句关系进行识别。但在搭配使用的关系标记一部分省略的关系词非充盈态复句中，剩下的关系词对应多种类别；还有一些跨类别的标记，这些标记对应多种类别。因此，需要综合考虑关系词出现的句子的各种特征，包括字面、词性、句法等特征。

本文根据语言学的研究成果以及用于关系词识别的规则，总结出用于判别有标复句类别的如下 7 类特征：

〈Word，FOB，endClause，FrontWords，BackWords，MatchWords，isSame〉

3.1 关系词 Word

关系词作为复句分句间的连接词，对复句的语义关系有显示、转化作用，因此关系词是复句关系类别识别不可忽视的特征。虽然有的关系标记和关系类别是一对多的映射，一种关系类别可能有多个关系标记对应（如图 1 所示），但将关系标记特征提取出来对一对多关系起到类别范围限定作用。

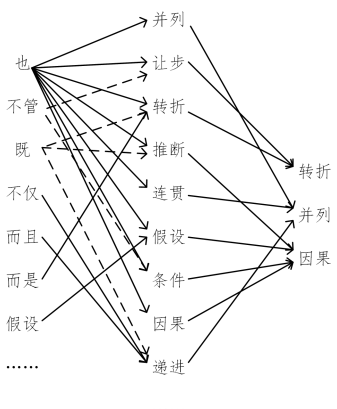

图 1 关系标记——关系类别对应

Fig. 1 Relation marking to relation category correspondence

3.2 前呼与后应 FOB

一个关系词与另一个关系词搭配使用时，出现的位置不同，将对应不同的关系类别。关系词有时作为“前呼”，有时作为“后应”，甚至有时单独出现。在例 1 中，“所以”作为前呼标记，与“是为了”进行搭配，标识句子为目的关系。在例 2 中，“所以”作为“后应”标记，与“因为”进行搭配，表示因果关系。在例 3 中，“所以”单独出现，表示因果关系。

例 1 过去在计划经济下，政府所以必须干预、控制农产品的生产、经营，是为了以低于市场均衡的价格从农民那里取得这些产品来支持重工业的优先发展。

例 2 贫农，因为最革命，所以他们取得了农会的领导权。

例 3 老炳父子住的地方偏僻，所以来晚了一步。

3.3 结束符 endClause

结束符有“！”“？”“。”等，当“？”作为结束符时，可表示推断关系。在例 4 中，前分句包含“既然如此”的意思，当句子中仅仅出现“又”关系标记时，无法确定它是连贯、因果、转折或是推断关系，因此当结束符为“？”时，显示句子为反问句，表推断。

例 4 你知道我的伤势不碍事，又何必担心？

3.4 前词 FrontWords

当复句中仅靠关系词无法识别关系类别，或者前一词使复句从某种关系转变为另一种关系时，关系词前一词这一特征对准确识别复句关系类别就显得尤为重要。在例 5 中，仅靠一个关系词“又”无法确定关系类别，当前一词为“就”时，可识别为连贯关系。在例 6 中，“既”“也”搭配表示并列，但“也”的前一词是“但”，这时复句就为转折关系。

例 5 梁家夫妇站到门房外边，对外看了一眼，就又要关上房门。

例 6 她既没有踊跃赞同，但也没有露出一丝不愿的样子。

3.5 后词 BackWords

同关系词的前一词一样，关系词的后一词也是识别复句关系的关键特征。在例 7 中，“一”“就”作为关系标记有时表连贯关系，有时表示因果关系，有时表示条件，当关系词“就”后面出现“会”“要”“能”等助动词时，确定为条件关系。

例 7 当然，一谈到这方面，就会遇到许多困难。

3.6 搭配词 MatchWords

当某一关系词出现时，大多数情况下复句中会有与之搭配的另一个关系词，不同的关系词进行搭配表示不同的关系。在例 8 中，“既”“更”搭配表示递进关系，在例子 9 中“既”“就”搭配表示推断关系。

例 8 咱们要找的既不是牛，更不是猪。

例 9 你既有这个意思，就没有什么说的了。

3.7 谓语动词相同 isSame

当字面特征不足以确定复句关系时，需要对句子进行句法分析。在例 10 中，只有“还”一个关系标记，其他字面特征也不明显，这时对句子进行句法分析，分析图如图 2 所示。前分句与后分句之间存在“COO”并列关系，且词语都为“请”，因此，利用这一特征可以判别句子为并列关系。

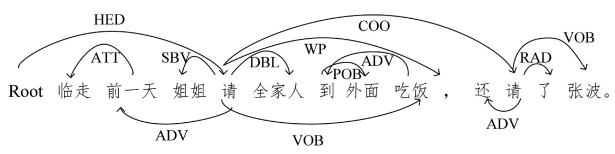

图 2 句法分析图

Fig. 2 Syntactic analysis diagram

例 10 临走前一天姐姐请全家人到外面吃饭，还请了张波。

使用上述7类特征对复句进行特征提取的过程如下。

1)使用LTP对复句进行分词处理,抽取复句中出现的关系标记,将Word特征值设置为关系标记,若复句中仅有一个关系标记,则设置FOB特征值为S,MatchWords特征值设置为null;若复句中有两个关系标记,根据关系标记出现的先后顺序,设置特征值为F/B,MatchWords特征值设置为搭配关系词。

2)根据1)中得到的Word特征和MatchWords特征获取FrontWords特征值和BackWords特征值,并根据关系词和搭配关系词所在分句的结束符号获取相应的endClause特征值。

3)判断关系词和搭配关系词所在分句间是否存在并列关系。若存在,判断支配词与被支配词是否相同,相同则isSame值为1,不同则其值为0。

4 条件约束的形式化表示

产生式知识表示法是常用的知识表示方式之一,它是依据人类大脑记忆模式中的各种知识之间大量存在的因果关系,并以"IF-THEN"的形式,即产生式规则表示出来的。这种形式的规则捕获了人类求解问题的行为特征,并通过认识——行动的循环过程求解问题。

产生式规则公式为:

$$R_k:\bigvee_{i=1}^{m}\left(\bigwedge_{j=1}^{n}E_{ijk}\right)\rightarrow C_k \tag{1}$$

其中,$m,n>1$,$k=1,2,\cdots,r$;R_k表示第k条规则;E_{ijk}表示第k条规则的第i类约束类型的第j个条件表达式;R_k左部表示条件表达式的组合;C_k表示第k条规则的结论。

为了便于计算机的处理,将每个规则具体表示为如下形式:

$\langle RuleType=\#, ID=\#, Specifications=\#\rangle$

其中,$RuleType$:规则的类别,对应某一对关系词;ID:该规则类型的规则集的某一条规则的编号;$Specifications$:假设样本有n个特征,对应有k个类别,定义为$C_1,C_2,\cdots,C_k$,则规则形式化定义如下:

$$\exists R_k\in R, R_k=(R_1^1=V_1^1\wedge R_1^2=V_1^2\wedge\cdots\wedge R_1^j=V_1^j)\vee(R_2^1=V_2^1\wedge R_2^2=V_2^2\wedge\cdots\wedge R_2^j=V_2^j)\vee\cdots\vee(R_i^1=V_i^1\wedge R_i^2=V_i^2\wedge\cdots\wedge R_i^j=V_i^j)\rightarrow Y=C_k \tag{2}$$

其中,R_k表示规则集合中某一条规则,R_i^j表示第i种约束类型的第j个特征函数,V_i^j表示具体特征值。

例12 当然,一谈到这方面,就会遇到许多困难。

经过LTP进行依存句法分析后的结果:

{'id':1,'word':'当然','pos':'d','parent':4,'relation':'ADV'}

{'id':2,'word':',','pos':'wp','parent':1,'relation':'WP'}

{'id':3,'word':'一','pos':'d','parent':4,'relation':'ADV'}

{'id':4,'word':'谈','pos':'v','parent':0,'relation':'HED'}

{'id':5,'word':'到','pos':'v','parent':4,'relation':'CMP'}

{'id':6,'word':'这','pos':'r','parent':7,'relation':'ATT'}

{'id':7,'word':'方面','pos':'n','parent':4,'relation':'VOB'}

{'id':8,'word':',','pos':'wp','parent':4,'relation':'WP'}

{'id':9,'word':'就','pos':'d','parent':10,'relation':'ADV'}

{'id':10,'word':'会','pos':'v','parent':11,'relation':'ADV'}

{'id':11,'word':'遇到','pos':'v','parent':4,'relation':'COO'}

{'id':12,'word':'许多','pos':'m','parent':13,'relation':'ATT'}

{'id':13,'word':'困难','pos':'a','parent':11,'relation':'VOB'}

{'id':14,'word':'。','pos':'wp','parent':4,'relation':'WP'}

例句12的特征表示为:

$V_w=\langle$就,B,"。","，","会","一",N$\rangle$

5 复句特征触发规则

如果规则R_i中关系词W的每个条件都能在关系词W对应的特征向量V_w中找到对应项,并且属性值都相等,则称特征向量W满足规则R,即$W\subset R_i$。

复句特征触发规则的流程如下:

1)首先判别有标复句中的关系词是否可以在一对一关系表中找到对应类别。

2)若使用一对一关系表(见表1)不能直接判断,则分析复句中关系词的特征向量V_w。

3)从一对多关系表(见表2)中取出与当前关系词有关的规则,记为R_i。

4)将特征向量V_w变换成一条临时规则的左部,记为$R_T=g(V_w)$。

5)判断临时规则R_T的左部与规则R_i的左部是否匹配,如果匹配,则按照规则R_i的右部进行操作(标识复句为12类关系中的一种);如果不匹配,则返回第3)步,与下一条规则进行匹配。

表1 一对一关系表

Table 1 One-to-one relationship

id	keymarks	fob	relation
37	不只是	f	dj
38	才能	b	tj
41	此后	nan	lg
42	此外	f	bl
43	但	b	zz
44	但凡	f	tj
45	但凡是	f	tj

表2 一对多关系表

Table 2 One-to-many relationship

id	keymarks	constraint Type	Constraints	relation
61	就	1+2	FOB(就)=F∧MatchWords(就)=也	rb
62	就	1+2	FOB(就)=B∧MatchWords(就)=不/一旦/果然/假定/假设/假使/假若/假如/如/如果/若/若是/倘/倘若/倘使	js
65	就	1+2+4	FOB(就)=B∧MatchWords(就)=一/不管/不论/一旦/一经/只要·BackWords(就)=会/要/能	tj

步骤 5)中的变换过程如下：

V_w=〈$Word\_w$，FOB_w，$endClause\_w$，$FrontWords\_w$，$BackWords\_w$，$MatchWords\_w$，$isSame\_w$〉

$R_T=g(V_w)=\exists V_w(V_w.Words=Word\_w \wedge V_w.FOB=FOB\_w \wedge V_w.BackWords=BackWords\_w \wedge V_w.MatchWords=MatchWords\_w)$

R_T的左部与规则R_i的左部匹配，是指对于R_i左部的任意一个条件，R_T中与之对应的项的值都满足。

例 12 的特征向量转换成临时规则为：

V_w=〈就，B，“。”，“，”，“会”，“一”，N〉

$R_T=g(V_w)=\exists V_w(V_w.Words=$就$\wedge V_w.FOB=B \wedge V_w.BackWords=$“会”$\wedge V_w.MatchWords=$一）

存在一条规则R_i：

$\exists W(W.Words=$“就”$\wedge$ $W.FOB=$“B”$\wedge W.MatchWords=$“一/不管/不论/一旦/一经/只要”$\wedge$ $W.BackWords=$“会/要/能”）

这里V_w与 W 对应。

对于R_i的条件 $W.Words=$“就”，将其对应的项 $V_w.Words=$“就”的值代入条件的属性 $W.Words$，有表达式“就”=“就”，此表达式的值为真。

对于R_i的条件 $W.FOB=$“B”，将其对应的项 $V_w.FOB=B$ 的值代入条件的属性 $W.FOB$，有表达式 “B”=“B”，此表达式的值为真。

对于R_i的条件 $W.BackWords=$“会/要/能”，将其对应的项 $V_w.BackWords=$“会” 的值代入条件的属性 $W.BackWords$，有表达式“会” ⊂“会/要/能”，此表达式的值为真。

对于R_i的条件 $W.MatchWords=$“一/不管/不论/一旦/一经/只要”，将其对应的项$V_w.MatchWords=$“一”的值代入条件的属性 $W.BackWords$，有表达式“一” ⊂“一/不管/不论/一旦/一经/只要”，此表达式的值为真。

对于R_i的所有条件，R_T都能使其表达式为真，所以例 2 中的复句关系可判别为条件关系。

6 规则的一致性检测

当规则设计完成后，需要对其中的规则进行一致性检测，规则库中可能存在从属规则、冲突规则、等价规则以及循环规则。

本文规则的推理过程是一步推理，不会产生中间结果，这里使用一致性检测模型对等价规则、冲突规则和从属规则进行检测。等价规则、冲突规则、从属规则的定义如下。

1)等价规则

$\exists P_1 \rightarrow Q_1$，$\exists P_2 \rightarrow Q_2$，

IF $P_1=P_2$ and $Q_1=Q_2$：

THEN 规则等价

2)冲突规则

$\exists P_1 \rightarrow Q_1$，$\exists P_2 \rightarrow Q_2$，

IF $P_1=P_2$ and $Q_1 \neq Q_2$：

THEN 规则冲突

3)规则从属

$\exists P_1 \rightarrow Q_1$，$\exists P_2 \rightarrow Q_2$，$\exists P_3 \rightarrow Q_3$

IF $P_1 \rightarrow Q_1 \subseteq P_2$ and $Q_2=Q_3$：

THEN 规则从属

为规则设计对象推理树，将关系词相同的规则集合对应一棵推理树，偶数层为特征函数，奇数层为特征值。表 2 中关于“就”的规则对象推理树如图 3 所示。

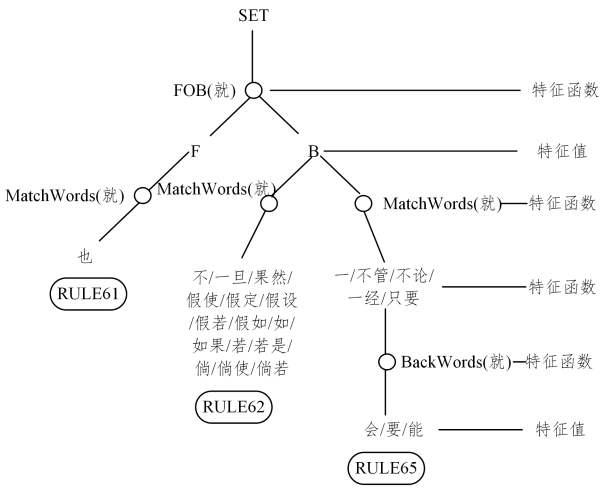

图 3 对象推理树

Fig. 3 Object inference tree

由推理树得出：“就”所在规则集中不存在矛盾规则、等价规则、从属规则。

7 实验结果与分析

从汉语复句语料库 CCCS(the Corpus of Chinese complex Sentences)中随机选择有标复句先进行人工标识，确定复句关系类别，得到每个类别复句 900 条，12 个类别共 10800 条；再利用规则库中的规则对复句关系进行标识，标识结果有可标识复句关系类别和不可标识复句关系类别(不能判断)两种，利用召回率(R)、精确率(P)、F1 评价实验结果。

实验整体结果如表 3 所列。

表 3 12 类有标复句识别结果

Table 3 Twelve types of marked complex sentence recognition results

（单位：%）

类型	R	P	$F1$
因果	87.65	93.03	90.26
推断	89.10	97.68	93.19
假设	91.23	97.51	94.27
条件	90.78	97.73	94.12
目的	96.67	99.89	98.25
并列	97.78	82.94	89.75
连贯	85.89	97.85	91.48
递进	95.33	98.85	97.06
选择	94.78	97.15	95.95
转折	91.55	95.81	93.63
让步	90.78	93.69	92.21
假转	92.77	98.35	95.48
平均	92.02	95.87	93.80

从以上实验结果可以看出，10800 条有标复句，可通过规则识别出来的有 10393 条，说明规则涵盖度还不够高，还需进一步完善与挖掘。从表 3 中可以看出，因果大类复句正确率普遍较低，分析测试结果可以发现，复句中通常存在跨大类关系标志，如“…就…”，“…也…”或者“又”等，而这些跨大类的关系标志，在某些情况下几乎都可以表示并列类关系；对连贯

类复句，正确率低，而精确率却很高，说明规则设计对于连贯类复句的覆盖还不够全面。

为了验证基于规则的识别方法的性能，选取了 SVM 和 textRCNN[20] 两个方法在同一数据集进行对比，表 4 列出了在 SVM 和 textRCNN 在 Precision，Recall 和 Macro-F1 上不同的表现结果。

表 4 不同方法的实验结果比较

Table 4 Comparison of experimental results of different methods

（单位：%）

方法	P	R	F1
SVM	72.44	61.65	64.80
TextRCNN	89.72	88.42	88.61
规则	92.02	95.87	93.80

结果表明，对于同样的数据集，传统的机器学习方法表现不佳。通常 SVM 适用于二分类，多分类问题就要通过多个二类支持向量机的组合来解决。当应用于复句的分类问题时，SVM 使用 TF-IDF 进行复句特征的提取，当所有类别中词语出现的频率越高，它的区分度则越小，权值也越低；而在某一类复句中，词语出现的频率越高，区分度越大，权重也越大。在数据集中，对于跨大类的关系标记，分类器给定的权重小，所以导致了 SVM 的分类效果不理想。TextRCNN 模型中，在 word embedding 层中对词进行词向量编码，之后将词向量输入到双向 LSTM 中，得到所有时刻的隐藏层状态（前向隐藏和后向隐藏拼接），并与 embedding 值拼接来表示一个词，然后用最大池化层筛选出复句中最重要的特征信息。TextRCNN 可以获得每条复句中词汇的上下文信息，并取得最重要的特征，在识别包含跨大类关系标记如“…就…”这样的复句时，它可以结合关系标记的上下文信息确定为因果并列还是转折类复句，但是当复句中同时出现两种关系标记如“但”“如果”连用就会错判，难以确定属于转折还是假设。而基于规则的方法中，复句中出现“但因为”“但如果”这样的连用标记时，根据规则，若连用标记为句子开头，则判断为因果、假设类关系；若不在句首则均为转折关系。由于关系标记连用的复句在数据集中分布比较少，导致 textRCNN“学习”得不够深入，从而误判类别。总的来说，虽然基于深度学习的方法可以自动提取特征信息，但是基于规则的方法使用了语言专家人工总结的知识，不仅取得了较好的分类效果，同时具有较强的说服力。

结束语 本文吸收复句研究的语言学成果，总结出了用于复句关系自动识别的关系标记、配位位置、搭配关系词等重要语句特征；根据产生式规则公式制定了一系列规则，建立了规则库；通过特征触发规则对有标复句进行关系类别识别；实验表明，本文提出的方法相比于机器学习和深度学习方法具有更高的识别准确率。

由于复句关系的自动识别对规则的完善程度有较大的依赖，在接下来的工作中，进一步利用语言学研究的成果，通过结合机器学习或是深度学习的方法进一步挖掘句内深层次的语义信息和句间的关联特征，以及利用 FP-tree 算法挖掘隐藏规则，完善规则库，从而提升复句关系自动识别的准确率。

参 考 文 献

[1] LIU Y, LI S, ZHANG X, et al. Implicit discourse relation classification via multi-task neural networks[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2016.

[2] ZOU B, ZHOU G, ZHU Q. Negation focus identification with contextual discourse information[C]// Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2014: 522-530.

[3] LIAKATA M, DOBNIK S, SAHA S, et al. A discourse-driven content model for summarising scientific articles evaluated in a complex question answering task[C]// Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing. 2013: 747-757.

[4] PAPINENI K. BLEU: A method for automatic evaluation of machine translation [C]// Proceedings of the 40 th Annual Meeting of the Association for Computational Linguistics (ACL). 2002.

[5] XING F Y. Research on Chinese Complex Sentences[M]. Beijing: The Commercial Press, 2001: 1-37.

[6] HU J Z, SHU J B, HU Q, et al. Research on the expression method of rules in the automatic recognition of relative words in complex sentences[J]. Computer Engineering and Applications, 2016, 52(1): 127-132.

[7] HU J Z, SHU J B, HU Q, et al. Research on the Restrictive Conditions of Rules in the Automatic Recognition of Relative Words in Chinese Complex Sentences[J]. Language and Writing Applications, 2015(1): 82-89.

[8] HU J Z, HU Q, SHU J B. Research on Inclusion Matching Algorithm of Rule Analysis in Automatic Recognition of Relative Words in Complex Sentences[J]. Journal of Central China Normal University (Natural Science Edition), 2014, 48(5): 643-649.

[9] JIA S M, LEI L L, HU M S. Rule-based automatic identification of relative words in complex sentences[J]. Journal of Chinese Information Processing, 2015, 29(1): 44-48, 66.

[10] LI Y C, SUN J, ZHOU G D, et al. Research on Recognition and Classification of Relative Words in Complex Sentences Based on Tsinghua Chinese Treebank[J]. Journal of Peking University (Natural Sciences Edition), 2014, 50(1): 118-124.

[11] HU J Z, CHEN J M, YANG J C, et al. Research on Automatic Marking of Joint Relation Marking Based on Rules[J]. Computer Science, 2012, 39(7): 190-194.

[12] YANG J C, XIE F, WANG Z H, et al. Conflict Detection and Processing in the Rule Generation System of Complex Sentence Relative Words[J]. Journal of Chinese Information Processing, 2015, 29(4): 8-15.

[13] YANG J C, XIE F, HU J Z. Research on Rule Engine in Automatic Identification of Relative Words in Chinese Complex Sentences[J]. Computer Science, 2014, 41(S2): 25-28.

[14] YANG J C, GUO K K, SHEN X J, et al. Automatic recognition and rule mining of complex sentences based on Bayesian model [J]. Computer Science, 2015, 42(7): 291-294, 319.

[15] HUANG H H, CHANG T W, CHEN H Y, et al. Interpretation of Chinese discourse connectives for explicit discourse relation recognition[C]// The 25th International Conference on Computational Linguistics: Technical Papers (COLING 2014). 2014: 632-643.

[16] HUANG H H,CHEN H H. Contingency and comparison relation labeling and structure prediction in Chinese sentences[C]// Proceedings of the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue. 2012:261-269.

[17] YANG J C,CHEN Z Z,SHEN X J,et al. Automatic identification of the relational category of marked complex sentences in the non-filling state of two sentences[J]. Application Research of Computers,2017,34(10):2950-2953.

[18] YANG J C,WANG Y Y,CAO Y,et al. Feature fusion CNN relation recognition method of relative words non-filled complex sentences[J]. Computer Systems & Applications,2020,29(6):224-229.

[19] SUN K L,DENG Z H,LI Y,et al. Recognition of Chinese Complex Sentence Relations Based on Multi-channel CNN Based on In-Sentence Attention Mechanism[J]. Journal of Chinese Information Processing,2020,34(6):9-17,26.

[20] LAI S,XU L,LIU K,et al. Recurrent convolutional neural networks for text classification[C]// Twenty-ninth AAAI Conference on Artificial Intelligence. 2015.

YANG Jin-cai, born in 1967, doctor, professor, doctoral supervisor, is a member of China Computer Federation. His main research interests include modern database and information system, Chinese information processing, artificial intelligence and natural language processing.

(上接第 101 页)

[21] KASHAN A H. League Championship Algorithm:A New Algorithm for Numerical Function Optimization[C]// International Conference of Soft Computing and Pattern Recognition. Malacca,2009:43-48.

[22] GHORBANI N,BABAEI E. Exchange market algorithm[J]. Applied Soft Computing Journal,2014,19:177-187.

[23] ALATAS B. ACROA:Artificial Chemical Reaction Optimization Algorithm for global optimization[J]. Expert Systems with Applications,2011,38(10):13170-13180.

[24] BECHIKH S,CHAABANI A,BEN SAID L. An Efficient Chemical Reaction Optimization Algorithm for Multiobjective Optimization[J]. IEEE Trans. Cybern,2015,45(10):2051-2064.

[25] TRUONG T K,LI K,XU Y. Chemical reaction optimization with greedy strategy for the 0-1 knapsack problem[M]. Elsevier Science Publishers B. V.,2013.

[26] MORADI P,GHOLAMPOUR M. A hybrid particle swarm optimization for feature subset selection by integrating a novel local search strategy[J]. Applied Soft Computing,2016,43(C):117-130.

[27] YU H,ZHAO N,WANG P,et al. Chaos-enhanced synchronized bat optimizer[J]. Applied Mathematical Modelling,2019,77.

[28] MAFARJA M M,MIRJALILI S. Hybrid Whale Optimization Algorithm with simulated annealing for feature selection[J]. Neurocomputing,2017,260:302-312.

[29] KENNEDY J,EBERHART R C. Particle Swarm Optimization [C]// Procedings of IEEE Conference on Neural Networks. Perth:IEEE,1995:1942-1948.

[30] YANG X S. Flower Pollination Algorithm for Global Optimization [C]// International Conference on Unconventional Computing and Natural Computation. Berlin:Springer,2012.

[31] MIRJALILI S,MIRJALILI S M,YANG X S. Binary bat algorithm[J]. Neural Computing & Applications,2014,25(3):663-618.

[32] MIRJALILI S,MIRJALILI S M,HATAMLOU A. Multi-Verse Optimizer:a nature-inspired algorithm for global optimization [J]. Neural Computing and Applications,2015,27(2):495-513.

[33] SOUZA R C T D,COELHO L D S,MACEDO C A D,et al. A V-Shaped Binary Crow Search Algorithm for Feature Selection [C]// 2018 IEEE Congress on Evolutionary Computation(CEC). IEEE,2018.

[34] HUSSIEN A G,HASSANIEN A E,HOUSSEIN E H,et al. S-shaped Binary Whale Optimization Algorithm for Feature Selection[M]// Recent Trends in Signal and Image Processing. Berlin:Springer,2019:79-87.

[35] ABDEL-BASSET M,EL-SHAHAT D,EL-HENAWY I,et al. A new fusion of grey wolf optimizer algorithm with a two-phase mutation for feature selection[J]. Expert Systems with Applications,2020,139:112824.

ZHANG Ya-chuan, born in 1996, postgraduate. Her main research interests include swarm intelligence optimization and data mining.

KANG Yan, born in 1972, Ph.D, associate professor. Her main research interests include transfer learning, deep learning and integrated learning.