

中文篇章级句间语义关系识别*

张牧宇¹, 宋原¹, 秦兵¹, 刘挺¹

(1. 哈尔滨工业大学, 黑龙江省 哈尔滨市 150001)

摘要: 篇章句间关系识别 (Discourse Relation Recognition) 是篇章分析的重要内容, 本文对中文篇章句间关系识别任务进行初步探索, 包括显式篇章句间关系识别与隐式篇章句间关系识别两类任务。针对显式篇章句间关系, 我们提出基于关联词规则的方法进行识别, 取得了很好的效果; 针对隐式篇章句间关系, 我们抽取词汇、句法、语义等特征, 采用有指导模型进行识别。本文的分析和实验结果为后续研究提供了参考和基本对照系统。

关键词: 中文篇章级语义分析; 显式篇章句间关系; 隐式篇章句间关系

中图分类号: TP391

文献标识码: A

Chinese Discourse Relation Recognition

Muyu Zhang¹, Yuan Song¹, Bing Qin¹, Ting Liu¹

(1. Harbin Institute of Technology, Harbin, Heilongjiang 150001, China)

Abstract: Discourse Relation Recognition is one important part of discourse analysis. This paper focused on Chinese discourse relation recognition, including explicit discourse relation recognition and implicit discourse relation recognition. For explicit discourse relation recognition, we proposed a statistical method based on discourse connectives rules which got rather good results. For implicit discourse relation recognition, we combined lexical, syntactic and semantic features in a supervised model to classify implicit relations. The detail analysis and experiment results are useful and provide a baseline for future work on this task.

Key words: Chinese discourse semantic analysis; explicit discourse relation recognition; implicit discourse relation recognition

1 引言

随着词汇语义、句子语义研究的逐渐成熟, 篇章级语义分析逐渐成为研究热点。作为篇章语义分析的重要内容, 篇章句间关系识别 (Discourse Relation Recognition) 也受到了越来越多的关注。该研究检测同一篇章内, 两个文本单元 (片段、分句、复句、句群、段落等) 之间的逻辑语义关联 (例如: 因果关系)。通过定义层次化的语义关系类型体系将句内的语义分析结果扩展为篇章级的语义信息, 从而成为语义分析的重要解决途径之一, 对自动文摘^[1]、自动问答^[2]、倾向性分析^[3-4]以及文本质量评价^[5]、文本连贯性评价^[6]等许多 NLP 任务起到了很大的帮助。

根据文本单元间是否存在篇章连接词 (也称作篇章关联词), 可将篇章句间关系分为显式篇章句间关系 (Explicit Discourse Relation, 简称显式关系) 与隐式篇章句间关系 (Implicit Discourse Relation, 简称隐式关系) 两类。其中显式关系包含篇章关联词, 如例 1 所示, 篇章关联词“因为”指示因果类型的关系实例; 隐式关系缺少显式关联词, 需要根据上下文推测语义类型, 如例 2 所示。

* 收稿日期: 2013.08.27

定稿日期: 2013.08.27

基金项目: 国家自然科学基金重点项目 (61133012); 国家自然科学基金面上项目 (61073126); 国家自然科学基金面上项目 (61273321);

作者简介: 张牧宇 (1987—), 男, 博士研究生, 主要研究方向为篇章语义分析, 文本挖掘; 宋原 (1989—), 女, 硕士研究生, 主要研究方向为篇章关系分析, 文本挖掘; 秦兵 (1968—), 女, 教授, 主要研究方向为自然语言处理、信息检索、信息抽取等; 刘挺 (1972—), 男, 教授, 主要研究方向为自然语言处理、文本挖掘、文本检索等。

例1: **因为** 我是你爸爸, 我愿意为你做所有一切。(显式因果关系)

例2: 他生病了, 今天没有来上课。(隐式因果关系)

已有篇章句间关系识别研究主要针对英文^[7], 印度语^[8]、土耳其语^[9]和阿拉伯语^[10]。虽然已有一些面向中文的研究^[11-13], 但主要集中在分析和语料标注, 对关系识别研究不足; 另外, 已有研究大都直接使用了英文关系类型体系, 忽略了中文本身的特点。

本文对中文篇章句间关系识别进行了探索, 包括显式关系识别和隐式关系识别两方面:

针对显式关系识别, 我们提出一种基于关联词的识别方案, 通过分析中文篇章句间关系语料获得关联词对关系类型的指示能力, 并根据关联词指示规则决定显式关系的语义关系类型。针对隐式关系识别, 由于缺少篇章关联词, 我们主要采用机器学习方法, 抽取词汇、句法和语义等特征训练分类模型, 根据模型输出判定最终的关系类型。以上识别研究均采用面向中文的篇章句间关系体系, 更好的适应中文特点。

实验结果显示, 基于关联词的显式关系识别方法取得了非常好的效果, 取得了 90%左右的识别准确率, F 值达到 80%; 此外, 我们的隐式关系识别方法也取得了较好的效果。文章余下内容组织如下: 第二部分介绍相关工作; 第三部分介绍显式关系识别方法, 给出实验结果与分析; 第四部分介绍隐式关系特征、识别方法及实验结果; 第五部分给出结论。

2 相关工作

篇章句间关系体系及语料: 作为有指导方法的基础, 英文中已经出现一些篇章句间关系语料^[14-16]。这些语料采用不同的关系类型体系^[14-17]描述文本单元之间的语义关系。典型的篇章句间关系语料包括以下两种: 基于 RST 理论^[17]的修辞结构理论树库 (Rhetorical Structure Theory Discourse Treebank)^[15]和基于 PDTB 体系的宾州篇章树库 (Penn Discourse Tree Bank)^[16], 它们采用不同的关系类型体系和标注标准^[18]。目前已有的语料和标注理论关注英语、印度语^[8]、土耳其语^[9]和阿拉伯语^[10]。Xue^[11]、Zhou and Xue^[12]、Huang and Chen^[13]在中文上做了部分分析工作, 不过这些研究直接将英文关系类型体系平移到中文, 忽略了中文本身的特点。本文采用了 Zhang^[19]在 2012 提出的面向中文的篇章句间关系类型体系, 更好的适应中文问题。

显式篇章句间关系识别: 显式篇章句间关系通常由篇章关联词作为指示, Pitler^[7]使用无指导方法, 仅仅利用关联词的统计特征识别显式篇章句间关系类型, 取得了较好的效果, 证明关联词对显式关系识别的重要性。除无指导方法之外, 有指导模型也被用于显式关系识别, Pitler^[20]使用关联词相关的标准句法特征帮助提高显式关系识别性能; Wellner and Pustejovsky^[21]采用有指导方法识别篇章句间关系元素范围; Elwell and Baldrige^[22]使用关联词排序器识别关系元素范围。本文提出基于中文关联词统计信息的识别方案探索显式关系识别, 并且取得比较好的效果。

隐式篇章句间关系识别: 隐式篇章句间关系通常存在于相邻句子之间, 同时缺少关联词。类似于显式关系识别, 隐式关系识别的相关研究最早出现在英文中, 主要关注词汇特征, 例如词汇之间的依存关系^[23-24]、词汇的语义类别^[20]和关联词预测^[25]。

由于隐式关系识别不同于显式关系^[26], 除了词汇特征之外, 一些额外信息被逐渐引入, 例如句法限制^{[27][20]}、核函数^[28]、实体特征^[29]以及事件配对特征^[30]。这些研究提高了隐式关系识别效果, 但到目前为止, 隐式关系识别效果依然不佳, 而且缺少面向中文的隐式关系识别研究。本文提出基于中文篇章句间关系体系的隐式关系识别模型, 通过引入词汇、句法和语义特征识别隐式篇章句间关系。

3 显式篇章句间关系识别

在显式关系识别中, 关联词通常作为关系类型的指示标志出现。本文提出基于关联词的中

文显式关系识别模型，利用关联词规则识别显式关系。

3.1 基于关联词的识别方案

显式关系的具体类型通常由关联词标识，如下面两例：

例3：如果大家都同意这个方案，咱们就按照它来执行；（条件关系）

例4：因为大家都同意这个方案，咱们就按照它来执行；（因果关系）

例3、例4中，除关联词外的句子成分完全一致，但不同的关联词使得两个句子具有不同的语义和关系类型。可以推测：关联词标识具体关系类型。基于这种想法，我们提出了基于关联词的显式篇章句间关系识别方案。据我们了解，这是首个利用中文篇章关联词识别显式关系类型的研究工作。

3.1.1 识别方案

我们将中文篇章句间关系语料分为两部分：Set 1 包含 996 篇文本，用于抽取篇章关联词和对应的关系类型；Set 2 包含 100 篇文本，用于测试识别方案。首先，我们从 Set 1 中抽取所有的篇章关联词和相应的关系类型；之后采用极大似然估计计算关联词对各关系类型的指示能力，获得“关联词-关系类型”矩阵：其中横轴对应某一篇章关联词，纵轴对应某一具体关系类型。具体的计算方法如公式 1 所示：

$$P(c_i, s_j) = \frac{Num(c_i, s_j)}{\sum_{s_j \in S} Num(c_i, s_j)} \quad (1)$$

其中 c_i 对应某一关联词； s_j 表示待计算的关系类型； S 是所有关系类型的集合。

对 Set 2 中的每一个测试实例，我们首先抽取篇章关联词；随后查找“关联词-关系类型”矩阵，获得该关联词对各关系类型的指示能力，从中选取最大值；并将该类型作为测试实例的最终标签。

3.2 实验设置

3.2.1 类型体系及语料获取

为了支持关联词分析和后续的有指导识别方法，我们采用 Zhang^[19]提出的中文篇章句间关系体系，并从 OntoNotes 4.0^[31]中随机筛选出 1096 篇文本并进行了人工标注。在这份语料中，三名标注人员独立标注了显式关系和隐式关系。为了验证标注质量，检验标注一致性，我们计算了用于统计多类、多标注人员标注一致性的 Fleiss Kappa 指标^[32]。

在最终的计算结果中，我们获得了 66.52% 的 Fleiss' Kappa 值，根据 Fleiss' Kappa 指标的性能分布区间，该数值反映了较好的标注一致性；此外，该结果包括显式关系和隐式关系在所有类别上的标注一致性，如果单独计算显式关系的标注一致性，我们会获得更好的结果。据我们所知，这是第一份中文篇章句间关系语料。

3.2.2 实验结果

训练语料中共标记出 1273 个不同的篇章关联词，利用这 1273 个关联词构成“关联词-关系类型”矩阵，并根据该矩阵对测试实例进行分类。对每一个测试实例，我们抽取相应的篇

表 1：基于关联词的显式关系识别方法实验结果

	时序关系	因果关系	条件关系	比较关系	扩展关系	并列关系
P	0.9512	0.9686	0.9890	0.9923	0.8683	0.6364
R	0.7156	0.8876	0.9045	0.8431	0.8224	0.5060
F	0.8167	0.9263	0.9449	0.9116	0.8447	0.5638

章关联词，之后检索矩阵，找到概率最大的关系类别作为最终结果。

我们在 Zhang^[19] 提出的中文篇章句间关系体系六个顶层类别进行实验，包括“时序关系”、“因果关系”、“条件关系”、“比较关系”、“扩展关系”、“并列关系”，采用标准 P、R、F 进行评价，结果如表 1 所示。

分析表 1，我们在“因果关系”、“条件关系”、“比较关系”三类获得了非常好的效果：准确率均高于 0.96，F 值均高于 0.91。效果最好的“条件关系”精确率达到 0.9890，召回率为 0.9045，F 值则是 0.9449。这意味着绝大多数情况下，“条件关系”对应的篇章关联词（例如：如果）都是无歧义的；一旦这些关联词出现，我们可以以非常高的概率将该关系实例判定为条件关系。类似的情况同样存在于“因果关系”和“比较关系”中。

“时序关系”的实验结果略有不同，我们获得了较高的准确率（0.9512），但召回率较低（0.7156）。高准确率说明“时序关系”对应的篇章关联词歧义性较小，低召回率说明统计信息的覆盖率较差。对于“扩展关系”和“并列关系”情况则比较复杂。在这两类中，准确率和召回率都相对较低，这意味着除了覆盖率问题外，两类关系对应的篇章关联词歧义性也比较高。对于歧义问题，很难单纯通过语料扩充或分析解决，需要后续工作的更多关注。

总的来说，基于关联词的识别方案在各个类别上的平均表现较好。但是，最高的 F 值（0.9449）和最低的 F 值（0.5638）之间差距较大，说明不同的关系类型之间差异非常明显，这提示我们：不同的关系类型适合不同的处理方法。

3.2.3 错误分析与讨论

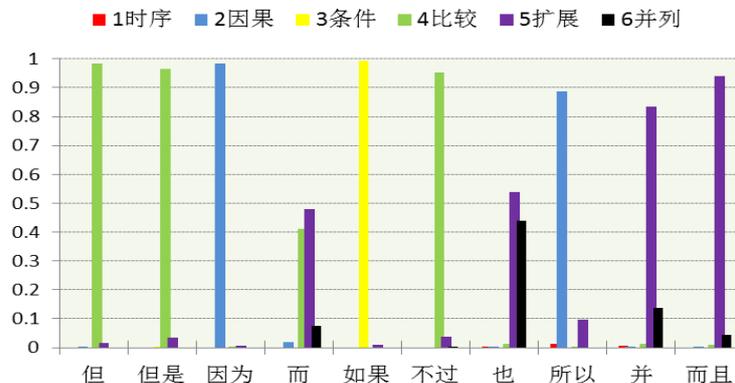
进一步分析实验结果，我们发现：大部分篇章关联词歧义较小，识别错误主要由少部分高歧义导致。这些关联词种类较少，但常用关联词较多（例如：而）。图 1 描述出现次数 Top 10 的篇章关联词在各关系类型上的分布情况：柱状图的不同颜色代表关联词对应的关系类型，不同的高度代表对应关系类型所占的比例；同一关联词对应的关系类型越少、类型越集中，该词的歧义性越小。从图中可知，大部分关联词（例如：因为）的歧义性较小，90%以上指示同一关系类型，但同时存在部分高歧义关联词。

以关联词“而”为例，它对应的关系类型分布情况包括以下几类：

- (1) 48.6% 对应“扩展关系”；
- (2) 41.8% 对应“比较关系”
- (3) 7.6% 对应“并列关系”
- (4) 2% 对应“因果关系”

根据 3.1.1 的计算公式，“扩展关系”对应的得分最高。在分类过程中，所有由“而”标识的篇章句间关系实例都被分为“扩展关系”类别。对于 48.6% 的实例而言，我们获得了正确结果；然而对于剩余的 51.4%，则发生了分类错误。实验分析发现，大部分分类错误都和该类关联词有关。这提示我们对于歧义性大，出现次数较多的关联词，需要特殊的处理方案。

图 1: Top 10 关联词的关系类型分布情况



4 隐式篇章句间关系识别

隐式篇章句间关系缺少篇章关联词，没有明显的语义类型标志，需要人类推理才能判断关系的存在和具体类型。这使得隐式篇章句间关系具有不同于显式关系的分布特点。

4.1 隐式关系分析

在很多情况下，关联词不仅仅起衔接作用，还会影响关系类型的分布，如下例：

例 5: **如果**你身体还没恢复，**就**先不用来上班了。（条件关系）

例 6: 你身体还没恢复，先不用来上班了。（因果关系）

例 5 首先描述某一假设条件，随后说明假设成立时的结果，属于“条件关系”；例 6 首先描述某一事实，随后指出事实引发的结果，属于“因果关系”。除关联词“如果…就…”之外两个例句内容完全相同，但却具有完全不同的语义类型。换言之，对某些关系类型来说（例如：条件关系），如果删除篇章句间关系关联词，句子语义会发生翻转。这种现象使得对应类型的隐式关系实例大大减少，形成和显式关系完全不同的分布特征。图 2 描述了隐式关系和显式关系在中文篇章句间关系体系^[19]中六个顶层类别上的分布情况，其中图 2-1 为显式关系分布图，图 2-2 为隐式关系分布图。

分析图 2 可知：相较于显式关系，隐式关系的分布非常不均衡，其中“扩展关系”的比例大大增加，占到了总数的 60.37%；而“条件关系”、“时序关系”、“比较关系”的数量则大大压缩，其中“条件关系”和“时序关系”分别只占 0.72%和 2.57%；只有“并列关系”和“因果关系”比例相对稳定。

分析原因，对“条件关系”和“时序关系”而言，由于关联词的省略导致了语义翻转，使得对应类型很少出现在隐式关系中；而“扩展关系”则非常适合用隐式关系来表达，这导致了图 2 中分布现象的出现。该特点提示我们，在隐式关系识别中，不同关系类型具有不同的分布特性，适合不同的识别方法。考虑到隐式关系中“条件关系”和“时序关系”数量极少，我们主要识别“扩展关系”、“因果关系”、“比较关系”、“并列关系”四类。

4.2 基于有指导方法的隐式关系识别模型

根据以上的分析，对隐式关系识别主要集中在“扩展关系”、“因果关系”、“比较关系”、“并列关系”四类。我们抽取了词汇、句法、语义等多层次的特征，采用最大熵和 SVM 两类学习方法训练四元分类模型，根据模型输出判定隐式篇章句间关系类型。

4.2.1 特征集合

核心动词：作为句子的主要成分，动词往往在语义表达中起很重要的作用，动词之间的关系常常反映了句子间的语义关系。如下例：

图 2：显式/隐式关系类型分布图

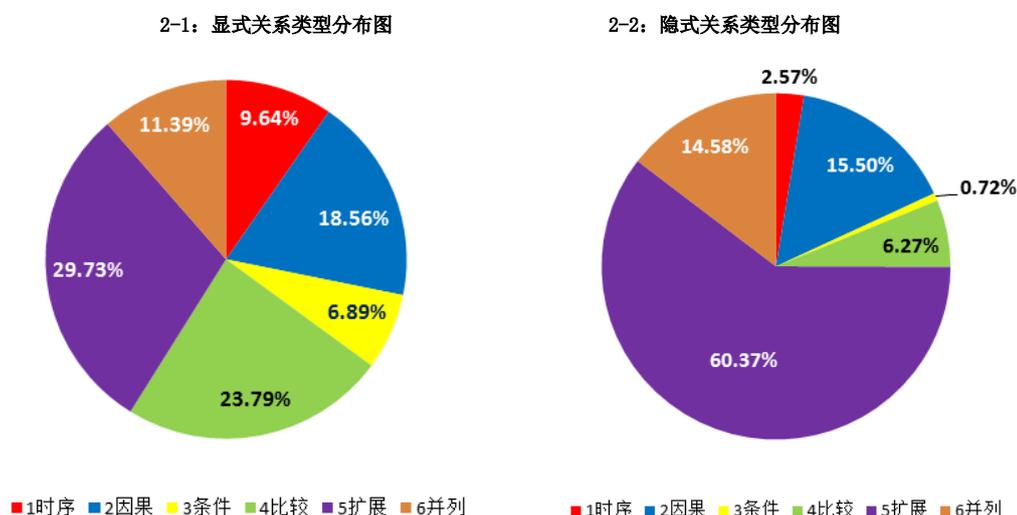


表 2: 隐式篇章句间关系识别结果

学习算法	评价指标	因果关系	比较关系	扩展关系	并列关系
最大熵	P	0.6875	0.6000	0.5849	0.3231
	R	0.0803	0.0390	0.9382	0.3182
	F	0.1438	0.0732	0.7206	0.3206
SVM	P	0.4627	0.7778	0.6324	0.3365
	R	0.2263	0.0909	0.8455	0.5303
	F	0.3039	0.1628	0.7236	0.4117

例7: 塔利班10日晚[袭击]了阿富汗北部一个村落, 导致18人[丧生]。(因果关系)

上例中, “袭击 - 丧生”之间存在因果联系, 同时也指示了两个分句之间的因果关系。通过挖掘动词之间的搭配特性, 有助于识别篇章句间关系类型。这两词在依存句法分析结果中均作为“SBV(主谓关系)”的谓语动词出现, 因此我们利用依存句法分析找到前后分句中的“SBV”关系, 抽取其中的谓语动词; 同时为了避免稀疏, 我们将两个谓语动词在同义词词林中泛化至第三层, 并将泛化结果配对构成核心动词特征。

极性特征: 不同的极性信息常常指示特定的篇章句间关系类型, 如下例:

例8: 他很[喜欢]苹果公司的产品, 遗憾的是[价格太高]了。(转折关系)

上例中, “喜欢”指示“Positive”的极性信息; “价格太高”指示“Negative”的极性信息, 前后分句的极性信息相反, 指示该实例属于“转折关系”。基于这种现象, 我们引入了篇章单元的极性特征, 采用极性词匹配的方法判定篇章单元极性, 并作为特征使用。

依存句法特征: 篇章单元的句法结构中, 最核心的关系包括“SBV(主谓)”和“VOB(动宾)”两类, 它们描述了文本单元的主要信息。本文将两个篇章单元中的“SBV”和“VOB”关系抽取出来, 并将对应词汇在同义词词林中泛化至第三层, 作为特征使用。

Unigram(句首): 在中文里, 句首词语通常起承上启下的作用, 对篇章句间关系类型具有一定的指示作用。本文分别抽取两个篇章单元中的第一个词, 作为识别特征使用。

Bigram(句首): 中文里承上启下的可以是单个词, 也可以是双词或短语。因此除 Unigram 特征之外, 本文还抽取两个篇章单元中的前两个词, 作为识别特征使用。

4. 3 实验结果

我们仍然采用 3.2.1 中提到的中文篇章句间关系语料库进行实验, 该语料库包含 1096 篇文本, 手工标注了显式篇章句间关系和隐式篇章句间关系两类信息。我们将其中 996 篇作为训练语料, 另外 100 篇作为测试语料, 抽取前文提出的词汇、句法、语义等特征, 分别训练最大熵和 SVM 两种模型进行分类。我们在中文篇章句间关系体系的四个顶层类别进行分类, 包括: “扩展关系”、“因果关系”、“比较关系”、“并列关系”, 结果如表 2 所示。

分析表2, 除“扩展关系”外, 其他类型存在高准确率、低召回率的特性。以最大熵模型下的“因果关系”为例, 识别准确率达到0.6875, 召回率却只有0.0803。而“扩展关系”情况恰恰相反。这说明数据不均衡性使得模型倾向于将测试实例分为“扩展关系”, 导致“扩展关系”类型召回率增加, 准确率下降; 同时使得其他类型召回率降低。系统的整体性能不佳, 很大一个原因是由于低召回率导致的。这提示我们在类别分布严重不均衡的情况下, 传统的统一识别思路存在很大的困难。

注意到 SVM 实验结果普遍高于最大熵, 这主要是由于隐式关系在各类型上分布不均衡, 而 SVM 模型对边界实例敏感, 但对数据不平衡有较强的容忍度, 因此取得了相对较好的效果。此外, 对比前文的显式关系识别结果, 我们发现“并列关系”识别效果始终不佳, 这在一定程度上反映该类型的特征不明显, 区分度较弱; 同时考虑图 1, 主要的关联词歧义集中在“扩

展关系”和“并列关系”，说明这两个类别特征接近。从语义体系定义上来说，是否有必要将“扩展关系”和“并列关系”区分开来，是值得考虑的一个问题。

5 结论与展望

本文首次探索面向中文的篇章句间关系识别任务，尝试了显式篇章句间关系识别和隐式篇章句间关系识别两方面研究。对于显式篇章句间关系识别，我们首次提出基于篇章关联词的显式关系识别方法，在关联词统计的基础上识别关系类型，取得了非常好的效果。对于隐式篇章句间关系识别，我们首先分析了隐式关系和显式关系在类型分布上的差别，指出隐式关系的特点，并在识别过程中进行了针对性处理；随后我们提出词汇、句法、语法等一系列特征，采用最大熵和SVM两种方案尝试了隐式篇章句间关系识别。本文的分析和实验结果为后续的工作提供了参考，推动了中文篇章分析研究，尤其是篇章句间关系分析的进一步发展。

6 致谢

本文获得国家自然科学基金重点项目（61133012）、国家自然科学基金面上项目（61073126）、以及国家自然科学基金面上项目（61273321）资助，特此感谢。

参考文献

- [1] D. Marcu. 2000. The rhetorical parsing of unrestricted texts: A surface-based approach [J]. *Computational Linguistics*, 26(3): 395 - 448.
- [2] R. Girju. 2003. Automatic detection of causal relations for Question Answering [C]. *Proceedings of the ACL 2003 workshop on Multilingual summarization and question answering*-Volume 12: 76 - 83.
- [3] S. Somasundaran, J. Wiebe, and J. Ruppenhofer. 2008. Discourse-level opinion interpretation [C]. *Proc. of Coling 2008*.
- [4] Zhou, L., Li, B., Gao, W., Wei, Z., Wong, K.F. 2011. Unsupervised Discovery of Discourse Relations for Eliminating Intra-sentence Polarity Ambiguities [C]. *Proceedings of the EMNLP 2011(Oral presentation)*, Edinburgh, Scotland, July: P27-31.
- [5] E. Pitler and A. Nenkova. 2008. Revisiting readability: A unified framework for predicting text quality [C]. *Proc. of EMNLP 2008*, pages 186 - 195.
- [6] Ziheng Lin, Hwee Tou NG, and Min-Yen Kan. 2011. Automatically Evaluating Text Coherence Using Discourse Relations [C]. s. *Proceedings of ACL-HLT*, pages 997 - 1006.
- [7] E. Pitler, M. Raghupathy, H. Mehta, A. Nenkova, A. Lee, and A. Joshi. 2008. Easily identifiable discourse relations [C]. *Proc. of COLING08*, short paper.
- [8] Rashmi Prasad, Samar Husain, Dipti Sharma, and Aravind Joshi. 2008b. Towards an annotated corpus of discourse relations in Hindi [C]. *Proc. of the IJCNLP 2008*, Hyderabad, India.
- [9] Deniz Zeyrek and Bonnie Webber. 2008. A Discourse Resource for Turkish: Annotating Discourse Connectives in theMETU Corpus [C]. *Proceedings of IJCNLP-2008*. Hyderabad, India.
- [10] A. AlSaif and K. Markert. 2010. The leeds arabic discourse treebank: Annotating discourse connectives for arabic [C]. *Proc. Of LREC 2010*.
- [11] Xue, Nianwen. 2005. Annotating discourse connectives in the Chinese Treebank [C]. *Proc. of The ACL Workshop in Frontiers in Annotation II: Pie in the Sky*. Ann Arbor, Michigan: ACL, .

- [12] Hen-Hsen Huang, Hsin-Hsi Chen. 2011. Chinese Discourse Relation Recognition [C]. *Proceedings of IJCNLP 2011*, pages 1442 - 1446.
- [13] Yuping Zhou, Nianwen Xue. PDTB-style Discourse Annotation of Chinese Text [C]. *Proc. Of ACL 2012*.
- [14] Hobbs, J. R. (1985). On the coherence and structure of discourse [M]. *CSLI*, page 85-37.
- [15] Carlson, L., Marcu, D., & Okurowski, M. E. (2003). Building a discourse-tagged corpus in the framework of rhetorical structure theory [M]. *Springer Netherlands*, page 85-112.
- [16] R. Prasad, N. Dinesh, A. Lee, E. Miltsakaki, L. Robaldo, A. Joshi, and B. Webber. 2008a. The Penn discourse treebank 2.0 [C].
- [17] William Mann and Sandra Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organization [J]. *Text*, 8(3):243 - 281.
- [18] A. AlSaif and K. Markert. 2010. The leeds arabic discourse treebank: Annotating discourse connectives for Arabic [C]. *Proc. Of LREC 2010*.
- [19] 张牧宇, 秦兵, 刘挺. 中文篇章级句间语义关系体系及标注 [C]. *Pro. Of CCIR 2012*.
- [20] Pitler E., Louis A. and Nenkova A. 2009. Automatic Sense Predication for Implicit Discourse Relations in Text [C]. *Proc. of ACL-IJCNLP 2009*.
- [21] Ben Wellner and James Pustejovsky. 2007. Automatically identifying the arguments of discourse connectives [C]. *Proc. of EMNLP-CoNLL 2007*, Prague, Czech Republic.
- [22] R. Elwell and J. Baldridge. 2008. Discourse connective argument identification with connective specific rankers [C]. *Proc. of the International Conference on Semantic Computing*.
- [23] D. Marcu and A. Echiabi. 2001. An unsupervised approach to recognizing discourse relations [C]. *Proc. Of ACL 2001*, pages 368 - 375.
- [24] S. Blair-Goldensohn, K. R. McKeown, and O. C. Rambow. 2007. Building and Refining Rhetorical-Semantic Relation Models [C]. *Proceedings of NAACL HLT*, pages 428 - 435.
- [25] Z. Zhou, Y. Xu, Z. Niu, M. Lan, . Su, and Tan. C. 2010. Predicting discourse connectives for implicit discourse relation recognition [C]. *Proc. of Coling 2010*, pages 1507 - 1514.
- [26] C. Sporleder and A. Lascarides. 2008. Using automatically labelled examples to classify rhetorical relations: an assessment [J]. *NLE 2008*, Volume 14(3).
- [27] Lin Z., Kan M. and Ng H. 2009. Recognizing Implicit Discourse Relations in the Penn Discourse Treebank [C]. *Proc. of EMNLP 2009*, Singapore, August.
- [28] W. Wang, J. Su, and C. Tan. 2010. Kernel-based discourse relation recognition with temporal ordering information [C]. *Proc. of ACL 2010*, pages 710 - 719.
- [29] A. Louis and A. Nenkova. 2010. Creating local coherence: An empirical assessment [C]. *Proc. of NAACL 2010*.
- [30] C. Chiarcos. Towards the Unsupervised Acquisition of Discourse Relations [C]. *Proc. Of ACL 2012*.

- [31] Eduard Hovy, Mitchell Marcus, Martha Palmer, Lance Ramshaw, and Ralph Weischedel. 2006. Ontonotes: The 90% solution [C]. *In Proceedings of the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers*, pages 57 - 60.
- [32] Fleiss, J. L. 1971. Measuring nominal scale agreement among many raters [J]. *Psychological Bulletin*, Vol. 76, No. 5 pp. 378 - 382

第一作者

姓名: 张牧宇 邮编: 150001

电话: +86 13796055168 电子邮箱: myzhang@ir.hit.edu.cn

地址: 哈尔滨市南岗区教化街 29 号六楼 社会计算与信息检索研究中心

第二作者

姓名: 宋原 邮编: 150001

电话: 0468 86413683 电子邮箱: ysong@ir.hit.edu.cn

地址: 哈尔滨市南岗区教化街 29 号六楼 社会计算与信息检索研究中心

第三作者

姓名: 秦兵 邮编: 150001

电话: 0468 86413683 电子邮箱: qinb@ir.hit.edu.cn

地址: 哈尔滨市南岗区教化街 29 号六楼 社会计算与信息检索研究中心

第四作者:

姓名: 刘挺 邮编: 150001

电话: 0468 86413683 电子邮箱: tliu72@vip.126.com

地址: 哈尔滨市南岗区教化街 29 号六楼 社会计算与信息检索研究中心