基于词汇链强化表征的篇章修辞结构分析研究

王金锋
苏州大学
计算机科学与技术学院
jfwang9@stu.suda.edu.cn

孔芳*
苏州大学
计算机科学与技术学院
kongfang@suda.edu.cn

摘要

篇章分析作为自然语言处理领域的基础问题一直广受关注。由于语料规模有限，绝大多数已有研究仍依赖于外部特征的加入。针对该问题，本文提出了提出一种通用的表征增强方法，借助图卷积神经网络将词汇链信息融入到基本篇章单元的表征中。在RST-DT和CDTB上的实验证明，本文提出的表征增强方法能够提升多种篇章解析器的性能。

关键词： 篇章分析；词汇链；图卷积神经网络

Lexical Chain Based Strengthened Representation for Discourse Rhetorical Structure Parsing

Jinfeng Wang
School of Computer
Science and Technology
Soochow University
jfwang9@stu.suda.edu.cn

Fang Kong *
School of Computer
Science and Technology
Soochow University
kongfang@suda.edu.cn

Abstract

As a basic problem in the field of Natural Language Processing, Discourse Parsing has been drawing more and more attention in recent years. Due to the limited scale of the corpus, existing studies still rely on handcraft features. To deal with this problem, we proposed a general representation enhancement method, which integrates lexical chain information into the EDU representation, with the help of Graph Convolutional Neural Network. Experiments on RST-DT and CDTB prove that the representation enhancement method can improve the performance of multiple discourse parsers on different corpus.

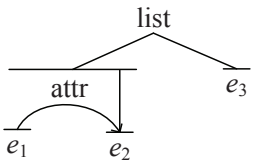
Keywords: Discourse Parsing , Lexical Chain , GCN

1 引言

篇章修辞结构分析任务旨在将一个篇章解析为一棵篇章树。如图1所示，在该示例中，篇章树的每个叶子节点对应一个基本篇章单元(EDU，Elementary Discourse Unit)，篇章单元之

基金项目：国家自然科学基金面上项目（61876118）面向篇章信息性的汉语篇章结构多层次联合分析研究；
*通信作者：kongfang@suda.edu.cn

第二十届中国计算语言学大会论文集，第467页-第476页，呼和浩特，中国，2021年8月13日至15日。



e_1: They don't think
e_2: it is contemplating a takeover, however,
e_3: and its officials couldn't be reached.

图 1: 篇章树示例

间以递归地方式，通过修辞关系构建成更大的篇章单元，每一个关系节点都会标注其两个子节点之间的核心卫星关系和修辞关系。篇章分析在许多下游任务中都有着重要的作用，例如开放领域对话系统(Ma et al., 2019)、情感分析(Nejat et al., 2017)和机器翻译(Joty et al., 2017; Sennrich, 2018)等。

篇章分析的早期研究主要致力于挖掘有效的人工特征(Feng and Hirst, 2014; Heilman and Sagae, 2015; Hernault et al., 2010; Joty et al., 2015)。近年来，随着神经网络在自然语言处理领域取得令人瞩目的成绩，许多基于神经网络的篇章解析器也被提出(Ji and Smith, 2017; Li et al., 2016; Yu et al., 2018; Zhang et al., 2020)。受限于语料规模，现有神经网络篇章解析器仍未能达到理想的性能。目前在篇章解析上的工作都致力于提出更为合理的解析方式，如自顶向下的解析方式，却忽略了在现有解析方式的基础上，加入篇章衔接信息来提高篇章解析的性能。前人的工作证明，在篇章结构解析的过程中融入篇章衔接信息能有效提升解析的效果。Joty et al. (2013)在进行句子间修辞关系检测时，加入了词汇衔接(词汇链) 特征，这种简单的特征能够有效提升句间关系识别的准确率。

篇章衔接主要有以下五个方面：省略、替代、指代、连接和词汇衔接。词汇衔接反映了词汇之间的语义关系，能够表示篇章中主题的变换(Morris and Hirst, 1991)。在本文中，我们使用词汇链来表示词汇衔接。本文认为位于同一词汇链上的EDU属于相同主题，基于此，提出了一种使用图卷积神经网络(Kipf and Welling, 2016)(Graph Convolutional Network, GCN)将词汇链信息融入EDU表征的方法。借助词汇链进行同主题下EDU表征之间的信息关联。该方法能够有效地将词汇链信息融入表征之中，丰富EDU表征包含的语义信息，提升EDU表征的质量。

实验结果证明，本文提出的基于词汇链的表征增强方法，在中英文数据集RST-DT和CDTB上，均能有效提升各类篇章解析器的性能。本文的组织结构如下：(1)第二节介绍了篇章修辞结构分析的相关工作；(2)第三节详细介绍了本文提出的基于GCN的表征增强模型；(3)第四节首先介绍实验数据集、评价指标和实验参数，然后对实验结果进行详细深入的分析；(4)第五章为总结和展望。

2 相关工作

早期的篇章结构分析工作主要致力于构建有效的特征来进行篇章树的构建(Hernault et al., 2010; Joty et al., 2013; Joty et al., 2015; Ji and Eisenstein, 2014; Feng and Hirst, 2014)。近年来，神经网络模型在各大自然语言处理任务中表现突出，在此基础上，篇章结构分析领域也有许多基于神经网络的篇章解析器被提出。

中文和英文的篇章结构分析工作主要在CDTB和RST-DT两个数据集上开展。现有的篇章解析器依据解析策略可以分为两种：自底向上 (Joty et al., 2013; Ji and Eisenstein, 2014; Li et al., 2014a; Li et al., 2016; Yu et al., 2018; Zhang et al., 2019; Kong and Zhou, 2017; Sun and Kong, 2018)和自顶向下 (Lin et al., 2019; Kobayashi et al., 2020; Zhang et al., 2020)。

自底向上的解析方式通过递归地合并相邻篇章单元来构建篇章树，并同时在合并的过程中识别孩子节点之间的修辞关系和核心卫星关系。Joty et al. (2013)使用CKY算法通过对句内和句间分别构建树结构的方式来获取最优的篇章树；Ji and Eisenstein (2014)提出了一种基于转移的篇章解析器，使用前馈神经网络学习分布式表征。Li et al. (2016)提出了一个CKY解析器，通过层次神经网络来学习篇章单元的表征，同时提出了一种基于张量的变换函数来进行篇章单元之间的信息关联。Yu et al. (2018)提出了一种基于转移的篇章解析器，并且使用了隐式的句法信息，同时还加入了句子和段落边界的特征。Zhang et al. (2019)提出了一种使用多种信息流

增强中间节点表征的方法，利用基于转移的方式进行篇章树的构建。Kong and Zhou (2017)构建了一个端到端的中文篇章解析器。Sun and Kong (2018)使用SPINN(Goyal and Eisenstein, 2016)（Stack-augmented Parser-Interpreter Neural Network）模型进行基于转移的篇章解析方法在汉语CDTB上的率先尝试，它从左到右扫描EDU序列，通过预测一系列的转移动作，完成树形结构的构建。

自顶向下的解析方式通过递归地查找中间节点的分割点的方式构建篇章树，并同时识别分割得到的孩子节点之间的修辞关系和核心卫星关系。Lin et al. (2019)首先提出了一个句子级的自顶向下的篇章解析器。Kobayashi et al. (2020)提出了一个基于多粒度（篇章、段落、句子）的自顶向下的篇章解析模型，同时加入了句子和段落边界特征。Zhang et al. (2020)提出了一个基于分割点优先级的中英文统一篇章解析框架，在不使用任何特征的情况下，取得了与基于特征的模型相近的性能。

随着深度学习的发展，尽管许多基于神经网络的篇章解析器被提出，但目前最好的篇章解析性能仍依赖于人工构建的特征。为解决这个问题，本文根据自动抽取的词汇链，构建EDU之间的关联矩阵，利用图卷积神经网络将词汇链信息融入到EDU表征得到强化表征，使用强化后的EDU表征参与篇章树的解析。

3 模型介绍

图2给出了本文提出的基于GCN的表征强化模型(**GREM**, **G**CN based **R**epresentation **E**nhancement **M**odel)，模型主要由两部分组成：(1)词汇链的自动获取；(2)基于词汇链的表征加强和融入。

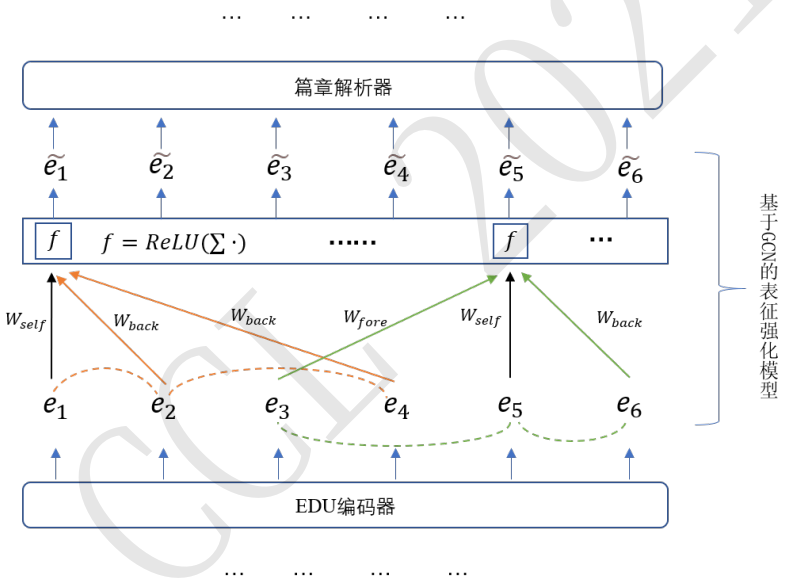


图 2: 基于GCN的表征强化模型

3.1 词汇链的自动获取

Joty et al. (2013)将词汇链作为特征拼接到EDU表征中，验证了词汇链能够有效提升篇章结构解析的效果。我们使用多条词汇链表示EDU之间的关系，并依据词汇链构建EDU之间的关联矩阵，利用GCN聚合同主题下EDU的表征，从而达到增强EDU表征的效果。

本文抽取的词汇链是EDU级别的，当两个EDU中的词汇相似度大于阈值时，则认为它们之间存在词汇链。我们通过计算词汇间的相似度来衡量它们之间的语义关系。在中文数据集CDTB上，基于中文同义词库——**Cilin**计算两个词之间的相似度；在英文数据集RST-DT上，则使用**WordNet**作为同义词库，计算两个词之间的相似度。

表1给出了具体的词汇链抽取过程。该方法能够判断任意两个EDU之间是否存在词汇链。本文抽取的词汇链仅由名词组成，包含普通名词(NN, NNS)和专属名词(NNP,NNPS)。首先，

对给定的EDU进行预处理，去除其中的非名词词汇。其次，对于给定的两个EDU_I和EDU_J，遍历EDU_I中的每个词w_{Ii}，计算w_{Ii}和EDU_J中的每个词的相似度，如果两个词的相似度高于预先设定的阈值$Thre$，则认为两个EDU之间存在词汇链，并将(I, J)添加到词汇链的集合中。另外，作为主题信息的一种体现，我们认为词汇链通常存在于在距离较近的基本篇章单元之间，所以在构建词汇链时对距离添加了限制$Dist$，当两个EDU的距离大于$Dist$之后，则不再判断EDU之间是否存在词汇链。

输入: EDU-Set=$\{EDU_1, EDU_2, \cdots, EDU_n\}$; EDU_I=$(w_{I1}, w_{I2}, \cdots, w_{In})$
初始化: Thre=, Dist=5, LC-set= {}
预处理:
去除掉EDU中的非名词
for w_i in EDU
  if w_i is not in (NN, NNS, NNP, NNPS):
    remove w_i from EDU
判断是否存在词汇链:
for w_{Ii} in EDU_I:
  for w_{Jj} in EDU_J:
    if sim(w_{Ii}, w_{Jj}) $\geq$ Thre and J-I < Dist:
      LC-Set += (J, I)
返回: LC-Set

表 1: 词汇链的自动抽取流程

3.2 基于词汇链的表征加强和融入

本文基于抽取到的词汇链，进行同主题下的EDU之间的信息关联。具体的，使用词汇链构建EDU之间的关联矩阵，对于任意两个基本篇章单元e_i和e_j，按照其在篇章中出现的先后次序，在关系图中存在以下四种关系类型:

- 前向相关($M(e_i, e_j) = fore, i < j$)，表示当前EDU与位于其前方的EDU之间存在关联关系;
- 自我相关($M(e_i, e_j) = self, i = j$)，为了在信息关联的过程中保留当前EDU的信息，添加一条指向当前EDU的边;
- 后向相关($M(e_i, e_j) = back, i > j$)，表示当前EDU与位于其后方的EDU之间存在关联关系;
- 无关($M(e_i, e_j) = none, i < j$)，表示两个EDU之间不存在关联关系。

	e_1	e_2	e_3	e_4	e_5	e_6
e_1	*self*	*fore*	*none*	*fore*	*none*	*none*
e_2	*back*	*self*	*none*	*fore*	*none*	*none*
e_3	*none*	*none*	*self*	*none*	*fore*	*fore*
e_4	*back*	*back*	*none*	*self*	*None*	*none*
e_5	*none*	*none*	*back*	*none*	*self*	*fore*
e_6	*none*	*none*	*back*	*none*	*back*	*self*

图 3: 基于词汇链生成的关联矩阵

如图2所示，在$(e_1, e_2, \cdots, e_6)$之间存在两条词汇链，分别为lc_1=(e_1, e_2, e_4)和lc_2=(e_3, e_5, e_6)。首先，将关联矩阵中(i, i)位置的元素设置为自我相关($self$)关系；其次，根据词汇链lc_1和lc_2中词汇之间的相对位置关系，将关联矩阵中的$(1, 2), (1, 4), (2, 4), (3, 5), (3, 6), (5, 6)$位

置的元素设置为前向相关($fore$)关系，$(2,1),(4,1),(4,2),(5,3),(6,3),(6,5)$位置的元素设置为后向相关($back$)关系。最终生成的关联矩阵如图3所示。

$$\tilde{e} = f\Big(\sum_{e' \in E(e)} W_{M(e,e')}e' + b\Big) \tag{1}$$

本文使用GCN(图2)来将词汇链信息融入到EDU表征中。公式1表示的是单个EDU在GCN中的表征增强过程。以e的强化过程为例，其中$\tilde{e}$为强化后的表征；$E(e)$为所有与e之间存在关联关系的EDU(e')的集合；$W_{M(e,e')}$是对e'进行线性变换的权重矩阵，由e和e'之间的具体关系决定；$f(\cdot)$表示ReLU激活函数。GCN模型根据$E(e)$中的EDU与e的关系，对其进行不同的线性变换，将线性变换后的表征进行加和，再通过ReLU激活函数对加和后的表征进行非线性变化，最终得到的$\tilde{e}$即为e的强化表征。

如图2所示，在篇章树的构建过程中，首先，篇章解析模型以EDU为单位进行编码，得到EDU的初步表征$(e_1, e_2, \ldots, e_6)$。其次，将依据词汇链构建的关联矩阵(图3)及EDU的初步表征作为输入，使用GCN对基本篇章单元进行信息关联。在GCN中，将EDU作为图中节点进行表征的聚合，将词汇链信息融入到表征之中。最后，将强化后的表征（$\tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_6$）作为篇章解析器的输入，进行篇章树的构建。

4 实验设置及结果分析

4.1 实验数据集

为了验证基于GCN融合词汇链信息的表征增强方法的效果，本文分别在英文篇章分析语料库RST-DT(Carlson and Marcu, 2001)和中文语料库CDTB(Li et al., 2014b)上进行了实验。

修辞结构篇章树库(RST-DT)依据修辞结构理论，使用385篇来自《华尔街日报》的新闻报道标注而成。首先将篇章切分为最小篇章单元(EDU)；其次，使用修辞关系来描述篇章单元之间的关系，依照篇章单元的重要性分为核心(Nucleus)和卫星(Satellite)两种成分。作为核心的篇章单元表达主要信息，而作为卫星的基本篇章单元起到补充信息的作用。其中，训练集包含347个篇章，测试集包含38个篇章，从训练集中随机选择34个篇章作为验证集。

汉语连接依存篇章树库（CDTB）(Li et al., 2014b)依据连接依存理论，以段落为单位，在宾州中文树库6.0(CTB6.0)的基础上标注了500个文档，包含2336个段落。篇章结构树的中间节点为关系节点，关系节点将下层的篇章子结构组合为更上层的篇章结构，从而形成一棵完整的篇章树。其中，训练集包含1991棵篇章树；验证集包含105棵篇章树；测试集包含215棵篇章树。

4.2 评价指标

本文从四个方面对模型的性能进行评价，包括结构(S)、核性(N)、修辞关系关系(R)和整体性能(F)。对于每一个指标，均使用F1-Score作为评价标准。在进行篇章树解析之前，所有EDU已经被划分好，并且将所有非二叉篇章树按照右分支优先的方式转化为二叉树。

为了真实反映篇章解析器的性能，我们使用了更为严格的评价标准。对于英文和中文语料上提出的篇章解析器，分别使用同Morey et al. (2017)和Sun and Kong (2018)相同的评价指标进行模型性能的评估。

4.3 实验参数

本文的主要目的是对比融入词汇链对模型性能的影响，实验主要参数如表2所示，我们使用Adam(Kingma and Ba, 2014)(Adaptive Moment Estimation)来优化所有可训练参数。中文使用(Qiu et al., 2018)提出的词向量；英文上使用GloVe(Pennington et al., 2014)来作为词向量。具体参数如表2所示。

4.4 强化表征在中英篇章解析器上的效果分析

本文提出了一种使用GCN融入词汇链信息的表征增强方法，该方法根据篇章中存在的词汇链，对存在于同一词汇链的EDU进行信息关联，能够丰富EDU表征所包含的语义信息。为了验证该方法的的有效性，我们选择了RST-DT和CDTB上已有的三个工作，基于其开源的代码进

Parameters-EN	Value	Parameters-CN	Value
word embedding dim	300	word embedding dim	300
POS embedding dim	30	POS embedding dim	30
hidden dim	256	hidden dim	300
learning rate	0.0001	learning rate	0.0005
L2 penalty	1e-8	L2 penalty	1e-8
dropout	0.3	dropout	0.5
Thre	0.5	Thre	0.5
distance	5	distance	3
GCN layer	1	GCN layer	1

表 2: 超参数设置

行复现，并在原有模型的基础上加入GREM，对比GREM在汉语和英文语境下，以及在自顶向下和自底向上的解析方式下的效果差异。

本文在中文和英文上分别选择了一个基于转移的自底向上模型，同时选择了一个能够同时适用于中英文的自顶向下的模型作为基准模型:

- Sun and Kong (2018)(Sun-18-CN)提出了一个基于转移的自底向上的篇章解析器，是基于转移的解析方式在中文上的率先尝试;
- Zhang et al. (2019)(Zhang-19-EN)提出了一种利用多种信息流增强中间节点表征的方法，利用基于转移的方式进行篇章树的构建。
- Zhang et al. (2020)(Zhang-20-CN and EN)提出了一个基于分割点优先级的自顶向下的篇章结构解析器，该框架能够同时适用于中文和英文。

Model	S	N	R	F
Sun_18_CN	83.1	55.5	50.3	47.1
Sun_18_CN + GREM	**84.8**	**55.9**	**51.7**	**47.6**
Zhang_20_CN	84.4	57.0	52.6	45.6
Zhang_20_CN + GREM	**85.0**	**57.6**	**53.0**	**45.9**

表 3: 强化表征在中文篇章解析器上的实验结果

Model	S	N	R	F
Zhang_19_EN	67.1	55.3	43.8	43.6
Zhang_19_EN + GREM	**68.1**	**56.5**	**44.7**	**44.5**
Zhang_20_EN	67.2	55.5	**45.3**	44.3
Zhang_20_EN + GREM	**67.7**	**55.9**	45.2	**44.9**

表 4: 强化表征在英文篇章解析器上的实验结果

融入词汇链信息的强化表征能够有效提升基准模型在四个指标上的效果，并且在结构预测上的提升最为明显。这表明根据词汇链进行EDU之间的信息关联和表征强化，能够将丰富的篇章衔接信息融入到篇章解析的过程之中，在不使用额外特征的情况下，有效提升篇章解析模型的解析效果。

对比加入词汇链后中文和英文上的模型性能（表3和表4），可以发现：(1)相比于自顶向下模型，自底向上模型能够从强化表征中获得更大的收益；(2)相比与英文篇章解析器，中文篇章解析器上在使用强化表征之后的提升更为明显。

对比自底向上和自顶向下模型的结果，可以发现:(1)融入词汇链信息的强化表征，在基于自底向上解析方式的模型上提升效果更加明显。这是因为在自底向上的解析框架下，强化表征能够提升较低高度上解析的性能，减少错误的向上传播，进而使更高层次的解析性能也得到提升。这一点在树的平均高度较低的CDTB上的自底向上模型中表现最为明显。(2)自顶向下的解析方式，最初的决策是在较高的节点处进行决策，在逐层向上构建上层节点的表征时，强化表征中包含的信息已经流失较多，因此提升效果有限。并且由于英文的篇章树高度更高，在英文上自顶向下模型的提升幅度最小，在**R**指标上甚至出现了下降。对比强化表征在中英文上的效果，可以发现：融入词汇链信息的强化表征对中文篇章解析器的性能提升更为明显。这是因为中文篇章树以段落为单位进行标注，篇章树的高度较低，所以自底向上和自顶向下的解析方式均能从中强化的表征中取得较大的收益。相比于CDTB中以段落为标注的低矮的篇章树，RST-DT中以篇章为单位进行标注，篇章树的高度更高，所以加入强化表征后，英文篇章解析器性能提升的幅度比中文篇章解析器更小。

4.5 词汇链选择策略的效果分析

在本文提出的词汇链抽取方法中，影响词汇链质量的因素主要的有两个：词汇链中元素间的距离、词汇链的构成。本章针对这两个变量进行了多组对比试验，分析两个变量对词汇链质量的影响。

Model		S	N	R	F
Sun_18_CN	- None	83.1	55.5	50.3	47.1
	- distance_3	**84.8**	**55.9**	**51.7**	**47.6**
	- distance_5	84.2	55.4	51.3	47.3
	- ∞	83.4	55.1	50.0	46.3
Zhang_20_CN	- None	84.4	57.0	52.6	45.6
	- distance_3	**85.0**	**57.6**	**53.0**	**45.9**
	- distance_5	84.7	57.0	52.7	45.7
	- ∞	84.2	56.7	51.7	45.2
Zhang_19_EN	- None	67.1	55.3	43.8	43.6
	- distance_3	67.4	56.2	44.4	44.2
	- distance_5	**68.1**	**56.5**	**44.7**	**44.5**
	- ∞	66.6	54.9	43.5	43.7
Zhang_20_EN	- None	67.2	55.5	**45.3**	44.3
	- distance_3	67.4	55.5	45.2	44.5
	- distance_5	**67.7**	**55.9**	45.2	**44.9**
	- ∞	66.5	55.1	44.8	44.1

表 5: 长度对词汇链质量的影响

对于长度对词汇链质量的影响，我们认为当两个EDU之间的距离太远时，对EDU之间关系判断的可靠性会明显降低。因此，本文在判断EDU之间是否存在词汇链时，对EDU之间的距离做出人工限制，当两个EDU之间的距离超过限制时，不再进行EDU之间关系的判断。

如表5所示，None表示不加入融入词汇链信息的强化表征；distance-3表示使用融入词汇链信息的强化表征，且对EDU之间的距离限制为3；distance-5表示对距离的限制为5；∞表示对距离无限制。此外，为保证变量的唯一性，上述实验结果中的词汇链包含所有类型的名词(NN、NNS、NNP、NNPS)。

根据表5中的实验结果，可以得出以下结论:

- 对比每个模型结果的第一行和第四行，可以发现，判断词汇链时不限制EDU之间的距离，词汇链的质量会有明显下降，不仅不能使表征得到强化，而且还会由于融入噪音而降低表征的质量，进而使模型性能降低。

- 对比每个模型结果的第一行和第二、三行，可以发现，在判断词汇链时对EDU之间的距离设置限制能够有效提升词汇链的质量，从而使表征得到有效强化，有助于模型性能的提高。这表明合适的长度限制，能够提高词汇链的质量，更充分地发挥本文提出的融入词汇链信息的表征强化方法的作用。

- 对比中英模型的第二、三行，可以发现，在英文语料库上抽取词汇链时，5是更加合适的长度限制；而在中文语料库上抽取词汇链时，3是更加合适的长度限制。这是因为英文语料库中篇章树是以篇章为单位进行标注的，包含更多的EDU；而中文语料库中篇章树是以段落为单位，包含的EDU个数更少，平均只有4.2个EDU，过大的距离限制对包含EDU较少的篇章树来说相当于不加限制。

Model		**S**	**N**	**R**	**F**
Sun_18_CN	- 4(ALL)	**84.8**	**55.9**	**51.7**	**47.6**
	- 2(NN,NNS)	84.2	55.4	51.3	47.1
	- None	83.1	55.5	50.3	47.1
Zhang_20_CN	- 4(ALL)	**85.0**	**57.6**	**53.0**	**45.9**
	- 2(NN, NNS)	84.6	56.9	52.7	45.8
	- None	84.4	57.0	52.6	45.6
Zhang_19_EN	- 4(ALL)	**68.1**	**56.5**	**44.7**	**44.5**
	- 2(NN,NNS)	67.5	55.9	44.2	44.0
	- None	67.1	55.3	43.8	43.6
Zhang_20_EN	- 4(ALL)	**67.7**	**55.9**	45.2	**44.9**
	- 2(NN,NNS)	67.3	55.4	**45.4**	44.4
	- None	67.2	55.5	45.3	44.3

表 6: 词汇构成对词汇链质量的影响

对于词汇构成对词汇链质量的影响，如表6所示，None表示模型不使用融入词汇链信息的强化表征；4(ALL)表示模型使用融入词汇链信息的强化表征，且词汇链包含所有的名词(NN、NNS、NNP、NNPS)；2(NN、NNS)表示词汇链中的词汇仅包含普通名词。为保证变量的唯一性，在探究词汇构成对词汇链质量的影响时，对于中英文语料，在构建词汇链时，距离限制分别使用前文实验分析得到的最优距离。具体的，在中文上，距离限制为3；在英文上，距离限制为5。

根据表6中的实验结果，可以得出以下结论:

- 对比每个模型性能的第三行和第一二行，可以发现使用融入词汇链信息的强化表征能够显著提升多个模型的性能。这表明我们提出的融入词汇链信息的表征增强方法是一种通用的方法，在不加入额外特征的情况下，提高中文和英文上多个解析模型的性能。
- 对比每个模型性能的第一行和第二行，可以发现当词汇链中包含所有类型的名词时，构建的关联矩阵更加可靠，能够使表征得大程度的强化。这是由于专属名词相比于普通名词，与主题更为相关，略掉专属名词会使词汇链无法充分反映EDU之间的信息关联，因此使强化表征的质量下降，从而提升效果相对不明显。

5 总结与展望

本文提出了一种通用的融入词汇链信息的表征加强方法，能够方便地应用与多种篇章解析模型。在中英文语料上的实验结果证明，在不使用人工特征的情况下，本文提出的表征增强方法，能够在有效提升多种篇章解析模型的性能。在未来的工作中，我们将致力于探究其他篇章衔接信息对篇章结构分析的辅助作用。

致谢

感谢匿名评审专家对本工作提出的建设性修改意见。感谢张龙印博士在前期实验设计和实验过程中给出的启发和指导。本文的工作受到国家自然科学基金面上项目(No. 61876118)支持。

参考文献

Lynn Carlson and Daniel Marcu. 2001. Discourse tagging reference manual. *ISI Technical Report ISI-TR-545*, 54:56.

Vanessa Wei Feng and Graeme Hirst. 2014. A linear-time bottom-up discourse parser with constraints and post-editing. In *Proceedings of the 52nd ACL*, volume 1, pages 511–521.

Naman Goyal and Jacob Eisenstein. 2016. A joint model of rhetorical discourse structure and summarization. In *Proceedings of the Workshop on Structured Prediction for NLP*, pages 25–34, Austin, TX, November. Association for Computational Linguistics.

Michael Heilman and Kenji Sagae. 2015. Fast rhetorical structure theory discourse parsing. *arXiv preprint arXiv:1505.02425*.

Hugo Hernault, Helmut Prendinger, Mitsuru Ishizuka, et al. 2010. Hilda: A discourse parser using support vector machine classification. *Dialogue & Discourse*, 1(3).

Yangfeng Ji and Jacob Eisenstein. 2014. Representation learning for text-level discourse parsing. In *Proceedings of the 52nd ACL*, volume 1, pages 13–24.

Yangfeng Ji and Noah Smith. 2017. Neural discourse structure for text categorization. *arXiv preprint arXiv:1702.01829*.

Shafiq Joty, Giuseppe Carenini, Raymond Ng, and Yashar Mehdad. 2013. Combining intra-and multi-sentential rhetorical parsing for document-level discourse analysis. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 486–496.

Shafiq Joty, Giuseppe Carenini, and Raymond T Ng. 2015. Codra: A novel discriminative framework for rhetorical analysis. *Computational Linguistics*, 41(3):385–435.

Shafiq Joty, Francisco Guzmán, Lluís Màrquez, and Preslav Nakov. 2017. Discourse structure in machine translation evaluation. *Computational Linguistics*, 43(4):683–722, December.

Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Thomas N Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.

Naoki Kobayashi, Tsutomu Hirao, Hidetaka Kamigaito, Manabu Okumura, and Masaaki Nagata. 2020. Top-down rst parsing utilizing granularity levels in documents. In *Association for the Advancement of Artificial Intelligence 2020, AAAI2020*.

F. Kong and G. Zhou. 2017. A cdt-styled end-to-end chinese discourse parser. *Acm Transactions on Asian Language Information Processing*, 16(4):26.1–26.17.

Jiwei Li, Rumeng Li, and Eduard Hovy. 2014a. Recursive deep models for discourse parsing. In *Proceedings of EMNLP 2014*, pages 2061–2069.

Yancui Li, Wenhe Feng, Jing Sun, Fang Kong, and Guodong Zhou. 2014b. Building Chinese discourse corpus with connective-driven dependency tree structure. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2105–2114, Doha, Qatar, October. Association for Computational Linguistics.

Qi Li, Tianshi Li, and Baobao Chang. 2016. Discourse parsing with attention-based hierarchical neural networks. In *Proceedings of EMNLP 2016*, pages 362–371.

Xiang Lin, Shafiq Joty, Prathyusha Jwalapuram, and M Saiful Bari. 2019. A unified linear-time framework for sentence-level discourse parsing. In *ACL*, pages 4190–4200, Florence, Italy, July. Association for Computational Linguistics.

Mingyu Derek Ma, Kevin Bowden, Jiaqi Wu, Wen Cui, and Marilyn Walker. 2019. Implicit discourse relation identification for open-domain dialogues. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 666–672, Florence, Italy, July. Association for Computational Linguistics.

Mathieu Morey, Philippe Muller, and Nicholas Asher. 2017. How much progress have we made on rst discourse parsing? a replication study of recent results on the rst-dt. In *Proceedings of EMNLP 2017*, pages pp–1330.

Jane Morris and Graeme Hirst. 1991. Lexical cohesion computed by thesaural relations as an indicator of the structure of text. *Computational linguistics*, 17(1):21–48.

Bita Nejat, Giuseppe Carenini, and Raymond Ng. 2017. Exploring joint neural model for sentence level discourse parsing and sentiment analysis. In *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*, pages 289–298, Saarbrücken, Germany, August. Association for Computational Linguistics.

Jeffrey Pennington, Richard Socher, and Christopher D Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543.

Yuanyuan Qiu, Hongzheng Li, Shen Li, Yingdi Jiang, Renfen Hu, and Lijiao Yang. 2018. Revisiting correlations between intrinsic and extrinsic evaluations of word embeddings. In Maosong Sun, Ting Liu, Xiaojie Wang, Zhiyuan Liu, and Yang Liu, editors, *Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data*, pages 209–221, Cham. Springer International Publishing.

Rico Sennrich. 2018. Why the time is ripe for discourse in machine translation. In *Invited talk at the 2nd Workshop on Neural Machine Translation and Generation*.

C. Sun and F. Kong. 2018. A transition-based framework for chinese discourse structure parsing. *Journal of Chinese Information Processing*.

Nan Yu, Meishan Zhang, and Guohong Fu. 2018. Transition-based neural rst parsing with implicit syntax features. In *Proceedings of COLING 2018*, pages 559–570.

Longyin Zhang, Xin Tan, Fang Kong, and Guodong Zhou. 2019. A recursive information flow gated model for rst-style text-level discourse parsing. In Jie Tang, Min-Yen Kan, Dongyan Zhao, Sujian Li, and Hongying Zan, editors, *Natural Language Processing and Chinese Computing - 8th CCF International Conference, NLPCC 2019, Dunhuang, China, October 9-14, 2019, Proceedings, Part II*, volume 11839 of *Lecture Notes in Computer Science*, pages 231–241. Springer.

Longyin Zhang, Yuqing Xing, Fang Kong, Peifeng Li, and Guodong Zhou. 2020. A top-down neural architecture towards text-level parsing of discourse rhetorical structure. In *Proceedings of 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics.