

基于Self-Attention的句法感知汉语框架语义角色标注*

王晓晖^{1,¶}, 李茹^{1,2,‡,†}, 王智强^{3,‡}, 柴清华^{4,‡}, 韩孝奇^{1,¶}

¹计算机与信息技术学院,山西大学

²计算智能与中文信息处理教育部重点实验室,山西大学

³智能信息处理研究所,山西大学

⁴外国语学院,山西大学

¶{1640051575,419432571}@qq.com ‡{liru,wangzq,charles}@sxu.edu.cn

摘要

框架语义角色标注 (Frame Semantic Role Labeling, FSRL) 是基于FrameNet标注体系的语义分析任务。语义角色标注通常对句法有很强的依赖性,目前的语义角色标注模型大多基于双向长短时记忆网络Bi-LSTM,虽然可以获取句子中的长距离依赖信息,但无法很好获取句子中的句法信息。因此,引入self-attention机制来捕获句子中每个词的句法信息。实验结果表明,该模型在CFN (Chinese FrameNet, 汉语框架网) 数据集上的F1达到83.77%,提升了近11%。

关键词: 语义角色标注; self-attention机制; Bi-LSTM; 汉语框架网

Syntax-Aware Chinese Frame Semantic Role Labeling Based on Self-Attention

Xiaohui Wang^{1,¶}, Ru Li^{1,2,‡}, Zhiqiang Wang^{3,‡}, Qinghua Chai^{4,‡}, Xiaoqi Han^{1,‡}

¹School of Computer and Information Technology, Shanxi University

²Key Laboratory of Ministry of Education for Computational Intelligence
and Chinese Information Processing, Shanxi University

³Institute of Intelligent Information Processing, Shanxi University

⁴School of Foreign Languages, Shanxi University

¶{1640051575,419432571}@qq.com ‡{liru,wangzq,charles}@sxu.edu.cn

Abstract

Frame semantic role labeling is a semantic analysis task based on the FrameNet. Semantic role labeling usually relies heavily on syntax. Most of the existing semantic role labeling models are based on Bi-LSTM. Although they can obtain the long-distance dependency information in sentences, they can not get the syntactic information well. Therefore, the self attention mechanism is introduced to capture the syntactic information of each word in the sentence. Experimental results show that the F1 of our model on Chinese FrameNet dataset reaches 83.77%, which is increased by nearly 11%.

Keywords: Semantic role labeling, Self-Attention, Bi-LSTM, Chinese FrameNet

1 引言

框架语义角色标注的主要任务是识别出句子中给定目标词所激起框架对应的框架元素,这些框架元素被赋予特定的含义,如施动者、受动者时间、地点等。简单来说,就是“谁”在“什

* 基金项目: 国家自然科学基金(No.61772324;No.61936012)

† 通讯作者

©2020 中国计算语言学大会

根据《Creative Commons Attribution 4.0 International License》许可出版

么时间”、“什么地点”对“谁”做了“什么”。图1为汉语框架语义角色标注示例，其中，例句下面为句法依赖标签，例句上面的为语义角色标签。例句的目标词为“参观”，该目标词激起的为框架为“拜访”，该框架有两个核心框架元素agt（施动者）和ent（实体），分别对应例句中的“我们”和“吴店镇”，而例句中的“实地”则标为manr（方式）。语义角色标注的应用非常广泛，如在信息抽取 (Surdeanu et al., 2003)、问答系统 (Yih et al., 2003)等领域取得了一定研究成果。

传统的语义角色标注方法多采用机器学习与特征工程相结合的方法。在这类方法中，通常依赖于人工抽取的特征，并且会带来模型复杂、特征稀疏等问题 (Zhang et al., 2019)。语义角色标注对句法有着较强的依赖性。如图1所示，对于目标词“参观”来说，通常其主语都被标注为“施动者”，宾语被标注为“实体”，而“方式”则通常为副词或介词短语作状语。因此，句法信息在一定程度上有助于语义角色标注。

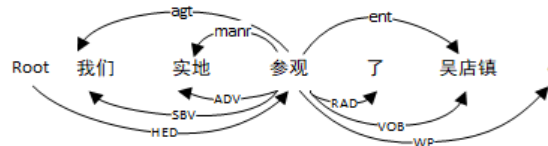


图 1. 汉语框架语义角色标注示例

近年来，深度学习方法逐渐被应用到自然语言处理任务中，它能自动学习文本中的特征，从而大大减少了特征工程中的工作量。特别地，(Zhou and Xu, 2015; He et al., 2017; Marcheggiani et al., 2017)提出了无句法的语义角色标注模型，并且取得了较好的效果。这似乎和句法信息是高性能语义角色标注的前提条件的观点相冲突。然而，句法信息被认为与语义关系密切，在语义角色标注任务中起着至关重要的作用 (Punyakankok et al., 2008)。(Marcheggiani and Titov, 2017)提出了图卷积网络模型，以相对较好的句法分析器作为输入，进一步提高了语义角色标注的性能。

由于Bi-LSTM可以有效地获取句子中的长距离依赖信息，在序列标注任务中有着天然优势。因此，现有的语义角色标注模型大多基于Bi-LSTM模型。但Bi-LSTM却无法很好的获取句子的句法信息，而近年来同样在自然语言处理领域广泛应用的self-attention机制却可以处理这个问题。因此，本文在Bi-LSTM的基础上引入self-attention机制。同时，在序列标注任务中标签之间是有依赖关系的，如在BIO标注模式中，标签I应出现在标签B之后，而不应该出现在O之后，所以本文利用CRF进行全局标签优化预测出最优标签序列。

2 相关工作

语义角色标注是由 (Gildea and Jurafsky, 2002)提出的，同时在人工标注的FrameNet语料上提出了基于统计的分类器。已有的汉语语义角色标注方法有两类，分别是基于特征工程的方法和基于神经网络的方法。

在早期的语义角色标注工作中，大多研究都致力于特征工程。(Gildea and Jurafsky, 2002)在没有大规模语义标注语料库的情况下对汉语语义角色标注进行了初步的研究，取得了很好的效果。(Xue, 2008)在中文PropBank (CPB) 上将最大熵分类器和特征工程相结合，同时将标准的句法分析和自动句法分析分别加入到特征工程中，实验表明汉语句法分析的性能是实现高效语义角色标注的关键；(Li et al., 2010)基于CRF在CFN数据集上进行语义角色标注研究，分别取得了63.65%和61.62%的F1值；(Wang, 2010)基于最大熵模型分别使用词层面和块层面特征实现了的汉语框架语义角色标注；(Tu et al., 2016)提出了一种基于主动学习的方法，当数据规模相同时，实验结果最高提升了4.83%，同时，达到同等F1值时最高可减少30%的人工标注量。

随着深度神经网络模型在自然语言处理领域的诸多任务上得到成功应用，一系列基于神经网络的语义角色标注模型被提出。(Ronan and Jason, 2008)等提出了使用单一的卷积神经网络结构进行包括语义角色标注任务在内的多任务学习模型，整个网络共享权重，代表了一种新的共享任务半监督学习形式；(Wang et al., 2014)将语义角色标注分为角色识别和角色分类两个步骤，利用分层输出的神经网络模型取得了64.19%的F1值。(Wang et al., 2015)提出了基于异构数据的Bi-LSTM模型，可以更方便地缓解单个标注语料库的可扩展性问题，

在CPB数据集上取得了77.59%的F1值；(Dang, 2015)基于词分布的汉语框架语义角色标注模型实现了72.89%的F1值；(Wang et al., 2017)基于分布式表示，提出了一种多特征融合的神经网络结构，同时使用Dropout正则化技术有效地缓解了模型的过拟合现象，使得模型的F1值提高了近7%。

最近，人们尝试构建基于span的无句法输入的端到端语义角色标注模型(Zhou and Xu, 2015; He et al., 2017; Tan et al., 2018; Ouchi et al., 2018)。尽管无句法信息的模型取得了成功，但是仍有许多研究致力于如何将句法优势应用到语义角色标注中。(Roth and Lapata, 2016)将复杂的句法结构和句法相关现象看作是句法依赖路径的子序列，并将深度Bi-LSTM应用到语义角色标注中，在PropBank数据集上取得较好效果；(Qian et al., 2017)提出了SA-LSTM，以一种结构工程的方式对整个句法依赖树结构进行建模。

目前，大部分性能较好的语义角色标注模型都是基于Bi-LSTM的，但Bi-LSTM无法很好的获取句子的句法信息。近年来，注意力机制在自然语言处理领域广泛应用，受到(He et al., 2018)和(Zhang et al., 2019)的启发，本文引入self-attention机制，将其加入到词表示和Bi-LSTM编码器之间，同时使用条件随机场进行标签预测。

3 基于Self-Attention的语义角色标注

本文将语义角色标注任务转换成序列标注问题，给定一个句子序列及其目标词，在目标词所激起框架已知的条件下，识别出句子中与目标词所搭配的语义角色。即：给定句子 $X = (w_1, w_2, w_3, \dots, w_n)$ ，和其目标词 $w_{target} (1 \leq target \leq n)$ ，输出角色标签序列 $Y^* = \arg \max p(Y|X, w_{target}, frame)$ 。图2为本文的模型结构，它由三个模块组成：

- (1) self-attention层：对句子中各词的语义重要性进行建模；
- (2) Bi-LSTM编码层；
- (3) CRF标签预测层。

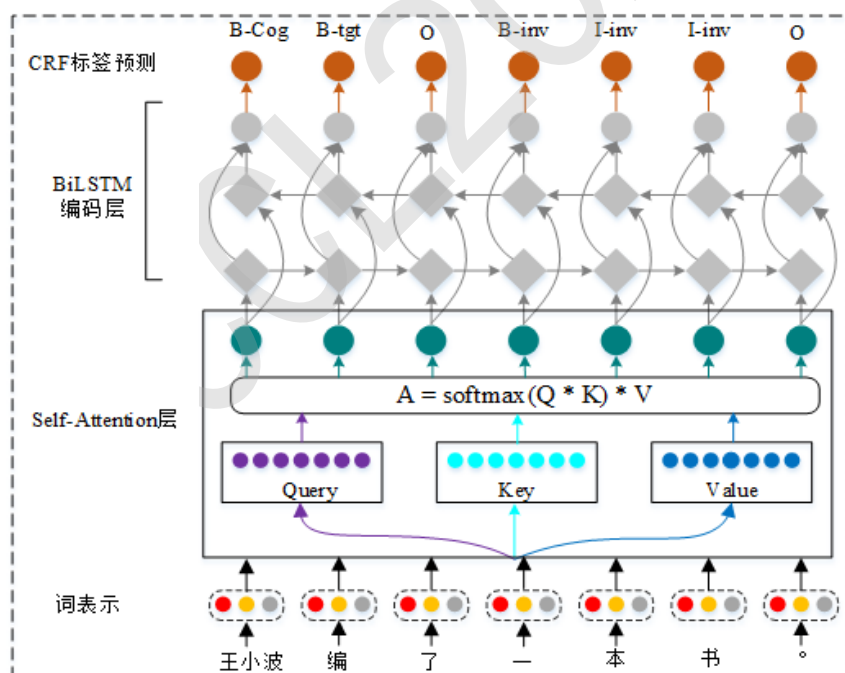


图 2. 语义角色标注模型图

3.1 词特征向量表示

本文为句子中的每个词 w_i 生成一个词表示 x_i ，其中， i 表示词在句子中的位置。每个 x_i 是由几个特征拼接起来的：预训练的词向量 x_i^{pe} 、随机初始化的词性表示 x_i^{pos} 、随机初始化的框架表示 x_i^{frame} 、随机初始化的目标词表示 x_i^{target} 、位置表示 x_i^{loc} 、以及谓词指示器 $x_i^{indicator}$ 。所以，最终的词表示为 $x_i = [x_i^{pe}, x_i^{pos}, x_i^{frame}, x_i^{target}, x_i^{loc}, x_i^{indicator}]$ 。

3.2 Self-Attention层

在语义角色标注任务中，通常对句法的依赖性较强，但是Bi-LSTM模型擅长获取句子中的长距离依赖信息，却不能很好地提取句法信息。因此本文在词表示和Bi-LSTM编码层之间引入self-attention机制捕获句子中每个词的句法信息，同时还可以进一步增强Bi-LSTM获取长距离依赖信息的能力。

将词表示矩阵 X 映射为不同的表示 K, Q, V ，首先对 Q 和 K 执行点积操作，并对其进行缩放操作，如公式(1)所示。再对其执行softmax操作进行归一化，得到句子中每个词之间的attention权重 a_i 。最后，将attention权重 a_i 点乘 V 并求和，得到每个词的表示矩阵 A 。

$$f(Q, K_i) = \frac{QK_i^T}{\sqrt{d_k}} \quad (1)$$

$$a_i = \text{softmax}(f(Q, K_i)) = \frac{\exp(f(Q, K_i))}{\sum_j \exp(f(Q, K_j))} \quad (2)$$

$$A = \sum_i a_i V_i \quad (3)$$

3.3 Bi-LSTM编码层

在汉语框架语义角色标注任务中，目标词与语义角色块并不会总是相邻的。因此，为了更加有效地获取长距离信息，在self-attention编码层之后，我们加入一个多层Bi-LSTM编码层，以取得更加丰富的表示。

在LSTM网络中，每个细胞单元有 $\tilde{C}, g_i, g_f, g_o, C_t, h_t$ ，其中 \tilde{C} 是当前细胞状态的候选值； g 是控制信息流动的的门； C_t 是当前细胞状态； h_t 是细胞隐状态，具体如下：

$$\tilde{C} = \tanh(W_c z_t + U_c h_{t-1} + b_c) \quad (4)$$

$$g_j = \text{sigmoid}(W_j z_t + U_j h_{t-1} + b_j), j \in \{i, f, o\} \quad (5)$$

$$C_t = g_i \odot \tilde{C} + g_f \odot C_{t-1} \quad (6)$$

$$h_t = g_o \odot f(C_{t-1}) \quad (7)$$

其中， \odot 代表按元素乘。最后将前向序列 \vec{h}_t 和反向序列 \overleftarrow{h}_t 拼接得到Bi-LSTM编码层的输出 $h_t = [\vec{h}_t; \overleftarrow{h}_t]$ 。

3.4 标签预测

我们在Bi-LSTM编码层之后，加上一层CRF进行标签预测。因为，在序列标注问题中，相邻词的标签间存在很强的依赖关系，单独考虑每个词的标签得分是不合适的。所以使用CRF对整个序列进行全局归一化，得到概率最大的最优序列。

假设输入序列为 $X = (x_1, x_2, \dots, x_n)$ ，输出序列为 $y = (y_1, y_2, \dots, y_n)$ ，约定 H 是Bi-LSTM层的输出矩阵，而 $H_{i,j}$ 对应于句子中第 i 个单词的第 j 个标签的得分。那么，它的得分定义如式(8)所示，其中， S 是标签的转移得分矩阵， $S_{i,j}$ 表示从标签 i 到标签 j 的转移得分。

$$s(X, y) = \sum_{i=0}^n S_{y_i, y_{i+1}} + \sum_{i=1}^n H_{i, y_i} \quad (8)$$

对于输出序列 y 所有可能的标签序列的概率的softmax形式如式(9)，其中， Y_X 表示输入序列 X 的所有可能的标签序列。在解码时，我们通过式(10)预测输出序列。

$$P(y|X) = e^{s(X,y)} / \sum_{\tilde{y} \in Y_X} e^{s(X, \tilde{y})} \quad (9)$$

$$y^* = \arg \max_{\tilde{y} \in Y_X} s(X, \tilde{y}) \quad (10)$$

4 实验设计及分析

4.1 实验数据

本文实验数据来自山西大学的汉语框架网的例句库，该例句库的标注数据由专业人士人工标注。本文选取其中的25个框架658个词元共19355条标注数据。本实验将上述数据按框架分割成训练集、验证集、测试集，分割比例为6:2:2。

4.2 评价指标

本文使用准确率（Precision）、召回率（Recall）、 $F1$ 值作为评价指标。具体如下：

$$P = CC/PC \quad (11)$$

$$R = CC/DC \quad (12)$$

$$F1 = 2PR/(P + R) \quad (13)$$

其中， CC 为模型识别出的正确的标签数； PC 为模型识别出的标签数； DC 为数据集中的标签数。

4.3 参数设置

使用Glove在CFN例句库上预训练的词向量，词向量维度为100。词性、句法路径特征等特征的维度为10，其中，使用LTP（语言技术平台）进行句法分析。此外，其他超参数设置为学习率 $learning_rate = 0.015$ ，丢弃率 $dropout = 0.5$ ，隐藏层维度 $hidden_dim = 200$ ，优化函数为SGD，正则化系数 $L2 = 1e - 8$ 。

4.4 实验结果与实验分析

表1为本文模型在CFN数据集上实验结果和已有模型的对比。表中第一栏为已有模型中基于神经网络模型的汉语框架语义角色标注模型，其中最好结果为72.89%。本文在基于Bi-LSTM的语义角色标注模型中融入self-attention机制，在CFN数据集上实现了83.77%的F1值，比已有模型中最好结果提高10.88个百分点，比未引入self-attention机制之前的实验结果提升了5.94个百分点。

模型	F1/%
(Lv, 2014)	60.51
(Yang, 2015)	69.91
(Wang et al., 2017)	70.54
(Dang, 2015)	72.89
ABLC	83.77

表 1. 语义角色标注在CFN数据集上的结果对比

为了验证self-attention机制的效果，设置了两个对比实验：BLC和BLAC，前者未引入self-attention机制，后者将self-attention机制加入到Bi-LSTM编码器之后，实验结果如表2所示。首先，无句法路径特征时，本文模型ABLC和BLAC较BLC分别提高了7.39%、4.49%，而加入路径特征时，ABLC、BLAC比BLC分别提高了5.94%、3.98%；其次，本文的ABLC模型比BLAC的性能分别高了2.9%和1.96%。因此，验证了本文模型引入self-attention机制在汉语框架语义角色标注任务中的有效性。此外，引入self-attention机制后，即BLAC和ABLC，加入句法路径特征对结果提升不是很明显，分别提高了0.92%和-0.02%，表明本文模型可以获取一定句法信息。

为了分析不同路径特征对实验结果的影响，本文分别使用了三种不同的路径特征进行实验，结果如表3所示。总体而言，使用一级路径特征的实验结果相对较好。另外，从表中可以看出，不同路径特征对BLC模型的影响较大，而对引入self-attention机制的BLAC和ABLC的影响则相对较小。进一步证明了本文模型可以提取一定的句法信息。

模型	无句法特征						有句法特征					
	dev			test			dev			test		
	<i>P</i>	<i>R</i>	<i>F1</i>	<i>P</i>	<i>R</i>	<i>F1</i>	<i>P</i>	<i>R</i>	<i>F1</i>	<i>P</i>	<i>R</i>	<i>F1</i>
BLC	81.67	68.60	74.12	83.56	71.70	76.40	81.20	70.03	74.46	83.83	73.30	77.83
BLAC	83.67	72.49	77.55	89.24	73.72	80.89	83.63	74.02	78.58	88.84	76.17	81.81
ABLC	86.16	77.88	81.53	89.21	79.35	83.79	83.03	74.53	78.15	89.93	79.00	83.77

表 2. 对比实验结果

模型	F1		
	全路径	二级路径	一级路径
BLC	77.83	76.91	81.41
BLAC	81.81	81.75	82.25
ABLC	83.77	84.32	83.96

表 3. 不同路径特征的实验结果对比

由于数据稀疏对实验的影响，本文对利用框架关系对实验数据的扩充进行初步探索。本文使用三种方式进行数据扩充。将目标框架的数据与关系数据的train、dev分别合并后作为扩充数据的train、dev，扩充后的数据称为Data1；为了验证扩充数据在无例句的框架的性能，将关系数据的train、dev作为扩充数据的train、dev，称为Data2；因为非核心框架元素具有稀疏性，会影响标注模型的性能，因此，将关系数据的train、dev去掉核心框架元素作为扩充数据的train、dev，称为Data3。

表4为本文模型在扩充数据上实验结果。从表2和表4可以得出，在Data1和Data3上的实验结果都比在目标框架的原始数据上的实验结果高；同时，在Data2上的实验结果可以达到与目标框架原始数据相当水平。结果表明，利用框架关系扩充数据可以提高汉语框架语义角色标注的性能。

模型	句法特征			有句法特征		
	<i>Data1</i>	<i>Data2</i>	<i>Data3</i>	<i>Data1</i>	<i>Data2</i>	<i>Data3</i>
BLC	81.79	77.36	84.79	82.43	76.58	85.83
BLAC	79.02	75.52	83.66	79.08	76.10	85.58
ABLC	82.48	75.05	83.11	87.71	76.40	84.38

表 4. 扩充数据上的实验结果

为了分析本文模型对句子中长距离依赖的影响，记录了不同的角色与目标词的距离集合的F1值，如图3所示。从图中可以看出，所有距离的F1值都得到了改进。这表明了我们的模型在引入self-attention机制后可以增强模型捕获长距离信息的能力，进而提升汉语框架语义角色标注性能。

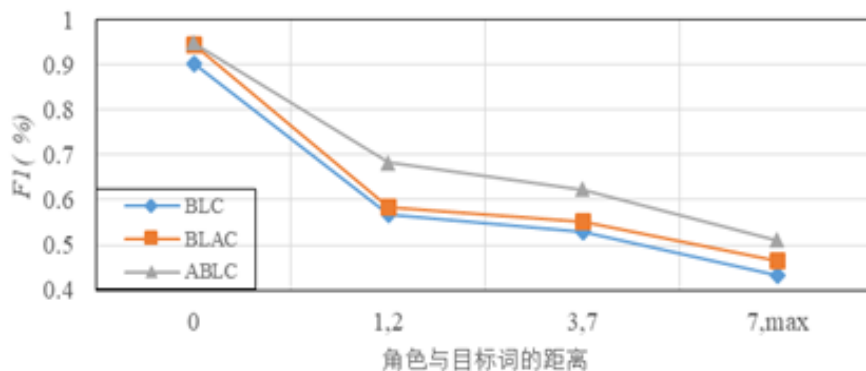


图 3. 角色与目标词距离和F1值

5 总结

本文提出了一种基于self-attention机制的句法感知汉语框架语义角色标注模型，同时，本文利用框架关系对实验数据进行扩充，以减少数据稀疏对语义角色标注的影响。实验结果表明，本文模型的实验结果比已有模型中的最好提高了超过10个百分点，并且验证了本文模型可以获取一定句法信息；同时，利用框架关系对数据进行扩充，有助于缓解SRL中的数据稀疏问题，尤其对无例句框架的SRL的帮助更大。此外，在未来工作中，在更好的融入句法信息以及数据稀疏方面还有很多工作需要完成。

参考文献

- Surdeanu Sanda, Williams John and Aarseth Paul. 2003. *Using Predicate-Argument Structures for Information Extraction, Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics*, pages:8-15. Sapporo, Japan.
- Yih Wen-tau, Richardson Matthew, Meek Chris, Chang Ming-Wei and Suh Jina. 2016. *The Value of Semantic Parse Labeling for Knowledge Base Question Answering, Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, pages:201-206. Berlin, Germany.
- 张苗苗, 张玉洁, 刘明童, 徐金安, 陈钰枫. 2018. 基于Gate机制与Bi-LSTM-CRF的汉语语义角色标注, 计算机与现代化, (4):1-6+31.
- Zhou Jie and Xu Wei. 2015. *End-to-end learning of semantic role labeling using recurrent neural networks, Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, pages:1127-1137. Beijing, China.
- He Luheng, Lee Kenton, Lewis Mike, Zettlemoyer Luke. 2017. *Deep Semantic Role Labeling: What Works and What's Next, Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, pages:473-483. Vancouver, Canada.
- Marcheggiani Diego, Frolov Anton, Titov Ivan. 2017. *A Simple and Accurate Syntax-Agnostic Neural Model for Dependency-based Semantic Role, Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017)*, pages:411-420. Vancouver, Canada.
- Punyakanok Vasin, Roth Dan, Yih Wen-tau. 2008. *The Importance of Syntactic Parsing and Inference in Semantic Role Labeling, Computational Linguistics*, 34(2):257-287.
- Marcheggiani Diego and Titov Ivanu. 2017. *Encoding Sentences with Graph Convolutional Networks for Semantic Role Labeling, Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages:1506-1515. Copenhagen, Denmark.
- Gildea Daniel and Jurafsky Daniel. 2002. *Automatic Labeling of Semantic Roles, Computational Linguistics*, 28(3):245-288.
- Sun Honglin and Jurafsky Daniel. 2004. *Shallow Semantic Parsing of Chinese, Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics: HLT-NAACL 2004*, pages:249-256. Boston, Massachusetts, USA.
- Xue Nianwen. 2008. *Labeling Chinese Predicates with Semantic Roles, Computational Linguistics*, 34(2):225-255.
- 李济洪, 王瑞波, 王蔚林, 李国臣. 2010. 汉语框架语义角色的自动标注, 软件学报, 21(4):597-611.
- 王蔚林. 基于最大熵模型的汉语框架语义角色自动标注, 山西大学, 2010.
- 屠寒非, 李茹, 王智强, 周铁峰. 2016. 一种基于主动学习的框架元素标注, 中文信息学报, 30(4):44-55.
- Ronan Collobert and Jason Weston. 2008. *A unified architecture for natural language processing: Deep neural networks with multitask learning, Proceedings of the Twenty-Fifth International Conference (ICML 2008)*, pages:249-256. Helsinki, Finland.
- 王臻, 常宝宝, 穗志方. 2014. 基于分层输出神经网络的汉语语义角色标注, 中文信息学报, 28(6):51-61.
- Wang Zhen, Jiang Tingsong, Chang Baobao, Sui Zhifang. 2015. *Chinese Semantic Role Labeling with Bidirectional Recurrent Neural Networks, Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages:1626-1631. Lisbon, Portugal.

党帅兵. 基于词分布表征的汉语框架语义角色识别研究, 山西大学,2015.

王瑞波, 李济洪, 李国臣, 杨耀文. 2017. 基于Dropout正则化的汉语框架语义角色识别, 中文信息学报,31(1):147-154.

Tan Zhixing , Wang Mingxuan, Xie Jun, Chen Yidong, Shi Xiaodong . 2018. *Deep Semantic Role Labeling With Self-Attention, Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18)*,pages:4929–4936. New Orleans, Louisiana, USA, February.

Ouchi Hiroki, Shindo Hiroyuki, Matsumoto Yuji. 2018. *A Span Selection Model for Semantic Role Labeling, Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*,pages:1630-1642. Brussels, Belgium.

Qian Feng, Sha Lei, Chang Baobao, Liu Luchen, Zhang Ming. 2017. *Syntax Aware LSTM model for Semantic Role Labeling, Proceedings of the 2nd Workshop on Structured Prediction for Natural Language Processing*.

Ouchi Hiroki, Shindo Hiroyuki, Matsumoto Yuji. 2016. *Neural Semantic Role Labeling with Dependency Path Embeddings, Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*,pages:1192-1202. Berlin, Germany.

He Luheng, Lee Kenton, Levy Omer, Zettlemoyer Luke. 2018. *Jointly Predicting Predicates and Arguments in Neural Semantic Role Labeling, Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*,pages:364-369. Melbourne, Australia.

He Luheng, Lee Kenton, Levy Omer, Zettlemoyer Luke. 2019. *Syntax-Enhanced Self-Attention-Based Semantic Role Labeling, Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*,pages:616-626. Hong Kong, China.

吕雷. 汉语框架语义角色自动标注研究, 山西大学,2014.

杨耀文. 基于神经网络模型的汉语框架语义角色识别, 山西大学,2016.