

# 融合知识的多目标词联合框架语义分析模型

陈旭东<sup>1,2</sup>, 郑策<sup>1</sup>, 常宝宝<sup>1\*</sup>

1.北京大学计算语言学教育部重点实验室, 北京100871

2.北京大学软件与微电子学院, 北京102600

{xdc,zce1112zslx,chbb}@pku.edu.cn

## 摘要

框架语义分析任务是自然语言处理领域的一项基础性任务。先前的研究工作大多针对单目标词进行模型设计,无法一次性完成多个目标词的框架语义结构提取。本文提出了一个面向多目标的框架语义分析模型,实现对多目标词的联合预测。该模型对框架语义分析的各项子任务进行交互性建模,实现子任务间的双向交互。此外,本文利用关系图网络对框架关系信息进行编码,将其作为框架语义学知识融入模型中。实验表明,本文模型在不借助额外语料的情况下相比之前模型都有不同程度的提高。消融实验证明了本文模型设计的有效性。此外我们分析了模型目前存在的局限性以及未来的改进方向。

**关键词:** 框架语义分析; 框架网络

## Knowledge-integrated Joint Model For Multi-target Frame Semantic Parsing

Xudong Chen<sup>1,2</sup>, Ce Zheng<sup>1</sup>, Baobao Chang<sup>1\*</sup>

1.The MOE Key Laboratory of Computational Linguistics, Peking University  
, Beijing 100871, China

2.School of Software and Microelectronics, Peking University  
, Beijing 100260, China

{xdc,zce1112zslx,chbb}@pku.edu.cn

## Abstract

Frame semantic parsing is a fundamental task in natural language processing. Most of the previous research work focuses on the design of the single-target model. Therefore, these models can't extract frame semantic structures of multiple targets in one time. This paper designs a frame semantic parsing model for multiple targets which jointly predicts the results of multiple targets. We model and achieve the bidirectional interaction among the subtasks of frame semantic parsing. Moreover, Relational Graph Convolution Network (R-GCN) is utilized to encode the frame relation information, which is a way to exploit frame semantic knowledge into our model. The experiments shows that our model maintains good performance without extra training corpus. Ablation Study proves the effectiveness of our model.

**Keywords:** Frame semantic parsing, FrameNet

\* 通讯作者

©2022 中国计算语言学大会

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

## 1 引言

框架语义分析任务是自然语言处理领域的基础性任务，被广泛应用于下游自然语言处理任务中，例如阅读理解(Guo et al., 2020)、问答系统(Shen and Lapata, 2007)等。



Figure 1: 框架语义结构的文本标注示例。

框架语义分析任务基于框架网络 (FrameNet) 标注体系，提取文本中的结构化的信息，包括语义框架以及语义框架下的一系列语义角色。Figure 1展示了一个包含框架语义结构标注的文本。在该文本中，can和write作为目标词分别触发了Capability和Text\_creation两个框架，此外，图中还标注了跟框架有关的框架语义角色。这里以Text\_creation为例，文本中的I、my name和on the deposit slip分别扮演了该框架下的作者 (Author)、文本 (Text) 和载体 (Form) 三种角色。

因此，框架语义分析任务由多个子任务组成。在给定目标词的情况下，需要完成对目标词的框架识别以及文本中与框架有关的角色位置和类型的识别，后者统称为角色识别。先前的研究工作提出的模型绝大多数是针对单目标词的框架语义分析模型，无法一次性处理文本中包含多个目标词及其对应的框架语义结构的情况。它们通常会对多目标词文本采用了一个目标词一个文本的切分方式进行处理，让模型多次预测。此外，同个文本中的不同目标词在语义上会存在一定的相关性，独立多次的预测方式下模型无法学习到目标词之间的语义联系。我们统计了FrameNet训练集中不同目标词个数区间的句子数量分布情况，如Figure 2所示。可以看到绝大多数的句子都包含两个以上的目标词，且大部分句子的目标词数目集中在2到9之间。可以推断出，自然文本中存在多个目标词以及它们对应的框架语义结构是非常普遍的。因此，对多目标词进行联合框架语义分析具有一定的应用价值和研究意义。

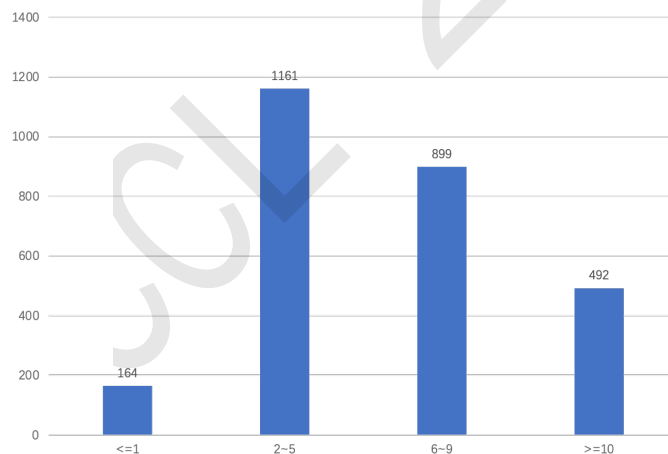


Figure 2: FN1.5训练集中不同目标词个数的句子数量分布情况。

本文提出了一个面向多目标词的框架语义分析模型，联合预测文本中所有目标词的框架语义结构。此外，我们在框架知识融入和子任务的交互性建模两个方面对模型进行改进。

在框架知识融入方面，FrameNet根据框架语言学理论定义了框架之间的多种关系，这些关系信息能够作为有价值的外部知识融入到框架识别任务中。对框架关系进行显式化建模，有利于模型学习到框架的语义。我们借助框架关系进行构图，并在此基础上进行扩展，将角色加入到框架语义关系中。我们利用图网络对扩展后的框架关系图进行编码，得到包含框架关系知识的框架和角色语义表示，将它们分别融入到框架语义分析的各项子任务中。

框架语义的各项子任务之间的关联性非常密切，加强子任务的交互能缓解误差传播，从而提升框架语义分析模型的表现。我们模型设计了一个两阶段的角色识别模块，分别采用序列标

注和生成式的方式进行角色位置识别，第一阶段的角色识别结果会传递到框架识别模块，框架识别模块的预测结果通过框架语义向量的方式传递到第二阶段的角色识别中，从而实现了子任务间的双向交互。在训练阶段，我们对所有子任务模块的参数进行联合优化。

我们模型在FN1.5数据集上进行实验，并且有了阶段性的结果。在不借助额外语料的情况下，模型在框架识别上有91.8%的准确率，在框架语义结构提取的F1值上达到了78.5%。消融实验证明了我们模型设计的有效性。此外，我们详细分析了模型目前存在的局限性，将在未来工作中作针对性的模型优化。

## 2 相关工作

框架语义分析由Daniel Gildea and Daniel Ju (2002)最先提出，并成为了2007年的SemEval会议的一项国际测评任务(Baker et al., 2007)。早期对于框架分析模型的研究主要基于传统的统计学模型，依赖人工构造一系列规则与特征作为模型的输入。Johansson (2007)利用支持向量机(Support Vector Machine)进行建模，构建出一系列分类器进行框架和角色的识别。选取的特征包括词形还原的目标谓词、目标词形以及句子的依存句法特征等，该方法构建的语义分析系统成为SemEval 2007国际测评任务上最优模型。Das(2010)利用基于统计学的条件对数线性模型进行框架语义分析，以该模型为基础的语义分析器SEMAFOR(2010)的框架语义分析效果超越了Johansson提出的模型。

近些年，随着深度学习技术在自然语言处理领域的广泛应用，有研究者将深度学习语言模型应用到框架语义分析任务中。FitzGerald(2015a)针对框架角色识别任务，将框架、角色与文本特征信息通过全连接神经网络映射到特定维度的向量空间，将得到的特征向量作为输入对角色标注进行打分。该模型可以看作是图网络模型，图节点是句子包含的所有文段以及框架中所有的角色，任意文段顶点与角色顶点存在一条边，边的值即为该文段被识别成对应框架角色的得分。Hermann(2014)最早提出了利用词向量进行框架识别任务，它通过挖掘上下文依存语法信息生成特征向量，并利用WSABIE算法(Weston et al., 2011)将特征向量映射到低维空间，同时训练得到低维的框架语义向量，通过计算特征向量与框架向量的相似性进行框架匹配。实验表明该方法相比之前的研究方法在结果上有显著提升。Hartmann(2017)针对框架语义分析中存在的领域依赖问题进行研究，提出了适用于跨领域的框架分析模型。该模型利用预训练词向量表示目标谓词与上下文信息，采用两层神经网络模型和基于WSABIE算法模型进行框架识别。其中上下文信息由SentBOW与DepBOW两种方式构建。其中，SentBOW是取句子中所有词的词向量均值作为上下文特征向量，DepBOW是取句法树中目标谓词依赖项的词向量均值作为上下文特征向量。在跨领域文本的框架识别任务中，相对于其他模型有更好的识别效果。Swayamdipta(2017)将循环神经网络运用到框架语义角色识别任务中，提出了基于SegRNN的框架角色识别模型，该模型能够在不借助语法解析的情况下（仅在训练阶段进行语法标注）完成框架角色识别任务。SegRNN模型是将双向长短时记忆神经网络(LSTM)(Hochreiter and Schmidhuber, 1997)模型与半马尔可夫随机条件场(Semi-CRF)(Sarawagi and Cohen, 2004)模型联合的联合模型。其中，Semi-CRF模型用于对文本进行切分，模型对文本的每个文段进行额外的编码，获得对应文段的特征向量。最后通过对文段特征向量进行分类，得到框架角色识别的结果。Yang(2017a)提出一种基于BIO标记方案的序列标注模型，该模型由多个LSTM网络堆叠，从而获得更深的上下文语义信息，此外该模型在输出层前添加了一层随机条件场(CRF)用于建模不同标签间的依赖关系。Peng(2018a)提出了一个基于文段的框架语义分析模型，该模型通过遍历句子中所有可能的文段并计算它们的匹配分数。此外，以上两个模型对框架语义分析的各项子任务采用联合训练策略，一定程度上加强了子任务的交互缓解误差传播问题。在此基础上，陈(Chen et al., 2021)通过在各任务模块间设计交互环节，进一步加强子任务之间的交互性。

预训练语言模型的提出，推动了自然语言技术的发展。以BERT(Devlin et al., 2019)为代表的预训练语言模型相比传统的深度学习语言模型，捕捉上下文语义信息的能力更强。目前有研究工作针对框架语义分析中的框架识别任务，提出了基于BERT的框架识别模型。Jiang(2021)利用框架网络语料库中的框架定义信息，将其与文本拼接，通过BERT模型进行编码，对编码后的特征向量进行二分类。该过程将框架识别任务建模成文本与框架的相似性判断任务。Su(2021)针对框架知识的融入进行研究，提出了融入框架知识的框架识别模型，证明了融入框架知识对框架识别任务的有效性。

以上的研究工作主要是面向单目标词的框架语义分析研究。本文模型实现了对多目标词的联合框架语义分析。此外，我们将框架关系知识融入到框架语义分析的各项子任务中，提升模型在整个框架语义分析任务上的表现。

### 3 模型介绍

本文模型主要由文本表示模块、框架关系图表示模块、角色识别模块和框架识别模块四部分组成。

文本表示模块用于编码文本的上下文信息。框架关系图表示模块基于R-GCN对框架关系进行建模，该模块的输出为包含框架关系信息的框架和角色向量表示。框架角色识别模块采用两阶段的识别方式。第一阶段为目标词感知的角色位置预测，该阶段仅利用目标词的语义信息（目标词所属的框架类别未知），采用序列标注的方式对角色位置进行识别，识别结果将送入框架识别模块中。第二阶段为框架感知的角色位置和类别预测，该阶段的输入除了目标词的语义信息外，还接收框架识别模块得到的预测框架所对应的框架向量表示（来自框架关系图表示模块），采用生成式的方式实现角色位置和类别的预测。该阶段的预测结果为最终的角色位置和类别的识别结果。框架识别模块包含一个角色位置感知的注意力机制和框架分类器，分别用于实现目标词的信息聚合和框架预测。

#### 3.1 文本表示模块

文本表示模块用于将一段自然文本 $S$ 进行编码表示，得到包含上下文语义表示的向量序列。该模块的输入为一段文本和一个目标词集合 $T = \{t_1, t_2, \dots, t_k\}$ ， $k$ 为该文本的目标词个数。

本文模型采用预训练好的英文版本的BERT语言模型作为编码器，从而得到包含上下文语义信息的表示序列。此外，常规的BERT语言模型的输入包含单词类型标记向量，用于区分不同的句子，这里我们针对框架识别任务的特点进行改进，将其作为识别某个单词是否属于目标词的标记向量。

此外，一些目标词可能是包含多个单词的词组。对于这种情况，我们模型采用平均所有单词的表示向量作为这一类目标词的语义表示向量。该过程用公式描述如下：

$$h_t = \text{Mean}(h_{w_i}), w_i \in t \quad (1)$$

最终得到所有目标词的语义表示 $H_t = \{h_{t_1}, h_{t_2}, \dots, h_{t_k}\}$

#### 3.2 框架关系图表示模块

框架关系图表示模块用于对框架关系进行建模，得到包含框架关系知识的框架和角色语义表示。由于框架网络定义了框架与框架之间的多种类型的语义关系，普通的GCN架构(Scarselli et al., 2008)只能建模单一类型边的图结构。因此我们采用R-GCN(Schlichtkrull et al., 2018)作为编码器，针对每个不同的关系，通过不同的参数进行图卷积计算。该模块的公式化描述如下：

$$g_i^l = \begin{cases} \text{ReLU} \left( \sum_{e \in E_f} \sum_{j \in N_i^e} \frac{1}{|N_i^e|} W_e^{(l)} g_j^{(l)} + \sum_{k \in N_i^r} \frac{1}{|N_i^r|} W_{r-f}^{(l)} g_k^{(l)} \right) + g_i^{(l)}, i \in F \\ \text{ReLU} \left( W_{f-r}^{(l)} g_k^{(l)} + \sum_{j \in N_i^r} \frac{1}{|N_i^r|} W_{r-r}^{(l)} g_j^{(l)} \right) + g_i^{(l)}, i \in R \cap k \in f_i \end{cases} \quad (2)$$

其中 $F$ 表示框架集合， $l$ 表示图卷积的层数。 $R$ 表示角色集合， $E_f$ 表示框架间的关系类型集合， $f_i$ 为角色 $i$ 对应的框架， $W_{r-f}$ 表示框架到角色的边的权重矩阵， $W_{f-r}$ 表示从角色到框架的边的权重矩阵， $N_i^e$ 表示框架 $i$ 的邻居框架集合， $N_i^r$ 表示框架或角色 $i$ 的邻居角色集合。

最终可以得到包含框架关系知识的框架语义表示 $G_f$ 和角色语义表示 $G_r$ 。

#### 3.3 角色识别模块

角色识别模块包括角色位置的识别和角色类型的识别。其中角色位置识别采用两阶段的识别策略。



### 3.3.1 目标词感知的角色位置识别

第一阶段的角色位置识别采用序列标注的方式，接收文本的上下文表示向量  $H = \{h_1, h_2, \dots, h_n\}$  和目标词向量  $H_t = \{h_{t_1}, h_{t_2}, \dots, h_{t_k}\}$  作为模块输入，这两部分向量都来源于文本表示模块。第一阶段的角色位置识别的目的是为了与框架识别模块中进行交互，指导目标词的上下文信息聚合，不作为最终的框架角色结果，所以这里从模型精简角度，并没有采用随机条件场（CRF）对标签约束进行建模。最终的预测标签通过以下公式得到：

$$y_i^t = \arg \max (\text{Linear}([h_i; h_{t_i}])) \quad (3)$$

其中线性层Linear的输出维度为3，分别代表B（角色开始位置）、I（角色中间或结束位置）和O（非角色）三个标签。第一阶段的序列标注预测结果将传入框架识别模块用于目标词表示的信息聚合。

### 3.3.2 框架感知的角色识别

第二阶段的角色位置识别采用陈(Chen et al., 2021)提出的生成式的角色位置识别方式。该模块采用层次化的指针网络识别角色在文本中的边界位置，同时建模了不同角色间的信息交互。具体的设计细节见他们的论文。该模块的输入包括目标词的语义表示和预测框架的语义表示。输出为所有角色在文本中的边界位置索引集合。公式简化如下：

$$\text{Span}_t^{f_t} = \text{Generator}(h_t, g_{f_t}) \quad (4)$$

其中  $\text{Span}_t^{f_t}$  为模块预测的角色边界位置下标集合，可以表示成集合  $\{(sta_1, end_1), \dots, (sta_m, end_m)\}$ ,  $m$  为角色的个数。

此外，还需要对角色进行分类，我们将角色语义表示  $G_r$  融入到角色分类过程中，利用框架关系知识指导模型对框架角色的识别。该过程如下：

$$h_{span} = \text{MLP}([h_i^{end} + h_i^{sta}; h_i^{end} - h_i^{sta}; g_f]) \quad (5)$$

$$\text{logit}_r = \text{Classifier}(h_{span}) + G_r \cdot h_{span} \quad (6)$$

$$P(r|f_t, t, i^{sta}, i^{end}) = \text{Softmax}(\text{logit}_r) \quad (7)$$

其中，MLP为多层感知机网络，由Relu激活函数和两层线性层组成。Classifier是一个可训练的线性分类器， $P(r|f_t, t, i^{sta}, i^{end})$  为该角色位置对应的角色类别预测概率分布。该阶段得到的框架位置和类别预测结果作为最终的角色识别结果输出。

## 3.4 框架识别模块

框架识别模块包括角色感知的注意力机制和一个基于多层感知机的分类器。角色感知注意力机制接收框架角色的序列标注信息，对标签为非O的词计算注意力系数，实现目标词的信息聚合。目标词  $t$  对其他非O标签的词的注意力系数计算如下，其他目标词的计算过程与之相同：

$$\alpha_i^t = \frac{\exp(h_i^\top W_1 h_t)}{\sum_{y_j^t \neq "O"} \exp(h_j^\top W_1 h_t)} \quad (8)$$

$$c_t = \sum_{y_j^t \neq "O"} \alpha_j^t h_j \quad (9)$$

其中  $\alpha_i^t$  表示目标词  $t$  对单词  $i$  的注意力系数， $c_t$  为信息聚合后的语义向量， $W_1$  为权重矩阵，下同。

聚合后的语义向量  $c_t$  向量与目标词语义向量  $h_t$  得到最终的目标词语义表示  $r_t$ ，并结合图网络编码得到的框架语义表示  $G_f$  进行框架识别，公式描述如下：

$$r_t = \text{Relu}(W_2 \cdot [h_t; c_t]) \quad (10)$$

$$P(f|t) = \text{softmax}(W_3 \cdot r_t + G_f \cdot r_t) \quad (11)$$

### 3.5 联合训练设置

在训练阶段，我们对模型各部分模块参数进行联合优化，主要包括框架识别损失 $L_{frame}$ ，第一、二阶段的角色位置识别损失 $L_{span}^{first}$ 和 $L_{span}^{second}$ ，和角色类别识别损失 $L_{role}$ ，各部分的损失计算公式如下：

$$L_{frame} = \frac{1}{|k|} \sum_{i=1}^k \log(P(\hat{f}_i | S, t_i)) \quad (12)$$

$$L_{span}^{first} = \frac{1}{|k|} \frac{1}{|length|} \sum_{n=1}^{length} \sum_{i=1}^k \log(p(\hat{g}_n^{t_i} | t_i, S)) \quad (13)$$

$$L_{span}^{second} = \frac{1}{|k|} \sum_{i=1}^k \sum_{\tau=1}^m \log(P(\hat{sta}_\tau | S, t_i, f_i)) + \frac{1}{|k|} \sum_{i=1}^k \sum_{\tau=1}^m \log(P(\hat{end}_\tau | S, t_i, f_i)) \quad (14)$$

$$L_{role} = \frac{1}{|k|} \sum_{i=1}^k \sum_{\tau=1}^m \log(P(\hat{r}_\tau | S, t_i, f_i)) + \frac{1}{|k|} \sum_{i=1}^k \log(P(r_{None} | S, t_i, f_i)) \quad (15)$$

由于各部分模块的训练难度与参数收敛速度不同，我们加入损失调节系数用于调整模型训练的梯度优化方向，平衡各部分模块的训练收敛速度，公式描述如下：

$$L = \alpha L_{frame} + \beta L_{span}^{first} + \gamma L_{span}^{second} + \delta L_{role} \quad (16)$$

在本文模型实验中，调节系数 $\alpha$ 、 $\beta$ 、 $\gamma$ 和 $\delta$ 分别取0.1、0.1、0.3和0.3。

我们采用预训练的BERT语言模型作为文本的编码器，输出向量的维度为768。框架关系表示模块采用两层R-GCN作为图编码器。在训练集上训练100轮，并保存在验证集上取得最优结果的模型参数。训练的Batch大小为4，选用的优化器为BertAdam，学习率为 $3 \times 10^{-5}$ 。

## 4 数据集与评测指标

### 4.1 数据集

目前，针对框架语义分析任务有两个版本的数据集，分别为FN1.5和FN1.7，这两个数据集都来源于框架网络语料库中的篇章标注语料。之前的框架语义分析模型主要是在FN1.5版本的数据集上进行评测，因此我们对比我们模型与现有模型在FN1.5数据集上的结果。

FN1.5包含了1019种框架类型，9634种框架角色。在FN1.5篇章标注语料中，篇章中的每个句子可能包含多个框架以及它们对应的一些角色。之前的工作将篇章语料划分成单句形式，并按照一句话一个目标词进行实例的切分，它们的模型也可以看作是针对于单目标词的框架语义分析模型。而我们模型能够同时处理一个句子中的所有目标词。此外FN还包含15万句的示例句子，每个示例句子只标注了单个目标词的框架语义结构。先前的一些工作会将这部分数据加入到模型的训练语料中，用于提升模型的表现。

### 4.2 评测指标

评测指标包括框架识别准确率和框架结构提取。分别用于评估模型在框架识别任务和框架语义分析整体任务的表现。

对于框架识别任务的评测，先前研究工作将能触发多个框架的目标词称为歧义(Ambiguous)目标词，并单独对歧义目标词这一集合进行评测。参照之前的相关研究，本文模型同样在所有目标词集合(All)和歧义目标词(Amb)集合上进行评测。

框架语义结构提取是一项评价框架语义分析整体性能的指标。该指标将框架本身视作特殊的框架角色，与其他框架角色统一计算精确率(Precision)、召回率(Recall)和F1值。此外该目标要求框架角色位置边界的精确匹配，但对于位置识别正确但类型识别错误的角色，会根据预测角色和正确角色在框架网络上的关联程度相应给与部分分数。该评测方式由Das[32]提出，并成为框架语义分析任务主流的评价指标。

## 5 实验结果与分析

我们在FN1.5测试集上对模型的框架识别准确率和整体框架分析表现进行评估。先前工作的模型指标数据来源于它们的论文。之前一些模型(Yang and Mitchell, 2017a; Peng et al., 2018a)采用集成学习策略, 通过多个相同或者不同架构的模型进行专家投票的方式对框架和框架角色识别, 从而进一步提升模型的整体表现。这里为公平衡量模型的真实性能, 只报告单模型的指标表现。

| Model                    | All  | Ambiguous |
|--------------------------|------|-----------|
| SEMAFOR(2010)            | 83.6 | 69.2      |
| Open-SESAME(2017)        | 87.0 | -         |
| Hartmann(2017)           | 87.6 | 73.8*     |
| Yang and Mitchell(2017a) | 88.2 | 75.7*     |
| Hermann(2014)            | 88.4 | 73.1      |
| Peng(2018a)(BASIC)       | 89.2 | 76.3      |
| Chen(2021)               | 89.4 | 76.7      |
| Chen (Bert)(2021)        | 90.5 | 79.1      |
| Jiang and Riloff(2021)   | 91.3 | 81.0      |
| Su et al(2021)           | 92.1 | 82.3      |
| Our model                | 91.9 | 82.1      |

Table 1: 各模型在FN1.5上的框架识别准确率。其中\*表示它们模型的歧义目标词集合与其他模型不同

### 5.1 实验结果

框架识别的结果如Table 1所示。其中第一行表示的是现有的框架语义分析模型, 能完整处理框架识别和角色识别任务。第二行表示的是单独的框架识别模型, 它们只针对框架识别任务进行优化。可以看到我们的模型超过了之前的框架语义分析模型, 接近目前在框架识别任务上表现最好的独立框架识别模型(Su et al., 2021)。我们分析认为, 我们模型之所以与目前最优的框架识别模型存在差距, 是因为我们模型在训练阶段对框架分析的各项子任务采用联合优化策略, 保存在综合指标上表现最好的模型。然而, 在训练阶段, 在框架识别任务和角色识别任务的优化上呈现不同步的情况。由于角色识别任务相比于框架识别任务训练难度更大, 因此需要更多的训练轮数。在整个训练过程中, 框架识别的训练通常会提前收敛, 最终导致过拟合的情况发生。虽然我们通过加入平衡系数在一定程度上缓解了收敛不同步问题, 但没有从根本上解决。

框架语义结构提取的结果如Table 2所示。其中, 第一行和第二行分别表示基于流水线策略和基于联合训练策略的框架语义分析模型, 同时也是针对单目标词的框架语义分析模型。†表示它们模型加入额外语料(例如15万句的示例句子)进行训练。除了陈的Bert版本的模型外, 我们模型在不加入额外语料的情况下, 显著高于其他所有框架语义分析模型。考虑到陈的Bert版本模型的效果提升一部分是通过加入15万句的示例句子得到的, 为消除这部分因素带来的模型性能增益, 将他们的模型仅在标准FN1.5训练集上进行训练, 最终得到它们的F1值为76.4。说明我们模型在不无外部数据依赖的情况下提取的框架语言结构更优。

### 5.2 消融实验

我们设置了两组消融实验探究模型设计的有效性: 一、框架关系知识的融入对框架语义分析任务的作用; 二、子任务的双向交互对对框架语义分析任务的作用。相应实验组设置如下, 用框架语义结构提取的指标在FN1.5测试集上进行评估。

| Model                            | P    | R    | F1   |
|----------------------------------|------|------|------|
| SEMAFOR(2010)                    | 69.2 | 65.1 | 67.1 |
| Framat(2010)                     | 71.1 | 63.7 | 67.2 |
| Framat+context(2015)             | 71.1 | 64.8 | 67.8 |
| Open-SESAME(2017)                | 71.0 | 67.8 | 69.4 |
| FitzGerald et al(2015b)          | 74.8 | 65.5 | 69.9 |
| Yang and Mitchell (2017b) (SEQ)† | 69.6 | 70.9 | 70.2 |
| Yang and Mitchell (2017b)(REL)†  | 77.1 | 68.7 | 72.7 |
| Peng (2018b)(BASIC)†             | 79.2 | 71.7 | 75.3 |
| Chen (2021)†                     | 75.1 | 76.9 | 76.0 |
| Chen (Bert) (2021)†              | 78.2 | 82.4 | 80.2 |
| Chen (Bert,w/o exemplar)(2021)   | 75.1 | 77.6 | 76.4 |
| Our Model                        | 76.2 | 80.8 | 78.5 |

Table 2: 各模型在FN1.5上的框架结构提取结果。

**实验组一（消去框架关系知识表示模块）**：移除框架关系图表示模块，框架和框架角色的特征向量由随机初始化生成。

**实验组二（消去角色-框架的信息传递）**：去除第一阶段框架角色识别的损失计算，角色识别模块向框架识别模块的信息传递不受监督。

从Table 3中可以看到实验组一和实验组二相比原模型在各项指标上都有不同程度的下降，证明了扩展后的框架关系知识和框架与角色识别的交互对框架语义分析任务具有正面作用。此外发现，实验一在各项指标上相比原模型都有较为明显的下降，证明了扩展了框架角色后的框架关系信息对模型的预测具有较为关键的指导作用。

| Model | P    | R    | F1   |
|-------|------|------|------|
| 我们模型  | 76.2 | 80.8 | 78.5 |
| 实验组一  | 75.1 | 78.1 | 76.6 |
| 实验组二  | 75.5 | 80.5 | 77.9 |

Table 3: 消融实验结果。

### 5.3 模型存在的局限性

从模型的训练过程和实验结果中，我们发现并总结了目前模型存在的一些问题。在未来工作中，我们将从这些问题入手进行研究，对模型作进一步的优化。

#### 5.3.1 子任务训练不同步

从模型的训练过程中发现，框架语义分析中的各项子任务的训练难度是不同的，导致它们在相同学习率下收敛所需要的训练轮数不同，这对于采用联合训练策略的框架语义分析模型而言是一个挑战。其中，框架识别任务相对简单，所以框架识别模块的参数收敛较快，但此时角色位置和角色类别的识别能还远未达到最优。为缓解这一问题，我们在各部分模块的损失计算前加入平衡系数，用于调节子任务的收敛速度。然而这种策略仍存在两个问题。一是调节系数是固定的，无法随着模型的训练进行动态调节。二是调节系数的设置依赖人为经验，增加了模型调参的难度。因此，我们目前的模型仍无法在训练阶段同时在框架识别准确率和框架语义结构提取的综合指标上达到最优。



### 5.3.2 示例语料的利用问题

FN1.5数据集包含15万句的示例语料，这部分语料每个句子只包含一个目标词的框架语义结构信息。先前的研究工作(Yang and Mitchell, 2017a; Peng et al., 2018a; Chen et al., 2021)会将这部分数据作为额外的训练语料加入到模型的训练当中。它们在实验中发现，加入这部分数据会使模型的框架语义结构提取的表现提升3到4个百分点。然而，我们在实验中发现，我们的模型加入示例语料作为额外语料并未带来明显提升（在验证集上提升不到1个百分点）。经过分析认为，先前的模型都是针对单目标词的框架语义分析模型，示例句子的输入形式刚好符合模型的输入。然而我们模型采用多目标词的输入形式，标准训练集和示例语料在模型看来是两种异质的语料文本，在模型的训练过程中可能会产生冲突，从而导致对示例语料的利用不够充分。在未来工作中，我们将针对示例语料的处理方式进行研究，例如在模型训练前对示例语料进行自动标注，得到形式与标准训练集更为接近但有噪声的数据，标注后的数据更适合多目标词模型的训练。

## 6 总结

本文提出了一个面向多目标的框架语义分析模型，实现对多目标词的联合预测。该模型对框架语义分析的各项子任务进行交互性建模，实现子任务间的双向交互。此外，本文利用关系图网络对框架关系信息进行编码，将其作为框架语义学知识融入模型中。实验表明，本文模型在不借助额外语料的情况下仍具有良好的表现，消融实验证明了本文模型设计的有效性。此外本文分析了我们模型存在的一些问题，将在未来工作中针对这部分问题对模型作进一步的优化。

## 致谢

本文工作得到国家自然科学基金（61936012, 61876004）支持，特此致谢。

## 参考文献

- Collin F Baker, Michael Ellsworth, and Katrin Erk. 2007. Semeval-2007 task 19: Frame semantic structure extraction. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 99–104.
- Anders Björkelund, Bernd Bohnet, Love Hafdell, and Pierre Nugues. 2010. A high-performance syntactic and semantic dependency parser. In *Coling 2010: Demonstrations*, pages 33–36.
- Desai Chen, Nathan Schneider, Dipanjan Das, and Noah A Smith. 2010. Semafor: Frame argument resolution with log-linear models. In *Proceedings of the 5th international workshop on semantic evaluation*, pages 264–267.
- Xudong Chen, Ce Zheng, and Baobao Chang. 2021. Joint multi-decoder framework with hierarchical pointer network for frame semantic parsing. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 2570–2578, Online, August. Association for Computational Linguistics.
- Dipanjan Das, Nathan Schneider, Desai Chen, and Noah A Smith. 2010. Probabilistic frame-semantic parsing. In *Human language technologies: The 2010 annual conference of the North American chapter of the association for computational linguistics*, pages 948–956.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: pre-training of deep bidirectional transformers for language understanding. In Jill Burstein, Christy Doran, and Thamar Solorio, editors, *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, pages 4171–4186. Association for Computational Linguistics.
- Nicholas FitzGerald, Oscar Täckström, Kuzman Ganchev, and Dipanjan Das. 2015a. Semantic role labeling with neural network factors. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 960–970, Lisbon, Portugal, September. Association for Computational Linguistics.

- Nicholas FitzGerald, Oscar Täckström, Kuzman Ganchev, and Dipanjan Das. 2015b. Semantic role labeling with neural network factors. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 960–970.
- Daniel Gildea and Daniel Jurafsky. 2002. Automatic labeling of semantic roles. *Comput. Linguistics*, 28(3):245–288.
- Shaoru Guo, Ru Li, Hongye Tan, Xiaoli Li, Yong Guan, Hongyan Zhao, and Yueping Zhang. 2020. A frame-based sentence representation for machine reading comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 891–896, Online, July. Association for Computational Linguistics.
- Silvana Hartmann, Iliia Kuznetsov, Teresa Martin, and Iryna Gurevych. 2017. Out-of-domain FrameNet semantic role labeling. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, pages 471–482, Valencia, Spain, April. Association for Computational Linguistics.
- Karl Moritz Hermann, Dipanjan Das, Jason Weston, and Kuzman Ganchev. 2014. Semantic frame identification with distributed word representations. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1448–1458, Baltimore, Maryland, June. Association for Computational Linguistics.
- Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation*, 9(8):1735–1780.
- Tianyu Jiang and Ellen Riloff. 2021. Exploiting definitions for frame identification. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 2429–2434, Online, April. Association for Computational Linguistics.
- Richard Johansson and Pierre Nugues. 2007. LTH: Semantic structure extraction using nonprojective dependency trees. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 227–230, Prague, Czech Republic, June. Association for Computational Linguistics.
- Hao Peng, Sam Thomson, Swabha Swayamdipta, and Noah A. Smith. 2018a. Learning joint semantic parsers from disjoint data. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1492–1502, New Orleans, Louisiana, June. Association for Computational Linguistics.
- Hao Peng, Sam Thomson, Swabha Swayamdipta, and Noah A. Smith. 2018b. Learning joint semantic parsers from disjoint data. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1492–1502.
- Michael Roth and Mirella Lapata. 2015. Context-aware frame-semantic role labeling. *Transactions of the Association for Computational Linguistics*, 3:449–460.
- Sunita Sarawagi and William W Cohen. 2004. Semi-markov conditional random fields for information extraction. *Advances in neural information processing systems*, 17.
- Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. 2008. The graph neural network model. *IEEE transactions on neural networks*, 20(1):61–80.
- Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne Van Den Berg, Ivan Titov, and Max Welling. 2018. Modeling relational data with graph convolutional networks. In *European semantic web conference*, pages 593–607. Springer.
- Dan Shen and Mirella Lapata. 2007. Using semantic roles to improve question answering. In *Proceedings of the 2007 joint conference on empirical methods in natural language processing and computational natural language learning (EMNLP-CoNLL)*, pages 12–21.
- Xuefeng Su, Ru Li, Xiaoli Li, Jeff Z. Pan, Hu Zhang, Qinghua Chai, and Xiaoqi Han. 2021. A knowledge-guided framework for frame identification. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 5230–5240, Online, August. Association for Computational Linguistics.

- Swabha Swayamdipta, Sam Thomson, Chris Dyer, and Noah A Smith. 2017. Frame-semantic parsing with softmax-margin segmental rnns and a syntactic scaffold. *arXiv preprint arXiv:1706.09528*.
- Jason Weston, Samy Bengio, and Nicolas Usunier. 2011. Wsabie: Scaling up to large vocabulary image annotation.
- Bishan Yang and Tom Mitchell. 2017a. A joint sequential and relational model for frame-semantic parsing. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1247–1256, Copenhagen, Denmark, September. Association for Computational Linguistics.
- Bishan Yang and Tom Mitchell. 2017b. A joint sequential and relational model for frame-semantic parsing. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1247–1256.

JCL 2022