

# CCL24-Eval 任务2系统报告：基于关系抽取的中文意合图语义解析方法研究

霍虹颖, 黄少平, 刘鹏远  
北京语言大学, 信息科学学院, 北京  
202111680974@stu.blcu.edu.cn  
www17379430207@163.com  
liupengyuan@pku.edu.cn

## 摘要

意合图是以事件为中心的单根有向语义表征图, 在语义计算与应用方面具有重要价值。在CCL-2024中文意合图语义解析评测任务中, 为克服意合图为单根有向图、意合图包含隐性事件词以及意合图的语义关系类型十分丰富, 导致关系类型过多等诸多方面的难点, 本文提出一种将该任务转换为关系抽取的方法。该方法首先对标签进行扩充, 分为正向标签和反向标签; 其次, 对输入进行扩充, 将隐性事件词添加到输入中, 无须额外对隐性事词进行预测; 最后, 细分为不带隐性事件词和带隐性事件词的关系抽取任务。实验结果表明, 本文方法在官方盲测集上的F1值为64.44%, 高出基线模型33.41%, 证明了本文方法的有效性。

**关键词:** 关系抽取; 语义分析; roBERTa; 意合图

## System Report for CCL24-Eval Task 2:A Study on Semantic Parsing Method of Chinese-Parataxis-Graph-Parsing Based on Relational Extraction

Hongying Huo,Shaoping Huang,Pengyuan Liu  
School of Information Science, Beijing Language and Culture University, Beijing  
202111680974@stu.blcu.edu.cn  
www17379430207@163.com  
liupengyuan@pku.edu.cn

## Abstract

Chinese Parataxis Graph is an event-centered single-root directed semantic representation graph, which is of great value in semantic computation and application. In the task of semantic analysis and evaluation of the Chinese Parataxis Graph in CCL-2024, in order to overcome the difficulties that the Chinese Parataxis Graph is a single-rooted directed graph, contains implicit event words, and its semantic relation types are very rich, resulting in too many relation types. This paper proposes a method to transform this task into relational extraction. In this method, the labels are expanded into forward labels and reverse labels. Secondly, the input is expanded by adding the implicit event words to the input, and there is no need to predict the implicit event words. Finally, it is subdivided into relational extraction tasks without implicit event words and with implicit event words. Experimental results show that the F1 value of the proposed method on the official blind test set is 64.44%, which is 33.41% higher than the baseline model, which proves the effectiveness of the proposed method.

**Keywords:** Relational Extraction, Semantic analysis, roBERTa, Chinese Parataxis Graph

## 1 引言

意合图是围绕事件的语义表征，事件词是事件的核心表达，事件词关联起事件内的各个成分，在事件间建立关系时，以事件词代表整个事件。意合图由事件结构和实体结构构成，均有内外结构之分，内部结构是构成要素间的关系，外部结构是整体间的关系。在CCL2024-CPG数据集中，定义了包含论元结构、关系事件、特殊标签等7个标签层级。例如，在句子“他哭肿了眼睛”中所包含的语义标签有：(他,哭,A0), (眼睛,肿,A0), (他,眼睛,EntityRel), (了,哭,Time), (了,肿,Time), (哭,因果关系,原因事件), (肿,因果关系,结果事件), (哭,ROOT,CoreWord)。将其看作关系抽取任务，则前两个元素相当于关系抽取中需要识别出来的实体，第三个元素则为这两个实体之间的关系。比如“他”和“哭”之间存在的语义关系为A0，A0代表着论元结构中的主体关系，即实体“他”和“哭”存在着动作发起的主体关系。

意合图理论的提出为解决自然语言处理中的语义难题提供了一种全新的思路，然而这一新兴领域，目前尚无成熟的研究方法和标注策略，本文的工作主要对意合图语义框架的识别提出了一种研究思路，我们通过关系抽取的方法，对语义标签识别并分类标注。

意合图语义分析与普通的关系抽取任务有所不同。意合图为单根有向图，意味着意合图的两个实体的先后位置固定，不可任意交换，实体在三元组中的第一个位置和第二个位置的标签含义并不相同。其次，意合图包含隐性事件词，例如当两个实体是并列关系，与其他事件词构成语义关系时，以“And”标签来指代这两个并列实体这一整体，“And”这样的标签就是隐性事件词，它并未在原始分词后的句子中出现。除此之外，意合图整体体系丰富，语义关系标签众多，这也为关系抽取带来了一定难度。

关系抽取是信息抽取的一项重要子任务，旨在从非结构化文本中识别并提取实体之间的语义关系，主要包括实体识别和关系分类。实体识别需要识别文本中的命名实体，如人名、地名等，关系分类则是判断实体之间是否存在某种特定的关系，并确定关系的类型，在涉及到更复杂的句子结构时，还需要识别事件及其参与者。通过我们的分析发现，意合图语义解析可以认为是一种关系抽取任务，即识别出句子中存在有向关系的实体并进行标签分类。

我们充分利用基于roBERTa的关系抽取方法，来从意合图中解析出事件与实体的语义关系。主要使用的是chinese-roBERTa-wwm-ext-large模型，利用roBERTa模型对文本进行编码，识别出文本中的实体，采用roBERTa模型对任意两个实体之间的关系进行识别和分类。与传统的基于特征工程的方法相比，我们的方法不需要手工设计特征，而是直接从原始文本中学习特征表示，具有更好的泛化能力和适应性。

总的来说，本文采用的方法是将意合图语义框架识别转换成关系抽取任务，针对意合图的单根有向性、包含隐性事件词、语义标签丰富的问题对数据进行了特殊处理，在评估中，我们的模型在盲测集上的F1得分为64.44%，在基线依存模型的基础上提升了33.41%。在chinese-roBERTa-wwm-ext-large模型的基础上，我们也尝试使用大模型进行标签标注，我们考虑到通过输入大模型能够理解的语义知识并提供示例的方法，来让大模型学习不同的语义标签并对句子进行标注。整合两种方法的结果最终得分为64.42%。

## 2 相关工作

在关系抽取领域的早期研究中，学者们主要探索了基于规则和基于监督学习的方法。基于规则的方法主要根据设定的规则以及固定模式来识别实体之间的关系，这种方法通常受限于规则的覆盖范围和适用性。

关系抽取的过程需要依赖大量标注的训练样本，随着深度学习技术的发展，传统的关系抽取方法逐渐被基于深度神经网络的关系抽取方法所取代。这样的方法可以改善传统方法中的人工选择部分，有效改善特征在抽取过程中的误差累积。监督学习的关系抽取模型以Zeng et al. (2014)提出的卷积神经网络CNN和Lin et al. (2018)提出的循环神经网络RNN为两大代表，Gormley et al. (2015)提出了一种基于因子的组合嵌入模型(FCM),该模型用依存树和命名实体从单词嵌入构建句子级和子结构嵌入。dos Santos et al. (2015a)提出了排序CNN (CR-CNN)，通过一个新的成对排名损失函数，减少人工类的影响，该方法比使用CNN+softMax classifier的模型更有效。Shen and Huang (2016)将CNN编码器与句子表示结合使用，该句子表示通过关注目标实体和句子中的单词之间的注意力对单词进行加权，以提高性能。Wang et

al. (2016) 为了捕获异构上下文中的模式以对关系进行分类, 提出了一种具有两层注意力级别的卷积神经网络体系结构。Lee et al. (2019) 开发了一个端到端的循环神经网络模型, 该模型结合了实体感知的注意力机制和潜在的实体类型以进行关系分类。然而训练数据的质量将大幅影响这两种模型的性能, 即模型依赖大量高质量的人工标注数据进行模型训练。之后Mintz et al. (2009) 等人提出了远程监督的关系抽取方法, 该方法通过将数据自动对齐远程知识库来解决大量无标签数据的自动标注问题, 在一定程度上能减少对人工的依赖。利用少量标注的样本进行学习也是一种方案。少样本学习属于监督学习的范畴, 通过给出每个类的少量标注信息的实例, 使得程序在处理该任务时能实现性能的提升, 在训练过程中不改变已有训练好的模型, 借助每类少数几个标注样本就可以快速学习并完成新类的分类。

这些方法利用深度神经网络模型来自动学习文本中的特征表示, 并在端到端的方式下进行关系抽取。尤其是Transformer模型的出现, 特别是其变种roBERTa的问世, 为关系抽取任务带来了新的突破, roBERTa模型通过预训练大规模语料库, 能够捕捉文本中丰富的语义信息, 使得关系抽取模型能够更好地理解文本语境, 从而提高了关系抽取的性能。Wu and He (2019) 提出了一个模型, 既利用预训练的bert语言模型, 又结合来自目标实体的信息来完成关系分类。基于R-BERT-CNN模型的实体关系抽取论文将实体级信息融入预训练模型获取目标实体的语义; 采用CNN提取句子级的语义信息; 连接句子向量、标签向量和目标实体向量, 获得全局信息; 通过softmax分类器抽取实体关系。

相较于GPT-3这样的超大型模型, roBERTa的参数规模较小, 因此在训练和推理时需要的计算资源更少。roBERTa模型通过优化模型结构和训练方法, 提高了计算效率, 在训练过程中使用了动态掩码的技术, 可以更好地利用计算资源, 从而加快速度, 所以在同样的数据集和硬件条件下, 相较于诸如GPT-3这样的模型, 训练roBERTa模型所需的时间更短。roBERTa在BERT的基础上进行了一系列训练技巧和超参数的调整, 以提升了模型在文本分类、命名实体识别等自然语言处理任务上的性能, 这也就意味着在相同的任务和数据集上, 使用roBERTa模型可以获得更好地表现而不需要使用更大的模型。

意合图首次参与评测活动, 前人的研究主要是在理论与资源构建方面的工作 (郭梦溪 et al., 2024a)。对于意合图语义分析, 目前没有太多可参考的方法。但是意合图属于语义分析的一种 (郭梦溪 et al., 2024b), 所以可以借鉴其他语义分析的计算方法, 比如抽象语义表示、语义依存等。这些方法在处理复杂的语义关系和结构方面已经取得了显著成果, 可以为意合图的进一步研究提供有价值的参考。本文采用的基线模型即为依存模型。

### 3 基于关系抽取的中文意合图语义解析方法

通过分析CCL2024-CPG数据集, 我们已经发现每条数据的输入与输出与传统的关系抽取任务十分类似, 抽取结果都为三元组 $\{word_1, word_2, Relval\}$ 。但相比之下, 意合图语义框架的难点在于: 首先, 中文意合图是单根有向图, 即 $\{word_1, word_2, Relval\}$ 的顺序不能随意更换。其次, 该任务中需要先识别“实体对”再对关系进行分类, 同时, 该任务需要识别的“实体对”不仅会出现在输入中, 由于意合图包括事件外结构, 所以还有隐性事件词需要额外预测。最后, 该数据集的关系标签分布也不平衡, 因为需要考虑“实体对”的顺序, 关系标签会加倍, 从而标签不平衡的现象更加明显。

为了解决上述问题, 我们分别提出了以下解决方法:

1. **针对 $word_1, word_2$ 的顺序问题:** 我们将分类标签加倍,  $word_1, word_2$ 在输入中的位置为从前到后, 则为正常标签, 反之,  $word_1, word_2$ 的位置在输入中为从后到前则为反标签, 记作“标签\_reverse”。
2. **针对隐性事件词的预测:** 我们对输入进行扩充, 转化为传统的关系抽取任务, 即我们整理了所有可能出现的隐性事件词, 将其全部以特定的格式“<隐性事件词>”添加到输入末尾。这时, 隐性事件词也出现在输入中, 不需要进行额外的预测, 从而转化成关系抽取任务。图 3是一个例子, 我们将所有隐性事件词加到原始句之后, 同时对标签中与前后顺序相反的标签添加“\_reverse”。
3. **针对标签不平衡的问题:** 考虑到标签加倍后, 会影响模型对其他标签的预测性能, 我们将任务划分成了两个子任务, 即不包含隐性事件词的关系抽取和包含隐性事件词的关系抽取。

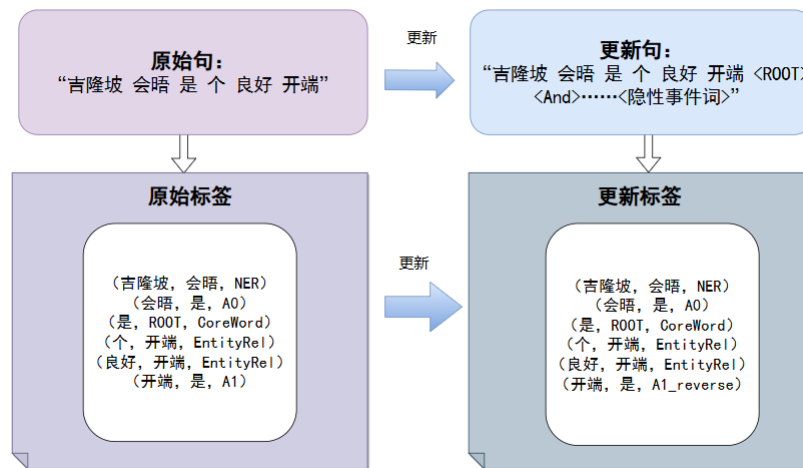


图 1: 展示原始句子标签和更新的句子标签

为了解决顺序问题和隐性事件词的出现，我们对数据进行预处理，以预处理后实体对是否包含有隐性事件词的标签为依据，划分为两个子任务。我们对这两个子任务分别使用对称的关系识别模型和关系分类模型（关系识别模型和关系分类模型均使用chinese-roBERTa-wwm-ext-large模型），从而得到两个子任务在关系抽取方法下的结果，通过对两个结果的整合，从而来得到最终的意合图语义分析结果。图2是该方法的整体流程图。

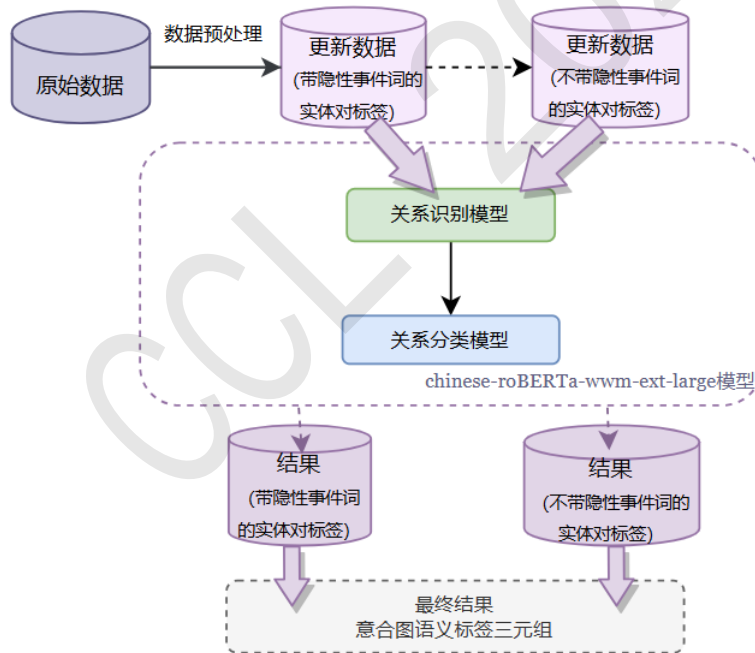


图 2: 整体流程图

首先，对于关系识别模型，我们将原数据集中的不带隐性事件词的实体对和带隐性事件词的实体对分别整理成两个数据集，再把两个数据集分别按7:3的比例划分成训练集和验证集。我们先对预处理后的数据进行关系识别，每条数据格式如表1所示，包含实体1、实体2、标签和句子。标签有类“yes”和“no”。“yes”代表两个实体之间有关系，“no”代表两个实体之间没有关系。

其次，对于关系分类模型，由于不带隐性事件词的实体对和带隐性事件词的实体对的关系

实体1	实体2	标签	句子
今天	发挥	yes	但今天中国队发挥较好，因此获得了胜利。
和	两端	no	佩雷斯说：“中国和以色列地处亚洲的两端，相隔万水千山。”

表 1: 关系识别数据集样例

类型有重叠，但大部分不一样，所以我们分别整理了所有可能出现的关系类型，加倍后，前者共64种，后者共94种。同时，所有可能出现的隐性事件词数量为13种，这些隐性事件词以“<隐性事件词>”的格式扩充到输入中。具体见表2。最后，将所有含答案数据按照不带隐性事件词的关系类型和带隐性事件词的关系类型分别整理成两个数据集，同样按7:3的比例划分成训练集和验证集。每条数据的格式与表1相同，标签部分为预处理后的更改标签。

chinese-roBERTa-wwm-ext-large模型的架构包括三个主要部分：roBERTa编码层、实体特征处理层和分类层。

### 3.1 roBERTa编码层

首先，将输入句子中的token通过roBERTa模型编码，得到每个token的隐藏状态和句子级别的表示（即[CLS] token的表示）。

假设输入句子表示为 $X = \{x_1, x_2, \dots, x_n\}$ ，其中 $n$ 为句子的长度。通过roBERTa模型编码后，得到每个token的隐藏状态 $H = \{h_1, h_2, \dots, h_n\}$ 和句子级别的表示 $h_{[CLS]}$ ：

$$H, h_{[CLS]} = \text{roBERTa}(X) \quad (1)$$

其中， $H \in \mathbb{R}^{n \times d}$ ， $h_{[CLS]} \in \mathbb{R}^d$ ， $d$ 为隐藏状态的维度。

### 3.2 实体特征处理层

对于关系识别模型中存在关系的两个实体，通过对实体的所有token的隐藏状态取平均值来表示实体的特征。具体来说，假设实体 $e_1$ 的token索引范围为 $[i, j]$ ，则实体 $e_1$ 的特征表示为：

$$e_1 = \frac{1}{j - i + 1} \sum_{k=i}^j h_k \quad (2)$$

为了简化计算，我们定义实体掩码 $e_1^{\text{mask}} \in \{0, 1\}^n$ ，当且仅当 $k \in [i, j]$ 时， $e_1^{\text{mask}}[k] = 1$ 。则上述平均操作可以表示为：

$$e_1 = \frac{e_1^{\text{mask}} \cdot H}{\sum_{k=1}^n e_1^{\text{mask}}[k]} \quad (3)$$

同理，实体 $e_2$ 的特征表示为：

$$e_2 = \frac{e_2^{\text{mask}} \cdot H}{\sum_{k=1}^n e_2^{\text{mask}}[k]} \quad (4)$$

为了增强实体特征的代表能力，将其通过一个线性层和非线性激活函数处理：

$$e'_1 = \tanh(W_e e_1 + b_e) \quad (5)$$

$$e'_2 = \tanh(W_e e_2 + b_e) \quad (6)$$

其中， $W_e \in \mathbb{R}^{d \times d}$ ， $b_e \in \mathbb{R}^d$ 。

### 3.3 分类层

将句子的[CLS]表示与两个实体的特征拼接起来，形成最终的特征向量：

$$h_{\text{concat}} = [h_{[\text{CLS}]}, e'_1, e'_2] \quad (7)$$

其中， $h_{\text{concat}} \in \mathbb{R}^{3d}$ 。

为了提高模型的泛化能力，对拼接后的特征向量进行层归一化和dropout处理：

$$h_{\text{concat\_norm}} = \text{LayerNorm}(h_{\text{concat}}) \quad (8)$$

$$h_{\text{concat\_drop}} = \text{Dropout}(h_{\text{concat\_norm}}) \quad (9)$$

最后，通过一个线性层将特征向量映射到关系类别空间：

$$\text{logits} = W_{\text{out}} h_{\text{concat\_drop}} + b_{\text{out}} \quad (10)$$

其中， $W_{\text{out}} \in \mathbb{R}^{k \times 3d}$ ， $b_{\text{out}} \in \mathbb{R}^k$ ， $k$  为关系类别的数量。

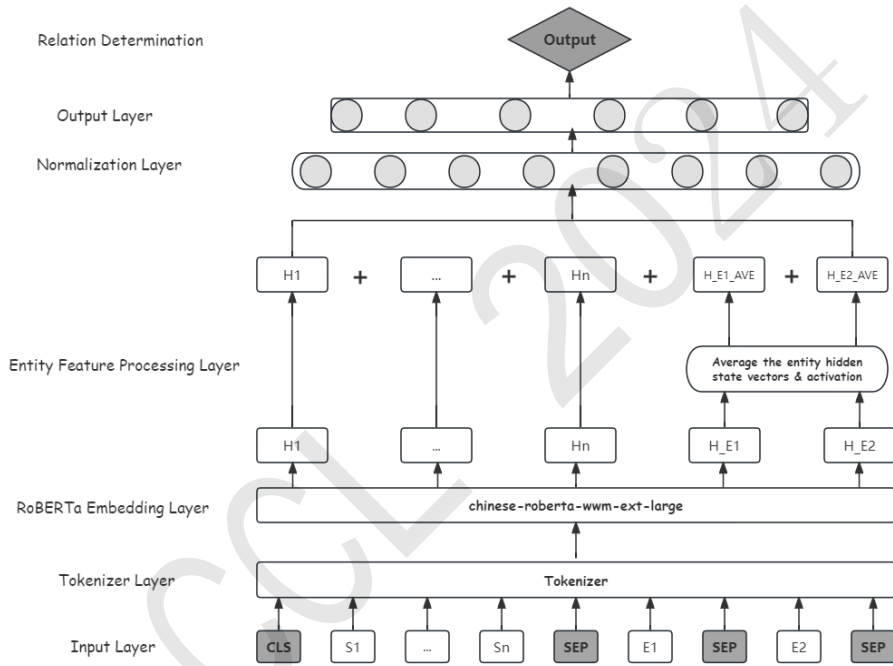


图 3: 模型框架图

通过以上三个主要部分的协同工作，chinese-roBERTa-wwm-ext-large模型能够有效地进行关系抽取任务，从而识别出文本中不同实体之间的关系。

## 4 实验

### 4.1 实验数据与评价方法

本文实验数据来自CCL2024-CPG数据集，包含来自国际中文教育阅读语料和中文宾州树库的新闻语料。每条数据包含分词后的句子、意合图语义标签、 $word_1, word_2$ 的idx，即特定序号等标注信息。本文使用数据集中的训练集和验证集共3000句语料来训练模型。本次评测采用F1值作为模型表现的评价标准，其计算方式如下：

$$Precision = \frac{\text{Matching Tuples}}{\text{Generated Tuples}} \quad (11)$$

$$Recall = \frac{Matching\ Tuples}{Gold\ Tuples} \quad (12)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (13)$$

其中，Generated Tuples 为模型预测的三元组集合数，Gold Tuples 为测试集/盲测集的三元组集合数，Matching Tuples 为模型预测的三元组集合与测试集/盲测集的三元组集合间的最大匹配个数。

## 4.2 实验参数设置

本文采用哈工大版本的chinese-roBERTa-wwm-ext-large模型，实验的相关参数设置如下表所示。

参数名	参数值
attention_probs_dropout_prob	0.1
bos_token_id	0
directionality	bidi
eos_token_id	2
hidden_act	gelu
hidden_dropout_prob	0.1
hidden_size	1024
initializer_range	0.02
intermediate_size	4096
layer_norm_eps	1e-12
max_position_embeddings	512
model_type	bert
num_attention_heads	16
num_hidden_layers	24
output_past	true
pad_token_id	0
pooler_fc_size	768
pooler_num_attention_heads	12
pooler_num_fc_layers	3
pooler_size_per_head	128
pooler_type	first_token_transform
type_vocab_size	2
vocab_size	21128

表 2: 实验参数设置表

## 4.3 实验结果

经过10轮的迭代训练后，关系识别模型在不含隐性事件词的任务中F1值最佳时为0.9653，关系识别模型在含隐性事件词的任务中F1最高值为0.9840。尽管F1值很高，但实际在盲测集进行关系识别时，是对所有实体进行一一枚举，让模型判断是否有关系，输出“yes”或者“no”，如果直接利用训练好的模型进行推理，包含隐性事件词的数据集预测的实体对会过多。因此我们对预测结果向量设置了一个最小阈值，该向量为二维向量，分别表示两个标签的置信度，由于该二维向量未经过Softmax以及归一化成概率，故将其称为logits，只有“yes”标签对应的logit大于2时，预测结果为“yes”的实体对才被模型输出，否则无效，其中最小阈值是一个超参数，经过多轮实验对比，最小阈值为2时，实验结果最优。

关系分类仍然利用基于roBERTa的关系抽取模型，为了缓解标签过多，造成对模型性能的影响，我们训练了两个关系分类模型，相对而言，由于不带隐性事件词的实体对数量多于带

隐性事件词的实体对数量，且后者中个别标签数量太少，所以从Marco F1值来说，不带隐性事件词的实体的关系分类模型整体性能较好，但从Weighted F1来看，带隐性事件词的关系分类模型性能好于不带隐性词的关系分类模型性能，前者为F1为0.9524，后者为0.9287。最后，对两个子任务的结果进行整合，对更新后的标签进行还原，并增加位置属性idx，得到最终结果，按照评价方法计算在盲测集上得到F1得分为64.44%。F1值在整合前高是由于在关系识别阶段，“yes”和“no”标签分布不均衡，大部分分类结果都是“no”，所以F1值会呈现过高的情况，在整合结果后，所有分类结果为“no”的情况已被筛出，所评价的均为分类结果为“yes”的数据，故F1值会下降。

#### 4.3.1 扩展研究

我们整理了chinese-roBERTa-wwm-ext-large模型的结果，发现对于数据集中数量少的语义标签，抽取效果并不理想。我们通过对验证集结果的整理分析，筛选出所有未被chinese-roBERTa-wwm-ext-large模型抽取出来的语义标签，基于大模型强大的语义理解和生成能力，考虑使用大模型来完成这一部分语义标签的标注任务。通过对意合图体系及标签定义的学习，以及对数据集的观察，我们发现不同的语义标签在意合图语义理论的定义（如，关系事件：关系事件由关系论元和关系事件词构成，关系论元即具有关系的实体或事件，表示实体间关系或事件间关系的抽象关系词，被视为关系事件的核心表达，作为关系事件的事件词，如“因果关系”“领属关系”等，属于关系型隐性事件词）之外，还可以与语法知识、上下文语境等来通俗地描述。由此，我们考虑到通过输入大模型能够理解的语义知识并提供示例的方法，来让大模型学习不同的语义标签并对句子进行标注。对于这些标签，我们使用大模型包括GPT-4、ChatGLM、deepseek等，查找相关定义以及在训练集中找到对应例子，基于语法、上下文语境等的理解同时使用思维链、上下文学习等方法来调用大模型API得到三元组，如：

prompt=

“对于句子“中国社会科学院的经济学者朱运法、张延群说：“住房分配货币化已是大势所趋。”，有一个三元组(And,说,A0)，把“中国社会科学院的经济学者朱运法、张延群”意思是这两个人作为一个整体，标注为And，这个整体和“说”是一种A0的关系，相当于说A0是一个主语与谓语之间的关系。

对于句子“中国古代在造纸的技术、设备、加工等方面为世界各国提供了一整套先进的工艺体系。”，三元组(And,提供,范围),将“一整套先进的工艺体系”And标签，在该句子中与“提供”构成“范围”标签。

当然，不是所有的句子都会出现这样的关系，不用标注其他关系，也不需要标注And的所指内容。

请你一步一步思考，识别以下分词了之后的句子中\*And和A0\*或者\*And和A1\*或者\*And和A2\*或者\*And和范围\*的关系，并按照这样的格式输出：

sent=[“分词后的句子”] results=[(“And”, “句子中的词”, “A0/A1/A2/范围”)]”’

#### 4.3.2 与基准系统相比

本文提出的基于关系抽取的意合图语义分析方法得分为64.44%，与基线依存模型相比提升了33.41%，与基线大模型GPT-4交互式相比提升了27.15%。在拓展研究中让大模型对在chinese-roBERTa-wwm-ext-large模型中F1得分低的标签如PN、CompPN、Conj、FW、TM、PF、SF等进行重新整理。整合大模型和小模型的结果，最终F1为64.42%。相较单独使用chinese-roBERTa-wwm-ext-large模型得分有所下降。

#### 4.3.3 错误分析

对实验结果进一步分析发现，本文提出的方法仍然存在不足之处。一方面，本文的标签翻倍虽然在一定程度上解决了意合图的有向问题，但是这会导致标签数量过多，从而影响模型性能。其次，对于标签位置idx的标注，本文在数据整合阶段单独标注，这会导致在某些出现多次的隐性事件词标签的位置idx标注只能有一个值，而这将会降低模型的F1值。另一方面，在大模型的应用上，本文的方法较为浅显，并没有很好地利用到数据集进行微调，所采用的提示工程方法也较为单一。

除此之外，可以观察到在扩展研究中使用大模型对chinese-roBERTa-wwm-ext-large模型未能抽取出来的标签重新单独抽取，得到的F1值相较不使用大模型并没有提高，对此，我们对大



实验系统	F1得分 (%)
基线依存模型	31.03
基线GPT4交互式	37.29
<b>chinese-roBERTa-wwm-ext-large模型</b>	<b>64.44</b>
混合模型 (GPT-4、ChatGLM、deepseek、chinese-roBERTa-wwm-ext-large模型)	64.42

表 3: 实验结果

模型标注的数据进行了错例分析。我们发现:

1. **大模型生成不可控**: 即使在提示语中约束的情况下, 除了对特定标签的标注, 大模型还会自己生成一些不在范围内的标签, 甚至还会生成二元组、四元组等, 大模型的生成不可控, 需要对格式、标签种类进行筛选。
2. **抽取标签数量过多**: 我们观察到, 这些特定标签本身出现概率较小, 而大模型几乎会在每一个句子都抽取特定标签, 其中大部分句子并不含有该标签, 因此会导致F1值下降。
3. **缺乏上下文示例**: 由于输入输出长度限制, 大模型无法学习每一个包含标签的句子, 所以缺乏上下文学习的示例, 这也导致大模型的抽取效果不佳。

## 5 结论

本文采用的基于关系抽取的中文意合图语义解析方法, 通过给分类标签加上顺序来解决意合图中的有向问题; 我们将所有可能出现的隐性事件词添加到输入末尾从而也将隐性事件词的预测转换成关系抽取任务。为了保持标签的平和, 我们划分成两个子任务, 即不包含隐性事件词的关系抽取和包含隐性事件词的关系抽取, 同时训练两个子任务模型并对于这一部分未能标注出的数据采用大模型进行尝试。实验结果证明, 本文提出的关系抽取方法不仅可以考虑到 $word_1$ ,  $word_2$ 的顺序, 还能通过扩充输入的方式统一非隐性事件词之间、隐性事件词之间、隐性事件词和非隐性事件词之间的关系抽取任务, 在基线依存模型的基础上F1提升了33.41%, 可见我们的方法有效且具有创新性。在下一步的研究工作中, 可以考虑通过微调大模型来提升大模型的效果, 也可以将关系抽取的方法运用到篇章的语义解析任务上。

## 参考文献

- Rui Cai, Xiaodong Zhang, and Houfeng Wang. 2016. Bidirectional recurrent convolutional neural network for relation classification. In Katrin Erk and Noah A. Smith, editors, *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 756–765, Berlin, Germany, August. Association for Computational Linguistics.
- Cícero dos Santos, Bing Xiang, and Bowen Zhou. 2015a. Classifying relations by ranking with convolutional neural networks. In Chengqing Zong and Michael Strube, editors, *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 626–634, Beijing, China, July. Association for Computational Linguistics.
- Cicero Nogueira dos Santos, Bing Xiang, and Bowen Zhou. 2015b. Classifying relations by ranking with convolutional neural networks.
- Matthew R. Gormley, Mo Yu, and Mark Dredze. 2015. Improved relation extraction with feature-rich compositional embedding models. In Lluís Màrquez, Chris Callison-Burch, and Jian Su, editors, *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1774–1784, Lisbon, Portugal, September. Association for Computational Linguistics.
- Joohong Lee, Sangwoo Seo, and Yong Suk Choi. 2019. Semantic relation classification via bidirectional lstm networks with entity-aware attention using latent entity typing.

- Chen Lin, Timothy Miller, Dmitriy Dligach, Hadi Amiri, Steven Bethard, and Guergana Savova. 2018. Self-training improves recurrent neural networks performance for temporal relation extraction. In Alberto Lavelli, Anne-Lyse Minard, and Fabio Rinaldi, editors, *Proceedings of the Ninth International Workshop on Health Text Mining and Information Analysis*, pages 165–176, Brussels, Belgium, October. Association for Computational Linguistics.
- Mike Mintz, Steven Bills, Rion Snow, and Daniel Jurafsky. 2009. Distant supervision for relation extraction without labeled data. In Keh-Yih Su, Jian Su, Janyce Wiebe, and Haizhou Li, editors, *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 1003–1011, Suntec, Singapore, August. Association for Computational Linguistics.
- Thien Huu Nguyen and Ralph Grishman. 2015. Relation extraction: Perspective from convolutional neural networks. In Phil Blunsom, Shay Cohen, Paramveer Dhillon, and Percy Liang, editors, *Proceedings of the 1st Workshop on Vector Space Modeling for Natural Language Processing*, pages 39–48, Denver, Colorado, June. Association for Computational Linguistics.
- Yatian Shen and Xuanjing Huang. 2016. Attention-based convolutional neural network for semantic relation extraction. In Yuji Matsumoto and Rashmi Prasad, editors, *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pages 2526–2536, Osaka, Japan, December. The COLING 2016 Organizing Committee.
- Linlin Wang, Zhu Cao, Gerard de Melo, and Zhiyuan Liu. 2016. Relation classification via multi-level attention CNNs. In Katrin Erk and Noah A. Smith, editors, *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1298–1307, Berlin, Germany, August. Association for Computational Linguistics.
- Shanchan Wu and Yifan He. 2019. Enriching pre-trained language model with entity information for relation classification.
- Mo Yu, Matthew R. Gormley, and Mark Dredze. 2015. Combining word embeddings and feature embeddings for fine-grained relation extraction. In Rada Mihalcea, Joyce Chai, and Anoop Sarkar, editors, *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1374–1379, Denver, Colorado, May–June. Association for Computational Linguistics.
- Daojian Zeng, Kang Liu, Siwei Lai, Guangyou Zhou, and Jun Zhao. 2014. Relation classification via convolutional deep neural network. In Junichi Tsujii and Jan Hajic, editors, *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pages 2335–2344, Dublin, Ireland, August. Dublin City University and Association for Computational Linguistics.
- 郭梦溪, 李梦, 荀恩东, 饶高琦, and 于钟洋. 2024a. 基于意合图语义理论的结构标注体系与资源建设. 第二十三届中国计算语言学大会.
- 郭梦溪, 荀恩东, 李梦, and 饶高琦. 2024b. 意合图: 中文多层次语义表示方法. 第二十三届中国计算语言学大会.