

# CCL23-Eval 任务2系统报告: 基于大型语言模型的中文抽象语义表示解析

杨逸飞\*, 程子鸣\*, 赵海†

上海交通大学计算机科学与工程系

{yifeiyang, kk.cheng}@sjtu.edu.cn, zhaohai@cs.sjtu.edu.cn

## 摘要

中文抽象语义表示解析旨在将自然语句转换为抽象语义表示, 是一个复杂的结构化预测任务。传统方法多利用抽象语义表示的图特征设计特殊模型或者多阶段解析来完成解析, 而这类方法通常需要设计复杂的神经网络模型。目前, 通用大型语言模型在已经多种自然语言处理任务上表现出惊人效果, 我们在本次测评中尝试直接利用大型语言模型进行零样本学习、少样本学习以及用LoRA和全参数的方式微调大型语言模型来完成解析。我们得到了一个较好的评测结果, 并对这些方案进行了讨论。

**关键词:** 中文抽象语义表示; 抽象语义表示解析; 大型语言模型

## System Report for CCL23-Eval Task 2: Chinese Abstract Meaning Representation Parsing based on Large Language Model

Yifei Yang\*, Ziming Cheng\*, Hai Zhao†

Department of Computer Science and Engineering, Shanghai Jiao Tong University

{yifeiyang, kk.cheng}@sjtu.edu.cn, zhaohai@cs.sjtu.edu.cn

## Abstract

Chinese Abstract Meaning Representation Parsing aims to convert natural language sentences into abstract semantic representations, which is a complex structure prediction task. Traditional approaches often utilize graph features of abstract semantic representations to design specialized models or employ multi-stage parsing. However, these methods typically require the design of complex neural network models. Currently, large language models have demonstrated astonishing performance on various natural language processing tasks. In this evaluation, we attempt to directly utilize a large language model for Zero-shot learning, Few-shot learning, and fine-tuning using LoRA and full-parameter approaches. We obtain promising evaluation results and discuss these approaches in detail.

**Keywords:** Chinese Abstract Meaning Representation, Abstract Meaning Representation Parsing, Large Language Model

## 1 引言

©2023 中国计算语言学大会

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

\*相同贡献

†通讯作者

语义理解是自然语言处理（NLP）中一个长期的研究问题，如果机器能够理解语义，并用机器适配的表示进行储存和转换，最后搭配生成模型，就能够完成大多数的自然语言任务。2013年Banarescu (2013)等人提出了一种领域无关且通用的句子语义表示方法，称为抽象语义表示（Abstract Meaning Representation, AMR）。这种语义表示只有一个根节点，构成有向无环图，且不受语句语法和表达的影响，能够抽象出句子中的语义成分并展示出成分之间的关系。经过多年的研究积累，英文上的文本与抽象语义表示的双向转换已经有了优秀的效果。

近年来，更多的人开始关注到了跨语言AMR解析与生成，在中文方面，Li等人(2019)根据汉语的特点优化了中文AMR（CAMR）的数据集，肖力铭等人(2022)创建了加入概念对齐指标和关系对齐指标的Align-smatch标准。这些工作奠定了中文AMR研究的基础。

聚焦于本次CAMR 2023评测任务中的中文AMR解析任务，前人的许多工作集中于利用图的特性来引导模型进行解析，如：利用多层次分类来生成图信息中的不同要素(Samuel and Straka, 2020)、利用迭代优化的方式一步一步生成图(Damonte et al., 2016)。这些模型大多结构复杂，且推理速度较慢，并且需要对训练数据集进行复杂的额外处理。受到英文AMR研究(Bevilacqua et al., 2021)的启发，通过深度遍历等特殊方式将AMR标注序列化，随后利用序列到序列（Seq2Seq）的建模方式来完成文本和抽象语义表示的转换，是AMR领域的一个崭新解决方案。该方案的模型结构简单，且能够直接满足双向转换。目前，大型语言模型（Large Language Model, LLM）也已经在众多NLP任务上展现出了惊人效果，我们探索了LLM是否能够借鉴Seq2Seq建模方式在复杂结构化预测任务上得以运用，并给出了LLM在零样本学习（Zero-shot），少样本学习（Few-shot）和用LoRA (Hu et al., 2021)、全参数微调场景下的不同表现，从而进行前瞻性分析与总结。

## 2 方法

### 2.1 思路总览

本团队考虑如下两种方案来完成CAMR这一复杂的结构化预测任务：

- 利用LLM进行Zero-shot和Few-shot设定下的预测。对于Zero-shot，直接让LLM解析一个给定的句子；对于Few-shot，利用上下文学习（in-context learning）(Brown et al., 2020)，给定若干个上下文样本（in-context sample），让模型仿造样本对给定句子进行解析。
- 利用LoRA (Hu et al., 2021)微调或全参数微调的方式，对预训练好的LLM进行领域微调，使其能够在输入给定句子时，输出其对应的解析结果。其中LoRA是现在流行的模型高效微调方法之一，它添加额外的低秩矩阵，在训练时仅优化这些低秩矩阵从而加速训练。

### 2.2 任务定义

AMR是一个复杂的结构化数据（如附录A所示），本质上为一个复杂的图结构。AMR解析则是对给定的自然语言文本生成其对应的AMR。传统的解析方法(Samuel and Straka, 2020; Damonte et al., 2016)采用迭代图生成或者二阶段生成的方式来进行解析。然而，现有的LLM拥有强大的Seq2Seq生成能力，将AMR数据转换为一个类文本序列数据后，可以尝试直接利用LLM的Seq2Seq能力完成解析。本工作涉及两种由AMR数据转换而成的类文本序列数据：

- AMR数据的文本表示（附录A图2）将图结构转换为括号分隔的层级结构来实现序列化。
- AMR数据的多元组（附录A图3）利用9个制表符隔开的元素所构成的文本行来表示AMR图中的节点和边关系，多行一起构成了序列化数据。

经过以上序列化方式后，我们可以将CAMR任务建模成自然语言中的Seq2Seq问题，即用LLM将数据集中的原始输入序列 $X$ 转换为序列化的AMR类文本序列 $Y$ 。由于LLM为生成式模型，只有单向的注意力机制，因此在训练过程中，我们会对 $X$ 与 $Y$ 序列进行拼接。同时，为了和预训练语料形成领域上的差异度，我们采用提示工程（Prompt Engineering）的方式对训练语料进行定制。

## 2.3 模型微调

我们在 $X$ 前后分别插入一段前缀文本 $pre$ 和一段后缀文本 $sub$ ，形成新的序列 $X' = [pre; X; sub]$ ，其中“;”代表拼接。

在微调模型过程中，假设 $Z = [X'; Y]$ ，定义 $P(\hat{z}_i) = \frac{e^{\hat{z}_i}}{\sum_{z \in V} e^z}$ 和交叉熵损失 $CE = \sum_{z_i \in V} -z_i \log P(\hat{z}_i)$ ，训练目标为最小化交叉熵损失 $CE$ 。

## 2.4 Zero-shot和Few-shot建模

对于Zero-shot，我们需要利用提示工程添加描述任务的语句来让模型完成CAMR解析。即，给定数据集中的原始序列 $X$ ，我们在其前面添加对CAMR任务的描述 $des$ ，在后面添加对输出的描述 $out$ 序列形成输入 $X'' = [des; X; out]$ ，将其输入给LLM后，希望模型输出对应的 $Y$ 。

对于Few-shot，在Zero-shot的基础上，我们利用上下文学习的思路，随机在训练集中采样 $n$ 个上下文样本 $[(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)]$ ，然后将其和原始序列 $X$ 一起构造成 $X''' = [des; X_1; out; Y_1; X_2; out; Y_2; \dots; X_n; out; Y_n; X; out]$ ，将其直接输入给LLM得到对应的输出。

# 3 实验和分析

## 3.1 模型选型

对于Zero-shot和Few-shot，我们选择ChatGPT (Ouyang et al., 2022)<sup>0</sup>进行评估。对于LoRA微调和全参数微调，我们使用ChatGLM-6B (Du et al., 2022)作为基底模型。其中ChatGPT大约有1750亿 (175B) 参数，ChatGLM具有约60亿 (6B) 参数。选型理由如下：

- Zero-shot和Few-shot能力通常被认为只有在超大规模（大于700亿参数）(Wei et al., 2022)的语言模型上才能突出表现。ChatGPT是目前公认表现最好的可以使用推理功能的LLM，在Zero-shot和Few-shot场景中表现惊人，且具有开放的推理API可供调用。
- ChatGLM-6B在高质量的中文数据上进行了预训练，并针对中文问答和对话进行了优化，是目前中文表现最好的LLM之一。且60亿的参数量足够我们在不具备大算力服务器的条件下进行微调。

## 3.2 数据集和实验实施

我们选取主办方提供的数据集进行训练和测试，参加开放测试的评测。由于评测结果需要在提供的多元组格式数据上测试，为了方便性，我们直接利用LLM输出多元组格式的预测。

在Zero-shot和Few-shot场景下使用ChatGPT完成CAMR任务时，我们购买了付费API，每分钟的最大请求数量是60次，单次最大输入长度的是4096个标记 (tokens)。对于Few-shot，我们设置上下文样本数量 $n = 5$ 。在微调ChatGLM-6B时，我们在一台8卡NVIDIA V100 32G高性能服务器上训练了约5个epoch。其中，使用LoRA微调时，我们将低秩矩阵的秩设为8。在全参数微调时，我们设置batch-size为8。训练使用AdamW优化器(Loshchilov and Hutter, 2017)并配合学习率 $1e-4$ 。我们采用deepspeed<sup>1</sup>的ZeRO-2策略加速训练并降低显存开销。

## 3.3 Zero-shot设定下的ChatGPT

在Zero-shot设定下时，我们定义 $des$ 为“你现在是中文抽象语义表示解析器，给定句子：”，定义 $out$ 为“请你直接输出它的多元组（节点编号1 概念1 同指节点1关系关系编号关系对齐词节点编号2 概念2 同指节点2）解析结果：”。我们在表1中展示了一个输入示例及对应的输出。

从表1中我们可以看出，由于CAMR任务输出端的结构复杂性，在没有上下文样本作为参照的情况下，模型甚至无法预测出正确的输出格式。这种现象说明对于类似于CAMR这种复杂的结构化预测任务来说，LLM在Zero-shot下直接解决仍有很大难度。

<sup>0</sup><https://chat.openai.com/chat>

<sup>1</sup><https://www.deepspeed.ai/>

原始序列	美国 批准 抗 除草剂 转基因 菊苣 商业化 种植
实际输入	你现在是中文抽象语义表示解析器，给定句子：美国 批准 抗 除草剂 转基因 菊苣 商业化 种植，请你直接输出它的多元组（节点编号1 概念1 同指节点1关系关系编号关系对齐词节点编号2 概念2 同指节点2）解析结果：
模型输出	(1, “美国”, “”, 0, “”, 0, “”) (2, “批准”, “”, 0, “”, 0, “”) (3, “抗除草剂”, “”, 0, “”, 0, “”) (4, “转基因菊苣”, “”, 0, “”, 0, “”) (5, “商业化种植”, “”, 0, “”, 0, “”)

Table 1: Zero-shot设置下的输入和输出示例

### 3.4 Few-shot设定下的ChatGPT

在Few-shot设定下时，我们定义 $des$ 为“你现在是中文抽象语义表示解析器，请你按照给定的例子进行补全：”，定义 $out$ 为“输出：”。我们在附录B表3中展示了一个示例。我们发现在给定上下文样本后，模型可以顺利地参照这些样本预测出CAMR的输出格式。但输出的答案和标准答案相比仍然差距很大。我们认为模型表现不佳的原因主要有两点：1. 现有的LLM在面临长文本输入时的生成效果欠(Anil et al., 2022)，而由于CAMR的特殊输出结构，在引入多个上下文样本后会无法避免地产生长文本输入，从而导致生成效果不理想；2. CAMR的输出序列中节点和节点之间的关系有多种类型，而仅靠给定的5个上下文样本很难穷尽到所有的关系类型，模型在预测的时候无法泛化到没有见过的关系。如附录B表3所示，由于上下文样本中没有“:name”这一关系，模型无法在输出中预测出“:name”关系。

### 3.5 微调ChatGLM-6B

在利用LoRA和全参数微调ChatGLM-6B时，根据2.3节，要对原始输入序列 $X$ 插入前缀和后缀文本克服领域差异性。我们设定前缀为 $pre$ 为“你是一个中文抽象语义表示解析器，给定一个分词后的句子：”，后缀 $sub$ 为“它的中文抽象语义表示解析结果为：”。我们在附录C表4中展示了一个由原始样本构造成的训练样本。在推理时，对于一个待预测的样本 $X_I$ ，我们构造输入样本 $X'_I = [pre; X_I; sub]$ ，让训练好的模型补全下文，从而预测出对应的CAMR多元组表示。

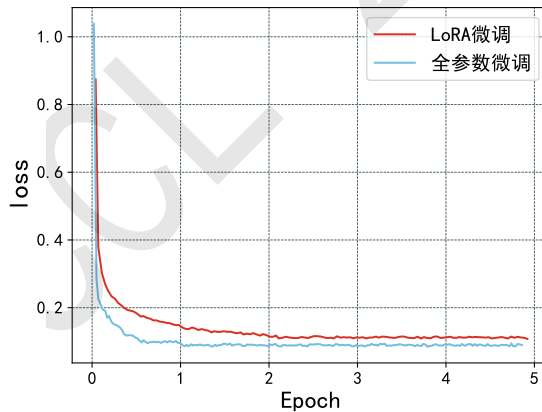


Figure 1: loss下降曲线

对于LoRA微调和全参数微调，我们在构造好的同一个数据集上进行训练，分别绘制了它们的损失（loss）下降曲线。如图1所示，两种微调方式在训练时都快速收敛，约2个epoch后趋近于稳定，但是全参数微调的损失明显低于LoRA微调的损失。我们也分别对两种优化方式微调好的模型进了粗略评估，如附录D表5所示了两个例子。从中可以看出LoRA微调出的模型甚至无法预测正确的CAMR多元组格式，但全参数微调的模型可以输出。因此我们选用全参数微调的模型对所有的盲测集进行了预测，并提交给主办方用于评测，结果如表2所示。

参考2022年的结果，我们的测评分数大致与排名第三的队伍相当，但与排名第一的队伍仍然差距明显。而2022年的方案都还未用到LLM，这说明LLM对于CAMR这种复杂结构化预测任务的处理能力仍有欠缺。



盲测集	P	R	F <sub>1</sub>
A	47.41	46.45	46.92
B	46.44	45.68	46.06
C	62.82	58.39	60.52

Table 2: 测评结果

## 4 讨论

### 4.1 全参数微调降低泛化性能的风险

我们发现在训练集上对模型进行全参数微调会严重降低模型泛化性能。ChatGLM-6B经过微调后会对任何输入的句子做CAMR解析而丧失了预训练阶段得到的通用对话能力，如附录E表6所示。这说明在实际应用中使用通用LLM在单一领域的数据集上微调不是一种非常合理的做法。

### 4.2 不同格式的预测序列

如附录A所示，主办方提供了中文AMR的文本表示和多元组表示。由于评测是在多元组表示上完成的，我们直接选用了多元组表示作为目标序列。而我们发现多元组表示其实存在大量冗余信息，如同指节点（coref）、关系编号（rid）这种属性存在大量缺省，不利于模型训练。我们已经训练了模型直接输出文本表示，再手动转换为多元组表示用于测评。但由于输出的文本表示可能存在不合AMR约束的情况，目前转换步骤还没有实现，我们将其作为后续工作。

## 5 总结与展望

本工作探索了LLM在CAMR解析这一结构化预测任务上的表现：

- 利用全参数微调可以使得LLM具备一定的CAMR解析能力，但会损伤模型泛化性。而LoRA微调不足以让LLM得到CAMR解析能力。
- 即使是公认最佳的大型语言模型ChatGPT也无法实现Zero-shot的CAMR解析。而Few-shot对导致输入序列太长并且无法穷举到所有边的类型，因此在CAMR上的表现也有欠缺。
- 我们选用经过全参数微调后的ChatGLM-6B模型的预测作为最终评测结果，达到了较为理想的评测结果，但仍与传统方法有一定差距。

作为探索性工作，我们的尝试具有一定启发意义。在未来，我们会继续延续本工作，如探究不同的提示方式带来的影响、继续完善采用AMR文本表示作为目标序列的实验等。

## 参考文献

- Cem Anil, Yuhuai Wu, Anders Andreassen, Aitor Lewkowycz, Vedant Misra, Vinay Ramasesh, Ambrose Slone, Guy Gur-Ari, Ethan Dyer, and Behnam Neyshabur. 2022. Exploring length generalization in large language models. *arXiv preprint arXiv:2207.04901*.
- Laura Banarescu, Claire Bonial, Shu Cai, Madalina Georgescu, Kira Griffitt, Ulf Hermjakob, Kevin Knight, Philipp Koehn, Martha Palmer, and Nathan Schneider. 2013. Abstract meaning representation for sembanking. In *Proceedings of the 7th linguistic annotation workshop and interoperability with discourse*, pages 178–186.
- Michele Bevilacqua, Rexhina Blloshmi, and Roberto Navigli. 2021. One spring to rule them both: Symmetric amr semantic parsing and generation without a complex pipeline. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 12564–12573.
- Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901.

- Marco Damonte, Shay B Cohen, and Giorgio Satta. 2016. An incremental parser for abstract meaning representation. *arXiv preprint arXiv:1608.06111*.
- Zhengxiao Du, Yujie Qian, Xiao Liu, Ming Ding, Jiezhong Qiu, Zhilin Yang, and Jie Tang. 2022. Glm: General language model pretraining with autoregressive blank infilling. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 320–335.
- Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2021. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*.
- Bin Li, Yuan Wen, Li Song, Weiguang Qu, and Nianwen Xue. 2019. Building a chinese amr bank with concept and relation alignments. In *Linguistic Issues in Language Technology, Volume 18, 2019-Exploiting Parsed Corpora: Applications in Research, Pedagogy, and Processing*.
- Ilya Loshchilov and Frank Hutter. 2017. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*.
- Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. 2022. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35:27730–27744.
- David Samuel and Milan Straka. 2020. \ufal at mrp 2020: Permutation-invariant semantic parsing in perin. *arXiv preprint arXiv:2011.00758*.
- Jason Wei, Yi Tay, Rishi Bommasani, Colin Raffel, Barret Zoph, Sebastian Borgeaud, Dani Yogatama, Maarten Bosma, Denny Zhou, Donald Metzler, et al. 2022. Emergent abilities of large language models. *arXiv preprint arXiv:2206.07682*.
- Liming Xiao, Bin Li, Zhixing Xu, Kairui Huo, Minxuan Feng, Junsheng Zhou, and Weiguang Qu. 2022. Align-smatch: A novel evaluation method for chinese abstract meaning representation parsing based on alignment of concept and relation. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 5938–5945.

## A 中文AMR不同表示案例

```

# ::id export_amr.2580 ::cid export_amr.2580 ::2017-02-02 17:03:12
# ::snt 这几天关于中俄战略合作伙伴关系成了大热点。
# ::wid x1_这 x2_几 x3_天 x4_关于 x5_中 x6_俄 x7_战略 x8_合作 x9_伙伴 x10_关系 x11_成
x12_了 x13_大 x14_热点 x15_。
(x11 / 成-01
  :aspect() (x12 / 了)
  :arg1() (x14 / 热点
    :arg0-of() (x13 / 大-01))
  :arg0(x4/关于) (x10 / 关系
    :mod() (x9 / 伙伴
      :mod() (x8 / 合作-01
        :arg0() (x26 / and
          :op1() (x33 / country
            :name() (x5 / name :op1 x5/中 ))
          :op2() (x35 / country
            :name() (x6 / name :op1 x6/俄 )))
        :mod() (x7 / 战略)))
  :duration() (x37 / temporal-quantity
    :quant() (x2 / 几)
    :unit() (x3 / 天)
    :mod() (x1 / 这)))

```

Figure 2: 中文AMR文本表示案例

句子编号	节点编号1	概念1	同指节点1	关系	关系编号	关系对齐词	节点编号2	概念2	同指节点2
sid	nid1	concept1	coref1	rel	rid	ralign	nid2	concept2	coref2
2580	x0	root	-	:top	-	-	x11	成-01	-
2580	x11	成-01	-	:aspect	-	-	x12	了	-
2580	x11	成-01	-	:arg1	-	-	x14	热点	-
2580	x11	成-01	-	:arg0	x4	关于	x10	关系	-
2580	x11	成-01	-	:duration	-	-	x37	temporal-quantity	-
2580	x14	热点	-	:arg0-of	-	-	x13	大-01	-
2580	x10	关系	-	:mod	-	-	x9	伙伴	-
2580	x9	伙伴	-	:mod	-	-	x8	合作-01	-
2580	x9	伙伴	-	:mod	-	-	x7	战略	-
2580	x8	合作-01	-	:arg0	-	-	x26	and	-
2580	x26	and	-	:op1	-	-	x33	country	-
2580	x26	and	-	:op2	-	-	x35	country	-
2580	x33	country	-	:name	-	-	x5	中	-
2580	x35	country	-	:name	-	-	x6	俄	-
2580	x37	temporal-quantity	-	:quant	-	-	x2	几	-
2580	x37	temporal-quantity	-	:unit	-	-	x3	天	-
2580	x37	temporal-quantity	-	:mod	-	-	x1	这	-

Figure 3: 中文AMR多元组表示案例

## B Few-shot示例



<b>原始序列</b>	美国 批准 抗 除草剂 转基因 菊苣 商业化 种植
<b>标准答案</b>	<pre>x0 root - :top - - x2 批准-01 - x2 批准-01 - :arg0 - - x12 country - x2 批准-01 - :arg1 - - x8 种植-01 - x12 country - :name - - x1 美国- x8 种植-01 - :mod - - x7 商业化- x8 种植-01 - :arg1 - - x6 菊苣- x6 菊苣- :mod - - x5 转基因- x6 菊苣- :arg0-of - - x3 抗-01 - x3 抗-01 - :arg1 - - x4 除草剂-</pre>
<b>实际输入</b>	<p>你现在是中文抽象语义表示解析器，请你按照给定例子进行补全：</p> <p>图 为 赣州市 新区 面貌 。</p> <p><b>输出：</b></p> <pre>x0 root - :top - - x5 面貌- x5 面貌- :poss - - x4 新区- x5 面貌- :domain x2 为x1 图- x4 新区- :location - - x11 city - x11 city - :name - - x3 赣州市-</pre> <p>那样 你 的 顾虑 ， 就 没有 了</p> <p><b>输出：</b></p> <pre>x0 root - :top - - x12 condition - x12 condition - :arg1 - - x1 那样- 12 condition - :arg2 x6 就x7.2 有-03 - x7.2 有-03 - :polarity - - x7.1 - - x7.2 有-03 - :arg1 - - x4 顾虑-01 - x7.2 有-03 - :aspect - - x8 了- x4 顾虑-01 - :arg0 x3 的x2 你-</pre> <p>...</p> <p>美国 准 抗 除草剂 转基因 菊苣 商业化 种植</p> <p><b>输出：</b></p> <pre>x0 root - :top - - x8 approve - x8 approve - :arg0 - - x1 美国- x8 approve - :arg1 x6 转基因x7 种植-01 - x6 转基因- :mod - - x5 除草剂- x7 种植-01 - :arg1 x3 抗x4 菊苣- x7 种植-01 - :arg0 x2 商业化- -</pre>

Table 3: Few-shot设置下的输入、输出和标准答案。为了简洁，我们只展示了2个上下文样本，在实际测试时我们使用了5个样本

### C 模型微调训练样本示例

---

**输入序列** 美国 批准 抗 除草剂 转基因 菊苣 商业化 种植

**标准答案** x0 root - :top - - x8 approve -  
 x8 approve - :arg0 - - x1 美国-  
 x8 approve - :arg1 x6 转基因x7 种植-01 -  
 x6 转基因- :mod - - x5 除草剂-  
 x7 种植-01 - :arg1 x3 抗x4 菊苣-  
 x7 种植-01 - :arg0 x2 商业化- -

---

**训练样本** 你现在是中文抽象语义表示解析器，给定句子：美国 批准 抗 除草剂 转基因 菊苣 商业化 种植，它的多元组（节点编号1 概念1 同指节点1关系关系编号关系对齐词节点编号2 概念2 同指节点2）解析结果为：

x0 root - :top - - x8 approve -  
 x8 approve - :arg0 - - x1 美国-  
 x8 approve - :arg1 x6 转基因x7 种植-01 -  
 x6 转基因- :mod - - x5 除草剂-  
 x7 种植-01 - :arg1 x3 抗x4 菊苣-  
 x7 种植-01 - :arg0 x2 商业化- -

---

Table 4: 模型微调时构造的训练样本

## D LoRA微调模型 vs. 全参数微调模型

<b>输入序列</b>	美国 批准 抗 除草剂 转基因 菊苣 商业化 种植
<b>标准答案</b>	x0 root - :top - - x8 approve - x8 approve - :arg0 - - x1 美国- x8 approve - :arg1 x6 转基因x7 种植-01 - x6 转基因- :mod - - x5 除草剂- x7 种植-01 - :arg1 x3 抗x4 菊苣- x7 种植-01 - :arg0 x2 商业化- -
<b>LoRA微调模型预测</b>	美国批准了转基因菊苣商业化种植。
<b>全参数微调模型预测</b>	x0 root - :top - - x8 approve - x8 approve - :arg0 - - x1 美国- x8 approve - :arg1 x6 转基因x7 种植-01 - x6 转基因- :mod - - x5 除草剂- x7 种植-01 - :arg1 x3 抗x4 菊苣- x7 种植-01 - :arg0 x2 商业化- -
<b>输入序列</b>	常言道：人间自有公道在。
<b>标准答案</b>	x0 root - :top - - x2 道-01 - x2 道-01 - :arg2 - - x1 常言- x2 道-01 - :arg1 - - x6 有-01 - x6 有-01 - :arg0 - - x4 人间- x6 有-01 - :mod - - x5 自- x6 有-01 - :arg1 x8 在x7 公道-
<b>LoRA微调模型预测</b>	人间有公道。 这句话的意思是,人们常说:“人间自有公道”,意思是在人类社会中,总会有公正和正义的存在。
<b>全参数微调模型预测</b>	x1 常言道- :arg1 - - x3 有-03 - x3 有-03 - :arg0 - - x4 公道- x3 有-03 - :arg1 - - x5 在-01 - x3 有-03 - :mod - - x2 自- x3 有-03 - :location - - x4 公道- x4 公道- :location - - x4 公道- x5 在-01 - :arg0 - - x4 公道- x5 在-01 - :arg1 - - x4 公道- x4 公道- :location - - x4 公道- x4 公道- :location - - x3 人间-

Table 5: LoRA微调模型的预测对比全参数微调模型的预测

## E 全参数微调导致的模型泛化性降低

输入	你好!
不微调的模型输出	你好! 你好!很高兴见到你,欢迎问我任何问题。
微调后的模型输出	你好! 它的多元组(节点编号1 概念1 同指节点1关系关系编号关系对齐词节点编号2 概念2 同指节点2)解析结果为: x0 root - :top - - x3 你好- x3 你好- :mode - - x4 expressive -

Table 6: 模型泛化性能降低, 对于任何输入都尝试做CAMR解析, 丧失了预训练得到的通用对话能力