doi:10.16112/j.cnki.53-1223/n.202511040002
引用格式:王晓聪,李英,李衍铎,等.基于多粒度知识对齐的低资源机器翻译方法[J].昆明理工大学学报(自然科学版),2026,51(3):96-108.
Citation:WANG Xiaocong,LI Ying,LI Yanduo,et al. Multi-granularity Knowledge Alignment for Low-Resource Machine Translation[J]. Journal of Kunming University of Science and Technology(Natural Science),2026,51(3):96-108.

基于多粒度知识对齐的低资源机器翻译方法

王晓聪[1,2],李 英[1,2],李衍铎[1,2],高盛祥[1,2],余正涛[1,2*]
(1. 昆明理工大学 信息工程与自动化学院,云南 昆明 650500; 2. 云南省人工智能重点实验室,云南 昆明 650500)

摘要:低资源语言中,复合实体的深层语义结构难以学习,导致其翻译准确率低,进而影响整体翻译质量.为此,本文提出基于多粒度知识对齐的低资源机器翻译方法.首先,利用词对齐模型提取词级、实体级的对应翻译对,并基于此对齐关系设计词对齐、实体对齐和思考过程词对齐三种规则式对齐奖励;其次,结合格式和整体语义的一致性形成奖励信号;最后,通过强化学习框架,将这些奖励信号整合到翻译模型的训练过程中,使模型能够专注于源文本中复合实体语义关键信息的准确翻译.在中英、中越、中老三个语言对的实验表明本文方法在 BLEU、COMET 和实体准确率指标上均优于基线模型.详细的对比实验证明多粒度知识对齐方法能更好地理解复合实体的语义信息,从而有效缓解复合实体翻译中的歧义问题,提升翻译质量.

关键词:低资源机器翻译;多粒度知识对齐;大语言模型;强化学习
中图分类号:TP391.2;H085;H3 **文献标识码**:A **文章编号**:1007-855X(2026)03-0096-13

Multi-granularity Knowledge Alignment for Low-Resource Machine Translation

WANG Xiaocong[1,2],LI Ying[1,2],LI Yanduo[1,2],GAO Shengxiang[1,2],YU Zhengtao[1,2*]
(1. Faculty of Information Engineering and Automation,Kunming University of Science and Technology,Kunming 650500,China; 2. Yunnan Key Laboratory of Artificial Intelligence,Kunming 650500,China)

Abstract: In low-resource languages,the deep semantic structures of compound entities are difficult to learn,leading to low translation accuracy and consequently affecting overall translation quality. To address this issue,this paper proposes a low-resource machine translation method based on multi-granularity knowledge alignment. First,a word alignment model is used to extract corresponding translation pairs at the word level and entity level,based on which three rule-based alignment rewards—word alignment,entity alignment,and thought process word alignment—are designed. Second,consistency in format and overall semantics is incorporated to form reward signals. Finally,these reward signals are integrated into the training process of the translation model through a reinforcement learning framework,enabling the model to focus on accurately translating semantically critical information of compound entities in the source text. Experiments on three language pairs,Chinese-Eng-

收稿日期:2025-11-04. **基金项目**:国家自然科学基金项目(U24A20334,U23A20388,62366027);云南省科技重大专项(202402AG050007,202302AD080003);云南省基础研究计划重大项目(202401BC070021).
作者简介:王晓聪(1993-),男,博士研究生.主要研究方向:自然语言处理,机器翻译. **E-mail**:wangxcai@foxmail.com
通信作者:余正涛(1970-),男,博士,教授,博士生导师.主要研究方向:自然语言处理、机器翻译、语音识别与合成等.
E-mail:ztyu@hotmail.com

lish, Chinese – Vietnamese, and Chinese – Lao, demonstrate that the proposed method outperforms baseline models in terms of BLEU, COMET, and entity accuracy metrics. Detailed comparative experiments show that the multi – granularity knowledge alignment method can better understand the semantic information of compound entities, thereby effectively alleviating ambiguity issues in compound entity translation and improving translation quality.

Keywords: low – resource machine translation; multi – granularity knowledge alignment; large language models; reinforcement learning

0 引 言

实体是句子中承载关键语义信息的语言单元,其翻译的准确性直接影响译文的可读性与专业性. 当前,基于大语言模型强大的泛化能力,机器翻译系统在整体翻译质量上取得了显著进展[1]. 然而,在低资源语言场景下,由于高质量平行语料的稀缺,模型对深层语义结构的学习仍面临严峻挑战,其中复合实体的准确翻译成为一个尤为突出的瓶颈问题. 复合实体(如“长城 SUV 哈弗 H6”“中国人民银行数字货币研究所”)是由多个语义单元构成的专有名词短语,其结构复杂、语义嵌套性强,且在不同语境下可能产生指代歧义. 低资源条件下,模型很难学习到复合实体的构成规律和映射关系,缺乏对实体深层语义的理解,对歧义复合实体的处理能力有限. 此外,低资源条件下模型对复合实体的特定指代含义理解不完整,语义边界模糊,易产生混淆. 这些问题的根源在于低资源条件下模型无法获得足够的训练数据,模型难以学习其深层语义结构,往往退化为字面直译、成分遗漏或错误替换,缺乏对实体深层语义的理解,严重损害译文的可用性.

针对实体翻译问题,研究者们已提出多种改进思路并取得一定进展. 一是通过引入外部知识补充语义映射信息. Wang 等[2]利用非结构化文档进行检索增强的机器翻译,并通过多任务训练方法教会大模型(Large Language Models, LLMs)在翻译过程中利用多语言文档中的信息. 二是通过扩充训练数据来提升模型的实体翻译能力. Moslem 等[3]通过合成含歧义实体的双语平行数据,利用监督微调(Supervised Fine – Tuning, SFT)让模型学习实体的语境依赖翻译规律;三是通过在生成阶段施加约束来修正实体翻译错误. Bogoychev 等[4]在解码阶段引入实体类型约束,限制模型输出与实体属性不匹配的译文,以此修正实体翻译错误;Kim 等[5]先检测源文本实体并构建领域实体库,再通过检索增强生成(Retrieval – Augmented Generation, RAG)机制,在翻译前为模型提供实体上下文信息.

尽管上述方法在高资源语言中取得了一定的成效,但它们在低资源场景下的适用性极为有限. 检索增强方法依赖于高质量外部资源,在低资源语言中其检索质量难以保证;数据合成方法需要大量高质量的实体标注或平行句对,成本高昂且难以覆盖所有实体类型;约束解码规则则往往将实体视为静态符号,忽略了其在动态上下文中的语义角色变化. 更为根本的是,现有方法大多缺乏对复合实体内部细粒度语义结构的显式建模,未能系统性地解决其多成分映射与语义消歧问题.

为解决上述问题,提出基于多粒度知识对齐的低资源机器翻译方法(Multi – Granularity Knowledge Alignment for Low – resource Machine Translation, MGKA – MT),旨在缓解基于 LLM 的机器翻译中复合实体的歧义问题,从而改善译文的质量. 多粒度知识对齐机制通过词对齐确保复合实体内部词的准确映射以维持结构完整性,通过实体对齐精准识别实体的语义边界与对应关系以保障语义准确性,同时通过思考对齐强化模型对关键实体的翻译策略思考,三者协同能有效缓解复合实体在不同语境下的多义性问题. 首先,建立多粒度对齐的强化学习奖励信号——利用多粒度对齐技术结合双语实体词典中的实体及注释建立跨语言映射关系,精准定位复合实体在不同语境下的翻译等价关系,从双语实体中提取可靠的翻译依据;该奖励不仅关注句子级翻译的流畅度,更在词汇与实体层级引入对齐一致性度量,促使模型在生成过程中感知实体语义并做出精准选择. 其次,基于该奖励信号直接对 LLM 进行强化训练,增强对实体多义性与上下文关联性的捕捉能力. 本文在中英、中越、中老三个语言对上进行了全面的实验验证,结果表明,本方法在 BLEU、COMET 和实体准确率等多个评估指标上均显著优于一系列强基线模型. 详细的消融研究进一步证实,多粒度知识对齐机制能有效提升模型对复合实体语义的理解能力,缓解翻译歧义问题.

1 相关工作

近年来,基于大语言模型的机器翻译在通用文本翻译中取得显著进展,但复合实体歧义问题仍未得到有效解决. 实体的准确翻译是保障机器翻译语义忠实性的关键. 近年来,研究者尝试通过外部知识注入与结构化建模提升模型对实体的感知能力. Huang 等[6]将通过目标语言对实体的解释作为上下文信息融合到模型中. 类似地,Zheng 等[7]利用少量含实体的示例进行提示学习,增强模型在专业领域的适应能力.

语义对齐是连接源语言与目标语言语义空间的核心机制. 传统神经机器翻译依赖注意力机制实现隐式对齐,而近年来的研究更注重显式的多粒度对齐建模. 词对齐技术通过分析平行语料中的共现模式[8],建立跨语言词汇间的等价关系,已被广泛用于术语提取与翻译一致性[9]. 近期工作尝试将对齐信息融入翻译模型训练过程. Artetxe 等[10]利用双语词典构建对比学习目标,增强模型对翻译等价词的敏感性;Hasler 等[11]则是在解码阶段引入对齐约束,确保关键术语的翻译一致性. Li 等[12]首次利用强化学习与词对齐在深度推理大型语言模型中探索术语翻译,但需要一定规模的平行数据进行模型训练,且工作主要聚焦于高资源语言. 然而,这些对齐机制多局限于词汇层级,缺乏对上下文语境中实体语义的建模能力. 尤其在低资源语言对中,双语实体词典覆盖有限,稀疏对齐数据难以提取可靠信号并有效指导模型训练.

为改善传统监督微调在翻译流畅性与语义一致性上的局限,强化学习(Reinforcement Learning,RL)被引入机器翻译以优化模型的长期生成行为. 受 DeepSeek - R1[13]推理增强模型的启发,研究者尝试通过 RL 激发大语言模型的深度语义的理解能力. He 等[14]采用两阶段训练策略:先基于含推理过程的合成数据进行监督微调,再以 COMET[15]作为自动评估指标构建奖励信号进行强化训练;Feng 等[16]提出的 MT - R1 - Zero 则借鉴 DeepSeek - R1 - Zero 的设计,直接在预训练模型上以 BLEU 与 COMETKiwi 为奖励进行端到端强化学习,避免了微调阶段的数据依赖. Wang 等[17]则通过 DeepTrans 进一步探索以大模型自身评分作为奖励,提升文学翻译的风格适配性. Cui 等[18]提出了一种置信度 - 奖励驱动的偏好优化方法,通过结合置信度估计和奖励信号来优化翻译质量. Dong 等[19]提出了基于语言感知参数检测和 LoRA 的知识迁移方法,通过选择性更新减少参数干扰. Nguefack 等[20]评估了各种预训练技术对低资源语言的有效性,通过结合单语和平行数据的预训练策略在低资源场景下取得了最佳性能. Ghanizadeh 等[21]提出了利用学习模型和预训练参考模型之间的协同效应来提高整体训练效果.

然而,这类方法的奖励信号仍基于整体译文质量,尽管这些方法在整体翻译质量上取得进展,其奖励机制仍主要聚焦于句子级流畅度或语义相似度,未能聚焦复合实体歧义这一特定问题,无法有效引导模型学习歧义实体的跨语言映射规则. 在实体歧义场景下,即使译文整体通顺,关键实体的误译仍可能导致严重语义偏差. 现有 RL 方法缺乏针对实体语义一致性的显式建模,难以引导模型在生成过程中感知并正确消解实体歧义.

不同于上述方法,本文提出的基于多粒度知识对齐的低资源机器翻译方法通过深度融合多粒度对齐技术与强化学习框架,构建一种面向实体语义一致性的新型奖励机制. 该方法利用双语实体词典中的注释信息提取跨语言实体映射关系,通过将词汇、实体短语与句子层级的对齐一致性奖励,引导大语言模型在生成过程中感知并准确翻译歧义实体. 实验表明,该方法在多个低资源语言对上显著提升实体翻译的准确性与语境适配性,有效缓解复合实体歧义偏差问题.

2 基于多粒度知识对齐的低资源机器翻译模型

不同于 Li 等[12]提出的 TAT - R1 模型,本文方法首先将源文本输入模型生成预测翻译,然后提取源语句中被误译的词和实体作为难词. 其次,针对实体利用检索增强技术从双语实体词典中检索出与之对应的相关的实体,如果有多条候选,进一步的,为了消除实体歧义,将源语句与检索到的实体注释进行相似度计算,根据相似度值进行重排序,最终确定与源语句中实体对应的目标语言的实体;对于普通词,则通过查询双语字典进行匹配. 再次,将对齐度、思考过程的质量、格式规范性(Format)和整体语义质量(COMET),形成总体奖励信号;最后,利用该信号指导模型进行强化训练,提升实体翻译的准确性与上下文适配性. 模型流程如图 1 所示.

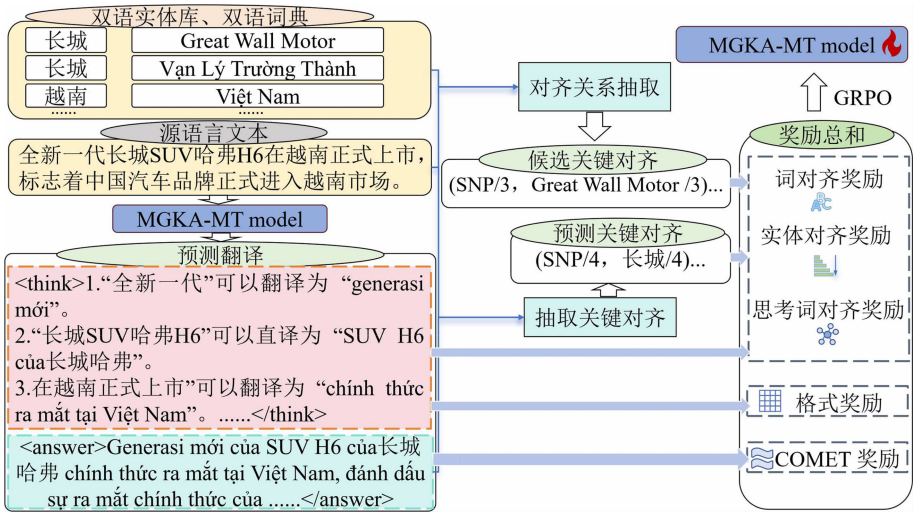


图 1　基于多粒度知识对齐的翻译模型架构图

Fig. 1　Architecture of multi – granularity knowledge alignment for machine translation

2.1　多粒度对齐

2.1.1　词对齐

词对齐作为机器翻译的核心技术之一,通过分析双语实体词典和双语词典确定跨语言词汇间的翻译等价关系. 有别于 TAT – R1 模型,所提方法利用词对齐技术设计了三种不同的奖励机制,旨在提升翻译质量,特别是实体翻译的准确性. 这三种基于词对齐的奖励机制将在下文中详细介绍,它们分别从不同角度评估翻译的准确性和一致性,为模型的强化训练提供更精准的反馈信号.

先对源文本进行分词处理. 接着,利用词对齐模型获取源文本与模型预测译文之间的词级对齐关系. 同时,基于双语词典构建源文本的对等翻译目标词集合,作为候选译文集. 假设源文本的分词序列为 $\boldsymbol{S}$,$\boldsymbol{s}_i$ 为 $\boldsymbol{S}$ 中的第 i 个词,则源文本分词序列可表示为:

$$\boldsymbol{S}=[\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{s}_3,\cdots,\boldsymbol{s}_i,\cdots,\boldsymbol{s}_N] \tag{1}$$

其中 N 为源序列长度. 同理,候选译文集与预测译文的分词序列可分别表示为:

$$\boldsymbol{C}=[\boldsymbol{c}_1,\boldsymbol{c}_2,\boldsymbol{c}_3,\cdots,\boldsymbol{c}_m,\cdots,\boldsymbol{c}_M] \tag{2}$$

$$\boldsymbol{P}=[\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3,\cdots,\boldsymbol{p}_k,\cdots,\boldsymbol{p}_K] \tag{3}$$

其中:M 为候选分词长度,K 为预测分词长度. 将源文本、候选译文集与预测译文的分词序列输入词对齐模型,可得到词级对齐关系:

$$A_{\text{can}}=Align(\boldsymbol{C},\boldsymbol{P})=[\cdots,A_{\text{can}_{mk}},\cdots] \tag{4}$$

$$A_{\text{pre}}=Align(\boldsymbol{S},\boldsymbol{P})=[\cdots,A_{\text{pre}_{ik}},\cdots] \tag{5}$$

其中:A_{can} 表示候选译文集与预测译文的对齐关系,A_{pre} 表示源文本与预测译文的对齐关系;

$A_{\text{can}_{mk}}$ 表示候选译文集第 m 个词与预测译文的 k 个词对齐,$A_{\text{pre}_{ik}}$ 表示源文本第 i 个词与预测译文第 k 个词对齐.

2.1.2　实体对齐

利用 LLaMA3 – 70B – Instruct 对输入的源语句进行分析,提取句子中的实体,提示样例如图 2 所示.

分析下面句子，标记出句子中的命名实体，不需要做任何解释：
全新一代长城SUV哈弗H6在越南正式上市，标志着中国汽车品牌正式进入越南市场。

图 2　基于 LLaMA3 – 70B – Instruct 的中文实体识别

Fig. 2　Chinese named entity recognition based on LLaMA3 – 70B – Instruct

针对识别出的实体利用检索双语实体词典中检索出与之匹配的实体,如果遇到一词多义的现象,则进一步地利用大模型获取源语句与实体的注释的相关性,记为 ρ,ρ 越大表示越相关,如图3所示. 假设源文本的实体序列为 $\boldsymbol{S}_t$,通过双语实体词典查询与 $\boldsymbol{S}_t$ 对应的目标语言实体,记为实体的候选集 $\boldsymbol{C}_t$,表示为:

$$\boldsymbol{C}_t = (\boldsymbol{C}_1, \boldsymbol{C}_2, \cdots, \boldsymbol{C}_n) \tag{6}$$

```
你是一个文本相关性评估专家。请严格评估以下源句子和候选集合中候选项的
相关性，并输出一个0到1之间的分数，1表示完全相关，0表示完全不相关。只
需输出分数，不要输出其他任何内容：
源句子：全新一代长城SUV哈弗H6在越南正式上市，标志着中国汽车品牌正式
进入越南市场。

候选集：
{1.[又称万里长城，是中国古代的军事防御工事，是一道高大、坚固而且连绵不
断的长垣，用以限隔敌骑的行动。]
 2.[长城汽车（Great Wall Motor）是成立于1984年的中国汽车品牌]
 3.[《长城》是由吉林动画学院投拍的神话题材动画电影。]
.....}
```

图3 中文句子相关性判断提示样例

Fig. 3 Prompt templates for Chinese sentence relevance judgment

首先对源文本抽取需要对齐的关键元素,表示为 $A_{\text{can}_{\text{key}}}$. 然后将源语句和初步翻译输入到 Qwen3 – Coder – 480B – A35B – Instruct 中,从而获取难词,同理,将难词作为需要对齐的关键元素,表示为 $A_{\text{pre}_{\text{key}}}$. 这是因为难词是模型翻译过程中的主要错误来源,对其进行针对性优化至关重要. 经过此步骤后,关键对齐关系是原始对齐关系的子集,且对齐的源文本词汇均为名词. 将 $A_{\text{can}_{\text{key}}}$ 和 $A_{\text{pre}_{\text{key}}}$ 定义为关键对齐关系集合. 随后,利用 $A_{\text{can}_{\text{key}}}$ 和 $A_{\text{pre}_{\text{key}}}$ 算对齐奖励.

2.2 强化学习奖励机制

如图1所示,在强化学习训练中使用的奖励包括四部分:格式奖励和思考过程奖励、COMET 奖励和对齐奖励.

2.2.1 格式奖励

给定源语句 s,模型首先思考如何用思维链(定义为 t_h)翻译 s 并提供相应的翻译(定义为 T). 生成格式可以定义为 < think > t_h </think > 、< Answer > T </Answer > ,其中 < think > 和 </think > 、< Answer > 和 </Answer > 分别是四个特殊 token,用于标记 t_h 和 T 的边界. 采用类似 TAT – R1 模板,要求模型在 < think > </think > 标签内输出推理过程,在 < Answer > </Answer > 标签内输出翻译结果. 其中,'源语言文本'表示输入文本. 翻译模板提示如图4所示.

```
你现在是一个[中文]到[越南语]的专业翻译助手。你需要先思考推理过程，并向
用户提供答案。
推理过程需包含在< think ></think >标签内，
结果需包含在< Answer ></Answer >标签内，
即： < think >推理过程</think >
     < Answer >翻译结果</Answer > 。
用户： 将以下文本翻译成[越南语]，无需额外解释：
[全新一代长城SUV哈弗H6在越南正式上市，标志着中国汽车品牌正式进入越南
市场。]
助手：
```

图4 翻译提示模版

Fig. 4 Translation prompt template

格式奖励和思考过程奖励:首先使用规则验证生成实体格式的正确性,并将格式奖励定义为:

$$r_{\text{format}} = \begin{cases} 1, \text{if format is correct} \\ 0, \text{if format is incorrect} \end{cases} \tag{7}$$

格式奖励是强化学习中一种基本且常用的奖励,实体要么是名词,要么是名词短语. 目的是确保策略模型输出实体的格式正确.

2.2.2 思考过程奖励

遵循 Wang 等[17]的方法,采用 DeepSeek - v3(671B)(Liu 等[22])评估思考过程(即 t_h)的质量. 通过精心设计的提示,DeepSeek - v3 可以将 t_h 分为三类:无分析、轻微分析和详细分析. 因此,将思考奖励定义为:

$$r_{\text{thought}}=\begin{cases}3, \text{if } v3^{\text{th}}(s,t_h)=\text{detailed analysis}\\ 2, \text{if } v3^{\text{th}}(s,t_h)=\text{slight analysis}\\ 1, \text{if } v3^{\text{th}}(s,t_h)=\text{no analysis}\end{cases} \tag{8}$$

其中 $v3^{\text{th}}$表示使用 DeepSeek - v3 作为思考奖励评判者. 正如 Wang 等人[17]所指出的,思考奖励可以鼓励策略模型生成有意义的思考内容,这对提高目标语言实体的消岐很重要.

2.2.3 COMET 奖励

COMET 是机器翻译中广泛使用的语义层面评估指标,其作为奖励的有效性已在 He 等[14]和 Feng 等[16]的研究中得到验证. 本文采用 COMET - 22 作为奖励函数的组成部分. 为保证训练稳定性,本文参考 He 等[14]的方法,公式如式(9):

$$R_{\text{comet}}(s,T)\in[0,1] \tag{9}$$

在机器翻译的评估体系中,传统的语义评估指标如 COMET 虽然能够衡量模型的整体翻译质量,但在捕捉局部信息特别是实体翻译准确性方面存在明显不足. COMET 指标侧重于句子级别的语义一致性,因此在评估特定实体的方面存在局限. 相比之下,基于 n - gram 的 BLEU 指标虽然通过计算参考翻译与模型输出的翻译准确性 n - gram 重叠度来评估翻译质量,但这种评估方式对翻译任务而言要求过于严格.

与数学或编码任务不同,翻译任务允许存在多种有效的表达形式,这使得 BLEU 等强制要求严格 n - gram 匹配的指标在应用中可能产生负面效果. 例如,一个中文句子可以改写成多个句法结构不同但语义一致的句子,那么与之对应的翻译则会有多个不同的表达,而 BLEU 可能会对这些合理的变体给出较低的评分. 当模型输出与参考译文存在合理的结构差异时,BLEU 评分常常低估翻译质量.

通常情况下,实体在句子结构上和语义表达上发挥着关键作用,因此实体的精准翻译与否影响着翻译的整体质量和用户的体验,尤其是在专业领域的翻译任务中,实体的翻译准确与否尤为关键. 通过在奖励机制中融入专门评估实体准确性的指标,可以在不损害整体语义质量的前提下显著提升实体翻译的准确性.

2.2.4 对齐奖励

如图 1 所示,设置了三种对齐奖励:词对齐奖励(R_{aaw})、实体对齐奖励(R_{eaw})和思考词对齐奖励(R_{taw}).

R_{aaw}反映参考关键对齐与预测关键对齐的词汇重叠率,该奖励鼓励模型在输出中准确翻译关键信息,计算公式如式(10):

$$R_{\text{aaw}}=\frac{len(A_{\text{can}_{\text{key}}}\cap A_{\text{pre}_{\text{key}}})}{len(S)+len(P)} \tag{10}$$

分母中加入模型输出分词序列长度,以防止模型通过生成过长输出“骗取”奖励.

R_{eaw}反映实体候选集与 < Answer > </Answer > 标签内文本的实体重叠率,该奖励鼓励模型生成最相关的实体,计算公式如式(11):

$$R_{\text{eaw}}=\frac{\text{答案中的出现在实体候选集中的实体个数}}{\text{实体候选集中实体个数}} \tag{11}$$

R_{taw}反映候选关键对齐与 < think > </think > 标签内文本的词汇重叠率,以及模型的思考过程是否充分;该奖励鼓励模型在输出最终答案前思考如何翻译关键信息,计算公式如式(12):

$$R_{\text{taw}}=0.6\,\frac{\text{候选关键对齐集中的词对在思考过程中同时出现的次数}}{\text{候选关键对齐集中词数}}+0.4r_{\text{thought}} \tag{12}$$

其中,“候选关键对齐集中出现在思考过程中的次数”是指参考关键对齐中的词对在 < think > </think > 标签内文本中同时出现的次数.

2.2.5 总奖励

综合上述奖励,总奖励的计算公式如式(13):

$$R_{\text{all}}=\begin{cases}0,\text{if } R_{\text{format}}=0\\ R_{\text{comet}}+\alpha\times R_{\text{aaw}}+\beta\times R_{\text{eaw}}+\gamma\times R_{\text{taw}},\text{if } R_{\text{format}}=1\end{cases} \tag{13}$$

其中,超参数 α、β、γ 用于平衡不同奖励的权重.

2.3 模型训练

借鉴 Li 等[12]方法,翻译模型采用 Shao 等[23]提出的分组相对策略优化(GRPO)方法进行训练. 训练过程中,对于每个输入问题 q,从当前策略模型 $\pi_{\theta_{\text{old}}}$ 生成一组候选输出 $\{o_1,o_2,\cdots,o_G\}$. 每个候选输出的优势值(A_i)通过其奖励(r_i)相对于组内均值和标准差的归一化计算得出:

$$A_i=\frac{r_i-\text{mean}(\{r_1,r_2,\cdots,r_G\})}{\text{std}(\{r_1,r_2,\cdots,r_G\})} \tag{14}$$

GRPO 通过最大化以下目标函数优化策略参数 θ:

$$J_{\text{GRPO}}(\theta)=E_{q\sim P(Q),\{o_i\}_{i=1}^{G}\sim\pi(\theta_{\text{old}}(O\mid q))}$$

$$\left[\frac{1}{G}\sum_{i=1}^{G}\min\left(\frac{\pi_\theta(o_i\mid q)}{\pi_{\theta_{\text{old}}}(o_i\mid q)}A_i,clip\,\frac{\pi_\theta(o_i\mid q)}{\pi_{\theta_{\text{old}}}(o_i\mid q)},1-\epsilon,1+\epsilon)A_i\right)-\beta D_{KL}(\pi_\theta\mid\pi_{\text{ref}})\right] \tag{15}$$

其中:ϵ 为控制策略更新的裁剪范围,通过限制剧烈变化保证训练稳定性;β 用于缩放 KL 散度惩罚项 D_{KL},约束策略与参考策略 π_{ref}(通常为初始策略)的偏差程度. KL 项近似为:

$$D_{KL}(\pi_\theta\mid\pi_{\text{ref}})=\frac{\pi_{\text{ref}}(o_i\mid q)}{\pi_\theta(o_i\mid q)}-\log\left(\frac{\pi_{\text{ref}}(o_i\mid q)}{\pi_\theta(o_i\mid q)}\right)-1 \tag{16}$$

这些超参数的设置遵循近端策略优化的相关研究[24-25],以平衡探索与稳定性.

3 实 验

3.1 实验数据

为了验证所提方法的有效性,本文在中英、中越、中老三个翻译任务上进行了实验.

强化学习训练数据:强化学习训练期间,双语实体词典,人工构建的实体词典,包含人名(PER)、地名(LOC)、组织机构名(ORG)、时间(TIME)、地缘政治实体(GPE)、军事装备实体(EQU)等 6 类实体. 对所有的中文实体,通过百度百科构建与之对应的注释. 与 Nguefack 等[20]的设置一样,中英、中越词典使用 LAMPLE 等[26]提供的真实双语词典. 这些词典是通过 Meta 的内部翻译工具精心制作,特别关注词汇的多义性处理,确保准确呈现词汇的多种含义. 中老词典是人工构建(30 万). 中 - 英监督微调数据来自于 IWSLT17 数据集;中 - 越监督微调数据是利用 Zhuang 等[27]的方法从 CCmatrix[28]随机抽取的 20 000 对平行句对,中老的监督微调数据是通过人工构建的 4 000 对平行句对,每个句对均由精通汉语 - 老挝语的专业人士进行校对.

测试数据:通过人工构建了每个语言对的实体评测数据;每个评测集均通过相应的母语专业人士进行校对. 评测集构建过程中,平行数据来自于高质量亚洲语言树库. 设定每个类型的实体数不少于 50 个,详细信息如表 1 所示.

表 1 双语实体数据及测试数据统计

Tab. 1 Bilingual entity data and test data statistics

测试数据	中 - 英	中 - 越	中 - 老
双语实体数据	464 677	305 588	208 896
评测数据	1 200	1 000	800

3.2 评价指标

在翻译质量评估方面,采用 sacreBLEU[29]和基于参考的 Comet 评分(Rei 等[15])作为评估指标,对于实体还对其准确率(P)进行评估.

3.3 实验设置

强化学习训练:总奖励中的超参数 α、β、γ 分别设为 1、1、1/10;采用 verl[30] 实现的 GRPO 强化学习算法,rollout 数量设为 16. 使用 2 卡 4090 GPU 训练模型. 设置批量大小为 128,最大生成长度为 4 096 个 token,学习率为 1e -6,采样温度设为 1.0,KL 散度损失系数为 1e -3. 训练轮数设置为 3.

为实现源文与译文间的精准词汇对应,在词对齐环节选用开源无监督词对齐工具 SimAlign[31]. 该工具在英语 -汉语等多语言对的对齐任务中表现优异,为跨语言词汇匹配提供了可靠基础. 在输入处理阶段,中文文本经由 Jieba[32] 分词器进行切分,英文文本则采用 NLTK[33] 分词工具完成分词处理,老挝语则采用实验室开发的工具进行分词[34].

选用 Qwen2.5 -7B -Instruct[35] 作为基础模型,主要因其在参数量相近的开源模型中展现出卓越的多语言处理能力,从而有效规避因模型基础性能局限可能引发的潜在缺陷.

3.4 与其他工作的对比实验

为验证所提方法的有效性,与现有的大模型做了比较.

LLaMA3.1 -8B -Instruct:具有 8B 参数的 LLaMA3 指令调优模型;Qwen2.5 -7B -instruct:阿里云发布的大语言模型系列之一,该 7B 模型在编码和数学等领域具有显著改进的能力,还提供了多语言支持;Qwen3 -8B:通义千问系列的最新大语言模型,拥有 8.2B 参数量.

首先在中英、中越、中老三个语言对的测试数据集上对比所提出的模型与多个基线模型的性能,包括通用大语言模型(LLaMA3.1 -8B -Instruct、Qwen3 -8B、Qwen2.5 -7B -Instruct)和监督微调模型(SFT),在此基础上进行的多粒度知识对齐强化学习模型(MGKA -MT). 实验结果如表 2 所示,MGKA -MT 在所有语言对上均取得不同程度的性能提升,充分验证了所提方法的有效性.

表 2　模型在中英、中越、中老翻译任务上的性能

Tab.2　Performance of various models on Chinese -English, Chinese -Vietnamese and Chinese -Lao translation tasks

模型	中 -英		中 -越		中 -老	
	BLEU	COMET	BLEU	COMET	BLEU	COMET
LLaMA3.1 -8B -Instruct						
-	22.30	59.20	13.54	53.70	8.78	32.71
SFT	23.97	60.17	14.45	54.43	40.13	40.11
MGKA -MT	27.08	64.39	18.92	61.22	**12.56**	**41.55**
Qwen3 -8B						
-	29.15	72.31	24.94	56.18	3.31	28.50
SFT	30.56	74.28	25.98	58.34	6.70	37.65
MGKA -MT	**33.65**	**75.00**	**28.34**	61.47	10.55	40.60
Qwen2.5 -7B -Instruct						
-	28.64	64.22	23.46	58.45	3.12	22.95
SFT	29.41	65.53	24.77	60.17	7.64	33.40
MGKA -MT	32.17	68.34	26.91	**62.29**	11.87	38.67

从不同基础模型的性能变化来看:

LLaMA3.1 -8B -Instruct 系列:该模型在未经微调的原始状态下表现相对较弱,特别是在低资源的中老翻译任务上,BLEU 仅为 8.78,COMET 为 32.71. 经过监督微调后,性能有所提升,中老翻译的 BLEU 提高至 10.13,COMET 达到 40.11. 而应用本文的 MGKA -MT 方法后,性能得到显著改善,BLEU 进一步提升至 12.56,COMET 达到 41.55.

Qwen3 -8B 系列:该模型展现出更强的多语言翻译能力,在中英和中越任务上表现优异. 原始模型在中英翻译上取得 29.15 的 BLEU 和 72.31 的 COMET,但在中老低资源任务上出现明显性能下降,BLEU 仅为 3.31. 经过 SFT 微调后,中老翻译性能有所改善,BLEU 达到 6.70. 应用 MGKA -MT 方法后,中老翻译

BLEU 大幅提升至 10.55,相比原始模型提升 7.24,充分证明了本方法在低资源场景下的适用性.

Qwen2.5 – 7B – Instruct 系列:该模型在参数量相对较小的情况下仍展现出较强竞争力的性能. 原始模型在中英和中越任务上分别取得 28.64 和 23.46 的 BLEU 分数. 经过 MGKA – MT 方法优化后,在所有语言对上均实现稳定提升,特别是在中老低资源任务上,BLEU 从 3.12 提升至 11.87,COMET 从 22.95 提升至 38.67.

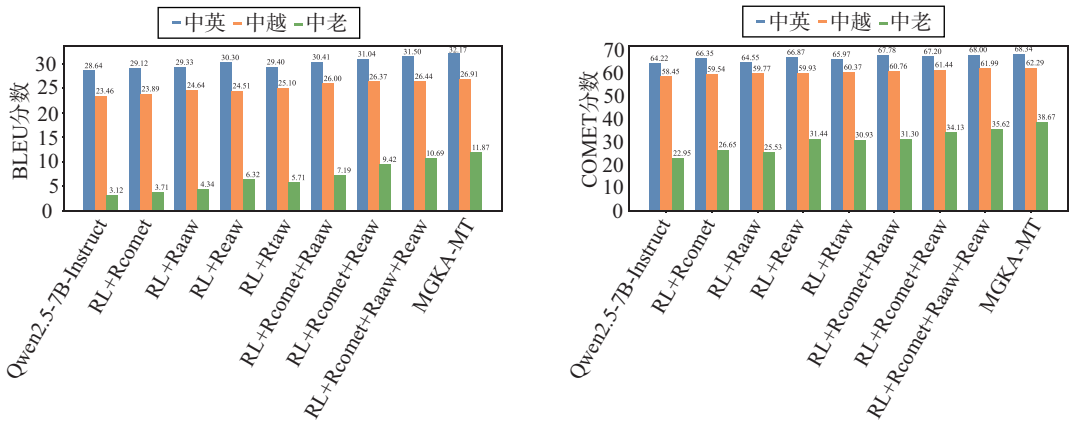
在所有基础模型上,监督微调均带来了一定的性能提升,但提升幅度有限. 以 Qwen2.5 – 7B – Instruct 为例,SFT 在中英、中越、中老任务上的 BLEU 提升分别为 0.77、1.31 和 4.52 个值. 这一结果表明,传统的监督微调虽然能够在一定程度上改善翻译质量,但在低资源语言对上难以实现突破性进展,特别是在处理实体歧义等复杂语义问题时表现有限. 传统基线模型的实体跨语言映射能力的提升主要来自于大规模平行语料,因此,当训练数据较少时模型的翻译性能会受到严重影响. 与监督微调相比,提出的多粒度知识对齐与强化学习方法展现出明显优势. 在 Qwen2.5 – 7B – Instruct 模型上,MGKA – MT 相比 SFT 在中英、中越、中老任务上的 BLEU 分别额外提升 2.76、2.14 和 4.23 个值,COMET 分别提升 2.81、2.12 和 5.27 个值. 这一结果充分证明了强化学习框架与多粒度知识对齐奖励机制的有效性.

分析认为,所提方法的主要优势在于:1)通过双语词典能为模型提供基础的跨语言对齐基础映射;2)通过双语实体词典及其中文注释,有效缓解了模型在生成译文过程的复合实体歧义问题,为模型提供了实体的精准跨语言映射;3)量化了词汇、实体及句子等不同级别粒度的差异,通过奖励机制显式引导模型聚焦于实体翻译的准确性;4)通过强化学习框架将有限的对齐知识转化为模型对语义的上下文理解能力,从而有效改善模型翻译性能的提升依赖大规模平行语料的现状. 从实验结果来看,BLEU 和 COMET 的提升验证了所提方法能有效改善低资源场景下因缺乏双语对齐数据导致的性能差的问题. 总体情况来看,所提方法通过多粒度对齐奖励机制将一定数量的双语实体词典、词典等外部知识量化到模型的强化学习过程中,可以有效缓解模型对大规模平行语料的高度依赖,提升模型的翻译性能.

4 分析实验

4.1 消融实验

为深入探究所提方法(MGKA – MT)中不同模块的有效性及相互协同关系,还进行了系统的消融实验,分别去除不同对齐奖励模块,对比不同消融模块在中英、中越、中老三个语言对上的翻译性能的好坏(以 Qwen2.5 – 7B – Instruct 为基础模型),结果如图 5 所示.



(a)各种方法在不同语言对的 BLEU 分数 (b)各种方法在不同语言对的 COMET 分数

图 5 不同对齐粒度对翻译质量的影响

Fig. 5 Impact of multi – granularity alignment on translation quality

相比之下,在基础模型中引入词对齐奖励($RL + R_{aaw}$)后,模型的 BLEU 分数有明显的改善(中 – 英 0.69,中 – 老 1.22),但 COMET 分数只有较小的提升. 这表明通过词对齐奖励,可以约束翻译模型在生成关键词时

保持翻译的一致性,能够有效改善了 n – gram 层面的匹配精度,但在整体语义连贯性上的改善则非常有限.

在基础模型中引入实体对齐奖励($RL + R_{eaw}$)后,模型的性能提升尤为明显,在三个语言对的翻译任务上,COMET 分数的提升是最明显的(中 – 英 2.65,中 – 越 1.48,中 – 老 8.49). 表明实体对齐奖励能有效改善模型处理复合实体歧义的问题. 通过显式的实体对齐知识融合到模型中,能够鼓励模型生成正确的实体.

在基础模型中引入思考词对齐奖励($RL + R_{taw}$)后,在中等资源的中英和中越任务上表现良好,但在低资源的中老任务上效果略逊于 $RL + R_{eaw}$. 这表明模型在资源充足时能够通过思维链推理有效提升翻译质量,而在资源极度稀缺时,直接约束实体输出的奖励机制更为有效.

当多个奖励模块组合使用时,模型性能呈现不同程度的提升趋势. 与单一奖励模块相比,$RL + R_{comet} + R_{eaw}$的组合模式,在中越和中老的翻译任务上取得了更好的效果(中老 COMET 达到 34.13),这表明句子级语义奖励与实体级对齐奖励具有很好的互补性:COMET 奖励确保了译文整体的流畅度,而实体对齐奖励是模型在生成准确关键信息的保障. 在此基础上加入词对齐奖励($RL + R_{comet} + R_{aaw} + R_{eaw}$)后,模型在所有语言对的翻译任务上的 BLEU 和 COMET 分数均达到新的高度,特别是中老翻译任务中,BLEU 提升至 11.87. 说明通过多个不同粒度的奖励的联合,能够从不同粒度提升模型的翻译质量.

消融实验结果表明:提出的 MGKA – MT 方法通过多粒度对齐奖励的协同组合,能够在一定程度上缓解翻译模型处理歧义复合实体的问题. 尤其是在低资源语言的翻译任务中,MGKA – MT 能够取得媲美高参数量大模型同水平的性能,不仅验证了多粒度对齐技术在改善翻译质量的有效性,更证明了其在低资源语言翻译任务中的鲁棒性.

4.2 不同奖励对实体准确率影响分析

为了进一步探究 MGKA – MT 模型中各个奖励模块对实体翻译准确率(P 值)的影响,通过对比基础大语言模型、单一奖励模块、多个奖励模块在中英、中越、中老语言对的翻译任务上的实体 P 值差异. 进一步分析所提方法在翻译过程中对模型性能提升的影响变化(以 Qwen2.5 – 7B – Instruct 为基础模型). 实验结果如表 3 所示.

表 3　各模型在不同语言对上的实验结果

Tab. 3　Experimental results of multiple models on various language pairs

模型	中 – 英	中 – 越	中 – 老
Qwen2.5 – 7B – Instruct	68.45	57.34	47.65
$RL + R_{comet}$	69.40	60.92	49.31
$RL + R_{aaw}$	70.01	61.27	51.23
$RL + R_{eaw}$	72.37	63.18	53.77
$RL + R_{taw}$	71.39	62.42	52.80
$RL + R_{comet} + R_{aaw}$	73.20	64.70	54.44
$RL + R_{comet} + R_{eaw}$	73.68	65.94	56.15
$RL + R_{comet} + R_{aaw} + R_{eaw}$	73.87	66.32	56.52
MGKA – MT	**74.30**	**67.54**	**57.21**

从实验结果来看,与基线模型 Qwen2.5 – 7B – Instruct 相比,通过引入强化学习与奖励机制,可以有效提升命名实体翻译的准确率,这验证了强化学习框架与多粒度知识对齐奖励在缓解复合实体歧义问题上的有效性.

从不同模块组合来看,$RL + R_{comet}$ 组合:仅使用 COMET 奖励时,三个语言对的实体准确率分别达到 69.4%、60.92% 和 49.31%,相比基线提升 0.95、1.58 和 1.66 个百分点. 这表明基于句子级语义的 COMET 奖励能够在一定程度上改善翻译质量,但对实体特异性问题的针对性不足. $RL + R_{aaw}$组合:引入词对齐奖励后,三个语言对的实体准确率分别达到 70.01%、61.27% 和 51.23%,相比单一 COMET 奖励分别提高 0.61、0.35 和 1.92 个百分点. 词对齐奖励通过强化词汇级翻译等价关系,为模型提供了更细粒度的对齐信号,特别在中老低资源语言对上表现尤为明显. $RL + R_{eaw}$组合:实体对齐奖励对模型的性能提升改善最大,三个语言对分别达到 72.37%、63.18% 和 53.77%,相比基线提升 3.92、3.84 和 6.12 个百分点. 这一结果充分验证了实体对齐奖励在解决复合命名实体歧义问题中的核心作用. 通过检索增强与实体消歧机制,R_{eaw}能够精准定位实体在特定语境下的正确指代. $RL + R_{taw}$组合:加入思考词对齐奖励后模型在三个语言对上分别达到了 71.39%、62.42% 和 52.80% 的准确率,性能介于 R_{aaw}和 R_{eaw}之间. 该奖励通过约束模型在思考过程中对关键实体的分析,增强了翻译决策的可解释性. $RL + R_{comet} + R_{aaw}$组合:将 COMET 奖励与词对齐奖励结合后,实体准确率分别达到 73.20%、64.70% 和 54.44%. 相比单一最佳奖励(R_{eaw}),在中英和中越任务上分别提升 0.83 和 1.52 个百分点,体现了句子级

语义奖励与词汇级对齐奖励的互补优势. $RL + R_{comet} + R_{eaw}$ 组合:COMET 与实体对齐奖励的组合取得了73.68%、65.94%和56.15%的优秀表现,特别是在中越和中老低资源语言对上,相比 $RL + R_{eaw}$ 单一奖励提升2.76和2.38个百分点. 这表明实体对齐奖励与整体语义奖励的结合能够产生协同效应,在保持译文流畅性的同时精准处理实体翻译. 这也验证了 MGKA-MT 在低资源语言场景中翻译性能的鲁棒性.

从奖励模块实验结果来讲,实体对齐奖励设置 R_{eaw} 是促使翻译模型处理复合实体歧义能力的核心关键,多个奖励模块的叠加机制为模型性能提升提供了更有效的组合参考. 单一奖励模型中,$RL + R_{eaw}$(实体对齐词奖励)表现最优,中英实体准确率达72.37%,较 $RL + R_{comet}$(69.4%)、$RL + R_{aaw}$(70.01%)分别提升2.97、2.36个百分点,说明针对实体层级的对齐奖励能有效缓解"歧义复合实体对齐难"的问题;而多个奖励组合使得性能随奖励模块的叠加持续提升,在中英、中老翻译任务中,$RL + R_{comet} + R_{aaw} + R_{eaw}$ 的实体准确率分别达73.87%、56.52%,较单一 R_{eaw} 模型提升1.5、2.75个百分点,这表明"语义级+词汇级+实体级"的多粒度对齐奖励通过合理组合,能协同提升模型的翻译性能. 同时保障翻译过程中实体的准确性以及译文的整体质量.

MGKA-MT 在低资源语言场景中有效缓解了实体翻译过程中的准确性. 中老语对中,MGKA-MT 实体准确率(57.21%)较基础模型 Qwen2.5-7B-Instruct(47.65%)提升效果十分明显;中越语对中,其准确率(67.54%)较 $RL + R_{comet} + R_{eaw}$(65.94%)提升1.6个百分点. 这一结果充分验证了 MGKA-MT 通过多粒度对齐奖励机制能够从有限的双语实体词典中提取映射关系,并为强化学习提供了精准的实体级奖励信号,促使模型在数据稀缺场景下仍能准确捕捉歧义复合实体的语义.

4.3 实例分析

为了能够更加清晰地说明所提方法的有效性,表4给出了基线模型和 MGKA-MT 模型在中译越翻译任务上生成的译文示例.

表4 实例分析

Tab.4 Case study

模型	示例
源语言	全新一代**长城 SUV 哈弗 H6** 在越南正式上市,标志着中国汽车品牌正式进入越南市场
参考译文	Dòng xe **SUV** thế hệmới **Haval H6** của thương hiệu **Great Wall Motor (GWM)** đã chính thức ra mắt ở Việt Nam,đánh dấu sự gia nhập của hãng ô-tô Trung Quốc tại thị trường Việt Nam.
LLaMA3.1-8B-Instruct	**Haval H6 SUV** mới nhất đã chính thức được bán ra tại Việt Nam, đánh dấu sự ra mắt của thương hiệu xehơi Trung Quốc tại Việt Nam.(最新款哈弗H6 SUV在越南正式上市,标志着中国汽车品牌在越南的首次亮相.)
Qwen3-8B	Thế hệ mới nhất của mẫu **SUV Haval H6** của **Changhe** chính thức ra mắt tại Việt Nam, đánh dấu sự gia nhập chính thức của thương hiệu xe Trung Quốc vào thịtrường Việt Nam.(昌河最新一代SUV车型哈弗H6在越南正式上市,标志着中国汽车品牌正式进入越南市场.)
Qwen2.5-7B-Instruct	Dòng **SUV** mới nhất của 长 城 **Harvester H6** chính thức ra mắt tại Việt Nam, đánh dấu sự ra mắt chính thức của các thương hiệu ô tô Trung Quốc tại thị trường Việt Nam.(**长城收割机 H6** 最新 SUV 系列在越南正式上市,标志着中国汽车品牌正式登陆越南市场.)
MGKA-MT	Thế hệ mới nhất của dòng xe **SUV Haval H6** của **Great Wall Motor** chính thức ra mắt tại Việt Nam,đánh dấu việc thương hiệu ô tô Trung Quố cchính thức gia nhập thị trường Việt Nam.(**长城汽车**最新一代**哈弗 H6** SUV 在越南正式上市,标志着中国汽车品牌正式进入越南市场.)

从表4的实例中加粗部分不难看出,LLaMA3.1-8B-Instruct 虽然翻译对了"SUV 哈弗 H6",但是将关键的品牌"长城"丢失了,导致关键信息缺失;Qwen3-8B 则发生了严重的品牌误译,虽然"SUV 哈弗

H6”翻译正确,却将关键的品牌“长城”错误的翻译成了“Changhe”;Qwen2.5 - 7B - Instruct 不仅将“哈弗 H6”错误的翻译成了“Harvester H6”,关键的“长城”没有进行翻译,直接复用汉字. MGKA - MT 通过实体注释及双语实体词典,精准锁定在这句话中,“长城”在这里代表的是中国的汽车品牌,因此可以获得与其对应翻译“Great Wall Motor”,这样的翻译结果能更好地保留复合实体结构的完整性,更加符合用户的需求. 从翻译结果来看,通过引入词级、实体级对齐奖励以及思考过程奖励,并在强化学习框架下进行优化,模型能够更好地利用有限的双语知识,精准识别并翻译复合实体中的深层语义结构. 这也充分展现了通过针对实体的多粒度对齐强化学习能够提升模型的翻译能力.

5 结 论

针对低资源语言复合实体翻译过程中存在歧义的问题,本文提出基于多粒度知识对齐的低资源机器翻译方法. 通过构建多粒度知识对齐的综合奖励体系,将实体层级语义一致性显式建模信息作为强化学习优化信号,不断提升模型挖掘复合实体语义的能力,使模型能够充分学习复合实体的完整语义和结构. 实验表明,在中英、中越、中老三个语言对的翻译任务中均优于其他基线模型. 这表明词对齐奖励能够确保复合实体内部词的准确映射以维持结构完整性,实体对齐则可以精准定位实体的语义边界与对应关系,同时思考过程对齐则可以强化模型对关键实体的翻译推理,三者协同能够有效缓解复合实体在不同语境下的多义性问题.

参考文献:

[1] ZHAO W X,ZHOU K,LI J Y,et al. A survey of large language models[J]. Frontiers of Computer Science,2026,20(12):1 - 40.

[2] WANG J A,MENG F D,ZHANG Y X,et al. Retrieval - augmented machine translation with unstructured knowledge[C]//Findings of the Association for Computational Linguistics:EMNLP 2025,Suzhou,China. Stroudsburg,PA,USA:ACL,2025:5858 - 5871.

[3] MOSLEM Y,ROMANI G,MOLAEI M,et al. Domain terminology integration into machine translation:Leveraging large language models[C]//Proceedings of the Eighth Conference on Machine Translation,Singapore. Stroudsburg,PA,USA:ACL,2023:902 - 911.

[4] BOGOYCHEV N,CHEN P Z. Terminology - aware translation with constrained decoding and large language model prompting[C]//Proceedings of the Eighth Conference on Machine Translation,Singapore. Stroudsburg,PA,USA:ACL,2023:890 - 896.

[5] KIM S,SUNG M,LEE J,et al. Efficient terminology integration for LLM - based translation in specialized domains[C]//Proceedings of the Ninth Conference on Machine Translation,Miami,Florida,USA. Stroudsburg,PA,USA:ACL,2024:636 - 642.

[6] HUANG Y C,LI B H,FENG X C,et al. Aligning translation - specific understanding to general understanding in large language models[C]//Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing,Miami,Florida,USA. Stroudsburg,PA,USA:ACL,2024:5028 - 5041.

[7] ZHENG J W,HONG H H,LIU F Y,et al. Fine - tuning large language models for domain - specific machine translation[EB/OL]. Arxiv,2024:2402.15061. https://arxiv.org/abs/2402.15061.

[8] ZHANG M S,LI Z H,FU G H,et al. Syntax - enhanced neural machine translation with syntax - aware word representations[C]//Proceedings of the 2019 Conference of the North,Minneapolis,Minnesota. Stroudsburg,PA,USA:ACL,2019:1151 - 1161.

[9] OCH F J,NEY H. A systematic comparison of various statistical alignment models[J]. Computational Linguistics,2003,29(1):19 - 51.

[10] ARTETXE M,SCHWENK H. Margin - based parallel corpus mining with multilingual sentence embeddings[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics,Florence,Italy. Stroudsburg,PA,USA:ACL,2019:3197 - 3203.

[11] HASLER E,GISPERT A,IGLESIAS G,et al. Neural machine translation decoding with terminology constraints[C]//Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics:Human LanguageTechnologies,Volume 2 (short Papers),New Orleans,Louisiana. Stroudsburg,PA,USA:ACL,2018:506 - 512.

[12] LI Z,ZHENG M,SONG M Y,et al. TAT - R1:Terminology - aware translation with reinforcement learning and word alignment[EB/OL]. Arxiv,2025:2505.21172. https://arxiv.org/abs/2505.21172.

[13] GUO D Y,YANG D J,ZHANG H W,et al. DeepSeek - R1 incentivizes reasoning in LLMs through reinforcement learning[J].

Nature,2025,645(8081):633 – 638.
[14] HE M G,LIU Y L,TAO S M,et al. R1 – T1:Fully incentivizing translation capability in LLMs via reasoning learning[EB/OL]. Arxiv,2025:2502. 19735. https://arxiv. org/abs/2502. 19735
[15] REI R,STEWART C,FARINHA A C,et al. COMET:A neural framework for MT evaluation[C]//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP),Online. Stroudsburg,PA,USA:ACL,2020:2685 – 2702.
[16] FENG Z P,CAO S S,REN J H,et al. MT – R1 – zero:Advancing LLM – based machine translation via R1 – zero – like reinforcement learning[C]//Findings of the Association for Computational Linguistics:EMNLP 2025,Suzhou,China. Stroudsburg,PA,USA:ACL,2025:18685 – 18702.
[17] WANG J A,MENG F D,ZHOU J. DeepTrans:Deep reasoning translation via reinforcement learning[EB/OL]. Arxiv,2025:2504. 10187. https://arxiv. org/abs/2504. 10187.
[18] CUI G F,WANG P C,LIU Y,et al. CRPO:Confidence – reward driven preference optimization for machine translation[C]//Findings of the Association for Computational Linguistics:ACL 2025,Vienna,Austria. Stroudsburg,PA,USA:ACL,2025:560 – 574.
[19] DONG T Y,LI B,LIU J S,et al. MLAS – LoRA:Language – aware parameters detection and LoRA – based knowledge transfer for multilingual machine translation[C]//Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (volume 1:Long Papers),Vienna,Austria. Stroudsburg,PA,USA:ACL,2025:15645 – 15660.
[20] NGUEFACK I N,FINKELSTEIN M,SAKAYO T S. Pretraining strategies using monolingual and parallel data for low – resource machine translation[C]//Proceedings of the Sixth Workshop on African Natural Language Processing (AfricaNLP 2025),Vienna,Austria. Stroudsburg,PA,USA:ACL,2025:31 – 38.
[21] GHANIZADEH M A,DOUSTI M J. Dynamic jointly batch selection for data efficient machine translation fine – tuning[EB/OL]. Arxiv,2025:2511. 04406. https://arxiv. org/abs/2511. 04406.
[22] LIU A X,FENG B,XUE B,et al. DeepSeek – V3 technical report[EB/OL]. arXiv,2024:2412. 19437. https://arxiv. org/abs/2412. 19437.
[23] SHAO Z H,WANG P Y,ZHU Q H,et al. DeepSeekMath:Pushing the limits of mathematical reasoning in open language models[EB/OL]. Arxiv,2024:2402. 03300. https://arxiv. org/abs/2402. 03300.
[24] WANG J A,MENG F D,ZHOU J. ExTrans:Multilingual deep reasoning translation via exemplar – enhanced reinforcement learning[EB/OL]. Arxiv,2025:2505. 12996. https://arxiv. org/abs/2505. 12996.
[25] SCHULMAN J,WOLSKI F,DHARIWAL P,et al. Proximal policy optimization algorithms[EB/OL]. Arxiv,2017:1707. 06347. https://arxiv. org/abs/1707. 06347.
[26] CONNEA U A,LAMPLE G,RANZATO M A,et al. Word translation without parallel data[J]. Arxiv,2017:1710. 04087.
[27] ZHUANG X H,GAO S X,YU Z T,et al. Low resource neural machine translation model optimization based on semantic confidence weighted alignment[J]. International Journal of Machine Learning and Cybernetics,2024,15(10):4325 – 4340.
[28] SCHWENK H,WENZEK G,EDUNOV S,et al. CCMatrix:Mining billions of high – quality parallel sentences on the web[C]//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),Online. Stroudsburg,PA,USA:ACL,2021:6490 – 6500.
[29] POST M. A call for clarity in reporting BLEU scores[C]//Proceedings of the Third Conference on Machine Translation:Research Papers,Belgium,Brussels. Stroudsburg,PA,USA:ACL,2018:186 – 191.
[30] VERL Team. VERL:Volcano engine reinforcement learning framework[EB/OL]. [2026 – 03 – 25]. https://github. com/verl – project/verl,2024.
[31] JALILI SABET M,DUFTER P,YVON F,et al. SimAlign:high quality word alignments without parallel training data using static and contextualized embeddings[C]//Findings of the Association for Computational Linguistics:EMNLP 2020,Online. Stroudsburg,PA,USA:ACL,2020:1627 – 1643.
[32] SUN F. Jieba:Chinese word segmentation tool[EB/OL]. [2026 – 03 – 25]. https://github. com/fxsjy/jieba,2024.
[33] NLTK Team. Natural language toolkit[EB/OL]. [2026 – 03 – 25]. https://www. nltk. org/,2024.
[34] 昆明理工大学自然语言处理实验室. 老挝语分词工具[EB/OL]. [2026 – 03 – 25]. http://222. 197. 219. 20:12080/.
[35] TEAM Q. Qwen2 technical report[EB/OL]. Arxiv,2024:2407. 10671. https://arxiv. org/abs/2407. 10671.

(编 辑:朱银周)