

面向时序知识图谱推理的本体-图谱双智能体强化学习方法

司悦航¹ 曾泽凡¹ 王博岩² 霍 静² 高 阳² 黄金才¹

¹(国防科技大学系统工程学院 长沙 410073)

²(计算机软件新技术全国重点实验室(南京大学) 南京 210023)

(siyuehang@126.com)

Ontology-Graph Dual-Agent Reinforcement Learning Method for Temporal Knowledge Graph Reasoning

Si Yuehang¹, Zeng Zefan¹, Wang Boyan², Huo Jing², Gao Yang², and Huang Jincai¹

¹(College of Systems Engineering, University of Electronic National University of Defense Technology, Changsha 410073)

²(National Key Laboratory of Novel Software Technology (Nanjing University), Nanjing 210023)

Abstract Temporal knowledge graph (TKG) reasoning aims to predict future events based on historical facts. However, existing reinforcement learning (RL)-based reasoning methods merely conduct path searches at the entity level, lacking high-level semantic and logical guidance for the reasoning process. Moreover, ontologies, as conceptual abstractions of graphs, have strong logicity, but their construction typically relies on static statistical rules and is difficult to adapt dynamically to varying data distributions. To effectively address these issues, we propose ODRT, an ontology-graph dual-agent reinforcement learning method for TKG reasoning. Specifically, it jointly initializes the TKG ontology by leveraging both large language model (LLM) semantic embeddings and pretrained TKG reasoning model structural embeddings. Furthermore, it dynamically optimizes the ontology structure through learnable frequency weights during the end-to-end training process. Within the proposed dual-agent RL framework, the upper-level ontology agent performs concept-level walks on the constructed ontology graph, providing effective semantic navigation for the lower-level graph agent through carefully designed shared historical states. Meanwhile, the lower-level graph agent conducts fine-grained entity-level searches on the temporal graph under the constraint of hierarchical soft reward functions. Extensive experiments on four publicly available reasoning datasets demonstrate that ODRT achieves state-of-the-art performance among existing RL-based baselines and reaches the current state-of-the-art level on two encyclopedic datasets. However, on event-based datasets with heterogeneous entity behavior patterns and predictions that rely on historical reuse, its performance still lags behind that of partially embedded methods. Comprehensive experimental analyses further verify the significant advantages of the proposed ontology-guided mechanism and the dual-agent reinforcement learning framework.

Key words knowledge graph; temporal knowledge graph reasoning; ontology; reinforcement learning; dual-agent

摘 要 时序知识图谱推理旨在基于历史事实预测未来事件,而现有基于强化学习的推理方法仅在实体级进行路径搜索,缺乏高层语义逻辑指导。本体作为图谱的概念级抽象具有很强的逻辑性,而本体构建时依赖静态统计,难以动态适应。为此,提出面向时序知识图谱推理的本体-图谱双智能体强化学习方法

收稿日期: 2026-06-04; 修回日期: 2026-07-16

基金项目: 江苏省科技重大专项(BG2024031); 教育部基础学科和交叉学科突破计划(JYB2025XDXM118)

This work was supported by the Jiangsu Science and Technology Major Project (BG2024031) and the Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China (JYB2025XDXM118).

通信作者: 黄金才(huangjincai@nudt.edu.cn)

ODRT。该方法通过大语言模型语义嵌入与预训练时序知识图谱推理模型结构嵌入联合初始化时序知识图谱本体,并在训练中通过可学习频次权重实现本体动态优化。双智能体强化学习框架中,上层本体智能体在本体图上执行概念级游走,通过历史状态共享为下层图谱智能体提供语义导航,下层智能体在时序图谱上执行实体级搜索并受分层软奖励约束。在4个公开的推理数据集上的广泛实验表明,ODRT在现有基于强化学习的时序推理方法中取得最优性能,并在2个百科型数据集上达到当前最优水平;而在实体行为模式异质且预测依赖历史复用的事件型数据集上,其性能与部分嵌入方法相比存在一定差距。深入实验分析证明了本体引导机制和双智能体强化学习框架的优势。

关键词 知识图谱;时序知识图谱推理;本体;强化学习;双智能体

中图法分类号 TP391

DOI: 10.7544/issn1000-1239.202660408

CSTR: 32373.14.issn1000-1239.202660408

知识图谱(knowledge graph, KG)以结构化三元组形式组织海量实体及其语义关联,已成为推荐系统、智能问答、信息检索等人工智能应用的重要知识基础设施^[1]。然而,现实世界中的事实持续动态演化,具有显著的时间依赖性。因此,时序知识图谱(temporal knowledge graph, TKG)在事实中引入时间戳,以(主语,关系,宾语,时间)四元组形式记录事件的发生时刻,为动态知识推理提供了更贴合实际的数据表示^[2]。图1展示了一个TKG的具体例子。

TKG推理旨在基于历史观测事实推断未知或未来发生的事件,要求模型预测未来时间戳下缺失的事实,对于事件预测、辅助决策等场景具有更高的实际应用价值。近年来,基于强化学习(reinforcement learning, RL)的路径推理方法因其具备显式推理过程与良好的可解释性,在TKG推理任务中受到广泛关

注。该类方法将推理过程建模为智能体在历史图谱快照上的序列决策过程,通过策略网络学习从多跳路径中搜索时序证据链,以回答关于未来的查询^[3]。

然而,现有基于强化学习的推理方法主要聚焦于实体级的单智能体游走,缺乏高层语义抽象,导致2个固有缺陷。其一,当目标答案需通过较长路径到达时,实体级动作空间随路径长度指数膨胀,智能体难以在庞大候选集合中持续做出正确决策,推理性能随路径增长急剧衰减。其二,智能体的推理路径仅为实体与关系的机械串联,缺乏可解释的概念级逻辑依据,无法映射到领域语义模式,难以满足实际应用中推理逻辑的需求。

现有工作尝试引入本体作为领域概念以辅助TKG推理。本体通过将实体泛化为概念,能够提取跨时间稳定的事件模式,如图1所示。然而,现有本

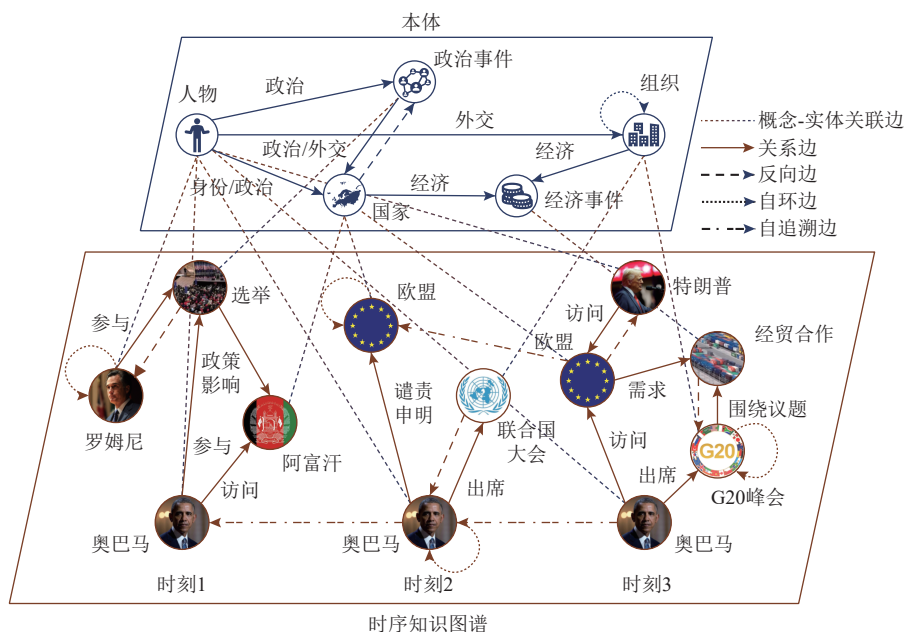


Fig. 1 Examples of ontology and temporal knowledge graphs

图1 本体与时序知识图谱示例

体构建方法仍存在明显局限:概念间的逻辑关系多基于训练集中的统计频次生成,容易因样本量悬殊而忽略低频但重要的语义模式,导致本体结构有偏。同时,本体一旦构建即为静态,无法根据下游推理任务的反馈进行动态修正,难以灵活应对事件演化过程。

针对上述问题,本文提出面向 TKG 推理的本体-图谱双智能体强化学习方法 ODRT(ontology-graph dual-agent reinforcement learning method for TKG reasoning)。其核心思想是:构建与 TKG 同构的语义抽象层,引入上层本体智能体对下层图谱智能体的推理过程进行逻辑导航。本体通过大语言模型(LLM)语义嵌入与预训练 TKG 推理模型结构嵌入联合初始化,并在训练中通过可学习频次权重实现动态调整。本体智能体与图谱智能体通过历史编码共享实现跨层协同训练,实现本体优化与时序推理。

本文的主要贡献包括:

1)提出一种结合 LLM 语义嵌入与 TKG 推理模型结构嵌入的本体建模方法,通过可学习频次权重在训练中动态优化本体,克服静态统计本体的偏差与固化问题;

2)设计本体-图谱双智能体强化学习框架 ODRT,上层本体智能体在概念级游走提供语义导航,下层图谱智能体在实体级游走执行时序推理,通过状态共享实现协同训练,同时完成本体学习与图谱推理;

3)在 4 个公开 TKG 推理数据集上进行大量实验,验证了 ODRT 在链接预测任务上的推理优势。

1 相关工作

本节对与本文工作相关的研究现状进行介绍,包括静态和 TKG 推理模型、基于强化学习的推理模型和结合本体的推理模型等内容。

1.1 知识图谱推理模型

知识图谱推理分为静态和时序 2 个方向。静态知识图谱推理旨在将实体与关系映射至低维连续向量空间,通过评分函数衡量三元组的语义合理性并推断缺失链接。TransE^[4]与 RotatE^[5]分别将关系视为头尾实体嵌入间的平移变换和旋转变换,以几何假设奠定嵌入学习的基础。ComplEx^[6]将表示空间拓展至复数域,通过复数内积区分关系的方向性来表达非对称语义。DistMult^[7]引入双线性对角化形式,利用元素级乘积刻画实体交互强度,并建模对称关系模式。ASFR^[8]利用注意力机制来提取局部关联信息和结构特征,并纳入位置信息来增强表示能力。

然而,上述静态方法假设事实不随时间变化,难以直接处理动态演化的知识场景。

TKG 推理在嵌入框架中引入时间维度,以刻画实体与关系的动态演化规律。根据预测目标的时间范围 TKG 可分为插值与外推 2 类。插值旨在补全历史时间窗口内的缺失事实,如通过张量分解评分函数实现图谱补全^[9],但无法回答未来的查询。

外推则要求基于已知历史预测未来事件,研究者从多个角度展开探索。一些模型通过引入时间偏差项到推理模型的得分函数来扩展到时序图谱的嵌入,如 TTransE^[10]、TADistMult^[11]与 T-Temp^[12]等。TANGO^[13]引入神经微分方程建立连续时间演化模型,突破离散时间快照的表示限制。此外,RE-NET^[14]、EvoKG^[15]等模型融合循环神经网络与图结构编码,从邻域聚合、演化表示等角度建模时序依赖。Coherence-Mode^[16]通过设计时间一致性评估单元挖掘深层时序关联。还有 ChapTER^[17]、PPT^[18]与 HFL^[19]等一些方法通过引入 LLM 来对复杂的时序模式进行编码,以提升时序推理能力。

然而,现有方法多聚焦于隐式表示学习或历史模式记忆,推理过程缺乏显式的逻辑化路径,难以提供可验证的决策依据。

1.2 基于强化学习的推理模型

强化学习将多跳链接预测建模为智能体在图结构上的序列决策过程,通过策略梯度优化使智能体学会逐步导航至目标实体的最优行为策略。

在静态知识图谱场景中,MINERVA^[20]将路径推理形式化为马尔可夫决策过程,利用策略网络指导智能体在离散关系空间中进行状态转移,以缓解传统嵌入方法的语义组合不足问题。CURL^[21]提出双智能体协同框架,部署负责簇级快速游走的宏观智能体与负责实体级精细搜索的微观智能体,通过共享历史状态与互增强奖励实现信息互补与长路径搜索。

面向 TKG 的动态特性,研究者将强化学习框架拓展至带时间戳的图结构。TPath^[22]将历史快照序列整合为统一搜索空间,使智能体能够在跨时间片的边上执行多步推理,并挖掘隐含于演化过程中的深层语义关联。TAgent^[23]针对时序链接预测任务设计了自适应步长调整机制,智能体根据查询的时间跨度动态决定历史回溯深度,增强对不同时间尺度事件的适应能力。TIter^[24]则引入相对时间编码刻画事实的时序距离,并基于历史分布构建时间塑形奖励,同时提出归纳均值表示以处理训练阶段未见的实体,为未来事件预测提供了可解释的时序证据链。

然而, 现有强化学习的研究在推理路径深度逻辑与奖励函数设计方面仍存在明显局限: 一方面, 推理路径局限于名称级搜索, 缺乏更深层次的概念支撑; 另一方面, 奖励的稀疏性导致长路径训练困难, 中间步骤缺乏可靠的语义反馈。

1.3 结合本体的推理模型

本体作为领域概念的形式化抽象, 能够通过层次化的类属关系与领域关系模式为实体提供语义约束。在静态知识图谱中, JOIE^[25]通过跨视图对齐机制同时学习实例层三元组结构与本体层概念关联, 建立了实体与概念联合嵌入的语义映射框架。ORT^[26]设计了本体引导的逆向思维策略, 以实现从目的到条件推理路径的构建, 并提升了LLM对于知识图谱的适应能力。JMOI^[27]利用自注意力机制挖掘概念和实例的语义关系, 并通过调整嵌入邻域范围实现不同概念信息的层次区分。

在TKG推理领域, OntoTKGE^[28]将本体知识引入TKG外推任务, 通过设计本体视图图谱, 并抽取实体在本体视图中的局部邻域子图, 捕获查询相关的语义上下文并改善稀疏实体的预测性能。

然而, 现有结合本体的推理方法存在以下局限: 一是本体结构在构建完成后即为静态, 其概念划分与关系频次无法根据下游推理反馈进行动态修正; 二是本体层与推理过程相互割裂, 隐式嵌入的约束无法在路径搜索阶段为智能体提供显式的语义导航信号。

因此, 本文重新定位本体在推理中的角色, 通过可学习的频次权重实现本体结构的动态演化, 并借助双智能体协同机制使本体层直接参与强化学习的序列决策, 从而增强推理的语义逻辑性与灵活性。

2 任务定义

本节详细定义了ODRT模型相关的概念, 包括TKG、TKG推理任务和本体等内容。

定义 1. TKG。设 $\mathcal{E}$ 、 $\mathcal{R}$ 、 $\mathcal{T}$ 分别表示实体集合、关系集合与时间戳集合。时序知识图谱 $\mathcal{G}$ 被定义为一组按时间顺序排列的离散知识图谱快照序列, 即 $\mathcal{G} = \{\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_T\}$ 。其中对于任意时间戳 $t \in \mathcal{T}$, 快照 $\mathcal{G}_t = (\mathcal{E}_t, \mathcal{R}_t, \mathcal{F}_t)$ 。这里 $\mathcal{E}_t \subseteq \mathcal{E}$ 表示在时间戳 t 出现的实体集合, $\mathcal{F}_t \subseteq \mathcal{E} \times \mathcal{R} \times \mathcal{E} \times \{t\}$ 表示在时间戳 t 观测到的事实集合, 每个事件以四元组形式表示为 (e_s, r, e_o, t) , 其中 $e_s, e_o \in \mathcal{E}$ 分别为主语实体、宾语实体, $r \in \mathcal{R}$ 为它们之间的关系, t 为事实发生的时间戳。

定义 2. TKG推理任务。本文聚焦于TKG外推

推理任务, 即基于历史观测信息预测未来时间戳发生的事实。具体地, 给定历史图谱快照序列 $\{\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_t\}$, 任务目标是回答关于未来时间戳 $t_q (t_q > t)$ 的查询, 查询形式包括主语给定查询 $(e_s, r_q, ?, t_q)$ 与宾语给定查询 $(?, r_q, e_o, t_q)$, 要求分别预测缺失的宾语实体 e_o 或主语实体 e_s , 其中 $e_s, e_o \in \mathcal{E}$, $r_q \in \mathcal{R}$, $t_q \in \mathcal{T}$ 。对于每个查询, 模型需从全体实体集合 $\mathcal{E}$ 中识别出最可能的答案实体。

定义 3. 本体。为捕获TKG中跨时间稳定的语义模式并缓解实体级推理的稀疏性与动作空间爆炸问题, 本文引入本体层对TKG进行语义抽象。设 $\mathcal{C}$ 为概念(类别)集合, 满足 $|\mathcal{C}| \ll |\mathcal{E}|$, 并定义概念映射函数 $\phi: \mathcal{E} \rightarrow \mathcal{C}$, 将每个实体 $e \in \mathcal{E}$ 映射至其所属概念 $c = \phi(e) \in \mathcal{C}$ 。基于该映射, 构建本体图 $\mathcal{G}_o = (\mathcal{C}, \mathcal{R}, \mathcal{F}_o)$, 其中 $\mathcal{C}$ 为本体节点集合, 每个节点代表一个语义概念, 由实体聚类得到; $\mathcal{R}$ 为关系集合, 与TKG中的关系集合一致; $\mathcal{F}_o \subseteq \mathcal{C} \times \mathcal{R} \times \mathcal{C}$ 为本体事实集合, 由TKG事实通过概念映射聚合得到, 即对于任意事实 $(e_s, r, e_o, t) \in \mathcal{F}_t$, 若 $\phi(e_s) = c_s$ 且 $\phi(e_o) = c_o$, 则对应生成本体事实 $(c_s, r, c_o) \in \mathcal{F}_o$ 。

值得注意的是, 本体层不含时间属性, $\mathcal{F}_o$ 中的事实不关联任何时间戳, 这种设计使得本体能够过滤掉TKG中随时间变化的细粒度噪声, 保留跨时间稳定的抽象事件模式, 从而为下游推理提供具有逻辑一致性的语义导航。由于本体关系直接来源于TKG原始关系, 本体图可视为对TKG在概念粒度上的同构投影。

3 ODRT 模型

本节首先介绍了ODRT的整体架构, 并对TKG本体构建过程和双智能体强化学习框架进行详细阐述。此外, 还对奖励函数进行重新设计, 并给出了参数学习与推理的具体细节。

3.1 ODRT 整体架构

ODRT的核心思想是通过构建与TKG同构的语义抽象层, 引入上层本体智能体对下层图谱智能体的推理过程进行逻辑导航, 从而在长路径、稀疏实体等复杂场景下实现可解释且鲁棒的时序推理。如图2所示, ODRT的整体架构由本体层与TKG层2个核心层级构成, 2层之间通过本体智能体和图谱智能体的状态传递、动作空间生成与奖励机制实现信息交互与联合优化。

在本体层, ODRT通过概念映射函数将TKG中

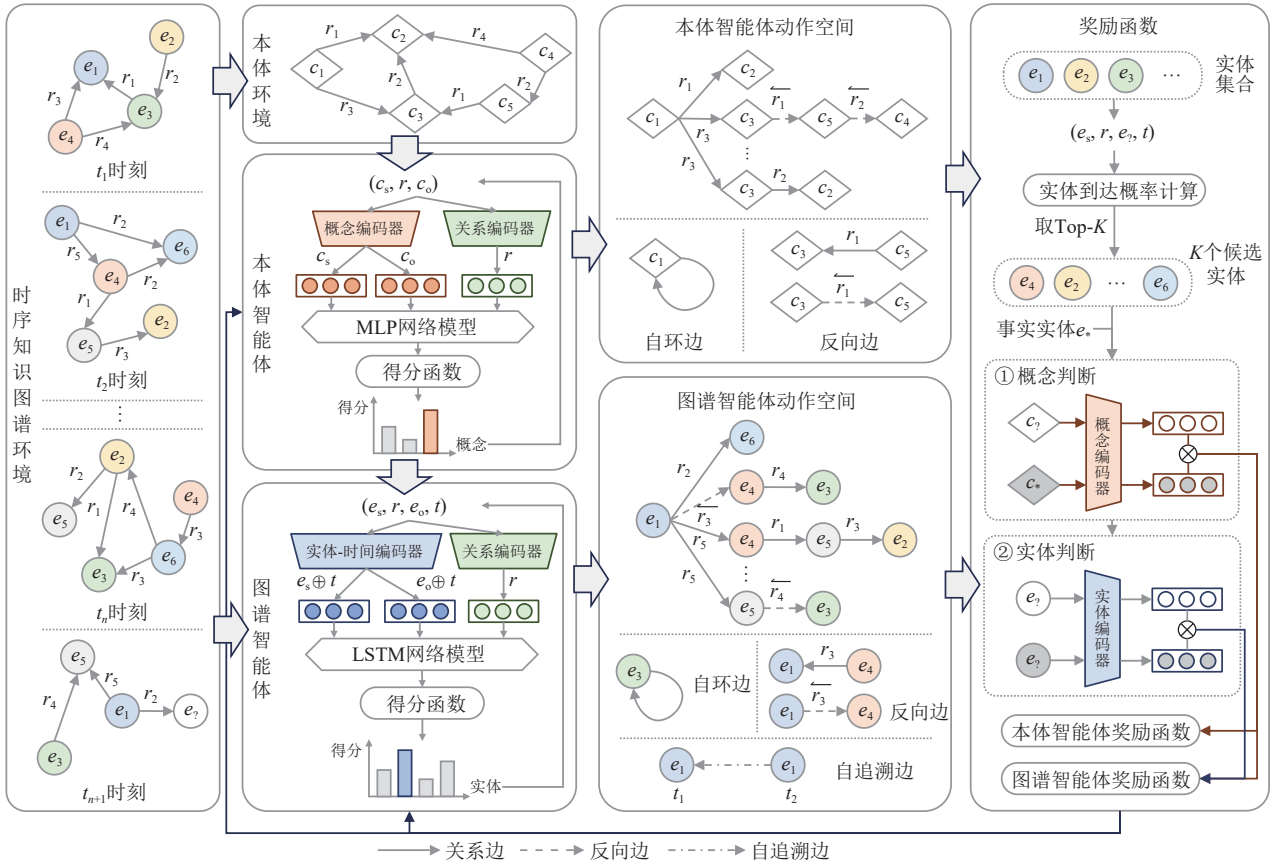


Fig. 2 Structure diagram of ODRT model

图2 ODRT 模型结构图

的实体集合抽象为规模远小于实体集的概念集合,并基于原始关系集合构建本体图。该本体层作为TKG在概念粒度上的同构投影,过滤了随时间波动的细粒度噪声,保留了跨时间稳定的抽象事件模式,从而为上层的语义导航提供了逻辑基础。值得注意的是,本体通过LLM与预训练TKG推理模型联合初始化生成,并在训练过程中通过本体智能体的训练进行动态调整和修正。

在TKG层,ODRT遵循基于路径搜索的时序推理范式,将图谱智能体在TKG历史快照上的游走过程建模为马尔可夫决策过程(MDP)。图谱智能体从查询主语实体出发,在历史事实构成的边空间中逐步选择下一跳关系-实体-时间三元组,最终在未来时间戳处停止于预测的答案实体。

与现有基于强化学习的推理方法不同,ODRT的上层本体智能体在本体图上执行并行的高层级游走,其每一步输出的目标概念与推荐关系类型作为软约束注入图谱智能体的动作选择过程,从而有效压缩候选动作空间并增强推理路径的语义逻辑性。在训练过程中,2层智能体实现了状态传递:本体智能体将其历史路径状态共享至图谱智能体的策略网络,

为实体级搜索提供阶段式的语义提示。

3.2 TKG 本体构建

ODRT的本体构建区别于传统静态本体工程与纯统计规则方法,采用“语义-结构联合初始化+训练动态优化”的两阶段策略,如图3所示。本体在初始化阶段通过LLM语义信息与TKG结构信息联合生成概念划分与本体图结构;在训练阶段,实体到概念的映射保持不变,仅通过本体智能体的学习动态调整概念间关系的频次权重,使本体能够适应不同数

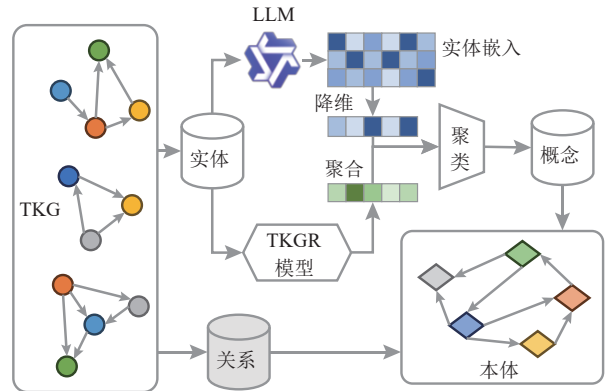


Fig. 3 Ontology construction process

图3 本体构建流程

据集的演化模式并修正初始化偏差。

1) 实体语义嵌入获取。设实体 $e \in \mathcal{E}$ 的文本描述为 D_e ，通过 LLM 编码器 f_{LLM} 将其映射为高维语义嵌入：

$$\mathbf{h}_e^{\text{sem}} = f_{\text{LLM}}(D_e) \in \mathbb{R}^{d_{\text{LLM}}}。 \quad (1)$$

对所有实体执行上述编码，得到语义嵌入矩阵 $\mathbf{H}^{\text{sem}} \in \mathbb{R}^{|\mathcal{E}| \times d_{\text{LLM}}}$ ，其中第 i 行对应实体 e_i 的语义表示。

2) 实体结构嵌入获取。采用预训练的 TKG 推理模型（如 TTransE）在历史快照上学习实体结构嵌入。通过 TKG 推理模型的实体编码器获得时间感知的结构嵌入：

$$\mathbf{s}_e^{\text{str}} = f_{\text{TKGR}}(e) \in \mathbb{R}^{d_{\text{TKGR}}}。 \quad (2)$$

提取实体结构嵌入矩阵 $\mathbf{S}^{\text{str}} \in \mathbb{R}^{|\mathcal{E}| \times d_{\text{TKGR}}}$ ，其中第 i 行为实体 e_i 的结构嵌入。

由于通常 $d_{\text{LLM}} \gg d_{\text{TKGR}}$ ，需对高维语义嵌入进行降维。对 $\mathbf{H}^{\text{sem}}$ 按列中心化后降维得到的语义嵌入为：

$$\hat{\mathbf{H}}^{\text{sem}} = (\mathbf{H}^{\text{sem}} - \bar{\mathbf{H}}^{\text{sem}}) \mathbf{P} \in \mathbb{R}^{|\mathcal{E}| \times d_{\text{TKGR}}}， \quad (3)$$

其中 $\bar{\mathbf{H}}^{\text{sem}}$ 为均值矩阵， $\mathbf{P} \in \mathbb{R}^{d_{\text{LLM}} \times d_{\text{TKGR}}}$ 是由特征分解得到的投影矩阵。将降维后的语义嵌入与结构嵌入拼接，得到联合实体表示：

$$\mathbf{J} = [\hat{\mathbf{H}}^{\text{sem}}; \mathbf{S}^{\text{str}}] \in \mathbb{R}^{|\mathcal{E}| \times 2d_{\text{TKGR}}}。 \quad (4)$$

3) 概念聚类。基于联合嵌入 $\mathbf{J}$ 执行 K-means 聚类。设聚类中心集合为 $\{\mu_1, \mu_1, \dots, \mu_K\}$ ，优化目标为最小化簇内平方和：

$$\mathcal{J} = \sum_{k=1}^K \sum_{e \in C_k} \|j_e - \mu_k\|_2^2， \quad (5)$$

其中 C_k 表示第 k 个簇包含的实体集合。迭代优化收敛后，每个实体被分至距离最近的概念：

$$\phi(e) = \arg \min_{k \in \{1, \dots, K\}} (\|j_e - \mu_k\|_2^2)。 \quad (6)$$

由此得到概念集合 $\mathcal{C} = \{c_1, c_2, \dots, c_K\}$ 与映射 $\phi: \mathcal{E} \rightarrow \mathcal{C}$ 。聚类数目 K 作为超参数按数据集规模设定，初步取值为 $K = \sqrt{|\mathcal{E}|/2}$ 。该映射在后续训练过程中固定不变。该启发式设定借鉴了聚类分析中基于样本量的经验法则，旨在平衡概念层的语义泛化能力与区分度： K 过小会导致概念过于宽泛，难以区分不同子空间内的实体行为模式； K 过大则会使概念粒度趋近实体级，削弱本体层的抽象收益。该映射在后续双智能体强化学习训练过程中固定不变，以保证训练稳定性并降低优化复杂度。

4) 本体图构建。基于概念集合 $\mathcal{C}$ 与 TKG 原始关系集合 $\mathcal{R}$ 构建本体图 $\mathcal{G}_0 = (\mathcal{C}, \mathcal{R}, \mathcal{F}_0)$ 。对于 TKG 中的所有历史事实 $(e_s, r, e_o, t) \in \mathcal{F}_{\text{train}}$ ，通过映射 ϕ 得到对应概念对 $(c_s, c_o) = (\phi(e_s), \phi(e_o))$ ，并直接保留所有出现过

的概念级事实：

$$(c_s, r, c_o) \in \mathcal{F}_0 \Leftrightarrow \exists (e_s, r, e_o, t) \in \mathcal{F}_{\text{train}}， \quad (7)$$

$$\text{s.t. } \phi(e_s) = c_s, \phi(e_o) = c_o。$$

每条本体边 $(c_s, r, c_o) \in \mathcal{F}_0$ 带一个可学习的频次权重 w_{c_s, r, c_o} ，其初始值由统计频次归一化得到：

$$w_{c_s, r, c_o}^{(0)} = \frac{\text{Count}(c_s, r, c_o)}{\sum_{(c'_s, r', c'_o) \in \mathcal{F}_0} \text{Count}(c'_s, r', c'_o)}， \quad (8)$$

其中，计数函数 $\text{Count}(c_s, r, c_o) = |\{(e_s, r, e_o, t) \in \mathcal{F}_{\text{train}} : \phi(e_s) = c_s \wedge \phi(e_o) = c_o\}|$ 。该权重作为本体图的核心组成部分，反映概念间在特定关系下的交互强度，是本体智能体进行高层级游走决策的重要依据。

5) 动态调整机制。在训练过程中，实体到概念的映射 ϕ 保持不变，动态调整的对象仅为本体图中各边的频次权重 w_{c_s, r, c_o} 。这种设计避免了引入规模庞大的隶属度矩阵并显著增加训练不稳定性，通过固定硬分配、优化边权重的方式实现了本体结构的动态演化，既保留了概念划分的一致性，又赋予本体对数据分布变化的适应能力。

具体地， w_{c_s, r, c_o} 作为本体智能体策略网络参数的一部分，通过双智能体联合优化目标进行端到端更新。设本体智能体策略参数为 θ_0 （包含所有频次权重），由奖励函数对应得到的联合损失函数为 $\mathcal{L}(\theta_0, \theta_G)$ ，则频次权重按梯度下降更新：

$$w_{c_s, r, c_o} \leftarrow w_{c_s, r, c_o} - \alpha \nabla w_{c_s, r, c_o} \mathcal{L}(\theta_0, \theta_G)， \quad (9)$$

其中 α 为学习率。通过该机制，模型能够根据双智能体协同推理的反馈自动强化与正确推理路径一致的高频本体模式，同时抑制由数据稀疏或聚类偏差导致的低频噪声边。

3.3 双智能体强化学习框架

本文将 TKG 推理建模为一个双智能体协同的马尔可夫决策过程。上层本体智能体在本体图 $\mathcal{G}_0$ 上执行高层级语义游走，输出概念级导航信号；下层图谱智能体在 TKG 历史快照序列上执行实体级时序游走，搜索未来答案。2 个智能体分别由独立的策略网络 π_θ^0 与 π_θ^G 参数化，通过状态共享实现跨层协同。

1) 环境与智能体

本体智能体感知的环境为无时间属性的本体图 $\mathcal{G}_0 = (\mathcal{C}, \mathcal{R}, \mathcal{F}_0)$ ，其以节点为概念，边为带可学习频次权重的概念-关系-概念三元组。本体智能体从查询源实体所属概念 $\phi(e_q)$ 出发，在概念集合间跳转，通过高层级语义推理压缩搜索空间。

图谱智能体感知的环境为时序知识图谱 $\mathcal{G}$ ，其状态转移发生在带时间戳的实体-关系边上。图谱智能

体从查询源实体 e_q ,在历史事实构成的边空间中逐步选择下一跳关系-实体-时间三元组,直至在未来时间戳 t_q 止于预测的答案实体。

2) 状态空间

本体智能体在时刻 t 的状态定义为:

$$s_t^O = (c_t, e_q, t_q, r_q, h_t^O), \quad (10)$$

其中 $c_t \in C$ 为当前访问的概念, (e_q, t_q, r_q) 为查询全局信息, h_t^O 为本体智能体历史路径的MLP编码。

图谱智能体在时刻 t 的状态定义为:

$$s_t^G = (e_t, t_{\text{now}}, e_q, t_q, r_q, h_t^G, h_t^O), \quad (11)$$

其中 (e_t, t_{now}) 为当前访问的实体-时间节点, (e_q, t_q, r_q) 为查询全局信息, h_t^G 为图谱智能体自身历史路径的LSTM编码, h_t^O 为从本体智能体共享而来的路径状态。

3) 动作空间

本体智能体在状态 s_t^O 下的候选动作空间 $\mathcal{A}_t^O$ 由当前概念 c_t 在本体图 $\mathcal{G}_O$ 中的所有邻接概念构成,每个候选动作表示为:

$$a_t^O = (r', c'). \quad (12)$$

对应边 $(c_t, r', c') \in \mathcal{F}_O$ 附带可学习频次权重 $w_{c_t, r', c'}$ 。动作空间包含自循环动作以实现与下层游走的阶段式对齐,执行时本体智能体保持当前概念不变。

图谱智能体在状态 s_t^G 下的候选动作空间 $\mathcal{A}_t^G$ 由当前实体 e_t 在历史快照中截至时间 t_{now} 的所有出边构成,每个候选动作表示为:

$$a_t^G = (r', e', t'), \quad (13)$$

其中,满足 $t' \leq t_{\text{now}}$ 且 $t' < t_q$ 。

4) 状态转移

本体智能体执行动作 $a_t^O = (r', c')$ 后,状态转移至:

$$s_{t+1}^O = (c', e_q, t_q, r_q, h_{t+1}^O), \quad (14)$$

其中 h_{t+1}^O 由MLP根据当前历史编码与上一动作更新得到;若执行自循环动作,则 $c_{t+1} = c_t$,仅更新历史编码与共享状态。

图谱智能体执行动作 $a_t^G = (r', e', t')$ 后,状态转移至:

$$s_{t+1}^G = (e', t', e_q, t_q, r_q, h_{t+1}^G, h_t^O), \quad (15)$$

即 $\delta^G(s_t^G, a_t^G) = s_{t+1}^G$,其中 h_{t+1}^G 由LSTM更新得到。

5) 策略网络

本体智能体的策略网络 $\pi_\theta^O(a_t^O|s_t^O)$ 建模为概念级动作的概率分布。输入为当前概念嵌入 c_t 、查询嵌入 $[e_q; r_q]$ 与历史编码 h_t^O ,通过MLP计算候选动作得分:

$$\pi_\theta^O(a_t^O|s_t^O) = \text{MLP}_O([c_t; h_t^O; e_q; r_q]). \quad (16)$$

图谱智能体的策略网络 $\pi_\theta^G(a_t^G|s_t^G)$ 建模为实体级动作的概率分布。输入为当前实体-时间嵌入 e_t^{now} (由实体

嵌入与相对时间编码拼接得到,即 $e_t^{\text{now}} = [e_t; \Phi(t_q - t_{\text{now}})]$ 、自身历史编码 h_t^G 、从本体智能体共享而来的路径状态 h_t^O 以及查询嵌入的拼接,通过LSTM计算候选动作得分:

$$\pi_\theta^G(a_t^G|s_t^G) = \text{LSTM}_G([e_t^{\text{now}}; h_t^G; h_t^O; e_q; r_q]). \quad (17)$$

6) 跨层协同机制

ODRT通过历史编码的跨层共享实现双智能体协同。图谱智能体在实体级决策时不仅依赖自身的实体级搜索历史 h_t^G ,还同步感知本体智能体的概念级推理意图。具体而言, h_t^O 与实体-时间嵌入、自身历史编码及查询嵌入拼接后输入LSTM,通过门控机制调节候选动作的概率分布:当本体智能体停留于当前概念时, h_t^O 持续累积语义证据,提升属于该概念邻域的候选实体的得分;直至本体跳转, h_t^O 的跃变驱动图谱智能体切换搜索子空间。值得注意的是,本体智能体的输出以软约束形式作用于图谱层,即图谱智能体仍保留完整的实体级动作空间, h_t^O 仅调节动作概率而非强制过滤,从而兼顾语义导航与局部搜索的灵活性。

当本体智能体与图谱智能体的推理路径出现不一致时,ODRT通过分层软奖励机制实现协调。若图谱智能体未命中目标实体,但所处概念与答案概念一致,仍可获得中等奖励以保证训练信号不中断;若概念级亦出现偏离,则依据概念相似度获得基础奖励,引导智能体逐步修正方向。这种设计使图谱智能体在实体级拥有自主探索空间,同时本体层的稳定导航可在图谱智能体陷入局部循环时提供容错与纠正。

3.4 奖励函数设计

本文采用分层软奖励机制实现双智能体训练优化。本体智能体与图谱智能体的终端奖励均基于嵌入空间中的语义相似度进行连续化评估,通过非稀疏的分段奖励结构缓解传统强化学习中的信号稀疏问题,同时使本体层的语义约束传递至图谱层。

本体智能体的终端奖励取决于其搜索终点概念与真实答案实体所属概念的匹配程度。设本体智能体搜索终点概念为 $c_{\text{ter}} \in C$,真实答案实体为 e_O ,所属概念为 $c^* = \phi(e_O)$,对应概念嵌入分别为 c_{ter} 和 c^* 。本体智能体的终端奖励定义为:

$$R_O(s_{\text{ter}}^O) = \begin{cases} 1, & c_{\text{ter}} = c^*, \\ \eta \cdot \frac{1 + \cos(c_{\text{ter}}, c^*)}{2}, & c_{\text{ter}} \neq c^*, \end{cases} \quad (18)$$

其中 $\cos(c_{\text{ter}}, c^*) = c_{\text{ter}}^T c^* / (\|c_{\text{ter}}\| \|c^*\|)$ 为余弦相似度, $\eta \in (0, 1)$ 为超参数。当终点概念与答案概念完全一致时,智能体获得完整奖励1;若不一致,则依据这2个概

念的语义相似程度在区间 $[0, \eta]$ 内获得连续奖励, 引导本体智能体向语义相近的概念靠近。

图谱智能体的终端奖励采用三段式分层设计, 直接体现本体层的语义结构约束。设图谱智能体搜索终点实体和所属概念分别为 e_{ter} 和 c_{ter} ; 真实答案实体为 e_o 和 c^* 。对应实体嵌入分别为 $\mathbf{e}_{\text{ter}}$ 和 $\mathbf{e}_o$, 概念嵌入分别为 $\mathbf{c}_{\text{ter}}$ 和 $\mathbf{c}^*$ 。图谱智能体的终端奖励定义为:

$$R_G(s_{\text{ter}}^G) = \begin{cases} 1, & e_{\text{ter}} = e_o, \\ \mu + (\nu - \mu) \frac{1 + \cos(\mathbf{e}_{\text{ter}}, \mathbf{e}_o)}{2}, & c_{\text{ter}} = c^* \wedge e_{\text{ter}} \neq e_o, \\ \xi \cdot \frac{1 + \cos(\mathbf{c}_{\text{ter}}, \mathbf{c}^*)}{2}, & c_{\text{ter}} \neq c^*, \end{cases} \quad (19)$$

其中 $\cos(\mathbf{e}_{\text{ter}}, \mathbf{e}_o)$ 与 $\cos(\mathbf{c}_{\text{ter}}, \mathbf{c}^*)$ 分别为实体对和概念对的余弦相似度, 超参数满足 $0 \leq \xi < \mu < \nu < 1$ 。当终点实体与答案实体完全一致时, 获得最高奖励 1; 当本体智能体已将搜索导向正确语义子空间(即 $c_{\text{ter}} = c^*$)但尚未命中精确实体时, 依据实体嵌入相似度在区间 $[\mu, \nu]$ 内获得中等奖励; 当终点实体偏离至不同概念时, 依据概念嵌入相似度在区间 $[0, \xi]$ 内获得较低奖励。上述奖励函数中的超参数 $\{\eta, \xi, \mu, \nu\}$ 均通过验证集网格进行搜索, 并选取验证集上的评价指标 *MRR* 最优的配置作为最终取值。

3.5 参数学习与推理

1) ODRT 中的智能体进行端到端优化训练。对于给定查询 $(e_q, r_q, ?, t_q)$, 设本体智能体在对应概念查询 $(c_q, r_q, ?)$ 上的轨迹为 $\tau_o = (c_0, c_1, \dots, c_{T_o})$, 其优化目标为最大化期望累积奖励:

$$\mathcal{J}_O(\theta_O) = \mathbb{E}_{\tau_o \sim \pi_{\theta}^O} \left[\sum_{t=0}^{T_o} R_O(c_t) \right]. \quad (20)$$

图谱智能体在相同查询上的轨迹为 $\tau_G = (e_o, e_1, \dots, e_{T_G})$, 其优化目标为最大化期望累积奖励:

$$\mathcal{J}_G(\theta_G) = \mathbb{E}_{\tau_G \sim \pi_{\theta}^G} \left[\sum_{t=0}^{T_G} R_G(e_t) \right]. \quad (21)$$

本文采用 REINFORCE 算法分别估计 2 个智能体的策略梯度, 计算累积奖励后分别更新策略参数。具体参数按以下梯度方向更新:

$$\nabla_{\theta_O} \mathcal{J}_O = \mathbb{E}_{\tau_o} \left[\sum_{t=0}^{T_o} R_O(c_t) \nabla_{\theta_O} \log \pi_{\theta}^O(a_t^O | c_t) \right], \quad (22)$$

$$\nabla_{\theta_G} \mathcal{J}_G = \mathbb{E}_{\tau_G} \left[\sum_{t=0}^{T_G} R_G(e_t) \nabla_{\theta_G} \log \pi_{\theta}^G(a_t^G | e_t) \right]. \quad (23)$$

本体图 $\mathcal{G}_O$ 中各边的可学习频次权重 $\mathbf{w}$ 作为 θ_O 的组成部分随本体智能体策略网络参数同步优化。频次权重通过梯度下降更新:

$$\mathbf{w}_{c_s, r, c_o} \leftarrow \mathbf{w}_{c_s, r, c_o} - \alpha \nabla_{\mathbf{w}_{c_s, r, c_o}} (-\mathcal{J}_O(\theta_O)), \quad (24)$$

其中 α 为学习率。

2) 测试阶段固定所有参数。对于给定查询 $(e_q, r_q, ?, t_q)$, 本体智能体基于历史路径提供语义导航, 图谱智能体在 TKG 历史快照中探索路径。模型遍历候选实体空间, 计算各实体的到达概率, 最终选择概率最高的实体作为预测结果:

$$e_{\text{final}} = \arg \max_{e \in \mathcal{E}} P(e | e_q, r_q, t_q), \quad (25)$$

其中 $P(e | e_q, r_q, t_q)$ 由图谱智能体到达该实体的轨迹概率确定。对于宾语给定查询 $(?, r_q, e_o, t_q)$, 将查询转换为逆关系形式 $(e_o, r_q^{-1}, ?, t_q)$ 后执行推理。

ODRT 的完整训练过程如算法 1 所示。

算法 1. ODRT 双智能体协同训练过程。

输入: 时序知识图谱 $\mathcal{G}$, 本体图 $\mathcal{G}_O$, 实体集 $\mathcal{E}$, 关系集 $\mathcal{R}$, 概念集 $\mathcal{C}$, 查询集 $\mathcal{Q}$, 最大路径长度 L , 学习率 α ;

输出: 训练好的策略网络 π_{θ}^O 与 π_{θ}^G 。

- ① For 回合 in $\{1, 2, \dots, P\}$ do
- ② 采样查询 $(e_q, r_q, ?, t_q)$ 及对应答案 e^* ;
- ③ 初始化本体起点 $c_0 \leftarrow \phi(e_q)$, 图谱起点 $(e_o, t_0) \leftarrow (e_q, t_q)$;
- ④ 初始化历史编码 $h_0^O \leftarrow 0, h_0^G \leftarrow 0$;
- ⑤ For 步骤 $k = 0, \dots, L - 1$ do
- ⑥ 构建本体智能体候选动作空间 $\mathcal{A}_k^O$: c_k 在 $\mathcal{G}_O$ 中的邻接概念;
- ⑦ 选择动作 $a_k^O \sim \pi_{\theta}^O(a_k^O | s_k^O)$, 其中 $s_k^O = (c_k, e_q, t_q, r_q, h_k^O)$;
- ⑧ 状态转移 $c_{k+1} \leftarrow \delta^O(c_k, a_k^O)$;
- ⑨ 更新历史编码 $h_{k+1}^O \leftarrow MLP_O(h_k^O, a_k^O)$;
- ⑩ 构建图谱智能体候选动作空间 $\mathcal{A}_k^G$: (e_k, t_k) 在 $\mathcal{G}$ 中的时序出边;
- ⑪ 选择动作 $a_k^G \sim \pi_{\theta}^G(a_k^G | s_k^G)$, 其中 $s_k^G = (e_k, t_k, e_q, t_q, r_q, h_k^G, h_k^O)$;
- ⑫ 状态转移 $(e_{k+1}, t_{k+1}) \leftarrow \delta^G((e_k, t_k), a_k^G)$;
- ⑬ 更新历史编码 $h_{k+1}^G \leftarrow LSTM_G(h_k^G, a_k^G)$;
- ⑭ End For
- ⑮ 依据式(18)(19)计算智能体奖励 R_O 和 R_G ;
- ⑯ 依据式(20)(23)和奖励更新智能体策略参数 θ_O 和 θ_G ;
- ⑰ 依据式(24)更新本体图 $\mathcal{G}_O$ 中各边的可学习频次权重 $\mathbf{w}$;
- ⑱ End For

⑬ Return 训练好的策略网络 $\pi_\theta^O, \pi_\theta^G$ 。

3.6 计算复杂度分析

首先对 ODRT 的时间复杂度进行分析。设本体智能体最大游走步长为 L_O , 图谱智能体为 L_G 。本体智能体在第 k 步的候选动作数为当前概念 c_k 在本体图 G_O 中的出度, 记平均概念出度为 $\bar{d}_O$; 图谱智能体在第 k 步的候选动作数由时序邻域 Top- N 采样确定, 固定为 N 。设单次策略网络前向传播的计算量与嵌入维度 d 线性相关, 则 ODRT 处理单条查询的推理时间复杂度为:

$$O_{\text{Time}} = O(L_O \cdot \bar{d}_O \cdot d + L_G \cdot N \cdot d)。(26)$$

由于本体层 $|C| \ll |\mathcal{E}|$, 通常 $\bar{d}_O$ 为较小常数, 且 $L_O \leq L_G$, 因此图谱层的局部搜索开销 $L_G \cdot N \cdot d$ 构成主要计算成本, 整体复杂度与路径长度 L_G 及固定采样数 N 线性相关。这表明 ODRT 在嵌入本体语义的同时控制了时间开销, 具备良好的可扩展性。

考虑空间复杂度, ODRT 的空间占用由静态参数存储与动态运行时缓存 2 部分组成。静态参数存储包括实体嵌入矩阵 $E \in \mathbb{R}^{|\mathcal{E}| \times d_e}$ 、关系嵌入矩阵 $R \in \mathbb{R}^{|\mathcal{R}| \times d_r}$ 、概念嵌入矩阵 $C \in \mathbb{R}^{|\mathcal{C}| \times d_c}$, 以及本体图邻接结构与可学习频次权重 $w \in \mathbb{R}^{|\mathcal{F}_O|}$ 。设双智能体策略网络的参数量分别为 $|\theta_O|$ 与 $|\theta_G|$, 则模型静态空间占用为:

$$O_{\text{Static}} = O(|\mathcal{E}|d_e + |\mathcal{R}|d_r + |\mathcal{C}|d_c + |\mathcal{F}_O| + |\theta_O| + |\theta_G|)。(27)$$

动态运行时空间主要为批量推理中的路径状态缓存。设批大小为 B , 则运行时额外空间为:

$$O_{\text{Dynamic}} = O(B \cdot L_G \cdot d), (28)$$

其中 $d = \max(d_e, d_c)$ 。由于 $|C| \approx \sqrt{|\mathcal{E}|/2}$, 本体层引入的额外存储 $O(|\mathcal{C}|d_c + |\mathcal{F}_O|)$ 远低于实体嵌入矩阵的存储开销, 表明 ODRT 以较小的空间增量换取了语义导航能力。

4 实 验

为了评估 ODRT 模型在 TKG 推理任务中的性能提升, 并证明其训练框架的合理性, 选择 4 个广泛使用的 TKG 推理数据集和 18 种先进基线模型展开全面的对比实验。实验主要围绕 4 个研究问题进行:

研究问题 1: ODRT 模型在 TKG 推理任务中是否能取得更优的性能?

研究问题 2: ODRT 中使用的多嵌入来源和本体层的设计是否不可或缺? 这些模块如何提升性能?

研究问题 3: ODRT 的性能是否能将不同 LLM 和 TKG 推理模型作为嵌入模型, 以及如何选择不同

聚类超参数敏感?

研究问题 4: ODRT 中本体层的引入能否在 TKG 推理过程中帮助智能体生成逻辑上更具可解释性的推理路径?

本节首先介绍了实验环境, 包括数据集与评价指标、基线模型与实验设置, 选择了 18 种经典的 TKG 推理模型进行对比分析, 并设计了消融实验、敏感性分析与效用分析以验证 ODRT 的性能。此外, 还通过具体案例研究进一步展示 ODRT 在可解释方面的贡献。

4.1 数据集与评价指标

本文选取事件型数据集 ICEWS14 与 ICEWS18^[29]、百科型数据集 WIKI^[10] 与 YAGO^[10] 等 4 个公开基准数据集进行实验, 涵盖事件型与百科型 2 类 TKG, 包括离散时间戳与时间区间 2 种时间表示形式, 具体统计信息如表 1 所示。

Table 1 Statistical Information of the Datasets Used in the Experiment

表 1 实验使用数据集的统计信息

数据集	实体	关系	时间	训练集	验证集	测试集	时间间隔
ICEWS14	23 033	256	365	373 018	45 995	49 545	24 小时
ICEWS18	12 498	260	365	323 895		341 409	24 小时
WIKI	12 554	24	232	539 286	67 538	63 110	1 年
YAGO	10 623	10	189	161 540	19 523	20 026	1 年

ICEWS14 与 ICEWS18 源自综合冲突预警系统 (integrated conflict early warning system, ICEWS), 该系统整合全球超过 100 个数据源的时政事件信息, 通过事件编码器自动识别并抽取事件, 时间粒度为 24 小时, 适用于评估模型对短期动态事件的预测能力。

WIKI 是由维基媒体基金会维护的多语言协作型知识库, 作为维基百科的核心基础设施, 其语义内容具有高度多样性, 实体与属性通过统一编码唯一标识, 形成 TKG 的节点与边。

YAGO 由马克斯-普朗克研究所开发, 是链接数据集成领域的重要成果, 通过规则自动抽取维基百科实体信息框中的结构化数据, 并融合时空维度与丰富的语言学知识, 实体类别体系完善。

本文遵循 TKG 推理任务的标准实验设置, 上述数据集均按时间顺序划分为训练集、验证集与测试集, 确保训练时间戳早于验证时间戳, 验证时间戳早于测试时间戳, 以严格模拟真实场景下的未来事件预测。

本文采用 TKG 推理任务的 2 个标准评价指标:

平均倒数排名 (mean reciprocal rank, MRR) 与 Hits@ K 。

设测试集为 $\mathcal{F}_{\text{test}}$, 对于其中每个测试事实 (e_s, r, e_o, t) , 模型需分别回答对象预测查询 $(e_s, r, ?, t)$ 与主体预测查询 $(?, r, e_o, t)$ 。令 $rank_o$ 与 $rank_s$ 分别表示真实答案实体 e_o 或 e_s 在候选实体排序中的位置。 MRR 定义为所有查询正确实体排名的倒数均值:

$$MRR = \frac{1}{2|\mathcal{F}_{\text{test}}|} \sum_{q \in \mathcal{F}_{\text{test}}} \sum_{e \in [s, o]} \left(\frac{1}{rank_e} \right). \quad (29)$$

Hits@ K ($K \in \{1, 3, 10\}$) 定义为正确实体排名进入前 K 位的查询占比:

$$\text{Hits}@K = \frac{1}{2|\mathcal{F}_{\text{test}}|} \sum_{q \in \mathcal{F}_{\text{test}}} \sum_{e \in [s, o]} (\mathbb{I}[rank_e]), \quad (30)$$

其中 $\mathbb{I}[\cdot]$ 为指示函数, 当条件成立时取值为 1, 否则为 0。上述指标取值越高, 表明模型预测性能越优。

4.2 基线模型与实验设置

为全面评估 ODRT 的推理性能, 本文选取静态知识图谱嵌入、时序知识图谱表示与基于强化学习的时序推理 3 类代表性方法作为基线, 分别对应嵌入学习、动态演化建模与可解释路径搜索 3 种研究范式。

1) 静态推理模型仅建模实体与关系的静态语义关联, 忽略时间维度, 本文选取 TransE^[4]、Complex^[6]、RotatE^[5] 与 DistMult^[7]。

2) 时序推理模型显式建模实体与关系随时间动态变化, 本文选取 TTransE^[10]、TATransE^[11]、TADistMult^[11]、RE-NET^[14]、TANGO^[13]、EvoKG^[15]、Coherence Mode^[16]、ChapTER^[17]、PPT^[18] 与 HFL^[19] 进行对比。

3) 可解释推理模型基于强化学习通过策略搜索实现可解释的时序路径推理, 本文选取 TPath^[22]、TAgent^[23] 与 TITer^[24] 作为对比基线。

在 ODRT 的奖励函数中, 超参数 $\{\eta, \xi, \mu, \nu\}$ 通过验证集网格搜索确定。具体地, 对于本体智能体奖励中的 η , 在 $\{0.1, 0.3, 0.5\}$ 范围内搜索; 对于图谱智能体奖励中的 ξ, μ, ν , 分别在 $\{0.1, 0.3, 0.5, 0.7, 0.9\}$ 中按约束 $\xi < \mu < \nu$ 的组合进行搜索。不同数据集上的最优取值略有差异, 以验证集 MRR 最优的配置作为选择依据。

4.3 主实验结果分析

为了进一步分析 ODRT 在不同数据集上的 TKG 推理能力, 我们在表 2 与表 3 中分别报告了 4 个数据集上的实验结果。

首先对 ODRT 与现有基于强化学习的 TKG 推理方法如 TPath、TAgent 与 TITer 在 4 个数据集上的性能进行对比。ODRT 在所有数据集的全部指标上均

Table 2 Reasoning Results on ICEWS14 and ICEWS18

表 2 在 ICEWS14 和 ICEWS18 上的推理结果

%

模型	ICEWS14				ICEWS18			
	MRR	Hits@1	Hits@3	Hits@10	MRR	Hits@1	Hits@3	Hits@10
TransE	18.65	1.12	31.34	47.07	17.56	2.48	26.95	43.87
Complex	24.47	16.13	27.49	41.09	30.09	21.88	34.15	45.96
RotatE	29.56	22.14	32.92	42.68	23.10	14.33	27.61	38.72
DistMult	19.06	10.09	22.00	36.41	22.16	12.13	26.00	42.18
TTransE	6.35	1.23	5.80	16.65	8.36	1.94	8.71	21.93
TADistMult	20.78	13.43	22.80	35.26	28.53	20.30	31.57	44.96
RE-NET	45.71	38.42	49.06	59.12	42.93	36.19	45.47	55.80
TANGO-TuckER	46.42	38.94	50.25	59.80	44.56	37.87	47.46	57.06
TANGO-DistMult	46.68	41.20	48.64	57.05	44.00	38.64	45.78	54.27
EvoKG	18.30	6.30	19.43	39.37	29.67	12.92	33.08	58.32
Co-CyGNet	49.57	43.04	53.64	60.03	47.93	43.21	51.38	58.99
Co-CENET	53.37	49.58	54.81	61.30	51.47	47.32	52.08	59.11
ChapTER	33.63	24.18	37.86	52.49	24.47	15.82	27.66	41.29
PPT	36.95	27.37	40.83	56.36	25.39	16.28	28.56	43.31
HFL	37.91	27.73	42.77	57.36	27.10	17.83	30.44	45.50
TPath	38.96	29.44	42.50	54.64	27.48	21.12	31.31	42.88
TAgent	39.12	30.78	43.24	55.27	27.21	20.10	30.73	42.37
TITer	41.73	32.74	46.46	58.44	29.98	22.05	33.46	44.83
ODRT (本文)	<u>41.90</u>	<u>32.92</u>	<u>47.04</u>	<u>58.78</u>	<u>30.13</u>	<u>22.16</u>	<u>33.57</u>	<u>45.04</u>

注: 黑体数值表示与全部推理模型相比最佳, 下划线数值表示与可解释性推理模型相比最佳。

优于或持平于现有最优的强化学习基线, 验证了本体-图谱双智能体协同机制对路径推理的增益效果。

在 ICEWS14 上, ODRT 的 MRR 为 41.90%, 相比 TPath 与 TAgent 分别提升 7.5% 与 7.1%, 相比 TITer 提升约 0.4%。在 ICEWS18 上, ODRT 在 MRR 上的提升则分别为 9.6%、10.7% 与 0.5%。在百科型数据集上, ODRT 的优势进一步放大。在 WIKI 上, ODRT 的 MRR 达到 75.58%, 较 TITer 提升 0.08 个百分点; Hits@3 与 Hits@10 分别为 77.59% 与 79.15%, 分别提升 0.10 个百分点与 0.13 个百分点。在 YAGO 上, ODRT 的 4 项指标相较 TITer 分别提升 1.24 个百分点、1.33 个百分点、1.14 个百分点与 1.18 个百分点。上述结果表明, 本体语义导航在关系模式稳定、概念层次清晰的数据集上能够为路径搜索提供更明确的逻辑约束, 从而带来更显著的性能增益。

对 ODRT 与其他类型的 TKG 推理基线进行对比, 可以发现 ODRT 在 WIKI 与 YAGO 两个百科型数据集上取得了当前最优性能 (SOTA)。具体而言, 在 WIKI 上 ODRT 的 MRR 显著优于次优基线 TITer 与 Co-CENET,

Table 3 Reasoning Results on WIKI and YAGO**表 3 在 WIKI 和 YAGO 上的推理结果** %

模型	WIKI				YAGO			
	<i>MRR</i>	Hits@1	Hits@3	Hits@10	<i>MRR</i>	Hits@1	Hits@3	Hits@10
TransE	46.68	36.19	49.71	51.71	48.97	46.23	62.45	66.05
ComplEx	47.84	38.15	50.08	51.39	61.29	54.88	62.28	66.82
RotatE	50.67	39.73	50.71	50.88	65.09	55.69	65.67	66.16
DistMult	46.12	37.24	49.81	51.38	49.47	52.97	60.91	65.26
TTransE	31.74	32.61	36.25	43.45	32.57	34.29	43.39	53.37
TADistMult	48.09	45.84	49.51	51.70	61.72	62.80	65.32	67.19
RE-NET	51.97	48.01	52.07	53.91	65.16	63.29	65.63	68.08
TANGO-TuckER	53.28	52.21	53.61	62.72	67.21	65.56	67.59	77.23
TANGO-DistMult	54.05	51.52	53.84	62.95	68.34	67.05	68.39	78.10
EvoKG	50.66	12.21	63.84	67.29	55.11	54.37	81.38	83.81
Co-CyGNet	46.31	52.19	52.34	54.11	63.83	65.14	66.49	70.03
Co-CENET	68.77	68.41	68.58	69.34	84.08	84.25	84.19	85.56
ChapTER					58.63	60.26	61.55	64.90
PPT	53.84	51.79	54.36	60.33	61.39	64.32	67.47	69.83
HFL					62.15	65.44	68.87	70.23
TPPath	27.62	17.18	31.63	47.04	45.06	41.73	48.15	56.36
TAgent	27.66	17.51	31.43	46.68	46.15	42.45	48.90	56.69
TITer	75.50	72.96	77.49	79.02	87.47	84.89	89.96	90.27
ODRT (本文)	75.58	73.04	77.59	79.15	88.71	86.22	91.10	91.45

注: 黑体数值表示与全部推理模型相比最佳, 下划线数值表示与可解释性推理模型相比最佳。

相对提升分别为 0.1% 与 9.9%; 在 YAGO 上 ODRT 的这一指标则同样领先, 相对提升分别为 1.4% 与 5.5%。

然而, 在事件型数据集 ICEWS14 与 ICEWS18 上, ODRT 却未能取得全局最优。ICEWS14 上 ODRT 的 *MRR* 为 41.90%, 低于 Co-CENET 的 53.37%、Co-CyGNet 的 49.57% 以及 TANGO 系列; ICEWS18 上 ODRT 的 4 项指标同样低于 Co-CENET 等基线。这一现象需从 ODRT 的设计定位与 ICEWS 数据特性的结构性差异加以理解。

首先, ODRT 的本体层通过概念聚合过滤了时序细粒度波动, 以提取跨时间稳定的抽象事件模式。ICEWS14 与 ICEWS18 的时间粒度为 24 小时, 事件密集且关系类型多达 256 种, 导致同一概念下的实体行为模式高度异质。例如, “Person” 概念可能同时包含政治家、军事人员与平民, 其在单日内的行为差异远大于年度粒度的百科数据, 使得概念级语义导航的区分度被稀释, 本体层的压缩收益相对有限。

其次, ICEWS 数据集中存在大量短期重复事件, 如同一国家连续多日的外交互动。这类事实的预测

高度依赖近期历史复用, 而非多跳路径推理。ODRT 作为路径搜索方法, 逐步推理的开销显著高于基于嵌入的模型; 在重复模式主导的场景中, 这种额外的计算投入难以转化为对应的性能回报。

最后, ICEWS 对应的本体图的边空间同样庞大, 本体智能体的概念级搜索面临与图谱智能体类似的决策负担, 双智能体的协同优势被部分抵消。而在实体规模适中、概念层次清晰的百科型数据集 (WIKI、YAGO) 上, 本体层的语义压缩与引导作用得以充分释放, 使 ODRT 达到了 SOTA 水平。

4.4 消融实验

为验证 ODRT 各组成模块的有效性, 本文在 ICEWS14 与 YAGO 上执行消融实验, 依次移除 LLM 语义嵌入、TKGR 结构嵌入、双通道嵌入聚类机制以及本体层, 通过性能衰减程度量化各组件贡献, 实验结果如表 4 所示。

Table 4 Ablation Experimental Results on ICEWS14 and YAGO**表 4 在 ICEWS14 和 YAGO 上的消融实验结果** %

模型	ICEWS14				YAGO			
	<i>MRR</i>	Hits@1	Hits@3	Hits@10	<i>MRR</i>	Hits@1	Hits@3	Hits@10
ODRT	41.90	32.92	47.04	58.78	88.71	86.22	91.10	91.45
w/o LLM 嵌入	41.10	31.80	46.11	58.39	88.16	85.73	90.68	91.15
w/o TKGR 模型嵌入	41.16	31.99	46.25	58.49	88.19	85.79	90.71	91.11
w/o 嵌入聚类	41.07	31.64	45.83	57.85	88.11	85.65	90.44	90.78
w/o 本体层	40.89	31.24	45.57	57.63	87.57	84.99	89.70	90.14

在 ICEWS14 上, 移除 LLM 嵌入后模型的 *MRR* 降至 41.10%, 较完整模型下降 1.9%, Hits@1 下降 1.12, 表明大语言模型提供的语义信息对概念边界划分具有补充作用, 尤其在区分语义相近但功能不同的实体时具有不可替代性。移除 TKGR 结构嵌入后 *MRR* 下降 0.74 个百分点, Hits@1 下降 0.93 个百分点, 说明时序图结构信息是确保聚类内实体行为模式一致性的基础。当进一步移除嵌入聚类机制, *MRR* 和 Hits@1 分别下降 0.83 个百分点与 1.28 个百分点, 验证了语义-结构双通道融合对概念质量的提升效应。最为显著的是移除本体层: 模型退化为单智能体路径搜索架构, *MRR* 和 Hits@1 分别下降 1.01 个百分点与 1.68 个百分点, 下降比例达到 2.4% 与 5.1%, 表明本体层是 ODRT 实现可解释语义导航与动作空间压缩的核心组件。

在 YAGO 上, 各模块的贡献趋势与 ICEWS14 一致, 但本体层的作用更为突出。移除本体层后 *MRR*、

Hits@1、Hits@3 与 Hits@10 分别下降 1.14 个百分点、1.23 个百分点、1.40 个百分点与 1.31 个百分点,性能衰减幅度在所有变体中最大。

上述结果表明,本体语义导航与联合嵌入初始化共同构成了推理性能的关键支撑,LLM 语义嵌入与 TKGR 结构嵌入分别从语义与结构维度保障了概

念划分的完备性,任何单一组件的缺失均会造成可度量的性能损失。

4.5 超参数敏感性分析

为探究概念划分数量 K 对 ODRT 性能的影响,本文在 4 个数据集上以 $K = \sqrt{|\mathcal{E}|/2}$ 为中心设置不同取值进行实验,结果如图 4 所示。

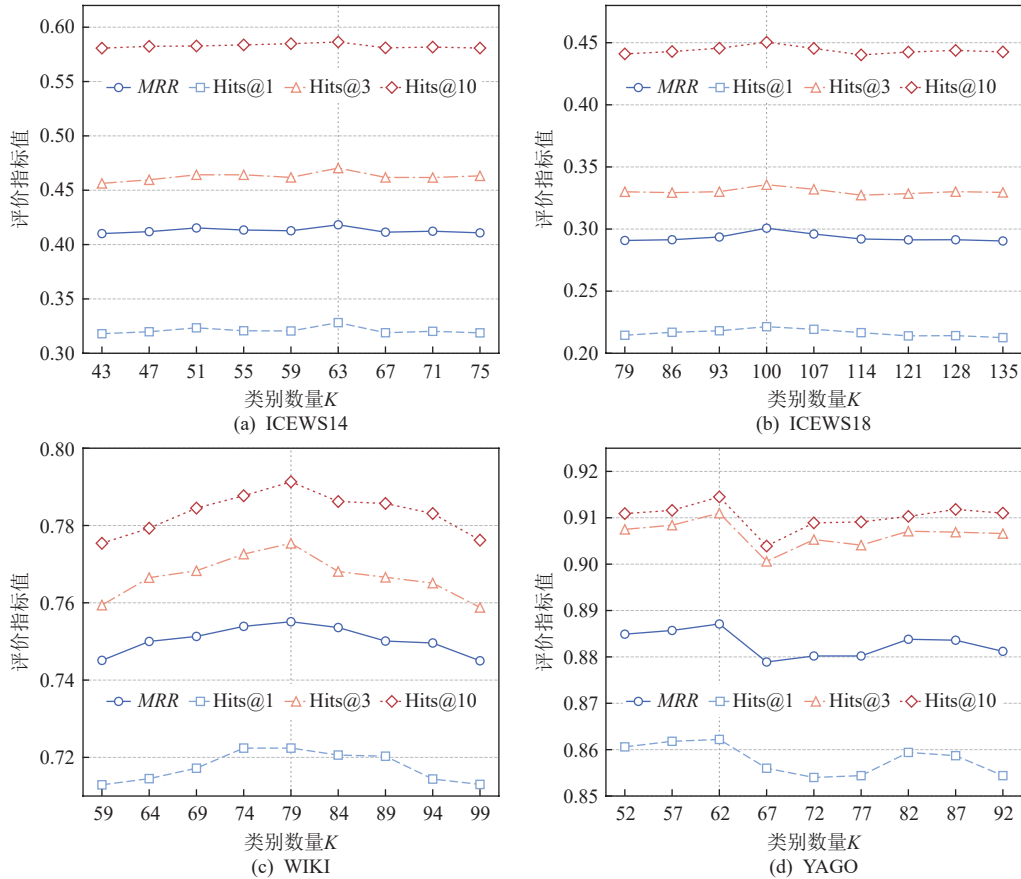


Fig. 4 Experimental results of cluster number hyperparameter sensitivity analysis on YAGO

图 4 在 YAGO 上的聚类数量超参数敏感性分析实验结果

在 ICEWS14 上,当 K 从 43 增至 75 时, MRR 在 0.410 1~0.418 2 波动,最优值 ($K = 63$),较最低值 ($K = 43$) 提升 0.81; Hits@3 与 Hits@10 同样呈现先升后降趋势,表明适中的聚类数目能够在语义泛化与概念区分度之间取得平衡。在 ICEWS18 上, K 从 79 变化至 135 时, MRR 介于 0.290 4 和 0.300 7 之间,最优值出现在 $K = 100$,对应 MRR 为 0.300 7, Hits@1 为 22.14,较 $K = 135$ 提升约 1.03 个百分点,说明该数据集对聚类数目相对敏感,过大的 K 会导致概念粒度过细、本体层泛化能力下降。

在百科型数据集上, ODRT 对 K 的敏感度呈现不同特征。WIKI 上 K 从 59 增至 99 时, MRR 从 0.745 1 升至峰值 0.755 1, $K=79$ 时,随后回落至 0.745 0; YAGO 上 K 从 52 变化至 92 时, MRR 处于 0.878 9~0.887 1,最

优值 ($K = 62$) 对应 MRR 为 0.887 1,较最低值提升 0.82 个百分点。

上述结果表明,聚类数目 K 直接影响本体层的语义抽象质量: K 过小(如 WIKI 上 $K = 59$)时,概念划分过于粗糙,异质实体被归入同一概念,导致本体智能体的指示信号模糊; K 过大(如 YAGO 上 $K = 92$)时,概念粒度接近实体级,本体层的动作空间压缩收益被削弱。4 个数据集的最优 K 值均集中在 $\sqrt{|\mathcal{E}|/2}$ 附近,验证了该启发式设定的合理性,同时表明 ODRT 在适度的聚类质量波动下仍具备鲁棒性。

4.6 嵌入模型敏感性分析

为评估本体初始化阶段各嵌入源的质量对最终推理性能的影响,本文在 YAGO 上分别替换语义嵌入模型与结构嵌入模型进行实验分析。

图5展示了采用3种主流LLM作为语义嵌入源的效果。Qwen3.5-9B在MRR与Hits@1上分别为0.8871与0.8622, Llama-3.1-8B与Deepseek-R1-7B则分别在MRR取得0.8780与0.8818,下降了0.91个百分点与0.53个百分点。Qwen3.5-9B获得最优表现,且三者性能差异相对较小,表明不同规模与架构的LLM能够提供不同的语义初始化。

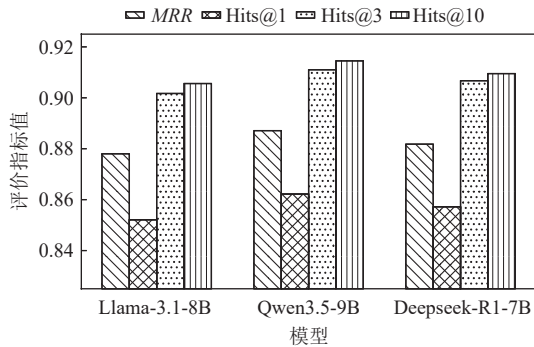


Fig. 5 Embedding effect analysis of different LLMs on YAGO

图5 在YAGO上的不同LLM的嵌入效果分析

图6展示了采用5种TKG推理模型作为结构嵌入源的效果。TADistMult取得最优性能, MRR与Hits@1分别为88.71%与86.22%,显著优于TTransE(MRR为86.83%)与TATransE(MRR为87.47%)。RE-NET(MRR为87.69%)与CyGNet(MRR为88.52%)则稍有提升,其中CyGNet较最优值下降0.19。上述结果表明,基于双线性交互的TADistMult能够更好地保留实体间的细粒度关系模式,而基于翻译假设的TTransE和TATransE在概念级聚合时信息损失较大。总体而言,任一嵌入源的质量波动能够对双通道嵌入融合产生一定影响。

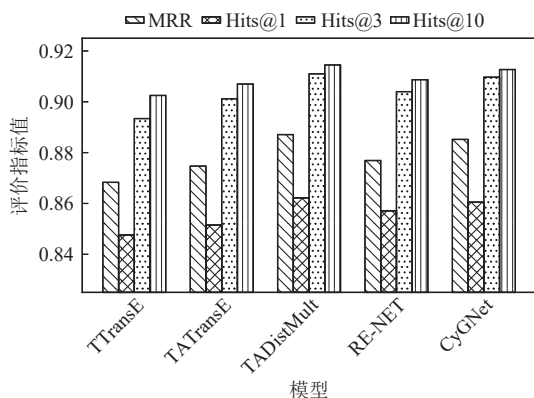


Fig. 6 Embedding effect analysis of different TKG reasoning models on YAGO

图6 在YAGO上的不同TKG推理模型的嵌入效果分析

4.7 本体动态调整机制效用分析

为验证本体的动态调整机制的实际效果,本文

从2个维度量化训练前后本体图中概念间关系权重的变化程度:一是“最高频边改变的数量”,即对于每一对概念,若其所有关联关系中权重最大的边在训练前后发生了改变,则计数1次,该指标反映本体智能体是否修正了静态统计阶段的“主导关系”假设;二是“边的频数分布变化”,以KL散度衡量每一对概念间所有关系边的权重分布在训练前后的差异,该指标反映整体关系模式的重新分配程度。

实验结果如表5所示。在事件型数据集ICEWS14上,共有791对概念的最高频边发生了改变,概念对关系权重的平均KL散度达到1.4063;ICEWS18上则为1510和0.9404。这表明在事件频繁演化、关系模式动态变化的数据集上,静态统计初始化产生的本体关系存在显著偏差,ODRT通过训练中的梯度优化大量修正了概念间的关联强度。

Table 5 Utility Analysis of Dynamic Adjustment Mechanism of the Ontology

表5 本体的动态调整机制的效用分析

数据集	ICEWS14	ICEWS18	WIKI	YAGO
最高频边改变的数量	791	1510	595	164
边的频数分布变化	1.4063	0.9404	0.3864	0.1891

相较之下,百科型数据集的本体调整幅度明显更小。WIKI上分别为595和0.3864;YAGO上则为164和0.1891。这说明在概念层次清晰、领域关系模式稳定的百科型数据中,LLM语义嵌入与TKGR结构嵌入联合初始化已能生成较为准确的本体结构,动态调整主要起局部调整作用。以上结果证明了ODRT动态调整本体的有效性:在不改变实体到概念硬分配的前提下,训练过程通过可学习频次权重有效修正本体的初始化偏差,以适应多变的事件模式。

4.8 对比TITer的推理效果和用时分析

为验证ODRT相比现有最优强化学习基线TITer的性能具有显著性,本文对这2个模型的实验效果提升进行了显著性检验。

如表2与表3中所示,ODRT在4个数据集的全部4项指标上均稳定优于TITer。符号检验表明,这一全正模式具有高度统计显著性。尤其在百科型数据集YAGO上,ODRT的MRR提升达1.42%,Hits@1提升达1.57%。跨数据集的配对分析显示,ODRT相对于TITer在4项指标上的平均绝对提升分别为0.41、0.43、0.48、0.47,效应量Cohen's d指标大于0.7。上述结果证明了ODRT推理效果提升的显著性。

为评估双智能体架构引入的计算开销,本文在

相同实验环境下对 TITer、ODRT(不含本体层)及完整 ODRT 的训练与推理效率进行了对比分析,结果如表 6 所示。

Table 6 Comparative Analysis of Training and Reasoning Time on ICEWS14 and YAGO

表 6 在ICEWS14和YAGO上的训练和推理用时对比分析

模型	ICEWS14		YAGO	
	训练	推理	训练	推理
TITer	28.52	3.98	67.73	5.90
ODRT w/o 本体层	28.39	3.91	67.56	5.83
ODRT (本文)	29.18	4.15	68.40	6.17

在 ICEWS14 上, ODRT 每百回合训练用时为 29.18 min, 较 TITer 的 28.52 min 增加约 2.3%; 测试集推理用时为 4.15 min, 增加约 4.3%。在 YAGO 上, 训练用时从 67.73 min 增至 68.40 min, 推理用时从 5.90 min 增至 6.17 min, 增幅分别约 1.0% 和 4.6%。值得注意的是, 移除本体层后的 ODRT 在训练与推理时间上均与 TITer 基本持平, 表明上述额外开销主要源于本体智能体的概念级游走与状态共享计算。综合来看, ODRT 以较小的时间增量换取了更强的语义导航能力与可解释性, 验证了其在实际应用中的可行性。

4.9 案例研究

为直观展示 ODRT 双智能体协同机制的工作方式, 本文从测试集中选取 3 类典型查询: 简单查询、复杂查询与纠错查询, 以展示本体智能体与图谱智能体的搜索路径, 如表 7 所示。

在简单查询案例中, 本体层将“live_in”这一直接关系拆解为“work_for → locate_in”的间接语义组合, 为图谱智能体提供了明确的子目标序列, 引导图谱智能体沿“任职机构→机构所在地”的时序事实链定位

答案, 避免了在国家等其他实体集合中的盲目搜索。

复杂查询案例涉及跨领域推理, 本体路径长达 4 跳, 图谱智能体在该语义框架引导下, 历经 5 跳实体转移。其中, 图谱路径中每一步的实体所属概念均与本体路径对应节点一致, 如“George_Wilams”和“Sanon_Parker”属于“Person(economy)”, “Jeff_Elliot”属于“Person(politics)”, 验证了本体导航对长路径搜索的阶段式约束作用。若缺乏本体层指导, 很可能导致智能体因动作空间膨胀而偏离答案。

纠错查询案例对于搜索路径中的错误进行实体回溯。图谱路径在第 3 步返回已访问实体“Thomas_Norie”, 形成局部循环, 这在传统单智能体强化学习中常被视为无效探索, 而本体路径始终保持正确方向。最终图谱智能体跳出循环, 通过“claim_responsibility”关系到达正确答案“Deal_Beginning”, 体现了本体层的概念级语义导航对图谱层的实体级探索误差具有容错与纠正能力。

上述案例表明, ODRT 的本体层通过概念级路径为实体级搜索提供可解释的导航框架: 简单查询中压缩搜索空间, 复杂查询中分解长程目标, 纠错查询中容忍局部噪声并维持全局方向一致性, 证明了 ODRT 在 TKG 推理场景中的可解释性优势。

5 结 论

本文针对现有基于强化学习的 TKG 推理方法缺乏高层语义指导、推理路径逻辑性不足, 以及本体构建静态僵化等问题, 提出了面向 TKG 推理的本体-图谱双智能体强化学习方法 ODRT。该方法通过 LLM 语义嵌入与预训练 TKG 推理模型结构嵌入联合初始化本体, 并引入可学习频次权重实现本体在训练中的动态优化; 在此基础上构建双智能体协同框架, 上

Table 7 Reasoning Path Examples for Three Types of Queries

表 7 3 种查询的推理路径示例

简单查询	Formen_Gamoz $\xrightarrow{\text{live_in}(2004)}$? (Berlin)			
本体路径	Person(education) $\xrightarrow{\text{work_for}}$ Affiliation(education) $\xrightarrow{\text{locate_in}}$ Country			
图谱路径	Fornen_Gamoz $\xrightarrow{\text{work_for}(1999)}$ HU_Berlin $\xrightarrow{\text{locate_in}(1983)}$ Berlin			
复杂查询	George_Wilams $\xrightarrow{\text{make_negative_comment}(2013)}$? (Threat_of_Force)			
本体路径	Affiliation(economy) $\xrightarrow{\text{criticize}}$ Person(economy) $\xrightarrow{\text{consult}}$ Person(politics) $\xrightarrow{\text{work_for}}$ Affiliation(politics) $\xrightarrow{\text{criticize}}$ Negative_Behavior(politics)			
图谱路径	Halen_Company $\xrightarrow{\text{employ}(1998)}$ George_Wilams $\xrightarrow{\text{work_with}(2002)}$ Sanon_Parker $\xrightarrow{\text{consult}(2009)}$ Jeff_Elliot $\xrightarrow{\text{work_for}(2009)}$ Justice&Development_Organization $\xrightarrow{\text{criticize}(2011)}$ Threat_of_Force			
纠错查询	Tinear_Group $\xrightarrow{\text{doubt_about}(2015)}$? (Deal_Beginning)			
本体路径	Affiliation(economy) $\xrightarrow{\text{investigate}}$ Person(economy) $\xrightarrow{\text{claim_responsibility}}$ Activity(economy)			
图谱路径	Tinear_Group $\xrightarrow{\text{investigate}(2008)}$ Thomas_Norie $\xrightarrow{\text{agree_with}(2010)}$ Raney_Alves $\xrightarrow{\text{agree_with}(2010)}$ Thomas_Norie $\xrightarrow{\text{claim_responsibility}(2013)}$ Deal_Beginning			

层本体智能体在概念级提供语义导航, 下层图谱智能体在实体级执行时序搜索, 通过历史状态共享与分层软奖励实现跨层协同。在 ICEWS14、ICEWS18、WIKI 与 YAGO 四个公开数据集上的实验结果表明, ODRT 全面优于现有基于强化学习的 TKG 推理基线, 并在百科型数据集上达到当前最优性能; 而在实体行为模式异质且预测依赖历史复用的事件型数据集上, ODRT 与部分嵌入方法相比存在一定性能差距。消融实验、敏感性分析、效用分析与案例分析进一步验证了本体引导机制与双智能体框架的有效性。

尽管 ODRT 取得了较好的实验效果, 仍存在一些值得深入探索的方向。首先, 本体初始化过程依赖外部预训练模型, 其计算开销与模型选型对最终性能有一定影响, 未来可探索更轻量级的本体自动生成与蒸馏策略。其次, 本文仅验证了 ODRT 在链接预测任务上的有效性, 未来可将其拓展至 TKG 上的多步推理、反事实预测与复杂逻辑查询等更具挑战性的任务。最后, 本体层的概念划分目前采用硬分配策略, 后续可研究软隶属度或自适应概念演化机制, 以进一步提升本体对动态知识的表征灵活性。

作者贡献声明: 司悦航提出方法思路和实验方案并撰写论文; 曾泽凡负责实验验证并修改论文; 王博岩和霍静指导方法设计和实验验证; 高阳和黄金才统筹指导整个研究过程和修改工作。

参 考 文 献

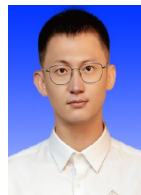
- [1] Cui Yuanning, Sun Zequn, Hu Wei. A pre-trained universal knowledge graph reasoning model based on rule prompts[J]. Journal of Computer Research and Development, 2024, 61(8): 2030–2044 (in Chinese)
(崔宁宇, 孙泽群, 胡伟. 基于规则提示的知识图谱通用推理预训练模型[J]. 计算机研究与发展, 2024, 61(8): 2030–2044)
- [2] Li Zixuan, Guan Saiping, Jin Xiaolong, et al. Two-stage temporal knowledge graph reasoning based on jointly evolutionary modeling of multiple history sequences[J]. Journal of Chinese Information Processing, 2024, 38(2): 46–53 (in Chinese)
(李紫宣, 官赛萍, 靳小龙, 等. 基于多历史序列联合演化建模的两阶段时序知识图谱推理[J]. 中文信息学报, 2024, 38(2): 46–53)
- [3] Ma Ang, Yu Yanhua, Yang Shengli, et al. Survey of knowledge graph based on reinforcement learning[J]. Journal of Computer Research and Development, 2022, 59(8): 1694–1722 (in Chinese)
(马昂, 于艳华, 杨胜利, 等. 基于强化学习的知识图谱综述[J]. 计算机研究与发展, 2022, 59(8): 1694–1722)
- [4] Bordes A, Usunier N, Garcia-Duran A, et al. Translating embeddings for modeling multi-relational data[C]// Proc of the 26th Int Conf on Neural Information Processing Systems. Cambridge, MA: MIT, 2013: 2787–2795
- [5] Sun Zhiqing, Deng Zhihong, Nie Jianyun, et al. Rotate: Knowledge graph embedding by relational rotation in complex space[J]. arXiv preprint, arXiv: 1902.10197, 2019
- [6] Trouillon T, Welbl J, Riedel S, et al. ComplEx embeddings for simple link prediction[C]//Proc of Int Conf on Machine Learning. New York: PMLR, 2016: 2071–2080
- [7] Yang Bishan, Yih S W, He Xiaodong, et al. Embedding entities and relations for learning and inference in knowledge bases[J]. arXiv preprint, arXiv: 1412.6575, 2014
- [8] Guo Wenbin, Wang Xin, Li Zhao, et al. Attention-enhanced semantic fusion knowledge graph representation learning framework[J]. Journal of Computer Research and Development, 2025, 62(11): 2687–2696 (in Chinese)
(郭文斌, 王鑫, 李钊, 等. 注意力增强的语义融合知识图谱表示学习框架[J]. 计算机研究与发展, 2025, 62(11): 2687–2696)
- [9] Wu Jiapeng, Cao Meng, Cheung J C K, et al. Temp: Temporal message passing for temporal knowledge graph completion[C]//Proc of the 2020 Conf on Empirical Methods in Natural Language Processing (EMNLP). Stroudsburg, PA: ACL, 2020: 5730–5746
- [10] Leblay J, Chekol M W. Deriving validity time in knowledge graph[C]//Companion Proc of the Web Conf 2018. New York: ACM, 2018: 1771–1776
- [11] García-Durán A, Dumančić S, Niepert M. Learning sequence encoders for temporal knowledge graph completion[C]//Proc of the 2018 Conf on Empirical Methods in Natural Language Processing. Stroudsburg, PA: ACL, 2018: 4816–4821
- [12] He Peng, Zhou Gang, Chen Jing, et al. Type-enhanced temporal knowledge graph representation learning model[J]. Journal of Computer Research and Development, 2023, 60(4): 916–929 (in Chinese)
(何鹏, 周刚, 陈静, 等. 类型增强的时态知识图谱表示学习模型[J]. 计算机研究与发展, 2023, 60(4): 916–929)
- [13] Han Zhen, Ding Zifeng, Ma Yunpu, et al. Learning neural ordinary equations for forecasting future links on temporal knowledge graphs[C]//Proc of the 2021 Conf on Empirical Methods in Natural Language Processing. Stroudsburg, PA: ACL, 2021: 8352–8364
- [14] Jin W, Qu Meng, Jin Xisen, et al. Recurrent event network: Autoregressive structure inference over temporal knowledge graphs[C]//Proc of the 2020 Conf on Empirical Methods in Natural Language Processing (EMNLP). Stroudsburg, PA: ACL, 2020: 6669–6683
- [15] Park N, Liu Fuchen, Mehta P, et al. EvoKG: Jointly modeling event time and network structure for reasoning over temporal knowledge graphs[C]//Proc of the 15th ACM Int Conf on Web Search and Data Mining. New York: ACM, 2022: 794–803
- [16] Si Yuehang, Hu Xingchen, Cheng Qing, et al. Coherence mode: Characterizing local graph structural information for temporal knowledge graph[J]. Information Sciences, 2025, 686: 121357
- [17] Peng Miao, Liu Ben, Xu Wenjie, et al. Deja vu: Contrastive historical

- modeling with prefix-tuning for temporal knowledge graph reasoning[C]//Findings of the Association for Computational Linguistics: NAACL 2024. Stroudsburg, PA: ACL, 2024: 1178–1191
- [18] Xu Wenjie, Liu Ben, Peng Miao, et al. Pre-trained language model with prompts for temporal knowledge graph completion[C]//Findings of the Association for Computational Linguistics: ACL 2023. Stroudsburg, PA: ACL, 2023: 7790–7803
- [19] Xu Wenjie, Liu Ben, Peng Miao, et al. Historical facts learning from long-short terms with language model for temporal knowledge graph reasoning[J]. *Information Processing & Management*, 2025, 62(3): 104047
- [20] Das R, Dhuliawala S, Zaheer M, et al. Go for a walk and arrive at the answer: Reasoning over paths in knowledge bases using reinforcement learning[J]. arXiv preprint, arXiv: 1711.05851, 2017
- [21] Zhang Denghui, Yuan Zixuan, Liu Hao, et al. Learning to walk with dual agents for knowledge graph reasoning[C]//Proc of the AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2022, 36(5): 5932–5941
- [22] Bai Luyi, Yu Wenting, Chen Mingzhuo, et al. Multi-hop reasoning over paths in temporal knowledge graphs using reinforcement learning[J]. *Applied Soft Computing*, 2021, 103: 107144
- [23] Tao Ye, Li Ying, Wu Zhonghai. Temporal link prediction via reinforcement learning[C]//Proc of IEEE Int Conf on Acoustics, Speech and Signal Processing (ICASSP 2021). Piscataway, NJ: IEEE, 2021: 3470–3474
- [24] Sun Haohai, Zhong Jialun, Ma Yunpu, et al. Timetraveler: Reinforcement learning for temporal knowledge graph forecasting[C]//Proc of the 2021 Conf on Empirical Methods in Natural Language Processing. Stroudsburg, PA: ACL, 2021: 8306–8319
- [25] Hao Junheng, Chen Muhao, Yu Wenchao, et al. Universal representation learning of knowledge bases by jointly embedding instances and ontological concepts[C]//Proc of the 25th ACM SIGKDD Int Conf on Knowledge Discovery & Data Mining. New York: ACM, 2019: 1709–1719
- [26] Liu Runxuan, Luo Bei, Li Jiaqi, et al. Ontology-guided reverse thinking makes large language models stronger on knowledge graph question answering[C]//Proc of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Stroudsburg, PA: ACL, 2025: 15269–15284
- [27] Qin Jing, Li Guanfeng, Chen Yuyin, et al. Embedding model of knowledge graph via jointly modeling ontology and instances[J]. *Computer Science*, 2026, 53(3): 331–340 (in Chinese)
(秦晶, 李贯峰, 陈昱胤, 等. 融合本体和实例的知识图谱嵌入模型[J]. *计算机科学*, 2026, 53(3): 331–340)
- [28] Lin Dongying, Liu Yinan, Wang Bin, et al. OntoTKGE: Ontology-enhanced temporal knowledge graph extrapolation[J]. arXiv preprint, arXiv: 2604.05468, 2026.
- [29] Si Yuehang, Zeng Zefan, Cheng Qing, et al. FRLT: Adaptive reinforcement learning via fuzzy search for temporal knowledge graph reasoning[J]. *Expert Systems with Applications*, 2026: 133398



Si Yuehang, born in 2000. PhD candidate. Student member of CCF. His main research interests include knowledge graph reasoning and large language models.

司悦航, 2000年生。博士研究生。CCF学生会员。主要研究方向为知识图谱推理、大语言模型。



Zeng Zefan, born in 1993. PhD. His main research interests include causal learning and strategic early warning.

曾泽凡, 1993年生。博士。主要研究方向为因果学习、战略预警。



Wang Boyan, born in 1990. PhD, assistant professor, PhD supervisor. Member of CCF. His main research interests include large scientific research models, multimodal alignment, and cognitive structure.

王博岩, 1990年生。博士, 助理教授, 博士生导师。CCF会员。主要研究方向为科学研究大模型、多模态对齐、认知结构。



Huo Jing, born in 1989. PhD, associate professor, PhD supervisor. Member of CCF. Her main research interests include generative artificial intelligence and embodied intelligence.

霍静, 1989年生。博士, 副教授, 博士生导师。CCF会员。主要研究方向为生成式人工智能、具身智能。



Gao Yang, born in 1972. PhD, professor, PhD supervisor. Senior member of CCF. His main research interests include reinforcement learning, multi-agent systems, computer vision, and big data analysis.

高阳, 1972年生。博士, 教授, 博士生导师。CCF高级会员。主要研究方向为强化学习、多智能体系统、计算机视觉、大数据分析。



Huang Jincai, born in 1973. PhD, professor, PhD supervisor. His main research interests include command and control, and artificial intelligence.

黄金才, 1973年生。博士, 研究员, 博士生导师。主要研究方向为指挥控制、人工智能。