

# 周报 2024/02/25

大论文整体架构：

1. 知识图谱的构建（实体关系联合抽取）（**完成**）
2. 时序知识图谱的推理研究（**进行中**）
3. 时序知识图谱的多跳推理研究（重点关注推理可解释性问题）

学期目标：投稿完成 **2** 篇文章（顶会或 **SCI 2** 区及以上）

完成进度： **0/2**

## 文献阅读

%使用标准文献引用格式，列出几篇本周精读的文献%

例：

- [1] Sun, H., Zhong, J., Ma, Y., Han, Z., & He, K. (2021, November). TimeTraveler: Reinforcement Learning for Temporal Knowledge Graph Forecasting. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (pp. 8306-8319).
- [2] Zhang, M., Xia, Y., Liu, Q., Wu, S., & Wang, L. (2023, July). Learning latent relations for temporal knowledge graph reasoning. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 12617-12631).
- [3] Dong, H., Wang, P., Xiao, M., Ning, Z., Wang, P., & Zhou, Y. (2024). Temporal inductive path neural network for temporal knowledge graph reasoning. Artificial Intelligence, 104085.

## 研究进展/文献感悟/科研想法

研究进展： 重点关注时序知识图谱的推理问题，目前时序知识图谱的推理问题大致可分为两类，一类是知识图谱时间范围内的推理问题，目的为补全知识图谱；另一类为时间范围外的知识图谱，目的为未来事件预测。根据目前课题目标，拟展开第一类研究，主要目的进行隐含传播链的推理。对于第二类问题也可进行地点与爆发病例的知识图谱构建的未来爆发事件预测（目前只是比较粗浅的想法）。

文献感悟：

文献 1： 文章重点关注未来事件预测问题，未来事件预测问题大概存在两个研究难点（如何有效建模时间信息以处理未来时间戳和如何进行归纳推理以处理随时间出现的以前未见实体）。文章利用强化学习解决时序知识图谱的推理问题，可以提供推理路径，在一定程度上提高了可解释性。通过将时间信息建模到 agent 观测到的 state 中来解决第一个问题，对于第二个问题，基于与相同的 relation 相连的实体有类似的性质的思想，利用之前出现过的与 relation 相连的实体的 embedding 的平均值的方法来表示新的实体。

文献 2：文章研究未来事件预测问题，重点关注了由于构建工具和数据源的限制而存在的知识图谱不完整问题。文章把这种不完整问题称为实体间存在潜在关系，主要包括两种，一种是同一个时间戳内的实体间的潜在关系，另一种是不同时间戳间的实体之间的隐藏关系。把知识图谱看作一组不同时间戳的静态知识图谱，首先利用图卷积神经网络得到每个时间戳内各个实体的 **embedding**。隐藏关系的确定由实体间的相似性决定的，作者认为存在关系的实体之间存在一定的相似性（这一点存疑），对于两种隐藏关系只选取相似性靠前的  $k$  个实体对生成隐藏关系，这些实体对共同组成一个图。对于时序信息的编码，在全局层面和局部层面分别采用 GRU 和 LSTM 来获得当前时间戳内某个实体的 **embedding**，最后通过一个门控 **vector** 来得到最终的 **entity embedding**，最后利用让所有的正确的三元组的分数最高的原则进行训练。

文献 3：文章关注未来时间预测问题，重点关注新加入的实体问题。文章认为关键挑战在于解释历史子图和时间模式中的结构依赖关系。由于目前的大部分方法依靠实体 **embedding** 来建模时序知识图谱，这种方法无法应对新出现的大量实体。文章提出的方法不依赖实体而是依赖实体间的路径来进行建模，与前面的一些方法不同，该方法没有把时序知识图谱看作一组静态知识图谱，而是把时序知识图谱看作一个整个的知识图谱，两个实体间的关系有 **relation** 和时间戳。训练的过程中最大化正确四元组对应的 **socre function**，同时加入一些副样本，同时最小化副样本的 **socre function**，为了加强 **relation embedding** 的正交性，在 **loss** 中加入正则项。

## 下周计划

利用目前的方法进行知识图谱的构建并进行可视化；

利用大模型的方法来实现病例相关的时序知识图谱构建；

进一步加强时序知识图谱时间内隐藏关系推理的理论研究，并对阅读过的论文中的实验进行复现。

## 存在问题

对于时序知识图谱的推理问题实验部分有一定的欠缺。