Online Rule Search Database

Tim Gutberlet and Janik Sauerbier

University of Mannheim, Mannheim, Germany
{tgutberl,jsauer}@mail.uni-mannheim.de

Abstract. Knowledge graph completion (KGC) received a lot of attention in recent years and is of high practical relevance inside and outside of academia. KGs are often incomplete and potentially noisy which hinders information extraction. Many different techniques are used to approach the knowledge graph completion problem including various embeddings and rule-based methods. Prior research indicates that inductive logic programming and embedding models can be used together to improve inference quality. This can be done by combining embedding models and constraints derived from rules as the objective function of an integer linear programming problem or by employing other methods focused on combining the results of different approaches. Another method (could be/is) the use of rules during the training of embedding models. Especially in the latter case, there is a significant need for a time-efficient database architecture which is used to identify rules that implicate the predictions of the embedding models. To contribute to this research area, we compare different database architectures using different indexing, hashing and pre-processing methods.

Keywords: Knowledge graph completion · Knowledge representation and reasoning · Database design and models.

1 Introduction

Explaining motivation based on popularity of knowledge graph completion, the indication that combining embeddings and rule-based approaches yielded solid results so far and the fact that a fast rule database would help.

2 Related Work

- AnyBURL, specificly the latest IJCAI paper.[1]
- Other work on combining embeddings and rule-based approaches. (e.g. Wang, Quan, Bin Wang, and Li Guo. "Knowledge base completion using embeddings and rules."[3] or Zhang, Wen, et al. "Iteratively learning embeddings and rules for knowledge graph reasoning."[4])
- Work specifically focused on building effective databases (e.g. Indexing, Hashing and preprocessing stratagies) -> What is should we look into her?

3 Problem statement

The problem we are trying to solve is identifying rules (explanations) that imply certain target facts that are not part of the KG. The potentially incomplete KG and a set of rules that describe similarities of the KG are given. The identified rules can be used to evaluate the predictions of embedding models and facilitate the integration of embedding models and rule-based approaches. The problem is based on abductive reasoning [2]. Abductive reasoning describes the problem of finding an explanation $\mathcal{E}$ for an observation t based on a theory Φ with $\Phi \not\models t$ and $\Phi \not\models \neg t$ such that $\Phi \cup \mathcal{E} \models t$ and $\Phi \cup \mathcal{E}$ is consistent. In our case, the explanation $\mathcal{E}$ refers to the rules implying the observation, the observation refers to the target facts and the theory Φ refers to the given set of rules on the KG. An explanation $\mathcal{E}$ is minimal if for each $\mathcal{E}' \subset \mathcal{E}$ we have that $\Phi \cup \mathcal{E}' \not\models t$. We refer to minimal explanations when mentioning explanations in the following paragraph.

As we are focusing on attaining as many good explanations as possible in a very short amount of time we aim to enhance computational efficiency. Therefore, we use the simplification proposed in [1] to simplify the computation while still achieving good results:

Definition 1 (One-step entailment $\models_1$). The fact t is one-step entailed by $\Phi \cup \mathcal{E}$ written as $\Phi \cup \mathcal{E} \models_1 t$, iff there is a rule $h \leftarrow b$ in Φ for which a grounding exists where the grounded rule body b is in $\mathcal{E}$ and the grounded head h is equal to t.

4 Database structure

We implemented the our solution using the open source object-relational database system PostgreSQL...

5 Experiments

- Comparison with AnyBURL given the limitation of AnyBURL to only find the following rule types:

$$h(X,Y) \leftarrow r(X,Y) \tag{1}$$

$$h(X,Y) \leftarrow r_1(X,Z) \wedge r_2(Z,Y) \tag{2}$$

$$h(X,e_1) \leftarrow r(X,e_2) \tag{3}$$

- XXX
- XXX

Acknowledgements ...

References

1. Betz, Patrick, Christian Meilicke, and Heiner Stuckenschmidt. "Adversarial explanations for knowledge graph embeddings." International Joint Conferences on Artificial Intelligence, 2022.
2. Mayer, Marta Cialdea, and Fiora Pirri. "First order abduction via tableau and sequent calculi." Logic Journal of the IGPL 1.1 (1993): 99-117.
3. Wang, Quan, Bin Wang, and Li Guo. "Knowledge base completion using embeddings and rules." Twenty-fourth international joint conference on artificial intelligence. 2015.
4. Zhang, Wen, et al. "Iteratively learning embeddings and rules for knowledge graph reasoning." The World Wide Web Conference. 2019.