Neural-Symbolic Entangled Framework for Complex Query Answering

Zezhong Xu[1,*], **Wen Zhang**[1,*], **Peng Ye**[1], **Hui Chen**[3], **Huajun Chen**[1,2†]
[1]Zhejiang University & AZFT Joint Lab for Knowledge Engine, China
[2]Hangzhou Innovation Center, Zhejiang University, [3]Alibaba Group
{xuzezhong, zhang.wen, yep, huajunsir}@zju.edu.cn,
weidu.ch@alibaba-inc.com

Abstract

Answering complex queries over knowledge graphs (KG) is an important yet challenging task because of the KG incompleteness issue and cascading errors during reasoning. Recent query embedding (QE) approaches embed the entities and relations in a KG and the first-order logic (FOL) queries into a low dimensional space, answering queries by dense similarity search. However, previous works mainly concentrate on the target answers, ignoring intermediate entities' usefulness, which is essential for relieving the cascading error problem in logical query answering. In addition, these methods are usually designed with their own geometric or distributional embeddings to handle logical operators like union($\vee$), intersection($\wedge$), and negation($\neg$), with the sacrifice of the accuracy of the basic operator – projection, and they could not absorb other embedding methods to their models. In this work, we propose a **Ne**ural and **Sy**mbolic **E**ntangled framework (**ENeSy**) [3]) for complex query answering, which enables the neural and symbolic reasoning to enhance each other to alleviate the cascading error and KG incompleteness. The projection operator in ENeSy could be any embedding method with the capability of link prediction, and the other FOL operators are handled without parameters. With both neural and symbolic reasoning results contained, ENeSy answers queries in ensembles. ENeSy achieves the SOTA performance on several benchmarks, especially in the setting of training model only with the link prediction task.

1 Introduction

People built different Knowledge Graphs, such as Freebase [4], YAGO [13], and Wordnet [11], to store complex structured information and knowledge. The facts in KG are usually represented in the form of triplets, e.g., *isCityOf(New York, USA)*. KGs have been widely applied in various intelligent systems such as question answering and natural language understanding. One of the key tasks on KG reasoning is complex query answering which involves answering FOL query with logical operators including existential quantification ($\exists$), conjunction($\wedge$), disjunction($\vee$), and negation($\neg$).

Given a question *"Who won the Turing Award in developing countries?"*, as illustrated in Figure 1, it could be converted to a FOL query, and a computation graph can be generated with the query. Each node in the computation graph represents an entity or an entity set, while each edge represents a logical operation. Answering these queries is challenging since not all the answers could be directly identified by traversing the KG because of the incompleteness of KG. To address this problem, several

[*]Equal Contribution.
[†]Corresponding Author.
[3]Source code of ENeSy is available at https://github.com/zjukg/ENeSy.

36th Conference on Neural Information Processing Systems (NeurIPS 2022).