CogKG: Unifying Symbolic and Continuous Knowledge Representations for Machine Reasoning

ZHEPEI WEI[†], YUE WANG[‡], JINNAN LI[†], YINGQI GUO[†], ZHINING LIU[†], ERXIN YU[†], and YI CHANG[†]

[†] *School of Artificial Intelligence, Jilin University;* [‡] *School of Information and Library Science, University of North Carolina at Chapel Hill*

Background: Symbolic, rule-based reasoning methods (e.g., expert systems [2]) are reliable and interpretable in solving complex inference problems in specialized domains, but are difficult to generalize because eliciting a comprehensive set of rules from human experts is costly and time-consuming. Recently, continuous knowledge representation models [1] utilize the distributed representations of a Knowledge Graph (KG) to generalize from known facts to unseen and probably true facts (knowledge graph completion). However, such models can only represent and reason about multi-relational data (subject-relation-object triples), not if-then rules. Therefore they are not suited to solve inference problems where observations lead to conclusions.

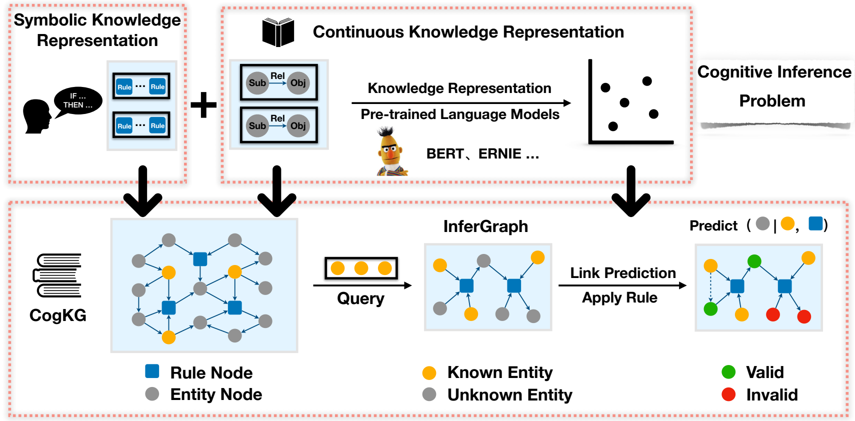

Fig. 1. The CogKG System for Machine Reasoning (Best view in color).

This work is motivated by one overarching question: can we unify symbolic and continuous knowledge representation techniques to perform complex inference tasks? More concretely, we study the following research question: given a large-scale KG with rich relational facts and a moderate set of if-then rules as the *cognitive knowledge*, can we reason about the most likely conclusion(s) given a set of observations? We formalize this question as the cognitive inference problem.

Problem Formulation *(Cognitive Inference Problem)*: Assume we are provided with cognitive knowledge that consists of (1) a KG with N relational facts, represented as $\mathcal{F} = \{f_i\}_{i=1}^{N}$, in which $f_i = (s, r, o)$ is a structured relational fact consisting of subject and object entities and the involved relation; (2) a collection of M expert rules is presented as $\mathcal{R} = \{r_i\}_{i=1}^{M}$, in which $r_i : A \to B$ is an unstructured expert rule that performs single-step if-then reasoning over a relatively small set of KG entities A and B. In a specific cognitive inference task, we are provided with a query that contains L entities, presented as $Q = (q_i)_{i=1}^{L}$ in which q_i is a registered entity in the KG. Our goal is to generate $(C, \mathcal{D})$ given $(\mathcal{F}, \mathcal{R}, Q)$. Specifically, it takes as input the query Q and cognitive knowledge $\mathcal{F}$ and $\mathcal{R}$, and outputs a query-specific conclusion set with K inferred entities $C = \{c_i \in \mathcal{F}\}_{i=1}^{K}$ as well as K decision paths $\mathcal{D} = \{d_i\}_{i=1}^{K}$ that lead to each inferred entity. Each decision path $d_i = (Q, ..., f_k, ..., r_m, ..., c_i)$ starts from the query Q and ends at the conclusion entity c_i, with relational facts $\{f_k\}$ or rules $\{r_m\}$ as intermediate reasoning steps.

Though recent pre-trained language models (e.g., ERNIE [3]) show promising performance in natural language understanding and reasoning tasks by incorporating prior knowledge from large-scale corpus and knowledge graph, they could only partially address the cognitive inference problem and they all require a large amount of training data and computational resources for fine-tuning. To solve the above cognitive inference problem, we first design and construct the **Cog**nitive **K**nowledge **G**raph (**CogKG**), which expresses relational facts and expert rules in a unified framework, and then develop a reasoning inference paradigm based on CogKG. CogKG is a graph-based structure with expert rules and entities as nodes, and the semantic relationship or logical consequence between nodes as edges. We define CogKG as $\mathcal{G} = (V, E)$ where $V = \{V^e, V^r\}$ and $E = \{E^e, E^r\}$. In particular, the entity set is V^e; the rule set is V^r; the relational edges and logical edges are E^e and E^r, respectively.

Proposed Method: To construct the CogKG, we perform relational fact acquisition and rule mining from text resources. We apply named entity recognition and relation extraction to construct V^e and E^e. We further generate expert rules for the construction of V^r and E^r through association rule mining followed by expert validation. With rich cognitive knowledge of expert rules and relational facts in CogKG, we then propose **CogInfer**, a novel paradigm performing machine reasoning based on the CogKG. Here we briefly introduce the instantiation of the CogInfer, and the reasoning procedure for the cognitive inference problem is presented in Fig. 1. **First**, we perform knowledge representation learning on (the entity part of) CogKG and obtain the distributed representations of nodes and edges, for example, $\mathcal{E}(\mathcal{G}) = TransE(V^e, E^e)$. **Then**, for each query Q, we create a corresponding InferGraph $\mathcal{G}^Q$ by iteratively identifying the closure of the involved rule nodes and entity nodes. $\mathcal{G}^Q \subset \mathcal{G}$ is a small sub-graph of the background CogKG while the entity nodes in this graph are particularly categorized into two sets, namely, KnownEntity Set and UnkownEntity Set. It refers to the certainty of the nodes carried out in the ensuing reasoning steps. **Lastly**, we conduct reasoning on the InferGraph through applying expert rules or link prediction with the learned distributed representations to predict the unknown entities in the InferGraph.

Ongoing Work: We are working on an extensive empirical study on clinical diagnosis benchmark to validate the advantage of the proposed method in inference compared to numerous strong baselines, including the time-tested clinical expert system MYCIN. We are also working on finding the upper bound performance for the prediction task produced by statistical machine learning methods such as k-nearest neighbors and support vector machine.

REFERENCES

[1] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko. 2013. Translating embeddings for modeling multi-relational data. *NIPS* 26 (2013).

[2] B. G Buchanan and E. H. Shortliffe. 1984. *Rule-based expert systems: the MYCIN experiments of the Stanford Heuristic Programming Project.* Addison-Wesley.

[3] Y. Sun, S. Wang, Y. Li, S. Feng, X. Chen, H. Zhang, X. Tian, D. Zhu, H. Tian, and H. Wu. 2019. Ernie: Enhanced representation through knowledge integration. *arXiv:1904.09223* (2019).