

---

# Integrating Logical Rules Into Neural Multi-Hop Reasoning for Drug Repurposing

---

Yushan Liu<sup>\*1,2</sup> Marcel Hildebrandt<sup>\*1,2</sup> Mitchell Joblin<sup>1</sup> Martin Ringsquandl<sup>1</sup> Volker Tresp<sup>1,2</sup>

## Abstract

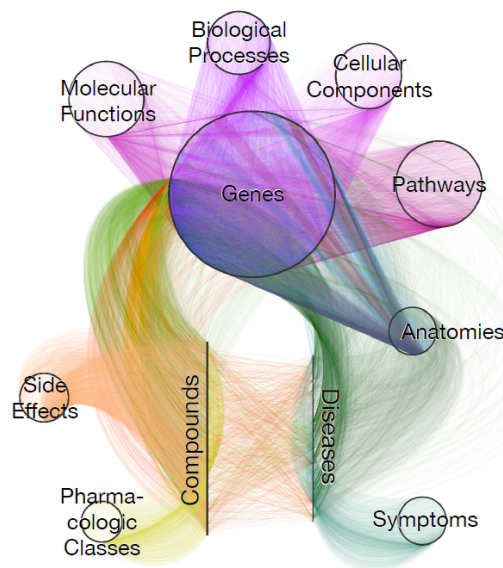
The graph structure of biomedical data differs from those in typical knowledge graph benchmark tasks. A particular property of biomedical data is the presence of long-range dependencies, which can be captured by patterns described as logical rules. We propose a novel method that combines these rules with a neural multi-hop reasoning approach that uses reinforcement learning. We conduct an empirical study based on the real-world task of drug repurposing by formulating this task as a link prediction problem. We apply our method to the biomedical knowledge graph Hetionet and show that our approach outperforms several baseline methods.

## 1. Introduction

Advancements in low-cost high-throughput sequencing and data acquisition technologies have given rise to a massive proliferation of data describing biological systems. Biomedical knowledge graphs (KGs) are becoming increasingly popular as backbones for artificial intelligence tasks such as personalized medicine, predictive diagnosis, and drug discovery (Dörpinghaus & Jacobs, 2019).

From a machine learning perspective, reasoning on biomedical KGs presents new challenges for existing approaches because of the unique structural characteristics of the graphs. One challenge arises due to the highly coupled nature of entities in biological systems that leads to many high-degree and densely interlinked entities. A second challenge is the requirement of information beyond second-order neighborhoods for reasoning about the relationship between two entities (Himmelstein et al., 2017) so that approaches where long-range interactions are incorporated only via node em-

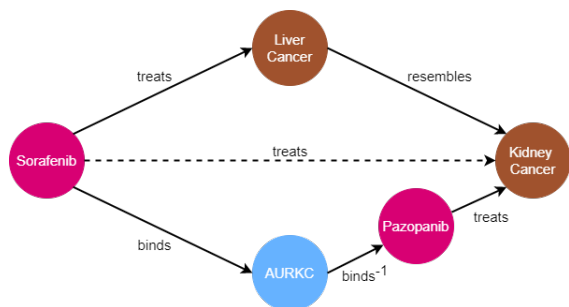
<sup>\*</sup>Equal contribution <sup>1</sup>Siemens AG, Corporate Technology, Munich, Germany <sup>2</sup>Ludwig Maximilian University of Munich, Munich, Germany. Correspondence to: Yushan Liu <yushan.liu@siemens.com>.



**Figure 1.** Visualization of the heterogeneous biomedical network Hetionet<sup>3</sup>.

beddings (e. g., RESCAL (Nickel et al., 2011), TransE (Bordes et al., 2013)) tend to underperform. Unfortunately, approaches that explicitly take the entire multi-hop neighborhoods into account (e. g., graph convolutional models, R-GCN (Schlichtkrull et al., 2018)), often have diminishing performance beyond two-hop neighborhoods (i. e., more than two convolutional layers). Furthermore, high-degree entities can cause the aggregation operations to smooth out the signals. Alternatively, symbolic reasoning approaches (e. g., RuleN (Meilicke et al., 2018), AnyBURL (Meilicke et al., 2019)) learn logical rules and employ them during inference. However, due to the massive scale and diverse topologies of many real-world KGs, combinatorial complexity often prevents the usage of symbolic approaches. Also, logical inference has difficulties handling noise in the data. Recently, path-based reasoning methods have become popular, and they present a seemingly ideal balance for combining information over multi-hop neighborhoods.

We propose a novel neuro-symbolic KG reasoning approach that combines path-based approaches with representation



**Figure 2.** Subgraph of Hetionet that illustrates the drug repurposing use case: The two paths that connect the chemical compound sorafenib and the disease kidney cancer can be used to predict a direct edge between the two entities.

learning and logical rules. These rules can be either mined from data or obtained from domain experts. Inspired by existing methods (Das et al., 2018; Lin et al., 2018; Hildebrandt et al., 2020a;b), we use reinforcement learning to train an agent to conduct policy-guided random walks on a KG. We propose a modification by introducing a reward function that allows the agent to leverage background knowledge formalized as metapaths. In summary, our paper makes the following contributions:

- We propose a novel neuro-symbolic approach that combines neural multi-hop reasoning based on reinforcement learning with logical rules.
- We conduct an empirical study of several state-of-the-art algorithms applied to a large biomedical KG.
- We show that our proposed approach outperforms state-of-the-art alternatives on a highly relevant biomedical prediction task (drug repurposing).

As an application of our method, we focus on the drug repurposing problem, which is characterized by finding new treatment targets for existing drugs. By repurposing existing drugs, available knowledge about drug-disease-interactions can be leveraged to reduce time and cost for developing new drugs significantly. A recent example is the repositioning of the medication remdesivir for the novel coronavirus disease COVID-19. We aim at generating candidates for the drug repurposing task with machine learning reasoning methods and formulate the task as a link prediction problem, where both compounds and diseases correspond to entities in a KG.

## 2. Notation

Let  $\mathcal{E}$  denote the set of entities in a KG and  $\mathcal{R}$  the set of binary relations. Elements in  $\mathcal{E}$  correspond to biomedical

entities including, e. g., chemical compounds, diseases, and genes. Each entity belongs to a unique type in  $\mathcal{T}$ , defined by the mapping  $\tau : \mathcal{E} \rightarrow \mathcal{T}$ . For example,  $\tau(AURKC) = Gene$  indicates that the entity *AURKC* has type *Gene*. We define a KG  $\mathcal{KG} \subset \mathcal{E} \times \mathcal{R} \times \mathcal{E}$  as a collection of triples of the form  $(h, r, t)$ , which consists of head, relation, and tail. Head and tail entities correspond to nodes in the graph, while the relation indicates the type of edge between them. For any relation  $r \in \mathcal{R}$ , we denote the corresponding inverse relation with  $r^{-1}$  (i. e.,  $(h, r, t)$  is equivalent to  $(t, r^{-1}, h)$ ). Triples in  $\mathcal{KG}$  are interpreted as true known facts. For example, the triple  $(Sorafenib, treats, Liver\ Cancer) \in \mathcal{KG}$  in Figure 2 corresponds to the fact that the kinase inhibitor drug sorafenib is approved for the treatment of liver cancer.

We further distinguish between two types of paths: instance paths and metapaths. An instance path of length  $T \in \mathbb{N}$  on  $\mathcal{KG}$  is given by a sequence

$$(e_1 \xrightarrow{r_1} e_2 \xrightarrow{r_2} \dots \xrightarrow{r_T} e_{T+1}),$$

where  $(e_i, r_i, e_{i+1}) \in \mathcal{KG}$ . Moreover, we call

$$(\tau(e_1) \xrightarrow{r_1} \tau(e_2) \xrightarrow{r_2} \dots \xrightarrow{r_T} \tau(e_{T+1}))$$

a metapath. For example,

$$(Sorafenib \xrightarrow{treats} Liver\ Cancer \xrightarrow{resembles} Kidney\ Cancer)$$

constitutes an instance path of length 2, where

$$(Compound \xrightarrow{treats} Disease \xrightarrow{resembles} Disease)$$

is the corresponding metapath.

Logical rules (e.g., the commonly used Horn clauses) are usually written in the form *head*  $\leftarrow$  *body*. The head can be written out as a triple, and the body can be expressed as a metapath. Define  $CtD := (Compound, treats, Disease)$ . Then, a rule with respect to edges of type *treats* is of the generic form

$$CtD \leftarrow (Compound \xrightarrow{r_1} Type_2 \xrightarrow{r_2} \dots \xrightarrow{r_T} Disease).$$

In particular, the body of a rule corresponds to a metapath starting at a compound and terminating at a disease. The goal is to find instance paths where the corresponding metapaths match the body of a rule to predict a new relation between the source and the target of the instance path. The confidence of a rule indicates how often a rule is correct and is defined as the rule support divided by the body support in the data.

## 3. Our Method

We pose the task of drug repurposing as a link prediction problem based on graph traversal. Starting at a query entity

(e.g., a compound to be repurposed), an agent performs a walk on the graph by sequentially transitioning to a neighboring node. The decision of which transition to make is determined by a stochastic policy. Each subsequent transition is added to the current path, extending the reasoning chain, until a finite number of transitions is reached. The general approach is inspired by the reinforcement learning method MINERVA (Das et al., 2018), with our primary contribution coming from the incorporation of logical rules into the training process.

The state of the environment consists of the entity  $e_t$  where the agent is located at time  $t$ , the source entity  $e_c$ , and the target entity  $e_d$ , where  $e_c$  and  $e_d$  correspond to the compound that we aim to repurpose and the target disease, respectively. Thus, a state  $S_t$  for time  $t \in \mathbb{N}$  is represented by  $S_t := (e_t, e_c, e_d)$ . The agent is given no information about the target disease so that the observed part of the state space is given by  $(e_t, e_c) \in \mathcal{E}^2$ . Let  $e \in \mathbb{R}^d$  denote the embedding of entity  $e$  and  $r \in \mathbb{R}^d$  the embedding of relation  $r$ . The set of available actions contains all outgoing edges from the node  $e_t$  with the corresponding target nodes and the option to stay at the current node with no transition. We denote with  $A_t \in \mathcal{A}_{S_t}$  the action that the agent performed at time  $t$ . The environment evolves deterministically by updating the state according to the previous action.

The agent encodes previous actions via a multi-layered LSTM (Hochreiter & Schmidhuber, 1997)

$$\mathbf{h}_t = \text{LSTM}([\mathbf{a}_{t-1}, \mathbf{e}_c]), \quad (1)$$

where  $\mathbf{a}_{t-1} := [\mathbf{r}_{t-1}, \mathbf{e}_t] \in \mathbb{R}^{2d}$  corresponds to the vector space embedding of the previous action (or the zero vector at time  $t = 0$ ). The action distribution is given by

$$\mathbf{d}_t = \text{softmax}(\mathbf{A}_t (\mathbf{W}_2 \text{ReLU}(\mathbf{W}_1 \mathbf{h}_t))), \quad (2)$$

where  $\mathbf{W}_1$  and  $\mathbf{W}_2$  are weight matrices and the rows of  $\mathbf{A}_t \in \mathbb{R}^{|\mathcal{A}_{S_t}| \times 2d}$  contain the latent representations of all admissible actions from  $S_t$ . An action  $A_t \in \mathcal{A}_{S_t}$  is sampled according to  $A_t \sim \text{Categorical}(\mathbf{d}_t)$ . Overall,  $T$  transitions are sampled, resulting in a path denoted by

$$P := (e_c \xrightarrow{r_1} e_2 \xrightarrow{r_2} \dots \xrightarrow{r_T} e_{T+1}),$$

where  $T$  is the maximum path length. Equations (1) and (2) induce a stochastic policy, represented by  $\pi_\theta$  where  $\theta$  denotes the set of all trainable parameters, including all entity and relation embeddings.

Furthermore, let  $\mathcal{M} = \{M_1, M_2, \dots, M_m\}$  be the set of metapaths, where each element corresponds to the body of a rule. For every metapath  $M$ , we assign a score  $S(M) \in \mathbb{R}$  that indicates a quality measure of the corresponding rule, such as the confidence or the support with respect to making a correct prediction. For a path  $P$ , we denote with  $\tilde{P}$  the corresponding metapath.

During training, a terminal reward is computed according to

$$R = \mathbb{I}_{\{e_{T+1}=e_d\}} \left( 1 + \lambda \sum_{i=1}^m S(M_i) \mathbb{I}_{\{\tilde{P}=M_i\}} \right).$$

The first term indicates whether the agent has reached the correct target disease. The second term checks whether the metapath corresponds to the body of a rule and adds to the score accordingly. Heuristically speaking, we want to reward the agent with a higher score for extracting a metapath that corresponds to a body. The hyperparameter  $\lambda \geq 0$  balances the two components of the reward. For  $\lambda = 0$ , we recover MINERVA.

We employ REINFORCE (Williams, 1992) to maximize the expected rewards. Thus, the agent’s maximization problem is given by

$$\arg \max_{\theta} \mathbb{E}_{e_c \sim \mathcal{E}_c} \mathbb{E}_{A_1, A_2, \dots, A_T \sim \pi_\theta} [R | e_c], \quad (3)$$

where  $\mathcal{E}_c$  denotes the true underlying distribution of the set of chemical compounds.

## 4. Experiments

### 4.1. Dataset

Hetionet (Himmelstein et al., 2017) is a biomedical KG that integrates data from 29 highly reputable and cited public databases. It consists of 47,031 entities with 11 different types and 2,250,197 edges with 24 different types. We aim to predict edges with type *treats* between entities that correspond to compounds and diseases. The goal is to perform candidate ranking according to the likelihood of successful drug repurposing in a novel treatment application. There are 1552 compounds and 137 diseases in Hetionet with 775 observed links of type *treats* between compounds and diseases.

### 4.2. Metapaths as Background Information

Himmelstein et al. (2017) compiled a list of 1206 metapaths corresponding to various pharmacological efficacy mechanisms that connect entities of type *Compound* with entities of type *Disease*. Through hypothesis testing and domain expertise, they identified 31 effective metapaths that served as features for a logistic regression model. Out of these metapaths, we select the 10 metapaths as background information that have at most path length 3 and exhibit positive regression coefficients, indicating their importance for predicting drug efficacy. The metapaths are included as rule bodies in  $\mathcal{M}$ , where the rule head is always (*Compound, treats, Disease*). We estimate the confidence score for each rule by sampling 10,000 paths whose metapaths correspond to the rule body and use the confidence for the score  $S(M)$  (see Section 3). Table 1 shows the three metapaths with the highest confidences.

**Table 1.** Three metapaths and their scores.

$S(M)$	Metapath $M$
0.446	$(\text{Compound} \xrightarrow{\text{includes}^{-1}} \text{Pharmacologic Class} \xrightarrow{\text{includes}} \text{Compound} \xrightarrow{\text{treats}} \text{Disease})$
0.265	$(\text{Compound} \xrightarrow{\text{resembles}} \text{Compound} \xrightarrow{\text{resembles}} \text{Compound} \xrightarrow{\text{treats}} \text{Disease})$
0.184	$(\text{Compound} \xrightarrow{\text{binds}} \text{Gene} \xrightarrow{\text{associates}^{-1}} \text{Disease})$

### 4.3. Experimental Setup

We apply our method, denoted by MINERVA+, to Hetionet and calculate hits@1, hits@3, hits@10, and the mean reciprocal rank (MRR). During inference, a beam search is carried out, and the entities are ranked by the probability of their corresponding paths. Moreover, we consider another evaluation scheme (MINERVA+ (pruned)) that retrieves and ranks only those paths from the test rollouts that correspond to one of the metapaths. All the other extracted paths are not considered in the ranking. We compare our approach with the path-based method MINERVA, the rule-based method AnyBURL, and the embedding-based methods TransE, RESCAL, and R-GCN.

### 4.4. Results

**Table 2.** Comparison with baseline methods.

Method	Hits@1	Hits@3	Hits@10	MRR
AnyBURL	0.139	0.254	0.358	0.210
AnyBURL (metapaths)	0.252	0.364	0.609	0.354
TransE	0.073	0.172	0.318	0.161
RESCAL	0.205	0.35	0.576	0.317
R-GCN	0.093	0.245	0.364	0.188
MINERVA	0.249	0.391	0.605	0.357
MINERVA+	0.294	0.437	0.615	0.396
MINERVA+ (pruned)	<b>0.319</b>	<b>0.468</b>	<b>0.628</b>	<b>0.416</b>

Table 2 displays the test results for the experiments. The reported values for MINERVA and MINERVA+ correspond to the mean across five independent training runs. The standard errors lie between 0.0072 and 0.0198. This indicates that the reported performance gains are highly significant.

AnyBURL only learns one rule for the relation *treats* that has a length of at least 2. To see the effect of applying a larger number of rules, we try a setting where we use the metapaths for the prediction step, which leads to significantly improved results. TransE and R-GCN show similar performance, and RESCAL performs best among the embedding-based methods. Applying the modified ranking scheme, our method yields performance gains of 26.6% for hits@1, 19.7% for hits@3, 3.1% for hits@10, and 16.5% for MRR with respect to best performing baseline method.

### 4.5. Discussion

Our method can act as a generic mechanism to inject domain knowledge into reinforcement learning-based reasoning methods on KGs (Lin et al., 2018; Xiong et al., 2017). While we employ rules that are extracted in a data-driven fashion, our method is agnostic towards the source of background information. The additional reward for extracting a rule (see Equation (3)) can be considered as a regularization that enforces the agent to walk along metapaths that generalize to unseen instances.

AnyBURL is strictly outperformed by both MINERVA and our method. Most likely, the large amount of high-degree nodes in Hetionet lead to the outcome that hardly any strong, predictive rules are extracted. Multi-hop reasoning methods contain a natural transparency mechanism by providing explicit inference paths. Surprisingly, our experimental findings show that path-based reasoning methods outperform existing black-box methods on the drug repurposing task without a trade-off between explainability and performance. Both TransE and RESCAL are trained to minimize the reconstruction error in the immediate first-order neighborhood, and our results indicate that these methods seem not to be suitable for the drug repurposing task. R-GCN is in principle capable of modeling long-term dependencies due to the receptive field containing the entire set of nodes in the multi-hop neighborhood. However, the aggregation and combination step of R-GCN essentially acts as a low-pass filter on the incoming signals, and in the presence of many high-degree nodes, the center nodes may receive an uninformative signal that smooths over the neighborhood embeddings.

To illustrate the applicability of our method, consider the compound sorafenib from Figure 2. The three highest predictions of our model for new target diseases include hematologic cancer, breast cancer, and Barrett’s esophagus. The database ClinicalTrials.gov (U. S. National Library of Medicine, 2000) lists 23 clinical studies for testing the effect of sorafenib on these three diseases, showing that the predictions are meaningful targets for further investigation.

### 5. Conclusion

We have proposed a novel neuro-symbolic knowledge graph reasoning approach that leverages path-based reasoning, representation learning, and logical rules. We apply our method to the highly relevant task of drug repurposing and compare our approach with both embedding-based and rule-based methods. We achieve better performance and an improvement of 26.6% for hits@1 and 16.5% for the mean reciprocal rank compared to popular baselines.



## Acknowledgements

This work has been supported by the German Federal Ministry for Economic Affairs and Energy (BMWi) as part of the project RAKI (no. 01MD19012C).

Supported by:



on the basis of a decision  
by the German Bundestag

## References

- Bordes, A., Usunier, N., Garcia-Durán, A., Weston, J., and Yakhnenko, O. Translating embeddings for modeling multi-relational data. In *Proceedings of the 26th International Conference on Neural Information Processing Systems*, volume 2 of *NIPS'13*, pp. 2787–2795, 2013.
- Das, R., Dhuliawala, S., Zaheer, M., Vilnis, L., Durugkar, I., Krishnamurthy, A., Smola, A., and McCallum, A. Go for a walk and arrive at the answer: reasoning over paths in knowledge bases using reinforcement learning. In *Proceedings of the 6th International Conference on Learning Representations*, 2018.
- Dörpinghaus, J. and Jacobs, M. Semantic knowledge graph embeddings for biomedical research: data integration using linked open data. In *Proceedings of the Posters and Demo Track of the 15th International Conference on Semantic Systems (SEMANTiCS)*, volume 2451 of *CEUR Workshop Proceedings*, 2019.
- Hildebrandt, M., Li, H., Koner, R., Tresp, V., and Günemann, S. Scene graph reasoning for visual question answering. *arXiv:2007.01072*, 2020a.
- Hildebrandt, M., Serna, J. A. Q., Ma, Y., Ringsquandl, M., Joblin, M., and Tresp, V. Reasoning on knowledge graphs with debate dynamics. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence*, 2020b.
- Himmelstein, D. S., Lizee, A., Hessler, C., Brueggeman, L., Chen, S. L., Hadley, D., Green, A., Khankhanian, P., and Baranzini, S. E. Systematic integration of biomedical knowledge prioritizes drugs for repurposing. *Elife*, 6: e26726, 2017.
- Hochreiter, S. and Schmidhuber, J. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.
- Lin, X. V., Socher, R., and Xiong, C. Multi-hop knowledge graph reasoning with reward shaping. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 3243–3253, 2018.
- Meilicke, C., Fink, M., Wang, Y., Ruffinelli, D., Gemulla, R., and Stuckenschmidt, H. Fine-grained evaluation of rule- and embedding-based systems for knowledge graph completion. In *The Semantic Web – ISWC 2018*, volume 11136 of *Lecture Notes in Computer Science*, pp. 3–20, 2018.
- Meilicke, C., Chekol, M. W., Ruffinelli, D., and Stuckenschmidt, H. Anytime bottom-up rule learning for knowledge graph completion. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pp. 3137–3143, 2019.
- Nickel, M., Tresp, V., and Kriegel, H.-P. A three-way model for collective learning on multi-relational data. In *Proceedings of the 28th International Conference on Machine Learning*, 2011.
- Schlichtkrull, M., Kipf, T. N., Bloem, P., Van Den Berg, R., Titov, I., and Welling, M. Modeling relational data with graph convolutional networks. In *The Semantic Web – ESWC 2018*, volume 10843 of *Lecture Notes in Computer Science*, pp. 593–607, 2018.
- U. S. National Library of Medicine. [clinicaltrials.gov](http://clinicaltrials.gov), 2000.
- Williams, R. J. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, 8(3-4):229–256, 1992.
- Xiong, W., Hoang, T., and Wang, W. Y. DeepPath: A reinforcement learning method for knowledge graph reasoning. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pp. 564–573, 2017.