Neural-Semantic Methods for Link Prediction and Explanation in Knowledge Graphs

Roberto Barile[a,*]

[a]Dipartimento di Informatica, Università degli Studi di Bari *Aldo Moro*, Italy
ORCID (Roberto Barile): https://orcid.org/0009-0007-3058-8692

1 Introduction

Knowledge Graphs (KGs) are symbolic and machine-understandable representations of knowledge [10]. They rely on graph data models for representing facts (triples) consisting of entities (nodes) and binary relationships (edges). KGs may also deliver schema level knowledge, typically via ontologies, that enables reasoning services by specifying the semantics of the symbols (e.g., entity and relationship names) in the KG. Despite KGs having proved useful in both academic and business initiatives [19, 24], they are incomplete and noisy [10] because they often result from a complex (semi-automatic) building process.

Among KG refinement tasks, *Link Prediction* (LP) is one of the most investigated; it aims at predicting missing facts. Neural/numerical LP methods based on *Knowledge Graph Embedding* (KGE) models are prominent as they have showed competitive effectiveness and scalability [20]. KGE models are representation learning solutions that encode the entities and relationships of a KG as low-dimensional embeddings (vectors) that reflect the structure of the KG and that can be leveraged to solve downstream tasks, such as LP, using efficient linear algebra operations. However, LP methods based on KGE models generally take into account solely facts and thus dismiss schema and reasoning services. This motivates the interest of this PhD project in pursuing *Neural-Semantic* (NeSem) LP methods, that is, LP methods that are semantically enriched by leveraging extensively schema and reasoning services. Specifically, we intend to investigate the following Research Question (**RQ**):

RQ 1 *Can we formalize NeSem LP methods showing improved effectiveness wrt SOTA LP methods?*

Furthermore, LP methods based on KGE models are *opaque boxes* because the interpretation of embeddings is implicit/latent and no evidence/explanation for the predictions can be obtained. Explanations and explicit interpretations are crucial, especially in fields where stakeholders need to understand facts (predictions) before relying on them for critical decisions. For example, in pharmacology, LP may be used for predicting the side effects of a drug [18]: stakeholders need to understand the prediction before relying on them for decisions about funding of research in the drug. In contrast, symbolic LP methods [13, 15] are *clear boxes*: the interpretation of the symbols is explicit. However, symbolic LP methods currently receive limited attention since they have often showed limited scalability. This motivates the introduction of *Neural-Symbolic* NeSy LP methods [25, 5] for preserving the interpretability of symbolic LP methods, while maintaining, as much as possible, the effectiveness and scalability of neural/numerical ones. However, current NeSy LP methods [25] dismiss schema and reasoning services. Hence, we intend to investigate the following RQ:

RQ 2 *Can we formalize NeSem LP methods showing improved interpretability wrt SOTA LP methods?*

An alternative approach for supplying explanations for predicted facts is to provide the explanans [1], i.e., the pieces of knowledge (e.g., facts) associated to the prediction. LP eXplanation (LP-X) methods [23] select the explanans often based on the influence of facts on the effectiveness of the LP methods. They are often *post-hoc* (after training) methods that work for any LP method. However, the existing LP-X methods generally disregard schema and reasoning services. Moreover, as KGs are incomplete, the explanantia are likely to be incomplete too. Hence, in this PhD project, we aim at NeSem LP-X methods that leverage schema and reasoning services for reasoning on the association of existing facts to predictions and for completing the explanantia by generating hypothesis, i.e., facts missing from the KG yet possibly true. This leads to the following **RQ**:

RQ 3 *Can we formalize NeSem LP-X methods showing improved effectiveness wrt SOTA LP-X methods?*

As also emerged in the AAAI 2025 Presidential Panel [2], *"principled empirical evaluation* [is] *more important than ever"*. However, a major problem of existing LP-X methods is the lack of a standard protocol [23, 14] for evaluating the resulting explanations, thus making hard the comparison of explanations for the same prediction, but coming from different LP-X methods. This leads to the last **RQ** of this PhD project:

RQ 4 *Can we formalize and standardize an algorithmic protocol for evaluating explanations?*

2 Contributions Provided

In this section, we illustrate the contributions that we made so far, specifically for **RQ 3** and **RQ 4**.

* Corresponding Author. Email: r.barile17@phd.uniba.it

[1] plural: explanantia

[2] https://aaai.org/about-aaai/presidential-panel-on-the-future-of-ai-research/

As for **RQ 3**, we investigated post-hoc LP-X methods, focusing specifically on methods based on combinatorial optimization, as these work for any LP method. They aim at finding the best explanans, according to an objective function, in a finite set of possible explanantia constructed from the KG. In this class of methods, the most effective one is KELPIE [21], which, however, shows significant computational complexity. In order to tackle this issue, we proposed KELPIE++ [2], that in principle, could enhance not only KELPIE but any LP-X method based on combinatorial optimization, with minor modifications. Specifically, KELPIE++ summarizes, via quotient graphs, the set of possible explanantia by grouping semantically related ones [22], thus making optimization more efficient and robust to noise. KELPIE++ also integrates a semantic similarity measure in the selection of the possible explanantia from the KG. Introducing quotient graphs largely increased efficiency and slightly increased effectiveness, while the semantic similarity measure slighlty increased effectiveness wrt KELPIE. Moreover, we formalized the complementary LP-X method IMAGINE [4] that computes explanations consisting of newly generated hypotheses. Specifically, while SOTA LP-X methods built explanantia starting from the facts (triples) in the KG, IMAGINE generates hypotheses by summarizing the KG and concretizing it back: the concretization introduces additional facts previously not present in the KG. IMAGINE outperformed both KELPIE and KELPIE++ when considering wrong predictions, but was less effective for correct predictions. This may be due to limitations of the adopted evaluation protocol (grounded on re-training) that is the one proposed by KELPIE for the sake of comparison and that is tailored to explanantia consisting of existing facts.

Concerning **RQ 4**, we formalized LP-DIXIT [3], that measures the quality of an explanans returned by a LP-X method. To the best of our knowledge, LP-DIXIT is the sole existing protocol that is user guided yet fully algorithmic and generic, i.e., working for explanans coming from any LP-X method and thus allowing to compare explanantia coming from different LP-X methods. LP-DIXIT grounds on a hypothesis formulated in cognitive sciences [9]: predictions are understandable when simulatable (predictable). Specifically, LP-DIXIT measures the *Forward Simulatability Variation* (FSV) induced by an explanans for a prediction (made by a LP method), that is it measures the variation between the simulatability (or predictability) of a prediction without and with an explanans. A prediction is simulatable (with an explanans) if a (human) verifier can correctly simulate the prediction, i.e., can hypothesize the output of the LP method given the same input provided to the LP method (and the explanans). LP-DIXIT bypasses the need for expert users by employing *Large Language Models* (LLMs) in order to mimic human users. We validated the hypothesis that LLMs can mimic human users by comparing the FSV measured with LP-DIXIT with measures in a ground-truth dataset [8]. Moreover, we employed LP-DIXIT for comparing different LP-X methods. The outcomes suggest that *less is more*: the most effective explanans are those consisting of exactly one fact.

3 Research Directions for the Remaining Work

The rest of the PhD (started on October 2024) will be devoted to answering **RQ 1** and **RQ 2** as well as to further investigate the solutions in reply to **RQ 3** and **RQ 4**.

As for **RQ 1**, the main goal is the formalization of NeSem LP methods that take into account schema and reasoning services. Being *Description Logics* (DLs) [1] the theoretical framework underlying OWL, that is the de facto standard ontology representation language, in order to target **RQ 1**, we aim at formalizing methods imposing logical (DL) requirements [7] within the main KGE architectural components. Specifically, we target both loss functions and the sampling of negative (false) facts required for learning KGE models. Indeed, since KGs adopts the *Open World Assumption*, which states that missing facts in the KG cannot be considered false unless derivable as such, negative facts can be hardly found within KGs. Hence, techniques for generating artificial (possibly false) negatives are adopted. Instead, we aim for a NeSem approach that would generate actual negatives. Moreover, we plan to investigate on injecting DL justifications [11] within the learning process. As for the experimental evaluation of the newly developed NeSem LP methods, we aim at showing that, despite the possible additional complexity, schema and reasoning services pay off in terms of effectiveness. Particularly, we plan to conduct comparative experimental studies grounded on established LP evaluation protocols [20].

Regarding **RQ 2**, the main goal is the formalization of *interpretable* NeSem LP methods. To this purpose, inspired by related works on neural theorem provers [17], we aim at formalizing extensions of the tableau algorithm (generally used for reasoning in DLs) over KGEs, by relaxing logical operators in order to handle numerical representations. We also plan to further extend the contribution in reply to **RQ 3** by exploring *abductive* reasoning and explanation based on DLs [12] for computing explanantia that consists of observed facts and new generated hypothesis. Moreover, we plan to explore other semantic similarity measures that take into account also the semantics of relationships, e.g., transitivity, symmetry. Furthermore, we target approaches to compute generalizations of multiple explanantia to be used for reasoning, e.g., through concept learning methods [6]. Also for **RQ 2** and **RQ 3** the experimental evaluation, we aim at showing that schema and reasoning services pay off in terms of effectiveness, despite the possible additional complexity. We plan to assess the solutions experimentally based on the protocols resulting from **RQ 4** or, alternatively, adopting the re-training evaluation protocol, proposed for evaluating KELPIE, that compares the LP accuracy of the KGE model used for the prediction to the one of a model trained on a perturbed KG, where the facts in the explanantia have been added, removed or isolated.

Concerning **RQ 4**, we plan to extend LP-DIXIT in order to assess also if explanations facilitate learning/generalizations as people ask for explanations not only for understanding, but also for learning [16]. Moreover, we target the usage of explanation methods for reasoning based on DLs as ground-truth for LP-X methods. Furthermore, we plan to perform empirical studies on the consistency between different evaluation protocols. We also aim at validating the proposed protocols wrt expert curated ground-truth explanations and evaluations.

Overall, the PhD project will offer the following contributions: 1) a class of effective NeSem LP methods; 2) a suite of NeSem *interpretable* LP methods; 3) a family of NeSem LP-X methods, providing semantic explanantia for LP tasks; 4) a new principled and standard protocol for evaluating explanations.

Acknowledgements

This work was partially supported by project *FAIR - Future AI Research* (PE00000013), spoke 6 - Symbiotic AI (https://future-ai-research.it/) and by PRIN project *HypeKG - Hybrid Prediction and Explanation with Knowledge Graphs* (Prot. 2022Y34XNM, CUP H53D23003700006), both under the PNRR MUR program funded by the European Union - NextGenerationEU

References

[1] F. Baader. The description logic handbook: Theory, implementation, and applications. *Cambridge University Press google schola*, 2:7–26, 2003.

[2] R. Barile, C. d'Amato, and N. Fanizzi. Explanation of link predictions on knowledge graphs via levelwise filtering and graph summarization. In *European Semantic Web Conference*, pages 180–198. Springer, 2024.

[3] R. Barile, C. d'Amato, and N. Fanizzi. Lp-dixit: Evaluating explanations for link predictions on knowledge graphs using large language models. In *Proceedings of the ACM on Web Conference 2025*, WWW '25, page 4034–4042, New York, NY, USA, 2025. Association for Computing Machinery. ISBN 9798400712746. doi: 10.1145/3696410.3714667. URL https://doi.org/10.1145/3696410.3714667.

[4] R. Barile, C. d'Amato, and N. Fanizzi. Additive counterfactuals for explaining link predictions on knowledge graphs. In M. Alam, M. Rospocher, M. van Erp, L. Hollink, and G. A. Gesese, editors, *Knowledge Engineering and Knowledge Management*, pages 346–363, Cham, 2025. Springer Nature Switzerland. ISBN 978-3-031-77792-9.

[5] L. N. DeLong, R. F. Mir, and J. D. Fleuriot. Neurosymbolic AI for reasoning over knowledge graphs: A survey. *IEEE Transactions on Neural Networks and Learning Systems*, 2024.

[6] N. Fanizzi, C. d'Amato, and F. Esposito. Dl-foil concept learning in description logics. In *Inductive Logic Programming: 18th International Conference, ILP 2008 Prague, Czech Republic, September 10-12, 2008 Proceedings 18*, pages 107–121. Springer, 2008.

[7] E. Giunchiglia, F. Imrie, M. vanáderáSchaar, and T. Lukasiewicz. Machine learning with requirements: A manifesto. *Neurosymbolic Artificial Intelligence*, (Preprint):1–13.

[8] N. Halliwell, F. Gandon, and F. Lecue. User Scored Evaluation of Non-Unique Explanations for Relational Graph Convolutional Network Link Prediction on Knowledge Graphs. In *K-CAP '21: Proceedings of the 11th Knowledge Capture Conference*, pages 57–64, New York, New York, USA, Dec. 2021. ACM. ISBN 978-1-4503-8457-5. doi: 10.1145/3460210.3493557.

[9] R. R. Hoffman, S. T. Mueller, G. Klein, and J. Litman. Metrics for Explainable AI: Challenges and Prospects, Feb. 2019.

[10] A. Hogan, E. Blomqvist, M. Cochez, C. D'amato, G. D. Melo, C. Gutierrez, S. Kirrane, J. E. L. Gayo, R. Navigli, S. Neumaier, A.-C. N. Ngomo, A. Polleres, S. M. Rashid, A. Rula, L. Schmelzeisen, J. Sequeda, S. Staab, and A. Zimmermann. Knowledge Graphs. *ACM Computing Surveys*, 54(4):1–37, May 2022. ISSN 0360-0300, 1557-7341. doi: 10.1145/3447772.

[11] A. Kalyanpur, B. Parsia, M. Horridge, and E. Sirin. Finding all justifications of owl dl entailments. In *Proceedings of the 6th International The Semantic Web and 2nd Asian Conference on Asian Semantic Web Conference*, ISWC'07/ASWC'07, page 267–280, Berlin, Heidelberg, 2007. Springer-Verlag. ISBN 3540762973.

[12] P. Koopmann, W. Del-Pinto, S. Tourret, and R. A. Schmidt. Signature-based abduction for expressive description logics. In *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning*, volume 17, pages 592–602, 2020.

[13] J. Lajus, L. Galárraga, and F. Suchanek. Fast and exact rule mining with amie 3. In *The Semantic Web: 17th International Conference, ESWC 2020, Heraklion, Crete, Greece, May 31–June 4, 2020, Proceedings 17*, pages 36–52. Springer, 2020.

[14] L. Longo, M. Brcic, F. Cabitza, J. Choi, R. Confalonieri, J. Del Ser, R. Guidotti, Y. Hayashi, F. Herrera, and A. Holzinger. Explainable artificial intelligence (XAI) 2.0: A manifesto of open challenges and interdisciplinary research directions. *Information Fusion*, page 102301, 2024. doi: 10.1016/j.inffus.2024.102301.

[15] C. Meilicke, M. W. Chekol, P. Betz, M. Fink, and H. Stuckeschmidt. Anytime bottom-up rule learning for large-scale knowledge graph completion. *The VLDB Journal*, 33(1):131–161, 2024.

[16] T. Miller. Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267:1–38, 2019.

[17] P. Minervini, M. Bosnjak, T. Rocktäschel, and S. Riedel. Towards neural theorem proving at scale. *ArXiv*, abs/1807.08204, 2018. URL https://api.semanticscholar.org/CorpusID:49904196.

[18] V. Nováček and S. K. Mohamed. Predicting polypharmacy side-effects using knowledge graph embeddings. *AMIA Summits on Translational Science Proceedings*, 2020:449, 2020.

[19] T. Pellissier Tanon, G. Weikum, and F. Suchanek. YAGO 4: A Reasonable Knowledge Base. In A. Harth, S. Kirrane, A.-C. Ngonga Ngomo, H. Paulheim, A. Rula, A. L. Gentile, P. Haase, and M. Cochez, editors, *The Semantic Web*, volume 12123, pages 583–596. Springer-Verlag, Berlin, Heidelberg, 2020. ISBN 978-3-030-49460-5 978-3-030-49461-2. doi: 10.1007/978-3-030-49461-2_34.

[20] A. Rossi, D. Barbosa, D. Firmani, A. Matinata, and P. Merialdo. Knowledge Graph Embedding for Link Prediction: A Comparative Analysis. *ACM Transactions on Knowledge Discovery from Data*, 15(2):1–49, Apr. 2021. ISSN 1556-4681, 1556-472X. doi: 10.1145/3424672.

[21] A. Rossi, D. Firmani, P. Merialdo, and T. Teofili. Explaining Link Prediction Systems based on Knowledge Graph Embeddings. In Z. Ives, editor, *SIGMOD/PODS '22: Proceedings of the 2022 International Conference on Management of Data, Philadelphia, Pennsylvania, USA, June 12-17, 2022*, pages 2062–2075, New York, New York, USA, June 2022. ACM. ISBN 978-1-4503-9249-5. doi: 10.1145/3514221.3517887.

[22] A. Scherp, D. Richerby, T. Blume, M. Cochez, and J. Rau. Structural summarization of semantic graphs using quotients. *Transactions on Graph Data and Knowledge*, 1(1):12, 2023.

[23] S. Schramm, C. Wehner, and U. Schmid. Comprehensible artificial intelligence on knowledge graphs: A survey. *Journal of Web Semantics*, 79:100806, 2023.

[24] A. Singhal. Introducing the Knowledge Graph: Things, not strings. https://blog.google/products/search/introducing-knowledge-graph-things-not/, May 2012.

[25] J. Zhang, B. Chen, L. Zhang, X. Ke, and H. Ding. Neural, symbolic and neural-symbolic reasoning on knowledge graphs. *AI Open*, 2:14–35, 2021. doi: 10/gn8nzc.