Combining Rule Learning and Nonmonotonic Reasoning for Link Prediction in Knowledge Graphs

Francesca A. Lisi[a], Daria Stepanova[b]

[b]Dipartimento di Informatica & Centro Interdipartimentale di Logica ed Applicazioni
Università degli Studi di Bari "Aldo Moro", Bari, Italy
francesca.lisi@uniba.it
[a]Max Planck Institute for Informatics, Saarbrücken, Germany
dstepano@mpi-inf.mpg.de

Abstract. Learning rules from knowledge graphs (KG) is a crucial task for KG completion, cleaning and curation. The majority of existing approaches are capable of learning only Horn rules from KGs, which are, however, insufficiently expressive for capturing exceptions and thus might make incorrect predictions on missing facts. In this paper we discuss our recent progress in addressing this limitation. More specifically, we report the challenges of learning rules with exceptions, briefly describe our recent approach which combines rule learning and nonmonotonic reasoning and outline the ongoing and future research directions.

1 Introduction

Recent advances in information extraction have led to the so-called *knowledge graphs* (KGs), *i.e.* huge collections of *triples* in the form of $\langle \mathit{subject\ predicate\ object} \rangle$ according to the RDF data model [1] such as DBpedia [2], and YAGO [15]. These triples encode facts about the world and can be straightforwardly represented by means of unary and binary first-order logic (FOL) predicates. The unary predicates are the objects of the RDF type predicate, while the binary ones correspond to all other RDF predicates, *e.g.*, $\langle \mathit{alice\ type\ researcher} \rangle$ and $\langle \mathit{bob\ isMarriedTo\ alice} \rangle$ from the KG in Fig. 1 correspond to the facts $\mathit{researcher}(\mathit{alice})$ and $\mathit{isMarriedTo}(\mathit{bob}, \mathit{alice})$.

Since KGs are automatically constructed, they are inherently *incomplete*. Therefore, they are naturally treated under the Open World Assumption (OWA). The task of *completion* (also known as *link prediction*) is of crucial importance for the curation of KGs. To this aim, rule mining techniques (*e.g.*, [3,9]) have been exploited to automatically build rules able to make predictions on missing links. However, they mine Horn rules, which are insufficiently expressive to capture exceptions, and might thus deduce incorrect facts. For example, the following rule

$$\mathit{r1} : \mathit{livesIn}(Y, Z) \leftarrow \mathit{isMarriedTo}(X, Y), \mathit{livesIn}(X, Z)$$

can be mined from the KG in Fig. 1 and used to produce the facts $\mathit{livesIn}(\mathit{alice},\mathit{berlin})$, $\mathit{livesIn}(\mathit{dave},\mathit{chicago})$ and $\mathit{livesIn}(\mathit{lucy}, \mathit{amsterdam})$. Observe that the first two predicted facts might actually be wrong. Indeed, both alice and dave are researchers, and the rule $\mathit{r1}$ could be suspected to have $\mathit{researcher}$ as a potential exception.

[1] https://www.w3.org/RDF/

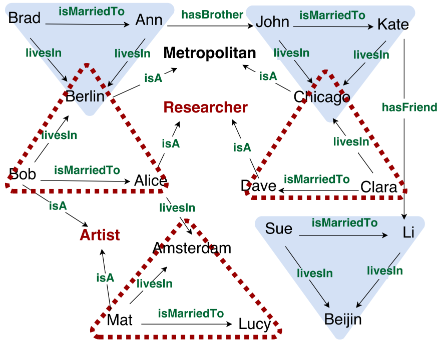

Fig. 1: Example of a Knowledge Graph

In this paper we summarize the challenges faced and the contributions made in a joint work [6] between the Max Planck Institute of Informatics in Saarbrücken, Germany, and the Department of Computer Science of the University of Bari, Italy, on the problem of link prediction in KGs. Moreover, we discuss the objectives of our ongoing and future research.

2 Challenges

Learning FOL rules for predictive purposes has been historically subject of investigation in Inductive Logic Programming (ILP). Exception handling has been traditionally faced in ILP by learning *nonmonotonic logic programs*, i.e., programs with negations (see, e.g., [18,17,4,10]). However, there are several important obstacles that prevent us from using the off-the-shelf nonmonotonic ILP algorithms. First, the *target predicates* can not be easily identified, since we do not know which parts of the considered KG need to be completed. A standard way of addressing this issue would be just to learn rules for all the different predicate names occurring in the KG. Unfortunately, this is unfeasible in our case given the huge size of KGs. Second, the *negative examples* are not available, and they can not be easily obtained from, *e.g.*, domain experts due to - once again - the huge size of KGs. A natural solution to cope with this issue is to learn rules from positive examples only. Third, the definition of a *language bias* turns out to be cumbersome since the schema of the KG is usually not available.

To overcome the obstacles mentioned above, it turns out to be appropriate to treat the KG completion problem as an unsupervised relational learning task, and exploit algorithms for relational association rule mining such as [9]. In [8] these techniques are applied to first learn a set of Horn rules, which subsequently can be revised by adding negated atoms to their bodies in order to account for exceptions. However, the proposed approach applies only to a flattened representation of a KG containing just unary facts.

3 Contributions

In [6] we have extended the results from [8] to KGs in their original relational form. More specifically, we have reformulated the KG completion problem as a *theory revision* problem, where, given a KG and a set of (previously learned) Horn rules, the task is to compute a set of *nonmonotonic rules*, such that the revised ruleset is more accurate for link prediction than the original one. Essentially, we are interested in tackling a theory revision problem, in which, as possible revision operations, we are only allowed to add negated atoms to the antecedents of the rules.

Our approach combines standard relational association rule mining techniques with a FOIL-like supervised learning algorithm, which is used to detect exceptions. More specifically, we propose a method that proceeds in four steps as follows: First, for every Horn rule we determine the *normal* and *abnormal* substitutions, i.e., substitutions that satisfy (resp. do not satisfy) the considered rule. Second, we compute the so-called *exception witness sets*, i.e., sets of predicates that are potentially involved in explaining why abnormal substitutions fail to follow the rule (e.g., *researcher* or *artist* in our example). Third, we construct candidate rule revisions by adding a single exception at a time. We devise quality measures for nonmonotonic rules to quantify their strength w.r.t the KG. We consider the crosstalk between the rules through the novel *partial materialization* technique instead of revising rules in isolation. Fourth, we rank rule revisions according to these measures to determine a ruleset that not only describes the data well but also shows a good predictive power by taking exceptions into account.

The contributions of our joint work up to now are:

- A theory revision framework, based on nonmonotonic relational rule learning, for capturing exceptions in rule-based approaches to KG completion.
- A methodology for computing exception candidates, measuring their quality, and ranking them taking into account the interaction among the rules.
- Implementation of our approach in a system prototype [2] and experiments with the YAGO3 and IMDB KGs, which demonstrate the gains of our method for rule quality as well as fact quality when performing KG completion.

4 Open Issues and Further Work

We have shortly presented an approach for predicting missing links in KGs which combines rule learning and nonmonotonic reasoning. More precisely, the approach allows for mining relational nonmonotonic rules from KGs under OWA by casting this problem into a theory revision task and by exploiting association rule mining methods to cope with the huge size of KGs. The approach is fully described in [6].

While we have made some progress in the context of learning nonmonotonic rules from KGs, some issues still remain to be explored.

Rich rule forms and complex exceptions. In [6] we have specified the language bias to learn rules of a predefined form. An interesting alternative to language bias is the use of a meta-querying language as suggested in [11]. Certainly, further extensions to

[2] `https://github.com/htran010589/nonmonotonic-rule-mining`

more complex combinations of exceptions as well as more general types of rules (e.g., with existentially quantified predicates in the head) are a natural future direction. Recent approaches to learning frequent graph patterns from a single labeled graph (e.g., [7]) could be exploited in this context. Another relevant stream of research concerns SPARQL reverse engineering [1], where given a set of positive and negative example entities, a SPARQL query is generated, whose answers include the majority of the positive examples and none of the negative examples. Indeed, from our abnormal and normal substitutions positive and negative examples can be constructed and methods from [1] could be further used for generating complex rule exception candidates.

Background knowledge. Another promising research direction is to learn rules from the data enriched with a schema represented in the form of a logical theory. The idea originates from the extensive work of Lisi on onto-relational rule learning (see, e.g, [14,12,13]) and has been more recently revised in [5]. However, the scalability of these algorithms should be improved and the revision of the induced rules to account for exceptions has not been targeted to the best of our knowledge.

Completeness meta-data. Since KGs are highly incomplete in the first place, rules mined from them are often spurious. For example, the faulty rule

$$isPoliticianOf(X, Z) \leftarrow hasChild(X, Y), isCitizenOf(Y, Z) \tag{1}$$

could be mined from an incomplete KG biased towards politicians. To avoid such incorrect correlations rule mining approaches may be improved by taking into account numerical meta-data about the KG's predicates and constants, e.g., "John has 3 children" or "Mary is a citizen of 2 countries", obtained from text with recent information extraction techniques like [16]. Such cardinality meta-data gives insight about the number of missing KG edges of a certain type, and could be effectively injected into the rule mining algorithms for the sake of avoiding erroneous rules like the one from above.

Mining constraints for inconsistency handling. Apart from incompleteness KGs obviously suffer from inaccuracies. Thus, another relevant research stream concerns the resolution of these inaccuracies by automatically extracting from the data constraints such as "a person can not graduate from the university before starting it unless this is his second degree".

Advanced evaluation strategies. On the practical side, we plan to develop advanced strategies for evaluating the rule learning approaches, which is very challenging due to the absence of the ideal graph (i.e., graph containing all true facts that hold in the world) and the large KG size.

References

1. Arenas, M., Diaz, G.I., Kostylev, E.V.: Reverse engineering SPARQL queries. In: Proceedings of the 25th International Conference on World Wide Web, WWW 2016, Montreal, Canada, April 11 - 15, 2016. pp. 239–249 (2016)
2. Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., Ives, Z.G.: DBpedia: A nucleus for a web of open data. In: Proc. of ISWC. pp. 722–735 (2007)
3. Chen, Y., Goldberg, S., Wang, D.Z., Johri, S.S.: Ontological Pathfinding: Mining First-Order Knowledge from Large Knowledge Bases. In: in Proc. of SIGMOD 2016. p. to appear (2016)

4. Corapi, D., Russo, A., Lupu, E.: Inductive logic programming as abductive search. In: Proceedings of ICLP. pp. 54–63 (2010)
5. d'Amato, C., Staab, S., Tettamanzi, A.G.B., Minh, T.D., Gandon, F.L.: Ontology enrichment by discovering multi-relational association rules from ontological knowledge bases. In: Proceedings of the 31st Annual ACM Symposium on Applied Computing, Pisa, Italy, April 4-8, 2016. pp. 333–338 (2016)
6. Dang Tran, H., Stepanova, D., Gad-Elrab, M., Lisi, F.A., Weikum, G.: Towards nonmonotonic relational learning from knowledge graphs. In: Cussens, J., Russo, A. (eds.) Inductive Logic Programming - 26th International Conference, ILP 2016, London, UK, September 4-6, 2016, Revised Selected Papers. Lecture Notes in Computer Science, vol. ? Springer (2017), under publication
7. Fan, W., Hu, C.: Big graph analyses: From queries to dependencies and association rules. Data Science and Engineering 2(1), 36–55 (2017)
8. Gad-Elrab, M.H., Stepanova, D., Urbani, J., Weikum, G.: Exception-enriched rule learning from knowledge graphs. In: Groth, P.T., Simperl, E., Gray, A.J.G., Sabou, M., Krötzsch, M., Lécué, F., Flöck, F., Gil, Y. (eds.) The Semantic Web - ISWC 2016 - 15th International Semantic Web Conference, Kobe, Japan, October 17-21, 2016, Proceedings, Part I. Lecture Notes in Computer Science, vol. 9981, pp. 234–251. Springer (2016)
9. Galrraga, L., Teflioudi, C., Hose, K., Suchanek, F.M.: Fast Rule Mining in Ontological Knowledge Bases with AMIE+. In: VLDB Journal (2015)
10. Law, M., Russo, A., Broda, K.: The ILASP system for learning answer set programs (2015)
11. Lisi, F.A.: Towards a metaquery language for mining the web of data. In: Cali, A., Poulovassilis, A., Wood, P.T. (eds.) Data Analytics: 31st British International Conference on Databases, BICOD 2017, London, UK, July 10-12, 2017, Proceedings. Lecture Notes in Computer Science, vol. 10365, pp. 1–4. Springer International Publishing AG
12. Lisi, F.A.: $\mathcal{AL}$-QuIn: An Onto-Relational Learning System for Semantic Web Mining. International Journal on Semantic Web and Information Systems 7(3), 1–22 (2011)
13. Lisi, F.A.: Learning onto-relational rules with inductive logic programming. In: Lehmann, J., Völker, J. (eds.) Perspectives on Ontology Learning, Studies on the Semantic Web, vol. 18, pp. 93–111. IOS Press/AKA (2014)
14. Lisi, F.A., Malerba, D.: Inducing Multi-Level Association Rules from Multiple Relations. Machine Learning 55(2), 175–210 (2004)
15. Mahdisoltani, F., Biega, J., Suchanek, F.M.: YAGO3: A knowledge base from multilingual wikipedias. In: Proceedings of CIDR (2015)
16. Mirza, P., Razniewski, S., Nutt, W.: Expanding wikidata's parenthood information by 178%, or how to mine relation cardinality information. In: Proceedings of the ISWC 2016 Posters & Demonstrations Track co-located with 15th International Semantic Web Conference (ISWC 2016), Kobe, Japan, October 19, 2016. (2016)
17. Ray, O.: Nonmonotonic abductive inductive learning. Journal of Applied Logic 3(7), 329–340 (2008)
18. Sakama, C.: Induction from answer sets in nonmonotonic logic programs. ACM Trans. Comput. Log. 6(2), 203–231 (2005)