Completeness-aware Rule Learning from Knowledge Graphs*

Thomas Pellissier Tanon[1,2], **Daria Stepanova**[1], **Simon Razniewski**[1], **Paramita Mirza**[1] and **Gerhard Weikum**[1]
[1] Max Planck Institute of Informatics
[2] Télécom ParisTech
{tpelliss,dstepano,srazniew,paramita,weikum}@mpi-inf.mpg.de

Abstract

Knowledge graphs (KGs) are huge collections of primarily encyclopedic facts that are widely used in entity recognition, structured search, question answering, and other tasks. Rule mining is commonly applied to discover patterns in KGs. However, unlike in traditional association rule mining, KGs provide a setting with a high degree of *incompleteness*, which may result in the wrong estimation of the quality of mined rules, leading to erroneous beliefs such as all artists have won an award. In this paper we propose to use (in-)completeness meta-information to better assess the quality of rules learned from incomplete KGs. We introduce completeness-aware scoring functions for relational association rules. Experimental evaluation both on real and synthetic datasets shows that the proposed rule ranking approaches have remarkably higher accuracy than the state-of-the-art methods in uncovering missing facts.

1 Introduction

Motivation. Advances in information extraction have led to general-purpose knowledge graphs (KGs) containing billions of positive facts about the world (e.g., [Carlson *et al.*, 2010; Bollacker *et al.*, 2007; Auer *et al.*, 2007; Mahdisoltani *et al.*, 2015]). KGs are widely applied in semantic web search, question answering, web extraction and many other tasks. Unfortunately, due to their wide scope, KGs are generally incomplete. To account for the incompleteness, they typically adopt the Open World Assumption (OWA) under which missing facts are treated as unknown rather than false.

An important task over KGs is rule learning [Galarraga *et al.*, 2015; Gad-Elrab *et al.*, 2016; Sazonau *et al.*, 2015; Wang and Li, 2015; Lisi, 2010; d'Amato *et al.*, 2016], which is relevant for a variety of applications ranging from knowledge graph curation (completion, error detection) [Paulheim, 2017] to data mining and semantic culturonomics [Suchanek and Preda, 2014]. However, since such rules are learned from incomplete data, they might be erroneous and might make incorrect predictions on missing facts. For example, from the KG in Fig. 1 the rule

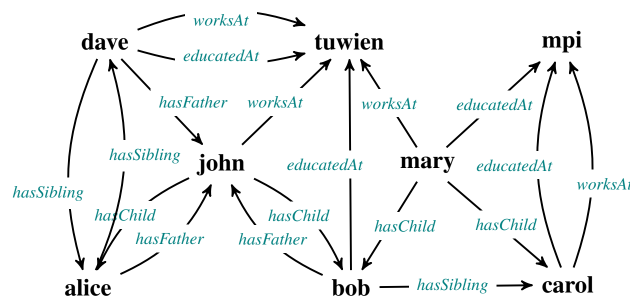

Figure 1: Example KG

$r_1 : hasChild(X, Y) \leftarrow worksAt(X, Z), educatedAt(Y, Z)$ could be mined, which states that workers of certain institutions often have children among the people educated there, as this is frequently the case for popular scientists. While r_1 is clearly not universal and should be ranked lower than the rule r_2 : *hasSibling(X,Z)* ← *hasFather(X, Y), hasChild(Y, Z)*, standard rule measures like confidence (i.e., conditional probability of the rule's head given its body) incorrectly favor r_1 over r_2 for the given KG.

Efforts have been put into adding completeness information to databases [Levy, 1996; Etzioni *et al.*, 1997] and recently to knowledge bases [Razniewski *et al.*, 2016; Darari *et al.*, 2013]. This could be done by detecting the concrete numbers of facts of certain types that hold in the real world (e.g., *"Einstein has 3 children"*) exploiting Web extraction or crowd-sourcing [Prasojo *et al.*, 2016; Darari *et al.*, 2016; Mirza *et al.*, 2017]. Such meta-data provides a lot of hints about the KG's topology, and reveals parts that should be especially targeted by rule learning methods. However, surprisingly, to date, no systematic way of making use of such information in rule learning exists.

In this work we exploit meta-data about the expected number of edges in KGs for better assessment of learned rules.

State of the art and its limitations. In [Galarraga *et al.*, 2015] a completeness-aware rule scoring based on the partial completeness assumption (PCA) was introduced. The idea of PCA is that whenever at least one object for a given subject and a predicate is in a KG (e.g., *"Eduard is Einstein's child"*), then all objects for that subject-predicate pair (*Einstein's children*) are assumed to be known. This assumption was taken into account in rule scoring, and empirically it turned out to

*Full version appeared at ISWC 2017.

be indeed valid in real-world KGs for some topics. However, PCA inappropriately treats cases, when edges in a graph are randomly missing. Similarly, whether to count absence of contradiction as confirmation for default rules was discussed in [Doppa *et al.*, 2011]. In [Galárraga *et al.*, 2017] new completeness data was learned from a KG by taking as ground truth completeness data obtained via crowd-sourcing. The acquired statements were used in a post-processing step of rule learning to filter out violating predictions. However, this kind of filtering does not have any impact on the quality of the mined rules and the incorrect predictions for instances about which no completeness information exists.

Contributions. In this shortened version of our work [Pellissier Tanon *et al.*, 2017] we present an approach that accounts for meta-data about the number of edges that should exist for given subject-predicate pairs in the ranking stage of rule learning. The novel rule ranking measures have been evaluated both on real-world and synthetic datasets, showing that they outperform existing ones both with respect to the quality of the mined rules and the predictions they produce.

2 Preliminaries

Knowledge graphs. Knowledge graphs (KG) represent interlinked collections of factual information, and they are often encoded using the RDF data model [Lassila and Swick, 1999]. The content of KGs is a set of $\langle subject\ predicate\ object \rangle$ triples, e.g., $\langle john\ hasChild\ alice \rangle$. For encyclopedic knowledge graphs on the semantic web, usually the open world assumption (OWA) is employed, i.e., these graphs contain only a subset of the true information.

In the following we write triples using binary predicates, like $hasChild(john, alice)$. A signature of a KG $\mathcal{G}$ is $\Sigma_{\mathcal{G}} = \langle \mathbf{R}, \mathcal{C} \rangle$, where $\mathbf{R}$ is the set of binary predicates and $\mathcal{C}$ is the set of constants appearing in $\mathcal{G}$. Following [Darari *et al.*, 2013], we define the gap between the available graph $\mathcal{G}^a$ and the ideal graph $\mathcal{G}^i$, which contains all correct facts over $\mathbf{R}$ and $\mathcal{C}$ that hold in the real world. An *incomplete data source* is a pair $G = (\mathcal{G}^a, \mathcal{G}^i)$ of two KGs, where $\mathcal{G}^a \subseteq \mathcal{G}^i$ and $\Sigma_{\mathcal{G}^a} = \Sigma_{\mathcal{G}^i}$. Note that the ideal graph $\mathcal{G}^i$ is an imaginary construct whose content is generally not known. What is known instead is to which extent the available graph approximates/lacks information wrt. the ideal graph, e.g., *"Einstein is missing 2 children and Feynman none"*). We formalize this knowledge as cardinality assertions in Sec. 3.

Rule learning. Association rule learning concerns the discovery of frequent patterns in a data set and the subsequent transformation of these patterns into rules. A *conjunctive query* Q over $\mathcal{G}$ is of the form $p_1(x_1, y_1), \ldots, p_m(x_m, y_m)$, where x_i and y_i are symbolic variables or constants and $p_i \in \mathbf{R}$ are binary predicates. The *answer* of Q on $\mathcal{G}$ is the set $Q(\mathcal{G}) = \{(\nu(x_1), \ldots, \nu(x_m), \nu(y_1), \ldots, \nu(y_m)) \mid \forall i : p_i(\nu(x_i), \nu(y_i)) \in \mathcal{G}\}$ where ν is a function that maps variables and constants to elements of $\mathcal{C}$. The *support* of Q in $\mathcal{G}$ is the number of distinct tuples in the answer of Q on $\mathcal{G}$ [Dehaspe and De Raedt, 1997].

An *association rule* is of the form $Q_1 \Rightarrow Q_2$, such that Q_1 and Q_2 are both conjunctive queries and $Q_1 \subseteq Q_2$, i.e., $Q_1(\mathcal{G}') \subseteq Q_2(\mathcal{G}')$ for any possible KG $\mathcal{G}'$. We call Q_2 the body of the rule and Q_1 its head. In this work we exploit association rules for reasoning purposes, and thus (with some abuse of notation) treat them as logical rules, i.e., for $Q_1 \Rightarrow Q_2$ we write $Q_2 \backslash Q_1 \leftarrow Q_1$, where $Q_2 \backslash Q_1$ refers to the set difference between Q_2 and Q_1 seen as sets of atoms.

Classical scoring of association rules is based on *rule support*, *body support* and *confidence*. For a rule $r : h(X, Y) \leftarrow B$ where B is a conjunctive query over the variables $\vec{Z} \supseteq X, Y$ and/or constants, they are defined in [Galarraga *et al.*, 2015] as:

$$supp(r) := \#(x, y) : \exists \vec{Z} : B \wedge h(x, y) \tag{1}$$

$$supp(B) := \#(x, y) : \exists \vec{Z} : B \tag{2}$$

$$conf(r) := \frac{supp(r)}{supp(B)} \tag{3}$$

where $\#\gamma : \Gamma$ denotes the number of γ that fulfill the condition Γ, and $conf(r) \in [0, 1]$.

Example 1. *Consider the KG $\mathcal{G}^a$ in Fig. 1 and the rules r_1 : $hasChild(X, Y) \leftarrow worksAt(X, Z), educatedAt(Y, Z)$ and r_2: $hasSibling(X, Z) \leftarrow hasFather(X, Y), hasChild(Y, Z)$ mined from it. The body and rule supports of r_1 over the KG are $supp(B) = 8$ and $supp(r_1) = 2$ respectively. Hence, we have $conf(r_1) = \frac{2}{8}$. Analogously, $conf(r_2) = \frac{1}{6}$.* □

Support and confidence were originally developed for scoring rules over complete data. If data is missing, their interpretation is not straightforward and they can be misleading. In [Galarraga *et al.*, 2015], *confidence under the partial completeness assumption* (PCA) has been proposed as a measure, which guesses negative facts by assuming that data is usually added to KGs in batches, i.e., if at least one child of John is known then most probably all John's children are present in the KG. Formally, the *PCA confidence* is defined as

$$conf_{pca}(r) := \frac{supp(r)}{\#(x, y) : \exists \vec{Z} : B \wedge \exists y' : h(x, y') \in \mathcal{G}^a} \tag{4}$$

Example 2. *We obtain $conf_{pca}(r_1) = \frac{2}{4}$. Indeed, since carol and dave are not known to have any children in the KG, four existing body substitutions are not counted in the denominator. Meanwhile, we have $conf_{pca}(r_2) = \frac{1}{6}$, since all people that are predicted to have siblings by r_2 already have siblings in the available graph.* □

Given a rule r and a KG $\mathcal{G}$ the application of r on $\mathcal{G}$ results in a rule-based graph completion. More formally,

Definition 1 (Rule-based KG completion)**.** *Let $\mathcal{G}$ be a KG over the signature $\Sigma_{\mathcal{G}} = \langle \mathbf{R}, \mathcal{C} \rangle$ and let $r : h(X, Y) \leftarrow B$ be a rule mined from $\mathcal{G}$, i.e. a rule over $\Sigma_{\mathcal{G}}$. Then the* completion *of $\mathcal{G}$ is a graph $\mathcal{G}_r := \mathcal{G} \cup \{h(x, y) \mid \exists \vec{Z} : B\}$.*

Example 3. *We have $\mathcal{G}^a_{r_1} = \mathcal{G}^a \cup \{$ hasChild(john, dave), hasChild(carol, mary), hasChild(dave, dave), hasChild(carol, carol), hasChild(dave, bob), hasChild(mary, dave) $\}$.* □

Note that $\mathcal{G}^i$ is the perfect completion of $\mathcal{G}^a$, i.e., it is supposed to contain all correct facts with entities and relations from $\Sigma_{\mathcal{G}^a}$ that hold in the current state of the world. The goal of rule-based KG completion is to extract from $\mathcal{G}^a$ a set of rules $\mathcal{R}$ such that $\cup_{r \in \mathcal{R}} \mathcal{G}^a_r$ is as close to $\mathcal{G}^i$ as possible.

3 Completeness-aware Rule Scoring

Scoring and ranking rules are core steps in association rule learning. A variety of measures for ranking rules have been proposed, with prominent ones being confidence, conviction and lift. The existing (in-)completeness-aware rule measure in the KG context (the PCA confidence (4) [Galarraga *et al.*, 2015]) has two apparent shortcomings: First, it only counts as counterexamples those pairs (x, y) for which at least one $h(x, y')$ is in $\mathcal{G}^a$ for some y' and a rule's head predicate h. Hence, it may incorrectly give high scores to rules predicting facts for very incomplete relations, e.g., *place of baptism*. Second, it is not suited for data in non-functional relations that is not added in batches, such as awards, where the important ones are added instantly, while others much slower or never.

Thus, in this work we focus on the improvements of rule scoring functions by making use of the extra (in-)completeness meta-data. Before dwelling into the details of our approach we discuss the formal representation of such meta-data.

Cardinality Statements. We represent the (in)completeness meta-data using cardinality statements by reporting (the numerical restriction on) the absolute number of facts over a certain relation in the ideal graph $\mathcal{G}^i$. More specifically, we define the partial function num that takes as input a predicate p and a constant s and outputs a natural number corresponding to the number of facts in $\mathcal{G}^i$ over p with s as the first argument:

$$num(p, s) := \#o : p(s, o) \in \mathcal{G}^i \tag{5}$$

These cardinality statements can be obtained using web extraction techniques [Mirza *et al.*, 2017]. With such statements, it is also possible to encode cardinalities on the number of subjects for given predicates and objects, provided that inverse relations can be expressed in a KG.

Naturally, the number of missing facts for a given p and s can be obtained as

$$miss(p, s) := num(p, s) - \#o : p(s, o) \in \mathcal{G}^a \tag{6}$$

Example 4. *Consider the KG in Fig. 1. and the following cardinality statements for it:*

- $num(hasChild, john) = num(hasChild, mary) = 3$; $num(hasChild, alice) = 1$; $num(hasChild, carol) = num(hasChild, dave) = 0$;
- $num(hasSibling, bob) = 3$; $num(hasSibling, alice) = num(hasSibling, carol) = num(hasSibling, dave) = 2$.

We then have:

- $miss(hasChild, mary) = miss(hasChild, john) = miss(hasChild, alice) = 1$; $miss(hasChild, carol) = miss(hasChild, dave) = 0$;
- $miss(hasSibling, bob) = miss(hasSibling, carol) = 2$; $miss(hasSibling, alice) = miss(hasSibling, dave) = 1$.

□

We are now ready to define the *completeness-aware rule scoring problem*. Given a KG and a set of cardinality statements, *completeness-aware rule scoring* aims to score rules not only by their predictive power on the known KG, but also wrt. the number of wrongly predicted facts in complete areas and newly predicted facts in known incomplete areas.

In the following we discuss and compare two novel approaches for completeness-aware rule scoring. These are (i) the *completeness confidence*, and (ii) *directional metric*. Henceforth, all examples consider the KG in Fig. 1, rules from Ex. 1, and cardinality statements described in Ex. 4.

3.1 Completeness Confidence

In this work we propose to explicitly rely on incompleteness information in determining whether to consider an instance as a counterexample for a rule at hand or not.

To do that, we first define two indicators for a given rule $r : h(X, Y) \leftarrow \vec{B}$, reflecting the number of new predictions made by r in incomplete ($npi(r)$) and, respectively, complete ($npc(r)$) KG parts:

$$npi(r) := \sum_x min(\#y\colon h(x, y) \in \mathcal{G}^a_r \backslash \mathcal{G}^a, miss(h, x)) \tag{7}$$

$$npc(r) := \sum_x max(\#y\colon h(x, y) \in \mathcal{G}^a_r \backslash \mathcal{G}^a - miss(h, x), 0) \tag{8}$$

Note that summation is done exactly over those entities for which $miss$ is defined. Exploiting these additional indicators for $r : h(X, Y) \leftarrow \vec{B}$ we obtain the following *completeness-aware confidence*:

$$conf_{comp}(r) := \frac{supp(r)}{supp(\vec{B}) - npi(r)} \tag{9}$$

Example 5. *The rule r_2 matches more the real world than r_1 and, so, should be preferred over it. For our novel completeness confidence, we get $conf_{comp}(r_1) = \frac{2}{6}$ and $conf_{comp}(r_2) = \frac{1}{2}$, resulting in the desired rule ordering, not achieved by existing measures (see Ex. 1 and 2).* □

Under the closed world assumption $conf_{comp}(r) = conf(r)$ because $\forall x : miss(h, x) = 0$. Similarly, under the partial completeness assumption $conf_{comp}(r) = conf_{pca}(r)$. Thus, our completeness confidence is a more general measure than both the standard and the PCA confidence.

3.2 Directional Bias

If rule mining does make use of completeness information, and both do not exhibit any statistical bias, then intuitively the rule predictions and the (in)complete areas should be statistically independent. On the other hand, correlation between the two indicates that the rule-mining is *(in)completeness-aware*.

Example 6. *Suppose in total a given KG stores 1 million humans, and we know that 10,000 (1%) of these are missing some children (incompleteness information), while we also know that 1000 of the persons are definitely complete for children (0.1%). Let the set of rules mined from a KG predict 50,000 new facts for the $hasChild$ relation. Assuming independence between predictions and (in)completeness statements, we would expect 1% out of 50,000, i.e., 500 facts to be predicted in the incomplete areas and 0.1%, i.e., 50 in the complete KG parts. If instead we find 1000 children predicted for people that are missing correspondingly many children, and 10 for people that are not missing these, the former deviates from the expected value by a factor of 2, and the latter by a factor of 5.* □

Following the intuition from the above example, we propose to look at the extent of the non-independence of completeness information and rule mining to quantify the (in)completeness-awareness of rule mining. Let us consider predictions made by rules in a given KG, and denote by *enpi(r)* and *enpc(r)* the expected numbers of facts added by the rule r in the incomplete and complete areas respectively. Then for $\alpha \in [0, 1]$ being the weight given to completeness versus incompleteness, the directional coefficient of r is defined as follows:

$$direct\_coef(r) := \alpha \cdot \frac{enpc(r)}{npc(r)} + (1 - \alpha) \cdot \frac{npi(r)}{enpi(r)} \quad (10)$$

Unlike the other measures that range from 0 to 1, the directional coefficient takes values between 0 and infinity, where 1 is the default. The higher the *directional coefficient* is, the more "*completeness-aware*" the rules are.

In practice, expected values might be difficult to compute, and statistical independence is a strong assumption. An alternative that does not require knowledge about expected values is to directly measure the proportion between predictions in complete and incomplete parts. We call this the *directional metric*, which is computed as

$$direct\_metric(r) := \frac{npi(r) - npc(r)}{2 \cdot (npi(r) + npc(r))} + 0.5 \quad (11)$$

The metric is similar to the directional coefficient, but does not require knowledge about the expected number of predictions in complete/incomplete KG parts. It is designed to range between 0 and 1 again, thus allowing convenient weighting with other $[0, 1]$ measures. The directional metric of a rule that predicts the same number of facts in incomplete as in complete parts is 0.5, a rule that predicts twice as many facts in incomplete parts has a value of 0.66, and so on.

We propose to consider the combination of a weighted existing association rule measure, e.g., confidence or lift and the directional metric, with the weighting factor $\beta = 0..1$. Using confidence, we obtain

$$wdm(r) = \beta \cdot conf(r) + (1 - \beta) \cdot direct\_metric(r) \quad (12)$$

Example 7. *It holds that* $direct\_metric(r_1) \approx$ *0.33 and* $direct\_metric(r_2) = 0.8$. *Moreover, for confidence we get* $wdm(r_1) \approx 0.29$ *and* $wdm(r_2) \approx 0.48$ *with* $\beta = 0.5$. □

4 Evaluation

We have implemented our **c**ompleteness-**a**ware **r**ule **l**earning approach in a C++ system prototype CARL[1], following a standard relational learning algorithm implementation such as [Goethals and den Bussche, 2002].

We restrict the search space by mining rules of the form

$$r(X, Z) \leftarrow p(X, Y), q(Y, Z) \quad (13)$$

We aim to compare the predictive quality of the top k rules mined by our completeness-aware approach with the ones learned by standard rule learning methods: (1) AMIE [Galarraga *et al.*, 2015] (PCA confidence) and (2) WarmeR [Goethals and den Bussche, 2002] (standard confidence).

[1]The source code and all the data are available at https://github.com/Tpt/CARL.

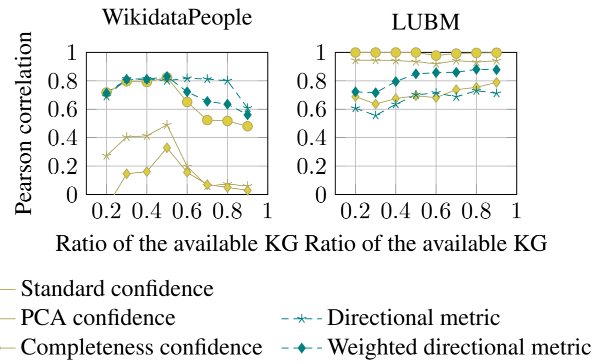

Figure 2: Evaluation results for WikidataPeople and LUBM datasets

Dataset. We used two datasets for the evaluation: (i) *WikidataPeople*, which is a dataset we have created from the Wikidata knowledge graph, containing 2.4M facts over 9 predicates about biographical information and family relationships of people; and (ii) *LUBM*, which is a synthetic dataset describing the structure of a university [Guo *et al.*, 2011].

For the WikidataPeople dataset, the approximation of the ideal KG ($\mathcal{G}^i$) is obtained by applying solid rules from the family domain. We acquired cardinality statements using data from Freebase [Pellissier Tanon *et al.*, 2016] and hand-crafted rules relying on the PCA [Galarraga *et al.*, 2015] assumption. This resulted in 10M cardinality statements.

LUBM $\mathcal{G}^i$, with 1.2M facts, was constructed by running the LUBM data generator. 464K cardinality statements were obtained by assuming the PCA on the whole dataset.

Experimental Setup. To assess the effect of our proposed measures, we first construct versions of the available KG ($\mathcal{G}^a$) by removing parts of the data from $\mathcal{G}^i$ and introducing a synthetic bias in the data (i.e., leaving many facts in $\mathcal{G}^a$ for some relations and few for others). The synthetic bias is needed to simulate our scenario of interest, where some parts of $\mathcal{G}^a$ are very incomplete, while others are fairly complete, which is indeed the case in real world KGs. In Wikidata, for instance, 2.7% of all humans have an assertion about their father, yet only 0.8% have an assertion about their mother.

We proceed in two steps: First, we define a *global ratio*, which determines a uniform percentage of data retained in the available graph. To further refine this, we then factor a *predicate ratio* individually for each predicate.

For a given predicate, the final ratio of facts in $\mathcal{G}^a$ retained from those in $\mathcal{G}^i$ is then computed as $min(1, 2 \cdot k \cdot n)$, where k is the predicate ratio and n is the global ratio.

The assessment of the rules learned from different versions of the available KG is performed by comparing rule predictions with the approximation of $\mathcal{G}^i$. More specifically, every learned rule is assigned a *quality score*, defined as:

$$quality\_score(r) = \frac{|\mathcal{G}^a_r \cap \mathcal{G}^i \setminus \mathcal{G}^a|}{|\mathcal{G}^a_r \setminus \mathcal{G}^a|} \quad (14)$$

This scoring naturally allows us to control the percentage of rule predictions that hit our approximation of $\mathcal{G}^i$, similar to standard precision estimation in machine learning.

Results. From every version of the available KG we have mined rules of the form (13) and kept only rules r with

$conf(r) \geq 0.001$ and $supp(r) \geq 10$, whose *head coverage* (ratio of the number of predicted facts that are in $\mathcal{G}^a$ over the number of facts matching the rule head) is greater than 0.001.

Evaluation results for WikidataPeople and LUBM datasets are in Figure 2. The horizontal axis displays the global ratio used for generating $\mathcal{G}^a$. The Pearson correlation factor (vertical axis) between each ranking measure and the rules quality score (14) is used to evaluate the measures' effectiveness.

For the WikidataPeople KG, directional metric, weighted directional metric and completeness confidence show the best results. For the LUBM KG, the completeness confidence outperforms the rest of the measures, followed by the standard confidence and the weighted directional metric.

5 Conclusion

We have defined the problem of learning rules from incomplete KGs enriched with the exact numbers of missing edges of certain types, and proposed several novel rule ranking measures that effectively make use of the meta-knowledge about complete and incomplete KG parts: *completeness confidence* and the (*weighted*) *directional metric*. Our measures have been injected in the rule learning prototype CARL and evaluated on real-world and synthetic KGs, demonstrating significant improvements both with respect to the quality of mined rules and predictions they produce.

References

[Auer *et al.*, 2007] Sören Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary G. Ives. DBpedia: A nucleus for a web of open data. In *ISWC*, pages 722–735, 2007.

[Bollacker *et al.*, 2007] Kurt D. Bollacker, Robert P. Cook, and Patrick Tufts. Freebase: A shared database of structured general human knowledge. In *AAAI*, pages 1962–1963, 2007.

[Carlson *et al.*, 2010] Andrew Carlson, Justin Betteridge, Bryan Kisiel, Burr Settles, Estevam R. Hruschka Jr., and Tom M. Mitchell. Toward an architecture for never-ending language learning. In *AAAI*, volume 5, page 3, 2010.

[d'Amato *et al.*, 2016] Claudia d'Amato, Steffen Staab, Andrea GB Tettamanzi, Tran Duc Minh, and Fabien Gandon. Ontology enrichment by discovering multi-relational association rules from ontological knowledge bases. In *SAC*, pages 333–338, 2016.

[Darari *et al.*, 2013] Fariz Darari, Werner Nutt, Giuseppe Pirrò, and Simon Razniewski. Completeness statements about RDF data sources and their use for query answering. In *ISWC*, pages 170–187, 2013.

[Darari *et al.*, 2016] Fariz Darari, Simon Razniewski, Radityo Eko Prasojo, and Werner Nutt. Enabling fine-grained rdf data completeness assessment. In *ICWE*, pages 170–187, 2016.

[Dehaspe and De Raedt, 1997] Luc Dehaspe and Luc De Raedt. Mining association rules in multiple relations. In *ILP*, pages 125–132, 1997.

[Doppa *et al.*, 2011] Janardhan Rao Doppa, Mohammad NasrEsfahani, Mohammad S Sorower, Jed Irvine, Thomas G Dietterich, Xiaoli Fern, and Prasad Tadepalli. Learning rules from incomplete examples via observation models. In *FAM-LbR/KRAQ*, 2011.

[Etzioni *et al.*, 1997] O. Etzioni, K. Golden, and D. S. Weld. Sound and efficient closed-world reasoning for planning. *AI*, 89(1-2):113–148, 1997.

[Gad-Elrab *et al.*, 2016] Mohamed H Gad-Elrab, Daria Stepanova, Jacopo Urbani, and Gerhard Weikum. Exception-enriched rule learning from knowledge graphs. In *ISWC*, pages 234–251, 2016.

[Galarraga *et al.*, 2015] Luis Galarraga, Christina Teflioudi, Katja Hose, and Fabian M. Suchanek. Fast rule mining in ontological knowledge bases with AMIE+. In *VLDB*, volume 24, pages 707–730, 2015.

[Galárraga *et al.*, 2017] Luis Galárraga, Simon Razniewski, Antoine Amarilli, and Fabian M Suchanek. Predicting completeness in knowledge bases. *WSDM*, pages 375–383, 2017.

[Goethals and den Bussche, 2002] Bart Goethals and Jan Van den Bussche. Relational association rules: Getting WARMeR. In *Pattern Det. and Disc. Workshop*, pages 125–139, 2002.

[Guo *et al.*, 2011] Yuanbo Guo, Zhengxiang Pan, and Jeff Heflin. LUBM: A benchmark for OWL knowledge base systems. *Web Semantics: Science, Services and Agents on the WWW*, 3(2-3), 2011.

[Lassila and Swick, 1999] Ora Lassila and Ralph R. Swick. Resource description framework (RDF) model and syntax specification. 1999.

[Levy, 1996] Alon Y. Levy. Obtaining complete answers from incomplete databases. In *VLDB*, pages 402–412, 1996.

[Lisi, 2010] Francesca A. Lisi. Inductive Logic Programming in Databases: From Datalog to DL+log. *TPLP*, 10(3):331–359, 2010.

[Mahdisoltani *et al.*, 2015] Farzaneh Mahdisoltani, Joanna Biega, and Fabian M. Suchanek. YAGO3: A knowledge base from multilingual wikipedias. In *CIDR*, 2015.

[Mirza *et al.*, 2017] Paramita Mirza, Simon Razniewski, Fariz Darari, and Gerhard Weikum. Cardinal virtues: Extracting relation cardinalities from text. *ACL*, page 347–351, 2017.

[Paulheim, 2017] Heiko Paulheim. Knowledge graph refinement: A survey of approaches and evaluation methods. *SW*, 8(3):489–508, 2017.

[Pellissier Tanon *et al.*, 2016] Thomas Pellissier Tanon, Denny Vrandečić, Sebastian Schaffert, Thomas Steiner, and Lydia Pintscher. From Freebase to Wikidata: The great migration. In *WWW*, pages 1419–1428, 2016.

[Pellissier Tanon *et al.*, 2017] Thomas Pellissier Tanon, Daria Stepanova, Simon Razniewski, Paramita Mirza, and Gerhard Weikum. Completeness-aware rule learning from knowledge graphs. In *ISWC*, pages 507–525, 2017.

[Prasojo *et al.*, 2016] Radityo Eko Prasojo, Fariz Darari, Simon Razniewski, and Werner Nutt. Managing and consuming completeness information for wikidata using COOL-WD. In *COLD@ISWC*, 2016.

[Razniewski *et al.*, 2016] Simon Razniewski, Fabian M Suchanek, and Werner Nutt. But what do we actually know. pages 40–44, 2016.

[Sazonau *et al.*, 2015] Viachaslau Sazonau, Uli Sattler, and Gavin Brown. General terminology induction in OWL. In *ISWC*, pages 533–550, 2015.

[Suchanek and Preda, 2014] Fabian M Suchanek and Nicoleta Preda. Semantic culturomics. *VLDB*, 7(12):1215–1218, 2014.

[Wang and Li, 2015] Zhichun Wang and Juan-Zi Li. RDF2Rules: Learning rules from RDF knowledge bases by mining frequent predicate cycles. *CoRR*, abs/1512.07734, 2015.