Rule Learning from Knowledge Graphs Guided by Embedding Models

Vinh Thinh Ho[1], Daria Stepanova[1], Mohamed Gad-Elrab[1], Evgeny Kharlamov[2], Gerhard Weikum[1]

hvthinh@mpi-inf.mpg.de, dstepano@mpi-inf.mpg.de, gadelrab@mpi-inf.mpg.de, evgeny.kharlamov@cs.ox.ac.uk, weikum@mpi-inf.mpg.de

[1]Max Planck Institute for Informatics, Saarbrücken, Germany
[2]University of Oxford, Oxford, United Kingdom

1. Motivation

- **Knowledge Graphs**: huge collections of positive unary and binary facts treated under **O**pen **W**orld **A**ssumption (e.g. $isMarriedTo(alice, bob), researcher(mat)$)
- Automatically constructed, thus **incomplete** $\Rightarrow$ **Knowledge Graph Completion**

Rule-based approach	Embedding-based approach
$livesIn(X,Z) \leftarrow spouse(X,Y), livesIn(Y,Z)$	$score(<Jack, livesIn, Chicago>) = 0.8$
+ Interpretable	- Hard to interpret
+ Limited training data	- A lot of training data
- Local patterns	+ Global patterns
- Not extendable	+ Extendable (e.g., text)

Our approach:
Mine rules with support from KG embeddings

- **Challenges:** **OWA**, **Huge size** of KGs, **how to combine** the 2 approaches
- **Contributions:**
  - Framework for rule learning guided by external sources
  - Concrete framework instantiation utilizing feedback from embedding models
  - Interactive non-monotonic rules mining mechanism
  - Experiments on real-world Knowledge Graphs

2. Rule Learning Guided by External Sources

Given:
- **Knowledge Graph** $\mathcal{G}$
- **External Source**
  - Probability function f reflecting trustfulness of $\mathcal{G}$'s missing facts

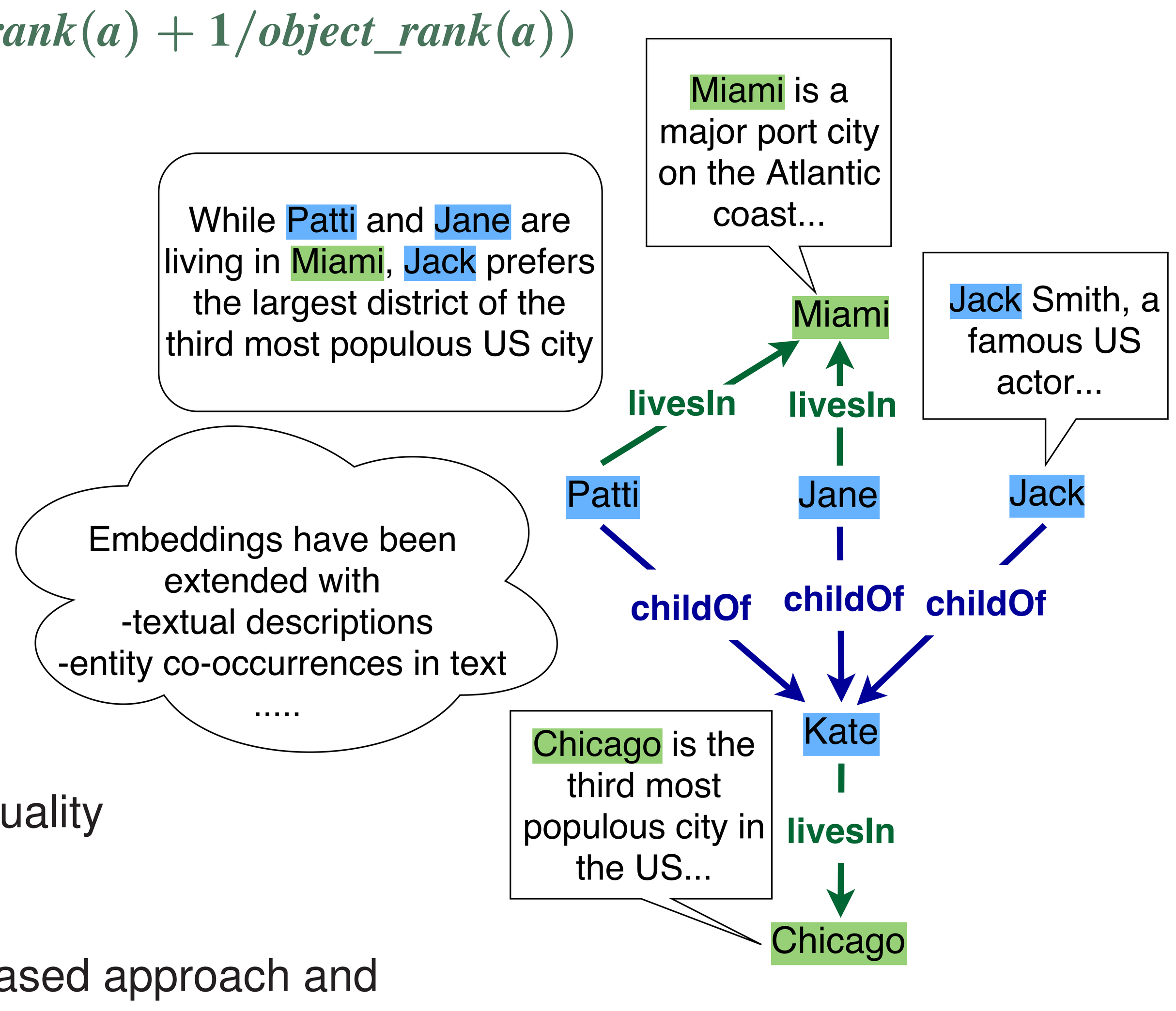
Find:
- **A ordered rule list** extracted from $\mathcal{G}$ with respect to the support from the quality function f

Our solution: Introduce the novel **Hybrid Rule Quality Function** μ as follows:

- Measure the **descriptive quality** μ_1 of rule r over $\mathcal{G}$:
$$\mu_1 : (r, \mathcal{G}) \mapsto \alpha \in [0,1]$$
$\Rightarrow$ any **classical quality measure** which describes $\mathcal{G}$ properly
- Measure the **predictive quality** μ_2 of $\mathcal{G}_r$ relying on $\mathcal{P} = (\mathcal{G}, f)$:
$$\mu_2 : (\mathcal{G}_r, \mathcal{P}) \mapsto \alpha \in [0,1]$$
where $\mathcal{G}_r$ is a extension of $\mathcal{G}$ by applying r
$\Rightarrow$ capture **information about missing facts** in $\mathcal{G}$ that are relevant for r
- Combine the result as **the weighted sum**:
$$\mu(r, \mathcal{P}) = (1-\lambda) \times \mu_1(r, \mathcal{G}) + \lambda \times \mu_2(\mathcal{G}_r, \mathcal{P})$$
weighting factor $\lambda \in [0,1]$ control the distribution of μ_1 and μ_2

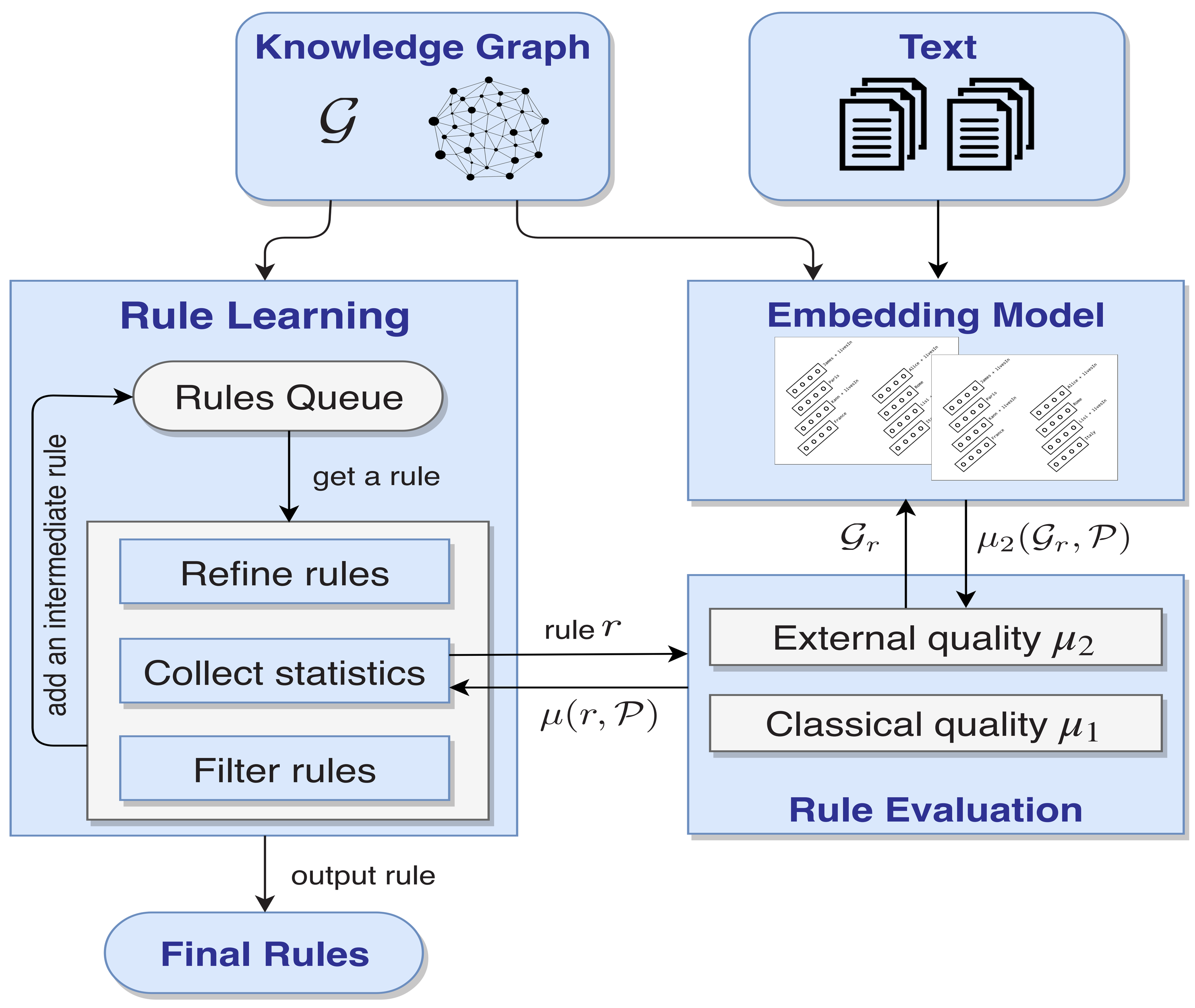
3. Embedding-based Rule Evaluation

- **Realization of** f **using embeddings [with additional textual data]:**
$$f(a) = 0.5 \times (1/subject\_rank(a) + 1/object\_rank(a))$$
- **Realization of** μ_1:
  - Standard Confidence ($conf$)
  - PCA Confidence ($conf_{pca}$)
  - Soft Confidence ($conf_{soft}$)
  - ...
- **Realization of** μ_2:
$$\mu_2(\mathcal{G}_r, \mathcal{P}) = \frac{\Sigma_{a \in \mathcal{G}_r \setminus \mathcal{G}} f(a)}{|\mathcal{G}_r \setminus \mathcal{G}|}$$
$\Rightarrow$ **average** embedding-based quality of **predicted facts** in $\mathcal{G}_r$



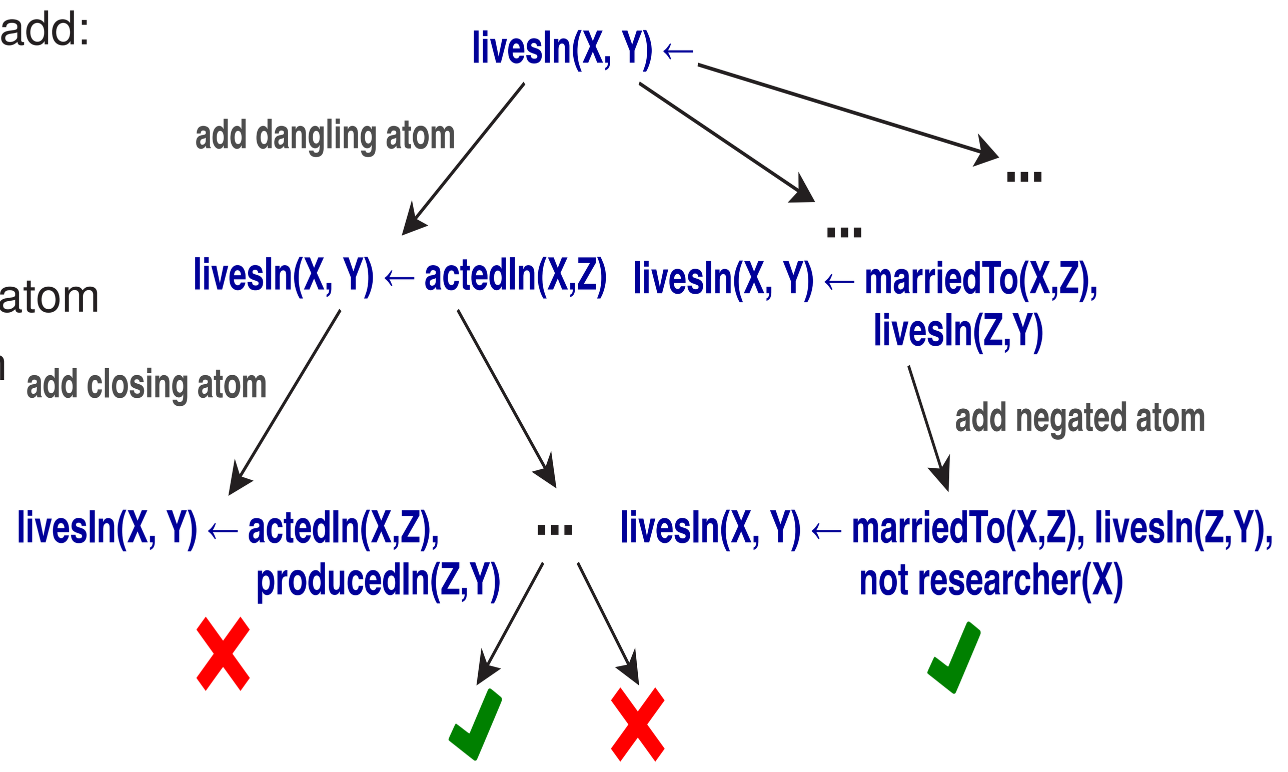
- Preserve the **accuracy** of rule-based approach and Enhance the **predicting power** of embedding-based approach

4. Non-monotonic Rule Mining with Embedding Support



5. Rule Refinement and Filters

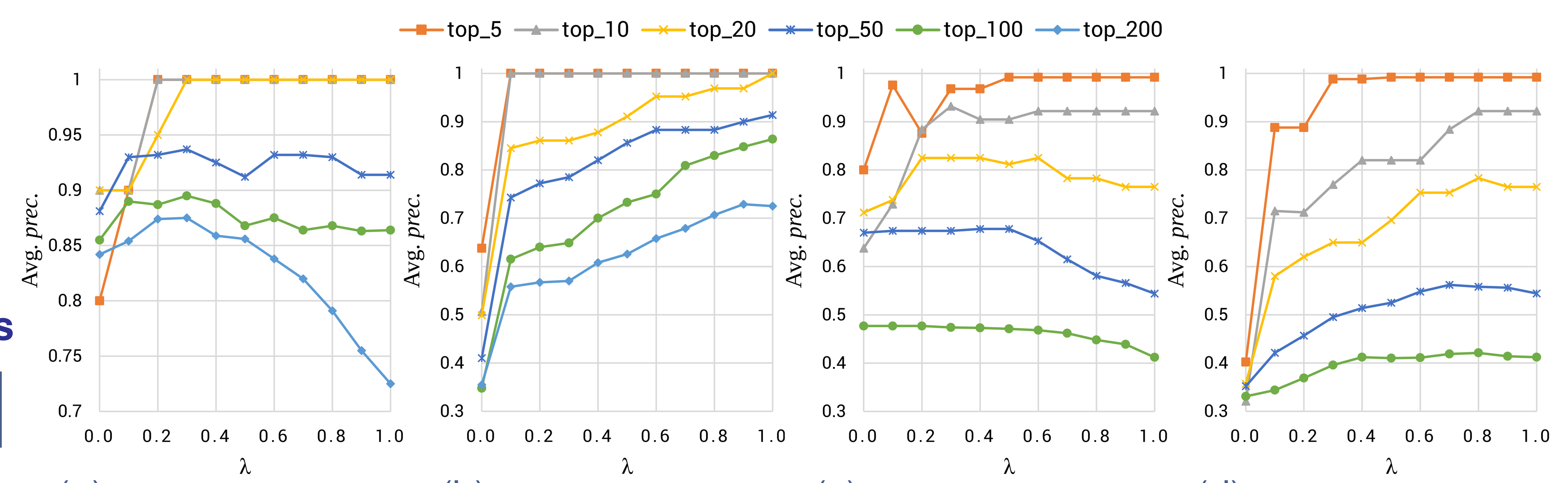
- **Refinement Operators** add:
  - dangling atom
  - instantiated atom
  - closing atom
  - **negated** instantiated atom
  - **negated** closing atom
- **Filtering rules** on:
  - rule form
  - support ($supp$)
  - head coverage (hc)
  - confidence ($conf$)
  - **the hybrid measure** (μ) guides the dynamic rules construction
  - **exception confidence**:



$$e\text{-}conf(r, \mathcal{G}) = conf(r', \mathcal{G})$$
where $r' : body^-(r) \leftarrow body^+(r), not\ head(r)$ $\Rightarrow$ ensure the **quality of exceptions**

6. Experiments

- **Data sets**: FB15K, Wiki44K
- **Approximated ideal KG**: original
- **Available KG**: random 80% of original KG, preserving the distribution of facts over predicates.
- **Embedding Models**: 2 settings
  - Without additonal textual data: TransE model, HolE model
  - With entities description text: SSP model
- **2 Evaluation settings**: *closed world* setting (CW) and *open world* setting (OW)



(a) Conf-SSP on FB15K (b) PCA-SSP on FB15K (c) Conf-SSP on Wiki44K (d) PCA-SSP on Wiki44K

Evaluation result on *closed world* setting (CW)

- **Examples of mined rules:**

r_1: $nationality(X,Y) \leftarrow graduated\_from(X,Z), in\_country(Z,Y), not\ research\_uni(Z)$
r_2: $scriptwriter\_of(X,Y) \leftarrow preceded\_by(X,Z), scriptwriter\_of(Z,Y), not\ tv\_series(Z)$

References

- V. T. Ho, D. Stepanova, M. Gad-Elrab, E. Kharlamov, G. Weikum. Rule Learning from Knowledge Graphs Guided by Embedding Models. In proc. *International Semantic Web Conference*, 2018.
- D. H. Tran., D. Stepanova, M. Gad-Elrab, Francesca A. Lisi, G. Weikum. Towards Nonmonotonic Rule Learning from Knowledge Graphs. *Inductive Logic Programming*, 2016.
- L. Galárraga, C. Teflioudi, K. Hose, F. M. Suchanek. Fast Rule Mining in Ontological Knowledge Bases with AMIE+. *VLBD journal*, 2015.