Towards Nonmonotonic Relational Learning from Knowledge Graphs

Hai Dang Tran
Advisor: Daria Stepanova
Supervisor: Gerhard Weikum

Max Planck Institute for Informatics
Saarland University, Saarbrücken, Germany

Motivation

- Knowledge Graphs (KGs): a large set of facts ⟨*subject predicate object*⟩
  ⟨*Brad isMarriedTo Ann*⟩, ⟨*Dave type Researcher*⟩
- Represent unary/binary facts with Open World Assumption (OWA)
  $isMarriedTo(Brad, Ann)$, $Researcher(Dave)$
- KGs are normally incomplete and contain inaccurate facts

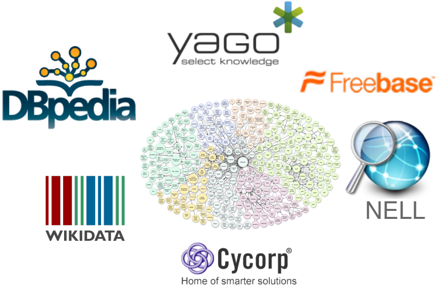
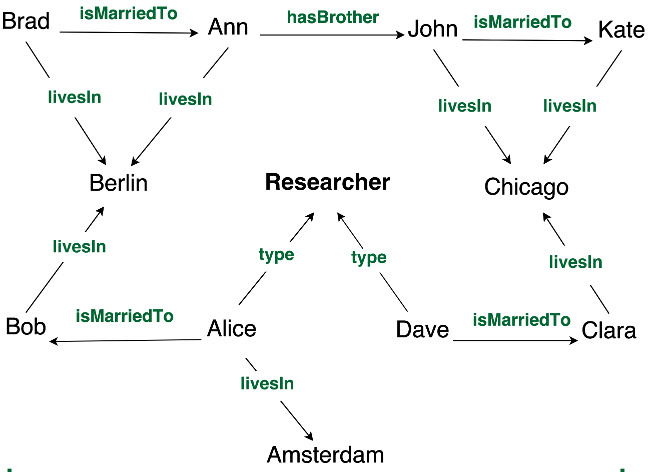


Figure: Typical Knowledge Graphs

Motivation

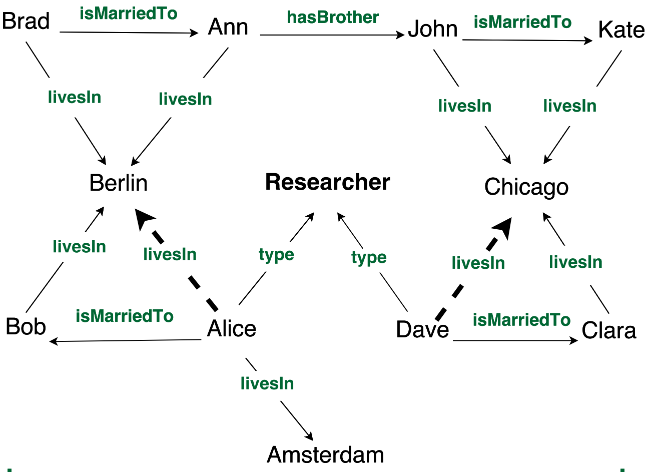
Complete KGs with positive (Horn) rule mining [Galárraga *et al.*, 2015]



$$r: \; livesIn(X, Z) \leftarrow isMarriedTo(X, Y), livesIn(Y, Z)$$

Motivation

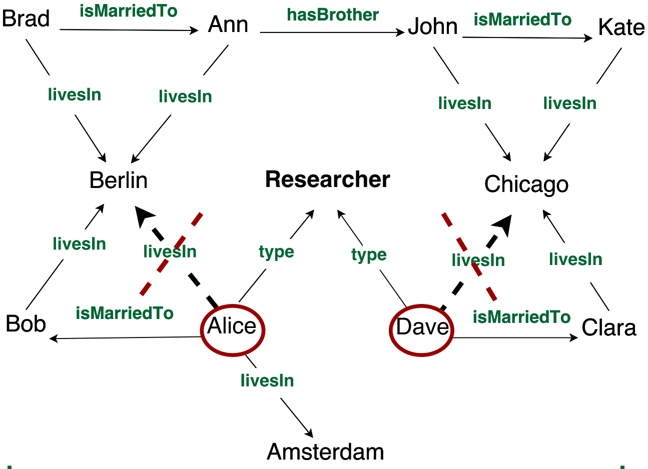
Complete KGs with positive (Horn) rule mining [Galárraga *et al.*, 2015]



$$r: \ livesIn(X, Z) \leftarrow isMarriedTo(X, Y), livesIn(Y, Z)$$

Motivation

In this thesis: nonmonotonic rule mining under OWA



$$r: \ livesIn(X,Z) \leftarrow isMarriedTo(X,Y), livesIn(Y,Z), not \ \ researcher(X)$$

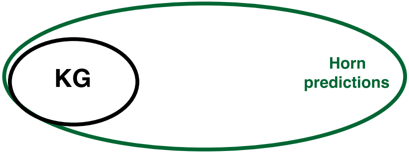
Outline

Problem Statement

Quality-based Horn Theory Revision (QHTR)

Given:

- KG
- Set of positive rules

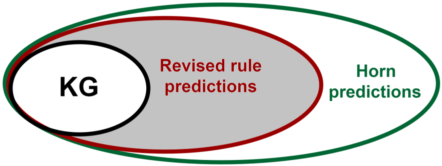


Problem Statement

Quality-based Horn Theory Revision (QHTR)

Given:

- KG
- Set of positive rules



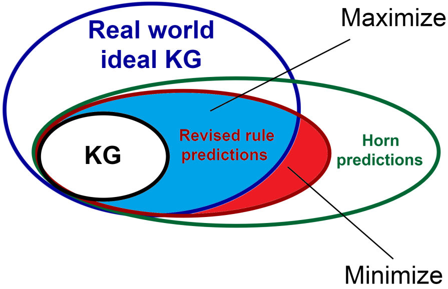
Find: Revision of Horn rules with negation to improve quality

Problem Statement

Quality-based Horn Theory Revision (QHTR)

Given:

- KG
- Set of positive rules



Find: Revision of Horn rules with negation to improve quality

Conflicting Predictions

Keep the number of conflicts as small as possible

$$\left\{\begin{array}{l} r1: livesIn(X,Z) \leftarrow isMarriedTo(X,Y), livesIn(Y,Z), not\ res(X) \\ r1^{aux}: not\_livesIn(X,Z) \leftarrow isMarriedTo(X,Y), livesIn(Y,Z), res(X) \\ r2: livesIn(X,Z) \leftarrow workIn(X,Z), not\ artist(X) \\ r2^{aux}: not\_livesIn(X,Z) \leftarrow workIn(X,Z), artist(X) \end{array}\right\}$$

possibly generate conflicting predictions $\{livesIn(a,b), not\_livesIn(a,b)\}$

Idea behind: researcher is a good exception for *r1*, but the occurrence of *r2* might break this; conflicts should be minimized

Problem Statement

Quality-based Horn Theory Revision (QHTR)

Given:

- KG
- Set of positive rules

Find: Revision of positive rules such that

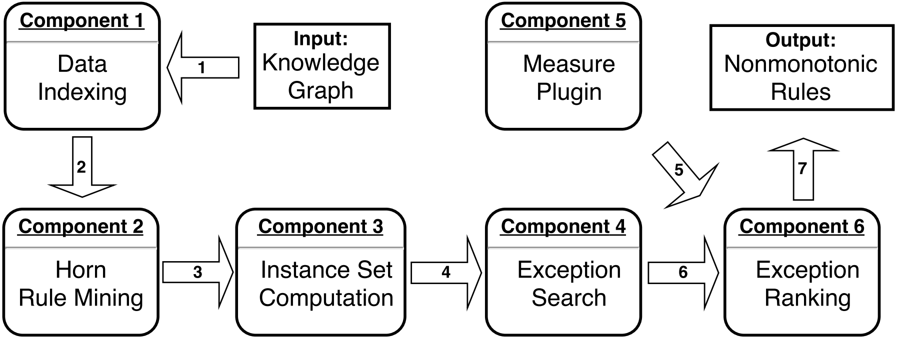
- number of conflicting predictions generated by the revision is as **small** as possible
- average conviction ($\frac{1 - support}{1 - confidence}$) is as **large** as possible [Azevedo and Jorge, 2007]

Related Work

- **Statistics-based Approaches**
  - Hidden features [Nickel *et al.*, 2011]
- **Text-based Approaches**
  - External methods [Ritze *et al.*, 2015]
- **Logic-based Approaches**
  - Inductive Logic Programming (ILP) systems: ALEPH, QuickFOIL [Zeng *et al.*, 2014]
  - Relational learning systems: AMIE(+) [Galárraga *et al.*, 2015]
  - Nonmonotonic rule mining system [Gad-Elrab *et al.*, 2016]

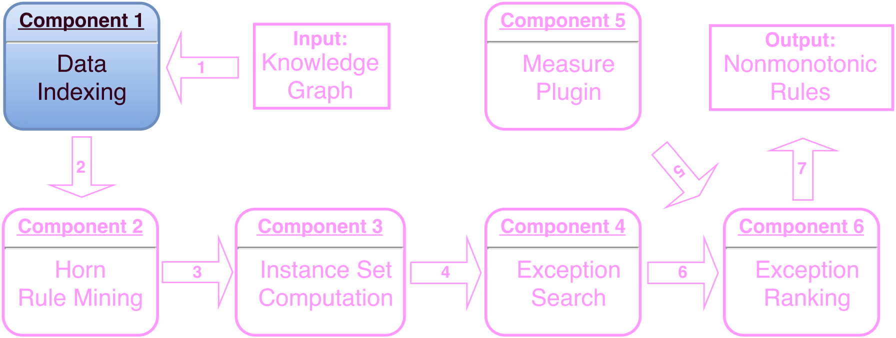
RUMIS Architecture

Nonmonotonic **RU**le **MI**ning **S**ystem in the relational setting



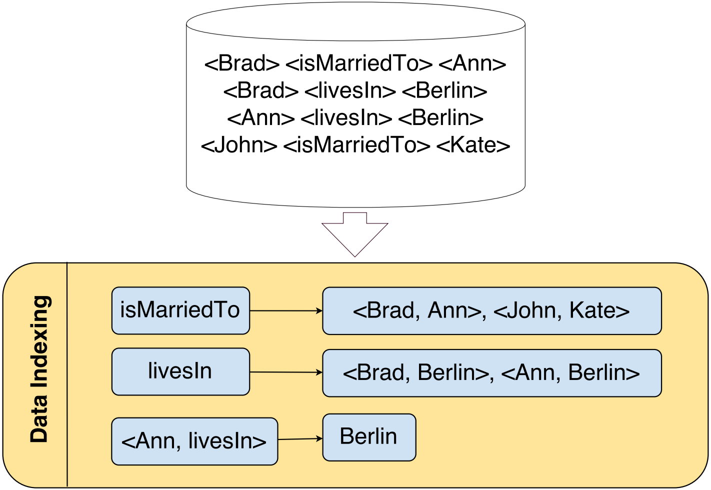
Data Indexing

Data Indexing Step



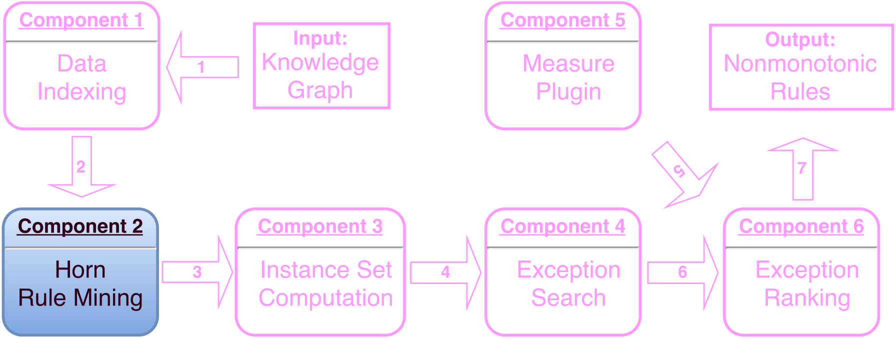
Data Indexing

Speed up searching over a dataset



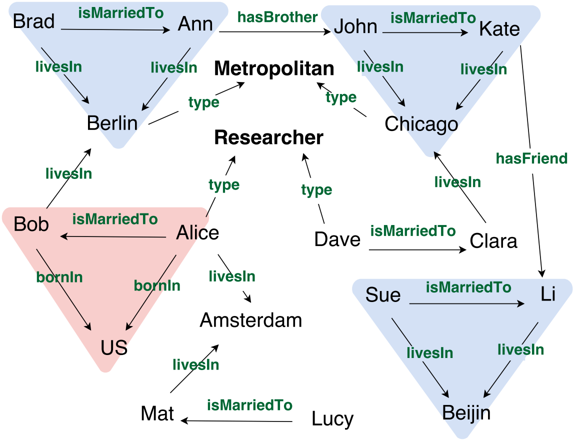
Horn Rule Mining

Horn Rule Mining Step



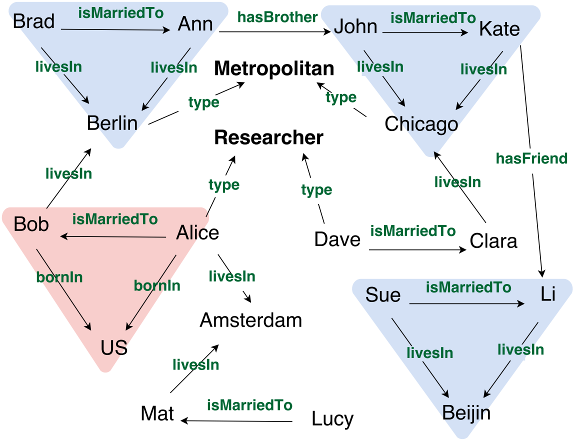
Horn Rule Mining

livesIn(X,Z), isMarriedTo(X,Y),livesIn(Y,Z): 3 substitutions
bornIn(X,Z), isMarriedTo(X,Y),bornIn(Y,Z): 1 substitution



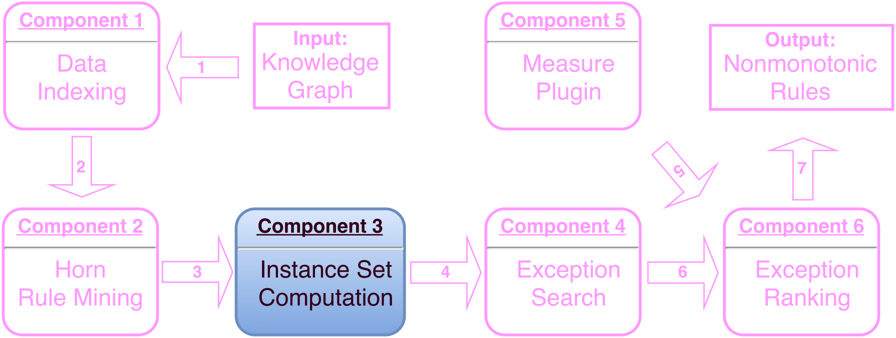
Horn Rule Mining

$$livesIn(X,Z) \leftarrow isMarriedTo(X,Y), livesIn(Y,Z)$$
$$bornIn(X,Z) \leftarrow isMarriedTo(X,Y), bornIn(Y,Z)$$



Instance Set Computation

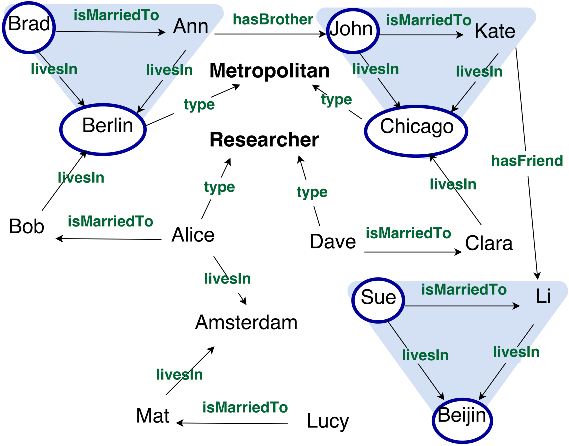
Instance Set Computation Step



Instance Set Computation

livesIn(X,Z) $\leftarrow$ *isMarriedTo(X,Y),livesIn(Y,Z)*

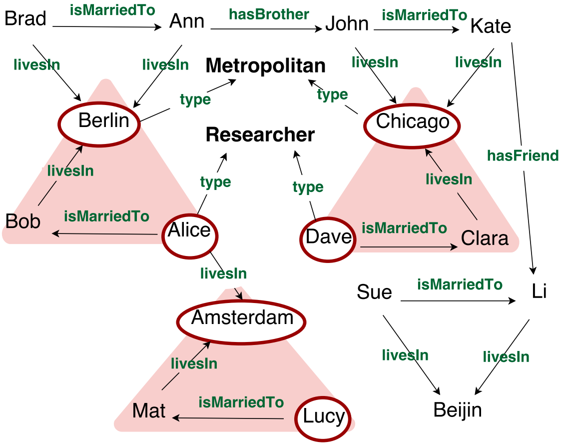
Normal set for $\langle$X, Z$\rangle$: $\{\langle$*Brad, Berlin*$\rangle$, $\langle$*John, Chicago*$\rangle$, $\langle$*Sue, Beijin*$\rangle\}$



Instance Set Computation

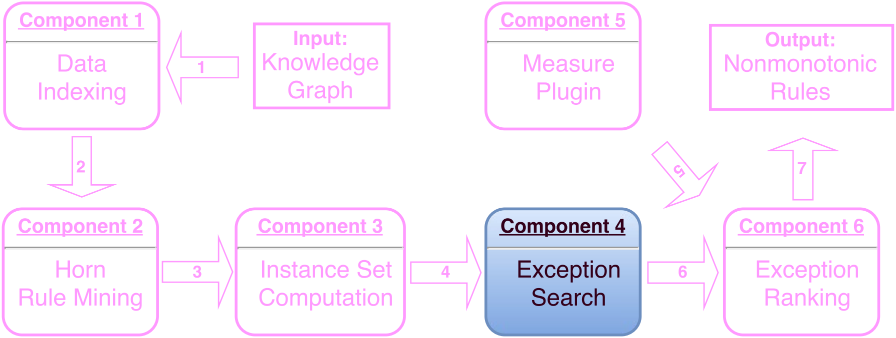
livesIn(X,Z) $\leftarrow$ *isMarriedTo(X,Y),livesIn(Y,Z)*

Abnormal set for $\langle$X, Z$\rangle$: $\{\langle$*Alice, Berlin*$\rangle$, $\langle$*Dave, Chicago*$\rangle$, $\langle$*Lucy, Ams*$\rangle\}$

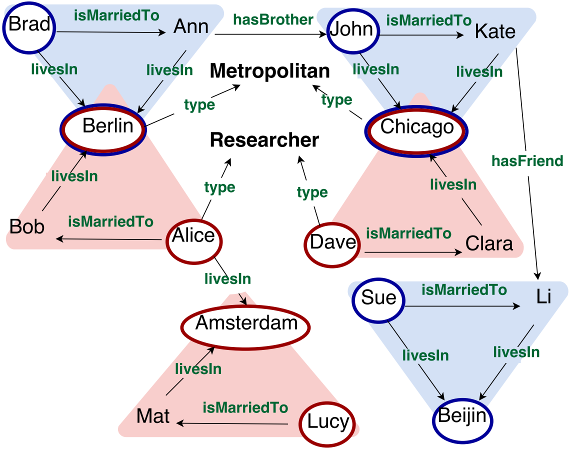


Exception Search

Exception Search Step



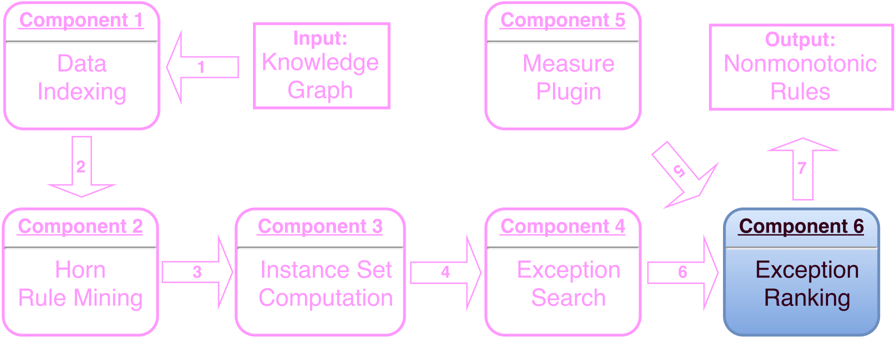
Exception Search



$$r: \ livesIn(X, Z) \leftarrow isMarriedTo(X, Y), livesIn(Y, Z), not \ \ researcher(X)$$

Exception Ranking

Exception Ranking Step



Exception Ranking

$$r1 \ldots\ldots\ldots \{\underline{\mathbf{e_1}}, e_2, e_3, \ldots\}$$
$$r2 \ldots\ldots\ldots \{e_1, \underline{\mathbf{e_2}}, e_3, \ldots\}$$
$$r3 \ldots\ldots\ldots \{\underline{\mathbf{e_1}}, e_2, e_3, \ldots\}$$

We propose a local optimum solution for the problem statement

- Naive ranking: among all revisions of $r \in \mathcal{R}_H$, r' having the largest $conv(r, \mathcal{G})$ is chosen
- Partial Materialization (PM): for each rule r, KG $\mathcal{G}'$ includes $\mathcal{G}$ and predicted facts of other rules. PM selects a revision having the largest $\dfrac{conv(r, \mathcal{G}') + conv(r^{aux}, \mathcal{G}')}{2}$
- Ordered Partial Materialization (OPM): similar to PM, but only previous rules of r are used to predict new facts

Experiments

- Two datasets for ideal KG:
  - IMDB [1]: 583 thousand triples, 112 thousand entities, 39 predicates
  - YAGO [Suchanek *et al.*, 2007]: 20.7 million triples, 1.8 million entities, 77 predicates
- Learning KG: delete 20% facts over each binary predicate from each ideal KG
- Horn rules with the form $h(X, Z) \leftarrow p(X, Y), q(Y, Z)$ are mined from the learning KG
- Three kinds of exception: $e_1(X), e_2(Z), e_3(X, Z)$
- Naive, PM, OPM ranking, dlv[2] is used to extend KGs

[1] `http://imdb.com`

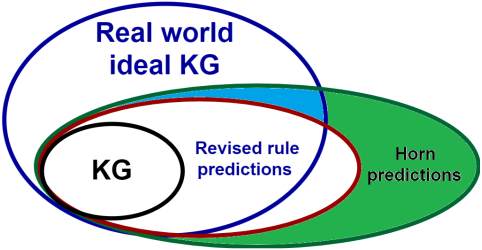
[2] `http://dlvsystem.com`

Experiments

predicate	predictions				outside ideal KG				corr. removed %		
	$\mathcal{R}_H$	$\mathcal{R}_N$	$\mathcal{R}_{PM}$	$\mathcal{R}_{OPM}$	$\mathcal{R}_H$	$\mathcal{R}_N$	$\mathcal{R}_{PM}$	$\mathcal{R}_{OPM}$	$\mathcal{R}_N$	$\mathcal{R}_{PM}$	$\mathcal{R}_{OPM}$
I:*actedIn*	1231	1214	1230	1214	1148	1131	1147	1131	90	100	90
I:*genre*	629	609	618	609	493	477	482	477	50	20	50
I:*hasLang*	173	102	125	102	163	92	115	92	60	100	60
I:*prodIn*	2489	2256	2327	2327	2488	2255	2326	2326	10	10	30
									52.50	45.16	57.75
Y:*direct*	41079	39174	39174	39174	41021	39116	39116	39116	100	100	100
Y:*grFrom*	3519	3456	3456	3456	3363	3300	3300	3300	100	100	70
Y:*citizOf*	3407	2883	2883	2883	3360	2836	2836	2836	50	50	70
Y:*bornIn*	110283	108317	109846	108317	109572	107607	109137	107607	90	90	100
									85	**85**	**85**

Table: New Facts Predicted by the Rulesets for IMDB (*I*) and YAGO (*Y*)

- Positive Horn rules ($\mathcal{R}_H$)
- Revisions ($\mathcal{R}_N, \mathcal{R}_{PM}, \mathcal{R}_{OPM}$)
- Blue area is nearly empty
- Last column for green part
- Sample 10 facts for each head relation



Experiments

predicate	predictions				outside ideal KG				corr. removed %		
	$\mathcal{R}_H$	$\mathcal{R}_N$	$\mathcal{R}_{PM}$	$\mathcal{R}_{OPM}$	$\mathcal{R}_H$	$\mathcal{R}_N$	$\mathcal{R}_{PM}$	$\mathcal{R}_{OPM}$	$\mathcal{R}_N$	$\mathcal{R}_{PM}$	$\mathcal{R}_{OPM}$
I:actedIn	1231	1214	1230	1214	1148	1131	1147	1131	90	100	90
I:genre	629	609	618	609	493	477	482	477	50	20	50
I:hasLang	173	102	125	102	163	92	115	92	60	100	60
I:prodIn	2489	2256	2327	2327	2488	2255	2326	2326	10	10	30
									52.50	45.16	57.75
Y:direct	41079	39174	39174	39174	41021	39116	39116	39116	100	100	100
Y:grFrom	3519	3456	3456	3456	3363	3300	3300	3300	100	100	70
Y:citizOf	3407	2883	2883	2883	3360	2836	2836	2836	50	50	70
Y:bornIn	110283	108317	109846	108317	109572	107607	109137	107607	90	90	100
									85	**85**	**85**

Table: New Facts Predicted by the Rulesets for IMDB (*I*) and YAGO (*Y*)

r_1 : $\mathit{writtenBy}(X,Z) \leftarrow \mathit{hasPredecessor}(X,Y), \mathit{writtenBy}(Y,Z),$ **not** $\mathit{americanFilm}(X)$

r_2 : $\mathit{actedIn}(X,Z) \leftarrow \mathit{isMarriedTo}(X,Y), \mathit{directed}(Y,Z),$ **not** $\mathit{silentFilmActor}(X)$

r_3 : $\mathit{isPoliticianOf}(X,Z) \leftarrow \mathit{hasChild}(X,Y), \mathit{isPoliticianOf}(Y,Z),$ **not** $\mathit{vicepresidentOfMexico}(X)$

Figure: Examples of the Revised Rules

Summary

Contributions:

- Theory revision framework for KG completion problem
- Partial materialization for rule revision
- RUMIS implementation in Java
- Experiments with IMDB and YAGO

Further Work:

- Support more kinds of exceptions
- Support more Horn rule forms
- Optimize data indexing

Appendix

Step: Horn Rule Mining [Tran *et al.*, 2016]

Given: KG $\mathcal{G}$

Goal: Find Horn rules ***h(X, Z)*** $\leftarrow$ ***p(X, Y), q(Y, Z)***

- Simplify *language bias* challenge
- $asupp(r) = |\{(X/a, Y/b, Z/c) : h(a, c), p(a, b), q(b, c)\}|$
- Top rules with the highest *asupp* are computed using Data Indexer
- Another tool can be used in this step

Appendix

Step: Instance Set Computation

Given: KG $\mathcal{G}$, Horn rules ***h(X, Z)*** $\leftarrow$ ***p(X, Y), q(Y, Z)***

Goal: Find (ab)normal sets exploiting Data Indexer

- Traverse all substitutions for X
- Find list of Y from pX based on indexing
- Find list of Z from qY based on indexing
- Check $\langle X\ h\ Z \rangle$ in the graph or not
- Classify $\langle X\ Z \rangle$ into $NS(r)$ and $ABS(r)$

Appendix

topk	YAGO				IMDB			
	$\mathcal{R}_H$	$\mathcal{R}_N$	$\mathcal{R}_{PM}$	$\mathcal{R}_{OPM}$	$\mathcal{R}_H$	$\mathcal{R}_N$	$\mathcal{R}_{PM}$	$\mathcal{R}_{OPM}$
5	1.3784	1.3821	1.3821	1.3821	2.2670	2.3014	2.3008	2.3014
30	1.1207	1.1253	1.1236	1.1237	1.5453	1.5644	1.5543	1.5640
50	1.0884	1.0923	1.0909	1.0913	1.3571	1.3749	1.3666	1.3746
60	1.0797	1.0837	1.0823	1.0829	1.3063	1.3221	1.3143	1.3219
70	1.0714	1.0755	1.0736	1.0744	1.2675	1.2817	1.2746	1.2814
80	1.0685	1.0731	1.0710	1.0720	1.2368	1.2499	1.2431	1.2497
100	1.0618	1.0668	1.0648	1.0659	1.3074	1.4100	1.3987	1.4098

Table: The Average Quality of the Top Positive and Nonmonotonic Rules

- Statistics for top Horn rules ($\mathcal{R}_H$) and their revisions ($\mathcal{R}_N, \mathcal{R}_{PM}, \mathcal{R}_{OPM}$) in each row
- Up trend for quality ratio between revision and Horn rules

Appendix

$$hasGenre(X, Z) \leftarrow hasPredecessor(X, Y), hasGenre(Y, Z)$$
$$actedIn(X, Z) \leftarrow actedIn(X, Y), hasSuccessor(Y, Z)$$
$$producedIn(X, Z) \leftarrow directed(X, Y), producedIn(Y, Z)$$
$$hasLanguage(X, Z) \leftarrow directedBy(X, Y), hasLanguage(Y, Z)$$

Figure: Selected Horn rules for IMDB

$$wasBornIn(X, Z) \leftarrow isCitizenOf(X, Y), hasCapital(Y, Z)$$
$$wasBornIn(X, Z) \leftarrow isPoliticianOf(X, Y), hasCapital(Y, Z)$$
$$graduatedFrom(X, Z) \leftarrow hasAcademicAdvisor(X, Y), worksAt(Y, Z)$$
$$directed(X, Z) \leftarrow isMarriedTo(X, Y), actedIn(Y, Z)$$
$$isCitizenOf(X, Z) \leftarrow hasChild(X, Y), isPoliticianOf(Y, Z)$$

Figure: Selected Horn rules for YAGO

References I

Paulo J. Azevedo and Alípio Mário Jorge.
Comparing Rule Measures for Predictive Association Rules.
In *Proceedings of ECML*, pages 510–517, 2007.

Mohamed H. Gad-Elrab, Daria Stepanova, Jacopo Urbani, and Gerhard Weikum.
Exception-Enriched Rule Learning from Knowledge Graphs, pages 234–251.
Springer International Publishing, Cham, 2016.

Luis Galárraga, Christina Teflioudi, Katja Hose, and Fabian M. Suchanek.
Fast Rule Mining in Ontological Knowledge Bases with AMIE+.
In *VLDB Journal*, 2015.

Maximilian Nickel, Volker Tresp, and Hans-Peter Kriegel.
A three-way model for collective learning on multi-relational data.
In Lise Getoor and Tobias Scheffer, editors, *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, ICML '11, pages 809–816, New York, NY, USA, June 2011. ACM.

Dominique Ritze, Oliver Lehmberg, and Christian Bizer.
Matching html tables to dbpedia.
In *Proceedings of the 5th International Conference on Web Intelligence, Mining and Semantics*, WIMS '15, pages 10:1–10:6, New York, NY, USA, 2015. ACM.

Fabian M. Suchanek, Gjergji Kasneci, and Gerhard Weikum.
Yago: A core of semantic knowledge.
In *Proceedings of the 16th International Conference on World Wide Web*, WWW '07, pages 697–706, New York, NY, USA, 2007. ACM.

Dang Hai Tran, Daria Stepanova, Francesca A. Lisi Mohamed Gad Elrab, and Gerhard Weikum.
Towards nonmonotonic relational learning from knowledge graphs.
In *Proc. of the International Conference on Inductive Logic Programming.*, 2016.

Qiang Zeng, Jignesh M. Patel, and David Page.
QuickFOIL: Scalable Inductive Logic Programming.
PVLDB, 8(3):197–208, 2014.