



Scalable Rule Learning in Probabilistic Knowledge Bases

Arcchit Jain	KU Leuven
Tal Friedman	UCLA
Ondřej Kuželka	KU Leuven
Guy Van den Broeck	UCLA
Luc De Raedt	KU Leuven





Takeaway

- ▶ Probabilistic rule learning has become much more SCALABLE and FASTER than before!

Motivation

KB Completion is like a pothole filling job on a large road network!



SafeLearner - A safe rule learner

researcher	paper	p
bob	plp	0.9
carl	plp	0.6
greg	plp	0.7
ian	db	0.9
harry	db	0.8

author/2

researcher	university	p
edwin	harvard	1.0
fred	harvard	0.9
alice	mit	0.6
dave	mit	0.7

location/2

researcher	researcher	p
alice	edwin	0.2
alice	fred	0.3
bob	carl	0.4
bob	greg	0.5
bob	harry	0.6
bob	ian	0.7
carl	greg	0.8
carl	harry	0.9
carl	ian	0.8
dave	edwin	0.7
dave	fred	0.6
edwin	fred	0.5
greg	harry	0.4
greg	ian	0.3
ian	ian	0.2

target: coauthor/2

AMIE+ Rules:

- ▶ $\text{coauthor}(A, B) :- \text{author}(A, C), \text{author}(B, C).$
- ▶ $\text{coauthor}(A, B) :- \text{location}(A, C), \text{location}(B, C).$

Query:

$\exists a, b \text{ s.t. } p_{h_1} \vee (p_{h_2} \wedge \text{location}(a, c) \wedge \text{location}(b, c)) \vee (p_{h_3} \wedge \text{author}(a, d) \wedge \text{author}(b, d))$

Learned Rules:

- ▶ $0.105::\text{coauthor}(A, B) :- \text{true}.$
- ▶ $0.687::\text{coauthor}(A, B) :- \text{location}(A, C), \text{location}(B, C).$
- ▶ $0.333::\text{coauthor}(A, B) :- \text{author}(A, C), \text{author}(B, C).$



What makes *SafeLearner* different?

01

Lifted Inference

02

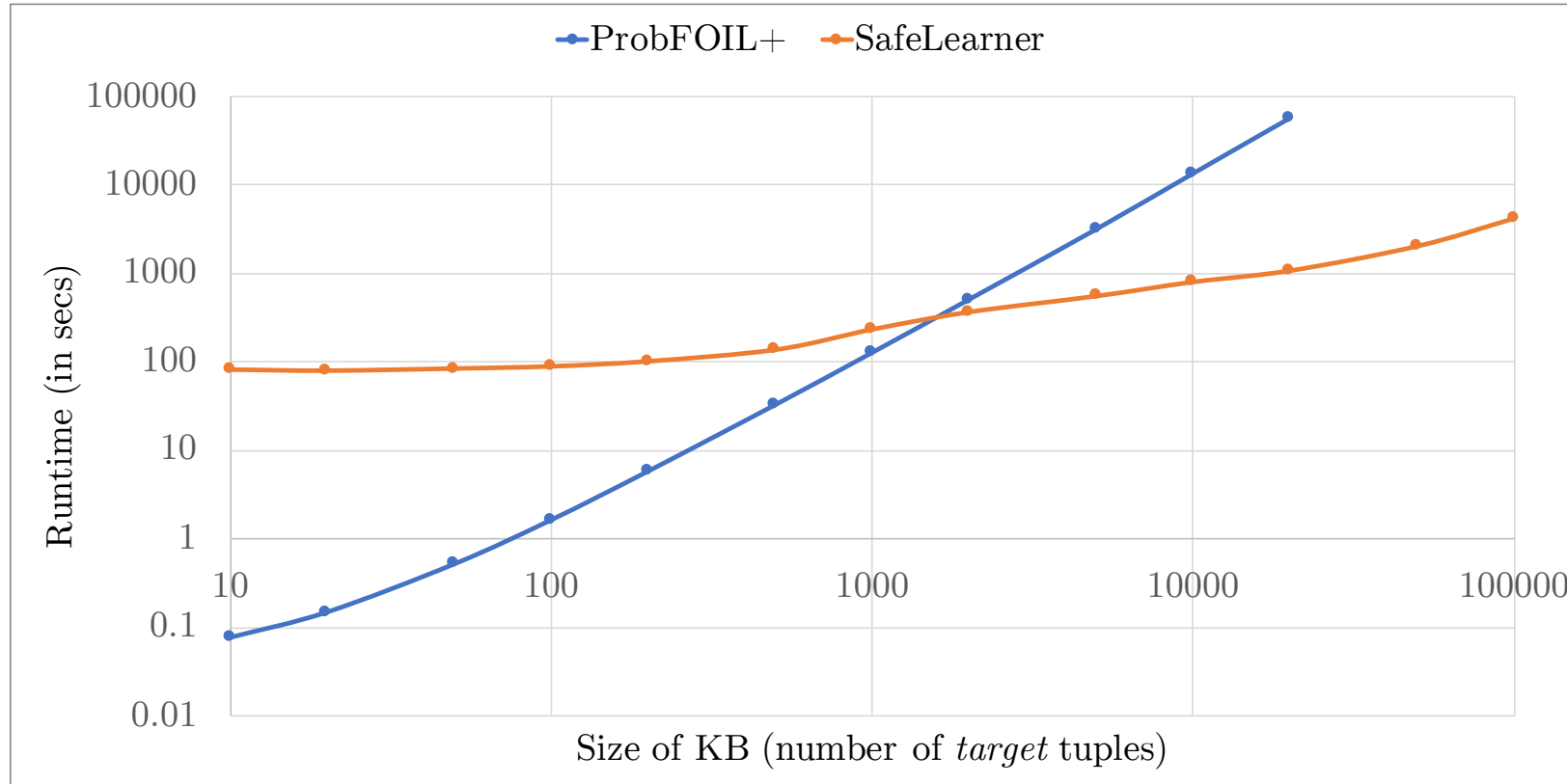
Probabilistic Expression

03

Caching

Does it really makes an impact?

- ▶ Significantly faster than ProbFOIL+, the state-of-the-art probabilistic rule learner



- ▶ Scales as good as AMIE+, the state-of-the-art Deterministic Rule Learner



SafeLearner can quickly fill some of these potholes! 😊



Thank You