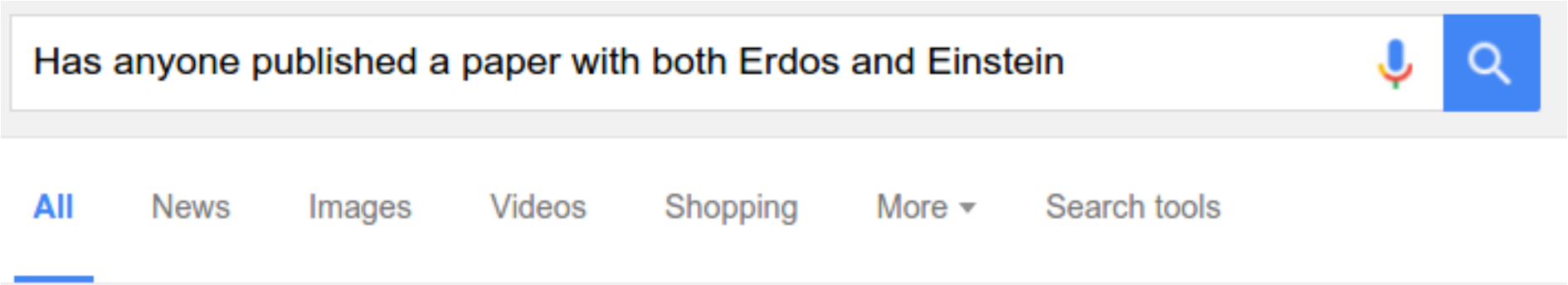
Outline

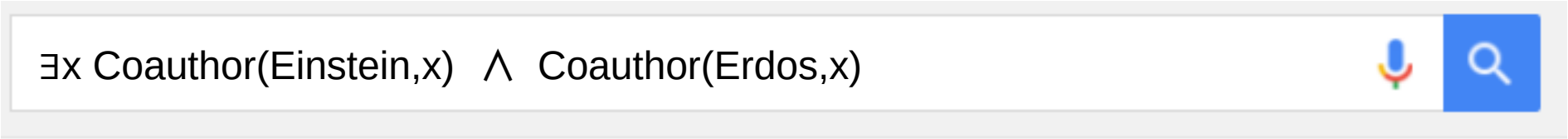
- Part 1: Motivation
- Part 2: Probabilistic Databases
- Part 3: Weighted Model Counting
- Part 4: Lifted Inference for WFOMC

- Part 5: Completeness of Lifted Inference
- Part 6: Query Compilation
- Part 7: Symmetric Lifted Inference Complexity
- Part 8: Open-World Probabilistic Databases
- Part 9: Discussion & Conclusions

What we'd like to do…



What we'd like to do…



What we’d like to do…

∃x Coauthor(Einstein,x) ∧ Coauthor(Erdos,x)

Ernst Straus

Kristian Kersting, …

Justin Bieber, …

Problem: Queries

- What if fact missing?
- Probability 0 for:

Coauthor

X	Y	P
Einstein	Straus	0.7
~~**Erdos**~~	~~**Straus**~~	~~**0.6**~~
Einstein	Pauli	0.9
Erdos	Renyi	0.7
Kersting	Natarajan	0.8
Luc	Paol	0.1
…	…	…

$Q1 = \exists x\ \text{Coauthor}(\text{Einstein}, \mathbf{x}) \wedge \text{Coauthor}(\text{Erdos}, \mathbf{x})$

Problem: Queries

- What if fact missing?
- Probability 0 for:

Coauthor

X	Y	P
Einstein	Straus	0.7
~~**Erdos**~~	~~**Straus**~~	~~**0.6**~~
Einstein	Pauli	0.9
Erdos	Renyi	0.7
Kersting	Natarajan	0.8
Luc	Paol	0.1
…	…	…

Q1 = ∃x Coauthor(Einstein,**x**) ∧ Coauthor(Erdos,**x**)

Q2 = Coauthor(Einstein,**Straus**) ∧ Coauthor(Erdos,**Straus**)

Problem: Queries

- What if fact missing?
- Probability 0 for:

Coauthor

X	Y	P
Einstein	Straus	0.7
~~**Erdos**~~	~~**Straus**~~	~~**0.6**~~
Einstein	Pauli	0.9
Erdos	Renyi	0.7
Kersting	Natarajan	0.8
Luc	Paol	0.1
…	…	…

Q1 = ∃x Coauthor(Einstein,**x**) ∧ Coauthor(Erdos,**x**)

Q2 = Coauthor(Einstein,**Straus**) ∧ Coauthor(Erdos,**Straus**)

Q3 = Coauthor(Einstein,**Kersting**) ∧ Coauthor(Erdos,**Kersting**)

Problem: Queries

- What if fact missing?
- Probability 0 for:

Coauthor

X	Y	P
Einstein	Straus	0.7
~~**Erdos**~~	~~**Straus**~~	~~**0.6**~~
Einstein	Pauli	0.9
Erdos	Renyi	0.7
Kersting	Natarajan	0.8
Luc	Paol	0.1
…	…	…

Q1 = ∃x Coauthor(Einstein,**x**) ∧ Coauthor(Erdos,**x**)

Q2 = Coauthor(Einstein,**Straus**) ∧ Coauthor(Erdos,**Straus**)

Q3 = Coauthor(Einstein,**Kersting**) ∧ Coauthor(Erdos,**Kersting**)

Q4 = Coauthor(Einstein,**Bieber**) ∧ Coauthor(Erdos,**Bieber**)

Problem: Queries

- What if fact missing?
- Probability 0 for:

Coauthor

X	Y	P
Einstein	Straus	0.7
~~**Erdos**~~	~~**Straus**~~	~~**0.6**~~
Einstein	Pauli	0.9
Erdos	Renyi	0.7
Kersting	Natarajan	0.8
Luc	Paol	0.1
…	…	…

Q1 = ∃x Coauthor(Einstein,**x**) ∧ Coauthor(Erdos,**x**)

Q2 = Coauthor(Einstein,**Straus**) ∧ Coauthor(Erdos,**Straus**)

Q3 = Coauthor(Einstein,**Kersting**) ∧ Coauthor(Erdos,**Kersting**)

Q4 = Coauthor(Einstein,**Bieber**) ∧ Coauthor(Erdos,**Bieber**)

Q5 = Coauthor(Einstein,**Bieber**) ∧ ¬Coauthor(**Einstein**,**Bieber**)

Intuition

X	Y	P
Einstein	Straus	0.7
~~**Erdos**~~	~~**Straus**~~	~~**0.6**~~
Einstein	Pauli	0.9
Erdos	Renyi	0.7
Kersting	Natarajan	0.8
Luc	Paol	0.1
…	…	…

Q1 = ∃x Coauthor(Einstein,**x**) ∧ Coauthor(Erdos,**x**)

Q2 = Coauthor(Einstein,**Straus**) ∧ Coauthor(Erdos,**Straus**)

Q3 = Coauthor(Einstein,**Kersting**) ∧ Coauthor(Erdos,**Kersting**)

Q4 = Coauthor(Einstein,**Bieber**) ∧ Coauthor(Erdos,**Bieber**)

Q5 = Coauthor(Einstein,**Bieber**) ∧ ¬Coauthor(**Einstein**,**Bieber**)

[Ceylan'16]

Intuition

X	Y	P
Einstein	Straus	0.7
~~**Erdos**~~	~~**Straus**~~	~~**0.6**~~
Einstein	Pauli	0.9
Erdos	Renyi	0.7
Kersting	Natarajan	0.8
Luc	Paol	0.1
…	…	…

Q1 = ∃x Coauthor(Einstein,**x**) ∧ Coauthor(Erdos,**x**)

Q2 = Coauthor(Einstein,**Straus**) ∧ Coauthor(Erdos,**Straus**)

Q3 = Coauthor(Einstein,**Kersting**) ∧ Coauthor(Erdos,**Kersting**)

Q4 = Coauthor(Einstein,**Bieber**) ∧ Coauthor(Erdos,**Bieber**)

Q5 = Coauthor(Einstein,**Bieber**) ∧ ¬Coauthor(**Einstein**,**Bieber**)

We know for sure that P(Q1) ≥ P(Q2), P(Q1) ≥ P(Q3), P(Q1) ≥ P(Q4)

[Ceylan'16]

Intuition

X	Y	P
Einstein	Straus	0.7
~~Erdos~~	~~Straus~~	~~0.6~~
Einstein	Pauli	0.9
Erdos	Renyi	0.7
Kersting	Natarajan	0.8
Luc	Paol	0.1
…	…	…

Q1 = ∃x Coauthor(Einstein,**x**) ∧ Coauthor(Erdos,**x**)

Q2 = Coauthor(Einstein,**Straus**) ∧ Coauthor(Erdos,**Straus**)

Q3 = Coauthor(Einstein,**Kersting**) ∧ Coauthor(Erdos,**Kersting**)

Q4 = Coauthor(Einstein,**Bieber**) ∧ Coauthor(Erdos,**Bieber**)

Q5 = Coauthor(Einstein,**Bieber**) ∧ ¬Coauthor(**Einstein**,**Bieber**)

We know for sure that P(Q1) ≥ P(Q2), P(Q1) ≥ P(Q3), P(Q1) ≥ P(Q4)

and P(Q2) ≥ P(Q5), P(Q3) ≥ P(Q5), P(Q4) ≥ P(Q5)

[Ceylan'16]

Intuition

X	Y	P
Einstein	Straus	0.7
~~**Erdos**~~	~~**Straus**~~	~~**0.6**~~
Einstein	Pauli	0.9
Erdos	Renyi	0.7
Kersting	Natarajan	0.8
Luc	Paol	0.1
…	…	…

Q1 = ∃x Coauthor(Einstein,**x**) ∧ Coauthor(Erdos,**x**)

Q2 = Coauthor(Einstein,**Straus**) ∧ Coauthor(Erdos,**Straus**)

Q3 = Coauthor(Einstein,**Kersting**) ∧ Coauthor(Erdos,**Kersting**)

Q4 = Coauthor(Einstein,**Bieber**) ∧ Coauthor(Erdos,**Bieber**)

Q5 = Coauthor(Einstein,**Bieber**) ∧ ¬Coauthor(**Einstein**,**Bieber**)

We know for sure that P(Q1) ≥ P(Q2), P(Q1) ≥ P(Q3), P(Q1) ≥ P(Q4)

and P(Q2) ≥ P(Q5), P(Q3) ≥ P(Q5), P(Q4) ≥ P(Q5) and P(Q5) = 0.

[Ceylan'16]

Intuition

X	Y	P
Einstein	Straus	0.7
~~**Erdos**~~	~~**Straus**~~	~~**0.6**~~
Einstein	Pauli	0.9
Erdos	Renyi	0.7
Kersting	Natarajan	0.8
Luc	Paol	0.1
…	…	…

Q1 = ∃x Coauthor(Einstein,**x**) ∧ Coauthor(Erdos,**x**)

Q2 = Coauthor(Einstein,**Straus**) ∧ Coauthor(Erdos,**Straus**)

Q3 = Coauthor(Einstein,**Kersting**) ∧ Coauthor(Erdos,**Kersting**)

Q4 = Coauthor(Einstein,**Bieber**) ∧ Coauthor(Erdos,**Bieber**)

Q5 = Coauthor(Einstein,**Bieber**) ∧ ¬Coauthor(**Einstein**,**Bieber**)

We know for sure that P(Q1) ≥ P(Q2), P(Q1) ≥ P(Q3), P(Q1) ≥ P(Q4)

and P(Q2) ≥ P(Q5), P(Q3) ≥ P(Q5), P(Q4) ≥ P(Q5) and P(Q5) = 0.

We have strong evidence that P(Q2) ≥ P(Q3) ≥ P(Q4).

[Ceylan'16]

Problem: Broken Learning Loop

Bayesian view on learning:

- Prior belief:

  Pr(HasStudent(Luc,Paol)) = 0.01

- Observe page

  Pr(HasStudent(Luc,Paol) | ) = 0.2

- Observe page

  Pr(HasStudent(Luc,Paol) |  , ) = 0.3

Principled and sound reasoning!

Problem: Broken Learning Loop

Current view on Knowledge Base Completion:

- Prior belief:

  Pr(HasStudent(Luc,Paol)) = 0

- Observe page

  Pr(HasStudent(Luc,Paol) | ) = 0.2

- Observe page

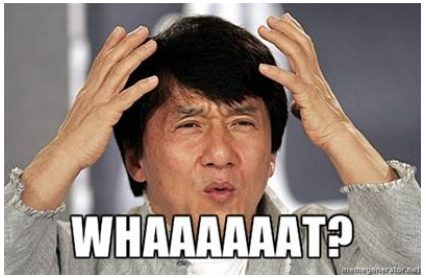
  Pr(HasStudent(Luc,Paol) |  , ) = 0.3

Problem: Broken Learning Loop

Current view on Knowledge Base Completion:

- Prior belief:

  Pr(HasStudent(Luc,Paol)) = 0



- Observe page

  Pr(HasStudent(Luc,Paol) | ) = 0.2

- Observe page

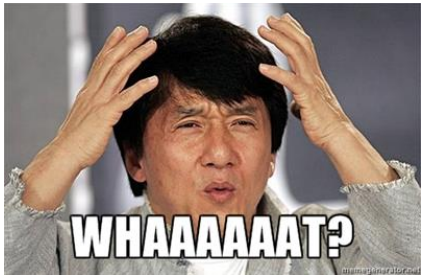
  Pr(HasStudent(Luc,Paol) |  , ) = 0.3

Problem: Broken Learning Loop

Current view on Knowledge Base Completion:

- Prior belief:

  Pr(HasStudent(Luc,Paol)) = 0



- Observe page

  Pr(HasStudent(Luc,Paol) | ) = 0.2

- Observe page

  Pr(HasStudent(Luc,Paol) |  , ) = 0.3

This is mathematical nonsense!

Knowledge Base Completion

Given:

LivesIn

X	Y
Luc	Belgium
Guy	USA
Kristian	Germany

LocatedIn

X	Y
Siemens	Germany
Siemens	Belgium
UCLA	USA
TUDortmund	Germany
KU Leuven	Belgium

WorksFor

X	Y
Luc	KU Leuven
Guy	UCLA
Kristian	TUDortmund
Ingo	Siemens

Learn:

0.8::LivesIn(x,y) :- WorksFor(x,z) ∧ LocatedIn(z,x).

How to measure success?

WorksFor

X	Y	P
Luc	KU Leuven	0.7
Guy	UCLA	0.6
Kristian	TUDortmund	0.3
Ingo	Siemens	0.3

LocatedIn

X	Y	P
Siemens	Germany	0.7
Siemens	Belgium	0.5
UCLA	USA	0.8
TUDortmund	Germany	0.6
KU Leuven	Belgium	0.7

0.8::LivesIn(x,y) :- WorksFor(x,z) ∧ LocatedIn(z,x).

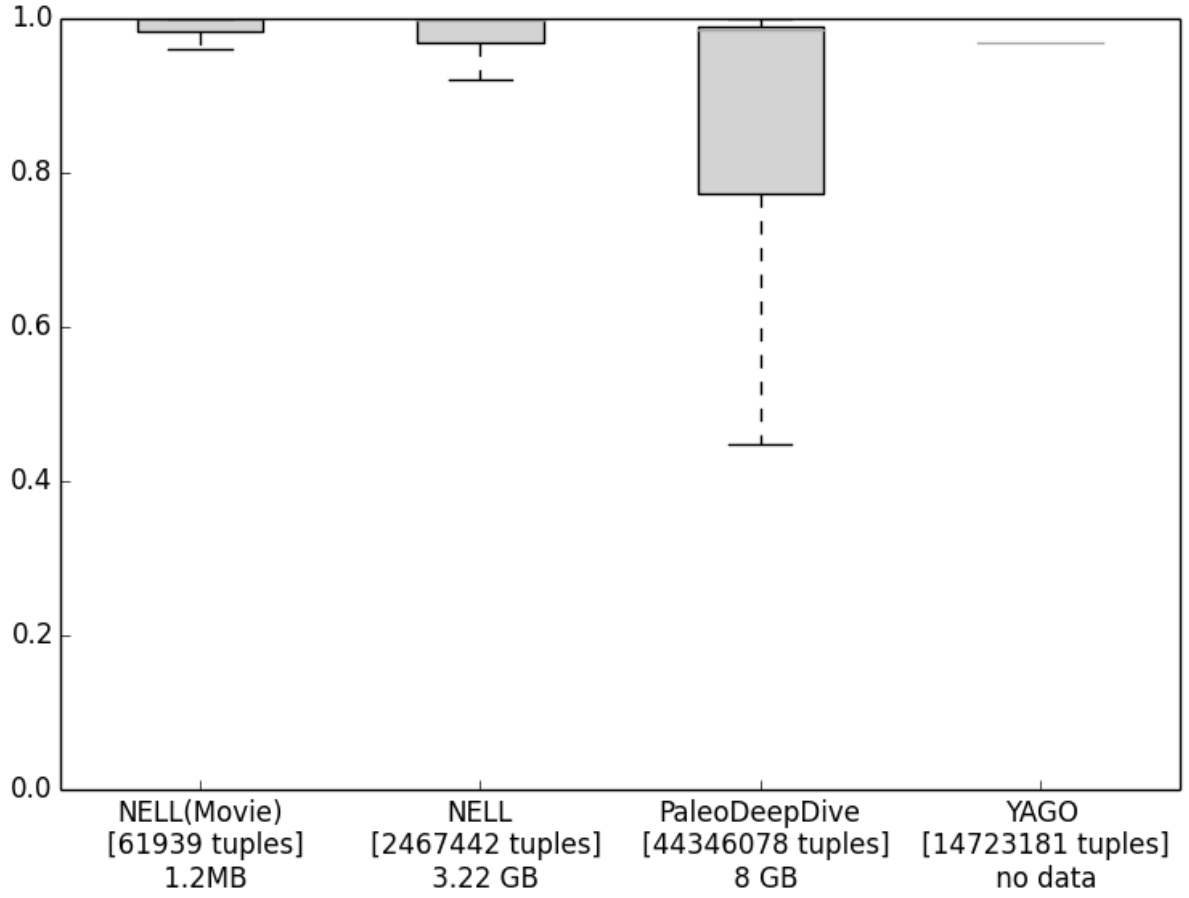
or

0.5::LivesIn(x,y) :- BornIn(x,y).

What is the likelihood, precision, accuracy, …?

Problem: Curse of Superlinearity

- Reality is worse!
- Tuples are intentionally missing!
- Every tuple has 99% pr.



Problem: Curse of Superlinearity

“This is all true, Guy,
but it’s just a temporary issue”

“No it’s not!”

Problem: Curse of Superlinearity

- A single table

Sibling

X	Y	P
…	…	…

- At the scale of facebook (billions of people)
- Real Bayesian belief about everyone

  I.e., all non-zero probabilities

⇒ 200 Exabytes of data

Problem: Curse of Superlinearity

All Google storage is a couple exabytes…

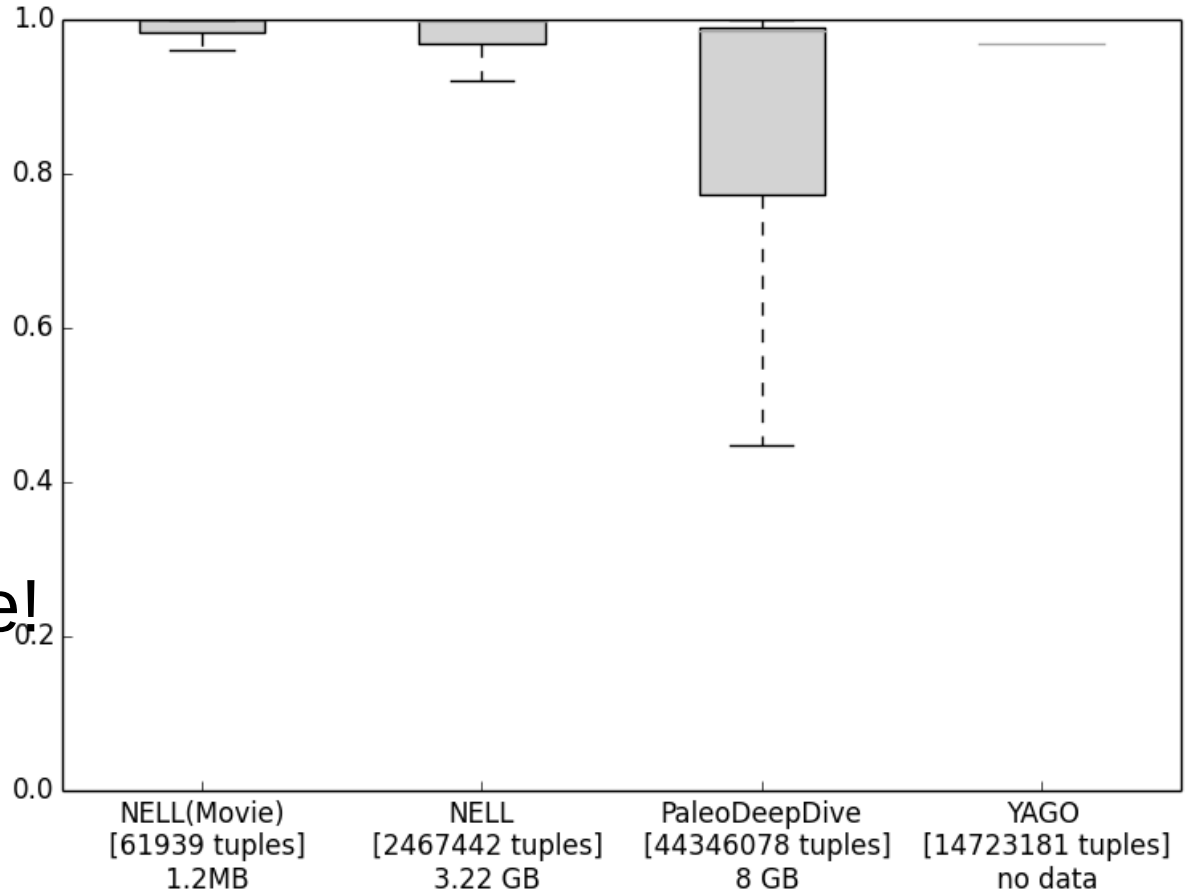
ing. In *Proc. of AAAI'15*. AAAI Press, 2015.

Randall Munroe. Google's datacenters on punch cards, 2015.

James D Park and Adnan Darwiche. Complexity Results and

Problem: Curse of Superlinearity

We should be here!



Closed-World Prob. Databases

A PDB $\mathcal{P}$ induces a *unique probability distribution* over worlds ω:

$$\mathrm{P}_{\mathcal{P}}(\omega) = \prod_{t \in \omega} \mathrm{P}_{\mathcal{P}}(t) \prod_{t \notin \omega} (1 - \mathrm{P}_{\mathcal{P}}(t)),$$

where for every tuple t, it holds that

$$\mathrm{P}_{\mathcal{P}}(t) = \begin{cases} p & \text{if } \langle t : p \rangle \in \mathcal{P} \\ 0 & \text{otherwise.} \end{cases}$$ [Probabilistic CWA]

Open-World Prob. Databases

An *OpenPDB* is a pair $\mathcal{G} = (\mathcal{P}, \lambda)$, where $\mathcal{P}$ is a PDB

$$\mathrm{P}_{\mathcal{G}}(t) = \begin{cases} p & \text{if } \langle t : p \rangle \in \mathcal{P} \\ [0, \lambda] & \text{otherwise.} \end{cases}$$

A *λ-completion* of $\mathcal{G}$ contains a tuple $\langle t : p \rangle$ for some $p \in [0, \lambda]$ for every $t \notin \mathcal{P}$. $\mathcal{G}$ induces *a set of probability distributions* $\mathrm{K}_{\mathcal{G}}$ [3]:

$$\underline{\mathrm{P}}_{\mathcal{G}}(\mathrm{Q}) = \min_{\mathrm{P} \in \mathrm{K}_{\mathcal{G}}} \mathrm{P}(\mathrm{Q}) \qquad \text{and} \qquad \overline{\mathrm{P}}_{\mathcal{G}}(\mathrm{Q}) = \max_{\mathrm{P} \in \mathrm{K}_{\mathcal{G}}} \mathrm{P}(\mathrm{Q}).$$

[Ceylan'16]

Open-World Prob. Databases

Intuition: tuples can be added with prob < λ

Q2 = Coauthor(Einstein,**Straus**) ∧ Coauthor(Erdos,**Straus**)

Coauthor

X	Y	P
Einstein	Straus	0.7
Einstein	Pauli	0.9
Erdos	Renyi	0.7
Kersting	Natarajan	0.8
Luc	Paol	0.1
…	…	…

Open-World Prob. Databases

Intuition: tuples can be added with prob < λ

Q2 = Coauthor(Einstein,**Straus**) ∧ Coauthor(Erdos,**Straus**)

Coauthor

X	Y	P
Einstein	Straus	0.7
Einstein	Pauli	0.9
Erdos	Renyi	0.7
Kersting	Natarajan	0.8
Luc	Paol	0.1
…	…	…

Coauthor

X	Y	P
Einstein	Straus	0.7
Einstein	Pauli	0.9
Erdos	Renyi	0.7
Kersting	Natarajan	0.8
Luc	Paol	0.1
…	…	…
Erdos	**Straus**	**λ**

Open-World Prob. Databases

Intuition: tuples can be added with prob < λ

Q2 = Coauthor(Einstein,**Straus**) ∧ Coauthor(Erdos,**Straus**)

Coauthor

X	Y	P
Einstein	Straus	0.7
Einstein	Pauli	0.9
Erdos	Renyi	0.7
Kersting	Natarajan	0.8
Luc	Paol	0.1
…	…	…

Coauthor

X	Y	P
Einstein	Straus	0.7
Einstein	Pauli	0.9
Erdos	Renyi	0.7
Kersting	Natarajan	0.8
Luc	Paol	0.1
…	…	…
Erdos	**Straus**	**λ**

$0.7 * \lambda \geq P(Q2) \geq 0$

Monotone Queries

- E.g., Unions of Conjunctive Queries (UCQ)

- Lower bound = closed world probability
- Upper bound = probability after adding all tuples with probability λ

- Quadratic blow-up ☹
- Lifted inference to the rescue!

Independent Project

$Q = \forall x \forall y \, (Smoker(x) \vee Friend(x,y))$

[Ceylan'16]

Independent Project

$Q = \forall x \forall y\, (Smoker(x) \vee Friend(x,y))$

$P(Q) = \prod_{A \in Domain} P(Smoker(A) \vee \forall y\, Friend(A,y))$

∀-Rule

Check independence:

Smoker(Alice) ∨ ∀y Friend(Alice,y)

Smoker(Bob) ∨ ∀y Friend(Bob,y)

[Ceylan'16]

Independent Project

$Q = \forall x \forall y\ (\text{Smoker}(x) \vee \text{Friend}(x,y))$

∀-Rule

$$P(Q) = \prod_{A \in \text{Domain}} P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y))$$

Check independence:

Smoker(Alice) ∨ ∀y Friend(Alice,y)

Smoker(Bob) ∨ ∀y Friend(Bob,y)

$$= P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y)) \times P(\text{Smoker}(B) \vee \forall y\ \text{Friend}(B,y)) \times P(\text{Smoker}(C) \vee \forall y\ \text{Friend}(C,y)) \times P(\text{Smoker}(D) \vee \forall y\ \text{Friend}(D,y)) \times P(\text{Smoker}(E) \vee \forall y\ \text{Friend}(E,y)) \times P(\text{Smoker}(F) \vee \forall y\ \text{Friend}(F,y)) \times \ldots$$

[Ceylan'16]

Independent Project

$Q = \forall x \forall y\ (Smoker(x) \vee Friend(x,y))$

$P(Q) = \prod_{A \in Domain} P(Smoker(A) \vee \forall y\ Friend(A,y))$

∀-Rule

Check independence:

Smoker(Alice) ∨ ∀y Friend(Alice,y)

Smoker(Bob) ∨ ∀y Friend(Bob,y)

$= P(Smoker(A) \vee \forall y\ Friend(A,y))$
$\times P(Smoker(B) \vee \forall y\ Friend(B,y))$
$\times P(Smoker(C) \vee \forall y\ Friend(C,y))$
$\times P(Smoker(D) \vee \forall y\ Friend(D,y))$
$\times P(Smoker(E) \vee \forall y\ Friend(E,y))$
$\times P(Smoker(F) \vee \forall y\ Friend(F,y))$
…

Complexity PTIME?

[Ceylan'16]

Closed-World Independent Project

$Q = \forall x \forall y\ (\text{Smoker}(x) \vee \text{Friend}(x,y))$

∀-Rule

$$P(Q) = \prod_{A \in \text{Domain}} P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y))$$

Check independence:

Smoker(Alice) ∨ ∀y Friend(Alice,y)

Smoker(Bob) ∨ ∀y Friend(Bob,y)

$$= P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y)) \times P(\text{Smoker}(B) \vee \forall y\ \text{Friend}(B,y)) \times P(\text{Smoker}(C) \vee \forall y\ \text{Friend}(C,y)) \times P(\text{Smoker}(D) \vee \forall y\ \text{Friend}(D,y)) \times P(\text{Smoker}(E) \vee \forall y\ \text{Friend}(E,y)) \times P(\text{Smoker}(F) \vee \forall y\ \text{Friend}(F,y)) \times \ldots$$

Closed-World Independent Project

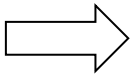
$Q = \forall x \forall y\ (\text{Smoker}(x) \vee \text{Friend}(x,y))$

∀-Rule

Check independence:
Smoker(Alice) ∨ ∀y Friend(Alice,y)
Smoker(Bob) ∨ ∀y Friend(Bob,y)

$$P(Q) = \prod_{A \in \text{Domain}} P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y))$$

$$= P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y)) \times P(\text{Smoker}(B) \vee \forall y\ \text{Friend}(B,y)) \times P(\text{Smoker}(C) \vee \forall y\ \text{Friend}(C,y)) \times P(\text{Smoker}(D) \vee \forall y\ \text{Friend}(D,y)) \times P(\text{Smoker}(E) \vee \forall y\ \text{Friend}(E,y)) \times P(\text{Smoker}(F) \vee \forall y\ \text{Friend}(F,y)) \times \ldots$$



No supporting facts in database!

Closed-World Independent Project

$Q = \forall x \forall y\ (\text{Smoker}(x) \vee \text{Friend}(x,y))$

∀-Rule

Check independence:
Smoker(Alice) ∨ ∀y Friend(Alice,y)
Smoker(Bob) ∨ ∀y Friend(Bob,y)

$$P(Q) = \prod_{A \in \text{Domain}} P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y))$$

$$= P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y)) \times P(\text{Smoker}(B) \vee \forall y\ \text{Friend}(B,y)) \times P(\text{Smoker}(C) \vee \forall y\ \text{Friend}(C,y)) \times P(\text{Smoker}(D) \vee \forall y\ \text{Friend}(D,y)) \times P(\text{Smoker}(E) \vee \forall y\ \text{Friend}(E,y)) \times P(\text{Smoker}(F) \vee \forall y\ \text{Friend}(F,y)) \times \ldots$$

⇒ No supporting facts in database!

⇒ Probability 0 in closed world

Closed-World Independent Project

$Q = \forall x \forall y\ (\text{Smoker}(x) \vee \text{Friend}(x,y))$

∀-Rule

$P(Q) = \prod_{A \in \text{Domain}} P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y))$

Check independence:
Smoker(Alice) ∨ ∀y Friend(Alice,y)
Smoker(Bob) ∨ ∀y Friend(Bob,y)

$= P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y))$
$\times P(\text{Smoker}(B) \vee \forall y\ \text{Friend}(B,y))$
$\times P(\text{Smoker}(C) \vee \forall y\ \text{Friend}(C,y))$
$\times P(\text{Smoker}(D) \vee \forall y\ \text{Friend}(D,y))$
$\times P(\text{Smoker}(E) \vee \forall y\ \text{Friend}(E,y))$
$\times P(\text{Smoker}(F) \vee \forall y\ \text{Friend}(F,y))$
…

⇒ No supporting facts in database!

⇒ Probability 0 in closed world

⇒ Ignore these queries!

Closed-World Independent Project

$Q = \forall x \forall y\ (\text{Smoker}(x) \vee \text{Friend}(x,y))$

∀-Rule

$P(Q) = \prod_{A \in \text{Domain}} P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y))$

Check independence:
Smoker(Alice) ∨ ∀y Friend(Alice,y)
Smoker(Bob) ∨ ∀y Friend(Bob,y)

$= P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y))$
$\times P(\text{Smoker}(B) \vee \forall y\ \text{Friend}(B,y))$
$\times P(\text{Smoker}(C) \vee \forall y\ \text{Friend}(C,y))$
$\times P(\text{Smoker}(D) \vee \forall y\ \text{Friend}(D,y))$
$\times P(\text{Smoker}(E) \vee \forall y\ \text{Friend}(E,y))$
$\times P(\text{Smoker}(F) \vee \forall y\ \text{Friend}(F,y))$
…

⇒ No supporting facts in database!

⇒ Probability 0 in closed world

⇒ Ignore these queries!

Complexity linear time!

Open-World Independent Project

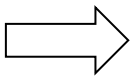
$Q = \forall x \forall y\ (\text{Smoker}(x) \vee \text{Friend}(x,y))$

∀-Rule

Check independence:
Smoker(Alice) ∨ ∀y Friend(Alice,y)
Smoker(Bob) ∨ ∀y Friend(Bob,y)

$$P(Q) = \prod_{A \in \text{Domain}} P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y))$$

$$= P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y)) \times P(\text{Smoker}(B) \vee \forall y\ \text{Friend}(B,y)) \times P(\text{Smoker}(C) \vee \forall y\ \text{Friend}(C,y)) \times P(\text{Smoker}(D) \vee \forall y\ \text{Friend}(D,y)) \times P(\text{Smoker}(E) \vee \forall y\ \text{Friend}(E,y)) \times P(\text{Smoker}(F) \vee \forall y\ \text{Friend}(F,y)) \times \ldots$$



No supporting facts in database!

Open-World Independent Project

$Q = \forall x \forall y\, (\text{Smoker}(x) \vee \text{Friend}(x,y))$

∀-Rule

Check independence:
Smoker(Alice) ∨ ∀y Friend(Alice,y)
Smoker(Bob) ∨ ∀y Friend(Bob,y)

$$P(Q) = \prod_{A \in \text{Domain}} P(\text{Smoker}(A) \vee \forall y\, \text{Friend}(A,y))$$

$$= P(\text{Smoker}(A) \vee \forall y\, \text{Friend}(A,y)) \times P(\text{Smoker}(B) \vee \forall y\, \text{Friend}(B,y)) \times P(\text{Smoker}(C) \vee \forall y\, \text{Friend}(C,y)) \times P(\text{Smoker}(D) \vee \forall y\, \text{Friend}(D,y)) \times P(\text{Smoker}(E) \vee \forall y\, \text{Friend}(E,y)) \times P(\text{Smoker}(F) \vee \forall y\, \text{Friend}(F,y)) \times \ldots$$

⇒ No supporting facts in database!

⇒ Probability p in closed world

Open-World Independent Project

$Q = \forall x \forall y\ (\text{Smoker}(x) \vee \text{Friend}(x,y))$

∀-Rule

$$P(Q) = \prod_{A \in \text{Domain}} P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y))$$

Check independence:

Smoker(Alice) ∨ ∀y Friend(Alice,y)

Smoker(Bob) ∨ ∀y Friend(Bob,y)

$$= P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y)) \times P(\text{Smoker}(B) \vee \forall y\ \text{Friend}(B,y)) \times P(\text{Smoker}(C) \vee \forall y\ \text{Friend}(C,y)) \times P(\text{Smoker}(D) \vee \forall y\ \text{Friend}(D,y)) \times P(\text{Smoker}(E) \vee \forall y\ \text{Friend}(E,y)) \times P(\text{Smoker}(F) \vee \forall y\ \text{Friend}(F,y)) \times \ldots$$

⇒ No supporting facts in database!

⇒ Probability p in closed world

Complexity PTIME!

Open-World Independent Project

$Q = \forall x \forall y\, (\text{Smoker}(x) \vee \text{Friend}(x,y))$

∀-Rule

Check independence:
Smoker(Alice) ∨ ∀y Friend(Alice,y)
Smoker(Bob) ∨ ∀y Friend(Bob,y)

$$P(Q) = \prod_{A \in \text{Domain}} P(\text{Smoker}(A) \vee \forall y\, \text{Friend}(A,y))$$

$$= P(\text{Smoker}(A) \vee \forall y\, \text{Friend}(A,y)) \times P(\text{Smoker}(B) \vee \forall y\, \text{Friend}(B,y)) \times P(\text{Smoker}(C) \vee \forall y\, \text{Friend}(C,y)) \times P(\text{Smoker}(D) \vee \forall y\, \text{Friend}(D,y)) \times P(\text{Smoker}(E) \vee \forall y\, \text{Friend}(E,y)) \times P(\text{Smoker}(F) \vee \forall y\, \text{Friend}(F,y)) \times \ldots$$

⇒ No supporting facts in database!

⇒ Probability p in closed world

Open-World Independent Project

$Q = \forall x \forall y\ (\text{Smoker}(x) \vee \text{Friend}(x,y))$

$$P(Q) = \prod_{A \in \text{Domain}} P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y))$$

∀-Rule

Check independence:

Smoker(Alice) ∨ ∀y Friend(Alice,y)

Smoker(Bob) ∨ ∀y Friend(Bob,y)

$$= P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y)) \times P(\text{Smoker}(B) \vee \forall y\ \text{Friend}(B,y)) \times P(\text{Smoker}(C) \vee \forall y\ \text{Friend}(C,y)) \times P(\text{Smoker}(D) \vee \forall y\ \text{Friend}(D,y)) \times P(\text{Smoker}(E) \vee \forall y\ \text{Friend}(E,y)) \times P(\text{Smoker}(F) \vee \forall y\ \text{Friend}(F,y)) \times \ldots$$

⇒ No supporting facts in database!

⇒ Probability p in closed world

⇒ All together, probability p^k

Do symmetric lifted inference

Open-World Independent Project

$Q = \forall x \forall y\ (\text{Smoker}(x) \vee \text{Friend}(x,y))$

∀-Rule

Check independence:
Smoker(Alice) ∨ ∀y Friend(Alice,y)
Smoker(Bob) ∨ ∀y Friend(Bob,y)

$$P(Q) = \prod_{A \in \text{Domain}} P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y))$$

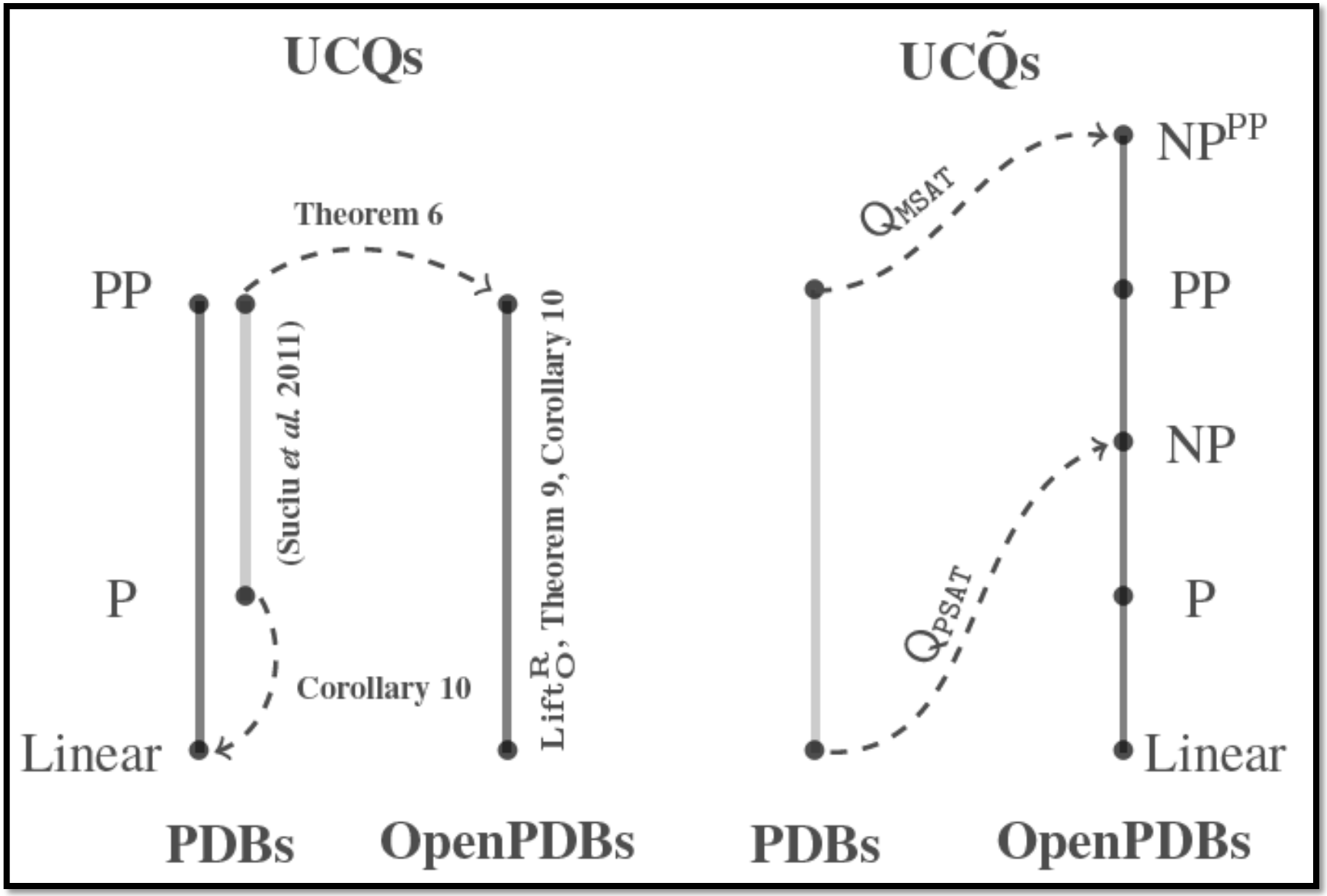
$$= P(\text{Smoker}(A) \vee \forall y\ \text{Friend}(A,y)) \times P(\text{Smoker}(B) \vee \forall y\ \text{Friend}(B,y)) \times P(\text{Smoker}(C) \vee \forall y\ \text{Friend}(C,y)) \times P(\text{Smoker}(D) \vee \forall y\ \text{Friend}(D,y)) \times P(\text{Smoker}(E) \vee \forall y\ \text{Friend}(E,y)) \times P(\text{Smoker}(F) \vee \forall y\ \text{Friend}(F,y)) \times \ldots$$

⇒ No supporting facts in database!

⇒ Probability p in closed world

⇒ All together, probability p^k
Do symmetric lifted inference

Complexity linear time!



$$\text{Linear} \subseteq \text{P} \subseteq \text{NP} \subseteq \text{PP} \subseteq \text{P}^{\text{PP}} \subseteq \text{NP}^{\text{PP}} \subseteq \text{PSpace} \subseteq \text{ExpTime}$$

[Ceylan'16]

Summary

- Open-world semantics make sense
- Matches how systems are employed
- Open-world reasoning is FREE for UCQs
- Beyond UCQs, can pay a hefty price
- Future work:
  More refined models of the open world
  E.g., (types, MLNs, additional statistics)