Probabilistic Reasoning by First-Order Model Counting

Guy Van den Broeck

UCLA

Simons Workshop on Uncertainty in Computation
Oct 5, 2016

What is this good for?

What we'd like to do…

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

All   News   Images   Videos   Shopping   More ▾   Search tools

About 82,400 results (0.73 seconds)

Erdős number - Wikipedia, the free encyclopedia
https://en.wikipedia.org/wiki/**Erdős**_number ▾   Wikipedia ▾
He **published** more **papers** during his lifetime (at least 1,525) than any other ...
Anybody else's **Erdős** number is k + 1 where k is the lowest **Erdős** number of any
coauthor. ... Albert **Einstein** and Sheldon Lee Glashow **have** an **Erdős** number of 2. ...
and mathematician Ruth Williams, **both** of whom **have** an **Erdős** number of 2.

Erdős–Bacon number - Wikipedia, the free encyclopedia
https://en.wikipedia.org/wiki/**Erdős**–Bacon_number ▾   Wikipedia ▾
This article possibly **contains** previously unpublished synthesis of **published** ... Her
paper gives her an **Erdős** number of 4, and a Bacon number of 2, **both** of ...

Einstein is in the Knowledge Graph

Erdős is in the Knowledge Graph

This guy is in the Knowledge Graph

This guy is in the Knowledge Graph

… and he published with both Einstein and Erdos!

Desired Query Answer

Has anyone published a paper with both Erdos and Einstein

⭐⭐⭐⭐⭐

Ernst Straus

⭐⭐☆☆☆

Barack Obama, …

⭐☆☆☆☆

Justin Bieber, …

Desired Query Answer

Has anyone published a paper with both Erdos and Einstein

⭐⭐⭐⭐⭐

Ernst Straus

⭐⭐☆☆☆

Barack Obama, …

⭐☆☆☆☆

Justin Bieber, …

1. Fuse uncertain information from web

⇒ **Embrace probability!**

2. Cannot come from labeled data

⇒ **Embrace query eval!**

Siri, Alexa and Other Virtual Assistants Put to the Test

Tech Fix

By **BRIAN X. CHEN**   JAN. 27, 2016

NAME: Google
3.1

NAME: Siri
2.9

NAME: Cortana
2.3

NAME: Alexa
1.7
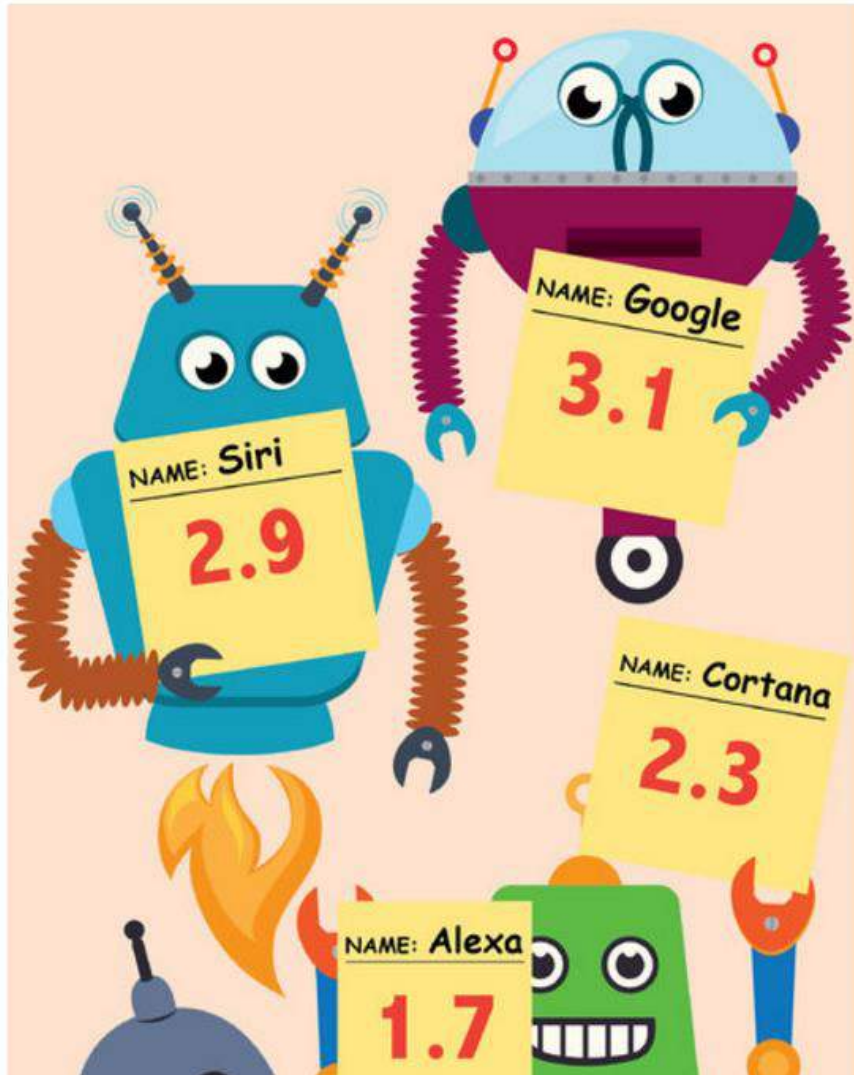
WHEN I asked Alexa earlier this week who was playing in the Super Bowl, she responded, somewhat monotonously, "Super Bowl 49's winner is New England Patriots."

"Come on, that's last year's Super Bowl," I said. "Even I can do better than that."

At the time, I was actually alone in my living room. I was talking to the virtual companion inside Amazon's wireless speaker, Echo, which was released last June. Known as Alexa, she has gained raves from Silicon Valley's tech-obsessed digerati and has become one of the newest members of the virtual assistants club.

All the so-called Frightful Five tech

[Chen'16]
(NYTimes)

Statistical Relational Learning
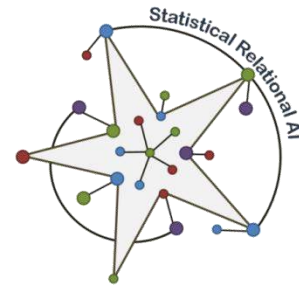
Augment graphical model with relations between entities (rows).

<u>Intuition</u>
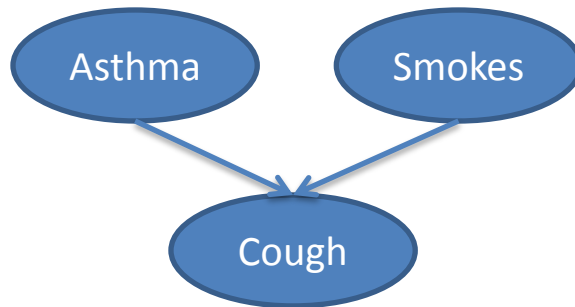
<u>Markov Logic</u>

+ Friends have similar smoking habits

+ Asthma can be hereditary

2.1 Asthma(x) ⇒ Cough(x)

3.5 Smokes(x) ⇒ Cough(x)

1.9 Smokes(x) ∧ Friends(x,y) ⇒ Smokes(y)

1.5 Asthma (x) ∧ Family(x,y) ⇒ Asthma (y)

*Why do we need first-order probabilistic **reasoning**?*

A Simple Reasoning Problem
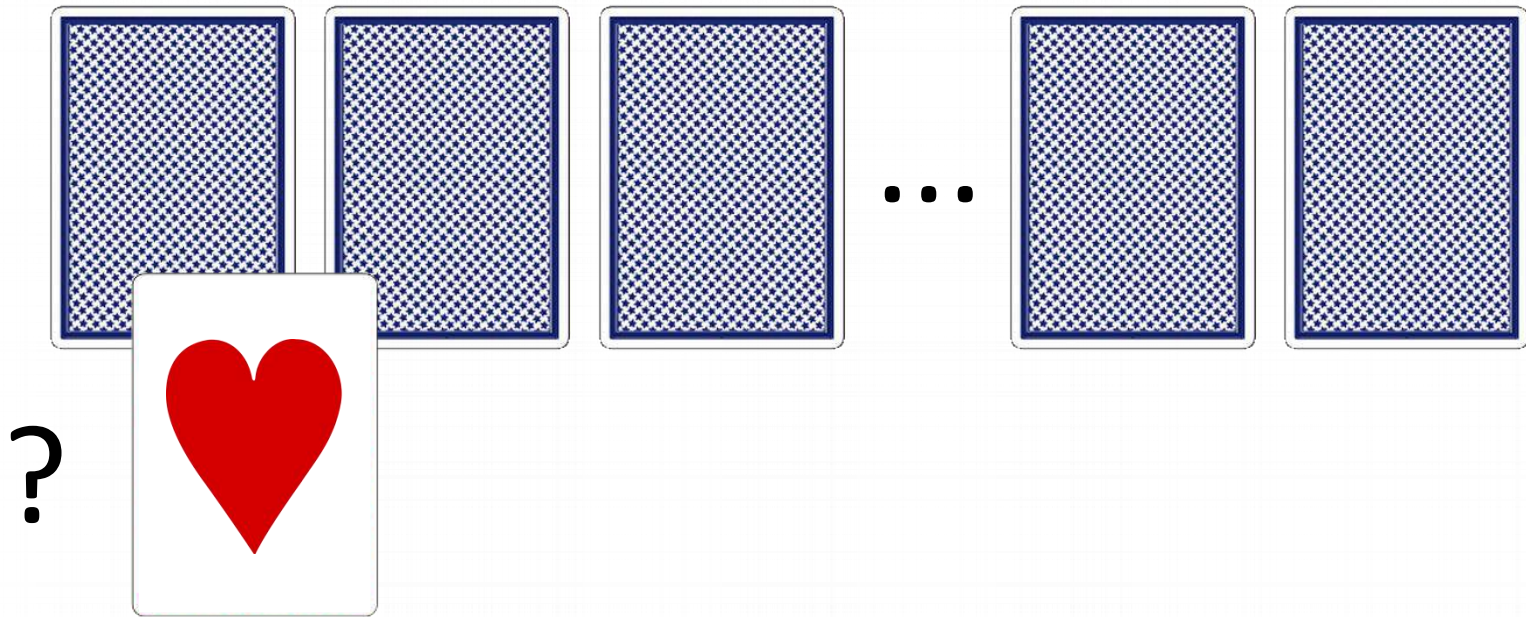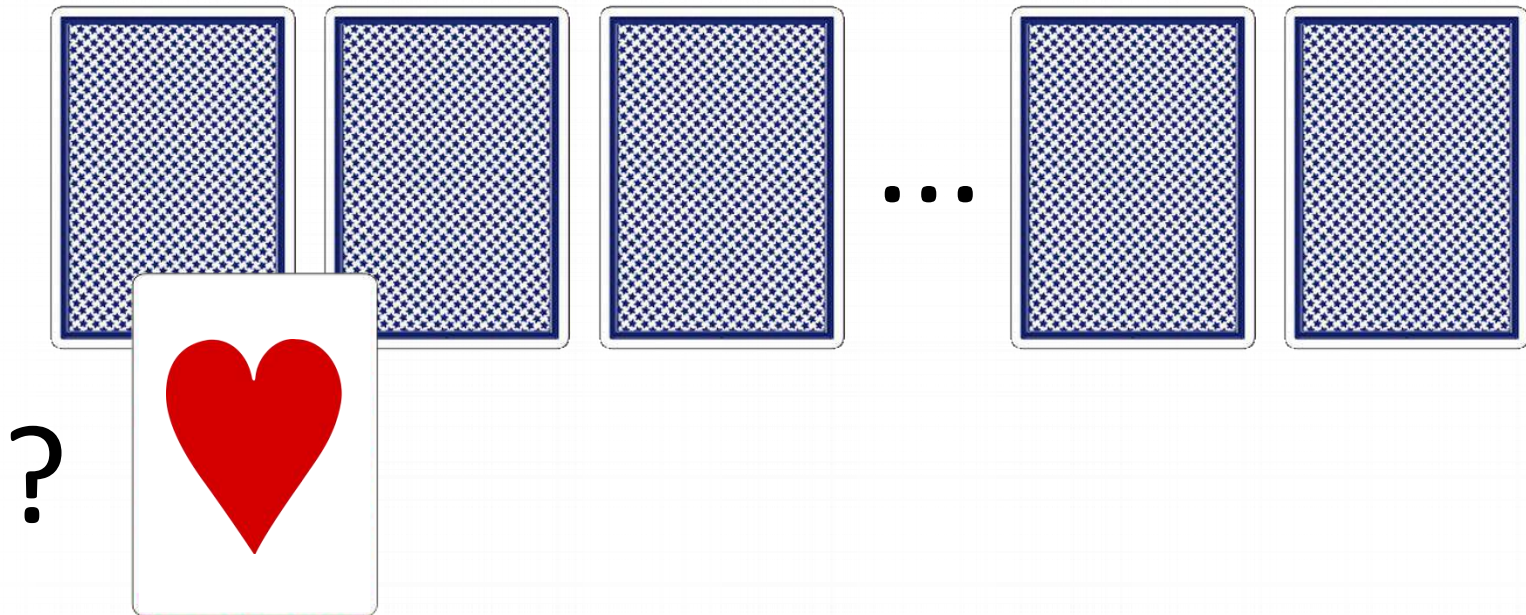
Probability that Card1 is Hearts?

[Van den Broeck; AAAI-KRR'15]

A Simple Reasoning Problem

?

Probability that Card1 is Hearts?   1/4
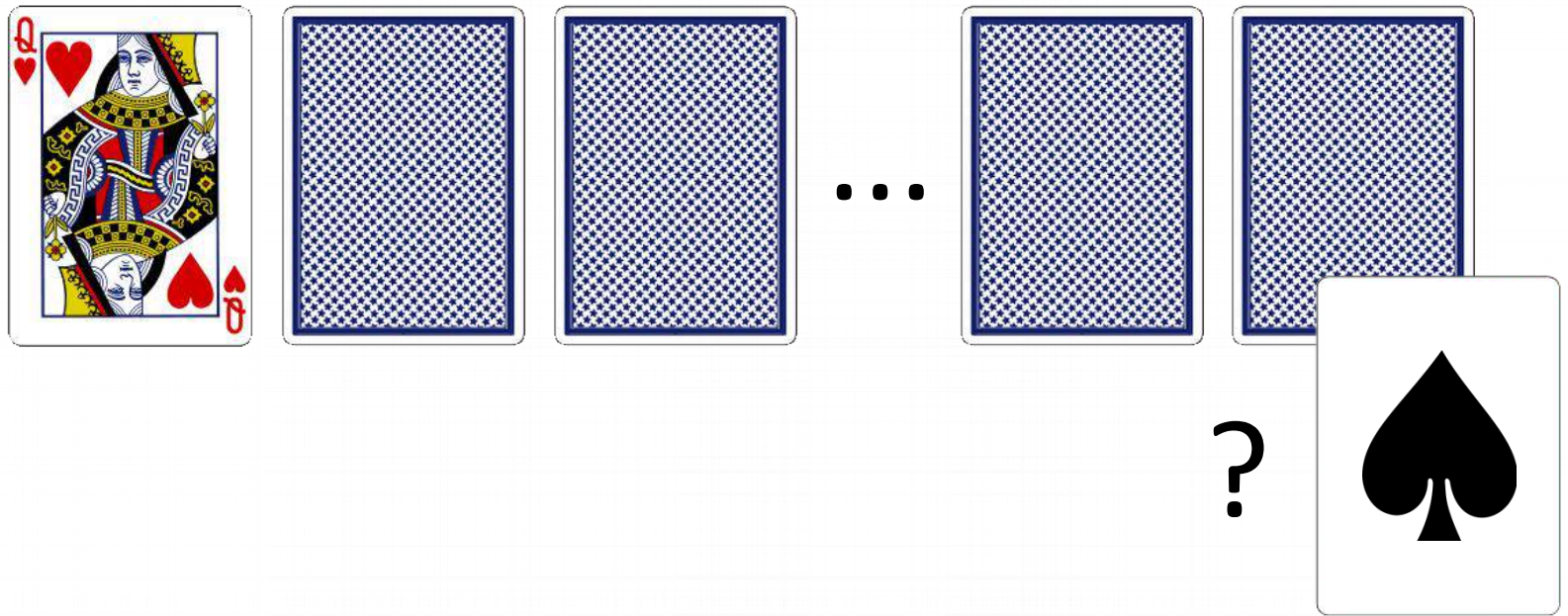
[Van den Broeck; AAAI-KRR'15]

A Simple Reasoning Problem

*Probability that Card52 is Spades
given that Card1 is QH?*

[Van den Broeck; AAAI-KRR'15]

A Simple Reasoning Problem
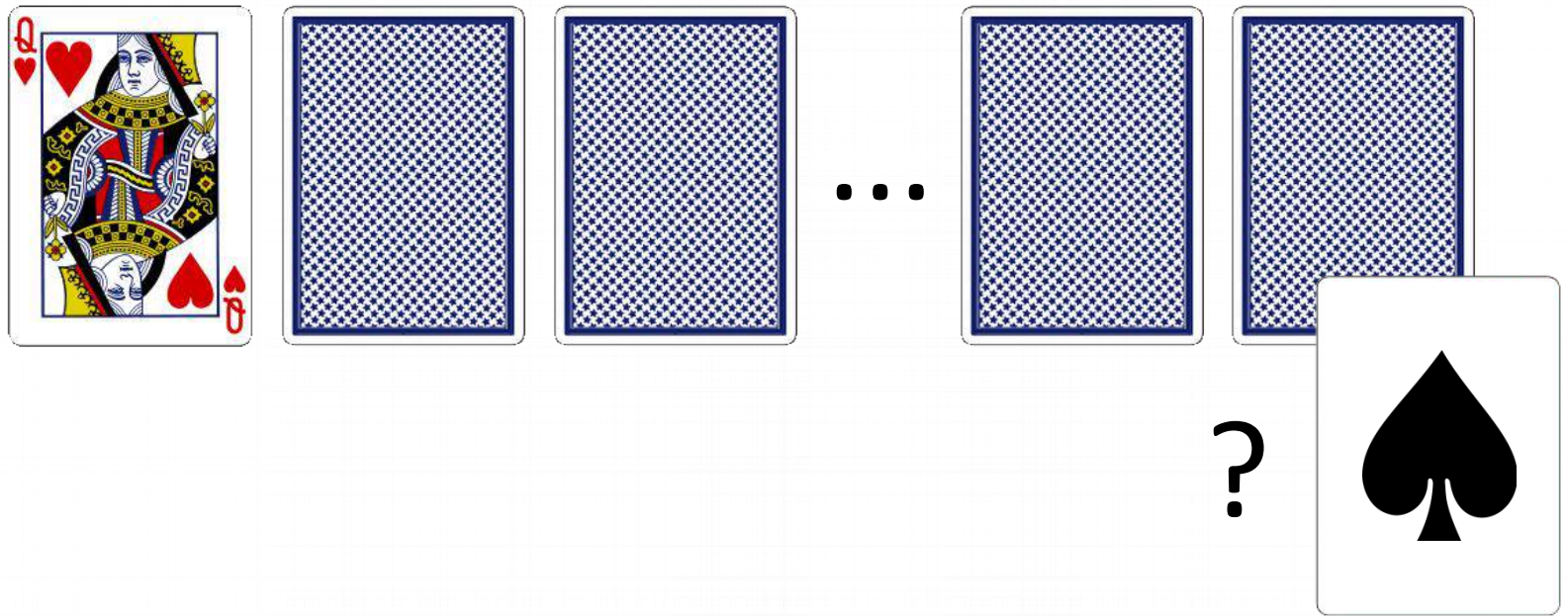
Probability that Card52 is Spades given that Card1 is QH?

13/51

Automated Reasoning

Let us automate this:

1. Probabilistic graphical model (e.g., factor graph)

2. Probabilistic inference algorithm
   (e.g., variable elimination or junction tree)

[Van den Broeck; AAAI-KRR'15]
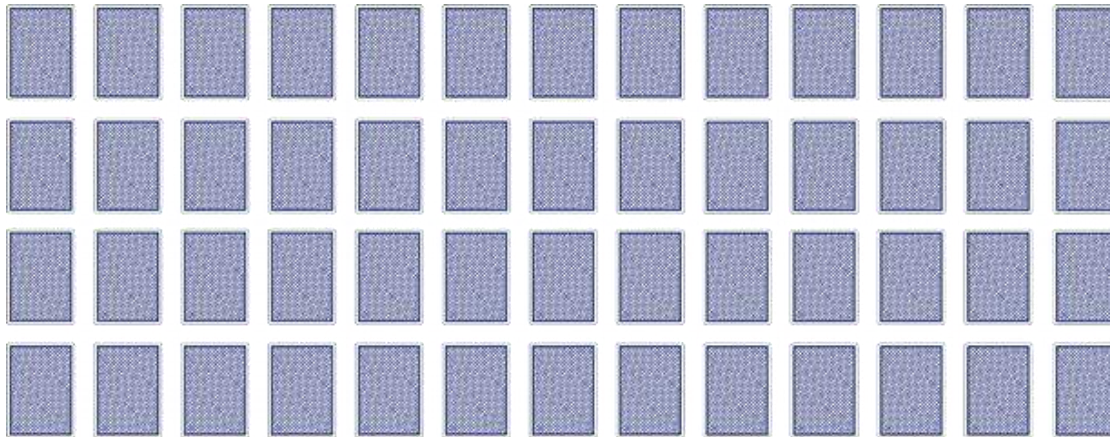
Automated Reasoning

Let us automate this:

1. Probabilistic graphical model (e.g., factor graph)
   is fully connected!

(artist's impression)
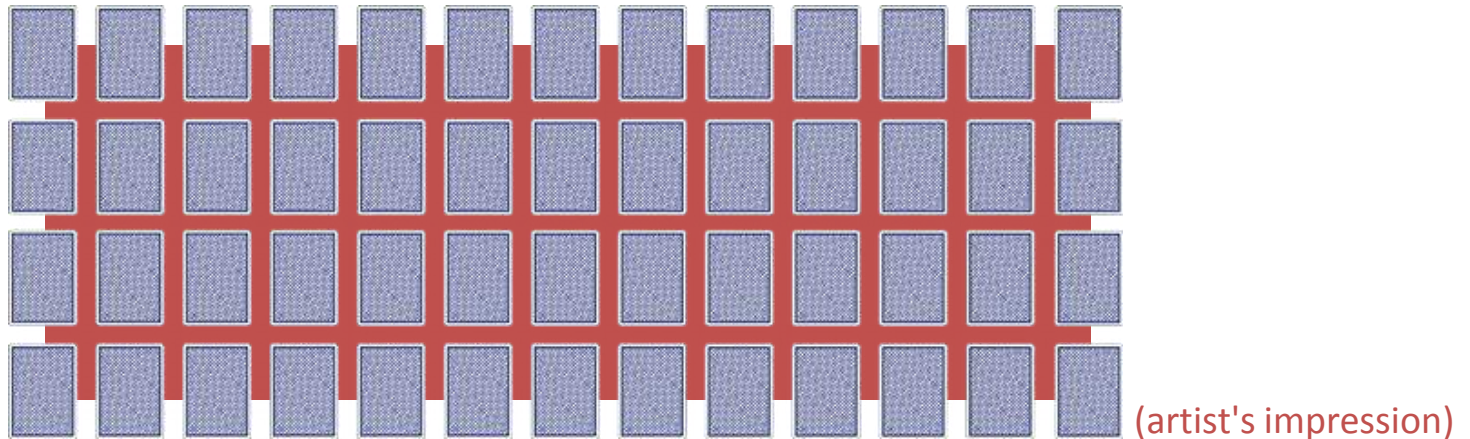
2. Probabilistic inference algorithm
   (e.g., variable elimination or junction tree)
   builds a table with **52^{52}** rows

[Van den Broeck; AAAI-KRR'15]

Statistical Relational Learning

- Statistical relational model (e.g., MLN)

  > 3.14 FacultyPage(x) ∧ Linked(x,y) ⇒ CoursePage(y)

- As a probabilistic graphical model:
  - 26 pages; 728 variables; 676 factors
  - 1000 pages; 1,002,000 variables; 1,000,000 factors

- Highly intractable?
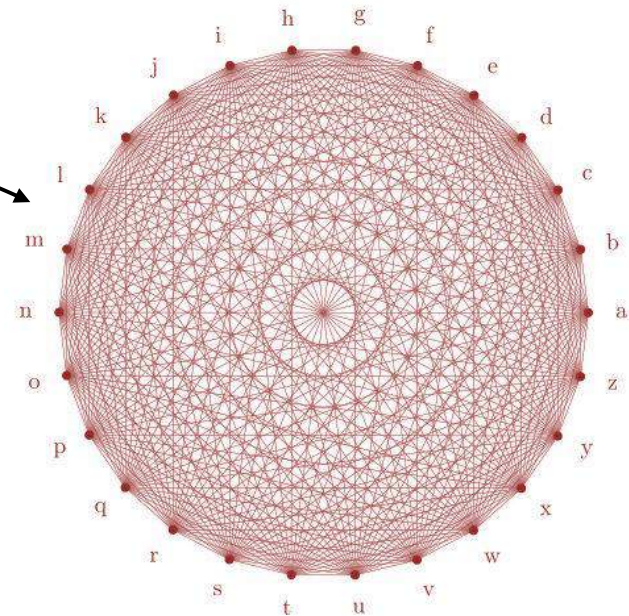
What's Going On Here?
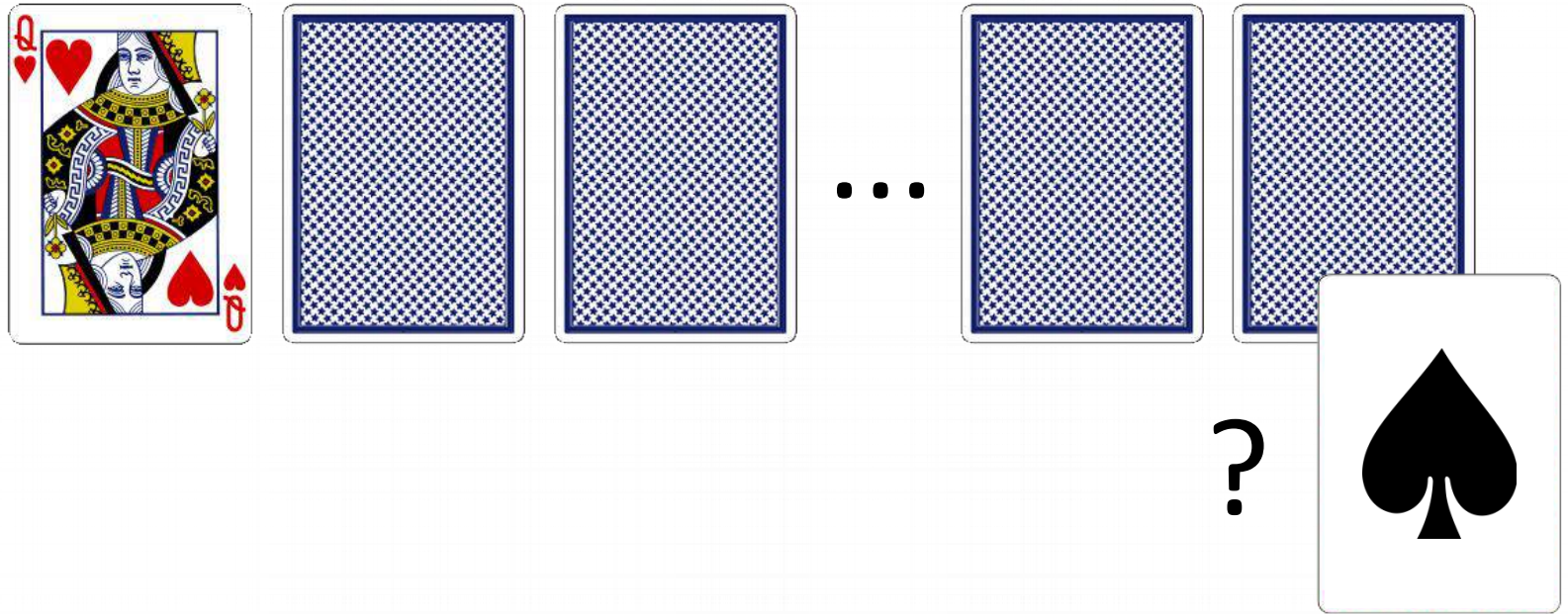
*Probability that Card52 is Spades
given that Card1 is QH?*

What's Going On Here?
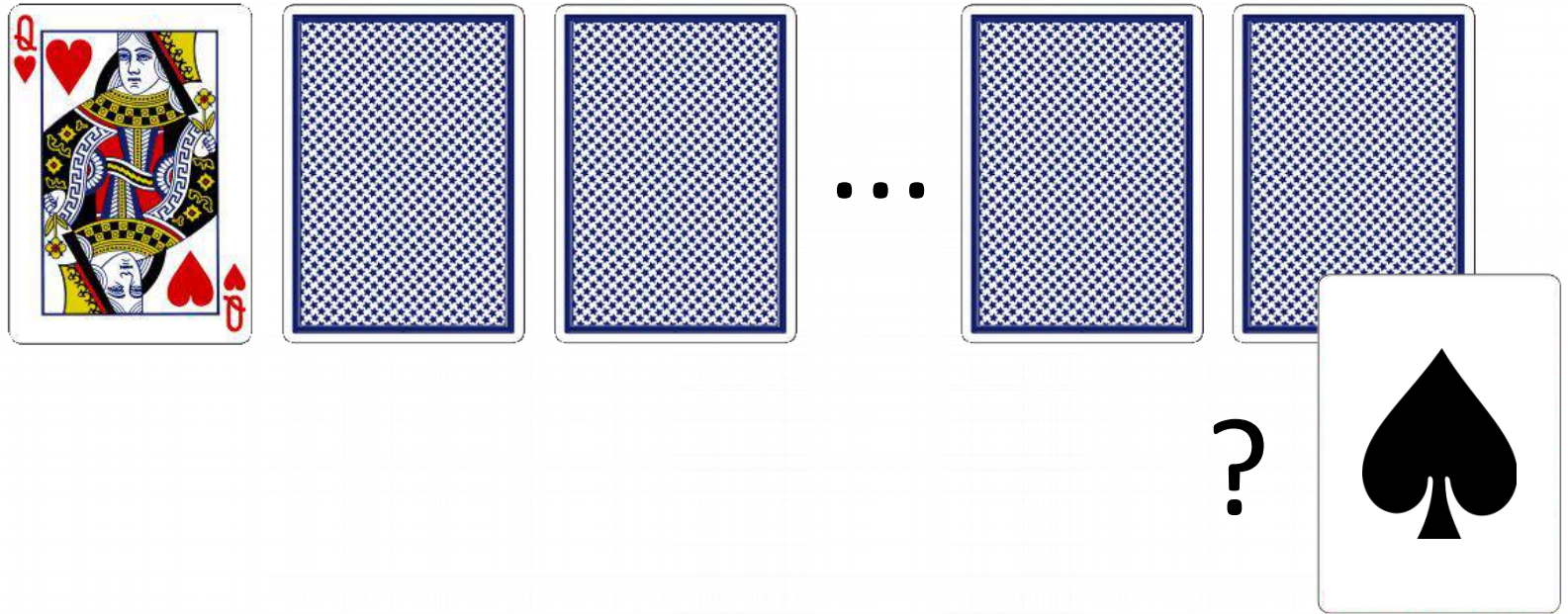
Probability that Card52 is Spades given that Card1 is QH?   13/51

[Van den Broeck; AAAI-KRR'15]
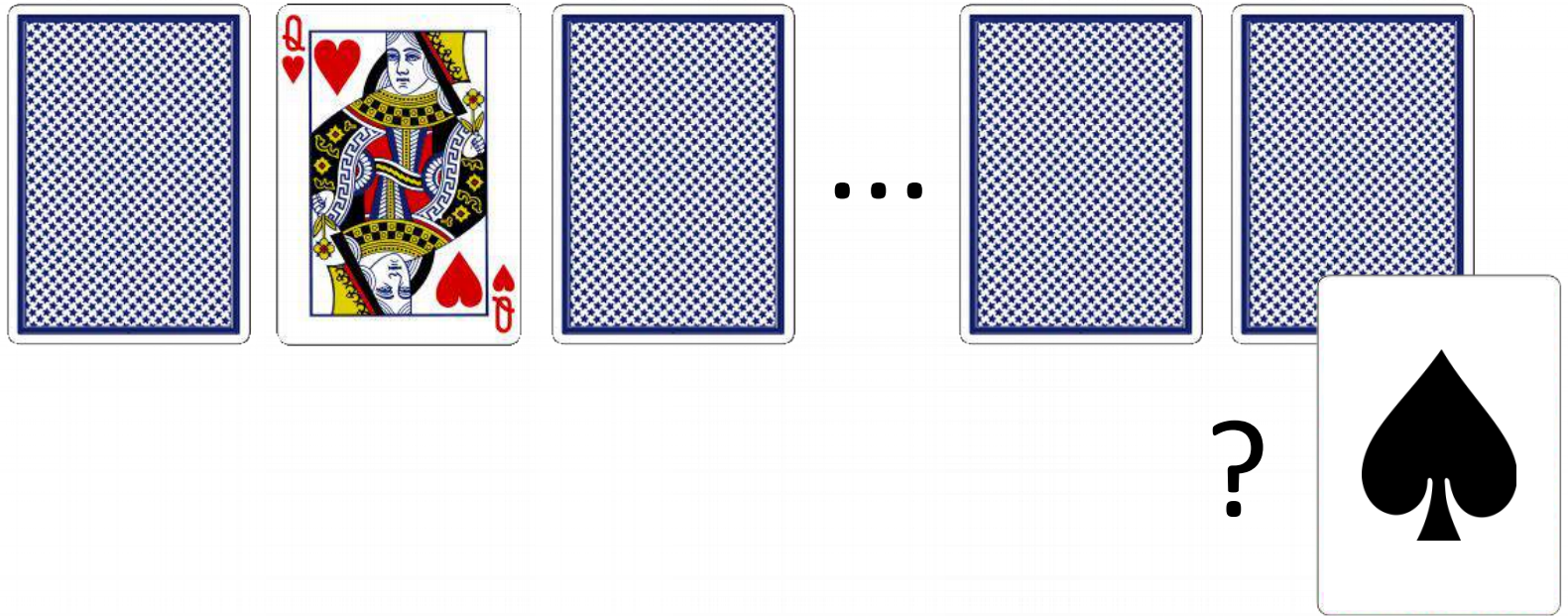
What's Going On Here?

Probability that Card52 is Spades given that Card2 is QH?

What's Going On Here?
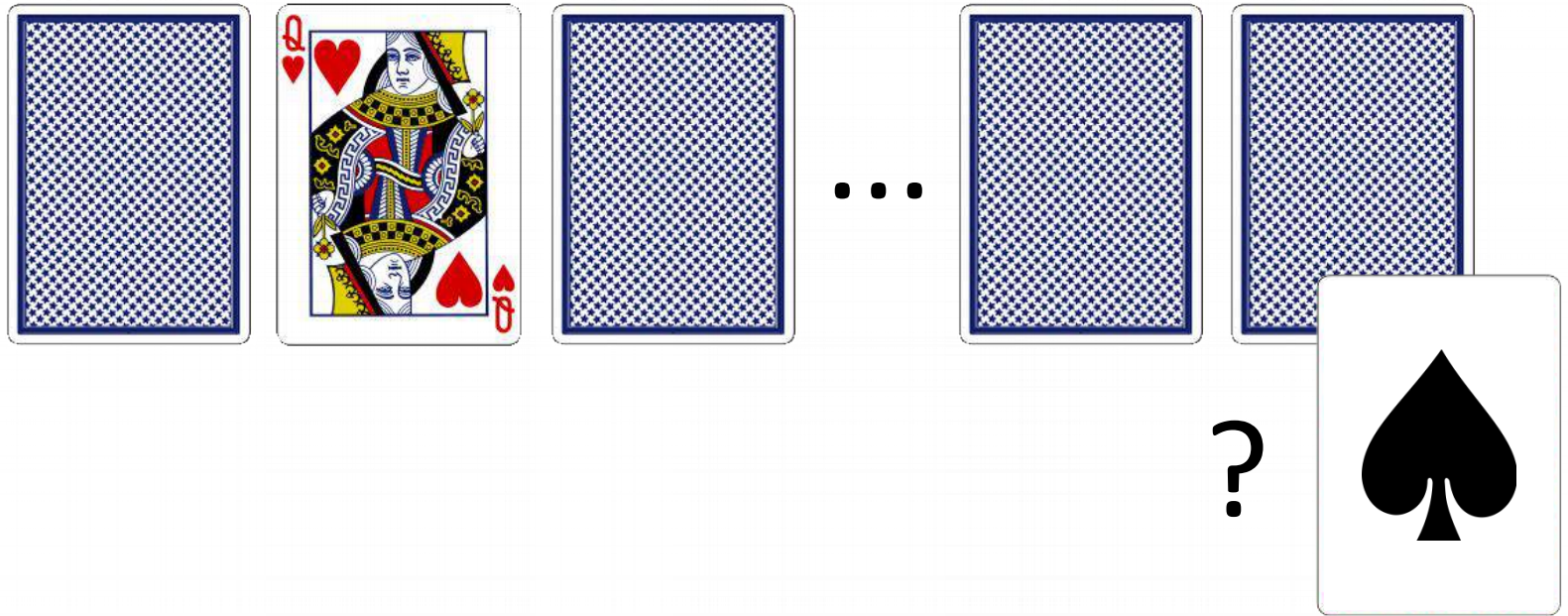
Probability that Card52 is Spades given that Card2 is QH?

13/51
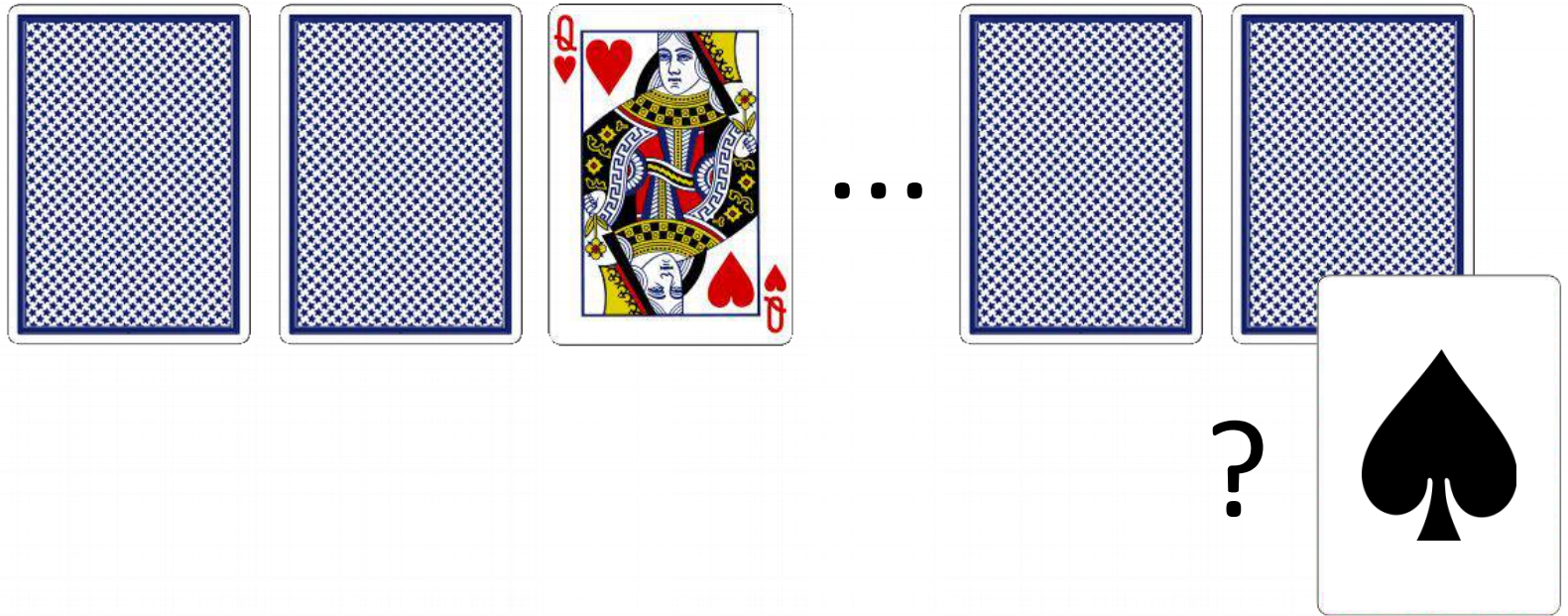
[Van den Broeck; AAAI-KRR'15]

What's Going On Here?

Probability that Card52 is Spades given that Card3 is QH?

[Van den Broeck; AAAI-KRR'15]

What's Going On Here?

Probability that Card52 is Spades given that Card3 is QH?

13/51

[Van den Broeck; AAAI-KRR'15]

Tractable Reasoning

What's going on here?

Which property makes reasoning tractable?

[Niepert and Van den Broeck, AAAI' 14], [Van den Broeck, AAAI-KRR'15]
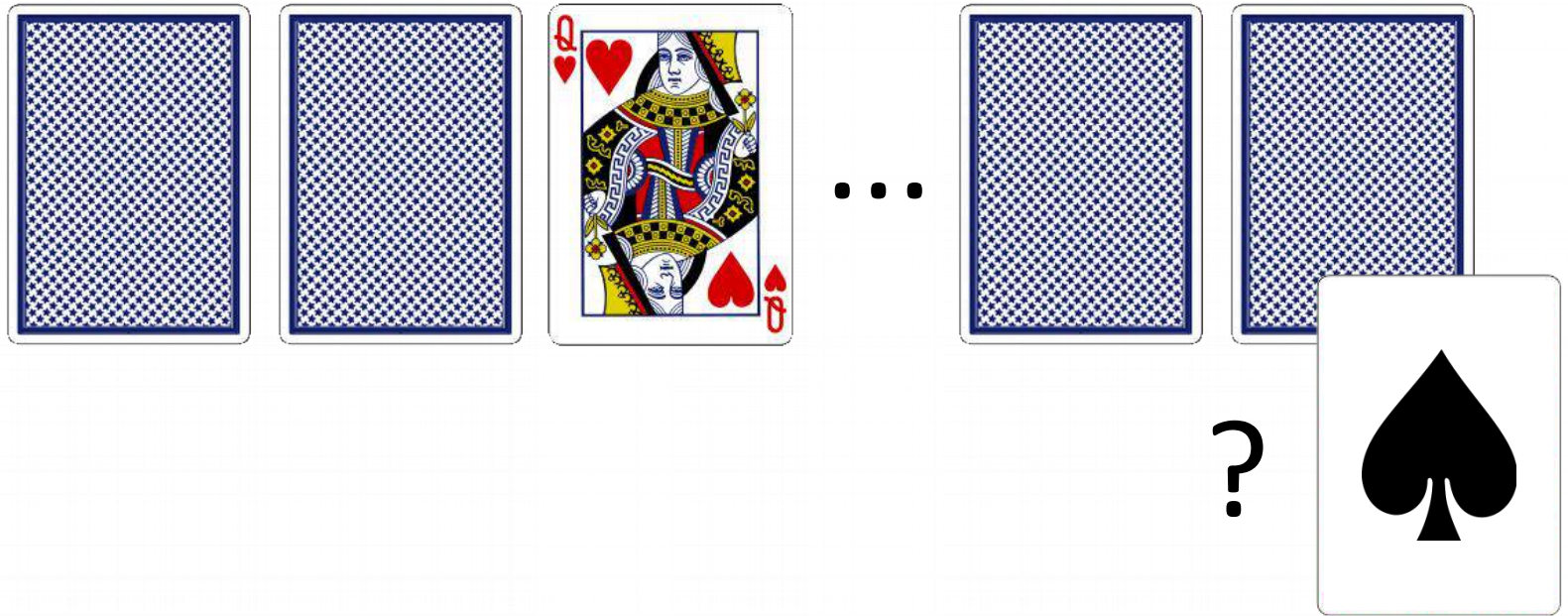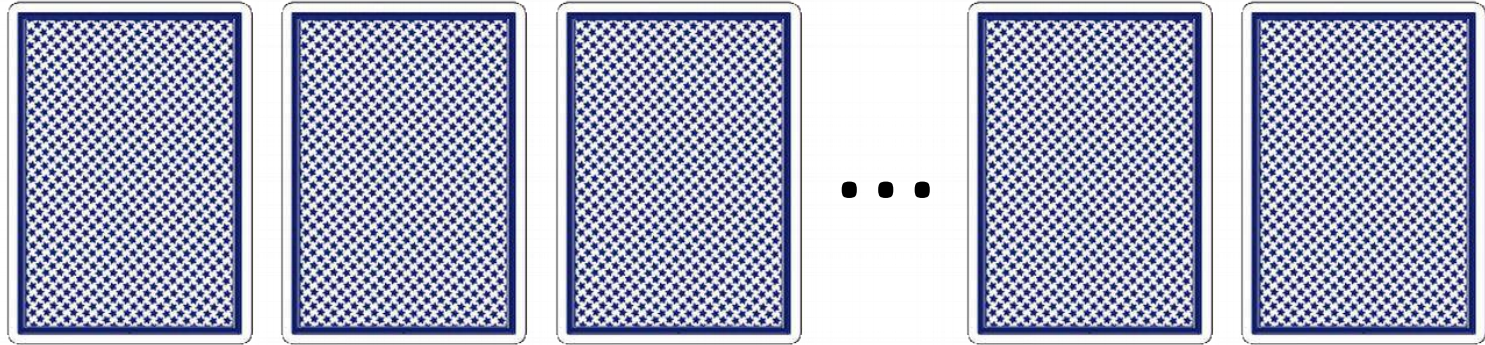
Tractable Reasoning
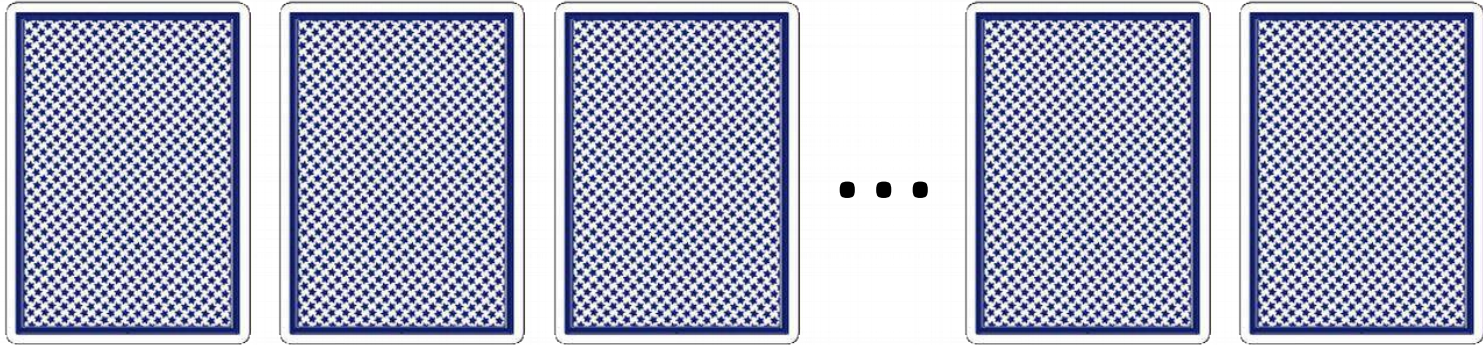
What's going on here?

Which property makes reasoning tractable?

- High-level (first-order) reasoning
- Relational structure
- Symmetry
- Exchangeability

⇒ **Lifted Inference**

[Niepert and Van den Broeck, AAAI' 14], [Van den Broeck, AAAI-KRR'15]

*How does lifted inference **work**?*

First-Order Model Counting

Model = solution to
first-order logic
formula Δ

Δ = ∀d (Rain(d)
        ⇒ Cloudy(d))

Days = {Monday}

First-Order Model Counting

Model = solution to
first-order logic
formula Δ

Δ = ∀d (Rain(d)
        ⇒ Cloudy(d))

Days = {Monday}

Rain(M)	Cloudy(M)	Model?
T	T	Yes
T	F	No
F	T	Yes
F	F	Yes

+ ——————

FOMC = 3

First-Order Model Counting

Model = solution to
first-order logic
formula Δ

Δ = ∀d (Rain(d)
⇒ Cloudy(d))

Days = {Monday
Tuesday}

First-Order Model Counting

Model = solution to first-order logic formula Δ

Δ = ∀d (Rain(d) ⇒ Cloudy(d))

Days = {Monday **Tuesday**}

Rain(M)	Cloudy(M)	Rain(T)	Cloudy(T)	Model?
T	T	T	T	Yes
T	F	T	T	No
F	T	T	T	Yes
F	F	T	T	Yes
T	T	T	F	No
T	F	T	F	No
F	T	T	F	No
F	F	T	F	No
T	T	F	T	Yes
T	F	F	T	No
F	T	F	T	Yes
F	F	F	T	Yes
T	T	F	F	Yes
T	F	F	F	No
F	T	F	F	Yes
F	F	F	F	Yes

+ _____

#SAT = 9

Weighted First-Order Model Counting

Model = solution to first-order logic formula Δ

Δ = ∀d (Rain(d)
    ⇒ Cloudy(d))

Days = {Monday
    Tuesday}

Rain

d	w(R(d))	w(¬R(d))
M	1	2
T	4	1

Cloudy

d	w(C(d))	w(¬C(d))
M	3	5
T	6	2

Rain(M)	Cloudy(M)	Rain(T)	Cloudy(T)	Model?
T	T	T	T	Yes
T	F	T	T	No
F	T	T	T	Yes
F	F	T	T	Yes
T	T	T	F	No
T	F	T	F	No
F	T	T	F	No
F	F	T	F	No
T	T	F	T	Yes
T	F	F	T	No
F	T	F	T	Yes
F	F	F	T	Yes
T	T	F	F	Yes
T	F	F	F	No
F	T	F	F	Yes
F	F	F	F	Yes

+ ——————

#SAT = 9

[VdB'11, Gogate'11]

Weighted First-Order Model Counting

Model = solution to first-order logic formula Δ

Δ = ∀d (Rain(d)
⇒ Cloudy(d))

Days = {Monday
Tuesday}

Rain

d	w(R(d))	w(¬R(d))
M	1	2
T	4	1

Cloudy

d	w(C(d))	w(¬C(d))
M	3	5
T	6	2

Rain(M)	Cloudy(M)	Rain(T)	Cloudy(T)	Model?	Weight
T	T	T	T	Yes	1 * 3 * 4 * 6 = 72
T	F	T	T	No	0
F	T	T	T	Yes	2 * 3 * 4 * 6 = 144
F	F	T	T	Yes	2 * 5 * 4 * 6 = 240
T	T	T	F	No	0
T	F	T	F	No	0
F	T	T	F	No	0
F	F	T	F	No	0
T	T	F	T	Yes	1 * 3 * 1 * 6 = 18
T	F	F	T	No	0
F	T	F	T	Yes	2 * 3 * 1 * 6 = 36
F	F	F	T	Yes	2 * 5 * 1 * 6 = 60
T	T	F	F	Yes	1 * 3 * 1 * 2 = 6
T	F	F	F	No	0
F	T	F	F	Yes	2 * 3 * 1 * 2 = 12
F	F	F	F	Yes	2 * 5 * 1 * 2 = 20

+ ——

#SAT = 9

[VdB'11, Gogate'11]

Weighted First-Order Model Counting

Model = solution to first-order logic formula Δ

Δ = ∀d (Rain(d)
⇒ Cloudy(d))

Days = {Monday
Tuesday}

Rain

d	w(R(d))	w(¬R(d))
M	1	2
T	4	1

Cloudy

d	w(C(d))	w(¬C(d))
M	3	5
T	6	2

Rain(M)	Cloudy(M)	Rain(T)	Cloudy(T)	Model?	Weight
T	T	T	T	Yes	1 * 3 * 4 * 6 = 72
T	F	T	T	No	0
F	T	T	T	Yes	2 * 3 * 4 * 6 = 144
F	F	T	T	Yes	2 * 5 * 4 * 6 = 240
T	T	T	F	No	0
T	F	T	F	No	0
F	T	T	F	No	0
F	F	T	F	No	0
T	T	F	T	Yes	1 * 3 * 1 * 6 = 18
T	F	F	T	No	0
F	T	F	T	Yes	2 * 3 * 1 * 6 = 36
F	F	F	T	Yes	2 * 5 * 1 * 6 = 60
T	T	F	F	Yes	1 * 3 * 1 * 2 = 6
T	F	F	F	No	0
F	T	F	F	Yes	2 * 3 * 1 * 2 = 12
F	F	F	F	Yes	2 * 5 * 1 * 2 = 20

+ ——— **#SAT = 9** + ——— **WFOMC = 608**

[VdB'11, Gogate'11]

Probabilistic Database Rules

$$P(\neg Q) = 1 - P(Q)$$

Negation

$$P(Q1 \wedge Q2) = P(Q1)\, P(Q2)$$
$$P(Q1 \vee Q2) = 1 - (1 - P(Q1))\, (1 - P(Q2))$$

Decomposable $\wedge, \vee$

$$P(\exists z\, Q) = 1 - \Pi_{A\,\in Domain}\, (1 - P(Q[A/z]))$$
$$P(\forall z\, Q) = \Pi_{A\,\in Domain}\, P(Q[A/z])$$

Decomposable $\exists, \forall$

$$P(Q1 \wedge Q2) = P(Q1) + P(Q2) - P(Q1 \vee Q2)$$
$$P(Q1 \vee Q2) = P(Q1) + P(Q2) - P(Q1 \wedge Q2)$$

Inclusion/exclusion

[Suciu'11]

Asymmetric WFOMC Rules

$$\text{WMC}(\neg\Delta) = Z\text{-WMC}(\Delta)$$

Normalization constants Z
(easy to compute)

Negation

$$\text{WMC}(\Delta_1 \wedge \Delta_2) = \text{WMC}(\Delta_1) * \text{WMC}(\Delta_2)$$
$$\text{WMC}(\Delta_1 \vee \Delta_2) = Z\text{-}(Z_1\text{-WMC}(\Delta_1))*(Z_2\text{-WMC}(\Delta_2))$$

Decomposable $\wedge,\vee$

$$\text{WMC}(\exists z\, \Delta) = Z - \Pi_{C\in\text{Domain}}\, (Z_C - \text{WMC}(\Delta[C/z])$$
$$\text{WMC}(\forall z\, \Delta) = \Pi_{C\in\text{Domain}}\, \text{WMC}(\Delta[C/z]$$

Decomposable $\exists,\forall$

$$\text{WMC}(\Delta_1 \wedge \Delta_2) = \text{WMC}(\Delta_1)+\text{WMC}(\Delta_2)\text{-WMC}(\Delta_1 \vee \Delta_2)$$
$$\text{WMC}(\Delta_1 \vee \Delta_2) = \text{WMC}(\Delta_1)+\text{WMC}(\Delta_2)\text{-WMC}(\Delta_1 \wedge \Delta_2)$$

Inclusion/
exclusion

Example MLN to WFOMC

3.75   Smoker(x) ∧ Friend(x,y) ⇒ Smoker(y)

1. **Formula Δ**

2. **Weight function w(.)**

Example MLN to WFOMC

3.75   Smoker(x) ∧ Friend(x,y) ⇒ Smoker(y)

1. <u>Formula Δ</u>

Δ = ∀x∀y (F(x,y) ⇔ [Smoker(x) ∧ Friend(x,y) ⇒ Smoker(y)])

2. <u>Weight function w(.)</u>

Example MLN to WFOMC

3.75   Smoker(x) ∧ Friend(x,y) ⇒ Smoker(y)

1. Formula Δ

Δ = ∀x∀y (F(x,y) ⇔ [Smoker(x) ∧ Friend(x,y) ⇒ Smoker(y)])

2. Weight function w(.)

F

x	y	w(F(x,y))	w(¬F(x,y))
A	A	exp(3.75)	1
A	B	exp(3.75)	1
A	C	exp(3.75)	1
B	A	exp(3.75)	1
	. . .	. . .	

Example MLN to WFOMC

3.75   Smoker(x) ∧ Friend(x,y) ⇒ Smoker(y)

1. Formula Δ

Δ = ∀x∀y (F(x,y) ⇔ [Smoker(x) ∧ Friend(x,y) ⇒ Smoker(y)])

2. Weight function w(.)

F

x	y	w(F(x,y))	w(¬F(x,y))
A	A	exp(3.75)	1
A	B	exp(3.75)	1
A	C	exp(3.75)	1
B	A	exp(3.75)	1
	. . .	. . .	

Z = WFOMC(Δ)

Symmetric WFOMC Rules

- Simplification to decomposable $\exists, \forall$ rules:

If $\Delta[C_1/x], \Delta[C_2/x], \ldots$ are independent

$$\mathrm{WMC}(\exists z\ \Delta) = Z - (Z_{C_1} - \mathrm{WMC}(\Delta[C_1/z])^{|Domain|}$$

$$\mathrm{WMC}(\forall z\ \Delta) = \mathrm{WMC}(\Delta[C_1/z])^{|Domain|}$$

Symmetric WFOMC Rules
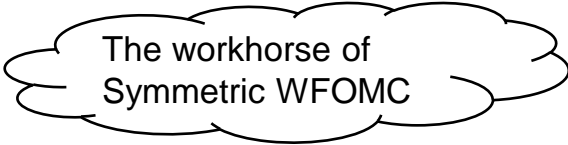
- Simplification to decomposable $\exists, \forall$ rules:

  If $\Delta[C_1/x]$, $\Delta[C_2/x]$, … are independent

  $$WMC(\exists z\ \Delta) = Z - (Z_{C_1} - WMC(\Delta[C_1/z])^{|Domain|}$$
  $$WMC(\forall z\ \Delta) = \ WMC(\Delta[C_1/z])^{|Domain|}$$

- A powerful new inference rule: *atom counting*
  Only possible with symmetric weights
  Intuition: **Remove unary relations**

  The workhorse of
  Symmetric WFOMC

FOMC Inference

$\Delta = \forall x,y, (\text{Smokes}(x) \land \text{Friends}(x,y) \Rightarrow \text{Smokes}(y))$

Domain = {n people}

FOMC Inference

Domain = {n people}

- If we know precisely who smokes, and there are k smokers?

Database:

Smokes(Alice) = 1
Smokes(Bob) = 0
Smokes(Charlie) = 0
Smokes(Dave) = 1
Smokes(Eve) = 0
...

Smokes   Friends   Smokes

k   k

n-k   n-k

FOMC Inference

$\Delta = \forall x,y, (\text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y))$

Domain = {n people}

- If we know precisely who smokes, and there are k smokers?

Database:

Smokes(Alice) = 1
Smokes(Bob) = 0
Smokes(Charlie) = 0
Smokes(Dave) = 1
Smokes(Eve) = 0
...

Smokes   Friends   Smokes

| k | → | k |
| n-k | | n-k |

FOMC Inference

Domain = {n people}

- If we know precisely who smokes, and there are k smokers?

Database:
Smokes(Alice) = 1
Smokes(Bob) = 0
Smokes(Charlie) = 0
Smokes(Dave) = 1
Smokes(Eve) = 0
...

FOMC Inference

$\Delta = \forall x,y, (Smokes(x) \wedge Friends(x,y) \Rightarrow Smokes(y))$

Domain = {n people}

- If we know precisely who smokes, and there are *k* smokers?

Database:

Smokes(Alice) = 1
Smokes(Bob) = 0
Smokes(Charlie) = 0
Smokes(Dave) = 1
Smokes(Eve) = 0
...

Smokes   Friends   Smokes

k   k

n-k   n-k

FOMC Inference

$\Delta = \forall x,y, (Smokes(x) \land Friends(x,y) \Rightarrow Smokes(y))$

Domain = {n people}

- If we know precisely who smokes, and there are *k* smokers?

Database:

Smokes(Alice) = 1
Smokes(Bob) = 0
Smokes(Charlie) = 0
Smokes(Dave) = 1
Smokes(Eve) = 0
...

FOMC Inference

$\Delta = \forall x,y, (\text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y))$

Domain = {n people}

- If we know precisely who smokes, and there are *k* smokers?

Database:
Smokes(Alice) = 1
Smokes(Bob) = 0
Smokes(Charlie) = 0
Smokes(Dave) = 1
Smokes(Eve) = 0
…

FOMC Inference

Domain = {n people}

- If we know precisely who smokes, and there are k smokers?

Database:

Smokes(Alice) = 1
Smokes(Bob) = 0
Smokes(Charlie) = 0
Smokes(Dave) = 1
Smokes(Eve) = 0
...

FOMC Inference

$\Delta = \forall x,y, (\text{Smokes}(x) \land \text{Friends}(x,y) \Rightarrow \text{Smokes}(y))$

Domain = {n people}

- If we know precisely who smokes, and there are k smokers?

Database:

Smokes(Alice) = 1
Smokes(Bob) = 0
Smokes(Charlie) = 0
Smokes(Dave) = 1
Smokes(Eve) = 0
...

FOMC Inference

$\Delta = \forall x,y, (Smokes(x) \wedge Friends(x,y) \Rightarrow Smokes(y))$

Domain = {n people}

- If we know precisely who smokes, and there are *k* smokers?

Database:

Smokes(Alice) = 1

Smokes(Bob) = 0

Smokes(Charlie) = 0
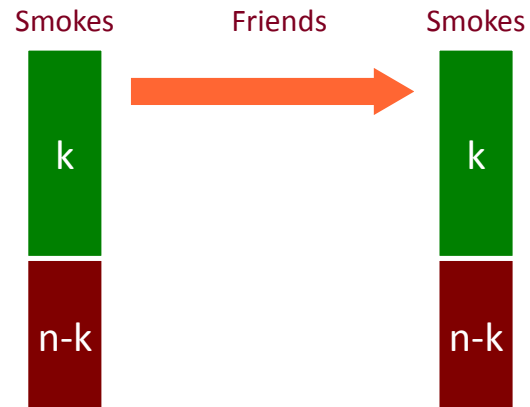
Smokes(Dave) = 1

Smokes(Eve) = 0

...

FOMC Inference

$\Delta = \forall x,y, (\text{Smokes}(x) \land \text{Friends}(x,y) \Rightarrow \text{Smokes}(y))$

Domain = {n people}

- If we know precisely who smokes, and there are *k* smokers?

Database:
Smokes(Alice) = 1
Smokes(Bob) = 0
Smokes(Charlie) = 0
Smokes(Dave) = 1
Smokes(Eve) = 0

...
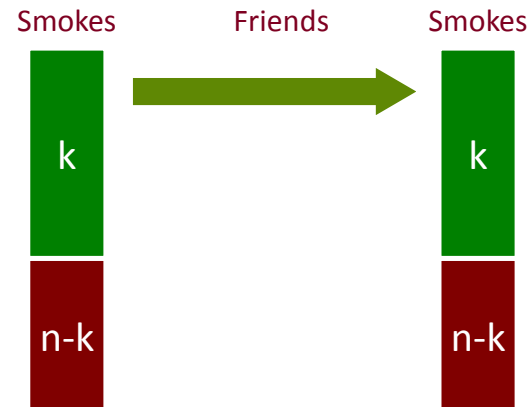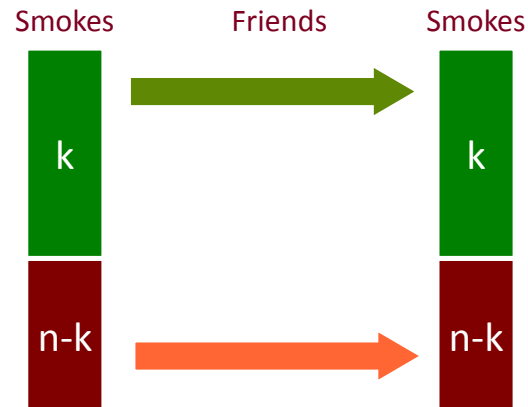
$\rightarrow 2^{n^2 - k(n-k)}$ models

FOMC Inference

$\Delta = \forall x,y, (\text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y))$

Domain = {n people}

- If we know precisely who smokes, and there are *k* smokers?

Database:

Smokes(Alice) = 1
Smokes(Bob) = 0
Smokes(Charlie) = 0
Smokes(Dave) = 1
Smokes(Eve) = 0

...

$\rightarrow 2^{n^2 - k(n-k)}$ models

Smokes   Friends   Smokes

k   k

n-k   n-k

- If we know that there are *k* smokers?

FOMC Inference

Δ = ∀x,y, (Smokes(x) ∧ Friends(x,y) ⇒ Smokes(y))

Domain = {n people}

- If we know precisely who smokes, and there are *k* smokers?

Database:
Smokes(Alice) = 1
Smokes(Bob) = 0
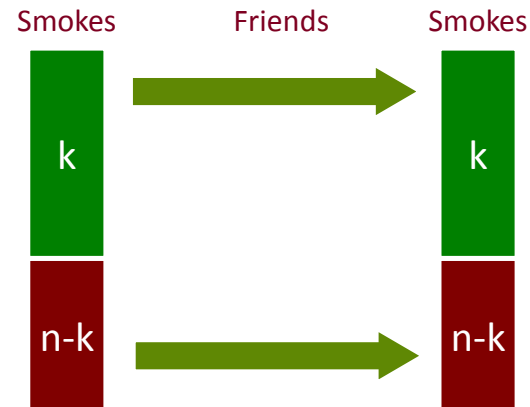Smokes(Charlie) = 0
Smokes(Dave) = 1
Smokes(Eve) = 0

...
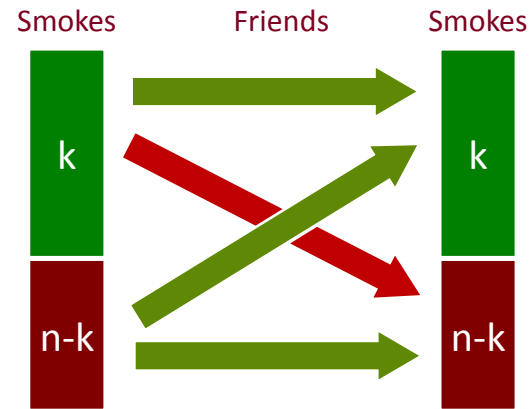
$\rightarrow 2^{n^2 - k(n-k)}$ models

- If we know that there are *k* smokers?

$\rightarrow \binom{n}{k} 2^{n^2 - k(n-k)}$ models

FOMC Inference

Δ = ∀x,y, (Smokes(x) ∧ Friends(x,y) ⇒ Smokes(y))

Domain = {n people}

- If we know precisely who smokes, and there are *k* smokers?

Database:
Smokes(Alice) = 1
Smokes(Bob) = 0
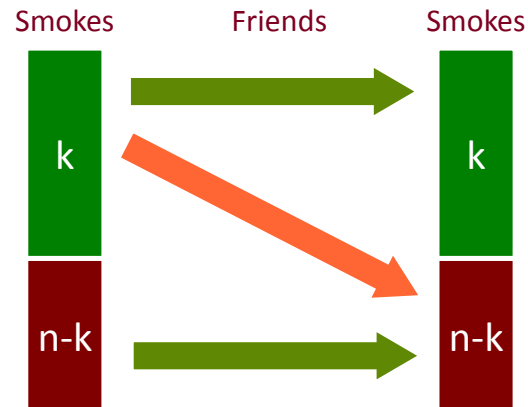Smokes(Charlie) = 0
Smokes(Dave) = 1
Smokes(Eve) = 0

...

→ $2^{n^2 - k(n-k)}$ models

- If we know that there are *k* smokers?

→ $\binom{n}{k} 2^{n^2 - k(n-k)}$ models

- In total...

FOMC Inference

$\Delta = \forall x,y, (Smokes(x) \wedge Friends(x,y) \Rightarrow Smokes(y))$

Domain = {n people}

- If we know precisely who smokes, and there are *k* smokers?

Database:
Smokes(Alice) = 1
Smokes(Bob) = 0
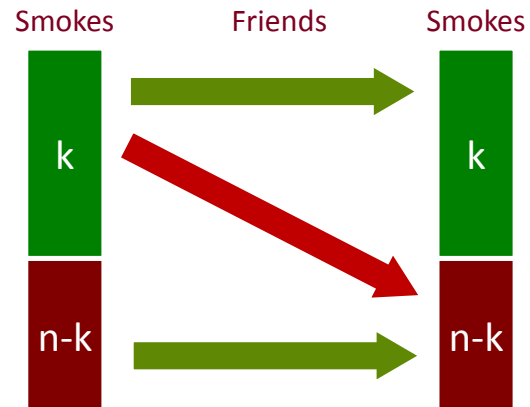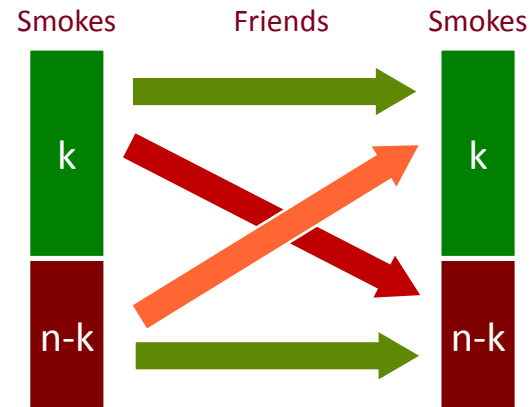Smokes(Charlie) = 0
Smokes(Dave) = 1
Smokes(Eve) = 0

...
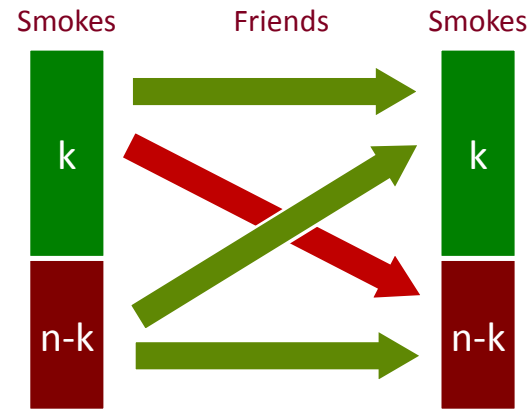
$\rightarrow 2^{n^2 - k(n-k)}$ models

- If we know that there are *k* smokers?

$\rightarrow \binom{n}{k} 2^{n^2 - k(n-k)}$ models

- In total...

$\rightarrow \sum_{k=0}^{n} \binom{n}{k} 2^{n^2 - k(n-k)}$ models

Playing Cards Revisited
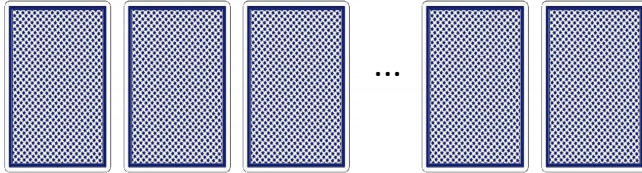
$\forall p, \exists c, Card(p,c)$
$\forall c, \exists p, Card(p,c)$
$\forall p, \forall c, \forall c', Card(p,c) \land Card(p,c') \Rightarrow c = c'$

Playing Cards Revisited
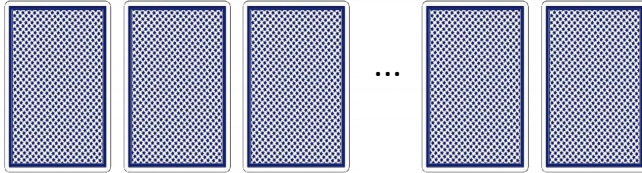
$\forall p, \exists c,\ Card(p,c)$

$\forall c, \exists p,\ Card(p,c)$

$\forall p, \forall c, \forall c',\ Card(p,c) \land Card(p,c') \Rightarrow c = c'$

$$\#SAT = \sum_{k=0}^{n} \binom{n}{k} \sum_{l=0}^{n} \binom{n}{l} (l+1)^k (-1)^{2n-k-l} = n!$$

Playing Cards Revisited
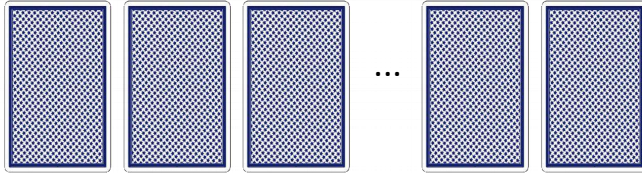
∀p, ∃c, Card(p,c)
∀c, ∃p, Card(p,c)
∀p, ∀c, ∀c', Card(p,c) ∧ Card(p,c') ⇒ c = c'

$$\#SAT = \sum_{k=0}^{n} \binom{n}{k} \sum_{l=0}^{n} \binom{n}{l} (l+1)^k (-1)^{2n-k-l} = n!$$

Computed in time polynomial in n

[Van den Broeck.; AAAI-KR'15]

When does it work?

Why not do propositional WMC?

Reduce to propositional model counting:

Why not do propositional WMC?

Reduce to propositional model counting:

Δ = Card(A♥,p_1) v … v Card(2♣,p_1)
   Card(A♥,p_2) v … v Card(2♣,p_2)
         …
   Card(A♥,p_1) v … v Card(A♥,p_{52})
   Card(K♥,p_1) v … v Card(K♥,p_{52})
         …
   ¬Card(A♥,p_1) v ¬Card(A♥,p_2)
   ¬Card(A♥,p_1) v ¬Card(A♥,p_3)
         …

Why not do propositional WMC?

Reduce to propositional model counting:

$\Delta = $ Card(A♥,p_1) v … v Card(2♣,p_1)
Card(A♥,p_2) v … v Card(2♣,p_2)
…
Card(A♥,p_1) v … v Card(A♥,p_{52})
Card(K♥,p_1) v … v Card(K♥,p_{52})
…
¬Card(A♥,p_1) v ¬Card(A♥,p_2)
¬Card(A♥,p_1) v ¬Card(A♥,p_3)
…

What will happen?

Deck of Cards Graphically

Deck of Cards Graphically

A♥

2♥

3♥

K♥
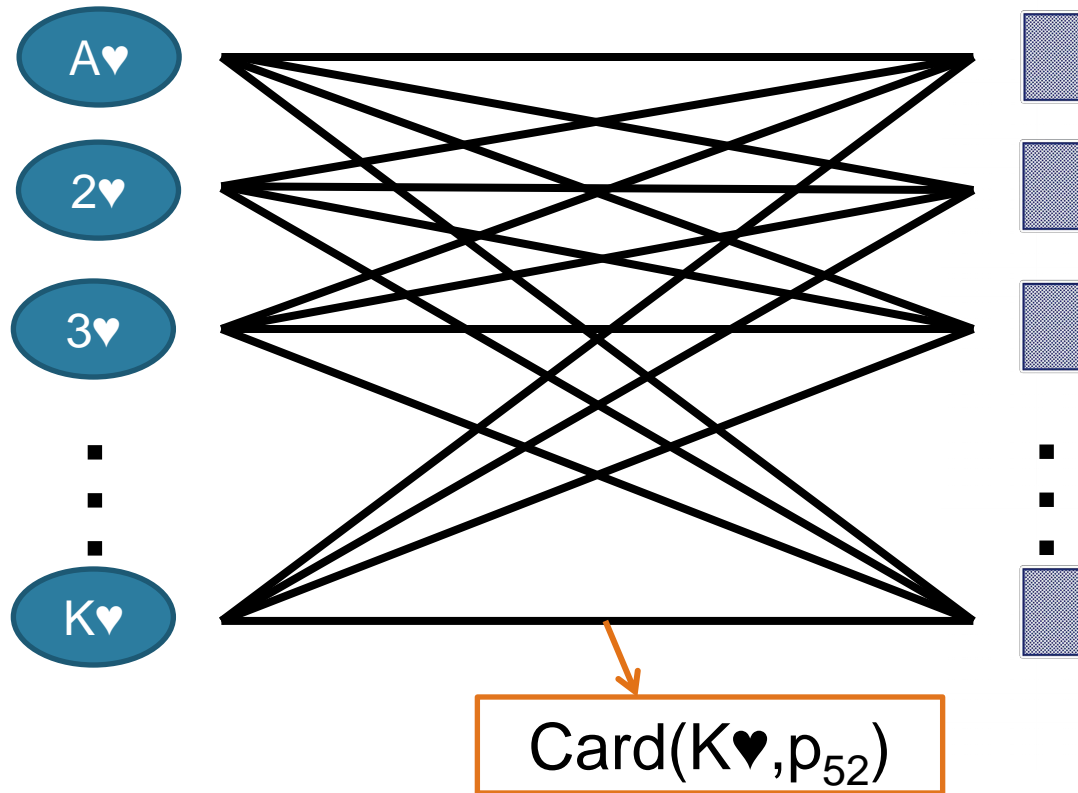
Card(K♥,p_{52})

Deck of Cards Graphically
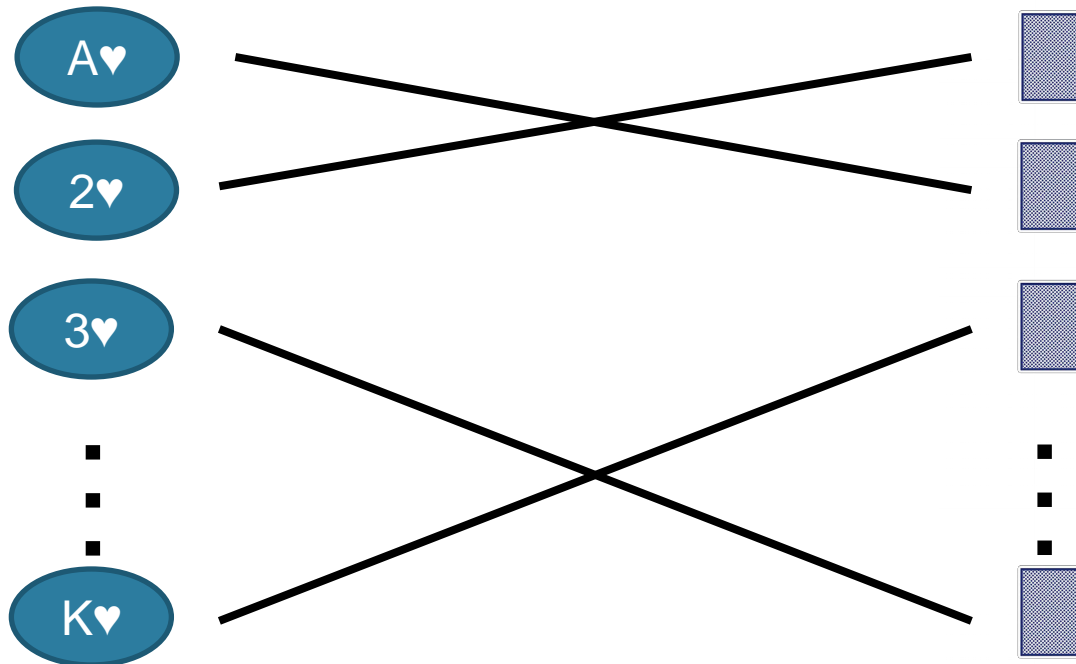
One model/*perfect matching*

Deck of Cards Graphically
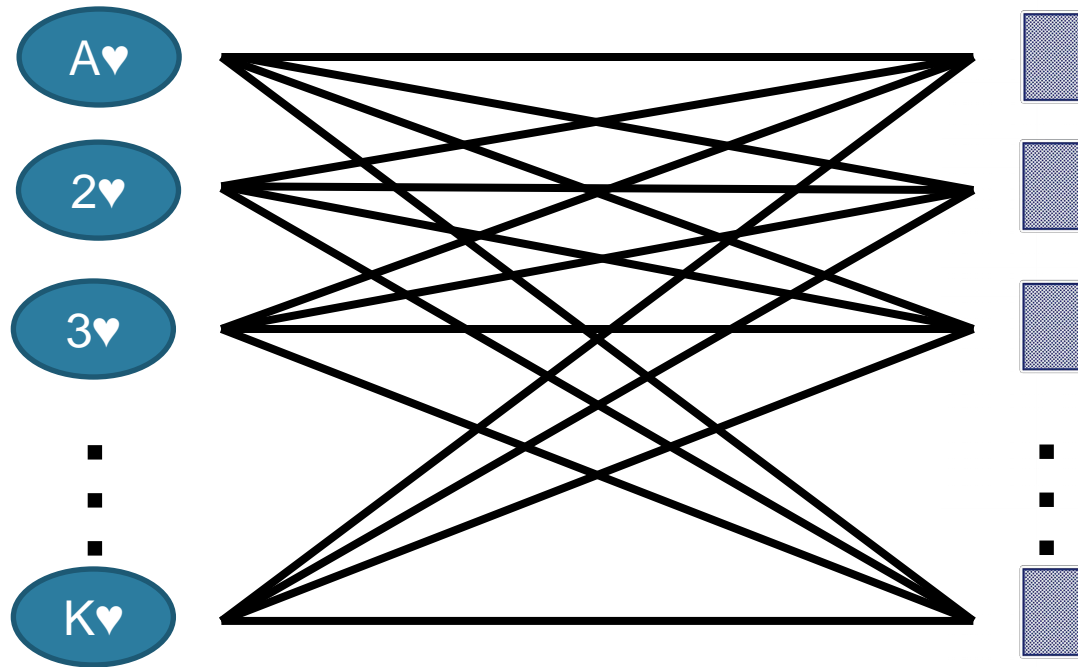
Deck of Cards Graphically
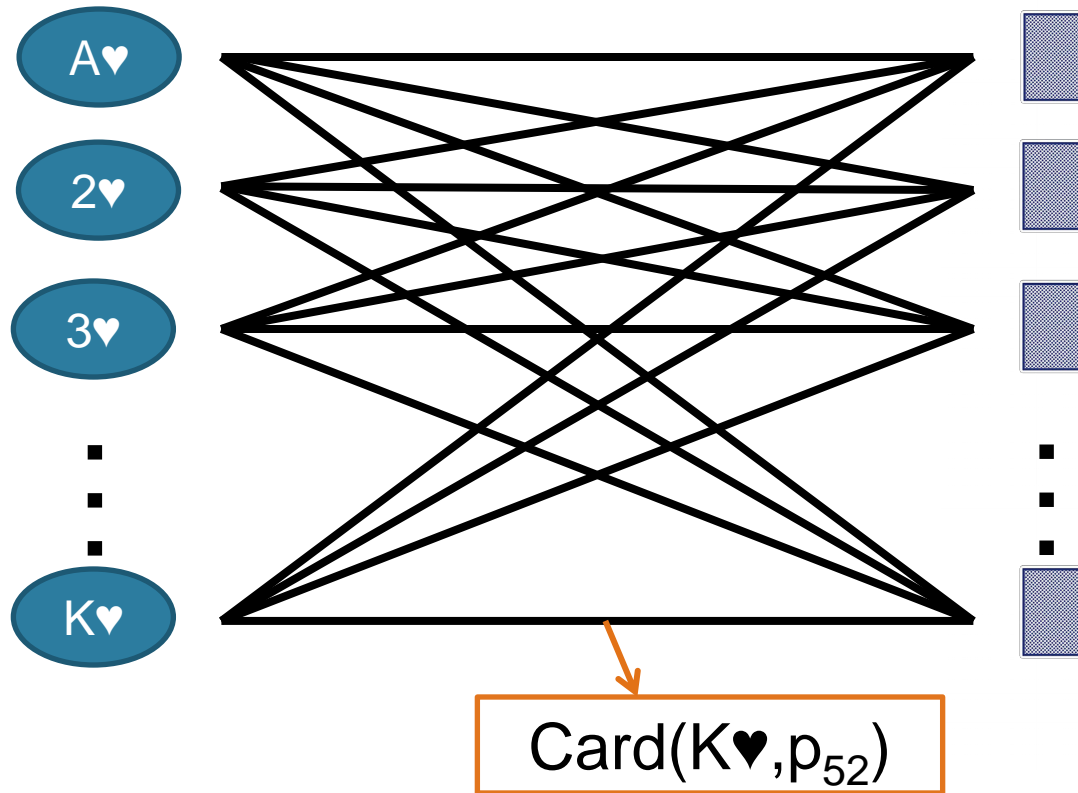
Card(K♥,p_{52})

Deck of Cards Graphically
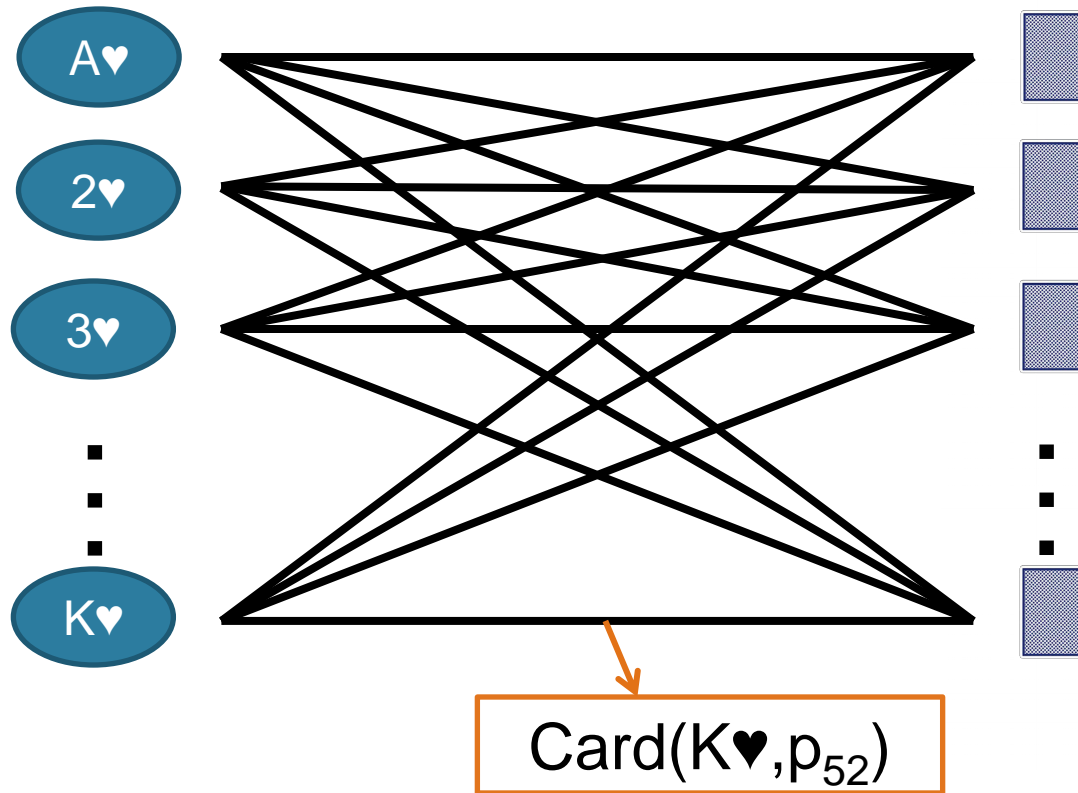
Card(K♥,p$_{52}$)

Model counting: How many *perfect matchings*?

Deck of Cards Graphically
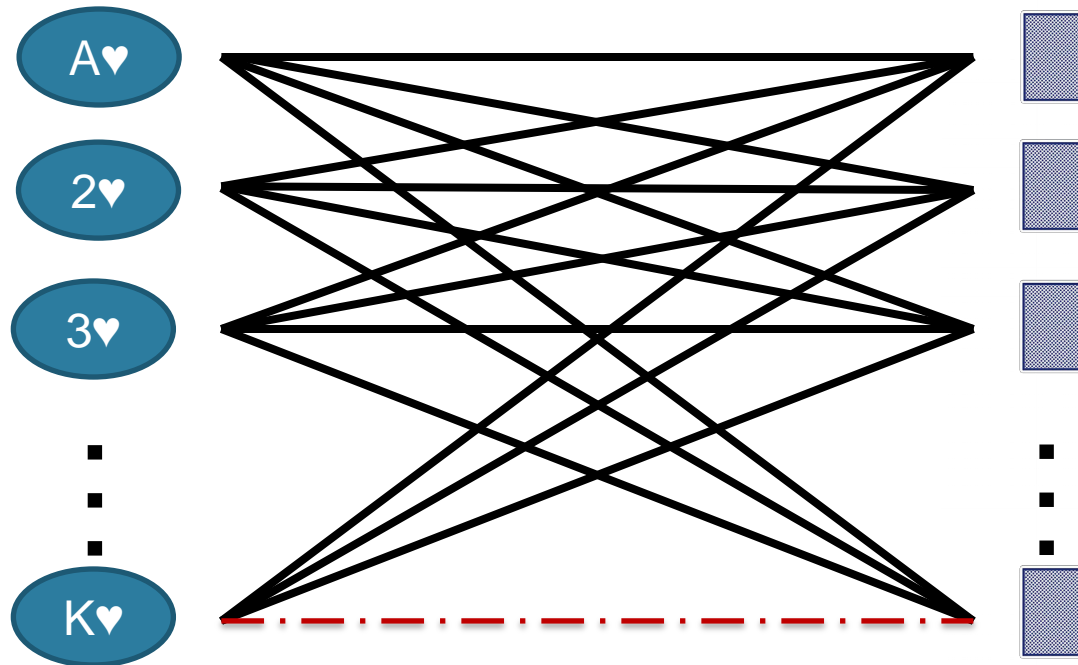
What if I set
$w(\text{Card}(K\heartsuit, p_{52})) = 0$?

Deck of Cards Graphically
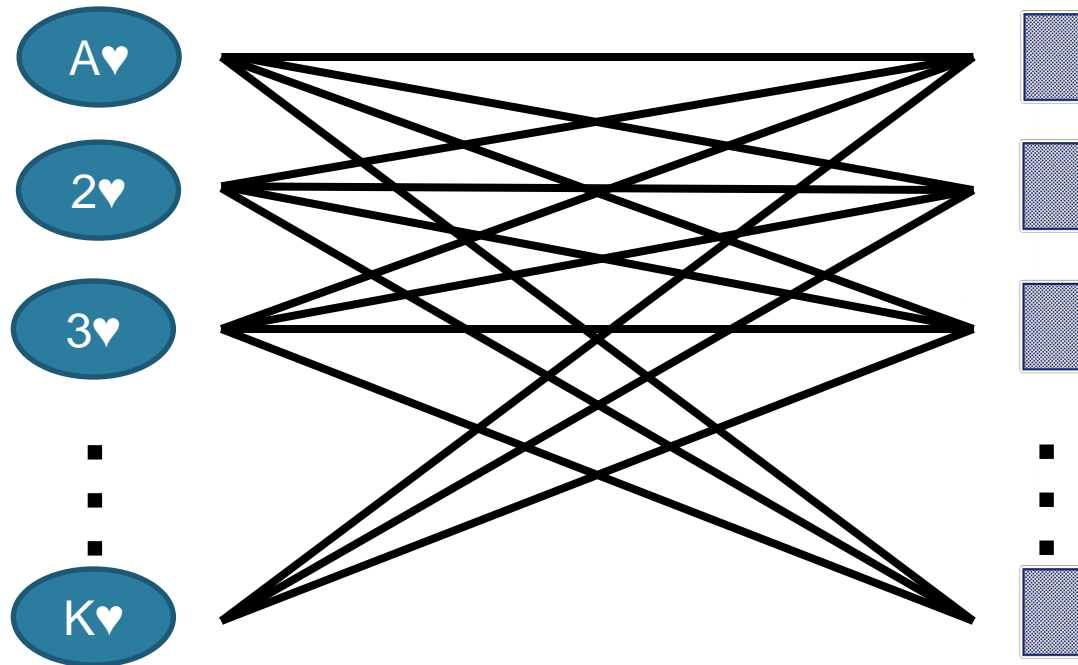
What if I set
$w(Card(K\heartsuit, p_{52})) = 0$?

Observations

- Asymmetric weight function = bigraph

- # models = #perfect matchings
- Problem is #P-complete! ☹

- All non-lifted WMC solvers efficiently handle asymmetric weights
- No solver does cards problem efficiently!

Tractable Classes

Safe monotone CNF

Safe type-1 CNF

[VdB; NIPS'11], [VdB et al.; KR'14], [Gribkoff, VdB, Suciu; UAI'15], [Beame, VdB, Gribkoff, Suciu; PODS'15], etc.

Tractable Classes

γ-acyclic CQs

FO²
CNF

Safe
monotone
CNF

FO²

Safe type-1 CNF

[VdB; NIPS'11], [VdB et al.; KR'14], [Gribkoff, VdB, Suciu; UAI'15], [Beame, VdB, Gribkoff, Suciu; PODS'15], etc.

Tractable Classes

γ-acyclic CQs

FO²
CNF

Safe
monotone
CNF

FO²

Safe type-1 CNF

$\Delta = \exists x,y, R(x) \wedge S(x,y) \wedge T(y)$

[VdB; NIPS'11], [VdB et al.; KR'14], [Gribkoff, VdB, Suciu; UAI'15], [Beame, VdB, Gribkoff, Suciu; PODS'15], etc.

Tractable Classes

γ-acyclic CQs

FO²
CNF

FO²

FO³

Safe
monotone
CNF

Safe type-1 CNF

[VdB; NIPS'11], [VdB et al.; KR'14], [Gribkoff, VdB, Suciu; UAI'15], [Beame, VdB, Gribkoff, Suciu; PODS'15], etc.

Tractable Classes

γ-acyclic CQs

FO^3

FO^2

FO^2 CNF

Safe monotone CNF

Safe type-1 CNF

FO^2

Cards

[VdB; NIPS'11], [VdB et al.; KR'14], [Gribkoff, VdB, Suciu; UAI'15], [Beame, VdB, Gribkoff, Suciu; PODS'15], etc.

Tractable Classes

[VdB; NIPS'11], [VdB et al.; KR'14], [Gribkoff, VdB, Suciu; UAI'15], [Beame, VdB, Gribkoff, Suciu; PODS'15], etc.

Tractable Classes

γ-acyclic CQs

CQs

$\#P_1$

$\#P_1$

FO^2 CNF

Safe monotone CNF

FO^3

FO^2

Safe type-1 CNF

$\#P_1$

?

$\Delta = \forall x, y, z, \text{Friends}(x,y) \wedge \text{Friends}(y,z) \Rightarrow \text{Friends}(x,z)$

[VdB; NIPS'11], [VdB et al.; KR'14], [Gribkoff, VdB, Suciu; UAI'15], [Beame, VdB, Gribkoff, Suciu; PODS'15], etc.

FO^2 is liftable!

Properties

Relations

Properties

Smokes(x)	Friends(x,y)	Smokes(y)
Gender(x)	Colleagues(x,y)	Gender(y)
Young(x)	Family(x,y)	Young(y)
Tall(x)	Classmates(x,y)	Tall(y)

X

Y
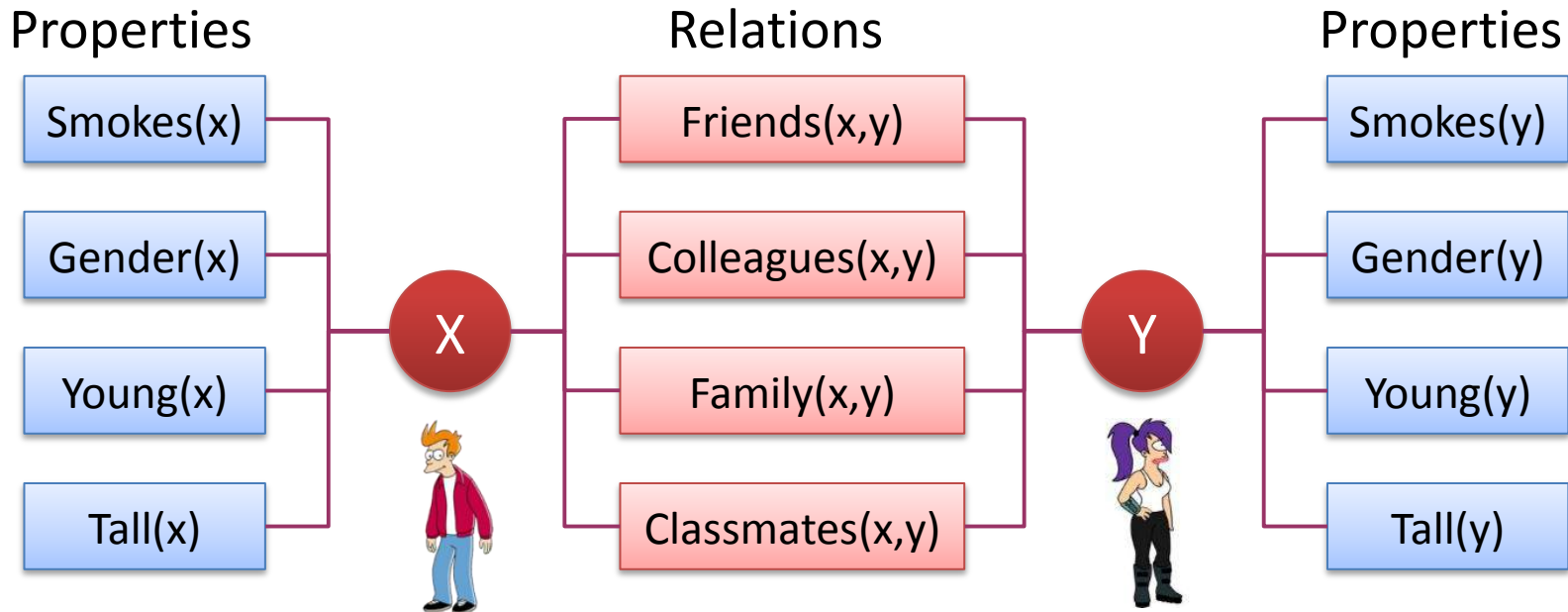
"Smokers are more likely to be friends with other smokers."
"Colleagues of the same age are more likely to be friends."
"People are either family or friends, but never both."
"If X is family of Y, then Y is also family of X."
"If X is a parent of Y, then Y cannot be a parent of X."

What is this really good for?

What we'd like to do…

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

⭐⭐⭐⭐⭐

Ernst Straus

⭐⭐☆☆☆

Kristian Kersting, …

⭐☆☆☆☆

Justin Bieber, …

Open World DB

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
...	...	...

- What if fact missing?

- Probability 0 for:

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

Open World DB

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
...	...	...

- What if fact missing?

- Probability 0 for:

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

Q2 = ∃x Coauthor(Bieber,**x**) ∧ Coauthor(Erdos,**x**)

Open World DB

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
...	...	...

- What if fact missing?

- Probability 0 for:

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

Q2 = ∃x Coauthor(Bieber,**x**) ∧ Coauthor(Erdos,**x**)

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

Open World DB

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
...	...	...

- What if fact missing?

- Probability 0 for:

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

Q2 = ∃x Coauthor(Bieber,**x**) ∧ Coauthor(Erdos,**x**)

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

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

Open World DB

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
...	...	...

- What if fact missing?

- Probability 0 for:

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

Q2 = ∃x Coauthor(Bieber,**x**) ∧ Coauthor(Erdos,**x**)

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

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**)

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

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

[Ceylan, Darwiche, Van den Broeck; KR'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**)

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

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

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

[Ceylan, Darwiche, Van den Broeck; KR'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**)

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

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

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

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

and P(Q3) ≥ P(Q5), P(Q4) ≥ P(Q5)

[Ceylan, Darwiche, Van den Broeck; KR'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**)

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

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

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

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

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

[Ceylan, Darwiche, Van den Broeck; KR'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 = ∃x Coauthor(Bieber,**x**) ∧ Coauthor(Erdos,**x**)

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

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

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

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

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

We have strong evidence that P(Q1) ≥ P(Q2).

[Ceylan, Darwiche, Van den Broeck; KR'16]

Bayesian Learning Loop

Bayesian view on learning:

1. Prior belief:

    P(Coauthor(Straus,Pauli)) = 0.01

2. Observe page

    P(Coauthor(Straus,Pauli|) = 0.2

3. Observe page

    P(Coauthor(Straus,Pauli)| ,) = 0.3

Principled and sound reasoning!

Problem: Broken Learning Loop

Bayesian view on learning:

1. Prior belief:

   P(Coauthor(Straus,Pauli)) = 0

2. Observe page

   P(Coauthor(Straus,Pauli |) = 0.2

3. Observe page

   P(Coauthor(Straus,Pauli) | ,) = 0.3

[Ceylan, Darwiche, Van den Broeck; KR'16]

Problem: Broken Learning Loop

Bayesian view on learning:

1. Prior belief:

   P(Coauthor(Straus,Pauli)) $= 0$

2. Observe page

   P(Coauthor(Straus,Pauli|) $= 0.2$

3. Observe page

   P(Coauthor(Straus,Pauli)| ,) $= 0.3$

[Ceylan, Darwiche, Van den Broeck; KR'16]

Problem: Broken Learning Loop

Bayesian view on learning:

1. Prior belief:

   P(Coauthor(Straus,Pauli)) = 0

2. Observe page

   P(Coauthor(Straus,Pauli|) = 0.2

3. Observe page

   P(Coauthor(Straus,Pauli)| ,) = 0.3

This is mathematical nonsense!

[Ceylan, Darwiche, Van den Broeck; KR'16]

Problem: Curse of Superlinearity

- Reality is worse!
- Tuples are **intentionally** missing!
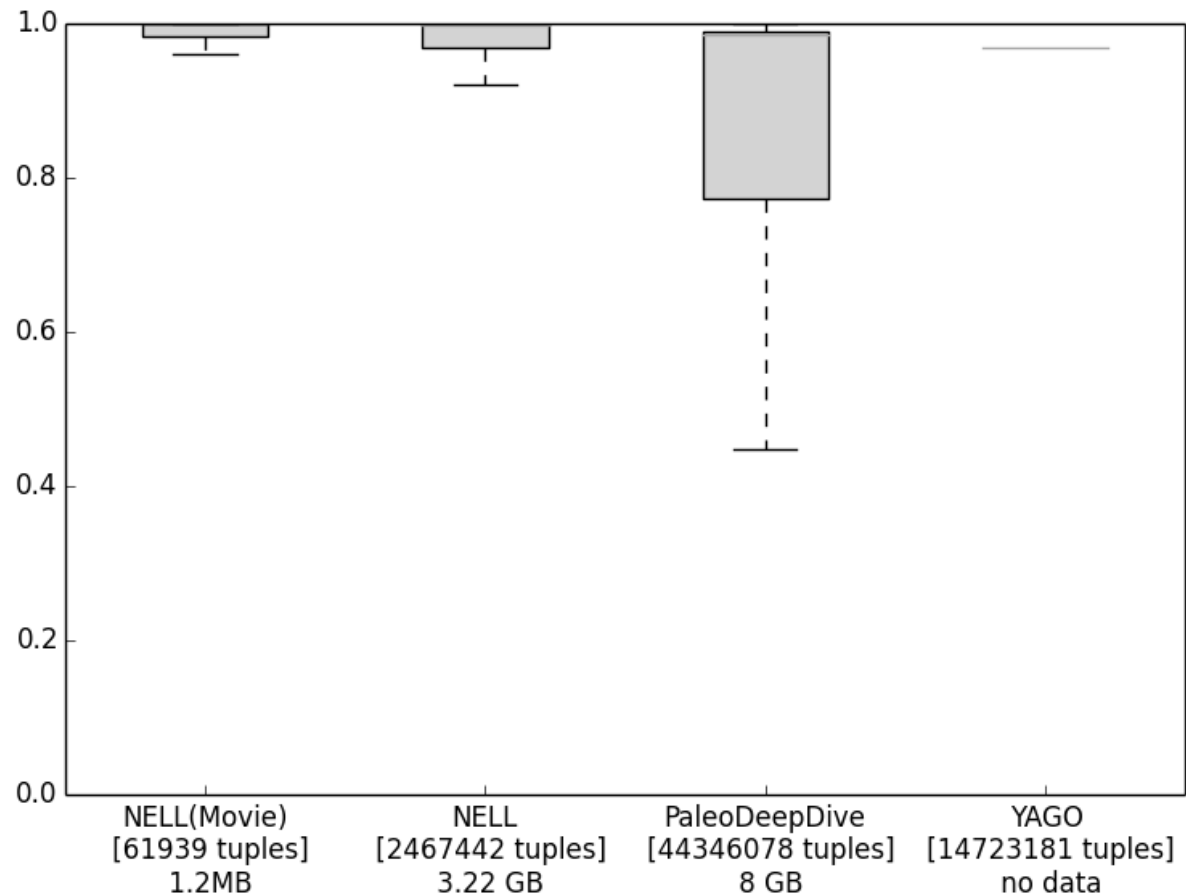
Problem: Curse of Superlinearity

Sibling

x	y	P
…	…	…

Facebook scale

[Ceylan, Darwiche, Van den Broeck; KR'16]

Problem: Curse of Superlinearity

Sibling

x	y	P
…	…	…

Facebook scale
⇒ 200 Exabytes of data"

[Ceylan, Darwiche, Van den Broeck; KR'16]

Problem: Curse of Superlinearity

Sibling

x	y	P
…	…	…

Facebook scale
⇒ 200 Exabytes of data"

All Google storage is a couple exabytes…

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

PUNCH CARDS

ICE SHEET

BOSTON

[Ceylan, Darwiche, Van den Broeck; KR'16]

Open-World Prob. Databases

Intuition: tuples can be added with $P < \lambda$

Q2 = Coauthor(Einstein,**Straus**) $\wedge$ Coauthor(Erdos,**Straus**)

$$P(Q2) \geq 0$$

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

<u>Intuition</u>: tuples can be added with P < λ

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

P(Q2) ≥ 0

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

<u>Intuition</u>: tuples can be added with $P < \lambda$

Q2 = Coauthor(Einstein,**Straus**) $\wedge$ Coauthor(Erdos,**Straus**)

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

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

UCQ

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

UCQ

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

- Polynomial time ☺

UCQ

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

- Polynomial time ☺

  Quadratic blow-up ☹

- 200 exabytes … again ☹

Closed-World Lifted Query Eval

$Q = \exists x\ \exists y\ \text{Scientist}(x) \wedge \text{Coauthor}(x,y)$

$P(Q) = 1 - \Pi_{A\ \in\ \text{Domain}}\ (1 - P(\text{Scientist}(A) \wedge \exists y\ \text{Coauthor}(A,y)))$

Closed-World Lifted Query Eval

$Q = \exists x \, \exists y \, \text{Scientist}(x) \land \text{Coauthor}(x,y)$

Decomposable ∀-Rule

$P(Q) = 1 - \Pi_{A \in \text{Domain}} (1 - P(\text{Scientist}(A) \land \exists y \, \text{Coauthor}(A,y)))$

Closed-World Lifted Query Eval

$Q = \exists x \, \exists y \, \text{Scientist}(x) \wedge \text{Coauthor}(x,y)$

Decomposable ∀-Rule

$P(Q) = 1 - \Pi_{A \in \text{Domain}} \, (1 - P(\text{Scientist}(A) \wedge \exists y \, \text{Coauthor}(A,y))$

$= 1 - (1 - P(\text{Scientist}(A) \wedge \exists y \, \text{Coauthor}(A,y))$
$ x \, (1 - P(\text{Scientist}(B) \wedge \exists y \, \text{Coauthor}(B,y))$
$ x \, (1 - P(\text{Scientist}(C) \wedge \exists y \, \text{Coauthor}(C,y))$
$ x \, (1 - P(\text{Scientist}(D) \wedge \exists y \, \text{Coauthor}(D,y))$
$ x \, (1 - P(\text{Scientist}(E) \wedge \exists y \, \text{Coauthor}(E,y))$
$ x \, (1 - P(\text{Scientist}(F) \wedge \exists y \, \text{Coauthor}(F,y))$
$ \ldots$

Closed-World Lifted Query Eval

$Q = \exists x \; \exists y \; \text{Scientist}(x) \land \text{Coauthor}(x,y)$

Decomposable ∀-Rule

$P(Q) = 1 - \Pi_{A \in \text{Domain}} (1 - P(\text{Scientist}(A) \land \exists y \; \text{Coauthor}(A,y))$

$$
\begin{aligned}
= 1 - &(1 - P(\text{Scientist}(A) \land \exists y \; \text{Coauthor}(A,y)) \\
&x \; (1 - P(\text{Scientist}(B) \land \exists y \; \text{Coauthor}(B,y)) \\
&x \; (1 - P(\text{Scientist}(C) \land \exists y \; \text{Coauthor}(C,y)) \\
&x \; (1 - P(\text{Scientist}(D) \land \exists y \; \text{Coauthor}(D,y)) \\
&x \; (1 - P(\text{Scientist}(E) \land \exists y \; \text{Coauthor}(E,y)) \\
&x \; (1 - P(\text{Scientist}(F) \land \exists y \; \text{Coauthor}(F,y)) \\
&\quad \ldots
\end{aligned}
$$

Complexity PTIME

Closed-World Lifted Query Eval

$Q = \exists x \; \exists y \; \text{Scientist}(x) \wedge \text{Coauthor}(x,y)$

$P(Q) = 1 - \prod_{A \in \text{Domain}} (1 - P(\text{Scientist}(A) \wedge \exists y \; \text{Coauthor}(A,y))$

$$
\begin{aligned}
= 1 - (1 - &P(\text{Scientist}(A) \wedge \exists y \; \text{Coauthor}(A,y)) \\
x \; (1 - &P(\text{Scientist}(B) \wedge \exists y \; \text{Coauthor}(B,y)) \\
x \; (1 - &P(\text{Scientist}(C) \wedge \exists y \; \text{Coauthor}(C,y)) \\
x \; (1 - &P(\text{Scientist}(D) \wedge \exists y \; \text{Coauthor}(D,y)) \\
x \; (1 - &P(\text{Scientist}(E) \wedge \exists y \; \text{Coauthor}(E,y)) \\
x \; (1 - &P(\text{Scientist}(F) \wedge \exists y \; \text{Coauthor}(F,y)) \\
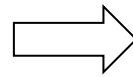&\ldots
\end{aligned}
$$

Closed-World Lifted Query Eval

$Q = \exists x\ \exists y\ \text{Scientist}(x) \wedge \text{Coauthor}(x,y)$

$P(Q) = 1 - \prod_{A\ \in\ \text{Domain}} (1 - P(\text{Scientist}(A) \wedge \exists y\ \text{Coauthor}(A,y))$

$= 1 - (1 - P(\text{Scientist}(A) \wedge \exists y\ \text{Coauthor}(A,y))$
$\quad\quad \times\ (1 - P(\text{Scientist}(B) \wedge \exists y\ \text{Coauthor}(B,y))$
$\quad\quad \times\ (1 - P(\text{Scientist}(C) \wedge \exists y\ \text{Coauthor}(C,y))$
$\quad\quad \times\ (1 - P(\text{Scientist}(D) \wedge \exists y\ \text{Coauthor}(D,y))$
$\quad\quad \times\ (1 - P(\text{Scientist}(E) \wedge \exists y\ \text{Coauthor}(E,y))$
$\quad\quad \times\ (1 - P(\text{Scientist}(F) \wedge \exists y\ \text{Coauthor}(F,y))$
$\quad\quad\quad \ldots$

⟹ No supporting facts in database!

Closed-World Lifted Query Eval

$Q = \exists x \; \exists y \; Scientist(x) \land Coauthor(x,y)$

$P(Q) = 1 - \Pi_{A \in \text{Domain}} (1 - P(Scientist(A) \land \exists y \; Coauthor(A,y))$

$= 1 - (1 - P(Scientist(A) \land \exists y \; Coauthor(A,y))$
$\quad x \; (1 - P(Scientist(B) \land \exists y \; Coauthor(B,y))$
$\quad x \; (1 - P(Scientist(C) \land \exists y \; Coauthor(C,y))$
$\quad x \; (1 - P(Scientist(D) \land \exists y \; Coauthor(D,y))$
$\quad x \; (1 - P(Scientist(E) \land \exists y \; Coauthor(E,y))$
$\quad x \; (1 - P(Scientist(F) \land \exists y \; Coauthor(F,y))$
$\quad \ldots$

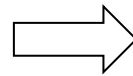No supporting facts in database!

Probability 0 in closed world

Closed-World Lifted Query Eval
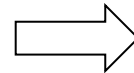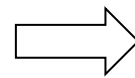
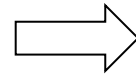$Q = \exists x\ \exists y\ \text{Scientist}(x) \wedge \text{Coauthor}(x,y)$

$P(Q) = 1 - \Pi_{A \in \text{Domain}}\ (1 - P(\text{Scientist}(A) \wedge \exists y\ \text{Coauthor}(A,y))$

$= 1 - (1 - P(\text{Scientist}(A) \wedge \exists y\ \text{Coauthor}(A,y))$
$\quad \times (1 - P(\text{Scientist}(B) \wedge \exists y\ \text{Coauthor}(B,y))$
$\quad \times (1 - P(\text{Scientist}(C) \wedge \exists y\ \text{Coauthor}(C,y))$
$\quad \times (1 - P(\text{Scientist}(D) \wedge \exists y\ \text{Coauthor}(D,y))$
$\quad \times (1 - P(\text{Scientist}(E) \wedge \exists y\ \text{Coauthor}(E,y))$
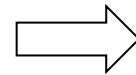$\quad \times (1 - P(\text{Scientist}(F) \wedge \exists y\ \text{Coauthor}(F,y))$
$\quad \quad \ldots$

$\Longrightarrow$ No supporting facts in database!

$\Longrightarrow$ Probability 0 in closed world

$\Longrightarrow$ Ignore these queries!

Closed-World Lifted Query Eval

Q = ∃x ∃y Scientist(x) ∧ Coauthor(x,y)

$$P(Q) = 1 - \Pi_{A \in \text{Domain}} (1 - P(\text{Scientist}(A) \land \exists y \text{ Coauthor}(A,y))$$

= 1 - (1 - P(Scientist(A) ∧ ∃y Coauthor(A,y))
    x (1 - P(Scientist(B) ∧ ∃y Coauthor(B,y))
    x (1 - P(Scientist(C) ∧ ∃y Coauthor(C,y))
    x (1 - P(Scientist(D) ∧ ∃y Coauthor(D,y))
    x (1 - P(Scientist(E) ∧ ∃y Coauthor(E,y))
    x (1 - P(Scientist(F) ∧ ∃y Coauthor(F,y))
    …

⟹ No supporting facts in database!

⟹ Probability 0 in closed world

⟹ Ignore these queries!

Complexity linear time!
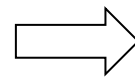
Open-World Lifted Query Eval
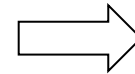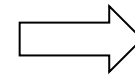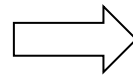
$Q = \exists x\ \exists y\ \text{Scientist}(x) \wedge \text{Coauthor}(x,y)$

$P(Q) = 1 - \Pi_{A \in \text{Domain}}\ (1 - P(\text{Scientist}(A) \wedge \exists y\ \text{Coauthor}(A,y))$

$= 1 - (1 - P(\text{Scientist}(A) \wedge \exists y\ \text{Coauthor}(A,y))$
$\quad\quad x\ (1 - P(\text{Scientist}(B) \wedge \exists y\ \text{Coauthor}(B,y))$
$\quad\quad x\ (1 - P(\text{Scientist}(C) \wedge \exists y\ \text{Coauthor}(C,y))$
$\quad\quad x\ (1 - P(\text{Scientist}(D) \wedge \exists y\ \text{Coauthor}(D,y))$
$\quad\quad x\ (1 - P(\text{Scientist}(E) \wedge \exists y\ \text{Coauthor}(E,y))$
$\quad\quad x\ (1 - P(\text{Scientist}(F) \wedge \exists y\ \text{Coauthor}(F,y))$
$\quad\quad\quad \ldots$

$\Longrightarrow$ No supporting facts in database!

Open-World Lifted Query Eval

Q = ∃x ∃y Scientist(x) ∧ Coauthor(x,y)

$P(Q) = 1 - \Pi_{A \in \text{Domain}} (1 - P(\text{Scientist}(A) \land \exists y\, \text{Coauthor}(A,y))$

= 1 - (1 - P(Scientist(A) ∧ ∃y Coauthor(A,y))
    x (1 - P(Scientist(B) ∧ ∃y Coauthor(B,y))
    x (1 - P(Scientist(C) ∧ ∃y Coauthor(C,y))
    x (1 - P(Scientist(D) ∧ ∃y Coauthor(D,y))
    x (1 - P(Scientist(E) ∧ ∃y Coauthor(E,y))
    x (1 - P(Scientist(F) ∧ ∃y Coauthor(F,y))
        …

⟹ No supporting facts in database!

⟹ Probability p in closed world

Open-World Lifted Query Eval

$Q = \exists x \, \exists y \, \text{Scientist}(x) \wedge \text{Coauthor}(x,y)$

$P(Q) = 1 - \prod_{A \in \text{Domain}} (1 - P(\text{Scientist}(A) \wedge \exists y \, \text{Coauthor}(A,y))$

$= 1 - (1 - P(\text{Scientist}(A) \wedge \exists y \, \text{Coauthor}(A,y))$
$\quad \times (1 - P(\text{Scientist}(B) \wedge \exists y \, \text{Coauthor}(B,y))$
$\quad \times (1 - P(\text{Scientist}(C) \wedge \exists y \, \text{Coauthor}(C,y))$
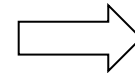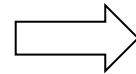$\quad \times (1 - P(\text{Scientist}(D) \wedge \exists y \, \text{Coauthor}(D,y))$
$\quad \times (1 - P(\text{Scientist}(E) \wedge \exists y \, \text{Coauthor}(E,y))$
$\quad \times (1 - P(\text{Scientist}(F) \wedge \exists y \, \text{Coauthor}(F,y))$
$\quad \quad \ldots$

→ No supporting facts in database!

→ Probability p in closed world

Complexity PTIME!

Open-World Lifted Query Eval

Q = ∃x ∃y Scientist(x) ∧ Coauthor(x,y)

$$P(Q) = 1 - \Pi_{A \in Domain} (1 - P(Scientist(A) \land \exists y\ Coauthor(A,y))$$

= 1 - (1 - P(Scientist(A) ∧ ∃y Coauthor(A,y))
    x (1 - P(Scientist(B) ∧ ∃y Coauthor(B,y))
    x (1 - P(Scientist(C) ∧ ∃y Coauthor(C,y))
    x (1 - P(Scientist(D) ∧ ∃y Coauthor(D,y))
    x (1 - P(Scientist(E) ∧ ∃y Coauthor(E,y))
    x (1 - P(Scientist(F) ∧ ∃y Coauthor(F,y))
       …

⟹ No supporting facts in database!

⟹ Probability p in closed world

Open-World Lifted Query Eval

$Q = \exists x \, \exists y \, \text{Scientist}(x) \land \text{Coauthor}(x,y)$

$P(Q) = 1 - \Pi_{A \in \text{Domain}} (1 - P(\text{Scientist}(A) \land \exists y \, \text{Coauthor}(A,y))$

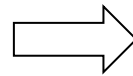$= 1 - (1 - P(\text{Scientist}(A) \land \exists y \, \text{Coauthor}(A,y))$
$ x \, (1 - P(\text{Scientist}(B) \land \exists y \, \text{Coauthor}(B,y))$
$ x \, (1 - P(\text{Scientist}(C) \land \exists y \, \text{Coauthor}(C,y))$
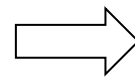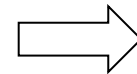$ x \, (1 - P(\text{Scientist}(D) \land \exists y \, \text{Coauthor}(D,y))$
$ x \, (1 - P(\text{Scientist}(E) \land \exists y \, \text{Coauthor}(E,y))$
$ x \, (1 - P(\text{Scientist}(F) \land \exists y \, \text{Coauthor}(F,y))$
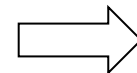$ \ldots$

→ No supporting facts in database!

→ Probability p in closed world

→ All together, probability $(1-p)^k$
Do symmetric lifted inference

Open-World Lifted Query Eval

Q = ∃x ∃y Scientist(x) ∧ Coauthor(x,y)

$$P(Q) = 1 - \Pi_{A \in \text{Domain}} (1 - P(\text{Scientist}(A) \land \exists y \text{ Coauthor}(A,y)))$$

= 1 - (1 - P(Scientist(A) ∧ ∃y Coauthor(A,y))
    x (1 - P(Scientist(B) ∧ ∃y Coauthor(B,y))
    x (1 - P(Scientist(C) ∧ ∃y Coauthor(C,y))
    x (1 - P(Scientist(D) ∧ ∃y Coauthor(D,y))
    x (1 - P(Scientist(E) ∧ ∃y Coauthor(E,y))
    x (1 - P(Scientist(F) ∧ ∃y Coauthor(F,y))
       …

⟹ No supporting facts in database!

⟹ Probability p in closed world

⟹ All together, probability $(1-p)^k$
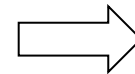Do symmetric lifted inference

Complexity linear time!

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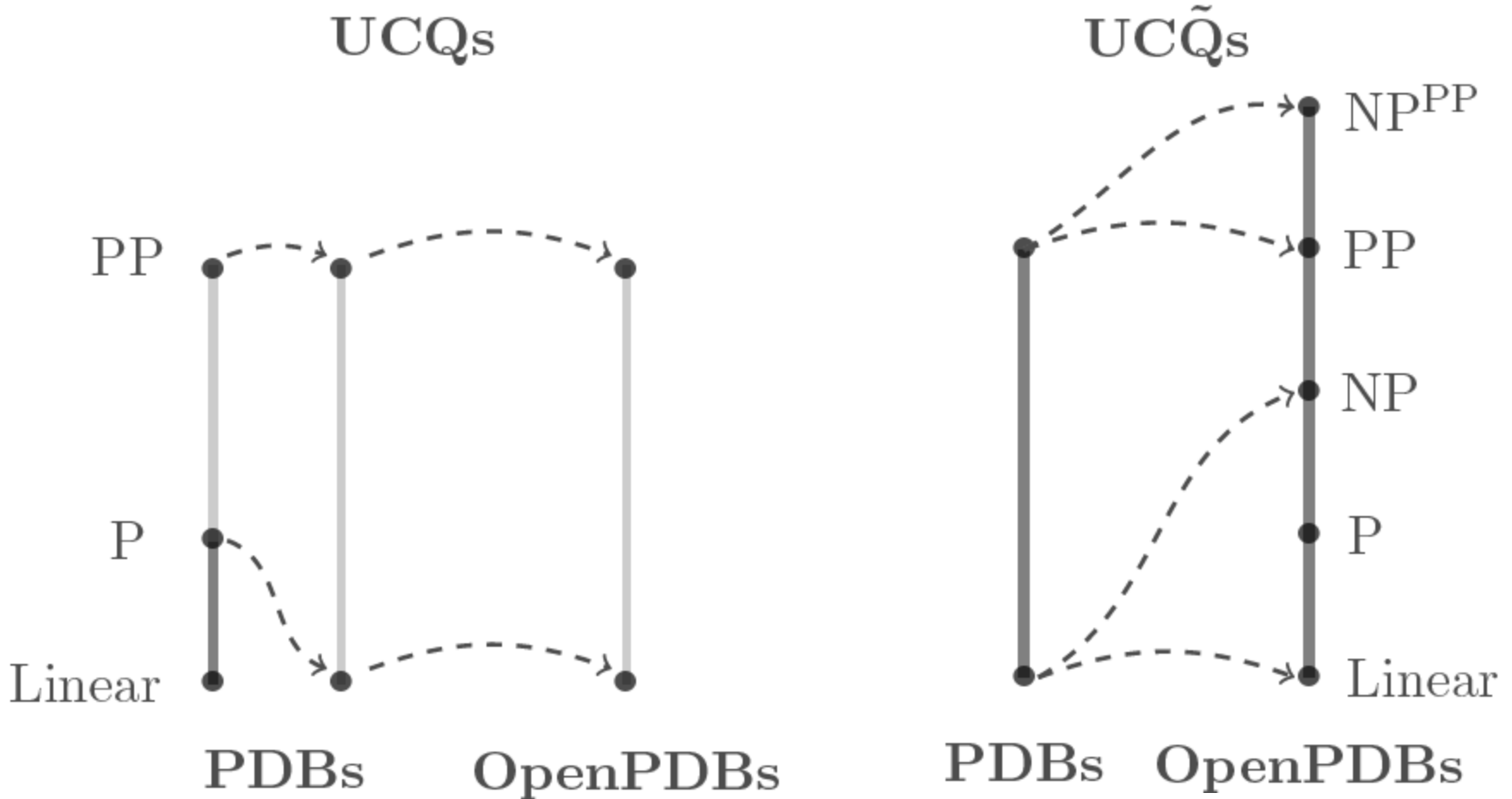
$$\text{Linear} \subseteq P \subseteq NP \subseteq PP \subseteq P^{PP} \subseteq NP^{PP} \subseteq \text{PSpace} \subseteq \text{ExpTime}$$

[Ceylan'16]

Parameter Learning

- **Given:** A set of first-order logic **formulas**

  w   FacultyPage(x) ∧ Linked(x,y) ⇒ CoursePage(y)

  A set of training **databases**

- **Learn:** The associated maximum-likelihood **weights**

$$\frac{\partial}{\partial w_j} \log \Pr_w(db) = n_j(db) - \mathbb{E}_w[n_j]$$

Count in databases
Efficient

Expected counts
Requires inference

$$\mathbb{E}_w[n_F] = \Pr(F\theta_1) + \cdots + \Pr(F\theta_m)$$

Lifted Parameter Learning

- **Given:** A set of first-order logic **formulas**
  A set of training **databases**
- **Learn:** Maximum-likelihood **weights**
- **Idea:** Lift the gradient computation

[Van Haaren et al.; MLJ'15]

Lifted Parameter Learning

- **Given:** A set of first-order logic **formulas**
  A set of training **databases**
- **Learn:** Maximum-likelihood **weights**
- **Idea:** Lift the gradient computation

900,030,000 random variables

[Van Haaren et al.; MLJ'15]
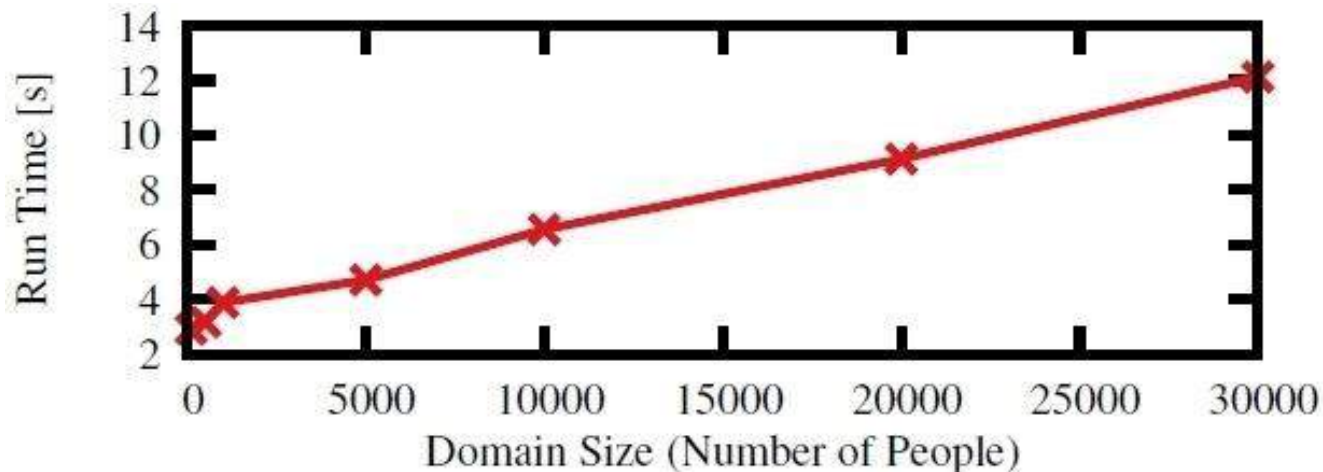
Lifted Structure Learning

- **Given:** A set of training **databases**

- **Learn:** A set of first-order logic **formulas**
  The associated maximum-likelihood **weights**

- **Idea:** Learn liftable models (regularize with symmetry)

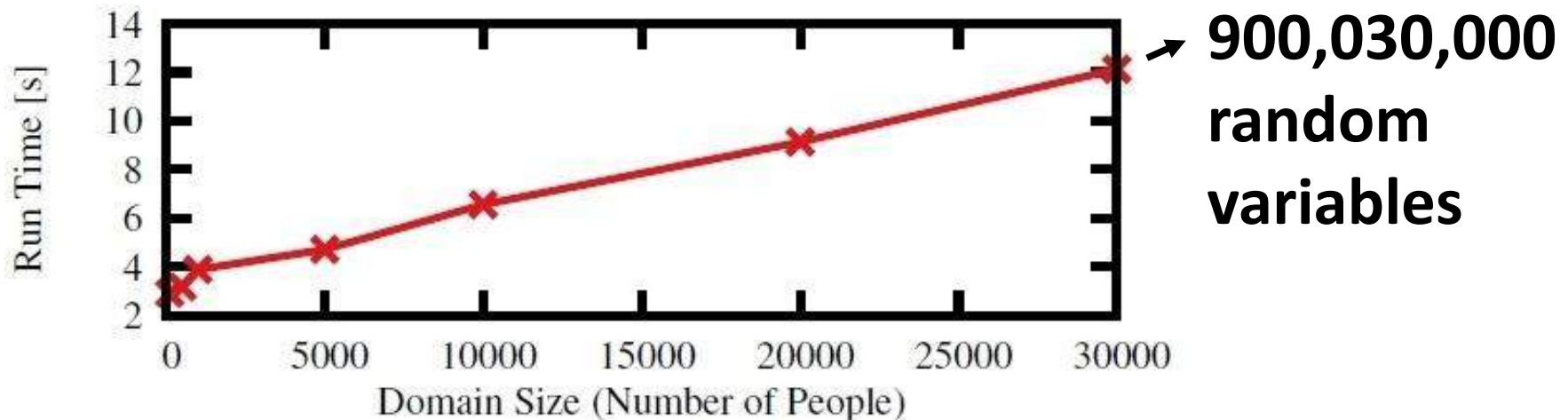|	IMDb			UWCSE		
	Baseline	Lifted Weight Learning	Lifted Structure Learning	Baseline	Lifted Weight Learning	Lifted Structure Learning
Fold 1	-548	-378	**-306**	-1,860	-1,524	**-1,477**
Fold 2	-689	-390	**-309**	-594	-535	**-511**
Fold 3	-1,157	-851	**-733**	-1,462	-1,245	**-1,167**
Fold 4	-415	-285	**-224**	-2,820	-2,510	**-2,442**
Fold 5	-413	-267	**-216**	-2,763	-2,357	**-2,227**

[Van Haaren et al.; MLJ'15]

What is the broader picture?

Uncertainty in AI

Probability Distribution

=

Qualitative

+

Quantitative
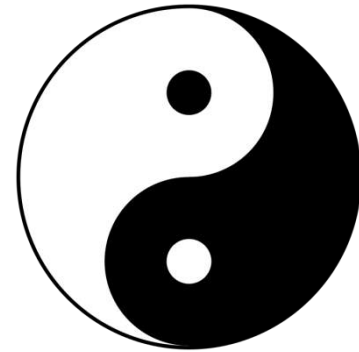
Generalized Model Counting

Probability Distribution

=

Logic

+

Weights

Generalized Model Counting

Probability Distribution	Logical Syntax
=	Model-theoretic Semantics
Logic	
+	**+**
Weights	Weight function w(.)

Weighted Model Integration

Probability Distribution

=

SMT(LRA)

+

Weights

[Belle et al. IJCAI'15, UAI'15]
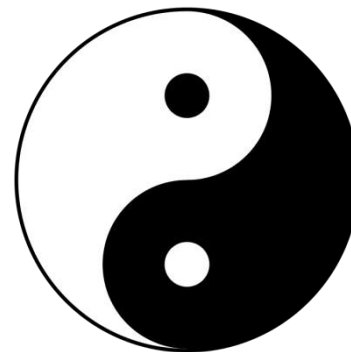
Weighted Model Integration

Probability Distribution

=

SMT(LRA)

+

Weights

$0 \leq$ height ≤ 200
$0 \leq$ weight ≤ 200
$0 \leq$ age ≤ 100
age $< 1 \Rightarrow$

  height+weight ≤ 90

+

w(height))=height-10
w(¬height)=$3*\text{height}^2$
w(¬weight)=5

…

[Belle et al. IJCAI'15, UAI'15]

Probabilistic Programming

Probability Distribution

=

Logic Programs

+

Weights

[Fierens et al., TPLP'15]

Probabilistic Programming

Probability Distribution

=
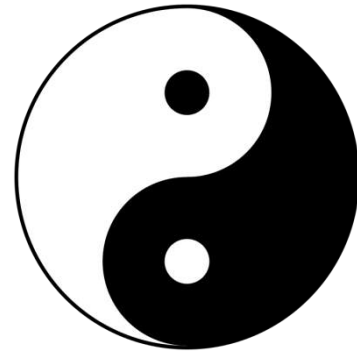
Logic Programs

+

Weights

```
path(X,Y) :-
        edge(X,Y).
path(X,Y) :-
        edge(X,Z), path(Z,Y).
```

+

[Fierens et al., TPLP'15]
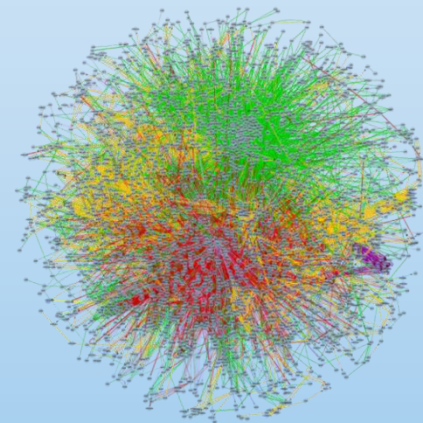
Probabilistic Functional Programming

Discrete probabilistic reachability program:

Logic Program (ProbLog) = Functional Program (Scala-like)

```
path(X,Y) :- edge(X,Y).
path(X,Y) :- edge(X,Z),
             path(Z,Y).
edge(X,Y) :- …random vars…
```

```
def path(start,end,visited=List())={
    if(start == end)
        return true
    if(visited.contains(start))
        return false
    return start.neighbors.exists{
        path(_,end,(visited+start))
    }
}
nodeA.neighbors = …random vars…
nodeB.neighbors = …random vars…
```

First-Order Knowledge Compilation

Markov Logic

| 3.14 | Smokes(x) ∧ Friends(x,y) ⇒ Smokes(y) |

[Van den Broeck,PhD'13]

First-Order Knowledge Compilation

Markov Logic | 3.14   Smokes(x) ∧ Friends(x,y) ⇒ Smokes(y)

Weight Function

w(Smokes)=1
w(¬Smokes)=1
w(Friends)=1
w(¬Friends)=1
w(F)=3.14
w(¬F)=1

FOL Sentence

∀x,y, F(x,y) ⇔ [Smokes(x) ∧ Friends(x,y) ⇒ Smokes(y)]

[Van den Broeck,PhD'13]

First-Order Knowledge Compilation

Markov Logic | 3.14   Smokes(x) ∧ Friends(x,y) ⇒ Smokes(y)

Weight Function

w(Smokes)=1
w(¬Smokes)=1
w(Friends)=1
w(¬Friends)=1
w(F)=3.14
w(¬F)=1

FOL Sentence

∀x,y, F(x,y) ⇔ [Smokes(x) ∧ Friends(x,y) ⇒ Smokes(y)]

Compile?

First-Order d-DNNF Circuit

First-Order Knowledge Compilation

Markov Logic $\quad$ 3.14 $\quad$ Smokes(x) ∧ Friends(x,y) ⇒ Smokes(y)

Weight Function

w(Smokes)=1
w(¬Smokes)=1
w(Friends)=1
w(¬Friends)=1
w(F)=3.14
w(¬F)=1

Domain

Alice
Bob
Charlie

FOL Sentence

∀x,y, F(x,y) ⇔ [Smokes(x) ∧ Friends(x,y) ⇒ Smokes(y)]

Compile?

First-Order d-DNNF Circuit

First-Order Knowledge Compilation

Markov Logic | 3.14   Smokes(x) ∧ Friends(x,y) ⇒ Smokes(y)

Weight Function

w(Smokes)=1
w(¬Smokes)=1
w(Friends)=1
w(¬Friends)=1
w(F)=3.14
w(¬F)=1

FOL Sentence

∀x,y, F(x,y) ⇔ [Smokes(x) ∧ Friends(x,y) ⇒ Smokes(y)]

Compile?

First-Order d-DNNF Circuit
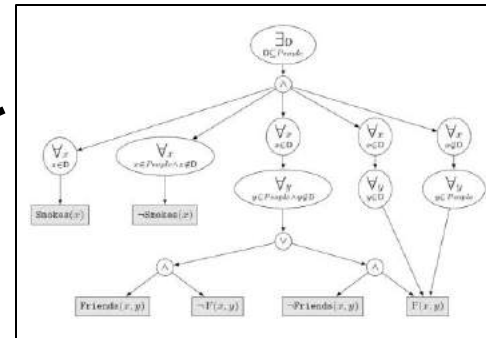
Domain

Alice
Bob
Charlie

Z = WFOMC = 1479.85

First-Order Knowledge Compilation

Markov Logic $\quad \boxed{3.14 \quad \text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y)}$
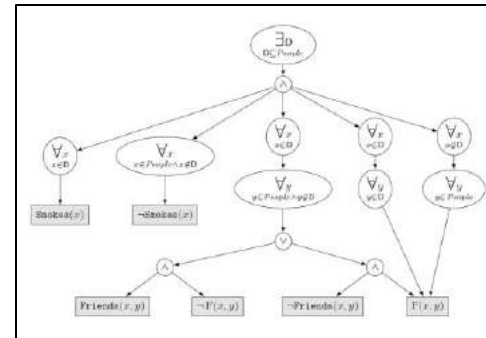
Weight Function

$\text{w}(\text{Smokes})=1$
$\text{w}(\neg\text{Smokes})=1$
$\text{w}(\text{Friends})=1$
$\text{w}(\neg\text{Friends})=1$
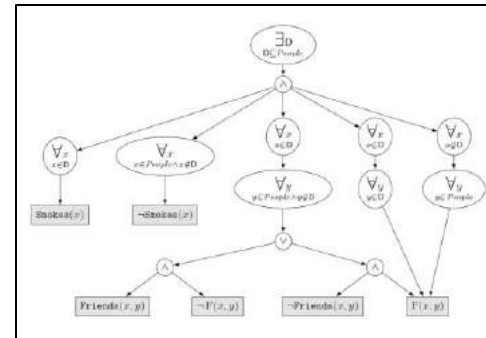$\text{w}(\text{F})=3.14$
$\text{w}(\neg\text{F})=1$

FOL Sentence

$\boxed{\forall x,y, \; F(x,y) \Leftrightarrow [\text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y)]}$

Compile?

First-Order d-DNNF Circuit
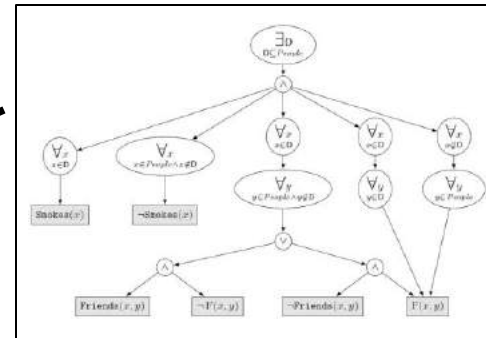
Domain

Alice
Bob
Charlie

$\boxed{\text{Z = WFOMC} = 1479.85}$

Evaluation in time polynomial in domain size!

First-Order Knowledge Compilation

Markov Logic | 3.14   Smokes(x) ∧ Friends(x,y) ⇒ Smokes(y)

Weight Function

w(Smokes)=1
w(¬Smokes)=1
w(Friends)=1
w(¬Friends)=1
w(F)=3.14
w(¬F)=1

FOL Sentence

∀x,y, F(x,y) ⇔ [Smokes(x) ∧ Friends(x,y) ⇒ Smokes(y)]

Compile?

First-Order d-DNNF Circuit
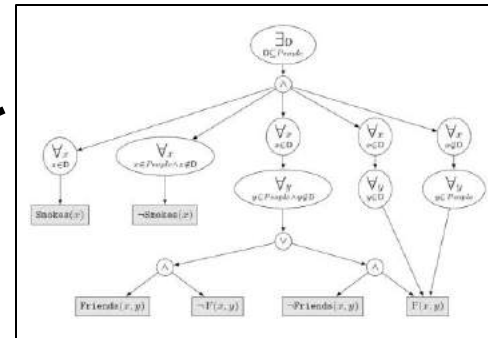
Domain

Alice
Bob
Charlie

Z = WFOMC = 1479.85

Evaluation in time polynomial in domain size!   = Lifted!

Approximate Symmetries in MCMC

- Exploit approximate symmetries:

  - Exact symmetry g
    $$Pr(\mathbf{x}) = Pr(\mathbf{x}^g)$$

  - Approximate symmetry g
    $$Pr(\mathbf{x}) \approx Pr(\mathbf{x}^g)$$

  $$Pr \begin{bmatrix} \end{bmatrix} \approx Pr \begin{bmatrix} \end{bmatrix}$$
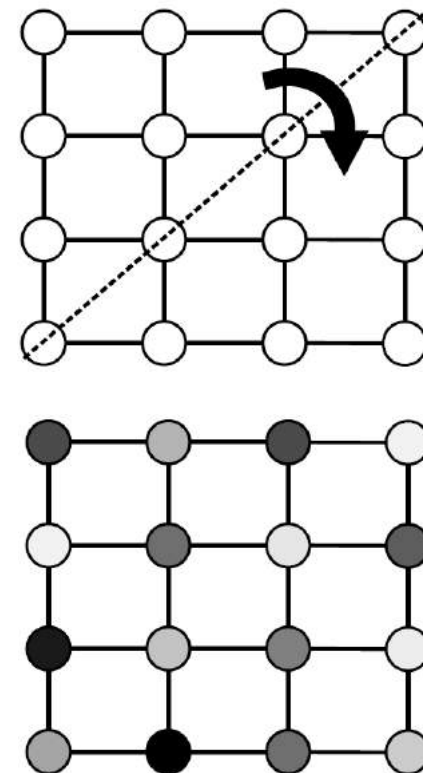
- Approximate lifted inference (MCMC)

[Van den Broeck, Darwiche; NIPS'13], [Van den Broeck, Niepert; AAAI'15]

Exchangeability

- FO2 characterized by a particular form of partial exchangeability

  You can construct a statistic that defines probability

  This is what our algorithms do

- Other forms of exchangeability open …

- Connection to computability?

Results for Symmetric Inference

Theorem
There exists Q in FO^3 s.t. FOMC(Q, n) is $\#P_1$ hard
There exists CQ Q s.t. WFOMC(Q, n) is $\#P_1$ hard

Theorem WFOMC(Q, n) is in PTIME
- For any Q in FO^2
- For any gamma-acyclic Q

- o Corresponding decision problem = the spectrum problem
- o Data complexity: { Spec(Q) | Q in FO} = NP_1 [Fagin'74]
- o Combined complexity: NP-complete for FO^2, PSPACE-complete for FO
- o 0-1 Laws connection

[Van den Broeck'14 ,Beame'15]

Conclusions

- Relational probabilistic reasoning is **frontier** and **integration** of AI, KR, ML, DB, TH, etc.
- We need
  - relational models and logic
  - probabilistic models and statistical learning
  - algorithms that scale
- Open-world data model
  - semantics make sense
  - FREE for UCQs
  - expensive otherwise

Long-Term Outlook

Probabilistic inference and learning exploit

~ 1988: conditional independence

~ 2000: contextual independence (local structure)

Long-Term Outlook

Probabilistic inference and learning exploit

~ 1988: conditional independence

~ 2000: contextual independence (local structure)

~ 201?: **symmetry & exchangeability & first-order**

Thanks!

References

- Chen, Brian X. "Siri, Alexa and Other Virtual Assistants Put to the Test" The New York Times (2016).

- Van den Broeck, Guy. "Towards high-level probabilistic reasoning with lifted inference." AAAI Spring Symposium on KRR (2015).

- Niepert, Mathias, and Guy Van den Broeck. "Tractability through exchangeability: A new perspective on efficient probabilistic inference." AAAI (2014).

- Van den Broeck, Guy, Nima Taghipour, Wannes Meert, Jesse Davis, and Luc De Raedt. "Lifted probabilistic inference by first-order knowledge compilation." In Proceedings of the Twenty-Second international joint conference on Artificial Intelligence, pp. 2178-2185. AAAI Press/International Joint Conferences on Artificial Intelligence, 2011.

- Gogate, Vibhav, and Pedro Domingos. "Probabilistic theorem proving." UAI (2011).

References

- Suciu, Dan, Dan Olteanu, Christopher Ré, and Christoph Koch. "Probabilistic databases." Synthesis Lectures on Data Management 3, no. 2 (2011): 1-180.

- Van den Broeck, Guy. "On the completeness of first-order knowledge compilation for lifted probabilistic inference." In Advances in Neural Information Processing Systems, pp. 1386-1394. 2011.

- Van den Broeck, Guy, Wannes Meert, and Adnan Darwiche. "Skolemization for weighted first-order model counting." In Proceedings of the 14th International Conference on Principles of Knowledge Representation and Reasoning (KR). 2014.

- Gribkoff, Eric, Guy Van den Broeck, and Dan Suciu. "Understanding the complexity of lifted inference and asymmetric weighted model counting." UAI, 2014.

References

- Beame, Paul, Guy Van den Broeck, Eric Gribkoff, and Dan Suciu. "Symmetric weighted first-order model counting." In Proceedings of the 34th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems, pp. 313-328. ACM, 2015.

- Ceylan, Ismail Ilkan, Adnan Darwiche, and Guy Van den Broeck. "Open-world probabilistic databases." Proceedings of KR (2016).

- Van Haaren, Jan, Guy Van den Broeck, Wannes Meert, and Jesse Davis. "Lifted generative learning of Markov logic networks." Machine Learning 103, no. 1 (2016): 27-55.

- Belle, Vaishak, Andrea Passerini, and Guy Van den Broeck. "Probabilistic inference in hybrid domains by weighted model integration." Proceedings of 24th International Joint Conference on Artificial Intelligence (IJCAI). 2015.

- Belle, Vaishak, Guy Van den Broeck, and Andrea Passerini. "Hashing-based approximate probabilistic inference in hybrid domains." In Proceedings of the 31st Conference on Uncertainty in Artificial Intelligence (UAI). 2015.

References

- Fierens, Daan, Guy Van den Broeck, Joris Renkens, Dimitar Shterionov, Bernd Gutmann, Ingo Thon, Gerda Janssens, and Luc De Raedt. "Inference and learning in probabilistic logic programs using weighted boolean formulas." Theory and Practice of Logic Programming 15, no. 03 (2015): 358-401.

- Van den Broeck, Guy. Lifted inference and learning in statistical relational models. Diss. Ph. D. Dissertation, KU Leuven, 2013.

- Van den Broeck, Guy, and Mathias Niepert. "Lifted Probabilistic Inference for Asymmetric Graphical Models." In Twenty-Ninth AAAI Conference on Artificial Intelligence. 2015.

- Van den Broeck, Guy, and Adnan Darwiche. "On the complexity and approximation of binary evidence in lifted inference." In Advances in Neural Information Processing Systems, pp. 2868-2876. 2013.