

First-Order Probabilistic Reasoning: Successes and Challenges

Guy Van den Broeck

UCLA

IJCAI Early Career Spotlight
Jul 14, 2016

Overview

1. *Why first-order probabilistic **models**?*
2. *Why first-order probabilistic **reasoning**?*
3. *How does lifted inference **work**?*
4. *What are the **successes**?*
5. *What are the **challenges**?*

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

Graphical Model Learning



Medical Records

Name	Cough	Asthma	Smokes
Alice	1	1	0
Bob	0	0	0
Charlie	0	1	0
Dave	1	0	1
Eve	1	0	0

Graphical Model Learning

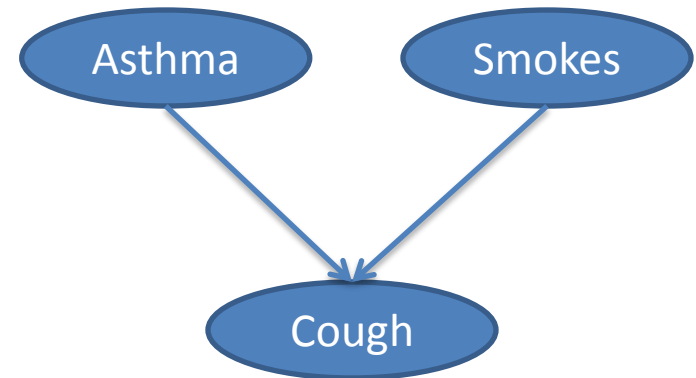


Medical Records



Bayesian Network

Name	Cough	Asthma	Smokes
Alice	1	1	0
Bob	0	0	0
Charlie	0	1	0
Dave	1	0	1
Eve	1	0	0



Graphical Model Learning



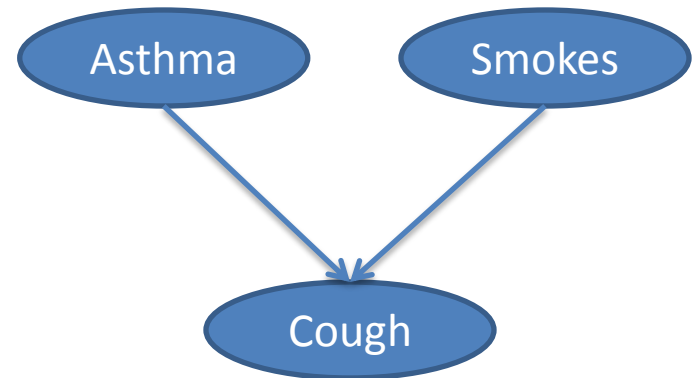
Medical Records



Bayesian Network

Name	Cough	Asthma	Smokes
Alice	1	1	0
Bob	0	0	0
Charlie	0	1	0
Dave	1	0	1
Eve	1	0	0

Frank	1	?	?
-------	---	---	---



Graphical Model Learning



Medical Records

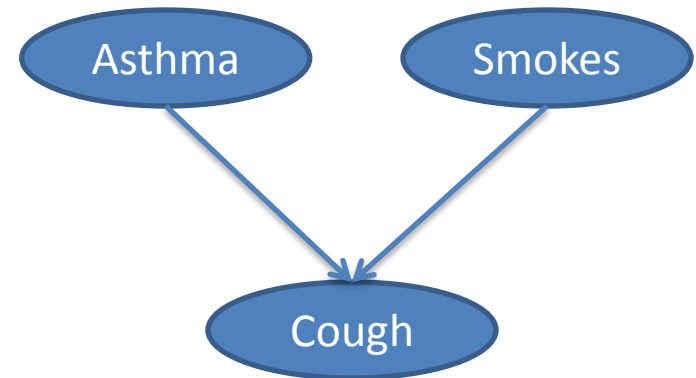


Bayesian Network

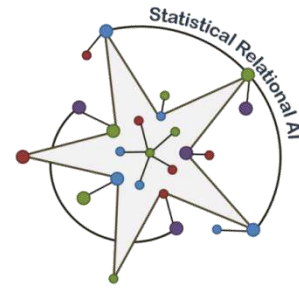
Name	Cough	Asthma	Smokes
Alice	1	1	0
Bob	0	0	0
Charlie	0	1	0
Dave	1	0	1
Eve	1	0	0

Frank	1	?	?
-------	---	---	---

Frank	1	0.3	0.2
-------	---	-----	-----



Statistical Relational Learning



Medical Records

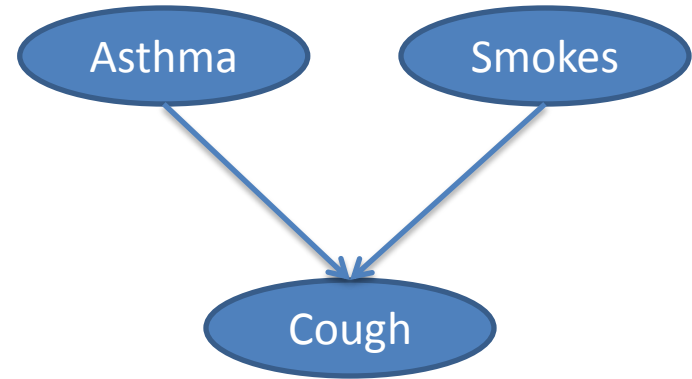
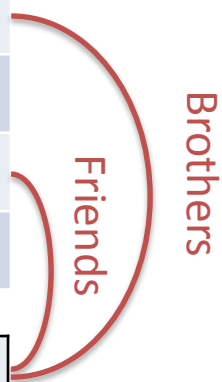


Bayesian Network

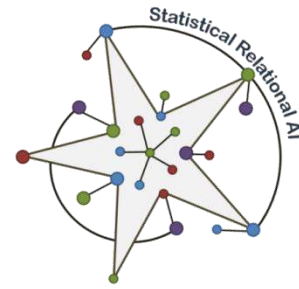
Name	Cough	Asthma	Smokes
Alice	1	1	0
Bob	0	0	0
Charlie	0	1	0
Dave	1	0	1
Eve	1	0	0

Frank	1	?	?
-------	---	---	---

Frank	1	0.3	0.2
-------	---	-----	-----



Statistical Relational Learning



Medical Records

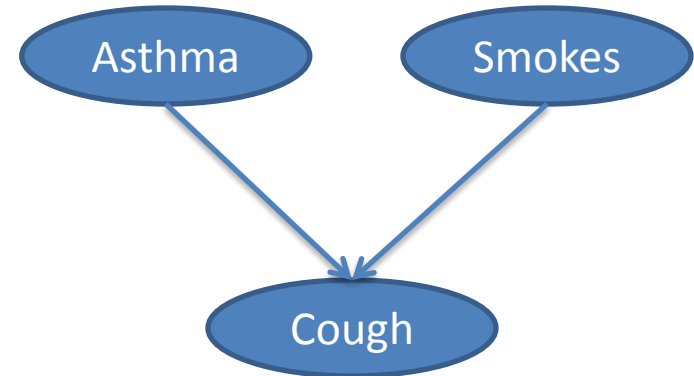
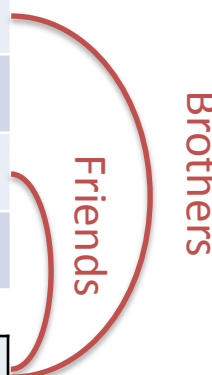


Bayesian Network

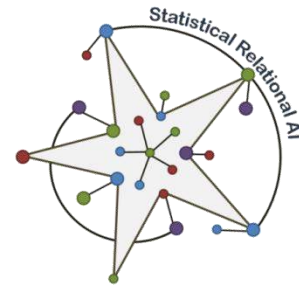
Name	Cough	Asthma	Smokes
Alice	1	1	0
Bob	0	0	0
Charlie	0	1	0
Dave	1	0	1
Eve	1	0	0

Frank	1	?	?
-------	---	---	---

Frank	1	0.3	0.2
-------	---	-----	-----



Statistical Relational Learning



Medical Records



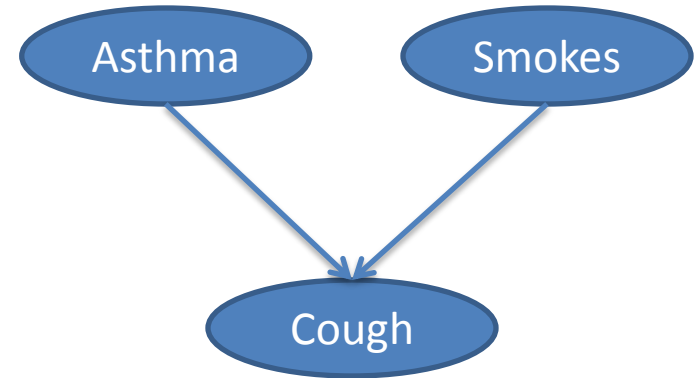
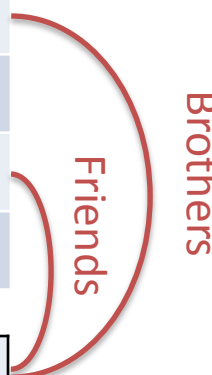
Bayesian Network

Name	Cough	Asthma	Smokes
Alice	1	1	0
Bob	0	0	0
Charlie	0	1	0
Dave	1	0	1
Eve	1	0	0

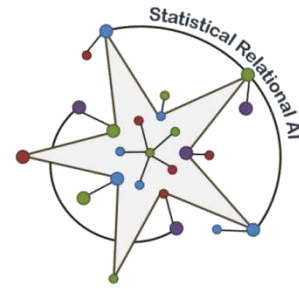
Frank	1	?	?
-------	---	---	---

Frank	1	0.3	0.2
-------	---	-----	-----

Frank	1	0.2	0.6
-------	---	-----	-----



Statistical Relational Learning

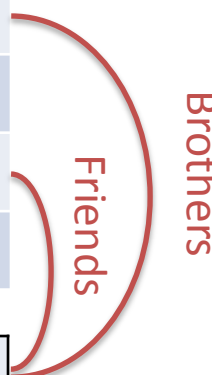


Medical Records



Bayesian Network

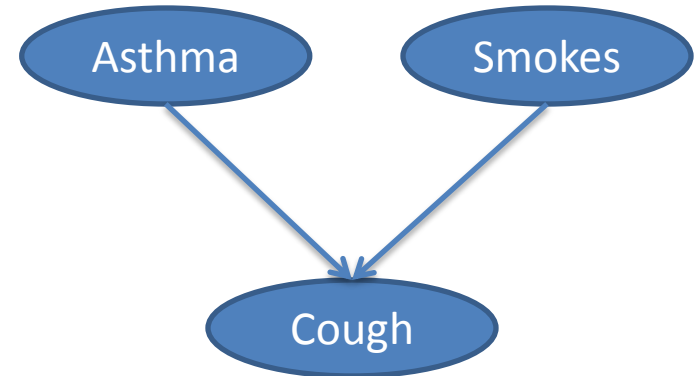
Name	Cough	Asthma	Smokes
Alice	1	1	0
Bob	0	0	0
Charlie	0	1	0
Dave	1	0	1
Eve	1	0	0



Frank	1	?	?
-------	---	---	---

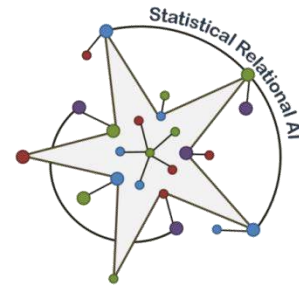
Frank	1	0.3	0.2
-------	---	-----	-----

Frank	1	0.2	0.6
-------	---	-----	-----



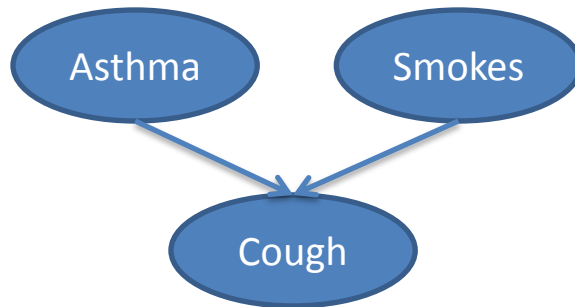
Rows are **independent** during learning and inference!

Statistical Relational Learning



Augment graphical model with relations between entities (rows).

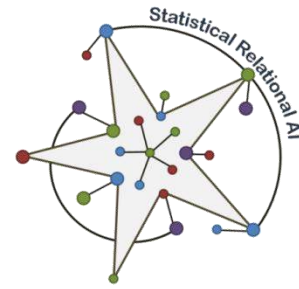
Intuition



Markov Logic

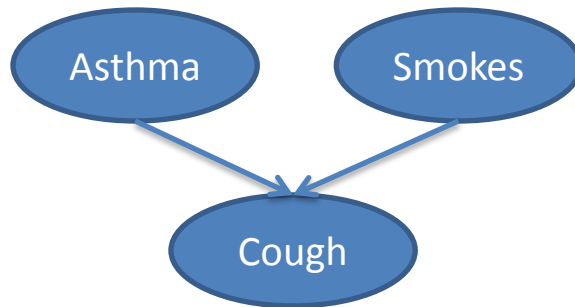
- + Friends have similar smoking habits
- + Asthma can be hereditary

Statistical Relational Learning



Augment graphical model with relations between entities (rows).

Intuition



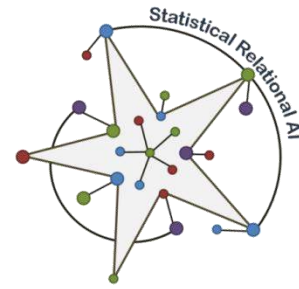
Markov Logic

2.1 Asthma \Rightarrow Cough

3.5 Smokes \Rightarrow Cough

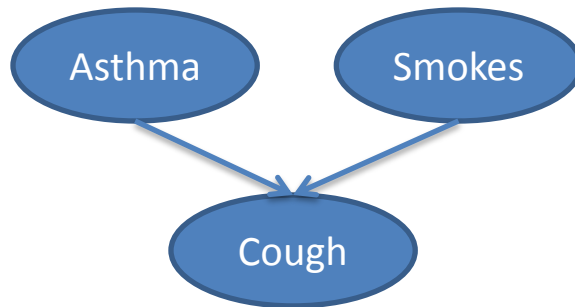
- + Friends have similar smoking habits
- + Asthma can be hereditary

Statistical Relational Learning



Augment graphical model with relations between entities (rows).

Intuition



- + Friends have similar smoking habits
- + Asthma can be hereditary

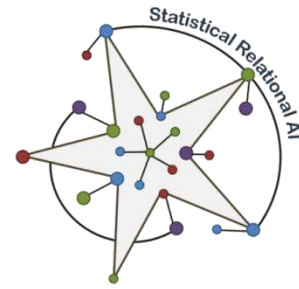
Markov Logic

2.1 $\text{Asthma}(x) \Rightarrow \text{Cough}(x)$

3.5 $\text{Smokes}(x) \Rightarrow \text{Cough}(x)$

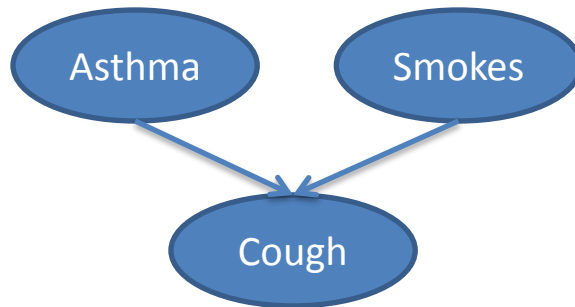
↑
Logical variables refer to entities

Statistical Relational Learning



Augment graphical model with relations between entities (rows).

Intuition



- + Friends have similar smoking habits
- + Asthma can be hereditary

Markov Logic

$$2.1 \text{ Asthma}(x) \Rightarrow \text{Cough}(x)$$

$$3.5 \text{ Smokes}(x) \Rightarrow \text{Cough}(x)$$

$$1.9 \text{ Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y)$$

$$1.5 \text{ Asthma}(x) \wedge \text{Family}(x,y) \Rightarrow \text{Asthma}(y)$$

Google Knowledge Graph

The image shows a Google search interface for 'Larry Page'. At the top, the search bar contains 'Larry Page' and the Google logo. Below the search bar, navigation tabs for 'Web', 'Images', 'Maps', 'Shopping', 'News', and 'More' are visible. The search results on the left include a snippet about 'Ubergizmo' and several links to profiles on Forbes, Google+, and Biography.com. On the right, a 'Knowledge Graph' panel is highlighted with a red box and arrows. This panel features a large portrait of Larry Page, a grid of smaller images, and a list of biographical facts such as his birth date (March 26, 1973), height (5' 11"), spouse (Lucinda Southworth), and education (East Lansing High School). Below the facts, there are sections for 'Recent posts' and 'People also search for' with small image thumbnails.

> 570 million entities
> 18 billion tuples

Knowledge Graph

Larry Page
6,606,633 followers on Google+

Lawrence "Larry" Page is an American computer scientist and Internet entrepreneur who is the co-founder of Google, alongside Sergey Brin. On April 4, 2011, Page succeeded Eric Schmidt as the chief executive officer of Google. *Wikipedia*

Born: March 26, 1973 (age 40), East Lansing, MI
Height: 5' 11" (1.80 m)
Spouse: Lucinda Southworth (m. 2007)
Siblings: Carl Victor Page, Jr.
Education: East Lansing High School (1987–1991), More
Awards: Marconi Prize, TR100

Recent posts
Just opened the new Android release, KitKat! Sep 3, 2013

People also search for

What we'd like to do...

Has anyone published a paper with both Erdos and Einstein



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 ...

What we'd like to do...

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



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 ...

Erdős is in the Knowledge Graph

Paul Erdos



All Images Videos Books News More Search tools

About 333,000 results (0.35 seconds)

Paul Erdős - Wikipedia, the free encyclopedia

https://en.wikipedia.org/wiki/Paul_Erdős - Wikipedia

Paul Erdős was a Hungarian Jewish mathematician. He was one of the most prolific mathematicians of the 20th century. He was known both for his social ...

Fan Chung - Ronald Graham - Béla Bollobás - Category:Paul Erdős

The Man Who Loved Only Numbers - The New York Times

<https://www.nytimes.com/books/.../hoffman-man.ht...> - The New York Times

Paul Erdős was one of those very special geniuses, the kind who comes along only once in a very long while yet he chose, quite consciously I am sure, to share ...

Paul Erdos | Hungarian mathematician | Britannica.com

www.britannica.com/biography/Paul-Erdos - Encyclopaedia Britannica

Paul Erdős, (born March 26, 1913, Budapest, Hungary—died September 20, 1996, Warsaw, Poland), Hungarian "freelance" mathematician (known for his work ...

Paul Erdős - University of St Andrews

www-groups.dcs.st-and.ac.uk/~history/Biographies/Erdos.html

Paul Erdős came from a Jewish family (the original family name being Engländer) although neither of his parents observed the Jewish religion. Paul's father ...

[PDF] Paul Erdős Mathematical Genius, Human - UnTruth.org

www.untruth.org/~josh/math/Paul%20Erdős%20bio-rev2.pdf

by J Hill - 2004 - Related articles



Paul Erdős

Mathematician

Paul Erdős was a Hungarian Jewish mathematician. He was one of the most prolific mathematicians of the 20th century. He was known both for his social practice of mathematics and for his eccentric lifestyle.

[Wikipedia](#)

Born: March 26, 1913, Budapest, Hungary

Died: September 20, 1996, Warsaw, Poland

Education: Eötvös Loránd University (1934)

Books: Probabilistic Methods in Combinatorics, [More](#)

Notable students: Béla Bollobás, Alexander Soifer, George B. Purdy, Joseph Kruskal

Einstein is in the Knowledge Graph

Albert Einstein



All News Images Books Videos More Search tools

About 82,800,000 results (0.45 seconds)

The Official Licensing Site of Albert Einstein

einstein.biz/

Welcome to the Official Licensing Site of **Albert Einstein**. Learn more about **Albert Einstein** and contact us today for any commercial licensing inquiries.

Albert Einstein - Wikipedia, the free encyclopedia

https://en.wikipedia.org/wiki/Albert_Einstein - Wikipedia

Albert Einstein (/ˈaɪnʃtaɪn/; German: [ˈalbɛʁt ˈaɪnʃtaɪn] (listen); 14 March 1879 – 18 April 1955) was a German-born theoretical physicist.

[Hans Albert Einstein](#) - [Mass–energy equivalence](#) - [Eduard Einstein](#) - [Elsa Einstein](#)

Albert Einstein (@AlbertEinstein) | Twitter

<https://twitter.com/AlbertEinstein>

16 hours ago - [View on Twitter](#)

ICYMI, Albert Einstein knew a thing or two about being romantic. Learn about the love letters he wrote. guff.com/didnt-know-einst...

20 hours ago - [View on Twitter](#)

An interesting read on Einstein's superstar status. What are your thoughts? twitter.com/aeonmag/status...

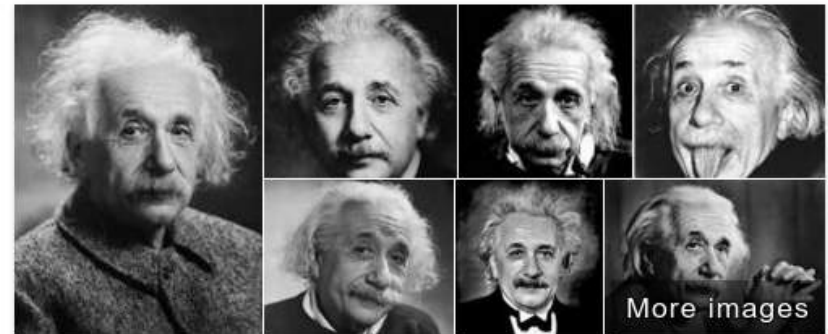


Albert Einstein - Biographical - Nobelprize.org

www.nobelprize.org/nobel_prizes/physics/.../einstein-bio.htm... - Nobel Prize

Albert Einstein was born at Ulm, in Württemberg, Germany, on March 14, 1879. ...

Later, they moved to Italy and Albert continued his education at Aarau



Albert Einstein

Theoretical Physicist

Albert Einstein was a German-born theoretical physicist. He developed the general theory of relativity, one of the two pillars of modern physics. Einstein's work is also known for its influence on the philosophy of science. [Wikipedia](#)

Born: March 14, 1879, [Ulm, Germany](#)

Died: April 18, 1955, [Princeton, NJ](#)

Influenced by: [Isaac Newton](#), [Mahatma Gandhi](#), [More](#)

Children: [Eduard Einstein](#), [Lieserl Einstein](#), [Hans Albert Einstein](#)

Spouse: [Elsa Einstein](#) (m. 1919–1936), [Mileva Marić](#) (m. 1903–1919)

This guy is in the Knowledge Graph

Ernst Straus

All Images News Maps Videos More Search tools

About 349,000 results (0.37 seconds)

Ernst G. Straus - Wikipedia, the free encyclopedia
https://en.wikipedia.org/wiki/Ernst_G._Straus Wikipedia
Ernst Gabor Straus (February 25, 1922 – July 12, 1983) was a German-American mathematician who helped found the theories of Euclidean Ramsey theory ...

Straus biography - University of St Andrews
www-groups.dcs.st-and.ac.uk/~history/Biographies/Straus.html
Ernst Straus's mother was Rahel Goitein who had the distinction of being one of the first women medical students officially studying at a German university.

Ernst G. Straus
Mathematician
Ernst Gabor Straus was a German-American mathematician who helped found the theories of Euclidean Ramsey theory and of the arithmetic properties of analytic functions. [Wikipedia](#)
Born: February 25, 1922, Munich, Germany
Died: July 12, 1983, Los Angeles, CA
Residence: United States of America

... 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, ...



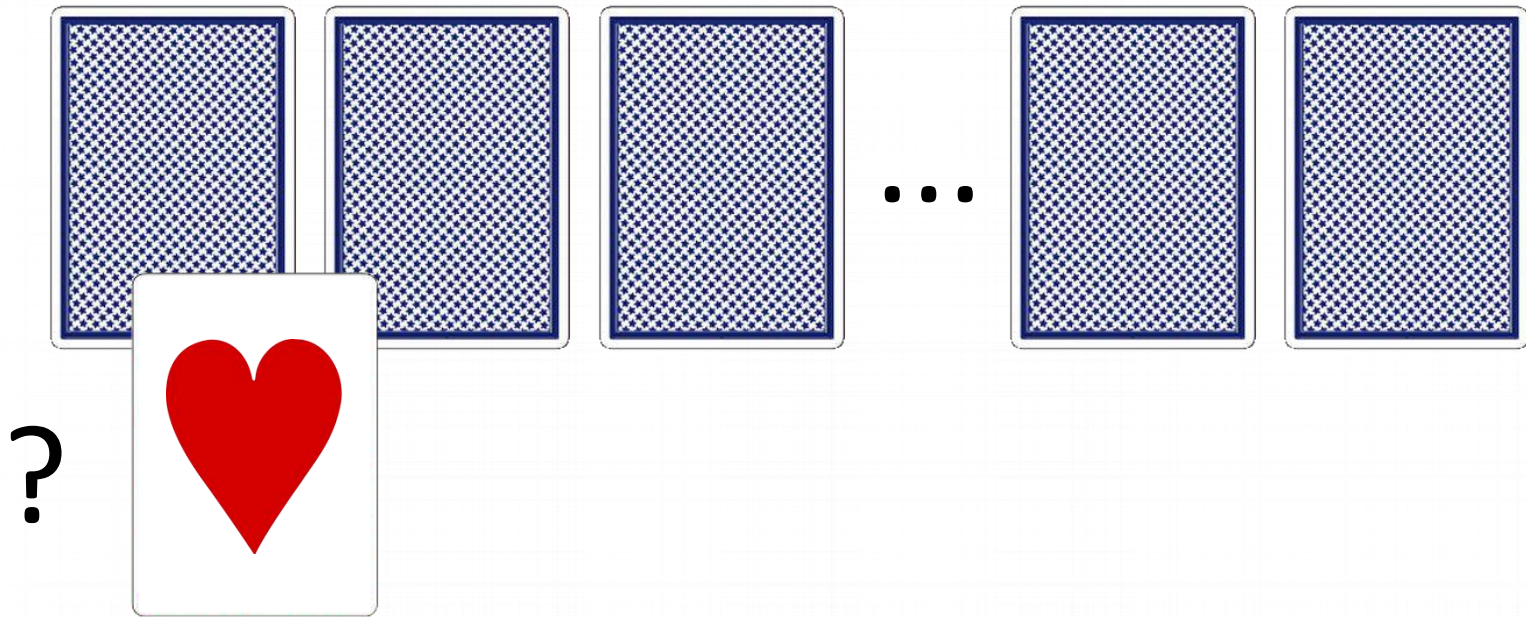
Justin Bieber, ...

- Cannot come from labeled data
- Fuse uncertain information from many web pages

⇒ Embrace probability!

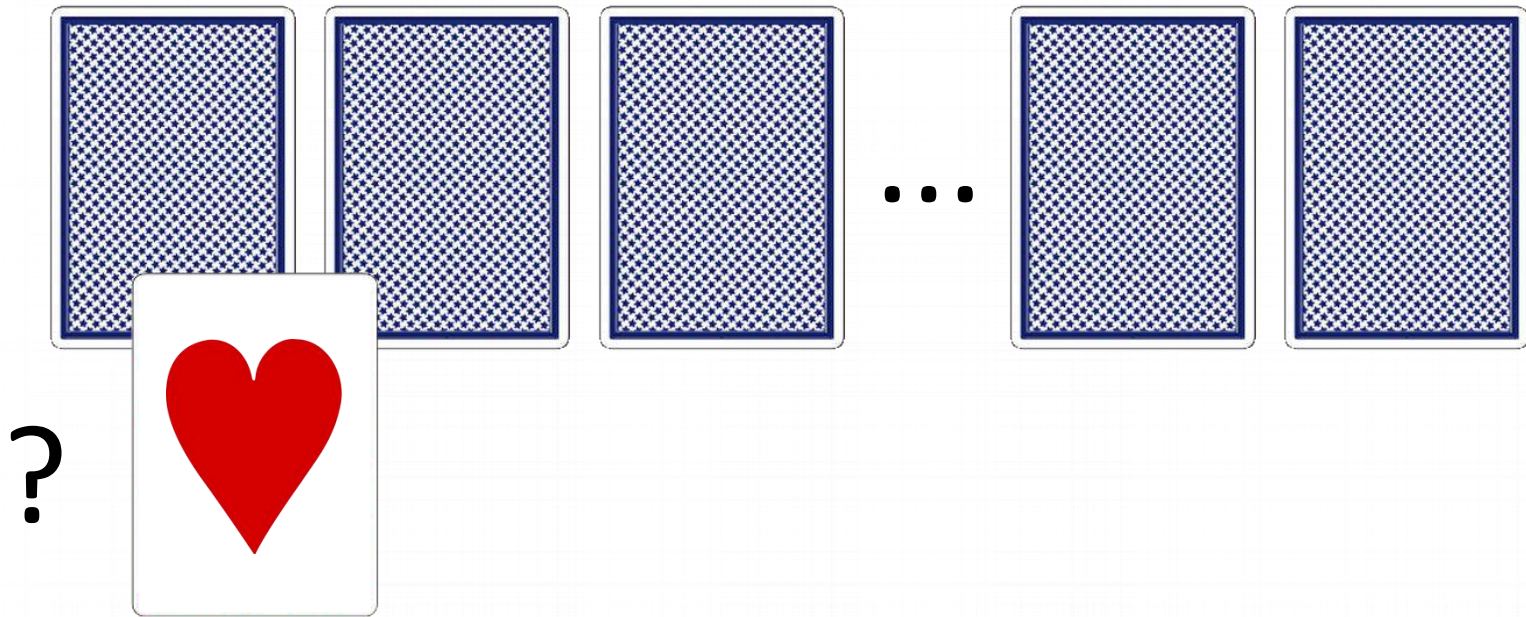
*Why do we need
first-order probabilistic
reasoning?*

A Simple Reasoning Problem



Probability that Card1 is Hearts?

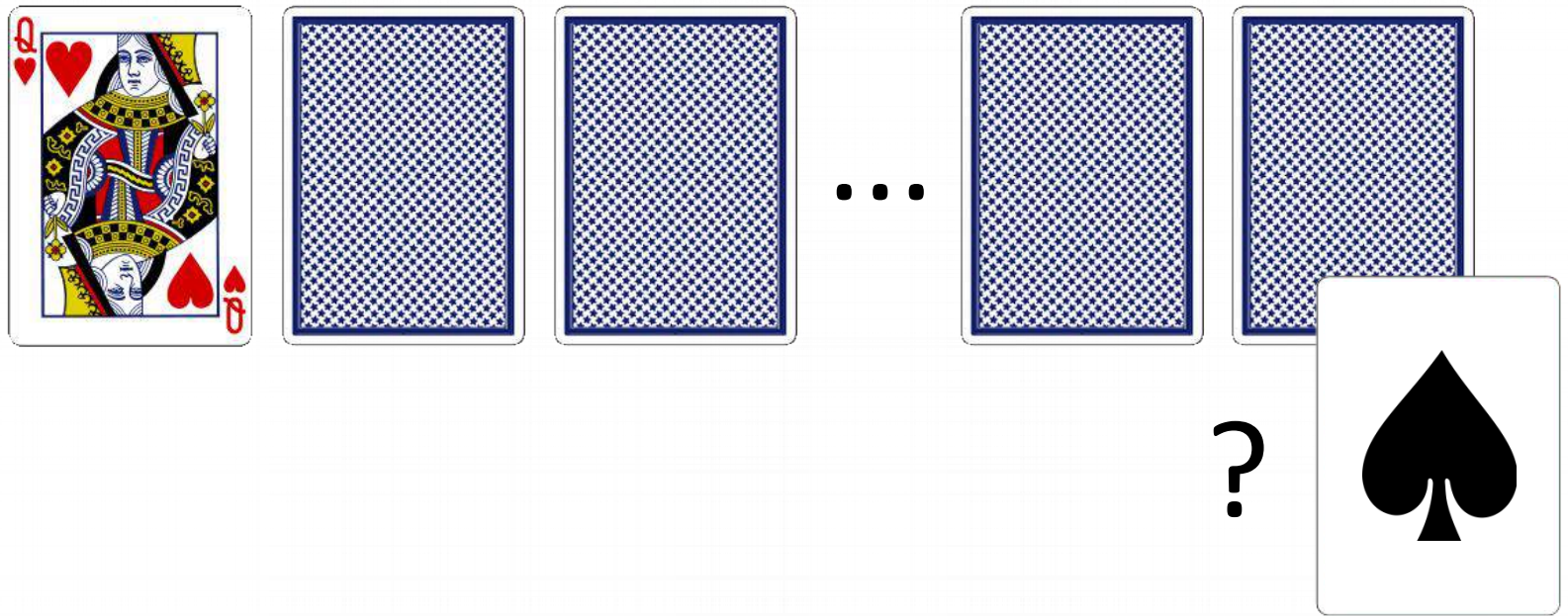
A Simple Reasoning Problem



Probability that Card1 is Hearts?

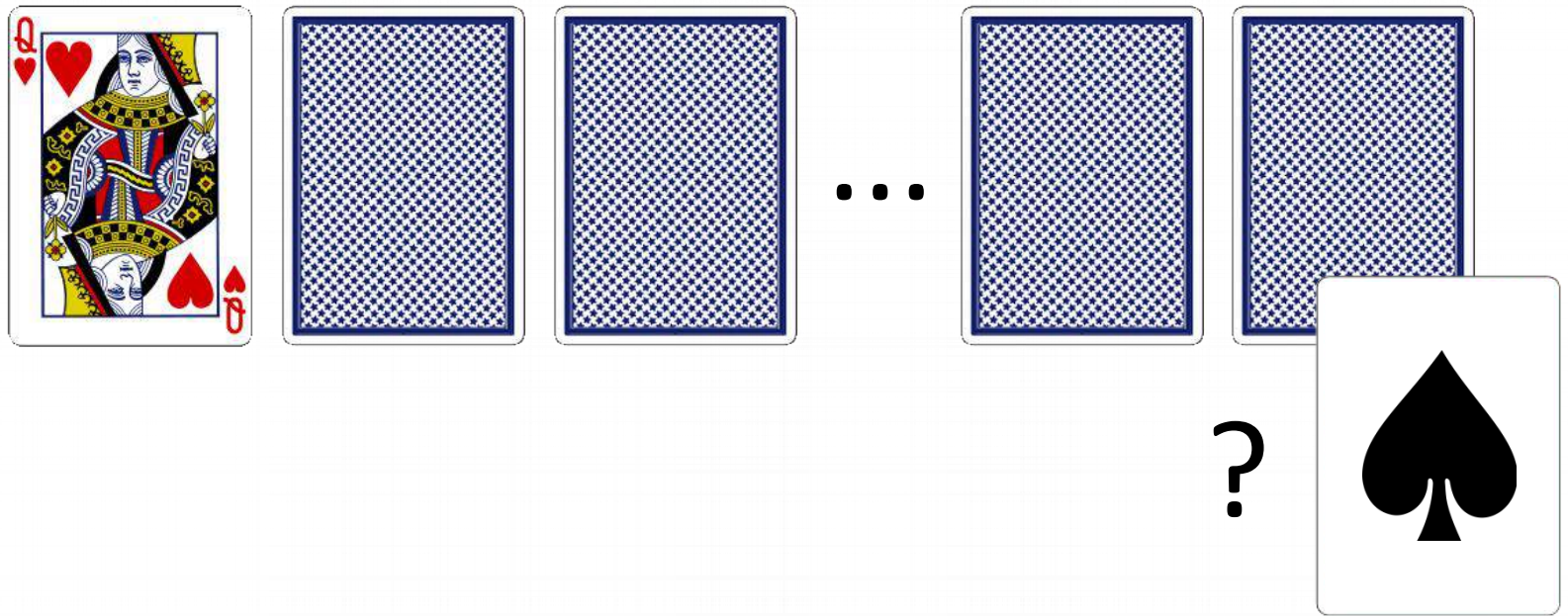
$1/4$

A Simple Reasoning Problem



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

A Simple Reasoning Problem



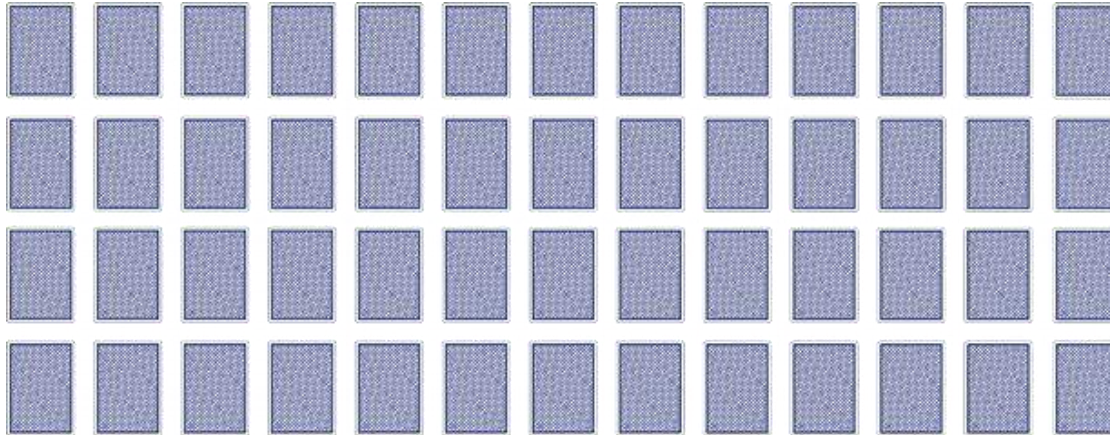
*Probability that Card52 is Spades
given that Card1 is QH?*

13/51

Automated Reasoning

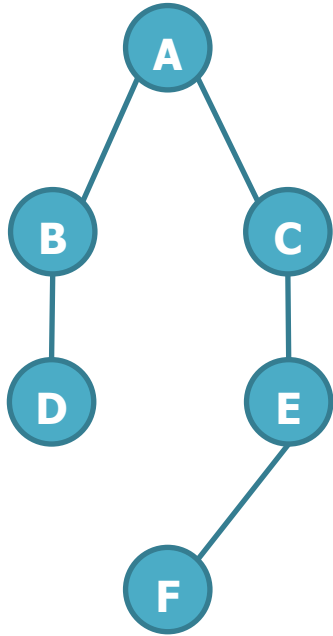
Let us automate this:

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

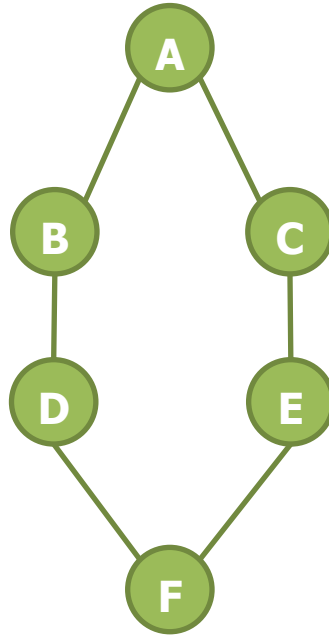


2. Probabilistic inference algorithm
(e.g., variable elimination or junction tree)

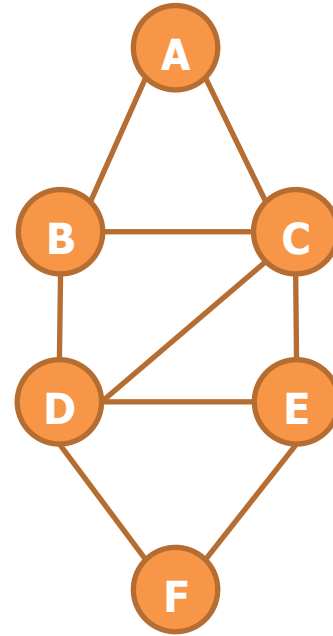
Classical Reasoning



Tree



Sparse Graph



Dense Graph

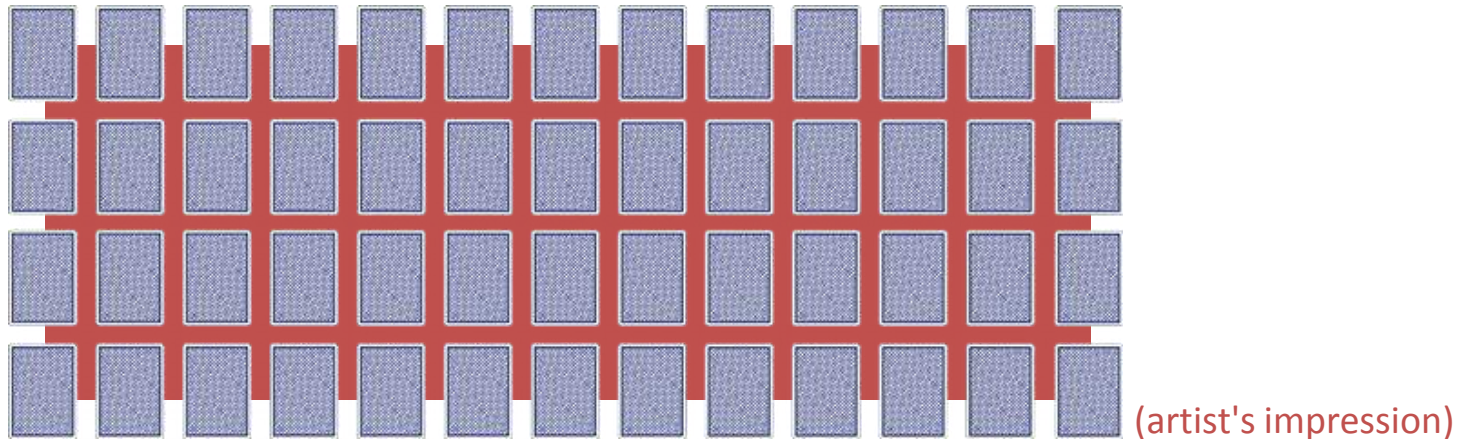


- Higher treewidth
- Fewer conditional independencies
- Slower inference

Automated Reasoning

Let us automate this:

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



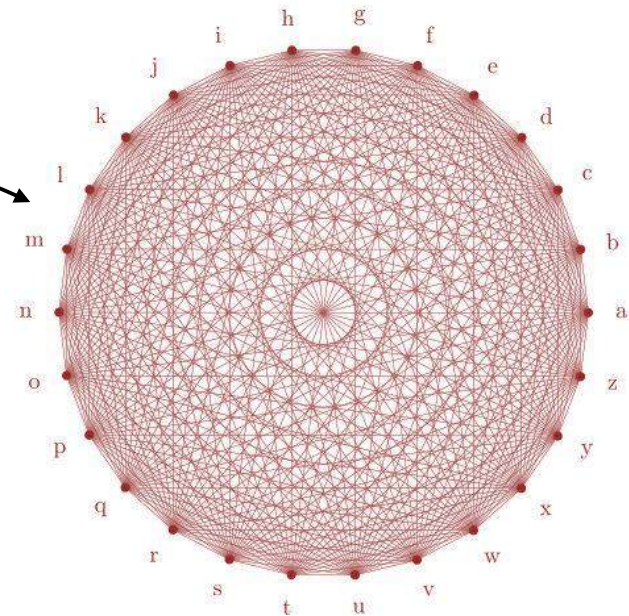
2. Probabilistic inference algorithm
(e.g., variable elimination or junction tree)
builds a table with 52^{52} rows

Lifted Inference in SRL

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

3.14 $\text{FacultyPage}(x) \wedge \text{Linked}(x,y) \Rightarrow \text{CoursePage}(y)$

- As a probabilistic graphical model:
 - 26 pages; 728 variables; 676 factors
 - 1000 pages; 1,002,000 variables; 1,000,000 factors
- Highly intractable?
 - **Lifted inference** in milliseconds!



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

Uncertainty in AI

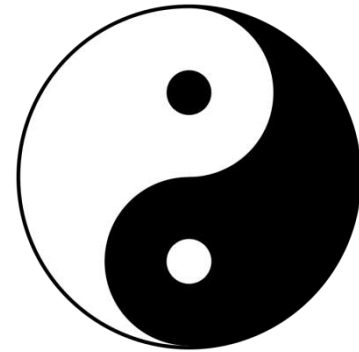
Probability Distribution

=

Qualitative

+

Quantitative



Probabilistic Graphical Models

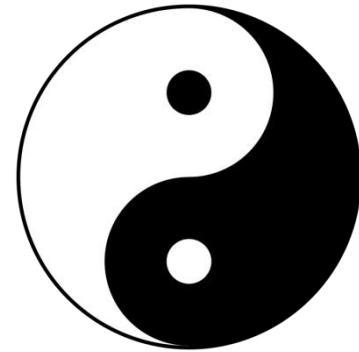
Probability Distribution

=

Graph Structure

+

Parameterization



Probabilistic Graphical Models

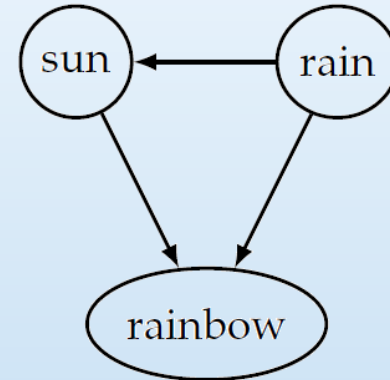
Probability Distribution

=

Graph Structure

+

Parameterization



+

rain	Pr(sun rain)
T	0.1
F	0.6

rain	sun	Pr(rainbow rain, sun)
T	T	0.9
T	F	0.05
F	T	0.05
F	F	0

Pr(rain)
0.2

Model Counting

- Model = solution to a propositional logic formula Δ
- Model counting = #SAT

$\Delta = (\text{Rain} \Rightarrow \text{Cloudy})$

Rain	Cloudy	Model?
T	T	Yes
T	F	No
F	T	Yes
F	F	Yes

+

 #SAT = 3

Weighted Model Counting

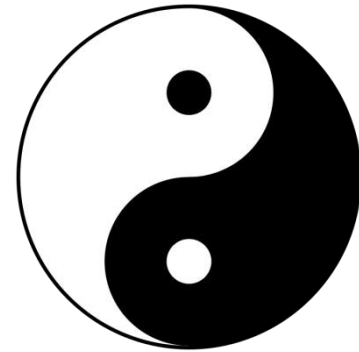
Probability Distribution

=

SAT Formula

+

Weights



Weighted Model Counting

Probability Distribution

=

SAT Formula

+

Weights

Rain \Rightarrow Cloudy
Sun \wedge Rain \Rightarrow Rainbow

+

$w(\text{Rain})=1$

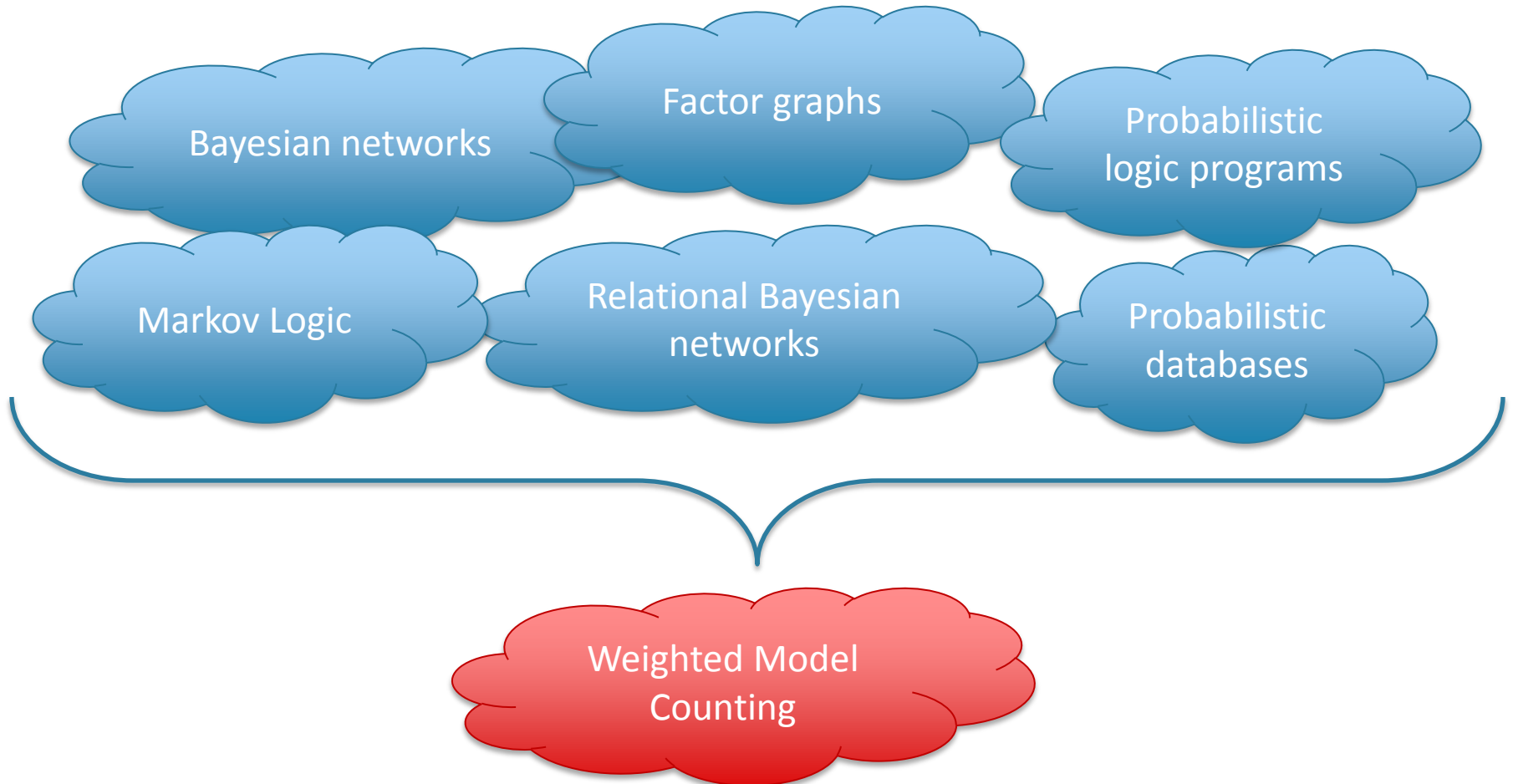
$w(\neg\text{Rain})=2$

$w(\text{Cloudy})=3$

$w(\neg\text{Cloudy})=5$

...

Assembly language for probabilistic reasoning



First-Order Model Counting

Model = solution to
first-order logic
formula Δ

$$\Delta = \forall d (\text{Rain}(d) \Rightarrow \text{Cloudy}(d))$$

$$\text{Days} = \{\text{Monday}\}$$

First-Order Model Counting

Model = solution to
first-order logic
formula Δ

$$\Delta = \forall d (\text{Rain}(d) \Rightarrow \text{Cloudy}(d))$$

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

$\Delta = \forall d (\text{Rain}(d)$
 $\Rightarrow \text{Cloudy}(d))$

Days = {Monday
Tuesday}

First-Order Model Counting

Model = solution to
first-order logic
 formula Δ

$$\Delta = \forall d (\text{Rain}(d) \Rightarrow \text{Cloudy}(d))$$

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

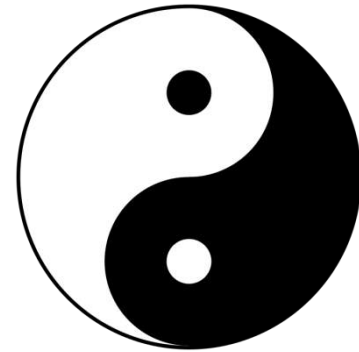
Probability Distribution

=

First-Order Logic

+

Weights



Weighted First-Order Model Counting

Probability Distribution

=

First-Order Logic

+

Weights

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

+

$w(\text{Smokes}(a))=1$

$w(\neg\text{Smokes}(a))=2$

$w(\text{Smokes}(b))=1$

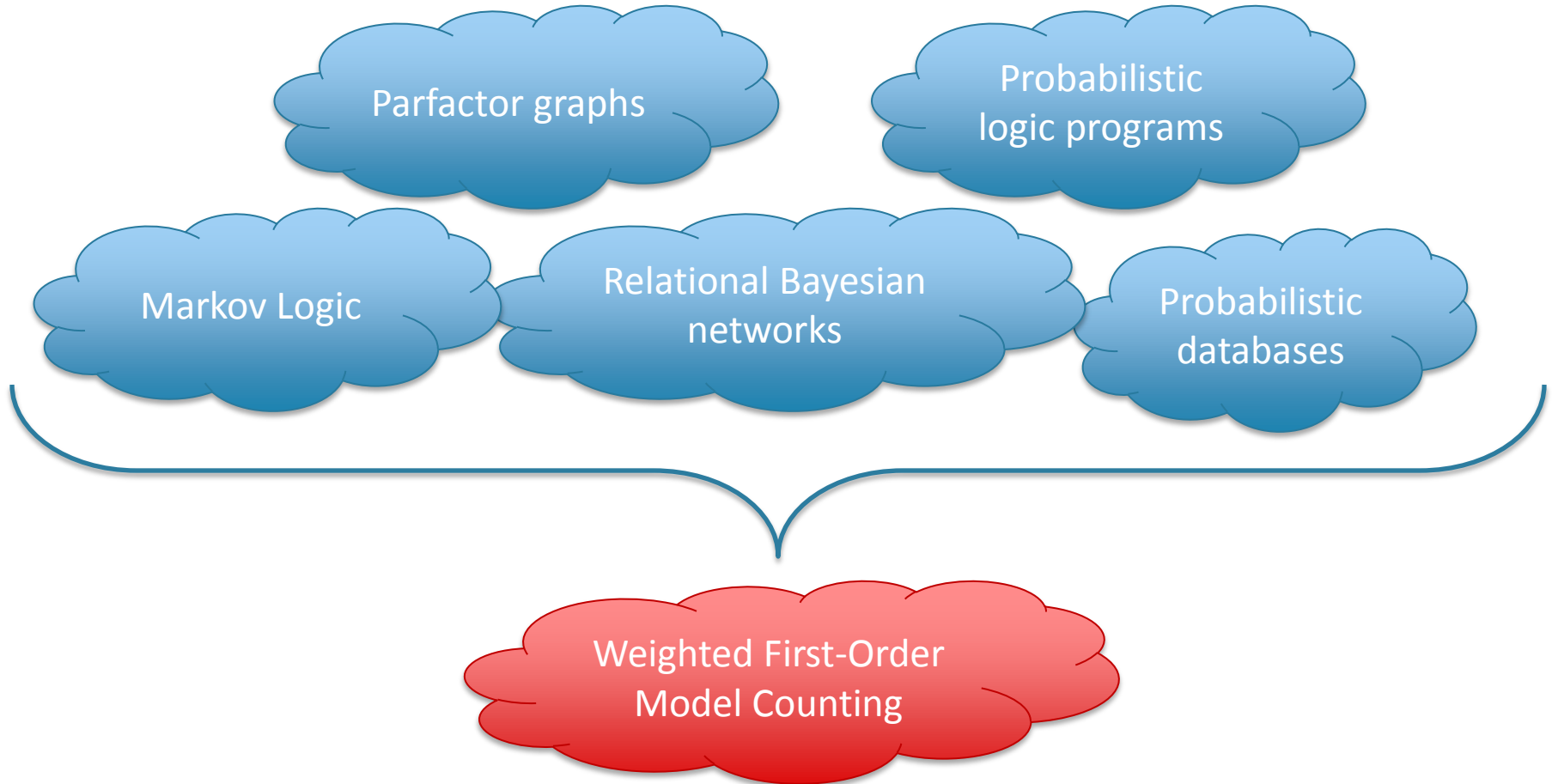
$w(\neg\text{Smokes}(b))=2$

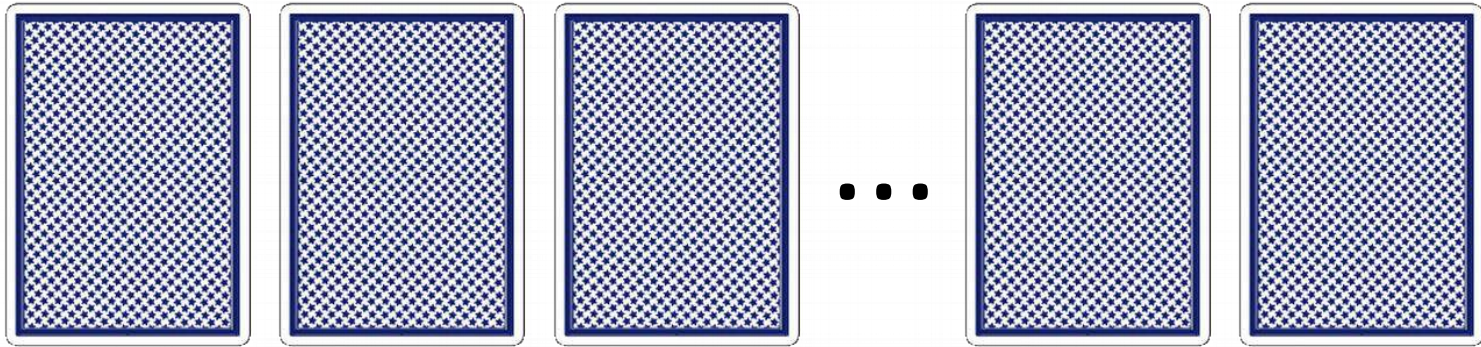
$w(\text{Friends}(a,b))=3$

$w(\neg\text{Friends}(a,b))=5$

...

Assembly language for **first-order** probabilistic reasoning





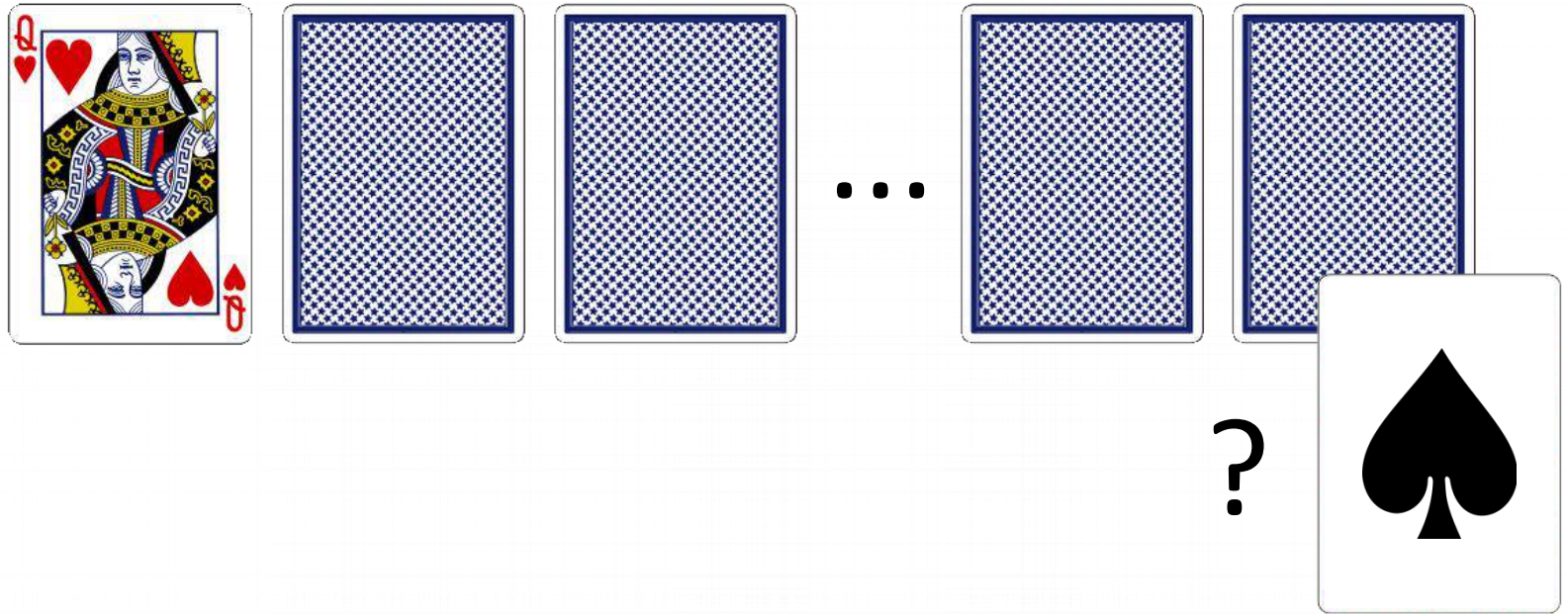
Let us automate this:

- **Relational** model

$$\begin{aligned} & \forall p, \exists c, \text{Card}(p,c) \\ & \forall c, \exists p, \text{Card}(p,c) \\ & \forall p, \forall c, \forall c', \text{Card}(p,c) \wedge \text{Card}(p,c') \Rightarrow c = c' \end{aligned}$$

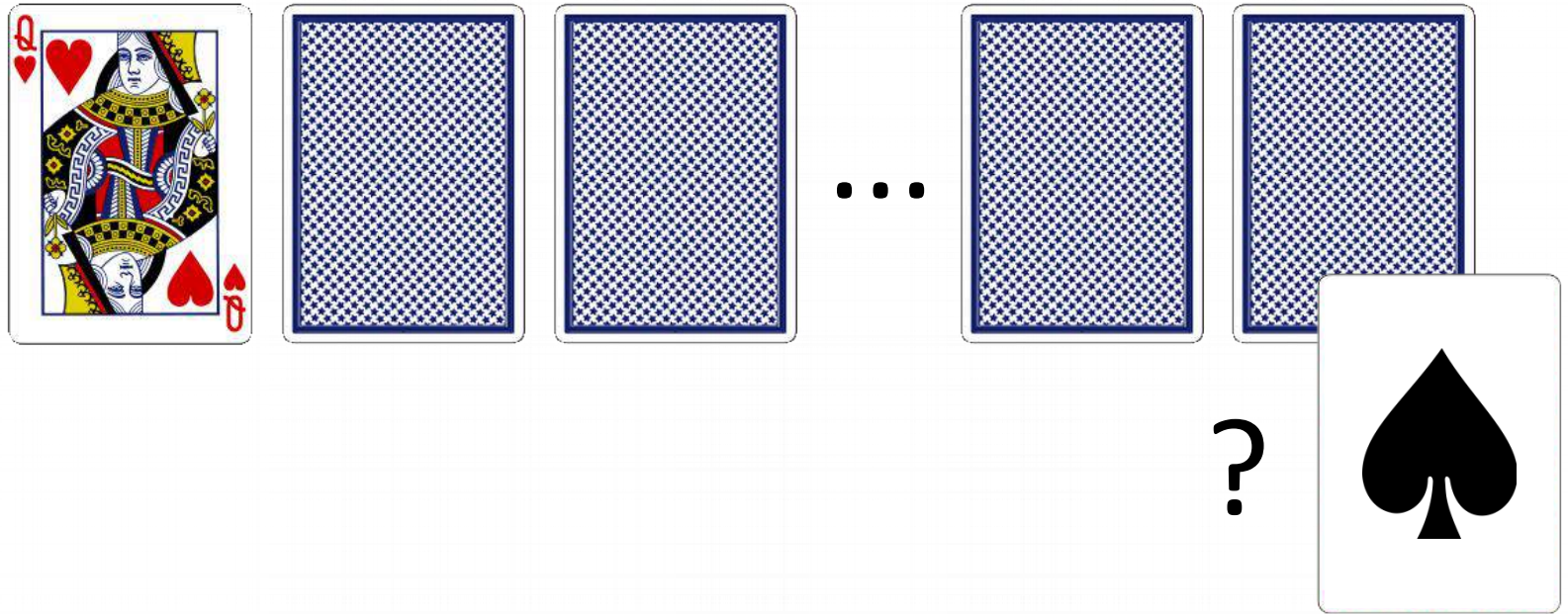
- **Lifted** probabilistic inference algorithm

What's Going On Here?



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

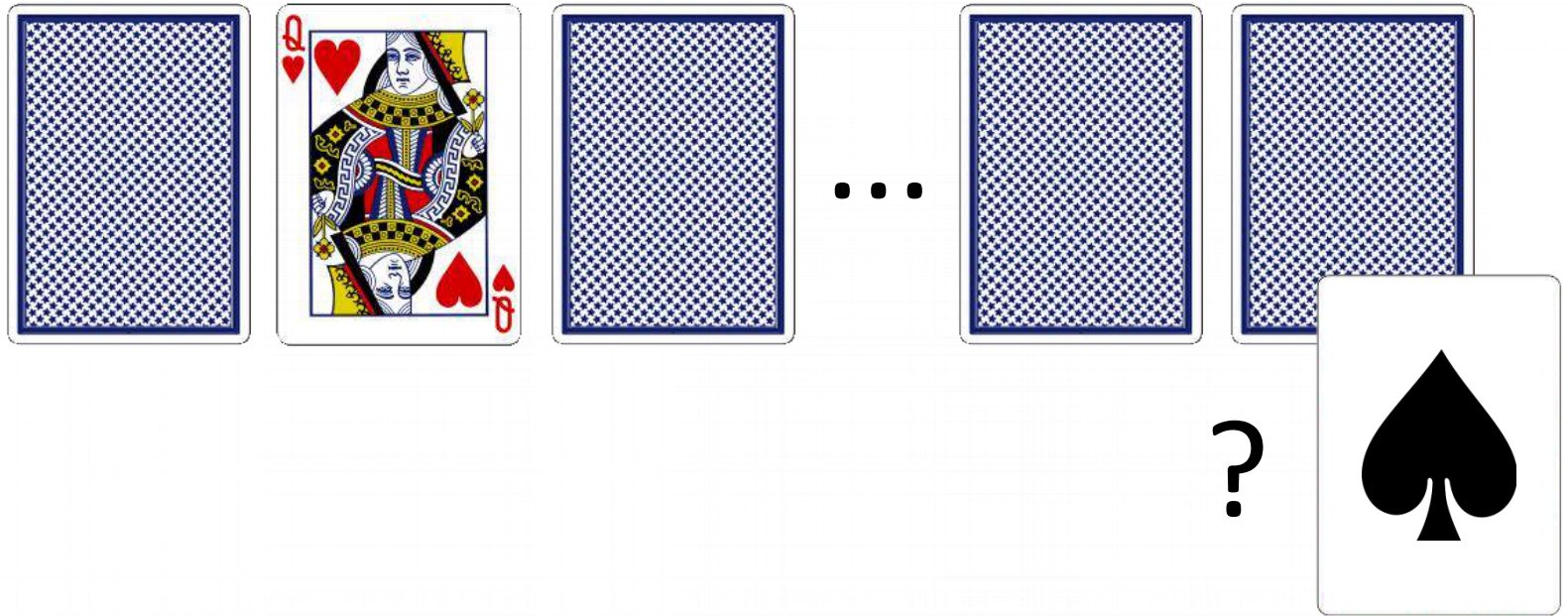
What's Going On Here?



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

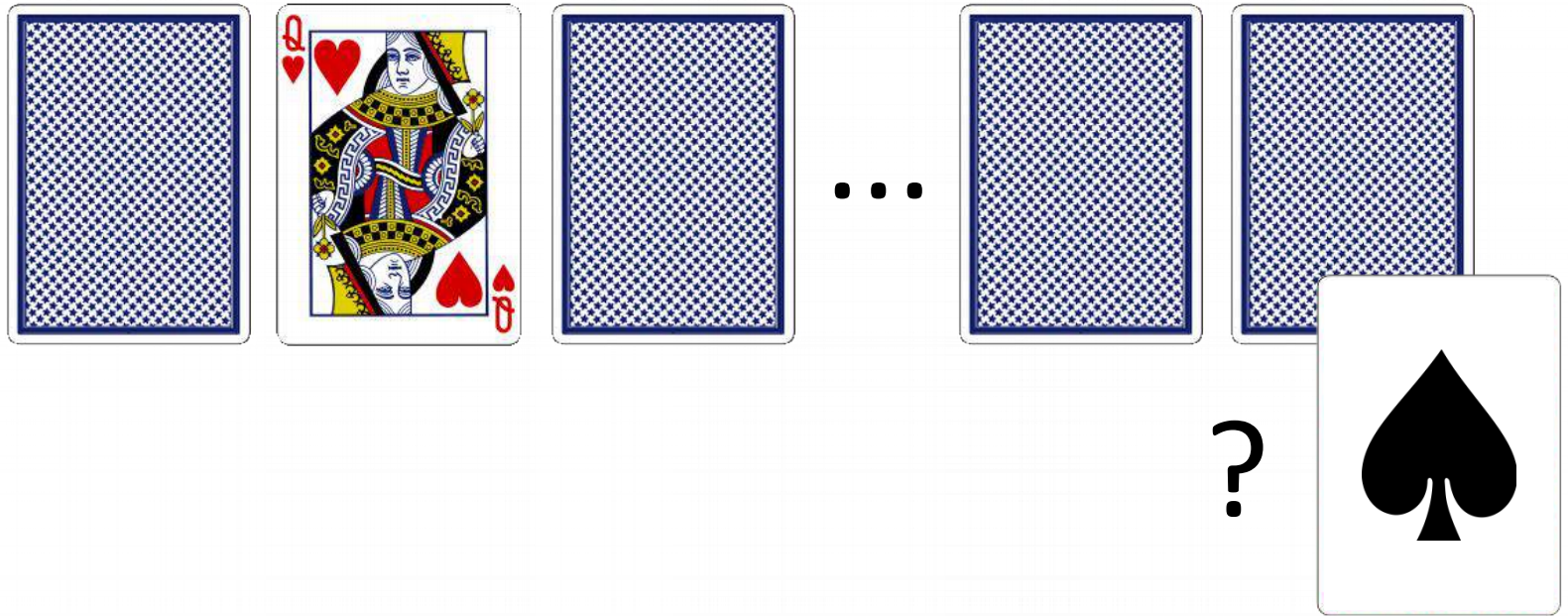
13/51

What's Going On Here?



*Probability that Card52 is Spades
given that Card2 is QH?*

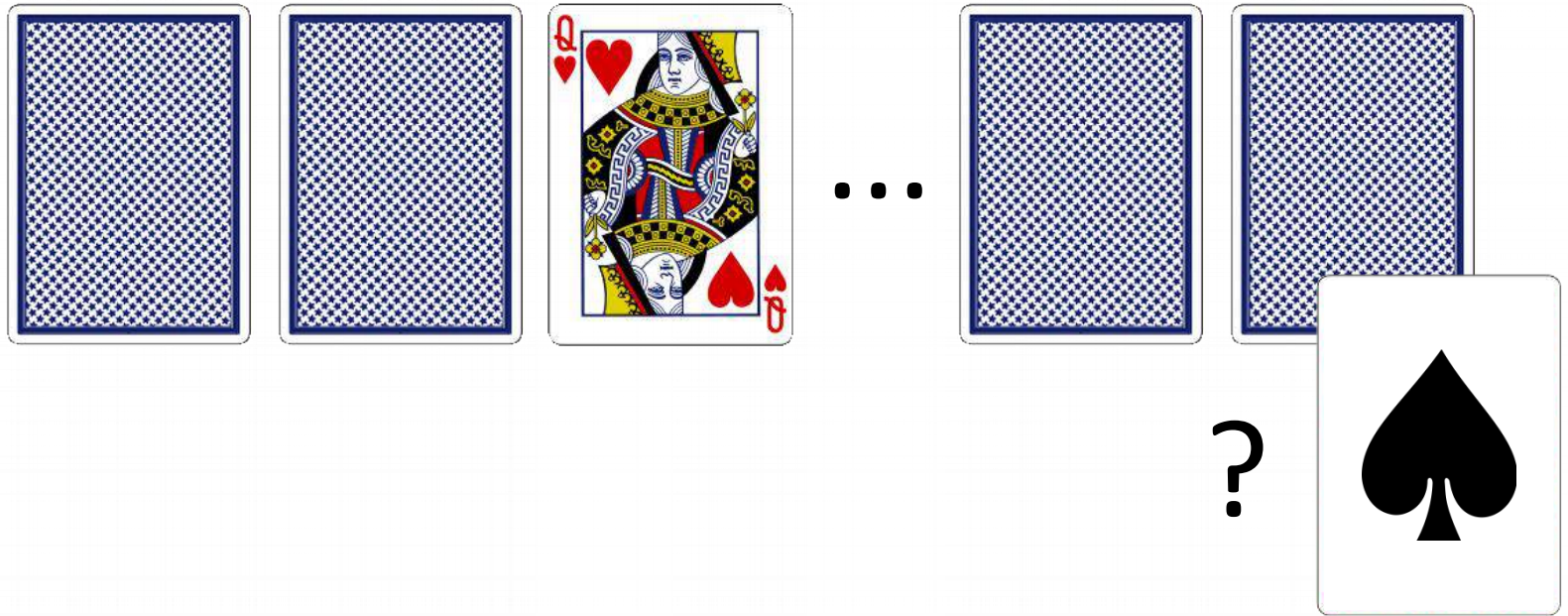
What's Going On Here?



*Probability that Card52 is Spades
given that Card2 is QH?*

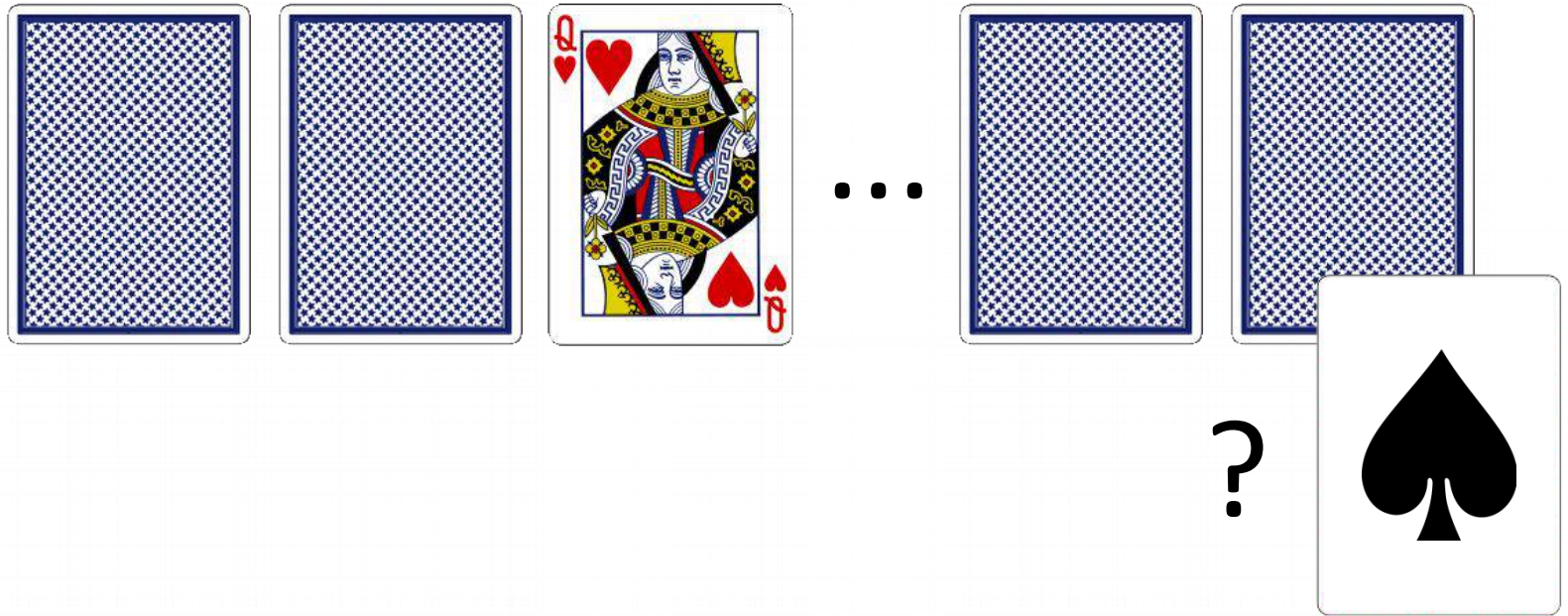
13/51

What's Going On Here?



*Probability that Card52 is Spades
given that Card3 is QH?*

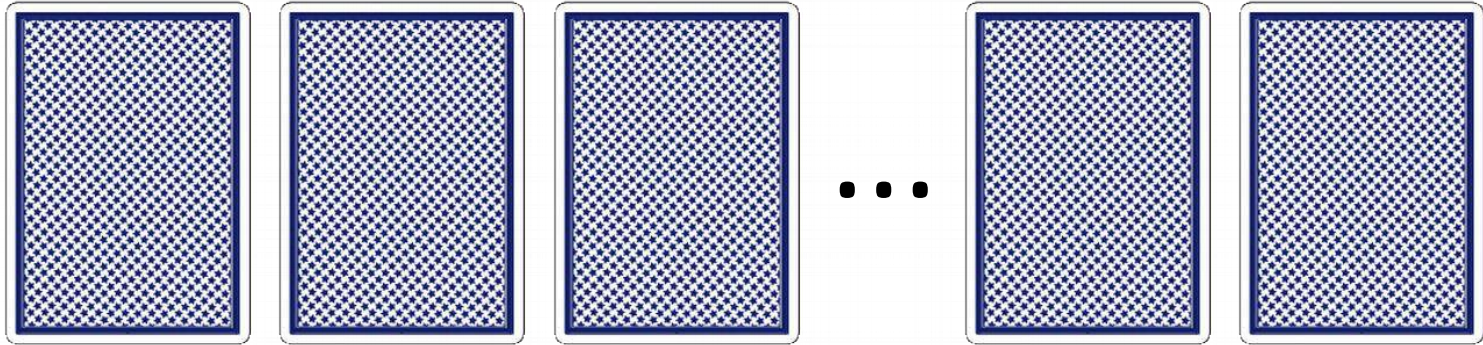
What's Going On Here?



*Probability that Card52 is Spades
given that Card3 is QH?*

13/51

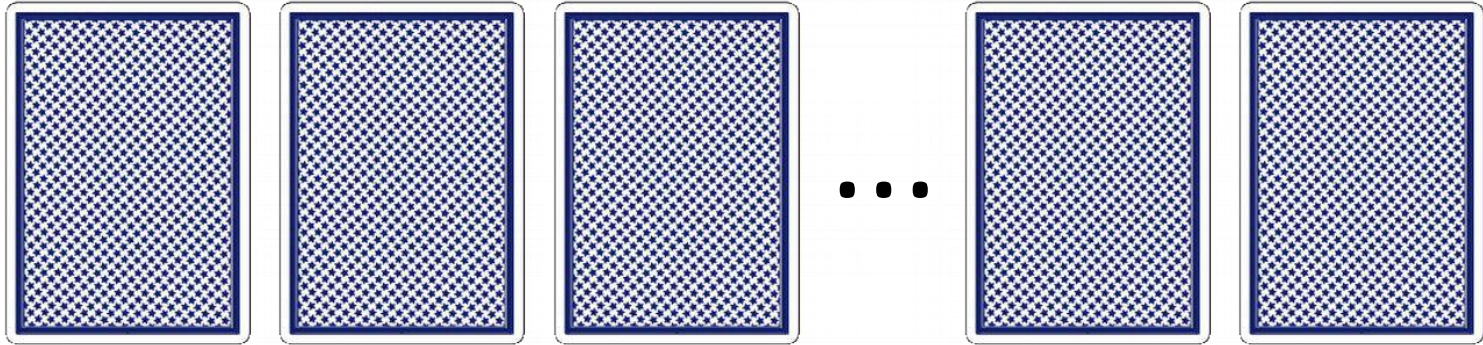
Tractable Reasoning



What's going on here?

Which property makes reasoning tractable?

Tractable Reasoning



What's going on here?

Which property makes reasoning tractable?

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

⇒ **Lifted Inference**

FOMC Inference

3. $\Delta = \forall x, (\text{Stress}(x) \Rightarrow \text{Smokes}(x))$

Domain = {n people}

FOMC Inference

3. $\Delta = \forall x, (\text{Stress}(x) \Rightarrow \text{Smokes}(x))$

Domain = {n people}

$\rightarrow 3^n$ models

2. $\Delta = \forall y, (\text{ParentOf}(y) \wedge \text{Female} \Rightarrow \text{MotherOf}(y))$

D = {n people}

FOMC Inference

3. $\Delta = \forall x, (\text{Stress}(x) \Rightarrow \text{Smokes}(x))$

Domain = {n people}

$\rightarrow 3^n$ models

2. $\Delta = \forall y, (\text{ParentOf}(y) \wedge \text{Female} \Rightarrow \text{MotherOf}(y))$

D = {n people}

If Female = true?

$$\Delta = \forall y, (\text{ParentOf}(y) \Rightarrow \text{MotherOf}(y))$$

$\rightarrow 3^n$ models

FOMC Inference

3. $\Delta = \forall x, (\text{Stress}(x) \Rightarrow \text{Smokes}(x))$

Domain = {n people}

$\rightarrow 3^n$ models

2. $\Delta = \forall y, (\text{ParentOf}(y) \wedge \text{Female} \Rightarrow \text{MotherOf}(y))$

D = {n people}

If Female = true?

$\Delta = \forall y, (\text{ParentOf}(y) \Rightarrow \text{MotherOf}(y))$

$\rightarrow 3^n$ models

If Female = false?

$\Delta = \text{true}$

$\rightarrow 4^n$ models

FOMC Inference

3. $\Delta = \forall x, (\text{Stress}(x) \Rightarrow \text{Smokes}(x))$

Domain = {n people}

$\rightarrow 3^n$ models

2. $\Delta = \forall y, (\text{ParentOf}(y) \wedge \text{Female} \Rightarrow \text{MotherOf}(y))$

D = {n people}

If Female = true?

$\Delta = \forall y, (\text{ParentOf}(y) \Rightarrow \text{MotherOf}(y))$

$\rightarrow 3^n$ models

If Female = false?

$\Delta = \text{true}$

$\rightarrow 4^n$ models

$\rightarrow 3^n + 4^n$ models

1. $\Delta = \forall x, y, (\text{ParentOf}(x, y) \wedge \text{Female}(x) \Rightarrow \text{MotherOf}(x, y))$

D = {n people}

FOMC Inference

3. $\Delta = \forall x, (\text{Stress}(x) \Rightarrow \text{Smokes}(x))$

Domain = {n people}

$\rightarrow 3^n$ models

2. $\Delta = \forall y, (\text{ParentOf}(y) \wedge \text{Female} \Rightarrow \text{MotherOf}(y))$

D = {n people}

If Female = true? $\Delta = \forall y, (\text{ParentOf}(y) \Rightarrow \text{MotherOf}(y))$

$\rightarrow 3^n$ models

If Female = false? $\Delta = \text{true}$

$\rightarrow 4^n$ models

$\rightarrow 3^n + 4^n$ models

1. $\Delta = \forall x, y, (\text{ParentOf}(x, y) \wedge \text{Female}(x) \Rightarrow \text{MotherOf}(x, y))$

D = {n people}

$\rightarrow (3^n + 4^n)^n$ models

FOMC Inference

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

Domain = {n people}

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



FOMC Inference

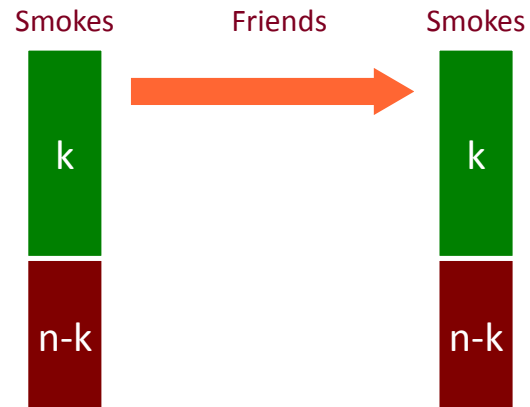
$$\Delta = \forall x,y, (\text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y))$$

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

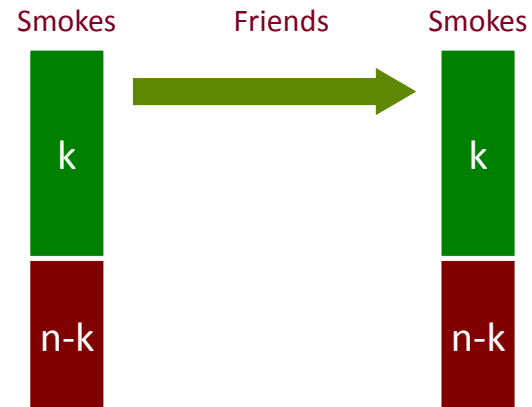
$$\Delta = \forall x,y, (\text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y))$$

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

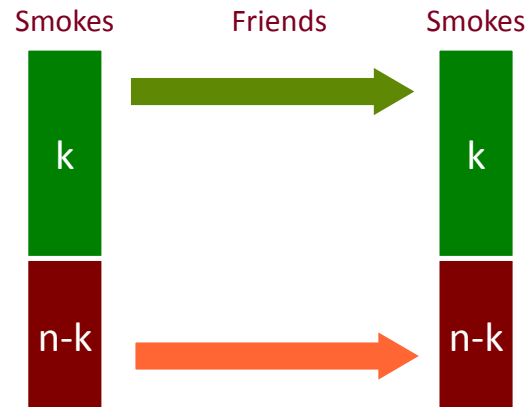
$$\Delta = \forall x,y, (\text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y))$$

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

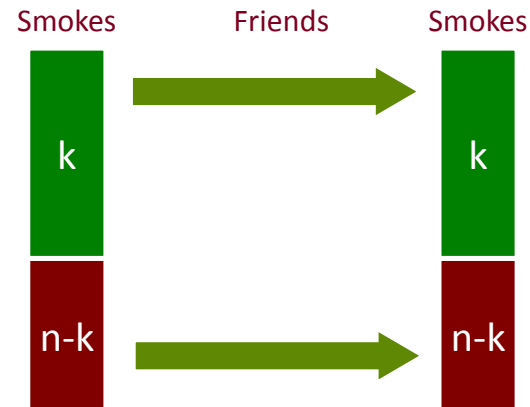
$$\Delta = \forall x,y, (\text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y))$$

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

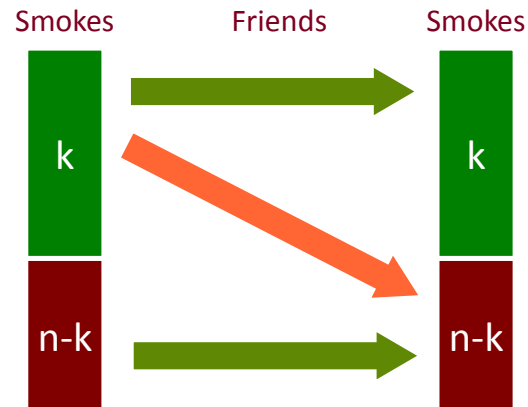
$$\Delta = \forall x,y, (\text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y))$$

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

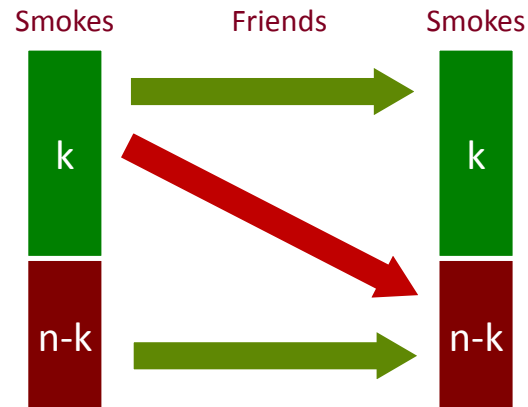
$$\Delta = \forall x,y, (\text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y))$$

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

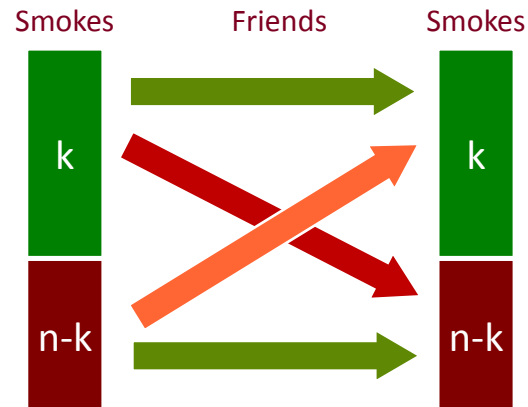
$$\Delta = \forall x,y, (\text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y))$$

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

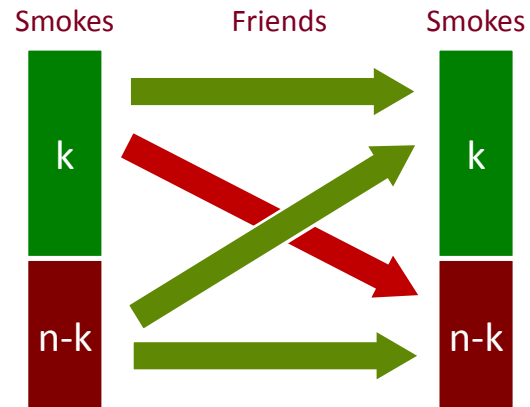
$$\Delta = \forall x,y, (\text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y))$$

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

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

$$\text{Domain} = \{n \text{ people}\}$$

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

Database:

Smokes(Alice) = 1

Smokes(Bob) = 0

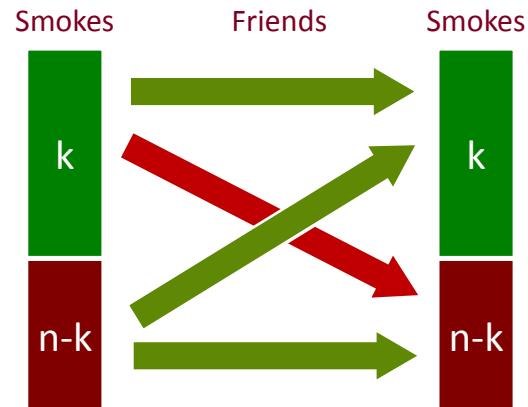
Smokes(Charlie) = 0

Smokes(Dave) = 1

Smokes(Eve) = 0

...

→ $2^{n^2 - k(n-k)}$ models



FOMC Inference

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

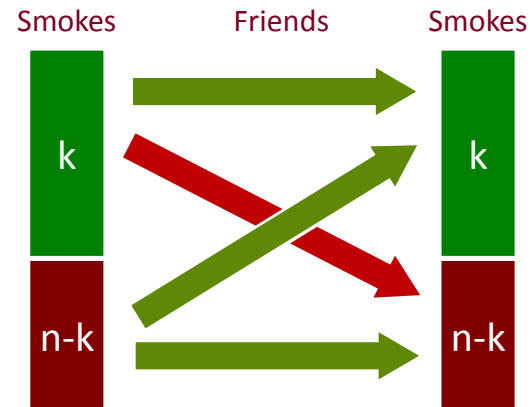
$$\text{Domain} = \{n \text{ people}\}$$

- 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?

FOMC Inference

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

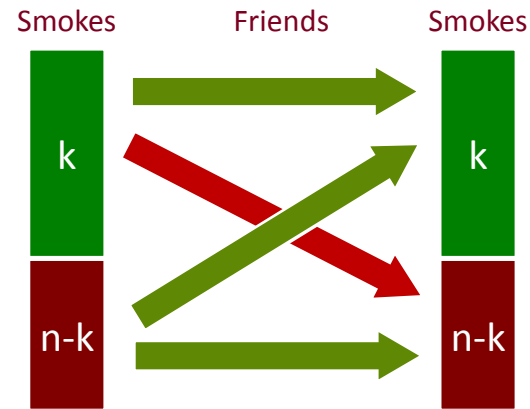
$$\text{Domain} = \{n \text{ people}\}$$

- 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

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

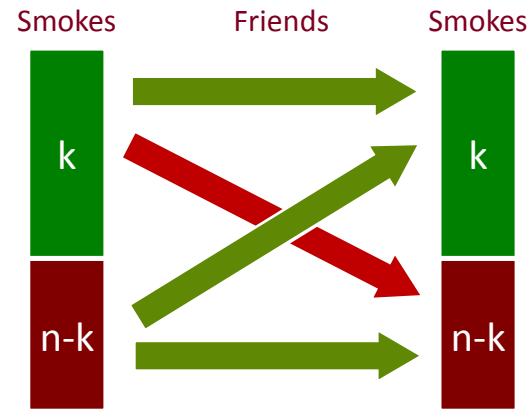
Domain = {n people}

- 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



- If we know that there are k smokers?

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

- In total...

FOMC Inference

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

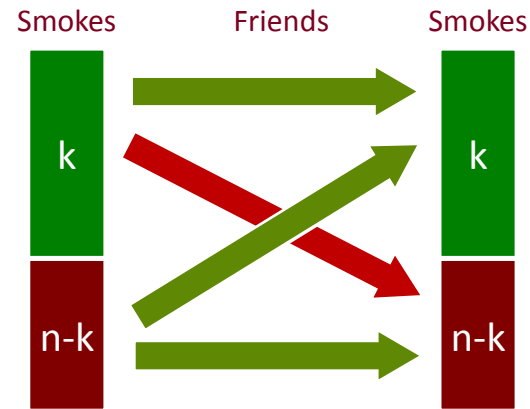
$$\text{Domain} = \{n \text{ people}\}$$

- 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
 ...

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



- If we know that there are k smokers?

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

- In total...

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

What are the
successes?

First-Order Knowledge Compilation

Markov Logic

3.14 $\text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y)$

First-Order Knowledge Compilation

Markov Logic

$$3.14 \quad \text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y)$$

Weight Function

$$\begin{aligned} w(\text{Smokes}) &= 1 \\ w(\neg \text{Smokes}) &= 1 \\ w(\text{Friends}) &= 1 \\ w(\neg \text{Friends}) &= 1 \\ w(F) &= 3.14 \\ w(\neg F) &= 1 \end{aligned}$$

FOL Sentence

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

First-Order Knowledge Compilation

Markov Logic

$$3.14 \quad \text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y)$$

Weight Function

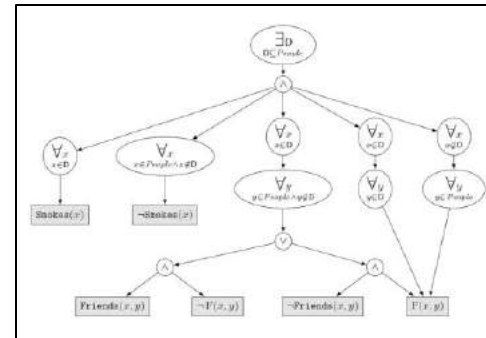
$$\begin{aligned} w(\text{Smokes}) &= 1 \\ w(\neg \text{Smokes}) &= 1 \\ w(\text{Friends}) &= 1 \\ w(\neg \text{Friends}) &= 1 \\ w(F) &= 3.14 \\ w(\neg F) &= 1 \end{aligned}$$

FOL Sentence

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

Compile?

First-Order d-DNNF Circuit



First-Order Knowledge Compilation

Markov Logic

$$3.14 \quad \text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y)$$

Weight Function

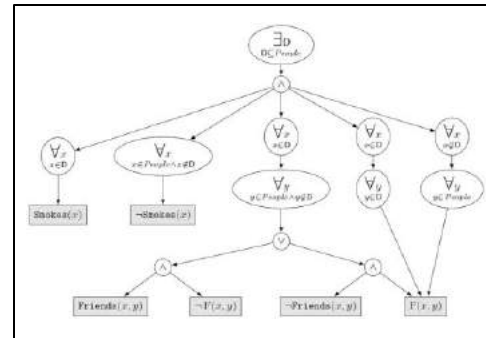
$$\begin{aligned} w(\text{Smokes}) &= 1 \\ w(\neg \text{Smokes}) &= 1 \\ w(\text{Friends}) &= 1 \\ w(\neg \text{Friends}) &= 1 \\ w(F) &= 3.14 \\ w(\neg F) &= 1 \end{aligned}$$

FOL Sentence

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

Compile?

First-Order d-DNNF Circuit



Domain

Alice
Bob
Charlie

First-Order Knowledge Compilation

Markov Logic

$$3.14 \quad \text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y)$$

Weight Function

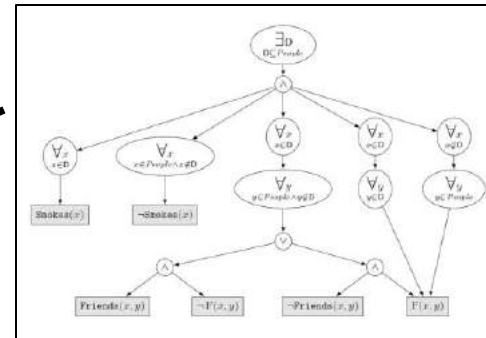
$$\begin{aligned} w(\text{Smokes}) &= 1 \\ w(\neg \text{Smokes}) &= 1 \\ w(\text{Friends}) &= 1 \\ w(\neg \text{Friends}) &= 1 \\ w(F) &= 3.14 \\ w(\neg F) &= 1 \end{aligned}$$

FOL Sentence

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

Compile?

First-Order d-DNNF Circuit



Domain

Alice
Bob
Charlie

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

First-Order Knowledge Compilation

Markov Logic

$$3.14 \quad \text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y)$$

Weight Function

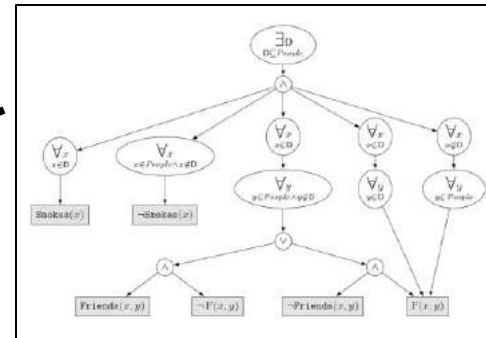
$$\begin{aligned} w(\text{Smokes}) &= 1 \\ w(\neg \text{Smokes}) &= 1 \\ w(\text{Friends}) &= 1 \\ w(\neg \text{Friends}) &= 1 \\ w(F) &= 3.14 \\ w(\neg F) &= 1 \end{aligned}$$

FOL Sentence

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

Compile?

First-Order d-DNNF Circuit



Domain

Alice
Bob
Charlie

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

Evaluation in time polynomial in domain size!

First-Order Knowledge Compilation

Markov Logic

$$3.14 \quad \text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y)$$

Weight Function

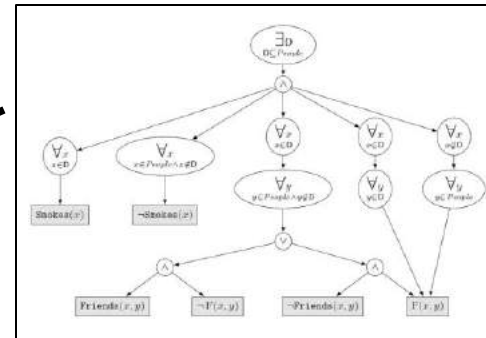
$$\begin{aligned} w(\text{Smokes}) &= 1 \\ w(\neg \text{Smokes}) &= 1 \\ w(\text{Friends}) &= 1 \\ w(\neg \text{Friends}) &= 1 \\ w(F) &= 3.14 \\ w(\neg F) &= 1 \end{aligned}$$

FOL Sentence

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

Compile?

First-Order d-DNNF Circuit



Domain

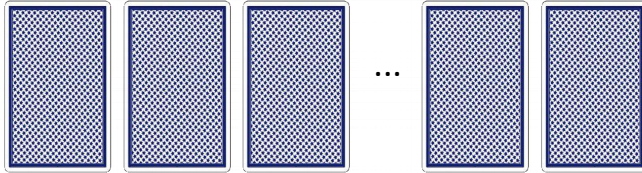
Alice
Bob
Charlie

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

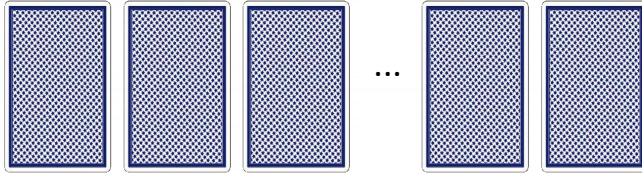
Evaluation in time polynomial in domain size!

= Lifted!

Playing Cards Revisited


$$\forall p, \exists c, \text{Card}(p,c)$$
$$\forall c, \exists p, \text{Card}(p,c)$$
$$\forall p, \forall c, \forall c', \text{Card}(p,c) \wedge \text{Card}(p,c') \Rightarrow c = c'$$

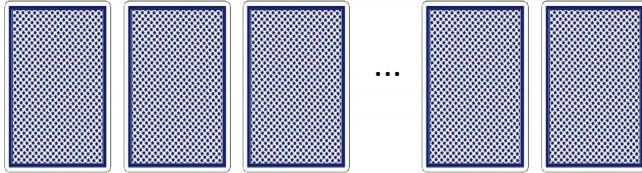
Playing Cards Revisited



$\forall p, \exists c, \text{Card}(p,c)$
 $\forall c, \exists p, \text{Card}(p,c)$
 $\forall p, \forall c, \forall c', \text{Card}(p,c) \wedge \text{Card}(p,c') \Rightarrow c = c'$

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

Playing Cards Revisited



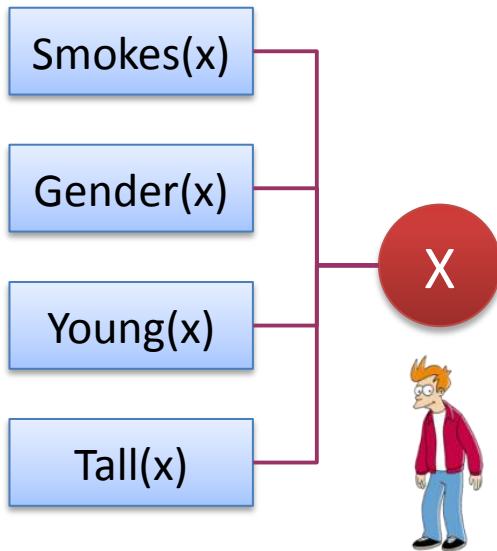
$\forall p, \exists c, \text{Card}(p,c)$
 $\forall c, \exists p, \text{Card}(p,c)$
 $\forall p, \forall c, \forall c', \text{Card}(p,c) \wedge \text{Card}(p,c') \Rightarrow c = c'$

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

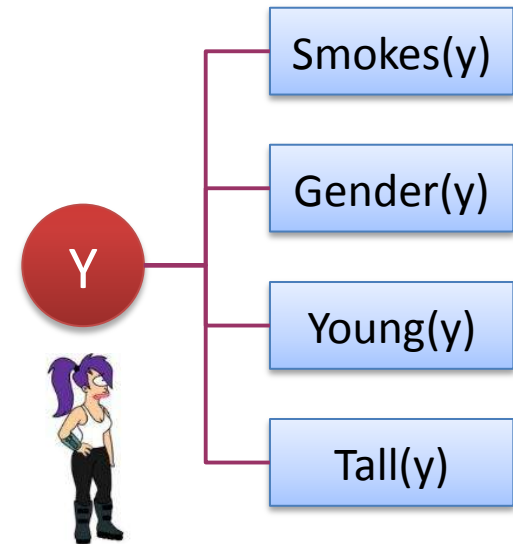
Computed in time polynomial in n

FO² is liftable!

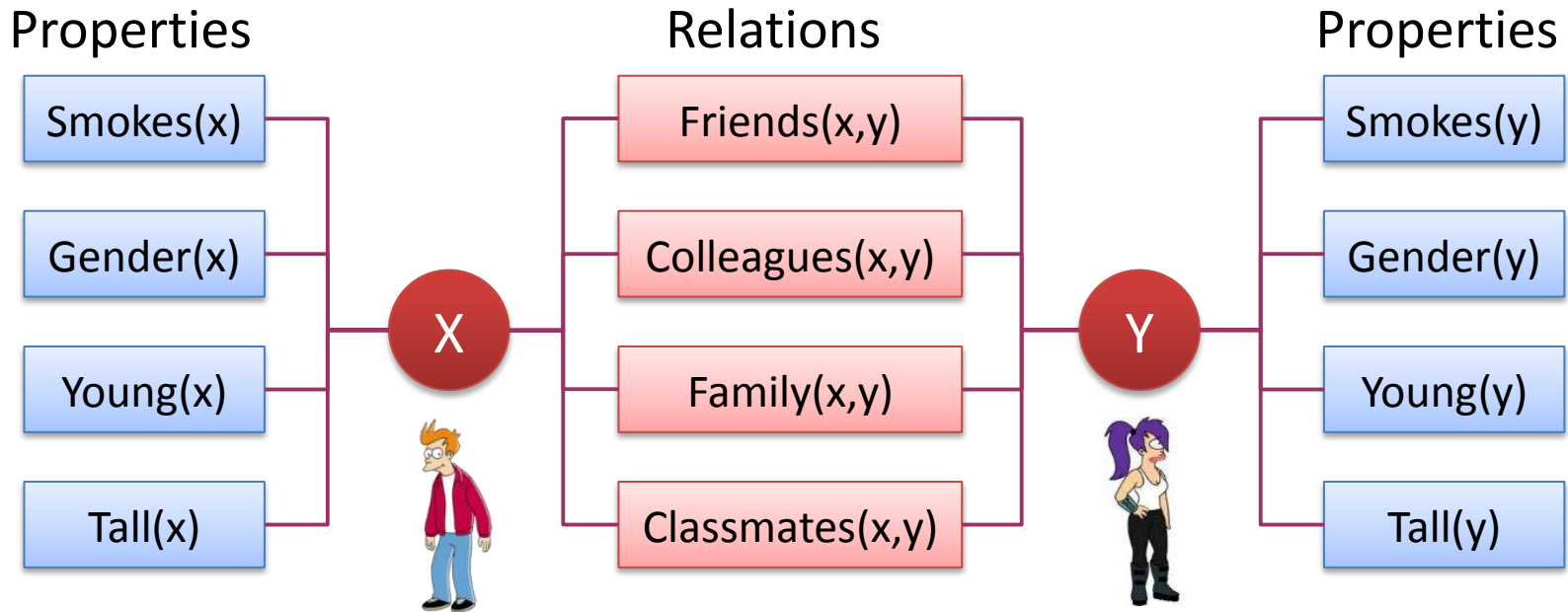
Properties



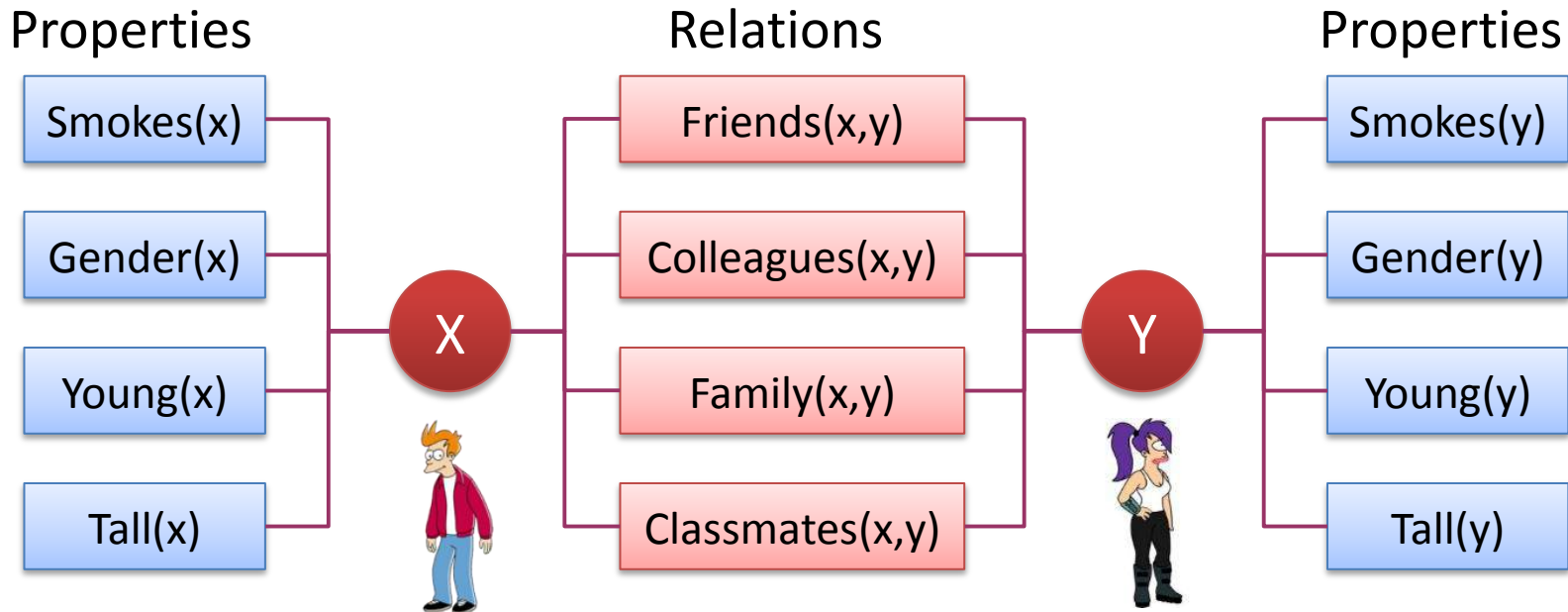
Properties



FO² is liftable!



FO² is liftable!



“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.”

FO² is liftable!



Medical Records

Name	Cough	Asthma	Smokes
Alice	1	1	0
Bob	0	0	0
Charlie	0	1	0
Dave	1	0	1
Eve	1	0	0

Frank	1	?	?
-------	---	---	---



Statistical Relational Model in FO²

- 2.1 $Asthma(x) \Rightarrow Cough(x)$
- 3.5 $Smokes(x) \Rightarrow Cough(x)$
- 1.9 $Smokes(x) \wedge Friends(x,y) \Rightarrow Smokes(y)$
- 1.5 $Asthma(x) \wedge Family(x,y) \Rightarrow Asthma(y)$

Frank	1	0.2	0.6
-------	---	-----	-----

FO² is liftable!



Medical Records

Name	Cough	Asthma	Smokes
Alice	1	1	0
Bob	0	0	0
Charlie	0	1	0
Dave	1	0	1
Eve	1	0	0

Frank	1	?	?
-------	---	---	---

Big data



Statistical Relational Model in FO²

- 2.1 $Asthma(x) \Rightarrow Cough(x)$
- 3.5 $Smokes(x) \Rightarrow Cough(x)$
- 1.9 $Smokes(x) \wedge Friends(x,y) \Rightarrow Smokes(y)$
- 1.5 $Asthma(x) \wedge Family(x,y) \Rightarrow Asthma(y)$

Frank	1	0.2	0.6
-------	---	-----	-----

Probabilistic Databases

Has anyone published a paper with both Erdos and Einstein



- Tuple-independent probabilistic database

Scientist	Name	Prob
	Erdos	0.9
	Einstein	0.8
	Straus	0.6

Coauthor	Actor	Director	Prob
	Erdos	Straus	0.6
	Einstein	Straus	0.7
	Obama	Erdos	0.1

- Learned from the web, large text corpora, ontologies, etc., using **statistical** machine learning.

Probabilistic Databases

- Query: SQL or First-order logic

```
SELECT Actor.name  
FROM Actor, WorkedFor  
WHERE Actor.name = WorkedFor.actor
```

$$Q(x) = \exists y \text{ Actor}(x) \wedge \text{WorkedFor}(x,y)$$

- Each UCQ query is either **#P-hard**, or **PTIME** in the size of the database.

Probabilistic Databases

- Query: SQL or First-order logic

```
SELECT Actor.name  
FROM Actor, WorkedFor  
WHERE Actor.name = WorkedFor.actor
```

$$Q(x) = \exists y \text{ Actor}(x) \wedge \text{WorkedFor}(x,y)$$

- Each UCQ query is either **#P-hard**, or **PTIME** in the size of the database.

Probabilistic query evaluation algorithm runs in linear time for all PTIME UCQ queries

Approximate Symmetries

- Exploit approximate symmetries:

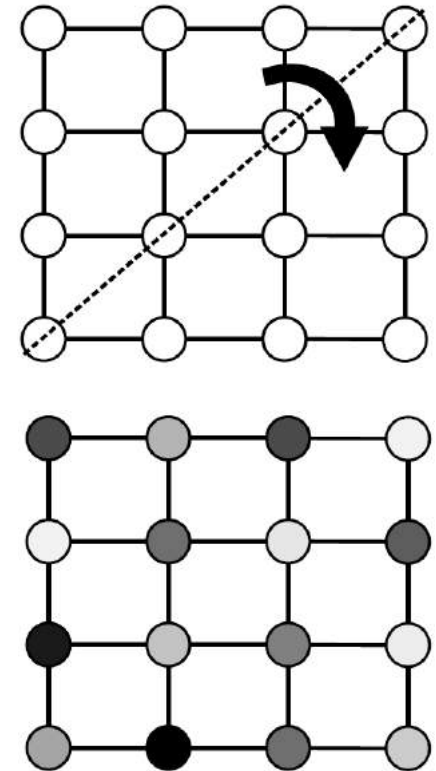
- Exact symmetry g

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

- Approximate symmetry g

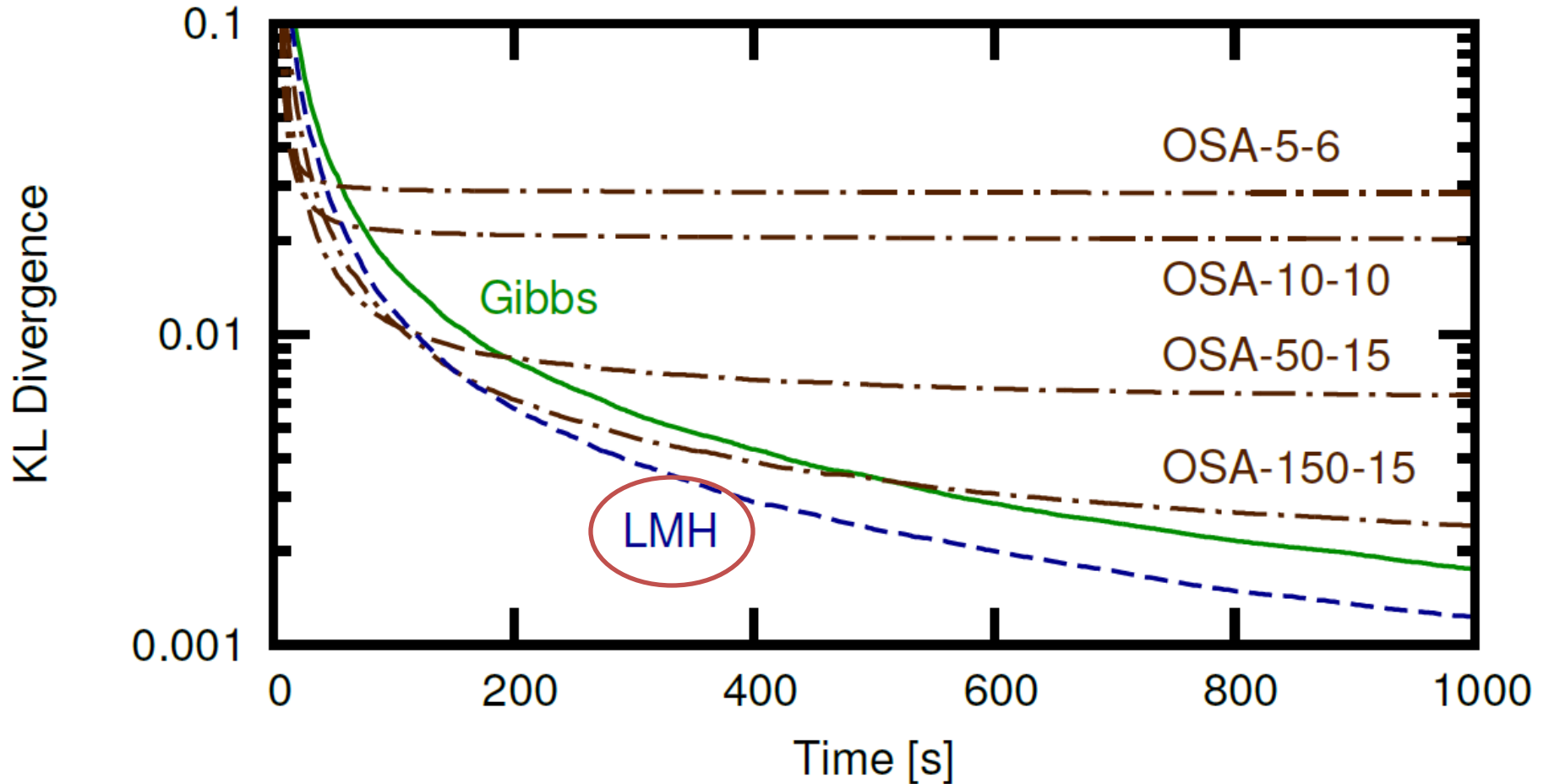
$$\Pr(\mathbf{x}) \approx \Pr(\mathbf{x}^g)$$

$$\Pr \left(\begin{array}{c} \text{Image of a woman's face} \end{array} \right) \approx \Pr \left(\begin{array}{c} \text{Image of a woman's face} \end{array} \right)$$



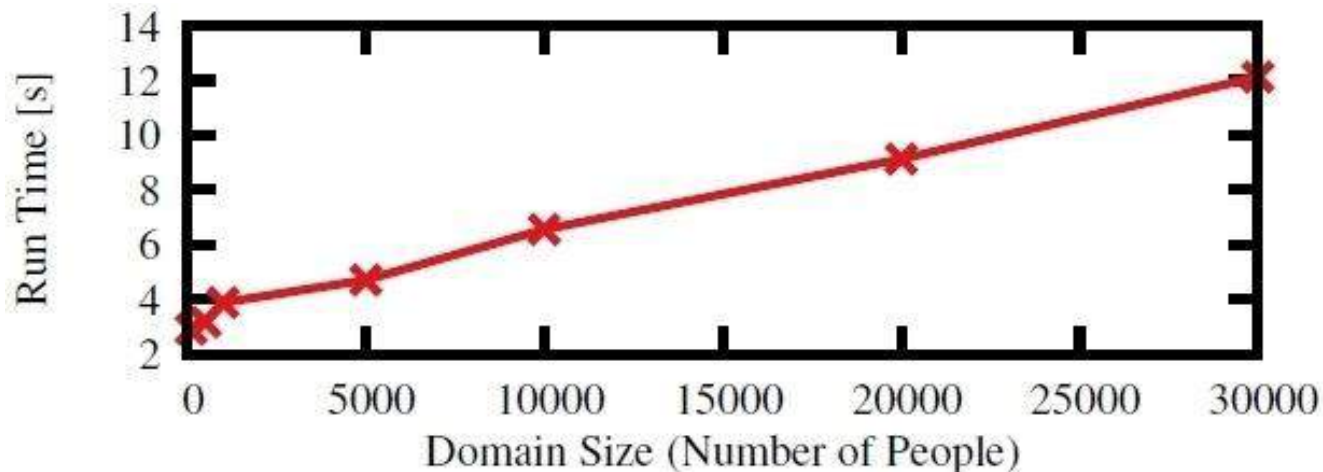
- Approximate lifted inference (MCMC)

Experiments: WebKB



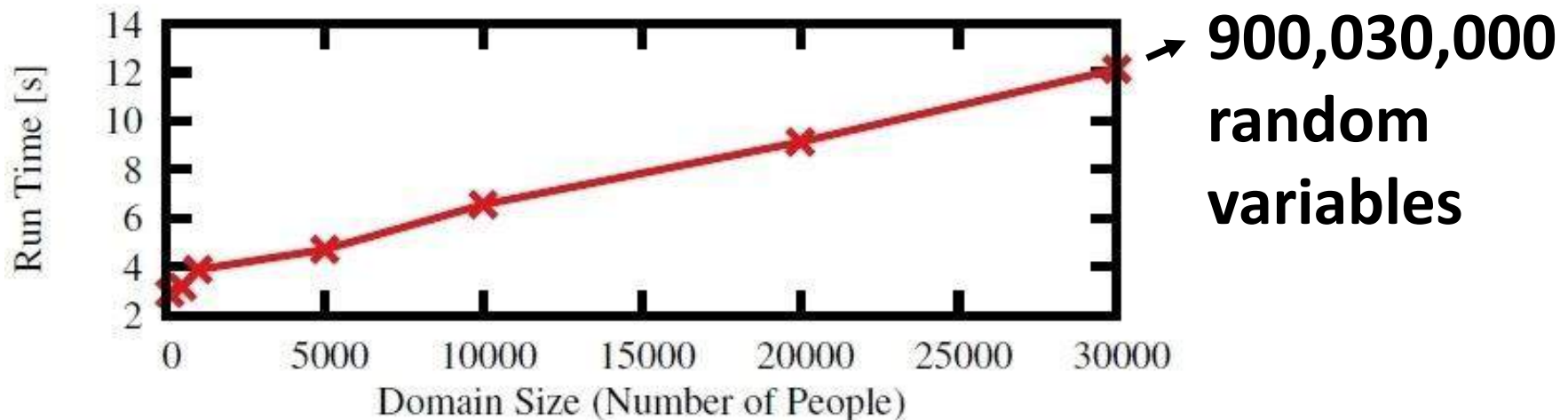
Lifted Parameter Learning

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



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



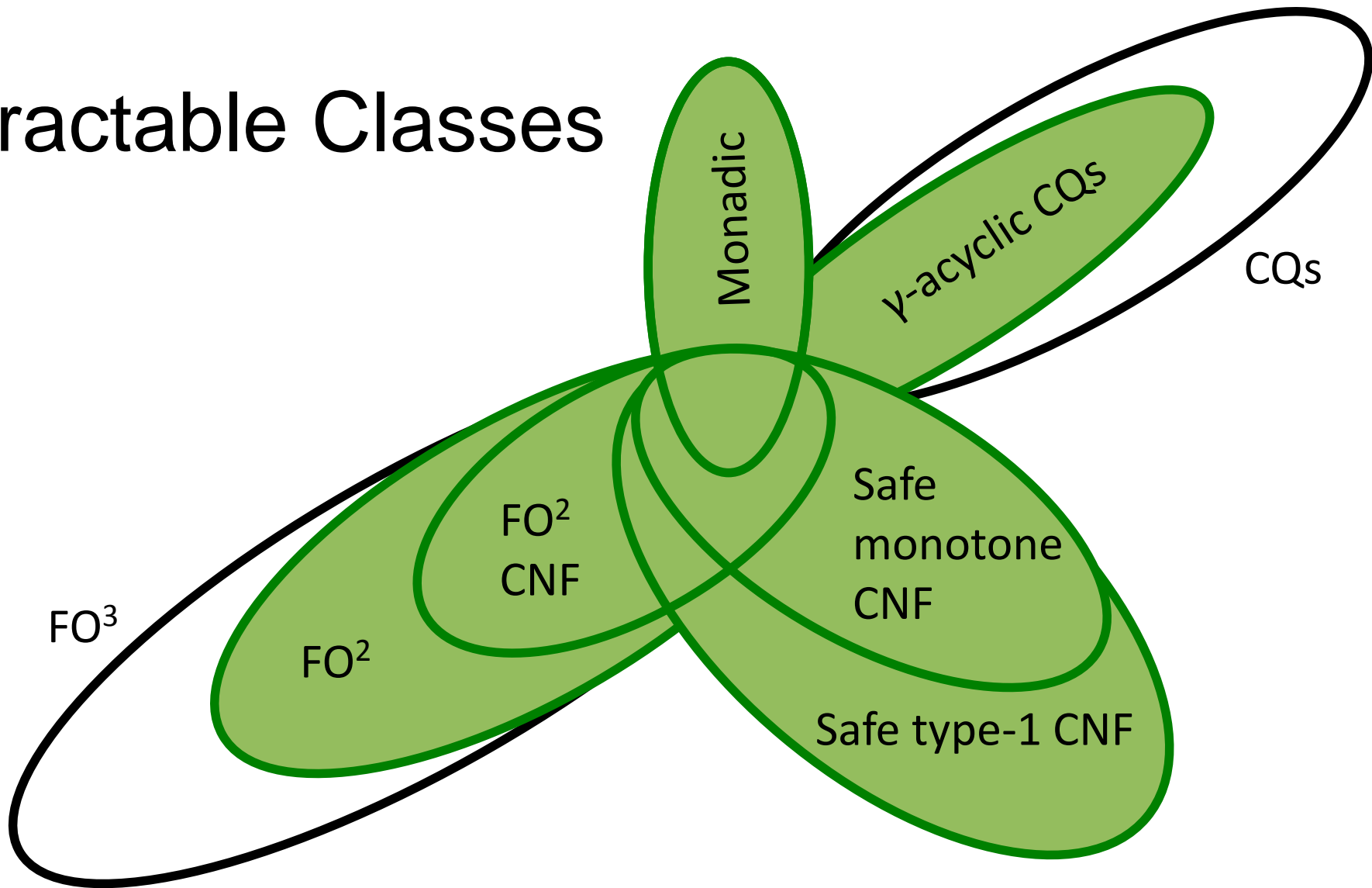
Lifted Structure Learning

- **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)

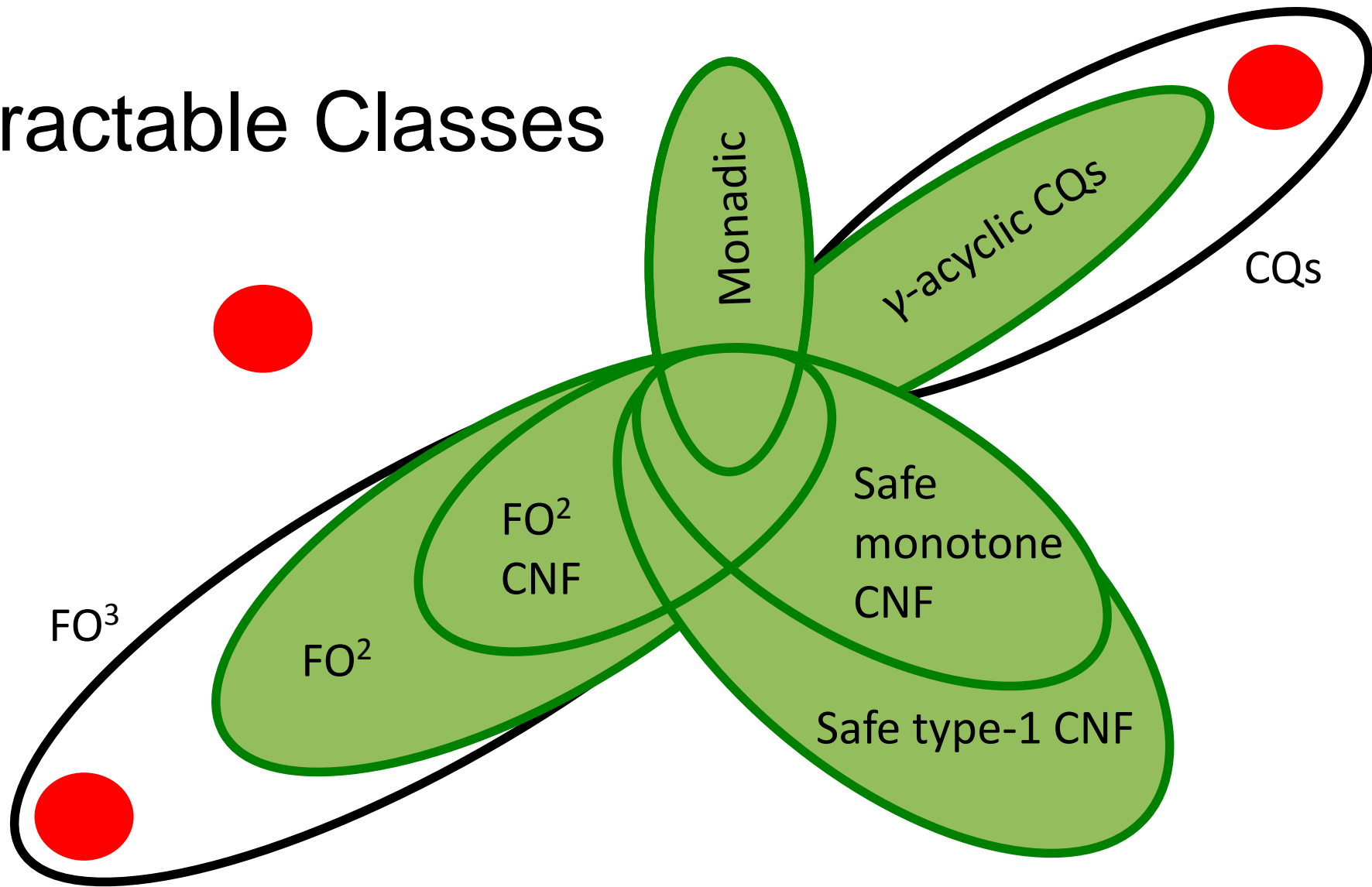
	<i>IMDb</i>			<i>UWCSE</i>		
	Baseline	Lifted Weight Learning	Lifted Structure Learning	Baseline	Lifted Weight Learning	Lifted Structure Learning
Fold 1	-548	-378	-306	-1,860	-1,524	-1,477
Fold 2	-689	-390	-309	-594	-535	-511
Fold 3	-1,157	-851	-733	-1,462	-1,245	-1,167
Fold 4	-415	-285	-224	-2,820	-2,510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

*What are the
challenges?*

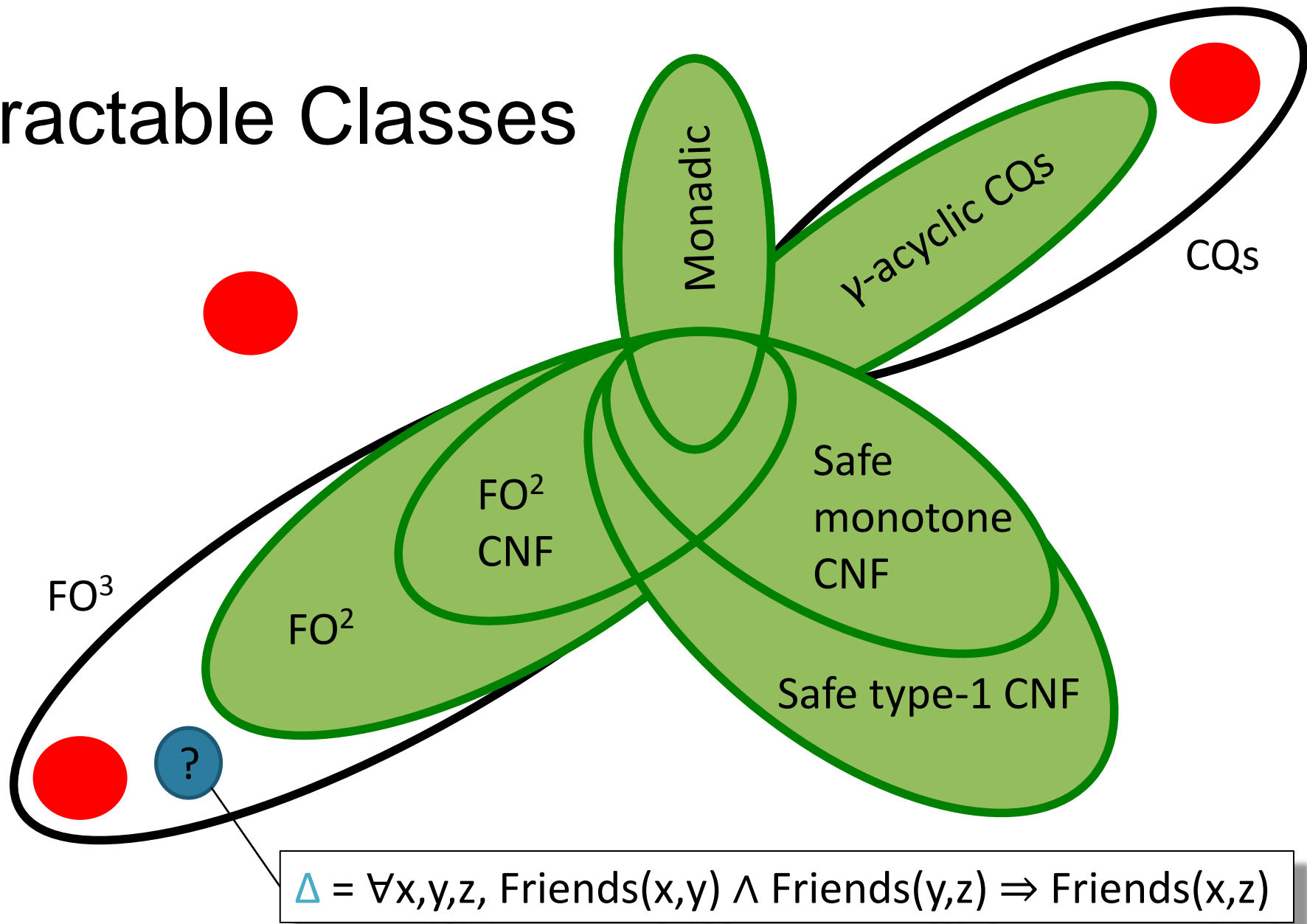
Tractable Classes



Tractable Classes



Tractable Classes



Generalized Model Counting

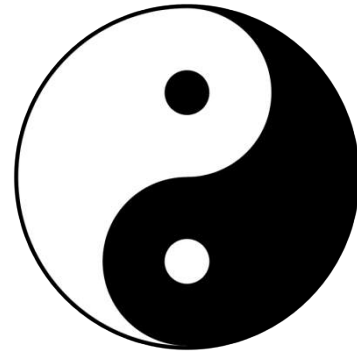
Probability Distribution

=

Logic

+

Weights



Generalized Model Counting

Probability Distribution

=

Logic

+

Weights

Logical Syntax

Model-theoretic
Semantics

+

Weight function $w(\cdot)$

Weighted Model Integration

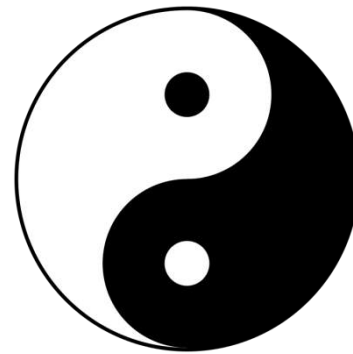
Probability Distribution

=

SMT(LRA)

+

Weights



Weighted Model Integration

Probability Distribution

=

SMT(LRA)

+

Weights

$0 \leq \text{height} \leq 200$

$0 \leq \text{weight} \leq 200$

$0 \leq \text{age} \leq 100$

$\text{age} < 1 \Rightarrow$

$\text{height} + \text{weight} \leq 90$

+

$w(\text{height}) = \text{height} - 10$

$w(\neg \text{height}) = 3 * \text{height}^2$

$w(\neg \text{weight}) = 5$

...

Probabilistic Programming

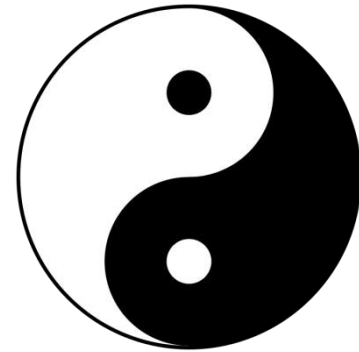
Probability Distribution

=

Logic Programs

+

Weights



Probabilistic Programming

Probability Distribution

=

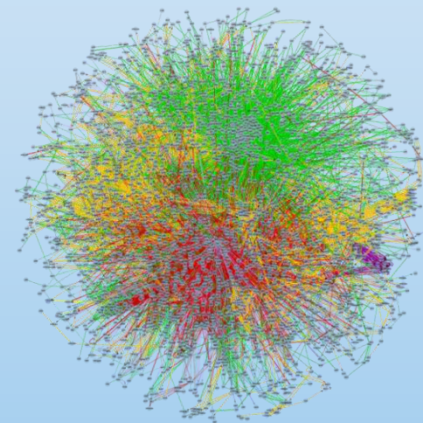
Logic Programs

+

Weights

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

+



Open World DB

- 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
...

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

Open World DB

- 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
...

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

$Q2 = \exists x \text{ Coauthor}(\text{Bieber}, x) \wedge \text{ Coauthor}(\text{Erdos}, x)$

Open World DB

- 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
...

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

$$Q2 = \exists x \text{ Coauthor}(\text{Bieber}, x) \wedge \text{Coauthor}(\text{Erdos}, x)$$

$$Q3 = \text{Coauthor}(\text{Einstein}, \text{Straus}) \wedge \text{Coauthor}(\text{Erdos}, \text{Straus})$$

Open World DB

- 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
...

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

$$Q2 = \exists x \text{ Coauthor}(\text{Bieber}, x) \wedge \text{Coauthor}(\text{Erdos}, x)$$

$$Q3 = \text{Coauthor}(\text{Einstein}, \text{Straus}) \wedge \text{Coauthor}(\text{Erdos}, \text{Straus})$$

$$Q4 = \text{Coauthor}(\text{Einstein}, \text{Bieber}) \wedge \text{Coauthor}(\text{Erdos}, \text{Bieber})$$

Open World DB

- 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
...

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

$$Q2 = \exists x \text{ Coauthor}(\text{Bieber}, x) \wedge \text{Coauthor}(\text{Erdos}, x)$$

$$Q3 = \text{Coauthor}(\text{Einstein}, \text{Straus}) \wedge \text{Coauthor}(\text{Erdos}, \text{Straus})$$

$$Q4 = \text{Coauthor}(\text{Einstein}, \text{Bieber}) \wedge \text{Coauthor}(\text{Erdos}, \text{Bieber})$$

$$Q5 = \text{Coauthor}(\text{Einstein}, \text{Bieber}) \wedge \neg \text{Coauthor}(\text{Einstein}, \text{Bieber})$$

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
...

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

$$Q3 = \text{Coauthor}(\text{Einstein}, \mathbf{\text{Straus}}) \wedge \text{Coauthor}(\text{Erdos}, \mathbf{\text{Straus}})$$

$$Q4 = \text{Coauthor}(\text{Einstein}, \mathbf{\text{Bieber}}) \wedge \text{Coauthor}(\text{Erdos}, \mathbf{\text{Bieber}})$$

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
...

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

$$Q3 = \text{Coauthor}(\text{Einstein}, \text{Straus}) \wedge \text{Coauthor}(\text{Erdos}, \text{Straus})$$

$$Q4 = \text{Coauthor}(\text{Einstein}, \text{Bieber}) \wedge \text{Coauthor}(\text{Erdos}, \text{Bieber})$$

We know for sure that $P(Q1) \geq P(Q3)$, $P(Q1) \geq P(Q4)$

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
...

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

$$Q3 = \text{Coauthor}(\text{Einstein}, \text{Straus}) \wedge \text{Coauthor}(\text{Erdos}, \text{Straus})$$

$$Q4 = \text{Coauthor}(\text{Einstein}, \text{Bieber}) \wedge \text{Coauthor}(\text{Erdos}, \text{Bieber})$$

$$Q5 = \text{Coauthor}(\text{Einstein}, \text{Bieber}) \wedge \neg \text{Coauthor}(\text{Einstein}, \text{Bieber})$$

We know for sure that $P(Q1) \geq P(Q3)$, $P(Q1) \geq P(Q4)$

and $P(Q3) \geq P(Q5)$, $P(Q4) \geq P(Q5)$

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
...

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

$$Q3 = \text{Coauthor}(\text{Einstein}, \text{Straus}) \wedge \text{Coauthor}(\text{Erdos}, \text{Straus})$$

$$Q4 = \text{Coauthor}(\text{Einstein}, \text{Bieber}) \wedge \text{Coauthor}(\text{Erdos}, \text{Bieber})$$

$$Q5 = \text{Coauthor}(\text{Einstein}, \text{Bieber}) \wedge \neg \text{Coauthor}(\text{Einstein}, \text{Bieber})$$

We know for sure that $P(Q1) \geq P(Q3)$, $P(Q1) \geq P(Q4)$

and $P(Q3) \geq P(Q5)$, $P(Q4) \geq P(Q5)$ because $P(Q5) = 0$.

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
...

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

$$Q2 = \exists x \text{ Coauthor}(\text{Bieber}, x) \wedge \text{Coauthor}(\text{Erdos}, x)$$

$$Q3 = \text{Coauthor}(\text{Einstein}, \mathbf{\text{Straus}}) \wedge \text{Coauthor}(\text{Erdos}, \mathbf{\text{Straus}})$$

$$Q4 = \text{Coauthor}(\text{Einstein}, \mathbf{\text{Bieber}}) \wedge \text{Coauthor}(\text{Erdos}, \mathbf{\text{Bieber}})$$

$$Q5 = \text{Coauthor}(\text{Einstein}, \mathbf{\text{Bieber}}) \wedge \neg \text{Coauthor}(\mathbf{\text{Einstein}}, \mathbf{\text{Bieber}})$$

We know for sure that $P(Q1) \geq P(Q3)$, $P(Q1) \geq P(Q4)$

and $P(Q3) \geq P(Q5)$, $P(Q4) \geq P(Q5)$ because $P(Q5) = 0$.

We have strong evidence that $P(Q1) \geq P(Q2)$.

Conclusions

- Integration of logic and probability is long-standing goal of AI
- First-order probabilistic reasoning is **frontier** and **integration** of AI, KR, ML, DBs, theory, PL, etc.
- We need
 - relational models and logic
 - probabilistic models and statistical learning
 - algorithms that scale

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**

QUESTIONS?



**THE
FIRST ORDER
NEEDS YOU**

References

- Van den Broeck, Guy. "Towards high-level probabilistic reasoning with lifted inference." AAAI Spring Symposium on KRR (2015).
- Chavira, Mark, and Adnan Darwiche. "On probabilistic inference by weighted model counting." *Artificial Intelligence* 172.6 (2008): 772-799.
- Sang, Tian, Paul Beame, and Henry A. Kautz. "Performing Bayesian inference by weighted model counting." *AAAI*. Vol. 5. 2005.
- Chavira, Mark, Adnan Darwiche, and Manfred Jaeger. "Compiling relational Bayesian networks for exact inference." *International Journal of Approximate Reasoning* 42.1 (2006): 4-20.
- 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.

References

- Van den Broeck, Guy, Nima Taghipour, Wannes Meert, Jesse Davis, and Luc De Raedt. "Lifted probabilistic inference by first-order knowledge compilation." AAI, 2011.
- Van den Broeck, Guy. Lifted inference and learning in statistical relational models. Diss. Ph. D. Dissertation, KU Leuven, 2013.
- Gogate, Vibhav, and Pedro Domingos. "Probabilistic theorem proving." UAI (2011).
- Gribkoff, Eric, Guy Van den Broeck, and Dan Suci. "Understanding the complexity of lifted inference and asymmetric weighted model counting." UAI (2014).
- Niepert, Mathias, and Guy Van den Broeck. "Tractability through exchangeability: A new perspective on efficient probabilistic inference." AAI (2014).

References

- Van den Broeck, Guy. "On the completeness of first-order knowledge compilation for lifted probabilistic inference." *Advances in Neural Information Processing Systems*. 2011.
- Van den Broeck, Guy, Wannes Meert, and Adnan Darwiche. "Skolemization for weighted first-order model counting." *Proceedings of the 14th International Conference on Principles of Knowledge Representation and Reasoning (KR)*. 2014.
- Ceylan, Ismail Ilkan, Adnan Darwiche, and Guy Van den Broeck. "Open-world probabilistic databases." *Proceedings of KR (2016)*.
- Van den Broeck, Guy, and Adnan Darwiche. "On the complexity and approximation of binary evidence in lifted inference." *Advances in Neural Information Processing Systems*. 2013.
- Van den Broeck, Guy, and Mathias Niepert. "Lifted probabilistic inference for asymmetric graphical models." *Proceedings of AAI (2015)*.

References

- 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.
- 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.
- 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.