

# Lifted Probabilistic Inference in Relational Models

Guy Van den Broeck  
UCLA

Dan Suciu  
U. of Washington

IJCAI Tutorial  
July 10, 2016

# About the Tutorial

Slides available at

<http://web.cs.ucla.edu/~guyvdb/talks/IJCAI16-tutorial/>

Extensive bibliography at the end.

Your speakers:



<http://web.cs.ucla.edu/~guyvdb/>

I work in AI



<https://homes.cs.washington.edu/~suciu/>

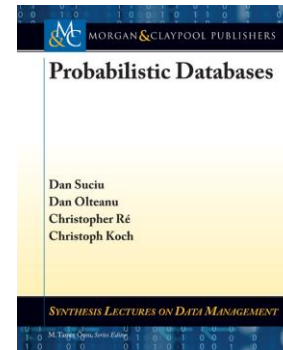
I work in DB

# About the Tutorial

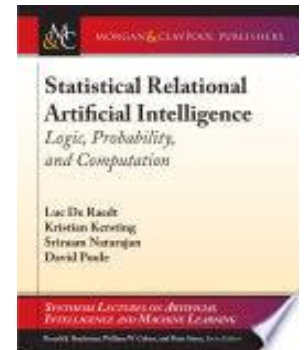
- The tutorial is about
  - deep connections between AI and DBs
  - a unified view on probabilistic reasoning
  - a logical approach to prob. reasoning
- The tutorial is NOT an exhaustive overview of lifted algorithms for graphical models (see references at the end)

# If you want more...

- Books
  - Probabilistic Databases
  - Statistical Relational AI
  - (Lifted Inference Book)

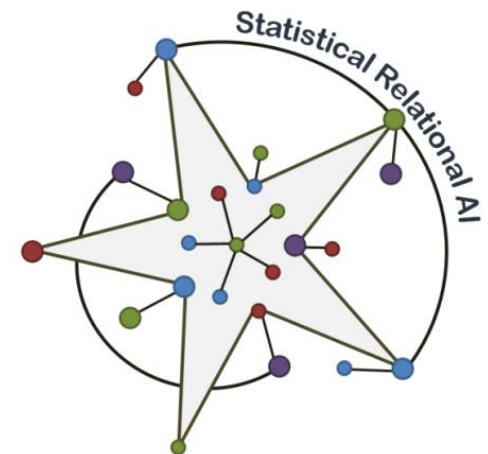


[Suciu'11]



[DeRaedt'16]

- StarAI workshop on Monday  
<http://www.starai.org>
- Main conference papers



# Outline

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



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

# Outline

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



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

# Part 1: Motivation

- Why do we need relational representations of uncertainty?
- Why do we need probabilistic queries?
- Why do we need lifted inference algorithms?

# Why Relational Data?

- Our data is already relational!
  - Companies run relational databases
  - Scientific data is relational:
    - Large Hadron Collider generated 25PB in 2012
    - LSST Telescope will produce 30TB per night
- Big data is big business:
  - Oracle: \$7.1BN in sales
  - IBM: \$3.2BN in sales
  - Microsoft: \$2.6BN in sales





# Why Probabilistic Relational Data?

- Relational data is increasingly probabilistic
  - NELL machine reading (>50M tuples)
  - Google Knowledge Vault (>2BN tuples)
  - DeepDive (>7M tuples)
- Data is inferred from unstructured information using statistical models
  - Learned from the web, large text corpora, ontologies, etc.
  - The learned/extracted data is relational

# Information Extraction

## PhD Students Luc De Raedt

- ✦ Laura-Andrea Antanas (co-promotor Tinne Tuytelaars)
- ✦ Dries Van Daele (co-promotor Kathleen Marchal)
- ✦ Thanh Le Van (co-promotor Kathleen Marchal)
- ✦ Bogdan Moldovan
- ✦ Davide Nitti (co-promotor Tinne De Laet)
- ✦ José Antonio Oramas Mogroycio (key supervisor Tinne Tuytelaars)
- ✦ Francesco Orsini (co-supervisor Paol Frasconi)
- ✦ Sergey Paramonov
- ✦ Joris Renkens
- ✦ Mathias Verbeke (with Bettina Berendt)
- ✦ Jonas Vlasselaer



## PublishedWith

X	Y	P
Luc	Laura	0.7
Luc	Hendrik	0.6
Luc	Kathleen	0.3
Luc	Paol	0.3
Luc	Paolo	0.1

## Alumni Luc De Raedt

1. Hendrik Blockeel, *Top-down induction of first order logical decision trees*, Ph.D. thesis, Department of Computer Science, K.U.Leuven, Leuven, Belgium, december 1998, 202+xv pages. (Co-promotor Maurice Bruynooghe)
2. Luc Dehaspe, *Frequent pattern discovery in first-order logic*, Ph.D. thesis, Department of Computer

# Extraction is so Noisy!

ebay Shop by category Search... All Categories Search Advanced

Back to home page | Listed in category: Books, Magazines > Non-Fiction Books > See more Probabilistic Inductive Logic Programming by S...

This is a private listing. Sign in to view your status or learn more about private listings.

**Probabilistic Inductive Logic Programming De Raedt, Luc (Editor)/ Frasconi, Paol**

Item condition: **Brand new**

Time left: 18d 13h (22 Feb, 2016 04:40:52 AEDST)

Quantity:  6 available

Price: **AU \$136.69** **Buy It Now** **Add to cart**

**Seller information**  
roxy\*books (236078 ★)  
99,1% Positive feedback  
Follow this seller  
Visit store: books\_music\_supers...

State-of-the-Art Survey  
UNAI 4911  
Luc De Raedt  
Paolo Frasconi  
Kristian Kersting  
Stephen Muggleton (Eds.)  
**Probabilistic Inductive Logic Programming**  
Theory and Applications

# Representation: Probabilistic Databases

- Tuple-independent probabilistic databases

Actor	Name	Prob
	Brando	0.9
	Cruise	0.8
	Coppola	0.1

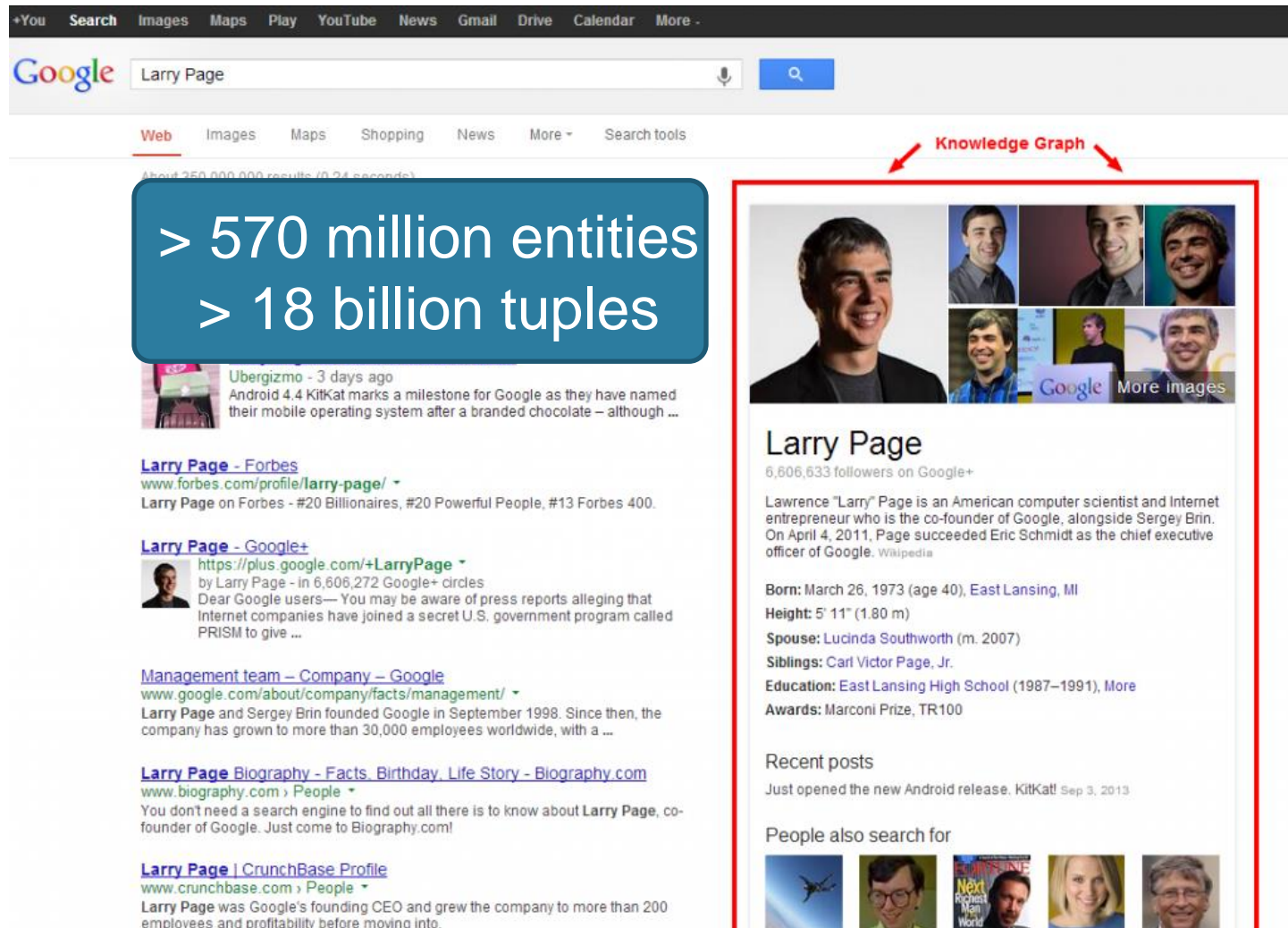
WorkedFor	Actor	Director	Prob
	Brando	Coppola	0.9
	Coppola	Brando	0.2
	Cruise	Coppola	0.1

- Query: SQL or First-order logic

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

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

# Why Probabilistic Queries?



The image shows a Google search interface with the query "Larry Page". The search results on the left include links to Larry Page's profile on Forbes, Google+, and Biography.com, along with a snippet about the Google management team. On the right, the Knowledge Graph panel for Larry Page is displayed, featuring a large portrait, several smaller images, and a list of biographical facts such as his birth date, height, spouse, siblings, education, and awards. A red box highlights the Knowledge Graph panel, and a red arrow points to it from the text "Knowledge Graph".

Google

Larry Page

Web Images Maps Shopping News More Search tools

About 250,000,000 results (0.24 seconds)

> 570 million entities  
> 18 billion tuples

Ubergizmo - 3 days ago  
Android 4.4 KitKat marks a milestone for Google as they have named their mobile operating system after a branded chocolate – although ...

[Larry Page - Forbes](#)  
[www.forbes.com/profile/larry-page/](http://www.forbes.com/profile/larry-page/) ▾  
Larry Page on Forbes - #20 Billionaires, #20 Powerful People, #13 Forbes 400.









[Larry Page - Google+](#)  
<https://plus.google.com/+LarryPage> ▾  
by Larry Page - in 6,606,272 Google+ circles  
Dear Google users— You may be aware of press reports alleging that Internet companies have joined a secret U.S. government program called PRISM to give ...

[Management team - Company - Google](#)  
[www.google.com/about/company/facts/management/](http://www.google.com/about/company/facts/management/) ▾  
Larry Page and Sergey Brin founded Google in September 1998. Since then, the company has grown to more than 30,000 employees worldwide, with a ...

[Larry Page Biography - Facts, Birthday, Life Story - Biography.com](#)  
[www.biography.com/people/larry-page](http://www.biography.com/people/larry-page) ▾  
You don't need a search engine to find out all there is to know about Larry Page, co-founder of Google. Just come to Biography.com!

[Larry Page | CrunchBase Profile](#)  
[www.crunchbase.com/people/larry-page](http://www.crunchbase.com/people/larry-page) ▾  
Larry Page was Google's founding CEO and grew the company to more than 200 employees and profitability before moving into.

**Knowledge Graph**







 More images

**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](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](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](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](http://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](http://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](http://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/ ▾](https://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 ▾](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...](https://guff.com/didnt-know-einst...)

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

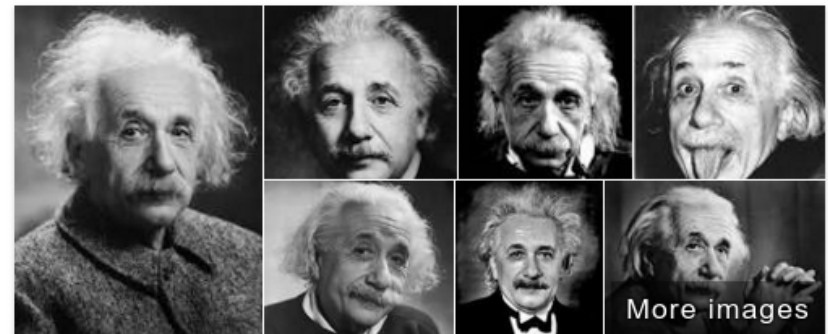
An interesting read on Einstein's superstar status. What are your thoughts? [twitter.com/aeonmag/status...](https://twitter.com/aeonmag/status...)



## Albert Einstein - Biographical - Nobelprize.org

[www.nobelprize.org/nobel\\_prizes/physics/.../einstein-bio.htm... ▾](https://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




[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](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](http://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



Kristian Kersting, ...



Justin Bieber, ...

# Observations

Has anyone published a paper with both Erdos and Einstein

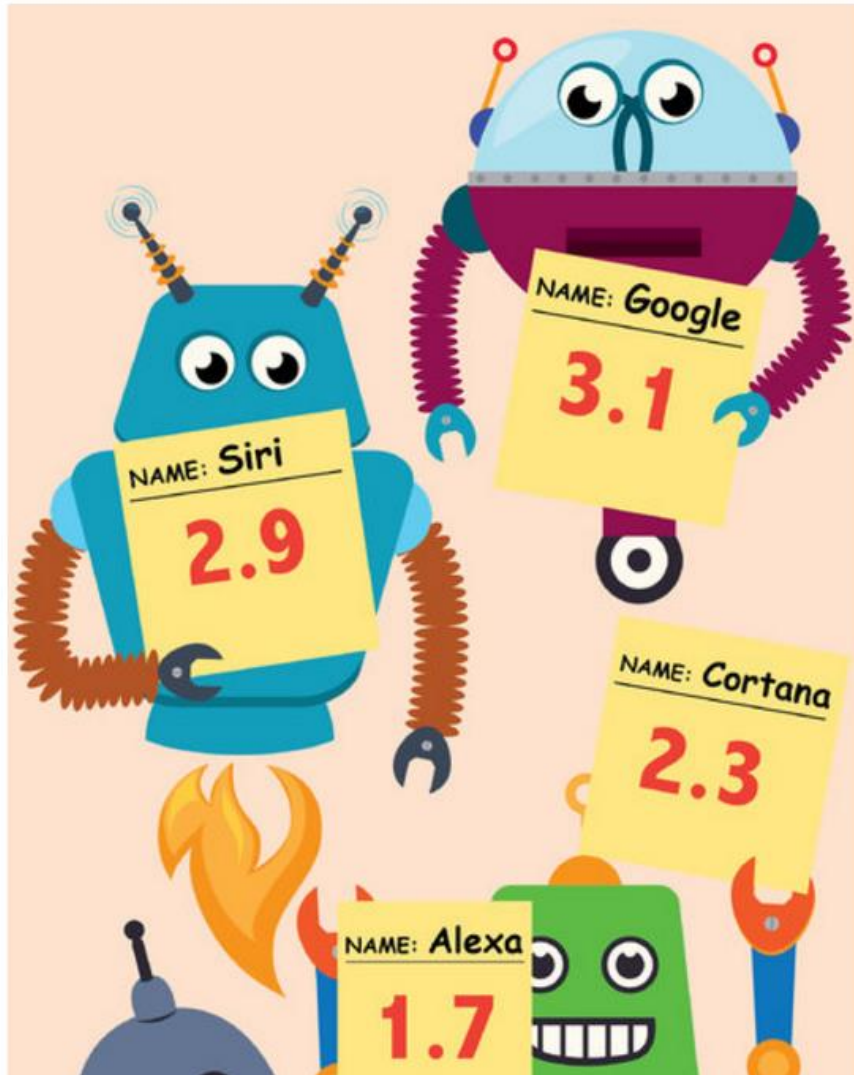


- Cannot come from labeled data
- Fuse uncertain information from many pages
- Expose uncertainty in query answers  
... and risk incorrect answers
- Embrace probability!

# *Siri, Alexa and Other Virtual Assistants Put to the Test*

## Tech Fix

By BRIAN X. CHEN JAN. 27, 2016



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.”

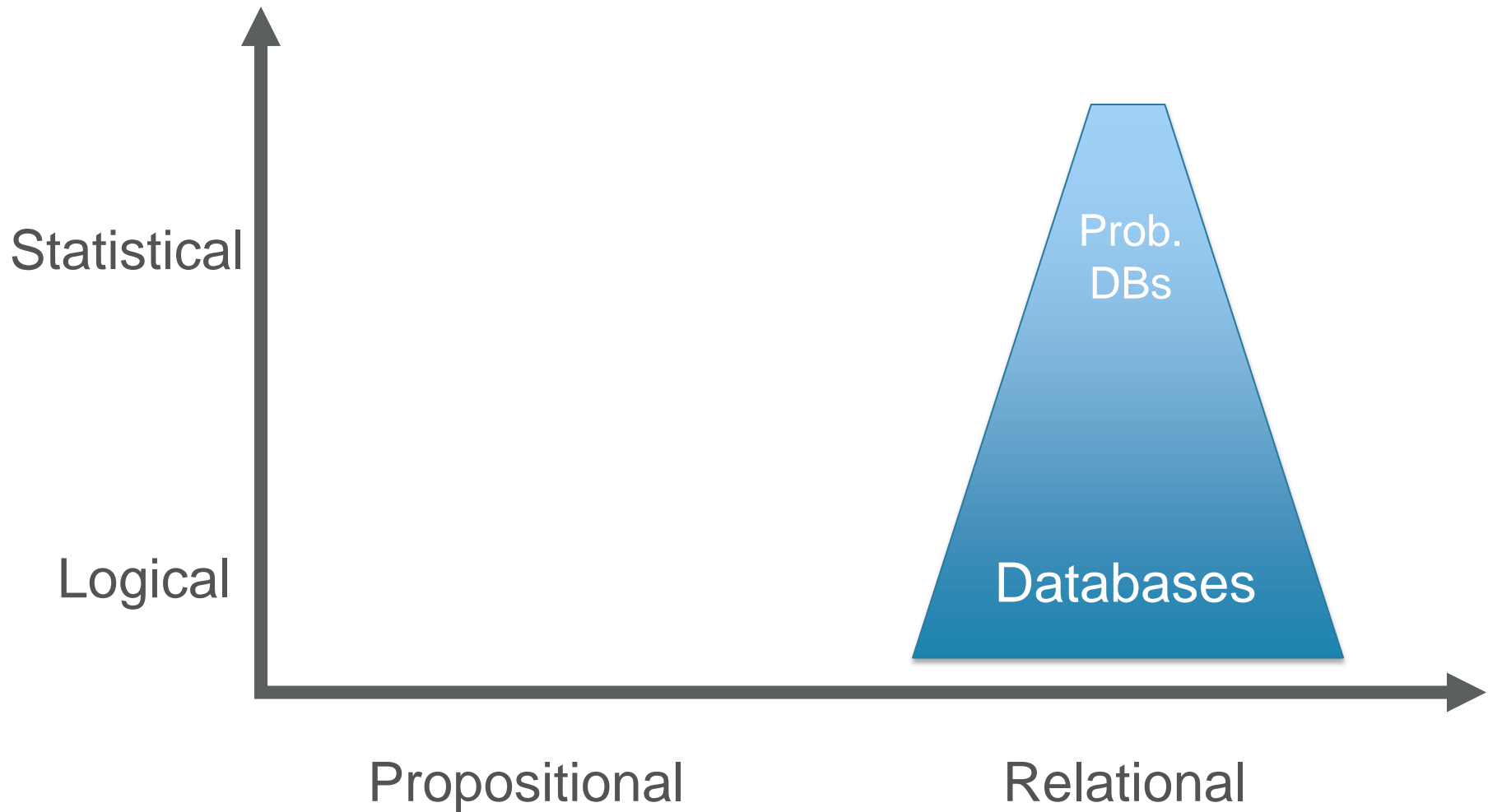
“Come on, that’s last year’s Super Bowl,” I said. “Even I can do better than that.”

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.

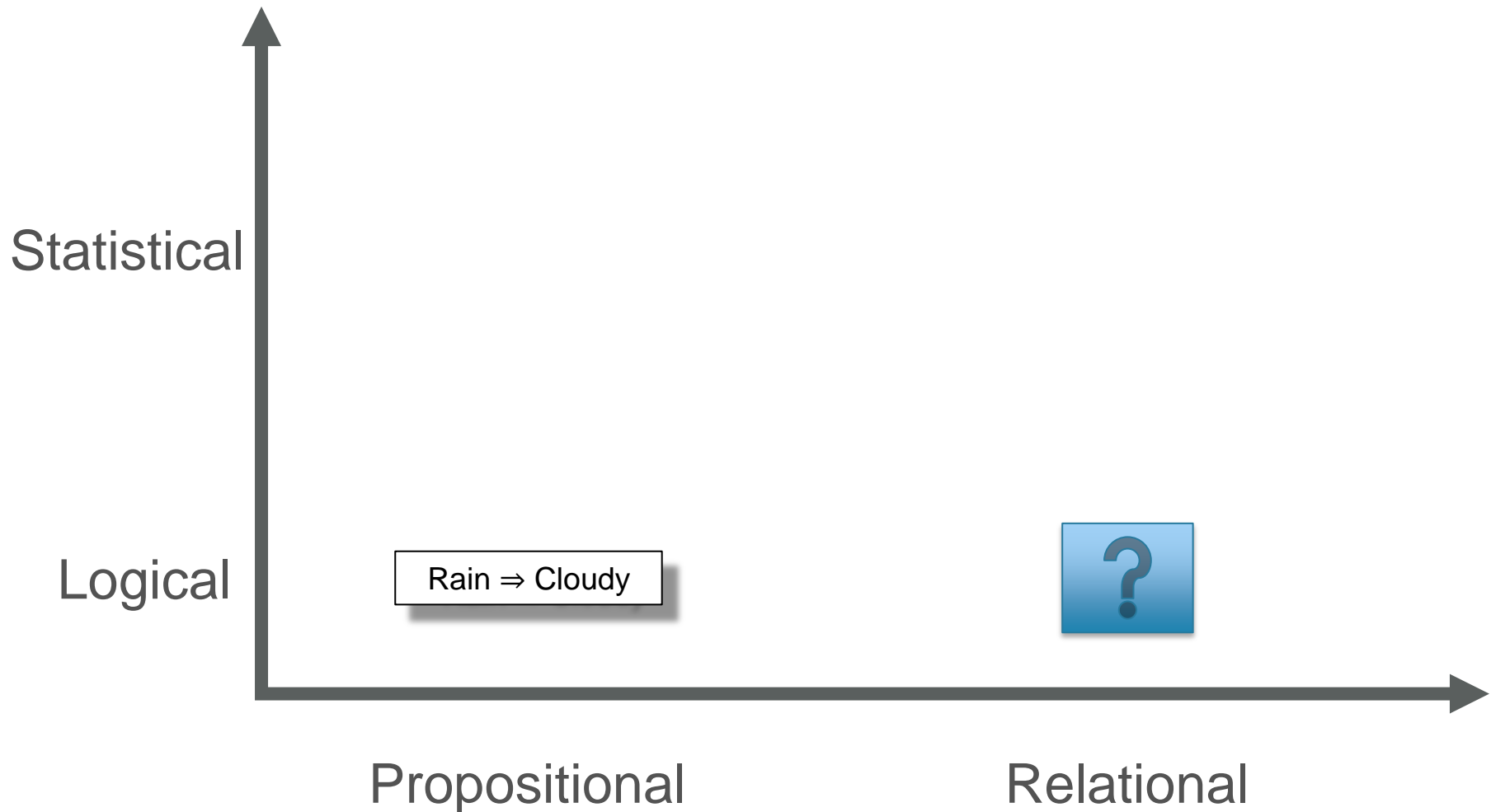
All the so-called [Frightful Five](#) tech

[Chen’16]  
(NYTimes)

# Summary



# Representations in AI and ML



# Graphical Model Learning

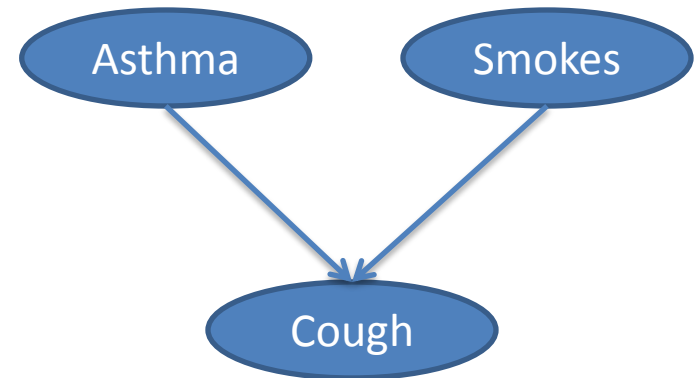


Medical Records



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

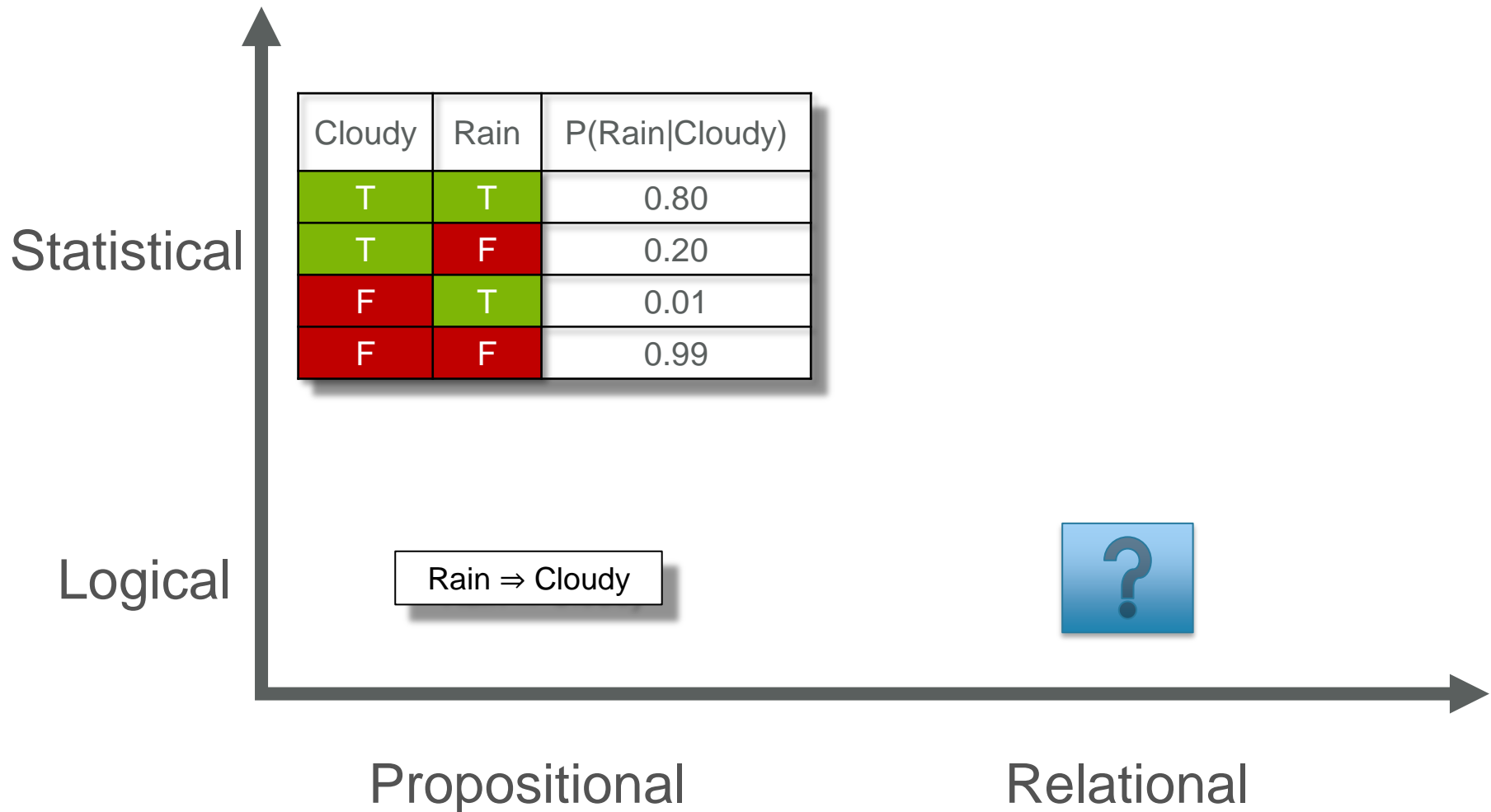


Frank	1	?	?
-------	---	---	---

Big data

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

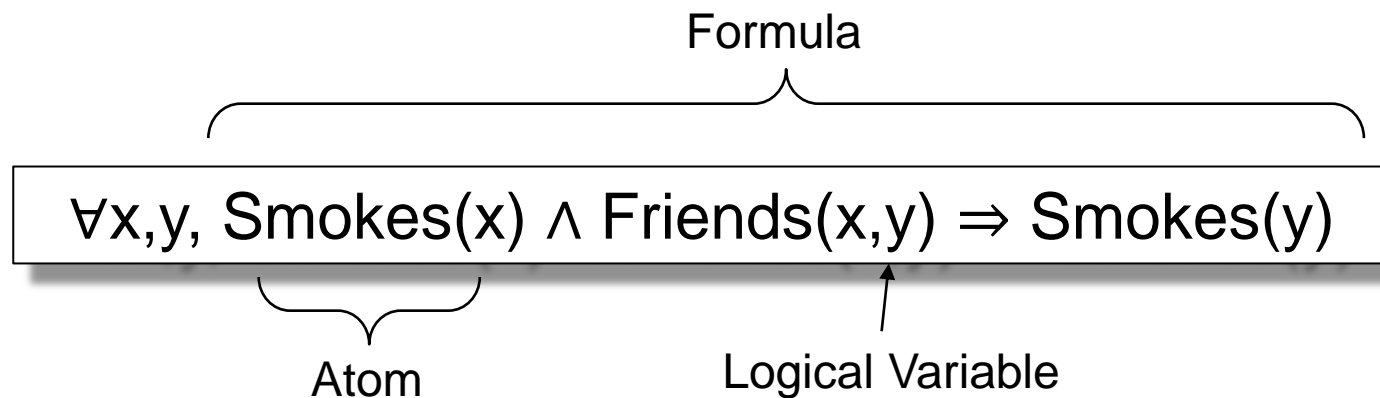
# Representations in AI and ML





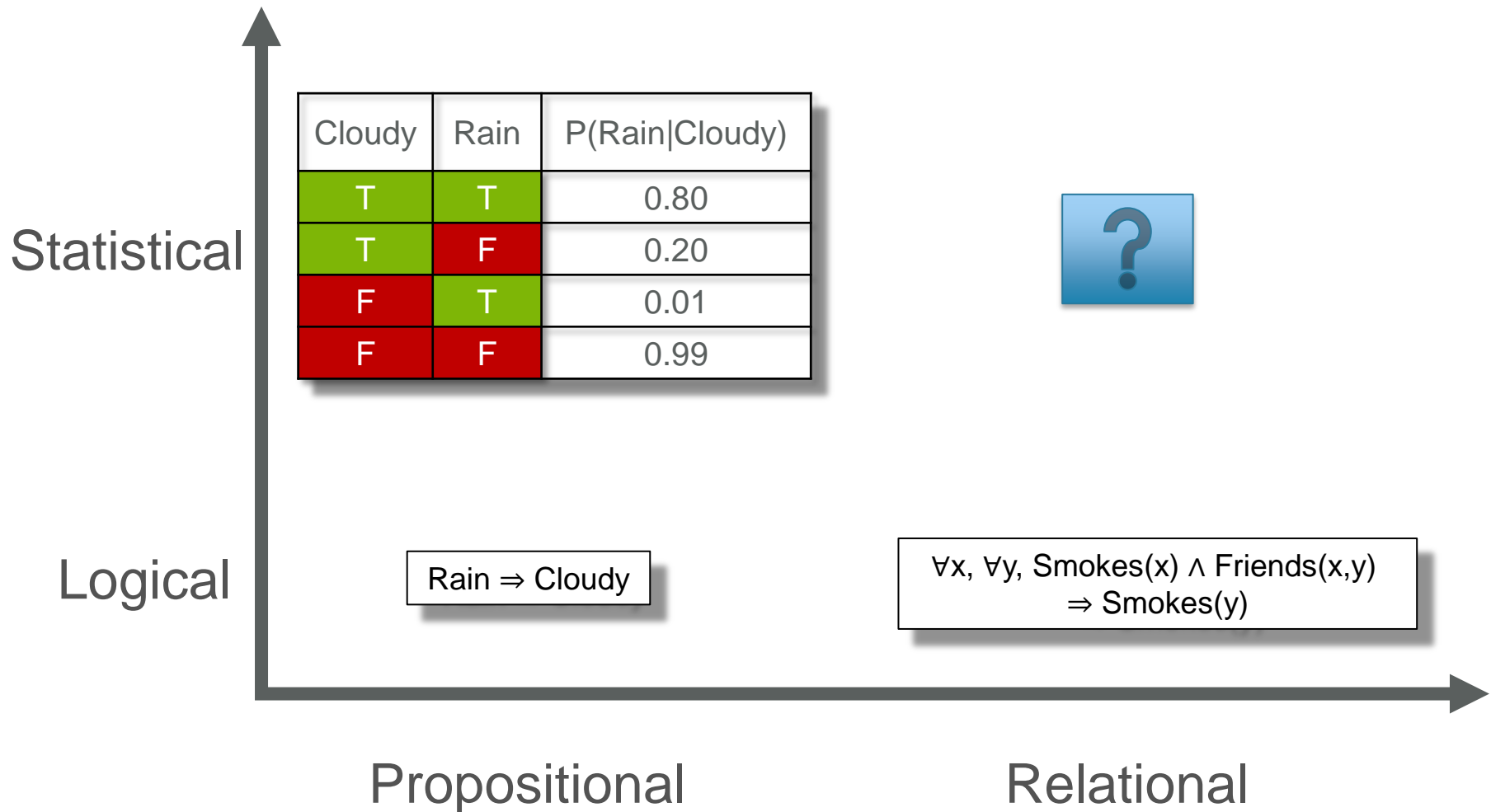
# Relational Representations

- Example: First-Order Logic



- Logical variables have domain of constants  
x,y range over domain People = {Alice,Bob}
- Ground formula has no logical variables  
 $\text{Smokes}(\text{Alice}) \wedge \text{Friends}(\text{Alice}, \text{Bob}) \Rightarrow \text{Smokes}(\text{Bob})$

# Representations in AI and ML



# Why Statistical Relational Models?

- Probabilistic graphical models
  - ✓ Quantify uncertainty and noise
  - ✗ Not very expressive
    - Rules of chess in ~100,000 pages*
- First-order logic
  - ✓ Very expressive
    - Rules of chess in 1 page*
  - ✓ Good match for abundant relational data
  - ✗ Hard to express uncertainty and noise

# Graphical Model Learning



Medical Records



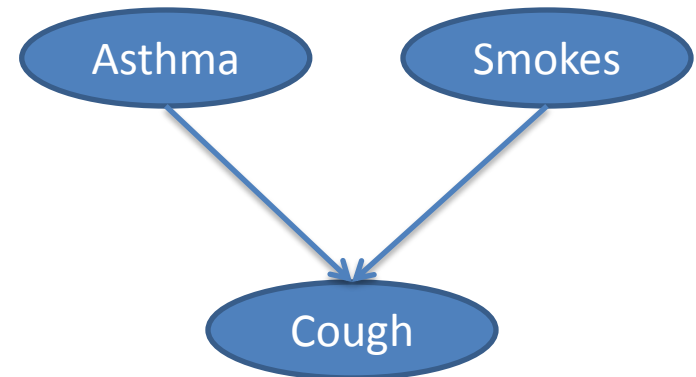
Bayesian Network

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
-------	---	-----	-----

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

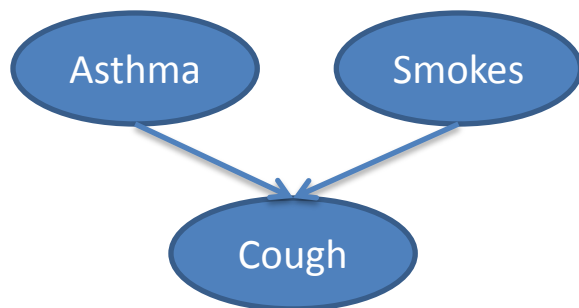


Rows are **independent**  
during learning and  
inference!

# Statistical Relational Representations

Augment graphical model with relations between entities (rows).

## Intuition



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

## Markov Logic

2.1  $\text{Asthma} \Rightarrow \text{Cough}$

3.5  $\text{Smokes} \Rightarrow \text{Cough}$

↑  
Logical variables refer to entities

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

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

# Classical Machine Learning



Purchases



Model

Name	Age	Product	Price
Dave	40	Android	€249
Alice	35	iPhone	€799
Bob	32	iPhone	€799
Charlie	22	iPhone	€699
Eve	17	Android	€299
Frank	15	Android	€199



People **older** than **27**  
probably buy **iPhone**.

People **younger** than **27**  
probably buy **Android**.

**Inference:** *Does Guy buy an iPhone?*

**Answer:** Yes, with probability 66%

# Statistical Relational Learning



Purchases

Name	Age	Product	Price
Dave	40	Android	€249
Alice	35	iPhone	€799
Bob	32	iPhone	€799
Charlie	22	iPhone	€699
Eve	17	Android	€299
Frank	15	Android	€199



Relationships

Person A	Person B	Type
Alice	Bob	Spouse
Alice	Charlie	Mother
Bob	Charlie	Father
Dave	Eve	Father
Dave	Frank	Father
Eve	Frank	Siblings

Family 1

Family 2



Model

**Family members** probably buy the **same type** of phone.

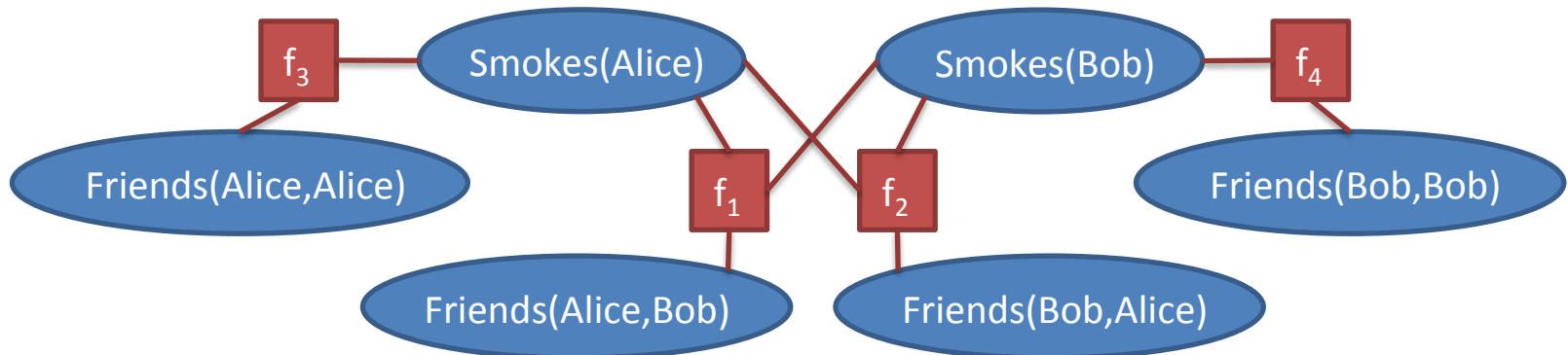
# Example: Markov Logic

- Weighted First-Order Logic

Weight or Probability FOL Formula

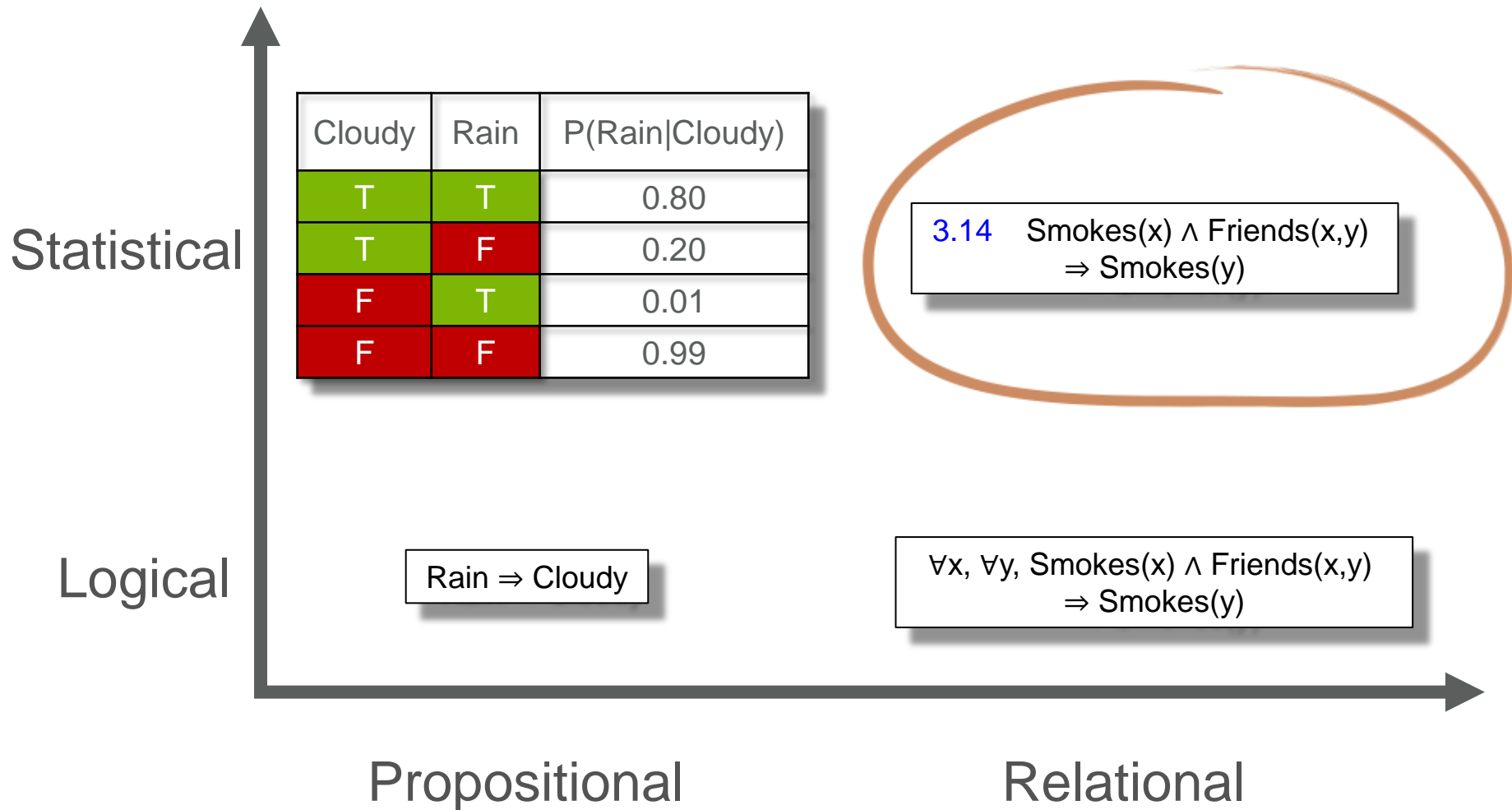
**3.14**  $\text{Smokes}(x) \wedge \text{Friends}(x,y) \Rightarrow \text{Smokes}(y)$

- Ground atom/tuple = **random variable** in {true,false}  
e.g.,  $\text{Smokes}(\text{Alice})$ ,  $\text{Friends}(\text{Alice},\text{Bob})$ , etc.
- Ground formula = **factor** in propositional factor graph

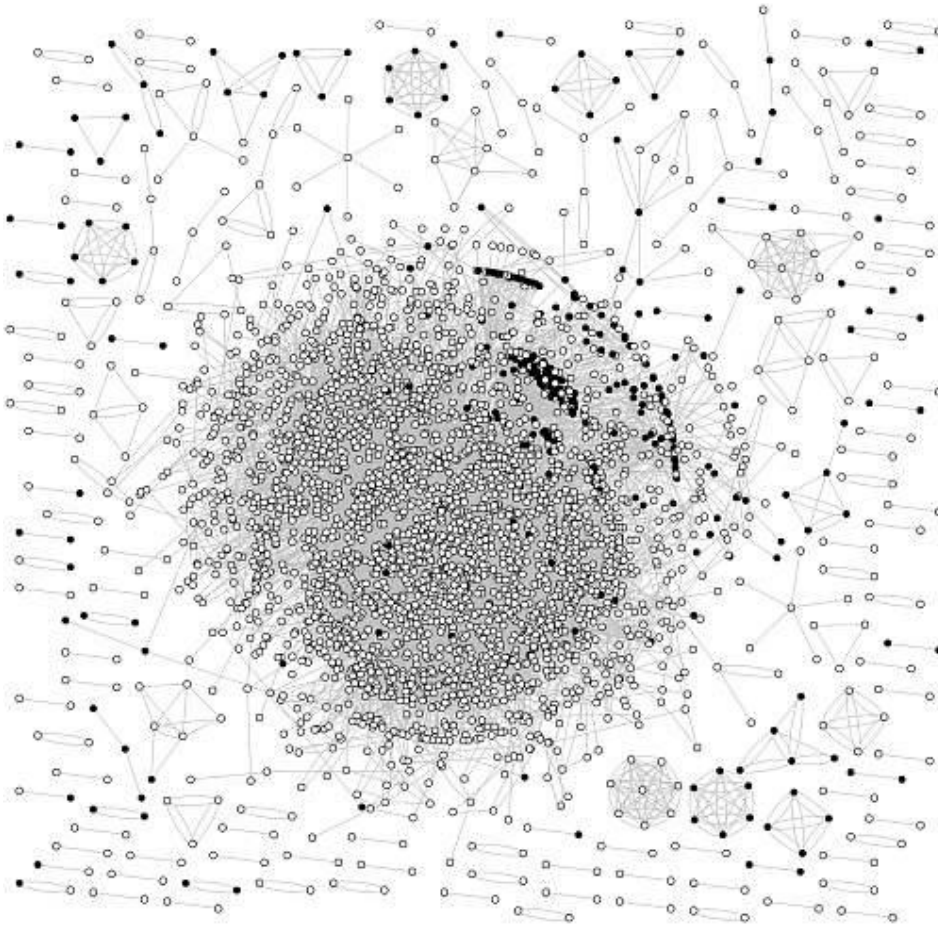




# Representations in AI and ML



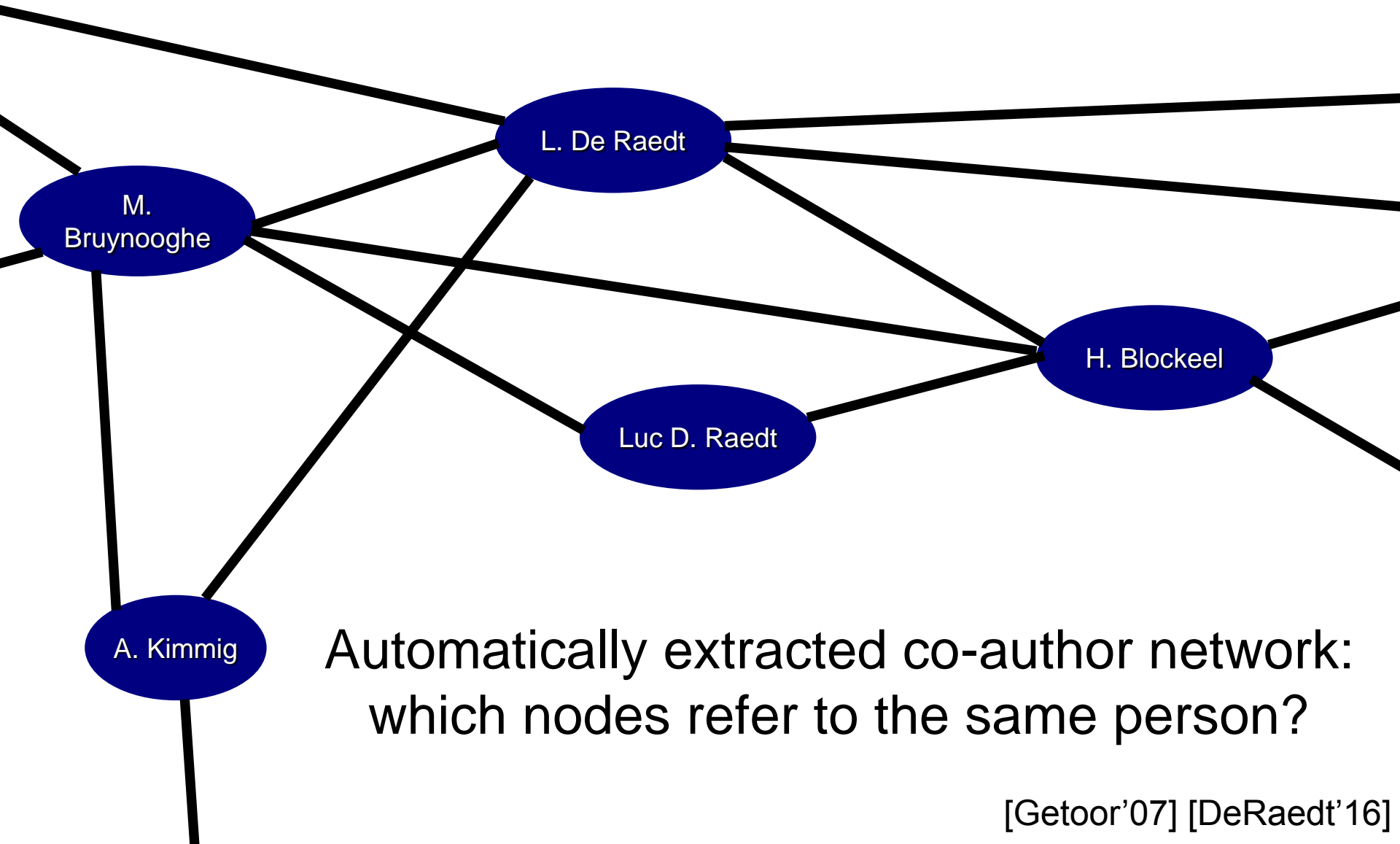
# Collective Classification



Can we predict the type of the nodes given information on its links and attributes?

E.g., the type of a webpage given its links and the words on the page?

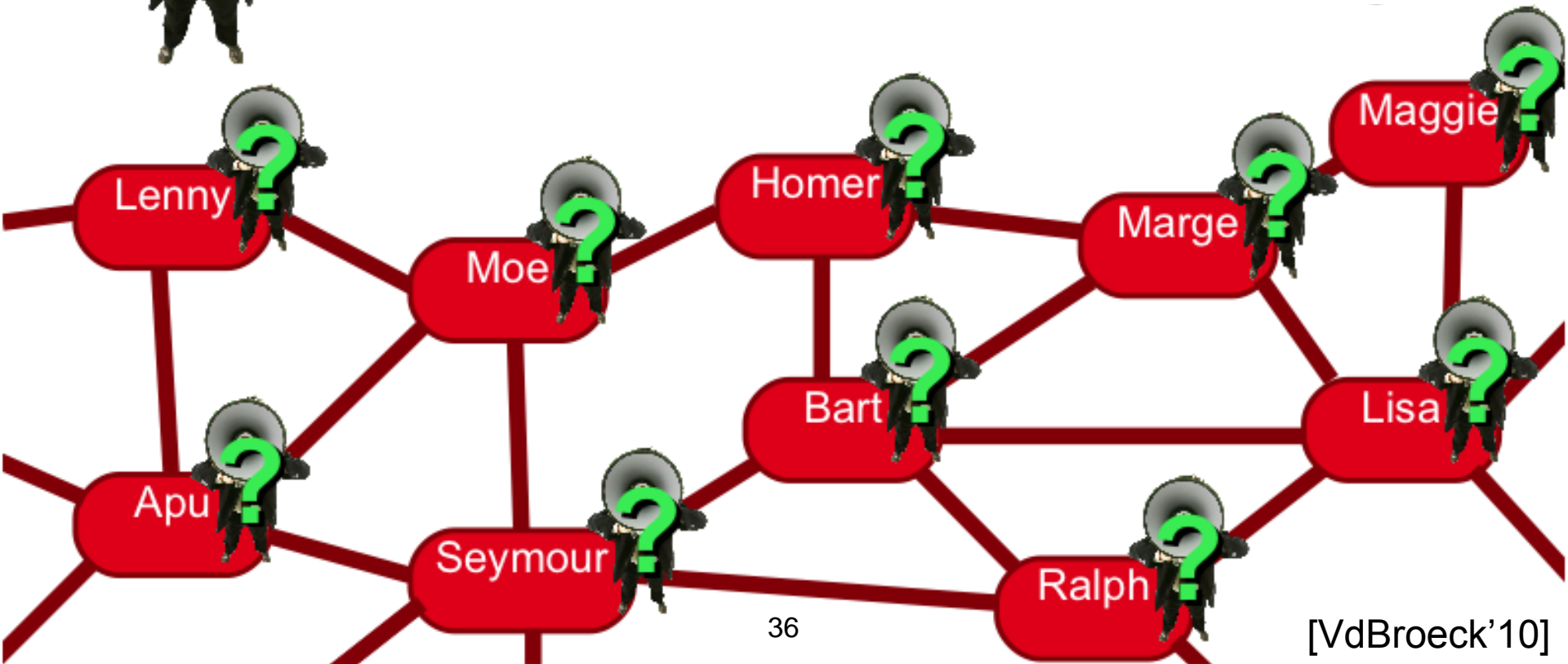
# Entity Resolution



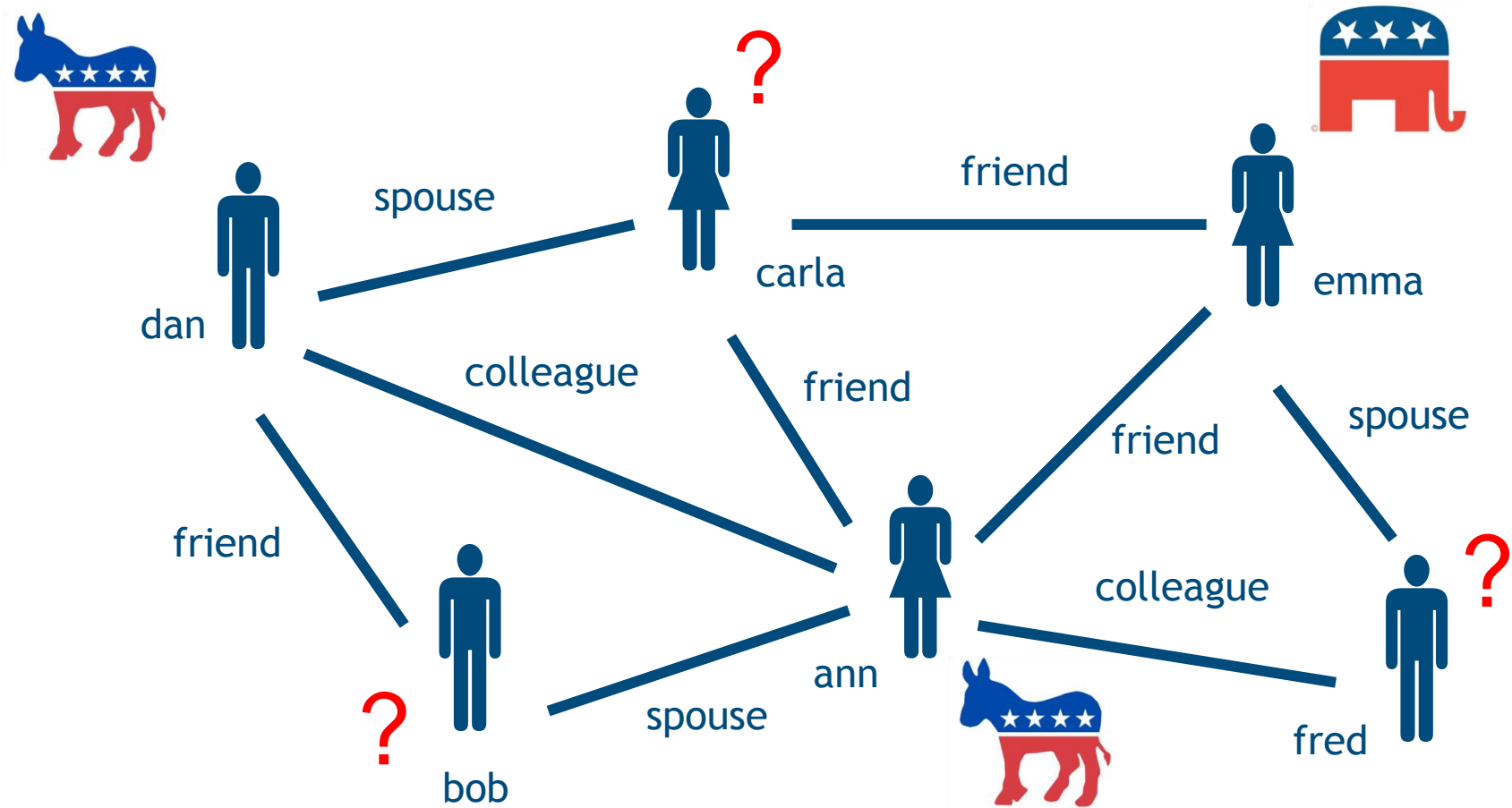
Automatically extracted co-author network:  
which nodes refer to the same person?

# Viral Marketing

Which advertising strategy maximizes expected profit?

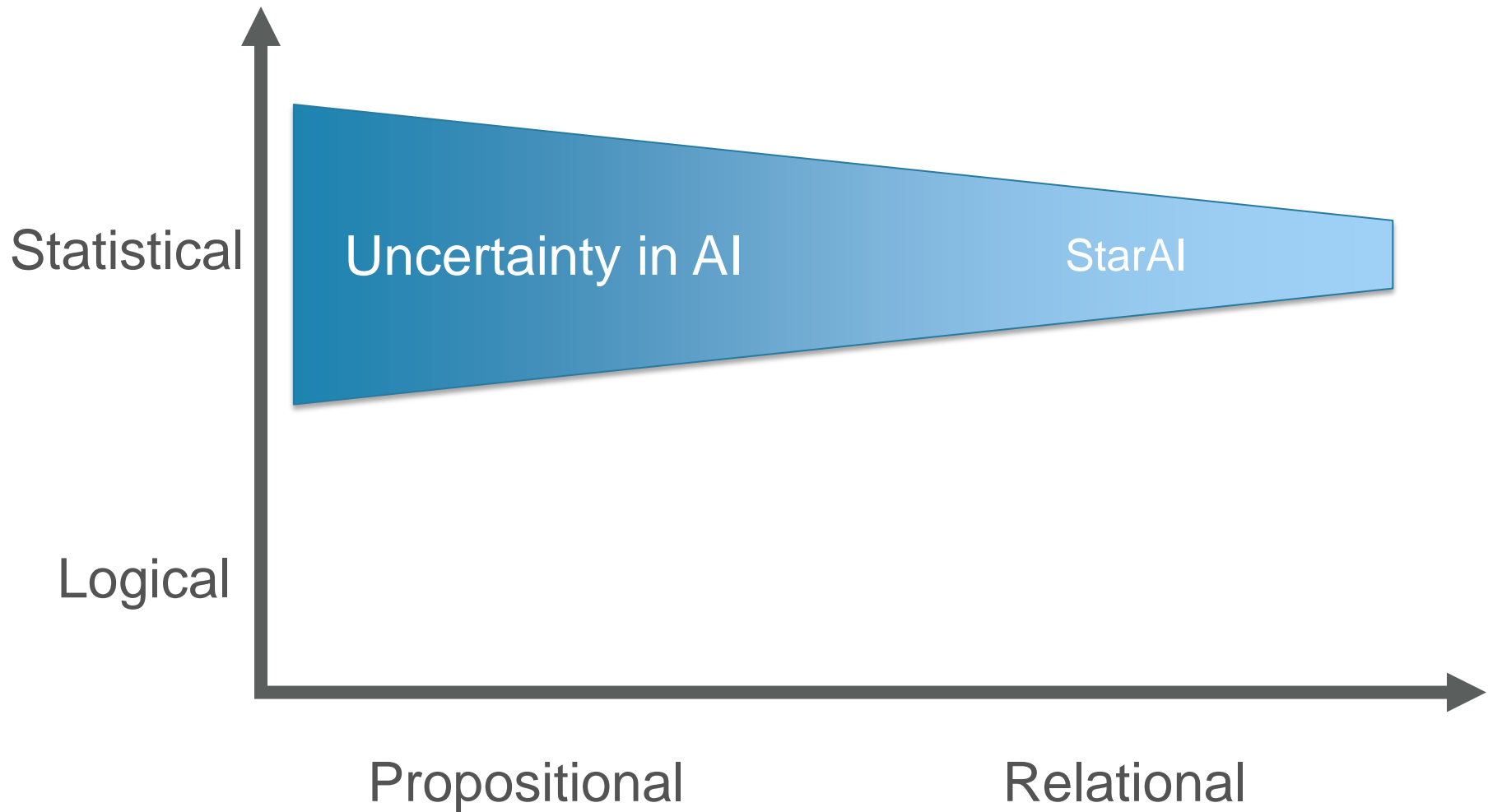


# Voter Opinion Modeling

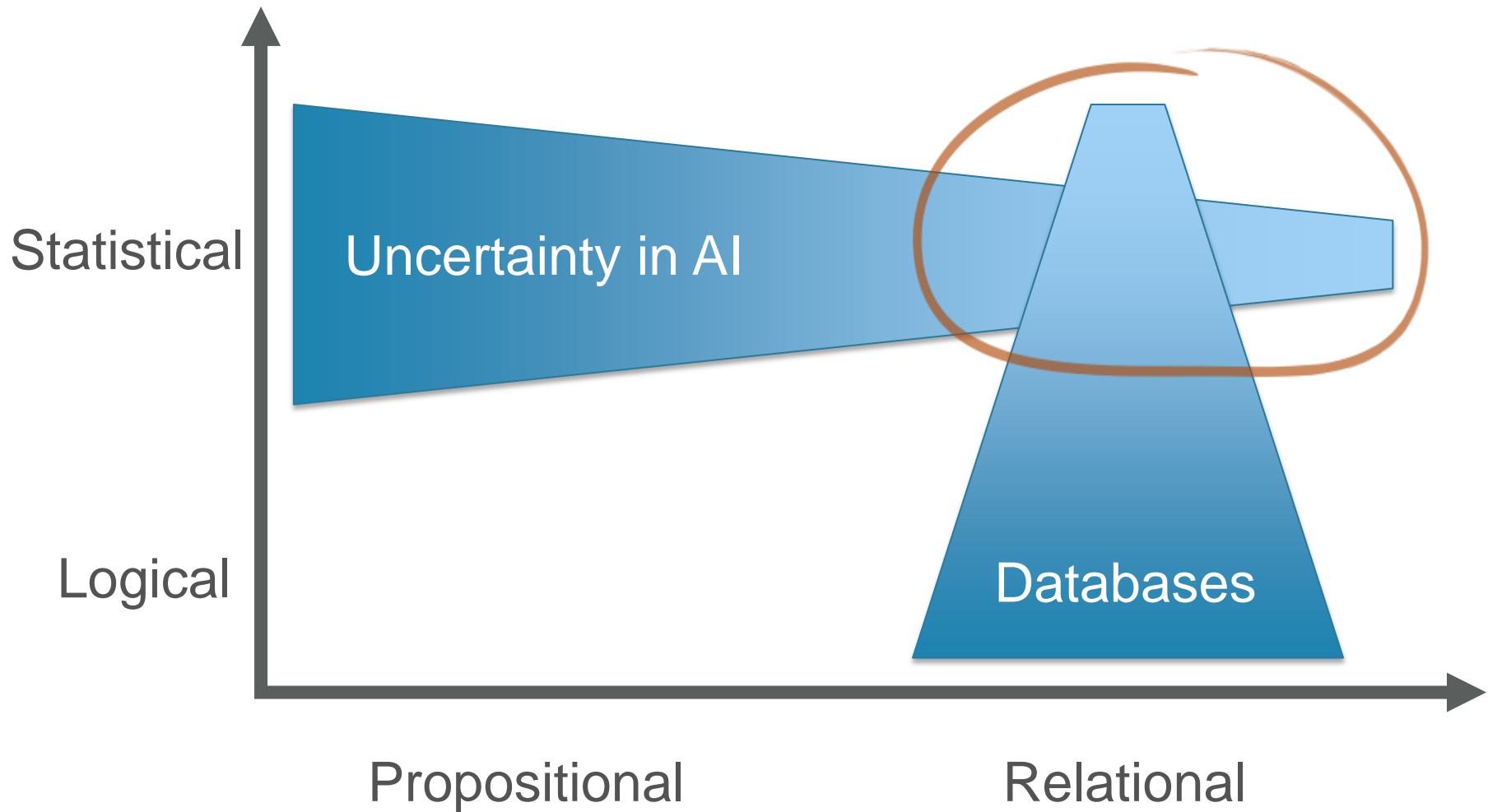


Can we predict preferences?

# Summary



# Summary

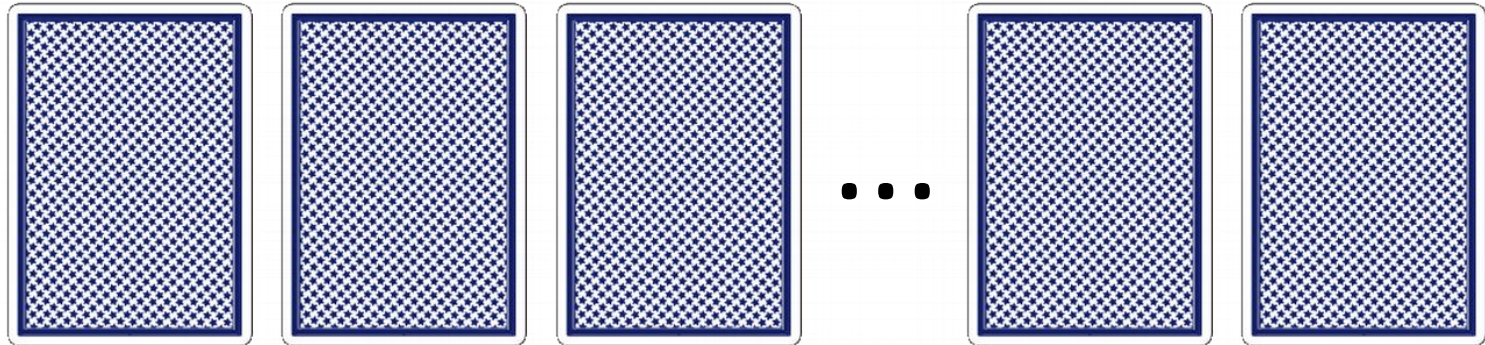


# Why Lifted Inference?

- Main idea: exploit high level relational representation to speed up reasoning
- Let's see an example...

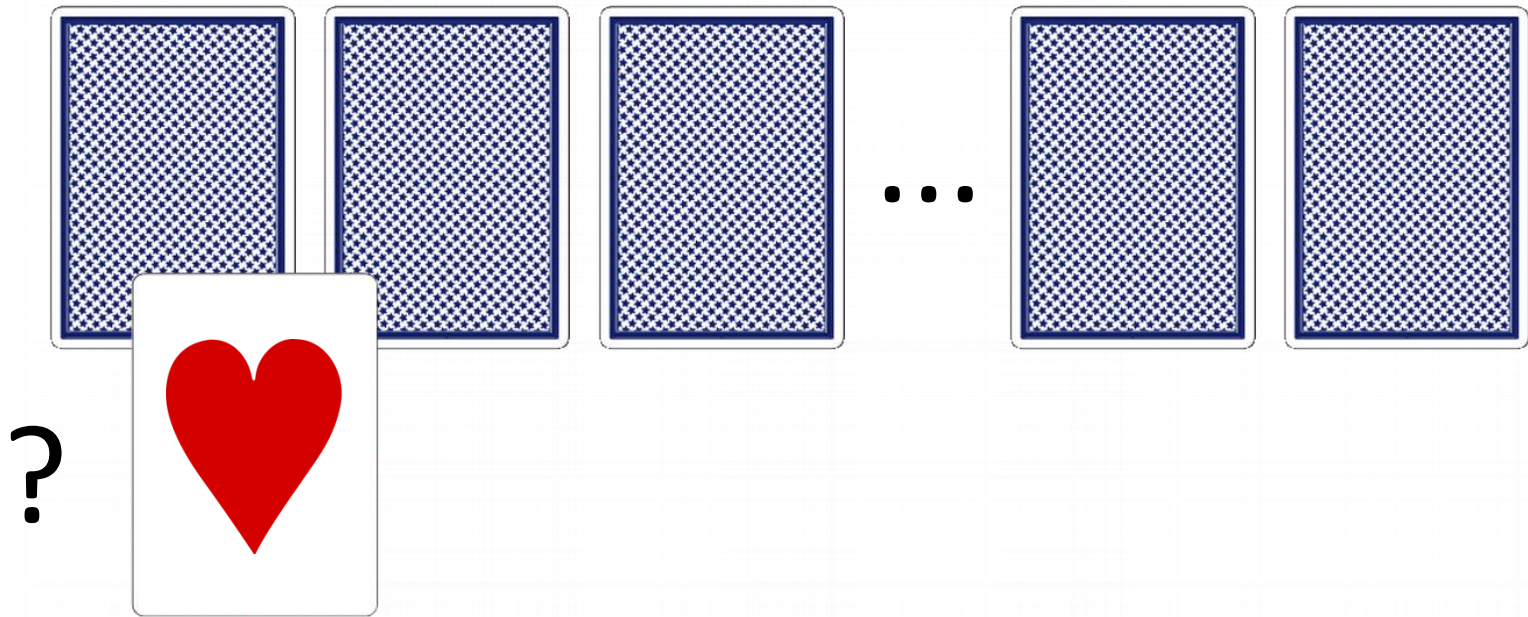


# A Simple Reasoning Problem



- 52 playing cards
- Let us ask some simple questions

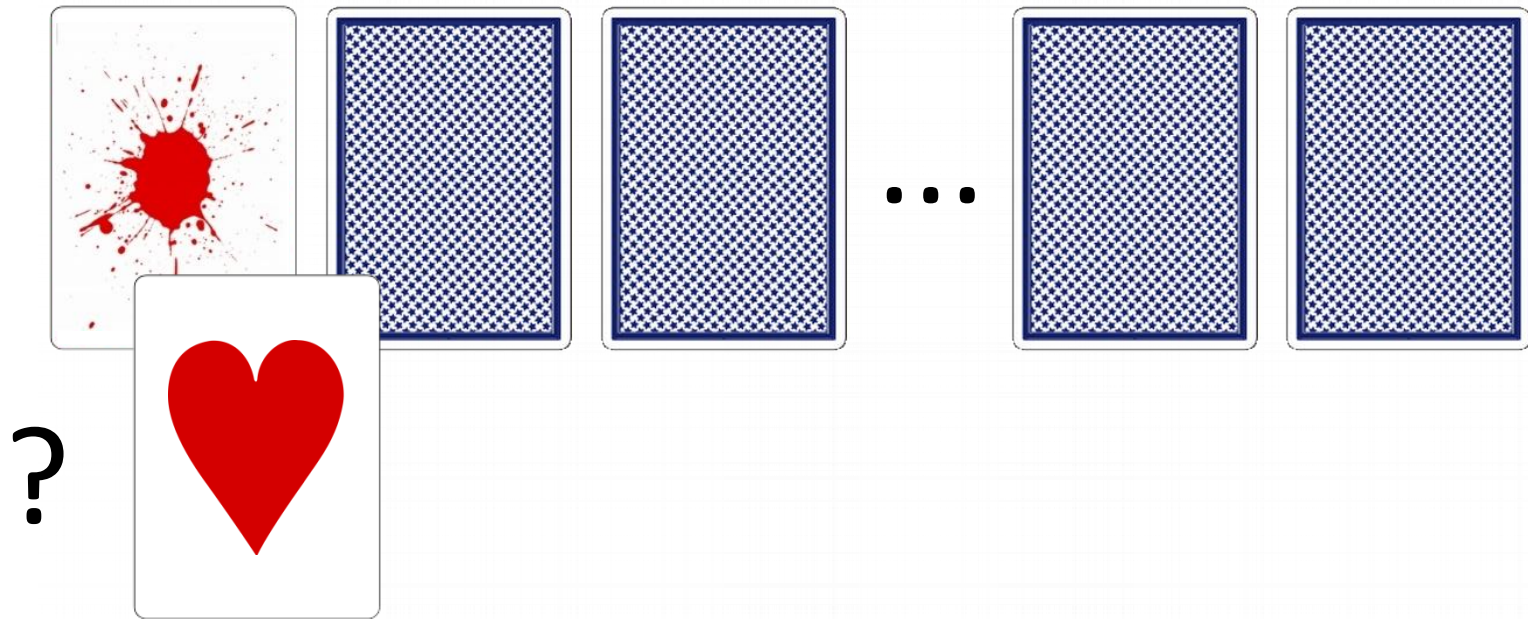
# A Simple Reasoning Problem



*Probability that Card1 is Hearts?*

$1/4$

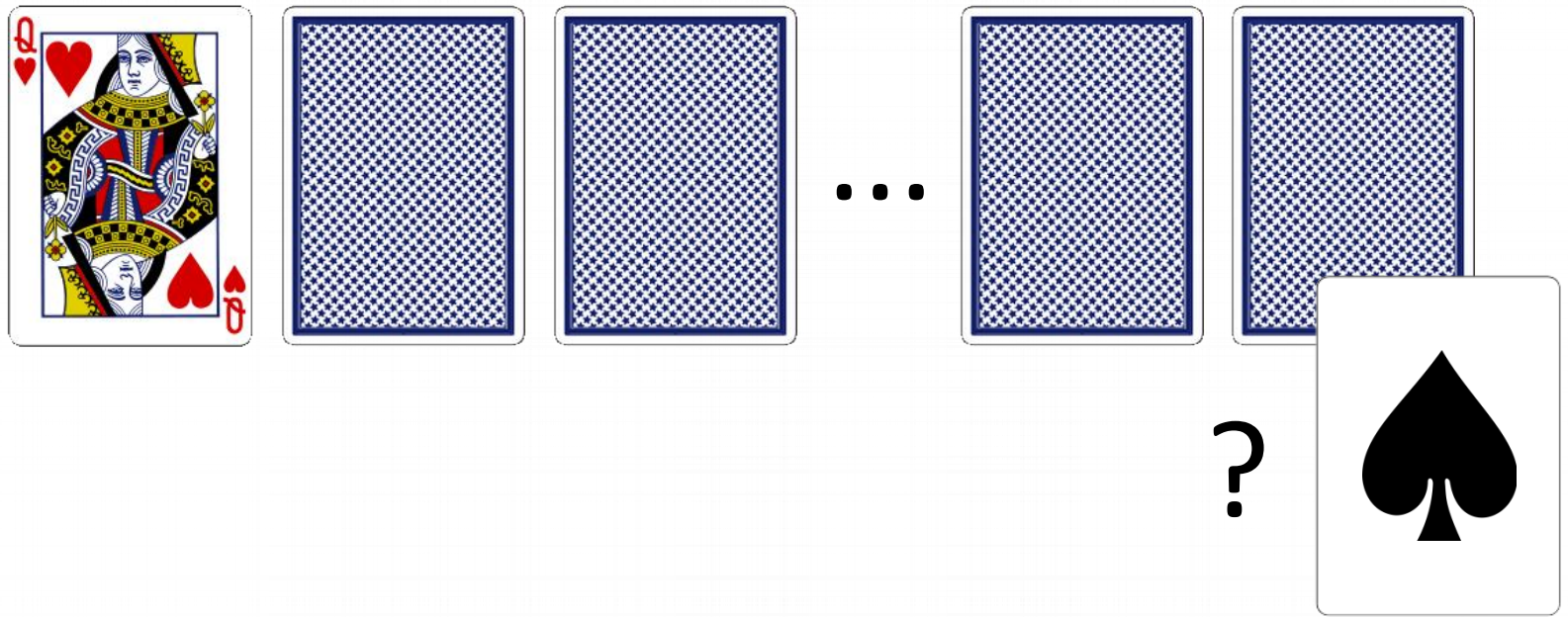
# A Simple Reasoning Problem



*Probability that Card1 is Hearts  
given that Card1 is red?*

$1/2$

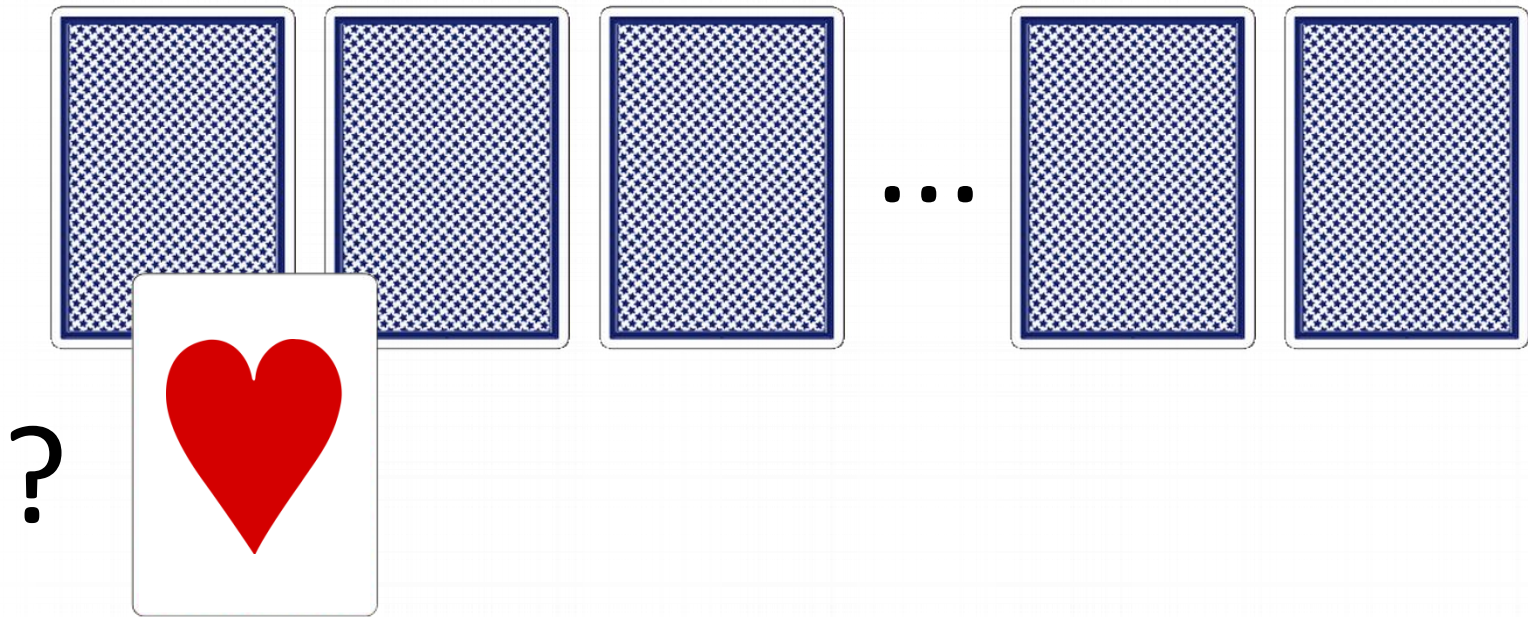
# A Simple Reasoning Problem



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

13/51

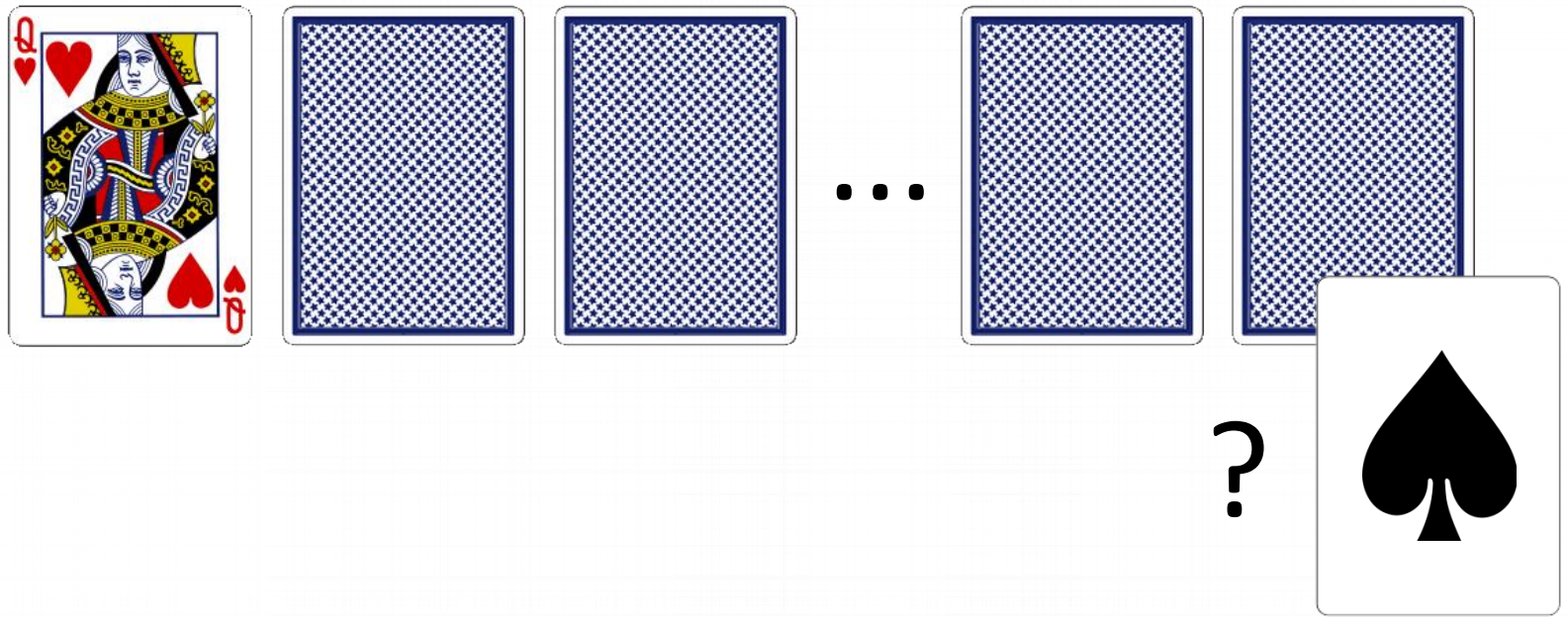
# A Simple Reasoning Problem



*Probability that Card1 is Hearts?*

$1/4$

# A Simple Reasoning Problem



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

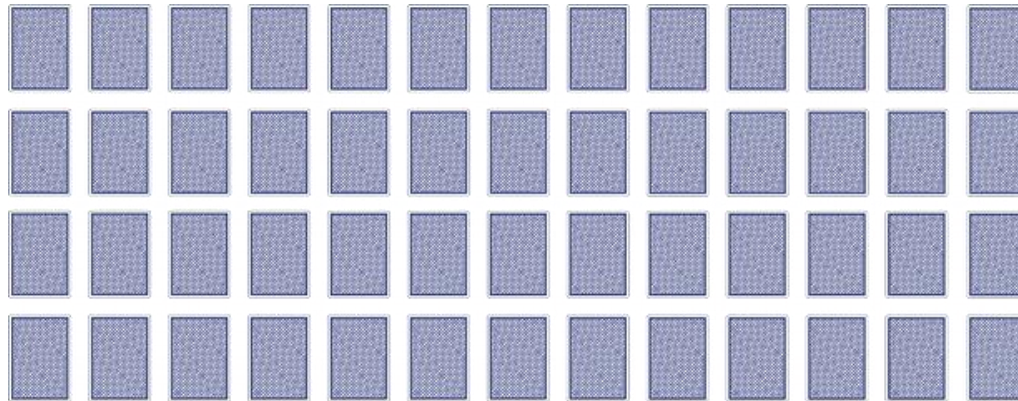
13/51



# Automated Reasoning

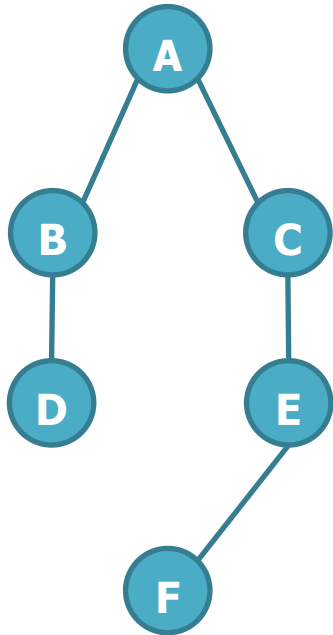
Let us automate this:

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

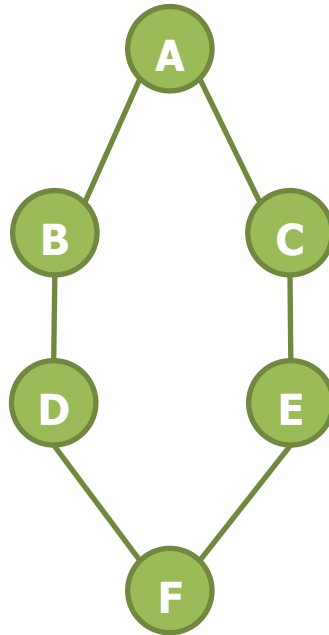


2. Probabilistic inference algorithm  
(e.g., variable elimination or junction tree)

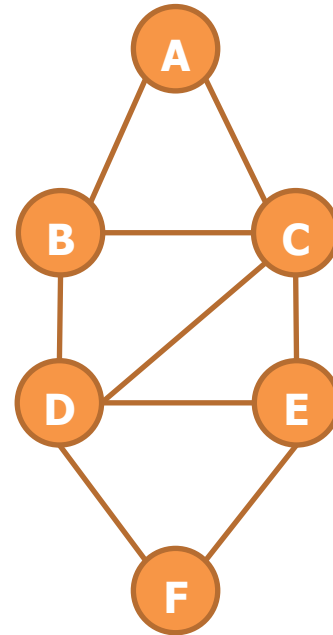
# Classical Reasoning



*Tree*



*Sparse Graph*



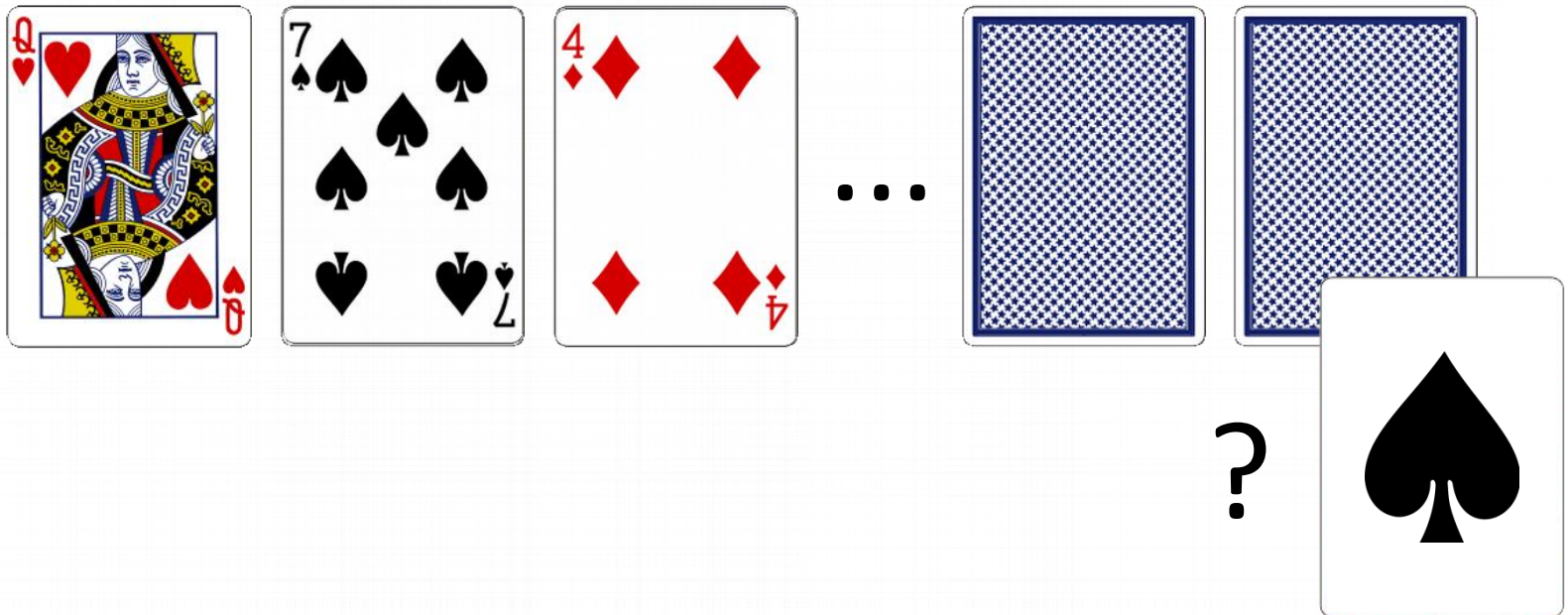
*Dense Graph*



- Higher treewidth
- Fewer conditional independencies
- Slower inference



# Is There Conditional Independence?



$$P(\text{Card52} \mid \text{Card1}) \neq P(\text{Card52} \mid \text{Card1}, \text{Card2})$$

$$13/51 \neq 12/50$$

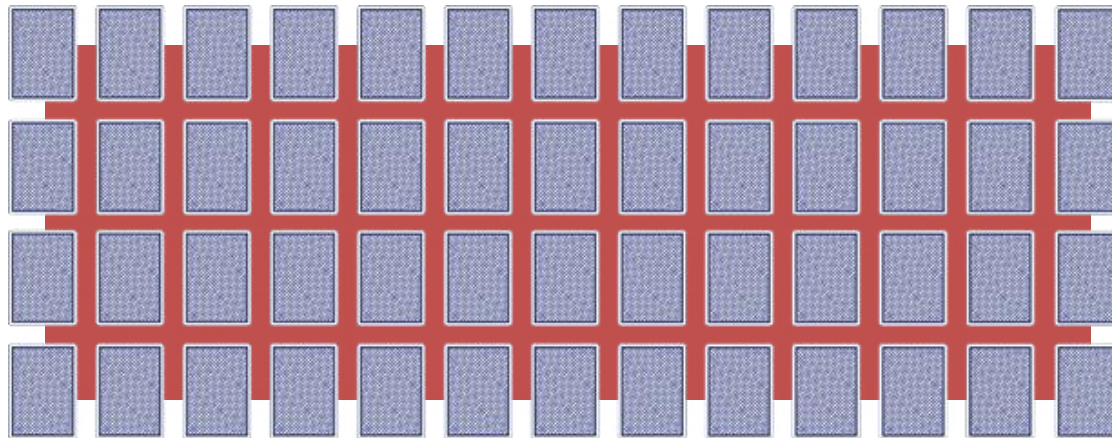
$$P(\text{Card52} \mid \text{Card1}, \text{Card2}) \neq P(\text{Card52} \mid \text{Card1}, \text{Card2}, \text{Card3})$$

$$12/50 \neq 12/49$$

# Automated Reasoning

Let us automate this:

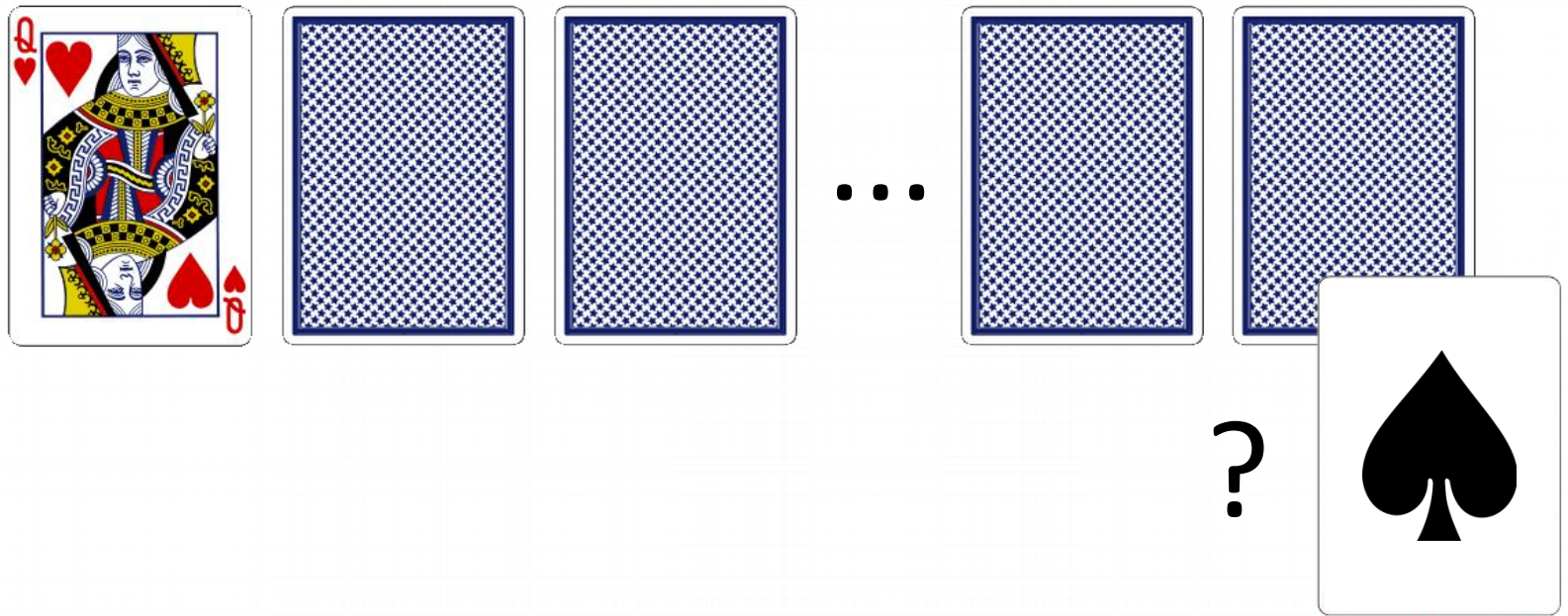
1. Probabilistic graphical model (e.g., factor graph)  
is fully connected!



(artist's impression)

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

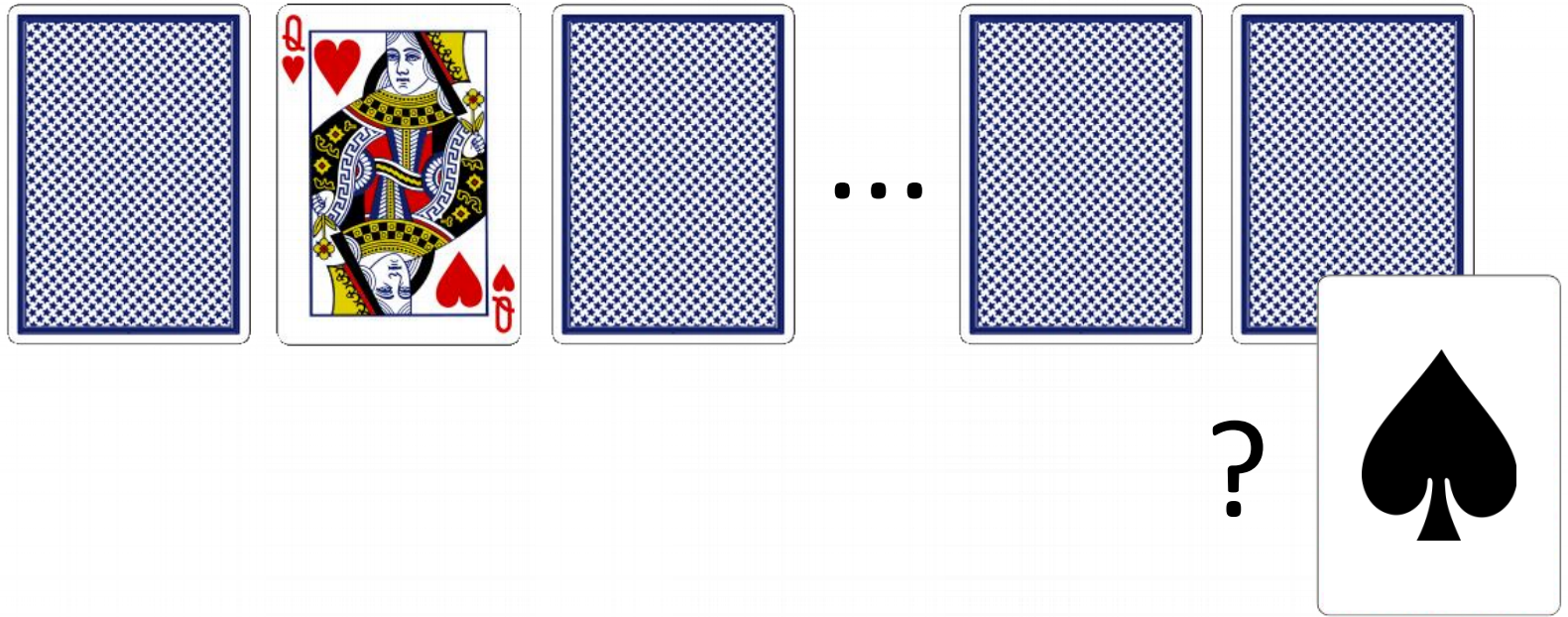
# What's Going On Here?



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

13/51

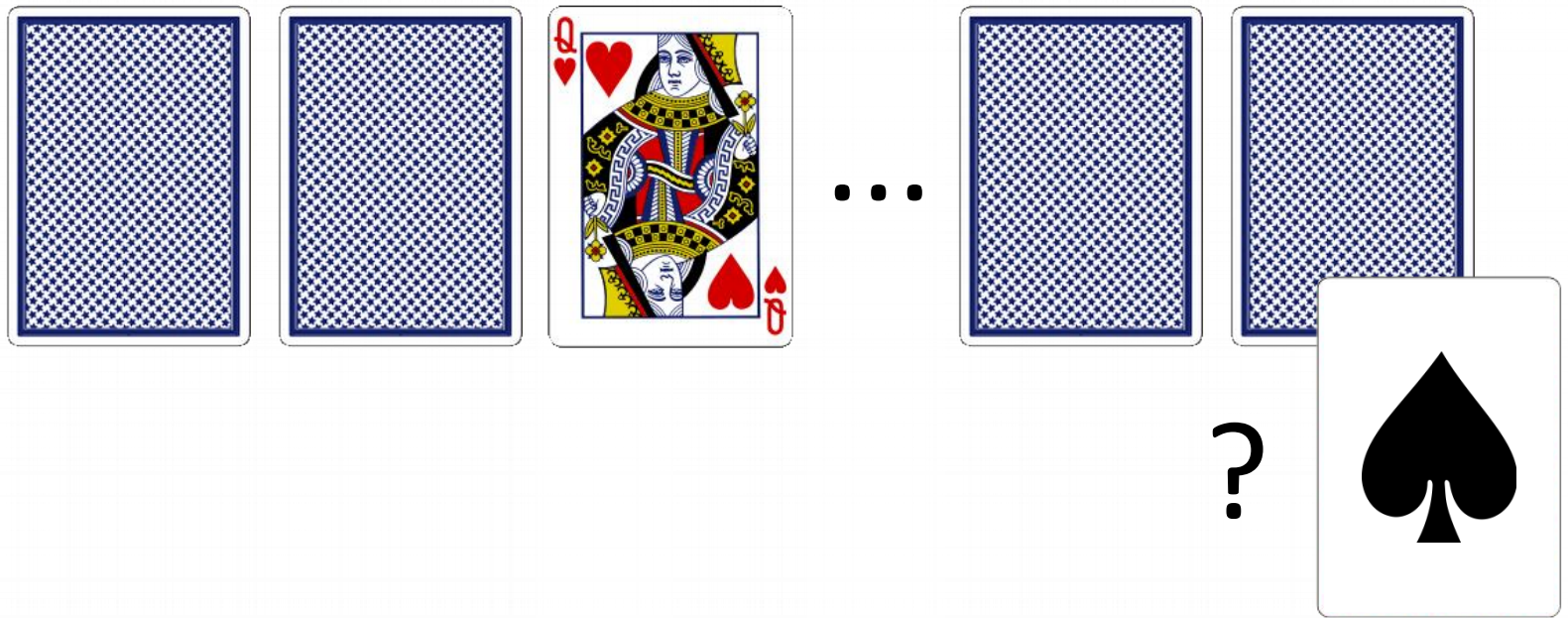
# What's Going On Here?



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

13/51

# What's Going On Here?

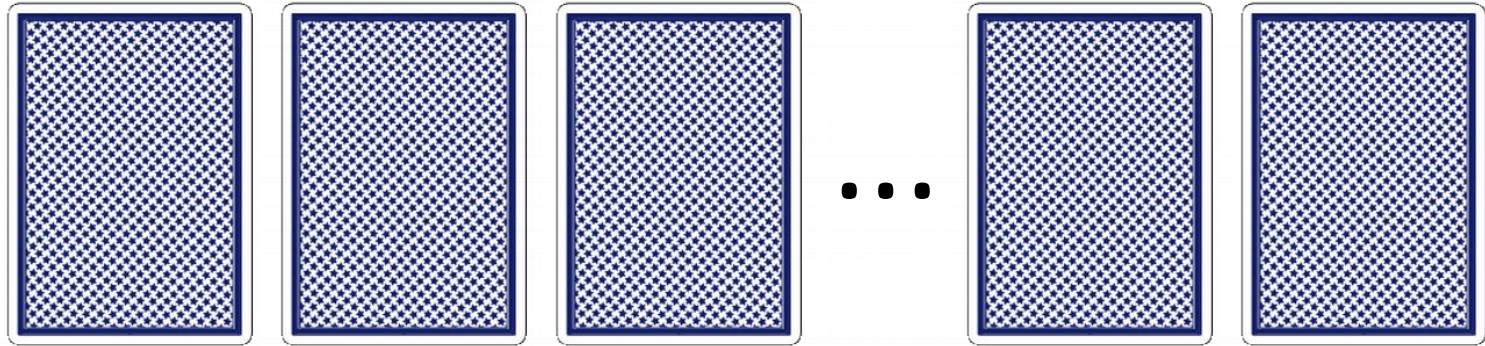


*Probability that Card52 is Spades  
given that Card3 is QH?*

13/51



# Tractable Reasoning



What's going on here?

Which property makes reasoning tractable?

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

⇒ **Lifted Inference**

# Automated 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

# Other Examples of Lifted Inference

- First-order resolution

$$\begin{array}{l} \forall x, \text{Human}(x) \Rightarrow \text{Mortal}(x) \\ \forall x, \text{Greek}(x) \Rightarrow \text{Human}(x) \end{array}$$

implies

$$\forall x, \text{Greek}(x) \Rightarrow \text{Mortal}(x)$$



# Other Examples of Lifted Inference

- First-order resolution
- Reasoning about populations

We are investigating a rare disease. The disease is more rare in women, presenting only in **one in every two billion women** and **one in every billion men**. Then, assuming there are **3.4 billion men** and **3.6 billion women** in the world, the probability that **more than five people** have the disease is

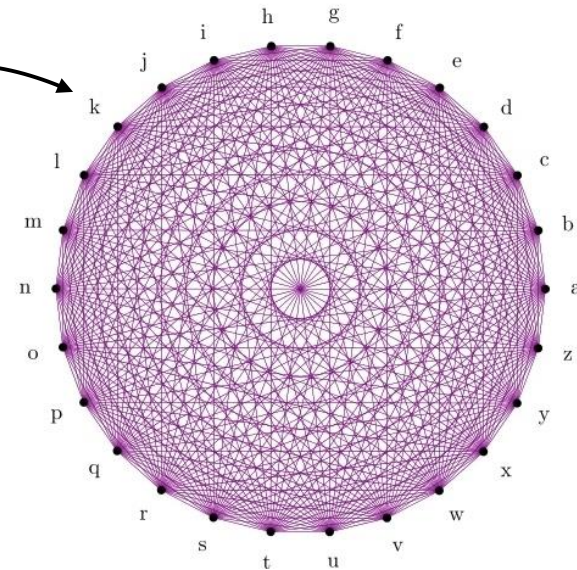
$$1 - \sum_{n=0}^5 \sum_{f=0}^n \binom{3.6 \cdot 10^9}{f} \left(1 - 0.5 \cdot 10^{-9}\right)^{3.6 \cdot 10^9 - f} \left(0.5 \cdot 10^{-9}\right)^f \\ \times \binom{3.4 \cdot 10^9}{(n-f)} \left(1 - 10^{-9}\right)^{3.4 \cdot 10^9 - (n-f)} \left(10^{-9}\right)^{(n-f)}$$

# 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!



# Statistical Properties

## 1. Independence

$$P\left( \begin{array}{|c|c|c|c|} \hline \text{Name} & \text{Cough} & \text{Asthma} & \text{Smokes} \\ \hline \text{Alice} & 1 & 1 & 0 \\ \hline \text{Bob} & 0 & 0 & 0 \\ \hline \text{Charlie} & 0 & 1 & 0 \\ \hline \end{array} \right) = P\left( \begin{array}{|c|c|c|c|} \hline \text{Alice} & 1 & 1 & 0 \\ \hline \end{array} \right) \times P\left( \begin{array}{|c|c|c|c|} \hline \text{Bob} & 0 & 0 & 0 \\ \hline \end{array} \right) \times P\left( \begin{array}{|c|c|c|c|} \hline \text{Charlie} & 0 & 1 & 0 \\ \hline \end{array} \right)$$

## 2. Partial Exchangeability

$$P\left( \begin{array}{|c|c|c|c|} \hline \text{Name} & \text{Cough} & \text{Asthma} & \text{Smokes} \\ \hline \text{Alice} & 1 & 1 & 0 \\ \hline \text{Bob} & 0 & 0 & 0 \\ \hline \text{Charlie} & 0 & 1 & 0 \\ \hline \end{array} \right) = P\left( \begin{array}{|c|c|c|c|} \hline \text{Name} & \text{Cough} & \text{Asthma} & \text{Smokes} \\ \hline \text{Charlie} & 1 & 1 & 0 \\ \hline \text{Alice} & 0 & 0 & 0 \\ \hline \text{Bob} & 0 & 1 & 0 \\ \hline \end{array} \right)$$

## 3. Independent and identically distributed (i.i.d.) = Independence + Partial Exchangeability

# Statistical Properties for Tractability

- Tractable classes independent of representation
- Traditionally:
  - Tractable learning from **i.i.d. data**
  - Tractable inference when **cond. independence**
- New understanding:
  - Tractable learning from **exchangeable data**
  - Tractable inference when
    - **Conditional independence**
    - **Conditional exchangeability**
    - **A combination**

# Summary of Motivation

- Relational data is everywhere:
  - Databases in industry and sciences
  - Knowledge bases
  - Probabilistically extracted/learned/queried
- Lifted inference:
  - Use relational structure during reasoning
  - Very efficient where traditional methods break

This tutorial: Lifted Inference in Relational Models