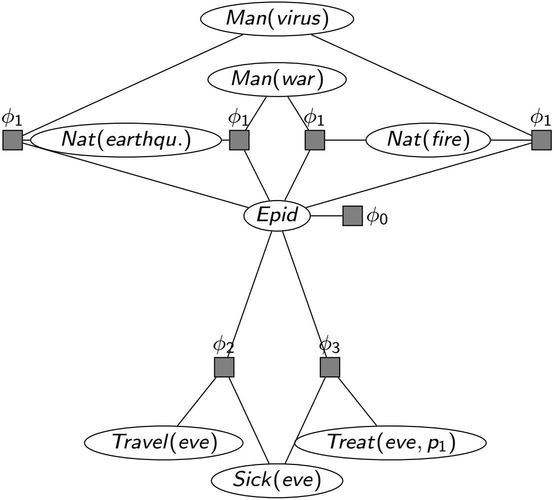
Uncertain Evidence for Probabilistic Relational Models

Marcel Gehrke, Tanya Braun, Ralf Möller

Institute of Information Systems
University of Lübeck

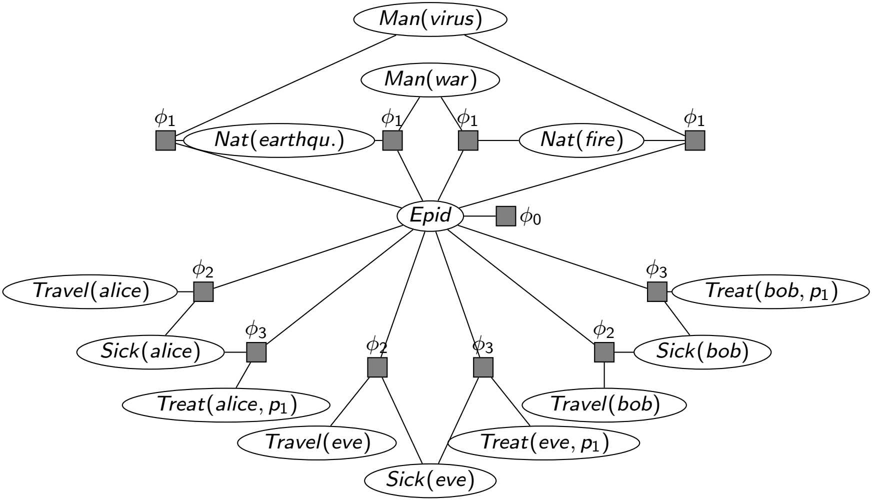
May 30, 2019

Probabilistic Graphical Models



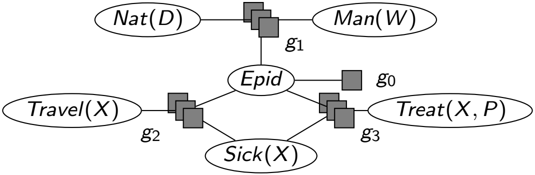
Query answering (QA): Eliminate all non-query variables

Probabilistic Graphical Models



Query answering (QA): Eliminate all non-query variables

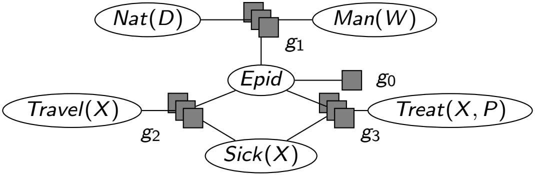
Probabilistic Relational Models



- Parameterisation
  - $\rightarrow$ Compact representation for isomorphic instances
  - $\rightarrow$ Identical observations for sets of random variables
- E.g., lifted variable elimination (LVE)[1] for query answering
  - $\rightarrow$ Elimination: $\sum$ over range values of random variables
  - $\rightarrow$ Lifting: eliminate once and account for isomorphic instances

[1] Poole (2003), de Salvo Braz et al. (2006), Milch et al. (2008), Apsel & Brafman (2011), Taghipour et al. (2013)

Probabilistic Relational Models



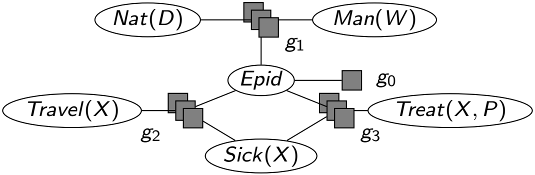
- Parameterisation
  - → Compact representation for isomorphic instances
  - → Identical observations for sets of random variables
- E.g., lifted variable elimination (LVE)[1] for query answering
  - → Elimination: $\sum$ over range values of random variables
  - → Lifting: eliminate once and account for isomorphic instances

Tractable inference w.r.t. domain sizes[2]

[1] Poole (2003), de Salvo Braz et al. (2006), Milch et al. (2008), Apsel & Brafman (2011), Taghipour et al. (2013)
[2] Niepert and Van den Broeck (2014)

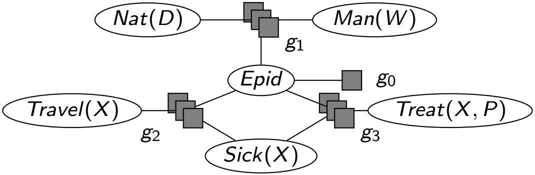
Query for a Conditional Probability Distribution

$P(Sick(eve) | Sick(alice) = true, Sick(bob) = true)$

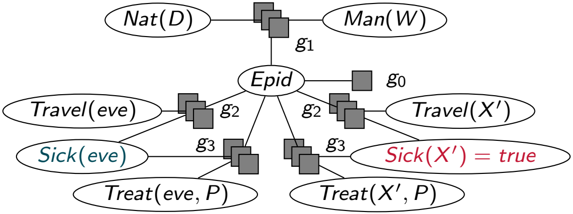


Query for a Conditional Probability Distribution

$P(Sick(eve) | Sick(alice) = true, Sick(bob) = true)$

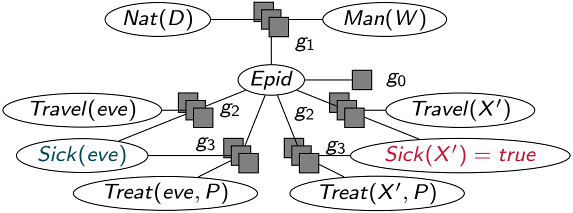


Shattering: Split factors for parts with and without evidence



Query for a Conditional Probability Distribution

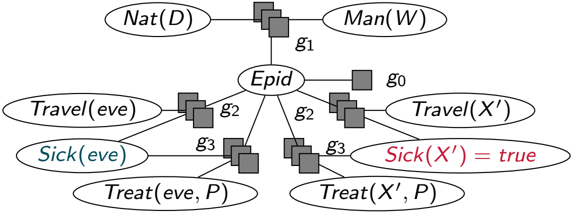
$P(Sick(eve) | Sick(alice) = true, Sick(bob) = true)$



Absorption: Absorb evidence at g_2 and g_3, e.g., for g_2:

Query for a Conditional Probability Distribution

$P(Sick(eve) | Sick(alice) = true, Sick(bob) = true)$

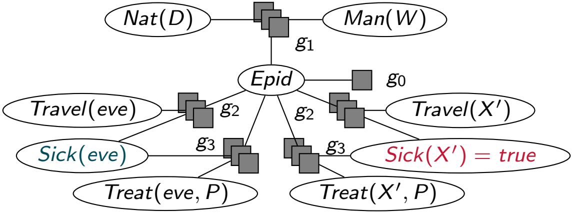


Absorption: Absorb evidence at g_2 and g_3, e.g., for g_2:

S	T	E	g_2
false	*false*	*false*	1
false	*false*	*true*	2
false	*true*	*false*	3
false	*true*	*true*	4
true	*false*	*false*	5
true	*false*	*true*	6
true	*true*	*false*	7
true	*true*	*true*	8

Query for a Conditional Probability Distribution

$P(Sick(eve) | Sick(alice) = true, Sick(bob) = true)$

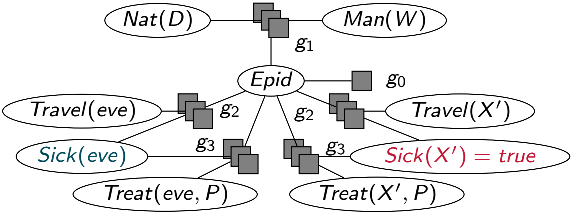


Absorption: Absorb evidence at g_2 and g_3, e.g., for g_2:

S	T	E	g_2
false	*false*	*false*	1·0
false	*false*	*true*	2·0
false	*true*	*false*	3·0
false	*true*	*true*	4·0
true	*false*	*false*	5·1
true	*false*	*true*	6·1
true	*true*	*false*	7·1
true	*true*	*true*	8·1

Query for a Conditional Probability Distribution

$P(Sick(eve) | Sick(alice) = true, Sick(bob) = true)$



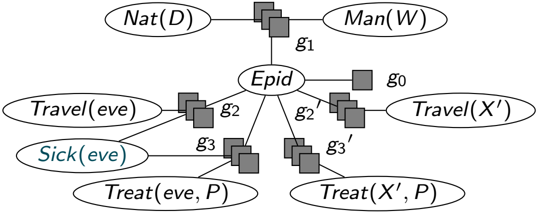
Absorption: Absorb evidence at g_2 and g_3, e.g., for g_2:

S	T	E	g_2
false	*false*	*false*	1·0
false	*false*	*true*	2·0
false	*true*	*false*	3·0
false	*true*	*true*	4·0
true	*false*	*false*	5·1
true	*false*	*true*	6·1
true	*true*	*false*	7·1
true	*true*	*true*	8·1

S	T	E	g_2
~~*false*~~	~~*false*~~	~~*false*~~	~~0~~
~~*false*~~	~~*false*~~	~~*true*~~	~~0~~
~~*false*~~	~~*true*~~	~~*false*~~	~~0~~
~~*false*~~	~~*true*~~	~~*true*~~	~~0~~
true	*false*	*false*	5
true	*false*	*true*	6
true	*true*	*false*	7
true	*true*	*true*	8

Query for a Conditional Probability Distribution

$P(Sick(eve) \mid Sick(alice) = true, Sick(bob) = true)$



Absorption: Absorb evidence at g_2 and g_3, e.g., for g_2:

S	T	E	g_2
false	*false*	*false*	1·0
false	*false*	*true*	2·0
false	*true*	*false*	3·0
false	*true*	*true*	4·0
true	*false*	*false*	5·1
true	*false*	*true*	6·1
true	*true*	*false*	7·1
true	*true*	*true*	8·1

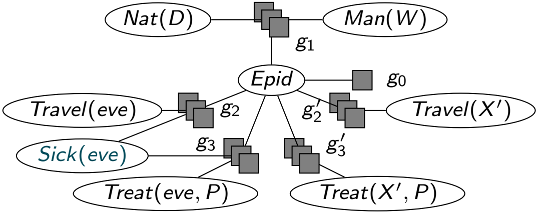
S	T	E	g_2
~~*false*~~	~~*false*~~	~~*false*~~	~~0~~
~~*false*~~	~~*false*~~	~~*true*~~	~~0~~
~~*false*~~	~~*true*~~	~~*false*~~	~~0~~
~~*false*~~	~~*true*~~	~~*true*~~	~~0~~
true	*false*	*false*	5
true	*false*	*true*	6
true	*true*	*false*	7
true	*true*	*true*	8

Dimension reduction:

T	E	g_2'
false	*false*	5
false	*true*	6
true	*false*	7
true	*true*	8

Query for a Conditional Probability Distribution

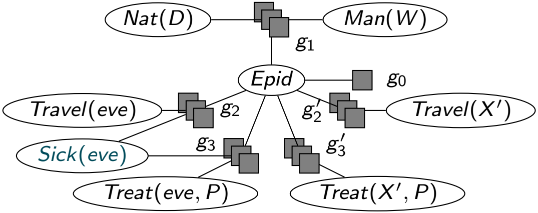
$P(Sick(eve) | Sick(alice) = true, Sick(bob) = true)$



Elimination: Sum-out non-query terms

Query for a Conditional Probability Distribution

$P(Sick(eve) | Sick(alice) = true, Sick(bob) = true)$

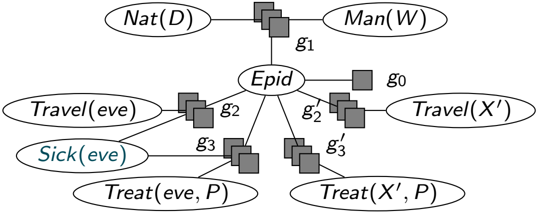


Elimination: Sum-out non-query terms

- $Treat(X', P)$ from g_3'
- $Travel(X')$ from g_2'
- $Treat(eve, P)$ from g_3
- $Travel(eve)$ from g_2
- $Nat(D)$ and $Man(W)$ from g_1
- *Epid* from the product of all factors

Query for a Conditional Probability Distribution

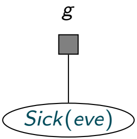
$P(Sick(eve) | Sick(alice) = true, Sick(bob) = true)$



Elimination: Sum-out non-query terms

- $Treat(X', P)$ from g_3'
- $Travel(X')$ from g_2'
- $Treat(eve, P)$ from g_3
- $Travel(eve)$ from g_2
- $Nat(D)$ and $Man(W)$ from g_1
- *Epid* from the product of all factors

Result



Observations

$P(Sick(eve) | Sick(alice) = true, Sick(bob) = true)$

Observation 1: Assumption about Certain Evidence

$Sick(alice) = true$ and $Sick(bob) = true$ assumed to be correct, i.e.,

$$P(Sick(alice) = true) = P(Sick(bob) = true) = 1.0$$

Observations

$P(Sick(eve) | Sick(alice) = true, Sick(bob) = true)$

Observation 1: Assumption about Certain Evidence

$Sick(alice) = true$ and $Sick(bob) = true$ assumed to be correct, i.e.,

$$P(Sick(alice) = true) = P(Sick(bob) = true) = 1.0$$

Observation 2: Existence of Uncertain Evidence

There exist noisy or faulty observations, i.e.,

the assumption from Observation 1 does not hold.

Conference Contribution

Uncertain Evidence

- To account for noise in observations
- Associate observations with a probability (distribution)
- If given a probability p for one range value, distribute the remaining probability $1 - p$ over the remaining range values (max-entropy style)

$\rightarrow$ Goal: Maintain tractable inference

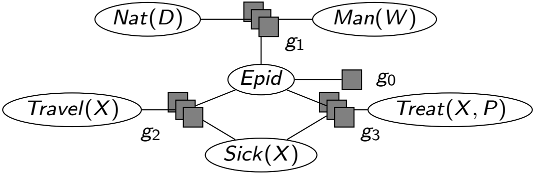
Certain evidence:
Sick(alice) = true
(implicit probability of 1.0)

$\Rightarrow$

Uncertain evidence:
Sick(alice) = true
with a probability of 0.8

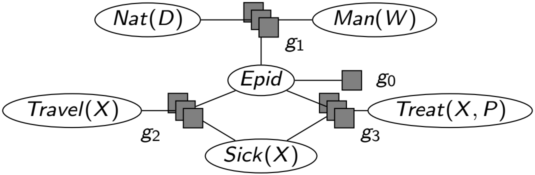
Query for a Conditional Probability Distribution

$P(Sick(eve) | Sick(alice) = true_{0.8}, Sick(bob) = true_{0.8})$

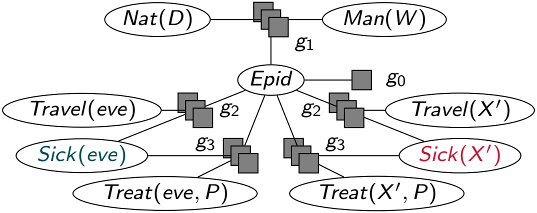


Query for a Conditional Probability Distribution

$P(Sick(eve) | Sick(alice) = true_{0.8}, Sick(bob) = true_{0.8})$

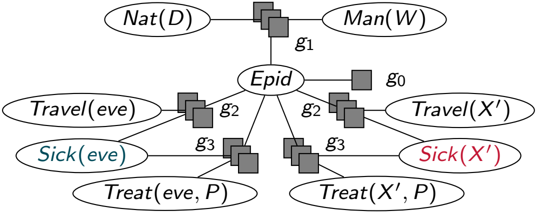


Shattering: Split factors for each observed evidence distribution



Query Answering with Uncertain Evidence

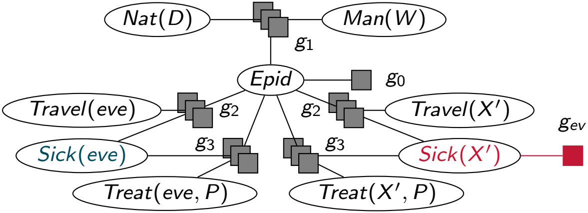
$P(Sick(eve) | Sick(alice) = true_{0.8}, Sick(bob) = true_{0.8})$



~~Absorption~~: Add evidence as factor (no dimension reduction)

Query Answering with Uncertain Evidence

$P(Sick(eve) | Sick(alice) = true_{0.8}, Sick(bob) = true_{0.8})$

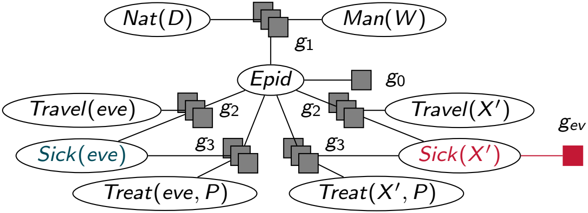


~~Absorption~~: Add evidence as factor (no dimension reduction)

S	g_{ev}
false	0.2
true	0.8

Query Answering with Uncertain Evidence

$P(Sick(eve) | Sick(alice) = true_{0.8}, Sick(bob) = true_{0.8})$



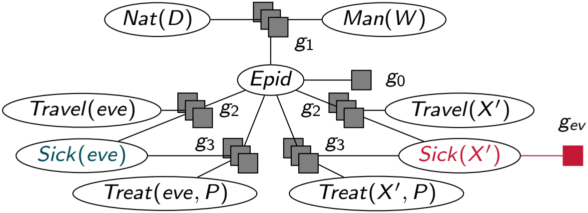
~~Absorption~~: Add evidence as factor (no dimension reduction)
Elimination: Sum-out non-query terms (as before)

S	g_{ev}
false	0.2
true	0.8

S	E	g_2
false	*false*	$4 \cdot 0.2$
false	*true*	$6 \cdot 0.2$
true	*false*	$12 \cdot 0.8$
true	*true*	$14 \cdot 0.8$

Query Answering with Uncertain Evidence

$P(Sick(eve) | Sick(alice) = true_{0.8}, Sick(bob) = true_{0.8})$

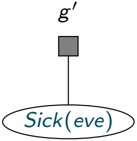


~~Absorption~~: Add evidence as factor (no dimension reduction)
Elimination: Sum-out non-query terms (as before)

S	g_{ev}
false	0.2
true	0.8

S	E	g_2
false	*false*	$4 \cdot 0.2$
false	*true*	$6 \cdot 0.2$
true	*false*	$12 \cdot 0.8$
true	*true*	$14 \cdot 0.8$

Result



Analysis: LVE for Uncertain Evidence

Algorithm steps

1. Build one evidence factor for each observed distribution
2. Add evidence factors to model
3. Eliminate non-query terms
4. Normalise

Lifted query answering

Distributions over evidence

→ Handle noise in observations

Max-entropy for unspecified parts

→ Compact input encoding

But: Variety of observed distributions

→ Faster grounding out

Theoretical Result

Completeness Results

If one distribution given for a set of random variables, runtime complexity of LVE still holds, i.e.,

goal achieved: tractable inference w.r.t. domain sizes ✓
$\Rightarrow$ Completeness results still holds

Theoretical Result

Completeness Results

If one distribution given for a set of random variables, runtime complexity of LVE still holds, i.e.,

goal achieved: tractable inference w.r.t. domain sizes ✓
$\Rightarrow$ Completeness results still holds

Practical Runtime Consequences

Absorption means a dimension reduction, which no longer occurs with uncertain evidence. Therefore, the question is

How much do runtimes suffer with uncertain evidence?

Test Run

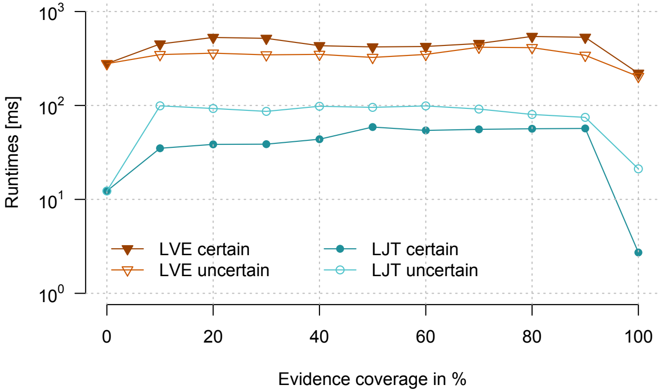


Figure: Grounded model size of 2,001,001

LVE: Implementation by Taghipour (2013) with lifted absorption of certain evidence and extended by us for uncertain evidence

LJT: A multi-query lifted algorithms by us (not part of the talk, but part of the paper)

Conference Contribution

Uncertain Evidence

- To account for noise in observations
- Associate observations with a probability (distribution)
- If given a probability p for one range value, distribute the remaining probability $1 - p$ over the remaining range values (max-entropy style)

$\rightarrow$ Goal: Maintain tractable inference ✓
with limited variety in observed distributions