Rescued from a Sea of Queries

Exact Inference in Probabilistic Relational Models

Colloquium

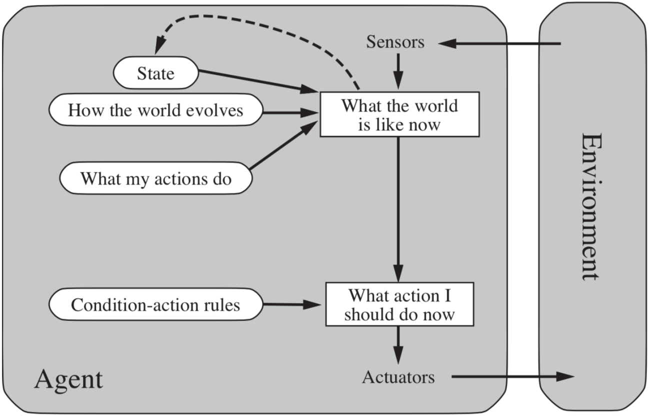
Tanya Braun

Institute of Information Systems
University of Lübeck

February 21, 2020

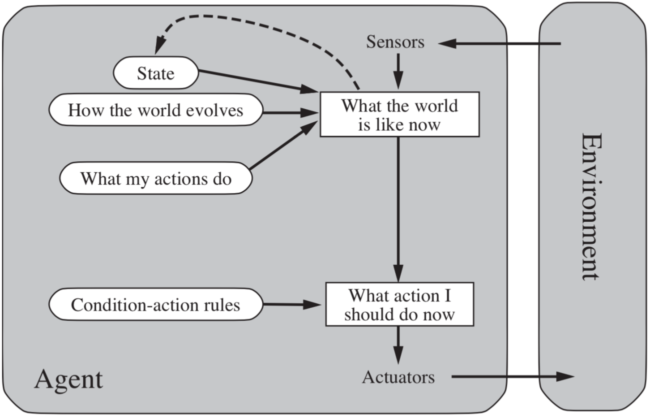
Artificial Intelligence: An Agent Perspective

Russell and Norvig (2010)



Artificial Intelligence: An Agent Perspective

Russell and Norvig (2010)

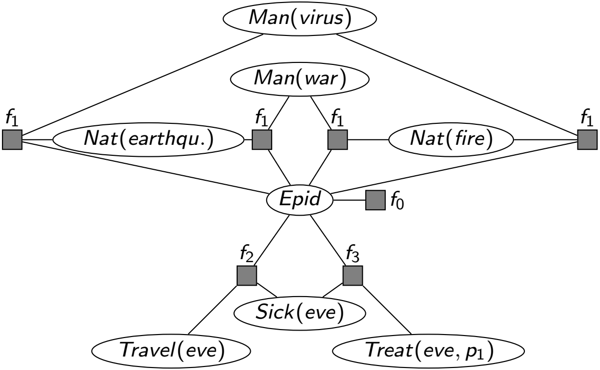


Knowledge representation and reasoning under uncertainty
→ Statistical Relational AI

Probabilistic Graphical Models

Hammersley and Clifford (1971), Kschischang et al. (2001)

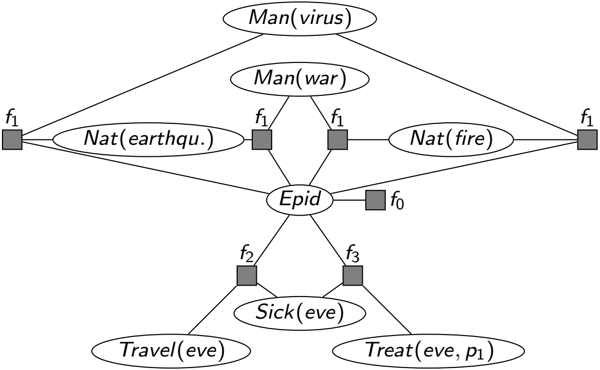
Factor graph F: Compact encoding of full joint distribution $P_F = \frac{1}{Z} \prod_i f_i$



Probabilistic Graphical Models

Hammersley and Clifford (1971), Kschischang et al. (2001)

Factor graph F: Compact encoding of full joint distribution $P_F = \frac{1}{Z}\prod_i f_i$

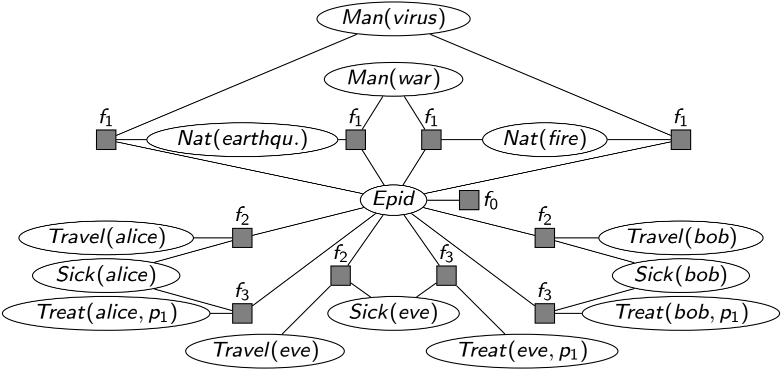


Query answering (QA): Eliminate all non-query variables **avoiding building** P_F

Probabilistic Graphical Models

Hammersley and Clifford (1971), Kschischang et al. (2001)

Factor graph F: Compact encoding of full joint distribution $P_F = \frac{1}{Z} \prod_i f_i$

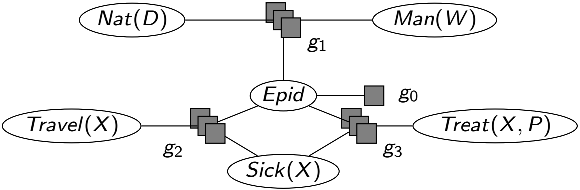


Query answering (QA): Eliminate all non-query variables **avoiding building** P_F

Probabilistic Relational and Lifted Models

Sato (1995), Poole (2003), Ahmadi et al. (2013)

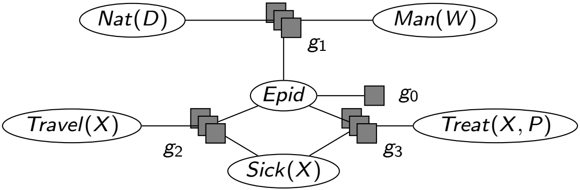
Parfactor graph G: Compact encoding of full joint d. $P_G = \frac{1}{Z} \prod_{f \in gr(G)} f$



Probabilistic Relational and Lifted Models

Sato (1995), Poole (2003), Ahmadi et al. (2013)

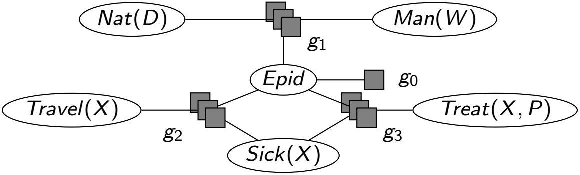
Parfactor graph G: Compact encoding of full joint d. $P_G = \frac{1}{Z} \prod_{f \in gr(G)} f$



QA: Eliminate all non-query variables
while **avoiding grounding G and building P_G**

QA: Lifted Variable Elimination (LVE)

Poole (2003), de Salvo Braz et al. (2005), Milch et al. (2008), Taghipour et al. (2013b)

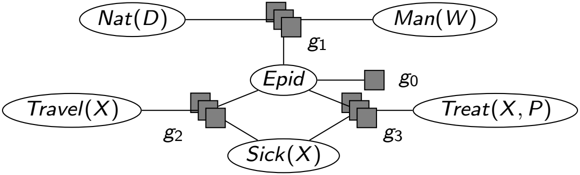


$P(Sick(eve))$

$\sum_V$ indicates a sum over the values of V, $|X|$ a domain size

QA: Lifted Variable Elimination (LVE)

Poole (2003), de Salvo Braz et al. (2005), Milch et al. (2008), Taghipour et al. (2013b)



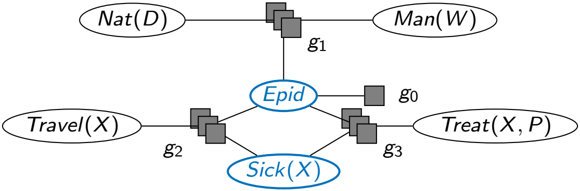
$$P(Sick(eve)) \propto \sum_{Epid} g_0 \left(\sum_{\substack{Sick(X)\\X \neq eve}} \sum_{Travel(X)} g_2 \left(\sum_{Treat(X,P)} g_3 \right)^{|P|} \right)^{|X|_{X \neq eve}}$$

$$\sum_{\#_D[Nat(D)]} \left(\sum_{Man(W)} g_1^{\#} \right)^{|W|}$$

$\sum_V$ indicates a sum over the values of V, $|X|$ a domain size

QA: Conditional Independences and Dynamic Programming

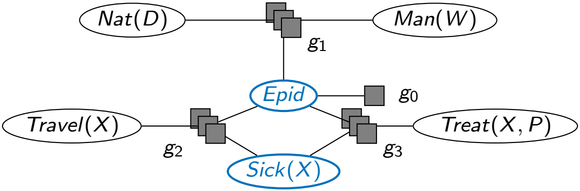
Lauritzen and Spiegelhalter (1988)



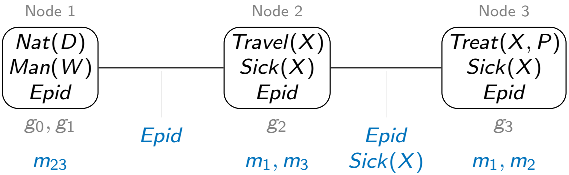
QA based on submodels
ensured to be independent

QA: Conditional Independences and Dynamic Programming

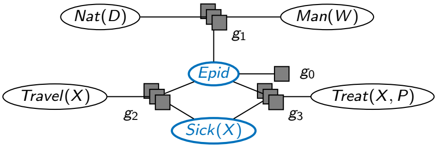
Lauritzen and Spiegelhalter (1988)



QA based on submodels
ensured to be independent



Lifting + Conditional Independences and Beyond

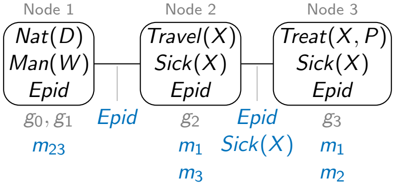


Lifted Junction Tree Alg. (LJT)

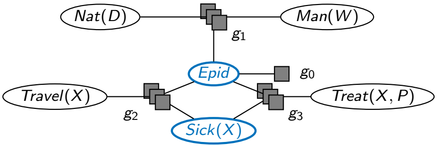
B and Möller (2016)

Answer multiple queries efficiently

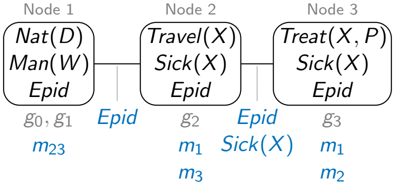
QA based on submodels
ensured to be independent



Lifting + Conditional Independences and Beyond



QA based on submodels
ensured to be independent



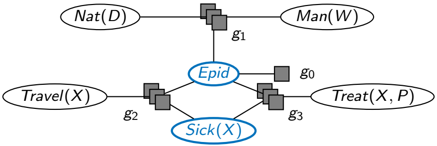
Lifted Junction Tree Alg. (LJT)

Liftability

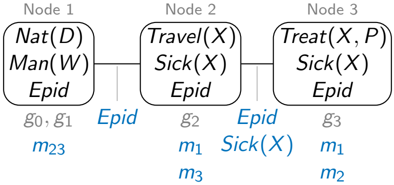
B and Möller (2017)

Avoid message-induced groundings

Lifting + Conditional Independences and Beyond



QA based on submodels ensured to be independent



Lifted Junction Tree Alg. (LJT)

Liftability

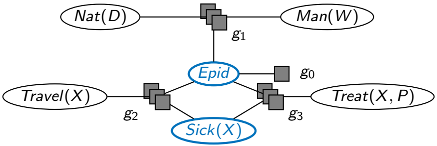
Marginal queries

B and Möller (2018a,c)

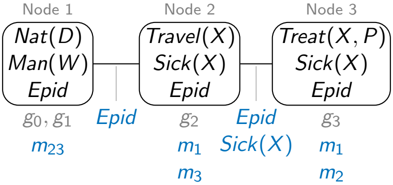
Conjunctive: $P(Sick(eve), Epid)$

Parameterised: $P(Sick(X))$

Lifting + Conditional Independences and Beyond



QA based on submodels
ensured to be independent



Lifted Junction Tree Alg. (LJT)

Liftability

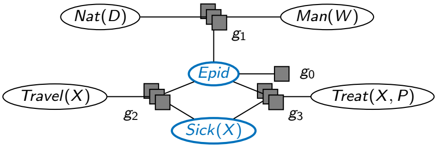
Marginal queries

Assignments queries

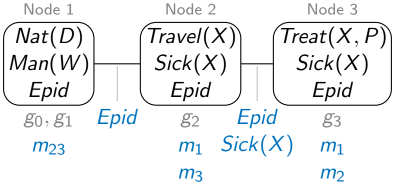
B and Möller (2018b)

LVE + LJT versions using arg max

Lifting + Conditional Independences and Beyond



QA based on submodels ensured to be independent



Lifted Junction Tree Alg. (LJT)

Liftability

Marginal queries

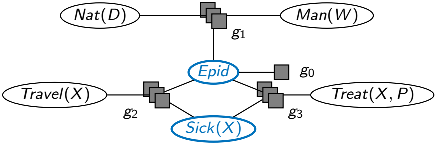
Assignments queries

Complexity & Completeness

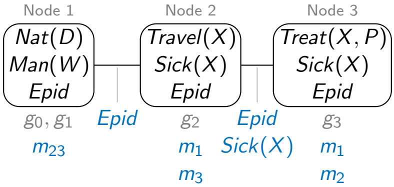
Polynomial w.r.t. domain size

Classes of liftable queries

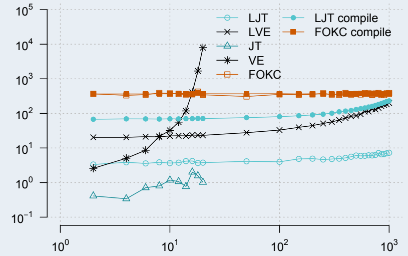
Lifted Inference Continued...



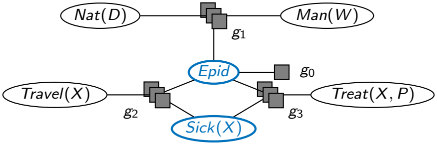
QA based on submodels
ensured to be independent



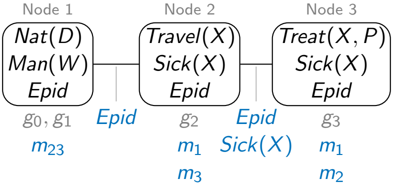
Empirical studies



Lifted Inference Continued...



QA based on submodels ensured to be independent



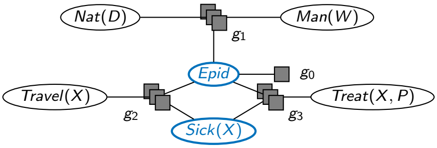
Empirical studies

Adaptive inference

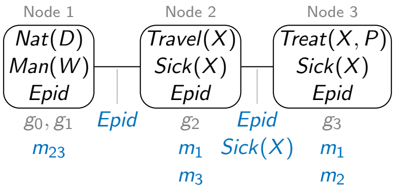
B and Möller (2018e)

Adapt to local changes

Lifted Inference Continued...



QA based on submodels
ensured to be independent



Empirical studies

Adaptive inference

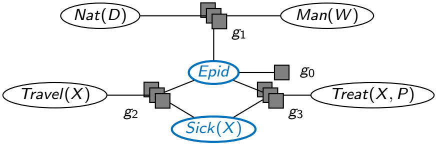
LJT inference framework

B and Möller (2018d)

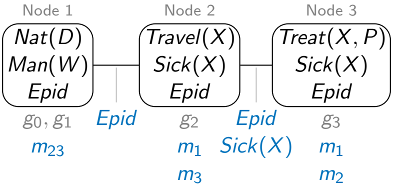
Use other QA algorithms

Conditions apply

Lifted Inference Continued...



QA based on submodels
ensured to be independent



Empirical studies

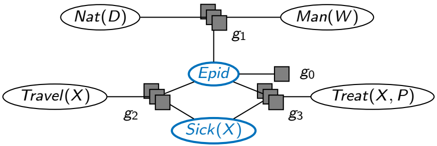
Adaptive inference

LJT inference framework

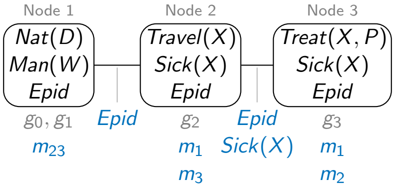
Unknown universe

B and Möller (2019)
Retain tractability

Lifted Inference Continued...



QA based on submodels
ensured to be independent



Empirical studies

Adaptive inference

LJT inference framework

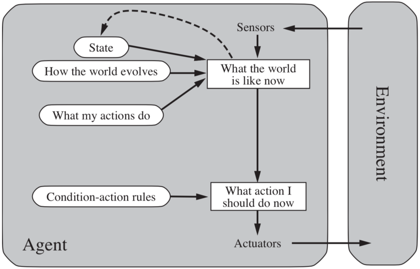
Unknown universe

Continued by/with colleagues

- lifted dynamic models
- lifted decision making
- lifted continuous models

Conclusion

Rescued from a Sea of Queries



Russell and Norvig (2010)

Knowledge representation and reasoning under uncertainty
→ Statistical Relational AI

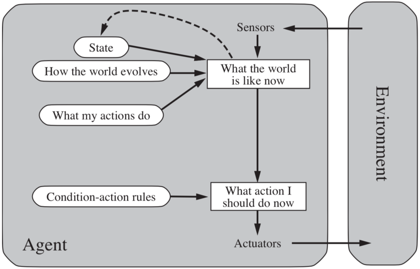
Exact Lifted Inference

- Lifting junction trees
- Lifting marginal and assignment queries
- Classes of liftable queries

Tractable inference for a variety of queries

Conclusion

Rescued from a Sea of Queries



Russell and Norvig (2010)

Knowledge representation and reasoning under uncertainty
$\rightarrow$ Statistical Relational AI

Exact Lifted Inference

- Lifting junction trees
- Lifting marginal and assignment queries
- Classes of liftable queries

Tractable inference for a variety of queries

Future Work

- Going beyond explanations
- Manoeuvring open universes
- Travelling between universes

Appendix

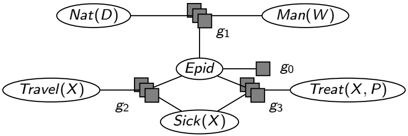
4 Construction
- First-order Decomposition Tree
- First-order Junction Tree
- Fusion

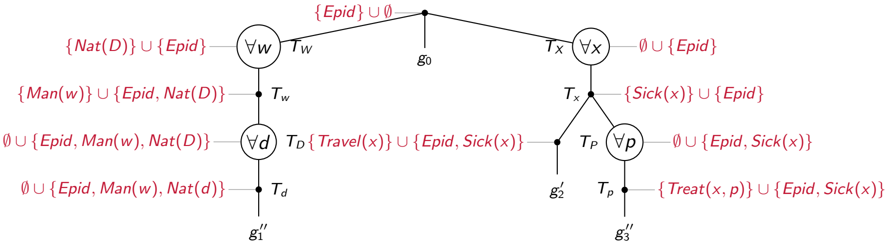
5 Queries
- Parameterised Queries
- Assigment Queries
- Variety of Queries
- Liftable Queries

6 Extensions
- Adaptive Inference
- LJT as a Backbone
- Unknown Universes

First-order Decomposition Tree (FO Dtree)

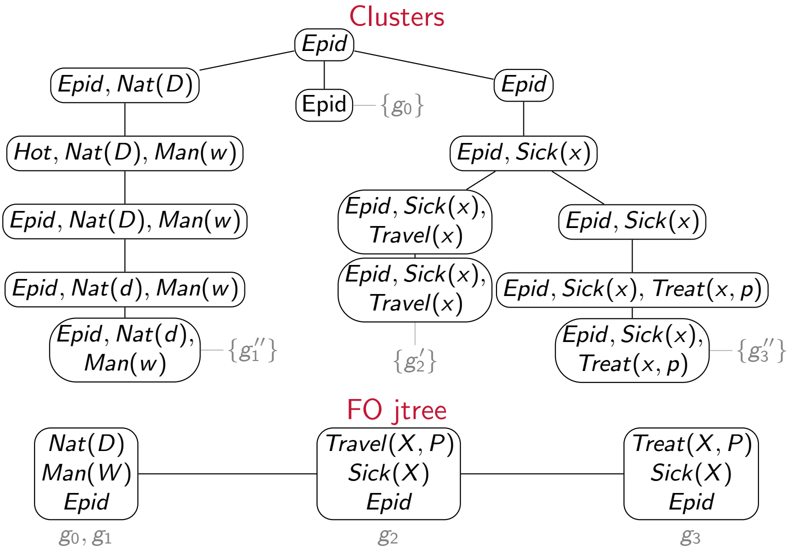
Darwiche (2001), Taghipour et al. (2013a)



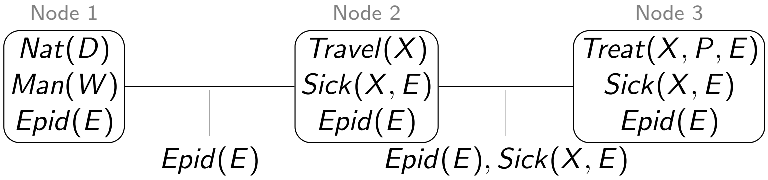


Labels: *cutset(N)* ∪ *context(N)*

First-order Junction Trees (FO Jtrees)



Fusion: Ensuring Lifted Calculations



Elimination order restricted by tree structure AND logical variables

- Lifted summing out of A in g:
  A has to contain all logical variables in g
- Message calculation:
  Terms not on edge need to be eliminated
- $Travel(X)$ has to be eliminated but does not contain X and E

$\rightarrow$ Merge nodes 2 and 3 to avoid elimination

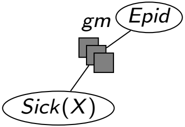
Parameterised Queries

$P(Sick(eve), Sick(alice), Sick(bob))$ vs. $P(Sick(X)_{|\top})$

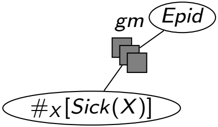
$Sick(alice)$	$Sick(eve)$	$Sick(bob)$	g'
0	0	0	1
0	0	1	2
0	1	0	2
1	0	0	2
1	1	0	3
1	0	1	3
1	1	0	3
1	1	1	4

$\#_X[Sick(X)]$	g'
$[0,3]$	1
$[1,2]$	2
$[2,1]$	3
$[3,0]$	4

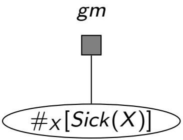
Elimination



Count conversion

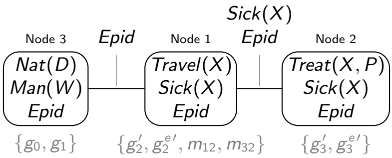


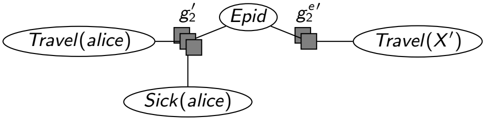
Elimination



Assignment Queries: Most Probable Explanation (MPE)

$\arg\max_{\mathbf{V}} P(\mathbf{V} | Sick(eve) = true, Sick(bob) = true)$





Node 2

LJT^{MPE}

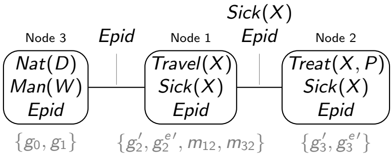
- Absorb evidence at nodes
- Pass messages (inward)
  - $1 \longrightarrow 2$
  - $3 \longrightarrow 2$
- At inner-most node
  - Eliminate variables
  - Output MPE

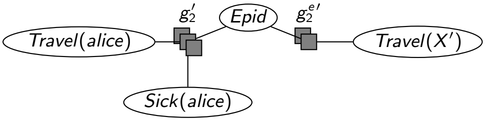
MPE

- $Epid = false$
- $\forall X' \in \{eve, bob\}, P :$ $Treat(X', P) = true,\ Travel(X') = false$
- $Sick(alice) = false,\ Travel(alice) = true,$ $\forall P : Treat(alice, P) = false$
- $\forall D : Nat(D) = false$
- $\forall W : Man(W) = false$

Assignment Queries: Maximum A Posterioi (MAP)

$\arg\max_{Travel(X')} \sum_{\mathbf{T}} P(Travel(X') | Sick(X') = true), X' \in \{eve, bob\}$





Node 1

LJT^{MAP}

- Absorb evidence at nodes
- Pass messages with LVE
- Answer query ($\mathbf{V} = \{Travel(X')\}$)
  - Find nodes covering $\mathbf{V}$
  - Eliminate remaining variables of $\mathbf{T}$
  - Eliminate $\mathbf{V}$

Answer query

- $Travel(eve)$, $Travel(bob)$: node 1
- Eliminate with LVE: $Travel(alice), Sick(alice), Epid$
- Eliminate with LVE^{MPE}: $Travel(X')$
- MAP: $\forall X' \in \{eve, bob\}$: $Travel(X') = false$

LJT for a Variety of Queries

procedure COM-LJT(Model G, Query terms and types $\{(\mathbf{Q}_k, t_k)\}_{k=1}^m$, Evidence $\mathbf{E}$)
  Construct an FO jtree J for G
  Enter $\mathbf{E}$ into J
  Pass messages on J ▷ LVE as subroutine
  for each $(\mathbf{Q}_k, t_k) \in \{(\mathbf{Q}_k, t_k)\}_{k=1}^m$ **do**
    if t_k = MPE **then** ▷ $\mathbf{Q}_k = \emptyset$
      J-MPE-LJT(J) ▷ Output or store result
    else
      Find a subtree J' s.t. $\mathbf{Q}_k \subseteq rv(J')$
      if t_k = MAP $\wedge$ $\mathbf{Q}_k = rv(J')$ **then**
        J-MPE-LJT(J') ▷ Output or store result
      else
        Extract a submodel G' from J'
        if t_k = MAP **then** ▷ $\mathbf{Q}_k \subset rv(J')$
          MAP-LVE(G', $\mathbf{Q}_k$, $\emptyset$) ▷ Output or store result
        else
          LVE(G', $\mathbf{Q}_k$, $\emptyset$) ▷ Output or store result
        end if
      end if
    end if
  end for
end procedure

Liftable Probability Queries

Conjunctive queries

For each logical variable, only one set of constants occurs.

$$P(Sick(eve), Treat(eve, p_1))$$
$$P(Sick(eve), Treat(alice, p_1))$$

Parameterised queries

Each query term contains at most one logical variable and one set of constants per logical variable.

$$P(Sick(X)_{X \in \{alice, eve\}}, Travel(X)_{X \in \{alice, eve\}})$$
$$P(Sick(X)_{X \in \{alice, eve\}}, Travel(X'))_{X' \in \{bob, eve\}})$$

Liftable Assignment Queries

Most probable explanation (MPE)

Liftability results from LVE transfer to MPE.

Maximum a posteriori (MAP)

Each MAP term contains at most one logical variable and one set of constants per logical variable.

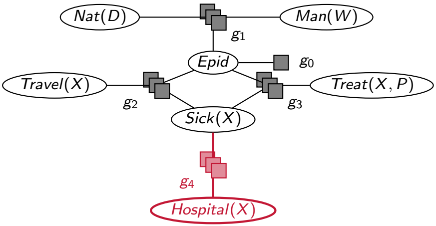
Bounded MAP queries

If the random variables of entire subtrees occur as MAP terms, then the MAP query does not lead to a higher tree width.

MAP query over $Sick(X), Travel(X), Epid$
MAP query over $Treat(X, P), Travel(X), Epid$

Adaptive Inference

Solve each query more efficiently than starting from scratch



Evidence changes (new/retracted)

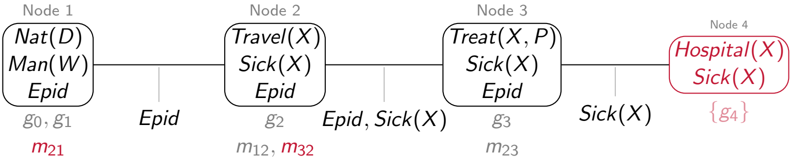
$Travel(alice) = true$

$Travel(bob) = true$

Domains change (new constants)

$X \in \{alice, eve, bob\}$

$\cup \{charlie\}$



LJT as a Backbone for Lifted Inference

Requirements for Subroutines

1. Lifted evidence handling
2. Lifted message calculation (conjunctive, parameterised query)

→ Expressiveness of the query language of the subroutine determines the expressivity of the query language of LJT.

LVE

1. Lifted absorption ✓
2. Eliminate all non-query terms with LVE ✓

→ Marginal or conditional distributions of conjunctions of random variables

FOKC (Van den Broeck, 2013)

1. Lifted conditioning ✓
2. Circuit determines queries easily answered ϟ

→ Marginal or conditional distributions of single random variables

LJT with LVE and FOKC

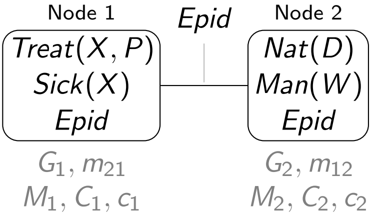
Parfactor model

$g_0(Epid)$

$g_1(Epid, Nat(D), Man(W))$

$g_2(Epid, Sick(X), Treat(X, P))$

First-order Junction Tree J



LVE in steps II + III

FOKC in step IV + V

Algorithm steps

- I. Construct J
- II. Enter evidence $\mathbf{E}$
- III. Pass messages
- IV. For each node i
  - i. Transform node into a Markov logic network M_i
  - ii. Transform M_i into d-DNNF.
  - iii. Build C_i for M_i.
  - iv. Compute WFOMC c_i in C_i.
- V. Answer queries $Q \in \mathbf{Q}$
  - i. Build C_q for $M_i \wedge q$
  - ii. Compute WFOMC c_q in C_q.
  - iii. Compute $\frac{c_q}{c_i}$

Probabilistic Inference with Unknown Universes

Syntactic components:

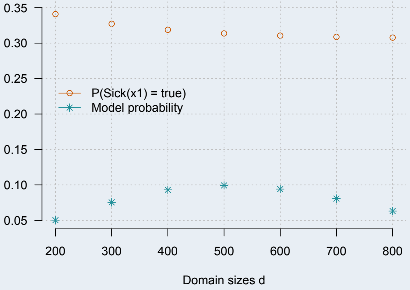
Template model, constraint program, domain worlds

→ Set of possible worlds: Expected values, runtime increases

→ Constraint meta-programming: Build oracle for algorithms

→ Transfer learning: Decoupling from specific domain

New queries emerging



Exploration and model checking, e.g., does the probability of

- an individual being sick decrease with larger domains?
- an epidemic rise if more people travel?

References I

Babak Ahmadi, Kristian Kersting, Martin Mladenov, and Sriraam Natarajan. Exploiting Symmetries for Scaling Loopy Belief Propagation and Relational Training. *Machine Learning*, 92(1):91–132, 2013.

Tanya Braun and Ralf Möller. Lifted Junction Tree Algorithm. In *Proceedings of KI 2016: Advances in Artificial Intelligence*, pages 30–42. Springer, 2016.

Tanya Braun and Ralf Möller. Preventing Groundings and Handling Evidence in the Lifted Junction Tree Algorithm. In *Proceedings of KI 2017: Advances in Artificial Intelligence*, pages 85–98. Springer, 2017.

Tanya Braun and Ralf Möller. Counting and Conjunctive Queries in the Lifted Junction Tree Algorithm - Extended. In *Postproceedings of the 5th International Workshop on Graph Structures for Knowledge Representation and Reasoning at the 26th International Joint Conference on Artificial Intelligence*, pages 54–72. Springer, 2018a.

References II

Tanya Braun and Ralf Möller. Lifted Most Probable Explanation. In *Proceedings of the International Conference on Conceptual Structures*, pages 39–54. Springer, 2018b.

Tanya Braun and Ralf Möller. Parameterised Queries and Lifted Query Answering. In *IJCAI-18 Proceedings of the 27th International Joint Conference on Artificial Intelligence*, pages 4980–4986. IJCAI Organization, 2018c.

Tanya Braun and Ralf Möller. Fusing First-order Knowledge Compilation and the Lifted Junction Tree Algorithm. In *Proceedings of KI 2018: Advances in Artificial Intelligence*, pages 24–37. Springer, 2018d.

Tanya Braun and Ralf Möller. Adaptive Inference on Probabilistic Relational Models. In *Proceedings of AI 2018: Advances in Artificial Intelligence*, pages 487–500. Springer, 2018e.

References III

Tanya Braun and Ralf Möller. Exploring Unknown Universes in Probabilistic Relational Models. In *Proceedings of AI 2019: Advances in Artificial Intelligence*. Springer, 2019.

Adnan Darwiche. Recursive Conditioning. *Artificial Intelligence*, 2(1–2): 4–51, 2001.

Rodrigo de Salvo Braz, Eyal Amir, and Dan Roth. Lifted First-order Probabilistic Inference. In *IJCAI-05 Proceedings of the 19th International Joint Conference on Artificial Intelligence*, pages 1319–1325. IJCAI Organization, 2005.

John Hammersley and Peter Clifford. Markov Fields on Finite Graphs and Lattices. unpublished, 1971.

Frank R. Kschischang, Brendan J. Frey, and Hans-Andrea Loeliger. Factor Graphs and the Sum-product Algorithm. *IEEE Transactions on Information Theory*, 47(2):498–519, 2001.

References IV

Steffen L. Lauritzen and David J. Spiegelhalter. Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems. *Journal of the Royal Statistical Society. Series B: Methodological*, 50:157–224, 1988.

Brian Milch, Luke S. Zettelmoyer, Kristian Kersting, Michael Haimes, and Leslie Pack Kaelbling. Lifted Probabilistic Inference with Counting Formulas. In *AAAI-08 Proceedings of the 23rd AAAI Conference on Artificial Intelligence*, pages 1062–1068. AAAI Press, 2008.

David Poole. First-order Probabilistic Inference. In *IJCAI-03 Proceedings of the 18th International Joint Conference on Artificial Intelligence*, pages 985–991. IJCAI Organization, 2003.

Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Pearson, 2010.

References V

Taisuke Sato. A Statistical Learning Method for Logic Programs with Distribution Semantics. In *Proceedings of the 12th International Conference on Logic Programming*, pages 715–729. MIT Press, 1995.

Nima Taghipour, Jesse Davis, and Hendrik Blockeel. First-order Decomposition Trees. In *NIPS-13 Advances in Neural Information Processing Systems 26*, pages 1052–1060. Curran Associates, Inc., 2013a.

Nima Taghipour, Daan Fierens, Jesse Davis, and Hendrik Blockeel. Lifted Variable Elimination: Decoupling the Operators from the Constraint Language. *Journal of Artificial Intelligence Research*, 47(1):393–439, 2013b.

Guy Van den Broeck. *Lifted Inference and Learning in Statistical Relational Models*. PhD thesis, KU Leuven, 2013.