WWU
MÜNSTER

The More, the Merrier

The Power of Relations for Probabilistic Graphical Models

Tanya Braun, University of Münster

Agenda

- Statistical Relational Artificial Intelligence
  - Probabilistic relational models
  - Grounding semantics
  - Context
- Using Relations in Inference
  - Lifted query answering and tractability
  - Privacy-preserving lifted query answering
  - Who did it? Identifying most likely sources of events
  - Agent types for multi-agent decision making
- Summary

Statistical Relational Artificial Intelligence (StaRAI)

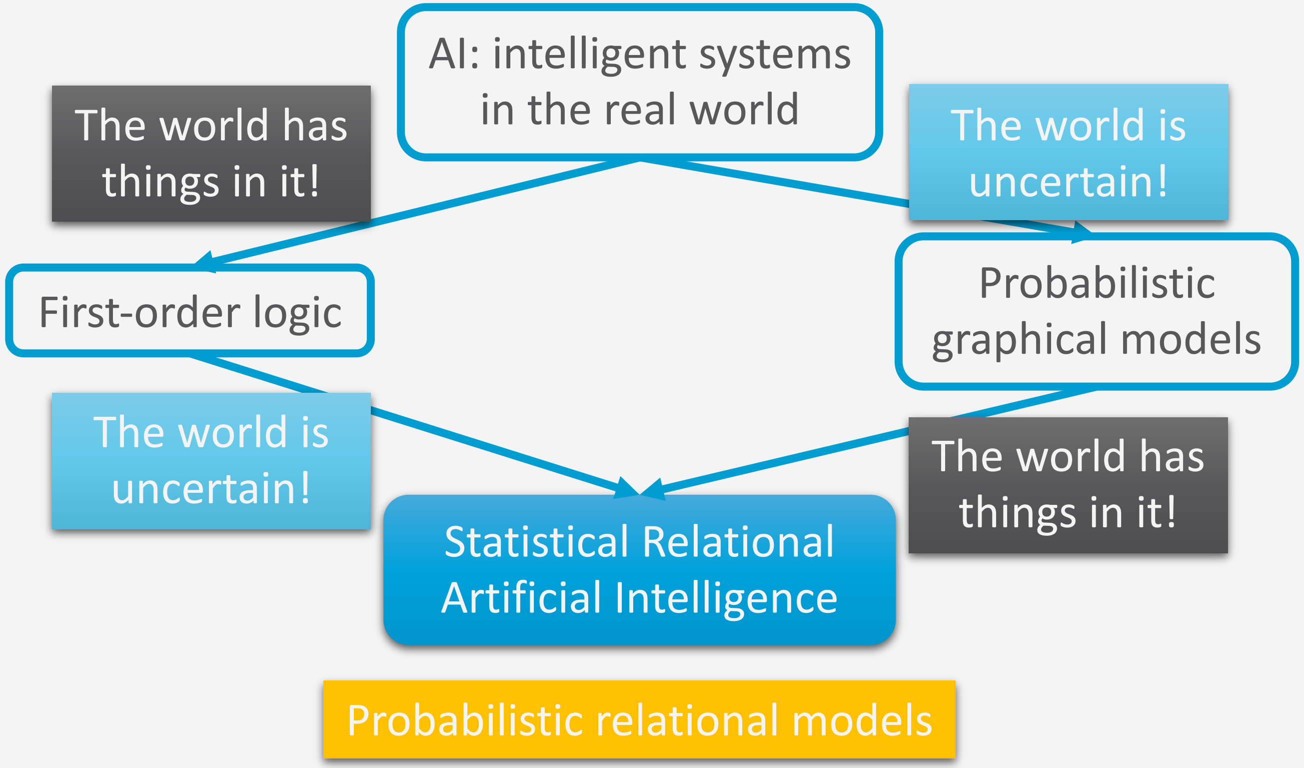
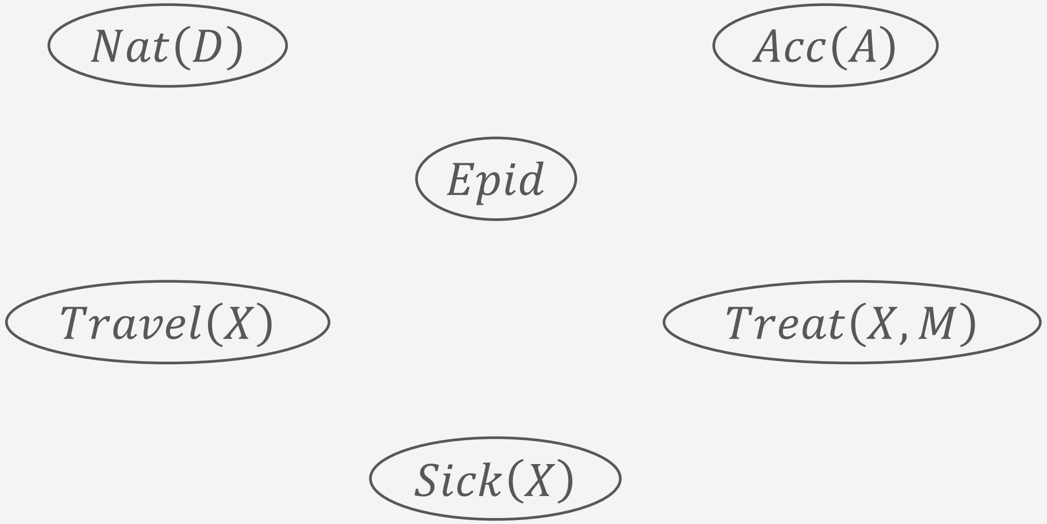


Figure based on Stuart Russell

Application: Epidemics

- Atoms: Parameterised random variables = PRVs
  - With logical variables
    - E.g., X, M
    - Possible values (domain):
      $$dom(X) = \{alice, eve, bob\}$$
      $$dom(M) = \{injection, tablet\}$$
  - With range
    - E.g., Boolean
    - $ran\big(Travel(X)\big) = \{true, false\}$
  - Represent sets of *indistinguishable* random variables

$$Nat(D) = natural\ disaster\ D$$
$$Acc(A) = accident\ A$$



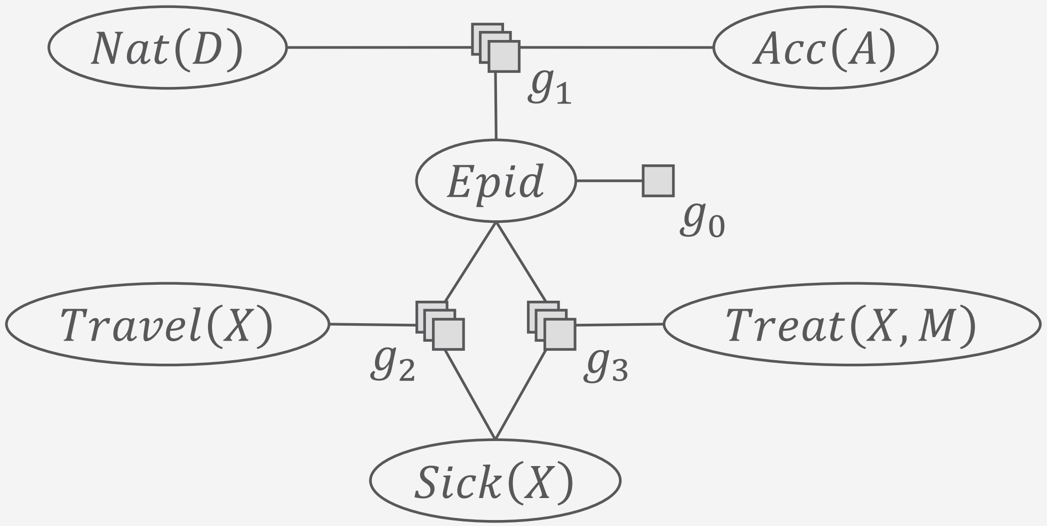
Encoding the Joint Distribution: Factorisation

- Factors with PRVs = parfactors
  - E.g., g_2

$Travel(X)$	$Epid$	$Sick(X)$	g_2
$false$	$false$	$false$	5
$false$	$false$	$true$	0
$false$	$true$	$false$	4
$false$	$true$	$true$	6
$true$	$false$	$false$	4
$true$	$false$	$true$	6
$true$	$true$	$false$	2
$true$	$true$	$true$	9

Potentials

- In parfactors, just like in factors, no probability distribution as factors required



Factors

- Grounding
  - E.g., $gr(g_2) = \{f_2^1, f_2^2, f_2^3\}$

$Travel(X)$	$Epid$	$Sick(X)$	g_2
$false$	$false$	$false$	5
$false$	$false$	$true$	0
$false$	$true$	$false$	4
$false$	$true$	$true$	6
$true$	$false$	$false$	4
$true$	$false$	$true$	6
$true$	$true$	$false$	2
$true$	$true$	$true$	9

$Travel(eve)$	$Epid$	$Sick(eve)$	g_2
$false$	$false$	$false$	5
$false$	$false$	$true$	0
$false$	$true$	$false$	4
$false$	$true$	$true$	6
$true$	$false$	$false$	4
$true$	$fals$		
$true$	tru		
$true$	tru		

$Travel(bob)$	$Epid$	$Sick(bob)$	g_2
$false$	$false$	$false$	5
$false$	$false$	$true$	0
$false$	$true$	$false$	4
		$true$	6
		$false$	4
		$true$	6
		$false$	2
		$true$	9

$Travel(alice)$	$Epid$	$Sick(alice)$	g_2
$false$	$false$	$false$	5
$false$	$false$	$true$	0
$false$	$true$	$false$	4
$false$	$true$	$true$	6
$true$	$false$	$false$	4
$true$	$false$	$true$	6
$true$	$true$	$false$	2
$true$	$true$	$true$	9

$reat(X, M)$

Encoding the Joint Distribution

- Set of parfactors = model
  - E.g., $G = \{g_1, g_2, g_3\}$
  - Semantics: Joint probability distribution P_G
    - Build by grounding, multiplying all grounded factors, and normalising the result
    - Grounding semantics [Sato 95, Fuhr 95]

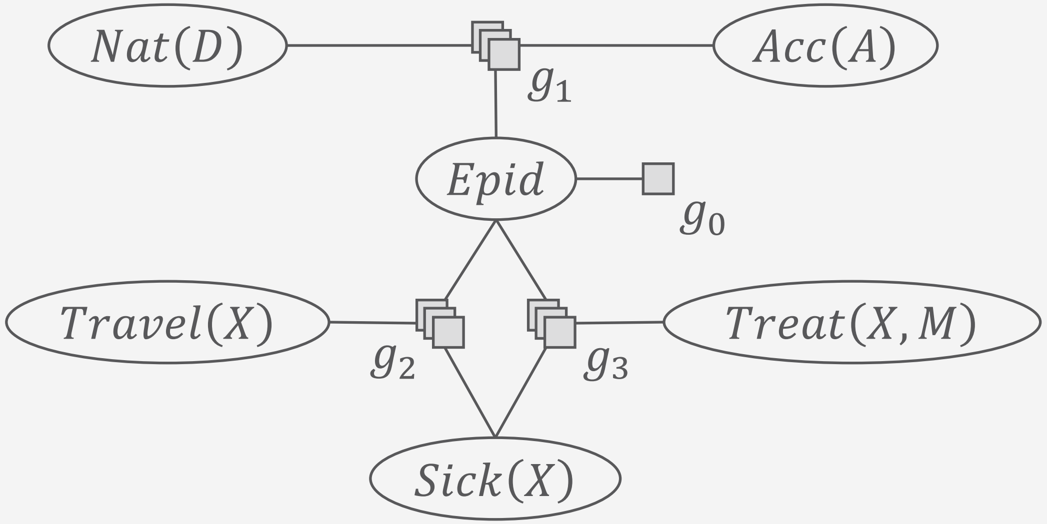
$$P_G = \frac{1}{Z} \prod_{f \in gr(G)} f$$

$$Z = \sum_{v \in r(rv(gr(G)))} \prod_{f \in gr(G)} f_i(\pi_{rv(f_i)}(v))$$

$\pi_{variables}(v)$ = projection of v onto $variables$

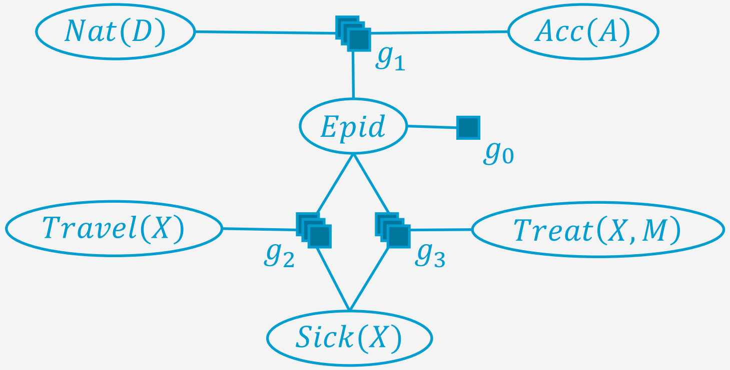
Sparse encoding of joint distribution

$3 \cdot 2^3 = 24$ entries in 3 parfactors, 6 PRVs

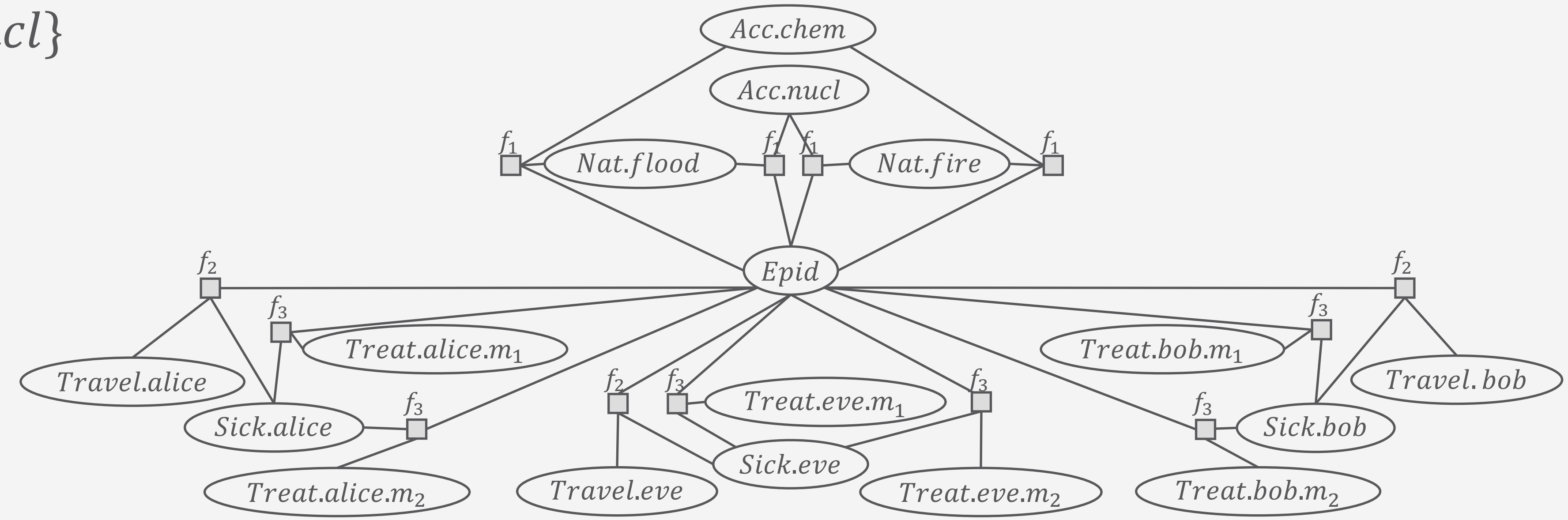


WWU MÜNSTER

Grounded Model



- Given domains
  - $dom(X) = \{alice, eve, bob\}$
  - $dom(M) = \{m_1, m_2\}$
  - $dom(D) = \{flood, fire\}$
  - $dom(A) = \{chem, nucl\}$



- Indistinguishability in
  - Graph structure
  - Factors

Probabilistic Relational Models and Variants

- Parfactors Models
  [Poole 2003, Taghipour et al. 2013, B 2020*, Gehrke 2021*]
- Markov Logic Networks (MLNs) [Richardson & Domingos 2006]
  - Use logical formulas to specify potential functions
- Probabilistic Soft Logic (PSL) [Bach et al. 2017]
  - Use density functions to specify potential functions
- Based on grounding semantics [Sato 1995, Fuhr 1995]

* and the papers cited therein (PhD thesis)

Agenda

- Statistical Relational Artificial Intelligence
  - Probabilistic relational models
  - Grounding semantics
  - Context
- Using Relations in Inference
  - Lifted query answering and tractability
  - Who did it? Identifying most likely sources of events
  - Privacy-preserving lifted query answering
  - Agent types for multi-agent decision making
- Summary

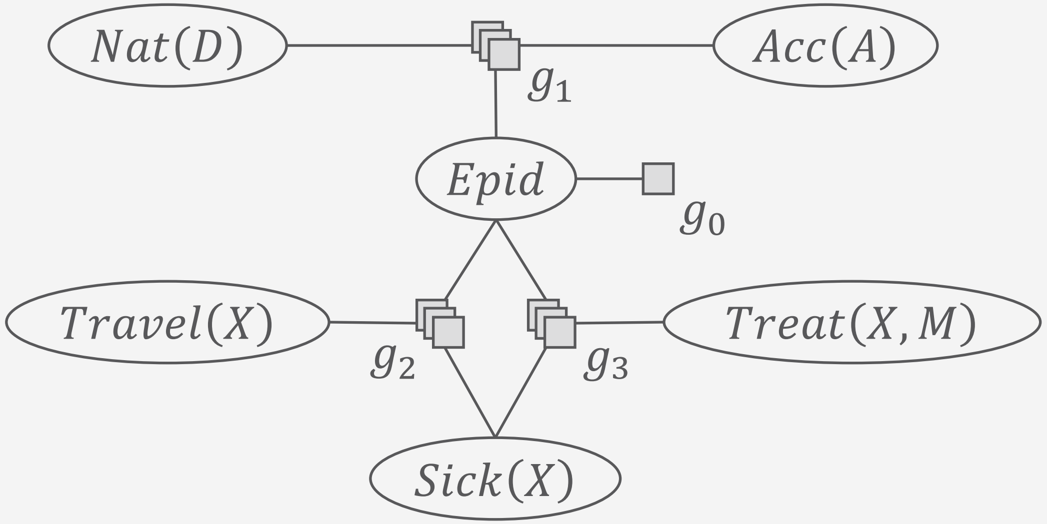
Lifted Query Answering and Tractability

Using Relations in Inference

Reasoning on Probabilistic Relational Models

- Inference task: query answering (QA)
- Queries:
  - Marginal distribution
    - $P(Sick(eve))$
    - $P(Travel(eve,)\ Treat(eve, m_1))$
  - Conditional distribution
    - $P(Sick(eve)|Epid)$
    - $P(Epid|Sick(eve) = true)$
  - Assignment queries: $\arg\max_{\boldsymbol{a} \in ran(\boldsymbol{A})} P(\boldsymbol{a}|\boldsymbol{e})$
    - MPE: $\boldsymbol{A} = rv(\boldsymbol{G}) \setminus rv(\boldsymbol{e})$
    - MAP: $\boldsymbol{A} \subseteq rv(\boldsymbol{G}) \setminus rv(\boldsymbol{e})$
      - What is not in $\boldsymbol{A}$ needs to be summed out

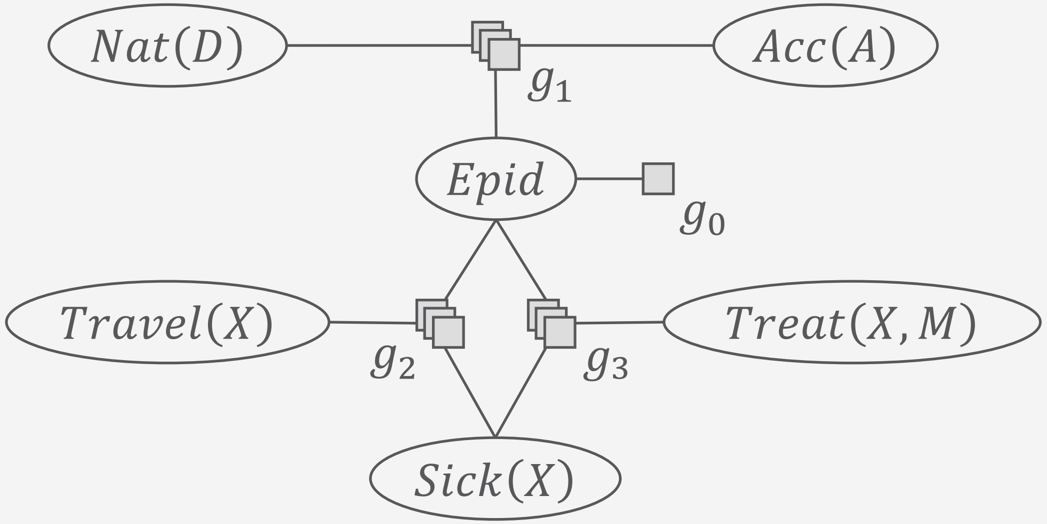
Goal: Avoid groundings!
→ *lifted* inference



QA: Lifted Variable Elimination (LVE)

- Eliminate all variables not appearing in query
- Lifted summing out
  - Sum out *representative* instance as in propositional variable elimination
  - Exponentiate result for indistinguishable instances

- Correctness: Equivalent ground operation
  - Each instance is summed out
  - Result: factor f that is identical for all instance
  - Multiplying indistinguishable results
    → exponentiation of one representative f

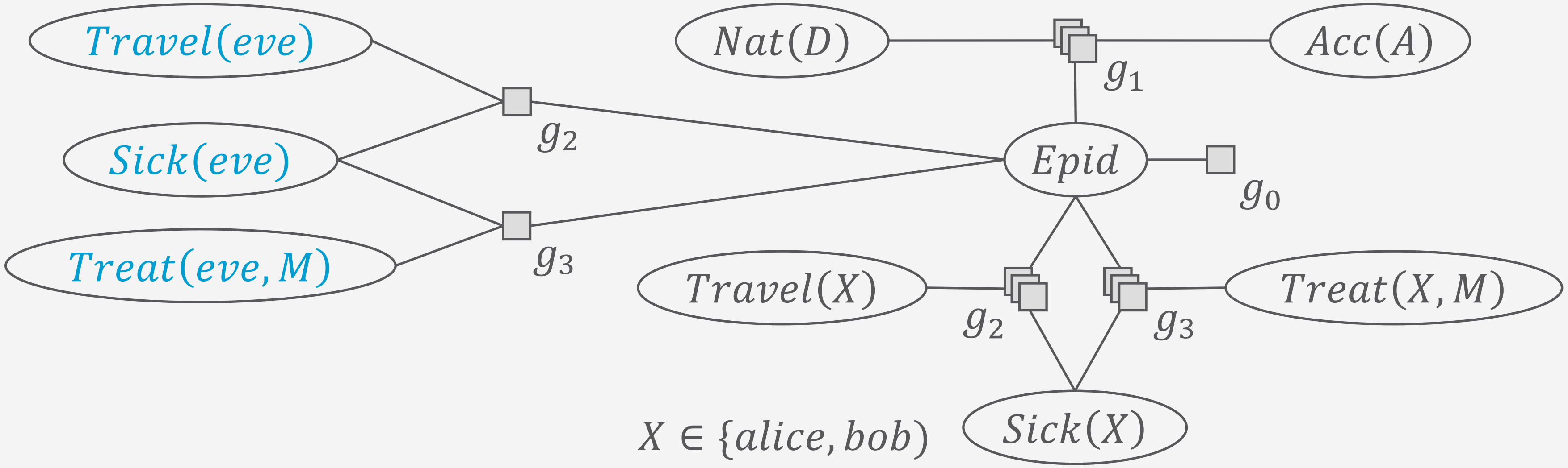


[Poole 2003, de Salvo Braz et al. 2005, 2006, Milch et al. 2008, Taghipour et al. 2013, 2013a, B & Möller 2018]

QA: LVE in Detail

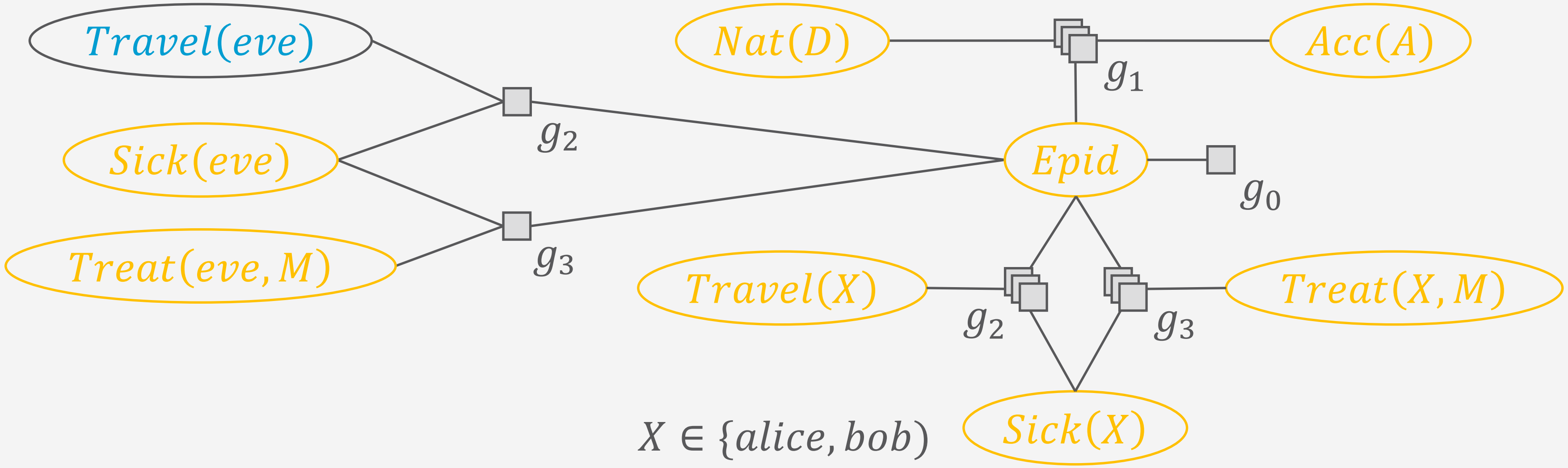
- E.g., marginal
  - $P(Travel(eve))$
  - Split atoms $R(\dots, X, \dots)$ w.r.t. eve if eve in $dom(X)$

Shattering



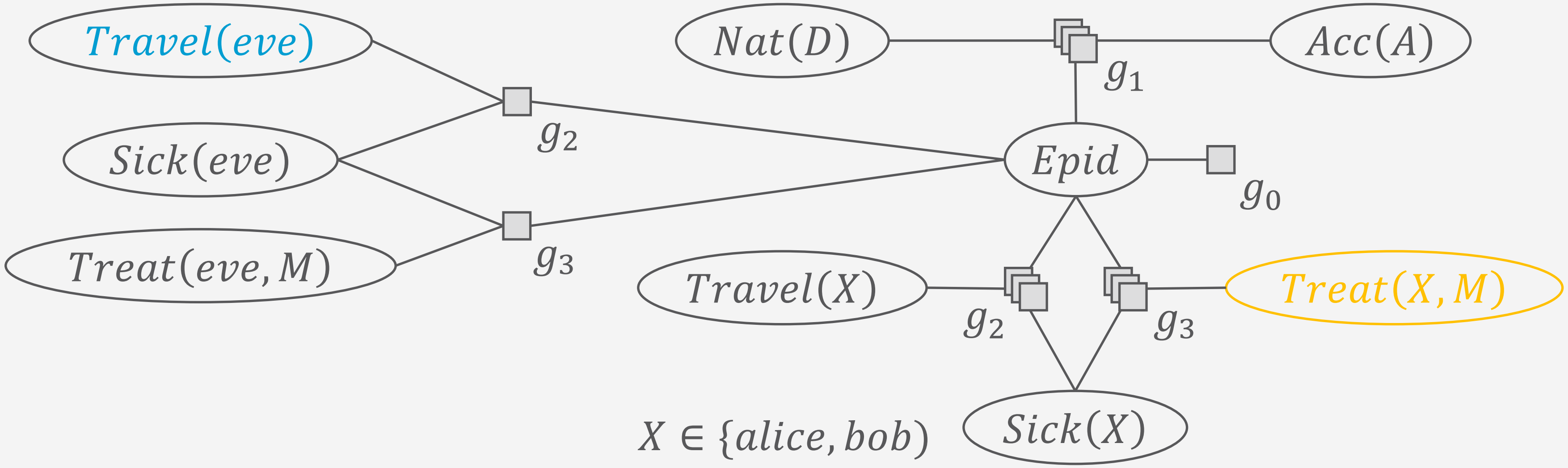
QA: LVE in Detail

- E.g., marginal
  - $P(Travel(eve))$
  - Split atoms $R(\dots, X, \dots)$ w.r.t. eve if eve in $dom(X)$
  - Eliminate all non-query variables



QA: LVE in Detail

- Eliminate $Treat(X, M)$
  - Appears in only one g: g_3
  - Contains all logical variables of g_3: X, M
  - For each X constant: the same number of M constants
  - ✓ Preconditions of lifted summing out fulfilled, lifted summing out possible

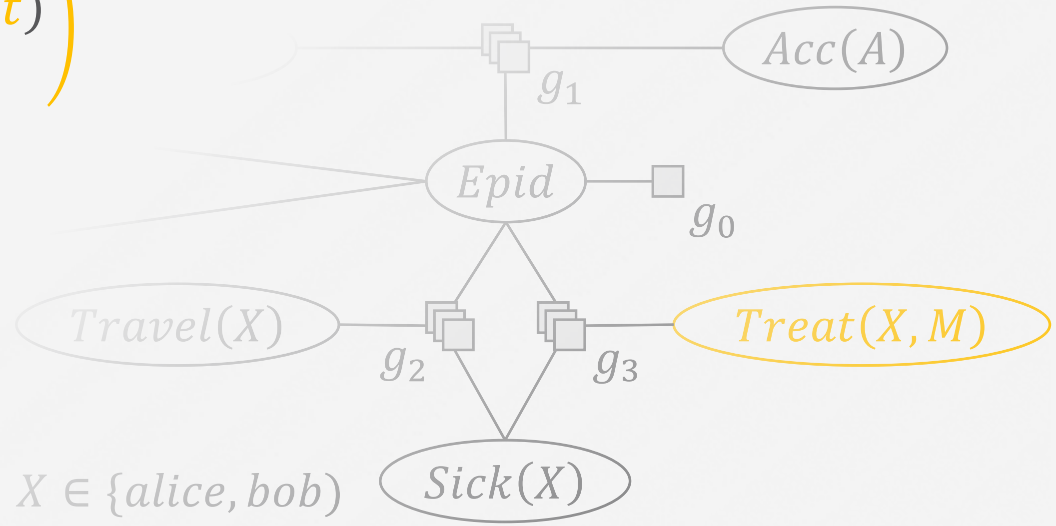


LVE in Detail: Lifted Summing Out

- Eliminate $Treat(X, M)$ by lifted summing out
  1. Sum out representative
  2. Exponentiate for indistinguishable objects

Only here, domain size comes into play
→ no change in graph / parfactor if domain size changes

$$\left(\sum_{t \in r(Treat(X,M))} g_3(Epid = e, Sick(X) = s, Treat(X, M) = t)\right)^{\#M|X}$$



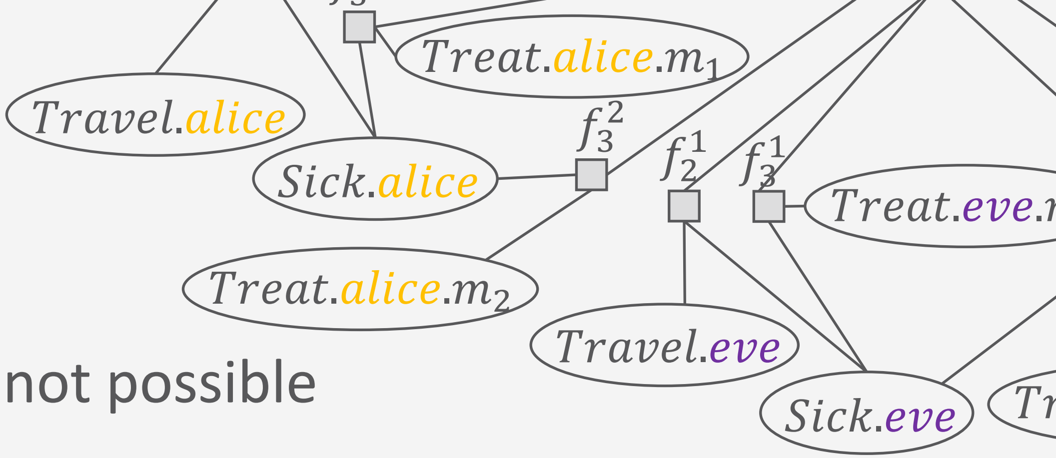
Tractability

- Given a model that allows for lifted calculations
  - I.e., no groundings during solving an instance of the problem
- Solving an instance of the problem is possible in time polynomial in domain sizes

  → The query answering algorithm is domain-lifted
- An query answering problem is tractable
  - when it is solved by an efficient algorithm, running in time polynomial in the number of random variables
- Assume that the number of random variables is characterised by domain sizes
  - Then, solving a query answering problem is tractable under domain-liftability
    - Runtime might still be exponential in other terms
    - More general results by Niepert & Van den Broeck (2014)

Privacy-preserving Lifted Inference

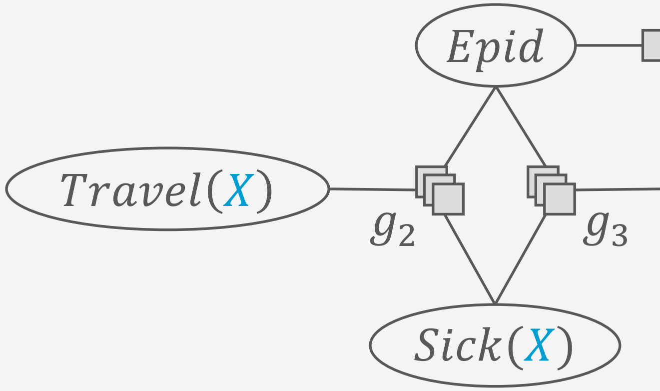
Using Relations in Inference



Privacy-preserving Lifted Query Answering

- Privacy preservation: Identifying individual instances not possible

- Propositional: All instances have explicit representations
  - Privacy preservation not possible!
- Relational: Instances as part of groups of indistinguishable instances
  - Privacy preservation *possible*?



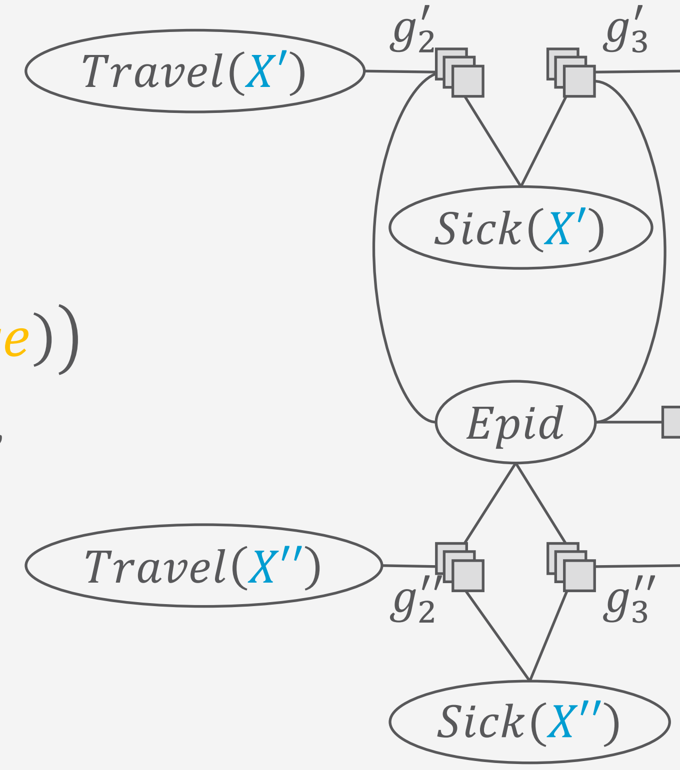
- Idea: Hide individuals in the groups of indistinguishable instances
  - Problem from a privacy perspective:

Groups of indistinguishable instances only hide instances that do not need protection

Preview of awesome new work with
Marcel Gehrke, IFIS
Esfandiar Mohammadi, ITS

Necessary Changes for Privacy Preservation

- Query terms
  - Concrete constants in query terms no longer allowed, e.g., $P\big(Sick(alice)\big)$
  - Instead: allow for representative constants per group, e.g., $P\big(Sick(x)\big)$, x representative for group X
- Evidence
  - Evidence over concrete constants no longer allowed as input, e.g., $sick(alice), sick(eve)$, …
  - Instead: Cluster evidence using a *differentially private* clustering algorithm
    - Uncertain evidence: $\phi\big(Sick(X')\big)$ with X' referring to a privacy-preserving cluster of evidence
- Assumes sufficiently large group sizes to begin with
- Assumes that the model is privacy-preserving to begin with



Preview of awesome new work with
Marcel Gehrke, IFIS
Esfandiar Mohammadi, ITS

Who did it? – Identifying the Most Likely Sources of Events

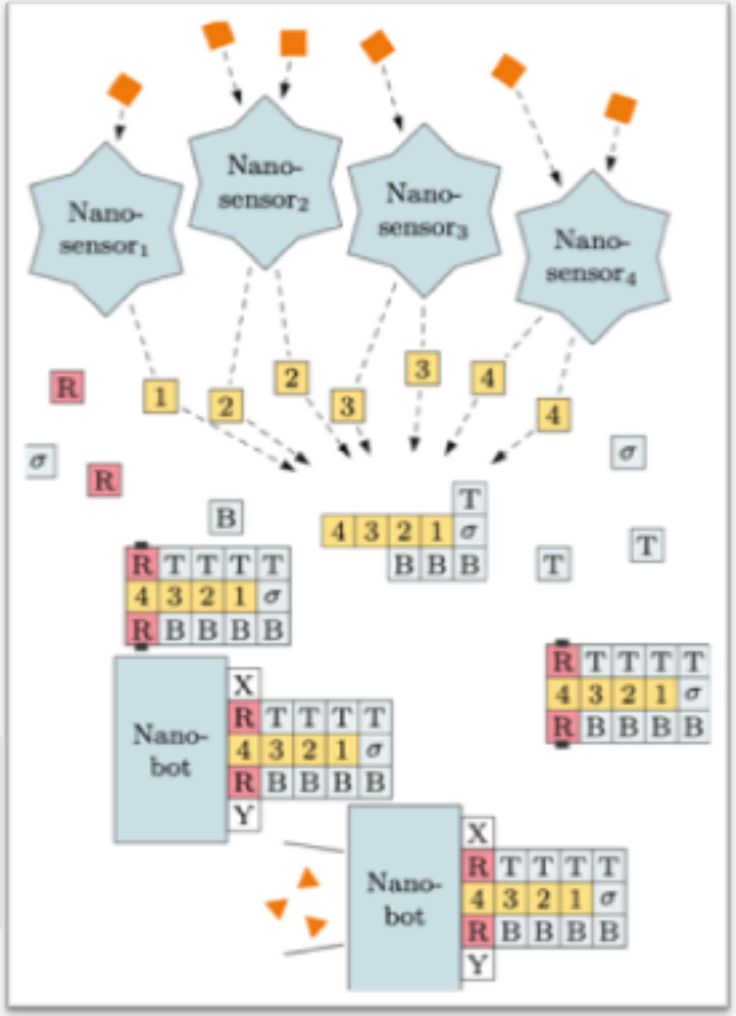
Using Relations in Inference

Awesome new work with
Marcel Gehrke & Ralf Möller, IFIS
Florian Lau, ITM

New Application Scenario: Nanoscale Medical System

- Nanoscale medical system
  - Set of nanoagents (very, very, very small entities)
    - Number of agents can be in the order of up to 10^9
  - Can receive *something*
    - E.g., receptive to certain markers in a patient's blood stream
  - Can release *something*
    - E.g., some form of medicine
- Interesting new problems
  - No longer necessarily attributable how many agents did what exactly
  - No longer necessarily known how many (functioning) agents there are

Both valuable information in medical systems to not poison a patient!



[B et al. 2021+22, Lau et al. 2022]

Awesome new work with
Marcel Gehrke & Ralf Möller, IFIS

Who did it? – Identifying the Most Likely Sources of Events

- Evidence, i.e., set of observations, without a known source: $R(X) = r$
  - $ObsMarker(X') = true, |dom(X')| = 1000$
  - NB: Not possible in propositional setting as an observation can only belong to a specific random variable without additional information about relations or types
- Optimisation problem for a single logical variable in the evidence:
  Given evidence $\boldsymbol{e}$ with known source, find a domain C for X such that the probability of the evidence without a source is maximal under the domain, written as

$$\arg\max_{C} P(R(X) = r|\boldsymbol{e})_{|C}$$

  - Use C as source for evidence
  - Example from above: If no further evidence and only one group of indistinguishable instances represented by X in model, then any 1000 instances represented by X will do

[Gehrke et al., 2022]

Awesome new work with
Marcel Gehrke & Ralf Möller, IFIS

Who did it? – Identifying the Most Likely Sources of Events

- Optimisation problem:

$$\arg\max_{C} P(R(X) = r|\boldsymbol{e})_{|C}$$

- However, it gets complicated once you have more sets of unknown sources or sets of known sources to consider as well
  - $ObsMarker(X) = true, |dom(X)| = 1000$
  - $ObsMarker(Y) = true, |dom(Y)| = 500$
  - $ObsMarker(Z) = true, dom(X) = \{x_1, \dots, x_{100}\}$
  - Various domain assignments possible from full overlap to complete disjoint sets
    - $dom(Z) = dom(Y') = dom(X') = \{x_1, \dots, x_{100}\}$, $dom(Y'') = dom(X'') = \{x_{101}, \dots, x_{500}\}$, $dom(X''') = \{x_{501}, \dots, x_{1000}\}$
    - …
    - $dom(Z) = \{x_1, \dots, x_{100}\}$ $dom(Y) = \{x_{101}, \dots, x_{600}\}$ $dom(X) = \{x_{601}, \dots, x_{1600}\}$

[Gehrke et al., 2022]

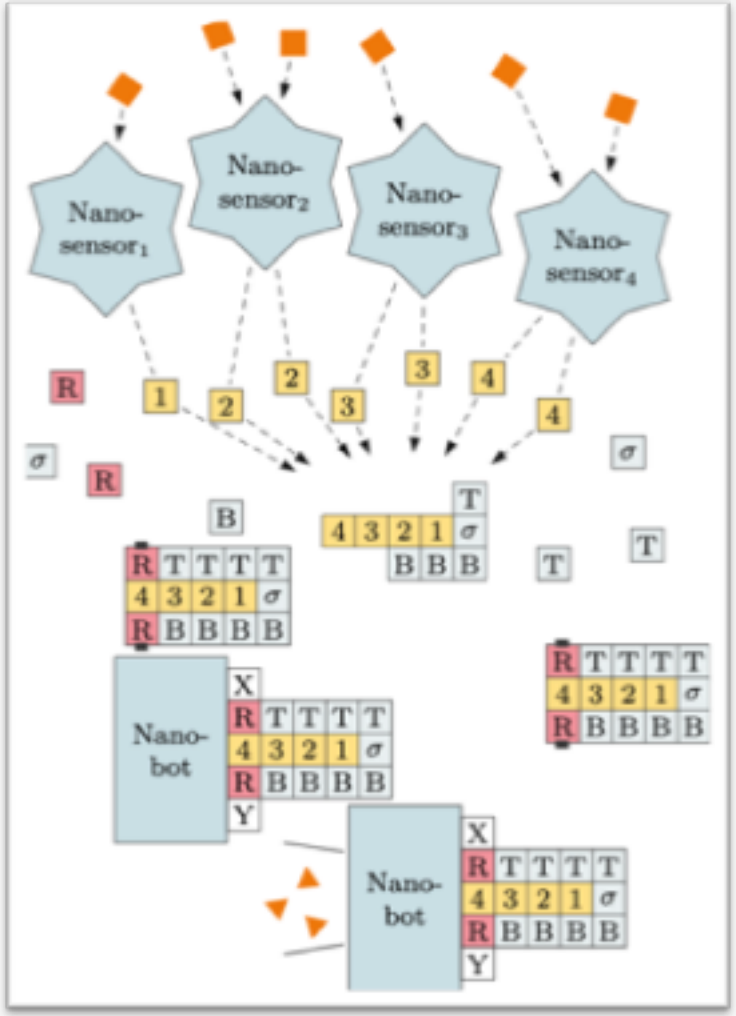
Agent Types for Multi-agent Decision Making

Using Relations in Inference

Awesome new work with
Marcel Gehrke & Ralf Möller, IFIS
Florian Lau, ITM

Multi-agent Decision making

- Formal model of a nanoscale system:
  Decentralised partially observable Markov decision process (decPOMDP) [see Oliehoek & Amato, 2016]
  - Set of agents with
    - Own set of available actions, observations
    - *Shared* state transition function, reward refunction, sensor model
  - Partial observability: full state not known to each agent
  - Markov: transition function does not change over time
- Complexity: exponential dependence on number of agents
  - Huge problem with agent numbers as large as 10^9

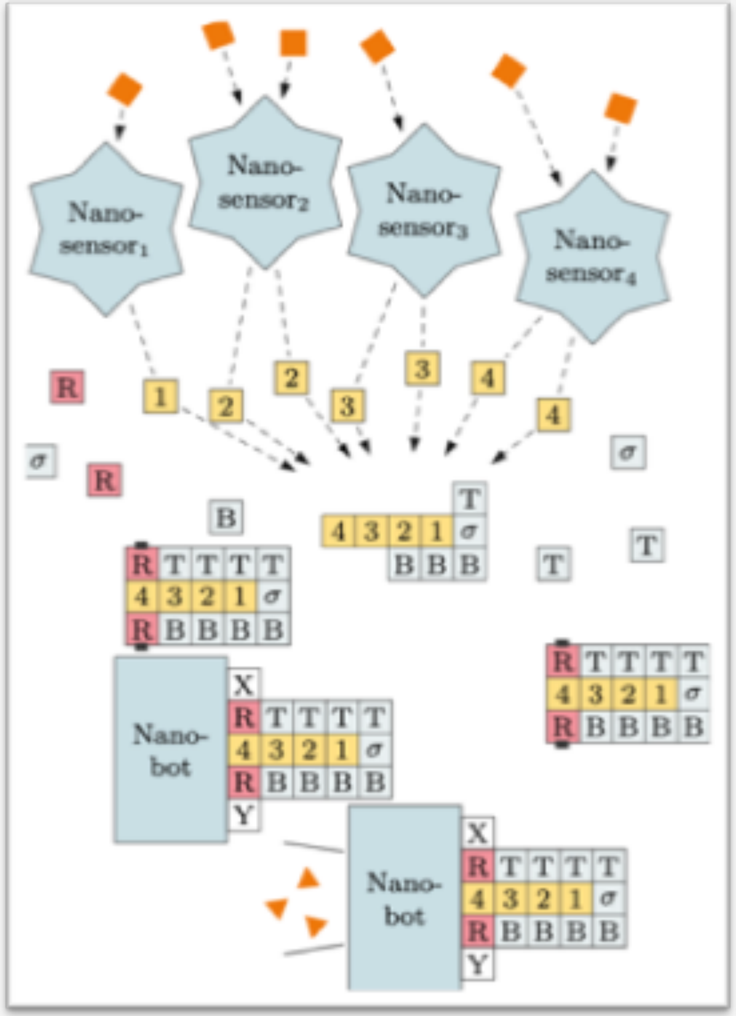


WWU MÜNSTER

Awesome new work with
Marcel Gehrke & Ralf Möller, IFIS
Florian Lau, ITM

Agent Types: Lifting for agents

- Agents with indistinguishable behaviour → partitions in the agent set, consider as agent types
  - Same sets of actions, observations available
- *If* certain independences in the reward function, state transition function, and sensor model hold:
  Same strategy / program applies to agents of same type
  - If independence among agents in partitions holds, then partitions can be treated by representatives
    - *Reduces exponential dependence on agent numbers to logarithmic dependence!*
    - *Enables optimisation problems asking for the necessary number of agents to reach a certain expected reward in a given number of time steps (horizon)*



[B et al., 2022]

Agenda

- Statistical Relational Artificial Intelligence
  - Probabilistic relational models
  - Grounding semantics
  - Context
- Using Relations in Inference
  - Lifted query answering and tractability
  - Who did it? Identifying most likely sources of events
  - Privacy-preserving lifted query answering
  - Agent types for multi-agent decision making
- Summary

The Finish Line: The Power of Relations

- Lifted query answering and tractability
  - Use information about indistinguishability to speed up inference
  - Tractability in terms of domain sizes through lifting
- Privacy-preserving query answering
  - Relational setting enables privacy preservation
  - Changes to query language and evidence necessary to actually have privacy preservation
- Unknown sources of evidence
  - Can be attributed to sources when solving a corresponding optimisation problem
- Agent types: same actions + observations
  - Treat through representatives if independences hold

What else is there to do? – Oh, so much…
- Approximating symmetries
- Generalising lifting operators
- More robust learning algorithms
- Privacy
- Ethical behaviour
- Explainability
- …

Thank you!

Bibliography

Ordered alphabetically

Bibliography

- Bach et al. (2017)
  Stephen H. Bach, Matthias Broecheler, Bert Huang, and Lise Getoor: Hinge-Loss Markov Random Fields and Probabilistic Soft Logic. In *Machine Learning Journal*, 2017.
- B & Möller (2018)
  Tanya Braun and Ralf Möller: Parameterised Queries and Lifted Query Answering. In *IJCAI-18 Proceedings of the 27th International Joint Conference on Artificial Intelligence*, 2018.
- B (2020)
  Tanya Braun: Rescued from a Sea of Queries: Exact Inference in Probabilistic Relational Models. *PhD thesis*, 2020.
- B et al. (2021)
  Tanya Braun, Marcel Gehrke, Florian Lau, and Ralf Möller: Lifting DecPOMDPs for Nanoscale Medical Systems. In *StaRAI-21 International Workshop on Statistical Relational Artificial Intelligence*, 2021.
- B et al. (2022)
  Tanya Braun, Marcel Gehrke, Florian Lau, and Ralf Möller: Lifting in Multi-agent Systems under Uncertainty. In *UAI-22 International Conference on Uncertainty in Artificial Intelligence*, 2022.

Bibliography

- De Salvo Braz et al. (2005)
  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*, 2005.
- De Salvo Braz et al. (2006)
  Rodrigo de Salvo Braz, Eyal Amir, and Dan Roth: MPE and Partial Inversion in Lifted Probabilistic Variable Elimination. In *AAAI-06 Proceedings of the 21st Conference on Artificial Intelligence*, 2006.
- Fuhr (1995)
  Norbert Fuhr: Probabilistic Datalog - A Logic for Powerful Retrieval Methods. In *SIGIR-95 Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1995.
- Gehrke (2021)
  Marcel Gehrke: Taming Exact Inference in Temporal Probabilistic Relational Models. *PhD thesis*, 2021.
- Gehrke et al. (2022)
  Marcel Gehrke, Ralf Möller, and Tanya Braun: Who did it? Identifying the Most Likely Origins of Events. In *PGM-22 Proceedings of the 11th International Conference on Probabilistic Graphical Models*, 2022.

Bibliography

- Gehrke et al. (2022) (under review)
  Marcel Gehrke, Esfandiar Mohammadi, and Tanya Braun: How Can Lifting Preserve Privacy in Probabilistic Graphical Models?. *Under review*, 2022.
- Lau et al. (2022)
  Florian Lau, Tanya Braun, Ralf Möller, and Stefan Fischer: An Extended Framework to Holistically Model Nanostructures and Nanonetworks as decPOMDPcoms. In ACM-*NanoCom-22 Proceedings of the 9th ACM International Conference on Nanoscale Computing and Communication*, 2022.
- Milch et al. (2008)
  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*, 2008.
- Niepert & Van den Broeck (2014)
  Mathias Niepert and Guy Van den Broeck. Tractability through Exchangeability: A New Perspective on Efficient Probabilistic Inference. In *AAAI-14 Proceedings of the 28th AAAI Conference on Artificial Intelligence*, 2014.
- Oliehoek & Amato (2016)
  Frans A. Oliehoek and Christopher Amato: A Concise Introduction to Decentralised POMDPs, 2019.

Bibliography

- Poole (2003)
  David Poole: First-order Probabilistic Inference. In *IJCAI-03 Proceedings of the 18th International Joint Conference on Artificial Intelligence*, 2003.
- Richardson & Domingos (2006)
  Matthew Richardson and Pedro Domingos. Markov Logic Networks. In *Machine Learning Journal*, 2006.
- Sato (1995)
  Taisuke Sato: A Statistical Learning Method for Logic Programs with Distribution Semantics. In *Proceedings of the 12th International Conference on Logic Programming*, 1995.
- Taghipour et al. (2013)
  Nima Taghipour, Daan Fierens, Guy Van den Broeck, Jesse Davis, and Hendrik Blockeel. Completeness Results for Lifted Variable Elimination. In *AISTATS-13 Proceedings of the 16th International Conference on Artificial Intelligence and Statistics*, 2013.
- Taghipour et al. (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, 2013.