Universität Münster

Oh, I forgot!

How human forgetting inspires AI research on uncertainty

Tanya Braun, Marcel Gehrke, Kai Sauerwald, Jan Speller

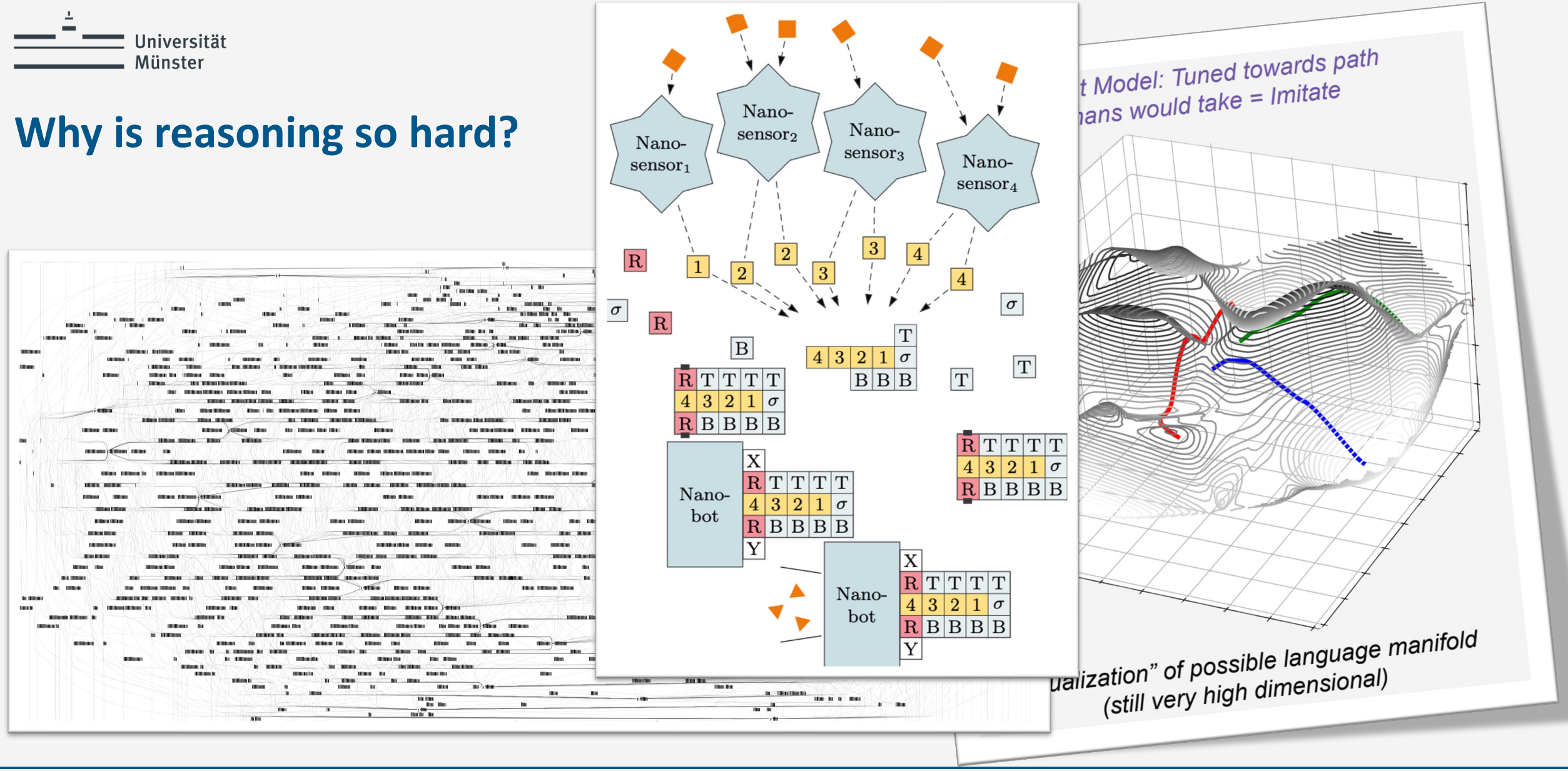


living.knowledge

Data Science Group
Computer Science Department

Universität Münster

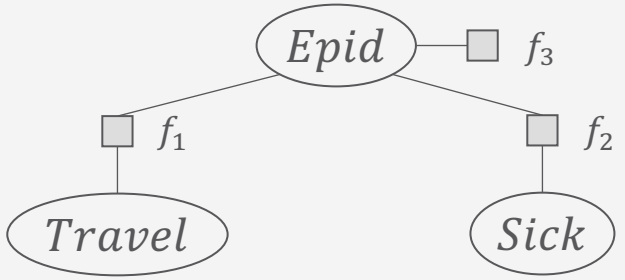
Why is reasoning so hard?



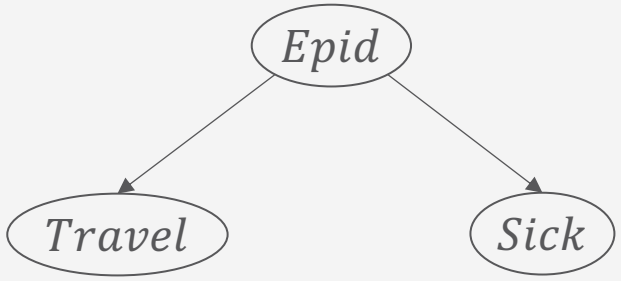
Right figure taken from a talk by Malte Schilling
https://www.dropbox.com/s/nsenp948uc93l5w/schilling_2023_06_LLM_Mechanisms.pdf?dl=0

Reasoning under uncertainty

- Probabilistic graphical model
  - Set of factors (local distributions) over random variables
    $$F = \{f_1, \dots, f_n\}$$
  - Represents a full joint probability distribution
    $$P_F = \prod_{f \in F} \phi_f(A_f)$$
- Reasoning
  - Query for a probability (distribution)
    $$P(R|\boldsymbol{e})$$



Factor graph (FG)



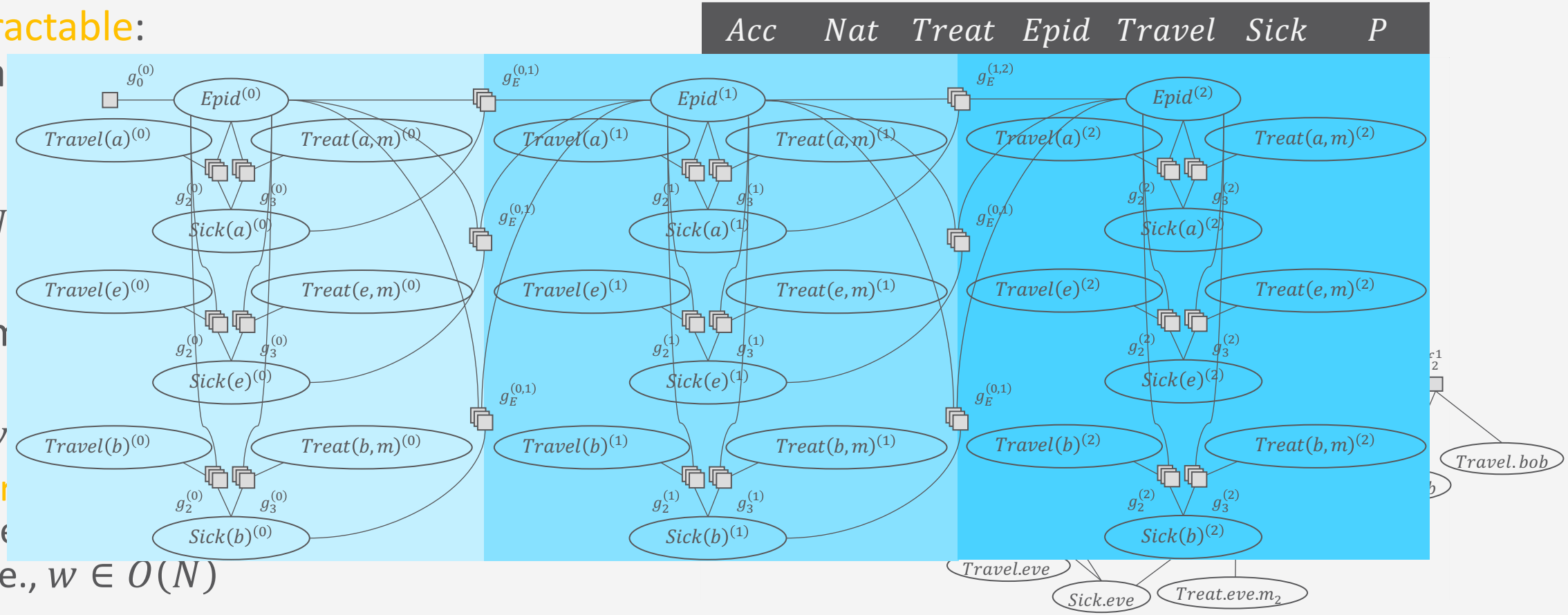
Bayesian network (BN)

Epid	*Travel*	*Sick*	*P*
false	false	false	0.20
false	false	true	0.24
false	true	false	0.28
false	true	true	0.08
true	false	false	0.05
true	false	true	0.06
true	true	false	0.07
true	true	true	0.02

Full joint probability distribution

Reasoning under uncertainty – The intractability problem

- Intractable:
  Rea
  the
  - N
  - r
- Com
  - w
  - Pr
    Re
    i.e., $w \in O(N)$



The world forgetting,
by the world forgot.
Eternal sunshine of the spotless mind!

Alexander Pope, Eloisa to Abelard

Universität Münster

Commonsense Forgetting

- Contraction
- Ignorance
- Revision / Update
- Abstraction
- Marginalisation
- Focussing
- Tunnelview
- Conditionalisation
- Fading out



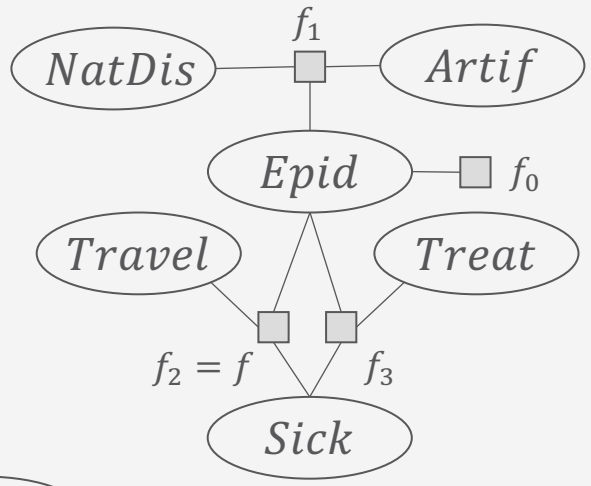
List based on Beierle et al.: Towards a General Framework for Kinds of Forgetting in Common-Sense Belief Management. In *Künstliche Intelligenz*, 2019.

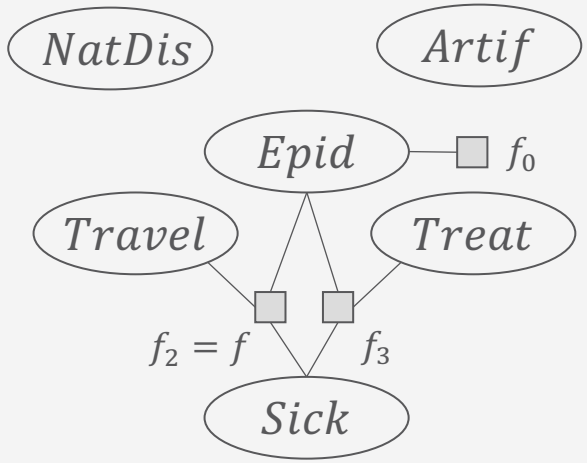
Contraction

- Direct intention to lose information
  - E.g., due to changes in privacy requirements

→ Change in syntax and semantics necessary
  - However, keep remainder of model for consistency purposes

→ Idea: Delete factors
  - Keeps remaining correlations as they are
  - Upside: Can make a model more sparse
  - Problem: How do we delete nodes, i.e., an input to a function?





Universität Münster

Commonsense Forgetting

- Contraction
- Ignorance
- Revision / Update
- Abstraction
- Marginalisation
- Focussing
- Tunnelview
- Conditionalisation
- Fading out



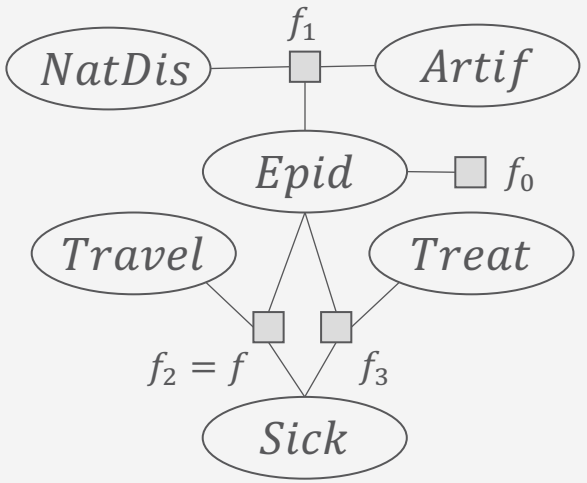
List based on Beierle et al.: Towards a General Framework for Kinds of Forgetting in Common-Sense Belief Management. In *Künstliche Intelligenz*, 2019.

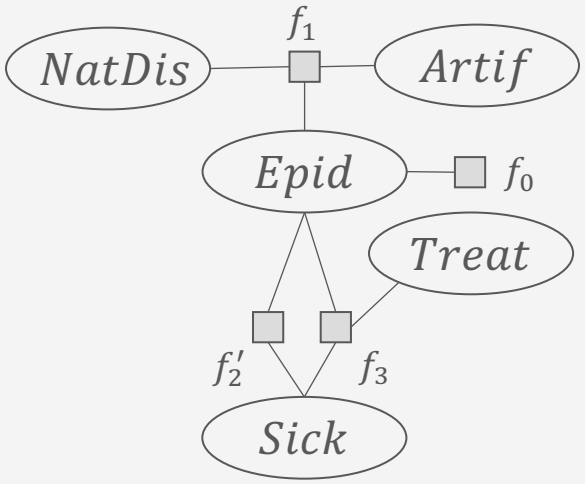
Marginalisation

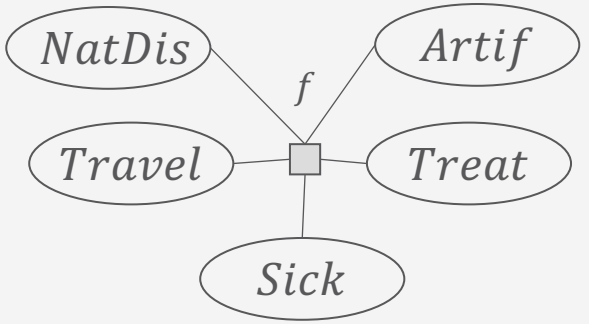
- Blinding out of information: Removal of certain aspects represented in a model
  - E.g., making decision by only taking certain aspects into account

→ Change in syntax necessary, but not semantics!

→ Idea: Sum-out random variables
  - Keeps semantics as is but removes variable from model
  - Downside: Can make a model much more interconnected







Universität Münster

Commonsense Forgetting

- Contraction
- Ignorance
- Revision / Update
- Abstraction
- Marginalisation
- Focussing
- Tunnelview
- Conditionalisation
- Fading out



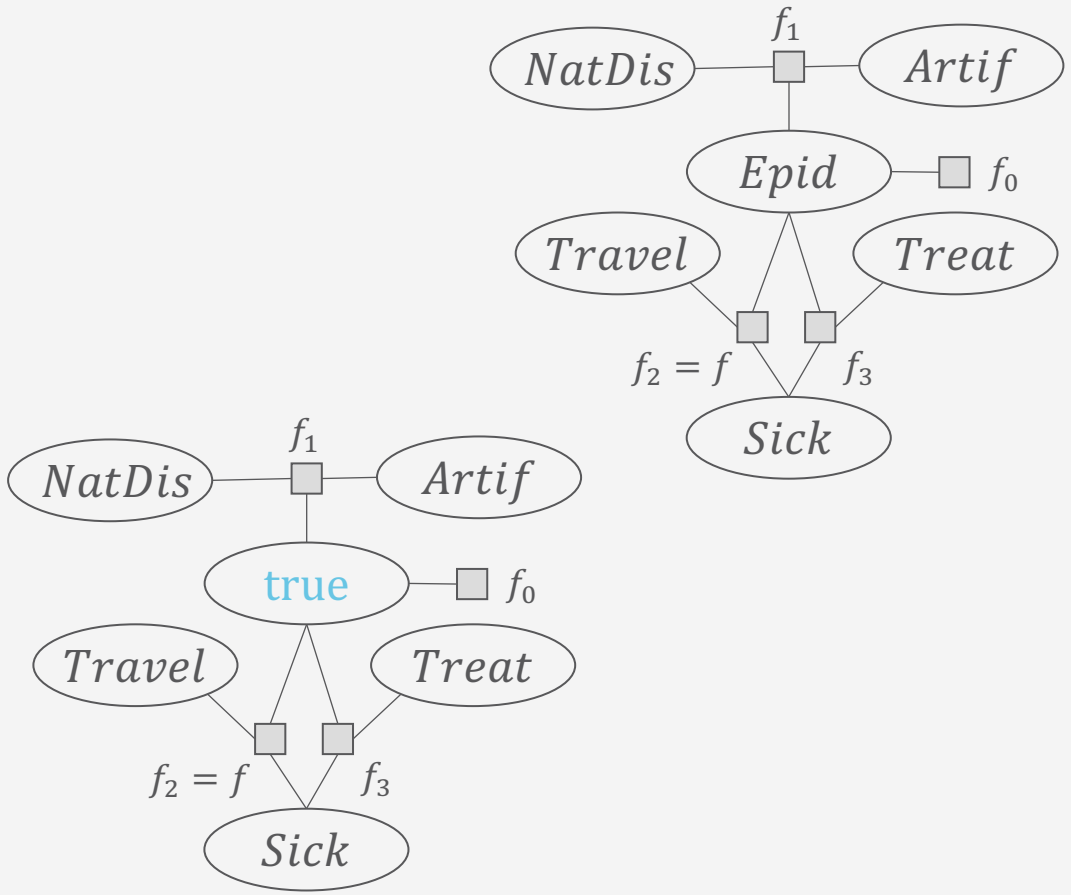
List based on Beierle et al.: Towards a General Framework for Kinds of Forgetting in Common-Sense Belief Management. In *Künstliche Intelligenz*, 2019.

Conditionalisation

- Restrict a belief to specific case or context
  - Interpret something under an assumption

→ Not necessarily a change in syntax nor semantics

- Idea: Set evidence
  - Evidence = set of observations
    → provide a context in which a query is asked
  - Also known as *conditioning* in probabilistic reasoning
    - Removes dimensions from reasoning process



Universität Münster

Commonsense Forgetting

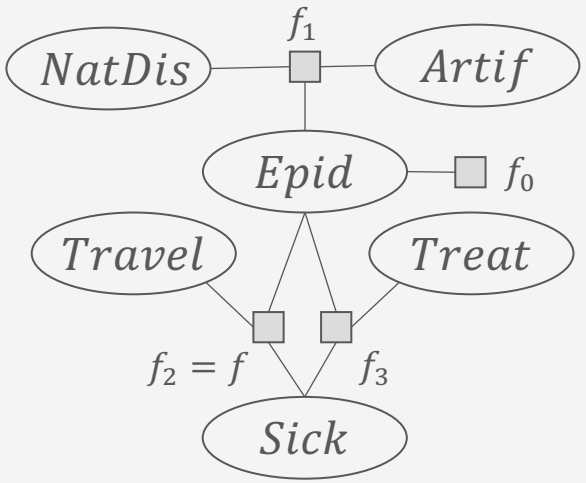
- Contraction
- Ignorance
- Revision / Update
- Abstraction
- Marginalisation
- Focussing
- Tunnelview
- Conditionalisation
- Fading out



List based on Beierle et al.: Towards a General Framework for Kinds of Forgetting in Common-Sense Belief Management. In *Künstliche Intelligenz*, 2019.

Ignorance

- Become deliberately unsure about your beliefs
  - E.g., one is unsure about the status of a person because the person is both a student and staff member
- Things you are certain about in a model
  - Evidence: the $\boldsymbol{e}$ in a query $P(R|\boldsymbol{e})$
  - The model itself, the connections and numbers



→Possible interpretations:

- Removal of evidence: $P(R|\boldsymbol{e}')$ or $P(R)$
- Add uncertainty to numbers: Credal nets
  - Interval around numbers in factors

A	B	ϕ	ϕ'
true	true	φ_1	$\varphi_1 \pm \varepsilon$
true	false	φ_2	$\varphi_2 \pm \varepsilon$
false	true	φ_3	$\varphi_3 \pm \varepsilon$
false	false	φ_4	$\varphi_4 \pm \varepsilon$

Universität Münster

Commonsense Forgetting

- Contraction
- Ignorance
- Revision / Update
- Abstraction
- Marginalisation
- Focussing
- Tunnelview
- Conditionalisation
- Fading out

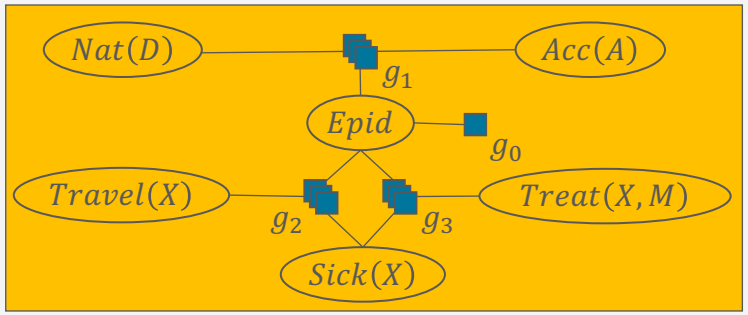


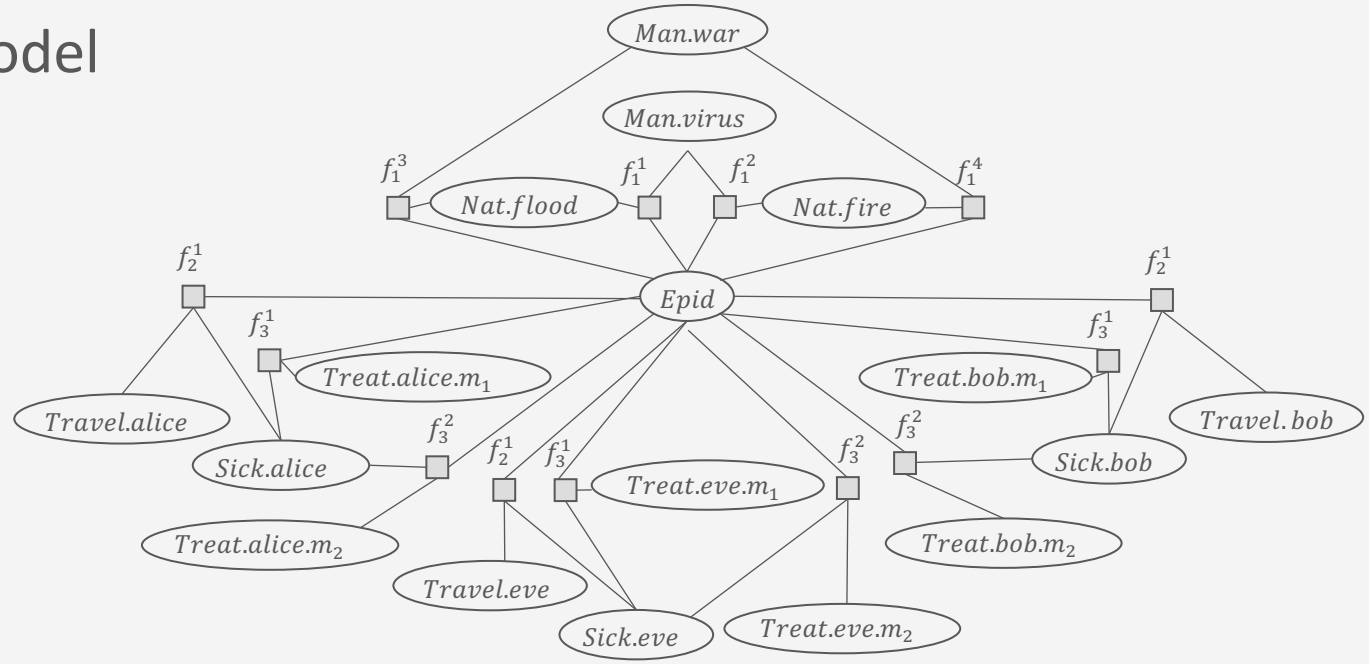
List based on Beierle et al.: Towards a General Framework for Kinds of Forgetting in Common-Sense Belief Management. In *Künstliche Intelligenz*, 2019.

Abstraction

- Build more general rules / knowledge bases from specific instances
  - E.g., one door with a handle is enough to learn how doors with handles work

- Find (approximate) regularities in a model
  - Among graph structure and functions
  - Idea: Use (approximate) colour passing to identify those regularities





Luttermann et al.: Approximate Lifted Model Construction. In *IJCAI-25*, 2025.

Commonsense Forgetting

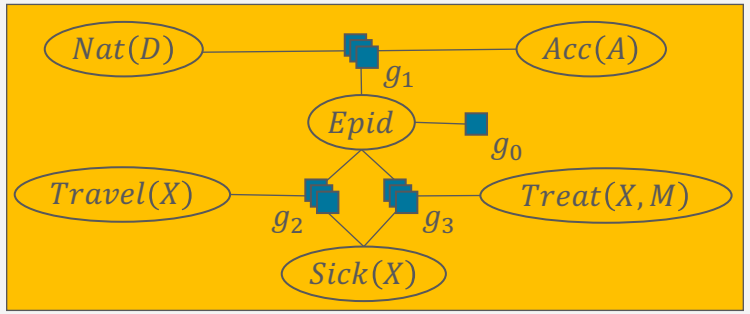
- Contraction
- Ignorance
- Revision / Update
- Abstraction
- Marginalisation
- Focussing
- Tunnelview
- Conditionalisation
- Fading out



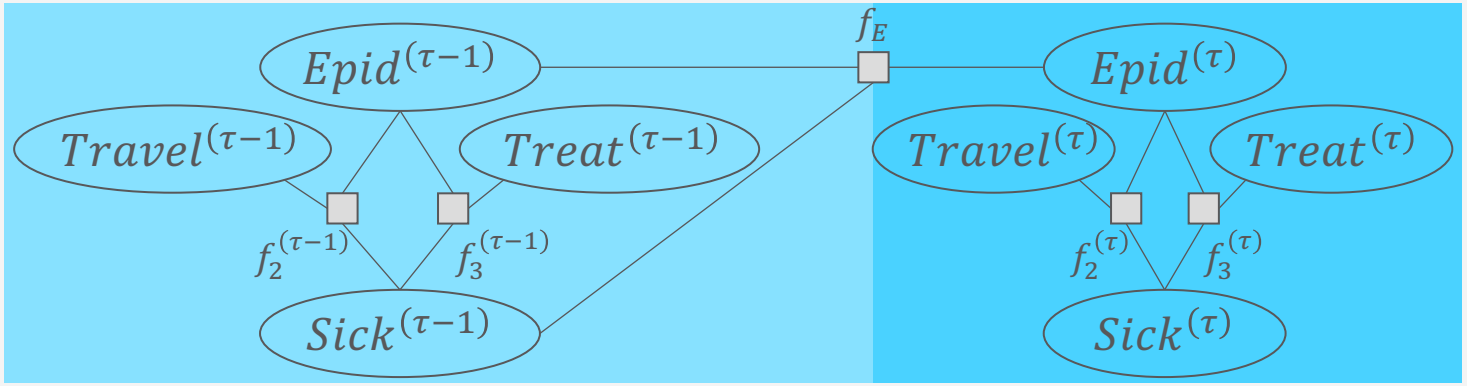
List based on Beierle et al.: Towards a General Framework for Kinds of Forgetting in Common-Sense Belief Management. In *Künstliche Intelligenz*, 2019.

Fading out

- Decay of knowledge
  - Prominent explanation for forgetting in cognitive psychology [Ebbinghaus, 1885]
  - Requires a notion of time!



- Temporal models
  - Describe what happens within a time step and between time steps
  - Evidence accumulates over time
  - Fade out evidence?
    - E.g., make evidence uncertain
  - Fade out numbers? → Ignorance



Ebbinghaus: Über das Gedächtnis. Untersuchungen zur experimentellen Psychologie. Duncker and Humblot, 1885

Commonsense Forgetting

- Contraction
- Ignorance
- Revision / Update
- Abstraction
- Marginalisation
- Focussing
- Tunnelview
- Conditionalisation
- Fading out



List based on Beierle et al.: Towards a General Framework for Kinds of Forgetting in Common-Sense Belief Management. In *Künstliche Intelligenz*, 2019.