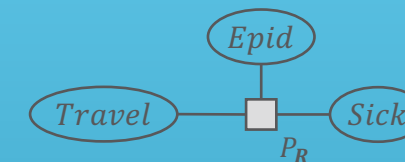
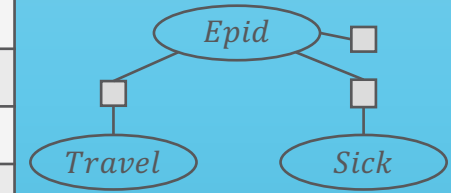
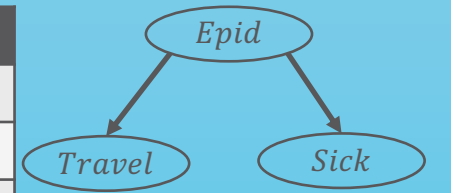


Foundations: Probabilistic Graphical Models

Statistical Relational Artificial Intelligence
(StaRAI)

<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
<i>false</i>	<i>false</i>	<i>false</i>	0.20
<i>false</i>	<i>false</i>	<i>true</i>	0.24
<i>false</i>	<i>true</i>	<i>false</i>	0.28
<i>false</i>	<i>true</i>	<i>true</i>	0.08
<i>true</i>	<i>false</i>	<i>false</i>	0.05
<i>true</i>	<i>false</i>	<i>true</i>	0.06
<i>true</i>	<i>true</i>	<i>false</i>	0.07
<i>true</i>	<i>true</i>	<i>true</i>	0.02



Contents

1. Introduction

- Artificial intelligence
- Agent framework
- StaRAI: context, motivation

2. Foundations

- Logic
- Probability theory
- Probabilistic graphical models (PGMs)

3. Probabilistic Relational Models (PRMs)

- Parfactor models, Markov logic networks
- Semantics, inference tasks

4. Lifted Inference

- Exact inference
- Approximate inference, specifically sampling

5. Lifted Learning

- Parameter learning
- Relation learning
- Approximating symmetries

6. Lifted Sequential Models and Inference

- Parameterised models
- Semantics, inference tasks, algorithm

7. Lifted Decision Making

- Preferences, utility
- Decision-theoretic models, tasks, algorithm

8. Continuous Space and Lifting

- Lifted Gaussian Bayesian networks (BNs)
- Probabilistic soft logic (PSL)

Overview: 2. Foundations

A. *Logic*

- Propositional logic: alphabet, grammar, normal forms, rules
- First-order logic: introducing quantifiers, domain constraints

B. *Probability theory*

- Modelling: (conditional) probability distributions, random variables, marginal and joint distributions
- Inference: axioms and basic rules, Bayes theorem, independence

C. ***Probabilistic graphical models***

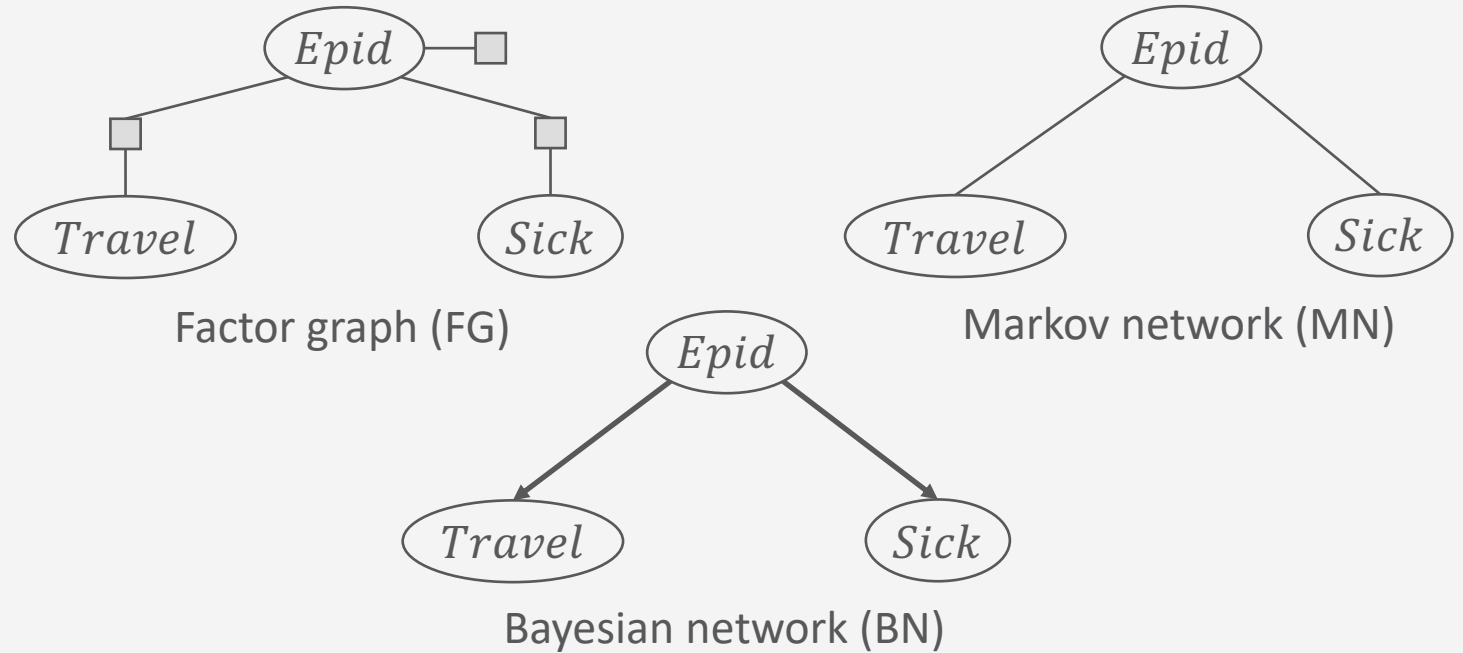
- Syntax, semantics
- Inference problems

Propositional Graphical Models (PGMs)

- Factorisation of a full joint according to (conditional) independences in the full joint

<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
<i>false</i>	<i>false</i>	<i>false</i>	0.20
<i>false</i>	<i>false</i>	<i>true</i>	0.24
<i>false</i>	<i>true</i>	<i>false</i>	0.28
<i>false</i>	<i>true</i>	<i>true</i>	0.08
<i>true</i>	<i>false</i>	<i>false</i>	0.05
<i>true</i>	<i>false</i>	<i>true</i>	0.06
<i>true</i>	<i>true</i>	<i>false</i>	0.07
<i>true</i>	<i>true</i>	<i>true</i>	0.02

Full joint probability distribution



Variants of PGMs

Random Variables

- Characterise scenario by set of **random variables**
 - $R = \{R_1, \dots, R_N\}$
 - Often depicted as ellipses
 - E.g.,

$$\{Epid, Travel.eve, Sick.eve\}$$
- Possible values a random variable can take = **range (or valuation)**
 - $ran(R) = Val(R) = \{v_1, \dots, v_m\}$
 - If $|ran(R)| = 2$, often called Boolean range
 - E.g.,

$$ran(Epid) = ran(Travel.eve) = ran(Sick.eve) = \{true, false\}$$

Epid

Travel

Sick

Events

- Observing or setting a random variable to a specific range value = **event**

- $R = r, r \in \mathcal{R}(R)$

- Shorthand:

- If R clear from context, we write r instead of $R = r$

- If $\text{ran}(R)$ Boolean, we write r for $R = \text{true}$ and $\neg r$ for $R = \text{false}$

- E.g.,

$$\begin{array}{ccc}
 \text{Epid} = e & | & \text{Epid} = \text{true} & \text{Epid} = \text{false} \\
 e & & \text{epid} & \neg \text{epid}
 \end{array}$$

- Setting range values for a set of random variables, one value for each variable = **compound event**

Epid

Travel

Sick

Full Joint Probability Distribution

- 1 world = compound event for R

$$\begin{aligned}
 & epid \\
 & \neg travel \\
 & \neg sick
 \end{aligned}$$

- Specify a probability for a world

$$P(
 \begin{array}{l}
 epid, \\
 \neg travel, \\
 \neg sick
 \end{array}
) = 0.05$$

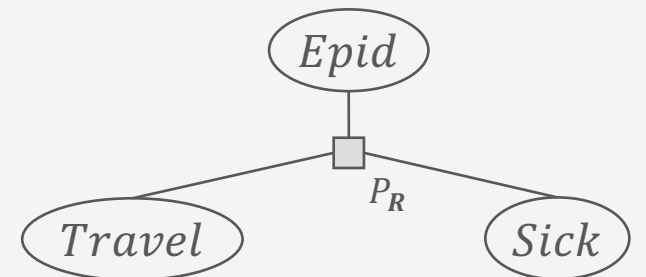
- Joint probability distribution P_R over all (l) possible worlds

- $\sum_{i=1}^l P(w_i) = 1$

- w_i : compound event for R

How large is l ?

<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
<i>false</i>	<i>false</i>	<i>false</i>	0.20
<i>false</i>	<i>false</i>	<i>true</i>	0.24
<i>false</i>	<i>true</i>	<i>false</i>	0.28
<i>false</i>	<i>true</i>	<i>true</i>	0.08
<i>true</i>	<i>false</i>	<i>false</i>	0.05
<i>true</i>	<i>false</i>	<i>true</i>	0.06
<i>true</i>	<i>true</i>	<i>false</i>	0.07
<i>true</i>	<i>true</i>	<i>true</i>	0.02



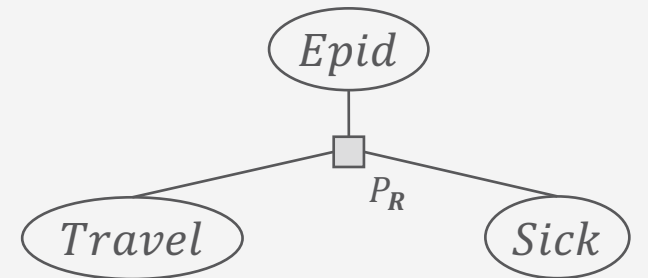
Space Complexity

- Joint probability distribution P_R over all (l) possible worlds
 - $\sum_{i=1}^l P(w_i) = 1$
 - w_i : compound event for $R \in \mathbf{R}$
- Space complexity: $O(r^N)$
 - $r = \max_{R \in \mathbf{R}} |\text{ran}(R)|$
 - $N = |\mathbf{R}|$
 - Derivation:

$$\underbrace{\prod_{R \in \mathbf{R}} |\text{ran}(R)|}_{\text{Exact size}} \leq \prod_{R \in \mathbf{R}} \max_{R \in \mathbf{R}} |\text{ran}(R)| = \prod_{R \in \mathbf{R}} r = r^{|\mathbf{R}|} = r^N$$

Exponential in N !

<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
<i>false</i>	<i>false</i>	<i>false</i>	0.20
<i>false</i>	<i>false</i>	<i>true</i>	0.24
<i>false</i>	<i>true</i>	<i>false</i>	0.28
<i>false</i>	<i>true</i>	<i>true</i>	0.08
<i>true</i>	<i>false</i>	<i>false</i>	0.05
<i>true</i>	<i>false</i>	<i>true</i>	0.06
<i>true</i>	<i>true</i>	<i>false</i>	0.07
<i>true</i>	<i>true</i>	<i>true</i>	0.02



Inference Tasks

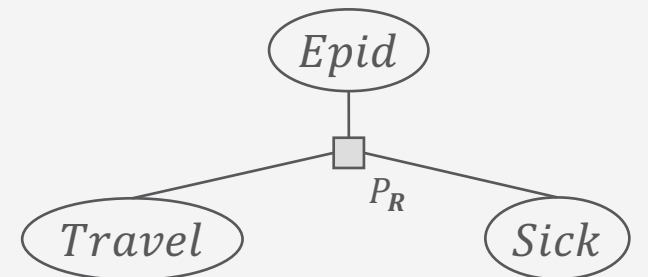
- Query Answering Problem

- Compute an answer to a query given full joint probability distribution P_R
 - Query for a marginal (conditional) probability (distribution)
 - Marginal probability of events: $P(epid)$
 - Marginal probability distribution of random variables: $P(Epid, Travel)$
 - Marginal **conditional** probability of events **given** random variables or events: $P(epid|sick)$
 - Marginal **conditional** probability distribution of random variables **given** random variables or events: $P(Sick|Epid)$

- Next slides

- Syntax of queries
- Solving an instance of a query answering problem
 - Preview: Eliminate all non-query terms

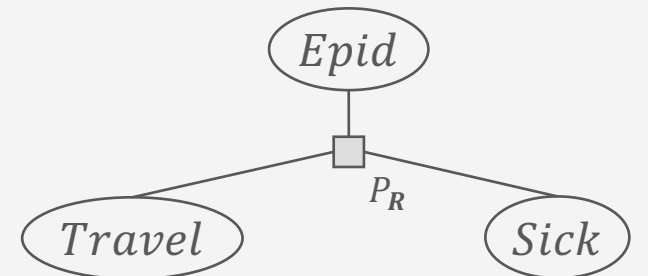
<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
<i>false</i>	<i>false</i>	<i>false</i>	0.20
<i>false</i>	<i>false</i>	<i>true</i>	0.24
<i>false</i>	<i>true</i>	<i>false</i>	0.28
<i>false</i>	<i>true</i>	<i>true</i>	0.08
<i>true</i>	<i>false</i>	<i>false</i>	0.05
<i>true</i>	<i>false</i>	<i>true</i>	0.06
<i>true</i>	<i>true</i>	<i>false</i>	0.07
<i>true</i>	<i>true</i>	<i>true</i>	0.02



Query Syntax

- Marginal probability (distribution) w.r.t. P_R : $P(\mathcal{S})$
 - $rv(\mathcal{S}) \subseteq R$
 - $rv(\cdot)$: shorthand notation to refer to random variables in the input
 - \mathcal{S} : random variables or events
 - Example: $P(Epid, Travel)$
- Conditional marginal probability distribution w.r.t. P_R : $P(\mathcal{S}|\mathcal{T})$
 - $rv(\mathcal{S}, \mathcal{T}) \subseteq R$
 - $\mathcal{S} \cap \mathcal{T} = \emptyset$
 - \mathcal{S} : random variables or events
 - \mathcal{T} : random variables or events t (considered observations, called **evidence**)
 - Example: $P(Travel|Epid)$, $P(Epid|sick)$

<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
<i>false</i>	<i>false</i>	<i>false</i>	0.20
<i>false</i>	<i>false</i>	<i>true</i>	0.24
<i>false</i>	<i>true</i>	<i>false</i>	0.28
<i>false</i>	<i>true</i>	<i>true</i>	0.08
<i>true</i>	<i>false</i>	<i>false</i>	0.05
<i>true</i>	<i>false</i>	<i>true</i>	0.06
<i>true</i>	<i>true</i>	<i>false</i>	0.07
<i>true</i>	<i>true</i>	<i>true</i>	0.02



Answering Marginal Queries

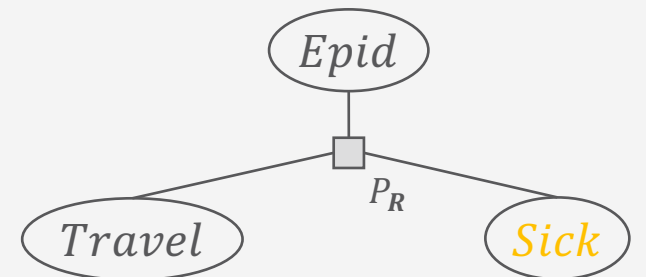
- Solving $P(\mathcal{S})$: Eliminate all non-query terms \mathbf{U}
 - $\mathbf{U} = \mathbf{R} \setminus rv(\mathcal{S})$

$$P(\mathcal{S}) = \sum_{\mathbf{u} \in \text{ran}(\mathbf{U})} P_{\mathbf{R}}(\mathcal{S}, \mathbf{U} = \mathbf{u})$$

- E.g., query $P(\text{Epid}, \text{Travel}) \rightarrow \mathbf{U} = \{\text{Sick}\}$

$$P(\text{Epid}, \text{Travel}) = \sum_{v \in \text{ran}(\text{Sick})} P(\text{Epid}, \text{Travel}, \text{Sick} = v)$$

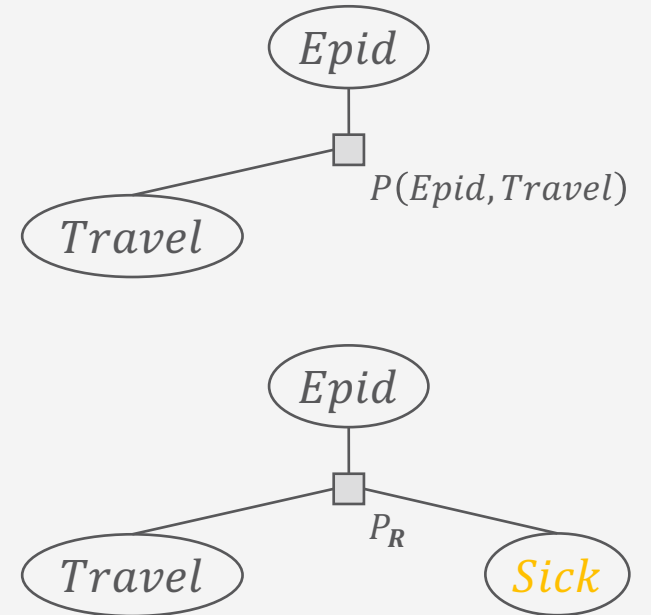
<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
<i>false</i>	<i>false</i>	<i>false</i>	0.20
<i>false</i>	<i>false</i>	<i>true</i>	0.24
<i>false</i>	<i>true</i>	<i>false</i>	0.28
<i>false</i>	<i>true</i>	<i>true</i>	0.08
<i>true</i>	<i>false</i>	<i>false</i>	0.05
<i>true</i>	<i>false</i>	<i>true</i>	0.06
<i>true</i>	<i>true</i>	<i>false</i>	0.07
<i>true</i>	<i>true</i>	<i>true</i>	0.02



Answering Marginal Queries

$$P(\text{Epid}, \text{Travel}) = \sum_{v \in \text{ran}(\text{Sick})} P(\text{Epid}, \text{Travel}, \text{Sick} = v)$$

<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>		<i>Epid</i>	<i>Travel</i>	<i>P</i>
<i>false</i>	<i>false</i>	<i>false</i>	0.20	+	<i>false</i>	<i>false</i>	0.44
<i>false</i>	<i>false</i>	<i>true</i>	0.24		<i>false</i>	<i>true</i>	0.36
<i>false</i>	<i>true</i>	<i>false</i>	0.28	+	<i>true</i>	<i>false</i>	0.11
<i>false</i>	<i>true</i>	<i>true</i>	0.08		<i>true</i>	<i>true</i>	0.09
<i>true</i>	<i>false</i>	<i>false</i>	0.05				
<i>true</i>	<i>false</i>	<i>true</i>	0.06				
<i>true</i>	<i>true</i>	<i>false</i>	0.07				
<i>true</i>	<i>true</i>	<i>true</i>	0.02				



Answering Marginal Queries

- If \mathcal{S} in $P(\mathcal{S})$ consists of **events**, consider only the worlds that are consistent with the events
 - E.g., query $P(\textit{epid})$ with $\mathbf{U} = \{\textit{Travel}, \textit{Sick}\}$

$$P(\textit{epid}) = \sum_{v_t \in \text{ran}(\textit{Travel})} \sum_{v_s \in \text{ran}(\textit{Sick})} P(\textit{epid}, \textit{Travel} = v_t, \textit{Sick} = v_s)$$

<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
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

No probability distribution at this point
(in the middle of the computation)

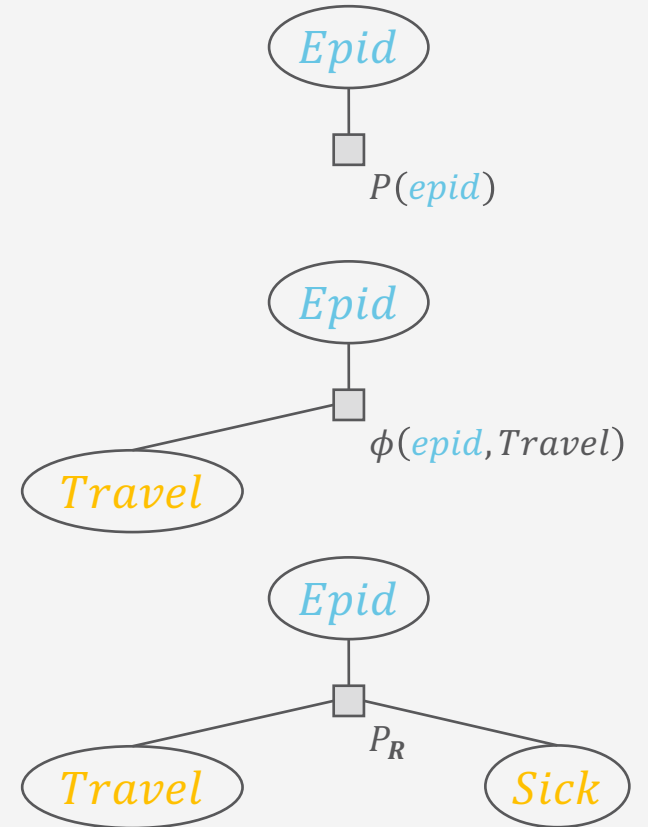
<i>Epid</i>	<i>Travel</i>	ϕ
true	false	0.11
true	true	0.09

true	false	0.11
------	-------	------

true	true	0.09
------	------	------

<i>Epid</i>	<i>P</i>
true	0.20

true	0.20
------	------



Answering Conditional Queries

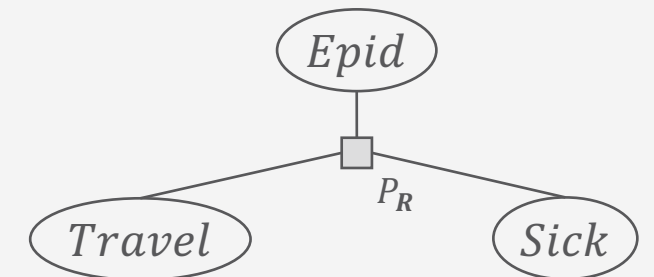
- Solving $P(\mathcal{S}|\mathcal{T})$ w.r.t. P_R :
 - $P(\mathcal{S}|\mathcal{T}) = \frac{P(\mathcal{S},\mathcal{T})}{P(\mathcal{T})}$
 - $P(\mathcal{T})$ normalising constant
 - Reduces to computing two marginal queries: $P(\mathcal{S},\mathcal{T})$, $P(\mathcal{T})$
 - Eliminate all non-query terms \mathcal{U} and **normalise**
 - $\mathcal{U} = \mathcal{R} \setminus rv(\mathcal{S},\mathcal{T})$
- $$P(\mathcal{S}|\mathcal{T}) = \frac{1}{P(\mathcal{T})} \sum_{u \in ran(\mathcal{U})} P_R(\mathcal{S},\mathcal{T}, \mathcal{U} = u)$$
- If $\mathcal{T} = t$ (evidence), drop the rows where $\mathcal{T} \neq t$ and columns \mathcal{T}
 - Called evidence *absorption*, reduces dimension of P_R as \mathcal{T} disappears
 - Equal to setting probabilities to 0 where $\mathcal{T} \neq t$ and summing out \mathcal{T}

- As \mathcal{T} often t , $P(\mathcal{T}) = P(t)$ is constant for $P(\mathcal{S}|t)$, thus called *normalising constant*.
- As $P(t)$ is a sum expression in itself, it is often abbreviated:

$$P(\mathcal{S}|\mathcal{T}) = \frac{1}{Z} \Sigma \dots \quad (\underline{Z} \text{ Zustandssumme})$$

$$P(\mathcal{S}|\mathcal{T}) = \alpha \Sigma \dots$$

$$P(\mathcal{S}|\mathcal{T}) \propto \Sigma \dots \quad (\text{proportional to})$$



Answering Conditional Queries: Normalisation

- Normalise by $P(\mathbf{T})$: for each $t \in \text{ran}(\mathbf{T})$, for each $s \in \text{ran}(\mathbf{S})$, compute $\frac{P(\mathbf{s}, t)}{P(t)}$

$$P(\text{Travel}|\text{Epid}) = \frac{1}{P(\text{Epid})} \sum_{v_s \in \text{ran}(\text{Sick})} \underbrace{P(\text{Epid}, \text{Travel}, \text{Sick} = v_s)}_{\phi}$$

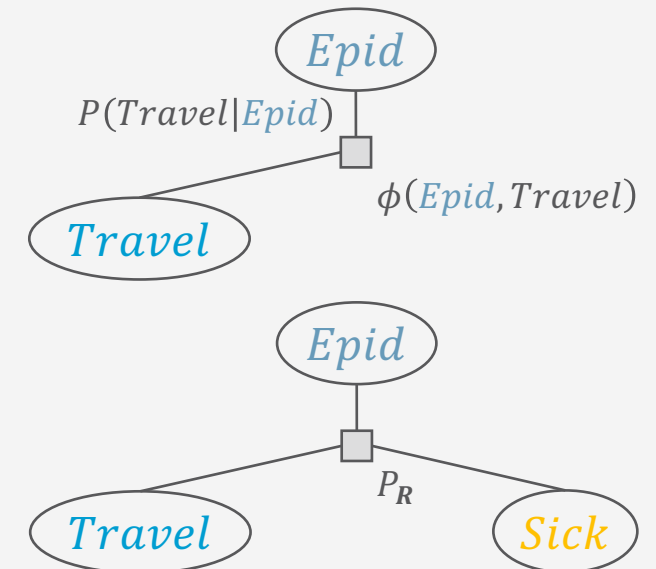
- for each $e \in \text{ran}(\text{Epid})$: for each $t \in \text{ran}(\text{Travel})$:
divide $\phi(e, t)$ by $P(e) = \phi(e, t) + \phi(e, \neg t)$

What do we need to do?

Epid	Travel	ϕ
false	false	0.44
false	true	0.36
true	false	0.11
true	true	0.09

$P(\text{Travel} \text{Epid})$	
$\frac{0.44}{0.44 + 0.36} = 0.55$	
$\frac{0.36}{0.44 + 0.36} = 0.45$	
$\frac{0.11}{0.11 + 0.09} = 0.55$	
$\frac{0.09}{0.11 + 0.09} = 0.45$	

Epid	$P(\text{travel} \text{Epid})$	$P(\neg \text{travel} \text{Epid})$
false	0.55	$1 - 0.55 = 0.45$
true	0.55	$1 - 0.55 = 0.45$



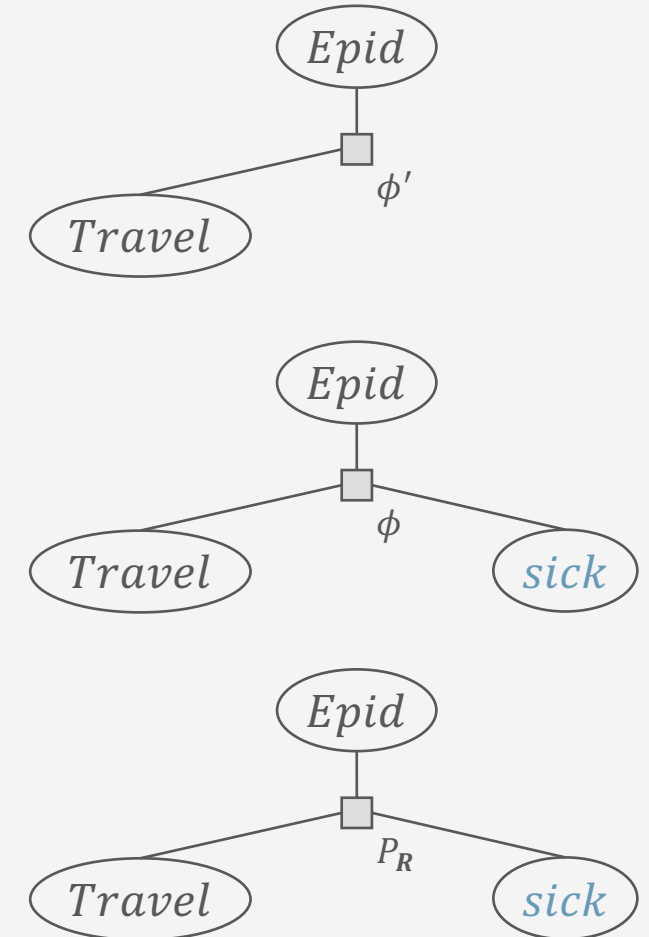
Answering Conditional Queries: Evidence – Absorption

$$P(\text{Epid}|\text{sick}) \propto \sum_{v \in \text{ran}(\text{Travel})} P(\text{Epid}, \text{Travel} = v, \text{sick})$$

Epid	Travel	Sick	P
false	false <td>false <td>0.20 </td></td>	false <td>0.20 </td>	0.20
false	false	true	0.24
false	true <td>false <td>0.28 </td></td>	false <td>0.28 </td>	0.28
false	true	true	0.08
true	false <td>false <td>0.05 </td></td>	false <td>0.05 </td>	0.05
true	false	true	0.06
true	true <td>false <td>0.07 </td></td>	false <td>0.07 </td>	0.07
true	true	true	0.02

Epid	Travel	ϕ'
false	false	0.24
false	true	0.08
true	false	0.06
true	true	0.02

ϕ	Σ_s
0	0 + 0.24
0.24	
0	0 + 0.08
0.08	
0	0 + 0.06
0.06	
0	0 + 0.02
0.02	



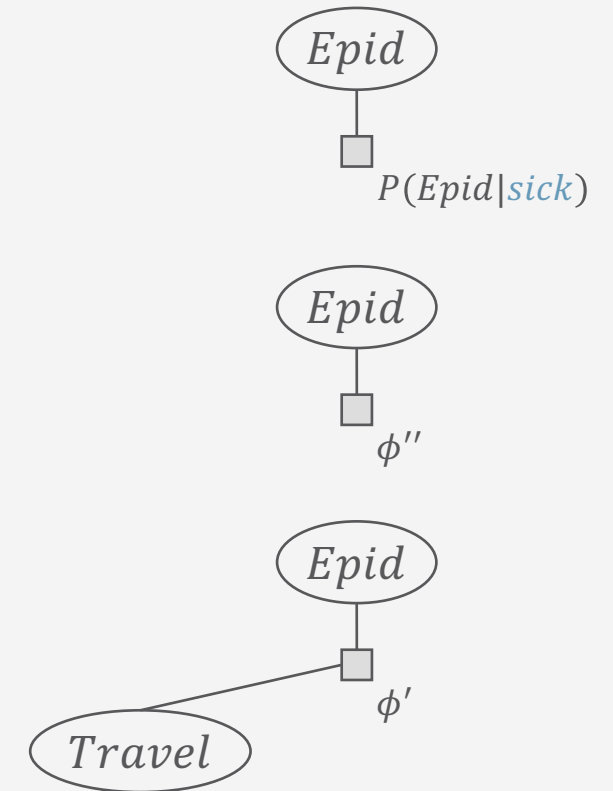
Answering Conditional Queries: Evidence – Normalisation

$$P(\text{Epid}|\text{sick}) \propto \sum_{v \in \text{ran}(\text{Travel})} P(\text{Epid}, \text{Travel} = v, \text{sick})$$

normalise

Epid	Travel	ϕ'		Epid	ϕ''		Epid	P
false	false	0.24	+	false	0.32	→	false	$\frac{0.32}{0.32 + 0.08} = 0.8$
false	true	0.08		true	0.08		true	$\frac{0.08}{0.32 + 0.08} = 0.2$
true	false	0.06	+			→		
true	true	0.02						

for each $t \in \text{ran}(\mathbf{T})$, for each $s \in \text{ran}(\mathbf{S})$, compute $\frac{P(s,t)}{P(t)}$
 with $\mathbf{T} = t$, only one case t



Runtime Complexity

$$P(\mathbf{S}|\mathbf{T}) = \frac{1}{P(\mathbf{T})} \sum_{\mathbf{u} \in \text{ran}(\mathbf{U})} P_{\mathbf{R}}(\mathbf{S}, \mathbf{T}, \mathbf{U} = \mathbf{u})$$

- Runtime complexity: $O(r^N)$

- $r = \max_{R \in \mathcal{R}} |\mathcal{R}(R)|$

- $N = |\mathbf{R}|$

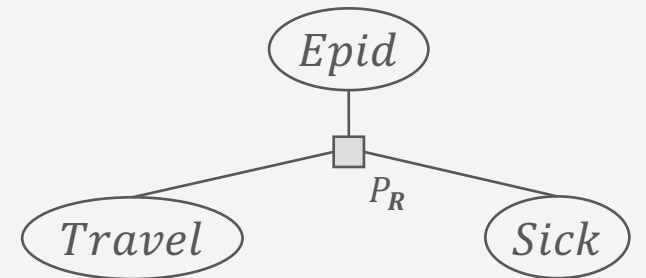
- Have to go through whole table; derivation as before

$$\prod_{R \in \mathbf{R}} |\text{ran}(R)| \leq \prod_{R \in \mathbf{R}} \max_{R \in \mathcal{R}} |\text{ran}(R)| = \prod_{R \in \mathbf{R}} r = r^{|\mathbf{R}|} = r^N$$

= Space complexity

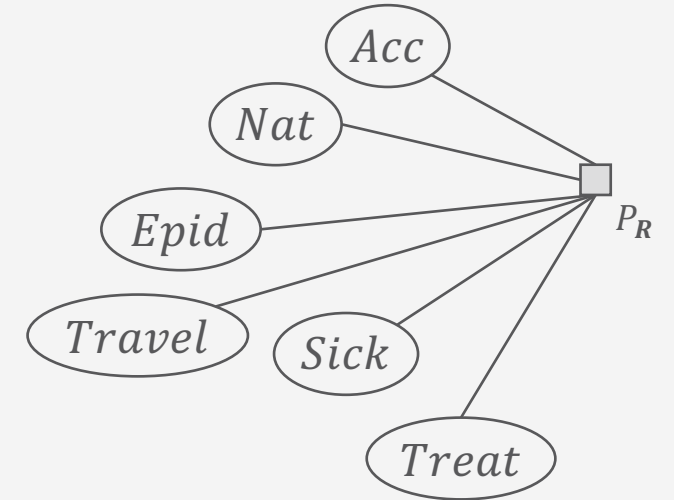
Exponential in N !

<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
<i>false</i>	<i>false</i>	<i>false</i>	0.20
<i>false</i>	<i>false</i>	<i>true</i>	0.24
<i>false</i>	<i>true</i>	<i>false</i>	0.28
<i>false</i>	<i>true</i>	<i>true</i>	0.08
<i>true</i>	<i>false</i>	<i>false</i>	0.05
<i>true</i>	<i>false</i>	<i>true</i>	0.06
<i>true</i>	<i>true</i>	<i>false</i>	0.07
<i>true</i>	<i>true</i>	<i>true</i>	0.02



Exponential Blowup!

<i>Acc</i>	<i>Nat</i>	<i>Treat</i>	<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	0.025
<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>true</i>	0.009
<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>true</i>	<i>false</i>	0.009
<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>true</i>	<i>true</i>	0.009
<i>false</i>	<i>false</i>	<i>false</i>	<i>true</i>	<i>false</i>	<i>false</i>	0.009
<i>false</i>	<i>false</i>	<i>false</i>	<i>true</i>	<i>true</i>	<i>false</i>	0.009
<i>false</i>	<i>false</i>	<i>true</i>	<i>false</i>	<i>false</i>	<i>false</i>	0.009
<i>false</i>	<i>false</i>	<i>true</i>	<i>true</i>	<i>false</i>	<i>false</i>	0.009
<i>false</i>	<i>true</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	0.009
<i>false</i>	<i>true</i>	<i>false</i>	<i>true</i>	<i>false</i>	<i>false</i>	0.009
<i>false</i>	<i>true</i>	<i>true</i>	<i>false</i>	<i>false</i>	<i>false</i>	0.009
<i>false</i>	<i>true</i>	<i>true</i>	<i>true</i>	<i>false</i>	<i>false</i>	0.009
<i>true</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	0.009
<i>true</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>true</i>	<i>false</i>	0.009
<i>true</i>	<i>false</i>	<i>false</i>	<i>true</i>	<i>false</i>	<i>false</i>	0.009
<i>true</i>	<i>false</i>	<i>false</i>	<i>true</i>	<i>true</i>	<i>false</i>	0.009
<i>true</i>	<i>true</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	0.009
<i>true</i>	<i>true</i>	<i>false</i>	<i>true</i>	<i>false</i>	<i>false</i>	0.009
<i>true</i>	<i>true</i>	<i>true</i>	<i>false</i>	<i>false</i>	<i>false</i>	0.009
<i>true</i>	<i>true</i>	<i>true</i>	<i>true</i>	<i>false</i>	<i>false</i>	0.009

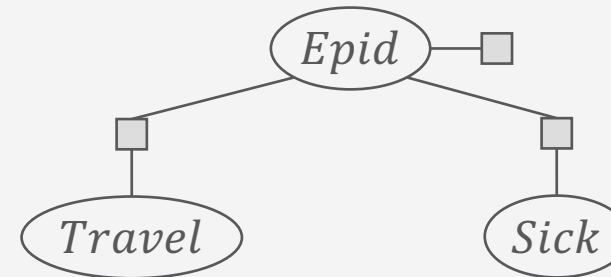
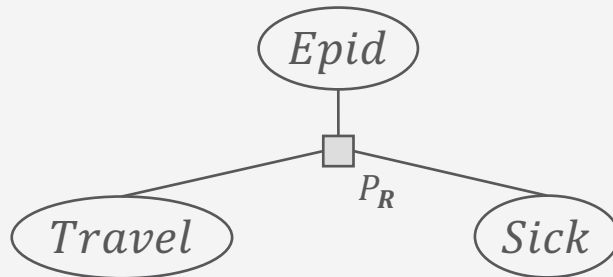


$2^6 = 64$ possible worlds

Adding relations means adding *Sick*, *Treat*, *Travel* variables for each person, blowing up the model further

Compact Encoding

- Full joint probability distribution:
Every random variable is connected with every other random variable!
- Factorise full joint probability distribution P_R using (conditional) independences
 - Independence: $P(\mathbf{R}_1, \mathbf{R}_2) = P(\mathbf{R}_1) \cdot P(\mathbf{R}_2) \rightarrow$ denoted $(\mathbf{R}_1 \perp \mathbf{R}_2)$
 - Conditional independence: $P(\mathbf{R}_1, \mathbf{R}_2 | \mathbf{R}_3) = P(\mathbf{R}_1 | \mathbf{R}_3) \cdot P(\mathbf{R}_2 | \mathbf{R}_3) \rightarrow$ denoted $(\mathbf{R}_1 \perp \mathbf{R}_2 | \mathbf{R}_3)$
- Hidden in P_R
 - Explicitly represent through factors and in graph
 - Full joint is then given by *product* of the factors



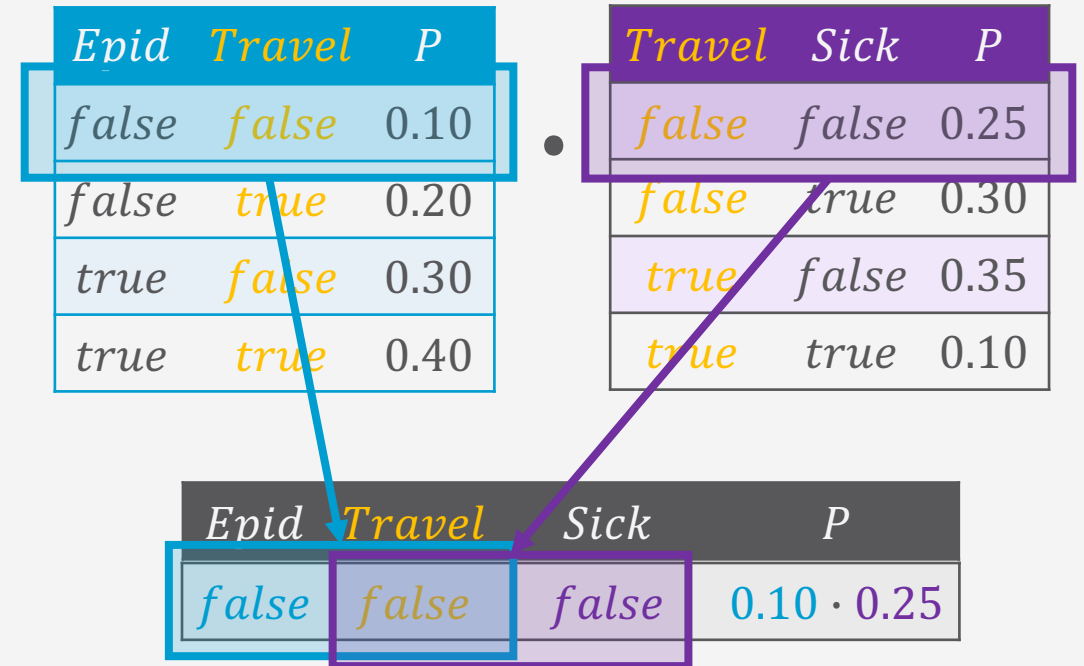
Excursion: Multiplication

- Join over arguments + product of probabilities:

$$\phi(R_1, \dots, R_l) = \phi_1(R_{11}, \dots, R_{1k}) \cdot \phi_2(R_{21}, \dots, R_{2m})$$

- $\{R_1, \dots, R_l\} = \{R_{11}, \dots, R_{1k}\} \cup \{R_{21}, \dots, R_{2m}\}$
- No common arguments
= cross product of ranges
- E.g.,

$$P(\text{Epid}, \text{Travel}) \cdot P(\text{Travel}, \text{Sick})$$



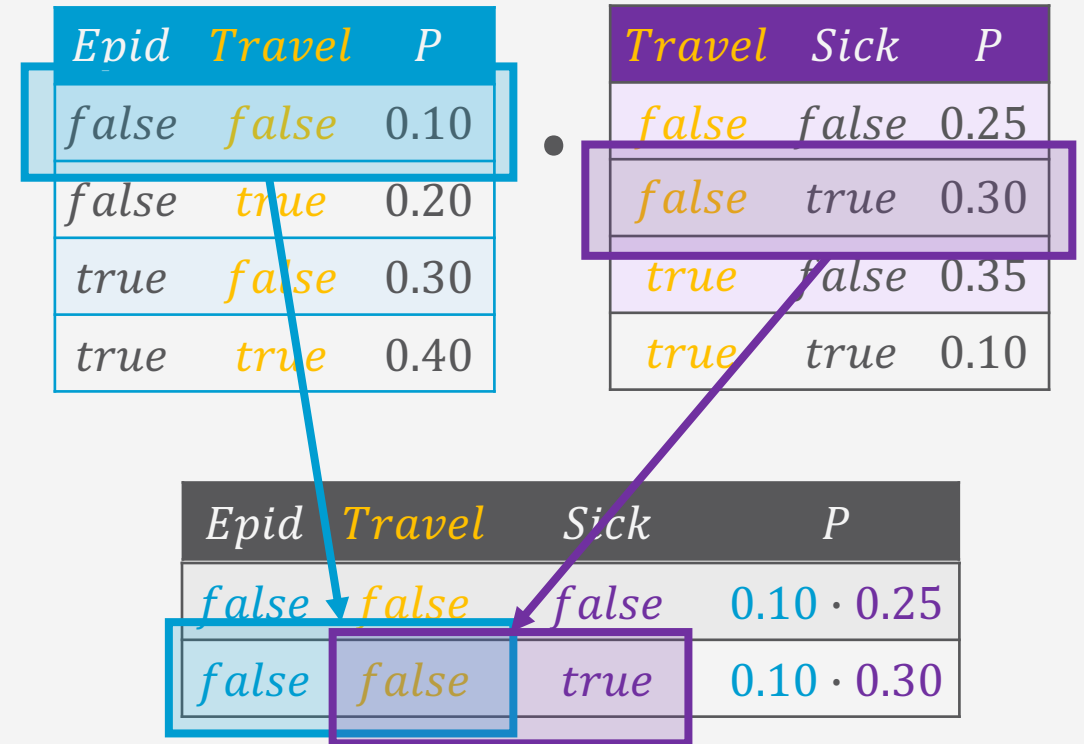
Excursion: Multiplication

- Join over arguments + product of probabilities:

$$\phi(R_1, \dots, R_l) = \phi_1(R_{11}, \dots, R_{1k}) \cdot \phi_2(R_{21}, \dots, R_{2m})$$

- $\{R_1, \dots, R_l\} = \{R_{11}, \dots, R_{1k}\} \cup \{R_{21}, \dots, R_{2m}\}$
- No common arguments
= cross product of ranges
- E.g.,

$$P(\text{Epid}, \text{Travel}) \cdot P(\text{Travel}, \text{Sick})$$



<i>Epid</i>	<i>Travel</i>	<i>P</i>
false	false	0.10
false	true	0.20
true	false	0.30
true	true	0.40

<i>Travel</i>	<i>Sick</i>	<i>P</i>
false	false	0.25
false	true	0.30
true	false	0.35
true	true	0.10

<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
false	false	false	0.10 · 0.25
false	false	true	0.10 · 0.30

Excursion: Multiplication

- Join over arguments + product of probabilities:

$$\phi(R_1, \dots, R_l) = \phi_1(R_{11}, \dots, R_{1k}) \cdot \phi_2(R_{21}, \dots, R_{2m})$$

- $\{R_1, \dots, R_l\} = \{R_{11}, \dots, R_{1k}\} \cup \{R_{21}, \dots, R_{2m}\}$
- No common arguments
= cross product of ranges
- E.g.,

$$P(\text{Epid}, \text{Travel}) \cdot P(\text{Travel}, \text{Sick})$$

<i>Epid</i>	<i>Travel</i>	<i>P</i>
false	false	0.10
false	true	0.20
true	false	0.30
true	true	0.40

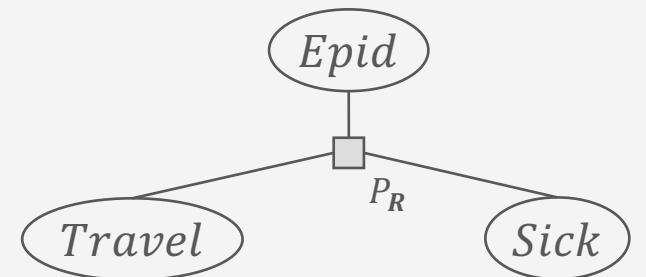
<i>Travel</i>	<i>Sick</i>	<i>P</i>
false	false	0.25
false	true	0.30
true	false	0.35
true	true	0.10

<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
false	false	false	0.10 · 0.25
false	false	true	0.10 · 0.30
false	true	false	0.20 · 0.35
			⋮

Independences: Examples

- Independence: ($Travel \perp Sick$)?
 - $P(Travel, Sick) \stackrel{?}{=} P(Travel) \cdot P(Sick)$
 \rightarrow only *true, true* case:
 - $P(travel, sick) = 0.08 + 0.02 = 0.1$
 - $P(travel) = 0.28 + 0.08 + 0.07 + 0.02 = 0.45$
 - $P(sick) = 0.24 + 0.08 + 0.06 + 0.02 = 0.4$
 - $P(travel) \cdot P(sick) = 0.45 \cdot 0.4 = 0.18 \neq 0.1$ ✗
- Conditional independence ($Travel \perp Sick | Epid$)?
 - $P(Travel, Sick | Epid) \stackrel{?}{=} P(Travel | Epid) \cdot P(Sick | Epid)$ ✗

<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
<i>false</i>	<i>false</i>	<i>false</i>	0.20
<i>false</i>	<i>false</i>	<i>true</i>	0.24
<i>false</i>	<i>true</i>	<i>false</i>	0.28
<i>false</i>	<i>true</i>	<i>true</i>	0.08
<i>true</i>	<i>false</i>	<i>false</i>	0.05
<i>true</i>	<i>false</i>	<i>true</i>	0.06
<i>true</i>	<i>true</i>	<i>false</i>	0.07
<i>true</i>	<i>true</i>	<i>true</i>	0.02



Independences: Examples

- Travel* \perp *Sick* | *Epid*?

- $P(\text{Travel}, \text{Sick} | \text{Epid}) \stackrel{?}{=} P(\text{Travel} | \text{Epid}) \cdot P(\text{Sick} | \text{Epid})$ ✓

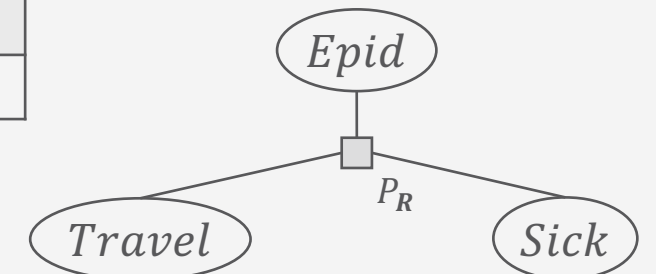
<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
false	false	false	0.6375
false	false	true	0.1125
false	true	false	0.2125
false	true	true	0.0375
true	false	false	0.45
true	false	true	0.15
true	true	false	0.3
true	true	true	0.1

<i>Epid</i>	<i>Travel</i>	<i>P</i>
false	false	0.75
false	true	0.25
true	false	0.6
true	true	0.4

<i>Epid</i>	<i>Sick</i>	<i>P</i>
false	false	0.85
false	true	0.15
true	false	0.75
true	true	0.25

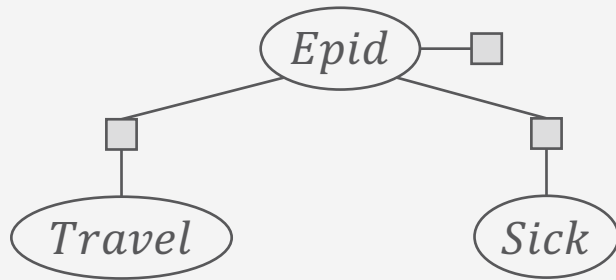
<i>P</i>
$0.75 \cdot 0.85 = 0.6375$
$0.75 \cdot 0.15 = 0.1125$
$0.25 \cdot 0.85 = 0.2125$
$0.25 \cdot 0.15 = 0.0375$
$0.6 \cdot 0.75 = 0.45$
$0.6 \cdot 0.25 = 0.15$
$0.4 \cdot 0.75 = 0.3$
$0.4 \cdot 0.25 = 0.1$

<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
false	false	false	0.51
false	false	true	0.09
false	true	false	0.17
false	true	true	0.03
true	false	false	0.09
true	false	true	0.03
true	true	false	0.06
true	true	true	0.02



Independences: Examples

- Factorise the full joint into its factors based on independences:
 - $P(\text{Epid}, \text{Travel}, \text{Sick}) = P(\text{Epid}) \cdot P(\text{Travel}|\text{Epid}) \cdot P(\text{Sick}|\text{Epid})$



Epid	P
false	0.8
true	0.2

Epid	Travel	P
false	false	0.75
false	true	0.25
true	false	0.6
true	true	0.4

Epid	Sick	P
false	false	0.85
false	true	0.15
true	false	0.75
true	true	0.25

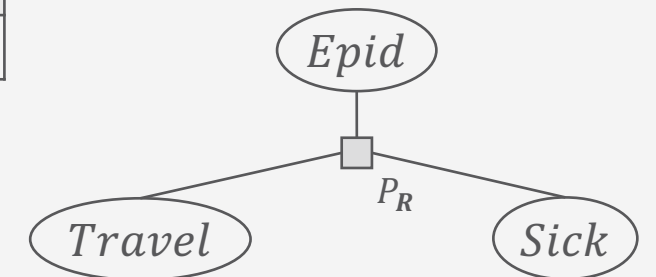
$P(\text{epid})$
0.2

Epid	$P(\text{travel} \text{Epid})$
false	0.25
true	0.4

Epid	$P(\text{sick} \text{Epid})$
false	0.15
true	0.25

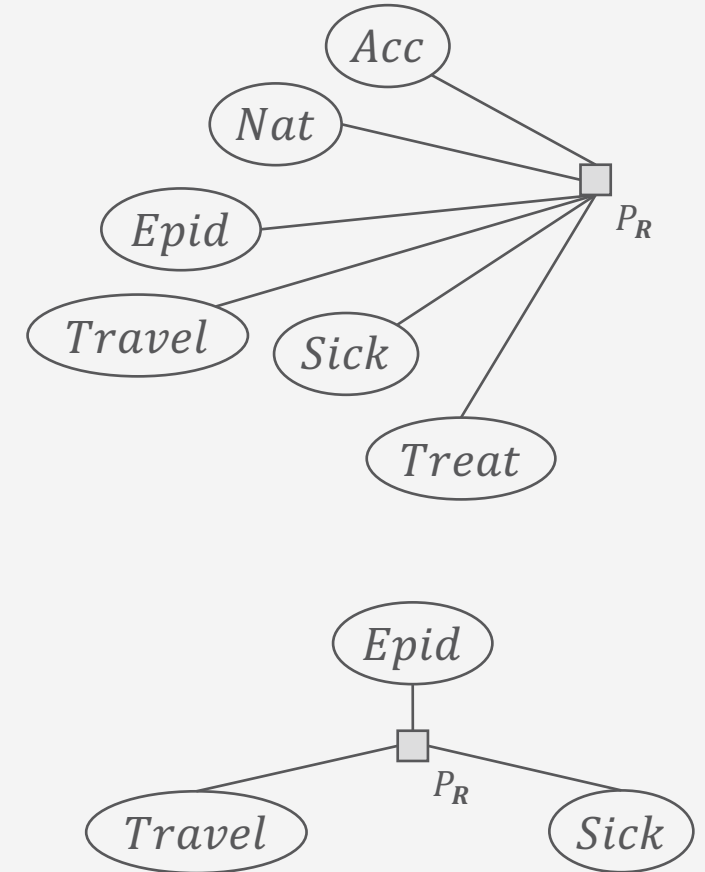
Epid	Travel	Sick	P
false	false	false	0.51
false	false	true	0.09
false	true	false	0.17
false	true	true	0.03
true	false	false	0.09
true	false	true	0.03
true	true	false	0.06
true	true	true	0.02

- Usually fewer entries to store



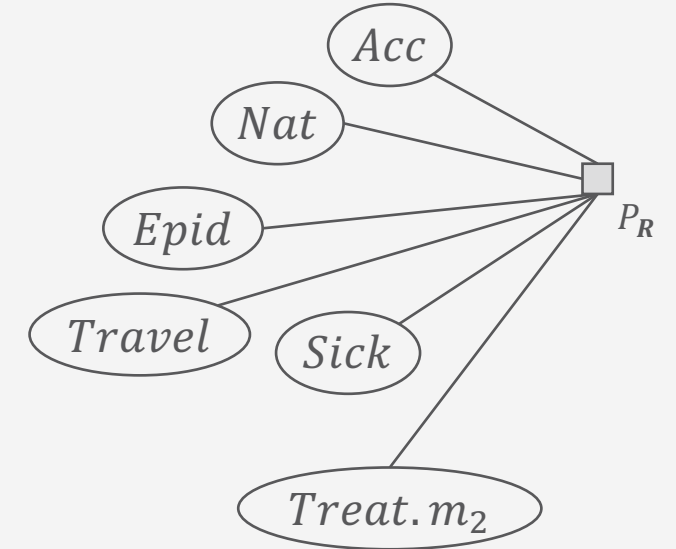
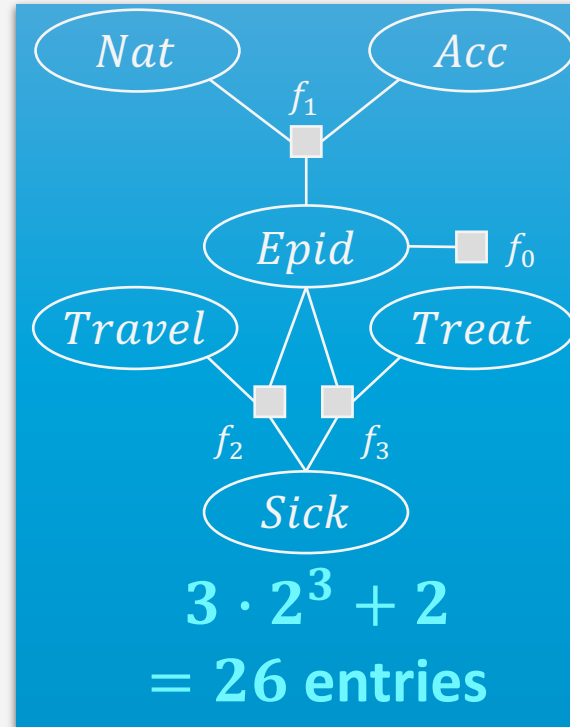
Finding a Compact Encoding

- At beginning: Everything connected with everything in full joint
- Find (conditional) independences in P_R
 - Check every possible combination → **Combinatorial explosion!**
 - E.g., (many more)
 - $(Epid \perp Travel, Sick)$, $(Travel \perp Epid, Sick)$, $(Sick \perp Travel, Epid)$
 - Partitions P_R into a **set of factors**
 - Deletes connections between random variables
- Alternative:
 - Start with no connections, add factors
 - More later (→ *Section 5: Lifted Learning*)
- For now, assume that we have a factorised model



Exponential Blowup! → Sparse Encoding

<i>Acc</i>	<i>Nat</i>	<i>Treat</i>	<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	0.025
<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>true</i>	0.009
<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>true</i>	<i>false</i>	0.009
<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>true</i>	<i>true</i>	0.009
<i>false</i>	<i>false</i>	<i>false</i>	<i>true</i>	<i>false</i>	<i>false</i>	0.009
<i>false</i>	<i>false</i>	<i>false</i>	<i>true</i>	<i>true</i>	<i>false</i>	0.009
<i>false</i>	<i>false</i>	<i>false</i>	<i>true</i>	<i>true</i>	<i>true</i>	0.009
<i>false</i>	<i>true</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	0.009
<i>false</i>	<i>true</i>	<i>false</i>	<i>false</i>	<i>true</i>	<i>false</i>	0.009
<i>false</i>	<i>true</i>	<i>false</i>	<i>false</i>	<i>true</i>	<i>true</i>	0.009
<i>false</i>	<i>true</i>	<i>true</i>	<i>false</i>	<i>false</i>	<i>false</i>	0.009
<i>false</i>	<i>true</i>	<i>true</i>	<i>false</i>	<i>true</i>	<i>false</i>	0.009
<i>false</i>	<i>true</i>	<i>true</i>	<i>false</i>	<i>true</i>	<i>true</i>	0.009
<i>false</i>	<i>true</i>	<i>true</i>	<i>true</i>	<i>false</i>	<i>false</i>	0.009
<i>false</i>	<i>true</i>	<i>true</i>	<i>true</i>	<i>true</i>	<i>false</i>	0.009
<i>false</i>	<i>true</i>	<i>true</i>	<i>true</i>	<i>true</i>	<i>true</i>	0.009
<i>true</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	0.009
<i>true</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>true</i>	0.009
<i>true</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>true</i>	<i>false</i>	0.009
<i>true</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>true</i>	<i>true</i>	0.009
<i>true</i>	<i>false</i>	<i>true</i>	<i>false</i>	<i>false</i>	<i>false</i>	0.009
<i>true</i>	<i>false</i>	<i>true</i>	<i>false</i>	<i>true</i>	<i>false</i>	0.009
<i>true</i>	<i>false</i>	<i>true</i>	<i>false</i>	<i>true</i>	<i>true</i>	0.009
<i>true</i>	<i>true</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>false</i>	0.009
<i>true</i>	<i>true</i>	<i>false</i>	<i>false</i>	<i>false</i>	<i>true</i>	0.009
<i>true</i>	<i>true</i>	<i>false</i>	<i>false</i>	<i>true</i>	<i>false</i>	0.009
<i>true</i>	<i>true</i>	<i>false</i>	<i>false</i>	<i>true</i>	<i>true</i>	0.009
<i>true</i>	<i>true</i>	<i>true</i>	<i>false</i>	<i>false</i>	<i>false</i>	0.009
<i>true</i>	<i>true</i>	<i>true</i>	<i>false</i>	<i>false</i>	<i>true</i>	0.009
<i>true</i>	<i>true</i>	<i>true</i>	<i>false</i>	<i>true</i>	<i>false</i>	0.009
<i>true</i>	<i>true</i>	<i>true</i>	<i>false</i>	<i>true</i>	<i>true</i>	0.009
<i>true</i>	<i>true</i>	<i>true</i>	<i>true</i>	<i>false</i>	<i>false</i>	0.009
<i>true</i>	<i>true</i>	<i>true</i>	<i>true</i>	<i>false</i>	<i>true</i>	0.009
<i>true</i>	<i>true</i>	<i>true</i>	<i>true</i>	<i>true</i>	<i>false</i>	0.009
<i>true</i>	<i>true</i>	<i>true</i>	<i>true</i>	<i>true</i>	<i>true</i>	0.009



$2^6 = 64$ possible worlds

Adding relations means adding *Sick*, *Treat*, *Travel* variables for each person, blowing up the model further

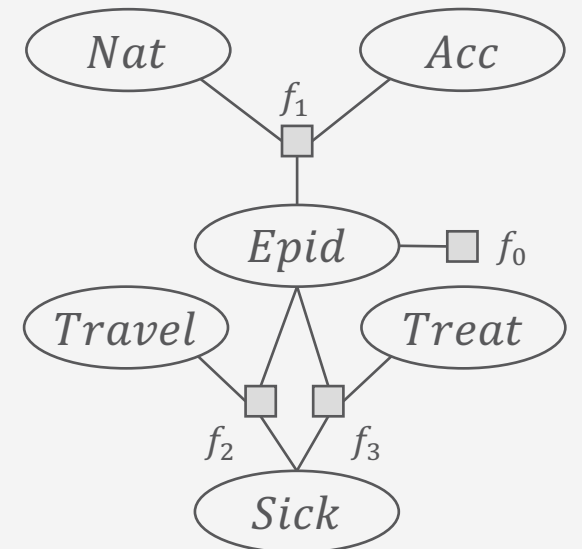
Model Representation: Factors

- Given set of random variables $\mathbf{R} = \{R_1, \dots, R_N\}$

Syntax:

- Set of factors $F = \{f_i\}_{i=1}^{n'}$
= model
- Factor $f = \phi(R_1, \dots, R_k)$
 - Arguments $R_1, \dots, R_k \in \mathbf{R}$
 - Potential function

$$\phi: \times_{i=1}^k \text{ran}(R_i) \rightarrow \mathbb{R}^{0,+}$$
 - At least one potential > 0
 - Write as table, list, ...
 - Not required to be a probability distribution



Model Representation: Factors

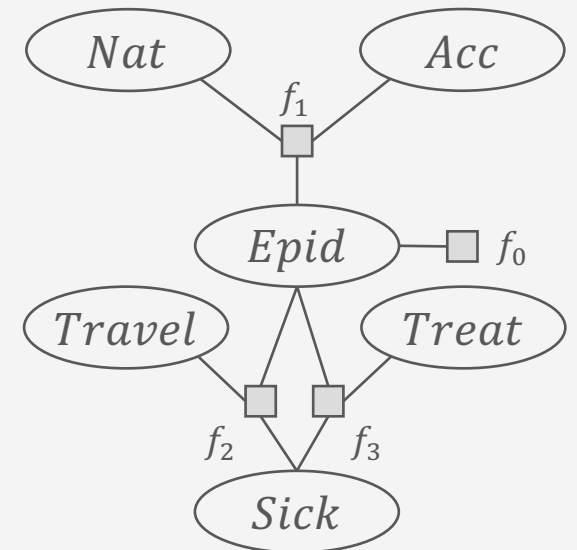
- Given model $F = \{f_i\}_{i=1}^n$ over random variables $\mathbf{R} = \{R_1, \dots, R_N\}$
 - $f_i = \phi_i(R_1, \dots, R_k)$

Semantics:

- Build full joint probability distribution P_F

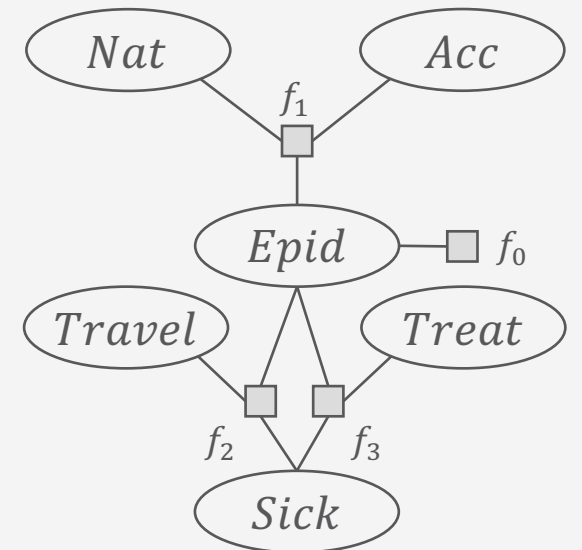
$$P_F = \frac{1}{Z} \prod_{i=1}^n \phi_i(R_1, \dots, R_k)$$

$$Z = \sum_{r_1 \in \text{ran}(R_1)} \sum_{r_n \in \text{ran}(R_n)} \prod_{i=1}^n \phi_i(r_1, \dots, r_k)$$



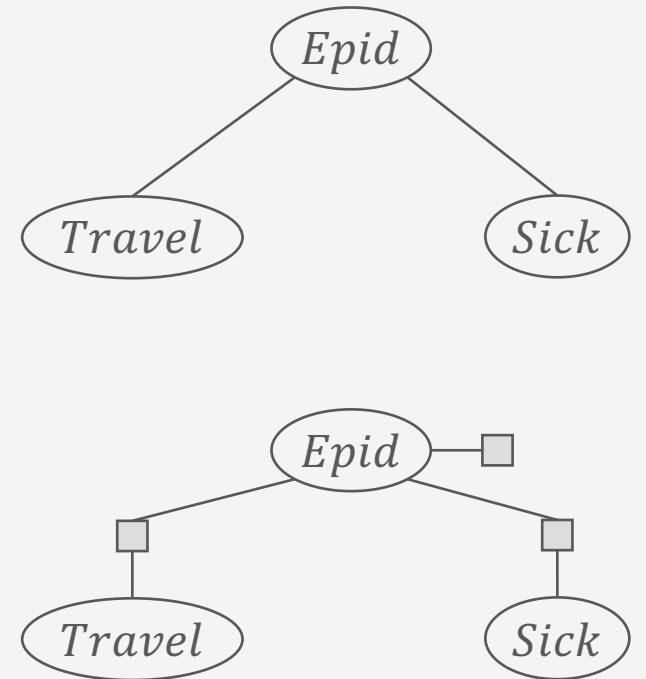
Model Representation: Factors

- Given model $F = \{f_i\}_{i=1}^n$ over random variables $\mathbf{R} = \{R_1, \dots, R_N\}$
 - $f_i = \phi_i(R_1, \dots, R_k)$
- Graphical representation: **Factor graph** (FG)
 - Each $R \in \mathbf{R}$: variable node in FG (ellipse)
 - Each $f \in F$: factor node in FG (box)
 - For each argument R in $f \in F$: edge between variable node for R and factor node for f



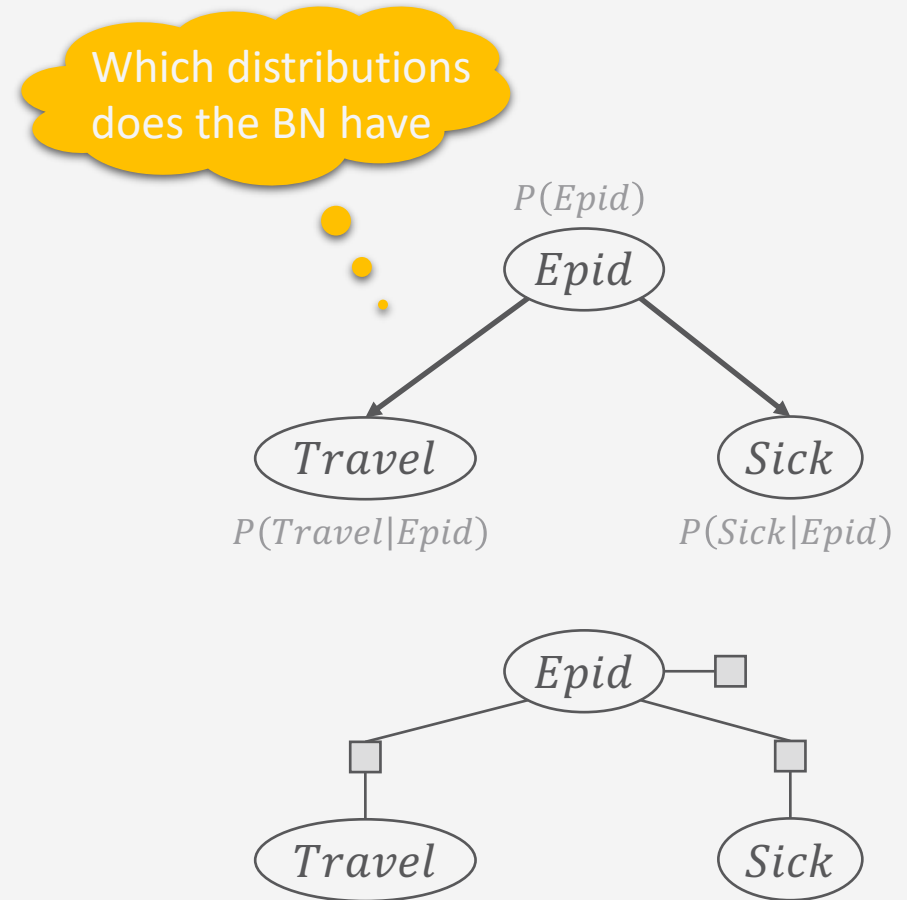
Other Model Representations

- **Markov network (MN)**
 - Alternative graphical representation of a factor-based model
 - Undirected graph
 - Factors: potential function for each clique in graph
 - If given a set of factors: add an edge between all random variables that occur together in a factor
 - With further information or the factors themselves, not clear what factors a model actually has just from an MN → **disadvantage!**
 - Semantics: Product of all factors, normalisation to get full joint
 - Neighbourhood directly defined between variables (not with factor nodes in between) → easier analysis (**advantage**)
- MNs and factor graphs have equivalent expressiveness



Other Model Representations

- **Bayesian network (BN)**
 - *Directed acyclic graph*
 - Explicit representation of (conditional) independences
 - Cannot model bidirectional influences → **disadvantage!**
 - Factors: set of probability distributions, one for each node
 - Prior probability tables for roots $R: P(R)$
 - Conditional probability tables (CPTs) for all other nodes N given its parents: $P(N|pa(N))$
 - Semantics
 - Product of tables, $Z = 1$ as tables are all probability distributions → **advantage!**
- Compared to undirected variants: independences readable in graph structure → **advantage!**

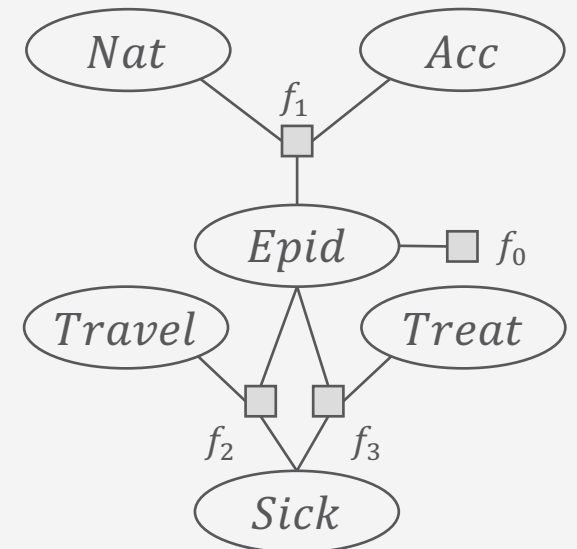


Space Complexity

- Given model $F = \{f_i\}_{i=1}^n$ over random variables \mathbf{R}
- Space complexity: $O(n \cdot r^k)$
 - $r = \max_{R \in \mathbf{R}} |\text{ran}(R)|$
 - $k = \max_{f \in F} |\text{rv}(f)|$
- Derivation:

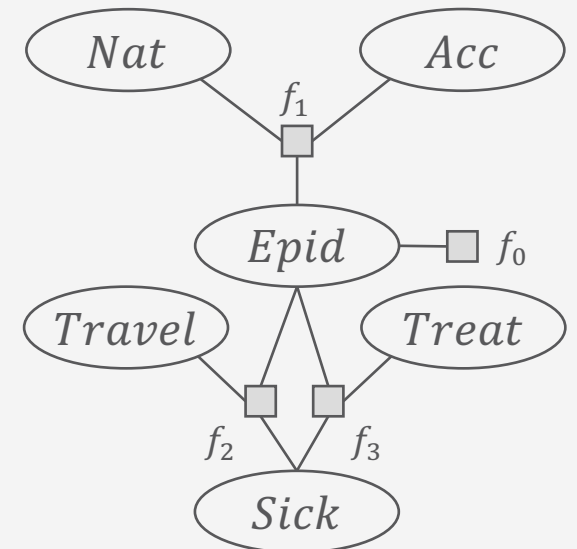
$$\sum_{f \in F} \prod_{R \in \text{rv}(f)} |\text{ran}(R)| \leq \sum_{f \in F} \prod_{R \in \text{rv}(f)} r = \sum_{f \in F} r^{|\text{rv}(f)|} \leq \sum_{f \in F} r^{\max_{f \in F} |\text{rv}(f)|} = \sum_{f \in F} r^k = n \cdot r^k$$

- No longer exponential in $N = |\mathbf{R}|$, but in k
 - If $k \ll N$, n not depending exponentially on N : $O(n \cdot r^k) \ll O(r^N)$



Compact Encoding for Faster Inference

- Use factorisation for query answering
 - Sum out all non-query variables from a product
 - Distributive law holds
 - Move factors from the inner sums outwards if inputs not affected by sum
 - Sum out variable from smaller sub-products
 - Basic idea of [variable elimination](#)
- Focus for the remainder of the lecture: queries $P(\mathbf{S} \mid \mathbf{t})$
 - Evidence \mathbf{t} = set of observations; may be empty: $\mathbf{t} = \emptyset \rightarrow P(\mathbf{S})$



Variable Elimination (VE)

- Outline:

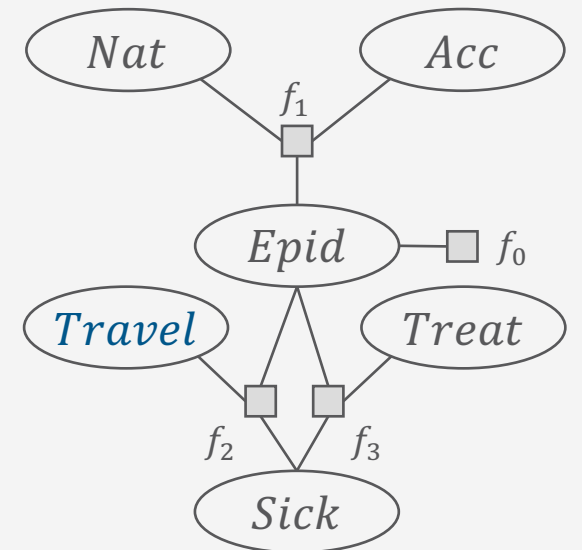
1. Absorb evidence \mathbf{t} in each factor covered by \mathbf{t} , i.e., $rv(f) \cap \mathbf{t} \neq \emptyset$,
2. Sum out non-query variables $\mathbf{U} = \mathbf{R} \setminus rv(\mathbf{S}, \mathbf{t})$ using factorisation in model F

$$\begin{aligned}
 P(\mathbf{S} \mid \mathbf{t}) &= \frac{1}{P(\mathbf{t})} \sum_{\mathbf{u} \in \text{ran}(\mathbf{U})} P_F(\mathbf{S}, \mathbf{t}, \mathbf{U} = \mathbf{u}) \\
 &= \frac{1}{P(\mathbf{t})} \sum_{\mathbf{u} \in \text{ran}(\mathbf{U})} \prod_{f \in F} \underbrace{\phi_f(R_1, \dots, R_k)}_{\pi_{rv(f)}(\mathbf{S}, \mathbf{t}, \mathbf{U} = \mathbf{u})}
 \end{aligned}$$

- Factor out factors from sums if arguments not covered by sum

3. Divide by $P(\mathbf{t}) = \text{Normalise } P(\mathbf{S}, \mathbf{t})$

- Example: $P(\textit{Travel})$ in $F = \{f_i\}_{i=0}^3$



Variable Elimination (VE): Example

 $P(\text{Travel})$

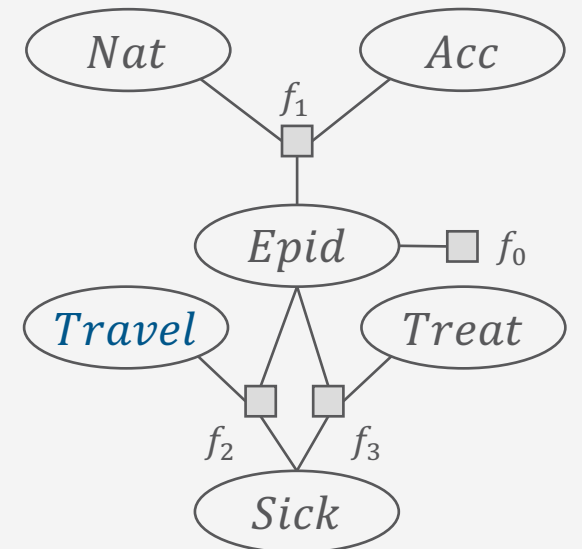
$$\propto \sum_{e \in \text{Val}(E)} \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \sum_{s \in \text{Val}(S)} \sum_{t \in \text{Val}(T)} P_R(E = e, N = n, A = a, S = s, \text{Travel}, T = t)$$

$$\propto \sum_{e \in \text{Val}(E)} \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \sum_{s \in \text{Val}(S)} \sum_{t \in \text{Val}(T)} \prod_{i=0}^3 \phi_i(\mathbf{R}_i = \mathbf{r}_i)$$

$$\propto \sum_{e \in \text{Val}(E)} \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \sum_{s \in \text{Val}(S)} \sum_{t \in \text{Val}(T)} \phi_0(e) \phi_1(e, n, a) \phi_2(\text{Travel}, e, s) \phi_3(e, s, t)$$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a) \sum_{s \in \text{Val}(S)} \phi_2(\text{Travel}, e, s) \sum_{t \in \text{Val}(T)} \phi_3(e, s, t)$$

Sums can be computed independently → could be done in parallel

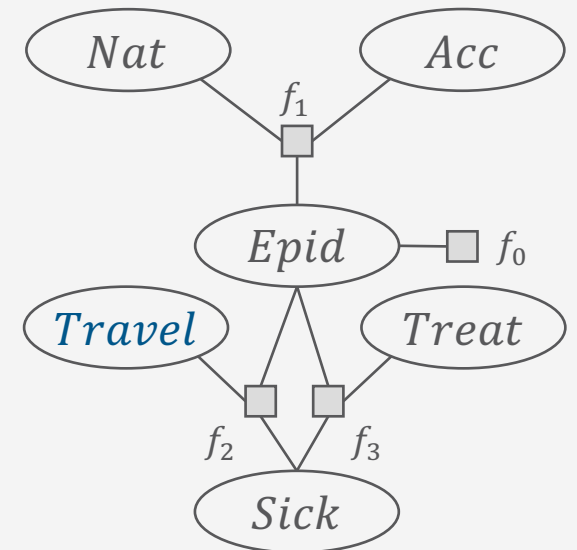
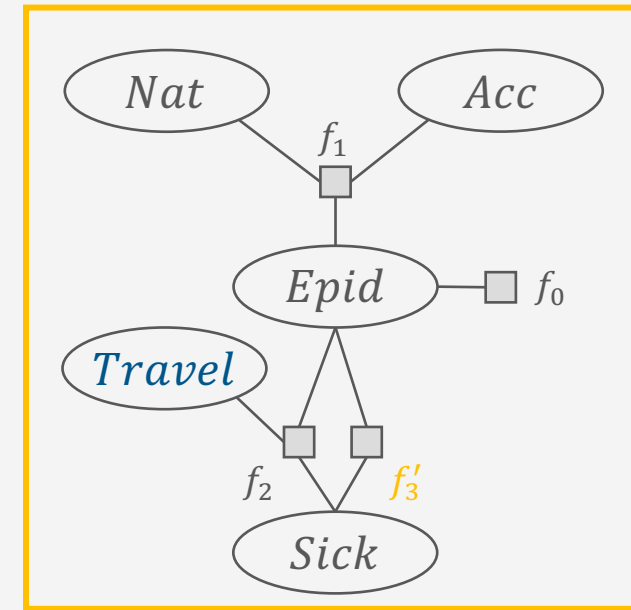


$P(\text{Travel})$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a) \sum_{s \in \text{Val}(S)} \phi_2(\text{Travel}, e, s) \sum_{t \in \text{Val}(T)} \phi_3(e, s, t)$$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a) \sum_{s \in \text{Val}(S)} \phi_2(\text{Travel}, e, s) \phi'_3(e, s)$$

<i>Epid</i>	<i>Sick</i>	<i>Treat</i>	ϕ_3		<i>Epid</i>	<i>Sick</i>	ϕ'_3
false	false	false	5	+	false	false	6
false	false	true	1		false	true	5
false	true	false	3	+	false	true	5
false	true	true	2		true	false	9
true	false	false	5	+	true	false	9
true	false	true	4		true	true	8
true	true	false	1	+	true	true	8
true	true	true	7				



$P(\text{Travel})$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a) \sum_{s \in \text{Val}(S)} \phi_2(\text{Travel}, e, s) \sum_{t \in \text{Val}(T)} \phi_3(e, s, t)$$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a) \sum_{s \in \text{Val}(S)} \phi_2(\text{Travel}, e, s) \phi'_3(e, s)$$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a) \sum_{s \in \text{Val}(S)} \phi_{23}(\text{Travel}, e, s)$$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \phi'_{23}(\text{Travel}, e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a)$$

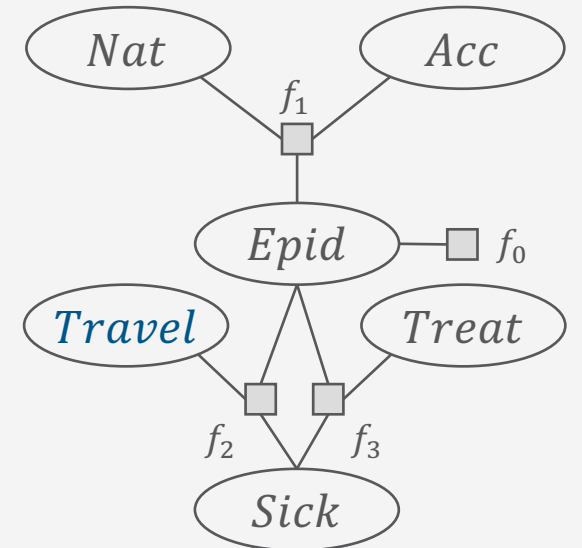
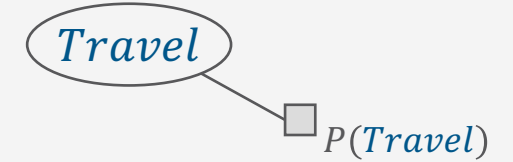
$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \phi'_{23}(\text{Travel}, e) \phi''_1(e) = \sum_{e \in \text{Val}(E)} \phi(\text{Travel}, e)$$

$$= \phi'(\text{Travel})$$

$$= \phi^n(\text{Travel}) = P(\text{Travel})$$

Intermediate results never larger than $2^3 < 2^6$

PGMs



VE with Evidence

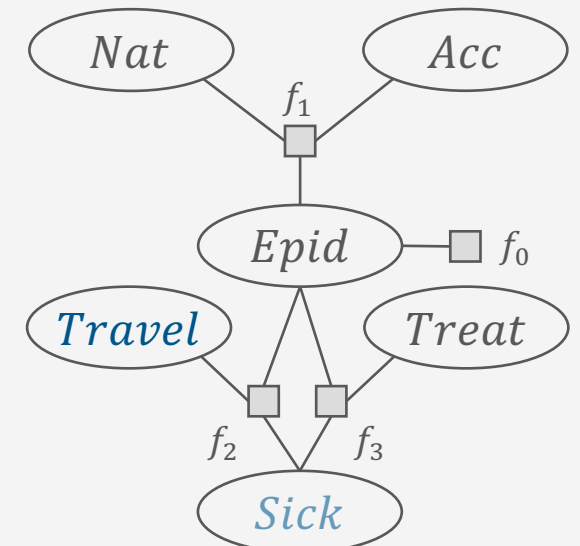
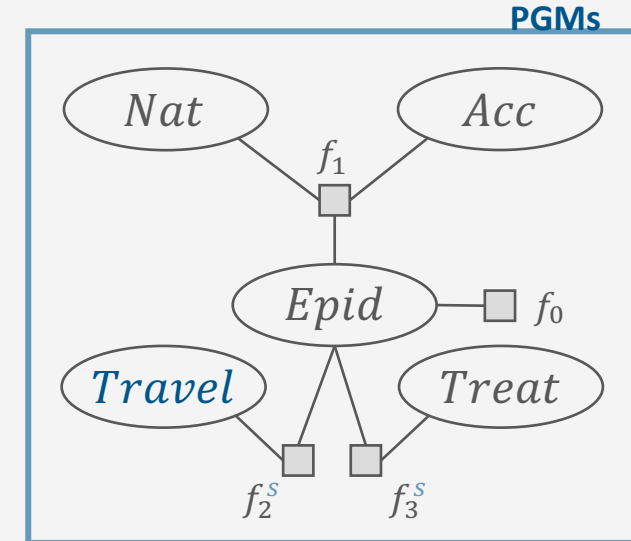
- Absorb each observation $t \in \mathbf{t}$ in each factor f , $t \cap rv(f) \neq \emptyset$
 - Drop rows with $T \neq t \rightarrow$ select all row with $T = t : f \leftarrow \sigma_{T=t}(f)$
 - Drop $T \rightarrow$ project result onto $rv(f) \setminus \{T\} : f \leftarrow \pi_{rv(f) \setminus \{T\}}(f)$
- Example: $P(\text{Travel}|\text{sick}) \rightarrow f_2, f_3$ have to absorb *sick*

Epid	Sick	Treat	ϕ_3
false	false	false	5
false	false	true	1
false	true	false	3
false	true	true	2
true	false	false	5
true	false	true	4
true	true	false	1
true	true	true	7

Epid	Treat	ϕ_3^S
false	false	3
false	true	2
true	false	1
true	true	7

Travel	Epid	Sick	ϕ_2
false	false	false	20
false	false	true	24
false	true	false	5
false	true	true	6
true	false	false	28
true	false	true	8
true	true	false	7
true	true	true	2

Travel	Epid	ϕ_2^S
false	false	24
false	true	6
true	false	8
true	true	2



VE with Evidence: Example

$P(\text{Travel}|\text{sick})$

$$\propto \sum_{e \in \text{Val}(E)} \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \sum_{t \in \text{Val}(T)} P_R(E = e, N = n, A = a, \text{sick}, \text{Travel}, T = t)$$

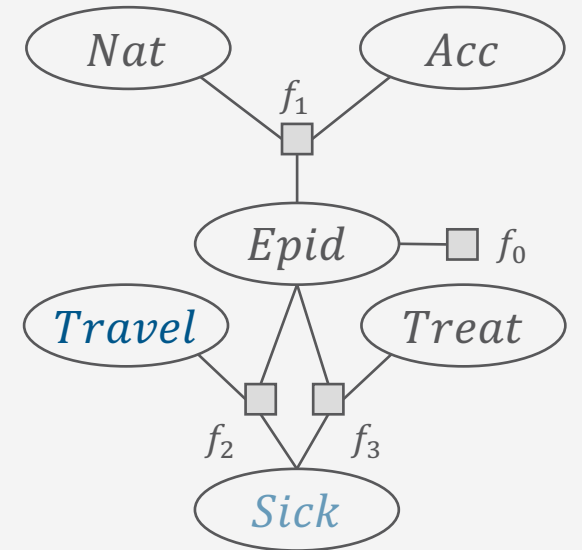
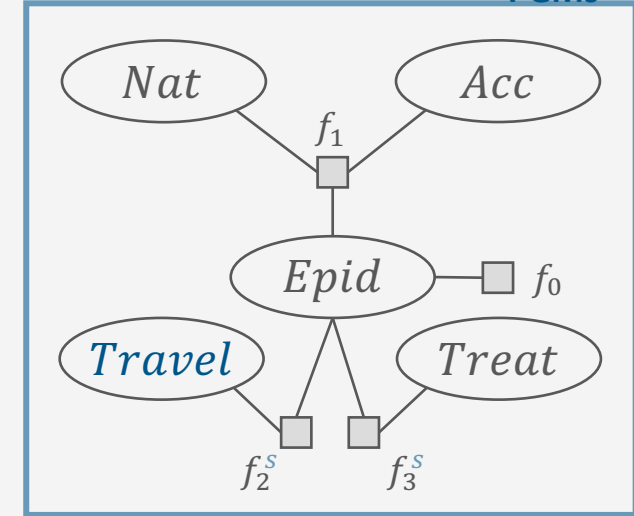
$$\propto \sum_{e \in \text{Val}(E)} \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \sum_{t \in \text{Val}(T)} \prod_{i=0}^3 \phi_i(R_i = r_i)$$

$$\propto \sum_{e \in \text{Val}(E)} \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \sum_{t \in \text{Val}(T)} \phi_0(e) \phi_1(e, n, a) \phi_2(\text{Travel}, e, \text{sick}) \phi_3(e, \text{sick}, t)$$

$$\propto \sum_{e \in \text{Val}(E)} \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \sum_{t \in \text{Val}(T)} \phi_0(e) \phi_1(e, n, a) \phi_2^S(\text{Travel}, e) \phi_3^S(e, t)$$

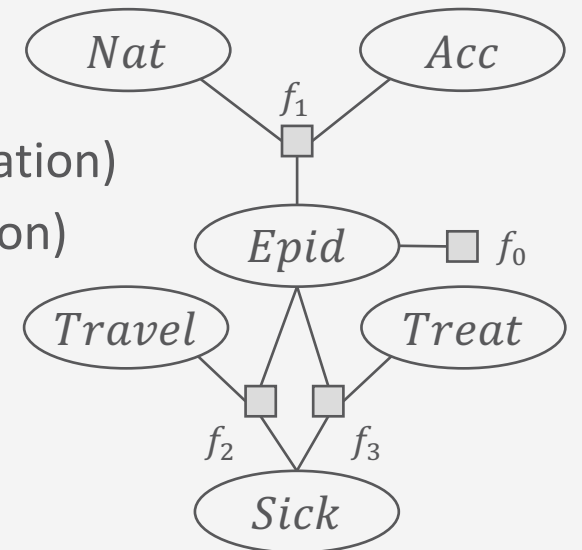
$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \phi_2^S(\text{Travel}, e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a) \sum_{t \in \text{Val}(T)} \phi_3^S(e, t)$$

PGMs



Elimination Order & Complexity

- Elimination order important
 - Wrong order \rightarrow large intermediate result: consider eliminating *Epid* first
 - Finding the best order not easier than inference using full joint
 - A lot of research has gone into finding a good order
- *Online greedy heuristic*: Variable R with smallest intermediate result
 - For each possible R to sum out
 - Collect all factors F_R containing R (would need to be multiplied before elimination)
 - Take number of arguments $|rv(F_R)|$ (Intermediate result size before elimination)
 - Decision criterion: $\arg \min_R |rv(F_R)| \rightarrow$ one-step VE simulation
- Complexity depends exponentially on largest intermediate result:
 - $O(N \cdot r^w)$, w called *tree width*



Interim Summary

- Inference tasks: Answer query for (conditional) marginal probability (distribution)
 - Full joint: Exponential dependence on number of random variables (space, runtime) $\rightarrow O(r^N)$
- Factorised model
 - Use (conditional) independences for factorisation: $P(\mathbf{R}, \mathbf{S} | \mathbf{T}) = P(\mathbf{R} | \mathbf{T}) \cdot P(\mathbf{S} | \mathbf{T})$
 - Independence: $\mathbf{T} = \emptyset$
 - Model = set of factors
 - Factor graphs, Markov network, Bayesian network, Markov properties (briefly)
 - Reduces space complexity $\rightarrow O(n \cdot r^k)$
- Variable elimination (VE): inference algorithm to solve query answering problems
 - Absorb evidence, multiply factors, sum-out variables
 - Good elimination order required (heuristics), complexity possibly reduced to $O(N \cdot r^w)$

Overview: 2. Foundations

A. *Logic*

- Propositional logic: alphabet, grammar, normal forms, rules
- First-order logic: introducing quantifiers, domain constraints

B. *Probability theory*

- Modelling: (conditional) probability distributions, random variables, marginal and joint distributions
- Inference: axioms and basic rules, Bayes theorem, independence

C. *Probabilistic graphical models*

- Syntax, semantics
- Inference problems

→ Probabilistic Relational Models (PRMs)

Appendix

Formal Definitions of Absorption, Multiplication, Summing out, as well as VE
Full VE Example Calculations

Absorption: Formal Definition

- Operator: **ABSORB**
 - Inputs:
 - Factor $f = \phi(R_1, \dots, R_n) \in F$
 - Variable $R \in \{R_1, \dots, R_n\}$ at position i
 - Factor $f^r = \phi(R)$ with mappings $r \mapsto 1$ and $\forall r' \neq r \in \text{ran}(R) : r' \mapsto 0$ for observation $R = r$,
 - Precondition: *none*
 - Output: Factor $\phi'(R_1, \dots, R_{i-1}, R_{i+1}, \dots, R_n)$
 - For all possible valuations $r_1, \dots, r_{i-1}, r_{i+1}, \dots, r_n$ of $R_1, \dots, R_{i-1}, R_{i+1}, \dots, R_n$
 - I.e., $r_1, \dots, r_{i-1}, r_{i+1}, \dots, r_n \in \text{ran}(R_1, \dots, R_{i-1}, R_{i+1}, \dots, R_n)$, with

$$\phi'(r_1, \dots, r_{i-1}, r_{i+1}, \dots, r_n) = \phi(r_1, \dots, r_{i-1}, r, r_{i+1}, \dots, r_n)$$

- Postcondition: $F \cup \{f^r\} \sim F \setminus \{f\} \cup \{f^r, \text{ABSORB}(f, R, f^r)\}$

Factor Multiplication: Formal Definition

- Operator: **MULTIPLY**
 - Inputs:
 - Factor $f_1 = \phi_1(R_1, \dots, R_n) \in F$
 - Factor $f_2 = \phi_2(S_1, \dots, S_m) \in F$
 - Precondition: *none*
 - Output: Factor $\phi(T_1, \dots, T_k)$
 - $\{(T_1, \dots, T_k)\} = \{(R_1, \dots, R_n)\} \bowtie \{(S_1, \dots, S_m)\}$ (ordered union)
 - For all possible valuations t_1, \dots, t_k of T_1, \dots, T_k , i.e., $t_1, \dots, t_k \in \text{ran}(T_1, \dots, T_k)$, with
 - $r_1, \dots, r_n = \pi_{R_1, \dots, R_n}(t_1, \dots, t_k)$ and $s_1, \dots, s_m = \pi_{S_1, \dots, S_m}(t_1, \dots, t_k)$ (select corresponding values from t_1, \dots, t_k)

$$\phi(t_1, \dots, t_k) = \phi_1(r_1, \dots, r_n) \cdot \phi_2(s_1, \dots, s_m)$$

- Postcondition: $F \sim F \setminus \{f_1, f_2\} \cup \text{MULTIPLY}(f_1, f_2)$

Summing out Variables: Formal Definition

- Operator: **SUM-OUT**
 - Inputs:
 - Factor $f = \phi(R_1, \dots, R_n) \in F$
 - Variable $R \in \{R_1, \dots, R_n\}$ at position i to sum out
 - Precondition: $\forall f' \in F \setminus \{f\}: R \notin \text{rv}(f')$
 - Output: Factor $\phi'(R_1, \dots, R_{i-1}, R_{i+1}, \dots, R_n)$
 - For each possible valuation $r_1, \dots, r_{i-1}, r_{i+1}, \dots, r_n$ of $R_1, \dots, R_{i-1}, R_{i+1}, \dots, R_n$
 - I.e., $r_1, \dots, r_{i-1}, r_{i+1}, \dots, r_n \in \text{ran}(R_1, \dots, R_{i-1}, R_{i+1}, \dots, R_n)$

$$\phi'(r_1, \dots, r_{i-1}, r_{i+1}, \dots, r_n) = \sum_{r \in \text{ran}(R)} \phi(r_1, \dots, r_{i-1}, r, r_{i+1}, \dots, r_n)$$

- Postcondition: $\sum_{r \in \text{ran}(R)} P_F \equiv P_{F \setminus \{f\} \cup \text{SUM-OUT}(f, R)}$

VE Algorithm Using a Heuristics h

VE($F, \mathcal{S}, \{\phi_t(T_t)\}_{t=1}^m, h$)

for $t = 1, \dots, m$ **do**

while $\exists f \in F : T_t \in rv(f)$ **do**

$F \leftarrow F \setminus \{f\} \cup \{\text{ABSORB}(f, T, \phi_t(T))\}$

Handle evidence

▸ Absorb $\phi_t(T_t)$ in F

while $rv(F) \setminus \mathcal{S} \neq \emptyset$ **do**

$U \leftarrow \arg \min_R h(F)$

while $\exists f_1, f_2 \in F : U \in rv(f_1) \wedge rv(f_2)$ **do**

$F \leftarrow F \setminus \{f_1, f_2\} \cup \{\text{MULTIPLY}(f_1, f_2)\}$

$F \leftarrow F \setminus \{f\} \cup \{\text{SUM-OUT}(f, U)\}$

Eliminate non-query variables

▸ Choose next U to eliminate

▸ Multiply f_1, f_2 in F

▸ Sum out U in F

while $\exists f_1, f_2 \in F$ **do**

$F \leftarrow F \setminus \{f_1, f_2\} \cup \{\text{MULTIPLY}(f_1, f_2)\}$

▸ Multiply f_1, f_2 in F , until $|F| = 1$

Normalise the potentials in the one remaining $f \in F$

Normalise

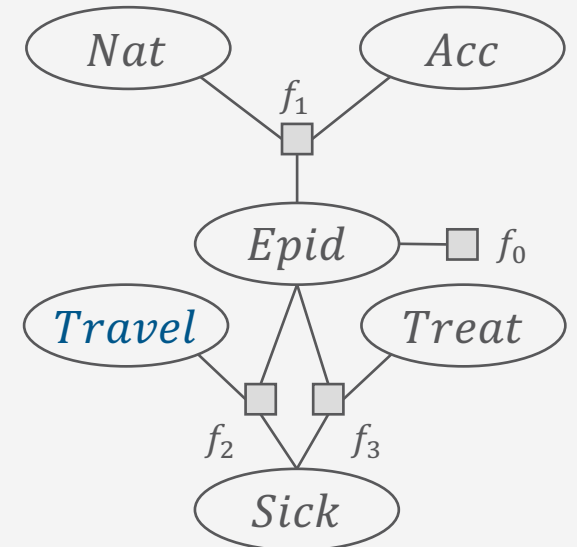
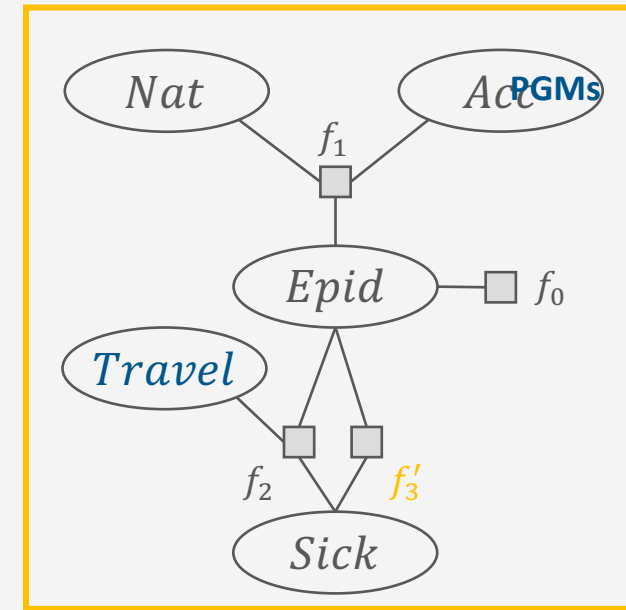
return f

$P(\text{Travel})$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a) \sum_{s \in \text{Val}(S)} \phi_2(\text{Travel}, e, s) \sum_{t \in \text{Val}(T)} \phi_3(e, s, t)$$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a) \sum_{s \in \text{Val}(S)} \phi_2(\text{Travel}, e, s) \phi'_3(e, s)$$

<i>Epid</i>	<i>Sick</i>	<i>Treat</i>	ϕ_3		<i>Epid</i>	<i>Sick</i>	ϕ'_3
false	false	false	5	+	false	false	6
false	false	true	1		false	true	5
false	true	false	3	+	false	true	5
false	true	true	2		true	false	9
true	false	false	5	+	true	false	9
true	false	true	4		true	true	8
true	true	false	1	+	true	true	8
true	true	true	7				

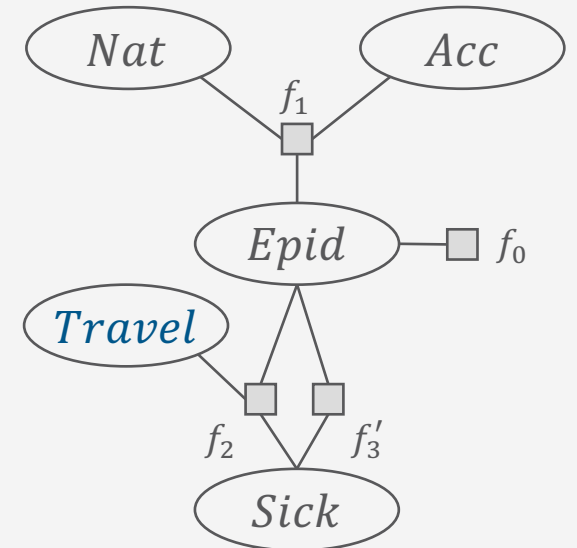
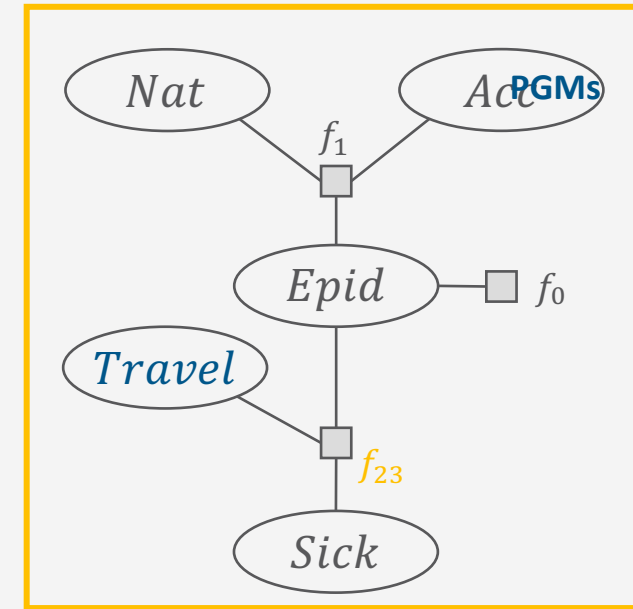


$P(\text{Travel})$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a) \sum_{s \in \text{Val}(S)} \phi_2(\text{Travel}, e, s) \phi'_3(e, s)$$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a) \sum_{s \in \text{Val}(S)} \phi_{23}(\text{Travel}, e, s)$$

<i>Travel</i>	<i>Epid</i>	<i>Sick</i>	ϕ_2	<i>Epid</i>	<i>Sick</i>	ϕ'_3	<i>Travel</i>	<i>Epid</i>	<i>Sick</i>	ϕ_{23}
false	false	false	20	false	false	6	false	false	false	$20 \cdot 6 = 120$
false	false	true	24	false	true	5	false	false	true	$24 \cdot 5 = 120$
false	true	false	5	true	false	9	false	true	false	$5 \cdot 9 = 45$
false	true	true	6	true	true	8	false	true	true	$6 \cdot 8 = 48$
true	false	false	28				true	false	false	$28 \cdot 6 = 168$
true	false	true	8				true	false	true	$8 \cdot 5 = 40$
true	true	false	7				true	true	false	$7 \cdot 9 = 63$
true	true	true	2				true	true	true	$2 \cdot 8 = 16$

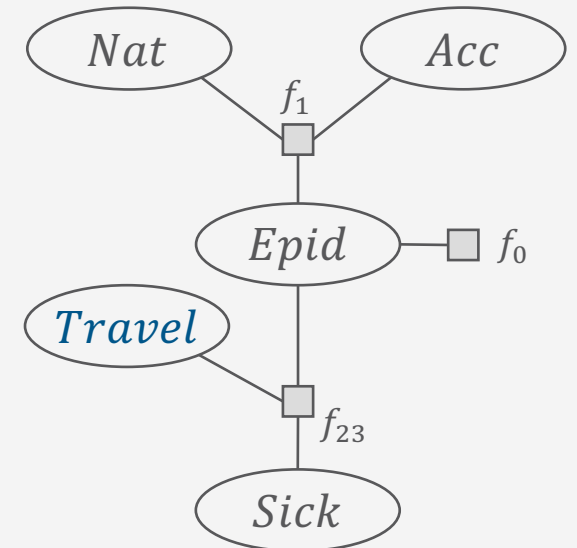
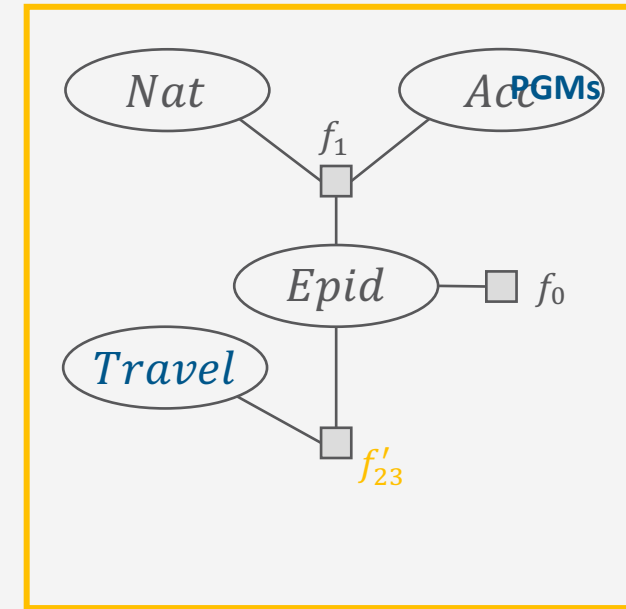


$P(\text{Travel})$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a) \sum_{s \in \text{Val}(S)} \phi_{23}(\text{Travel}, e, s)$$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a) \phi'_{23}(\text{Travel}, e)$$

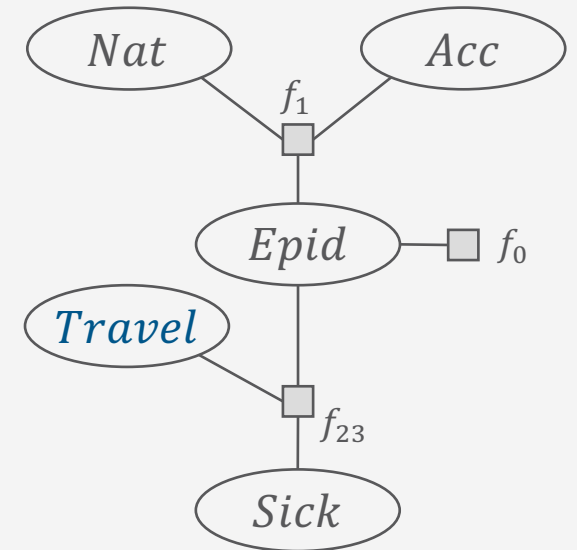
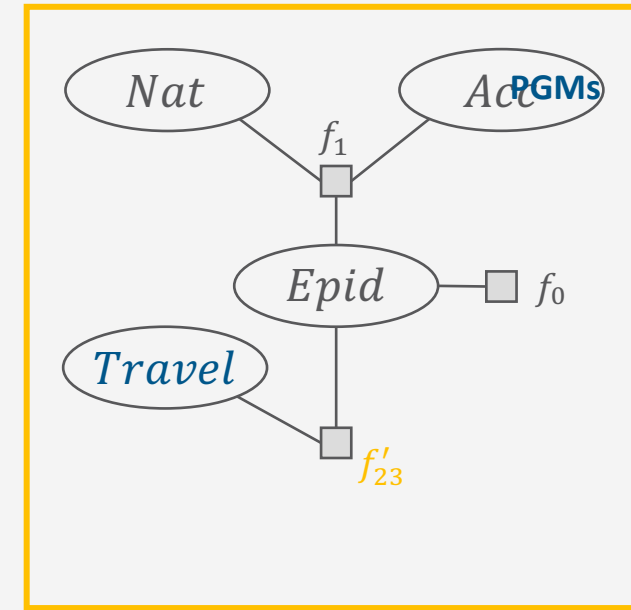
<i>Travel</i>	<i>Epid</i>	<i>Sick</i>	ϕ_{23}		<i>Travel</i>	<i>Epid</i>	ϕ'_{23}
false	false	false	120	+	false	false	240
false	false	true	120				
false	true	false	45	+	false	true	93
false	true	true	48				
true	false	false	168	+	true	false	208
true	false	true	40				
true	true	false	63	+	true	true	79
true	true	true	16				



$P(\text{Travel})$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a) \phi'_{23}(\text{Travel}, e)$$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \phi'_{23}(\text{Travel}, e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a)$$

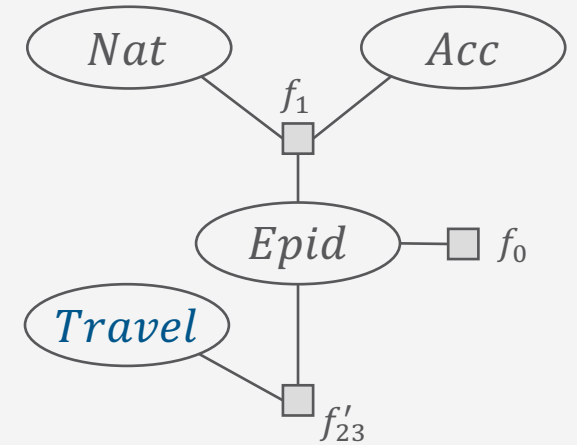
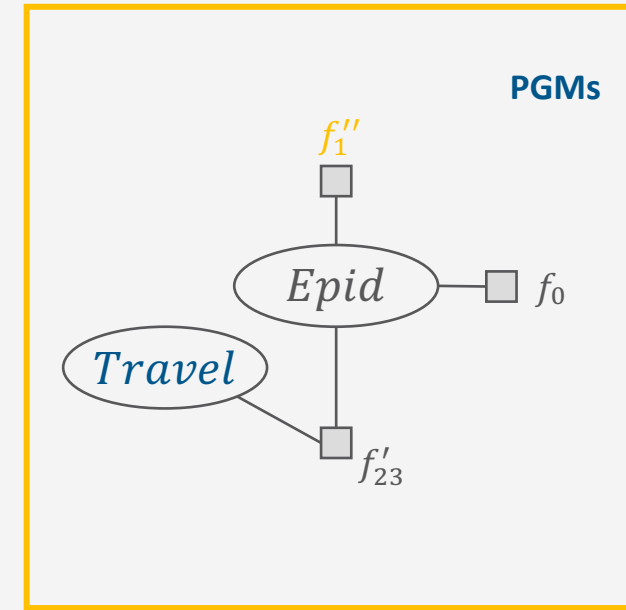


$P(\text{Travel})$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \phi'_{23}(\text{Travel}, e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a)$$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \phi'_{23}(\text{Travel}, e) \phi''_1(e)$$

Epid	NatDis	Artif	ϕ_1		Epid	NatDis	ϕ'_1		Epid	ϕ''_1
false	false	false	12	+	false	false	14	+	false	18
false	false	true	2		false	true	4			
false	true	false	3	+	true	false	11	+	true	17
false	true	true	1		true	true	6			
true	false	false	7							
true	false	true	4							
true	true	false	5							
true	true	true	1							

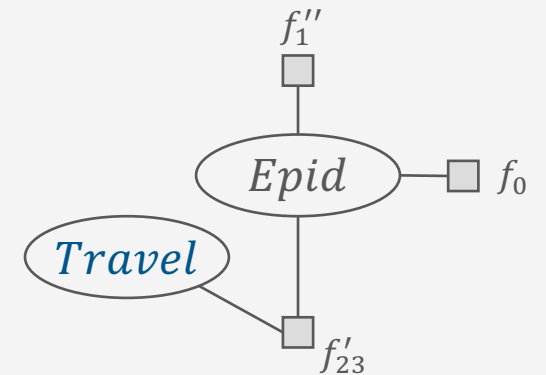
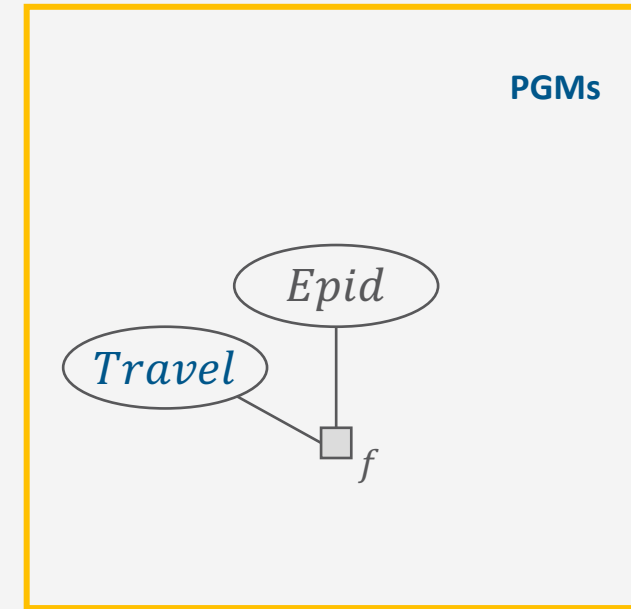


$P(\text{Travel})$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \phi'_{23}(\text{Travel}, e) \phi''_1(e)$$

$$\propto \sum_{e \in \text{Val}(E)} \phi(\text{Travel}, e)$$

Travel	Epid	ϕ'_{23}	Epid	ϕ''_1	Epid	ϕ_0	Travel	Epid	ϕ
false	false	240	false	18	false	50	false	false	$240 \cdot 18 \cdot 50 = 216,000$
false	true	93	true	17	true	1	false	true	$93 \cdot 17 \cdot 1 = 1,581$
true	false	208					true	false	$208 \cdot 18 \cdot 50 = 187,200$
true	true	79					true	true	$79 \cdot 17 \cdot 1 = 1,343$

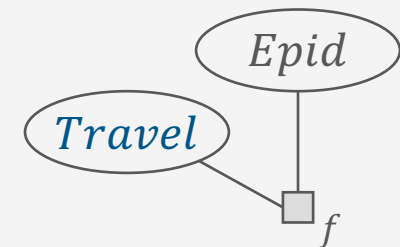
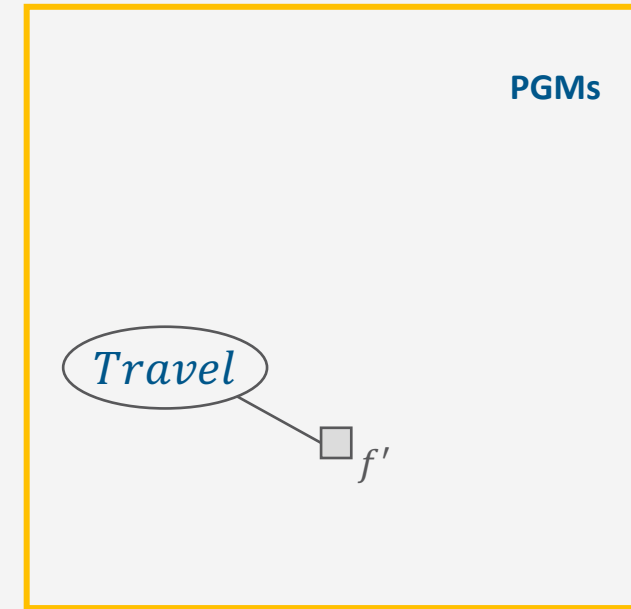


$P(\text{Travel})$

$$\propto \sum_{e \in \text{Val}(E)} \phi(\text{Travel}, e)$$

$$\propto \phi'(\text{Travel})$$

<i>Travel</i>	<i>Epid</i>	ϕ		<i>Travel</i>	ϕ'
<i>false</i>	<i>false</i>	216,000	+	<i>false</i>	217,581
<i>false</i>	<i>true</i>	1,581		<i>false</i>	217,581
<i>true</i>	<i>false</i>	187,200	+	<i>true</i>	188,543
<i>true</i>	<i>true</i>	1,343		<i>true</i>	188,543



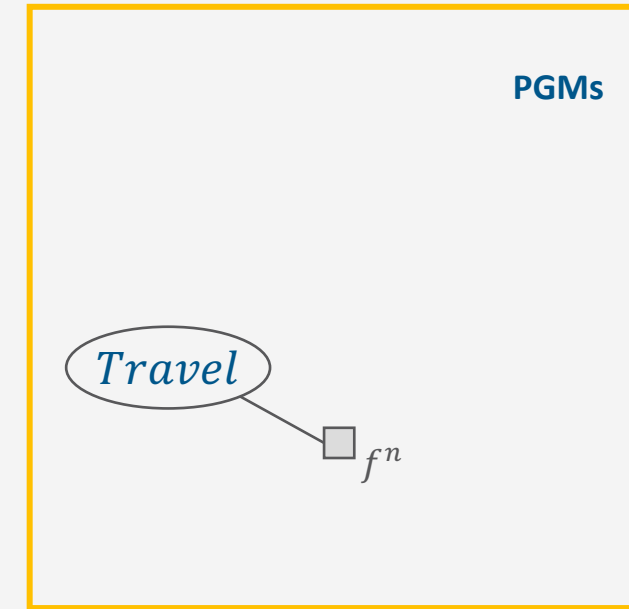
$P(\text{Travel})$

$$\propto \phi'(\text{Travel})$$

$$= \phi^n(\text{Travel})$$

$$= P(\text{Travel})$$

<i>Travel</i>	ϕ	<i>Travel</i>	ϕ^n
<i>false</i>	217,581	<i>false</i>	$\frac{217,581}{217,581 + 188,543} = \frac{217,581}{406,124} = 0.54$
<i>true</i>	188,543	<i>true</i>	$\frac{188,543}{217,581 + 188,543} = \frac{188,543}{406,124} = 0.46$



VE with Evidence: Example

$P(\text{Travel} \mid \text{sick})$

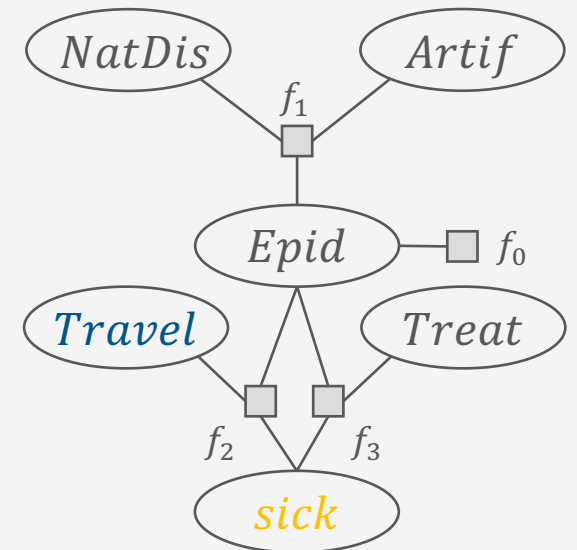
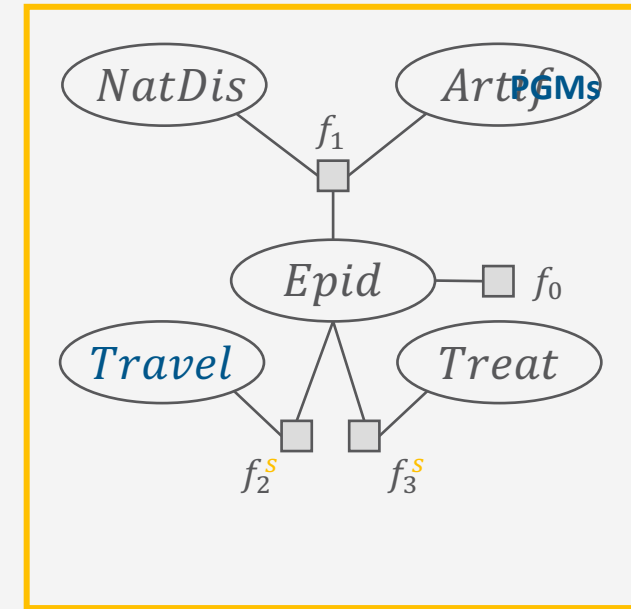
$$\propto \sum_{e \in \text{Val}(E)} \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \sum_{t \in \text{Val}(T)} P_R(E = e, N = n, A = a, \text{sick}, \text{Travel}, T = t)$$

$$\propto \sum_{e \in \text{Val}(E)} \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \sum_{t \in \text{Val}(T)} \prod_{i=0}^3 \phi_i(R_i = r_i)$$

$$\propto \sum_{e \in \text{Val}(E)} \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \sum_{t \in \text{Val}(T)} \phi_0(e) \phi_1(e, n, a) \phi_2(\text{Travel}, e, \text{sick}) \phi_3(e, \text{sick}, t)$$

$$\propto \sum_{e \in \text{Val}(E)} \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \sum_{t \in \text{Val}(T)} \phi_0(e) \phi_1(e, n, a) \phi_2^S(\text{Travel}, e) \phi_3^S(e, t)$$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \phi_2^S(\text{Travel}, e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a) \sum_{t \in \text{Val}(T)} \phi_3^S(e, t)$$



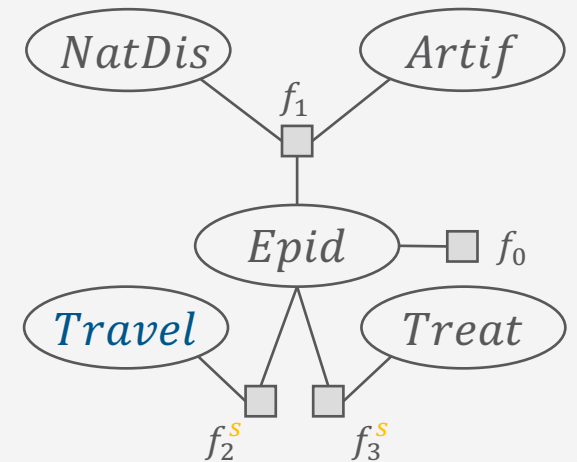
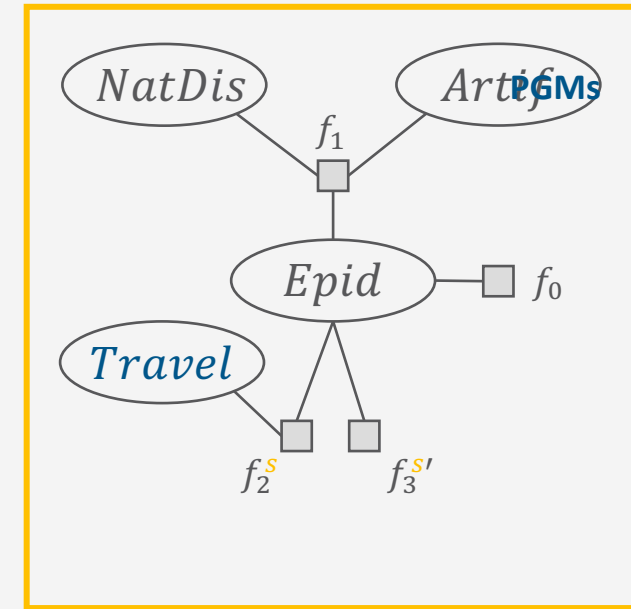
VE with Evidence: Example

$P(\text{Travel} \mid \text{sick})$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \phi_2^S(\text{Travel}, e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a) \sum_{t \in \text{Val}(T)} \phi_3^S(e, t)$$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \phi_2^S(\text{Travel}, e) \phi_3^{S'}(e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a)$$

Epid	Treat	ϕ_3^S		Epid	$\phi_3^{S'}$
false	false	3	+	false	5
false	true	2	+	true	8
true	false	1			
true	true	7			



VE with Evidence: Example

$$P(\text{Travel} \mid \text{sick})$$

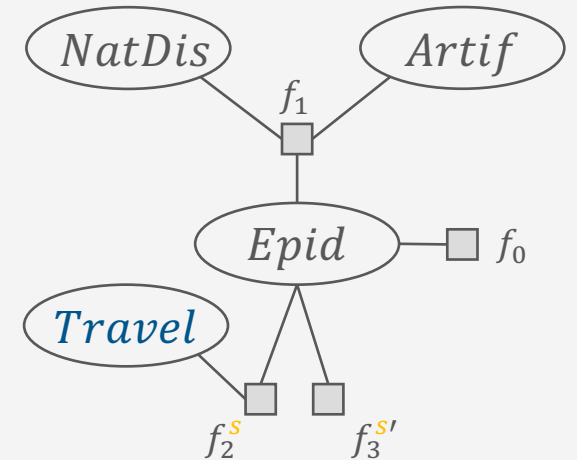
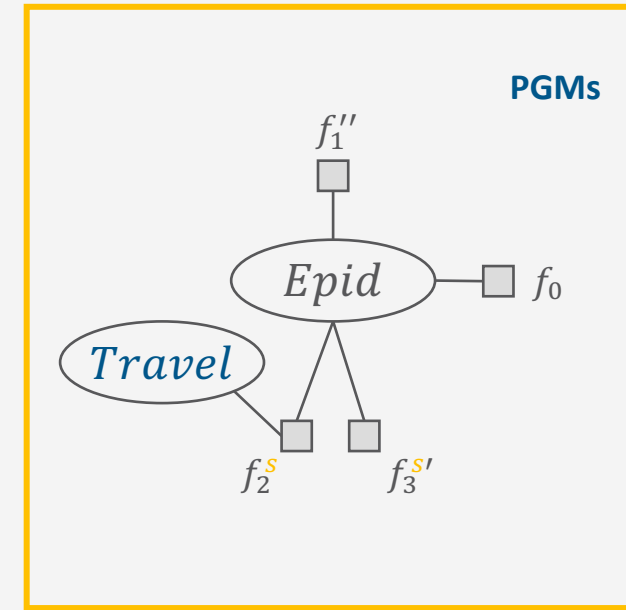
$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \phi_2^S(\text{Travel}, e) \phi_3^{S'}(e) \sum_{n \in \text{Val}(N)} \sum_{a \in \text{Val}(A)} \phi_1(e, n, a)$$

Wie vorher

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \phi_2^S(\text{Travel}, e) \phi_3^{S'}(e) \phi_1(e'')$$

Epid	NatDis	Artif	ϕ_1
false	false	false	12
false	false	true	2
false	true	false	3
false	true	true	1
true	false	false	7
true	false	true	4
true	true	false	5
true	true	true	1

Epid	ϕ_1''
false	18
true	17



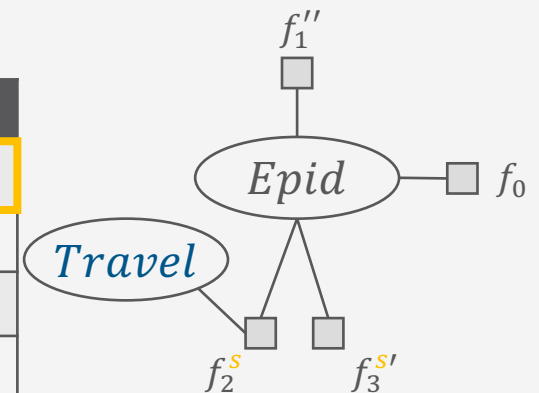
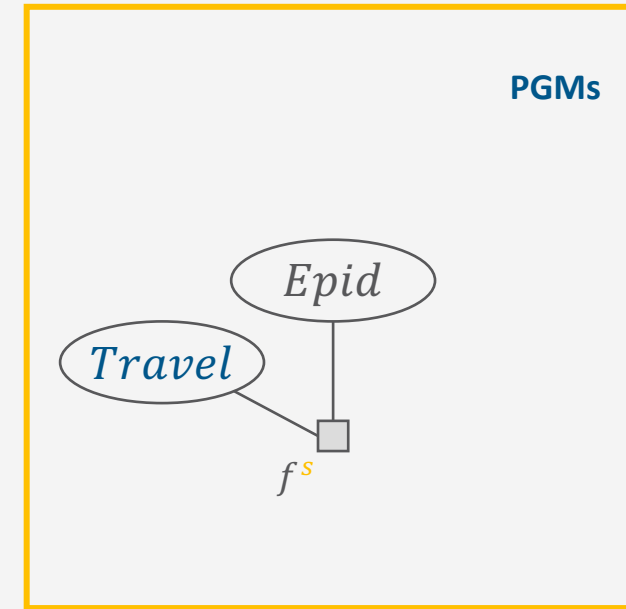
VE with Evidence: Example

$$P(\text{Travel} \mid \text{sick})$$

$$\propto \sum_{e \in \text{Val}(E)} \phi_0(e) \phi_2^S(\text{Travel}, e) \phi_3^{S'}(e) \phi_1''(e)$$

$$\propto \sum_{e \in \text{Val}(E)} \phi^S(\text{Travel}, e)$$

Travel	Epid	ϕ_2^S	Epid	ϕ_0	Epid	$\phi_3^{S'}$	Epid	ϕ_1''	Travel	Epid	ϕ
false	false	24	false	50	false	5	false	18	false	false	$24 \cdot 50 \cdot 5 \cdot 18 = 108.000$
false	true	6	true	1	true	8	true	17	false	true	$6 \cdot 1 \cdot 8 \cdot 17 = 816$
true	false	8							true	false	$8 \cdot 50 \cdot 5 \cdot 18 = 36.000$
true	true	2							true	true	$2 \cdot 1 \cdot 8 \cdot 17 = 272$



VE with Evidence: Example

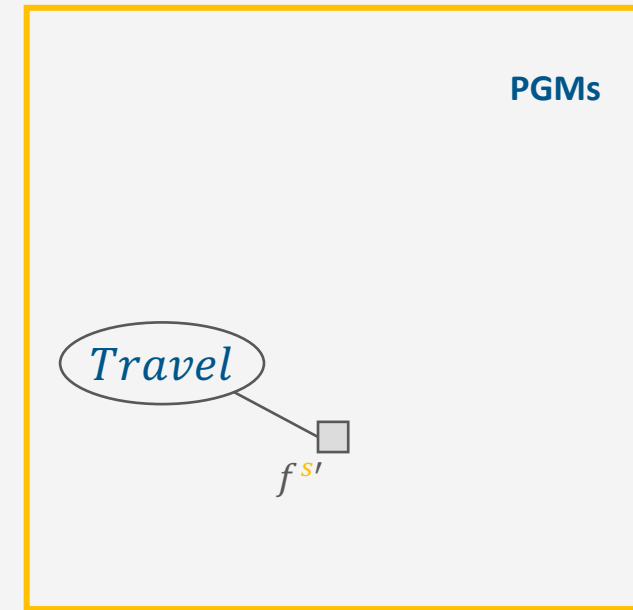
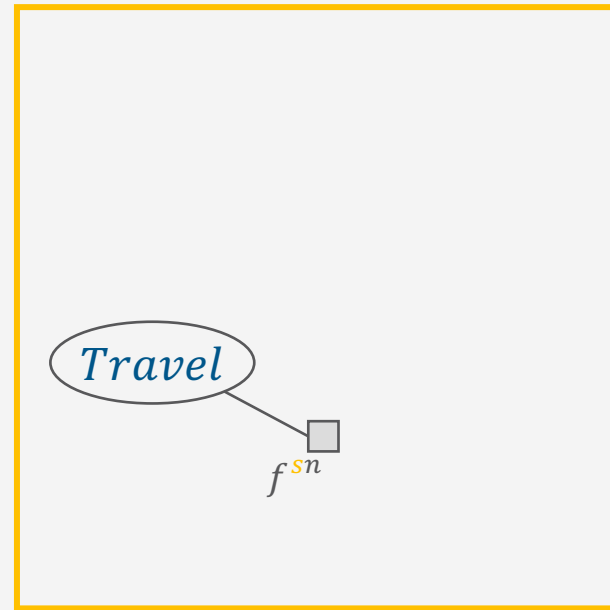
$$P(\text{Travel} \mid \text{sick})$$

$$\propto \sum_{e \in \text{Val}(E)} \phi^s(\text{Travel}, e)$$

$$\propto \phi^{s'}(\text{Travel})$$

$$= \phi^{sn}(\text{Travel})$$

$$= P(\text{Travel} \mid \text{sick})$$



Travel	Epid	ϕ^s
false	false	108.000
false	true	816
true	false	36.000
true	true	272

Travel	$\phi^{s'}$
false	108.816
true	36.272

Travel	ϕ^{sn}
false	$\frac{108.816}{108.816 + 36.272} = \frac{108.816}{145.088} = 0.75$
true	$\frac{36.272}{108.816 + 36.272} = \frac{36.272}{145.088} = 0.25$

