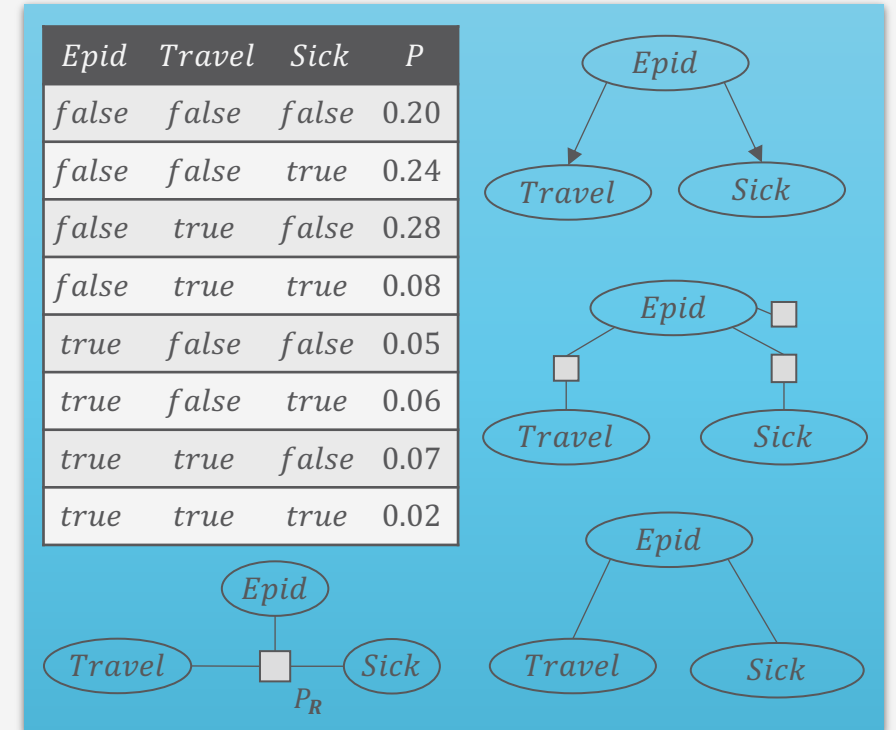


Modelling Formalisms: Probabilistic Graphical Models (PGMs)

Statistical Relational Artificial Intelligence
(StaRAI)



Contents

1. Introduction

- Notions of AI, agents, StaRAI

2. Modelling Formalisms

- a. *Logic*
- b. *Probabilistic Graphical Models (PGMs)*
- c. *Probabilistic Relational Models (PRMs)*
 - Syntax, semantics, inference tasks

3. Inference

- a. *Exact Inference*
 - i. *(Weighted) Model Counting*
 - ii. *Variable Elimination*
 - iii. *Lifted Inference*

b. *Approximate Inference*

- Inference algorithms

4. Learning

- Relation learning in logic
- Parameter / structure learning for PGMs
- From PGMs to PRMs, PRM learning

5. Sequential Models and Inference

- Syntax, semantics, inference tasks, algorithms

6. Decision-theoretic Models and Inference

- Preferences, utility
- Syntax, semantics, inference task, algorithm

Overview: 2. Modelling Formalisms

a. Logic

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

b. Probabilistic Graphical Models (PGMs)

- Syntax, semantics, inference problems
- Correctness

c. Probabilistic Relational Models (PRMs)

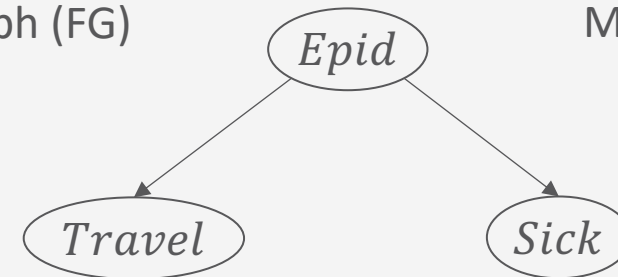
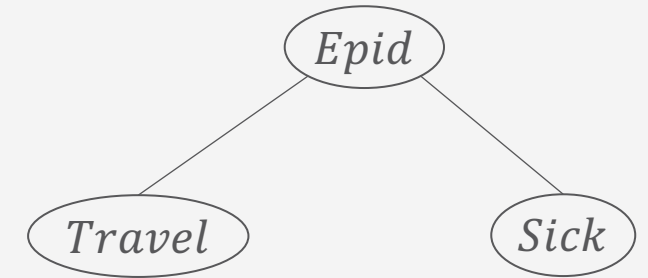
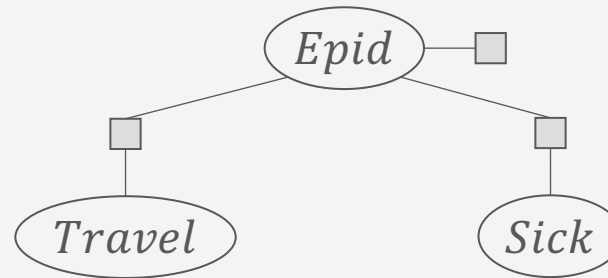
- Parfactor models (PMs): motivation, syntax, semantics, graphical representation, inference tasks
- Markov logic networks (MLNs): Syntax, semantics of MLNs, graphical representation
- Turning MLNs into PMs and vice versa

Probabilistic 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>
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

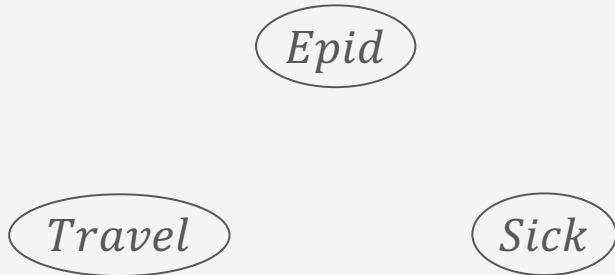


Bayesian network (BN)

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 or binary 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**

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

- 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.,

$Epid = e$	$Epid = \text{true}$	$Epid = \text{false}$
e	$epid$	$\neg epid$

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

Epid

Travel

Sick

Full Joint Probability Distribution

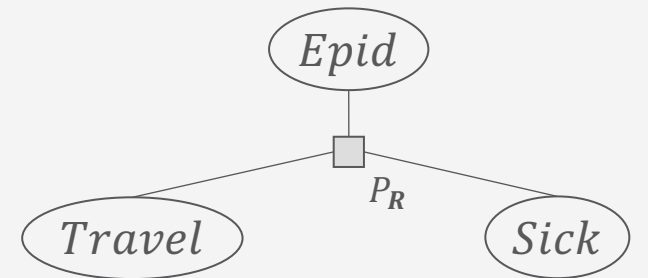
- 1 **world** = compound event for R

$$\begin{array}{l} epid \\ \neg travel \\ \neg sick \end{array}$$
- Specify a **probability** for a world

$$P(epid, \neg travel, \neg sick) = 0.05$$
- **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



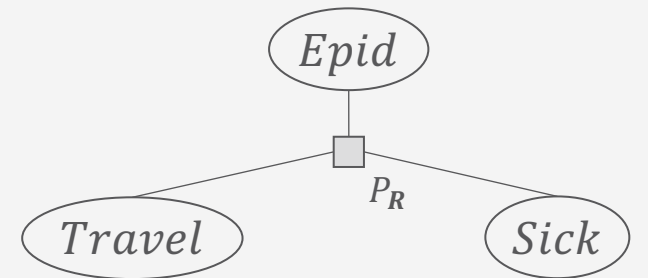
Space Complexity

- 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

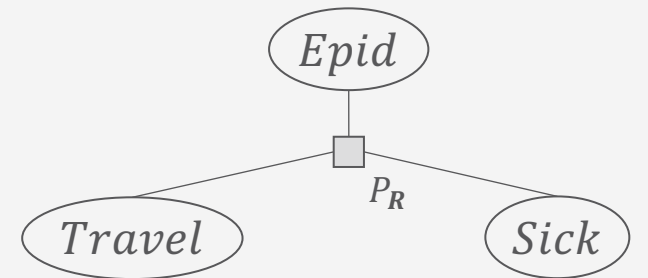
• Query Answering Problem

- 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

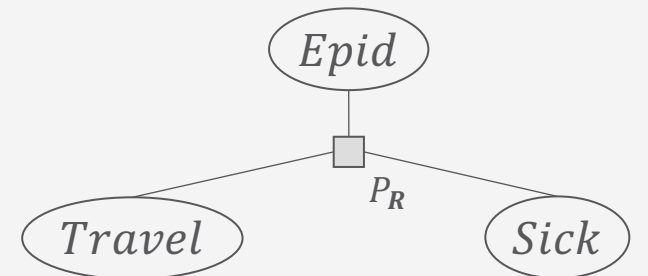
<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
false	false	false	0.20
false	false	true	0.24
false	true	false	0.28
false	true	true	0.08
true	false	false	0.05
true	false	true	0.06
true	true	false	0.07
true	true	true	0.02



Query Syntax

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

<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



Answering Marginal Queries

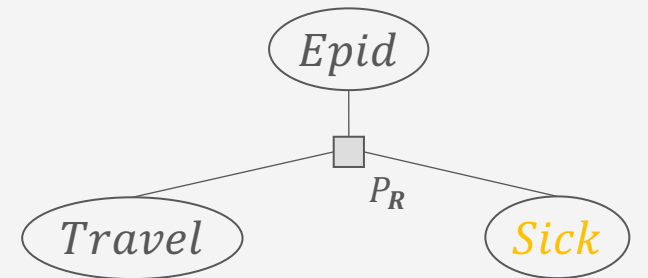
- Solving $P(\mathcal{S})$: Eliminate all non-query terms $\mathbf{U}$
 - $\mathbf{U} = \mathbf{R} \setminus \text{rv}(\mathcal{S})$

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

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

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

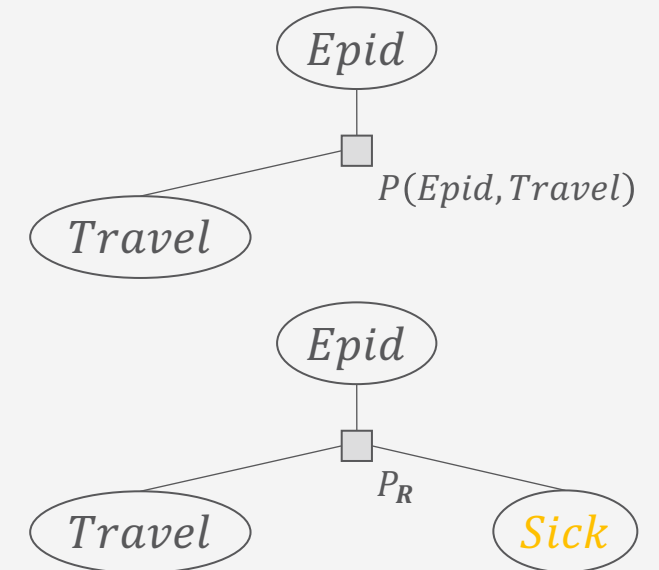
<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
false	false	false	0.20
false	false	true	0.24
false	true	false	0.28
false	true	true	0.08
true	false	false	0.05
true	false	true	0.06
true	true	false	0.07
true	true	true	0.02



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>
false	false	false	0.20	+	false	false	0.44
false	false	true	0.24				
false	true	false	0.28	+	false	true	0.36
false	true	true	0.08				
true	false	false	0.05	+	true	false	0.11
true	false	true	0.06				
true	true	false	0.07	+	true	true	0.09
true	true	true	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 $\mathbf{U} = \{\textit{Travel}, \textit{Sick}\}$

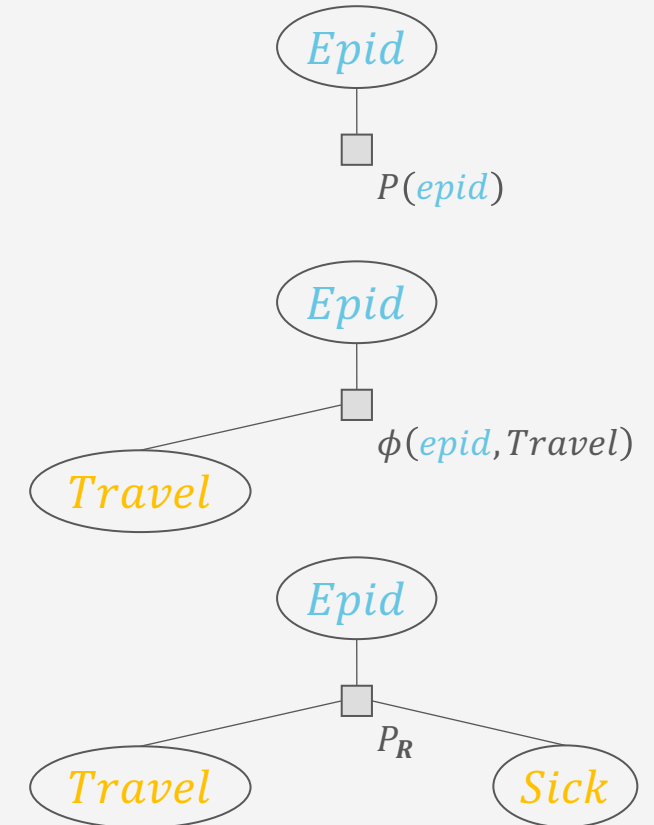
$$P(\textit{epid}) = \sum_{v_t \in \text{ran}(\textit{Travel})} \sum_{v_s \in \text{ran}(\textit{Sick})} P(\textit{epid}, \textit{Travel} = v_t, \textit{Sick} = v_s)$$

<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

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



Answering Conditional Queries

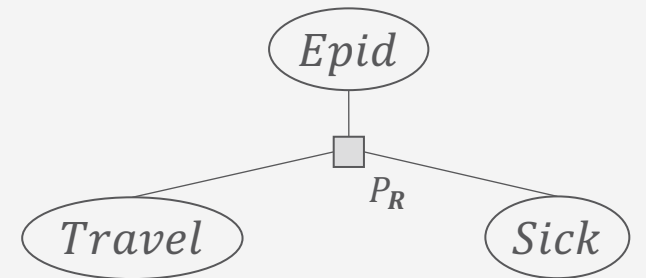
- Solving $P(\mathcal{S}|\mathcal{T})$ w.r.t. P_R :
 - $P(\mathcal{S}|\mathcal{T}) = \frac{P(\mathcal{S}, \mathcal{T})}{P(\mathcal{T})}$
 - $P(\mathcal{T})$ normalising constant
 - Reduces to computing two marginal queries: $P(\mathcal{S}, \mathcal{T})$, $P(\mathcal{T})$
 - Eliminate all non-query terms $\mathcal{U}$ and **normalise**
 - $\mathcal{U} = \mathcal{R} \setminus \text{rvs}(\mathcal{S}, \mathcal{T})$
- $$P(\mathcal{S}|\mathcal{T}) = \frac{1}{P(\mathcal{T})} \sum_{u \in \text{ran}(\mathcal{U})} P_R(\mathcal{S}, \mathcal{T}, \mathcal{U} = u)$$
- 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:

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

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

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



Answering Conditional Queries: Normalisation

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

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

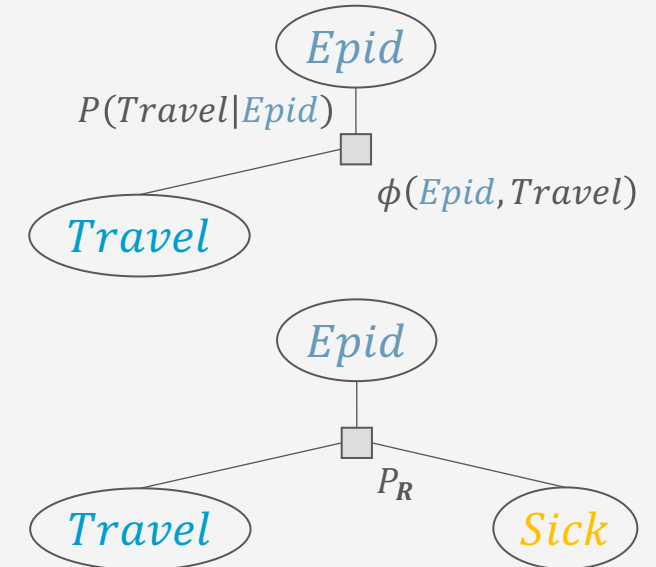
- for each $e \in \text{ran}(\text{Epid})$: for each $t \in \text{ran}(\text{Travel})$:
divide $\phi(e, t)$ by $P(e) = \phi(e, t) + \phi(e, \neg t)$

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	$1 - 0.55 = 0.45$



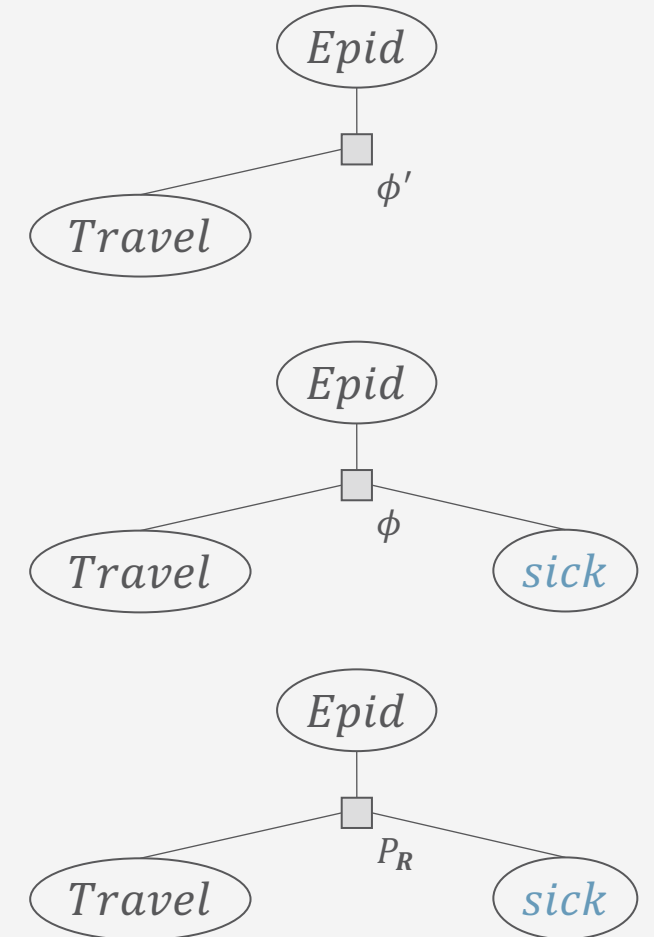
Answering Conditional Queries: Evidence – Absorption

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

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

<i>Epid</i>	<i>Travel</i>	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	



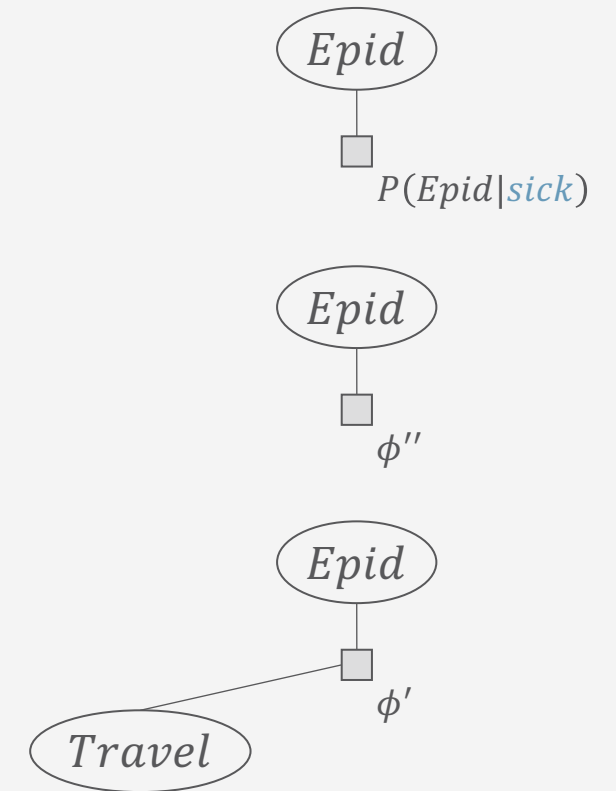
Answering Conditional Queries: Evidence – Normalisation

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

normalise

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

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



Runtime Complexity

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

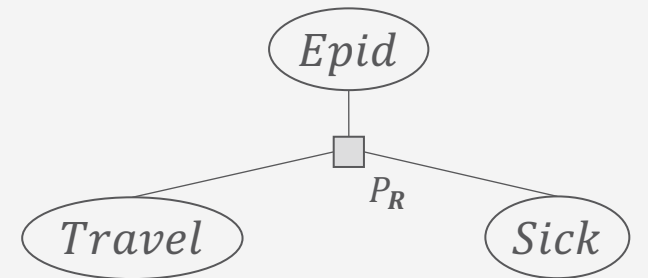
- Runtime complexity: $O(r^N)$
 - $r = \max_{R \in \mathcal{R}} |\text{ran}(R)|$
 - $N = |\mathcal{R}|$
 - Have to go through whole table; derivation as before

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

= Space complexity

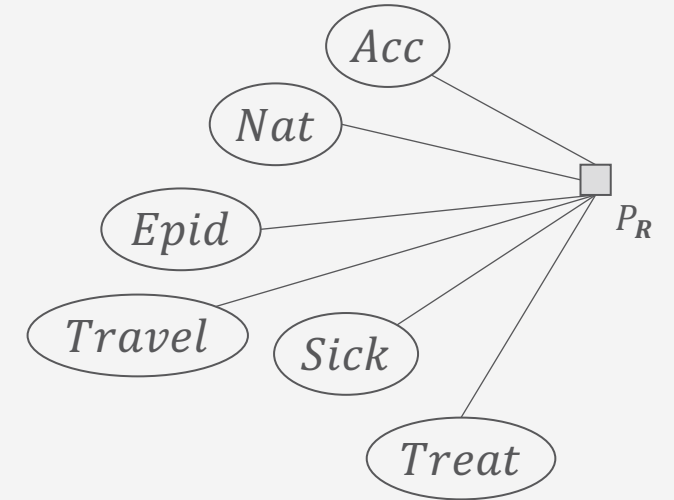
Exponential in N !

<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>
false	false	false	false	false	false	0.025
false	false	false	false	false	true	0.009
false	false	false	false	true	false	0.0
false	false	false	false	true	true	0.0
false	false	false	true	false	false	0.0
false	false	false	true	true	false	0.0
false	false	true	false	false	false	0.0
false	false	true	true	false	false	0.0
false	true	false	false	false	false	0.0
false	true	false	false	true	false	0.0
false	true	false	true	false	false	0.0
false	true	false	true	true	false	0.0
false	true	true	false	false	false	0.0
false	true	true	true	false	false	0.0
false	true	true	true	true	false	0.0
true	false	false	false	false	false	0.0
true	false	false	false	true	false	0.0
true	false	false	true	false	false	0.0
true	false	false	true	true	false	0.0
true	false	true	false	false	false	0.0
true	false	true	true	false	false	0.0
true	false	true	true	true	false	0.0
true	true	false	false	false	false	0.0
true	true	false	false	true	false	0.0
true	true	false	true	false	false	0.0
true	true	false	true	true	false	0.0
true	true	true	false	false	false	0.0
true	true	true	true	false	false	0.0
true	true	true	true	true	false	0.0

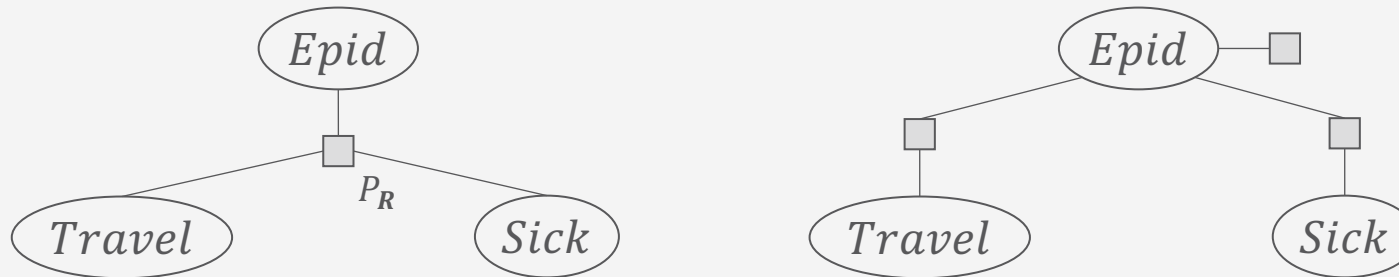


$2^6 = 64$ possible worlds

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

Compact Encoding

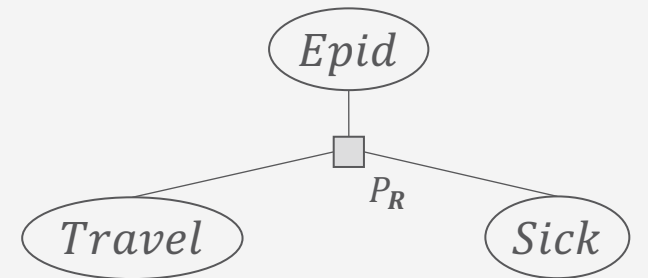
- Full joint probability distribution:
Every random variable is connected with every other random variable!
- Factorise full joint probability distribution P_R using (conditional) independences
 - Independence: $P(R_1, R_2) = P(R_1) \cdot P(R_2) \rightarrow$ denoted $(R_1 \perp R_2)$
 - Conditional independence: $P(R_1, R_2 | R_3) = P(R_1 | R_3) \cdot P(R_2 | R_3) \rightarrow$ denoted $(R_1 \perp R_2 | R_3)$
- Hidden in P_R
 - Explicitly represent through factors and in graph
 - Full joint is then given by *product* of the factors



Independences: Examples

- Independence: ($Travel \perp Sick$)?
 - $P(Travel, Sick) \stackrel{?}{=} P(Travel) \cdot P(Sick)$
 → Consider [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>
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



Independences: Examples

- $Travel \perp Sick \mid Epid$?

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

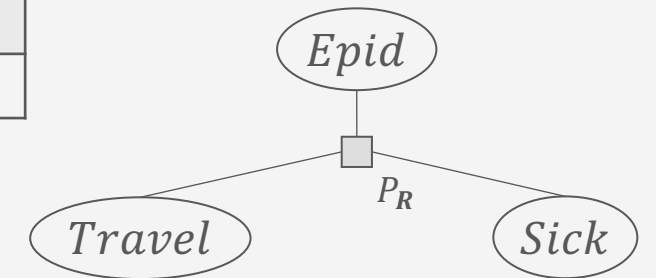
<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
false	false	false	0.6375
false	false	true	0.1125
false	true	false	0.2125
false	true	true	0.0375
true	false	false	0.45
true	false	true	0.15
true	true	false	0.3
true	true	true	0.1

<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

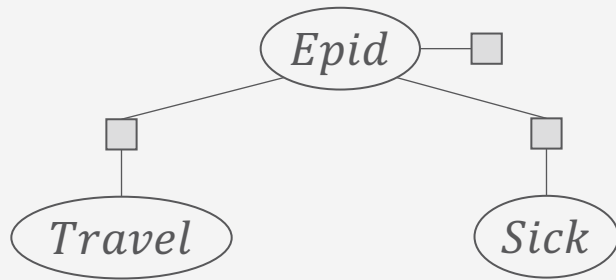
<i>P</i>
$0.75 \cdot 0.85 = 0.6375$
$0.75 \cdot 0.15 = 0.1125$
$0.25 \cdot 0.85 = 0.2125$
$0.25 \cdot 0.15 = 0.0375$
$0.6 \cdot 0.75 = 0.45$
$0.6 \cdot 0.25 = 0.15$
$0.4 \cdot 0.75 = 0.3$
$0.4 \cdot 0.25 = 0.1$

<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:
 - $P(Epid, Travel, Sick) = P(Epid) \cdot P(Travel|Epid) \cdot P(Sick|Epid)$



<i>Epid</i>	<i>P</i>
false	0.8
true	0.2

$P(epid)$
0.2

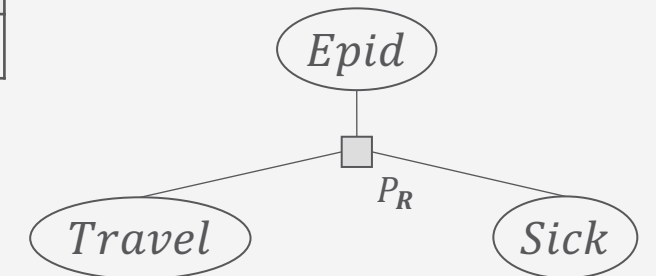
<i>Epid</i>	<i>Travel</i>	<i>P</i>
false	false	0.75
false	true	0.25
true	false	0.6
true	true	0.4

<i>Epid</i>	$P(travel Epid)$
false	0.25
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>Epid</i>	$P(sick Epid)$
false	0.15
true	0.25

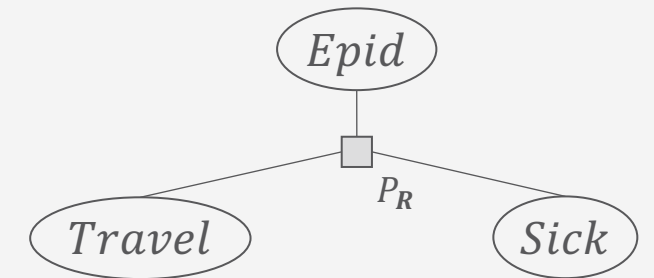
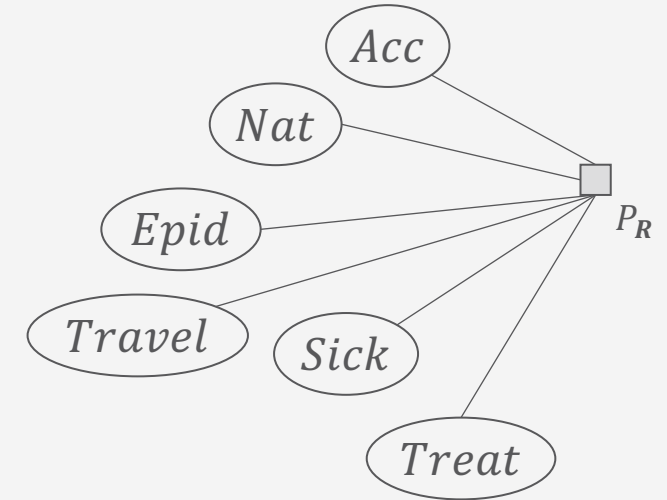
<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
false	false	false	0.51
false	false	true	0.09
false	true	false	0.17
false	true	true	0.03
true	false	false	0.09
true	false	true	0.03
true	true	false	0.06
true	true	true	0.02



- 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 4: Learning*)
- For now, assume that we have a factorised model



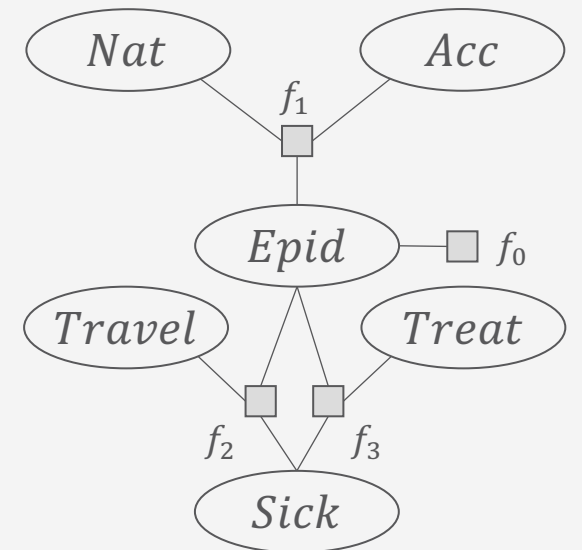
Model Representation: Factors

- 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

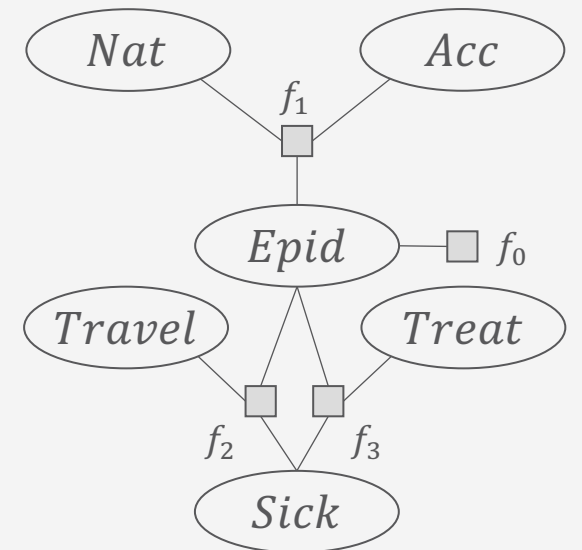
- Given model $F = \{f_i\}_{i=1}^n$ over random variables $\mathbf{R} = \{R_1, \dots, R_N\}$
 - $f_i = \phi_i(R_1, \dots, R_k)$

Semantics:

- Build full joint probability distribution P_F

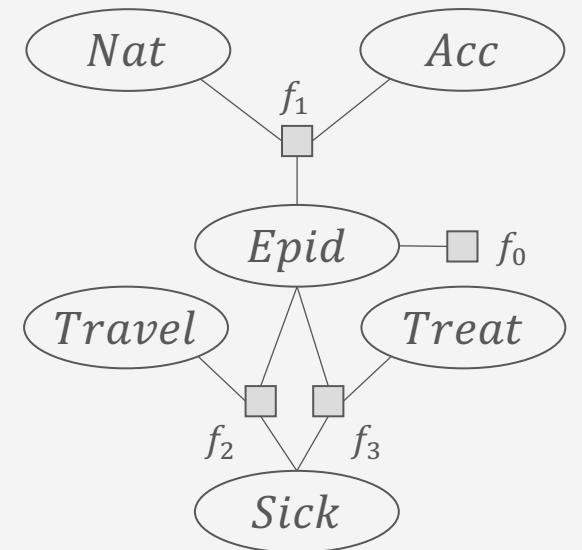
$$P_F = \frac{1}{Z} \prod_{i=1}^n \phi_i(R_1, \dots, R_k)$$

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



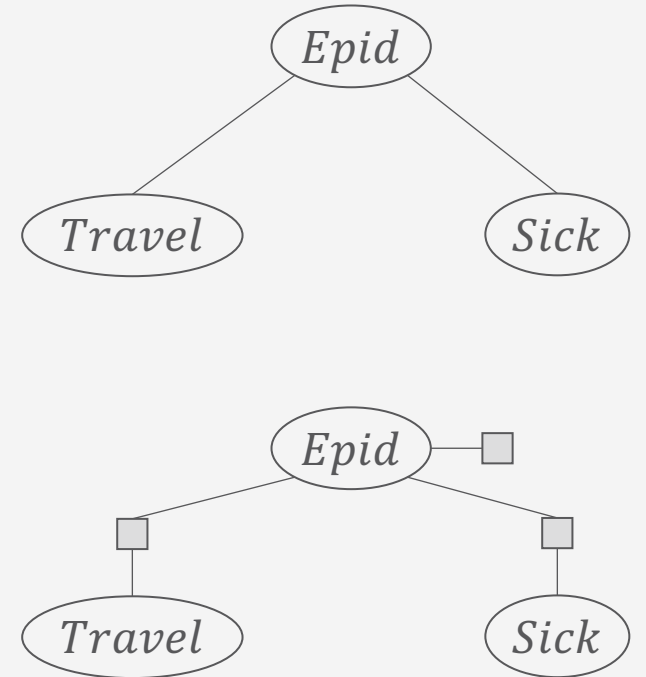
Model Representation: Factors

- 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



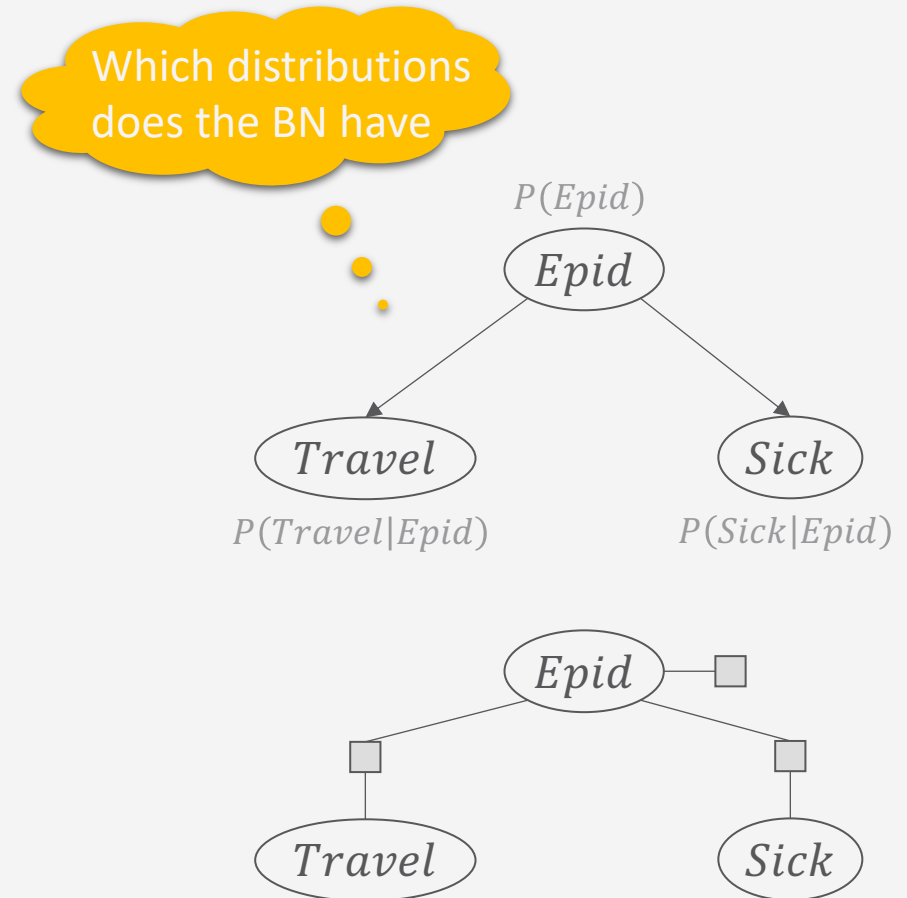
Other Model Representations

- **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 $\text{Pa}(N)$: $P(N|\text{Pa}(N))$
 - **Semantics**
 - Product of tables, $Z = 1$ as tables are all probability distributions → **advantage!**

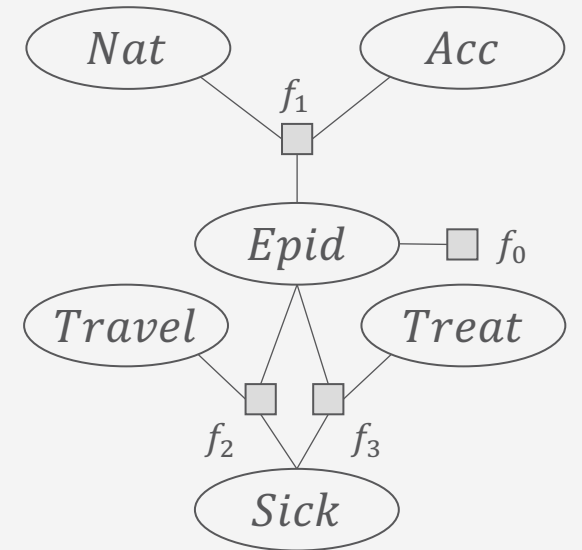


Space Complexity

- 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:

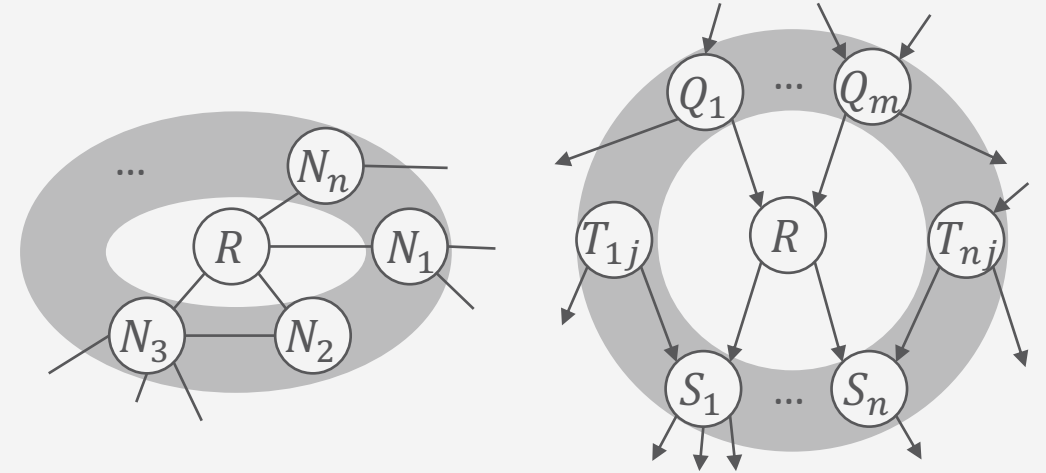
$$\sum_{f \in F} \prod_{R \in \text{rv}(f)} |\text{ran}(R)| \leq \sum_{f \in F} \prod_{R \in \text{rv}(f)} r = \sum_{f \in F} r^{|\text{rvs}(f)|} \leq \sum_{f \in F} r^{\max_{f \in F} |\text{rvs}(f)|} = \sum_{f \in F} r^k = n \cdot r^k$$

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



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 $\rightarrow O(n \cdot r^k)$



Correctness

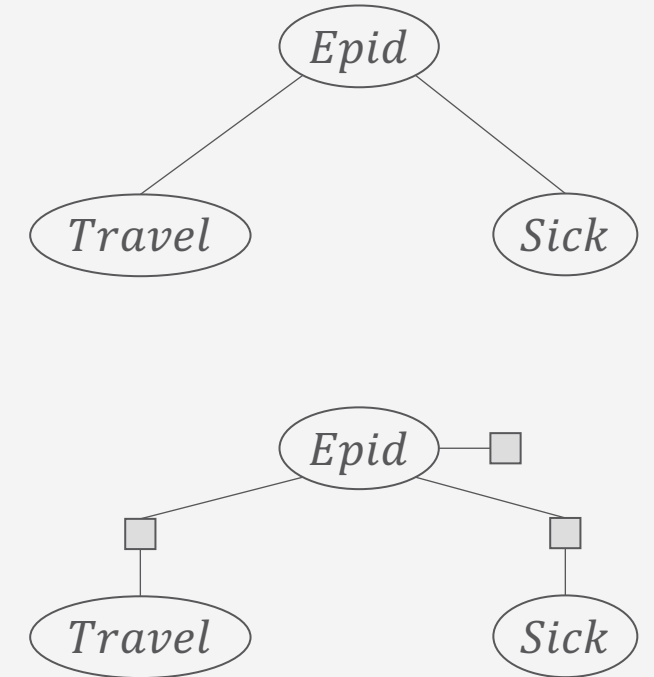
Why can we represent the factorised distribution in a graph structure with cliques over factors?

Conditional Independences in the Graph Structure

- Graphs allows for reading of (conditional) independences
- Given a model F over random variables $\mathbf{R}$
- Local independences in F
Each $R \in \mathbf{R}$ is independent of all other random variables given its neighbourhood, i.e.,

$$\{(R \perp R' \mid \text{Nb}(R)) \mid R \in \mathbf{R} \wedge R' = (\mathbf{R} \setminus \{R\}) \setminus \text{Nb}(R)\}$$
- Neighbourhood blocks all paths to non-neighbours
- Neighbourhood also called Markov blanket: $\text{MB}(R) = \text{Nb}(R)$
 - With analogue definition of local independences:

$$\{(R \perp R' \mid \text{MB}(R)) \mid R \in \mathbf{R} \wedge R' = (\mathbf{R} \setminus \{R\}) \setminus \text{MB}(R)\}$$

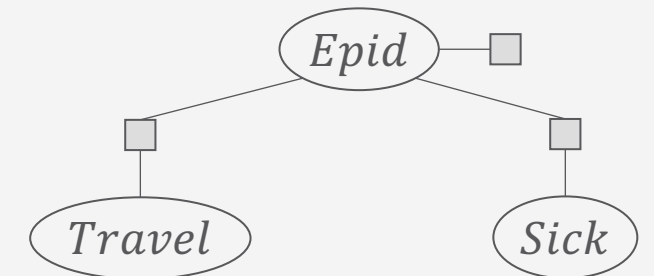
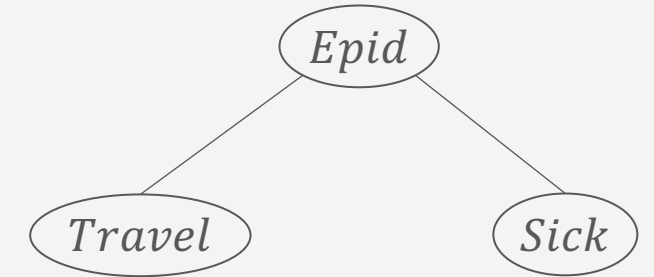


Further Conditional Independences in the Graph Structure

- Given a model F over random variables $\mathbf{R}$
- Pairwise independences in F
Each pair of non-neighbours $R_i, R_j \in \mathbf{R}$ is independent given all other random variables, i.e.,

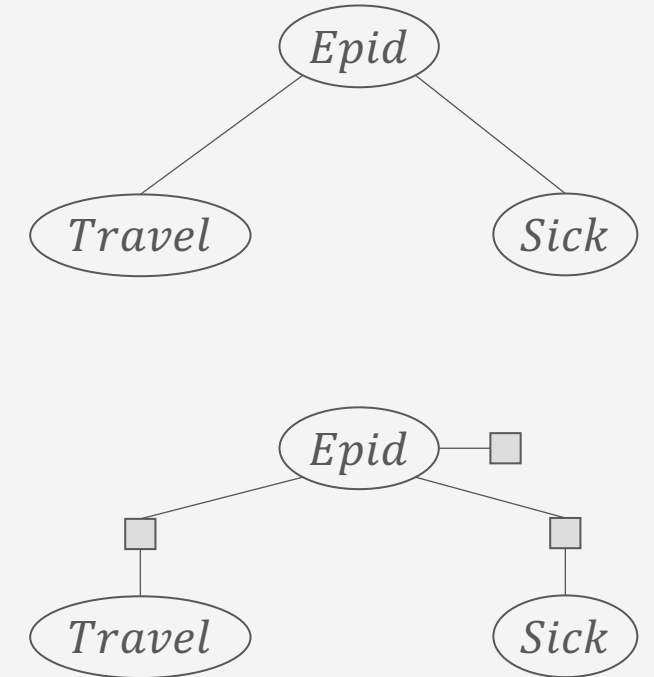
$$\left\{ (R_i \perp R_j \mid \mathbf{R}') \mid R_i, R_j \in \mathbf{R} \wedge \{R_i, R_j\} \notin E \wedge \mathbf{R}' = \mathbf{R} \setminus \{R_i, R_j\} \right\}$$
- If conditioning on all remaining variables, all paths between R_i, R_j must be blocked
- Weaker statement than local independences • •

Why?



Further Conditional Independences in the Graph Structure

- Given a model F over random variables $\mathbf{R}$ with MN (V, E)
- Active paths
 - Given a set $\mathbf{S} \subseteq \mathbf{R}$, path $R_1 - \dots - R_k$ in (V, E) is **active** given $\mathbf{S}$, if no $R_i, i \in \{1, \dots, k\}$ in $\mathbf{S}$ is, i.e.: $\forall i \in \{1, \dots, k\} : R_i \notin \mathbf{S}$
 - That is, a path is inactive given $\mathbf{S}$ if at least one R_i is in $\mathbf{S}$
- Separation (also called m -Separation after Markov)
 - A set $\mathbf{S}$ **separates** two sets $\mathbf{T}, \mathbf{U}$ in (V, E) if there is no active path between any variable $T \in \mathbf{T}$ and any variable $U \in \mathbf{U}$ given $\mathbf{S}$
 - That is, all paths between the variables in $\mathbf{T}, \mathbf{U}$ have to be inactive given $\mathbf{S}$
 - On each path between T, U , there must be an R_i that is in $\mathbf{S}$
 - Denoted by $sep(\mathbf{T}; \mathbf{U} \mid \mathbf{S})$
 - $\mathbf{S}$ also called **separator**



Further Conditionanl Independences in the Graph Structure

- Given a model F over random variables $\mathbf{R}$ with $\text{MN}(V, E)$

- Global independences in F :

$$\{(\mathbf{T} \perp \mathbf{U} \mid \mathbf{S}) \mid \text{sep}(\mathbf{T}; \mathbf{U} \mid \mathbf{S})\}$$

- Proof: Given $\text{sep}(\mathbf{T}; \mathbf{U} \mid \mathbf{S})$, does $(\mathbf{T} \perp \mathbf{U} \mid \mathbf{S})$ hold?

- Initially: $\mathbf{T} \cup \mathbf{U} \cup \mathbf{S} = \mathbf{R}$

- As $\mathbf{S}$ is a separator between $\mathbf{T}$ and $\mathbf{U}$, there are no direct edges between $\mathbf{T}$ and $\mathbf{U}$
- As such, each clique in (V, E) (i.e., each factor in F) either in $\mathbf{T} \cup \mathbf{S}$ or in $\mathbf{U} \cup \mathbf{S}$
- With F_T the set of factors, whose cliques are in $\mathbf{T} \cup \mathbf{S}$, and F_U the remaining factors

$$P_F = \frac{1}{Z} \prod_{f \in F} f = \frac{1}{Z} \prod_{f \in F_T} f \cdot \prod_{g \in F_U} g = \frac{1}{Z} f(\mathbf{T}, \mathbf{S}) \cdot g(\mathbf{U}, \mathbf{S})$$

- As factors in F_T do not contain variables from $\mathbf{U}$ and factors in F_U not from $\mathbf{T}$, the factor product can be split into a factor over $\mathbf{T} \cup \mathbf{S}$ and a factor over $\mathbf{U} \cup \mathbf{S}$
- Given the factorisation $f(\mathbf{T}, \mathbf{S}) \cdot g(\mathbf{U}, \mathbf{S})$, $(\mathbf{T} \perp \mathbf{U} \mid \mathbf{S})$ holds (Bayes Theorem, def. of cond. indep.)

Further Conditionanl Independences in the Graph Structure

- Given a model F over random variables $\mathbf{R}$ with $\text{MN}(\mathbf{V}, E)$

- Global independences in F :

$$\{(T \perp U \mid S) \mid \text{sep}(T; U \mid S)\}$$

- Proof: Given $\text{sep}(T; U \mid S)$, does $(T \perp U \mid S)$ hold?

- In general: $T \cup U \cup S \subset \mathbf{R}$ with $V = \mathbf{R} \setminus (T \cup U \cup S)$

- V can be split into two sets V_1, V_2 such that S is a separator between $T \cup V_1$ and $U \cup V_2$

- Afterwards, the initial argument holds:

With F_T the set of factors with its cliques in $T \cup V_1 \cup S$, and F_U the remaining factors

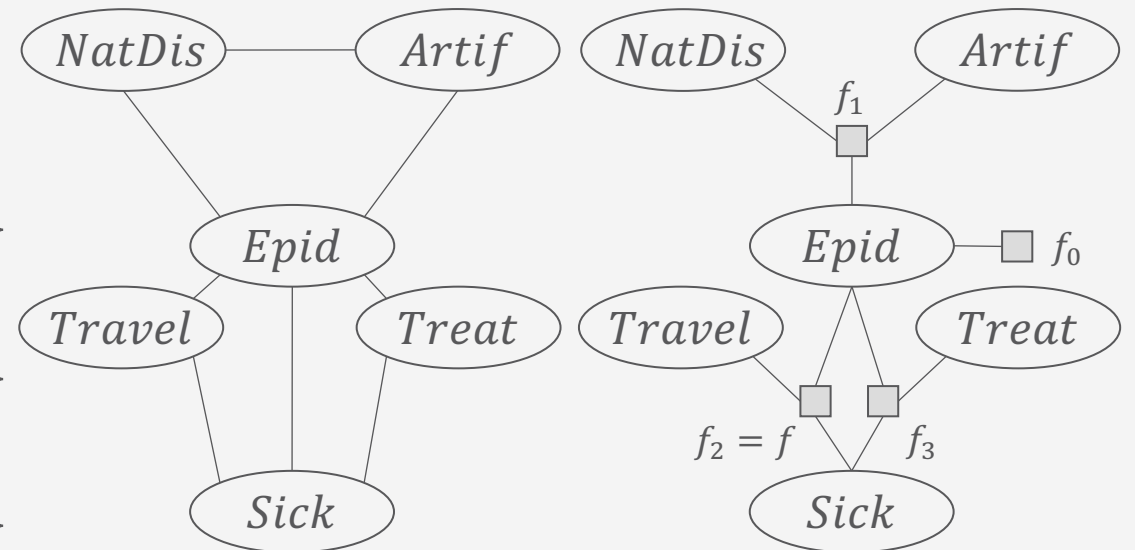
$$P_F = \frac{1}{Z} \prod_{f \in F} f = \frac{1}{Z} \prod_{f \in F_T} f \cdot \prod_{g \in F_U} g = \frac{1}{Z} f(T, V_1, S) \cdot g(U, V_2, S)$$

- Decomposition rule of conditional independences $(A \perp B, C \mid D) \Rightarrow (A \perp B \mid D) \rightarrow (T \perp U \mid S)$ holds
- Monotone*: if $\text{sep}(T; U \mid S)$ holds, then $\text{sep}(T; U \mid S')$ holds for all $S' \supseteq S$

Example: Independences in Factor Models

- Pairwise independences:
 - $NatDis \perp Travel \mid \{Artif, Epid, Treat, Sick\}$
 - $NatDis$ with $Artif$ exchangeable
 - $Travel$ with $Treat, Sick$ exchangeable
 - $Treat \perp Travel \mid \{NatDis, Artif, Epid, Sick\}$
 - $Epid$ has no non-neighbours
- Local independences
 - $NatDis \perp \{Travel, Treat, Sick\} \mid \{Artif, Epid\}$
 - $NatDis$ with $Artif$ exchangeable
 - $Treat \perp \{NatDis, Artif, Travel\} \mid \{Epid, Sick\}$
 - $Treat$ with $Travel$ exchangeable
 - $Sick \perp \{NatDis, Artif\} \mid \{Epid, Travel, Treat\}$

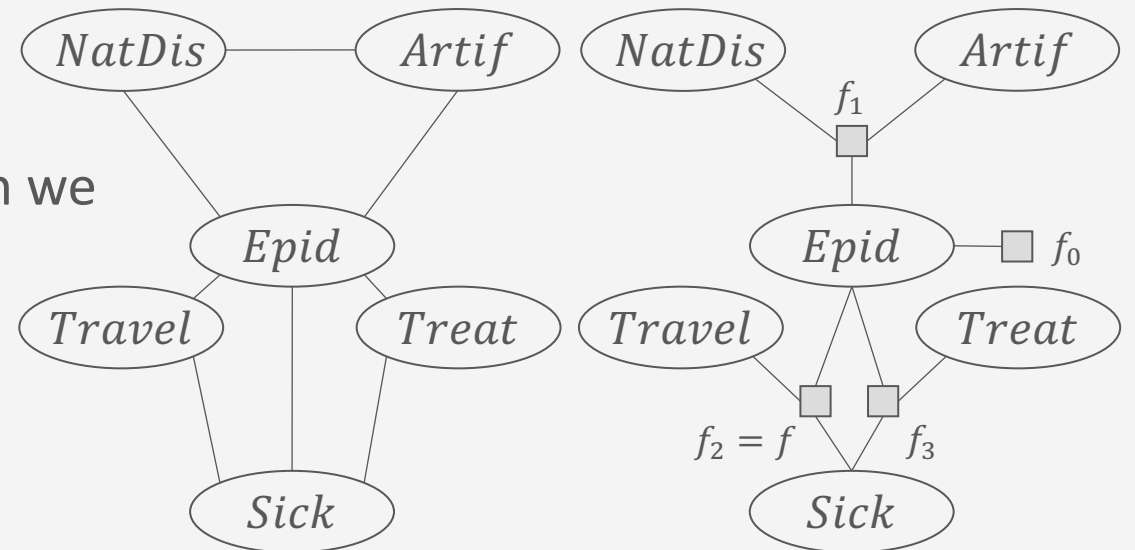
Pairwise: $(R_i \perp R_j \mid R'), R' = R \setminus \{R_i, R_j\}$
 Local: $(R \perp R' \mid MB(R)), R' = R \setminus MB(R) \setminus \{R\}$
 Global: $\{(T \perp U \mid S) \mid sep(T; U \mid S)\}$



Example: Independences in Factor Models

- Global independences:
 - Pairwise and local independences can be mapped on $\{(T \perp U \mid S) \mid \text{sep}(T; U \mid S)\}$
 - $\{NatDis, Artif\} \perp \{Travel, Treat, Sick\} \mid Epid$
 - Decomposition: Subsets of $\{NatDis, Artif\}$ and $\{Travel, Treat, Sick\}$ possible
 - Monotone $\rightarrow$ Set right of $\mid$ extendable
 - Set of representative independences, from which we can infer all other independences:
 - $\{NatDis, Artif\} \perp \{Travel, Treat, Sick\} \mid Epid$
 - $Treat \perp Travel \mid \{Epid, Sick\}$

Pairwise: $(R_i \perp R_j \mid R'), R' = R \setminus \{R_i, R_j\}$
 Local: $(R \perp R' \mid MB(R)), R' = R \setminus MB(R) \setminus \{R\}$
 Global: $\{(T \perp U \mid S) \mid \text{sep}(T; U \mid S)\}$

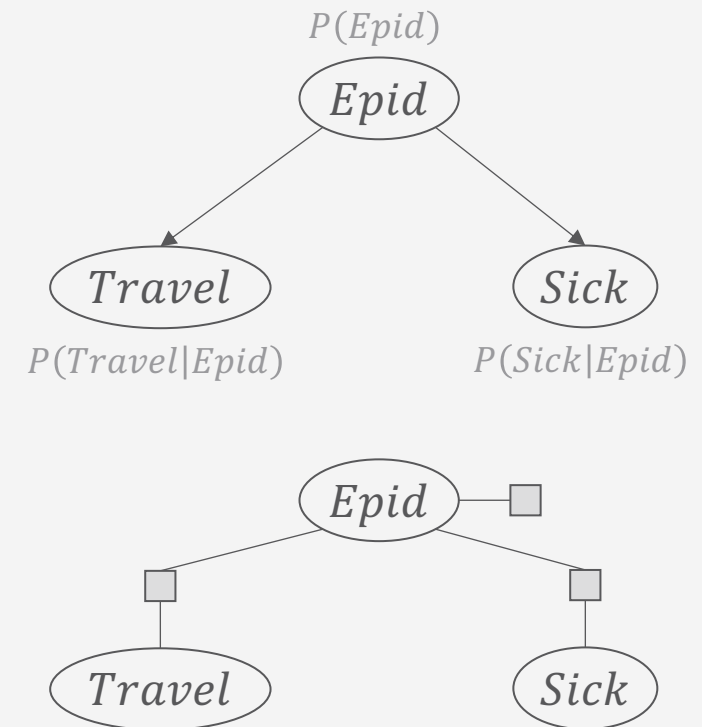


Conditional Independences in BNs

- BNs encode local independences as follows:
 - For each random variable $R \in \mathbf{R}$, R is independent from its non-descendants given its parents, i.e.,

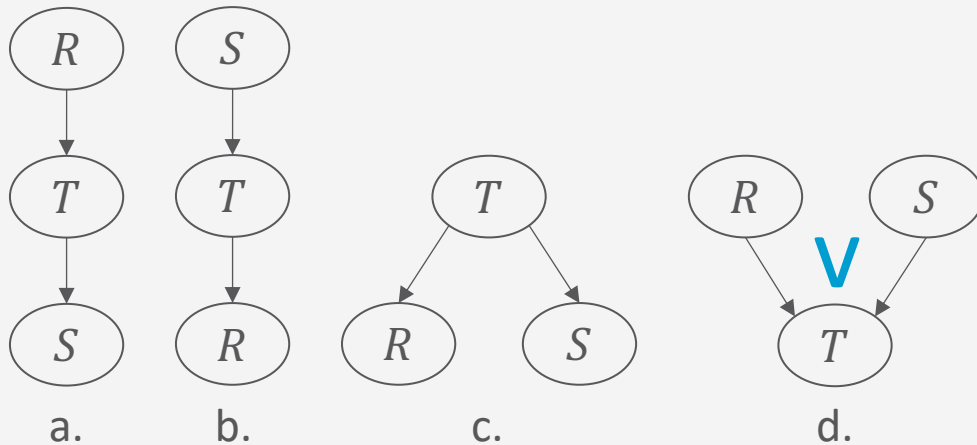
$$(R \perp \text{Ndesc}(R) \mid \text{Pa}(R))$$
- Why is it different from factor models?
 - Compare: In a factor model, each $R \in \mathbf{R}$ is independent of all other random variables given its neighbourhood, i.e.,

$$(R \perp \mathbf{R}' \mid \text{Nb}(R)) \text{ or } (R \perp \mathbf{R}' \mid \text{MB}(R))$$
 - $R \in \mathbf{R}, \mathbf{R}' = (\mathbf{R} \setminus \{R\}) \setminus \text{Nb}(R)$



Influences in BNs

- Direct edge: Direct influence
- Indirect connection: Case distinction
 - a. Indirect causal effect
 - b. Indirect evidential effect
 - c. Common cause
 - d. Common effect → **v-structure**



- Given by local independences
 - Case a-c: R influences S via T , but only if T is not observed
 - Case d: If T not observed, R and S independent
 - If T observed, R and S are dependent
 - Value z has to happen through R or S
- Statements also hold in larger nets with more nodes on the path between R , S
 - Case d: neither T nor a descendent observed

Statements for BNs regarding pairwise, local, and global independences need to consider **v-structures**

Conditional Independences in BNs

- Given a BN B over random variables $\mathbf{R}$

- Markov blanket** in BNs:

Parents, children and all other parents of children, i.e.,

$$\text{MB}(R) = \text{Pa}(R) \cup \text{Ch}(R) \cup \bigcup_{R' \in \text{Ch}(R)} \text{Pa}(R') \setminus \{R\}$$

Includes all *v-structures*

- Local independences** in B analogue to an MN:

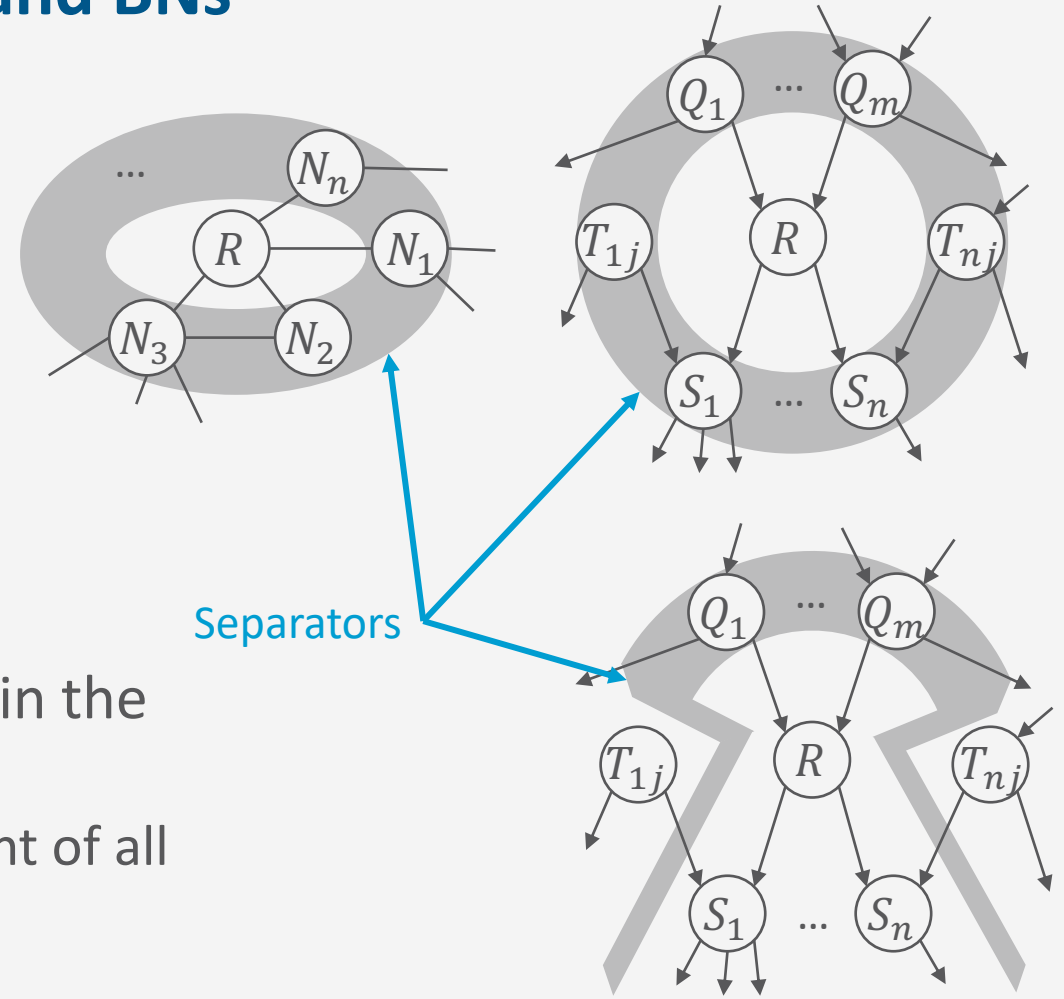
Each $R \in \mathbf{R}$ is independent of all other random variables given its Markov blanket, i.e.,

$$\{(R \perp \mathbf{R}' \mid \text{MB}(R)) \mid R \in \mathbf{R} \wedge \mathbf{R}' = \mathbf{R} \setminus \{R\} \setminus \text{MB}(R)\}$$

- Markov blanket blocks all paths, including v-structures

Illustration of Local Independences in MNs and BNs

- Markov blanket of a random variable R
 - in a factor model:
 - Set of direct neighbours N_i of R in MN
 - in a BN:
 - Union of
 - Parents Q_k of R
 - Children S_i of R
 - Parents T_{ij} of children S_i that are not R
- R conditionally independent of all other variables in the model given its Markov blanket as separator
 - In BNs, given parents as separators, then independent of all non-descendants



Conditional Independences in BNs

- Given a BN B over random variables $\mathbf{R}$
- Pairwise independences in B :
 - Usually not considered for BNs
 - Usual definition: two non-neighbours $R_i, R_j \in \mathbf{R}$ independent given all other variables $\mathbf{R}'$
 - Common descendant R' between two non-neighbours $R_i, R_j \in \mathbf{R}$ would make R_i, R_j dependent and would be included in $\mathbf{R}'$
 - That is, definition would need to exclude v-structures
 - Statement would be possible if
 - Only considering R_i, R_j that do not have a common descendant or
 - Excluding common descendants from $\mathbf{R}'$

Conditional Independences in BNs

- Given disjoint sets $\mathcal{S}, \mathcal{T}, \mathcal{U}$ of random variables B
- Active connection between two random variables T, U
 - Connection: *undirected* path between T, U
 - Connection between T, U is **active** given $\mathcal{S}$ if

- If there is a v-structure $R_{i-1} \rightarrow R_i \leftarrow R_{i+1}$ on the connection, then R_i or a descendant of R_i is in $\mathcal{S}$
 - No further node of the connection is in $\mathcal{S}$
- Directed separation (*d-separation*)
 - $\mathcal{T}$ and $\mathcal{U}$ are **d-separated** given $\mathcal{S}$ if there is no active connection between any variable $T \in \mathcal{T}$ and any variable $U \in \mathcal{U}$ given $\mathcal{S}$, denoted by $d\text{-sep}(\mathcal{T}; \mathcal{U} \mid \mathcal{S})$
- Global independences that hold in B and P_B respectively*

$$\{(\mathcal{T} \perp \mathcal{U} \mid \mathcal{S}) \mid d\text{-sep}(\mathcal{T}; \mathcal{U} \mid \mathcal{S})\}$$

Includes all
v-structures

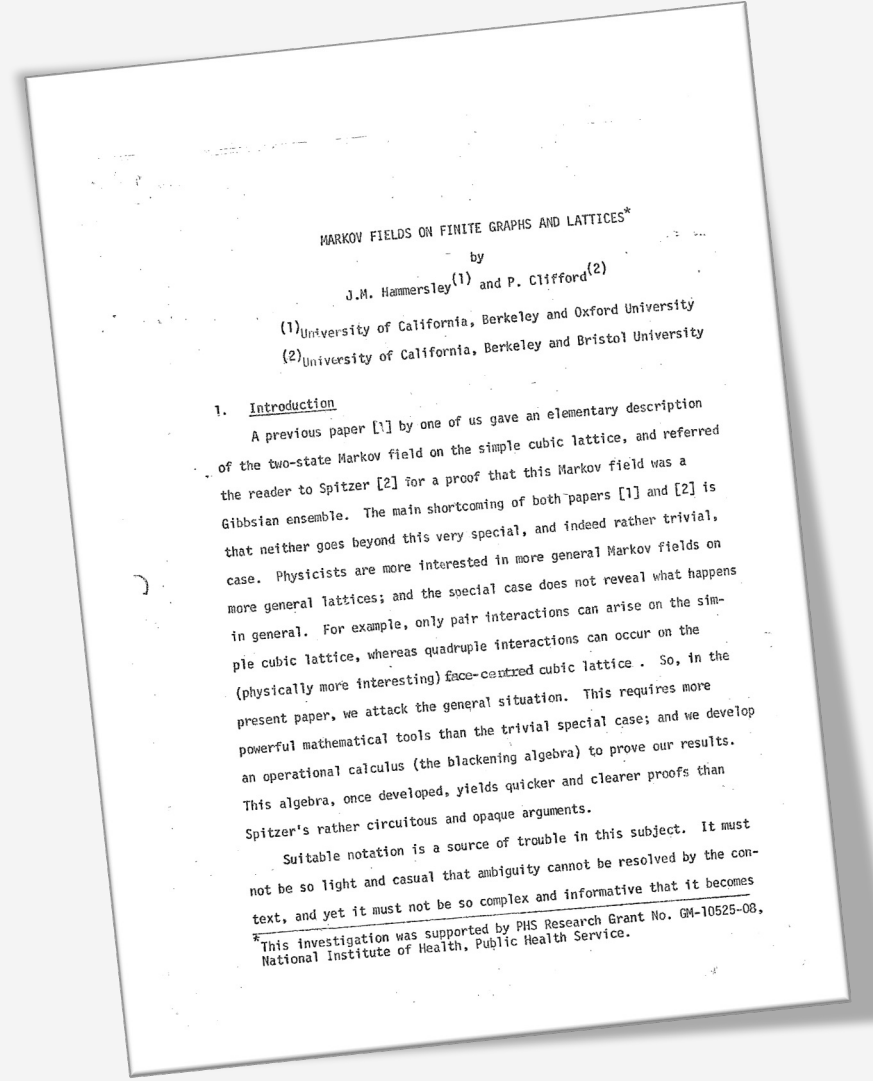


Conditional Independences

- Connection between the independences
 - Pairwise: $(R_i \perp R_j \mid \mathbf{R}'), \mathbf{R}' = \mathbf{R} \setminus \{R_i, R_j\}$ in MNs
 - Local: $(R \perp \mathbf{R}' \mid \text{MB}(R)), \mathbf{R}' = \mathbf{R} \setminus \text{MB}(R) \setminus \{R\}$
 - Global: $\{(\mathbf{T} \perp \mathbf{U} \mid \mathbf{S}) \mid \text{sep}(\mathbf{T}; \mathbf{U} \mid \mathbf{S})\}$ or $\{(\mathbf{T} \perp \mathbf{U} \mid \mathbf{S}) \mid d\text{-sep}(\mathbf{T}; \mathbf{U} \mid \mathbf{S})\}$
 - ❖ Equivalent in strictly positive distributions
 - Strictly positive: $P(\omega) > 0$ for all $\omega \in \text{Val}(\mathbf{R})$
 - Otherwise, a conditional probability might not be defined because the divisor might be 0
- Pairwise, local, and global independences sometimes also called pairwise, local, and global **Markov property**

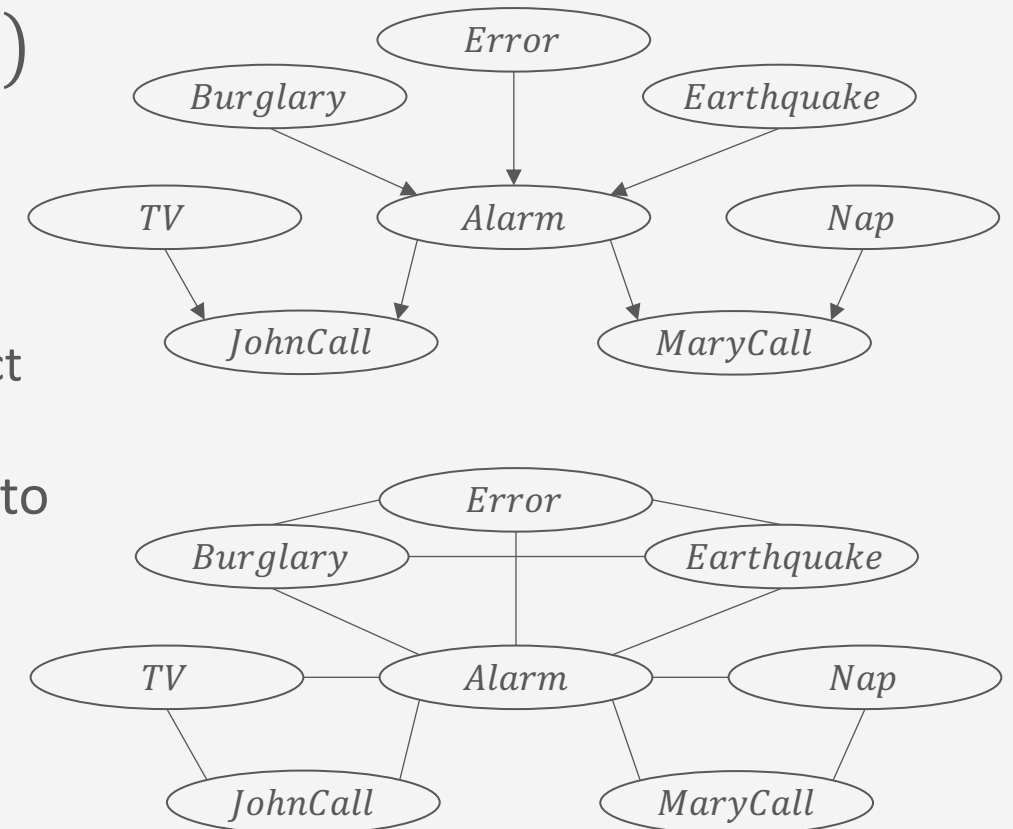
Conditional Independences

- Hammersley-Clifford Theorem
 - P fulfils one of the Markov properties with respect to an undirect graph *if and only if* P can be factorised over the cliques of the graph
 - If P fulfils a Markov property, then P can be factorised over the cliques
 - If the factorisation of P can be represented using the cliques of the graph, then the graph fulfils the Markov properties
- The theorem does not state that the factorisation is unique nor how to find such a factorisation



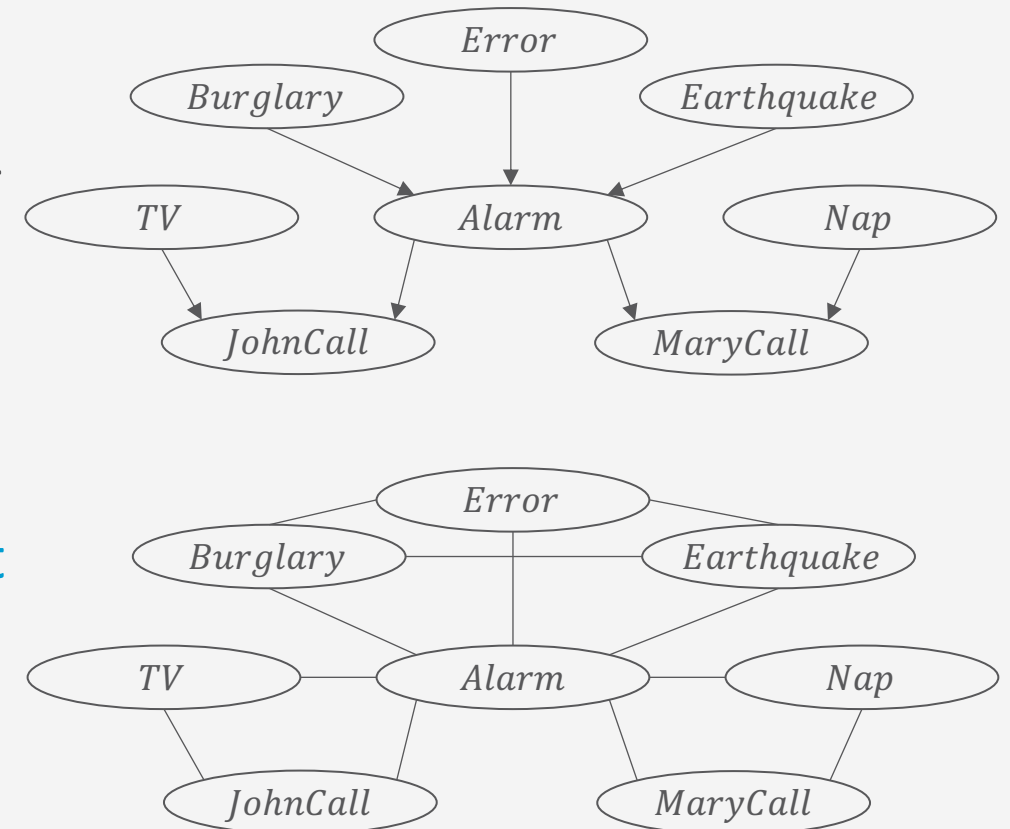
Turning BNs into Factor Models

- Each CPD $P(R|\text{Pa}(R))$ becomes a factor $\phi(R, \text{Pa}(R))$
 - Effect: Conditional probability hidden in factor
- Effect on graph structure
 - All parent nodes get directly connected
 - **Moralising** the graph: Make edges undirected and connect all former parents through direct edges
 - As such, explicit representation of independences due to v-structures gets lost
 - In example,
 - Independence of *Burglary*, *Error*, *Earthquake* if *Alarm*, *JohnCall*, *MaryCall* not given
 - Independence of *Alarm*, *TV* if *JohnCall* not given and of *Alarm*, *Nap* if *MaryCall* not given



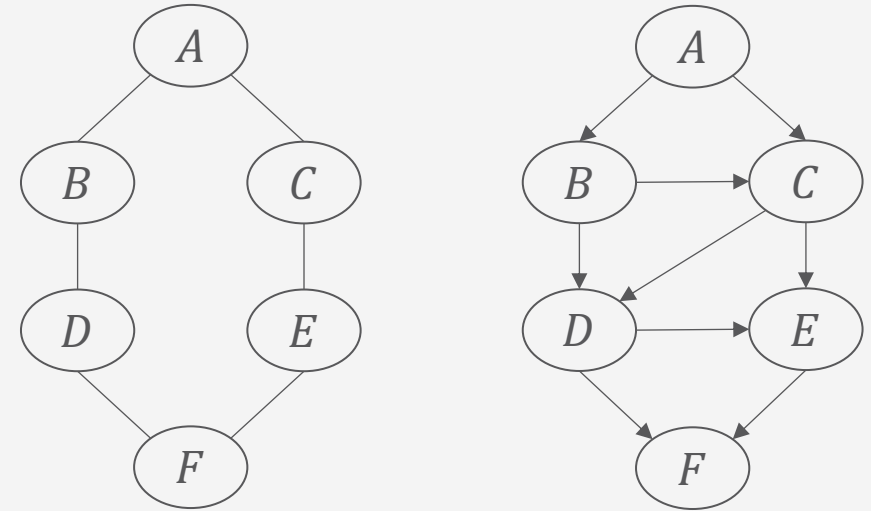
Turning BNs into Factor Models

- No explicit representation of independences gets lost if BN is a **moral** graph
 - BN $B = (V, E)$ is a **moral** graph if all pairs of parents of a child node are already connected by an edge
 - Formally, for each $R_i, R_j \in V$ with $R_i, R_j \in \text{Pa}(R_k)$ for any $R_k \in V$, it holds that $(R_i, R_j) \in E \vee (R_j, R_i) \in E$
- If a BN $B = (V, E)$ is a **moral** graph, then there exists a factor model F with an MN (V, E') such that the independences are equal
 - Because then connecting all parents in the graph does not add any edge and thus no independences get lost



The Other Direction: Turning Factor Models into BNs

- Informally: Given an ordering, add nodes to the BN and add an edge whenever there is a dependence in the MN between the already added nodes and the new one
 - Example: A, B, C, D, E, F
 - Add A
 - Add B , B dependent on A in MN $\rightarrow$ edge
 - Add C , C dependent on A in MN $\rightarrow$ edge
 - But C also dependent of B through $E, F, D \rightarrow$ edge
 - Add D , D dependent on $B, C \rightarrow$ edges
 - Add E , E dependent on $C, D \rightarrow$ edges
 - Add F , F dependent on $D, E \rightarrow$ edges

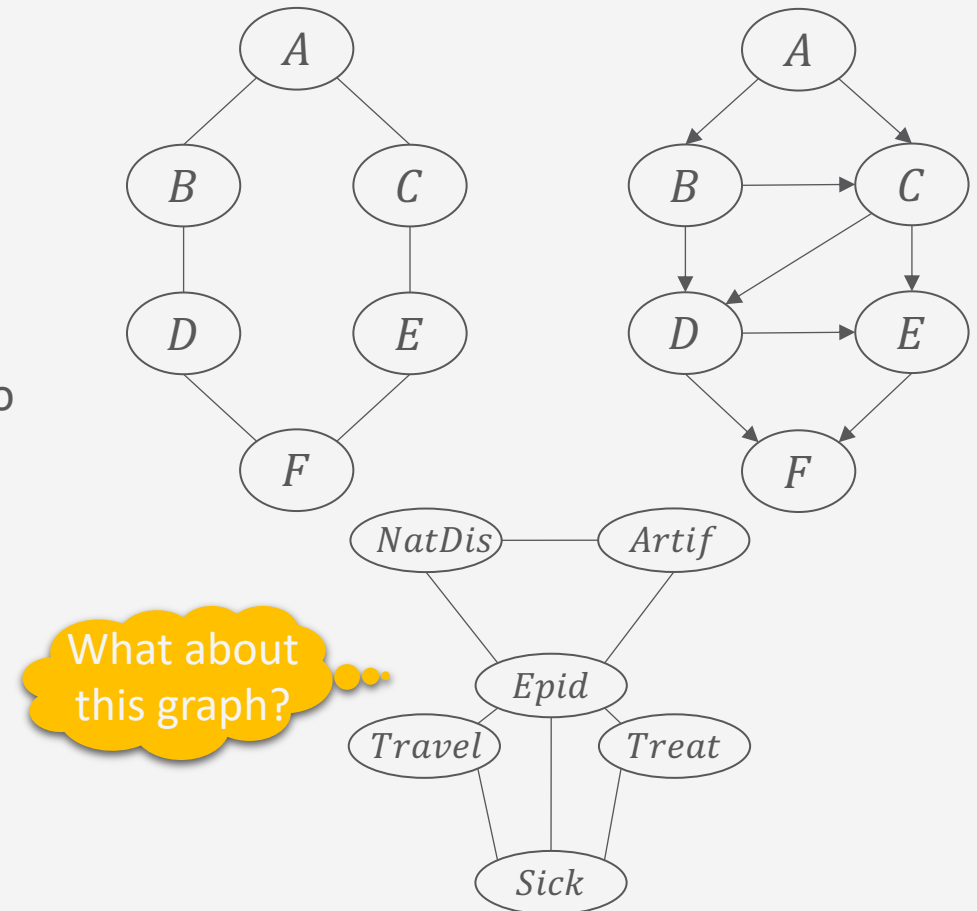


Transforming an MN into a BN usually leads to adding much more edges to the BN than the MN contained and as such, leads to losing the explicit representation of independences

The Other Direction: Turning Factor Models into BNs

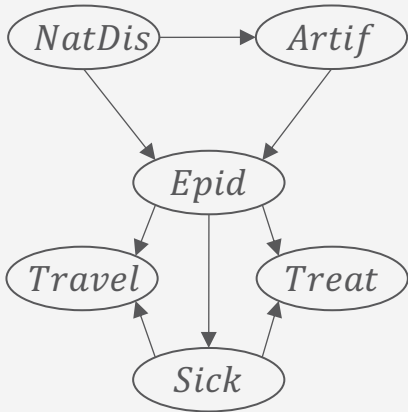
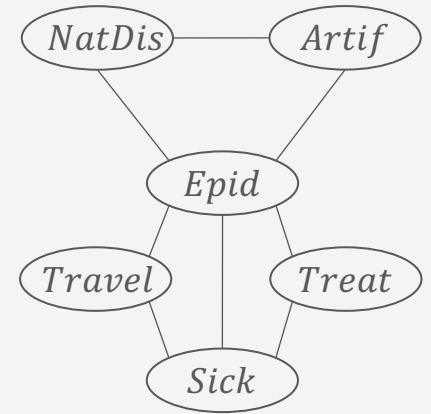
- Transformation is called **triangulation** of the MN
 - All cycles are partitioned into triangles
 - Result is always a chordal graph
 - Chordal / triangulated**: All cycles with at least four nodes have a chord
 - Chord*: Edge that is not part of the cycle but connects two nodes of the cycle
- If an MN is chordal, then there is a BN that represents the same independences

Stronger statement: An MN and a BN represent the same independences if and only if the undirected graph is chordal.

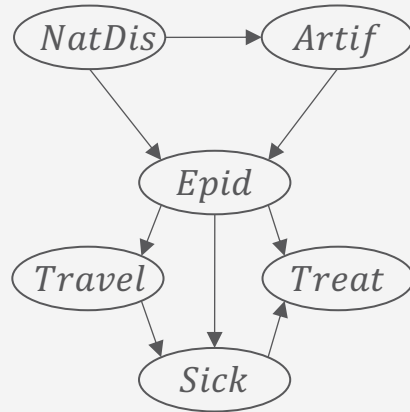


The Other Direction: Turning Factor Models into BNs

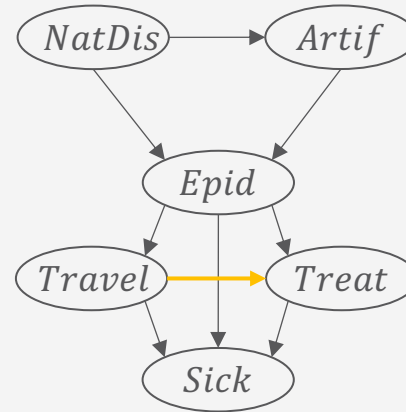
- Example: Chordal MN
 - Results given different orderings
 - Not all orders possible as some add more edges, then independences get lost: $Treat \perp Travel \mid Epid, Sick$



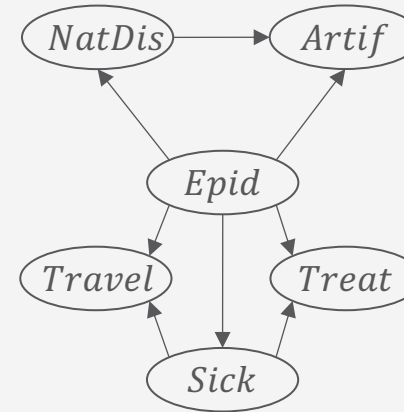
N, A, E, S, Tl, Tt



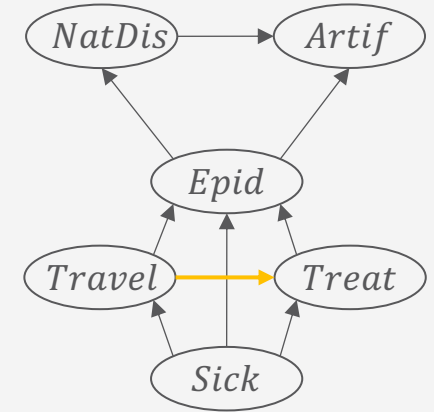
N, A, E, Tl, S, Tt



N, A, E, Tl, Tt, S



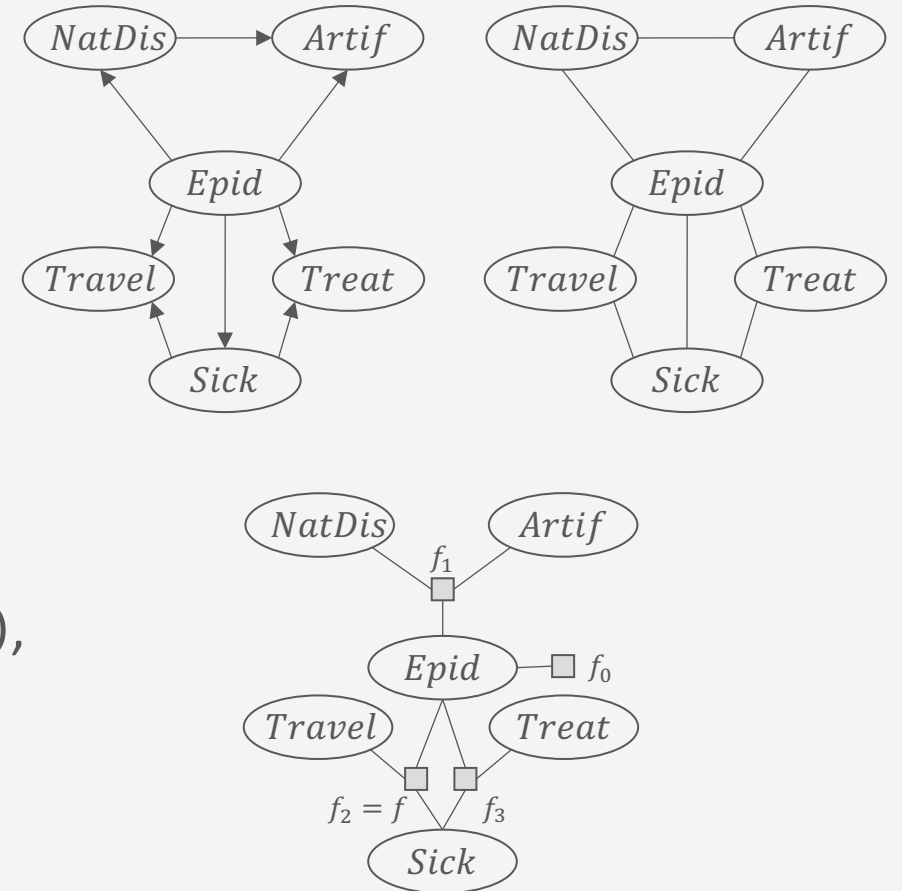
E, N, A, S, Tl, Tt



S, Tl, Tt, E, N, A

Factors to CPDs

- Turning a chordal MN into a BN relatively simple
- However, number of CPDs can increase compared to number of factors:
 - Example:
 - In factor model, *four* factors f_0, f_1, f_2, f_3
 - In BN, *six* CPDs $P(E), P(N | E), P(A | N, E), P(S | E), P(Tl | E, S), P(Tt | E, S)$
- Get CPDs by asking queries in factor model
 - Queries: $P(E), P(N | E), P(A | N, E), P(S | E), P(Tl | E, S), P(Tt | E, S)$



Interim Summary

- Pairwise, local, and global independences / Markov properties
 - Equivalent in strictly positive distributions
 - v-structures in BNs
 - Changes to Markov blanket definition
 - Result in losing explicit representation of independences if turning a BN into a factor model
- Hammersley-Clifford Theorem: there is a factorisation if and only if one of the Markov properties holds
- Turning BNs into MNs by moralisation, turning MNs into BNs by triangulation
 - Equivalent representation of independences if undirected graph is chordal
 - If no edges get added during transformation, then represented independences are identical

Overview: 2. Modelling Formalisms

a. Logic

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

b. Probabilistic Graphical Models (PGMs)

- Syntax, semantics, inference problems
- Correctness

c. Probabilistic Relational Models (PRMs)

- Parfactor models (PMs): motivation, syntax, semantics, graphical representation, inference tasks
- Markov logic networks (MLNs): Syntax, semantics of MLNs, graphical representation
- Turning MLNs into PMs and vice versa

Appendix

Multiplication example

<i>Epid</i>	<i>Travel</i>	<i>P</i>		<i>Travel</i>	<i>Sick</i>	<i>P</i>
false	false	0.10	•	false	false	0.25
false	true	0.20		false	true	0.30
true	false	0.30		true	false	0.35
true	true	0.40		true	true	0.10

<i>Epid</i>	<i>Travel</i>	<i>Sick</i>	<i>P</i>
false	false	false	$0.10 \cdot 0.25$
false	false	true	$0.10 \cdot 0.30$

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.,

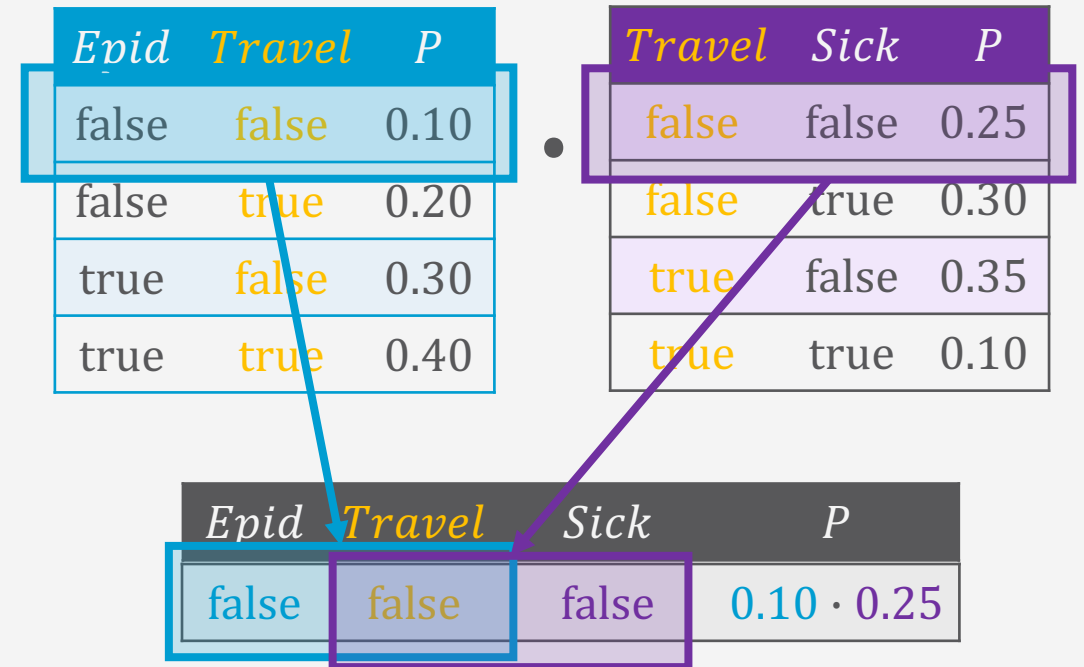
$$P(\text{Epid}, \text{Travel}) \cdot P(\text{Travel}, \text{Sick})$$

<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



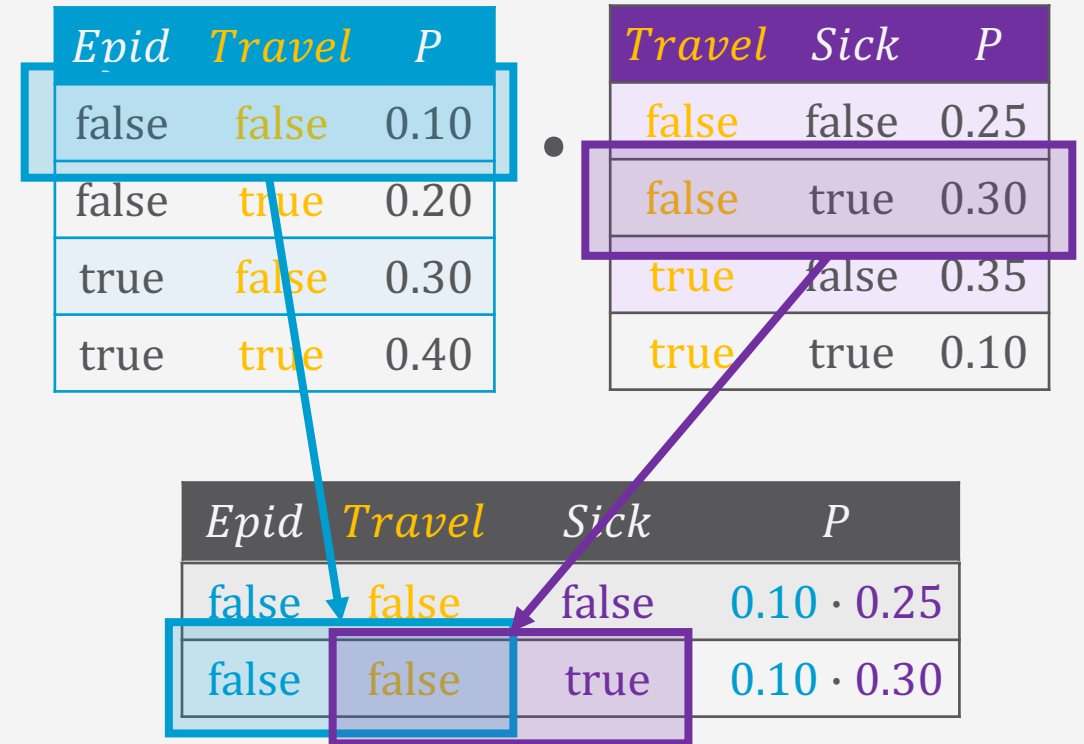
Excursion: Multiplication

- 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	$0.10 \cdot 0.25$
false	false	true	$0.10 \cdot 0.30$

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
			⋮

Diagram illustrating the multiplication of two probability distributions. The first table (left) shows the joint distribution for *Epid* and *Travel*. The second table (right) shows the joint distribution for *Travel* and *Sick*. The third table (bottom) shows the resulting joint distribution for *Epid*, *Travel*, and *Sick*, where the probability is the product of the probabilities from the first two tables. Arrows indicate the mapping from the first two tables to the third table.