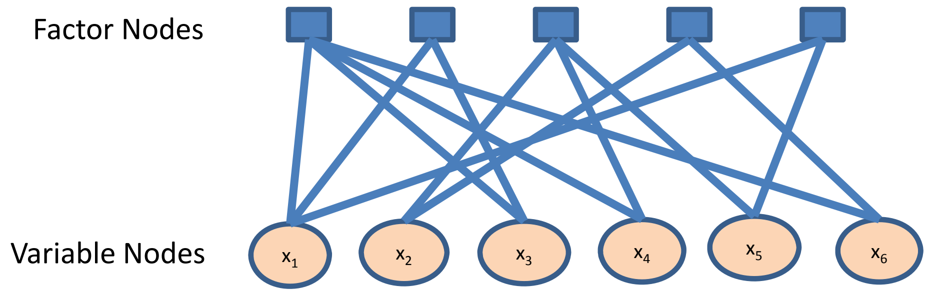
iPMR

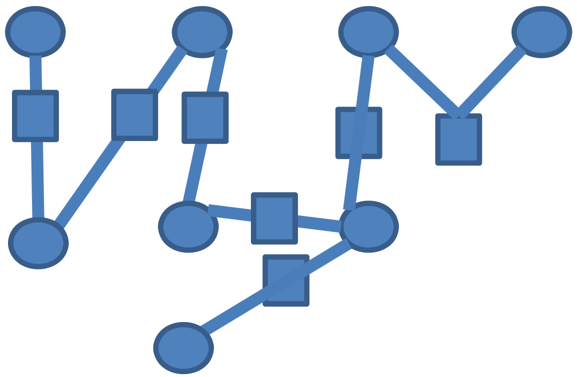
Recap

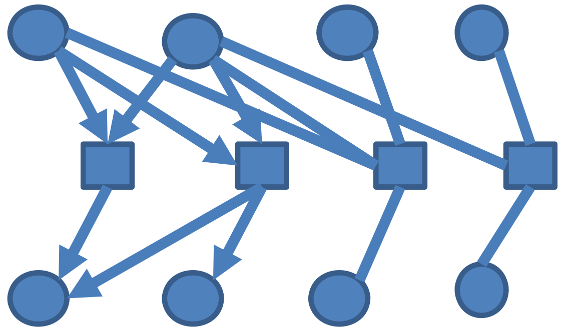
- What did we learn from the last lecture?

Summary

$$P(\mathbf{x}) = \frac{1}{Z}\prod_{i=1}^{N_f}\psi_i(\chi_i)$$







- Write down the distribution forms for the last two graphs.

Decorating

- Build example factor graph for colours of:
  - SOFA, ARMCHAIR, BED, TELEVISION, WARDROBE, LIVING ROOM CARPET, BEDROOM CARPET.

Conditional Independence

- Rule of independence extends to conditional probabilities,

$$P(\mathbf{x}|\mathbf{y},\mathbf{z}) = \frac{P(\mathbf{x}|\mathbf{z})P(\mathbf{y}|\mathbf{z})}{P(\mathbf{y}|\mathbf{z})} = P(\mathbf{x}|\mathbf{z})$$

- This is conditional independence and is notated by $I(\mathbf{x},\mathbf{y}|\mathbf{z})$
- Some comments on handling conditional probabilities…
  - Each variable must appear either on the right hand side or left hand side of | not both.
  - Conditional independence means you can drop a variable from the right side.

In Factor Graphs

- Factor graph ➔ conditional independence relationships.
  - Directed ✓
  - Undirected ✓
  - Partially directed ✓

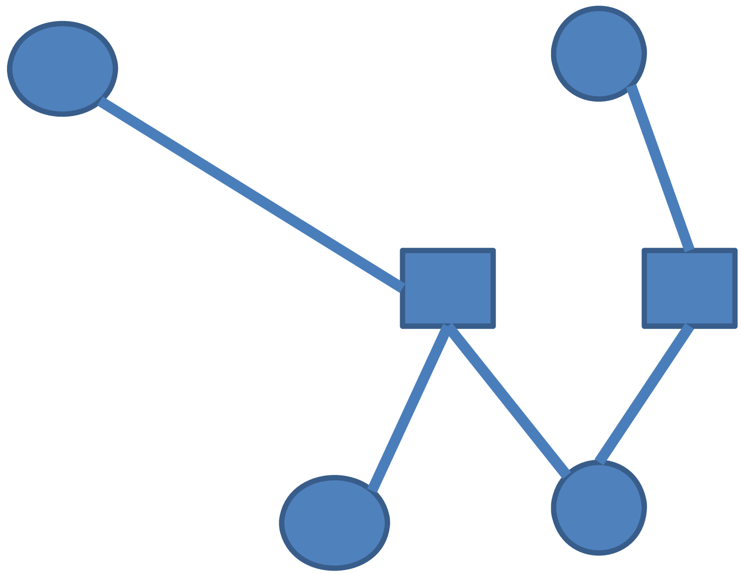
Two variables are guaranteed conditionally independent **given** a set of conditioned variables **if** all paths connecting the two variables are blocked.

A path is blocked if one or more of the following conditions is satisfied:

- One of the variables in the path is in the conditioning set.
- One of the variables or factors in the path has two incoming edges that are part of the path, and neither the variable/factor nor any of its descendants are in the conditioning set.

- Notes: if but not only if. Incoming: must be directed. Descendants: must all be directed.
- Frey 2003 Extending Factor Graphs so as to Unify Directed and Undirected Graphical Models

Examples

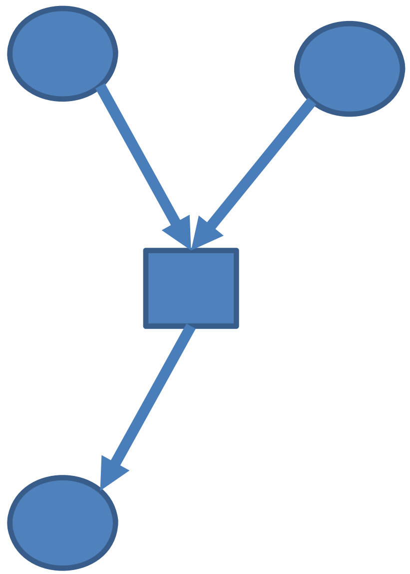


Two variables are guaranteed conditionally independent **given** a set of conditioned variables **if** all paths connecting the two variables are blocked.

A path is blocked if one or more of the following conditions is satisfied:

- One of the variables in the path is in the conditioning set.
- One of the variables or factors in the path has two incoming edges that are part of the path, and neither the variable/factor nor any of its descendants are in the conditioning set.

Examples

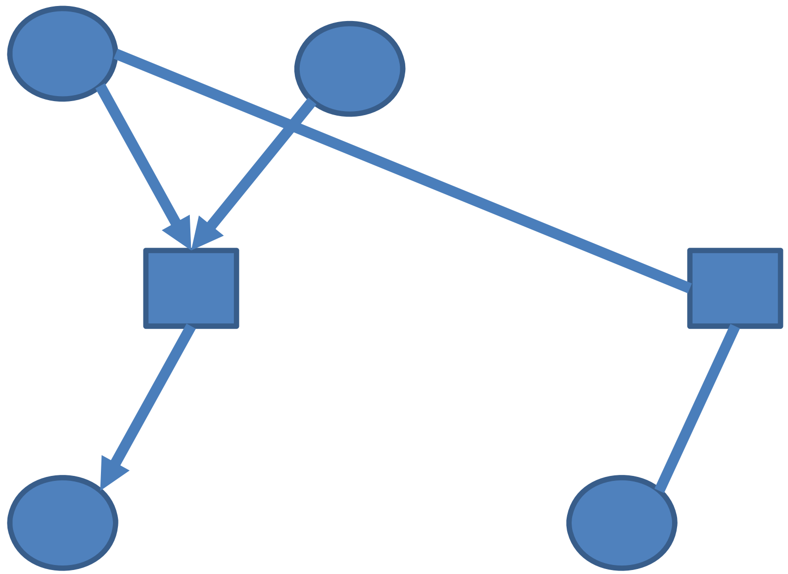


Two variables are guaranteed conditionally independent **given** a set of conditioned variables **if** all paths connecting the two variables are blocked.

A path is blocked if one or more of the following conditions is satisfied:

- One of the variables in the path is in the conditioning set.
- One of the variables or factors in the path has two incoming edges that are part of the path, and neither the variable/factor nor any of its descendants are in the conditioning set.

Examples

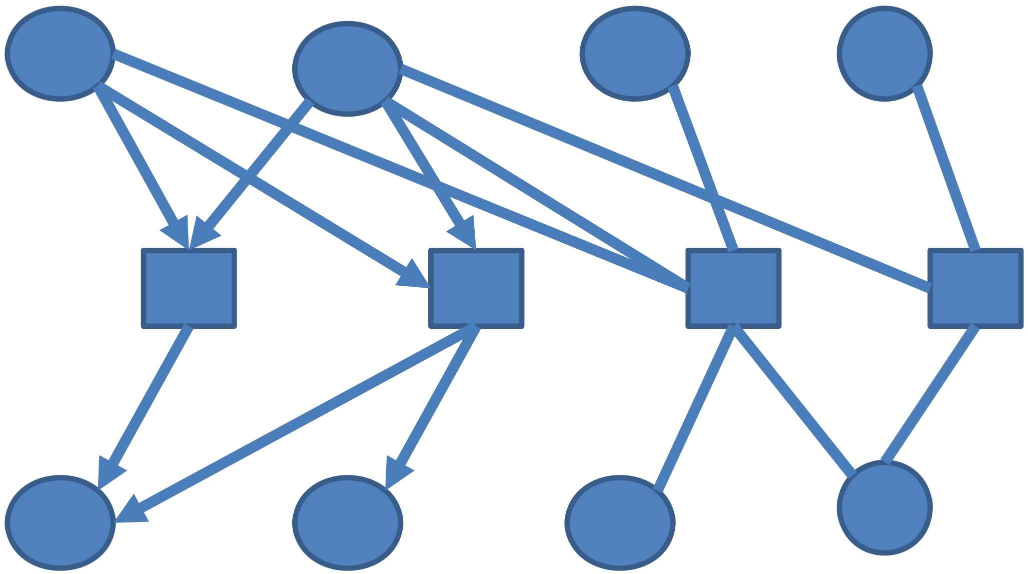


Two variables are guaranteed conditionally independent **given** a set of conditioned variables **if** all paths connecting the two variables are blocked.

A path is blocked if one or more of the following conditions is satisfied:

- One of the variables in the path is in the conditioning set.
- One of the variables or factors in the path has two incoming edges that are part of the path, and neither the variable/factor nor any of its descendants are in the conditioning set.

Examples

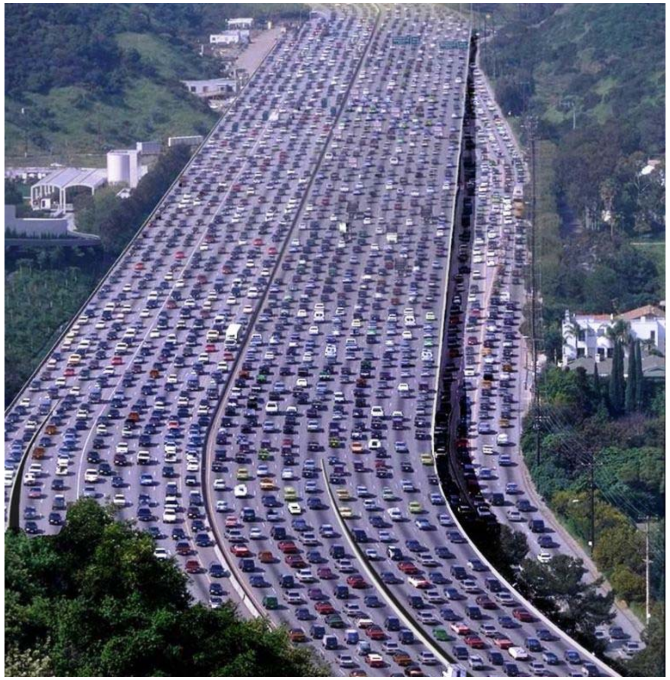


Two variables are guaranteed conditionally independent **given** a set of conditioned variables **if** all paths connecting the two variables are blocked.

A path is blocked if one or more of the following conditions is satisfied:

- One of the variables in the path is in the conditioning set.
- One of the variables or factors in the path has two incoming edges that are part of the path, and neither the variable/factor nor any of its descendants are in the conditioning set.

Revisit: Markov Blanket



Break

Factor Graphs

But what do Factor Graphs buy us?

Need to look at inference.

What is Inference?

- If we *know* a probabilistic model for something, how do we use it?
- Usually we ask questions we care about.
- They take the form
  - If THIS and THAT happen then what might happen to THOSE?
  - This is a conditional probability

$$P(\mathrm{THOSE}|\mathrm{THIS},\mathrm{THAT}) = \sum_{\mathrm{OTHER}} P(\mathrm{THOSE},\mathrm{OTHER}|\mathrm{THIS},\mathrm{THAT})$$

- Given a probabilistic model (which represents a joint distribution), what we want out are marginalised conditional distributions.
- Finding these is called inference.

Remember: Marginalising is summing out over unwanted variables

Example

$$P(\text{fun}|\text{barbeque}) = \sum_{\text{weather}} P(\text{fun}, \text{weather}|\text{barbeque})$$

Why is Inference Hard?

- In a general distribution, when we condition or marginalise a variable we have to worry about the effects on every combination of all the other variables.
- We can do inference by *enumeration* of all possibilities
- This becomes infeasible in large systems.
- *Sometimes* graphical models help.

- Aside: computing probabilities over many variables is also problematic, as finding the normalisation constant is costly.
  - We leave this issue for later. Here we assume we want marginal distributions over a small number of variables (we can condition on as many variables as we want).

Inference in Factor Graphs

- Consider marginalisation and conditioning operations on a tree.
- Conditioning
  - Look at all neighbours. Replace factors at all neighbours to be conditional factors. This is called *absorbing.*
- Marginalising
  - Find all the factors containing the node to be marginalised. Replace all these factors with one big factor produced by marginalising over those factors only.
  - All other factors stay the same.
- This is the basis of the elimination algorithm.

Next time: Sum-Products

- Sum distribution in sum-products

$$
\begin{aligned}
P(x_1, x_2, x_4, x_5) &= \sum_{x_3} \frac{1}{Z} \psi_1(x_1, x_4)\psi_2(x_1, x_5)\psi_3(x_2, x_3)\psi_4(x_3, x_5) \\
&= \frac{1}{Z} \psi_1(x_1, x_4)\psi_2(x_1, x_5) \sum_{x_3} \psi_3(x_2, x_3)\psi_4(x_3, x_5) \\
&= \frac{1}{Z} \psi_1(x_1, x_4)\psi_2(x_1, x_5)\psi_*(x_2, x_5)
\end{aligned}
$$

where

$$\psi_*(x_2, x_5) = \sum_{x_3} \psi_3(x_2, x_3)\psi_4(x_3, x_5)$$

- Order matters. Do cheap eliminations first.

Our Journey

Graphical Models

- Lecture 1: Introduce PMR
  - Focus: Distributions, Joint Models, Unsupervised.
- Lecture 2: Introduce Factor Graphs
  - Distributions → Factor Graphs
  - Content v Form
    - Structure of distributions
- Lecture 3: Conditional Independence in Factor Graphs.
  - Introduction to elimination.
- Next Lecture: More on Inference in Factor Graphs.