Learning & Inference in Graphical Models
Friday, April 1, 2011

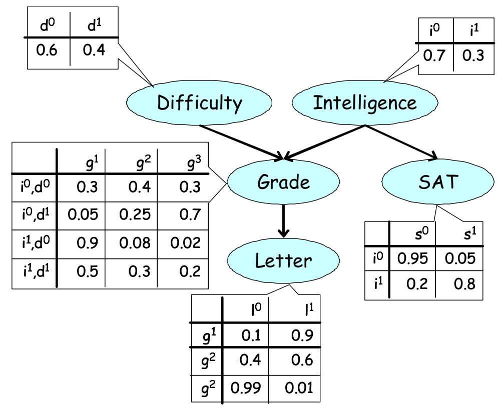
Instructors: Raquel Urtasun (rurtasun@ttic.edu)
Tamir Hazan (tamir@ttic.edu)

- Monday, Wednesday, Friday 1:30-2:20

http://ttic.uchicago.edu/~rurtasun/courses/GraphicalModels/graphical_models.html

Reminder

- Bayesian network is a way to describe a joint probability



$$p(I,D,G,S,L) \stackrel{def}{=} p(I)P(D)p(G|I,D)p(S|I)p(L|G)$$

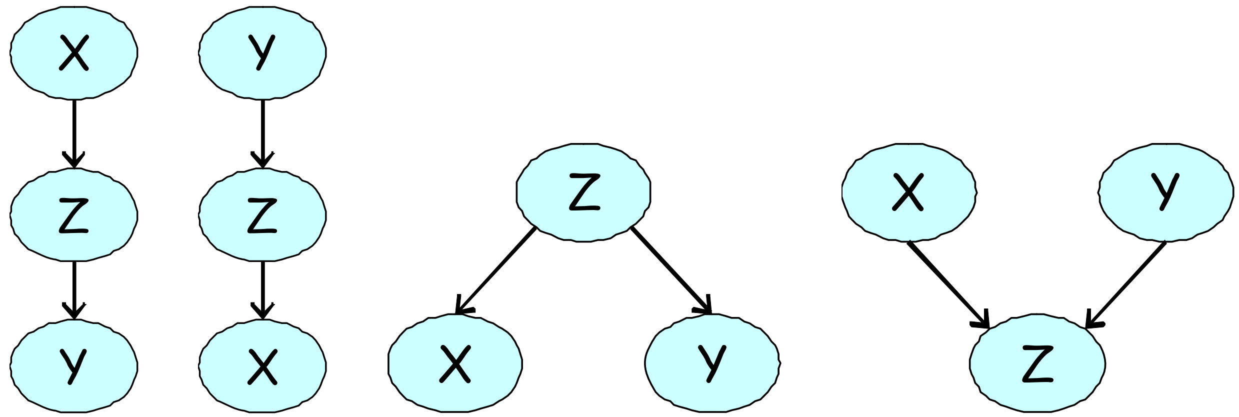
- Implicit independence statements!
- For example: $I \perp D$ (proof sketch, chain rule)

$$p(I,D,G,S,L) = p(I)P(D|I)p(G|I,D)p(S|I,D,G)p(L|I,D,G,S)$$

- independence from connectivity

$$x_i \perp x_{non-descendants(i)} \,\Big|\, x_{parents(i)}$$

d-separation



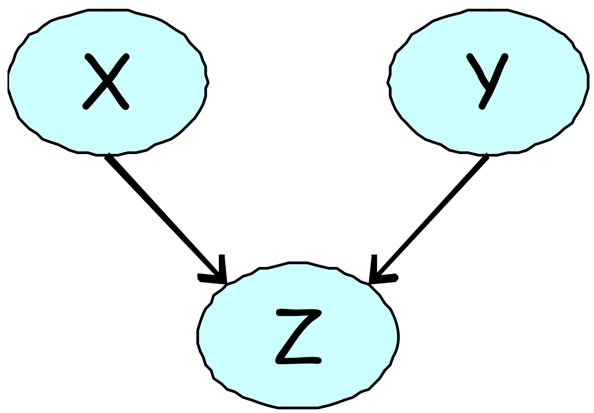
(a) (b) (c) (d)

$$X \perp Y | Z \qquad X \not\perp Y | Z$$

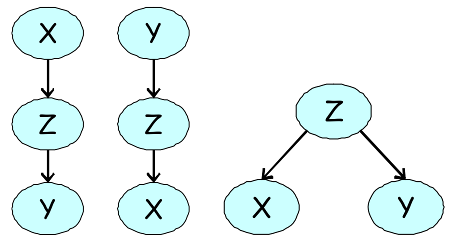
- (a,b,c): "dependence" is active if Z is not observed
- (d): "dependence" is active if Z is observed.

d-separation

- A path in the graph is active given evidence if any of the v-structure (or one of its descendants) is observed



and no other node is observed.



d-separation

- direct-separation: X_I, X_J are d-separated given X_K if there is no active path between nodes in $i \in I, j \in J$ given that nodes in K are observed.
- The independency map of the graph is the set of independencies of is d-sep

$$I(G) = \{X_I \perp X_J | X_K \quad : \quad d - sep(X_I, X_J; X_K)\}$$

- Future (next week): $I(G) \subset I(P)$

d-separation

$I(G) \subset I(P)$

- The converse is not true. We can construct a graph $A \rightarrow B$ without separation but with statistical independence:

$$p(a, b) = p(a)p(b|a)$$

$$p(b = 1|a = 1) = p(b = 1|a = 0) = p$$

$$p(b = 0|a = 1) = p(b = 0|a = 0) = 1 - p$$

d-separation

- d-separation reduces statistical independencies (hard) to connectivity in graphs (easy)
- Sets are independent is there are no active paths between them (there are exponentially many paths).

Algorithm for d-separation

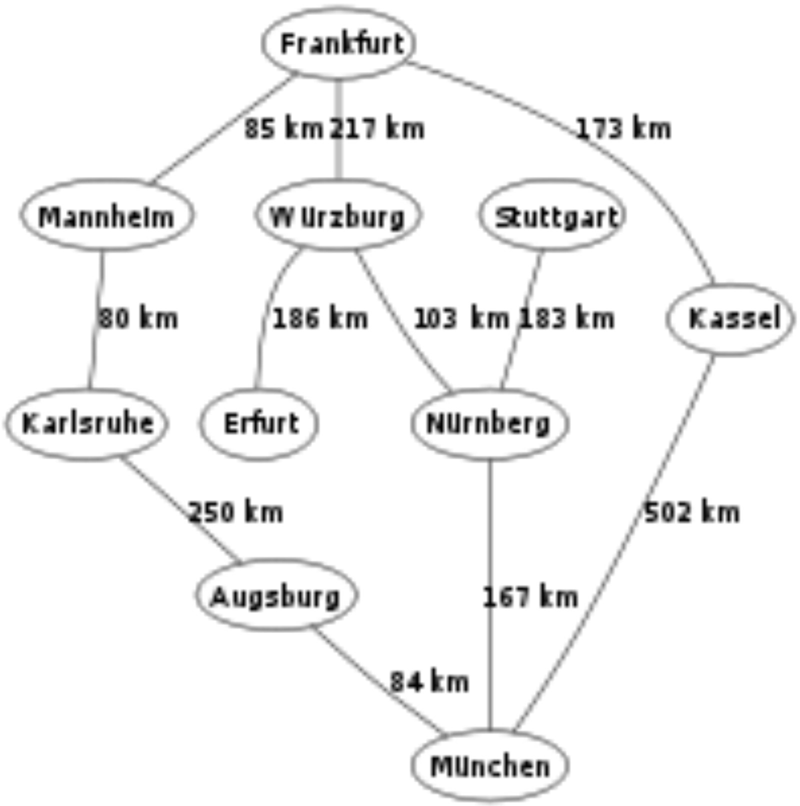
- Find candidates for v-structures:

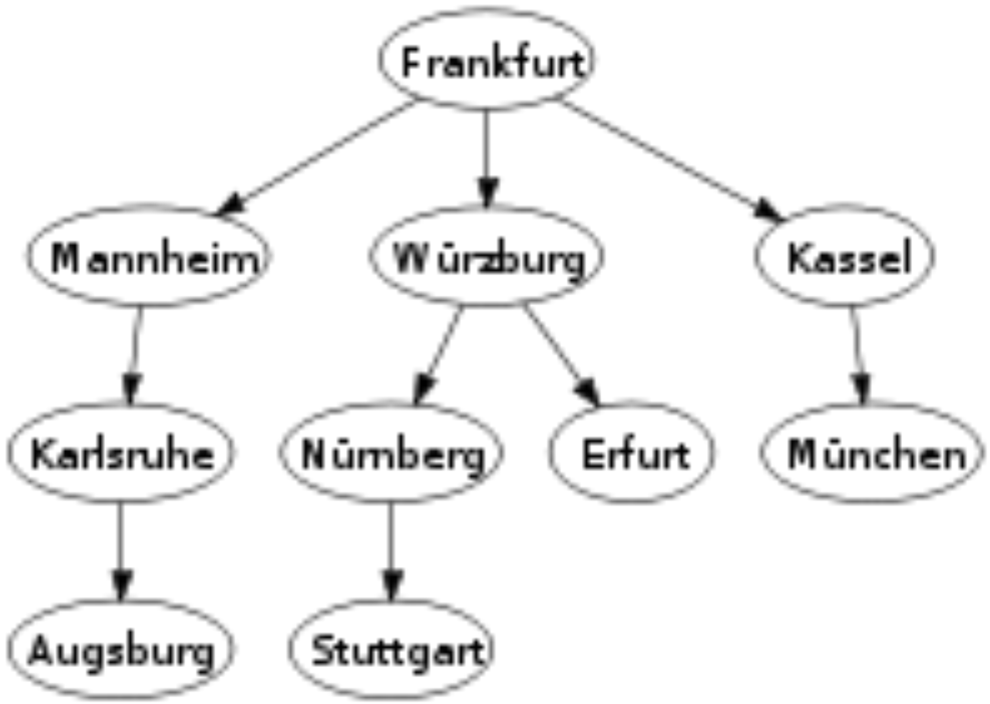
Mark (from leaves to root) all nodes that are observed or that have observed descendants

- Go over paths from X to Y (BFS) stopping when encountering "blocked" node.

"Blocked" is unmarked and in the middle of a v-structure or Observed and not in the middle of a v-structured.

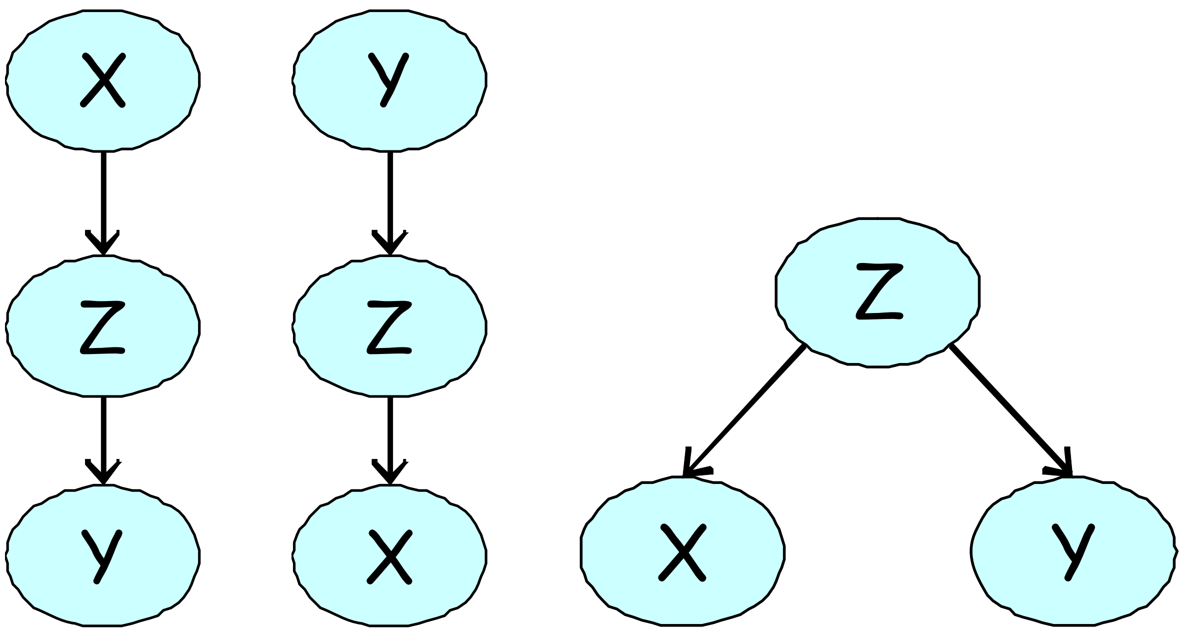
BFS (Breadth-first search)





I-Equivalence

- Two directed graphs G_1, G_2 are I-equivalent if their d-sep induce the same independencies, $I(G_1) = I(G_2)$



I-Equivalence

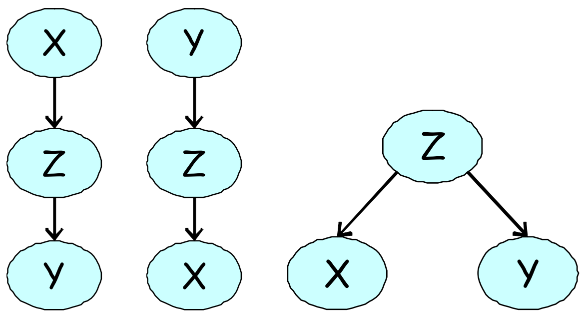
- Claim: If $I(G_1) = I(G_2)$ and

$$p(x_1, ..., x_n) = \prod_{i=1}^{n} p(x_i | x_{parents_{G_1}(i)})$$

then

$$p(x_1, ..., x_n) = \prod_{i=1}^{n} p(x_i | x_{parents_{G_2}(i)})$$

- Proof:



$$p(x, y, z) = p(x)p(z|x)p(y|z)$$

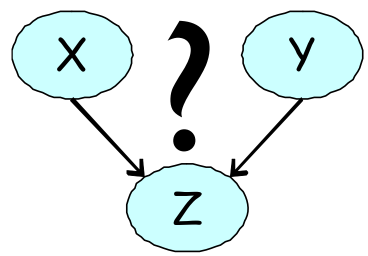
$$= p(y)p(z|y)p(x|z) = p(z)p(x|z)p(y|z)$$

Distributions to Graphs

- Every distribution can be represented by a direct graph (i.e. the chain rule).
- Not every distribution independencies can be exactly described by a direct graph

$$P(x,y,z) = \begin{cases} 1/12 & x \oplus y \oplus z = \mathit{false} \\ 1/6 & x \oplus y \oplus z = \mathit{true} \end{cases}$$

$$X \perp Y, Z \not\perp X | Y, Z \not\perp Y | X$$



however $X \perp Z$

Distributions to Graphs

- Can we learn / find a good / exact graphical model for a probability? Chapter 18