Alin Tomescu, http://people.csail.mit.edu/~alinush
6.867 Machine learning | Prof. Tommi Jaakkola | Week 13, Tuesday, November 26th, 2013| Lecture 23

Lecture 23: Inference in Bayesian networks

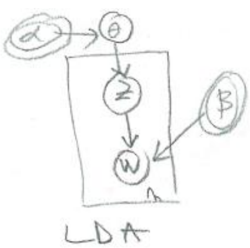
Final will be on December 10, 7-10pm.

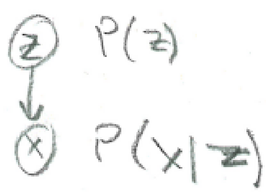
Today we'll try to infer more on *arbitrary graphs*.

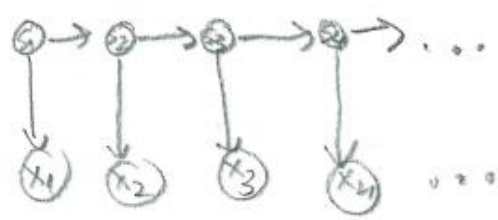
Directed acyclic graph (DAG)

Such graphs capture **independence properties**.

- They have nodes from $1, \ldots, n$
- Random variables are associated with the nodes $x_1, \ldots, x_n$
- Each node has a parent $pa_i, i = 1, \ldots, n$. $X_{pa_i} = \{X_j\}_{j \in pa_i}$
- If the graph is acyclic then there exists a node with no parent: $\exists k$ s.t. $X_{pa_k} = \emptyset$

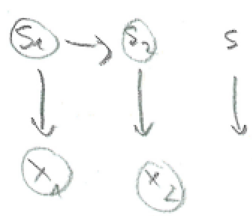




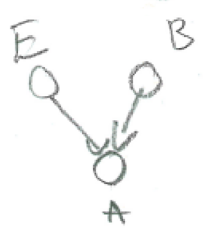


Probabilities on any such graph can be computed as follows:

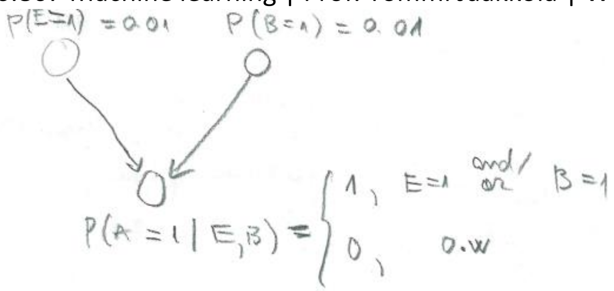
$$P(x_1, \ldots, x_n) = \prod_{i=1 \ldots n} P_i(x_i | x_{pa_i})$$



- s_1 is a a parent of x_1
- s_1 and s_2 are ancestors of x_2
- s_2 and x_2 are descendants of s_1



- This is an *"open V-structure,"* $E =$ earthquake, $B =$ burglary, $A =$ alarm
- $E \perp B$ (E and B are independent of each other)



If we apply our graph inference rule:

$$P(x_1, \dots, x_n) = \prod_{i=1 \dots n} P_i(x_i | x_{pa_i})$$

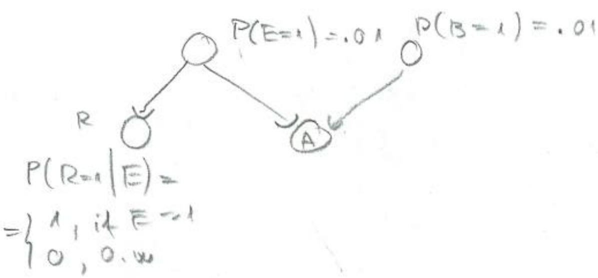
We get:

$$P(E, B) = \sum_A P(E, B, A) = \sum_A P(A|E, B)P(E)P(B) = P(E)P(B) \sum_A P(A|E, B) = P(E)P(B)$$

Are $E \perp B \mid A$? If I know there was an alarm, that implies either an earthquake or a burglary occurred so I cannot set the variables independently. Thus, they are not independent given A. This is called **induced dependence.**

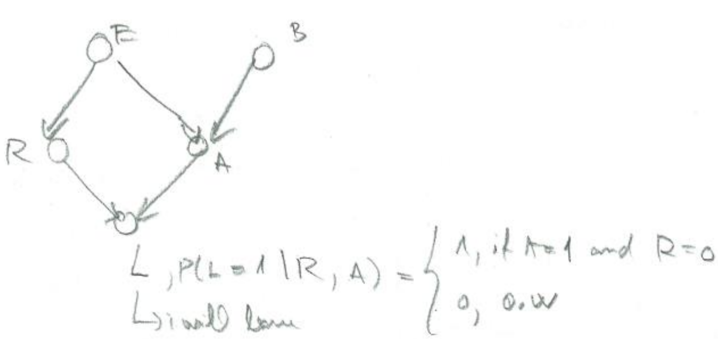
$$P(B = 1 \mid A = 1) = .5$$

Now we can include a "*radio report*" event in the graph, such that if an earthquake occurred, a radio report will be released with probability 1. Now we can ask if $R \perp B$? Yes.



What is $P(B = 1 \mid R = 1, A = 1) = .01 = P(B = 1)$

Why is that? "Explaining away" phenomenon (See http://www.cs.ubc.ca/~murphyk/Bayes/bnintro.html). Now, we can add a *"will leave"* event in the graph.



Is $L \perp B \mid A$? No.

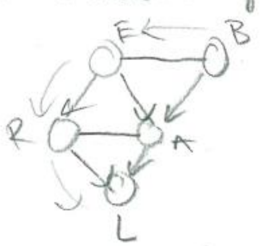
How to read off independence statements from the graph

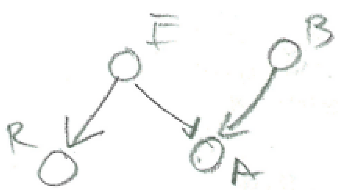
D-separation (and independence), says $x_i \perp x_j \mid x_k$, if i and j are separated (no path between them) by k in the *moralized* ancestral graph.

How can we answer such independence questions?

1. Keep only x_i, x_j, x_k and their ancestors (prune the graph)
2. "marry" the parents of all the nodes (in my initial notes I had just "of the 3 nodes"): you draw an edge between any two pair of parents
    a. After this point you can think of it as an undirected graph
3. $x_i \perp x_j \mid x_k$ is true if k separates i and j in the resulting graph

ancestral graph for $L \perp\!\!\!\perp B | A$





$$P(E, B, R, A) = P(E)P(B|E)P(R|E, B)P(A|E, B, R)$$

$$B \perp E$$

$$R \perp B \mid E$$

$$A \perp R \mid E, B$$

$$P(E, B, R, A) = P(E)P(B|E)P(R|E, B)P(A|E, B, R) = P(E)P(B)P(R|E)P(A|E, B)$$

What does this mean: Variable is independent of its preceding non-parents given the parents.

A variable is conditionally independent of its non-descendants given its immediate parents.

Alin Tomescu, http://people.csail.mit.edu/~alinush

Equivalence of graphs

$x_1 \to x_2 \to x_3$ $\quad x_1 \perp\!\!\!\perp x_3 \mid x_2$

$x_1 \leftarrow x_2 \to x_3$ $\quad x_1 \perp\!\!\!\perp x_3 \mid x_2$

$x_1 \to x_2 \leftarrow x_3$ $\quad x_1 \perp\!\!\!\perp x_3$

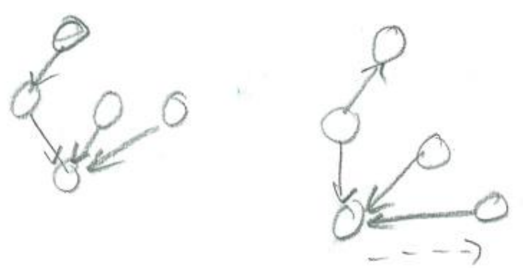
but $x_1 \perp\!\!\!\perp x_3 \mid x_2$ is false

Definition: Graph G and G' are equivalent iff they make the same independence assumptions.

Two graphs are equivalent if:

- they have the same set of edges (undirected)
- they have the same open v-structures

are equivalent. neither graph has an open-v structure



equivalent.
not equiv. if we flip the ----> arrow