d-separation: How to determine which variables are independent in a Bayes net

(This handout is available at http://web.mit.edu/jmn/www/6.034/d-separation.pdf)

The Bayes net assumption says:

"Each variable is conditionally independent of its non-descendants, given its parents."

It's certainly possible to reason about independence using this statement, but we can use **d-separation** as a more formal procedure for determining independence. We start with an independence question in one of these forms:

- "Are X and Y conditionally independent, given {givens}?"
- "Are X and Y marginally independent?"

For instance, if we're asked to figure out "Is P(A|BDF) = P(A|DF)?", we can convert it into an independence question like this: "Are A and B independent, given D and F?"

Then we follow this procedure:

1. Draw the ancestral graph.

Construct the "ancestral graph" of all variables mentioned in the probability expression. This is a reduced version of the original net, consisting only of the variables mentioned and all of their ancestors (parents, parents' parents, etc.)

2. "Moralize" the ancestral graph by "marrying" the parents.

For each pair of variables with a common child, draw an undirected edge (line) between them. (If a variable has more than two parents, draw lines between every pair of parents.)

3. "Disorient" the graph by replacing the directed edges (arrows) with undirected edges (lines).

4. Delete the givens and their edges.

If the independence question had any given variables, erase those variables from the graph and erase all of their connections, too. Note that "given variables" as used here refers to the question "Are A and B conditionally independent, given D and F?", not the equation "P(A|BDF) =? P(A|DF)", and thus does not include B.

5. Read the answer off the graph.

- If the variables are **disconnected** in this graph, they are guaranteed to be independent.
- If the variables are **connected** in this graph, they are not guaranteed to be independent.*
  Note that "are connected" means "have a path between them," so if we have a path X-Y-Z, X and Z are considered to be connected, even if there's no edge between them.
- If one or both of the variables are missing (because they were givens, and were therefore deleted), they are independent.

* We can say "the variables are dependent, as far as the Bayes net is concerned" or "the Bayes net does not require the variables to be independent," but we cannot guarantee dependency using d-separation alone, because the variables can still be numerically independent (e.g. if P(A|B) and P(A) happen to be equal for all values of A and B).

Practicing with the d-separation algorithm will eventually let you determine independence relations more intuitively. For example, you can tell at a glance that two variables with no common ancestors are marginally independent, but that they become dependent when given their common child node.

Here are some examples of questions we can answer about the Bayes net below, using d-separation:

1. Are A and B conditionally independent, given D and F?

(Same as "P(A|BDF) =? P(A|DF)" or "P(B|ADF) =? P(B|DF)")

2. Are A and B marginally independent? (Same as "P(A|B) =? P(A)" or "P(B|A) =? P(B)")

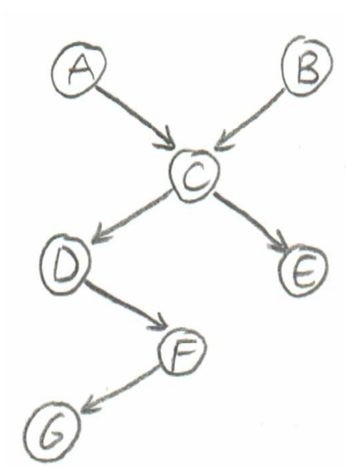
3. Are A and B conditionally independent, given C?

4. Are D and E conditionally independent, given C?

5. Are D and E marginally independent?

6. Are D and E conditionally independent, given A and B?

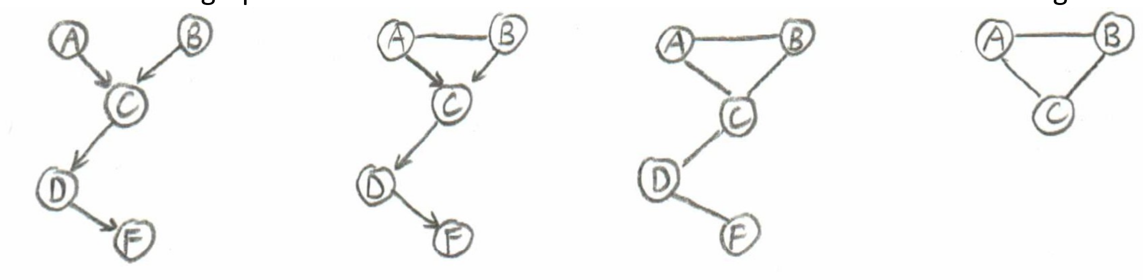
7. P(D|BCE) =? P(D|C)



Solutions are on the following pages.

1. Are A and B conditionally independent, given D and F?
(Same as “P(A|BDF) =? P(A|DF)” or “P(B|ADF) =? P(B|DF)”)

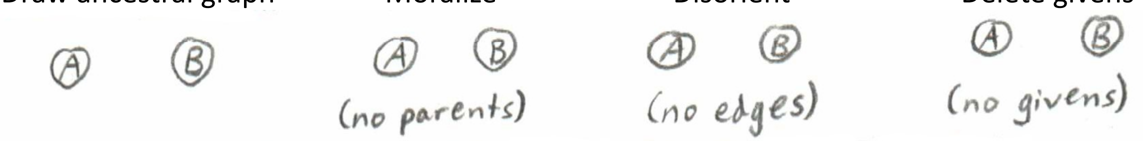
Draw ancestral graph | Moralize | Disorient | Delete givens



Answer: No, A and B are connected, so they are not required to be conditionally independent given D and F.

2. Are A and B marginally independent? (Same as “P(A|B) =? P(A)” or “P(B|A) =? P(B)”)

Draw ancestral graph | Moralize | Disorient | Delete givens



Answer: Yes, A and B are not connected, so they are marginally independent.

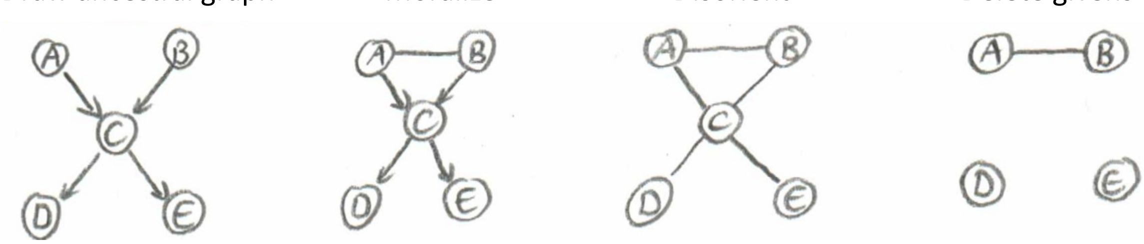
3. Are A and B conditionally independent, given C?

Draw ancestral graph | Moralize | Disorient | Delete givens

Answer: No, A and B are connected, so they are not required to be conditionally independent given C.

4. Are D and E conditionally independent, given C?

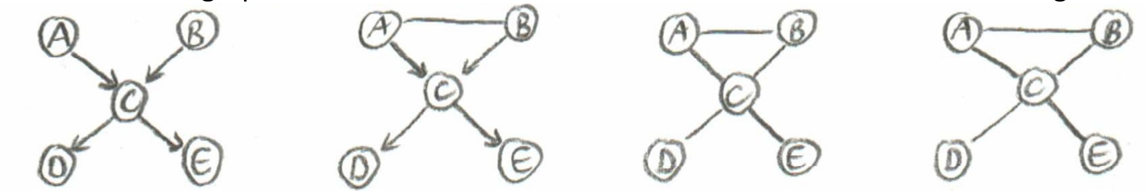
Draw ancestral graph | Moralize | Disorient | Delete givens



Answer: Yes, D and E are not connected, so they are conditionally independent given C.

5. Are D and E marginally independent?

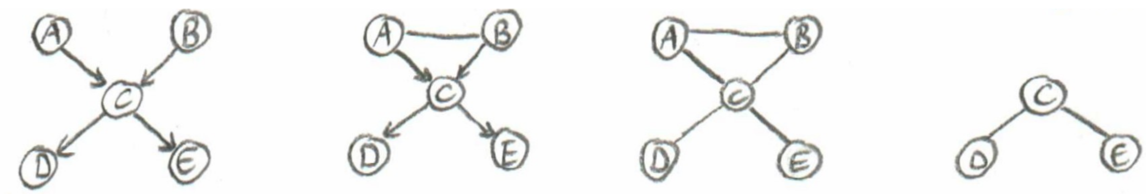
Draw ancestral graph | Moralize | Disorient | Delete givens



Answer: No, D and E are connected (via a path through C), so they are not required to be marginally independent.

6. Are D and E conditionally independent, given A and B?

Draw ancestral graph | Moralize | Disorient | Delete givens



Answer: No, D and E are connected (via a path through C), so they are not required to be conditionally independent given A and B.

7. P(D|CEG) =? P(D|C)

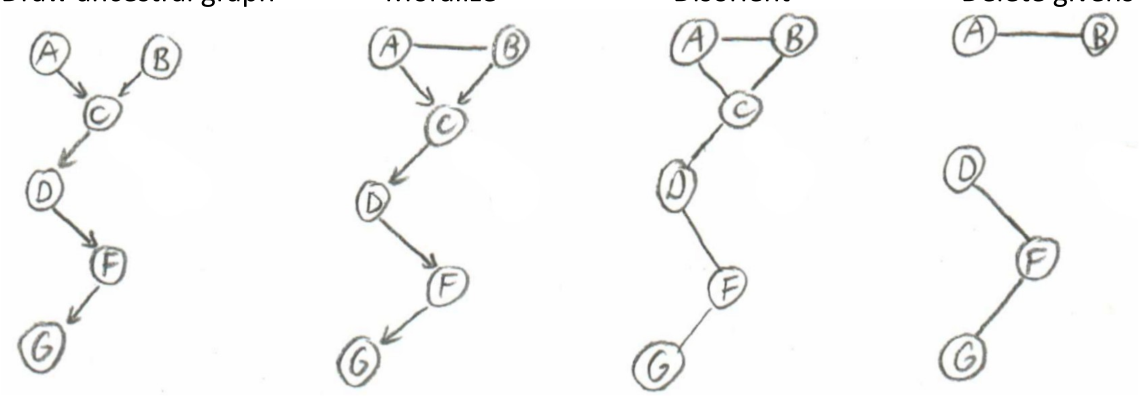
Rewrite as independence questions "Are X and Y conditionally independent, given {givens}?":

- Are D and E conditionally independent, given C? AND
- Are D and G conditionally independent, given C?

(a) Are D and E conditionally independent, given C? Yes; see example 4.

(b) Are D and G conditionally independent, given C? No, because they are connected (via F):

Draw ancestral graph | Moralize | Disorient | Delete givens



Overall answer: No. D and E are conditionally independent given C, but D and G are not required to be. Therefore we cannot assume that P(D|CEG) = P(D|C).