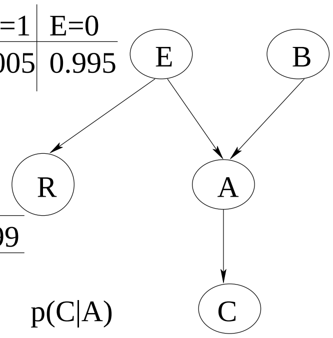
Lecture 3: Belief networks. Bayes ball

- An example
- Conditional independencies implied by a belief network
- The Bayes ball algorithm

Recall from last time

- Conditional independence is an important tool for making probability distributions tractable
- Two random variables X and Z are conditionally independent *given* Y if, once we know Y, knowing Z does not reduce our uncertainty in the value of X, and vice versa.
- Bayesian networks are a graphical representation of conditional independence using directed acyclic graphs (DAGs)

Example: A Belief (Bayesian) Network



p(E)

E=1	E=0
0.005	0.995

p(B)

B=1	B=0
0.01	0.99

p(R|E)

	R=1	R=0
E=0	0.0001	0.9999
E=1	0.65	0.35

p(A|B,E)

	A=1	A=0
B=0,E=0	0.001	0.999
B=0,E=1	0.3	0.7
B=1,E=0	0.8	0.2
B=1,E=1	0.95	0.05

p(C|A)

	C=1	C=0
A=0	0.05	0.95
A=1	0.7	0.3

- The nodes represent random variables
- The arcs represent “influences”
- At each node, we have a conditional probability table (CPD) for the corresponding variable *given its parents*

Using a Bayes net for reasoning (1)

Computing any entry in the joint probability table is easy:

$$p(B, \neg E, A, C, \neg R) = p(B)p(\neg E)p(A|B, \neg E)p(C|A)p(\neg R|lnotE)$$
$$= 0.01 \cdot 0.995 \cdot 0.8 \cdot 0.7 \cdot 0.9999 \approx 0.0056$$

What is the probability that a neighbor calls?

$$p(C = 1) = \sum_{e,b,r,a} p(C = 1, e, b, r, a) = ...$$

What is the probability of a call in case of a burglary?

$$p(C = 1|B = 1) = \frac{p(C = 1, B = 1)}{p(B = 1)} = \frac{\sum_{e,r,a} p(C = 1, B = 1, e, r, a)}{\sum_{c,e,r,a} p(c, B = 1, e, r, a)}$$

This is **causal reasoning** or **prediction**

Using a Bayes net for reasoning (2)

Suppose we got a call. What is the probability of a burglary? What is the probability of an earthquake?

$$p(B|C) = \frac{p(C|B)p(B)}{p(C)} = ...$$
$$p(E|C) = \frac{p(C|E)p(E)}{p(C)} = ...$$

This is **evidential reasoning** or **explanation**

What happens to the probabilities if the radio announces an earthquake?

$p(E|C,R) \gg p(E|C)$ and $p(B|C,R) \ll p(B|C)$

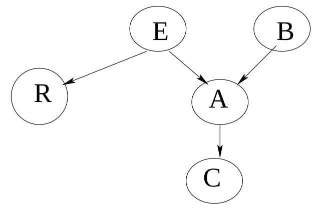
This is called **explaining away**

Using DAGs to represent independencies

- Graphs have been proposed as models of human memory and reasoning on many occasions (e.g. semantic nets, inference networks, conceptual dependencies)
- There are many efficient algorithms that work with graphs, and efficient data structures

DAGs and independencies

- Given a graph G, what sort of independence assumptions does it imply? E.g. Consider the alarm network:

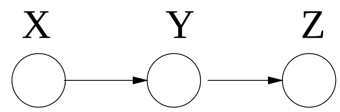


- In general the *lack of an edge* corresponds to lack of a variable in the conditional probability function.
- But there are other independencies between variables as well E.g. In the alarm network, we have $E \perp\!\!\!\perp B$, $R \perp\!\!\!\perp \{B, A, C\} | E$ and $C \perp\!\!\!\perp \{E, B, R\} | A)$. How about node A?

Implied independencies

- Independencies are important because they can help us answer queries more efficiently
- E.g. Suppose that we want to know the probability of a radio report given that there was a burglary. Do we really need to sum over all values of A, C, E?
- Given a Bayes net structure G, and given values for evidence variable Z, what can we say about the sets of variables X and Y?
- Intuitively, the evidence will propagate along paths in the graph, and if it reaches both X and Y, then they are <u>*not*</u> independent.

A simple case: Indirect connection



- We interpret the lack of an edge between X and Z as a conditional independence, $X \perp\!\!\!\perp Z | Y$. Is this justified?
- Based on the graph structure, we have:

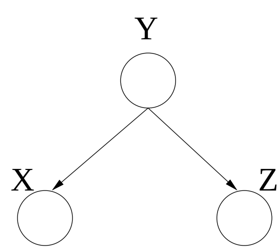
$$p(X, Y, Z) = p(X)p(Y|X)p(Z|Y)$$

- Hence, we have:

$$p(Z|X, Y) = \frac{p(X, Y, Z)}{p(X, Y)} = \frac{p(X)p(Y|X)p(Z|Y)}{p(X)p(Y|X)} = p(Z|Y)$$

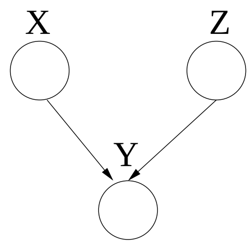
- Note that the edges that are *present* do not imply dependence. But the edges that are *missing* do imply independence.

A more interesting case: Common cause



- Again, we interpret the lack of edge between X and Z as $X \perp\!\!\!\perp Z | Y$. Why is this true?
- This is a “hidden variable” scenario: if Y is unknown, then X and Z could appear to be dependent on each other

The most interesting case: V-structure



- In this case, the lacking edge between X and Z is a statement of *marginal independence*: $X \perp\!\!\!\perp Z$.
- In this case, once we know the value of Y, X and Z might depend on each other.
- This is the case of “explaining away” when there are multiple, competing explanations.
- Note that in this case, X is *not* independent of Z given Y!

Bayes ball algorithm

- Suppose we want to decide whether $X \perp\!\!\!\perp Z | Y$ for a general Bayes net with corresponding graph G.
- We shade all nodes in the evidence set, Y
- We put balls in all the nodes in X, and we let them bounce around the graph according to rules inspired by these three base cases
- Note that the balls can go in any direction along an edge!
- If any ball reaches any node in Z, then the conditional independence assertion is *not* true.