Cognitive Modeling

Lecture 15: Bayes Nets

Sharon Goldwater

School of Informatics
University of Edinburgh
sgwater@inf.ed.ac.uk

March 5, 2010

Reading: Charniak (1991).

Slides are based on a tutorial held by J. Tenenbaum and T. Griffiths at the 26th Annual Conference of the Cognitive Science Society, 2004.

Motivation

Many tasks humans perform involve reasoning and prediction in complex domains with many variables.

- Simple inference: given results of a single test and no other information about patient, does patient have disease X?
- Complex inference: given several observed symptoms, test results, and history, which disease does patient have?

Also, we make judgments about causation:

- 3 out of the last 5 times I ate chocolate, I got a headache. Does chocolate give me a headache?

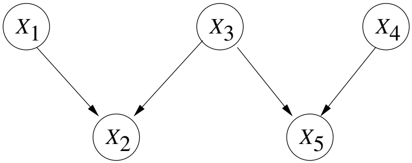
Bayesian networks are a way of representing complex probabilistic relationships and reasoning about causation.

Graphical Models

Bayesian networks are a type of *graphical model* consisting of:

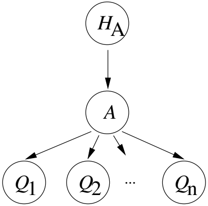
- a set of nodes, corresponding to variables;
- a set of directed edges, indicating dependencies;
- a *conditional probability distribution* for each node conditioned on its parents; multiplied together, these yield the joint distribution over all variables.

Bayes nets take the form of directed acyclic graphs (DAGs):



Bayes Nets

Bayes nets are a way to represent probabilistic models. E.g,



represents Anderson's (1990) rational model of memory:

$$\text{need probability of A} = P(A|H_A) \prod_i P(Q_i|A)$$

Bayes Nets and Bayesian Statistics

Bayes nets and Bayesian statistics *solve two different problems:*

- Bayesian statistics is a method of inference;
- Bayes nets are a form of representation.

There is *no necessary connection* between the two:

- many users of Bayes nets rely upon frequentist statistical methods;
- many Bayesian inferences cannot be easily represented using Bayes nets.

Properties of Bayes Nets

Properties of Bayes nets (Pearl 1988):

- *efficient representation and inference:* exploiting dependency structure makes it easier to represent and compute with probabilities;
- *explaining away:* pattern of probabilistic reasoning characteristic of Bayes nets.

The efficiency of Bayes net is due to the *Markov assumption* they make: conditioned on its parents, the value of each node is independent of all other ancestors.

$$P(child|parents, grandparents, \ldots) = P(child|parents)$$

Conditional Independence

Let's assume we have three binary variables:

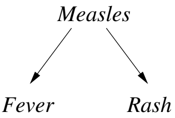
- M: patient has measles;
- R: patient has rash;
- F: patient has fever.
- We'll use m for $M = 1$, $\neg m$ for $M = 0$, etc.

All three variables are dependent, but R and F are independent once we know the value of M: a *conditional independence assumption:*

$$P(R, F|M) = P(R|M)P(F|M)$$

Joint Distribution

A Bayes net is a graphical representation of the (in)dependencies among a set of random variables. The Bayes net for the previous example is:



The Bayes net specifies a *factorization* of the joint distribution of all the variables:

$$P(V_1 \dots V_n) = \prod_{V_i} P(V_i|\text{parents}(V_i))$$

In our example:

$$P(M, F, R) = P(M)P(F|M)P(R|M)$$

Efficient Representation

A factorized distribution requires fewer parameters to specify.

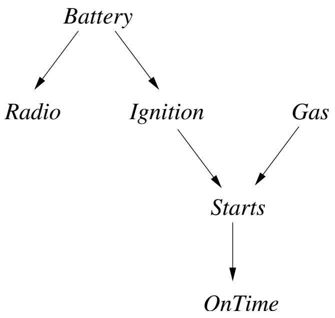
- Specifying $P(M, F, R)$ requires 7 parameters: one for each set of values, minus one because distribution sums to 1.
- Using Bayes net and conditional independencies, requires only 5 parameters: $P(m), P(r|m), P(r|\neg m), P(f|m), P(f|\neg m)$.
- In general, a distribution with n binary variables has $2^n - 1$ parameters, while a Bayes net may have as few as $2n - 1$.

This efficiency is useful in *expert systems*, which aim to capture human knowledge in complex domains.

Example

A more complex example:

Battery	battery is charged
Radio	radio works
Ignition	ignition works
Gas	there's gas in the tank
Starts	car starts
OnTime	I'm on time for work



$$P(B, R, I, G, S, O) = P(B)P(R|B)P(I|B)P(G)P(S|I, G)P(O|S)$$

Example

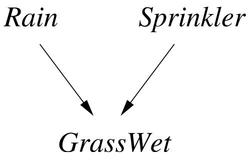
Knowing the joint distribution is sufficient for any inference in the Bayes net. For example, we would like to compute $P(O|G)$:

$$\begin{aligned} P(O|G) &= \frac{P(O,G)}{P(G)} \\ &= \sum_{B,R,I,S} \frac{P(B)P(R|B)P(I|B)P(G)P(S|I,G)P(O|S)}{P(G)} \\ &= \sum_{B,I,S} P(B)P(I|B)P(S|I,G)P(O|S) \end{aligned}$$

- $P(R|B)$ can be eliminated based on rules of *d-separation* (see Charniak, 1991)
- Often in larger nets, most terms can be eliminated.

Explaining Away

Given the following Bayes net:



The joint probability distribution is:

$$P(R, S, W) = P(R)P(S)P(W|R, S)$$

Assume grass will be wet if and only if it rained last night or the sprinklers were left on:

$$P(w|s, r) = P(w|\neg s, r) = P(w|s, \neg r) = 1$$

$$P(w|\neg s, \neg r) = 0$$

Explaining Away

Compute probability it rained last night, given that the grass is wet:

$$
\begin{aligned}
P(r|w) &= \frac{P(w|r)P(r)}{P(w)} = \frac{P(w|r)P(r)}{\sum_{R,S} P(w|R,S)P(R,S)} \\
&= \frac{P(r)}{P(r,s) + P(r,\neg s) + P(\neg r, s)} \\
&= \frac{P(r)}{P(r) + P(\neg r, s)} = \frac{P(r)}{P(r) + P(\neg r)P(s)}
\end{aligned}
$$

The term $P(r) + P(\neg r)P(s)$ varies between 1 and $P(s)$, therefore $P(r|w) > P(r)$.

The probability that it rained given that the grass is wet is larger than the probability that it rained.

Explaining Away

Now compute probability it rained last night, given that the grass is wet and the sprinklers were left on:

$$P(r|w,s) = \frac{P(w|r,s)P(r|s)}{P(w|s)}$$

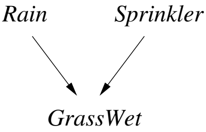
Since $P(w|r,s) = 1$ and $P(w|s) = 1$:

$$P(r|w,s) = P(r|s) = P(r)$$

The probability that it rained given that the grass is wet and the sprinklers were left on is the same as the probability that it rained.

Knowing that s occurred *explains away* the occurrence of w, so the alternative cause is no longer necessary as an explanation.

Comparison with Production Rules



Formulate production rules for reasoning from *Wet* to *Rain*:

IF *Rain* THEN *Wet*

But how do we reason from effects to causes? Maybe add:

IF *Wet* THEN *Rain*

This fails to distinguish the direction of the inference. Instead we could use:

IF *Wet* AND NOT *Sprinkler* THEN *Rain*

But this leads to a combinatorial explosion of rules.

Causation vs. Correlation

Graphical models represent statistical dependencies among variables (conditional probabilities):

- this models *correlations* in the data;
- allows us to answer questions about *observations.*

Causal graphical models represent causal dependencies among variables (Pearl 2000):

- this models the underlying causal structure;
- allows us to answer questions about *interventions.*

The two kinds of models may look the same, but interpretation of arrows is different.

Interventions

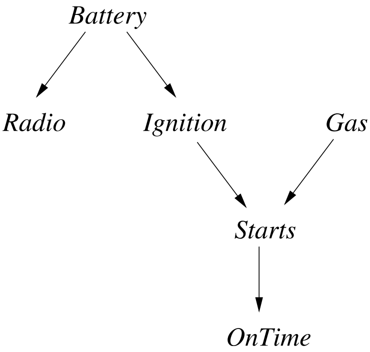
Intervention: change the value of a variable from the outside:

- if two variables A and B are causally related, then intervening to change the value of A will also change the value B;
- causal Bayes nets predict the effects of interventions on a causal structure;
- causes Bayes nets capture evidence from observations and interventions in a single structure.

Technically, interventions work by changing the graph structure of the Bayes net.

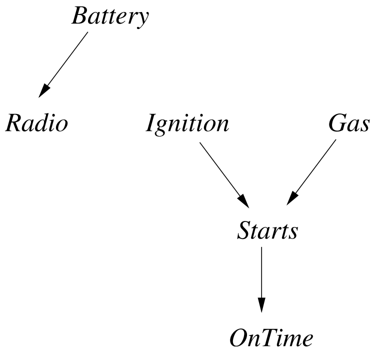
Interventions

Graphical model: $P(Radio|Ignition)$



Interventions

Graphical model: $P(Radio|Ignition)$
Causal graphical model: $P(Radio|\mathrm{do}(Ignition))$



Intervention is *"graph surgery":* it produces a "mutilated" graph that we can then reason with.

Assessing Interventions

Intervention as graph surgery:

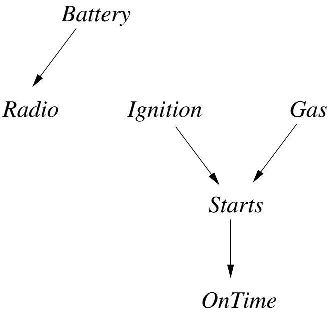
- model an intervention on variable X, remove all edges into X and leave all other edges intact;
- to determine whether an intervention on X changes Y, check whether there is a path from X to Y in the mutilated graph.

More formally:

- intervention probability $P(Y|\mathrm{do}(X = x))$: the probability of Y given that we intervene to set variable X to value x;
- to compute $P(Y|\mathrm{do}(X = x))$, delete all edges coming into X and compute $P(Y|X = x)$ for resulting Bayes net.

This makes it possible to use a single structure to make predictions about *both observations and interventions.*

Assessing Interventions



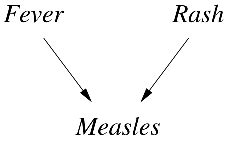
We intervene to start the ignition. The edge leading to *Ignition* is deleted:

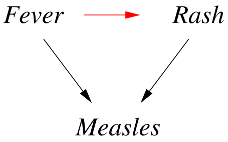
- path from *Ignition* to *Starts* and *OnTime*: *Ignition* causally affects these two variables;
- no path from *Battery* to *Ignition*: *Battery* doesn't causally affect *Ignition*;
- other causal links (e.g., from *Battery* to *Radio*) are preserved.

Causality Simplifies Inference

Causality simplifies inference:

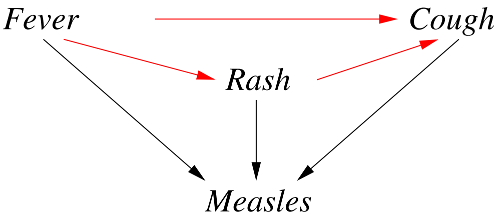
- using a representation in which the direction of causality is correct produces sparser graphs;
- suppose we get the direction of causality wrong, thinking that symptoms causes diseases;
- the model doesn't capture the correlation between symptoms; we can fix this by adding a *new arrow;*
- but the new model is too complex; also, no more explaining away is possible.





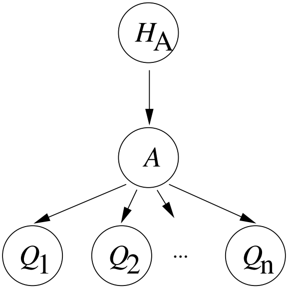
Causality Simplifies Inference

New symptoms require a combinatorial proliferation of new arrows. This reduces efficiency of inference:



Limitations of Causal Models

Not all Bayes nets can be easily modified into causal graphs:



However, non-causal models can still be useful.

Summary

- Bayes nets are directed graphical models in which the edges represent dependencies;
- Markov assumption (conditional independence) allows efficient representation and inference;
- explaining away: $P(a|b) > P(a|b, c)$;
- causal graphical models assume edges represent causation, with interventions as graph surgery;
- causality simplifies model structure but not always possible.

Next class: more on causal models and humans.

References

Anderson, John R. 1990. *The Adaptive Character of Thought*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Charniak, Eugene. 1991. Bayesian networks without tears. *AI Magazine* 12(4):50–63.

Pearl, Judea. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, CA.

Pearl, Judea. 2000. *Causality: Models, Reasoning and Inference*. Cambridge University Press, Cambridge.