Bayes Nets

Independence

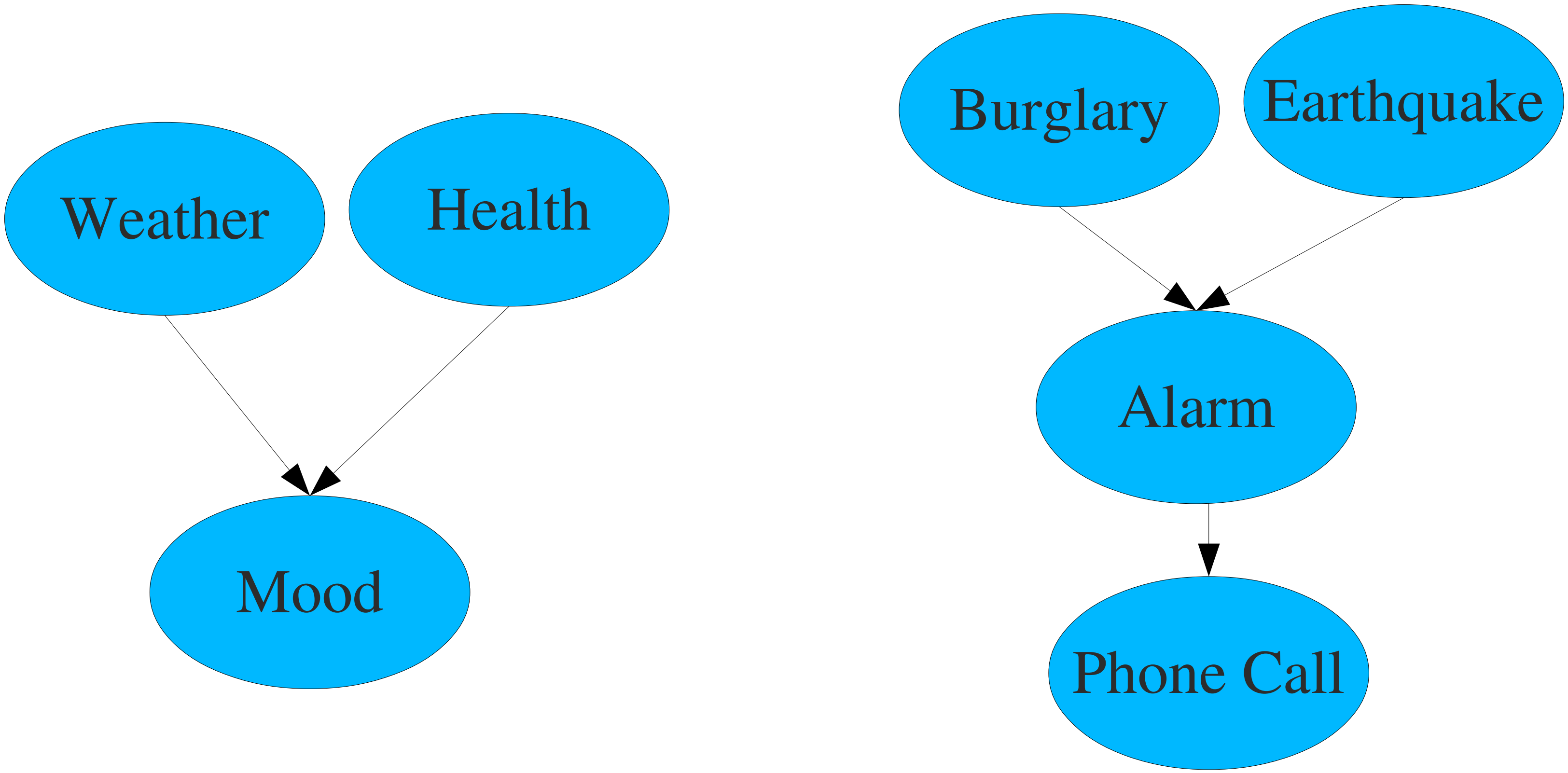
- With joint probability distributions we can compute many useful things, but working with joint PD's is often intractable.
- The naïve Bayes' approach represents one (boneheaded?) solution – assume every piece of evidence is independent given a hypothesis.
- Maybe we could do better if we could *really* characterize which variables are independent of which other variables.
- For example my health is independent of the weather, but my mood depends on both weather and health.

Graphical Models

- This is the idea behind graphical models.
- We use graph structures to encode our assumptions about independence relations between random variables.
- One example is the Bayesian Network...

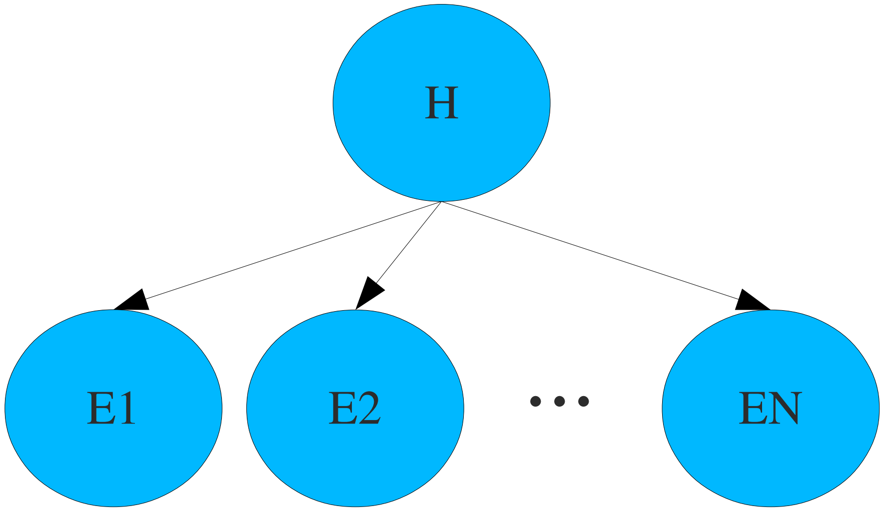
Bayesian Networks

- A Bayesian network is a directed acyclic graph that represents causal relationships between random variables.



An Aside: Naïve Bayes'

- Our naïve Bayes' classifier can be represented as a Bayes' net.



Bayesian Networks

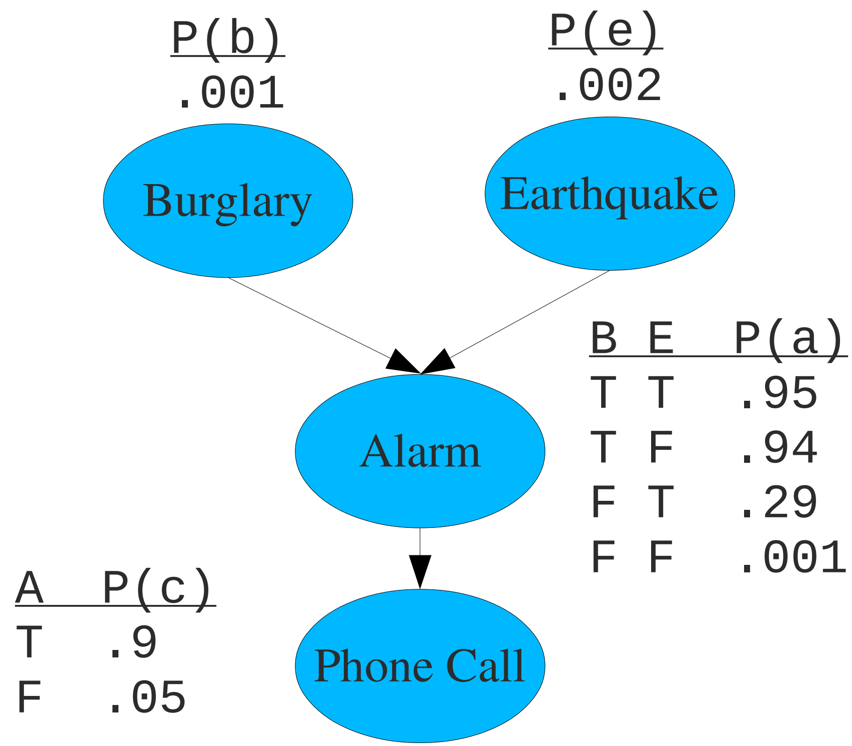
- Bayes' nets have the following property:
  - Each variable is conditionally independent of all it's non-descendants in the graph, given the value of it's parents.
- This implies that:

$$P(X_1, X_2, \ldots, X_N) = \prod_{i=1}^{N} P(X_i | parents(X_i))$$

- In other words, the complete joint probability distribution can be reconstructed from the *N* conditional distributions.
- For *N* binary valued variables with *M* parents each
  - 2^N vs. $N * 2^M$

Specifying a Bayes' Net

- We need to specify:
  - The topology of the network.
  - The conditional probabilities.



Inference In Bayes' Nets

- We build Bayes' nets to answer questions.
- For example, if there is a phone call what is the probability that there was a burglary, P(b | c)?
- We already know how to do this.
  - Generate all entries in the joint PD.
  - Compute:

$$P(b|c) = \frac{\sum_{\text{entries matching } b \text{ and } c} P(entries)}{\sum_{\text{entries matching } c} P(entries)}$$

- Remember that:

$$P(b|c) = \frac{P(b \wedge c)}{P(c)}$$

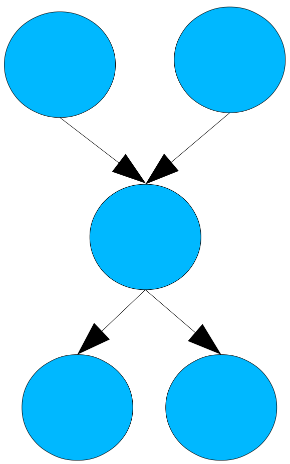
Time Complexity of Inference

- This is not very efficient.
- The bad news:
  - We can save some space.
  - We can save some time.
  - But, in general, inference in Bayes' nets is NP-Complete.

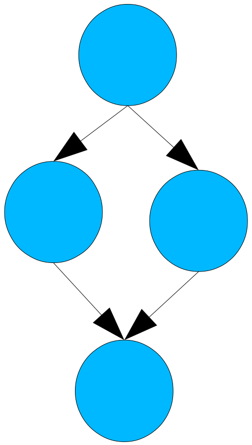
The Good News

- If our network is a polytree inference is efficient – linear in the size of the network.
- Polytree – at most one undirected path between any two nodes.

polytree

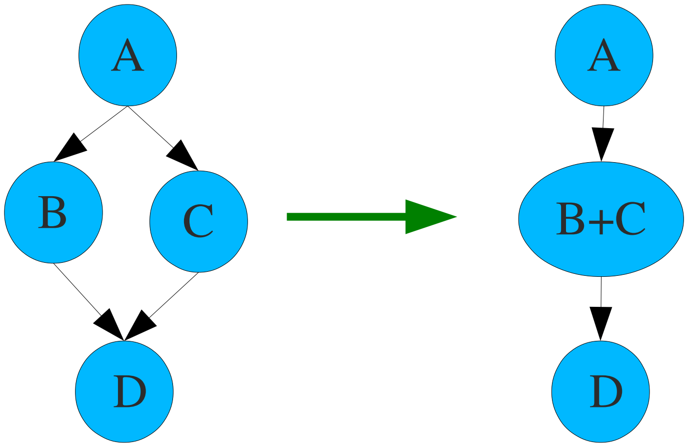


non-polytree



More Good News

- If a network is *almost* a polytree we can convert it to a polytree and regain efficient inference:



- If B and C were boolean random variables, B+C is a new random variable that has four possible values.

Generative Models

- We have seen the process of learning probability distributions from data.
- If we have a PD, we can use it to generate data.
- Models that have this property are called generative models.
- Example use: approximate inference in Bayes' nets.
- If exact inference is too expensive, generate samples, and use them to estimate the desired probability.

Sampling From a Bayes' Net

- Let's say we have a Bayes' net, and we want to generate data that is consistent with the implicit joint probability distribution.
- 1 sample = 1 assignment to all variables.
- Sort nodes topologically
- For variables with no parents:
  - Assign a value to the variable according to the prior distribution.
- For nodes with parents:
  - once the parents have been assigned values, assign values to children according to the conditional probability table.

Rejection Sampling

- The preceding does not apply to the case where some variables are given an assignment.
- For example P(Burglary=TRUE | Call=TRUE).
- The simplest fix: use the same procedure, but throw out all instances where Call=FALSE.
- Estimate P(Burglary=TRUE | Call=TRUE) by counting the number of times Burglary=TRUE in the remaining instances.
- Problem: wastes many samples.

More Efficient Alternatives

- Likelihood weighting
- Markov Chain Monte Carlo (MCMC)/Gibbs Sampling
- We may talk about these at some point in the future.