Graphical Models: Bayesian Networks

Ira A. Fulton Schools of Engineering
Arizona State University

Objectives

Objective

Describe Bayesian Networks

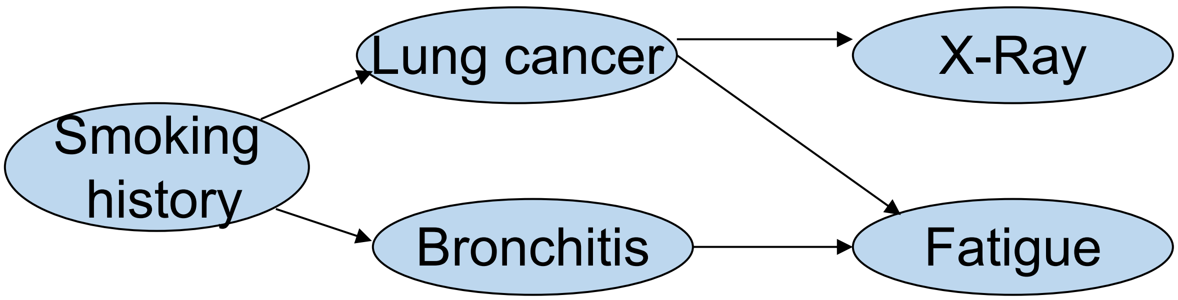
Objective

Illustrate key tasks in implementing Bayesian Networks

Why do we use graphical models?

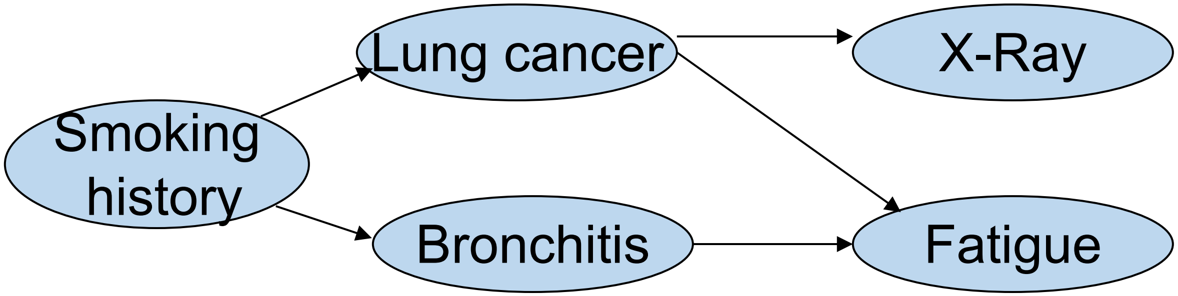
- In machine learning, we are often concerned with joint distributions of many random variables.

- A graph may provide an intuitive way of representing or visualizing the relationships of the variables.
  - Making it easier for domain experts to build a model



Graphical models for Causal Relations

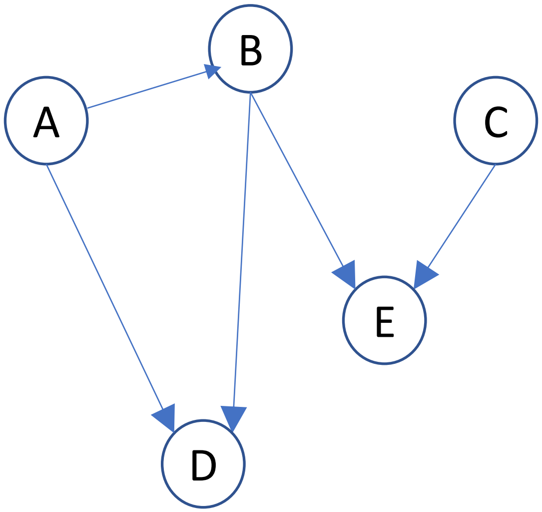
- Graphical models arise naturally from, often causal, independency relations of physical events.



- Caveat: probabilistic relationship does not imply causality.

Bayesian Networks

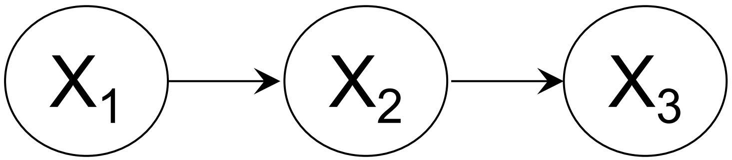
- A BN is directed acyclic graph (DAG), where
  - Nodes (vertices) represent random variables.
  - Directed edges represent immediate dependence of nodes.



- Other names: Belief networks, Bayes nets, etc.

Conditional Independence

- E.g., given the following graph, check the relationship between X_3 and X_1



  - X_3 is dependent of X_2, and X_2 is dependent of X_1
  - Thus X_3 is dependent of X_1
  - But given X_2, X_3 is dependent of X_1

➔ Conditional Independence

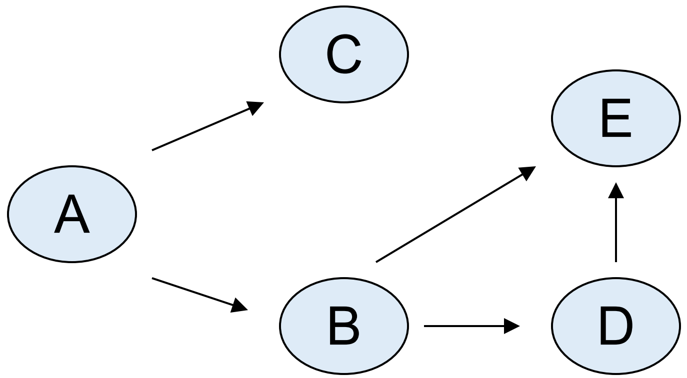
BN for General Conditional Dependency

- A BN can be used to model given conditional dependencies
  - For example, using the *chain rule of probability*, we have

$$P(A,B,C,D,E)=P(A)P(B|A)P(C|A,B)P(D|A,B,C)P(E|A,B,C,D)$$

- If we know that, given A, C won't rely on B, and so forth, we may have

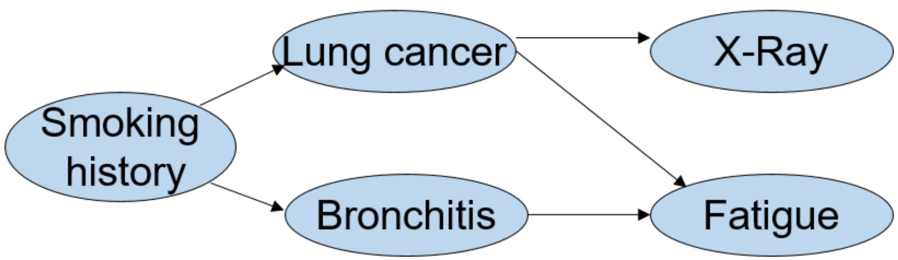
$$P(A,B,C,D,E)=P(A)P(B|A)P(C|A)P(D|B)P(E|B,D)$$



➢ We could represent joint distributions more compactly in BN → Efficient computation

Inference in Bayesian Networks

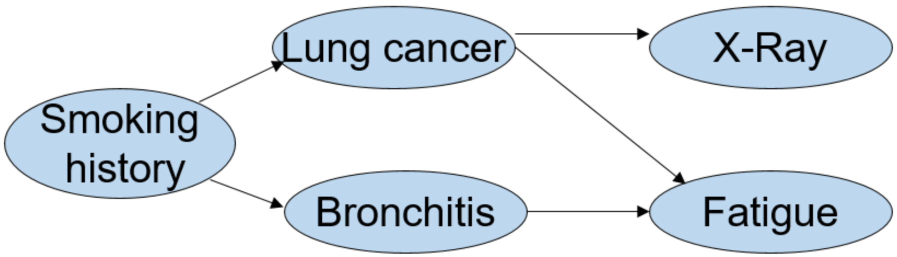
- Given a model and some data ("evidence"), how to update our belief?



What are the model parameters?

Inference in Bayesian Networks (cont'd)

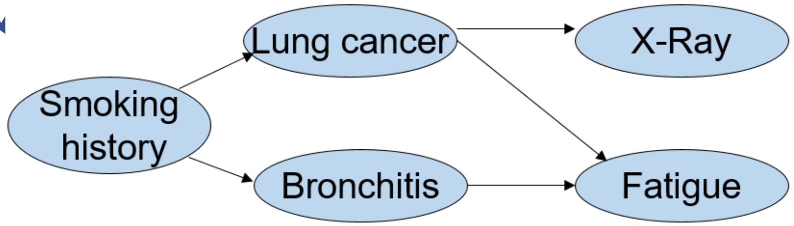
- Given a model and some data ("evidence"), how to update our belief?



- E.g., for a patient with certain smoking history (non-smoker), whose X-ray result is positive, and who does not experience fatigue:
  - What is probability of having lung cancer?

Inference in Bayesian Networks (cont'd)

- In a simple BN like this, we can compute the exact probabilities.



- In general, for a tree-structured BN, we may use **belief propagation** for the inference problem.

- For general structures, sometimes it is possible to generalize the above method (e.g., the *junction tree algorithm*). More often, we must resort to approximation methods
  - E.g. Variational methods, Sampling (Monte Carlo) methods.

Learning in Bayesian Networks

- Learning parameters (probabilities) for a given BN (the graph is given).
  - Estimate the (conditional) probabilities from past data.

- Learning both the structure and the parameters for a BN
  - A more challenging task beyond the scope of this discussion.

Learning the Probabilities

- Basic ideas
  - Use relative frequency for estimating probability.
  - A prior distribution is typically assumed.
  - The prior is then updated by the data into posterior.
  - Using the MLE principle
- The so-called “Expectation-Maximization (EM) Algorithm” is often used.
  - Iteratively update our guess for the parameter and each step attempts to apply the MLE principle.