

Introduction to Artificial Intelligence

Unit # 10

Recap

- Joint Distribution: A joint probability gives us a full picture about how random variables are related.
- Marginal Distribution: Marginalization lets us to focus one aspect of the picture.
- Prior probability: Belief about a hypothesis h before obtaining observations.
 - Example: Suppose 10% of people suffer from Hepatitis B. A doctor's prior probability about a new patient suffering from Hepatitis B is 0.1.
- Posterior probability: Belief about a hypothesis after obtaining observations.

Bayes Theorem

- $P(A | B) = \frac{P(B | A) P(A)}{P(B)}$

$$= \frac{P(B | A) P(A)}{P(B | A)P(A) + P(B | \neg A)P(\neg A)}$$
- $P(A)$ is the prior probability and $P(A | B)$ is the posterior probability.
- Suppose events A_1, A_2, \dots, A_k are mutually exclusive and exhaustive; i.e., exactly one of the events must occur. Then for any event B :

$$P(A_i | B) = \frac{P(B | A_i) P(A_i)}{\sum P(B | A_i) P(A_i)}$$

Probabilistic Reasoning

- Logic is used to represent and reason knowledge about the world with facts and rules.
 - `bird(tweety).`
 - `fly(X) :- bird(X).`
- In many cases, certain rules are hard to find.
- Probability theory is the most popular way of representing uncertainty.
 - `lung_cancer(X) :- smoking(X).` is replaced by
 - $P(\text{Lung Cancer} | \text{Smoking}) = 0.6$

Bayes Theorem Example

- Suppose that Bob can decide to go to work by one of three modes of transportation, car, bus, or commuter train. Because of high traffic, if he decides to go by car, there is a 50% chance he will be late. If he goes by bus, which has special reserved lanes but is sometimes overcrowded, the probability of being late is only 20%. The commuter train is almost never late, with a probability of only 1%, but is more expensive than the bus.
- Suppose that Bob is late one day, and his boss wishes to estimate the probability that he drove to work that day by car. Since he does not know which mode of transportation Bob usually uses, he gives a prior probability of $1/3$ to each of the three possibilities. What is the boss' estimate of the probability that Bob drove to work?

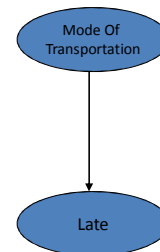
Sajjad Haider

Spring 2010

5

Bayes Theorem Example

- Suppose that a coworker of Bob's knows that he almost always takes the commuter train to work, never takes the bus, but sometimes, 10% of the time, takes the car. What is the coworker's probability that Bob drove to work that day, given that he was late?



Sajjad Haider

Spring 2010

6

Earthquake Example (Pearl)

- You have a new burglar alarm installed.
- It is reliable about detecting burglary, but responds to minor earthquakes.
- Two neighbors (John, Mary) promises to call you at work when they hear the alarm.
 - John always call when hears alarm, but confuses alarm with phone ringing (and calls then also)
 - Mary likes loud music and sometimes misses alarm
- Given evidence about who has and hasn't called, estimate the probability of a burglary.

Sajjad Haider

Spring 2010

7

Cause and Effect Relationships

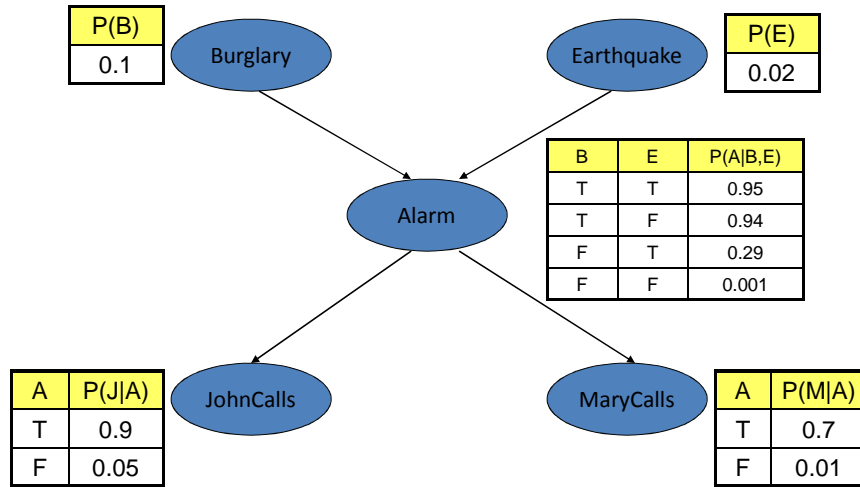
- Burglary → Alarm Goes Off
- Earthquake → Alarm Goes Off
- Alarm Goes Off → John Calls
- Alarm Goes Off → Mary Calls

- Number of Variables
 - 5
- Number of Links
 - 4

Sajjad Haider

8

Earthquake Example (Pearl)



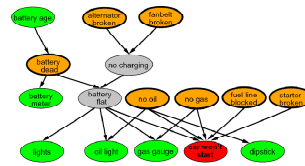
Sajjad Haider

Spring 2010

9

Graphical Models

- Models are descriptions of how (a portion of) the world works
- **Models are always simplifications**
 - May not account for every variable
 - May not account for all interactions between variables
- What do we do with probabilistic models?
 - We need to reason about unknown variables, given evidence
 - Example: explanation (diagnostic reasoning)
 - Example: prediction (causal reasoning)
 - Example: value of information



Sajjad Haider

Spring 2010

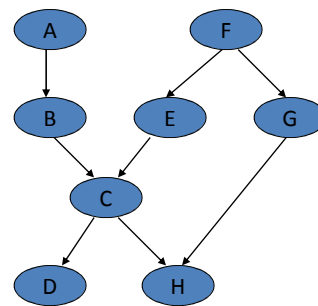
10

Bayesian Networks

- Two problems with using full joint distributions for probabilistic models:
 - Unless there are only a few variables, the joint is WAY too big to represent explicitly
 - Hard to estimate anything empirically about more than a few variables at a time
- **Bayesian Networks** are a technique for describing complex joint distributions using a bunch of simple, local distributions
 - We describe how variables locally interact
 - Local interactions chain together to give global, indirect interactions

Bayesian Networks

- A problem domain is modeled by a list of variables X_1, X_2, \dots, X_n .
- Knowledge about the problem domain is represented by a joint probability $P(X_1, X_2, \dots, X_n)$.
- General probability distribution of 8 variables with 2 states each has $2^8 = 256$ possible values and $2^8 - 1$ probabilities need to be specified.
 - Exponential model size
 - Knowledge acquisition difficult
 - Exponential storage and inference
- A Bayes net is a graphical representation of the joint distribution.
- Assumes that each node is conditionally independent of all its non-descendants given its parents.
- Product of all conditional probabilities is the joint probability of all variables.
 - $P(X_1, X_2, \dots, X_n) = \prod P(X_i \mid \text{parents}(X_i))$



18 probabilities are required to specify the joint distribution

Bayesian Networks

- A BN is a Directed Acyclic Graph (DAG) in which:
 - A set of random variables makes up the nodes in the network.
 - A set of directed links or arrows connects pairs of nodes.
 - Each node has a conditional probability table that quantifies the effects the parents have on the node.
- The intuitive meaning of an arrow from a parent to a child is that the parent directly influences the child.
- The direction of this influence is often taken to represent casual influence.
- These influences are quantified by conditional probabilities.

Inference Problem

- The canonical inference problem: find the posterior probability distribution for some variable(s) given direct or virtual evidence about other variable(s)

Cancer Example (Pearl)

- Metastatic cancer is a possible cause of a brain tumor and is also an explanation for increased total serum calcium. In turn, either of these could explain a patient falling into a coma. Severe headache is also possibly associated with a brain tumor.

Cancer Example (Pearl)

