Introduction to Bayesian Networks

Alice Gao
Lecture 14

Based on work by K. Leyton-Brown, K. Larson, and P. van Beek

Outline

Learning Goals

Examples of Bayesian Networks

Semantics of Bayes Net
- Representing the joint distribution
- Encoding the conditional independence assumptions

Constructing Bayes Nets

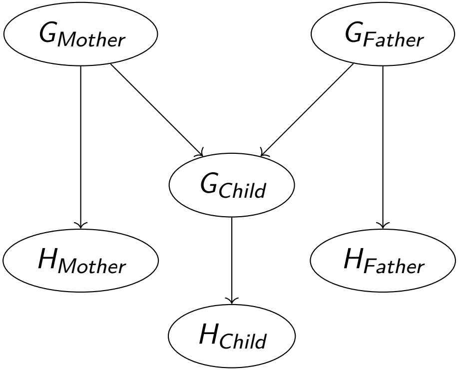
Revisiting the Learning goals

Learning Goals

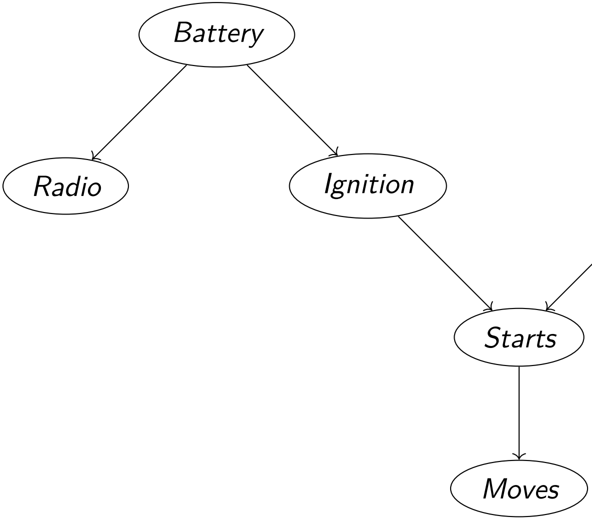
By the end of the lecture, you should be able to

- Compute a joint probability given a Bayesian network.
- Identify the conditional independence assumptions required by a Bayesian network.
- Given a Bayesian network, determine if two variables are independent or conditionally independent given a third variable.
- Given a scenario with independent assumptions and a given order of the variables, construct a Bayesian network by adding the variables to the network based on the given order.

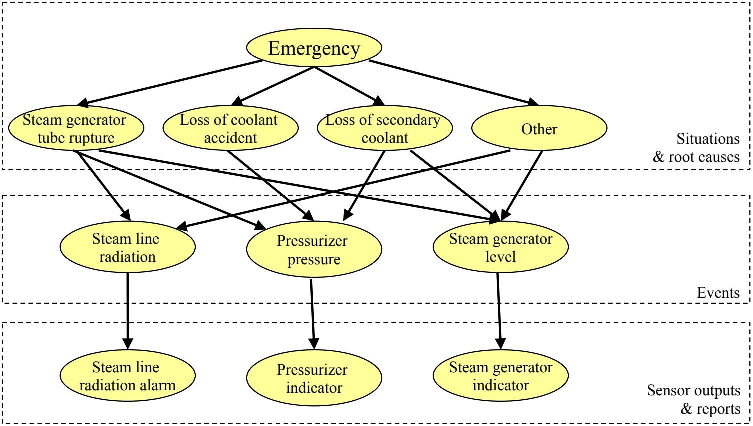
Inheritance of Handedness



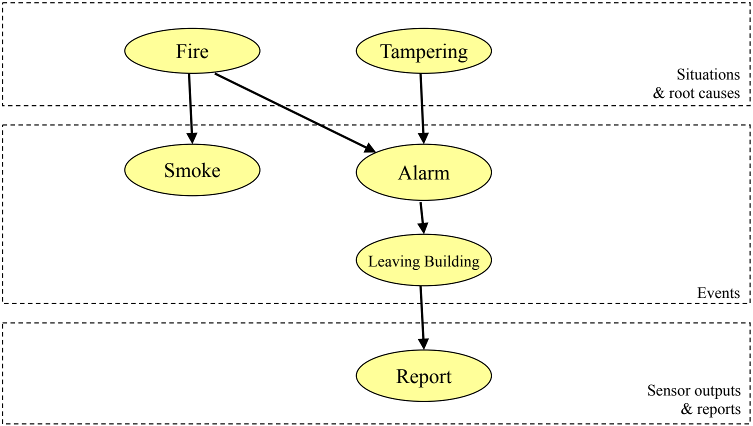
Car Diagnostic Network



Example: Nuclear power plant operations

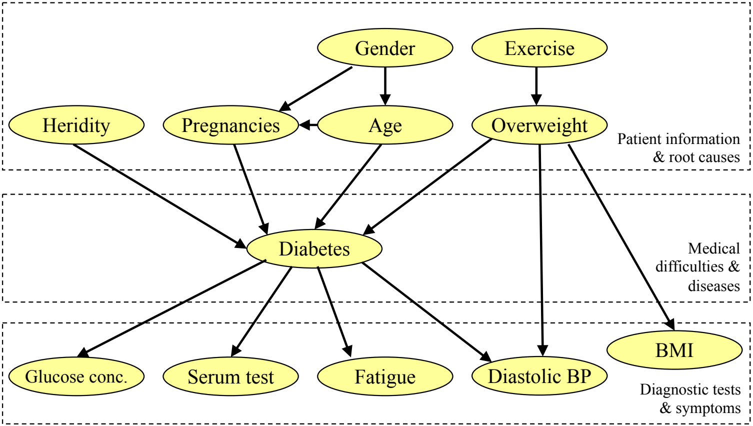


Example: Fire alarms



Report: “report of people leaving building because a fire alarm went off”

Example: Medical diagnosis of diabetes



Why Bayesian Networks?

A probabilistic model of the Holmes scenario:

- The random variables: Earthquake, Radio, Burglary, Alarm, Watson, and Gibbon.
- # of probabilities in the joint distribution: $2^6 = 64$.
- For example,
  $P(E \wedge R \wedge B \wedge A \wedge W \wedge G) = ?$
  $P(E \wedge R \wedge B \wedge A \wedge W \wedge \neg G) = ?$
  ... etc ...

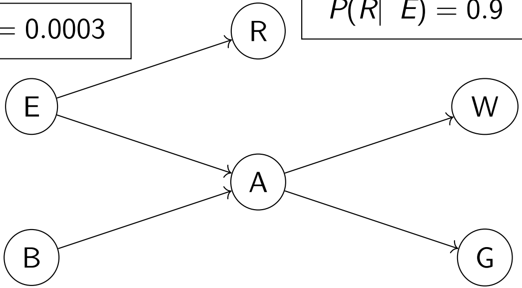
We can answer any question about the domain using the joint distribution, but

- Quickly become intractable as the number of variables grows.
- Unnatural and tedious to specify all the probabilities.

Why Bayesian Networks?

A Bayesian Network is a compact version of the joint distribution and it takes advantage of the unconditional and conditional independence among the variables.

A Bayesian Network for the Holmes Scenario



$$P(E) = 0.0003$$

$$P(B) = 0.0001$$

$$P(R|\neg E) = 0.0002$$
$$P(R|\ \ E) = 0.9$$

$$P(W|\neg A) = 0.4$$
$$P(W|\ \ A) = 0.8$$

$$P(G|\neg A) = 0.04$$
$$P(G|\ \ A) = 0.4$$

$$P(A|\neg B \wedge \neg E) = 0.01$$
$$P(A|\neg B \wedge \ \ E) = 0.2$$
$$P(A|\ \ B \wedge \neg E) = 0.95$$
$$P(A|\ \ B \wedge \ \ E) = 0.96$$

Bayesian Network

A Bayesian Network is a directed acyclic graph.

- Each node corresponds to a random variable.
- X is a parent of Y if there is an arrow from node X to node Y. The graph has no directed cycles.
- Each node X_i has a conditional probability distribution $P(X_i|Parents(X_i))$ that quantifies the effect of the parents on the node.

The Semantics of Bayesian Networks

Two ways to understand Bayesian Networks:

- A representation of the joint probability distribution
- An encoding of the conditional independence assumptions

Representing the joint distribution

We can calculate every entry in the joint distribution using the Bayesian Network. How do we do this?

1. Choose an order of the variables that is consistent with the partial ordering of the nodes in the Bayesian Network.
2. Compute the joint probability using the following formula.

$$P(X_n \wedge \cdots \wedge X_1) = \prod_{i=1}^{n} P(X_i | Parents(X_i))$$

Representing the joint distribution

Example: What is the probability that

- The alarm has sounded,
- Neither a burglary nor an earthquake has occurred,
- Both Watson and Gibbon call and say they hear the alarm, and
- There is no radio report of an earthquake?

CQ: Calculating the joint probability

CQ: What is the probability that

- NEITHER a burglary NOR an earthquake has occurred,
- The alarm has NOT sounded,
- NEITHER of Watson and Gibbon is calling, and
- There is NO radio report of an earthquake?

(A) 0.5699

(B) 0.6699

(C) 0.7699

(D) 0.8699

(E) 0.9699

Encoding the Conditional Independence Assumptions

By modeling a domain using a Bayesian network,
we are making the following key assumption.

*For a given ordering of the nodes,
each node is conditionally independent of its predecessors
given its parents.*

$$P(X_i | Parents(X_i)) = P(X_i | X_{i-1} \wedge \cdots \wedge X_1), \forall i = 1, \ldots, n$$

Identifying The Conditional Independence Assumptions

Given a Bayesian Network,

- Consider all orderings of the variables that are consistent with the partial ordering in the Bayesian network.
- Based on the Bayesian Network,

$$P(X_n \wedge \cdots \wedge X_1) = \prod_{i=1}^{n} P(X_i | Parents(X_i))$$

  Based on the chain rule,

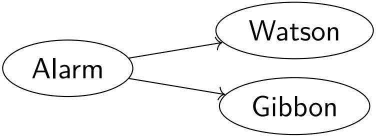
$$P(X_n \wedge \cdots \wedge X_1) = \prod_{i=1}^{n} P(X_i | X_{i-1} \wedge \cdots \wedge X_1)$$

- The difference between the RHS of the equations give the conditional independence assumptions.

Proving Independence and Conditional Independence

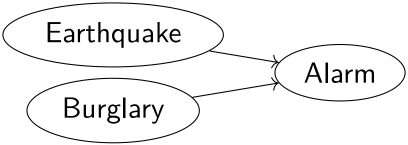
Prove that Watson and Burglary are conditionally independent given Alarm.

Proving Independence and Conditional Independence



Prove that Watson and Gibbon are conditionally independent given Alarm.

Proving Independence and Conditional Independence



Prove that Burglary and Earthquake are independent.

CQ: Conditional Independence

CQ: Are Radio and Gibbon conditionally independent given Earthquake?

(A) Yes

(B) No

(C) I don't know.

Independence Relationships Required by a Bayesian Network

Every variable is conditionally independent of its non-descendants given its parents.

The Markov Blanket

The Markov blanket of a node consists of

- its parents
- its children
- the other parents of its children

Given a node's Markov blanket, a node is conditionally independent of all other nodes in the network.

Constructing Bayes Nets

A Bayesian network should

- make the correct independence assumptions.
- require as few probabilities as possible.

Constructing a Correct Bayesian Network

1. Determine the set of variables that are required to model the domain.
2. Order the variables, $\{X_1, ..., X_n\}$.
3. For $i = 1$ to n, do the following
   3.1 Choose a minimum set of parents from X_1, ..., X_{i-1} such that $P(X_i|Parents(X_i)) = P(X_i|X_{i-1} \wedge \cdots \wedge X_1)$ is satisfied.
   3.2 Create a link from each parent of X_i to X_i.
   3.3 Write down the conditional probability table $P(X_i|Parents(X_i))$.

Example: Construct a Bayes Net

Construct a correct Bayesian network using the following ordering. (Let's drop Radio.)

$$B, E, A, W, G$$

Revisiting the Learning Goals

By the end of the lecture, you should be able to

- Compute a joint probability given a Bayesian network.
- Identify the conditional independence assumptions required by a Bayesian network.
- Given a Bayesian network, determine if two variables are independent or conditionally independent given a third variable.
- Given a scenario with independent assumptions and a given order of the variables, construct a Bayesian network by adding the variables to the network based on the given order.