Belief Network

Why?

- Bayesian networks have been the most important contribution
  - to the field of AI in the last few years
- Provide a way to represent knowledge in
  - an uncertain domain and
  - a way to reason about this knowledge
- Many applications: medicine, factories, help desks, spam filtering, etc.

Belief Network

- Find material such a copper, uranium after digging. We know
  - The probability of certain material
  - Linkage of certain material with other material
  - Physical characteristics
- Systems made using these concepts, architecure are as follows
  - CAS-Net
  - INTERNIST/CADCEUS

Definition

A Bayesian network is a

- Probabilistic graphical model
  - which represents a set of variables and
  - their conditional dependencies
  - using a directed acyclic graph.

Bayes Belief Network

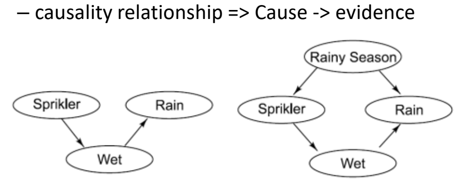
A Bayes net (also called a belief network) is an augmented directed acyclic graph, represented by the pair V , E where:

- V is a set of vertices.
- E is a set of directed edges joining vertices. No loops of any length are allowed.

Each vertex in V contains the following information:

- The name of a random variable
- A probability distribution table indicating how the probability of this variable's values depends on all possible combinations of parental values.

- Causality relationship using
  - Directed Acyclic Graph (DAG)
  - causality relationship => Cause -> evidence



(a) (b)

Bayesian Network

- Reduce the complexity of Bayesian reasoning system by
  - making approximation to formalism.
  - Preserve the formalism
  - Rely on modularity of world
- No Use of joint distribution of probability of all conceivable events
- Most events are conditionally independent
  - Interaction need not be considered
- Cluser of events that interact
  - Local representation

Two ways to view Bayes nets:

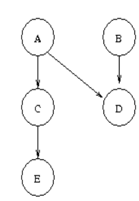
- A representation of a joint probability distribution
- An encoding of a collection of conditional independence statements
- It is intractable to update the probability consistently throughout the network.
  - However it is possible due to limited influence.
  - Nodes have limited domain of influence
  - Hence they can be computed as independant

- Systems made using these concepts, architecure are as follows
  - CAS-Net
  - INTERNIST/CADCEUS

Bayesian Example

Consider the following Bayesian network:

- A and B are (absolutely) independent.
- C is independent of B given A.
- D is independent of C given A and B.
- E is independent of A, B, and D given C.



Prob(A=T) = 0.3
Prob(B=T) = 0.6
Prob(C=T|A=T) = 0.8
Prob(C=T|A=F) = 0.4
Prob(D=T|A=T,B=T) = 0.7
Prob(D=T|A=T,B=F) = 0.8
Prob(D=T|A=F,B=T) = 0.1
Prob(D=T|A=F,B=F) = 0.2
Prob(E=T|C=T) = 0.7
Prob(E=T|C=F) = 0.2

Find

- P(D=T)
- P(D=F,C=T)
- P(A=T|C=T)
- P(A=T|D=F)
- P(A=T,D=T|B=F)
- P(C=T | A=F, E=T)

P(D=T)

= P(D=T,A=T,B=T) + P(D=T,A=T,B=F) + P(D=T,A=F,B=T) + P(D=T,A=F,B=F)

= P(D=T|A=T,B=T) P(A=T,B=T) + P(D=T|A=T,B=F) P(A=T,B=F) + P(D=T|A=F,B=T) P(A=F,B=T) + P(D=T|A=F,B=F) P(A=F,B=F)
(since A and B are independent absolutely)

= P(D=T|A=T,B=T) P(A=T) P(B=T) + P(D=T|A=T,B=F) P(A=T) P(B=F) + P(D=T|A=F,B=T) P(A=F) P(B=T) + P(D=T|A=F,B=F) P(A=F) P(B=F)

=0.7*0.3*0.6 + 0.8*0.3*0.4 + 0.1*0.7*0.6 + 0.2*0.7*0.4 = 0.32

P(D=F,C=T) =P(D=F,C=T,A=T,B=T) + P(D=F,C=T,A=T,B=F) + P(D=F,C=T,A=F,B=T) + P(D=F,C=T,A=F,B=F) =

P(D=F,C=T|A=T,B=T) P(A=T,B=T) + P(D=F,C=T|A=T,B=F) P(A=T,B=F) + P(D=F,C=T|A=F,B=T) P(A=F,B=T) + P(D=F,C=T|A=F,B=F) P(A=F,B=F) =
(since C and D are independent given A and B)

P(D=F|A=T,B=T) P(C=T|A=T,B=T) P(A=T,B=T) + P(D=F|A=T,B=F) P(C=T|A=T,B=F) P(A=T,B=F) + P(D=F|A=F,B=T) P(C=T|A=F,B=T) P(A=F,B=T) + P(D=F|A=F,B=F) P(C=T|A=F,B=F) P(A=F,B=F) =
(since C is independent of B given A and A and B are independent absolutely)

P(D=F|A=T,B=T) P(C=T|A=T) P(A=T) P(B=T) + P(D=F|A=T,B=F) P(C=T|A=T) P(A=T) P(B=F) + P(D=F|A=F,B=T) P(C=T|A=F) P(A=F) P(B=T) + P(D=F|A=F,B=F) P(C=T|A=F) P(A=F) P(B=F) =

0.3*0.8*0.3*0.6 + 0.2*0.8*0.3*0.4 + 0.9*0.4*0.7*0.6 + 0.8*0.4*0.7*0.4 = 0.3032

- **P(A=T|C=T)**

P(A=T|C=T)
= P(C=T|A=T)P(A=T) / P(C=T).

Now P(C=T) = P(C=T,A=T) + P(C=T,A=F)
= P(C=T|A=T)P(A=T) + P(C=T|A=F)P(A=F)
= 0.8*0.3+ 0.4*0.7 = 0.52
So P(C=T|A=T)P(A=T) / P(C=T)
= 0.8*0.3/0.52
= 0.46.

P(A=T|D=F)

P(A=T|D=F)
=P(D=F|A=T) P(A=T)/P(D=F).
Now P(D=F)
= 1-P(D=T) = 0.68 from the first question above.

P(D=F|A=T)
= P(D=T,B=T|A=T) + P(D=F,B=F|A=T)
= P(D=F|B=T,A=T) P(B=T|A=T) + P(D=F|B=F,A=T) P(B=F|A=T)
(since B is independent of A)
= P(D=F|B=T,A=T) P(B=T) + P(D=F|B=F,A=T) P(B=F)
= 0.3*0.6 + 0.2*0.4 = 0.26.

So P(A=T|D=F) = P(D=F|A=T) P(A=T)/P(D=F)
= 0.26 * 0.3 / 0.68
= 0.115

P(A=T,D=T|B=F).

P(A=T,D=T|B=F)
=P(D=T|A=T,B=F) P(A=T|B=F)
(since A and B are independent)
=P(D=T|A=T,B=F) P(A=T)
= 0.8*0.3
= 0.24.

P(C=T | A=F, E=T)

P(C=T | A=F, E=T) = P(E=T|C=T,A=F) *P(C=T|A=F) / P(E=T|A=F)
= P(E=T|C=T) * P(C=T|A=F) / P(E=T|A=F).
Now P(E=T|A=F) = P(E=T,C=T|A=F) + P(E=T,C=F|A=F)
=P(E=T|C=T,A=F) P(C=T|A=F) + P(E=T|C=F,A=F) P(C=F|A=F)
= P(E=T|C=T) * P(C=T|A=F) + P(E=T|C=F) * P(C=F|A=F).
So we have
P(C=T | A=F, E=T) = P(E=T|C=T)*P(C=T|A=F) / (P(E=T|C=T) * P(C=T|A=F) + P(E=T|C=F) * P(C=F|A=F)) =
0.7*0.4 / (0.7 * 0.4 + 0.2 * 0.6) = 0.7

Bayes Example