Belief Networks / Bayesian Networks

Tuesday, October 29, 2019 2:24 PM

Why probability:- Reason about unceirtanty.

$$p(\text{wall} = 2m \mid \text{sonar} = 1.2m) = ?$$

$$p(\text{cat} = \text{true} \mid \text{whisker} = \text{true}, \text{fur} = \text{true}, \text{legs} = 4) = ?$$

- Conditional probability has a weak-point.

  We require the "Probability measure."

  suppose $X_0, X_1, X_2, \ldots X_n$ boolean. 2^n worlds

  Its inpractical to store complete Probability measure

Idea:  given a variable X there is usually few variables that affect X let's call them Zs

Any other variable is irrelevant to knowing X if we are given Z's

formally: $p(X \mid Z\text{'s}) = p(X \mid Y, Z\text{s})$ Independence

X is conditionally independent of Y given Z's

e.i. $x \in \text{domain}(X)$, $y, y' \in \text{domain}(Y)$ $z \in \text{domain}(Z)$

$$p(X = x \mid Z = z) = p(X = x \mid Y = y \wedge Z = z)$$
$$= p(X = x \mid Y = y' \wedge Z = z)$$

Side Note:

X and Y are unconditionally independent when:

$$P(X, Y) = P(X) * P(Y)$$

Why:-

$$P(X_0, X_1, X_2, X_3 \ldots X_n) = \prod^n P(X_i \mid X_0 \ldots X_{i-1})$$ by the chain rule

Why:-

$$P(X_0, X_1, X_2, X_3 \dots X_n) = \prod_{i=0}^{n} P(X_i \mid X_1, \dots X_{i-1})$$ by the chain rule.

joint probability distribution.

- suppose for variable X, X depends parents(X)
- Let's order the variables so that for every X, parents(X) are predecessors of X

X_j parents$(X_j) \subseteq (X_0, X_1, \dots X_{j-1})$

$$P(X_0, X_1, X_2, \dots, X_n) = \prod_{i=0}^{n} P(X_i \mid \text{parents}(X_i))$$

Example: Domain of student Bob:-

Bob attandance to class, Bob revews his notes

Bob answers to the test, Bobs Grades.

Variables A, R, T, G

$P(A, R, T, G)$

- A : is independent — parents$(A) = \{\}$
- R : is independent — parents$(R) = \{\}$
- T : is dependent on A and R — parents$(T) = \{A, R\}$
- G : is dependent T — parents$(G) = \{T\}$

$$P(A,R,T,G) = P(A) \cdot P(R \mid A) \cdot P(T \mid A,R) \cdot P(G \mid A,R,T)$$

$$= P(A) \cdot P(R) \cdot P(T \mid A,R) \cdot P(G \mid T)$$ $\Leftarrow$ easier to store and to obtain.

$P(A \mid \text{Grade} = A)$

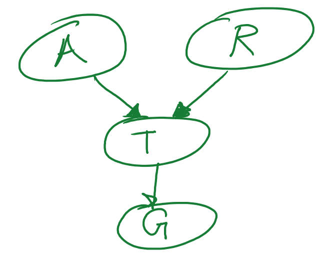
$P(R \mid \text{Grade} = \text{'A'} \wedge A = \text{Low})$

Bayesian Network / Belief Network:

- A Directed Acyclic Graph: represent our assumptions of variable dependency

- A Directed Acyclic Graph: represent our assumptions of variable dependency
  - Nodes are laveled by variables
  - There is an arc from every member of parents(x) to X
- A domain for each variable.
- A set of conditional probability tables.

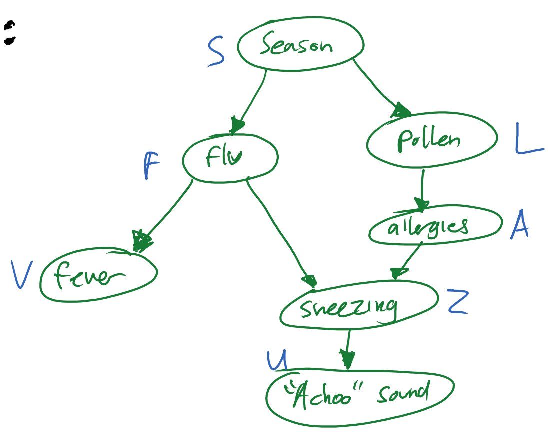
$$p((X \mid \text{parents}(X)))$$



Example: Alexa, Google home, "flu-app"

- Listen to "Achoo" sound. depends on sneezing.
- sneezing could be couse by allergies and flu.
- the flu. couses fever
- allergies are dependent on pollen.
- both pollen and flu are dependent on seasons.

Bayesion Network:



$$P(S,F,V,L,A,Z,U) = \prod \left\{ \begin{array}{ll} P(S) & P(V \mid F) \\ P(L \mid S) & P(A \mid L) \\ P(F \mid S) & P(Z \mid A, F) \\ & P(U \mid Z) \end{array} \right\}$$

Computing new probabilities ≡ Probabilistic Inference ≡

$$P(\text{flu} = \text{true} \mid \text{sneeze} = \text{true}) = ?$$

$$P(\text{season} = \text{winter} \mid \text{Pollen} = \text{true}) = ?$$

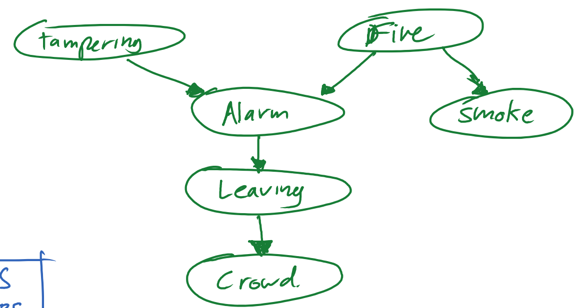
$$P(\text{flu} = \text{true} \mid \text{fever} = \text{false} \wedge \text{allergies} = \text{false}) = ?$$

Example #2: Smart house.-

• fire alarm.- external doorbell camera:

- the fire alarm can be tampered, ring when there is a fire
- Fire produces smoke.
- When the fire alarm rings, people leave the building.
- When people leave the building, you see peope in the doorbell camera.

Belief / Bayesian Network:



$P(\text{Tampering}) = 0.02$

$P(\text{Fire}) = 0.01$

$P(\text{Alarm} \mid \text{Tampering}, \text{Fire}) =$

t	f	0.5
t	¬f	0.85
¬t	f	0.99
¬t	¬f	0.0001

$P(\text{Smoke} \mid \text{Fire}) = 0.9$

$P(\text{Leaving} \mid \text{Alarm}) = 0.88$

$P(\text{Crowd} \mid \text{Leaving}) = 0.75$

The camera detects a crowd at the door: Crowd = true

$P(\text{fire} \mid \text{Crowd}) = 0.23$

$P(\text{Tampering} \mid \text{Crowd}) = 0.39$

$P(\text{Smoke} \mid \text{Crowd}) = 0.21$

• Suppose smoke = true

$P(\text{fire} \mid \text{smoke}) = 0.476$

$P(\text{tampering} \mid \text{smoke}) = 0.02$

• Suppose crowd = true but smoke = false

$P(\text{fire} \mid \text{crowd} = \text{true} \wedge \text{smoke} = \text{false}) = 0.029$

$P(\text{tampering} \mid \text{crowd} = \text{true} \wedge \text{smoke} = \text{false}) = 0.50$

Effects of observation on a belief network.

Effects of observation on a belief network.

- Observe a variable Y; which probabilities change:
  - The descendants of Y change
  - The ancestors of Y change.