Bayesian Networks 1

Wednesday, March 21, 2018 8:51 AM

- Part of larger body of techniques
    - Graphical models
    - Model is a graph $G = (V, E)$

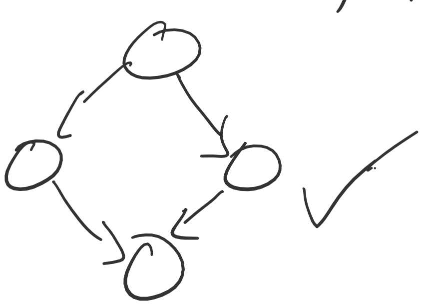
vertices (nodes) edges

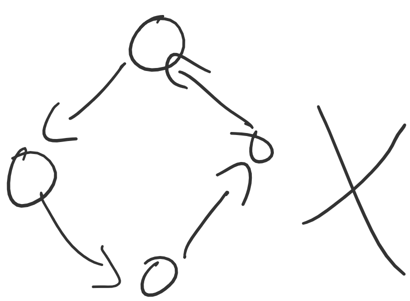
- Node is a random variable
    - take on discrete number of values
- Edges between nodes indicate conditional dependence
    - Fancy way to say that one variable has an impact on another

Variables R, P

P is conditionally dependent on R

Directed Acyclic Graph (DAG)

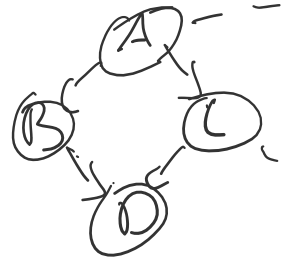




(V, E) - Defines network structure

- Conditional probability table

assume { binary variables



$P(A = \text{True})$
$P(A = \text{False})$

$P(C = \text{True} \mid A = \text{True})$
$P(C = \text{False} \mid A = \text{True})$
$P(C = \text{True} \mid A = \text{False})$
$P(C = \text{False} \mid A = \text{False})$

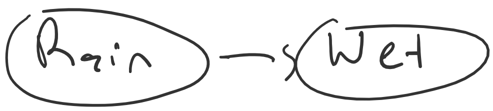
$P(D = \text{true} \mid B = \text{true}, C = \text{true})$
$P(D = \text{true} \mid B = \text{False}, C = \text{true})$
$\vdots$

This network defines the joint distribution over a set of variables

$$P(A, B, C, D) = P(A) * P(B \mid A) * P(C \mid A) * P(D \mid B, C)$$

Joint distribution

Ex. Following network



$P(R) = 0.4$

$P(W \mid R) = 0.9$
$P(W \mid \neg R) = 0.2$

- We can ask what is the probability of $W = \text{true}$

$$P(W = \text{true}) = \sum_R P(R, W) = P(W \mid R) P(R) + P(W \mid \neg R) P(\neg R)$$

$$0.9 * 0.4 + 0.2 * 0.6 = \underline{0.48}$$

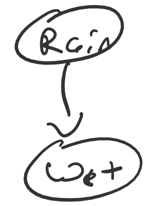
Marginalization over values of R

- Network structure
  - usually given, not always
  - Bayesian Information Criteria (BIC)
    - criteria of network "goodness"
- Ex. Bioinformatics
  - Vertices = genes
    values = expression information
  - Randomize edges
  - Evaluate BIC
  - Add / remove edges if BIC would increase
  - Start with empty network
    - greedily add edges according to BIC
  - Optimize using Genetic Algorithm
- Back to the other thing
  - Diagnosis: invert conditional dependence
  - Recall grass example.

- Grass is wet ($W = \text{true}$)
- What is the probability that it rained?

- $P(R|W) = \dfrac{P(W|R)\,P(R)}{P(W)}$

Bayes Rule



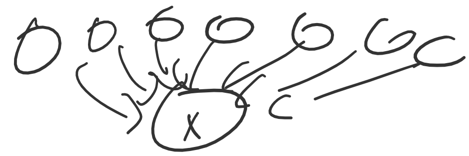
$P(R) = 0.4$

$P(W|R) = 0.9$

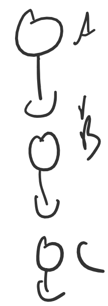
$P(W|\neg R) = 0.2$

$$\frac{0.9 * 0.4}{0.48} = 0.75$$

- This is classification
  - R is class label
  - W is some feature



$P(X|A, B, C, D, \ldots, Z)$



$P(C|B)$

Independence Formally

- 2 events (variable) are independent if

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

- 2 variables are conditionally independent given a third variable if

$$P(X,Y|Z) = P(X|Z)\,P(Y|Z)$$

rewritten

$$P(X|Y,Z) = P(X|Z)$$