Probabilistic Graph Models

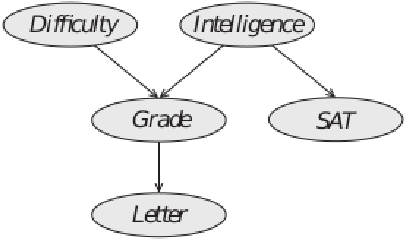
Liu Xuefeng

June 3, 2019

Bayesian Networks

- The core of Bayesian network representation is a directed acyclic graph (DAG) G, whose nodes are the random variables in our domain and whose edges correspond, intuitively, to direct influence of one node on another.

Student Example



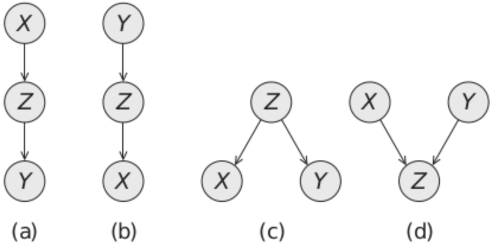
$$P(I, D, G, S, L) = P(I)P(D)P(G|I, D)P(S|I)P(L|G)$$

Bayesian Network Structure

A Bayesian network structure G is a directed acyclic graph whose nodes represent random variables $x_1, x_2, \cdots, x_n$.
Let $Pa_{x_i}^G$ denote the parents of x_i in G, and $NonDescendants_{x_i}$ denote the variables in the graph that are not descendants of x_i.
Then G encodes the following set of conditional independence assumptions, called the local independencies:

$$(x_i \perp\!\!\!\perp NonDescendants_{x_i} | Pa_{x_i}^G)$$

Indirect Connection



a Indirect causal effect.

b Indirect evidential effect.

c Common cause.

d Common effect.

d-separation

Definition

Let G be a BN structure, and $x_1 - x_2 - \cdots - x_n$ a trail in G. Let Z be a subset of observed variables. The trail is active given Z if

- Whenever we have a v-structure $x_{i-1} \rightarrow x_i \leftarrow x_{i+1}$, then x_i or one of its descendants are in Z;
- no other node along the trail is in Z.

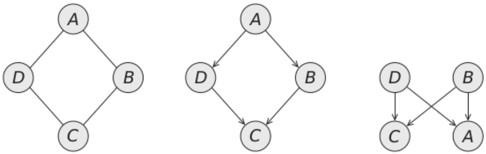
d-separation

Definition

Let X, Y, Z be three sets of nodes in G. We say that X and Y are d-separated given Z, denoted $d - sep_G(X; Y|Z)$, if there is no active trail between any node $x \in X$ and $y \in Y$ given Z. We use $L(G)$ to denote the set of independencies that corresponds to d-separation:

$$L(G) = \{(X \perp\!\!\!\perp Y|Z) : d - sep_G(X; Y|Z)\}$$

Misconception Example



(a) (b) (c)

- $(B \perp\!\!\!\perp D | \{A, C\})$
- $(A \perp\!\!\!\perp C | \{B, D\})$

Markov Network

Definition

Let D be a set of random variables. We define a factor ϕ to be a function from $Val(D)$ to R. The set of variables D is called the scope of the factor and denoted $Scope[\phi]$.

Definition

Let X, Y, Z be three disjoint sets of variables, and let $\phi_1(X, Y)$ and $\phi_2(Y, Z)$ be two factors. We define the factor product $\phi_1 \times \phi_2$ to be factor $\phi : Val(X, Y, Z) \to R$ as follows:

$$\phi(X, Y, Z) = \phi_1(X, Y)\phi_2(Y, Z).$$

Misconception Example

$\phi_1(A,B)$		
a^0	b^0	30
a^0	b^1	5
a^1	b^0	1
a^1	b^1	10

$\phi_2(B,C)$		
b^0	c^0	100
b^0	c^1	1
b^1	c^0	1
b^1	c^1	100

$\phi_3(C,D)$		
c^0	d^0	1
c^0	d^1	100
c^1	d^0	100
c^1	d^1	1

$\phi_4(D,A)$		
d^0	a^0	100
d^0	a^1	1
d^1	a^0	1
d^1	a^1	100

Assignment				*Unnormalized*	*Normalized*
a^0	b^0	c^0	d^0	300,000	0.04
a^0	b^0	c^0	d^1	300,000	0.04
a^0	b^0	c^1	d^0	300,000	0.04
a^0	b^0	c^1	d^1	30	$4.1 \cdot 10^{-6}$
a^0	b^1	c^0	d^0	500	$6.9 \cdot 10^{-5}$
a^0	b^1	c^0	d^1	500	$6.9 \cdot 10^{-5}$
a^0	b^1	c^1	d^0	5,000,000	0.69
a^0	b^1	c^1	d^1	500	$6.9 \cdot 10^{-5}$
a^1	b^0	c^0	d^0	100	$1.4 \cdot 10^{-5}$
a^1	b^0	c^0	d^1	1,000,000	0.14
a^1	b^0	c^1	d^0	100	$1.4 \cdot 10^{-5}$
a^1	b^0	c^1	d^1	100	$1.4 \cdot 10^{-5}$
a^1	b^1	c^0	d^0	10	$1.4 \cdot 10^{-6}$