BNs: Independence

$X \perp\!\!\!\perp Y \mid Z$

$P(x,y|z) = P(x|z)P(y|z)$

$P(x|y,z) = P(x|z)$

Bayes' Net joint distribution

$$P(X_1, X_2, \cdots, X_n) = \prod_{i=1}^{n} P(X_i \mid parent(X_i))$$

query $P(+b,-e,+a,-j,+m)$

$= P(+b)P(-e)P(+a|+b,-e)P(-j|+a)P(+m|+a)$

D-separation

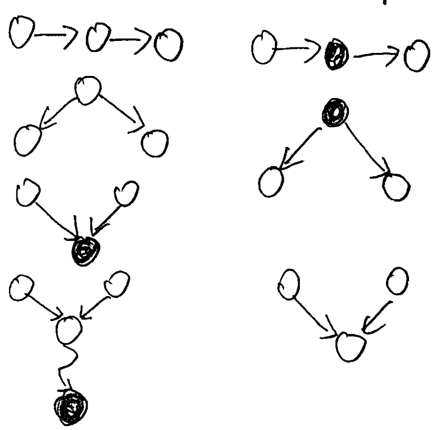
Are X and Y conditionally independent given evidence variables $\{Z\}$

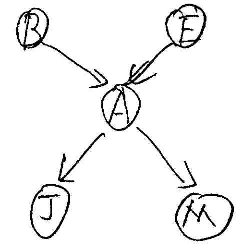
No active path = independense

A path is active if each triple is active

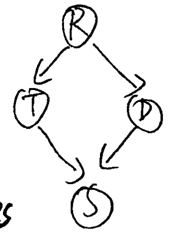
All it takes to block a path is a single inactive segment

Active triples   Inactive triples





Example



$T \perp\!\!\!\perp D$ No

$T \perp\!\!\!\perp D \mid R$ Yes

$T \perp\!\!\!\perp D \mid R, S$ No

BNs: Inference

Variable elimination

We want $P(Q|e_1 \cdots e_k)$

Evidence viable	$E_1 \cdots E_k = e_1 \cdots e_k$	$X_1 \cdots X_n$ all variables
Query	Q	
Hidden	$H_1 \cdots H_r$	

Example

$P(B|j,m) \propto P(B,j,m)$

$P(B)\ P(E)\ P(A|B,E)\ P(j|A)\ P(m|A)$

choose A

$P(A|B,E)$, $P(j|A)$, $P(m|A)$ $\xrightarrow{\times}$ $P(j,m,A|B,E)$ $\xrightarrow{\Sigma}$ $P(j,m|B,E)$

left with $P(B)\ P(E)\ P(j,m|B,E)$

$$P(A|B) = \sum_c P(A,c|B)$$

$$P(A) = \sum_b P(A,b)$$

$$P(B|A,C) = \frac{P(A)P(B|A)P(C|A,B)}{\sum_b P(A)P(b|A)P(C|A,B)}$$

BNs: sampling

prior sampling

Rejection sampling

Likelihood sampling

Gibbs $P(S|+r)$

1 fix evidence $R = +r$

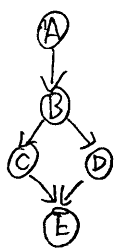
2. initialize other variables randomly

3.1 choose a non-evidence variable X

3.2 Resample X from $P(X|\text{all other variables})$

Then repeat

weight of the likelihood sample:



We want to estimate $P(C=1|B=1,E=1)$

Fix $B=1$, $E=1$

$\text{Weight} = P(B=1|A=1) * P(E=1|C=0,D=0)$

Suppose we get a sample $A=1$ $B=1$ $C=0$ $D=0$ $E=1$

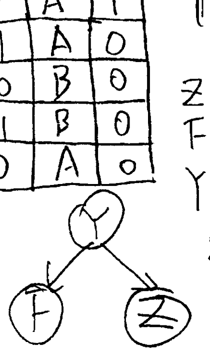
estimate $P(C=1|B=1,E=1)$

$$= \frac{\text{sum of weight where } C=1}{\text{total weight of samples}}$$

F	Y	Z
1	A	1
0	B	1
1	B	1
0	A	1
1	A	0
0	B	0
1	B	0
0	A	0

Suppose we know $P(Y)$ from samples, then we can calculate $P(F|Y)$, $P(Z|Y)$

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



Z	1	0	1	0	0	1
F	0	1	1	1	1	1
Y	A	A	A	A	B	B

Samples

given a new observation $F=1, Z=0$, predict Y.

We calculate $P(Y=A, F=1, Z=0)$

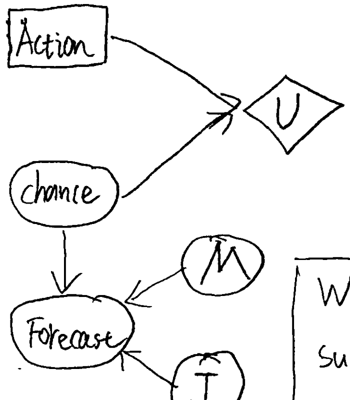
$P(Y=B, F=1, Z=0)$

Then posterior prob of Y is <u>normalize</u> joint distribution

$$P(Y=A|F=1,Z=0) = \frac{P(Y=A,F=1,Z=0)}{P(Y=A,F=1,Z=0)+P(Y=B,F=1,Z=0)}$$

Naive Bayes

Decision Network



A	W	U(A,W)
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70

$EU(leave) = \sum_{w} P(w)\, U(leave, w) = 70$

$EU(take) = \sum_{w} P(w)\, U(take, w) = 35$

$MEU(\phi) = \max_{a} EU(a) = 70$

Now observe Forecast = bad

W	P(W)	P(W|F=bad)
sun	0.7	0.34
rain	0.3	0.66

$EU(leave|bad) = 0.34 \times 100 + 0.66 \cdot 0 = 34$

$EU(take|bad) = 0.34 \times 20 + 0.66 \cdot 70 = 53$

$MEU(F=bad) = \max_{a} EU(a|bad) = 53$

VPI

$MEU(\phi) = 70$ no evidence

$MEU(F=bad) = 53$

$MEU(F=good) = 95$

P(F)	F
0.59	good
0.41	bad

$VPI(E'|e) = \left(\sum_{e'} P(e'|e) MEU(e, e')\right) - MEU(e)$

$0.59 \cdot 95 + 0.41 \times 53 - 70 = 7.8$

properties:

1. Nonnegative
   $\forall E', e : VPI(E'|e) \geq 0$
2. $VPI(E_i, E_j | e) \neq VPI(E_i|e) + VPI(E_j|e)$ (think observe E_j twice)
3. order-independence
   $VPI(E_j, E_k|e) = VPI(E_j|e) + VPI(E_k|e, E_j)$
   $= VPI(E_k|e) + VPI(E_j|e, E_k)$

Note that M is independent with chance, hence $\perp\!\!\!\perp$ Utility node

so $VPI(M) = 0$.

Markov Chain stationary distribution

X_{t-1}	X_t	$P(X_t \mid X_{t-1})$
sun	sun	0.9
sun	rain	0.1
rain	sun	0.3
rain	rain	0.7

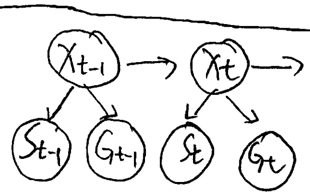
$P_\infty(sun) + P_\infty(rain) = 1$

$P_\infty(sun) = P(sun|sun) P_\infty(sun) + P(sun|rain) P_\infty(rain) = 0.9 P_\infty(sun) + 0.3 P_\infty(rain)$

$P_\infty(rain) = P(rain|sun) P_\infty(sun) + P(rain|rain) P_\infty(rain) = 0.1 P_\infty(sun) + 0.7 P_\infty(rain)$

HMMs

Dynamic update



$P(X_t | S_{1:t-1}, g_{1:t-1}) = \sum_{X_{t-1}} P(X_{t-1}, X_t | S_{1:t-1}, g_{1:t-1})$

$= \sum_{X_{t-1}} P(X_t | X_{t-1}, S_{1:t-1}, g_{1:t-1}) P(X_{t-1} | S_{1:t-1}, g_{1:t-1})$

$X_t \perp\!\!\!\perp S_{1:t-1}, G_{1:t-1}$ given X_{t-1}

$= \sum_{X_{t-1}} P(X_t | X_{t-1}) P(X_{t-1}, X_t | S_{1:t-1}, g_{1:t-1})$

Observation update

$P(X_t | S_{1:t}, g_{1:t}) = P(X_t | S_t, g_t, S_{1:t-1}, g_{1:t-1})$

$= \frac{1}{P(S_t, g_t | S_{1:t-1}, g_{1:t-1})} P(X_t, S_t, g_t | S_{1:t-1}, g_{1:t-1})$

$= * \cdot P(X_t | S_{1:t-1}, g_{1:t-1}) \cdot P(S_t, g_t | X_t, S_{1:t-1}, g_{1:t-1})$

$= * \cdot P(X_t | S_{1:t-1}, g_{1:t-1}) \cdot P(S_t | X_t) P(g_t | X_t)$

Forward update

$P(X_t | S_{1:t}, g_{1:t}) = \sum_{X_{t-1}} P(X_{t-1}, X_t) | S_t, g_t, S_{1:t-1}, g_{1:t-1})$

$S_t, G_t \perp\!\!\!\perp S_{1:t-1}, G_{1:t-1} | X_t$, $S_t \perp\!\!\!\perp G_t | X_t$

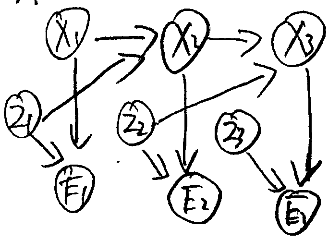
lottery $[0.6, \$0; 0.4, \$100] = L$

what is $\$X$ s.t. $U(\$X) = U(L)$

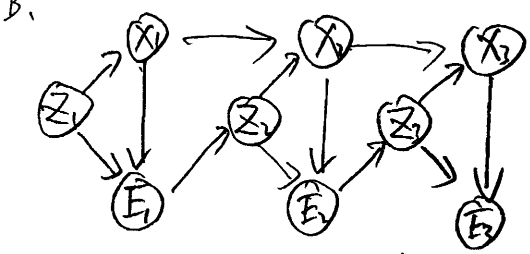
Hint: $U(L) = 0.6U(0) + 0.4U(100) = 400$

$U(X) = 400$

A



B.



observe A, B

$P(X_t, Z_t | e_{1:t}) \propto P(X_t, Z_t, e_t | e_{1:t-1})$

$\propto P(X_t, Z_t | e_{1:t-1}) P(e_t | X_t, Z_t, e_{1:t-1})$

$\propto P(X_t, Z_t | e_{1:t-1}) P(e_t | X_t, Z_t)$

elapse time update A

$$P(X_t, Z_t | e_{1:t-1}) = \sum_{X_{t-1}, Z_{t-1}} P(X_t, Z_t, X_{t-1}, Z_{t-1} | e_{1:t-1})$$

$$= \sum_{X_{t-1}, Z_{t-1}} P(Z_t | X_t, X_{t-1}, Z_{t-1}, e_{1:t-1}) P(X_t | X_{t-1}, Z_{t-1}, e_{1:t-1}) P(X_{t-1}, Z_{t-1} | e_{1:t-1})$$

$$= \sum_{X_{t-1}, Z_{t-1}} P(X_{t-1}, Z_{t-1} | e_{1:t-1}) P(X_t | X_{t-1}, Z_{t-1}) P(Z_t)$$

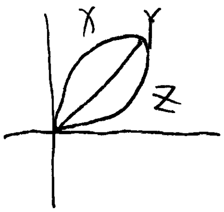
spread out as possible as it can

Particle Filtering, w(x) = P(e|x)

B $$P(X_t, Z_t | e_{1:t-1}) = \sum_{X_{t-1}, Z_{t-1}} P(X_t, Z_t, X_{t-1}, Z_{t-1} | e_{1:t-1})$$

$$= \sum_{X_{t-1}, Z_{t-1}} P(X_t | Z_t, X_{t-1}, Z_{t-1}, e_{1:t-1}) P(Z_t | Z_{t-1}, X_{t-1}, e_{1:t-1}) P(Z_{t-1}, X_{t-1} | e_{1:t-1})$$

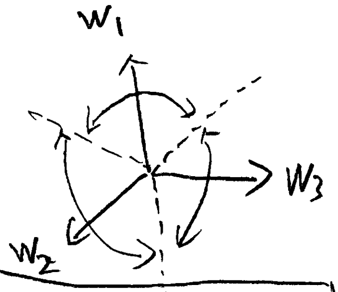
$$= \sum_{X_{t-1}, Z_{t-1}} P(X_t | Z_t, X_{t-1}) P(Z_t | e_{1:t-1}) P(Z_{t-1}, X_{t-1} | e_{1:t-1})$$



X: risk averse

Y: risk neutral

Z: risk taker



Start with all $W_i = 0$

Pick training sample one by one

predict

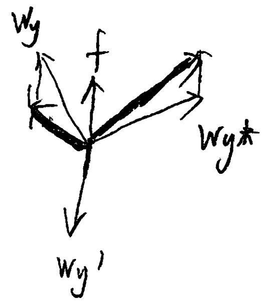
$y = \arg\max_y w_y f(x)$

If correct, no change

otherwise:

$W_y = W_y - f(x)$

$W^* = W_{y^*} + f(x)$



1. calculate $P(\text{sensor-reading} = H \mid \text{position} = H)$

$P(S\text{-}r = T \mid \text{position} = H)$

usually given in question

positions $\begin{Bmatrix} H, & C, \\ T, & D \end{Bmatrix}$

2. initialize:

3. tell you at $t=1$, sensor reading is $E_1 = D$

particles: P_1 P_2 ... P_{10}

H C ... D

(calculate $P(X_1 = \text{positions})$ based on particles)

The weight of each particle is $P(E_1 = D \mid \text{particle position})$

Then normalize each particles weight, so that we can resample based on new particle's distribution P_1 P_2 P_3 ... P_{10}

0.2 0.1 0.1 0.3

4. new particles - $r_1 = 0.23$ $r_2 = 0.15$

$P_1 =$ old $P_2 = C$ new $P_2 =$ old $P_1 = H$