2/18/14

BN: INFERENCE

① Setup:

→ Random Vars: $\vec{X} = \{X_1, \ldots, X_n\}$ (all categorical for now)

→ Given BN Structure: DAG $G = (V, E)$

→ Given BN Parameters CPTs $P(X_i \mid Pa_{X_i})$

→ Some "evidence" variables $\vec{E} \subseteq \vec{X} = \{X_1 \ldots X_n\}$
& their states $\vec{E} = \vec{e}$ or $E_1 = e_1, E_2 = e_2 \ldots$

→ Some Query variables of interest $\vec{Y} \subseteq \vec{X}$
$\{Y_1, Y_2, \}$

→ CONDITIONAL PROB. / MARGINAL Inference
Find $P(\vec{Y} = \vec{y} \mid \vec{E} = \vec{e})$

→ MAP Inference [Also called MPE]
Find argmax $\underbrace{P(\vec{X} \setminus \vec{E}}_{} \mid \vec{E} = \vec{e})$

"Find the best setting of everything else"

→ MARGINAL-MAP Inference [Also called MAP]
Find $\underset{\vec{y}}{\text{argmax}} \; P(\vec{Y} = \vec{y} \mid \vec{E} = \vec{e})$

$$= \underset{\vec{y}}{\text{argmax}} \sum_{\vec{x}} P(\vec{Y} = \vec{y}, \vec{X} \setminus \{\vec{Y} \cup \vec{E}\} = \vec{x} \mid \vec{E} = \vec{e})$$

"Marginalize out things I don't care about, then tell me best setting of variables I do care about"

② Is MAP "consistent" with Marginals?

Consider $S \longrightarrow N$

$$P(S) = \begin{bmatrix} 0.6 & 0 \\ 0.4 & 1 \end{bmatrix} \qquad P(N \mid S) = \begin{array}{c|cc} N \backslash S & 0 & 1 \\ \hline 0 & 0.5 & 0 \\ 1 & 0.5 & 1 \end{array}$$

$P(S, N) =$

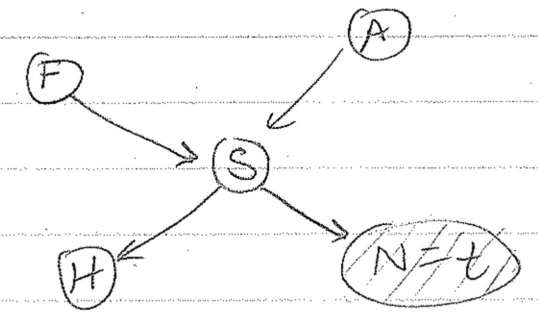
S	N	Prob
0	0	0.3
0	1	0.3
1	0	0
1	1	0.4

$$\text{MAP} = \underset{S,N}{\text{argmax}}\ P(S,N) = \overset{S\ \ N}{(1, 1)}$$

Marginal $P(S) = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}$

$$\text{So} \quad \underset{s}{\text{argmax}}\ P(S=s) \neq \underset{s,n}{\text{argmax}}\ P(S=s, N=n)$$

3 Marginal Inference Example



Compute: $P(F \mid N=t)$

First Notice that $P(F=f \mid N=t) = \dfrac{P(F=f, N=t)}{P(N=t)}$

$$= \frac{P(F=f, N=t)}{\underbrace{\sum_{f'} P(F=f', N=t)}_{\text{Normalization constant}}}$$

So lets do this; let's compute

$$\begin{bmatrix} P(F=f, N=t) \\ P(F=t, N=t) \end{bmatrix}$$

Then lets "normalize" this vector to get $P(F \mid N=t)$

In general $P(Y \mid \vec{E}=\vec{e}) \propto P(Y, \vec{E}=\vec{e})$

So need to compute $P(\vec{Y}=\vec{y}, \vec{E}=\vec{e}) \; \forall \vec{y}$

Now,

$$P(F=f, N=t) = \underbrace{\sum_a \sum_s \sum_h}_{} P(F=f, A=a, S=s, H=s, N=t)$$

Naively : 2^8 sums

In general n vars $\Rightarrow$ 2^n sums

$\Rightarrow$ Exponential!

Can we do better?

In worst case, no. #P-hardness.

But sometime ($\equiv$ special cases), Yes!

For example, notice

$$\sum_a \sum_s \sum_h P(f) P(a) P(s \mid f, a) P(h \mid s) P(N=t \mid s)$$

terms that depend on A: 2

S: 3

H: 1 } Interesting

Let's "Eliminate" H first $\Rightarrow$ push in $\sum_h$

$$\sum_a \sum_s P(f) P(a) P(s \mid f,a) P(N=t \mid s) \cancel{\sum_h P(h \mid s)} \to 1$$

Very Nice!

Let's "Eliminate" S next

$$\sum_a P(f) P(a) \underbrace{\sum_s P(s \mid f, a) P(N=t \mid s)}_{\text{New "Factor" } g(F,A)}$$

$$g(F=f, A=a) = \begin{array}{c|cc} F \backslash A & 0 & 1 \\ \hline 0 & g(0,0) & g(0,1) \\ 1 & g(1,0) & g(1,1) \end{array}$$

$$\underbrace{g(f,a)} = \sum_s P(s \mid f,a)\, P(N=t \mid s)$$

2^2 entries (1 sum per entry)

Now, let's Eliminate A

$$P(f) \underbrace{\sum_a P(a)\, g(f,a)}_{\text{New factor } h(F)} \quad \begin{bmatrix} h(0) \\ h(1) \end{bmatrix}$$

2 entries (1 sum per entry)

Finally,

$$P(F=f, N=t) = P(f) \cdot h(f)$$

Normalize to get $P(F \mid N=t)$

Recap: Variable Elimination Algorithm in BNs

Elim Ordering: H, S, A

time/ iteration	List of Factors	Operation/ Elimination	New Factor
1	$P(F)\,P(A)\,P(S \mid F,A)\,P(H \mid S)\,P(N=t \mid S)$	$\sum_h P(h \mid S)$	ϕ
2	$P(F)\,P(A)\,P(S \mid F,A)\,P(N=t \mid S)$	$\sum_s P(s \mid f,a)\,P(N=t \mid s)$	$g(F,A)$ size 2^2
3	$P(F)\,P(A)\,g(F,A)$	$\sum_a P(a)\,g(f,a)$	$h(F)$ size 2^1

Thus $P(F \mid N=t) \propto \underbrace{P(F)\,h(F)}_{\text{left-over factors}}$

Normalize to get answer