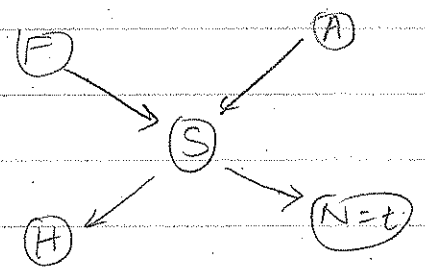


2/20/14

①

BN: VARIABLE ELIMINATION

① Extending VE to MAP Inference



Compute: $\text{argmax}_{f, a, s, h} P(F=f, A=a, S=s, H=h | N=t)$

Notice $\text{argmax}_{f, a, s, h} P(f, a, s, h | N=t) = \text{argmax}_{f, a, s, h} P(f, a, s, h, N=t)$

Let's focus on finding max value first. We'll find argmax next.

$$\max_f \max_a \max_s \max_h P(f)P(a)P(s|f, a)P(h|s)P(N=t|s)$$

2^4 settings. Can we do better like the marginal case? Yes!

$$\boxed{\max_i \{a \cdot b_i\} = a \cdot \max_i \{b_i\}}$$

if $a \geq 0$. So let's use this

Elim Ordering: H, S, A, F

Ite	Current Factors	Elimination	New Factors
①	$P(F) P(A) P(S F, A) P(H S) P(N=1 S)$	$\max_h P(h s)$	$g_1(s)$
②	$P(F) P(A) P(S F, A) P(N=1 S) g_1(s)$	$\max_s P(N=1 s) g_1(s) P(S F, a)$	$g_2(f, a)$
③	$P(F) P(A) g_2(f, a)$	$\max_a P(a) g_2(f, a)$	$g_3(f)$
④	$P(F) g_3(f)$	$\max_f P(f) g_3(f)$	Proportional to MAP value

Now, to find argmax, let's back track

We know $f^* = \operatorname{argmax}_f P(f) g_3(f)$ ← what setting of f led to MAP value

Now $a^* = \operatorname{argmax}_a P(a) g_2(f^*, a)$

← what setting of a (consistent with MAP setting of f) lead to MAP value?

$$s^* = \operatorname{argmax}_s P(N=1|s) g_1(s) P(s|f^*, a^*)$$

$$h^* = \operatorname{argmax}_h P(h|s^*)$$

② Can we generalize VE further?
What else can VE (or dynamic programming) do?

Consider an abstract "operator" \otimes
and another " " \oplus

If \oplus distributes over \otimes we can save computation.
Meaning?

$$\text{If } a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c)$$

Then

$$(a \otimes b_1) \oplus (a \otimes b_2) \oplus \dots \oplus (a \otimes b_k)$$

$$\uparrow = a \otimes [b_1 \oplus \dots \oplus b_k]$$

$k + (k-1)$
operations

\uparrow
 $(k-1) + 1$ operations \Rightarrow Savings!

In our case \otimes is "product"
 \oplus is "sum" for sum-product VE
 \oplus is "max" for max-product VE

What else?

How about $\oplus \equiv$ top-k elements

This will result in a VE alg that returns top-K Most probable configurations

