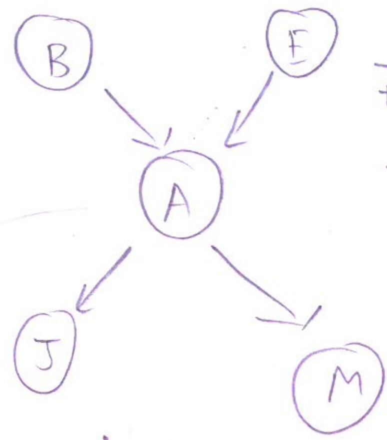
① Bayesian Network :

Oct 17th lecture
Class notes

B	P(B)
+b	
-b	



E	P(E)
+e	
-e	

A B E	P(A|B,E)
+a +b +e	
+a +b -e	
+a -b +e	
+a -b -e	
⋮	

A, J	P(A|J)
a j	
a -j	
-a j	
-a -j	

A, M	P(M|A)
a m	
a -m	
-a m	
-a -m	

→ Any query can be answered w/ these tables.

~~$P(j|e) = \frac{P(A,B,E,J,M)}{P(}$~~

~~$P(A,B,E,J,M|E)$~~

marginalize

$$P_r(j|e) = \frac{P(j,e)}{P(e)} = \frac{\sum_{A,B,M} P(A,B,e,j,M)}{\sum_{ABJM} P(A,B,e,j,M)}$$

factorize

$$P(A,B,e,j,M) = P(B)\,P(e)\,P(A|B,e)\,P(j|A)\,P(M|A)$$

> compute values for f_2 once and store.

$$f_2(c) = P(c \mid b)\, f_1(b) + P(c \mid -b)\, f_1(-b)$$

$$f_2(-c) = P(-c \mid b)\, f_1(b) + P(-c \mid -b)\, f_1(-b)$$

> $$P(d) = \sum_{C} P(d \mid c)\, f_2(c)$$

$$P(d) = P(d \mid c)\, f_2(c) + P(d \mid -c)\, f_2(-c)$$

Phew! Done.

Variable Elimination for chain networks

> Chain networks:

$$X_1 \to X_2 \to \cdots \to X_n$$

where, say ~~each~~ $|X_i| \le k$

> ~~BA~~ A sample query, say, $P(X_{i+1}) = ?$

$$P(X_{i+1}) = \sum_{x_j \in X_i} P(X_{i+1} \mid x_j)\, P(x_j)$$

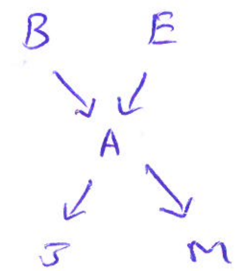
or

$$= \sum_{x_i} P(X_{i+1} \mid x_i)\, P(x_i) \checkmark$$

or

$$= \sum_{x_i} P(X_{i+1} \mid X_i = x_i)\, P(X_i = x_i)$$

④ Variable Elimination : Notion of Factors.



$P(B \mid j, m)$

$$= \alpha \cdot P(B) \sum_{e} P(e) \sum_{a} \left\{ \begin{array}{l} P(a \mid B, e) \\ \times P(j \mid a) \\ \times P(m \mid a) \end{array} \right\}$$

$f_1(B) = P(b) \quad \forall b \in B$

$f_2(E) = P(e) \quad \forall e \in E$

$f_3(A, B, E) = P(a, b, e) \quad \forall \begin{array}{l} a \in A \\ e \in E \\ b \in B \end{array}$

$f_4(A) = P(j \mid a) \quad \forall a \in A$ // no J since j is evidence.

$f_5(A) = P(m \mid a) \quad \forall a \in A$

Factor:
These $f_1 \ldots f_5$ are factors. Denote a mapping of values of the arguments to a real number

$$f_i : X_1 \times \ldots \times X_m \longrightarrow \mathbb{R}$$

(e.g.)

A	P(A)
a	0.2
-a	0.8

$f(A)$
↳ scope = A

A	B	P(B/A)
a	b	0.2
a	-b	0.8
-a	b	0.3
-a	-b	0.7

$f(A, B)$
↳ scope = A, B.

A	B=b	P(A|b)
a	b	0.3
-a	b	0.7

$f(A)$
↳ scope =

$$P(B|j,m) = \alpha \cdot f_1(B) \times \sum_e f_2(E) \times f_6(B,E)$$

A has been eliminated!

$$\sum_e f_2(E) \times f_6(B,E)$$

$$f_7(B) = f_2(e) \times f_6(B,e) + f_2(-e) \times f_6(B,-e)$$

B	$f_7(B)$
b	$f_2(e) \times f_6(b,e) + f_2(-e) \times f_6(b,-e)$
$-b$	$f_2(e) \times f_6(-b,e) + f_2(-e) \times f_6(-b,-e)$

And $P(B|j,m) = \alpha \cdot f_1(B) \times f_7(B)$

And E is eliminated!!

B	$P(B|j,m)$
b	$\alpha \cdot f_1(b) \times f_7(b)$
$-b$	$\alpha \cdot f_1(-b) \times f_7(-b)$

In general if output factor has scope of $\boxed{q}$ variables

→ each w/ $\boxed{r}$ values : $\boxed{r^{q}}$ ← N_K

If input has m_k factors.
Each row requires ~~multiplying~~ $m_k - 1$ multiplication.

∴ complexity of join : $O\left(\cancel{(l-1)k^{m}}\right)$ $O\left((m_k - 1) N_K\right)$

↳ Marginalization

A, B, C	
a b c	
a b -c	
a -b c	
a -b -c	
-a b c	
-a b -c	
-a -b c	
-a -b -c	

A, B	
a, b	
a, -b	
-a b	
-a -b	

one addition for each row

(or)

each row gets used in one addition

$< \sim O(N_k)$

Total complexity

8

- Starts w/ n variables = n factors

- Each step we eliminate one variable (but that generates one factor)

- At most 'n' eliminations

- Total factors $\leq$ initial n + generated n $\leq 2n$

- Let $N = \max(N_k)$ ie) the size of the largest ~~sized~~ factor that we see.

- Let m_k be the # of input factors for forming factor N_k

Step 1: we eliminate H_1 w/ new factor f_1 of size N_1. The # of input factors were m_1

join cost = $O((m_1 - 1) N_1)$

Step 2: $O((m_2 - 1) N_2)$

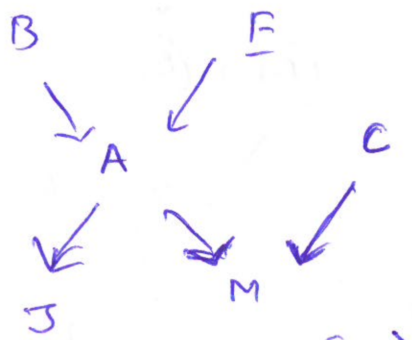
$\vdots$

$\sum_k (m_k - 1) N_k$. $O(m_n - 1) N_n)$

Variable ordering

9 -1

9-1
9-2
9-3



$$\sum_{B,A,E,J,M,C} f_1(B)\, f_2(E)\, f_3(C)\, f_4(A,B,E)\, f_5(J,A)\, f_6(M,A,C)$$

- if we choose to eliminate A first.

$$f_4(A,B,E)\; f_5(J,A)\; f_6(M,A,C)$$

$$f_7\,(\underbrace{A, B, C, E, J}_{6 \text{ vars.}})$$

- if we choose to eliminate B or E first then we can reduce this max size!

9-3

↳ In general finding the best ordering is NP-~~complete~~ hard!

↳ A greedy choice works ok for most cases

Oct 19th Lecture

Plan:

VE

(1) Recap of VE

(2) complexity of VE

- → Two operations: Join & Marginalize
- → Complexities of operations
- → Total complexity

(3) Ordering of variables

Sampling

(1) Estimating probabilities via sampling

(2) Rejection Sampling

(3) Likelihood-weighted

(4) Gibbs sampling.

② (b) Marginalize

$f_2(E)$ & $f_6(B,E)$
$\Rightarrow f_7(B)$

$$\sum_E f_7'(B) = f_7(B)$$

B	E	$f_7'(B)$
b	e	$x_1 y_1$
b	-e	$x_2 y_2$
-b	e	$x_1 y_3$
-b	-e	$x_2 y_4$

$\Rightarrow$

	$f_7(B)$
b	$x_1 y_1 + x_2 y_2$
-b	$x_1 y_3 + x_2 y_4$

③ Complexities of Join & Marginalize

Join : If output factor has scope q variables each w/ $\leq \gamma$ values then size of o/p factor $N = \gamma^{q}$
(# rows)

If input factors are m in number, In each row we have $(m-1)$ multiplicans

$\Rightarrow$ Join complexity : $O(N(m-1))$

Marginalize : One addition for one rows in o/p factor.

$\Rightarrow$ Marg. complexity : $O(N)$
$\hookrightarrow$ size of the o/p factor.

(4) Marginalization cost:

$\forall k \; N_k \leq N$

Each marg step $= O(N_k)$

Total marg cost $= \sum_{k=1}^{n} N_k \leq \sum_{k=1}^{n} N$

$\leq N \cdot n$

n → # of vars.

N → # size of the largest factor

Total cost of VE $= \underbrace{O(Nn)}_{\text{join}} + \underbrace{O(Nn)}_{\text{marg}}$

$= O(Nn)$

Good news? No

- Linear in N & n but

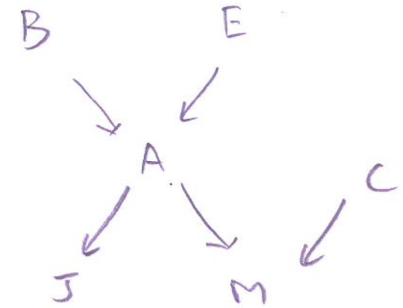
$N = r^{q}$ → size of the largest domain

q → scope of the o/p factor

$0, \ldots 0_q$

$r \times r \times \ldots r$

(5) Variable Ordering



$$\sum_{BAEJMC} f_1(B)\, f_2(E)\, f_3(C)\, f_4(A,B,E)\, f_5(J,A)\, f_6(M,A,C)$$

- if we choose to eliminate A first then:

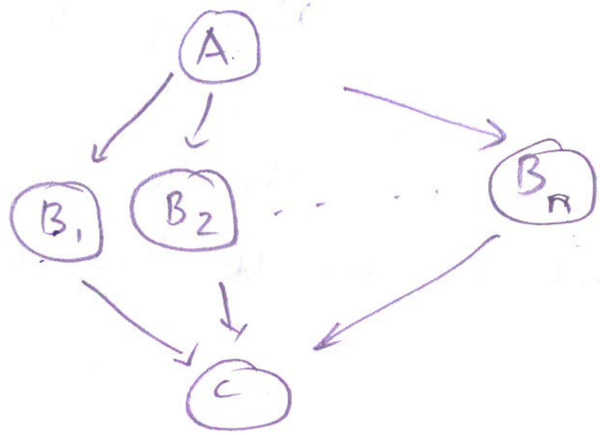
$$\sum_{BEJMC} f_1(B) f_2(E) f_3(C) \underbrace{\sum_{A} f_4(A,B,E)\, f_5(J,A)\, f_6(M,A,C)}_{f_7(\cancel{A},\ \underline{B, C, E, J, M})}$$

$\hookrightarrow$ 5 vars.

- if we choose B or E then:

you can have reduced size factors! $\rightarrow$ Exercise.

⑤ - contd Variable Ordering : Extreme example.



N - max size can be exponential in size of n/w

How?

→ Try to eliminate A first

$$\sum_A f_1(A, B_1)\, f_2(A, B_2) \ldots f_n(A, B_n)$$

$$= f_{n+1}(B_1 \ldots B_n) \quad N_{n+1} = \gamma^{\underline{\underline{n}}}$$

size of BN

→ Try eliminating B_i's one at a time

B_1 : $f_1(B_1) = \sum_{B_1} P(B_1 | A)$

$$\Downarrow$$

$$f_{n+1}(A)$$

B_2 ... $f_2(B_2) \ldots f_{n+1}(A)$ ...

$$f_{n+2}(A)$$