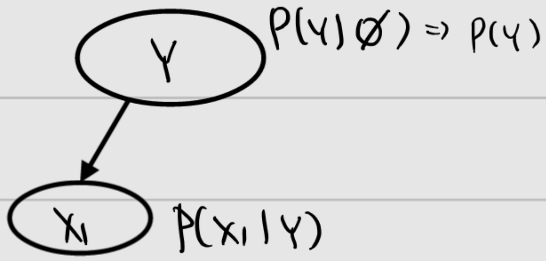
기존 Naive Bayes → feature은 conditional independence 가정

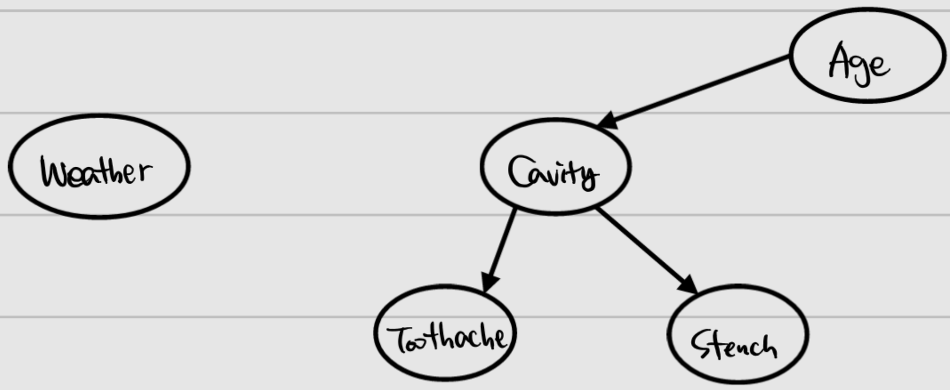
$$f_{NB}(x) = \operatorname{argmax}_{\forall y} P(Y=y)\prod_{1\le i\le d} P(X_i=x_i \mid Y=y)$$

Bayesian Network

↳ graphical notation of random variables
conditional independence
obtain represetation of full joint distributions

- acyclic, directed graph
- $P(X_i \mid parent(X_i))$
- direct influence from P → C
  (parent) (child)

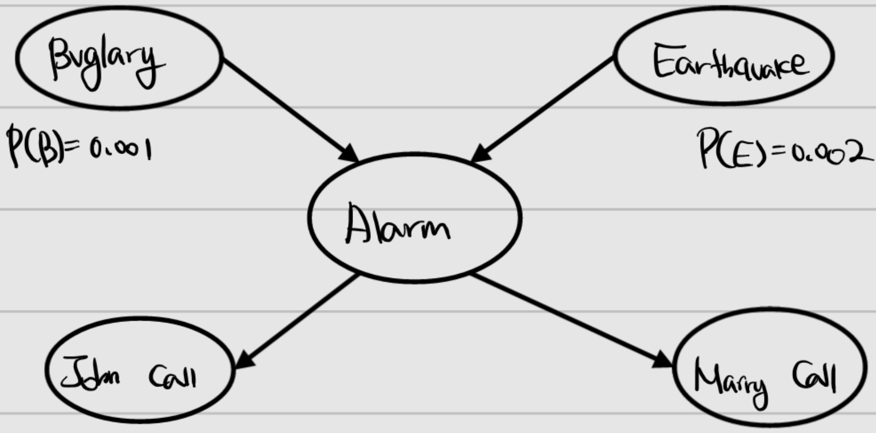




Weather ⊥ Cavity (independent)

Toothache, Stench conditionally independent given Cavity

Cavity influences probability of Toothache, Stench



$P(A|B,E)$

B	E	A
T	T	0.95
T	F	0.94
F	T	0.29
F	F	0.001

$P(M|A)$

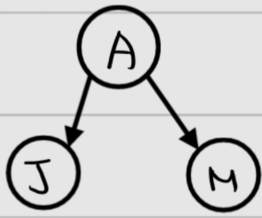
A	M
T	0.7
F	0.01

$P(J|A)$

A	J
T	0.9
F	0.05

Typical local structures

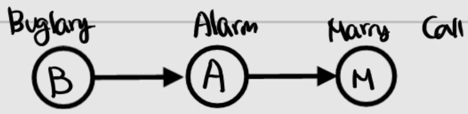
① Common Parent



→ $J \perp M|A$

* A가 given → J, M은 independent $P(J,M|A) = P(J|A) \cdot P(M|A)$

② Cascading

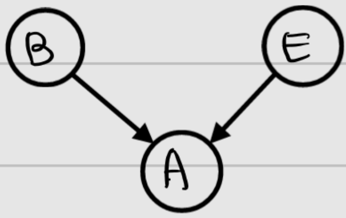


$B \perp M|A$

* A가 given → B와 M은 independent $P(M|B,A) = P(M|A)$

알람이 울려 Mary Call O → 도둑의 영향은? X

③ V-structure



$\sim(B \perp E|A)$

알람O → 지진이 없다? 도둑일 확률↑

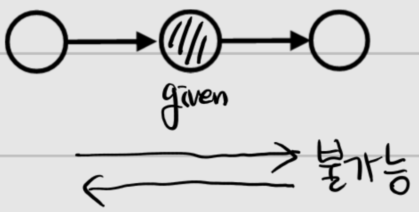
영향을 미침

< Bayes Ball Algorithm >

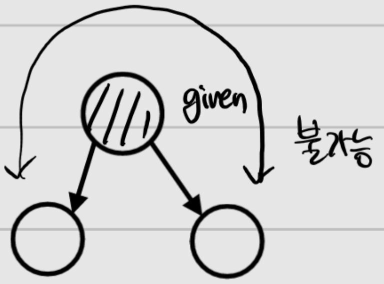
공을 굴려 갈수 있다? → dependent

없다 → independent

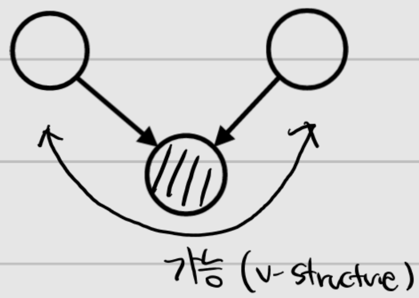
①

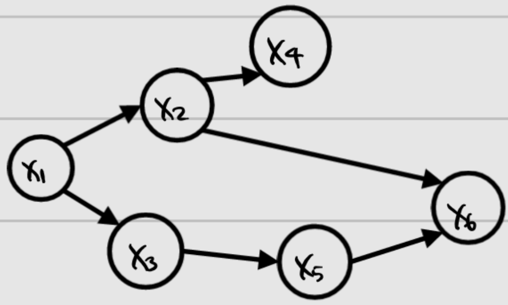


②

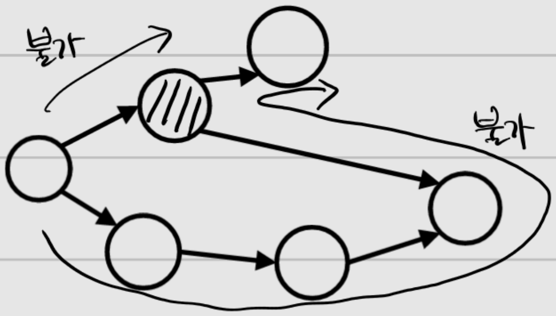


③



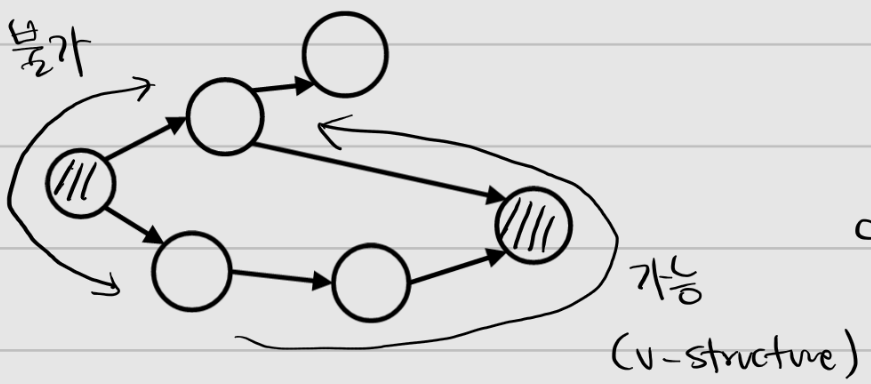


① $X_1 \perp X_4 \ \overset{\text{given}}{\{X_2\}}$



independent

② $X_2 \perp X_3 \ \{X_1, X_6\}$

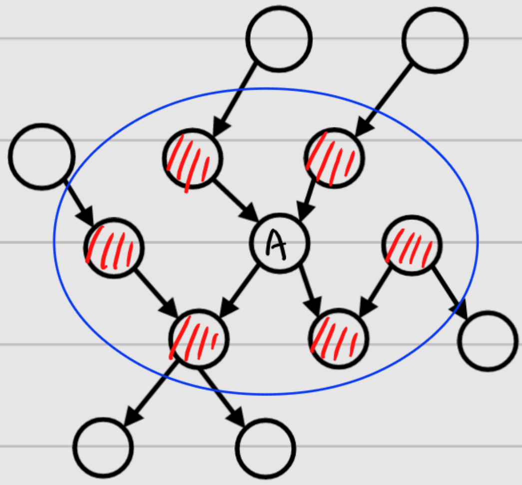


dependent

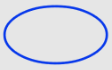
< Markov blanket >

$P(A \mid \text{Blanket}, B) = P(A \mid \text{Blanket})$

$\text{Blanket} = \{\text{parents, child, child parents}\}$



: $A \perp B \mid \text{Blanket}$

 내부를 알면 외부와 independent

Parents, children → Cascading
Child parents → V-structure 막아줌

< D-seperation > Directly separated

$X \perp Z \mid Y$

Y가 given → $X \perp Z$

blanket 밖

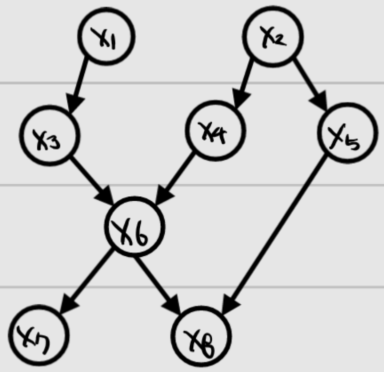
위의 $A \perp B \mid \text{Blanket}$ 과 유사 개념

Factorization

$$P(X_1, X_2, X_3)$$
$$= P(X_1 \mid X_2, X_3) \cdot P(X_2 \mid X_3) \cdot P(X_3)$$

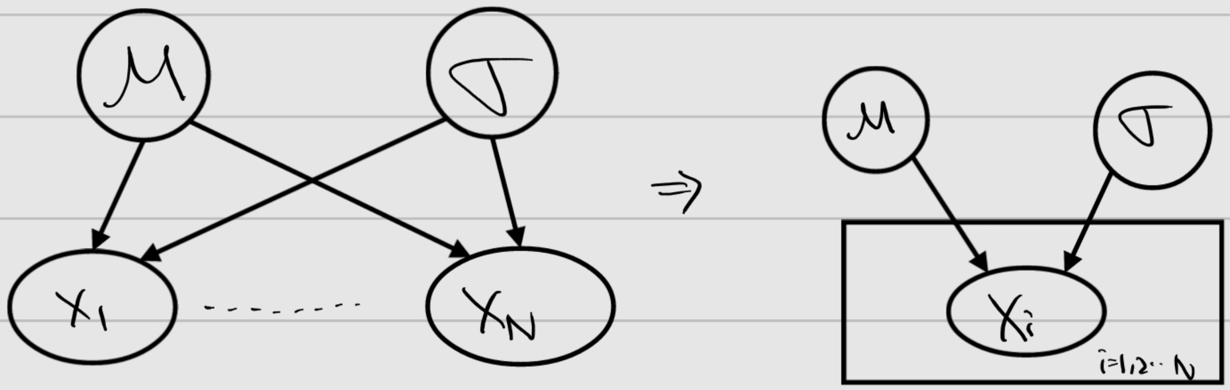
만약 조건 $X_1 \perp X_2 \mid X_3 \Rightarrow P(X_1 \mid X_2, X_3) = P(X_1 \mid X_3)$

parameter 수 $\downarrow$

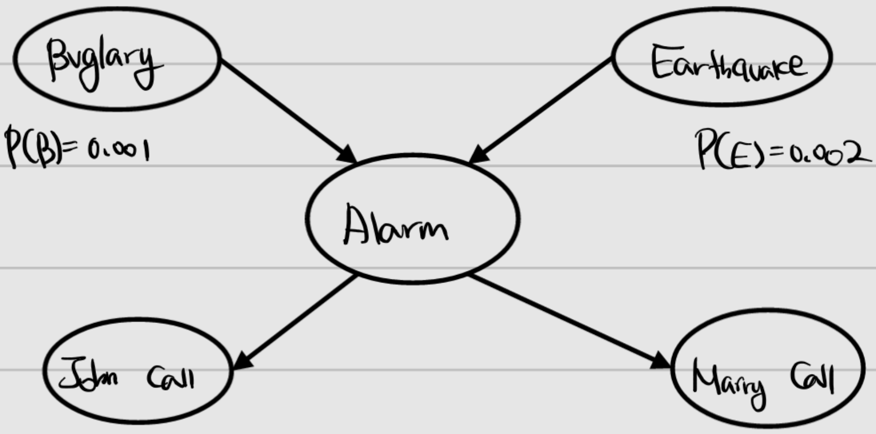


$$P(X_1, X_2, X_3 \cdots X_8)$$
$$= P(X_1) \cdot P(X_2) \cdot P(X_3 \mid X_1) \cdot P(X_4 \mid X_2) \cdot P(X_5 \mid X_2) \cdot P(X_6 \mid X_3, X_4) \cdot P(X_7 \mid X_6) \cdot P(X_8 \mid X_5, X_6)$$

Plate Notation



$$P(D \mid \theta) = P(X_1 X_2 \cdots X_N \mid \mu, \sigma)$$



$P(A|B,E)$

B	E	A
T	T	0.95
T	F	0.94
F	T	0.29
F	F	0.001

$P(M|A)$

A	M
T	0.7
F	0.01

$P(J|A)$

A	J
T	0.9
F	0.05

X: $\{X_1, X_2 \cdots X_N\}$ all random variable

X_V: evidence variable, 관측한 내용, 알아볼 내용 ...

X_H: $X - X_V$, hidden variable

Q1. Likelihood → $P(B, \text{Marry call}) = ?$

John, Alarm, Earthquake → hidden

B, M → evidence

factorization : $P(B)P(E)P(A)P(M)P(J)$

$= P(B)P(E)P(A|B,E)\,P(M|A)\,P(J|A)$

$$P(X_V) = \sum_{X_1} P(\underbrace{X_H, X_V}_{}) = \sum_{X_1}\sum_{X_2} \cdots \sum_{X_K} P(X_1, X_2 \cdots X_K, X_V)$$

$'' X$

위를 보면 $\sum_{E}\sum_{J}\sum_{A} P(B, E, A, M, J)$

$$= \sum_{E}\sum_{J}\sum_{A} P(B)P(E)P(A|B,E)P(M|A)P(J|A)$$

계산

Q2. Conditional Probability

$$P(A \mid B=T, MC=T) = ?$$

$X_H \rightarrow \{Y, Z\}$ 　 Y: interested hidden

Z: uninterested hidden

$$P(Y \mid X_v) = \sum_{z} P(Y, Z=z \mid X_v)$$

$$= \sum_{z} \frac{P(Y, Z, X_v)}{P(X_v)}$$ → fully joint $P(X_H, X_V) = P(X)$

$$= \sum_{z} \frac{P(X)}{\sum_{y}\sum_{z} P(Y=y, Z=z, X_v)}$$

Q3. Assignment

→ Given set of evidence → most probable assignment

Prediction : $B, E \rightarrow A$ 　 $P(A|B,E)$

Diagnosis 　 $A \rightarrow B, E$ 　 $P(B,E|A)$

Marginalization, Elimination

$$P(A=T, B=T, Mc=T) = \sum_{J}\sum_{E} P(A, B, E, Jc, Mc)$$

hidden : 2개
(Jc, E)

$$= \sum_{Jc}\sum_{E} P(A|B,E)\,P(E)\,P(B)\,P(Jc|A)\,P(Mc|A)$$

E에 대해 영향X

노드 수↑ ⇒ 계산 복잡

Σ 영향 X

Σ 밖으로 뺀 후 계산 O

∴ 변형식 ⇒ $P(B)\cdot P(Mc|A)\cdot \sum_{Jc} P(Jc|A)\sum_{E} P(A|B,E)\,P(E)$

- Preliminary

$$P(E|Jc, Mc) = \alpha\, P(E, Jc, Mc)$$

$$\alpha = \frac{1}{P(Jc, Mc)}$$

- Joint probability $P(E=T, Mc=T, Jc=T)$

$\Rightarrow \sum_{A}\sum_{B} P(B, E, A, Mc, Jc)$

↳ tree 층 대로 order
(topological order)

$$= \alpha\, P(E)\cdot \sum_{B} P(B) \sum_{A} P(A|B,E)\, P(Jc|A)\, P(Mc|A)$$

기존

A	P(J|A)	P(M|A)
T	0.9	0.7
F	0.05	0.01

⇒

변형

A	$f_J(A)$
T	0.9
F	0.05

×

A	$f_M(A)$
T	0.7
F	0.01

⇒

A	$f_{JM}(A)$
T	0.63
F	0.0005

$$\alpha P(E) \cdot \sum_B P(B) \sum_A P(A|B,E)\, P(J_c|A)\, P(M_c|A)$$

$$= \alpha f_E(e) \sum_B f_B(b) \sum_A f_A(a,b,e) \underbrace{f_J(a)\, f_M(a)}_{f_{JM}(a)}$$

A	B	E	$f_A(a,b,e)$
T	T	T	0.95
T	T	F	0.94
T	F	T	0.29
T	F	F	0.01
F	T	T	0.05
F	T	F	0.06
F	F	T	0.71
F	F	F	0.95

$\times$

A	$f_{JM}(a)$
T	0.63
F	0.0005

$\Rightarrow$

A	B	E	$f_{AJM}(a)$
T	T	T	0.95 × 0.63
⋮			

$$\Rightarrow \alpha f_E(e) \sum_B f_B(b) \sum_A f_{AJM}(a,b,e)$$

⋮

계산

Potential function

$$A \leftarrow B \leftarrow C \leftarrow D$$

$$P(A,B,C,D) = \underline{P(A|B)\cdot P(B|C)\cdot P(C|D)\cdot P(D)}$$

potential function : pdf X, once normalized → it can be pdf

Clique : Sub graph, fully connected

ex) $B \rightarrow A$, $C \rightarrow B$, $D \rightarrow C$

$$(A,B) - \boxed{B} - (B,C) - \boxed{C} - (C,D)$$

Clique

↑ Separator

→ Clique 연결

Potential function on nodes

$\psi(A,B)$, $\psi(B,C)$, $\psi(C,D)$

Potential function on links

$\phi(B)$, $\phi(C)$

ex1)

$$P(A,B,C,D) = P(U) = \frac{\prod_N \psi(N)}{\prod_L \phi(L)} = \frac{\psi(A,B)\,\psi(B,C)\,\psi(C,D)}{\phi(B)\,\phi(C)}$$

$$\psi(A,B) = P(A|B)$$
$$\psi(B,C) = P(B|C)$$
$$\psi(C,D) = P(C|D)\cdot P(D)$$

$$\phi(B) = 1$$
$$\phi(C) = 1$$

ex2)

$$P(A,B,C,D) = P(U) = \frac{\prod_N \psi(N)}{\prod_L \phi(L)} = \frac{\psi^*(A,B)\,\psi^*(B,C)\,\psi^*(C,D)}{\phi(B)\,\phi(C)}$$

$$\psi^*(A,B) = P(A,B) \qquad \phi(B) = P(B)$$
$$\psi^*(B,C) = P(B,C) \qquad \phi(C) = P(C)$$
$$\psi^*(C,D) = P(C,D)$$

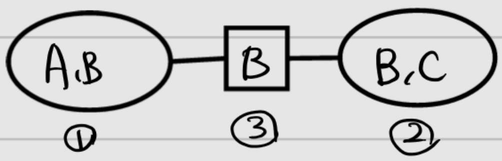
< P(B) 표현 >

① $\sum_A \psi(A,B) = \sum_A P(A,B) = P(B)$

② $\sum_C \psi(B,C) = \sum_C P(B,C) = P(B)$

③ $\phi(B) = P(B)$

①, ②, ③의 P(B)는 모두 동일해야함.



만약 $P(A,B) \rightarrow$ A=1 임이 관찰된다면? $\rightarrow P(A=1, B)$

ψ 가 변하게 된다.

ψ 하나 변해도 뒤의 ψs 들의 변화 이끌어냄

⇒ Belief Propagation

How to propagate the belief?

↳ Absortion Rule

Assume $\psi^*(A,B)$, $\psi(B,C)$, $\phi(B)$

$$\phi^*(B) = \sum_A \psi^*(A,B)$$

$\phi(B)$의 update는 쉬움

marginalize 해주기

$\Rightarrow \psi(B,C)$의 update 필요

$$\psi^*(B,C) = \psi(B,C) \cdot \frac{\phi^*(B)}{\phi(B)}$$

$$\sum_C \psi^*(B,C) = \sum_C \psi(B,C) \cdot \frac{\phi^*(B)}{\phi(B)}$$

$$= \frac{\phi^*(B)}{\phi(B)} \cdot \sum_C \psi(B,C)$$

$$= \frac{\phi^*(B)}{\phi(B)} \cdot \phi(B) = \phi^*(B) = \sum_A \psi^*(A,B)$$

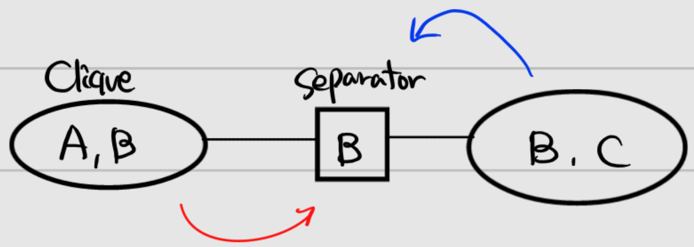
$$\therefore \sum_C \psi^*(B,C) = \sum_A \psi^*(A,B)$$

$$A \leftarrow B \leftarrow C$$

Initialize potential function

$$\psi(A,B) = P(A|B)$$
$$\psi(B,C) = P(B|C) \cdot P(C)$$
$$\phi(B) = 1$$

Q1 $P(b) = ?$



update

→ $$\phi^{*}(B) = \sum_{A} \psi(A,B) = 1$$

$$\psi^{*}(B,C) = \psi(B,C) \cdot \frac{\phi^{*}(B)}{\phi(B)} = \psi(B,C)$$
$$= P(B|C) \cdot P(C) = P(B,C)$$

→ local consistency maintain을 위해

$$\phi^{**}(B) = \sum_{C} \psi^{*}(B,C) = \sum_{C} P(B,C) = P(B)$$

$$\psi^{*}(A,B) = \psi(A,B) \cdot \frac{\phi^{**}(B)}{\phi^{*}(B)} = \psi(A,B) \cdot P(B)$$
$$= P(A|B) \cdot P(B) = P(A,B)$$

→ 다시 update 방향으로 가보면?

$$\phi^{***}(B) = \sum_{A} \psi^{*}(A,B) = \sum_{A} P(A,B) = P(B)$$

$\phi^{**}(B) = \phi^{***}(B)$ 유지됨

local consistency maintain O

Q2. $P(B \mid A=1, C=1)$

$\rightarrow$

조건

A=1 인 상황에서 potential만 활용 marginalize out

$$\phi^*(B) = \sum_A \psi(A,B)\, \delta(A=1) = P(A=1 \mid B)$$

기존 $\sum_A \psi(A,B) = P(A \mid B)$

$$\psi^*(B,C) = \psi(B,C) \cdot \frac{\phi^*(B)}{\phi(B)}$$

$$= P(B \mid C=1) \cdot P(C=1) \cdot \frac{P(A=1 \mid B)}{1}$$

$\leftarrow$

$$\phi^{**}(B) = \sum_C \psi^*(B,C)\, \delta(C=1) = P(B \mid C=1) \cdot P(C=1)\, P(A=1 \mid B)$$

$$\psi^*(A,B) = \psi(A,B) \cdot \frac{\phi^{**}(B)}{\phi^*(B)} = P(A=1 \mid B) \cdot \frac{P(B \mid C=1) \cdot P(C=1) \cdot P(A=1 \mid B)}{P(A=1 \mid B)}$$

$$= P(B \mid C=1) \cdot P(C=1) \cdot P(A=1 \mid B)$$

$\rightarrow$

$$\phi^{***}(B) = \sum_A \psi^*(A,B)\, \delta(A=1) = P(B \mid C=1) \cdot P(C=1) \cdot P(A=1 \mid B)$$

$$\therefore P(B \mid A=1, C=1) = P(B \mid C=1) \cdot P(C=1) \cdot P(A=1 \mid B)$$