Bayes' Rule:

$$\underbrace{p(a|b)}_{\text{posterior}} = \frac{\overbrace{p(a,b)}^{\text{joint}}}{\underbrace{p(b)}_{\text{normalizing const.}}} = \frac{\overbrace{p(b|a)}^{\text{likelihood}}\,\overbrace{p(a)}^{\text{prior}}}{\underbrace{p(b)}_{\text{normalizing const.}}}$$

Marginalization: $p(A=a) = \sum_{b \in B} P(A=a, B=b)$

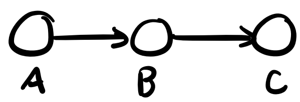
for binary vars, $P(A=1) = P(A=1, B=0) + P(A=1, B=1)$

might be useful for HW: $P(A=1) = P(A=1|B=0)P(B=0) + P(A=1|B=1)P(B=1)$

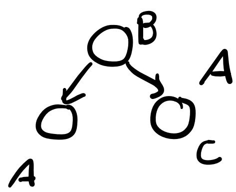
Independence: $P(A,B) = P(A)P(B)$ if $A \perp\!\!\!\perp B$

$P(A,B|C) = P(A|C)P(B|C)$ if $A \perp\!\!\!\perp B \mid C$

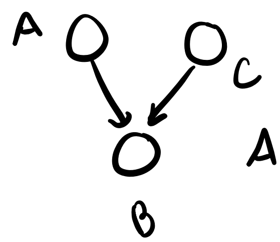
Bayesian Network: Represents conditional independencies.



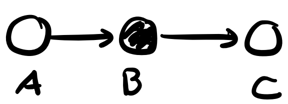
$A \not\!\perp\!\!\!\perp C$



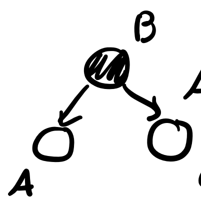
$A \not\!\perp\!\!\!\perp C$



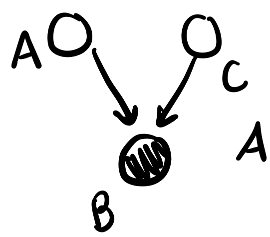
$A \perp\!\!\!\perp C$



$A \perp\!\!\!\perp C \mid B$



$A \perp\!\!\!\perp C \mid B$

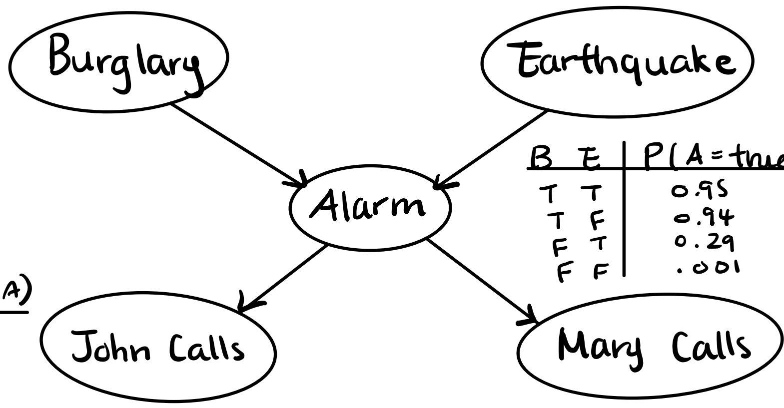


$A \not\!\perp\!\!\!\perp C \mid B$

Example

P(B=true)
.001

P(E=true)
.002



B	E	P(A=true | B,E)
T	T	0.95
T	F	0.94
F	T	0.29
F	F	.001

CPT

A	P(J=true|A)
T	0.90
F	0.05

A	P(M=true|A)
T	0.70
F	0.01

Joint distribution:

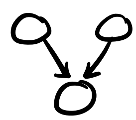
$$P(B, E, A, J, M) = P(B)\,P(E)\,P(A|B,E)\,P(J|A)\,P(M|A)$$

ex) $P(B=F, E=F, A=T, J=T, M=T) = (0.999)(0.998)(.001)(.90)(.70)$

$$P(A=T) = P(A=T, B=T, E=T) + P(A=T, B=T, E=F) + P(A=T, B=F, E=T) + P(A=T, B=F, E=F)$$

$$= P(A=T|B=T,E=T)\,P(B=T,E=T) + P(A=T|B=T,E=F)\,P(B=T,E=F) + P(A=T|B=F,E=T)\,P(B=F,E=T) + P(A=T|B=F,E=F)\,P(B=F,E=F)$$

$B \perp\!\!\!\perp E$ but $B \not\perp\!\!\!\perp E | A$: why?



explaining away

can substitute $P(B,E) = P(B)P(E)$ above.

$$P(A=T) = P(A=T|B=T,E=T)\,P(B=T)\,P(E=T) + P(A=T|B=T,E=F)\,P(B=T)\,P(E=F) + P(A=T|B=F,E=T)\,P(B=F)\,P(E=T) + P(A=T|B=F,E=F)\,P(B=F)\,P(E=F)$$