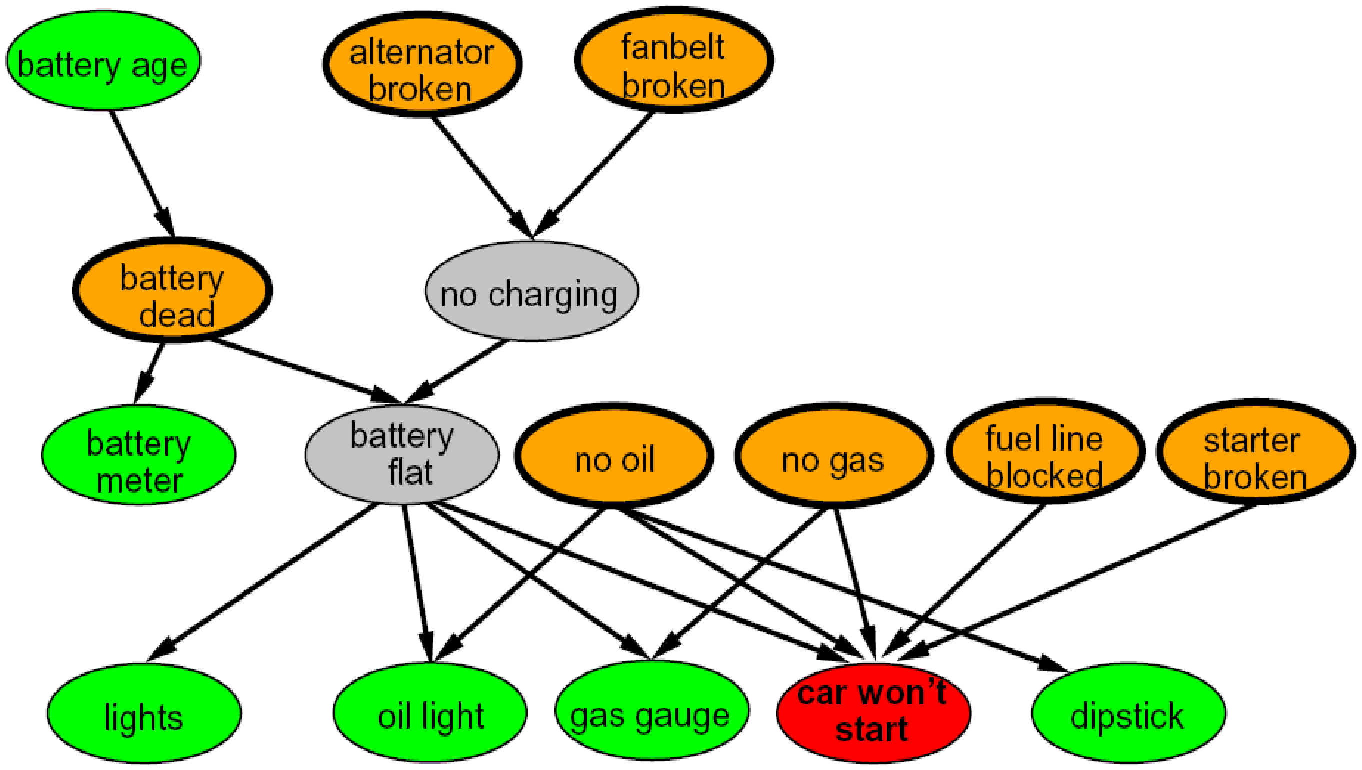
CSCI 3202: Intro to Artificial Intelligence

Bayesian Networks, Part II

Rhonda Hoenigman
Department of Computer Science

Exam grades posted. Canvas messed up. Avg: 93.9%

Homework 6 posted later today.



Pick up your exam from Deans office front desk.

Bayesian Networks – example

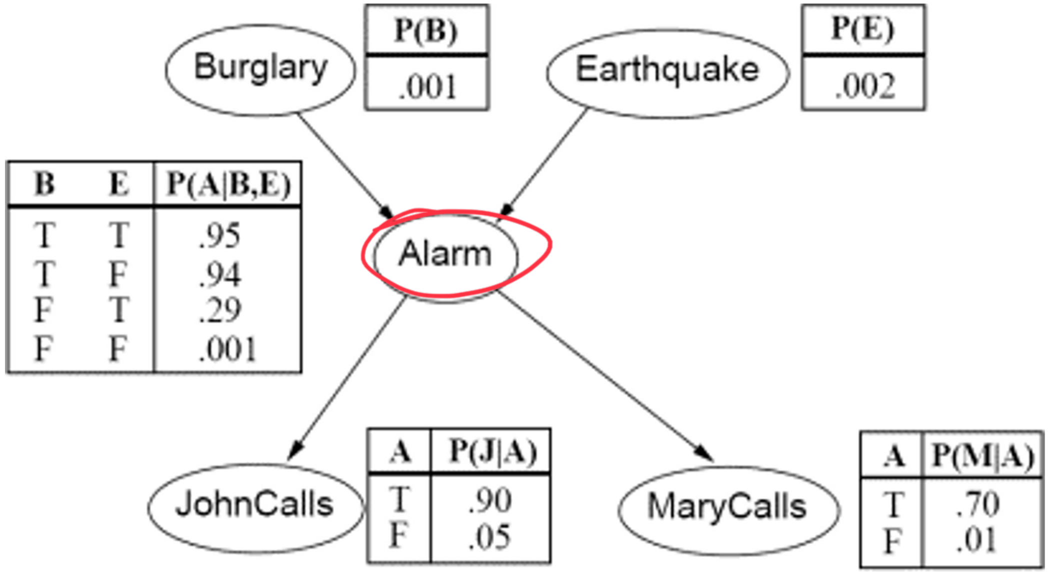
Bayes nets implicitly encode joint distributions as a product of the local conditional distributions:

$$P(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid \mathrm{parents}(X_i))$$

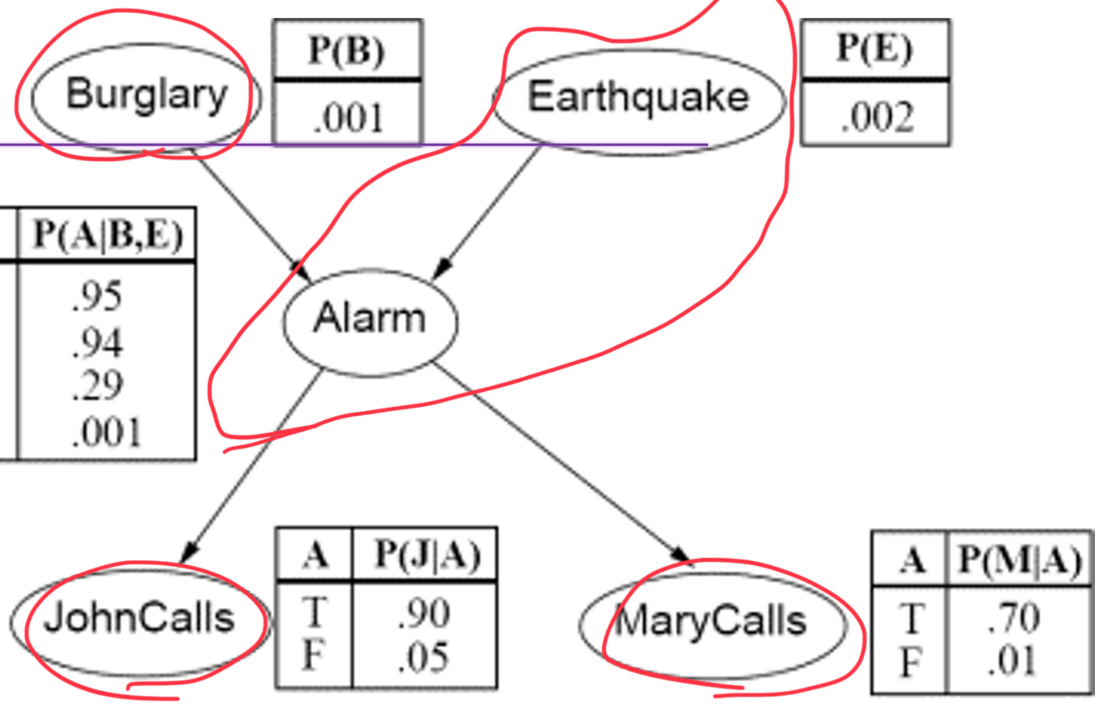
P(b | j, m) =

Burglary given J and M call

$$\frac{P(B, J, M)}{P(J, M)}$$



Bayesian Networks



P(B)
.001

P(E)
.002

B	E	P(A|B,E)
T	T	.95
T	F	.94
F	T	.29
F	F	.001

A	P(J|A)
T	.90
F	.05

A	P(M|A)
T	.70
F	.01

Example: S'pose we know the alarm has gone off, and both John and Mary have called to warn us. What is the probability that we have been burgled?

→ $P(+b \mid +j, +m) = ?$

Query variables: what we want the **posterior** probability of X, **given** some **evidence**

X = B   Burglary

Evidence variables: the variables we are given an assignment of (the **data**)

E = [+j, +m]

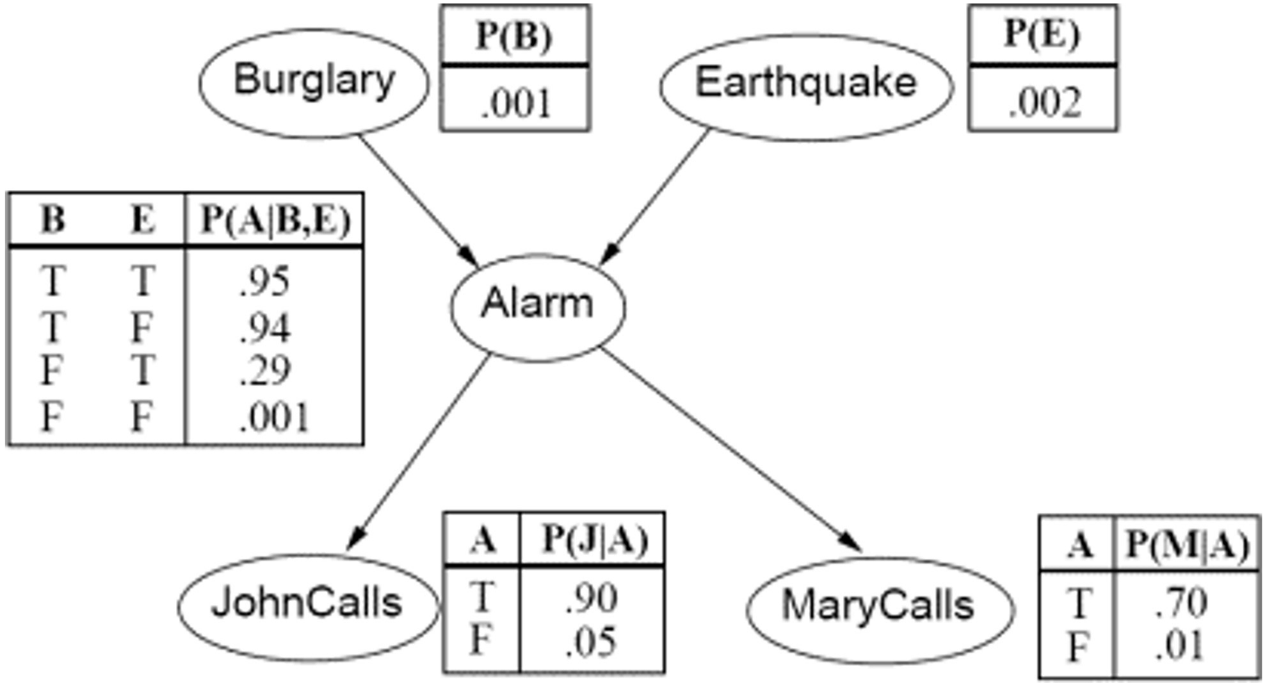
Hidden variables: the non-evidence, non-query variables

y = [E, A]   Earthquake, Alarm

Bayesian Networks

Example: *S'pose we know the alarm has gone off, and both John and Mary have called to warn us. What is the probability that we have been burgled?*

→ $P(+b \mid +j, +m) = ?$

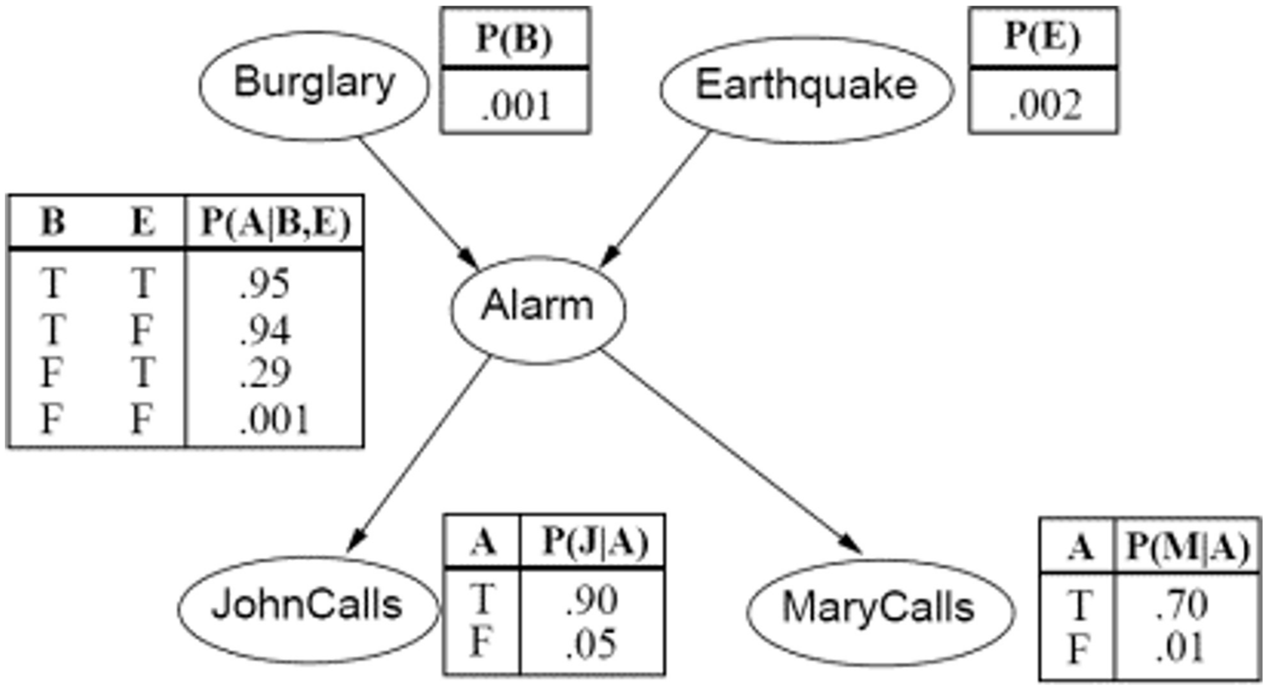


Calculation by **enumeration:**

Bayesian Networks

Example: S'pose we know the alarm has gone off, and both John and Mary have called to warn us. What is the probability that we have been burgled?

→ *P*(+b | +j, +m) = ?



Calculation by **enumeration:**

$$P(B \mid j, m) = \frac{P(B, j, m)}{P(j, m)} \Rightarrow \alpha = \frac{1}{P(j, m)} \Rightarrow P(B \mid j, m) = \alpha P(B, j, m)$$

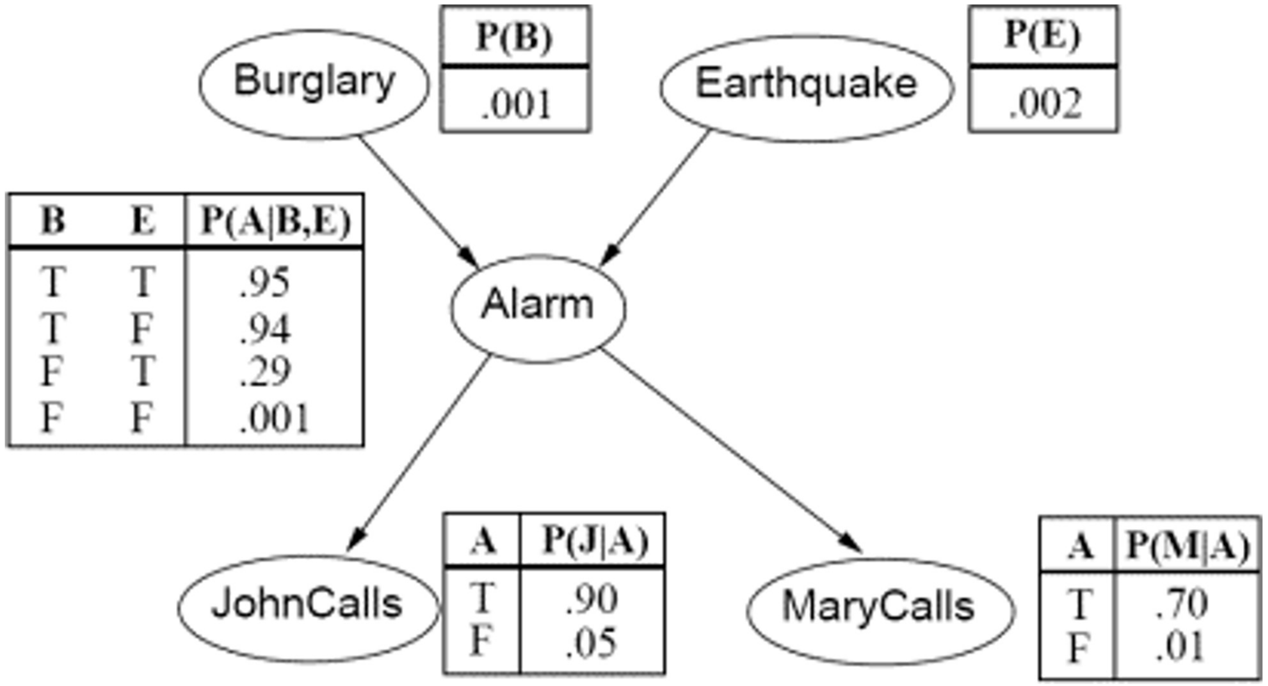
numerator for now

We'll do our thing, then figure out the normalizing constant α later

Bayesian Networks

Example: Suppose we know the alarm has gone off, and both John and Mary have called to warn us. What is the probability that we have been burgled?

→ *P*(+b | +j, +m) = ?



P(B)
.001

P(E)
.002

B	E	P(A|B,E)
T	T	.95
T	F	.94
F	T	.29
F	F	.001

A	P(J|A)
T	.90
F	.05

A	P(M|A)
T	.70
F	.01

Alarm = T, F

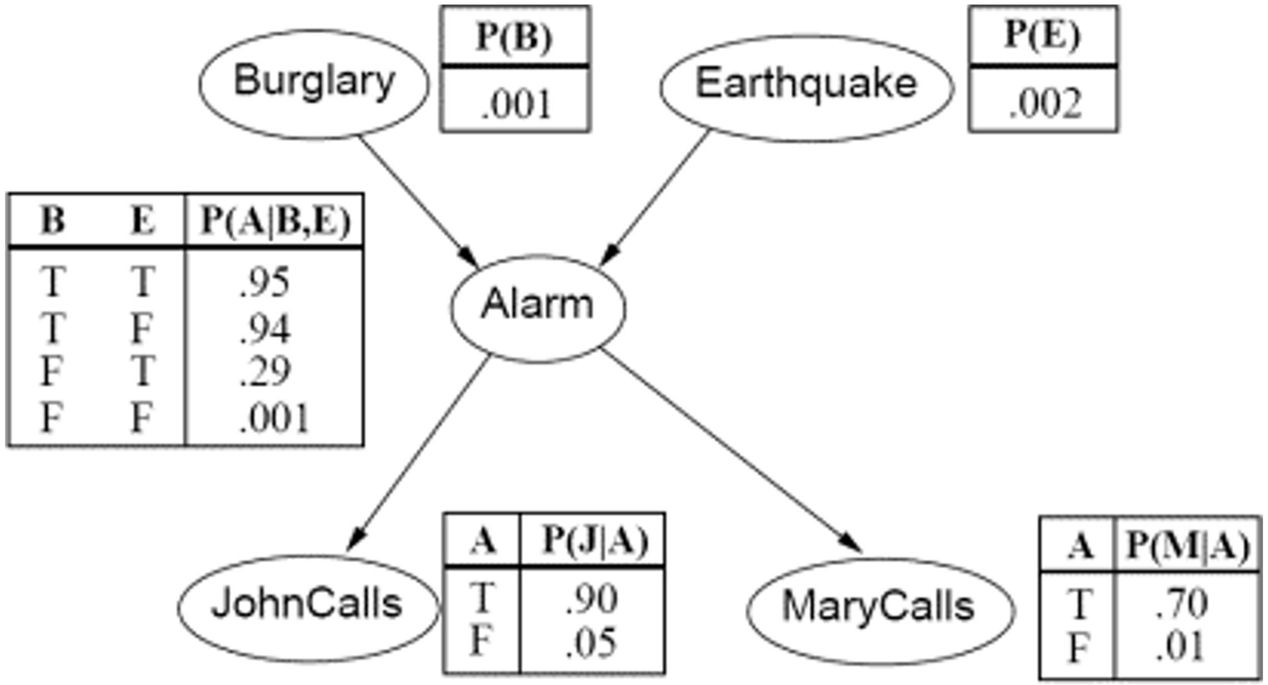
Calculation by **enumeration:**

$$\begin{aligned} P(B \mid j, m) &= \alpha P(B, j, m) \\ &= \alpha \sum_a P(B, j, m \mid a) P(a) \\ &= \alpha \sum_a P(B, j, m, a) \end{aligned}$$

Bayesian Networks

Example: Suppose we know the alarm has gone off, and both John and Mary have called to warn us. What is the probability that we have been burgled?

→ $P(+b \mid +j, +m) = ?$



P(B)
.001

P(E)
.002

B	E	P(A|B,E)
T	T	.95
T	F	.94
F	T	.29
F	F	.001

A	P(J|A)
T	.90
F	.05

A	P(M|A)
T	.70
F	.01

Earthquake
T or F

Calculation by **enumeration:**

$$P(B \mid j, m) = \alpha \sum_a P(B, j, m, a)$$
$$= \alpha \sum_e \sum_a P(B, j, m, a \mid e) P(e)$$
$$= \alpha \sum_e \sum_a P(B, j, m, a, e)$$

Bayesian Networks

<u>Example</u>: Suppose we know the alarm has gone off, and both John and Mary have called to warn us. What is the probability that we have been burgled?

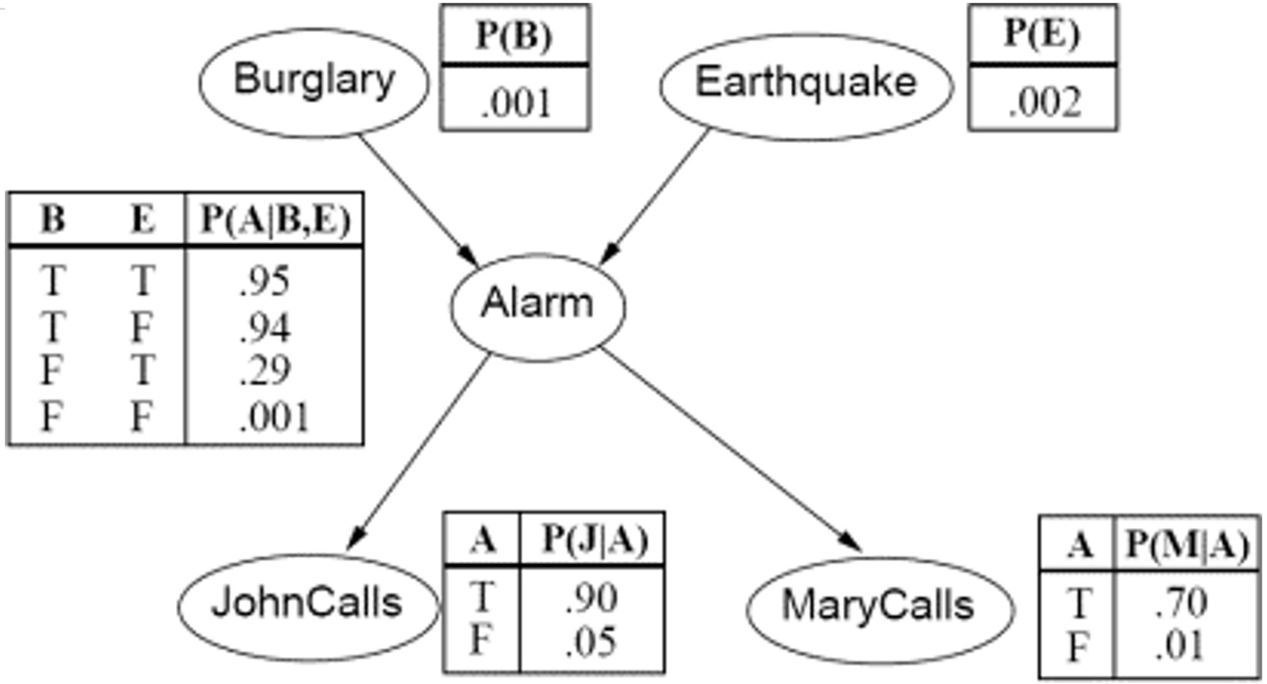
→ *P*(+b | +j, +m) = ?

FINALLY we have:

$$P(B \mid j, m) = \alpha \sum_e \sum_a P(B, j, m, a, e)$$

From the conditional independence of the Bayes net:

$$P(B \mid j, m) = \alpha \sum_e \sum_a \prod_{i=1}^{n} P(x_i \mid \text{parents}(X_i))$$



Bayesian Networks

Example: Suppose we know the alarm has gone off, and both John and Mary have called to warn us. What is the probability that we have been burgled?

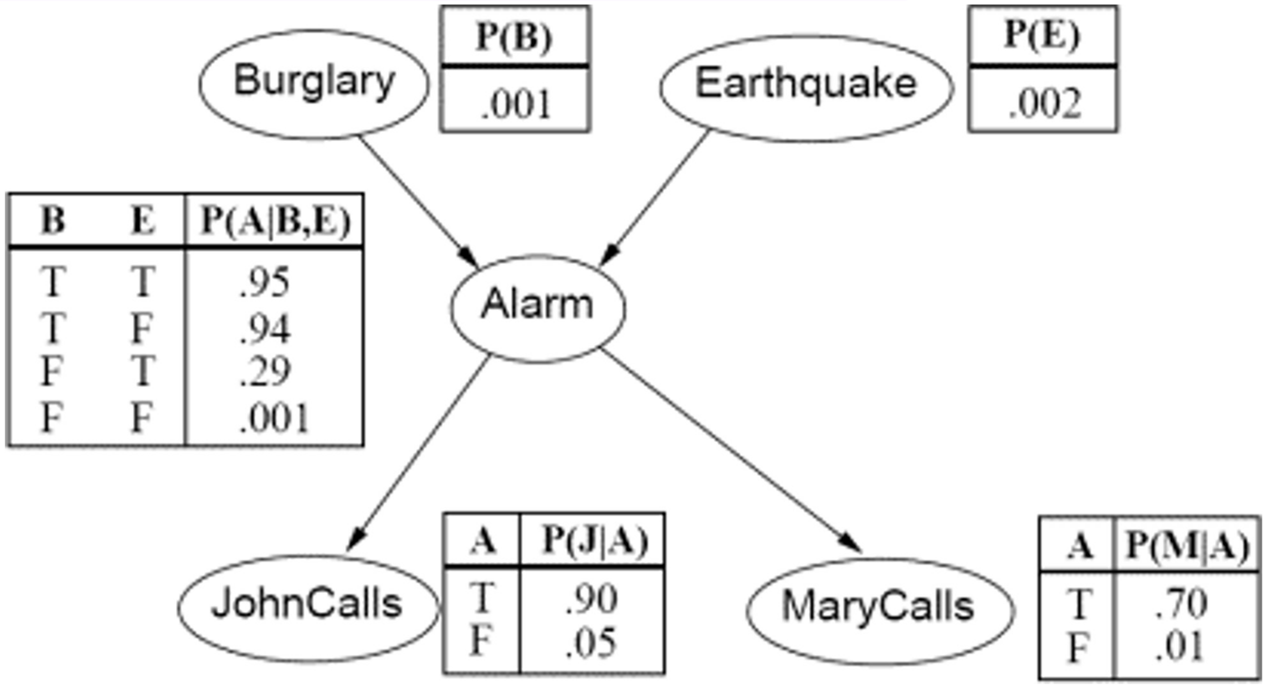
→ *P*(+b | +j, +m) = ?

So for this problem…

$$P(B \mid j, m) = \alpha \sum_{e} \sum_{a} \prod_{i=1}^{n} P(x_i \mid \text{parents}(X_i))$$

$$= \alpha \sum_{e} \sum_{a} P(B)P(e)P(a \mid B, e)P(j \mid a)P(m \mid a)$$

$$= \alpha P(B) \sum_{e} P(e) \sum_{a} P(a \mid B, e)P(j \mid a)P(m \mid a)$$



Bayesian Networks

How many terms in this summation?

Example:

$$P(B \mid j, m) = \alpha P(B) \sum_e P(e) \sum_a P(a \mid B, e) P(j \mid a) P(m \mid a)$$

P(B) P(E=T,F) P(A=T,F) P(J|A) P(M|A)

E=T, A=T, B=T, J=T, M=T

P(B) P(E) P(A|B,E) P(J|A) P(M|A) +

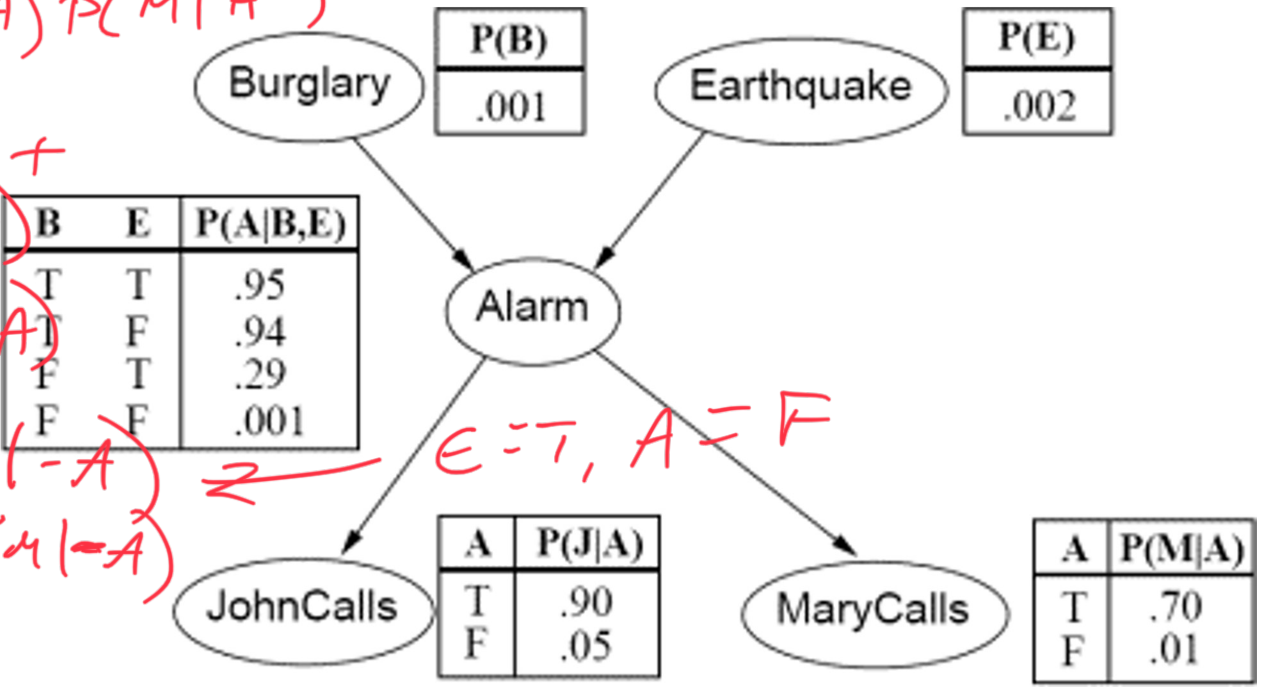
E=F, A=T

P(B) P(-E) P(A|B,-E) P(J|A) P(M|A)

E=T, A=F

P(B) P(E) P(-A|B,E) P(J|-A) P(M|-A)

P(B) P(-E) P(-A|B,-E) P(J|-A) P(M|-A)



Bayesian Networks

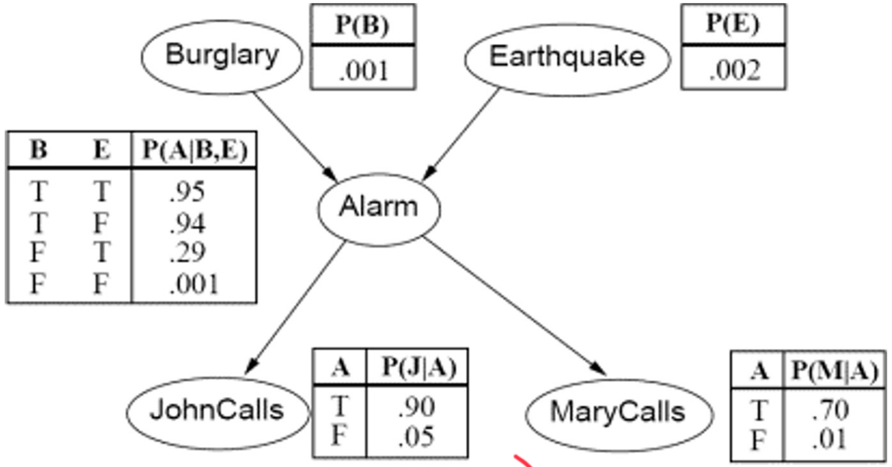
$P(B,J,M) =$

$.001 \times .002 \times .95 \times .90 \times .7 +$

$.001 \times .998 \times .94 \times .9 \times .7 +$

$.001 \times .002 \times .05 \times .05 \times .01 +$

$.001 \times .998 \times .06 \times .05 \times .01$



P(B)
.001

P(E)
.002

B	E	P(A|B,E)
T	T	.95
T	F	.94
F	T	.29
F	F	.001

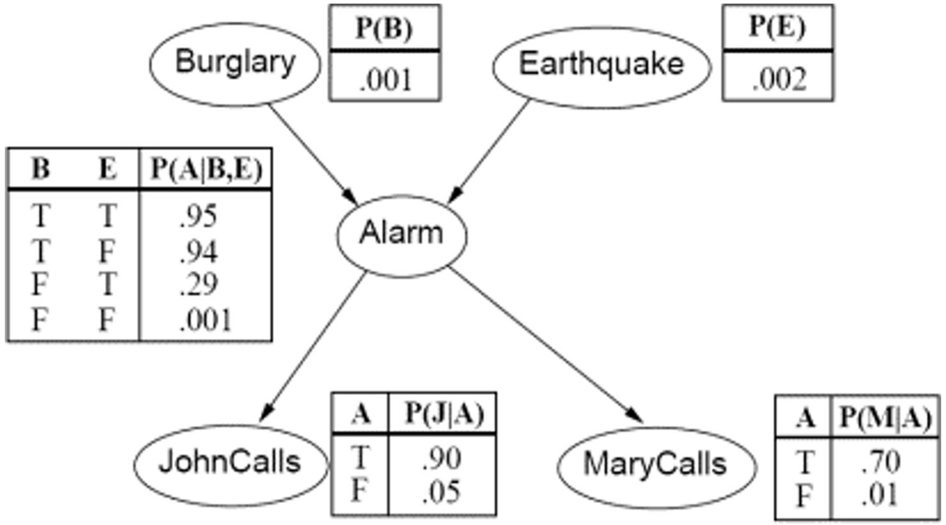
A	P(J|A)
T	.90
F	.05

A	P(M|A)
T	.70
F	.01

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

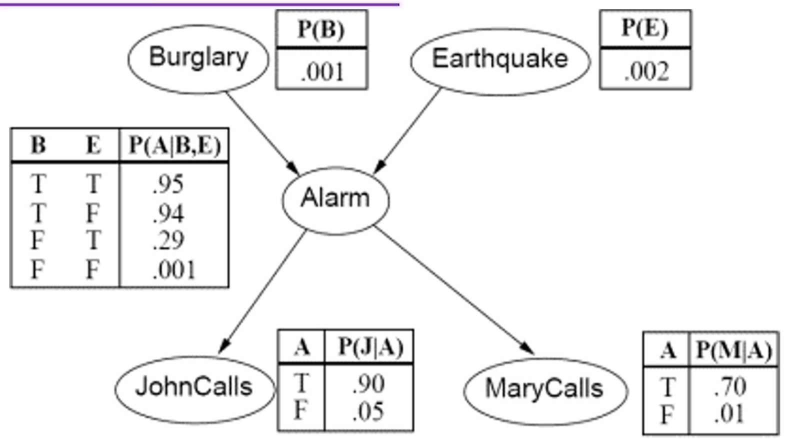
$$P(B|J,M) = \underline{P(B,J,M)}$$

Bayesian Networks



Bayesian Networks

Example: Find $P(Burglary = T \mid JohnCalls = T, MaryCalls = T)$



Entire Joint Probability Distribution:

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

$$P(Burglary = T \mid JohnCalls = T, MaryCalls = T) = \frac{P(Burglary = T, JohnCalls = T, MaryCalls = T)}{P(JohnCalls = T, MaryCalls = T)}$$

$$= \frac{P(Burglary = T, \boldsymbol{Earthquake = T\ or\ F, Alarm = T\ or\ F}, JohnCalls = T, MaryCalls = T)}{P(\boldsymbol{Burglary = T\ or\ F, Earthquake = T\ or\ F, Alarm = T\ or\ F}, JohnCalls = T, MaryCalls = T)}$$

$$= \frac{P(+B,+E,+A,+J,+M) + P(+B,+E,-A,+J,+M) + \boldsymbol{P(+B,-E,+A,+J,+M)} + P(+B,-E,-A,+J,+M)}{P(+B,+E,+A,+J,+M) + P(+B,+E,-A,+J,+M) + P(+B,-E,+A,+J,+M) + P(+B,-E,-A,+J,+M) + P(-B,+E,+A,+J,+M) + P(-B,+E,-A,+J,+M) + P(-B,-E,+A,+J,+M) + P(-B,-E,-A,+J,+M)}$$

$$= \frac{\boldsymbol{p(+B)p(+E)p(+A|+B,+E)p(+J|+A)p(+M|+A) + p(+B)p(+E)p(-A|+B,+E)p(+J|-A)p(+M|-A) + p(+B)p(-E)p(+A|+B,-E)p(+J|+A)p(+M|+A) + p(+B)p(-E)p(-A|+B,-E)p(+J|-A)p(+M|-A)}}{\begin{gathered}\boldsymbol{p(+B)p(+E)p(+A|+B,+E)p(+J|+A)p(+M|+A) + p(+B)p(+E)p(-A|+B,+E)p(+J|-A)p(+M|-A) + p(+B)p(-E)p(+A|+B,-E)p(+J|+A)p(+M|+A) + p(+B)p(-E)p(-A|+B,-E)p(+J|-A)p(+M|-A)} \\ + p(-B)p(+E)p(+A|-B,+E)p(+J|+A)p(+M|+A) + p(-B)p(+E)p(-A|-B,+E)p(+J|-A)p(+M|-A) + p(-B)p(-E)p(+A|-B,-E)p(+J|+A)p(+M|+A) + p(-B)p(-E)p(-A|-B,-E)p(+J|-A)p(+M|-A)\end{gathered}}$$

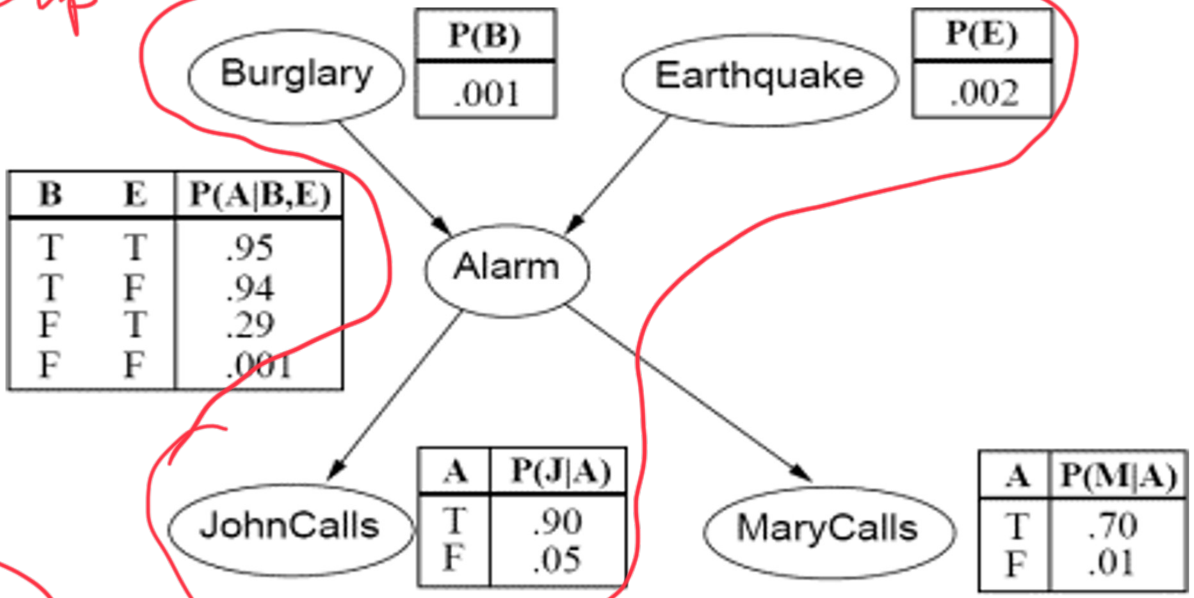
$$= \frac{\boldsymbol{(0.001)(0.002)(0.95)(0.90)(0.70) + (0.001)(0.002)(1-0.95)(0.05)(0.01) + (0.001)(1-0.002)(0.94)(0.90)(0.70) + (0.001)(1-0.002)(1-0.94)(0.05)(0.01)}}{\begin{gathered}\boldsymbol{(0.001)(0.002)(0.95)(0.90)(0.70) + (0.001)(0.002)(1-0.95)(0.05)(0.01) + (0.001)(1-0.002)(0.94)(0.90)(0.70) + (0.001)(1-0.002)(1-0.94)(0.05)(0.01)} \\ + (1-0.001)(0.002)(0.29)(0.90)(0.70) + (1-0.001)(0.002)(1-0.29)(0.05)(0.01) + (1-0.001)(1-0.002)(0.001)(0.90)(0.70) + (1-0.001)(1-0.002)(1-0.001)(0.05)(0.01)\end{gathered}}$$

$$= \frac{\boldsymbol{0.0000012 + 0.000000001 + 0.0005910156 + 0.00000002994}}{\boldsymbol{0.0000012 + 0.000000001 + 0.0005910156 + 0.00000002994} + 0.0003650346 + 0.00000070929 + 0.00062811126 + 0.000498002499} = \frac{0.00059224654}{0.002084104189} = \boldsymbol{0.284173}$$

Bayesian Networks

Focus on set up

Example: $P(+b\,|\,+j)$



Query: Burglary

Hidden: Alarm, Earthquake

Evidence: John calls

Is Mary part of calculation? No

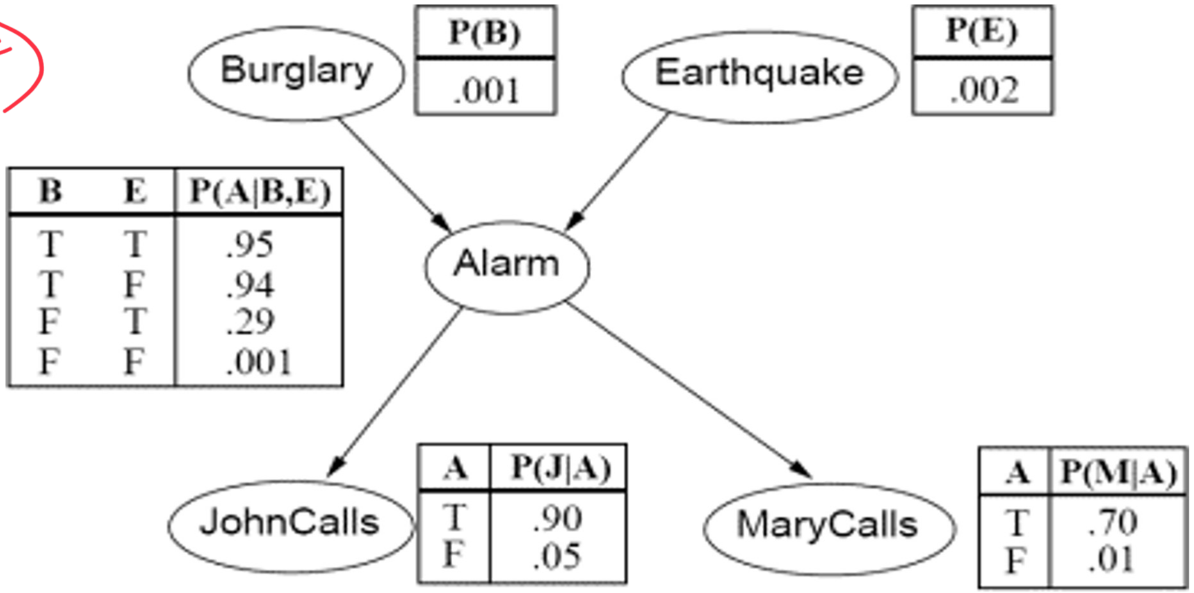
$$\frac{P(B,J)}{P(J)} = .016$$

$$= \frac{P(B)\sum_{e} P(E) \sum_{a} P(A|B,E)\,P(J|A)}{\sum_{b} P(B) \sum_{e} P(E) \sum_{a} P(A|B,E)\,P(J|A)}$$

Bayesian Networks

$$P(A) = \sum_{e} \sum_{b} P(A \mid B, E)\, P(B)\, P(E)$$

$$= .0025$$



P(B)
.001

P(E)
.002

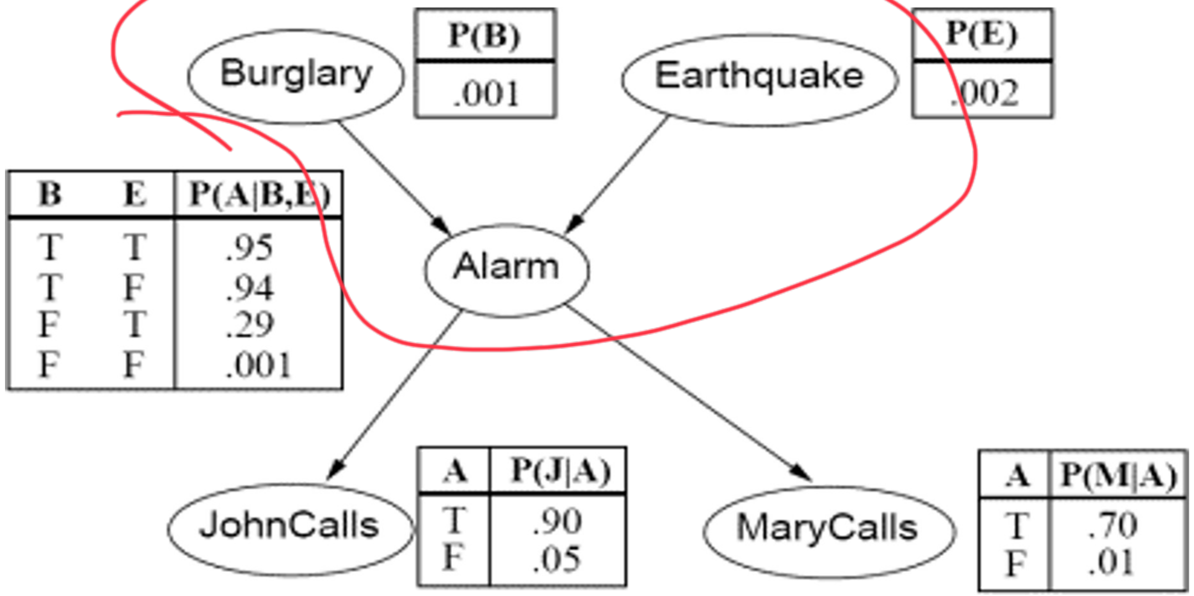
B	E	P(A|B,E)
T	T	.95
T	F	.94
F	T	.29
F	F	.001

A	P(J|A)
T	.90
F	.05

A	P(M|A)
T	.70
F	.01

Bayesian Networks

Example: $P(+b \mid +a)$



P(B)
.001

P(E)
.002

B	E	P(A|B,E)
T	T	.95
T	F	.94
F	T	.29
F	F	.001

A	P(J|A)
T	.90
F	.05

A	P(M|A)
T	.70
F	.01

Query: B

Evidence: A

Hidden: E

$$\frac{P(B,A)}{P(A)} = \frac{P(A|B)P(B)}{P(A)}$$

$$= \frac{P(B)\sum_e P(A|B,e)P(e)}{P(A)}$$

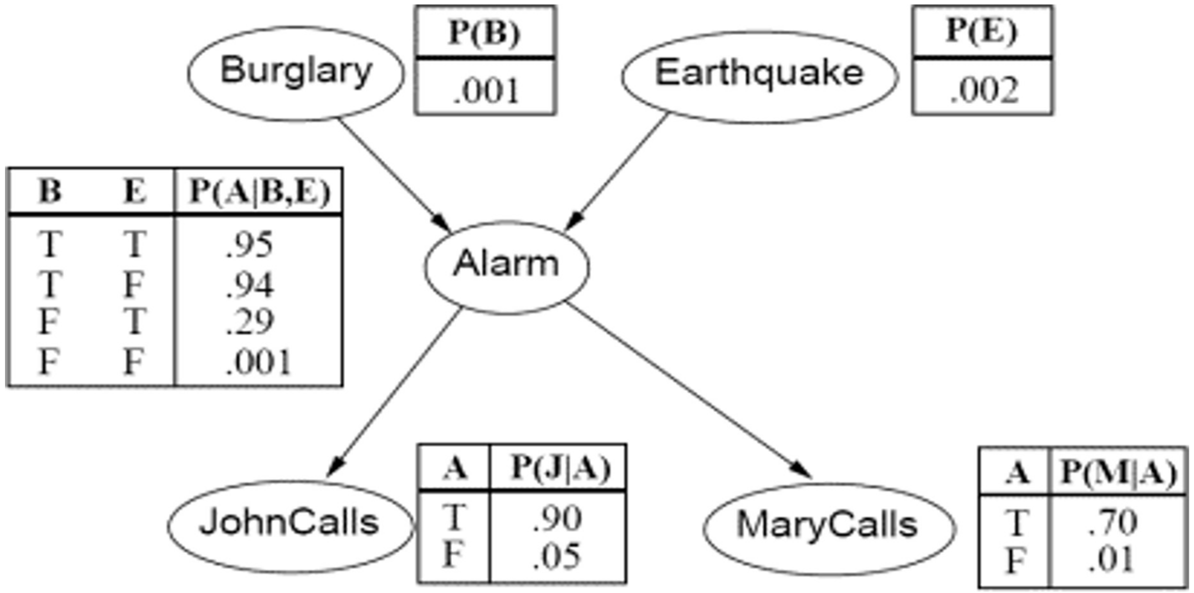
$$P(A) = \sum_b \sum_e P(A|b,e)P(b)P(e)$$

$$P(A) = .0025$$

$$= .001 \times .95 \times .002 + .001 \times .94 \times .998$$

$$= .0000019 + .000938 = \frac{.00094}{.0025} = .376$$

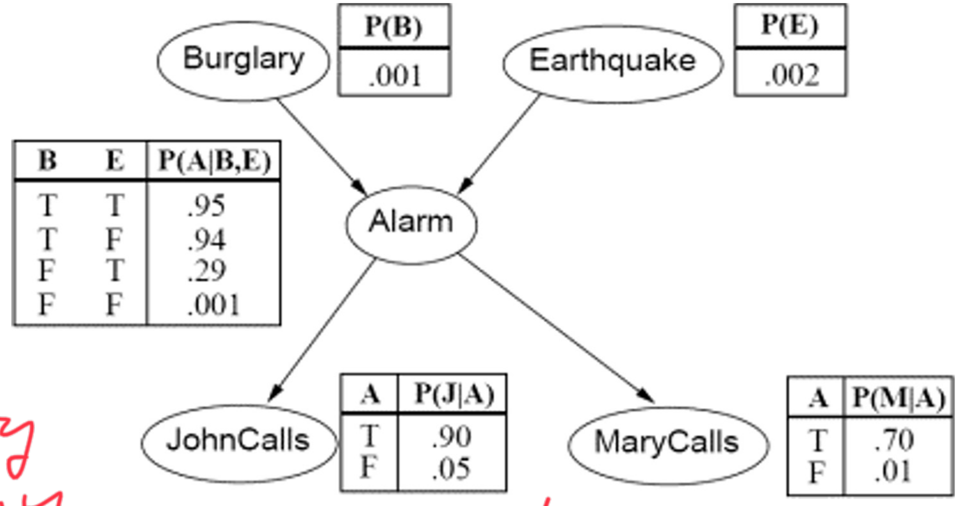
Bayesian Networks



Bayesian Networks

Example: Find $P(Earthquake = T \mid Burglary = T, Alarm = T)$
And $P(Earthquake = T \mid Alarm = T)$

Is it true that Burglary = T explains away Earthquake = T?



$P(E|A)$

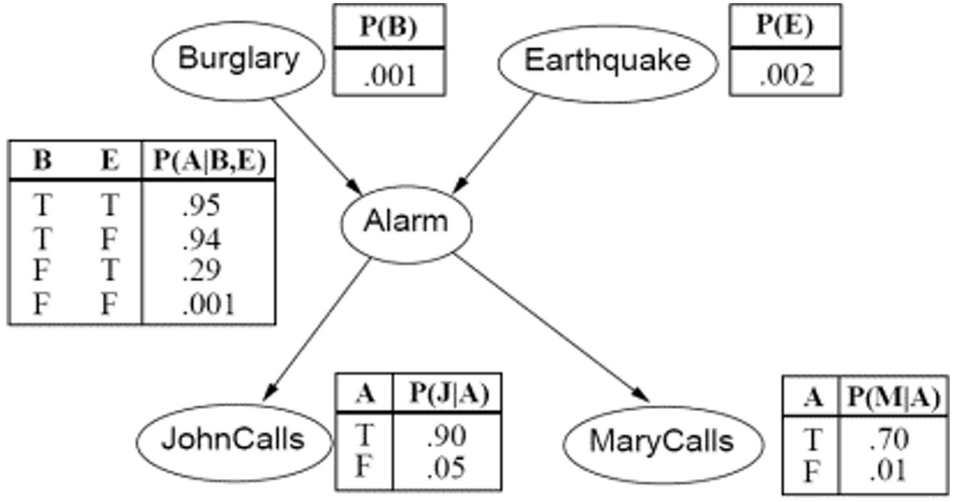
$P(E|A,B)$

Learning that there has been a burglary explains away earthquake as cause of alarm.

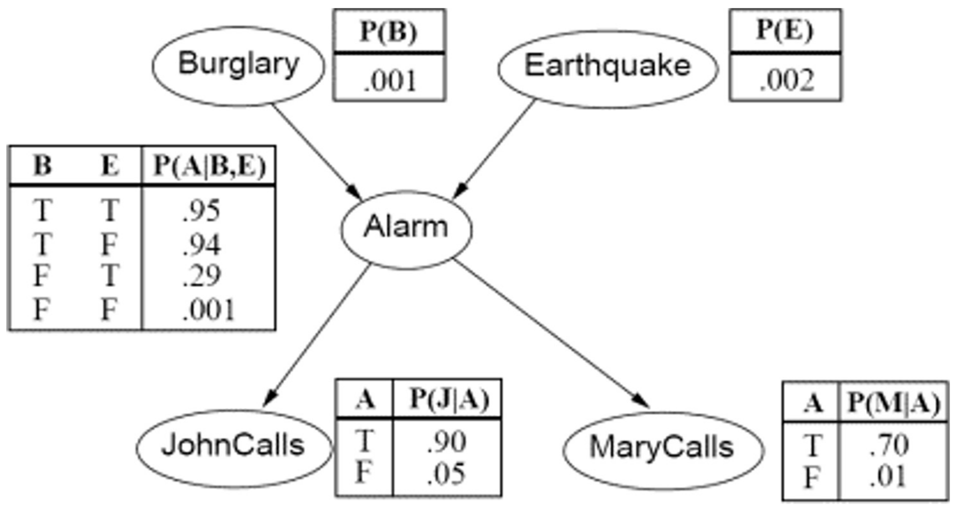
$$P(E|A) = \frac{P(A|E)P(E)}{P(A)} = .23 = \frac{\sum_b P(A|B,E)P(B)P(E)}{P(A)}$$

$$P(E|A,B) = \frac{P(A|B,E)P(B)P(E)}{\sum_e P(A|B,E)P(B)P(E)} = .021$$

Bayesian Networks



Bayesian Networks



Bayesian Networks – construction

Bayes nets implicitly encode joint distributions as a product of the local conditional distributions:

$$P(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid x_{i-1}, x_{i-2}, \ldots, x_2, x_1) = \ldots ?$$

Node ordering: write in such a way that

$$\text{parents}(X_i) \subseteq \{X_{i-1}, X_{i-2}, \ldots, X_2, X_1\}$$

$\rightarrow$

$$\prod_{i=1}^{n} P(x_i \mid x_{i-1}, x_{i-2}, \ldots, x_2, x_1) = \prod_{i=1}^{n} P(x_i \mid \text{parents}(X_i))$$

This statement:

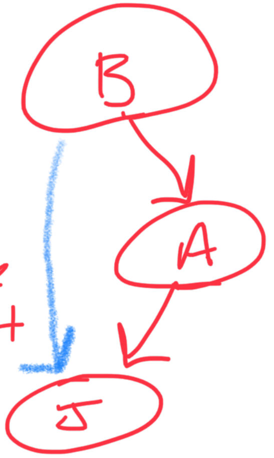
$$P(X_i \mid X_{i-1}, X_{i-2}, \ldots, X_2, X_1) = P(X_i \mid \text{parents}(X_i))$$

is key: **each node is conditionally independent**

of its other predecessors, given its parents

Blue edge is redundant

B would be parent and



Bayesian Networks: Construction

grandparent of J

Show a "flow" from cause to effect: **Pearl's Network Construction Algorithm**

Nodes: What is the set of variables we need to model?

Order them: $\{X_1, X_2, X_3, \ldots, X_n\}$

Best if ordered such that **causes precede effects**

Links: For each node X_i , do:

- Choose a minimal set of parents $\text{parents}(X_i) \subseteq \{X_{i-1}, X_{i-2}, \ldots, X_2, X_1\}$ such that $P(x_i \mid x_{i-1}, x_{i-2}, \ldots, x_1) = P(x_i \mid \text{parents}(X_i))$
- For each parent, insert arcs (links) from parent to X_i
- Write down CPT $P(X_i \mid \text{parents}(X_i))$

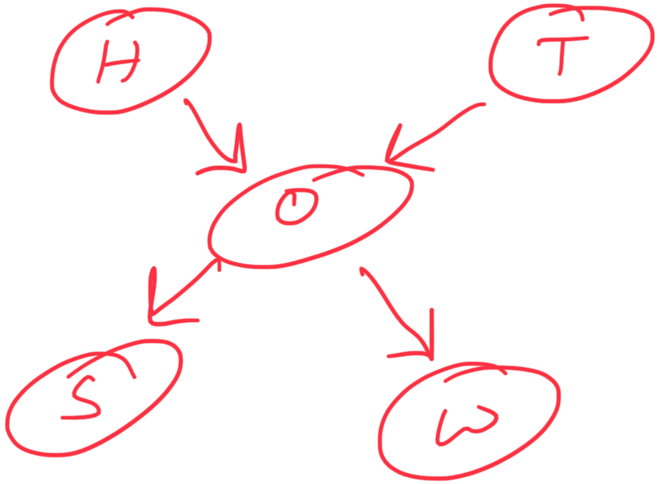
Bayesian Networks: Construction

Example: Suppose we have an old motorcycle that might either blow a head gasket (H) or have a broken thermometer (T). Either one would cause the bike to overheat (O). If the bike overheats, then it might blow smoke (S) and/or run weak (W).

Construct a Bayesian network for this situation.

x_1 x_2 x_3 4 5

1. Node ordering: {H, T, O, W, S}
2. Insert arcs



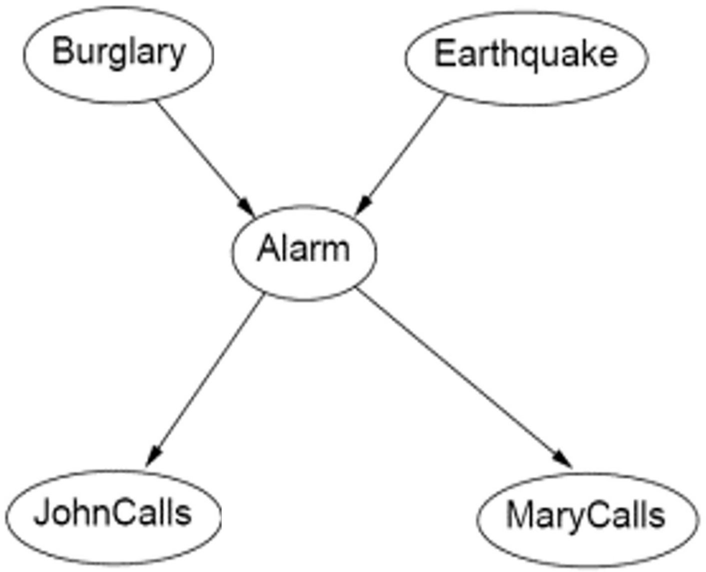
Bayesian Networks: Construction

Here, we chose to put the causes before effects:

{Burglary, Earthquake, Alarm, JohnCalls, MaryCalls}

What if instead we did the following?

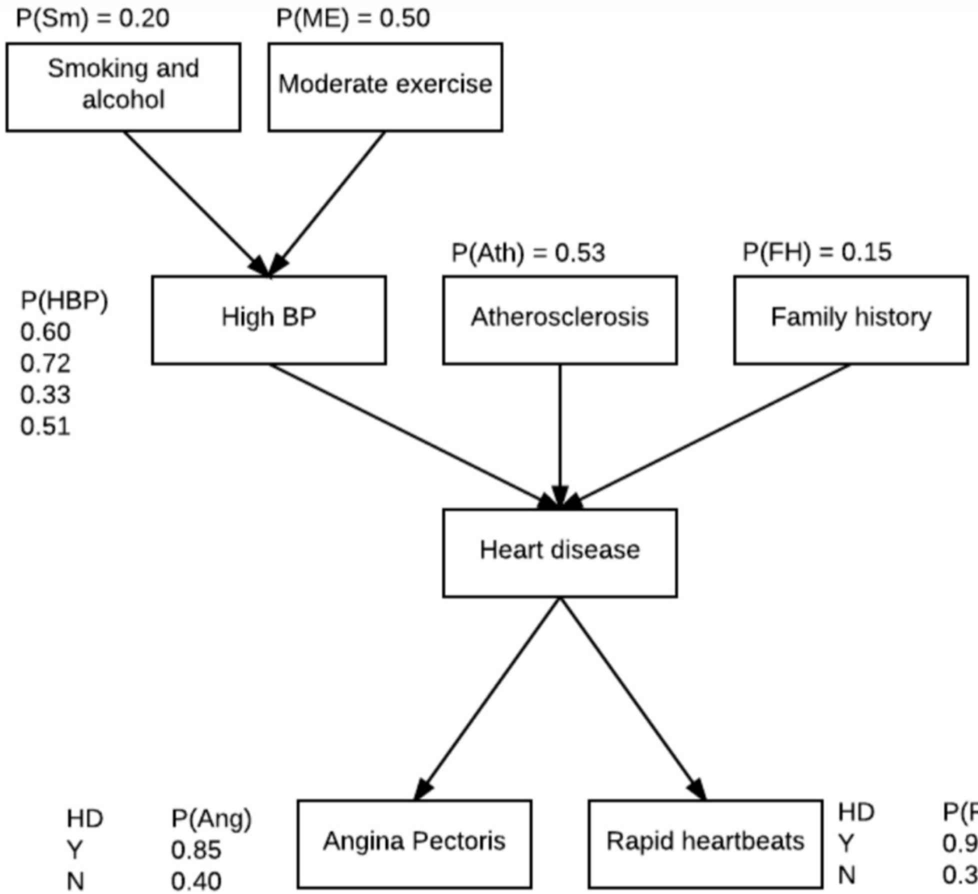
{MaryCalls, JohnCalls, Alarm,
Burglary, Earthquake}



Bayesian Networks: Types of reasoning – diagnostic

We'll start here on Wednesday

The following Bayesian network is based loosely on a study that examined heart disease risk factors in 167 elderly individuals in South Carolina. Note that this figure uses Y and N to represent Yes and No, whereas in class we used the equivalent T and F to represent True and False Boolean values.



SM	ME	P(HBP)
Y	Y	0.60
Y	N	0.72
N	Y	0.33
N	N	0.51

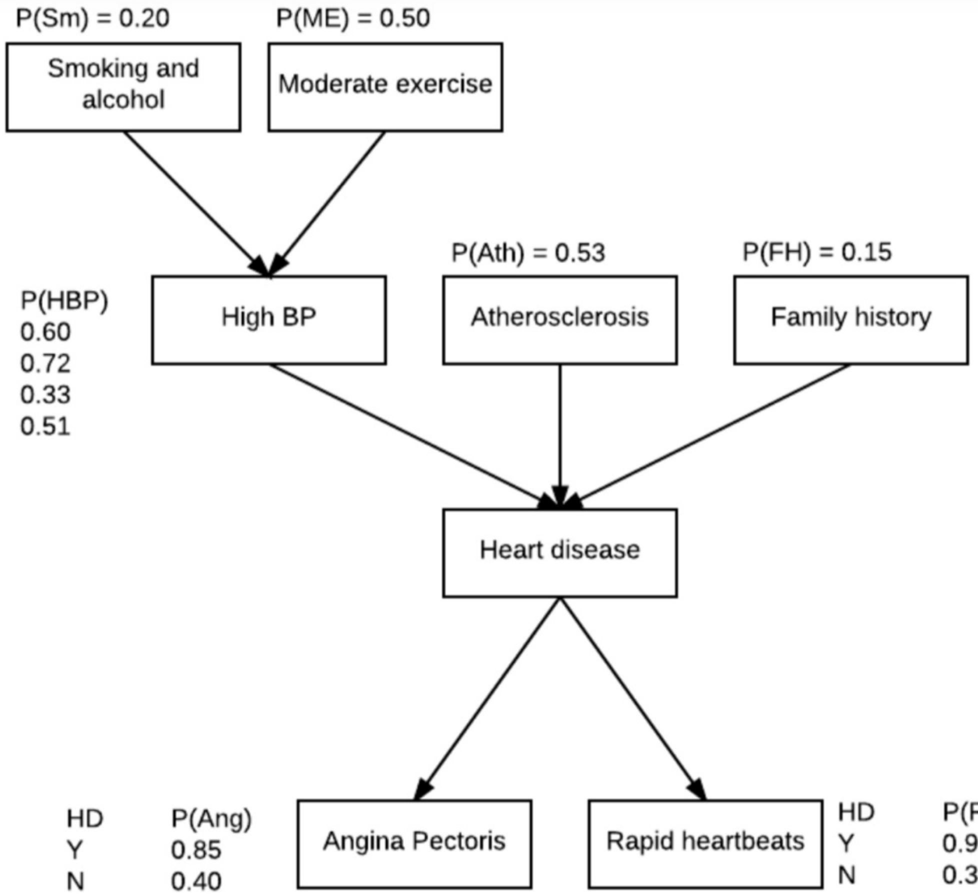
Ath	HBP	FH	P(HD)
Y	Y	Y	0.92
Y	Y	N	0.91
Y	N	Y	0.81
Y	N	N	0.77
N	Y	Y	0.75
N	Y	N	0.69
N	N	Y	0.38
N	N	N	0.23

HD	P(Ang)
Y	0.85
N	0.40

HD	P(Rapid)
Y	0.99
N	0.30

Bayesian Networks: Types of reasoning – predictive

The following Bayesian network is based loosely on a study that examined heart disease risk factors in 167 elderly individuals in South Carolina. Note that this figure uses Y and N to represent Yes and No, whereas in class we used the equivalent T and F to represent True and False Boolean values.



SM	ME	P(HBP)
Y	Y	0.60
Y	N	0.72
N	Y	0.33
N	N	0.51

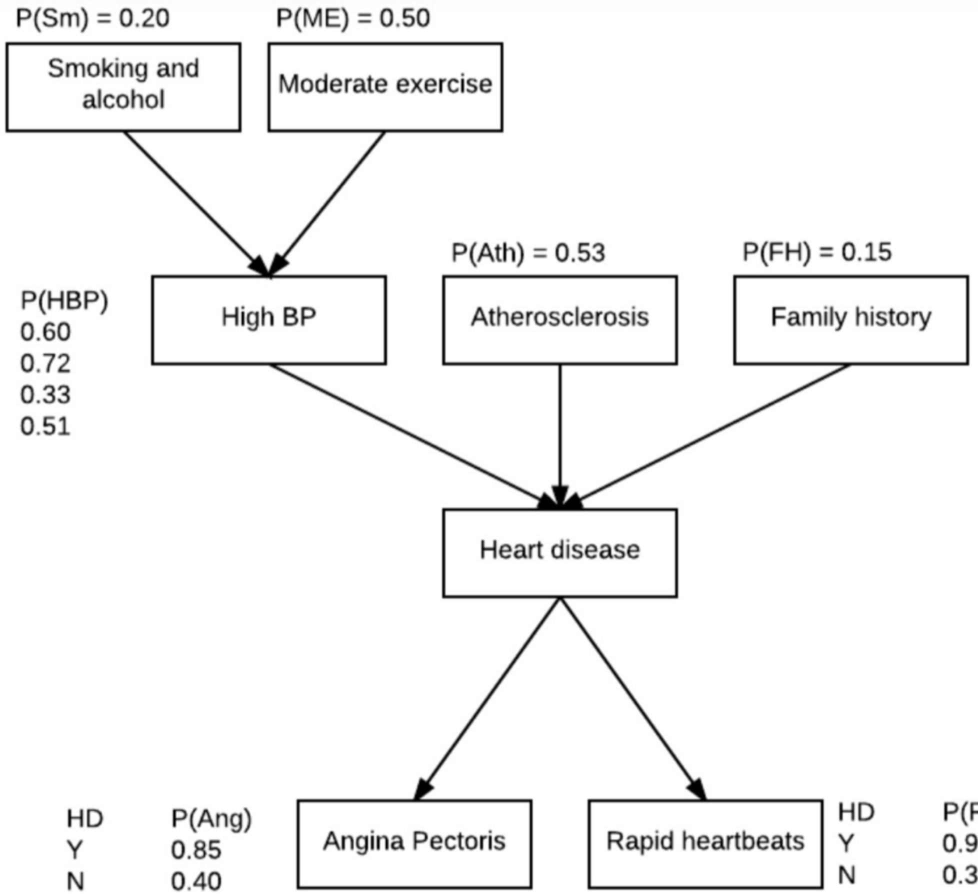
Ath	HBP	FH	P(HD)
Y	Y	Y	0.92
Y	Y	N	0.91
Y	N	Y	0.81
Y	N	N	0.77
N	Y	Y	0.75
N	Y	N	0.69
N	N	Y	0.38
N	N	N	0.23

HD	P(Ang)
Y	0.85
N	0.40

HD	P(Rapid)
Y	0.99
N	0.30

Bayesian Networks: Types of reasoning – intercausal

The following Bayesian network is based loosely on a study that examined heart disease risk factors in 167 elderly individuals in South Carolina. Note that this figure uses Y and N to represent Yes and No, whereas in class we used the equivalent T and F to represent True and False Boolean values.



SM	ME	P(HBP)
Y	Y	0.60
Y	N	0.72
N	Y	0.33
N	N	0.51

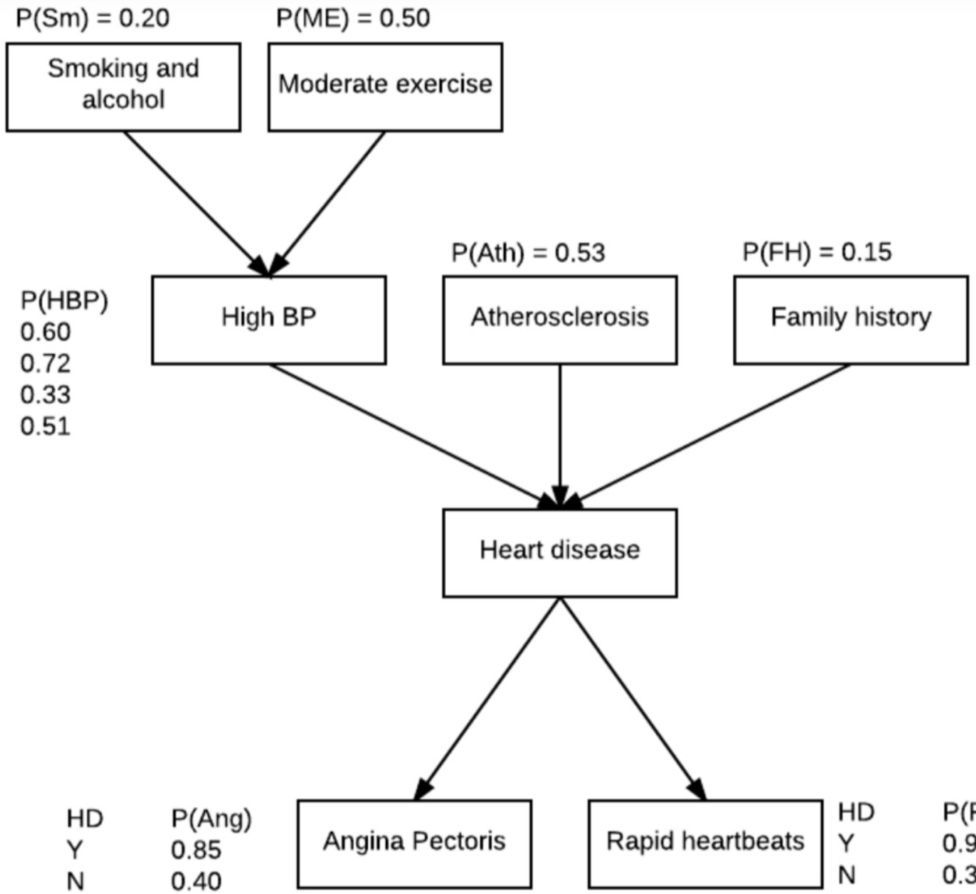
Ath	HBP	FH	P(HD)
Y	Y	Y	0.92
Y	Y	N	0.91
Y	N	Y	0.81
Y	N	N	0.77
N	Y	Y	0.75
N	Y	N	0.69
N	N	Y	0.38
N	N	N	0.23

HD	P(Ang)
Y	0.85
N	0.40

HD	P(Rapid)
Y	0.99
N	0.30

Bayesian Networks: Types of reasoning – combined

The following Bayesian network is based loosely on a study that examined heart disease risk factors in 167 elderly individuals in South Carolina. Note that this figure uses Y and N to represent Yes and No, whereas in class we used the equivalent T and F to represent True and False Boolean values.



SM	ME	P(HBP)
Y	Y	0.60
Y	N	0.72
N	Y	0.33
N	N	0.51

Ath	HBP	FH	P(HD)
Y	Y	Y	0.92
Y	Y	N	0.91
Y	N	Y	0.81
Y	N	N	0.77
N	Y	Y	0.75
N	Y	N	0.69
N	N	Y	0.38
N	N	N	0.23

HD	P(Ang)
Y	0.85
N	0.40

HD	P(Rapid)
Y	0.99
N	0.30