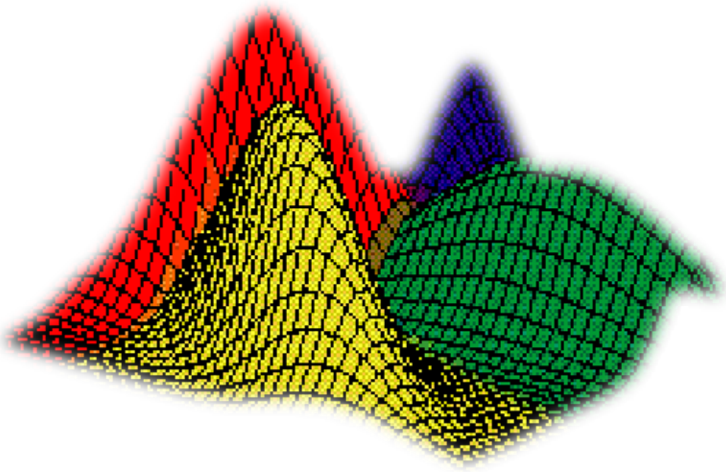


Bayesian Network

Graphical Models

- Key Idea:
  - Conditional independence assumptions useful
  - but Naïve Bayes is extreme!
  - Graphical models express sets of conditional independence assumptions via graph structure
  - Graph structure plus associated parameters define *joint probability distribution over set of variables*

Graphical Models – Why Care?

- Among most important ML developments of the decade

- Graphical models allow combining:
  - Prior knowledge in form of dependencies/independencies
  - Prior knowledge in form of priors over parameters
  - Observed training data

- Principled and ~general methods for
  - Probabilistic inference
  - Learning

- Useful in practice
  - Diagnosis, help systems, text analysis, time series models, ...

Conditional Independence

Definition: X is conditionally independent of Y given Z, if the probability distribution governing X is independent of the value of Y, given the value of Z

$$(\forall i, j, k) P(X = x_i | Y = y_j, Z = z_k) = P(X = x_i | Z = z_k)$$

Which we often write $P(X|Y, Z) = P(X|Z)$

E.g., $P(Thunder|Rain, Lightning) = P(Thunder|Lightning)$

Marginal Independence

Definition: X is marginally independent of Y if

$$(\forall i, j) P(X = x_i, Y = y_j) = P(X = x_i) P(Y = y_j)$$

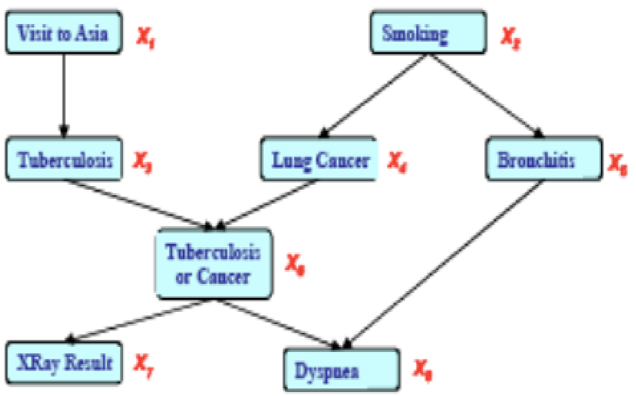
Equivalently, if

$$(\forall i, j) P(X = x_i | Y = y_j) = P(X = x_i)$$

Equivalently, if

$$(\forall i, j) P(Y = y_i | X = x_j) = P(Y = y_i)$$

Bayes Nets define Joint Probability Distribution in terms of this graph, plus parameters



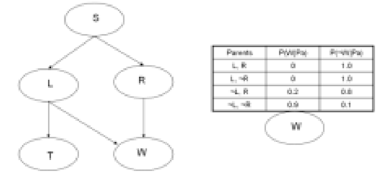
$$P(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8)$$

$$= P(X_1)\, P(X_2)\, P(X_3 | X_1)\, P(X_4 | X_2)\, P(X_5 | X_2)\, P(X_6 | X_3, X_4)\, P(X_7 | X_6)\, P(X_8 | X_5, X_6)$$

Benefits of Bayes Nets:

- Represent the full joint distribution in fewer parameters, using prior knowledge about dependencies
- Algorithms for inference and learning

Bayesian Networks Definition



Parents	P(W|Pa)	P(¬W|Pa)
L, R	0	1.0
L, ¬R	0	1.0
¬L, R	0.2	0.8
¬L, ¬R	0.9	0.1

A Bayes network represents the joint probability distribution over a collection of random variables

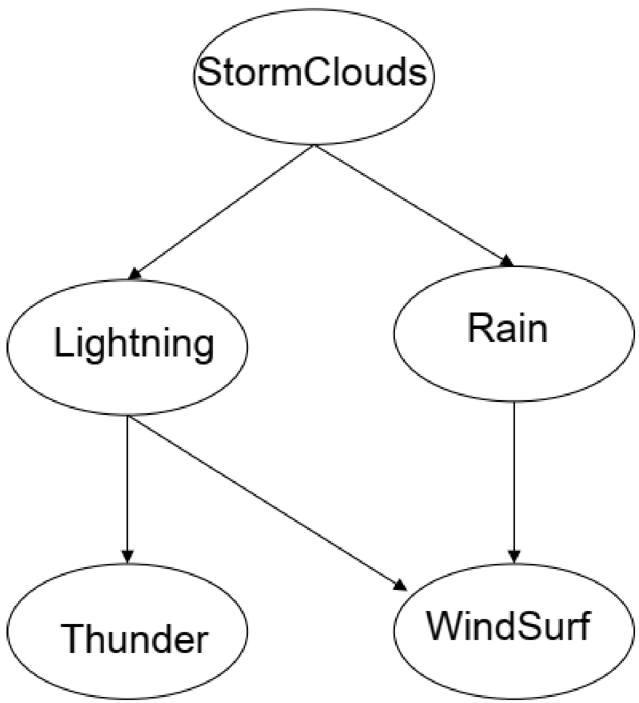
A Bayes network is a directed acyclic graph and a set of conditional probability distributions (CPD's)

- Each node denotes a random variable
- Edges denote dependencies
- For each node X_i its CPD defines $P(X_i \mid Pa(X_i))$
- The joint distribution over all variables is defined to be

$$P(X_1 \ldots X_n) = \prod_i P(X_i | Pa(X_i))$$

Pa(X) = immediate parents of X in the graph

Bayesian Network



Nodes = random variables

A conditional probability distribution (CPD) is associated with each node N, defining P(N | Parents(N))

Parents	P(W|Pa)	P(¬W|Pa)
L, R	0	1.0
L, ¬R	0	1.0
¬L, R	0.2	0.8
¬L, ¬R	0.9	0.1

The joint distribution over all variables:

$$P(X_1 \dots X_n) = \prod_i P(X_i | Pa(X_i))$$

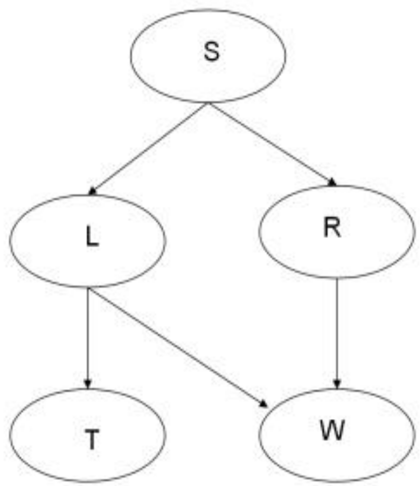
Some helpful terminology

Parents = Pa(X) = immediate parents

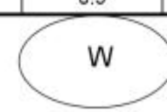
Antecedents = parents, parents of parents, ...

Children = immediate children

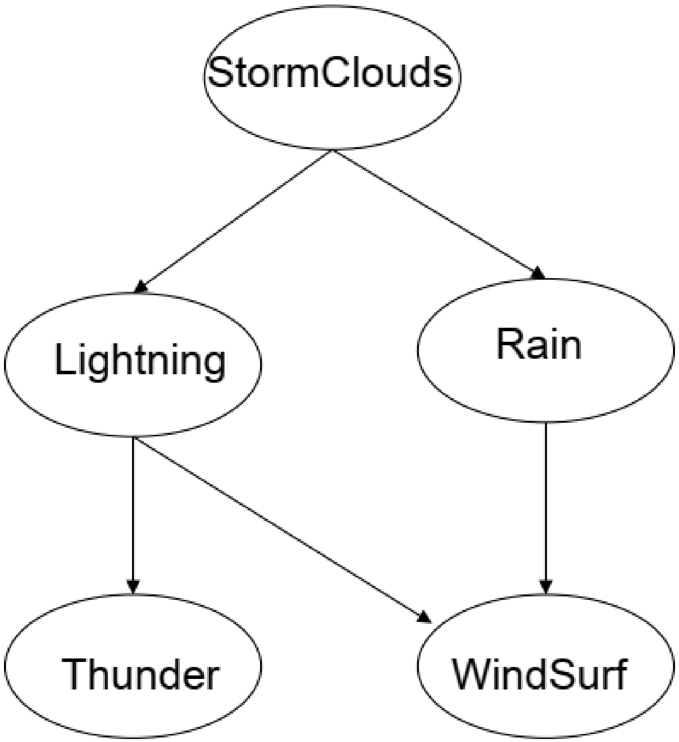
Descendents = children, children of children, ...



Parents	P(W|Pa)	P(¬W|Pa)
L, R	0	1.0
L, ¬R	0	1.0
¬L, R	0.2	0.8
¬L, ¬R	0.9	0.1



Bayesian Network



What can we say about conditional independencies in a Bayes Net?

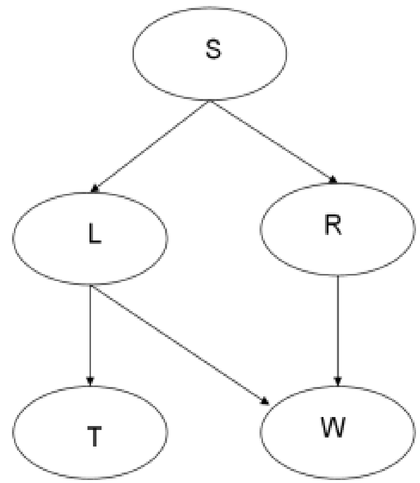
One thing is this:

Each node is conditionally independent of its non-descendents, given only its immediate parents.

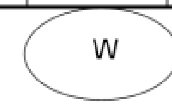
Parents	P(W|Pa)	P(¬W|Pa)
L, R	0	1.0
L, ¬R	0	1.0
¬L, R	0.2	0.8
¬L, ¬R	0.9	0.1

Bayesian Networks

- CPD for each node X_i describes $P(X_i \mid Pa(X_i))$



Parents	P(W|Pa)	P(¬W|Pa)
L, R	0	1.0
L, ¬R	0	1.0
¬L, R	0.2	0.8
¬L, ¬R	0.9	0.1



Chain rule of probability says that in general:

$$P(S,L,R,T,W) = P(S)P(L|S)P(R|S,\cancel{L})\underbrace{P(T|\cancel{S},L,\cancel{R})}_{8 \text{ params}}P(W|\cancel{S},L,R,\cancel{T})$$

But in a Bayes net: $P(X_1 \ldots X_n) = \prod_i P(X_i|Pa(X_i))$

$$P(S,L,R,T,W) = P(S)\,P(L|S)\,P(R|S)\,\underbrace{P(T|L)}_{2}\,P(W|L\,R)$$

Bayesian networks

- a BN consists of a Directed Acyclic Graph (DAG) and a set of conditional probability distributions

- in the DAG
  - each node denotes random a variable
  - each edge from X to Y represents that X *directly influences* Y
  - formally: each variable X is independent of its non-descendants given its parents

- each node X has a *conditional probability distribution* (CPD) representing $P(X \mid Parents(X))$

Bayesian networks

- using the chain rule, a joint probability distribution can be expressed as

$$P(X_1,\ \ldots,\ X_n) = \mathrm{P}(X_1)\prod_{i=2}^{n} P(X_i \mid X_1,\ \ldots,\ X_{i-1}))$$

- a BN provides a compact representation of a joint probability distribution

$$P(X_1,\ \ldots,\ X_n) = \prod_{i=1}^{n} P(X_i \mid Parents(X_i))$$

Bayesian network example

- Consider the following 5 binary random variables:

  B = a burglary occurs at your house

  E = an earthquake occurs at your house

  A = the alarm goes off

  J = John calls to report the alarm

  M = Mary calls to report the alarm

- Suppose we want to answer queries like what is $P(B \mid M, J)$?

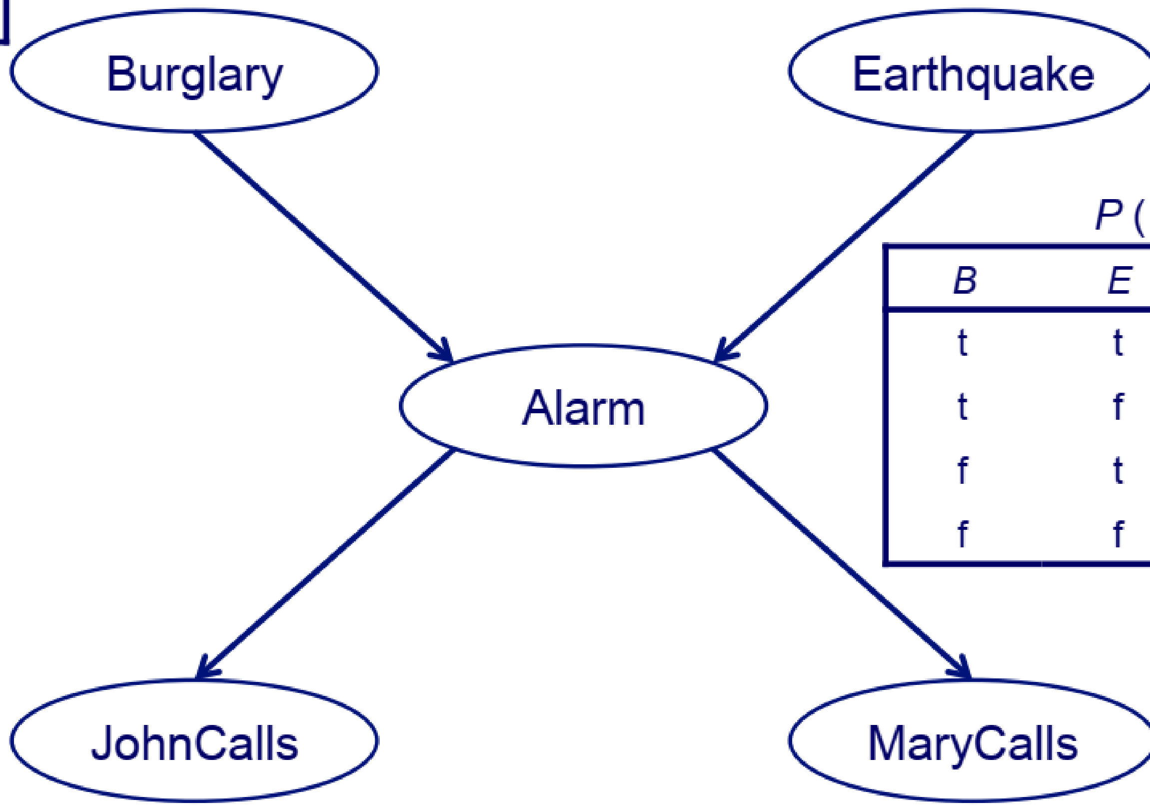
Bayesian network example

P (B)

t	f
0.001	0.999

P (E)

t	f
0.001	0.999



P (A | B, E)

B	*E*	t	f
t	t	0.95	0.05
t	f	0.94	0.06
f	t	0.29	0.71
f	f	0.001	0.999

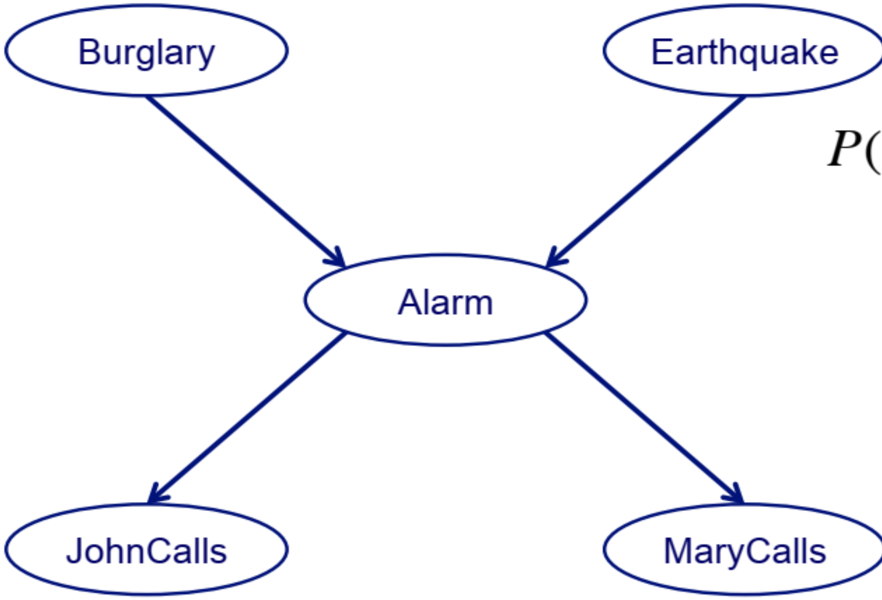
P (J | A)

A	t	f
t	0.9	0.1
f	0.05	0.95

P (M | A)

A	t	f
t	0.7	0.3
f	0.01	0.99

Bayesian networks



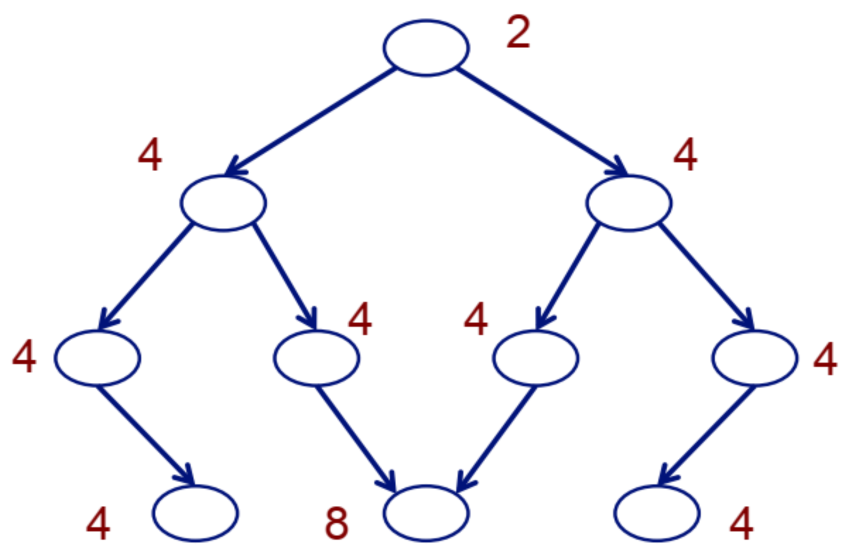
$$
\begin{aligned}
P(B,E,A,J,M) = \; & P(B) \times \\
& P(E) \times \\
& P(A \mid B,E) \times \\
& P(J \mid A) \times \\
& P(M \mid A)
\end{aligned}
$$

- a standard representation of the joint distribution for the Alarm example has $2^5 = 32$ parameters
- the BN representation of this distribution has 20 parameters

Bayesian networks

- consider a case with 10 binary random variables

- How many parameters does a BN with the following graph structure have?



= 42

- How many parameters does the standard table representation of the joint distribution have? = 1024

Advantages of the Bayesian network representation

- Captures independence and conditional independence where they exist
- Encodes the relevant portion of the full joint among variables where dependencies exist
- Uses a graphical representation which lends insight into the complexity of inference

Bayesian Belief Networks

- Let we *l* random variables
- The joint probability is given by,

$$p(x_1, x_2, \ldots, x_\ell) = p(x_\ell \mid x_{\ell-1}, \ldots, x_1) \cdot p(x_{\ell-1} \mid x_{\ell-2}, \ldots, x_1) \cdot \ldots$$
$$\ldots \cdot p(x_2 \mid x_1) \cdot p(x_1)$$

Bayesian Belief Networks

The formula

$$p(x_1, x_2, \ldots, x_\ell) = p(x_\ell \mid x_{\ell-1}, \ldots, x_1) \cdot p(x_{\ell-1} \mid x_{\ell-2}, \ldots, x_1) \cdot \ldots$$
$$\ldots \cdot p(x_2 \mid x_1) \cdot p(x_1)$$

can be written as

$$p(x_1, x_2, \ldots, x_\ell) = p(x_1) \cdot \prod_{i=2}^{\ell} p(x_i \mid A_i)$$

where

$$A_i \subseteq \{x_{i-1}, x_{i-2}, \ldots, x_1\}$$

– For example, if ℓ=6, then we could assume:

$$p(x_6 \mid x_5, \dots, x_1) = p(x_6 \mid x_5, x_4)$$

Then:

$$A_6 = \{x_5, x_4\} \subseteq \{x_5, \dots, x_1\}$$

– Similalrly, if we assume

$$p(x_5|x_4, \ldots, x_1) = p(x_5|x_4)$$

$$p(x_4|x_3, x_2, x_1) = p(x_4|x_2, x_1)$$

$$p(x_3|x_2, x_1) = p(x_3|x_2)$$

$$p(x_2|x_1) = p(x_2)$$

Then:

$$A_5 = \{x_4\},\ A_4 = \{x_2, x_1\},\ A_3 = \{x_2\}, A_2 = \emptyset$$

– Simialrly, if we assume

$$p(x_5|x_4, \ldots, x_1) = p(x_5|x_4)$$

$$p(x_4|x_3, x_2, x_1) = p(x_4|x_2, x_1)$$

$$p(x_3|x_2, x_1) = p(x_3|x_2)$$

$$p(x_2|x_1) = p(x_2)$$

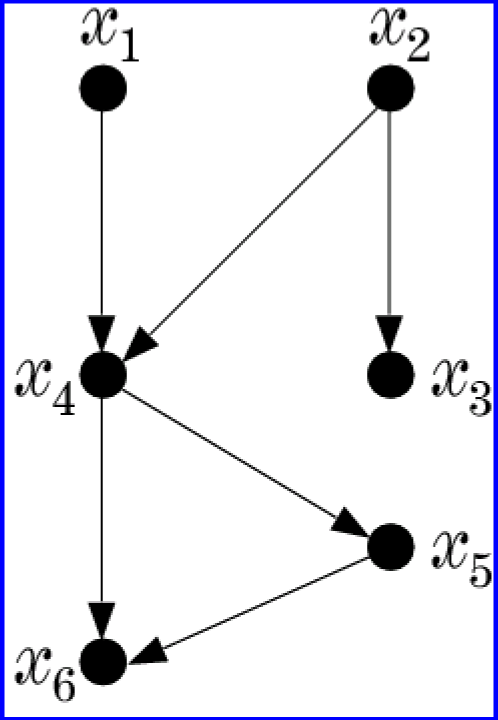
Then:

$$A_5 = \{x_4\},\ A_4 = \{x_2, x_1\},\ A_3 = \{x_2\}, A_2 = \emptyset$$

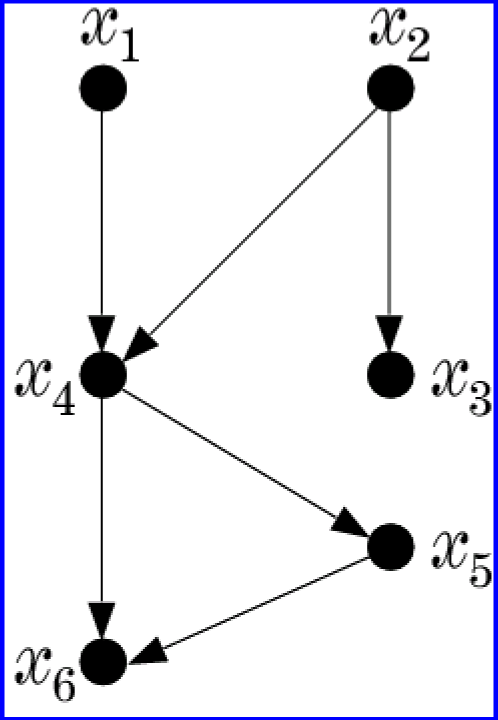
– The above is a generalization of the Naïve – Bayes. For the Naïve – Bayes the assumption is:

$$A_i = \emptyset, \text{ for } i=1, 2, \ldots, \ell$$

– A graphical way to portray conditional dependencies

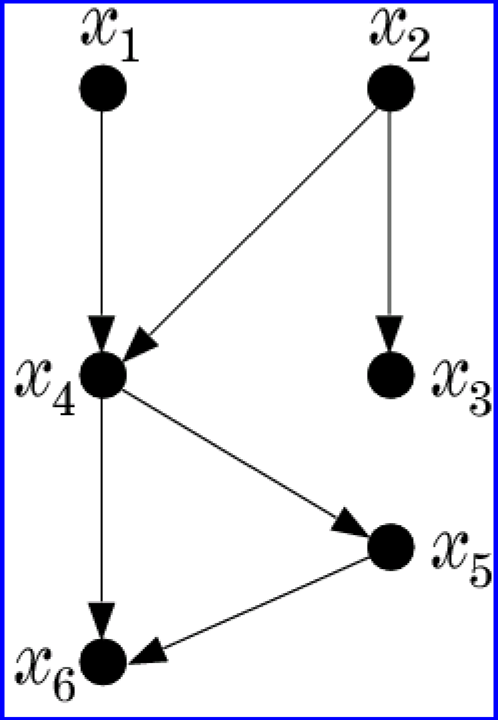


– A graphical way to portray conditional dependencies



- According to this figure we have that:
  - x_6 is conditionally dependent on x_4, x_5.
  - x_5 on x_4
  - x_4 on x_1, x_2
  - x_3 on x_2
  - x_1, x_2 are conditionally independent on other variables.

– A graphical way to portray conditional dependencies



- According to this figure we have that:
  - x_6 is conditionally dependent on x_4, x_5.
  - x_5 on x_4
  - x_4 on x_1, x_2
  - x_3 on x_2
  - x_1, x_2 are conditionally independent on other variables.

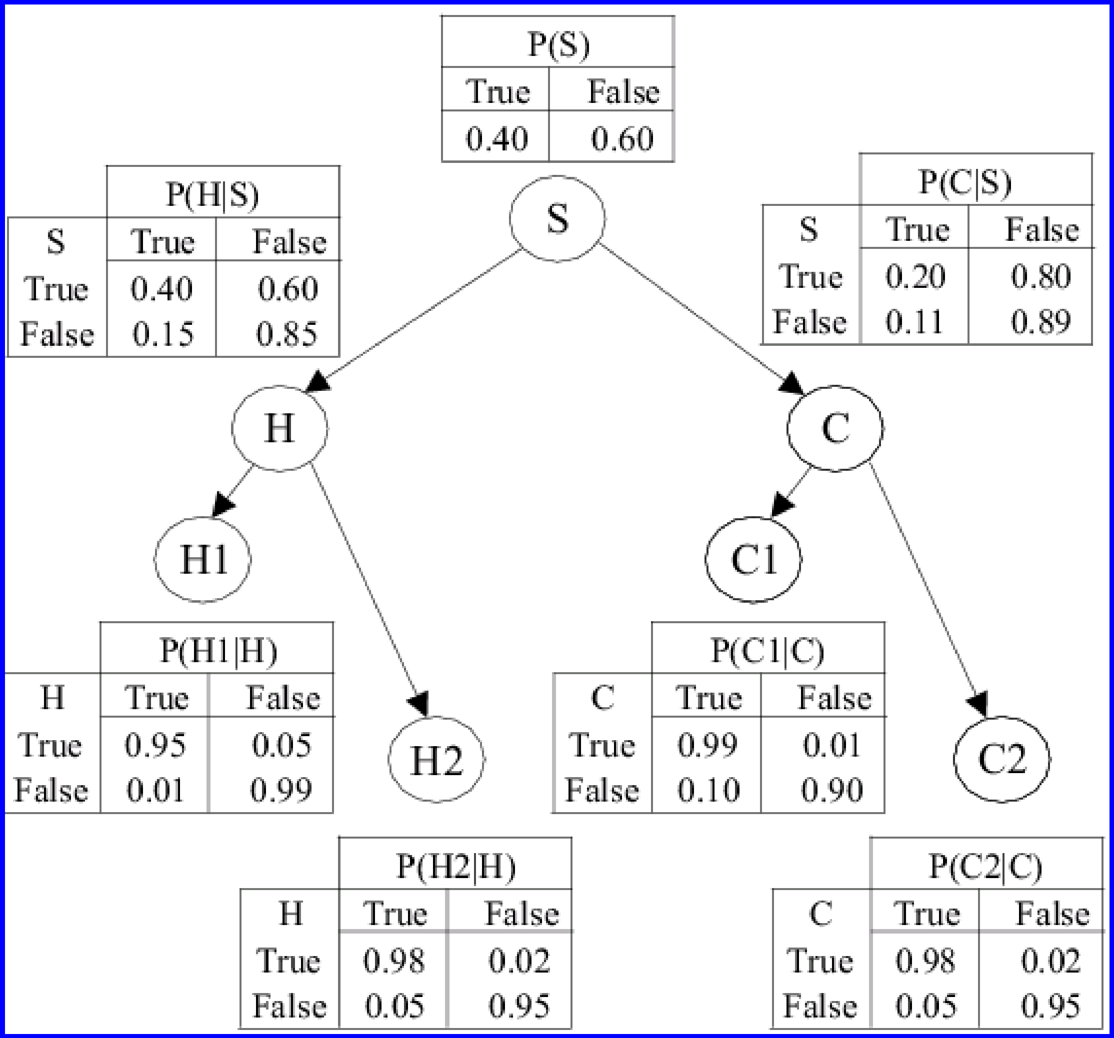
- For this case:

$$p(x_1, x_2, \ldots, x_6) = p(x_6 \mid x_5, x_4) \cdot p(x_5 \mid x_4) \cdot p(x_3 \mid x_2) \cdot p(x_2) \cdot p(x_1)$$

- Bayesian Networks
  - a directed acyclic graph (DAG)
  - the nodes correspond to random variables
  - arc represents parent-child (*dependence*) relationship

- A Bayesian Network is specified by:
  - The prior probabilities of its root nodes.
  - The conditional probabilities of the non-root nodes, given their parents, for ALL possible combinations.

– A Bayesian Network from a medical application



P(S)	
True	False
0.40	0.60

	P(H|S)	
S	True	False
True	0.40	0.60
False	0.15	0.85

	P(C|S)	
S	True	False
True	0.20	0.80
False	0.11	0.89

	P(H1|H)	
H	True	False
True	0.95	0.05
False	0.01	0.99

	P(C1|C)	
C	True	False
True	0.99	0.01
False	0.10	0.90

	P(H2|H)	
H	True	False
True	0.98	0.02
False	0.05	0.95

	P(C2|C)	
C	True	False
True	0.98	0.02
False	0.05	0.95

- BBN models conditional dependencies concerning smokers (S), tendencies to develop cancer (C) and heart disease (H), together with variables corresponding to heart (H1, H2) and cancer (C1, C2) medical tests.

- any joint probability can be obtained by multiplying the prior (root nodes) and the conditional (non-root nodes) probabilities.

- Training: given a topology, probabilities are estimated from training data. There are also methods that learn the topology.

- Probability Inference: Given an pattern (evidence), the goal is to compute the conditional probabilities for some of the other variables (class)

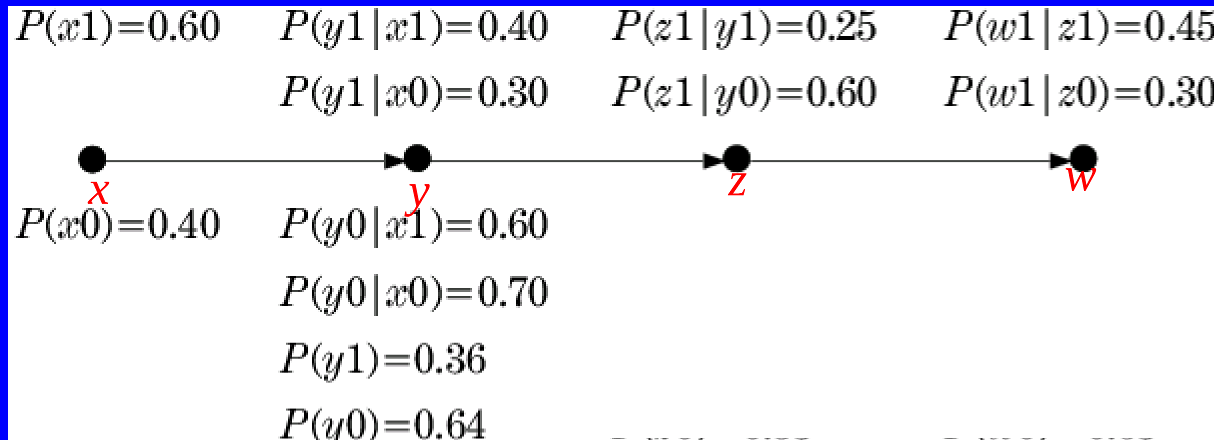
- Example: Consider the Bayesian network of the figure:

$P(x1) = 0.60 \quad P(y1|x1) = 0.40 \quad P(z1|y1) = 0.25 \quad P(w1|z1) = 0.45$

$P(y1|x0) = 0.30 \quad P(z1|y0) = 0.60 \quad P(w1|z0) = 0.30$

- Random variables: *x*, *y*, *w*, *z*
- *x*0 means *x* = 0
- *x*1 means *x* = 1

- We can calculate the other probabilities

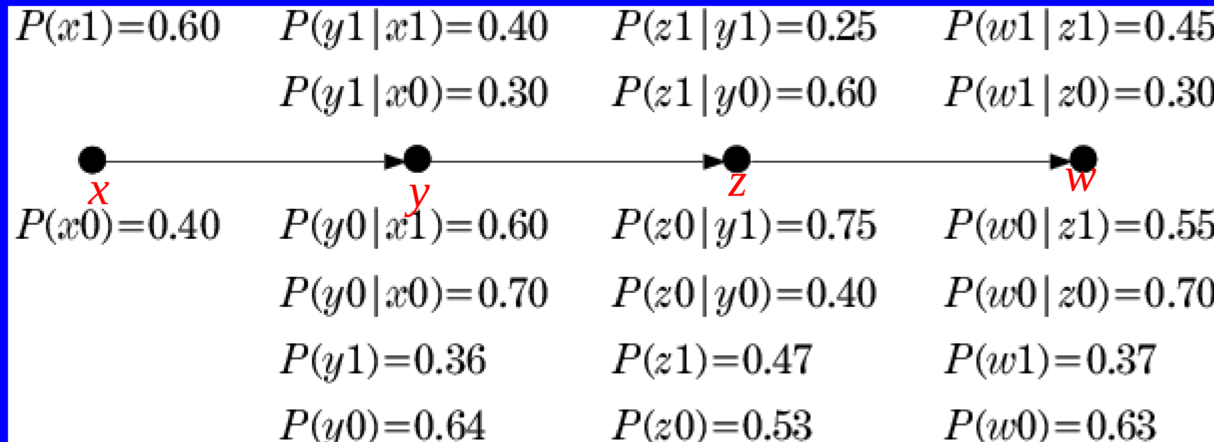


Example: $p(y_1)$:

$$P(y1) = \sum_x P(y1, x) = P(y1, x1) + P(y1, x0)$$

$$P(y1) = P(y1|x1)P(x1) + P(y1|x0)P(x0) = (0.4)(0.6) + (0.3)(0.4) = 0.36$$

- We can calculate the other probabilities



- Given this info, we can answer any probabilistic query:

$P(x1)=0.60$	$P(y1\mid x1)=0.40$	$P(z1\mid y1)=0.25$	$P(w1\mid z1)=0.45$
	$P(y1\mid x0)=0.30$	$P(z1\mid y0)=0.60$	$P(w1\mid z0)=0.30$

$P(x0)=0.40$	$P(y0\mid x1)=0.60$	$P(z0\mid y1)=0.75$	$P(w0\mid z1)=0.55$
	$P(y0\mid x0)=0.70$	$P(z0\mid y0)=0.40$	$P(w0\mid z0)=0.70$
	$P(y1)=0.36$	$P(z1)=0.47$	$P(w1)=0.37$
	$P(y0)=0.64$	$P(z0)=0.53$	$P(w0)=0.63$

a) If x is measured to be $x=1$ ($x1$), compute $P(z1|x1)$ *and* $P(w0|x1)$.

b) If w is measured to be $w=1$ ($w1$) compute $P(z1|w1)$].

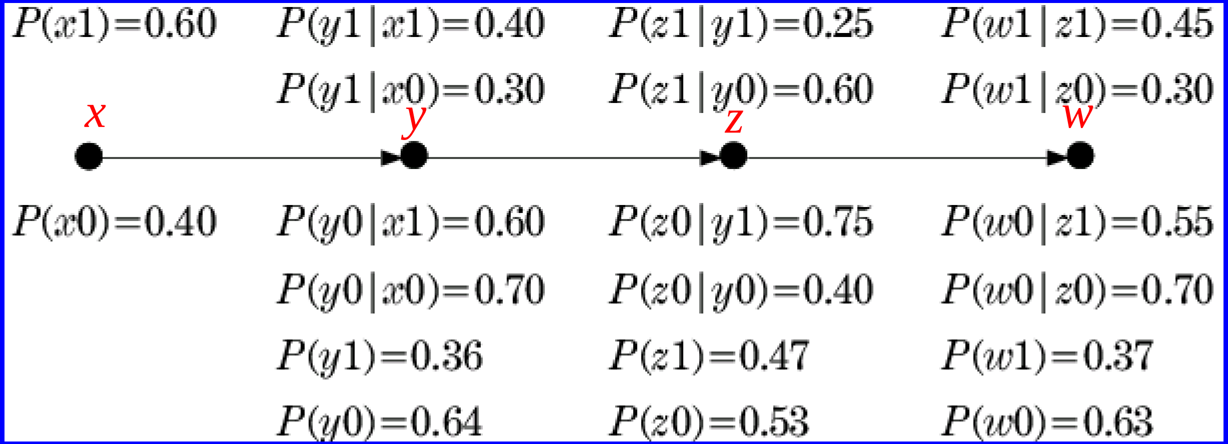
a) If x is measured to be $x=1$ ($x1$), compute P(z1|x1) and $P(w0|x1)$.

$P(x1)=0.60$	$P(y1\|x1)=0.40$	$P(z1\|y1)=0.25$	$P(w1\|z1)=0.45$
	$P(y1\|x0)=0.30$	$P(z1\|y0)=0.60$	$P(w1\|z0)=0.30$

$P(x0)=0.40$	$P(y0\|x1)=0.60$	$P(z0\|y1)=0.75$	$P(w0\|z1)=0.55$
	$P(y0\|x0)=0.70$	$P(z0\|y0)=0.40$	$P(w0\|z0)=0.70$
	$P(y1)=0.36$	$P(z1)=0.47$	$P(w1)=0.37$
	$P(y0)=0.64$	$P(z0)=0.53$	$P(w0)=0.63$

$$\begin{aligned} P(z1|x1) &= P(z1|y1, x1)P(y1|x1) + P(z1|y0, x1)P(y0|x1) \\ &= P(z1|y1)P(y1|x1) + P(z1|y0)P(y0|x1) \\ &= (0.25)(0.4) + (0.6)(0.6) = 0.46 \end{aligned}$$

a) If x is measured to be $x=1$ ($x1$), compute P(z1|x1) and $P(w0|x1)$.



$$
\begin{aligned}
P(w0|x1) &= P(w0|z1,\, x1)P(z1|x1) + P(w0|z0,\, x1)P(z0|x1) \\
&= P(w0|z1)P(z1|x1) + P(w0|z0)P(z0|x1) \\
&= (0.55)(0.46) + (0.7)(0.54) = 0.63
\end{aligned}
$$

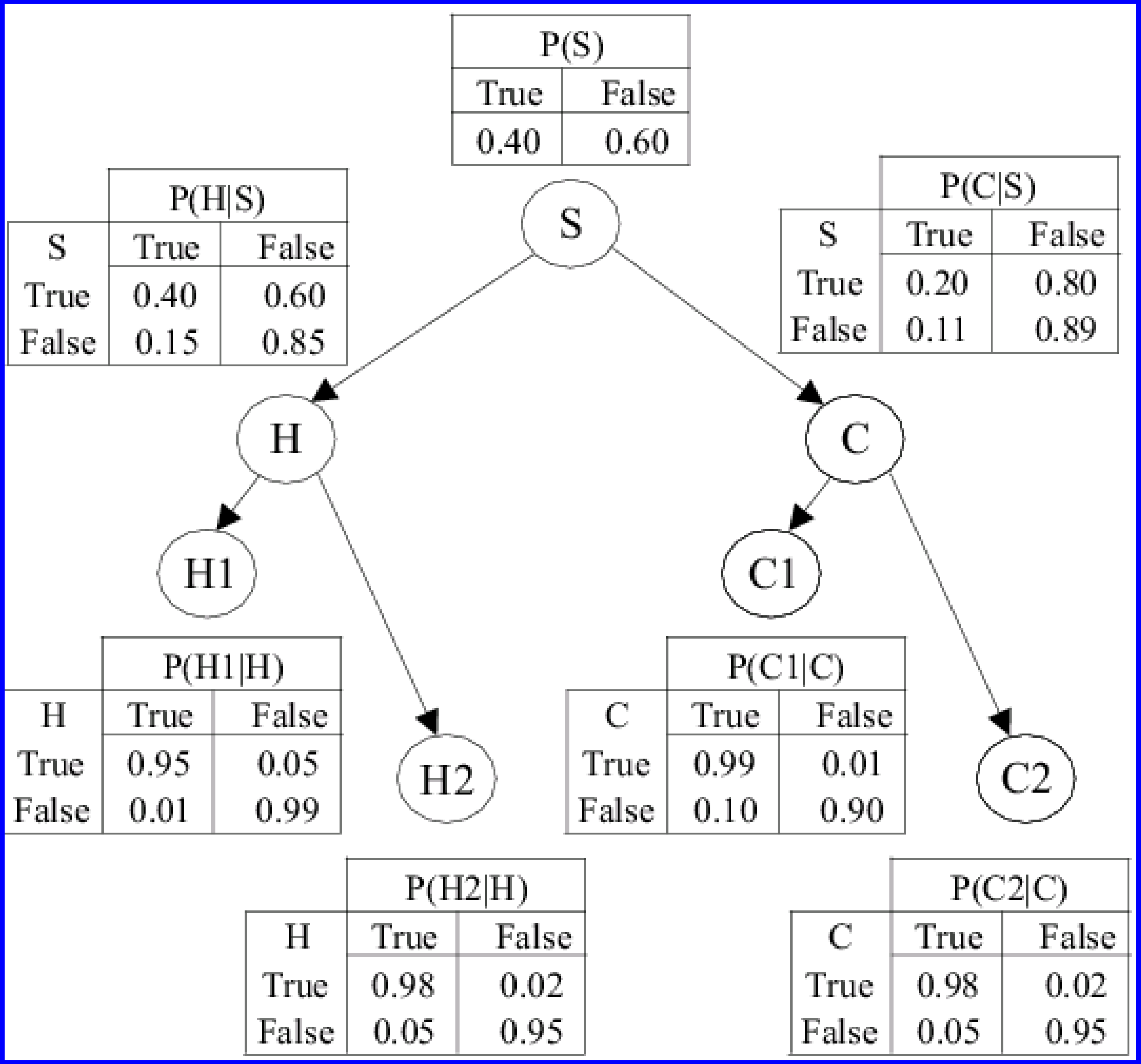
b) If w is measured to be w=1 (w1) compute P(z1|w1|)].

$P(x1)=0.60$	$P(y1\|x1)=0.40$	$P(z1\|y1)=0.25$	$P(w1\|z1)=0.45$
	$P(y1\|x0)=0.30$	$P(z1\|y0)=0.60$	$P(w1\|z0)=0.30$

$P(x0)=0.40$	$P(y0\|x1)=0.60$	$P(z0\|y1)=0.75$	$P(w0\|z1)=0.55$
	$P(y0\|x0)=0.70$	$P(z0\|y0)=0.40$	$P(w0\|z0)=0.70$
	$P(y1)=0.36$	$P(z1)=0.47$	$P(w1)=0.37$
	$P(y0)=0.64$	$P(z0)=0.53$	$P(w0)=0.63$

$$P(z1|w1) = \frac{P(w1|z1)P(z1)}{P(w1)} = \frac{(0.45)(0.47)}{0.37} = 0.57$$

What's about more complex networks?



P(S)	
True	False
0.40	0.60

	P(H|S)	
S	True	False
True	0.40	0.60
False	0.15	0.85

	P(C|S)	
S	True	False
True	0.20	0.80
False	0.11	0.89

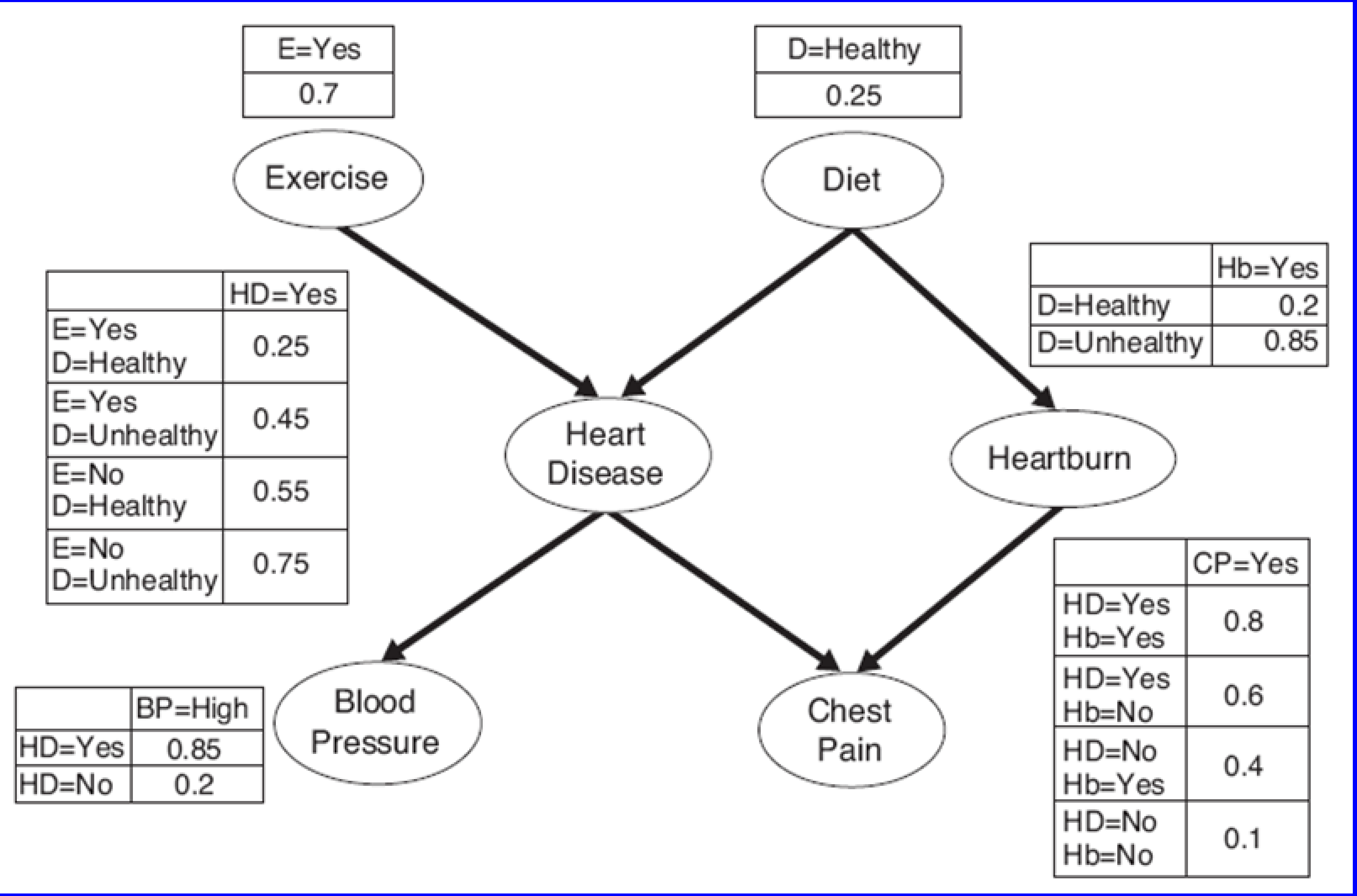
	P(H1|H)	
H	True	False
True	0.95	0.05
False	0.01	0.99

	P(C1|C)	
C	True	False
True	0.99	0.01
False	0.10	0.90

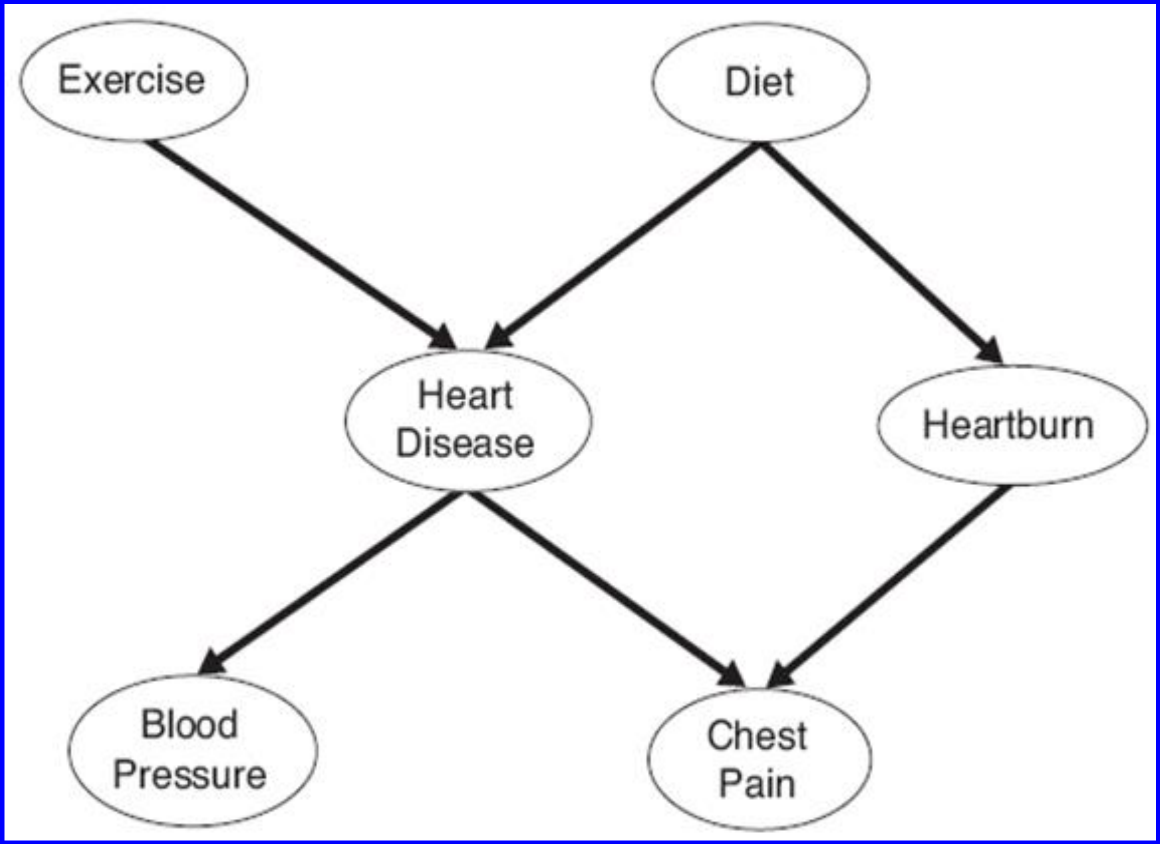
	P(H2|H)	
H	True	False
True	0.98	0.02
False	0.05	0.95

	P(C2|C)	
C	True	False
True	0.98	0.02
False	0.05	0.95

What's about more complex networks?

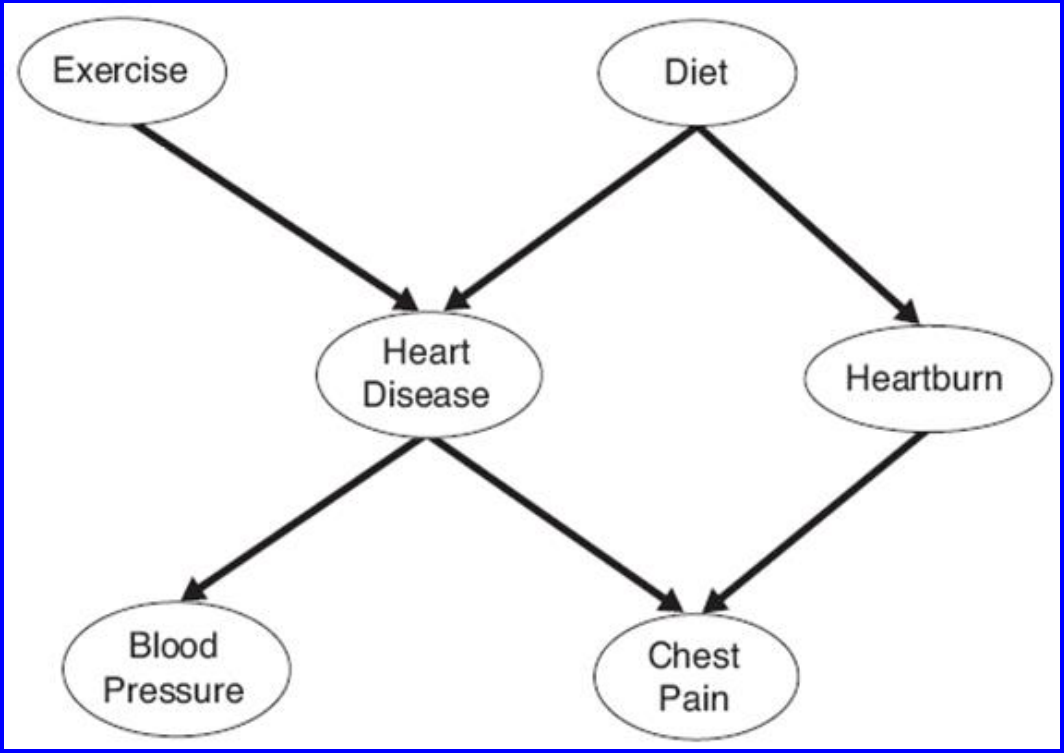


We will study this graph



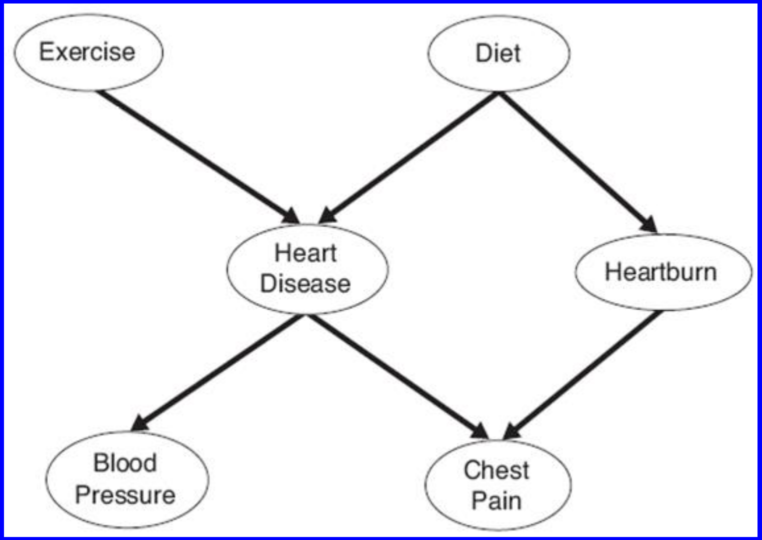
We can show:

- $P(D|E)=P(D)$
- $P(Hb|HD, E, D)= P(Hb|D)$
- $P(CP|Hb, HD, E, D)= P(CP|Hb, HD)$
- $P(BP|CP, Hb, HD, E, D)= P(BP|HD)$
- However, $P(HD|E,D)$ cannot be implified



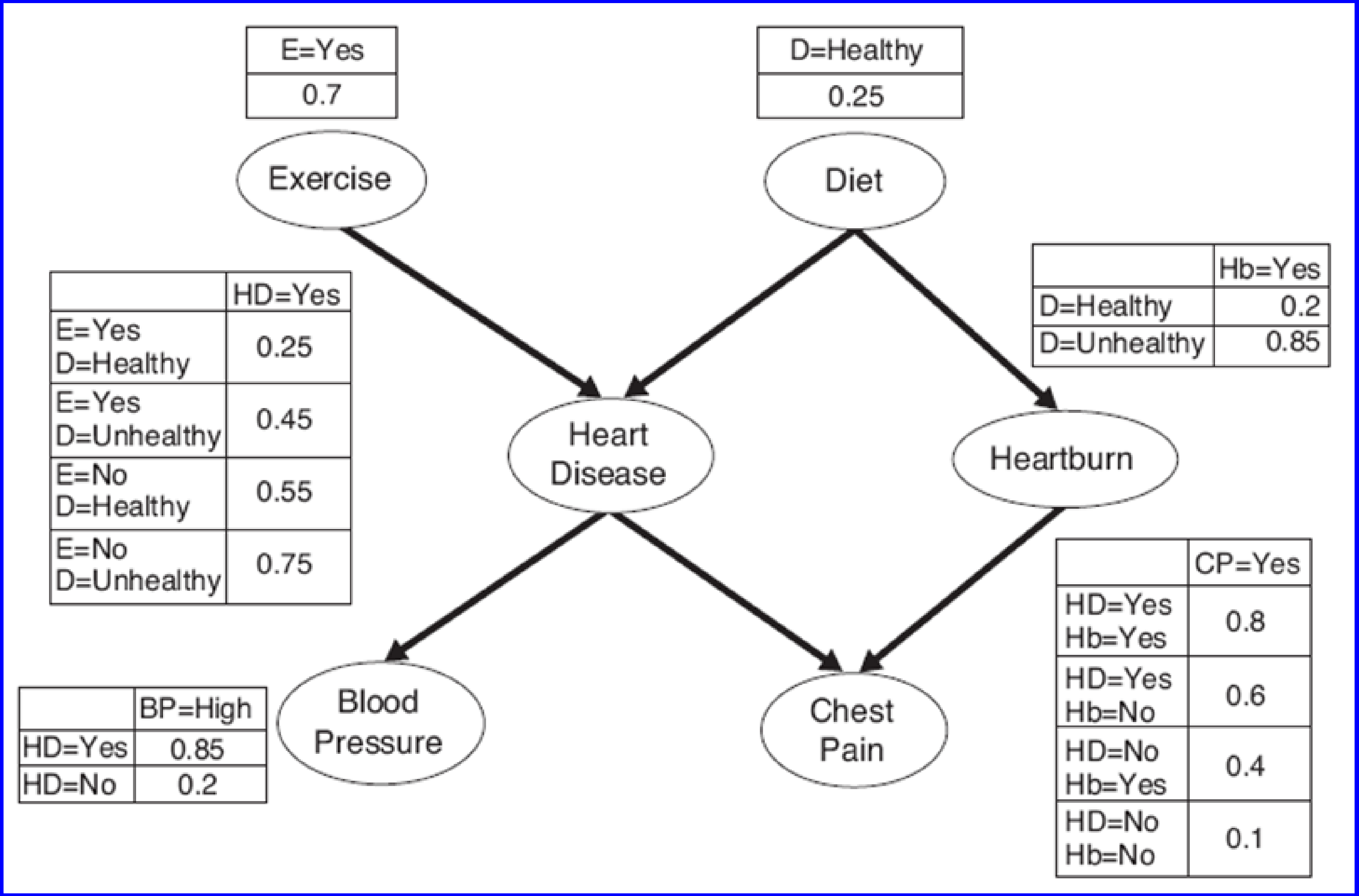
Exercise:

- $P(CP|HD, BP, E, D) = ?$



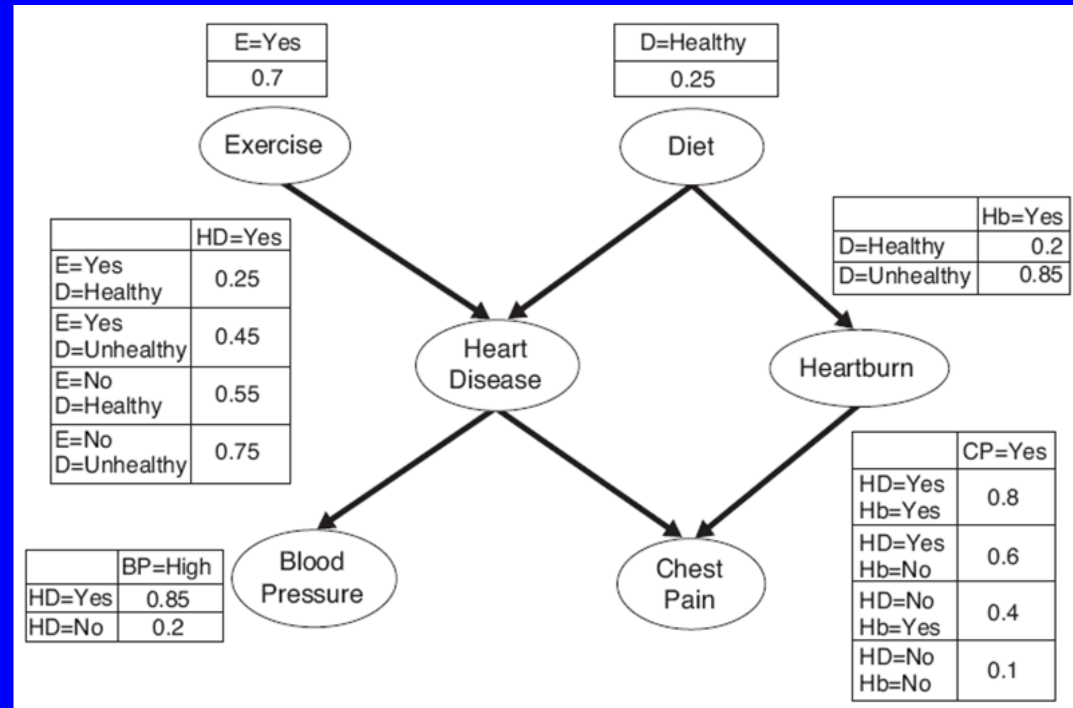
Exercise:

- *P (CP|HD, BP, E, D)*= No simplification



Calculate P(HD=yes) ?

Calculate P(HD=yes) ?



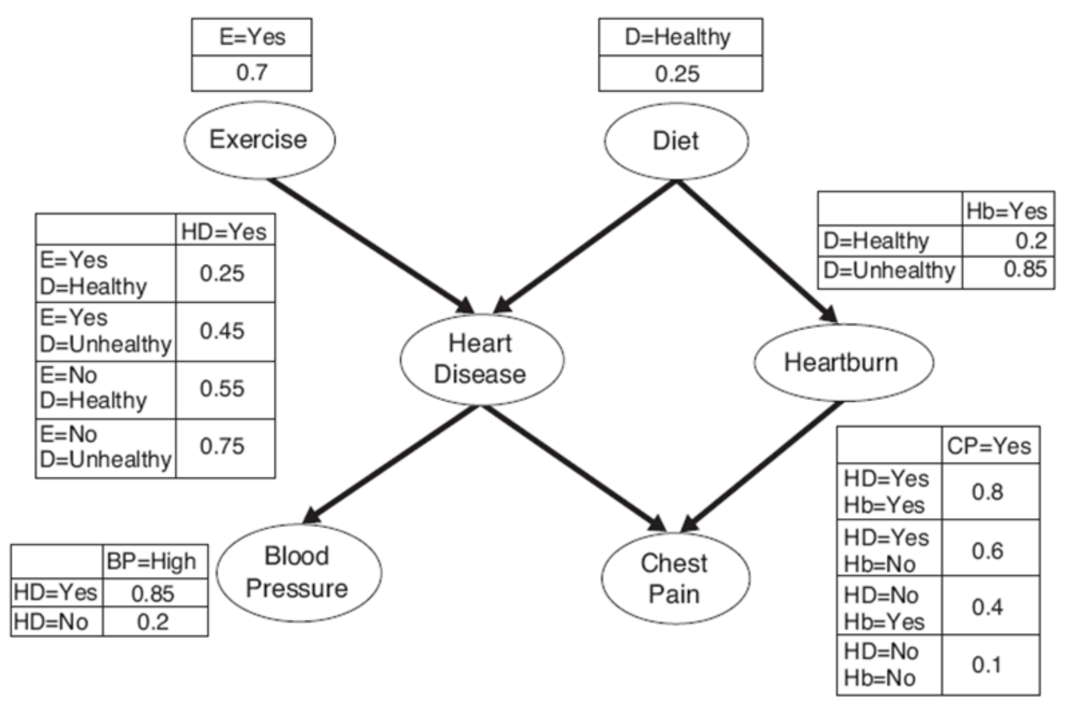
$$P(HD=Yes)=\sum_{\alpha}\sum_{\beta}P(HD=yes\,|\,E=\alpha,D=\beta)P(E=\alpha,D=\beta)$$

where,

α = Set of Values of Exercise(E) = {Yes, No}

β = Set of Values of Diet(D) = {Healthy, Not Healthy}

Calculate P(HD=yes) ?



	HD=Yes
E=Yes D=Healthy	0.25
E=Yes D=Unhealthy	0.45
E=No D=Healthy	0.55
E=No D=Unhealthy	0.75

	Hb=Yes
D=Healthy	0.2
D=Unhealthy	0.85

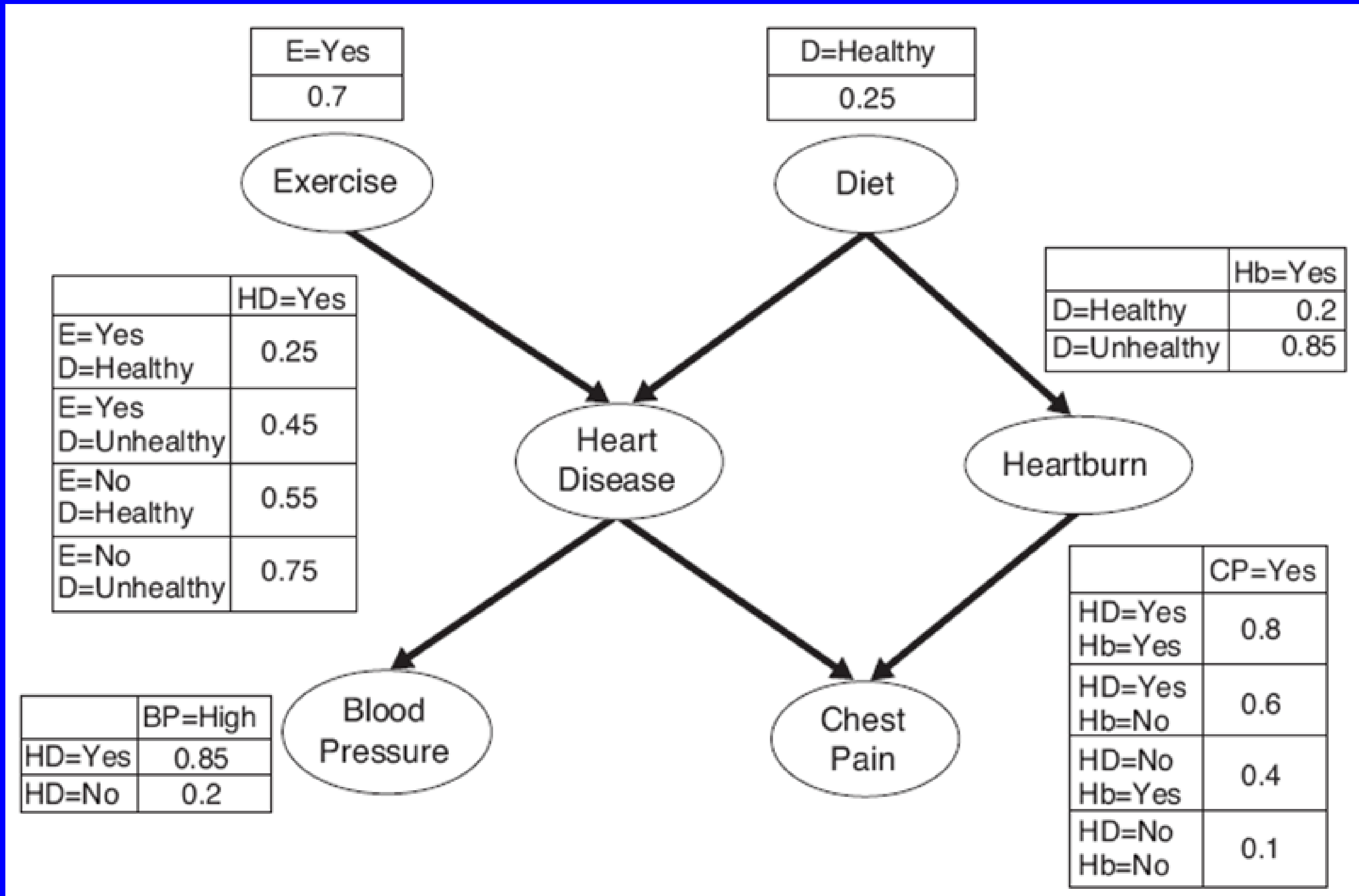
	BP=High
HD=Yes	0.85
HD=No	0.2

	CP=Yes
HD=Yes Hb=Yes	0.8
HD=Yes Hb=No	0.6
HD=No Hb=Yes	0.4
HD=No Hb=No	0.1

$$P(HD=Yes)=\sum_{\alpha}\sum_{\beta}P(HD=yes\,|\,E=\alpha,D=\beta)P(E=\alpha,D=\beta)$$

$$=\sum_{\alpha}\sum_{\beta}P(HD=yes\,|\,E=\alpha,D=\beta)P(E=\alpha)P(D=\beta)$$

$$=0.25\times0.7\times0.25+0.45\times0.7\times0.75+0.55\times0.3\times0.25$$
$$+0.75\times0.3\times0.75$$
$$=0.49$$



E=Yes
0.7

D=Healthy
0.25

	HD=Yes
E=Yes D=Healthy	0.25
E=Yes D=Unhealthy	0.45
E=No D=Healthy	0.55
E=No D=Unhealthy	0.75

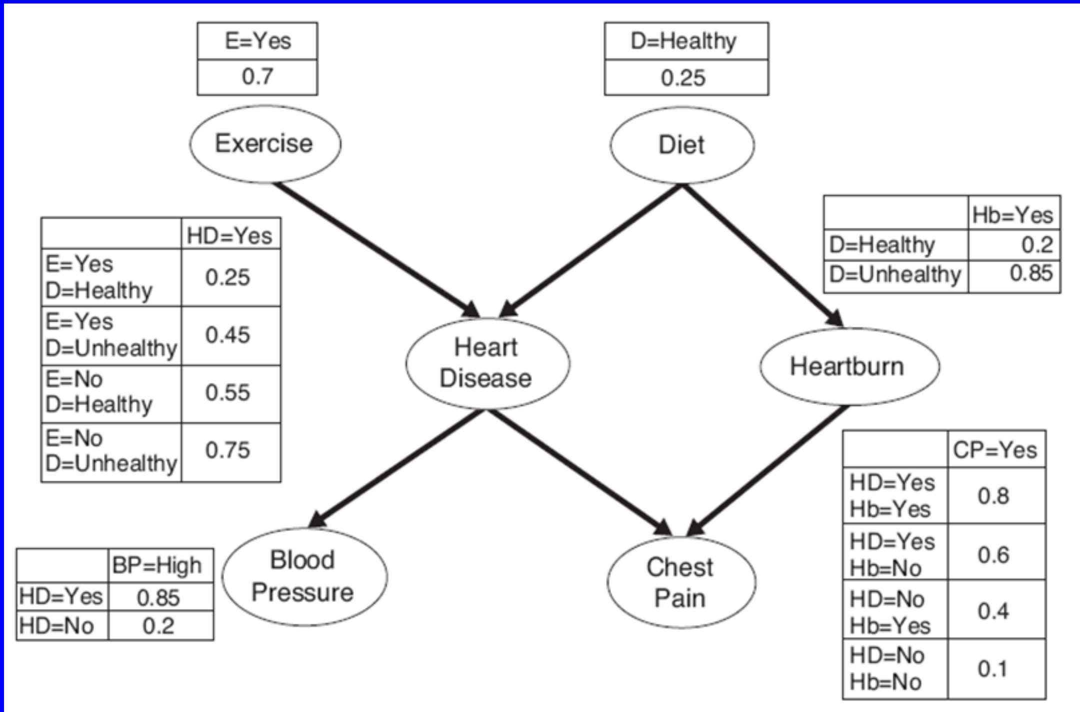
	Hb=Yes
D=Healthy	0.2
D=Unhealthy	0.85

	BP=High
HD=Yes	0.85
HD=No	0.2

	CP=Yes
HD=Yes Hb=Yes	0.8
HD=Yes Hb=No	0.6
HD=No Hb=Yes	0.4
HD=No Hb=No	0.1

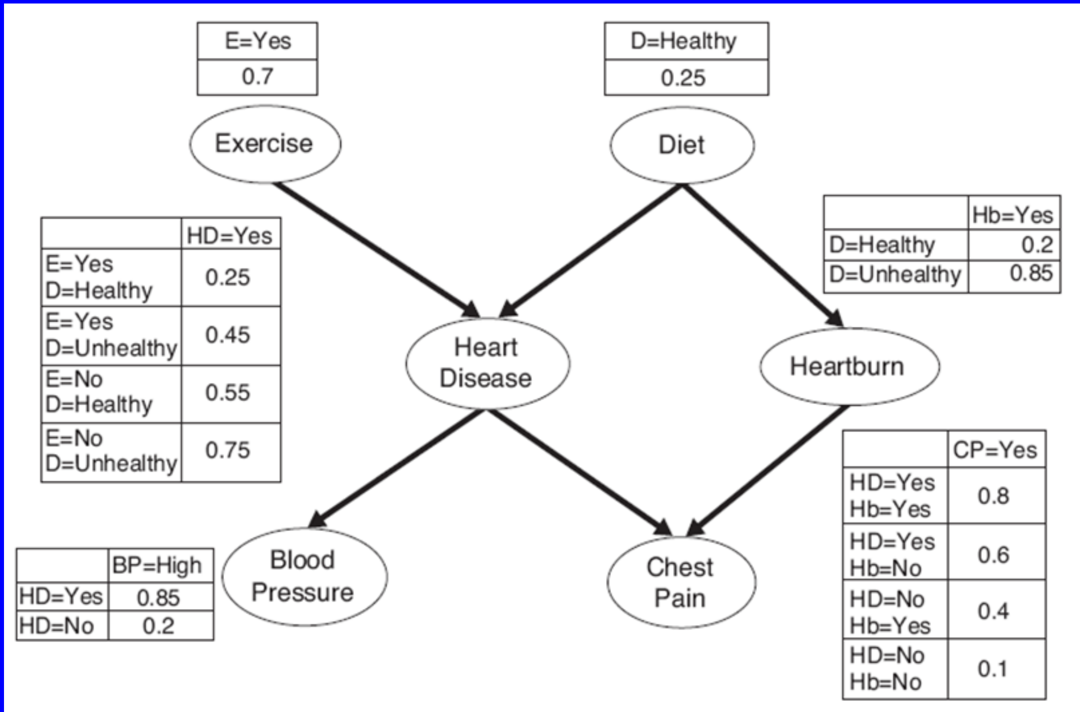
Calculate P(HD=yes|BP=High)?

Calculate P(HD=yes|BP=High)



$P(HD = yes \mid BP = High)$ can be written as $\dfrac{P(BP = High \mid HD = yes)P(HD = yes)}{P(BP = High)}$

Calculate P(HD=yes|BP=High)

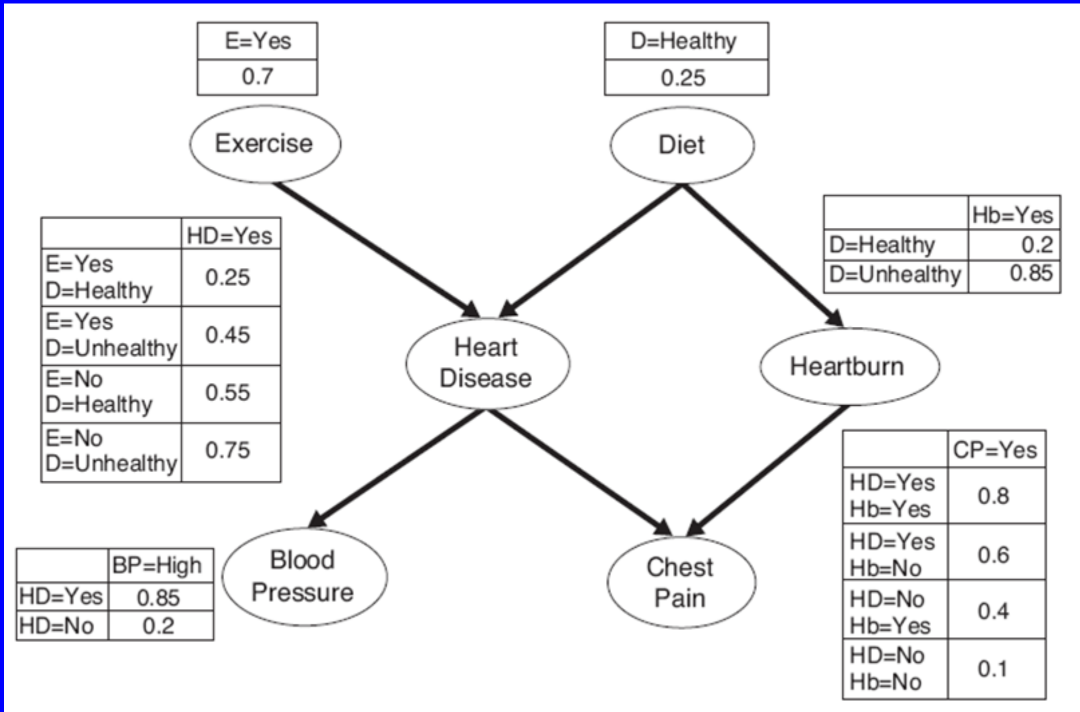


$$P(BP = High) = \sum_{\gamma} P(BP = high \mid HD = \gamma)P(HD = \gamma)$$

where,

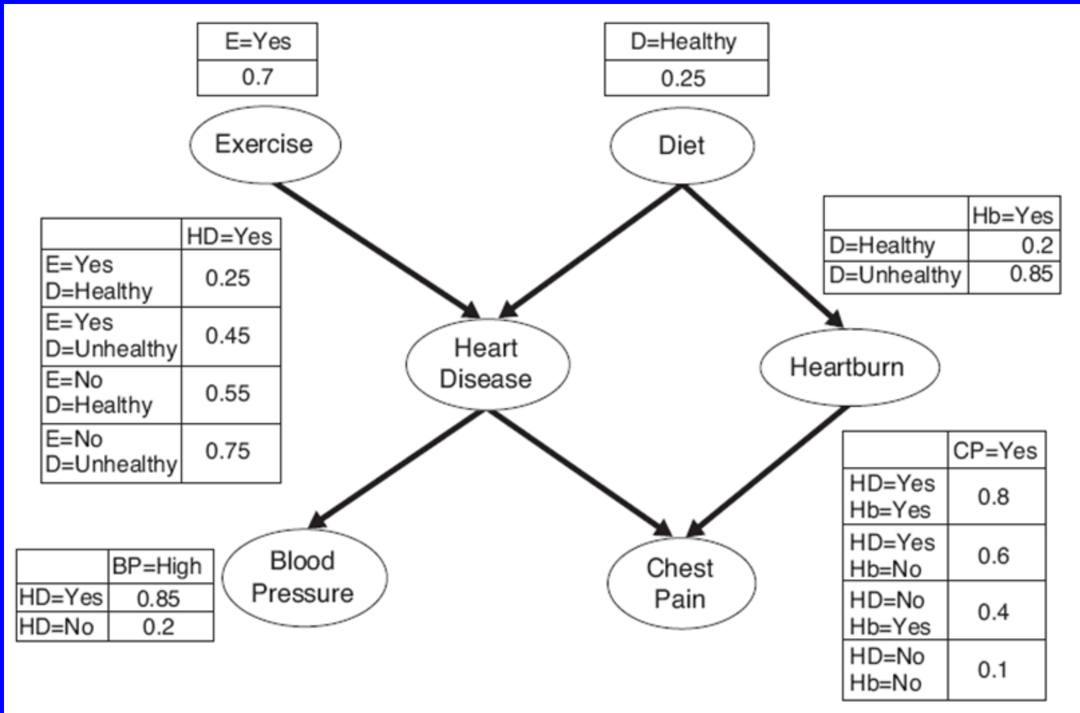
γ = Set of Values of Heart Disease (HD) = {Yes, No}

Calculate P(HD=yes|BP=High)



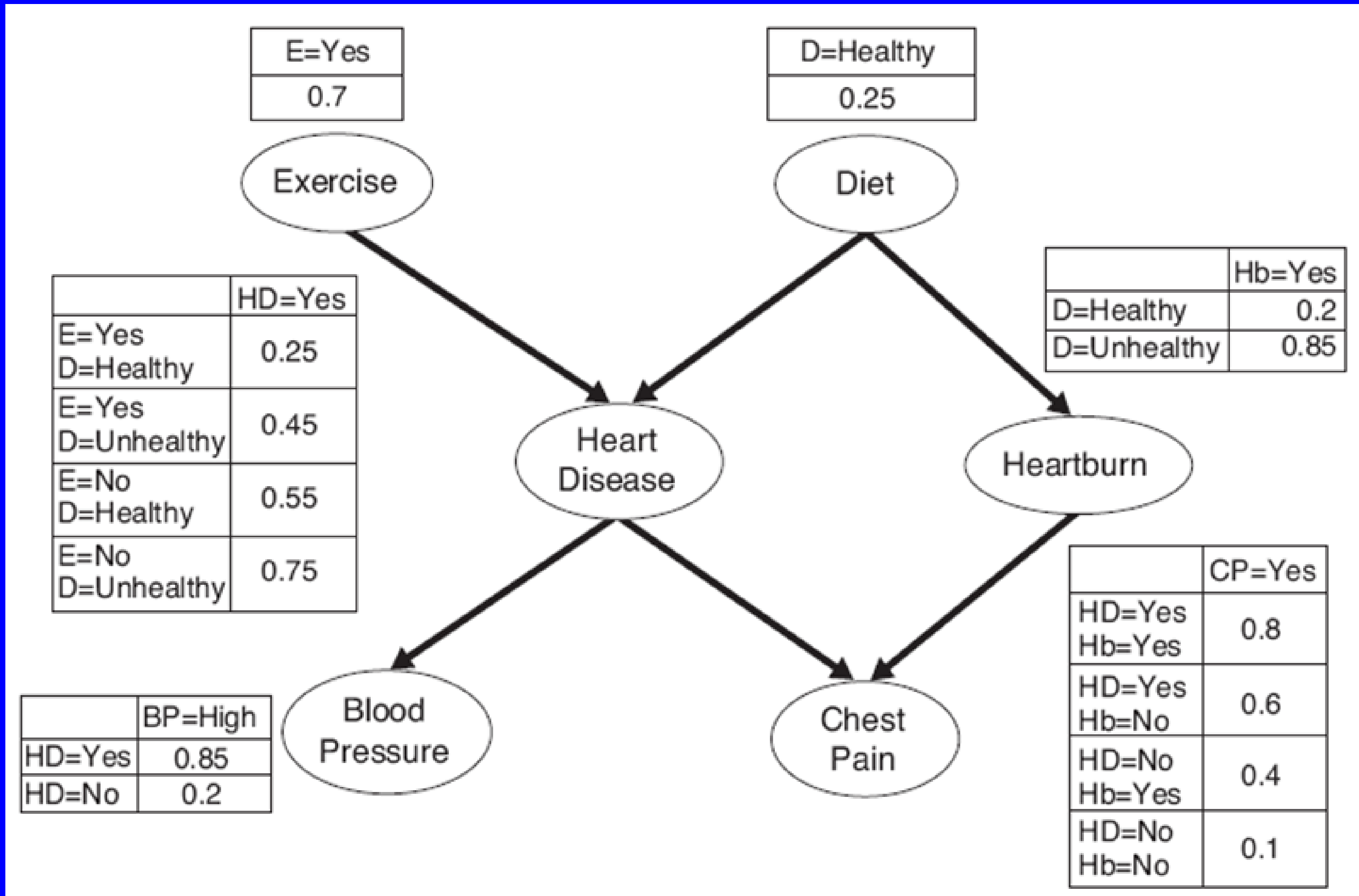
$$P(BP = High) = \sum_{\gamma} P(BP = high \mid HD = \gamma) P(HD = \gamma)$$
$$= 0.85 \times 0.49 + 0.2 \times 0.51 = 0.5185$$

Calculate P(HD=yes|BP=High)



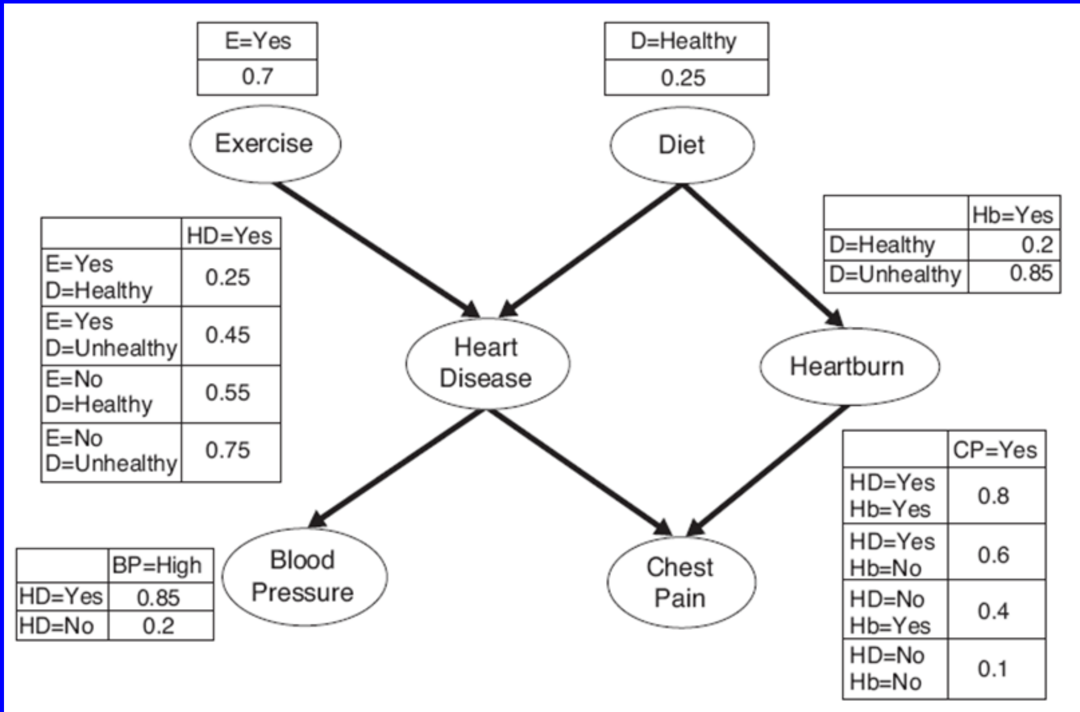
$$P(HD = yes \mid BP = High) = \frac{P(BP = High \mid HD = yes)P(HD = yes)}{P(BP = High)}$$

$$= \frac{0.85 \times 0.49}{0.5185} = 0.8033$$



Calculate P(HD=yes|BP=high, D=Healthy, E=yes)?

Calculate P(HD=yes|BP=high, D=Healthy, E=yes) ?



	HD=Yes
E=Yes D=Healthy	0.25
E=Yes D=Unhealthy	0.45
E=No D=Healthy	0.55
E=No D=Unhealthy	0.75

	Hb=Yes
D=Healthy	0.2
D=Unhealthy	0.85

	BP=High
HD=Yes	0.85
HD=No	0.2

	CP=Yes
HD=Yes Hb=Yes	0.8
HD=Yes Hb=No	0.6
HD=No Hb=Yes	0.4
HD=No Hb=No	0.1

$$P(HD = yes \mid BP = high, D = Healthy, E = Yes)$$

$$= \frac{P(BP = high \mid HD = yes, D = Healthy, E = Yes)}{P(BP = high \mid D = Healthy, E = Yes)} \times P(HD = yes \mid D = Healthy, E = Yes)$$

How is this formula true?

$$P(HD = yes \mid BP = high, D = Healthy, E = Yes)$$
$$= \frac{P(BP = high \mid HD = yes, D = Healthy, E = Yes)}{P(BP = high \mid D = Healthy, E = Yes)} \times P(HD = yes \mid D = Healthy, E = Yes)$$

Let

$$P(X \mid Y) = \frac{P(Y \mid X)}{P(Y)} \times P(X)$$

Now add *Z* and *W* as condition

$$P(X \mid Y, Z, W) = \frac{P(Y \mid X, Z, W)}{P(Y \mid Z, W)} \times P(X \mid Z, W)$$

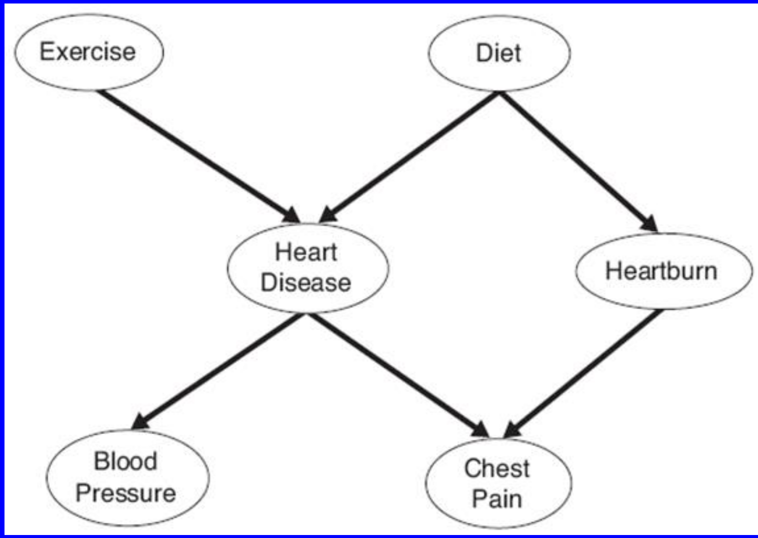
Similarly,

$$P(HD = yes \mid BP = high) = \frac{P(BP = high \mid HD = yes)}{P(BP = high)} \times P(HD = yes)$$

Now add conditions *D= Healthy* and *E= Yes* to above formula

Similarly,

$$P(BP = high \mid D = Healthy, E = Yes)$$
$$= \sum_{\gamma} P(BP = high \mid HD = \gamma, D = Healthy, E = Yes) \times P(HD = \gamma \mid D = Healthy, E = Yes)$$



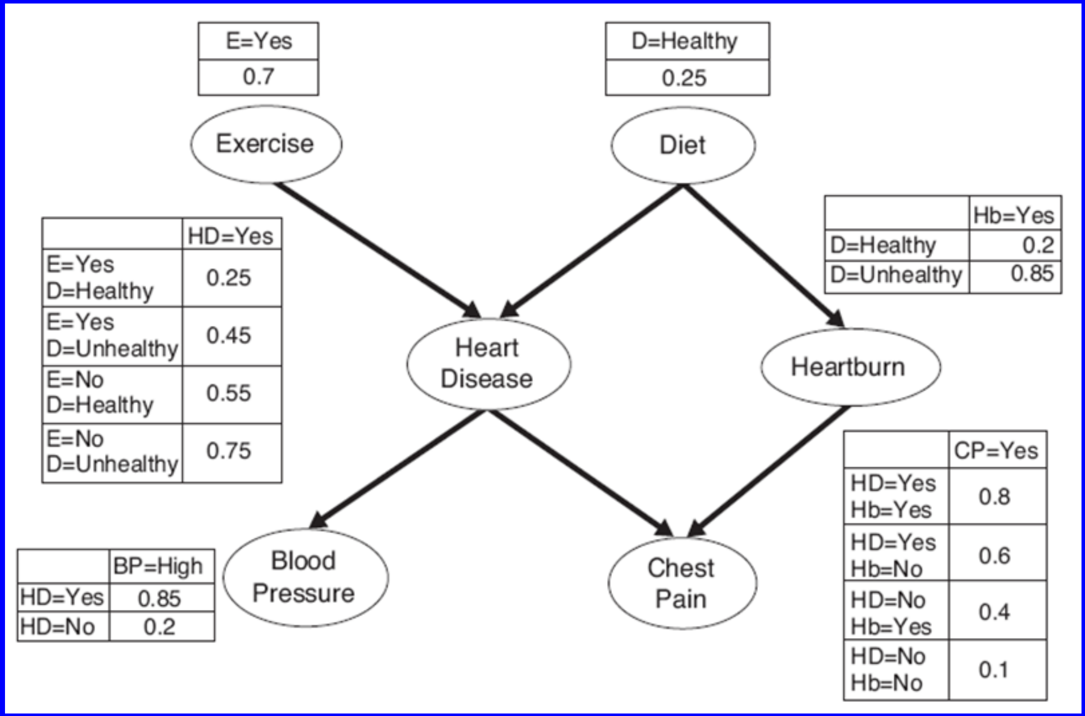
Proof:

$$P(BP = high) = \sum_{\gamma} P(BP = high \mid HD = \gamma) \times P(HD = \gamma)$$

Adding conditions *D= Healthy* and *E= Yes*

we get,

$$P(BP = high \mid D = Healthy, E = Yes)$$
$$= \sum_{\gamma} P(BP = high \mid HD = \gamma, D = Healthy, E = Yes) \times P(HD = \gamma \mid D = Healthy, E = Yes)$$
$$= \sum_{\gamma} P(BP = high \mid HD = \gamma) \times P(HD = \gamma \mid D = Healthy, E = Yes)$$



$$P(HD = yes \mid BP = high, D = Healthy, E = Yes)$$

$$= \frac{P(BP = high \mid HD = yes, D = Healthy, E = Yes)}{P(BP = high \mid D = Healthy, E = Yes)} \times P(HD = yes \mid D = Healthy, E = Yes)$$

$$= \frac{P(BP = high \mid HD = yes)}{P(BP = high \mid D = Healthy, E = Yes)} \times P(HD = yes \mid D = Healthy, E = Yes)$$

$$= \frac{P(BP = high \mid HD = yes)}{\sum_{\gamma} P(BP = high \mid HD = \gamma) P(HD = \gamma \mid D = Healthy, E = Yes)} \times P(HD = yes \mid D = Healthy, E = Yes)$$

$$= \frac{0.85 \times 0.25}{0.85 \times 0.25 + 0.2 \times 0.75} = 0.5862$$