Bayesian Belief Network

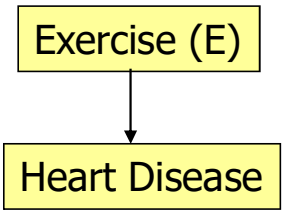
- Naïve Bayes Classifier
  - Independent Assumption
- Bayesian Belief Network
  - Do not have independent assumption

Yes/No | Healthy/ Unhealthy | Yes/No | High/Low | Yes/No | Yes/No

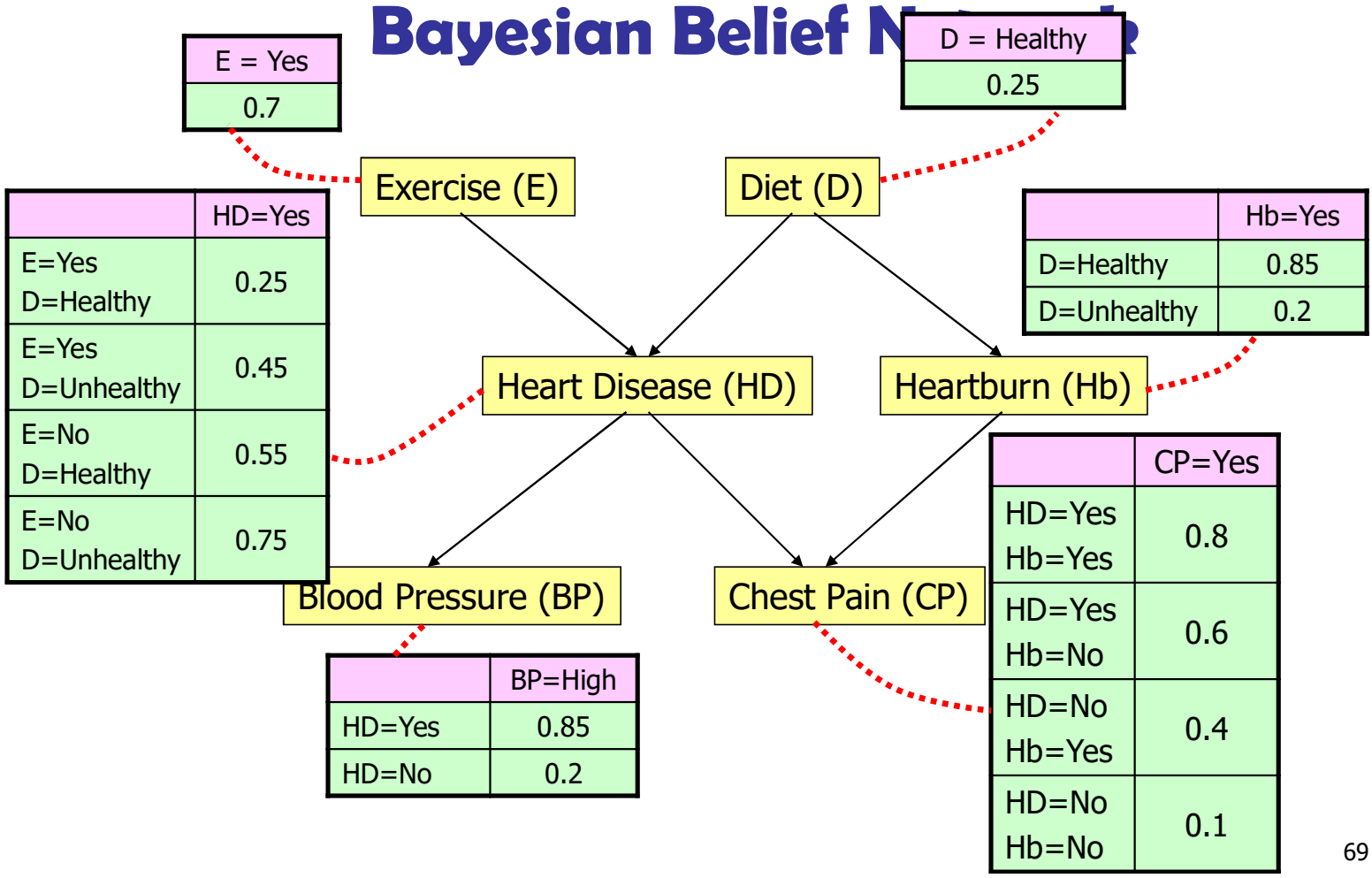
Exercise	Diet	Heartburn	Blood Pressure	Chest Pain	Heart Disease
Yes	Healthy	No	High	Yes	No
No	Unhealthy	Yes	Low	Yes	No
No	Healthy	Yes	High	No	Yes
…	…	…	…	…	…

Some attributes are dependent on other attributes.

e.g., doing exercises may reduce the probability of suffering from Heart Disease



Bayesian Belief N



Let X, Y, Z be three random variables.
X is said to be **conditionally independent** of Y given Z if the following holds.

$$P(X \mid Y, Z) = P(X \mid Z)$$

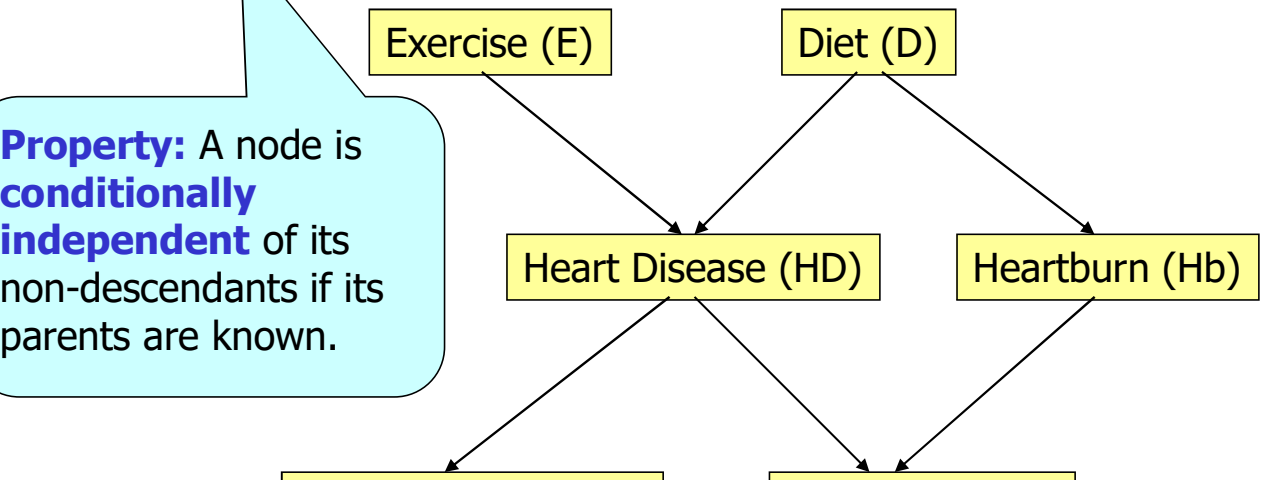
Lemma:
If X is conditionally independent of Y given Z,

$$P(X, Y \mid Z) = P(X \mid Z) \times P(Y \mid Z) \ ?$$

Let X, Y, Z be three random variables.
X is said to be **conditionally independent** of Y given Z if the following holds.

$$P(X \mid Y, Z) = P(X \mid Z)$$

Property: A node is **conditionally independent** of its non-descendants if its parents are known.



e.g., P(BP = High | HD = Yes, D = Healthy) = P(BP = High | HD = Yes)

"BP = High" is **conditionally independent** of "D = Healthy" given "HD = Yes"

e.g., P(BP = High | HD = Yes, CP=Yes) = P(BP = High | HD = Yes)

"BP = High" is **conditionally independent** of "CP = Yes" given "HD = Yes"

Yes/No | Healthy/ Unhealthy | Yes/No | High/Low | Yes/No | Yes/No

Exercise	Diet	Heartburn	Blood Pressure	Chest Pain	Heart Disease
Yes	Healthy	No	High	Yes	No
No	Unhealthy	Yes	Low	Yes	No
No	Healthy	Yes	High	No	Yes
…	…	…	…	…	…

Suppose there is a new person and I want to know whether he is likely to have Heart Disease.

Exercise	Diet	Heartburn	Blood Pressure	Chest Pain	Heart Disease
?	?	?	?	?	?

Exercise	Diet	Heartburn	Blood Pressure	Chest Pain	Heart Disease
?	?	?	High	?	?

Exercise	Diet	Heartburn	Blood Pressure	Chest Pain	Heart Disease
Yes	Healthy	?	High	?	?

Bayesian Belief Network

Suppose there is a new person and I want to know whether he is likely to have Heart Disease.

Exercise	Diet	Heartburn	Blood Pressure	Chest Pain	Heart Disease
?	?	?	?	?	?

$$P(HD = Yes) = \sum_{x \in \{Yes, No\}} \sum_{y \in \{Healthy, Unhealthy\}} P(HD=Yes|E=x, D=y) \times P(E=x, D=y)$$

$$= \sum_{x \in \{Yes, No\}} \sum_{y \in \{Healthy, Unhealthy\}} P(HD=Yes|E=x, D=y) \times P(E=x) \times P(D=y)$$

$$= 0.25 \times 0.7 \times 0.25 + 0.45 \times 0.7 \times 0.75 + 0.55 \times 0.3 \times 0.25 + 0.75 \times 0.3 \times 0.75$$

$$= 0.49$$

$$P(HD = No) = 1 - P(HD = Yes)$$

$$= 1 - 0.49$$

$$= 0.51$$

Bayesian Belief Network

Suppose there is a new person and I want to know whether he is likely to have Heart Disease.

Exercise	Diet	Heartburn	Blood Pressure	Chest Pain	Heart Disease
?	?	?	High	?	?

$$P(BP = High) = \sum_{x \in \{Yes, No\}} P(BP = High|HD=x) \times P(HD = x)$$
$$= 0.85x0.49 + 0.2x0.51$$
$$= 0.5185$$

$$P(HD = Yes|BP = High) = \frac{P(BP = High|HD=Yes) \times P(HD = Yes)}{P(BP = High)}$$
$$= \frac{0.85 \times 0.49}{0.5185}$$
$$= 0.8033$$

$$P(HD = No|BP = High) = 1 - P(HD = Yes|BP = High)$$
$$= 1 - 0.8033$$
$$= 0.1967$$

Bayesian Belief Network

Suppose there is a new person and I want to know whether he is likely to have Heart Disease.

Exercise	Diet	Heartburn	Blood Pressure	Chest Pain	Heart Disease
Yes	Healthy	?	High	?	?

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

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

$$= \frac{P(BP = High \mid HD = Yes)\, P(HD = Yes \mid D = Healthy, E = Yes)}{\sum_{x \in \{Yes, No\}} P(BP=High \mid HD=x)\, P(HD=x \mid D=Healthy, E=Yes)}$$

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

$$= 0.5862$$

$$P(HD = No \mid BP = High, D = Healthy, E = Yes)$$

$$= 1 - P(HD = Yes \mid BP = High, D = Healthy, E = Yes)$$

$$= 1 - 0.5862$$

$$= 0.4138$$