SC4000/CZ4041/CE4041: Machine Learning

Solutions to L4 Tutorial Questions

Kelly KE
School of Computer Science and Engineering,
NTU, Singapore

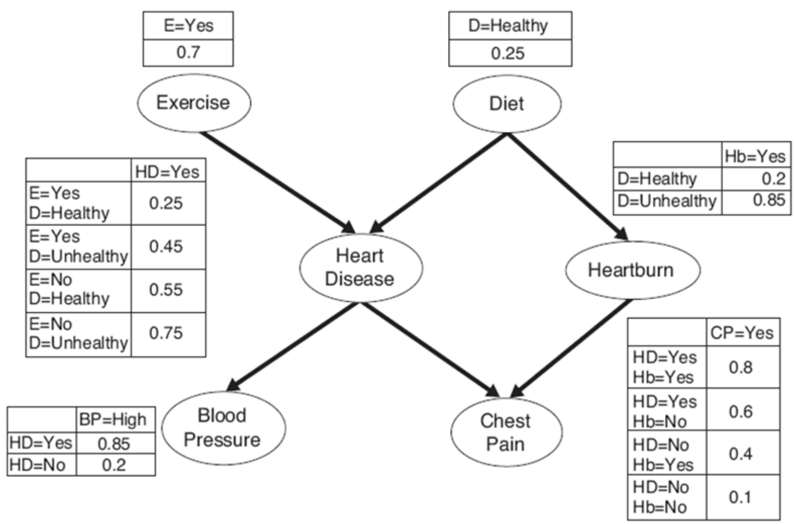
Question 1

- If the person has high blood pressure, but exercises regularly and eats a healthy diet, to diagnose about heart disease (estimate the probabilities)

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

v.s.

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



$$P(\text{HD=Yes}|\text{BP=High, D=Healthy, E=Yes})$$

Denote by $\mathbf{U} = \{\text{BP, D, E}\}$

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

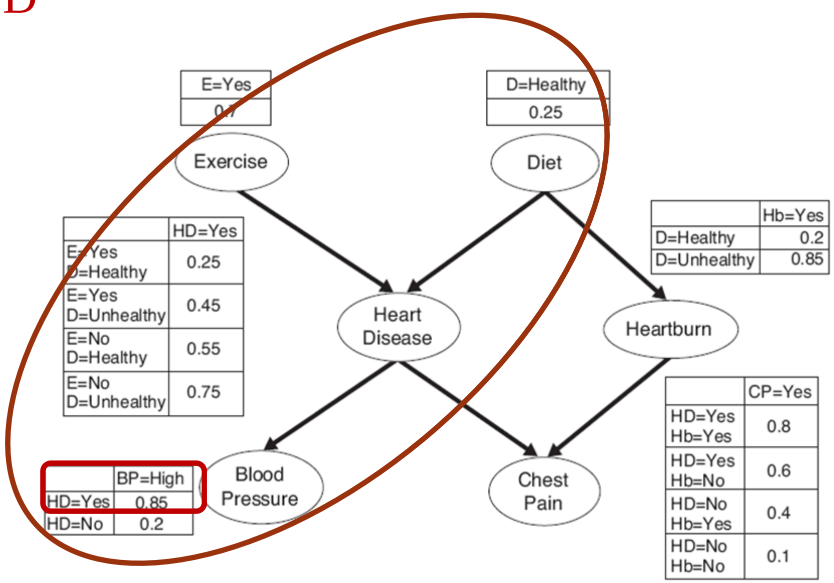
$$P(\text{HD}|\mathbf{U}) = \frac{P(\text{HD}, \mathbf{U})}{P(\mathbf{U})}$$

$$= \frac{\boxed{P(\text{BP=High}|\text{HD=Yes, D=Healthy, E=Yes})}\,\boxed{P(\text{HD=Yes, D=Healthy, E=Yes})}}{P(\text{BP=High, D=Healthy, E=Yes})}$$

Next slide

When HD is observed, BP is conditionally independent of E and D

$$P(\text{BP=High}|\text{HD=Yes}) = 0.85$$



	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

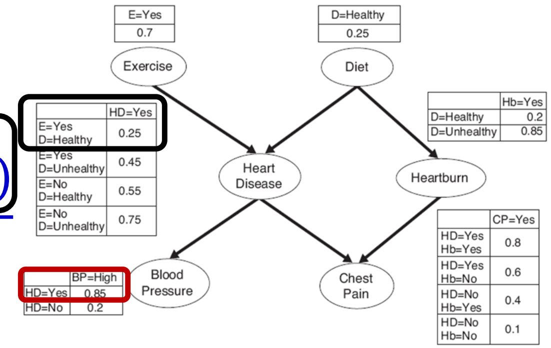
	BP=High
HD=Yes	0.85
HD=No	0.2

	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

$$P(\text{HD=Yes}|\text{BP=High, D=Healthy, E=Yes})$$

$$= \frac{\overset{0.85}{\boxed{P(\text{BP=High}|\text{HD=Yes})}}\,\overset{1}{\boxed{P(\text{HD=Yes, D=Healthy, E=Yes})}}}{\boxed{P(\text{BP=High, D=Healthy, E=Yes})}\;{}^{2}}$$

$$= \frac{\overset{0.85}{\boxed{P(\text{BP=High}|\text{HD=Yes})}}\,\overset{0.25}{\boxed{P(\text{HD=Yes}|\text{D=Healthy, E=Yes})}}}{P(\text{BP=High}|\text{D=Healthy, E=Yes})}$$



We can use the method used in Example 1 (Lecture notes) to estimate the exact probability of P(HD=Yes, D=Healthy, E=Yes). Or alternatively

1

$$P(\text{HD=Yes, D=Healthy, E=Yes})$$

$$= \quad P(\text{HD=Yes}|\text{D=Healthy, E=Yes})\cancel{P(\text{D=Healthy, E=Yes})}$$ Product rule

2

$$P(\text{BP=High, D=Healthy, E=Yes})$$

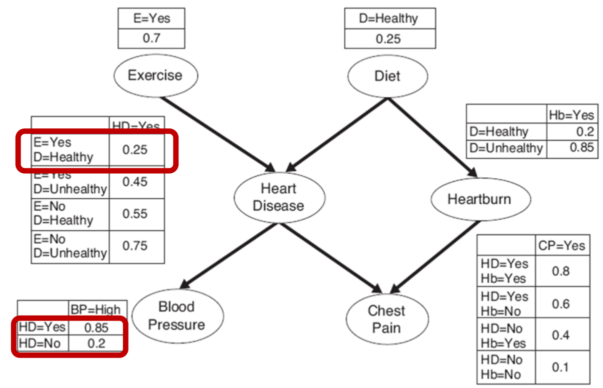
Canceled out

$$= \quad P(\text{BP=High}|\text{D=Healthy, E=Yes})\cancel{P(\text{D=Healthy, E=Yes})}$$ Product rule

$$P(\text{HD=Yes}|\text{BP=High}, \text{D=Healthy}, \text{E=Yes})$$

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

$$= \frac{0.85 \times 0.25}{P(\text{BP=High}|\text{D=Healthy}, \text{E=Yes})}$$



$$P(\text{HD=No}|\text{BP=High}, \text{D=Healthy}, \text{E=Yes})$$

$$= \frac{P(\text{BP=High}|\text{HD=No})P(\text{HD=No}|\text{D=Healthy}, \text{E=Yes})}{P(\text{BP=High}|\text{D=Healthy}, \text{E=Yes})}$$

$$= \frac{0.2 \times 0.75}{P(\text{BP=High}|\text{D=Healthy}, \text{E=Yes})}$$

$$P(\text{HD=No}|\text{D=Healthy}, \text{E=Yes})$$
$$= 1 - P(\text{HD=Yes}|\text{D=Healthy}, \text{E=Yes})$$
$$= 0.75$$

How to estimate $P(\text{BP=High}|\text{D=Healthy}, \text{E=Yes})$?

$$\boxed{\begin{array}{c} P(\text{HD=Yes}|\text{BP=High},\text{D=Healthy},\text{E=Yes}) \\ \parallel \\ \dfrac{0.85 \times 0.25}{P(\text{BP=High}|\text{D=Healthy},\text{E=Yes})} \end{array}} + \boxed{\begin{array}{c} P(\text{HD=No}|\text{BP=High},\text{D=Healthy},\text{E=Yes}) \\ \parallel \\ \dfrac{0.2 \times 0.75}{P(\text{BP=High}|\text{D=Healthy},\text{E=Yes})} \end{array}} = 1$$

$$\frac{0.85 \times 0.25}{P(\text{BP=High}|\text{D=Healthy},\text{E=Yes})} + \frac{0.2 \times 0.75}{P(\text{BP=High}|\text{D=Healthy},\text{E=Yes})} = 1$$

$$\Downarrow$$

$$\frac{0.85 \times 0.25 + 0.2 \times 0.75}{P(\text{BP=High}|\text{D=Healthy},\text{E=Yes})} = 1$$

$$\Downarrow$$

$$P(\text{BP=High}|\text{D=Healthy},\text{E=Yes}) = 0.85 \times 0.25 + 0.2 \times 0.75$$

Therefore

$$P(\mathrm{HD{=}Yes}|\mathrm{BP{=}High, D{=}Healthy, E{=}Yes})$$

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

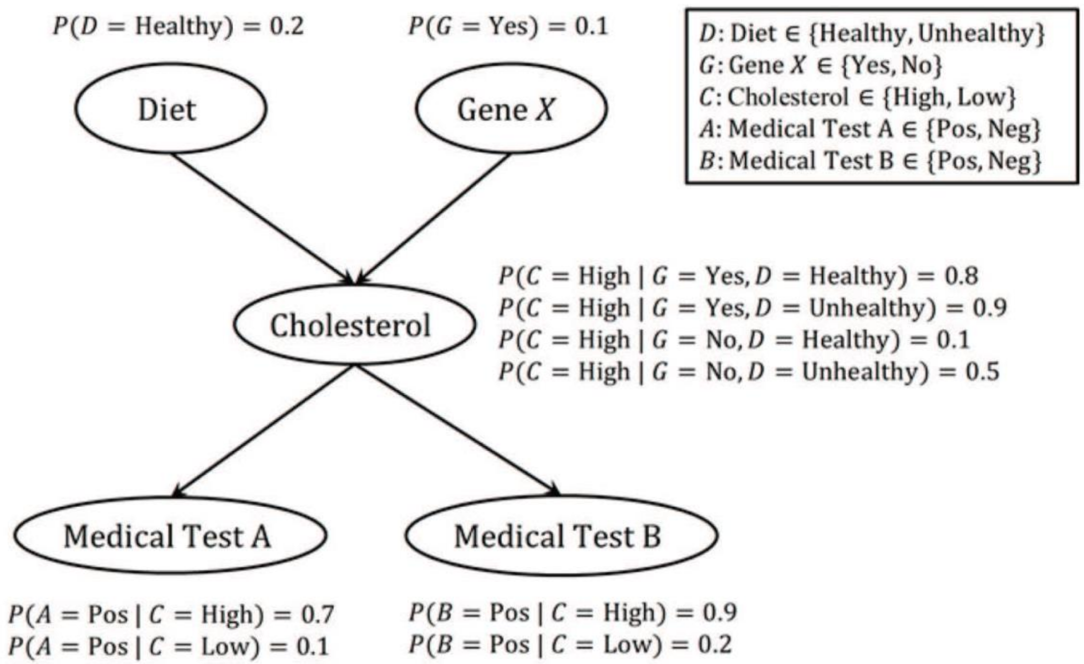
$$P(\mathrm{HD{=}No}|\mathrm{BP{=}High, D{=}Healthy, E{=}Yes})$$

$$= \frac{0.2 \times 0.75}{0.85 \times 0.25 + 0.2 \times 0.75} = 0.4138$$

Question 2

- Consider the BBN below. Given that the outcomes of medical test A and medical test B for a specific patient are positive and negative, respectively. That is, A = Pos and B = Neg. Predict the probability that the patient has high cholesterol.

$$P(C = \text{High} | A = \text{Pos}, B = \text{Neg})$$



$$P(C=\text{High}|A=\text{Pos},B=\text{Neg})$$

C is the parent node of both *A* and *B*

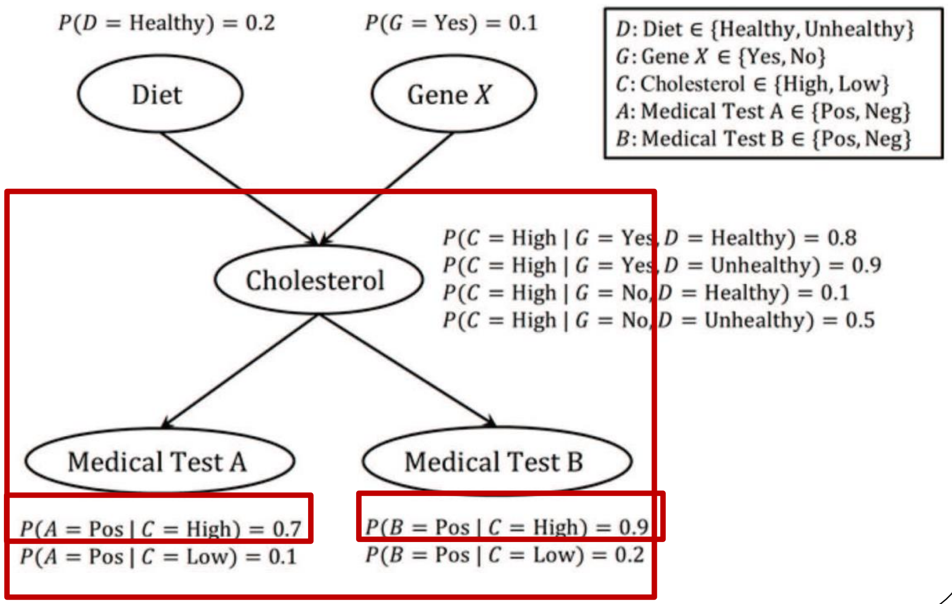
$$=\frac{P(A=\text{Pos},B=\text{Neg}|C=\text{High})P(C=\text{High})}{P(A=\text{Pos},B=\text{Neg})}$$

$$=\frac{P(A=\text{Pos}|C=\text{High})P(B=\text{Neg}|C=\text{High})P(C=\text{High})}{P(A=\text{Pos},B=\text{Neg})}$$

$$=\frac{0.7\times 0.1\times P(C=\text{High})}{P(A=\text{Pos},B=\text{Neg})}$$

$$P(B=\text{Pos}|C=\text{High})=0.9$$

$$P(B=\text{Neg}|C=\text{High})=0.1$$



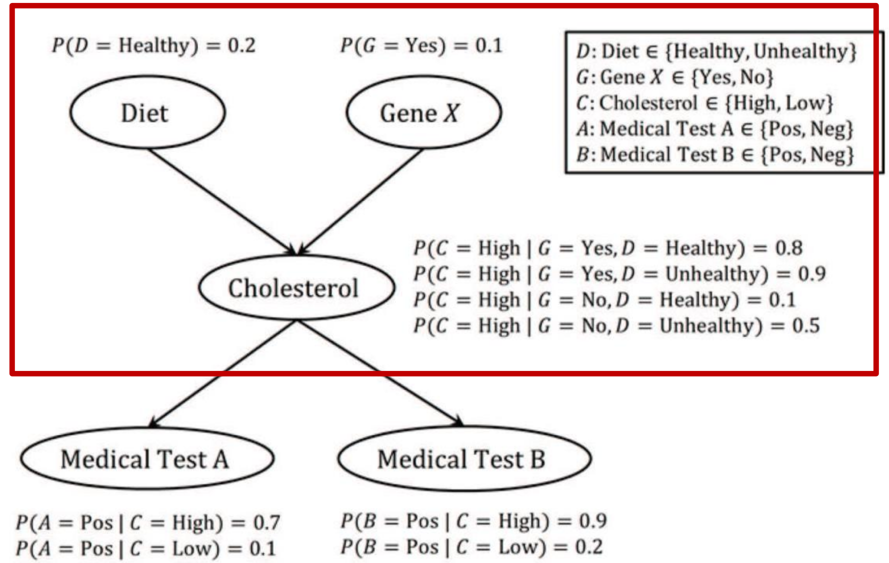
$$P(C = \text{High})$$

$$\alpha \in \{\text{Healthy, Unhealthy}\}$$
$$\beta \in \{\text{Yes, No}\}$$

$$= \sum_{\alpha,\beta} P(D = \alpha, G = \beta, C = \text{High})$$

$$= \sum_{\alpha,\beta} P(C = \text{High} \mid D = \alpha, G = \beta) \boxed{P(D = \alpha, G = \beta)}$$

$$= \sum_{\alpha,\beta} P(C = \text{High} \mid D = \alpha, G = \beta) \boxed{P(D = \alpha) P(G = \beta)}$$



$$P(C = \text{High}) = \sum_{\alpha,\beta} P(C = \text{High} \mid D = \alpha, G = \beta) P(D = \alpha) P(G = \beta)$$

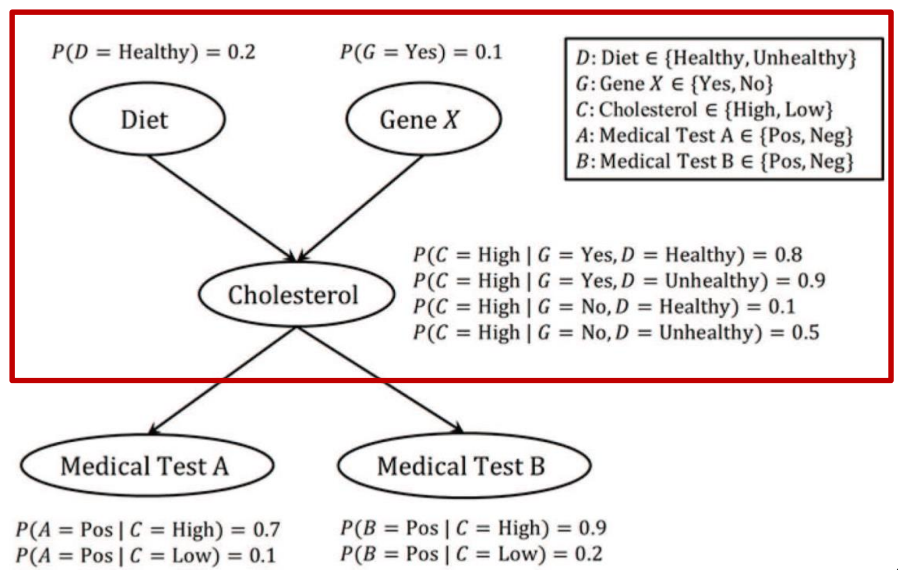
$$\begin{aligned}
&= P(C = \text{High} \mid D = \text{Healthy}, G = \text{Yes}) P(D = \text{Healthy}) P(G = \text{Yes}) \\
&\quad + P(C = \text{High} \mid D = \text{Healthy}, G = \text{No}) P(D = \text{Healthy}) P(G = \text{No}) \\
&\quad + P(C = \text{High} \mid D = \text{Unhealthy}, G = \text{Yes}) P(D = \text{Unhealthy}) P(G = \text{Yes}) \\
&\quad + P(C = \text{High} \mid D = \text{Unhealthy}, G = \text{No}) P(D = \text{Unhealthy}) P(G = \text{No}) \\
&= 0.8 \times 0.2 \times 0.1 + 0.1 \times 0.2 \times 0.9 + 0.9 \times 0.8 \times 0.1 + 0.5 \times 0.8 \times 0.9 \\
&= 0.47
\end{aligned}$$

Thus, $P(C = \text{High} \mid A = \text{Pos}, B = \text{Neg})$

$$= \frac{0.7 \times 0.1 \times P(C = \text{High})}{P(A = \text{Pos}, B = \text{Neg})}$$

$$= \frac{0.7 \times 0.1 \times 0.47}{P(A = \text{Pos}, B = \text{Neg})}$$

$$= \frac{0.0329}{P(A = \text{Pos}, B = \text{Neg})}$$



$$P(C = \text{Low} | A = \text{Pos}, B = \text{Neg})$$

$$P(C = \text{High}) = 0.47$$

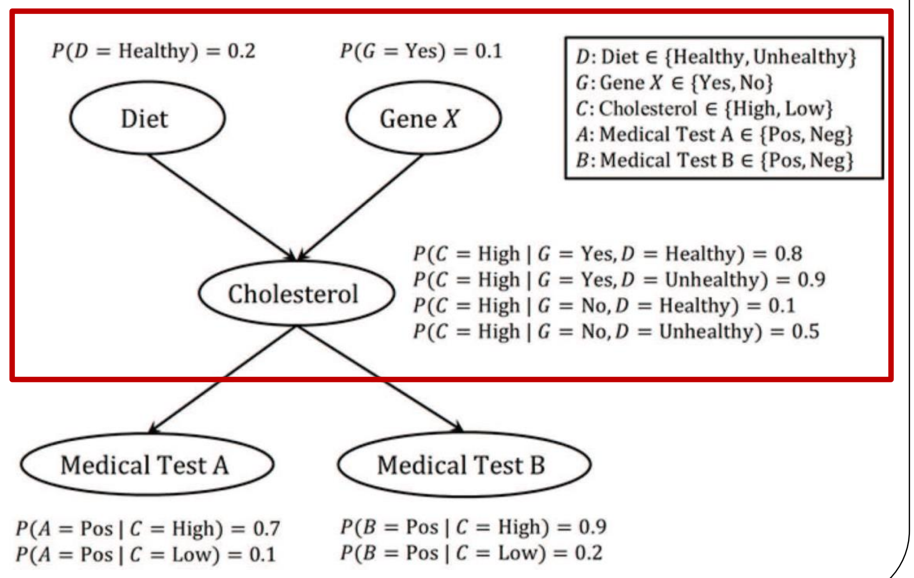
$$= \frac{P(A = \text{Pos}, B = \text{Neg} | C = \text{Low}) P(C = \text{Low})}{P(A = \text{Pos}, B = \text{Neg})}$$

$$P(C = \text{Low}) = 1 - 0.47 = 0.53$$

$$= \frac{P(A = \text{Pos} | C = \text{Low}) P(B = \text{Neg} | C = \text{Low}) \boxed{P(C = \text{Low})}}{P(A = \text{Pos}, B = \text{Neg})}$$

$$= \frac{0.1 \times 0.8 \times 0.53}{P(A = \text{Pos}, B = \text{Neg})}$$

$$= \frac{0.0424}{P(A = \text{Pos}, B = \text{Neg})}$$



$$P(C = \text{Low} | A = \text{Pos}, B = \text{Neg}) + P(C = \text{High} | A = \text{Pos}, B = \text{Neg}) = 1$$

$$\begin{array}{ccc} \| & & \| \\ \dfrac{0.0424}{P(A = \text{Pos}, B = \text{Neg})} & & \dfrac{0.0329}{P(A = \text{Pos}, B = \text{Neg})} \end{array}$$

$$\frac{0.0424 + 0.0329}{P(A = \text{Pos}, B = \text{Neg})} = 1 \quad \Rightarrow \quad P(A = \text{Pos}, B = \text{Neg}) = 0.0753$$

$$\text{Thus, } P(C = \text{High} | A = \text{Pos}, B = \text{Neg}) = \frac{0.0329}{0.0753} = 0.44$$

Thank you!