Probabilistic Graphical Models

Lectures 9

How to represent CPDs?

Table representation

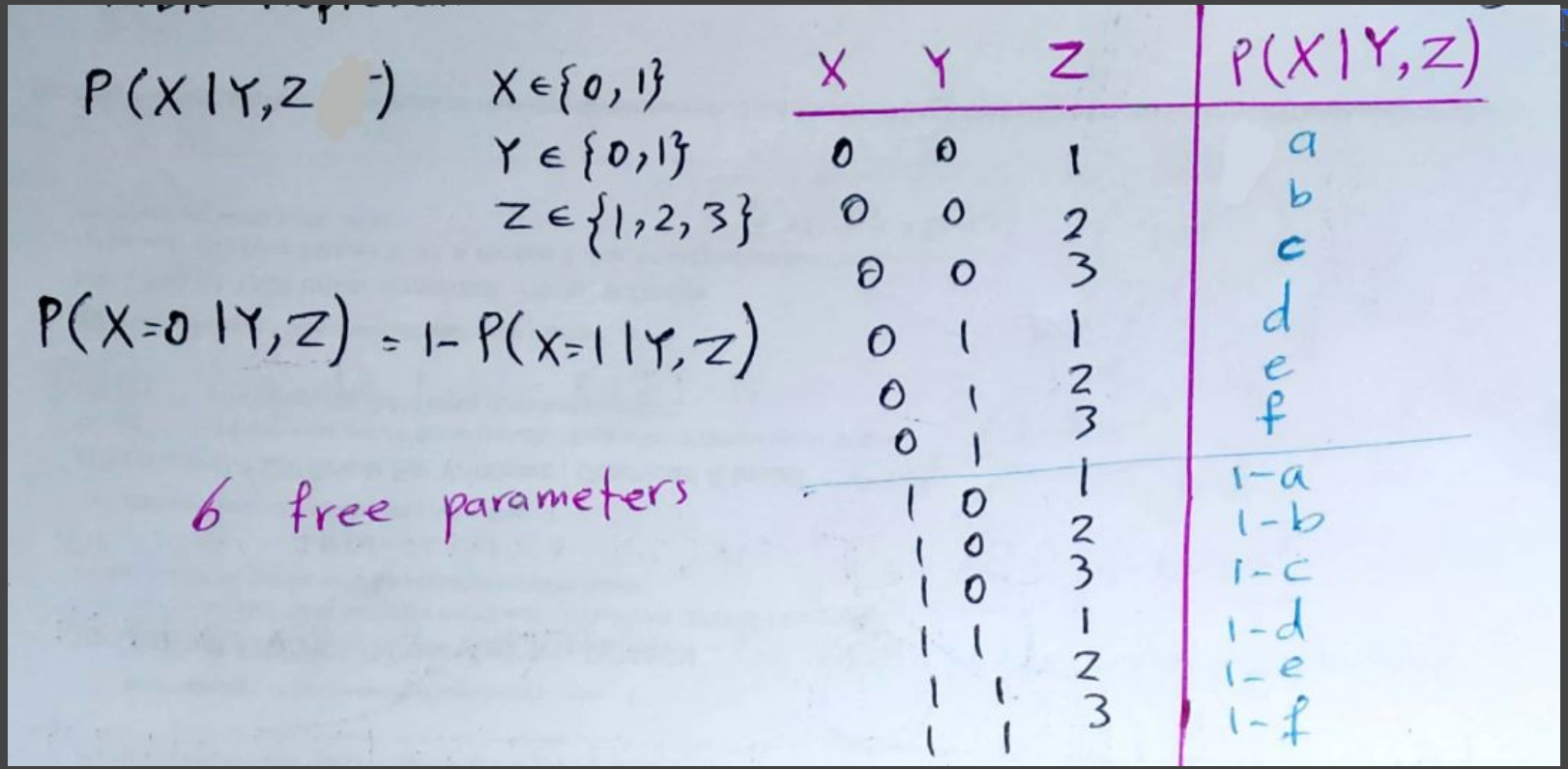
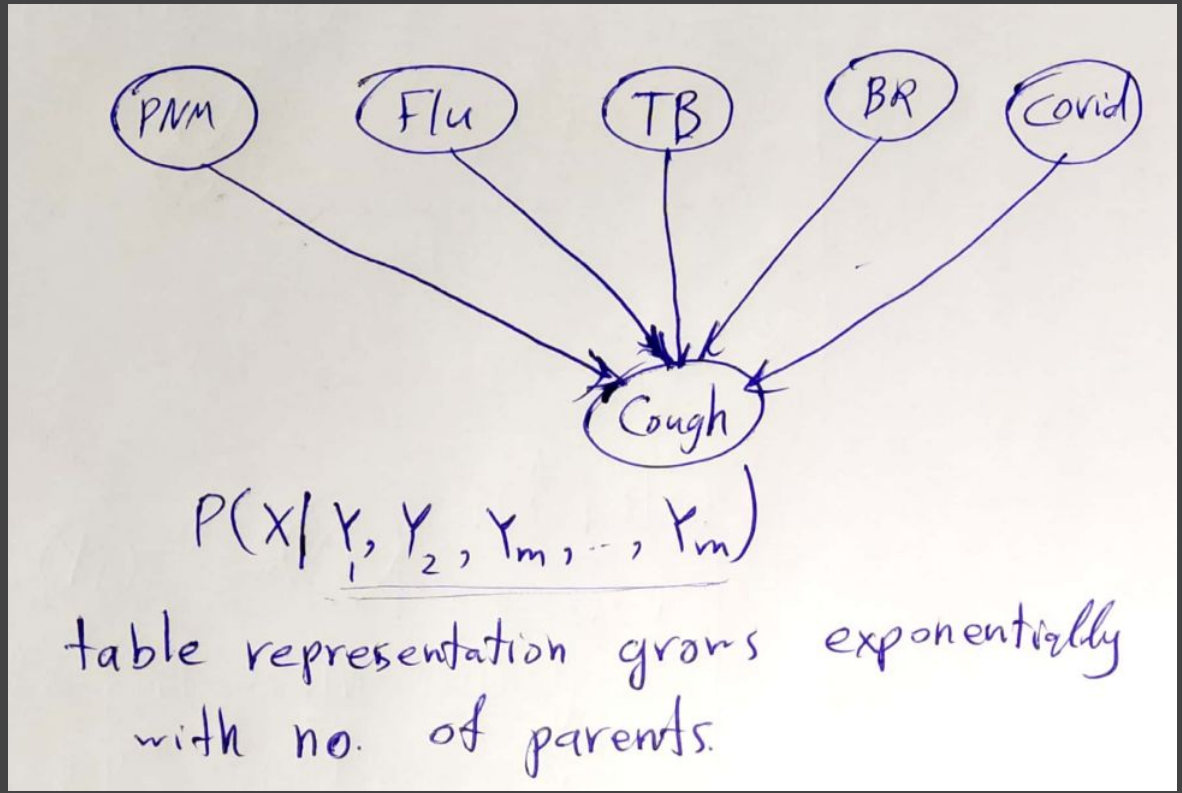
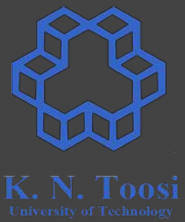


Table representation what's wrong?



K. N. Toosi
University of Technology

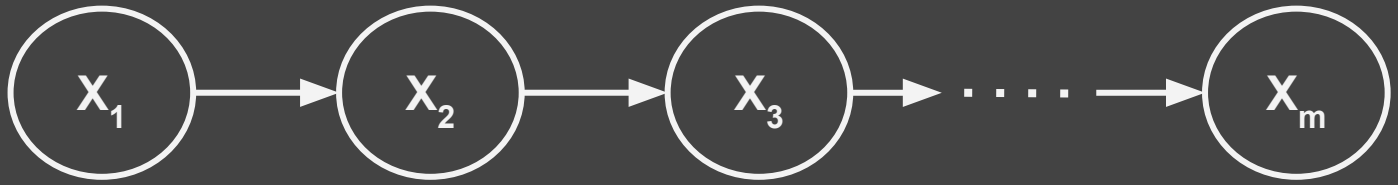
Table representation



- Many Parents
- Variables can get countably many values
- true model naturally needs less parameters
- Variables are continuous

Remember: faulty push-button

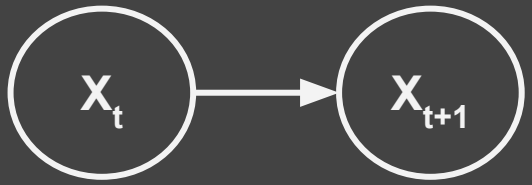
K. N. Toosi
University of Technology



$$p(x_1, x_2, \ldots, x_m) = p(x_1)\ p(x_2|x_1)\ p(x_3|x_2)\ \ldots\ p(x_m|x_{m-1})$$

Remember: faulty push-button

K. N. Toosi
University of Technology

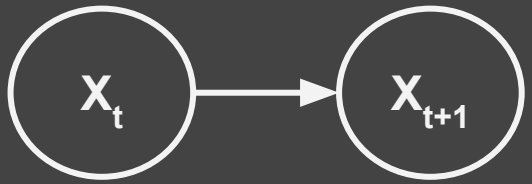


- works with probability p if device is on and with probability q if device is off
- works with probability p

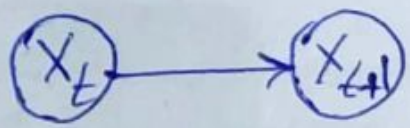
$p(x_{t+1} \mid x_t)$

Remember: faulty push-button

K. N. Toosi
University of Technology



works with probability p if device is on and with probability q if device is off



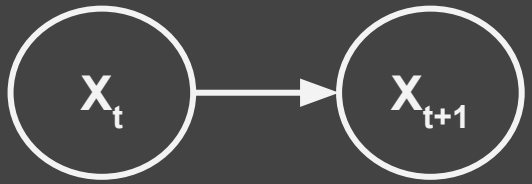
$P(X_{t+1} \mid X_t)$

2 free parameter

X_t	X_{t+1}	$P(X_t \mid X_{t-1})$
0	0	1-q
→ 0	1	q
→ 1	0	p
1	1	1-p

Remember: faulty push-button

K. N. Toosi
University of Technology



works with probability p

button work with probability p

1 free parameter

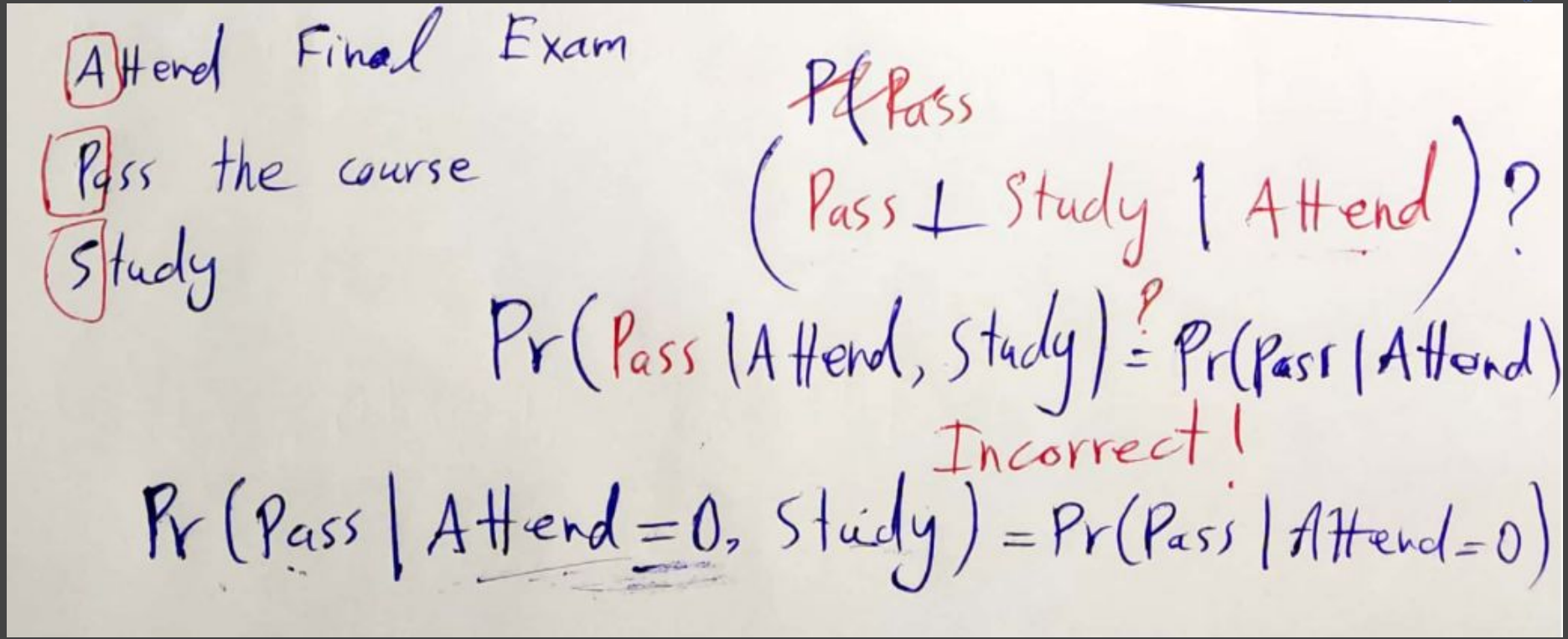
$P(X_{t+1} = x' \mid X_t = x)$

$= p\ 1(x \neq x') + (1-p)\ 1(x = x')$

↳ indicator function

X_t	X_{t+1}	$P(X_t \mid X_{t-1})$
0	0	1-p
0	1	p
1	0	p
1	1	1-p

Context specific independence



Context specific independence

K. N. Toosi
University of Technology

P: Pass the course
S: Study
A: Attend Final Exam

$P(\text{Pass} \mid \text{Study}, \text{Attend})$

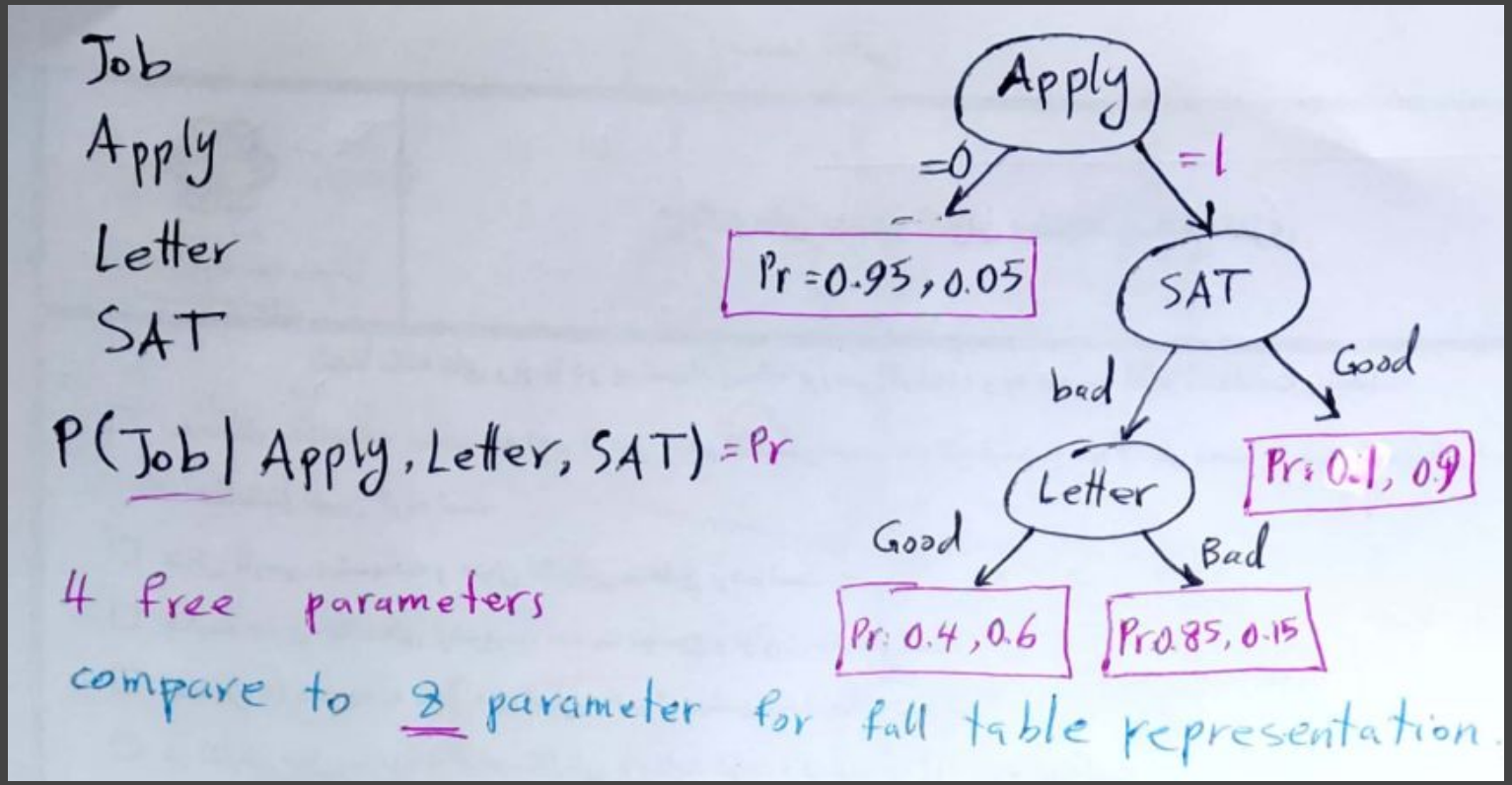
You will fail if not attend final exam

$P(\text{Pass} \mid \text{Study}, \text{Attend}) \stackrel{?}{=} P(\text{pass} \mid \text{Attend})$

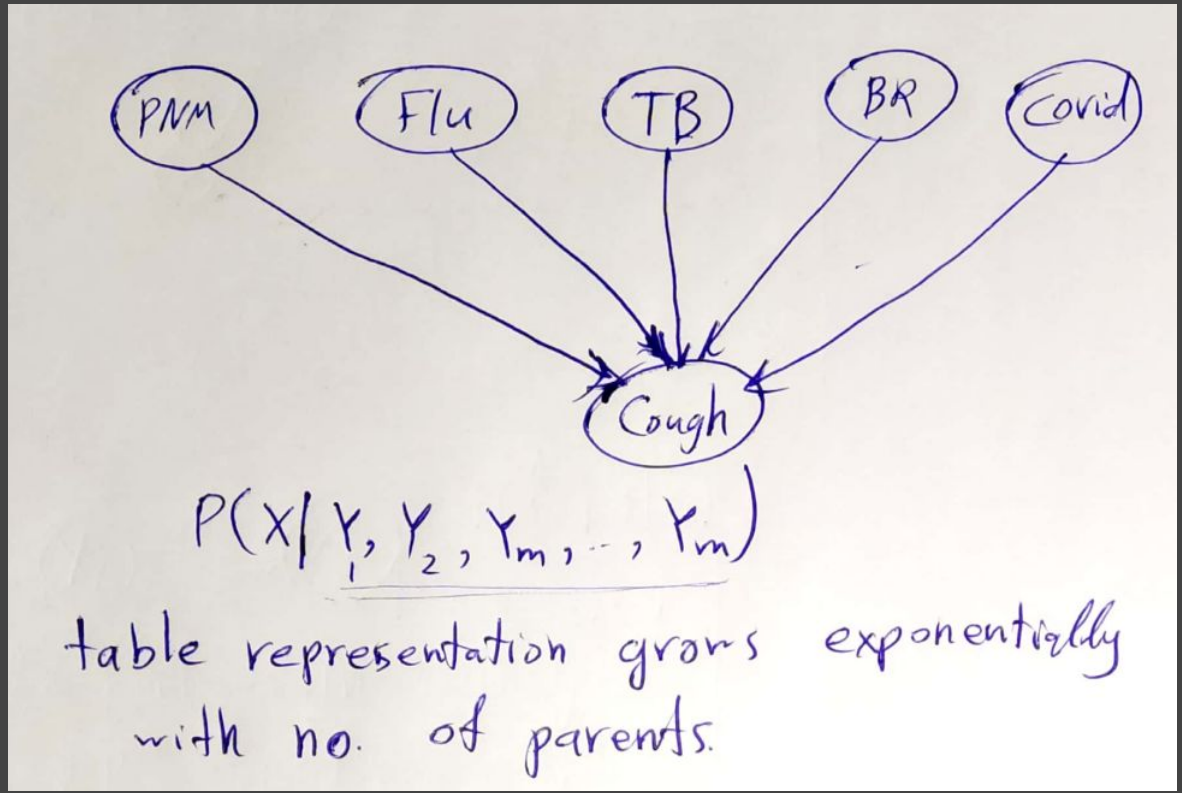
$P(\text{Pass} \mid \text{Study}, \text{Attend}=0) = P(\text{Pass} \mid \text{Attend}=0) = \mathbb{1}(\text{pass}=0)$

$P(\text{Pass} \mid \text{Study}, \text{Attend}=1) \neq P(\text{Pass} \mid \text{Attend}=1)$

Example (Koller)



What about many parents?



K. N. Toosi
University of Technology

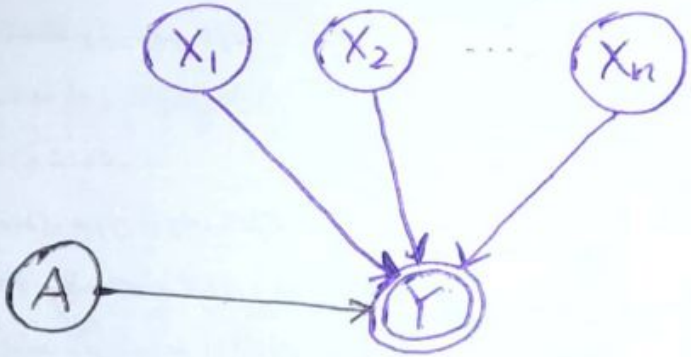
Multiplexer

Multiplexer

$X_i \in \{0, 1\}$

$Y \in \{0, 1\}$

$A \in \{1, 2, 3, \dots, n\}$

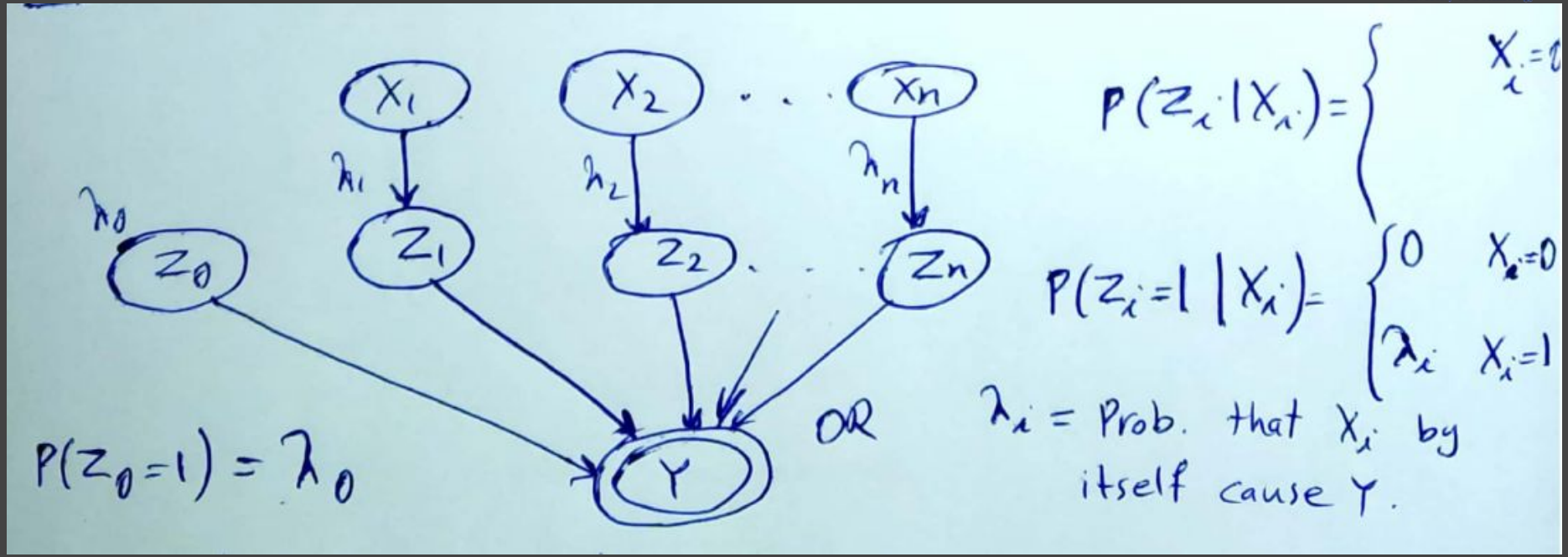


(O) means Y is deterministic given parents $(A, X_1 \cdots X_n)$

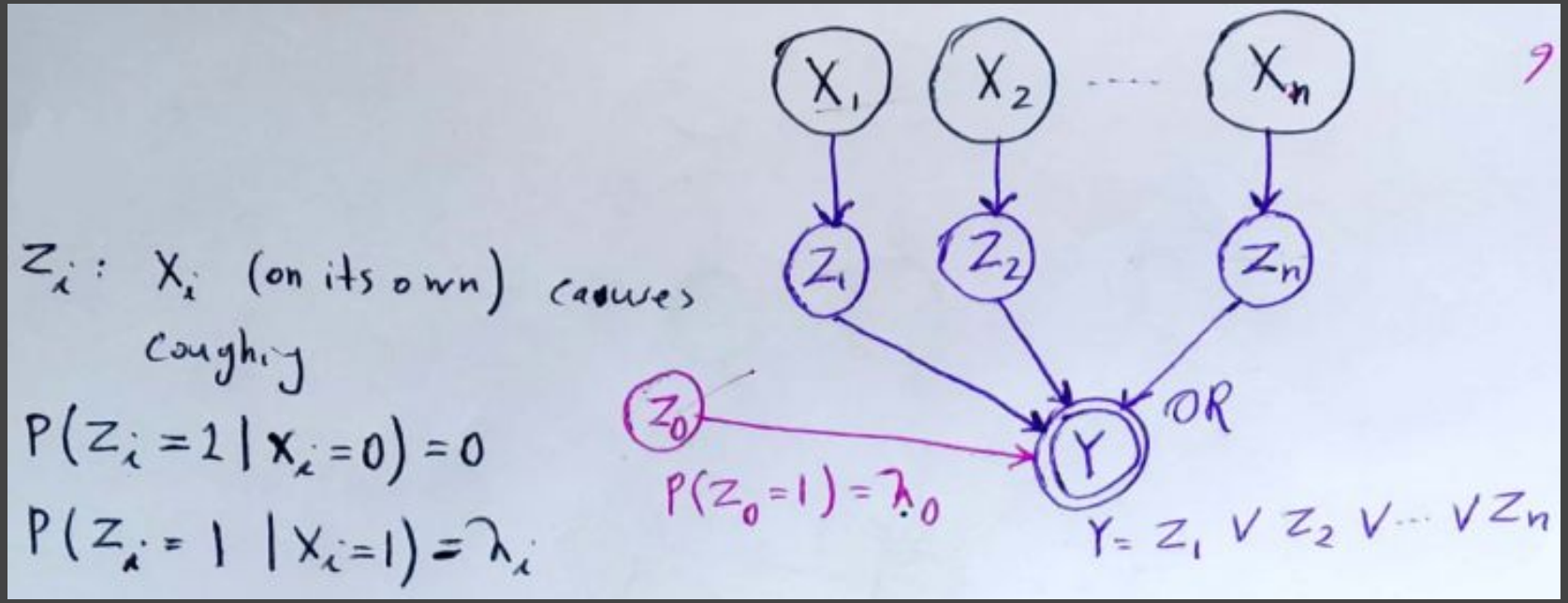
$$P(Y = 1 \mid A, X_1, X_2, \dots, X_n) = 1(X_A = 1)$$

$$Y = X_A \quad (\text{if } A = 2 \Rightarrow Y = X_2)$$

Noisy OR

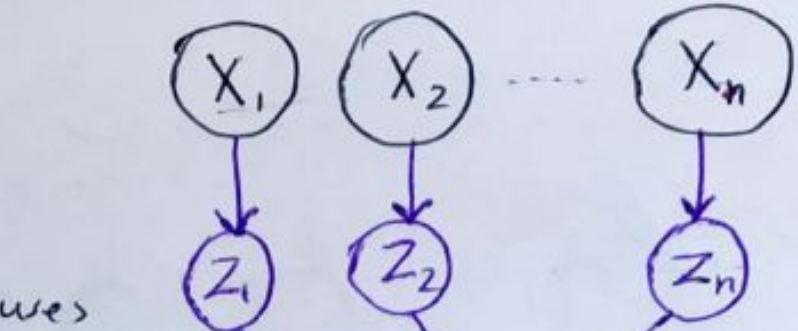


Noisy OR



How many parameters?

Noisy OR



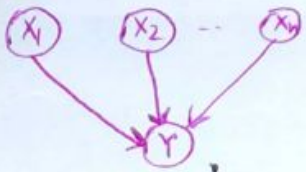
Z_i: X_i (on its own) causes coughing

$P(Z_i=1 \mid X_i=0)=0$

$P(Z_i=1 \mid X_i=1)=\lambda_i$

a total of $n+1$ parameters: $\lambda_0, \lambda_1, \ldots, \lambda_n$

Noisy OR



What is the original CPD

$P(Y \mid X_1, X_2, \dots, X_n)$

$$P(Y=0 \mid X_1 \dots X_n) = Pr(Z_1=0 \wedge Z_2=0, \dots Z_n=0, Z_0=0 \mid X_1 \dots X_n)$$

$$= Pr(Z_1=0 \mid X_1)\, Pr(Z_2=0 \mid X_2) \dots P(Z_n=0 \mid X_n)\, P(Z_0=0)$$

$$= (1-\lambda_1)^{1(X_1=1)} (1-\lambda_2)^{1(X_2=1)} \dots (1-\lambda_n)^{1(X_n=1)} (1-\lambda_0)$$

$$= \prod_{i=1}^{n} (1-\lambda_i)^{\underbrace{1(X_i=1)}_{=X_i}} (1-\lambda_0)$$

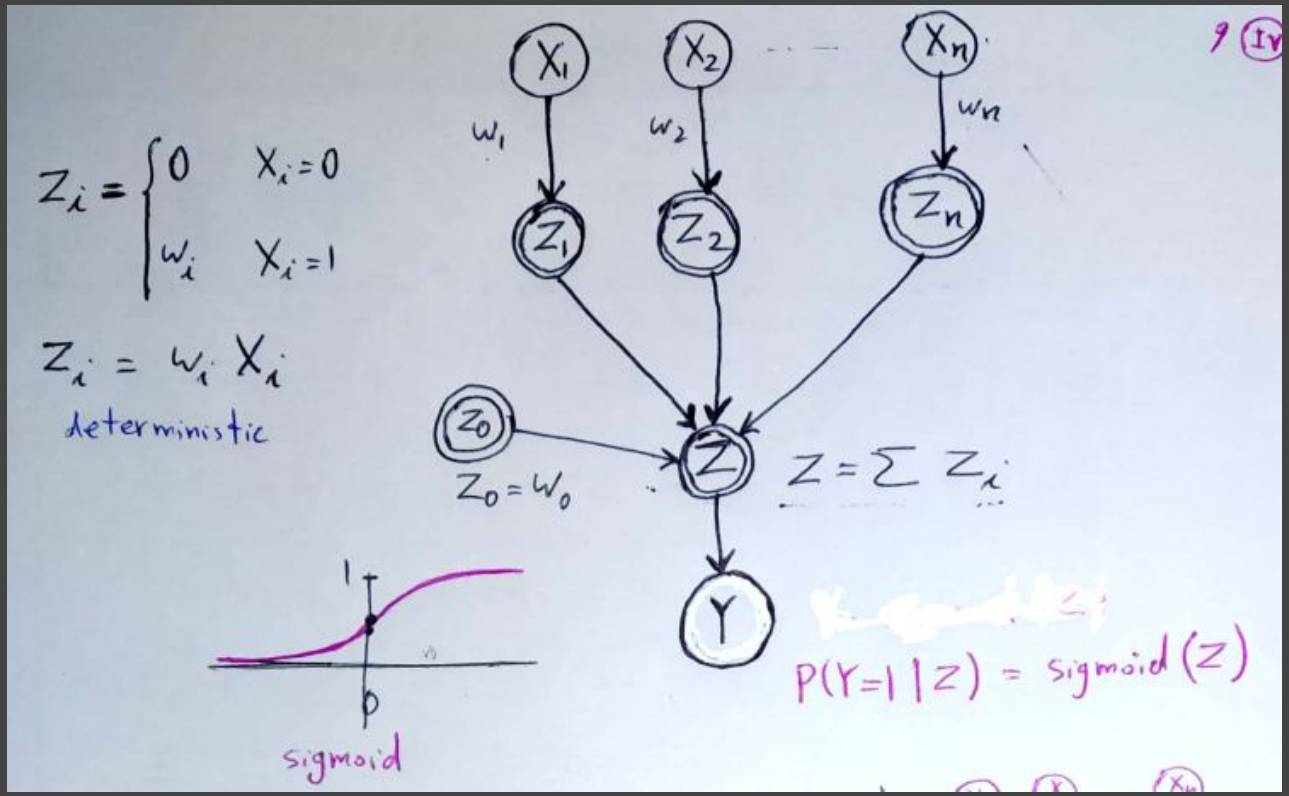
$$\Rightarrow P(Y=1 \mid X_1 \dots X_n) = 1 - (1-\lambda_0) \prod_{i=1}^{n} (1-\lambda_i)^{X_i}$$

↳ How many parameters?

$n+1$: $\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_n$

Noisy AND/MAX/Sum

Adding the effects up

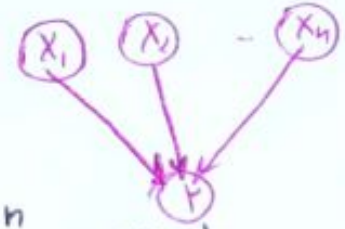


Adding the effects up

K. N. Toosi
University of Technology

sigmoid

what is the original CPD $P(Y \mid X_1, X_2, \dots, X_n)$?



$$P(Y=1 \mid X_1, X_2, \dots X_n) = \text{sigmoid}\left(w_0 + \sum_{i=1}^{n} w_i X_i\right)$$

How many parameters? $n+1 \rightarrow w_0, w_1, w_2, \dots, w_n$