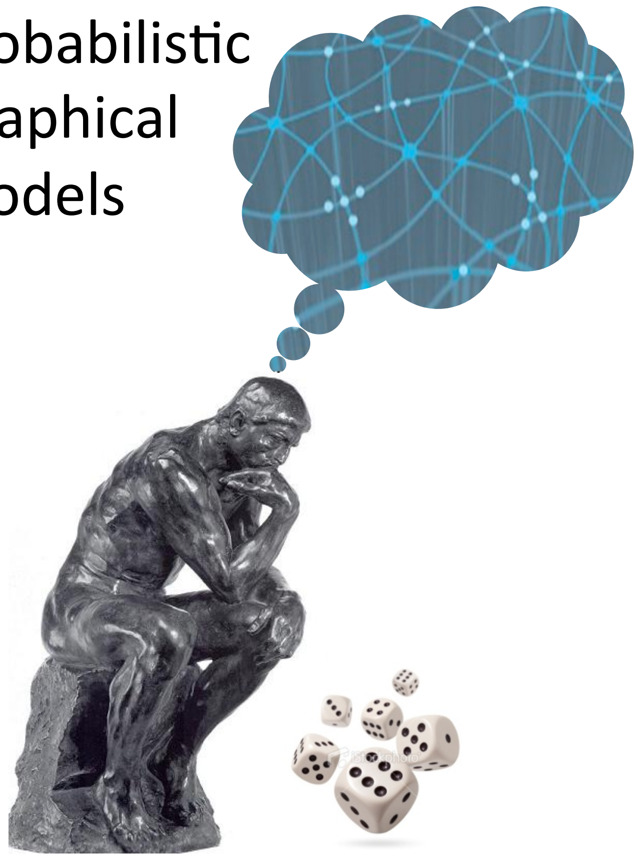
Probabilistic
Graphical
Models

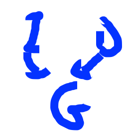


Representation

Local Structure

Overview

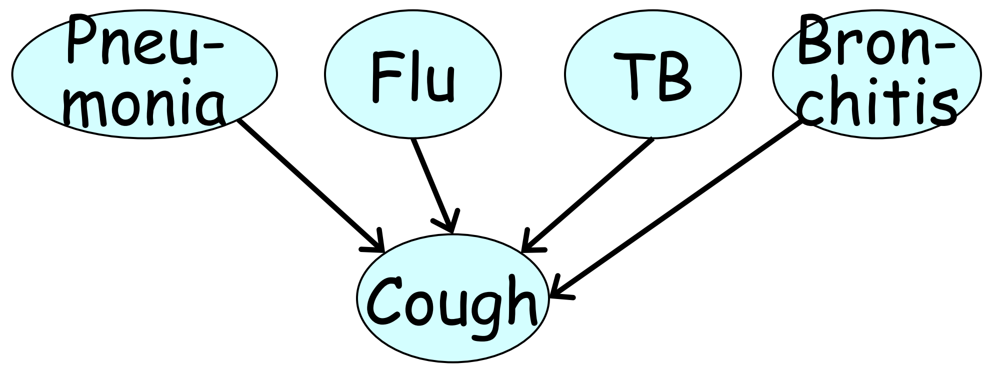
Tabular Representations



	g^1	g^2	g^3
i^0,d^0	0.3	0.4	0.3
i^0,d^1	0.05	0.25	0.7
i^1,d^0	0.9	0.08	0.02
i^1,d^1	0.5	0.3	0.2

k parents

$O(2^k)$ entries



General CPD

- CPD $P(X \mid Y_1, \ldots, Y_k)$ specifies distribution over X for each assignment $y_1, \ldots, y_k$
- Can use any function to specify a factor $\phi(X, Y_1, \ldots, Y_k)$ such that

$$\sum_x \phi(x, y_1, \ldots, y_k) = 1 \text{ for all } y_1, \ldots, y_k$$

Many Models

- Deterministic CPDs
- Tree-structured CPDs
- Logistic CPDs & generalizations
- Noisy OR / AND
- Linear Gaussians & generalizations

Context-Specific Independence

$$P \models (X \perp_c Y \mid \boldsymbol{Z}, \boldsymbol{c})$$

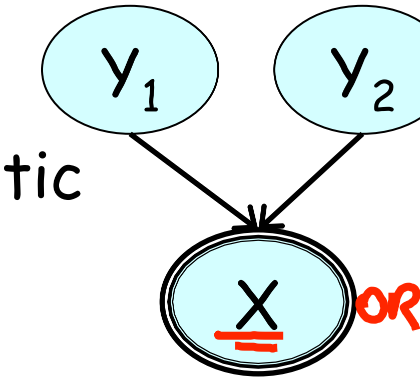
assignment to C

$$P(\boldsymbol{X}, \boldsymbol{Y} \mid \boldsymbol{Z}, \boldsymbol{c}) = P(\boldsymbol{X} \mid \boldsymbol{Z}, \boldsymbol{c}) P(\boldsymbol{y} \mid \boldsymbol{Z}, \boldsymbol{c})$$

$$P(\boldsymbol{X} \mid \boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{c}) = P(\boldsymbol{X} \mid \boldsymbol{Z}, \boldsymbol{c})$$

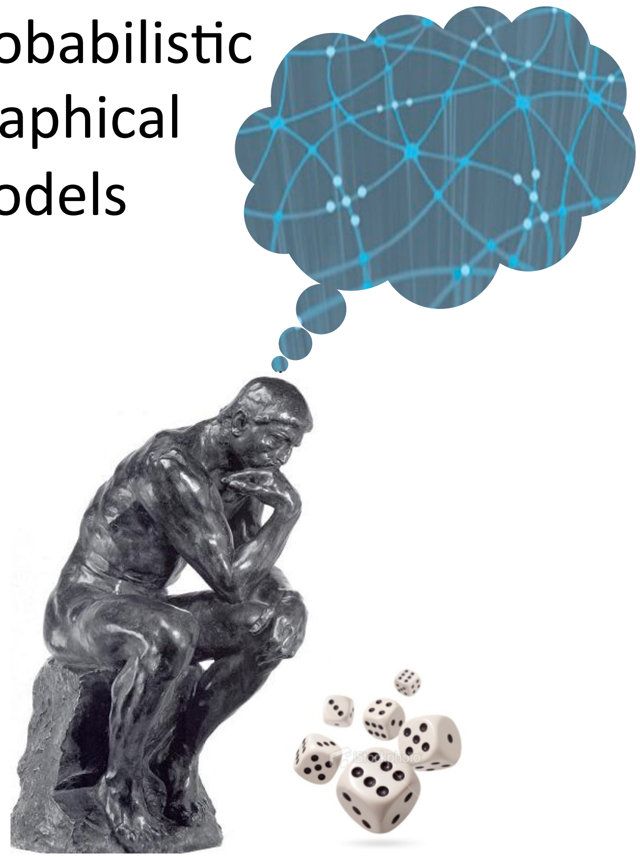
$$P(\boldsymbol{Y} \mid \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{c}) = P(\boldsymbol{Y} \mid \boldsymbol{Z}, \boldsymbol{c})$$

Which of the following context-specific independences hold when X is a deterministic OR of Y_1 and Y_2? (Mark all that apply.)



- [ ] $(X \perp Y_1 \mid y_2^0)$
- [x] $(X \perp Y_1 \mid y_2^1)$
- [x] $(Y_1 \perp Y_2 \mid x^0)$
- [ ] $(Y_1 \perp Y_2 \mid x^1)$

Probabilistic
Graphical
Models

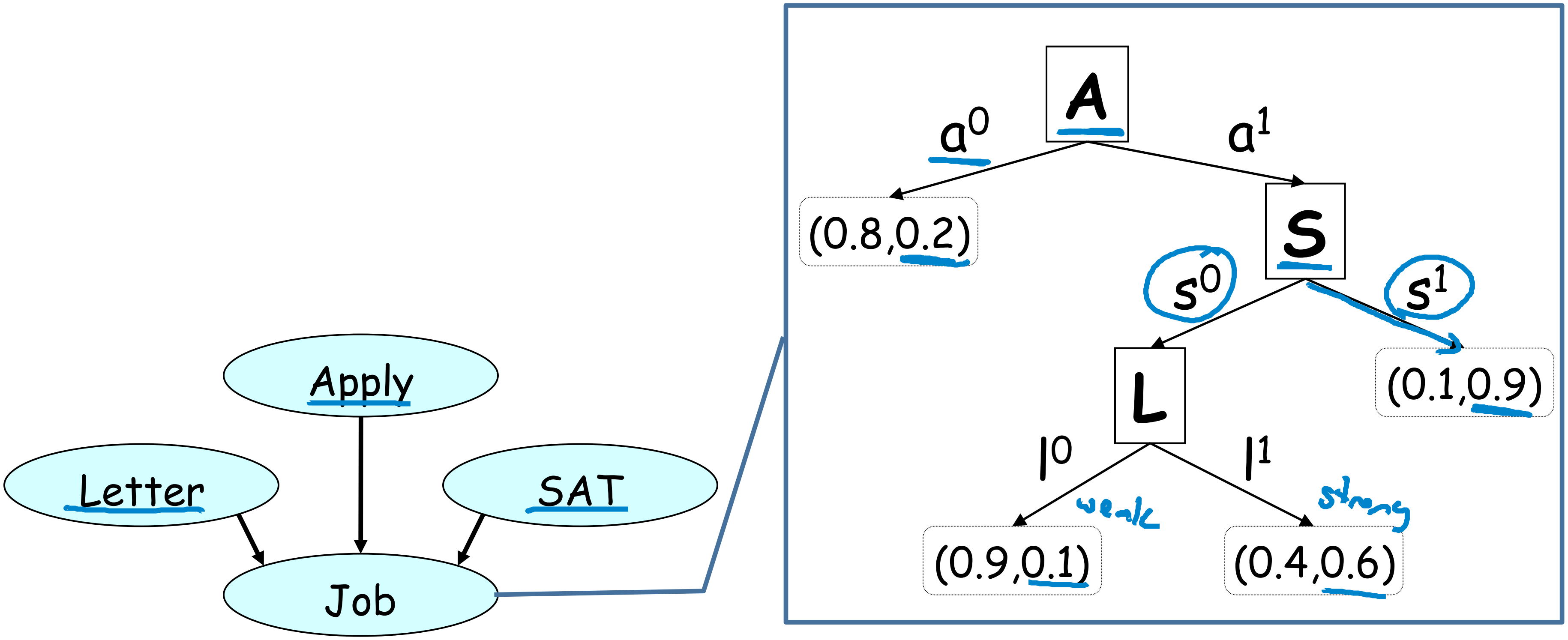


Representation

Local Structure

Tree-structured CPDs

Tree CPD



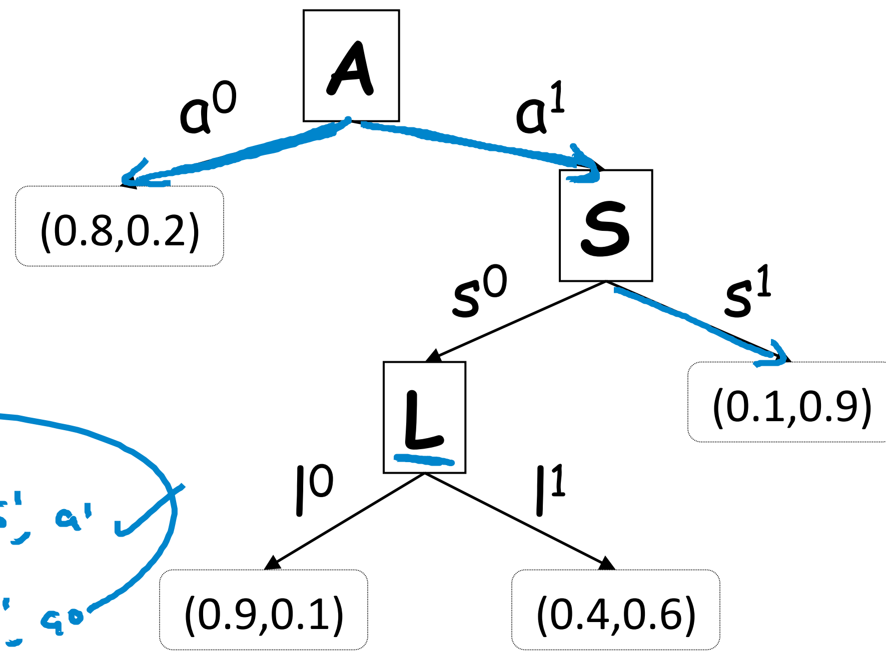
Daphne Koller

Which context-specific independencies are implied by the structure of this CPD? (Mark all that apply.)

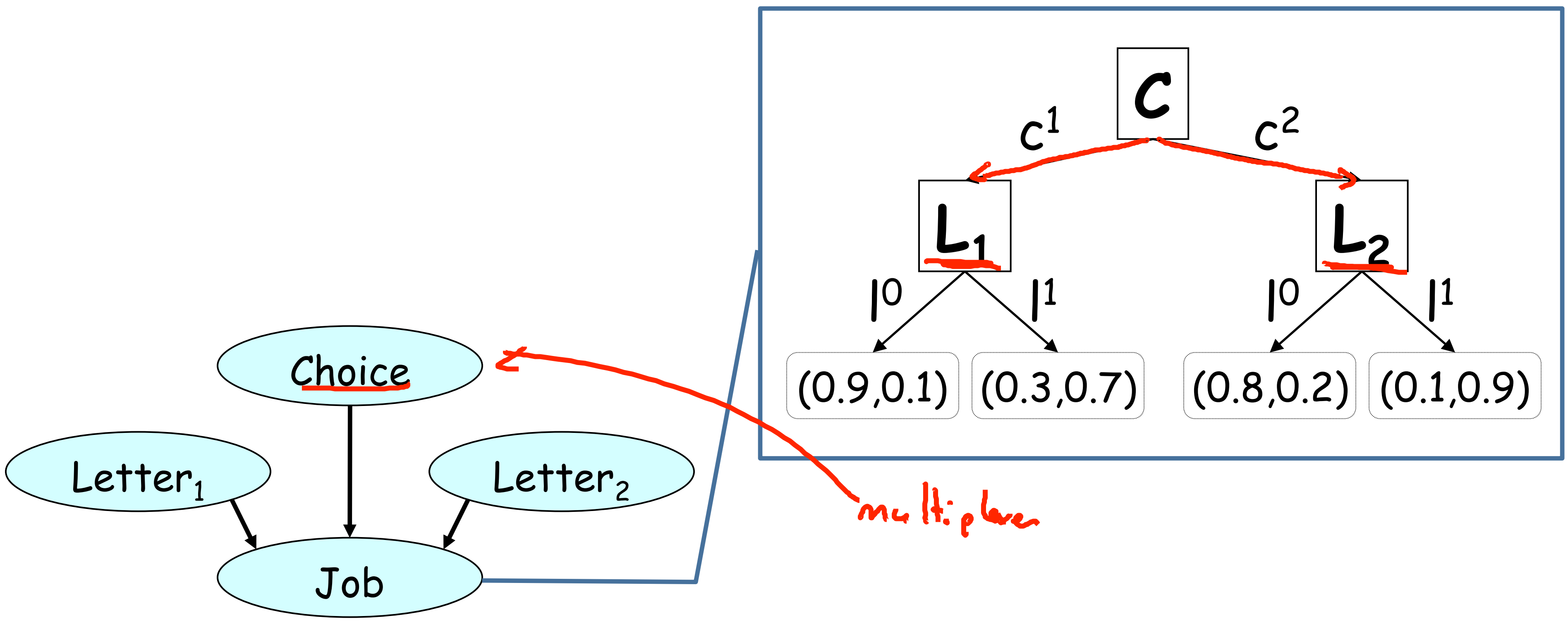
- [x] $(J \perp_c L \mid a^1, s^1)$
- [ ] $(J \perp_c L \mid a^1)$
- [x] $(J \perp_c L, S \mid a^0)$
- [x] $(J \perp_c L \mid s^1, A)$

$J \perp_c L \mid s^1, a^1$ ✓

$J \perp_c L \mid s^1, a^0$



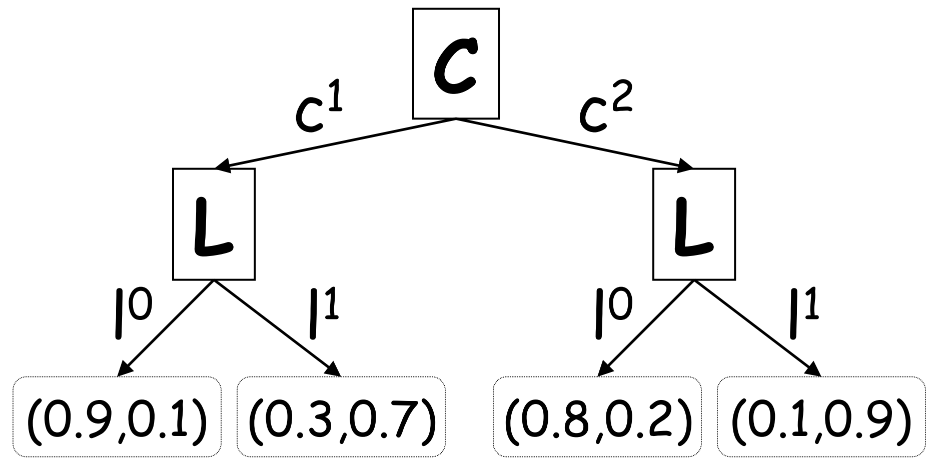
Tree CPD

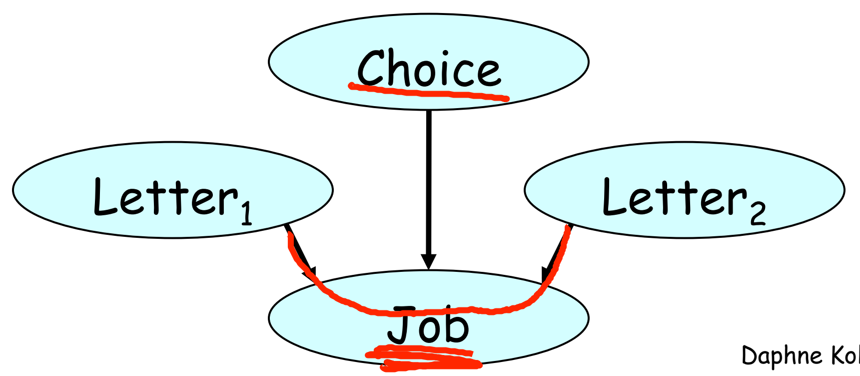


Daphne Koller

$$(L_1 \perp L_2 \mid J, C)$$

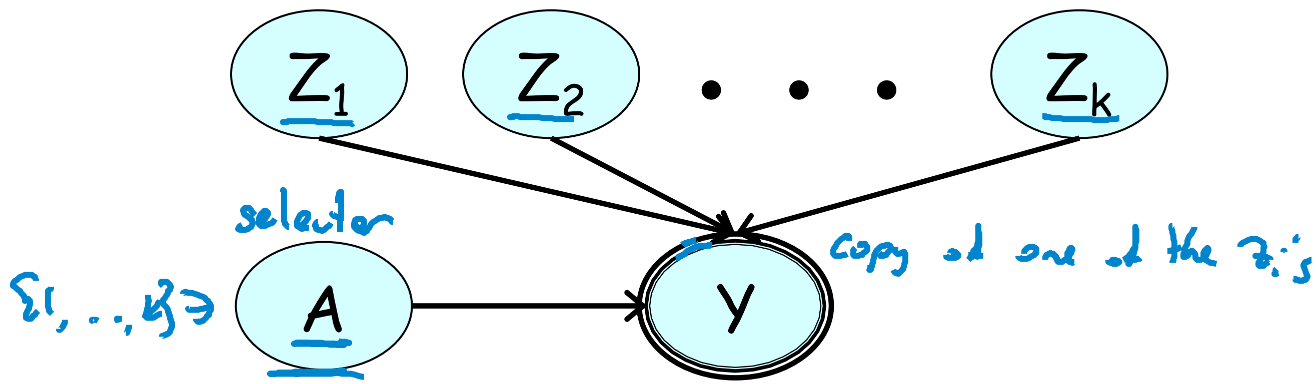
$$(L_1 \perp_c L_2 \mid J, c_1)$$
$$(L_1 \perp_c L_2 \mid J, c_2)$$





Daphne Koller

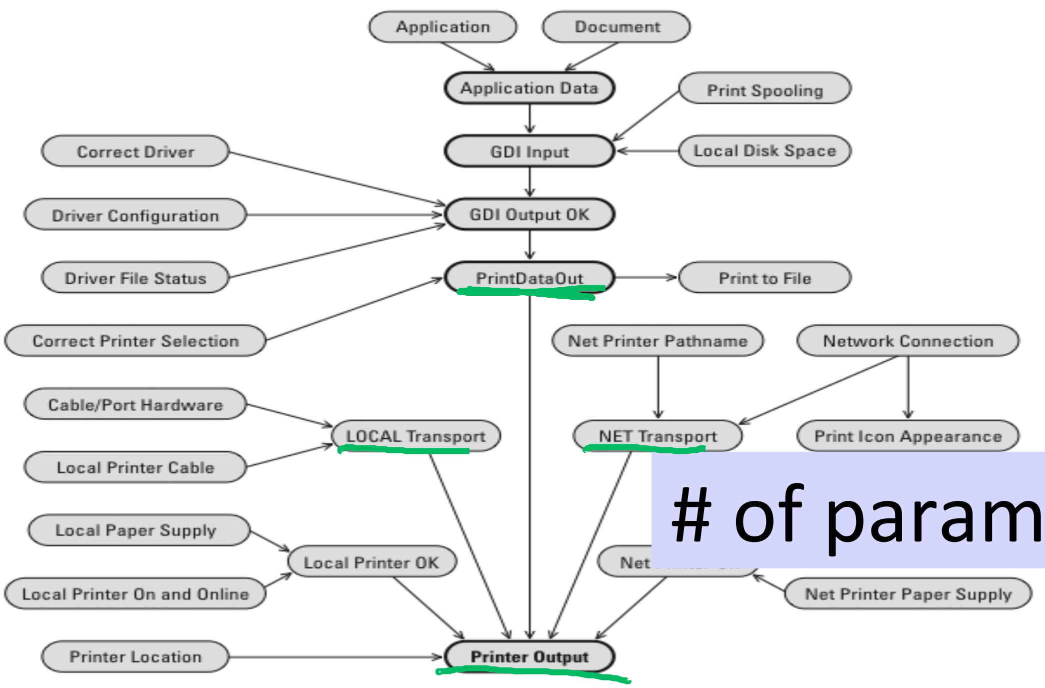
Multiplexer CPD

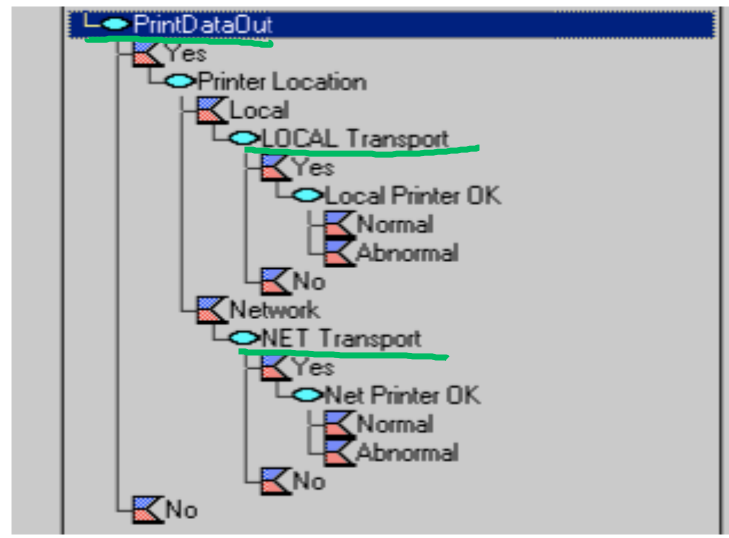


$A = a \Rightarrow Y = Z_a$

$$P(Y \mid A, Z_1, \ldots, Z_k) = \begin{cases} 1 & Y = Z_A \\ 0 & \text{otherwise} \end{cases}$$

Microsoft Troubleshooters





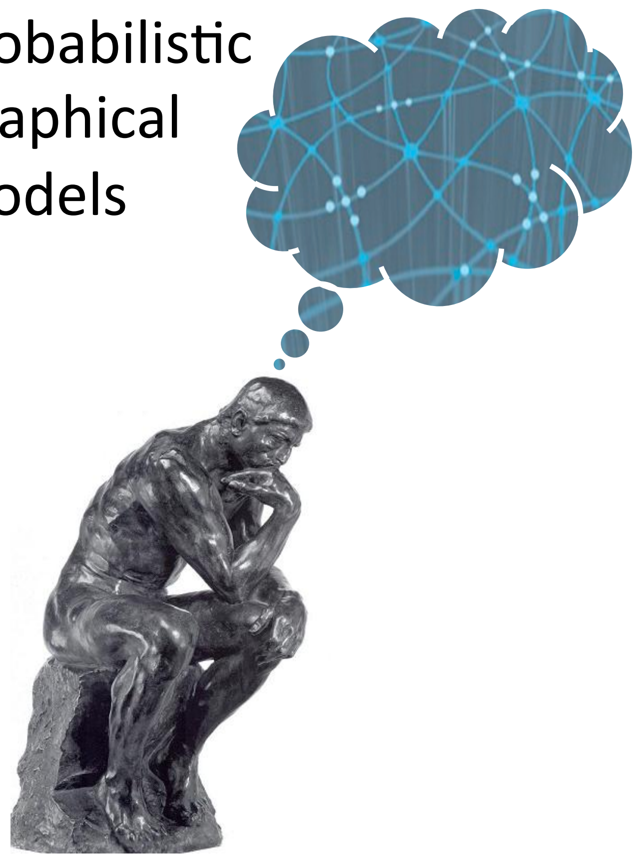
# of parameters: 145 to 55

Daphne Koller

Summary

- Compact CPD representation that captures context-specific dependencies
- Relevant in multiple applications:
  - Hardware configuration variables
  - Medical settings
  - Dependence on agent's action
  - Perceptual ambiguity

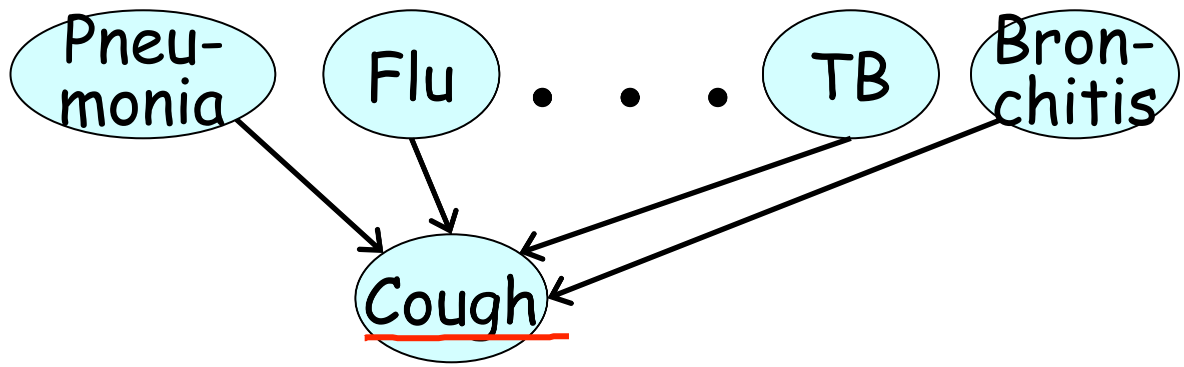
Probabilistic
Graphical
Models



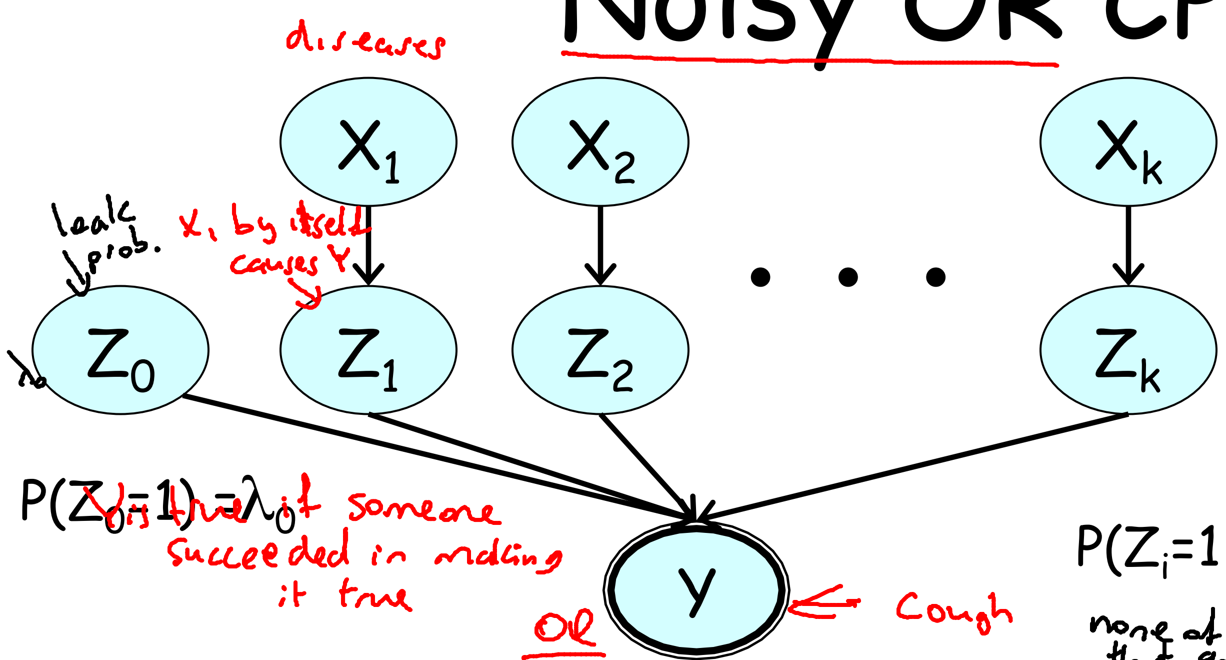
Representation

Local Structure

Independence of Causal Influence


Pneu-
monia
Flu
. . .
TB
Bron-
chitis
Cough

Noisy OR CPD



$P(Z_0=1) = \lambda_0$

$$P(Z_i=1 \mid X_i) = \begin{cases} 0 & X_i=0 \\ \lambda_i & X_i=1 \end{cases}$$

$\lambda_i \in [0,1]$

none of causes that are on fail

$$P(Y=0 \mid X_1,\ldots,X_k) = (1-\lambda_0) \prod_{i:X_i=1} (1-\lambda_i)$$

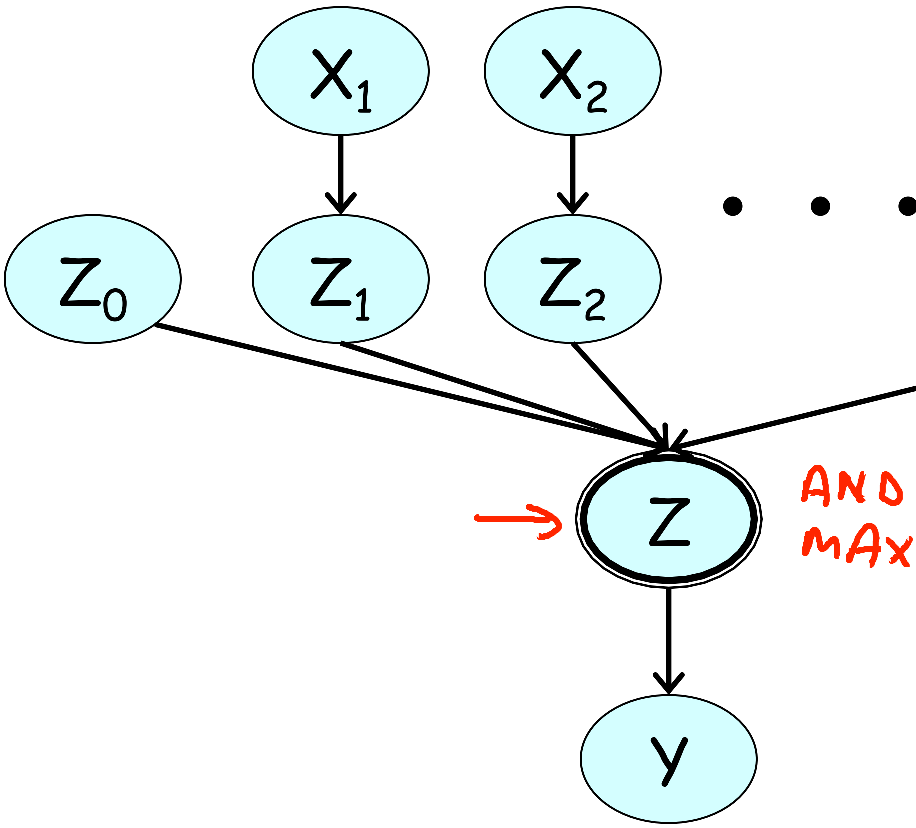
$$P(Y=1 \mid X_1,\ldots,X_k) = 1 - P(Y=0 \mid X_1,\ldots,X_k)$$

Daphne Koller

What context-specific independencies are induced by a noisy OR CPD?

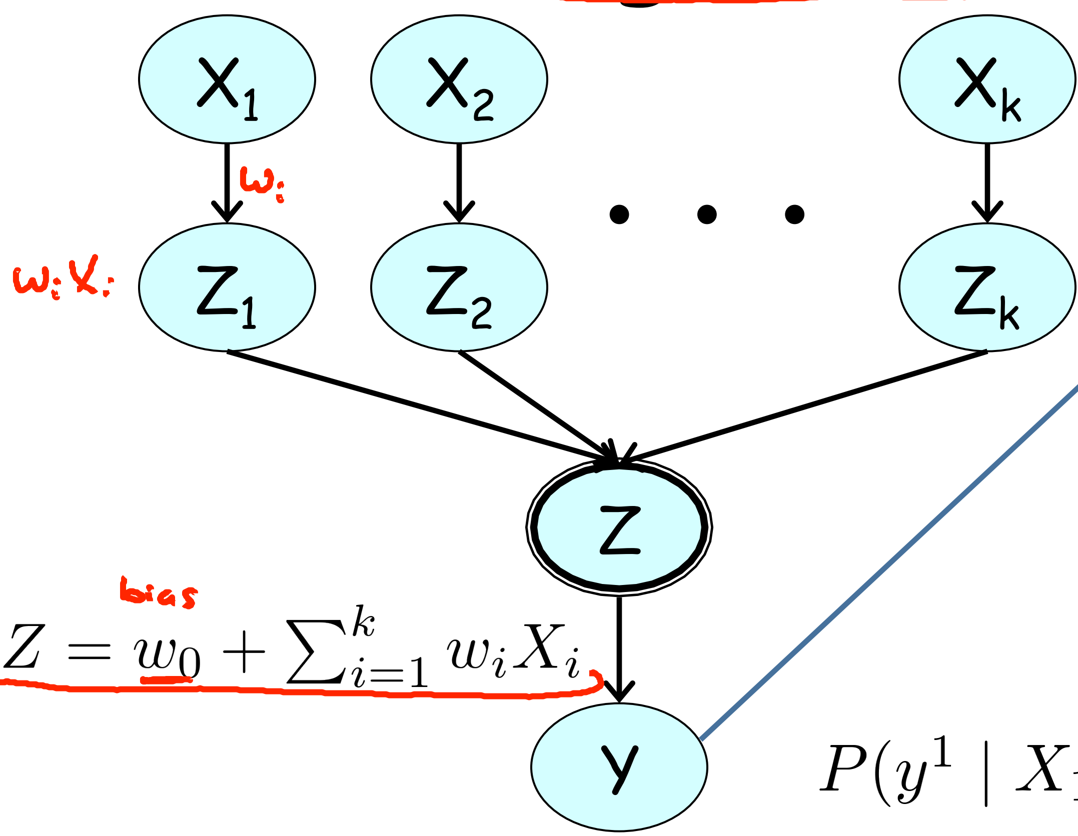
- $(Y \perp_c X_2 \mid x_1^1)$
- $(X_1 \perp_c X_2 \mid y^1)$
- $(X_1 \perp_c X_2 \mid y^0)$
- A noisy OR CPD induces no context-specific independencies

Independence of Causal Influence



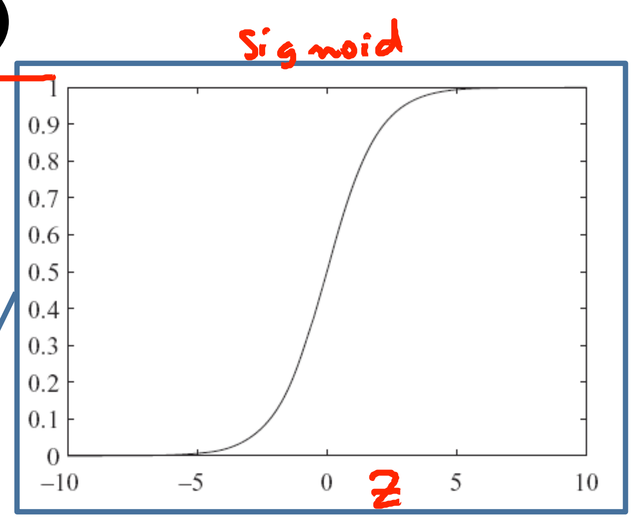
Daphne Koller

Sigmoid CPD



bias

$$Z = w_0 + \sum_{i=1}^{k} w_i X_i$$

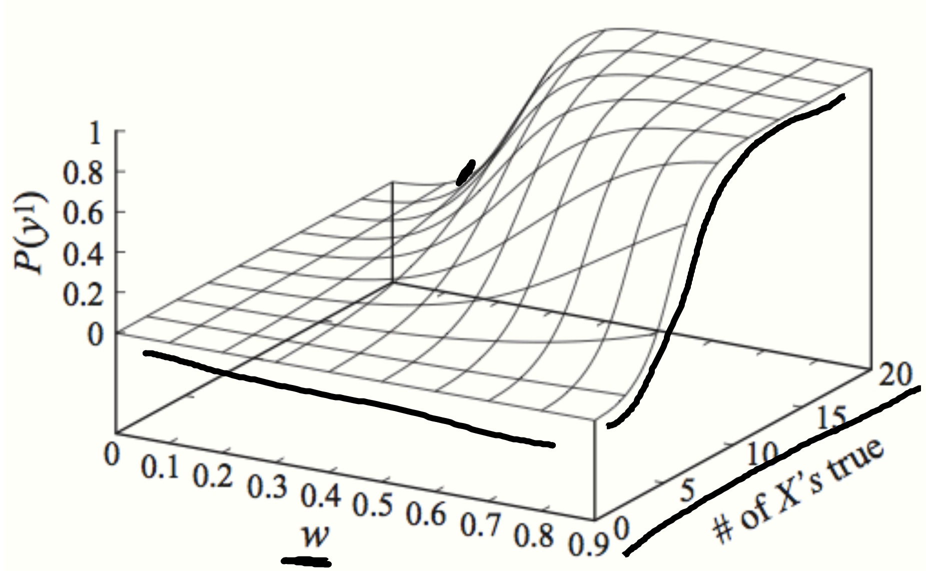


$$\text{sigmoid}(z) = \frac{e^z}{1 + e^z} \quad (0,1)$$

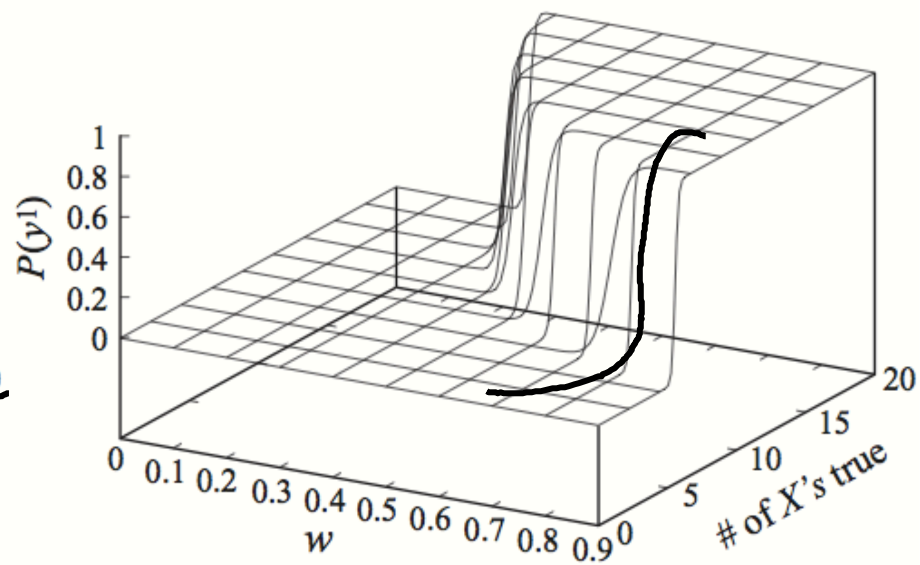
$$P(y^1 \mid X_1, \ldots, X_k) = \text{sigmoid}(Z)$$

Daphne Koller

Behavior of Sigmoid CPD



w_0 = -5



multiply w and w_0 by 10

Daphne Koller

$$P(y^1 \mid X_1, \ldots, X_k) = \text{sigmoid}(w_0 + \sum_{i=1}^{k} w_i X_i)$$

The odds ratio of Y is: $O(\boldsymbol{x}) = \frac{P(y^1|\boldsymbol{x})}{P(y^0|\boldsymbol{x})}$

It captures the relative likelihood of the two values of Y

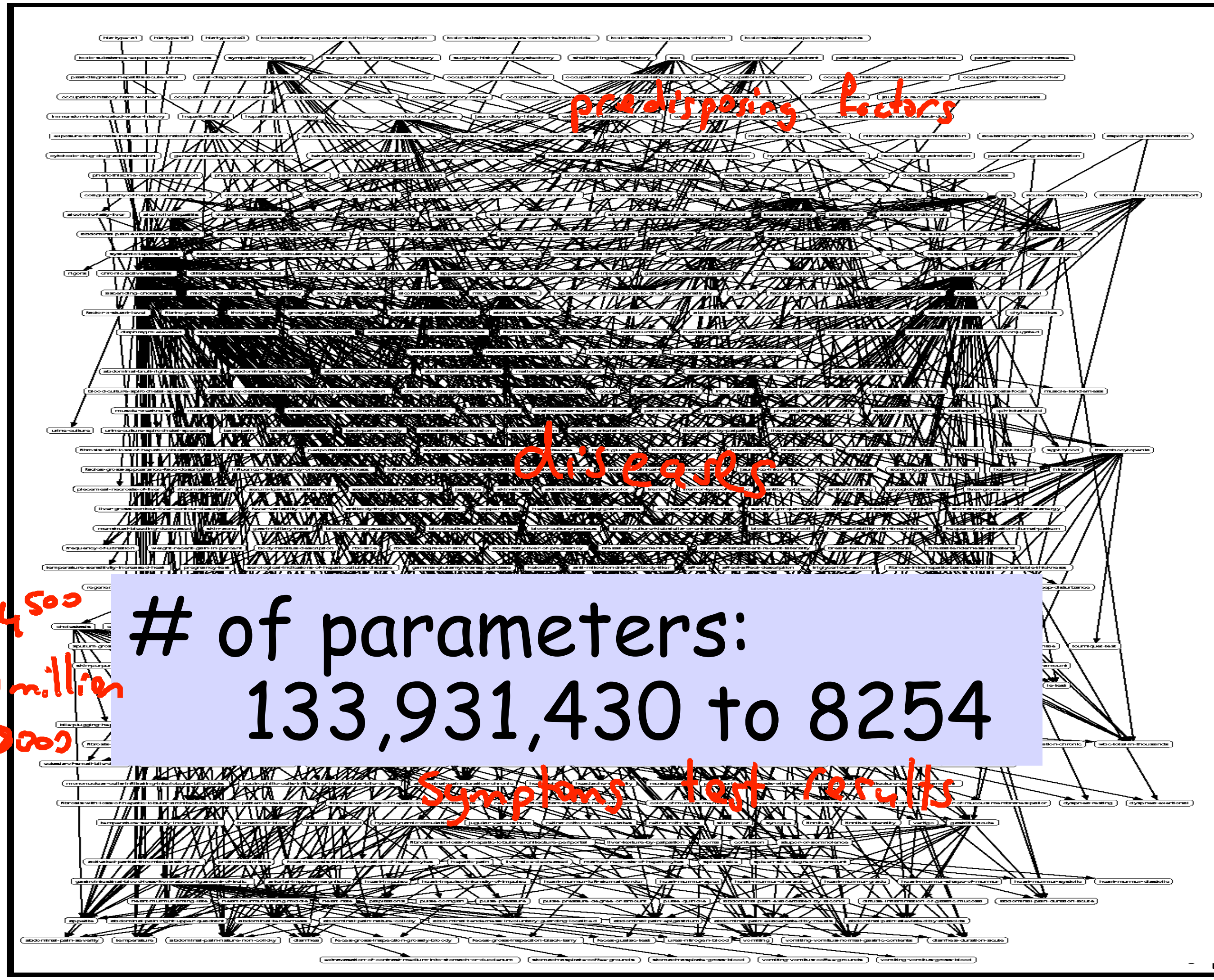
By what factor does O(**x**) change if the value of X_i goes from 0 to 1?

- $e^{w_i}/(1 + e_i^{w})$
- w_i
- e^{w_i}
- It depends on the values of the other X_i's

CPCS

M. Pradhan
G. Provan
B. Middleton
M. Henrion
UAI 1994

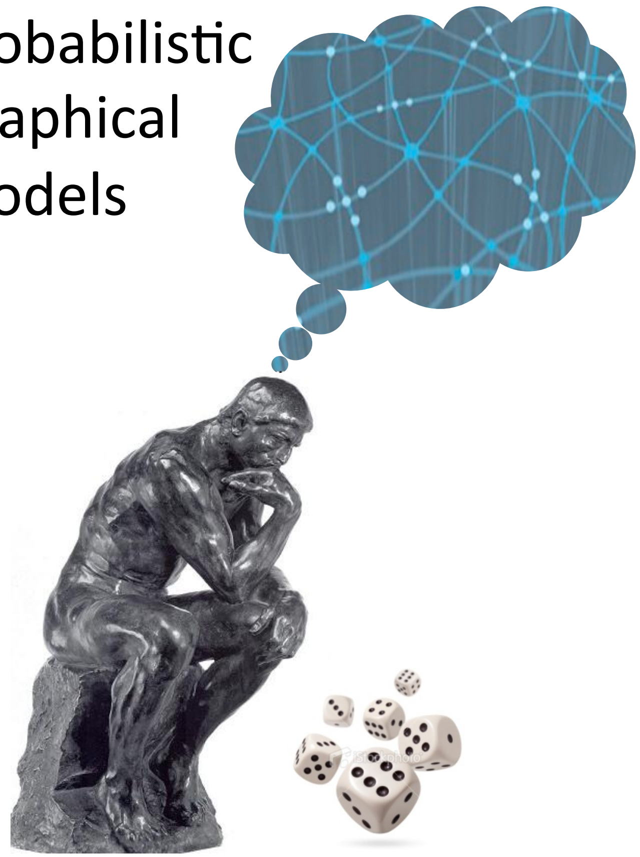
joint dist $\approx 4^{500}$
factorized $\approx$ 134 million
noisy max CPD $\approx$ 8000



# of parameters:
133,931,430 to 8254

hne Koller

Probabilistic
Graphical
Models

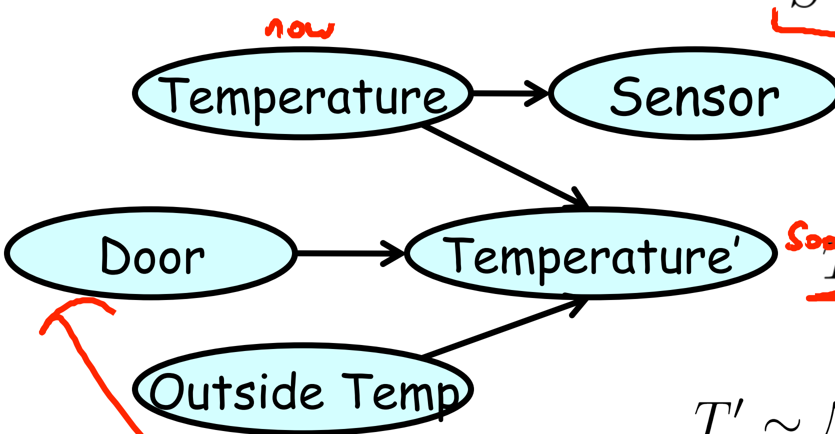


Representation

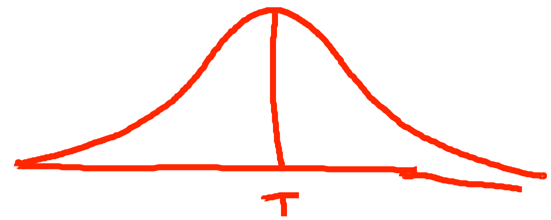
Local Structure

Continuous Variables

Continuous Variables



$$S \sim \mathcal{N}\left(T; \sigma_S^2\right)$$



linear Gaussian

soon $$T' \sim \mathcal{N}\left(\alpha T + (1-\alpha)O; \sigma_T^2\right)$$

Gaussian around

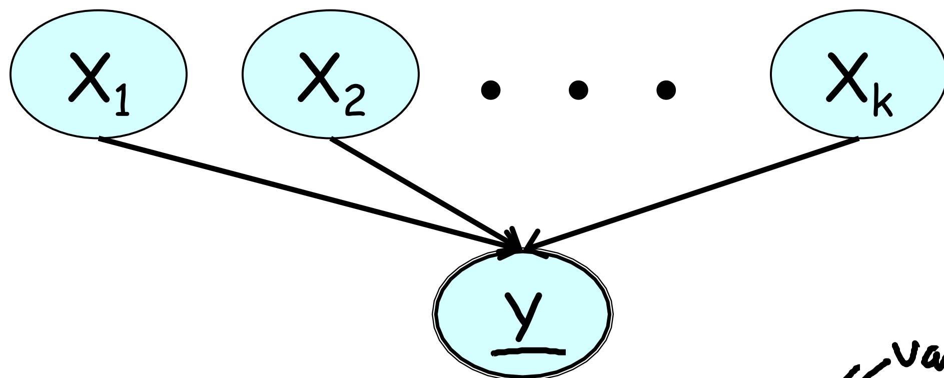
$$T' \sim \mathcal{N}\left(\alpha_0 T + (1-\alpha_0)O; \sigma_{0T}^2\right) \quad \text{if D=0}$$

$$T' \sim \mathcal{N}\left(\alpha_1 T + (1-\alpha_1)O; \sigma_{1T}^2\right) \quad \text{if D=1}$$

conditional linear Gaussian

Daphne Koller

Linear Gaussian



$$Y \sim \mathcal{N}\left(\underbrace{w_0 + \sum w_i X_i}_{\text{mean}}; \sigma^2\right)$$

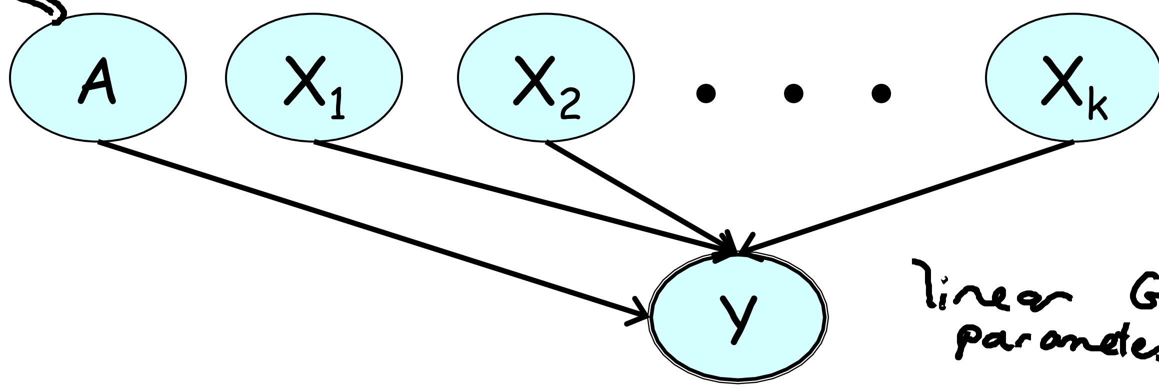
variance doesn't depend on parents

linear function of parents

Daphne Koller

Conditional Linear Gaussian

discrete

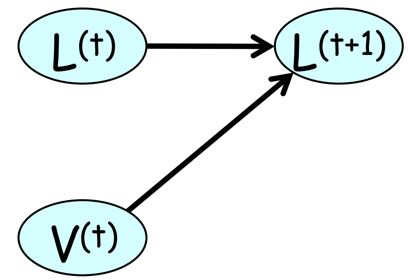


linear Gaussian whose parameters depend on A

$$Y \sim \mathcal{N}\left(w_{\underline{a}0} + \sum w_{\underline{a}i} X_i; \sigma^2_{\underline{a}}\right)$$

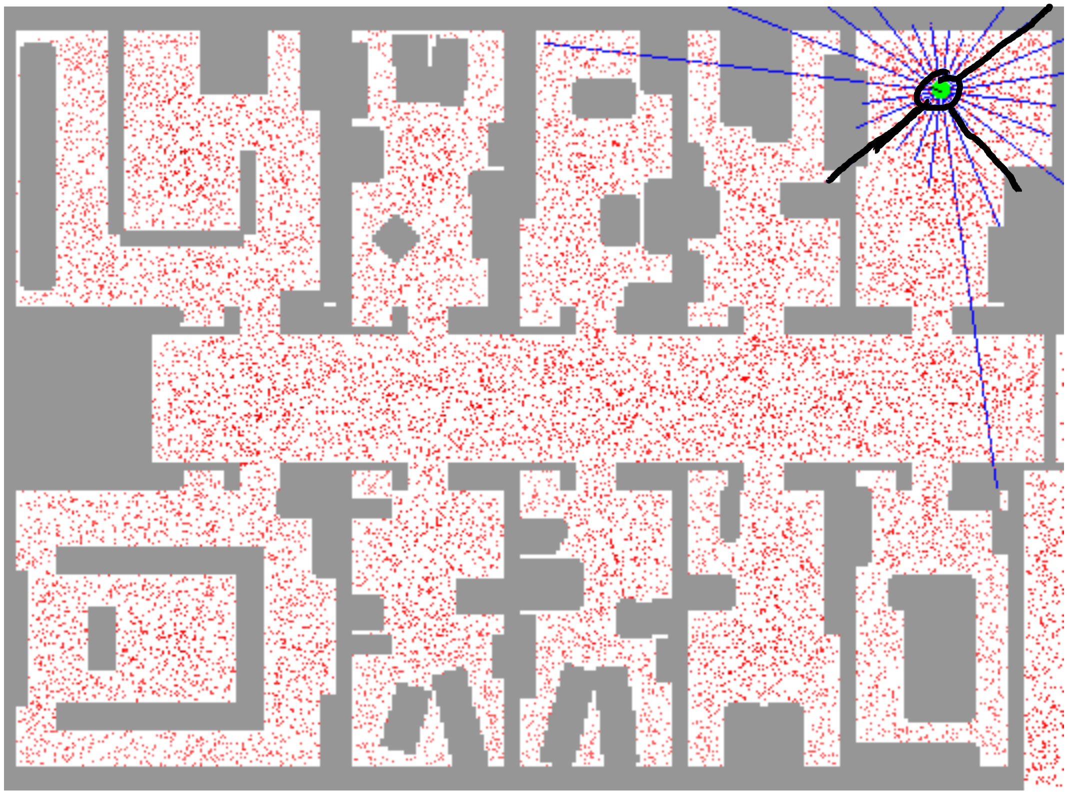
variance can depend on A

Let L and V be the location and velocity of a car. Assume that the CPD on the right is a linear Gaussian. Which of the following statements could possibly be consistent with that CPD? Mark all that apply.



- ☐ Cars that move faster skid more and have greater variance in position.
- ☐ Due to friction, the single most likely value for $L^{(t+1)}$ is $L^{(t)} + 0.9 * V^{(t)}\Delta t$.
- ☐ The distance moved, $|L^{(t+1)} - L^{(t)}|$, will never be more than $2 * V^{(t)}\Delta t$.
- ☐ $L^{(t+1)}$ might possibly end up far from its expected position.

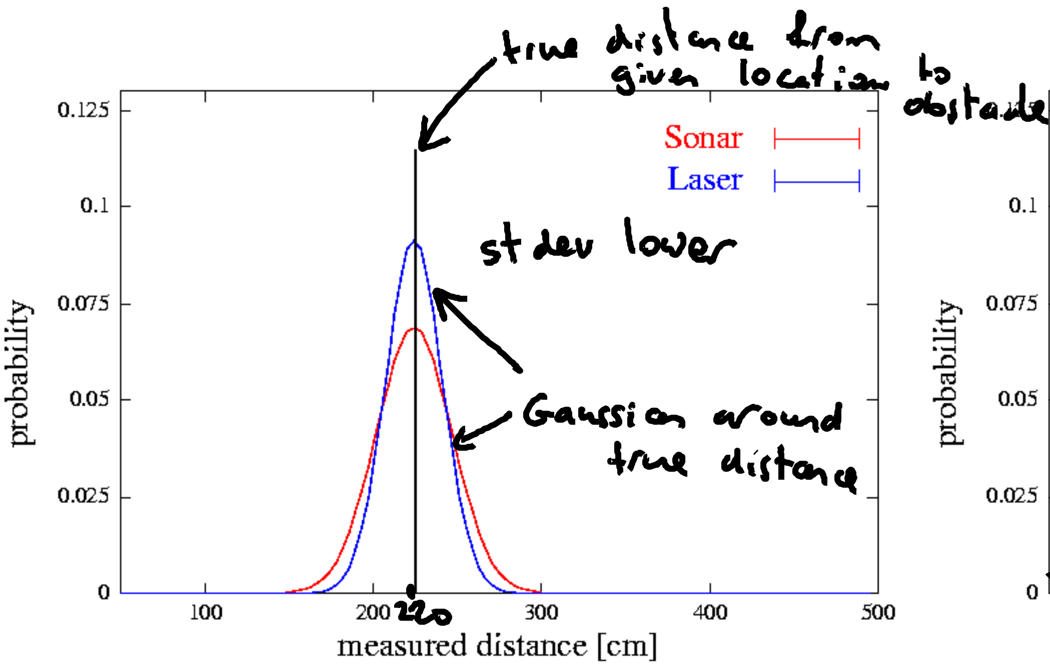
Robot Localization

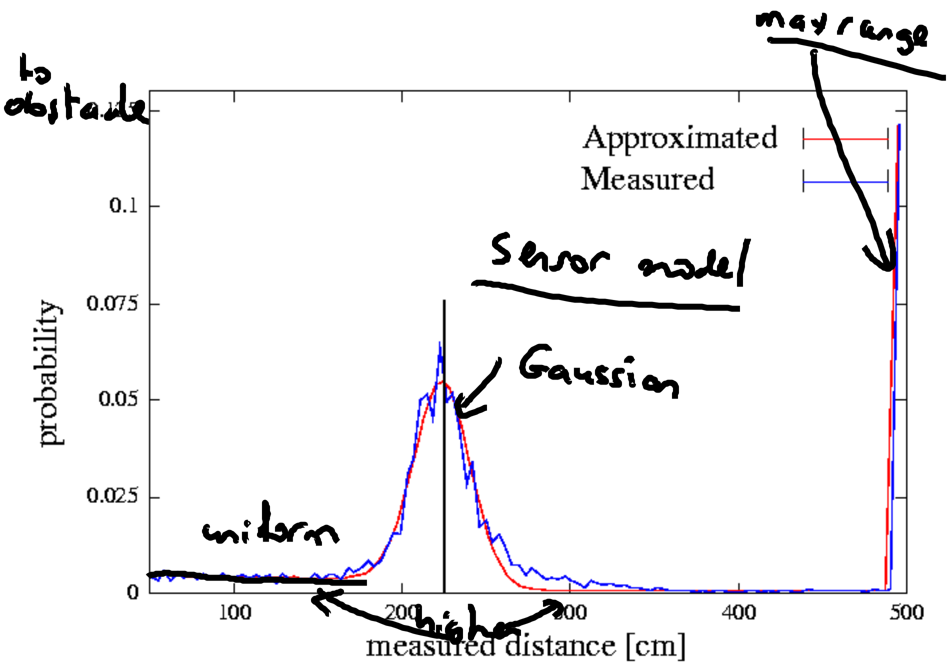


Fox, Burgard, Thrun

Daphne Koller

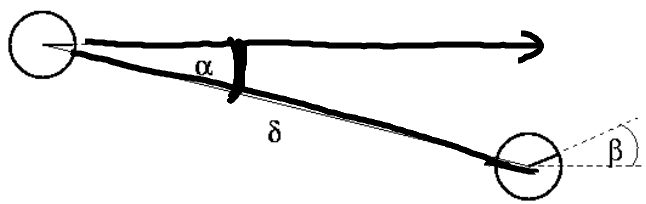
Nonlinear Gaussians

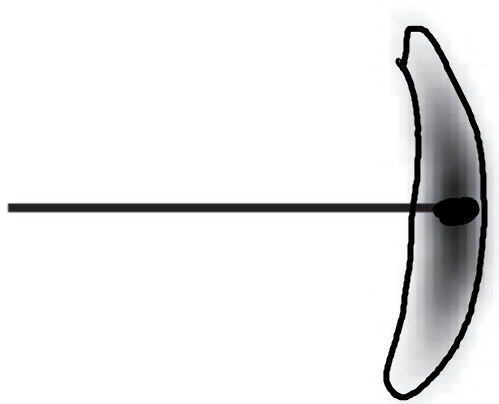


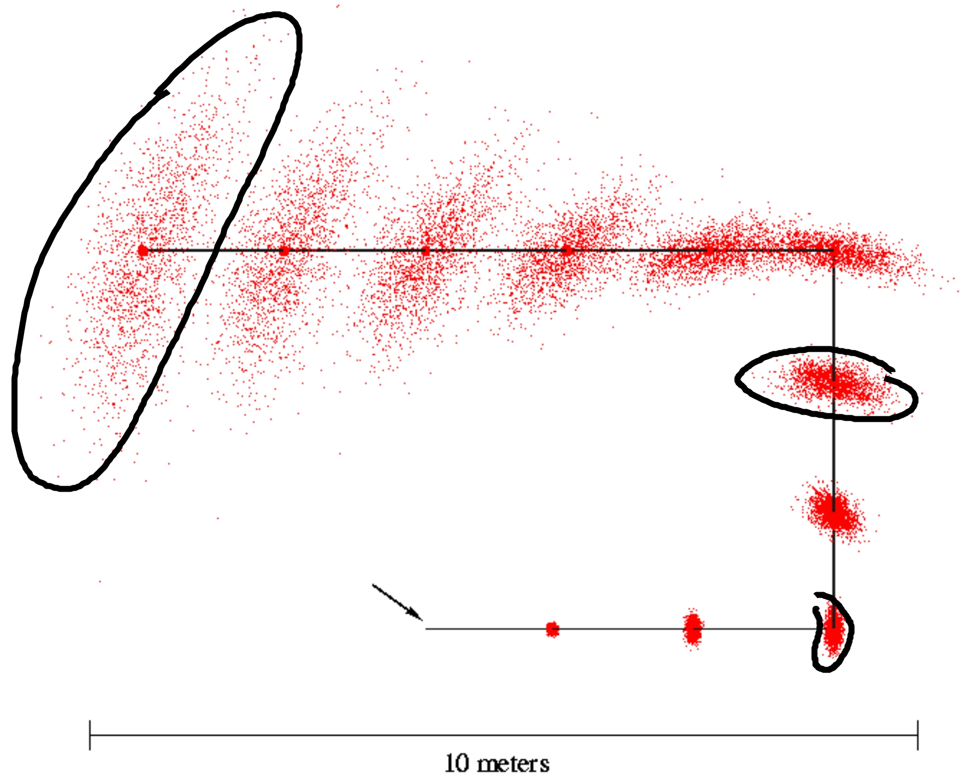


Daphne Koller

Robot Motion Model







Daphne Koller