Artificial Intelligence

Uncertainty

Wen-Huang Cheng (鄭文皇)
National Taiwan University
wenhuang@csie.ntu.edu.tw

CVPR
JUNE 17-21, 2024
SEATTLE, WA

Let's Think Outside the Box !

Exploring Leap-of-Thought in Large Language Models with Creative Humor Generation

Shanshan Zhong*, Zhongzhan Huang*, Shanghua Gao, Wushao Wen, Liang Lin,
Marinka Zitnik, Pan Zhou

https://zhongshsh.github.io/CLoT/

你可真会抢镜
@ You sure know how to steal the spotlight.

突然发现野餐地点是墓地
@ Just realized the picnic spot is a graveyard all of a sudden.

又到了给主人买饭的时间
@ It's time to grab some food for the owner again.

宇宙旅行前的最后祈祷
The final prayer before a journey through the cosmos.

被狗咬了，但只能吃热狗泄愤
@ Got bit by a dog, but all I can do is vent my frustration by eating a hot dog.

情侣车出门约会
@ Couple's car heading out for a date

AI Weekly

今天是情人节......
@ Today is Valentine's Day...

抹茶奥利奥
@ Matcha Oreo

你是不是在外面有狗了？！
@ Is there another dog in your life?!

狗界最好的演唱会！
@ The ultimate dog concert!

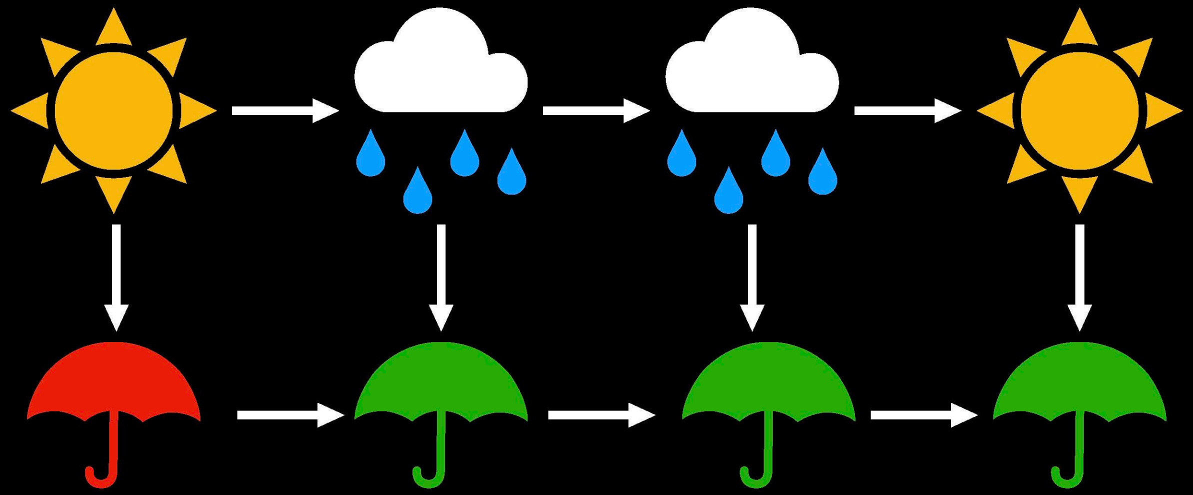
如果我这么做的时候，会被别人误会成小偷吗?
@ Would people mistake me for a thief if I do this?

我今天有作业要交，可以借我一下吗?
@ I have an assignment due today. Can I borrow yours for a bit?

This Lecture – Agenda

- **Bayer’s Net**
  - Representation
  - Conditional Independences
  - Probabilistic Inference
  - Learning Bayes’ Nets from Data

Uncertainty



10 Day Weather - Sanli, Taipei

As of 2:17 pm CST

Thu 11 | Day

87°

47%

NNE8 mph

Cloudy with occasional rain showers. High 87F. Winds NNE at 5 to 10 mph. Chance of rain 50%.

Humidity **70%**	UV Index **6 of 11**
Sunrise **5:35 am**	Sunset **6:14 pm**

Thu 11 | Night

70°

13%

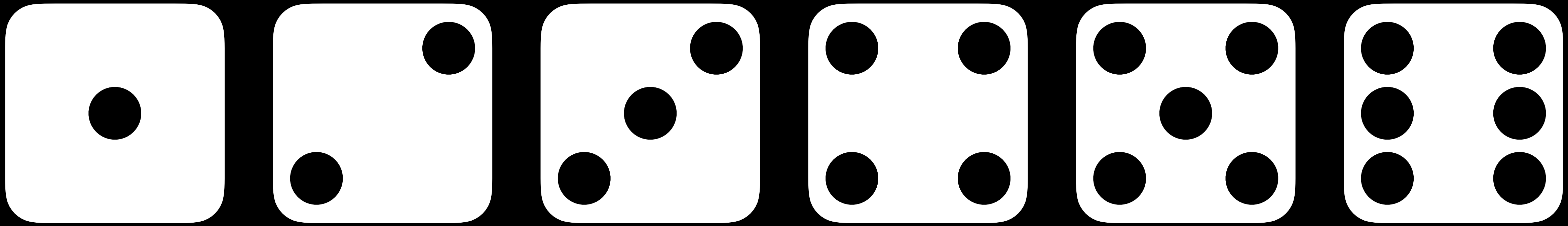
SE6 mph

A few clouds. Low around 70F. Winds light and variable.

Humidity **87%**	UV Index **0 of 11**
Moonrise **7:07 am**	Moonset **9:08 pm**

Waxing Crescent

Fri 12	**85°**/71°	PM Thunderstorms	39%	WNW 9 mph
Sat 13	**84°**/72°	Scattered Thunderstorms	47%	E 9 mph
Sun 14	**87°**/73°	Partly Cloudy	16%	SW 10 mph
Mon 15	**91°**/74°	Partly Cloudy	19%	SSW 10 mph



Review of Probability Theory

Probability theory is the study of uncertainty.

Probabilistic Inference

- Probabilistic inference: compute a desired probability from other known probabilities (e.g. conditional from joint)

- We generally compute conditional probabilities
  - P(on time | no reported accidents) = 0.90
  - These represent the agent's *beliefs* given the evidence

- Probabilities change with new evidence:
  - P(on time | no accidents, 5 a.m.) = 0.95
  - P(on time | no accidents, 5 a.m., raining) = 0.80
  - Observing new evidence causes *beliefs to be updated*

Inference by Enumeration

- General case:
  - Evidence variables: $E_1 \dots E_k = e_1 \dots e_k$
  - Query* variable: Q
  - Hidden variables: $H_1 \dots H_r$

$$\left.\begin{array}{l} E_1 \dots E_k = e_1 \dots e_k \\ Q \\ H_1 \dots H_r \end{array}\right\} X_1, X_2, \dots X_n$$

All variables

- We want:

** Works fine with multiple query variables, too*

$$P(Q|e_1 \dots e_k)$$

Review: Inference by Enumeration

- P(W)?

- P(W | winter)?

→W is query variable, S is evidence variable,
T is hidden variable

$\mathrm{P(sun|winter)} = 0.25\ (0.10 + 0.15)$

$\mathrm{P(rain|winter)} = 0.25\ (0.05 + 0.20)$

$\mathrm{P(sun|winter)} = 0.5$
$\mathrm{P(rain|winter)} = 0.5$

- P(W | winter, hot)?

Season Temperature Weather

S	T	W	P
summer	hot	sun	0.30
summer	hot	rain	0.05
summer	cold	sun	0.10
summer	cold	rain	0.05
winter	hot	sun	0.10
winter	hot	rain	0.05
winter	cold	sun	0.15
winter	cold	rain	0.20

Inference by Enumeration

- General case:
  - Evidence variables: $E_1 \dots E_k = e_1 \dots e_k$
  - Query* variable: Q
  - Hidden variables: $H_1 \dots H_r$

$\left.\begin{matrix} E_1 \dots E_k = e_1 \dots e_k \\ Q \\ H_1 \dots H_r \end{matrix}\right\} \quad X_1, X_2, \dots X_n$

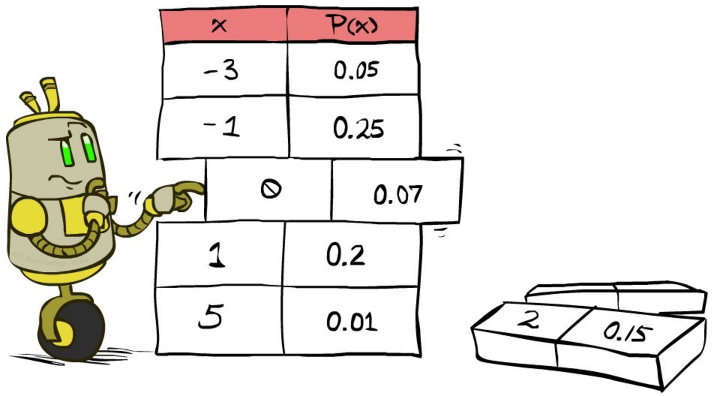
All variables

- We want:

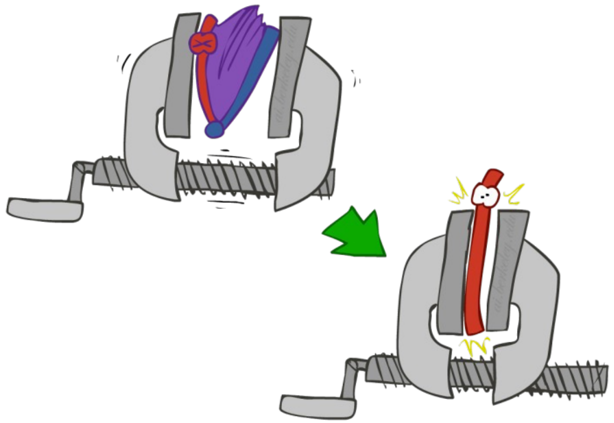
** Works fine with multiple query variables, too*

$$P(Q|e_1 \dots e_k)$$

- Step 1: Select the entries consistent with the evidence



- Step 2: Sum out H to get joint of Query and evidence



$$P(Q, e_1 \dots e_k) = \sum_{h_1 \dots h_r} \underbrace{P(Q, h_1 \dots h_r, e_1 \dots e_k)}_{X_1, X_2, \dots X_n}$$

- Step 3: Normalize

$$\times \frac{1}{Z}$$

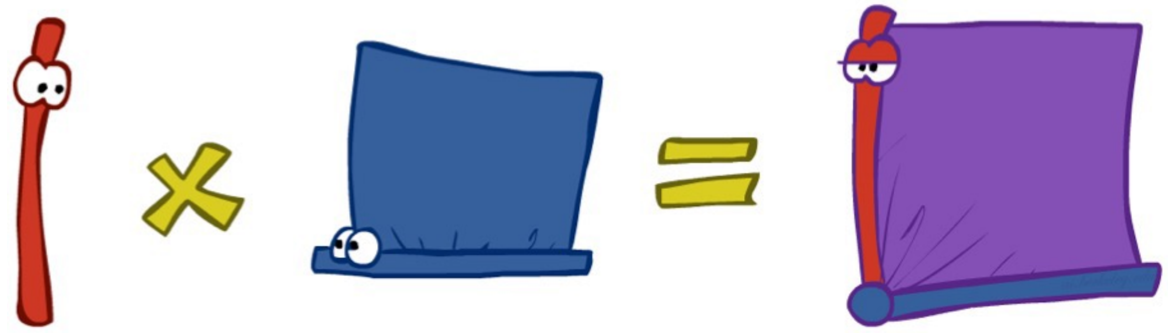
$$Z = \sum_q P(Q, e_1 \cdots e_k)$$

$$P(Q|e_1 \cdots e_k) = \frac{1}{Z} P(Q, e_1 \cdots e_k)$$

The Product Rule

- Sometimes have conditional distributions but want the joint

$$P(y)P(x|y) = P(x, y) \qquad \Longleftrightarrow \qquad P(x|y) = \frac{P(x, y)}{P(y)}$$



The Product Rule

$$P(y)P(x|y) = P(x,y)$$

- Example:

$P(W)$

R	P
sun	0.8
rain	0.2

$P(D|W)$

D	W	P
wet	sun	0.1
dry	sun	0.9
wet	rain	0.7
dry	rain	0.3

$\Leftrightarrow$

$P(D,W)$

D	W	P
wet	sun	0.08
dry	sun	0.72
wet	rain	0.14
dry	rain	0.06

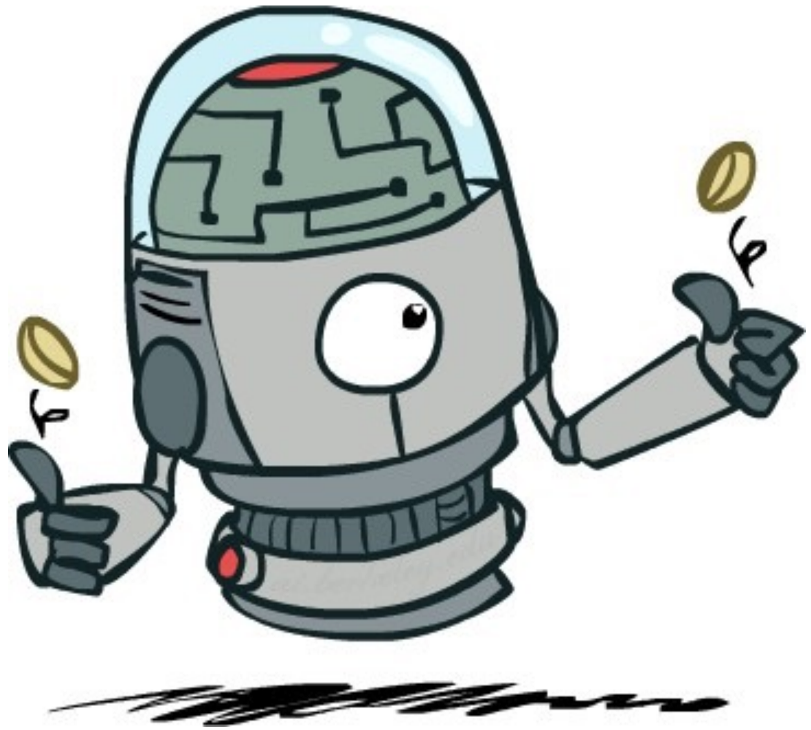
Probabilistic Models

- Models describe how (a portion of) the world works

- Models are always simplifications
  - May not account for every variable
  - May not account for all interactions between variables
  - "All models are wrong; but some are useful."
    – George E. P. Box: "one of the great statistical minds of the 20th century"

- What do we do with probabilistic models?
  - We (or our agents) need to reason about unknown variables, given evidence
  - Example: explanation (diagnostic reasoning)
  - Example: prediction (causal reasoning)
  - Example: value of information

Independence



Independence

- Two variables are *independent* if:

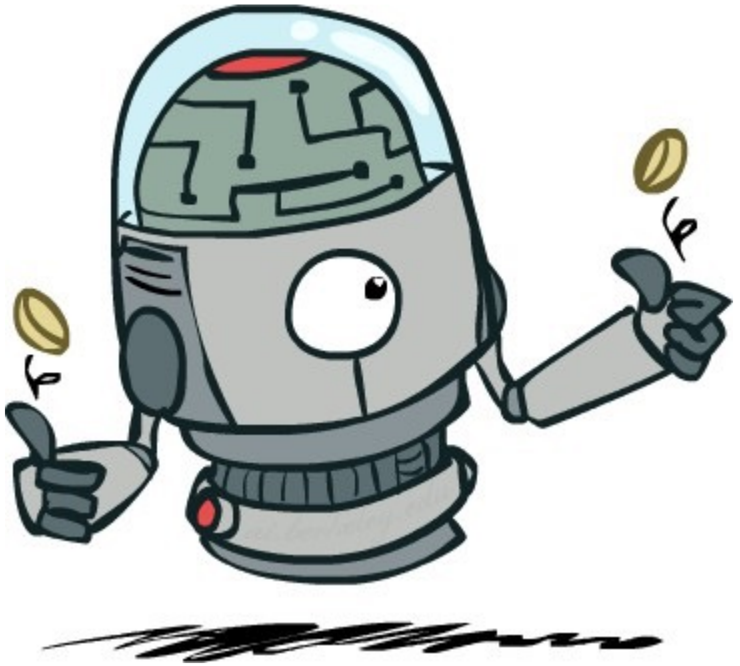
$$\forall x, y : P(x, y) = P(x)P(y)$$

  - This says that their joint distribution *factors* into a product two simpler distributions
  - Another form:

$$\forall x, y : P(x|y) = P(x)$$

  - We write: $X \perp\!\!\!\perp Y$

- Independence is a simplifying *modeling assumption*
  - *Empirical* joint distributions: at best "close" to independent
  - What could we assume for {Traffic, Toothache}?



Example: Independence?

$P_1(T,W)$!= $P(T)P(W)$

$P_2(T,W) = P(T)P(W)$

$P(T)$

T	P
hot	0.5
cold	0.5

$P_1(T,W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

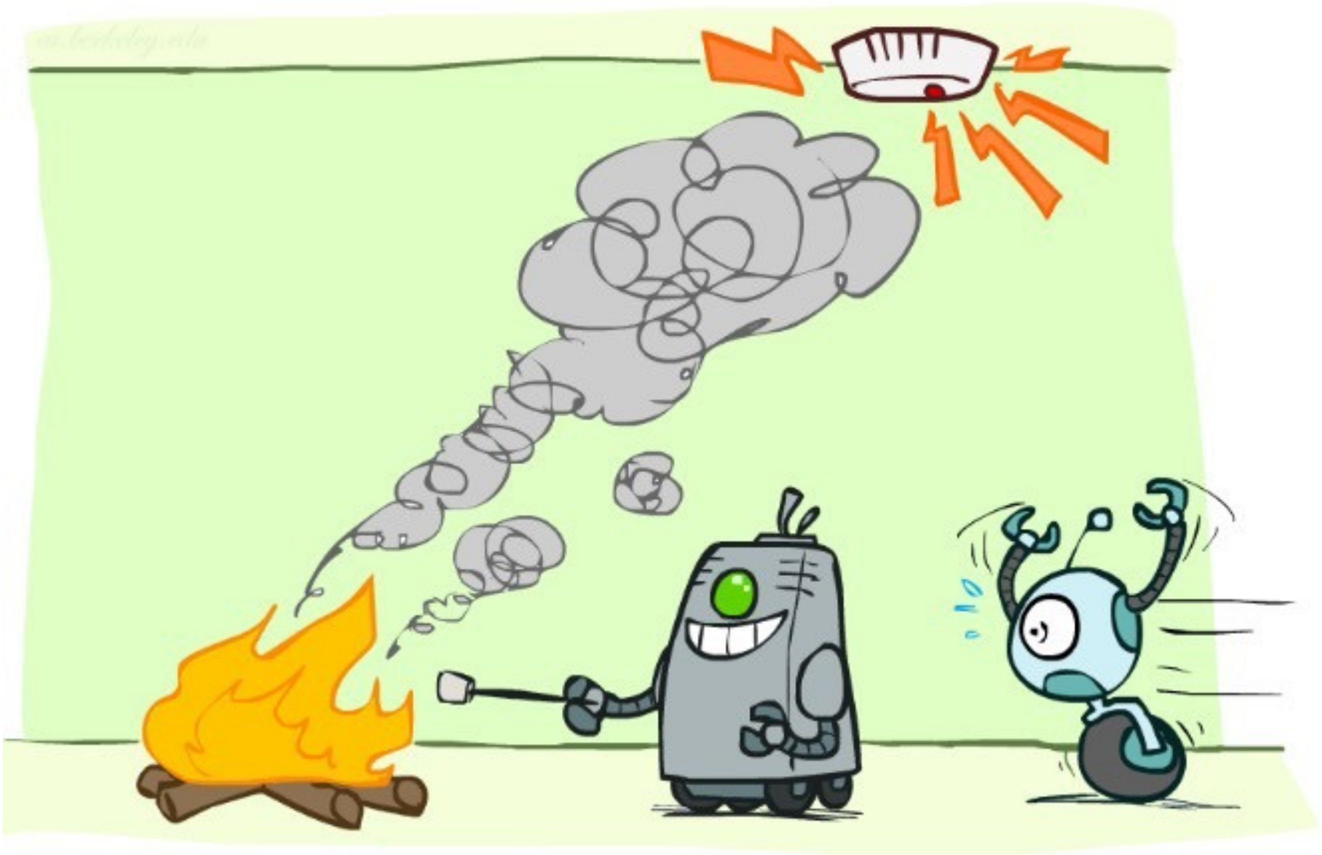
$P_2(T,W)$

T	W	P
hot	sun	0.3
hot	rain	0.2
cold	sun	0.3
cold	rain	0.2

$P(W)$

W	P
sun	0.6
rain	0.4

Conditional Independence

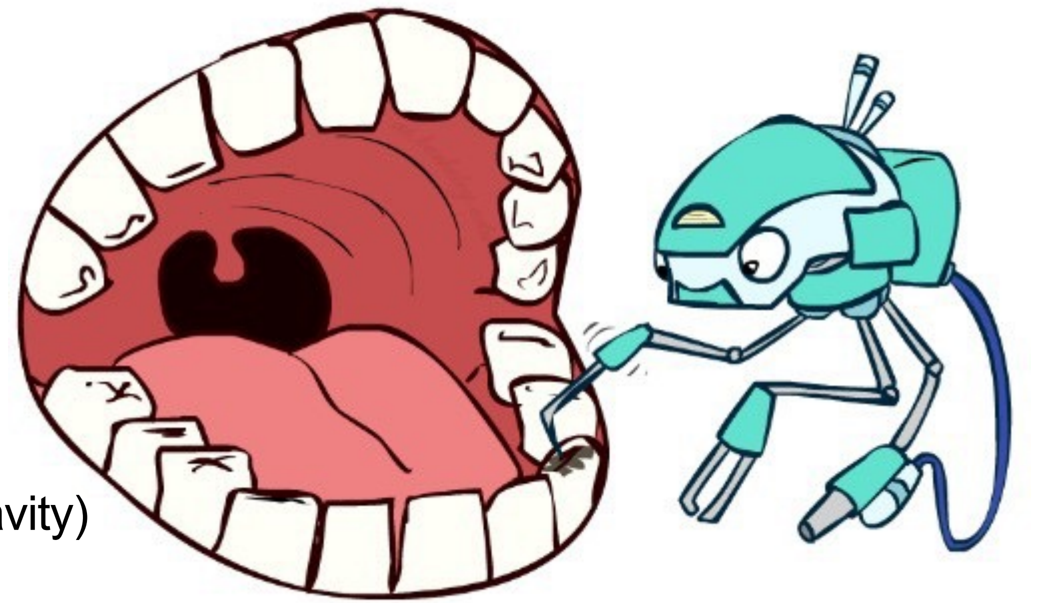


Conditional Independence

- P(Toothache, Cavity, Catch)

- Catch is *conditionally independent* of Toothache given Cavity:
  - P(Catch | Toothache, Cavity) = P(Catch | Cavity)

- Equivalent statements:
  - P(Toothache | Catch , Cavity) = P(Toothache | Cavity)
  - P(Toothache, Catch | Cavity) = P(Toothache | Cavity) P(Catch | Cavity)
  - One can be derived from the other easily



Conditional Independence

- Unconditional (absolute) independence very rare
- *Conditional independence* is our most basic and robust form of knowledge about uncertain environments.
- X is conditionally independent of Y given Z

$$X \perp\!\!\!\perp Y \mid Z$$

  if and only if:

$$\forall x, y, z : P(x, y|z) = P(x|z)P(y|z)$$

  or, equivalently, if and only if

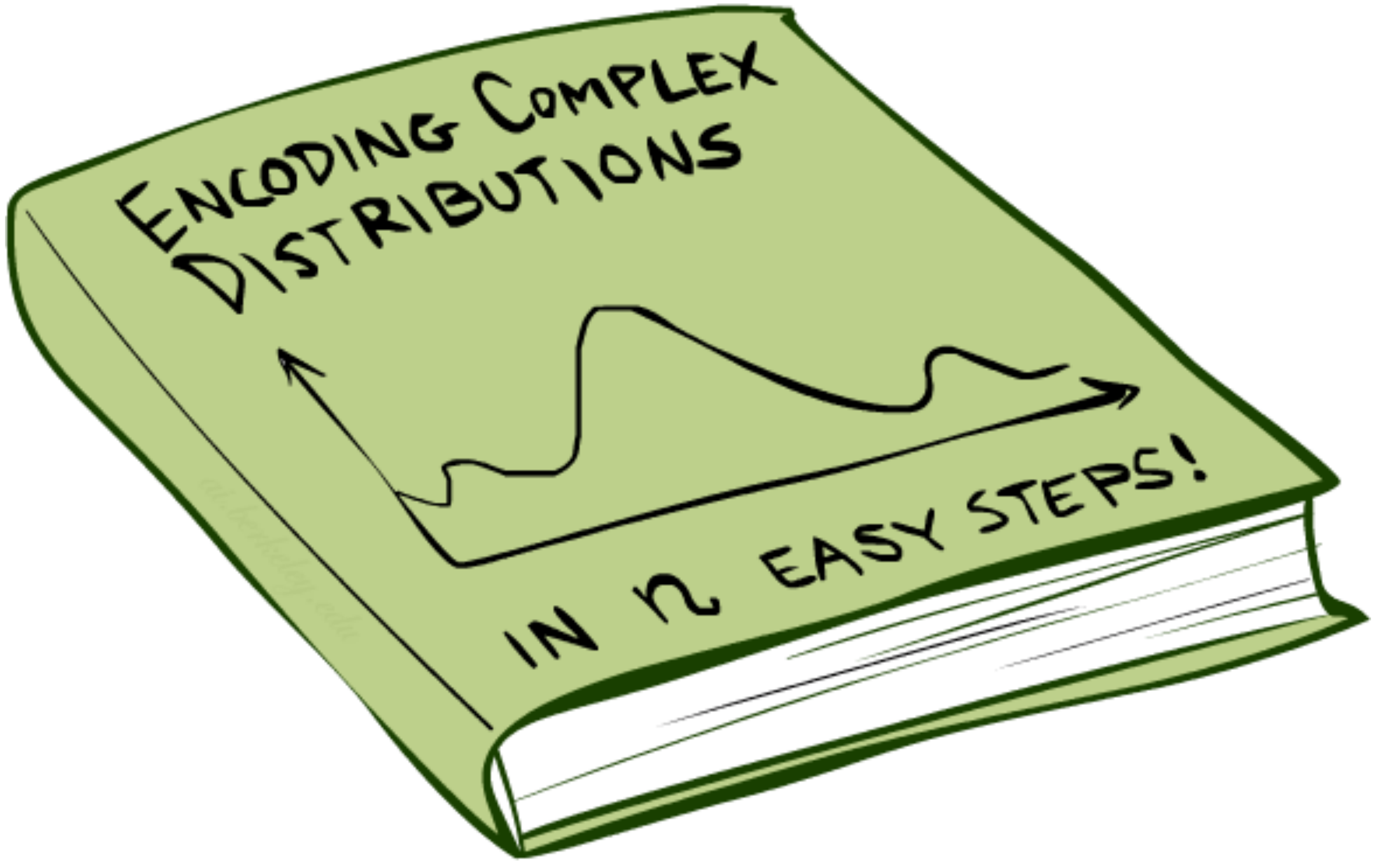
$$\forall x, y, z : P(x|z, y) = P(x|z)$$

Bayes' Nets

- Representation
- Conditional Independences
- Probabilistic Inference
- Learning Bayes' Nets from Data

Bayes'Nets: Big Picture

- Bayes'nets models help us express conditional independence assumptions



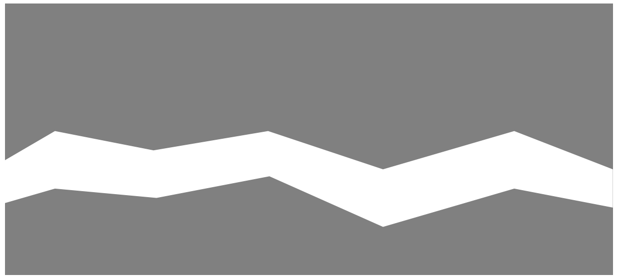
Bayes' Nets: Big Picture

- Two problems with using full joint distribution tables as our probabilistic models:
  - Unless there are only a few variables, the joint is WAY too big to represent explicitly
  - Hard to learn (estimate) anything empirically about more than a few variables at a time

if independence

$P(X_1, X_2, \ldots X_n)$

2^n



$P(X_1)$

H	0.5
T	0.5

$P(X_2)$

H	0.5
T	0.5

$\ldots$

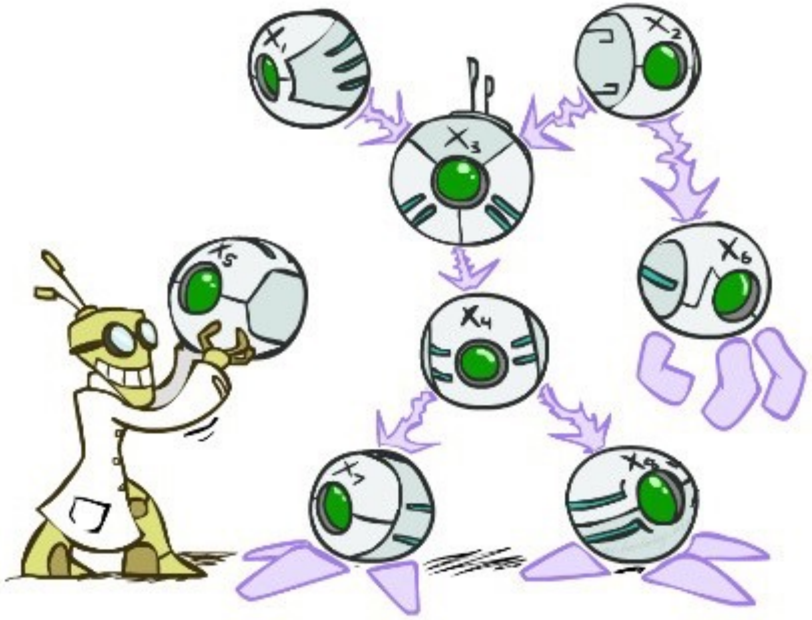
$P(X_n)$

H	0.5
T	0.5

Bayes' Nets: Big Picture

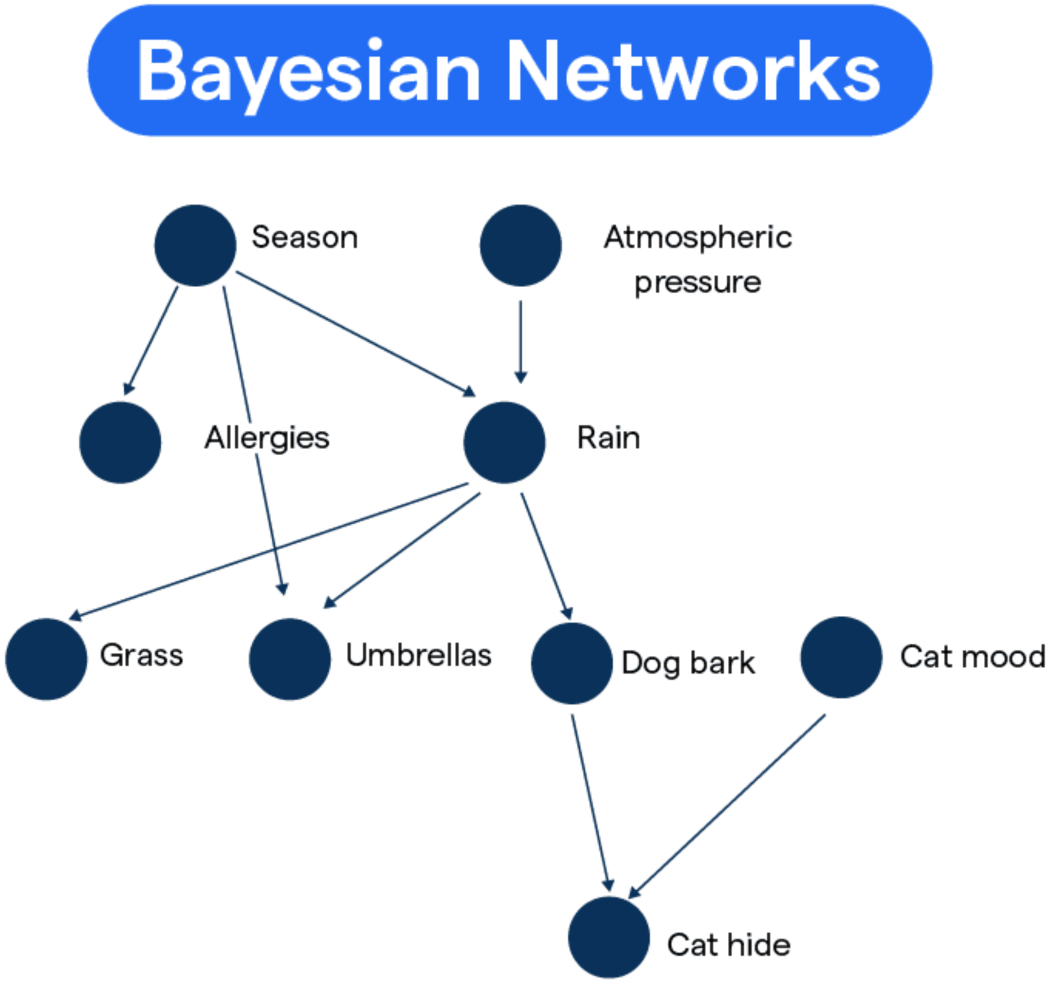
- Two problems with using full joint distribution tables as our probabilistic models:
  - Unless there are only a few variables, the joint is WAY too big to represent explicitly
  - Hard to learn (estimate) anything empirically about more than a few variables at a time

- Bayes' nets: a technique for describing complex joint distributions (models) using simple, local distributions (conditional probabilities)
  - More properly called graphical models
  - We describe how variables locally interact
  - Local interactions chain together to give global, indirect interactions



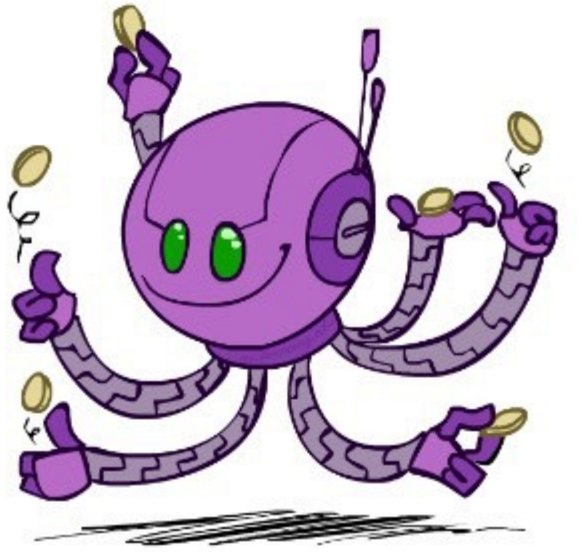
Graphical Model Notation

- Nodes: variables (with domains)
  - Can be assigned (observed) or unassigned (unobserved)
- Arcs: interactions
  - Indicate “direct influence” between variables



Example: Coin Flips

- N independent coin flips



$$X_1 \quad X_2 \quad \cdots \quad X_n$$

- No interactions between variables: absolute independence

Example: Traffic

- Variables:
  - R: It rains
  - T: There is traffic

- Model 1: independence

- Model 2: rain causes traffic