Marginal Independence and Conditional Independence

CPSC 322 Lecture 25

Lecture Overview

- **Recap with Example and Bayes' Rule**
- Marginal Independence
- Conditional Independence

Recap Joint Distribution

- 3 binary random variables: **P(H,S,F)**
  - **H dom(H)={h, ¬h}** has heart disease, does not have…
  - **S dom(S)={s, ¬s}** smokes, does not smoke
  - **F dom(F)={f, ¬f}** high fat diet, low fat diet

Recap Joint Distribution (JPD)

- 3 binary random variables: **P(H,S,F)**
  - **H dom(H)={h, ¬h}** has heart disease, does not have…
  - **S dom(S)={s, ¬s}** smokes, does not smoke
  - **F dom(F)={f, ¬f}** high fat diet, low fat diet

	f, **s**	**f**, **¬ s**	**¬ f**, **s**	**¬ f**, **¬ s**
h	.015	.007	.005	.003
¬ h	.21	.51	.07	.18

Recap Marginalization

P(H,S,F)

	f		¬ f	
	s	**¬ s**	**s**	**¬ s**
h	.015	.007	.005	.003
¬ h	.21	.51	.07	.18

$$P(H,S) = \sum_{x \in dom(F)} P(H,S,F=x)$$

P(H,S)

	s	**¬ s**
h	.02	.01
¬ h	.28	.69

P(H)?

P(S)?

Recap Conditional Probability

P(H,S)	**s**	**¬ s**	P(H)
h	.02	.01	.03
¬ h	.28	.69	.97
P(S)	.30	.70	

$$P(S \mid H) = \frac{P(S,H)}{P(H)}$$

P(S|H)

h	.666	.333
¬ h	.29	.71

multiple probability distributions

do P(H|S) as an exercise

Recap Conditional Probability (cont.)

$$P(S \mid H) = \frac{P(S,H)}{P(H)}$$

Two key points

- We derived this equality from a “possible world” semantics of probability
- It is not a probability distribution but a **set of probability distributions**
  - One for each configuration of the conditioning variable(s)

Recap Chain Rule

$$P(H,S,F) =$$

Bayes Theorem

$$P(S \mid H) = \frac{P(S,H)}{P(H)}$$

$$P(H \mid S) = \frac{P(S,H)}{P(S)}$$

$$P(S \mid H) = \frac{P(H \mid S)P(S)}{P(H)}$$

Learning Goals for today's class

- **You can:**

- Define and use **Marginal Independence**

- Define and use **Conditional Independence**

Lecture Overview

- Recap with Example and Bayes Theorem
- **Marginal Independence**
- Conditional Independence

Do you always need to revise your beliefs?

NO, not when your knowledge of **Y**'s value doesn't affect your belief in the value of **X**

DEF. A random variable **X** is marginally independent of random variable **Y** if, for all $x_i \in dom(X)$ and all $y_k \in dom(Y)$,

$$P(X= x_i \mid Y= y_k) = P(X= x_i)$$

Marginal Independence: Example

- *X* and *Y* are independent iff:

$\mathbf{P}(X|Y) = \mathbf{P}(X)$ or $\mathbf{P}(Y|X) = \mathbf{P}(Y)$ or $\mathbf{P}(X, Y) = \mathbf{P}(X)\,\mathbf{P}(Y)$

- That is, new evidence Y (or X) does not affect current belief in X (or Y)
- Ex: **P**(*Toothache, Catch, Cavity, Weather*)
  = **P**(*Toothache, Catch, Cavity*) **P**(*Weather*)
- JPD requiring ___ entries is reduced to two smaller ones (___ and ___)

In our example are Smoking and Heart Disease marginally Independent ?

iclicker

What are our probabilities telling us….?

X and *Y* are independent iff:

$\mathbf{P}(X|Y) = \mathbf{P}(X)$ or $\mathbf{P}(Y|X) = \mathbf{P}(Y)$ or $\mathbf{P}(X, Y) = \mathbf{P}(X)\,\mathbf{P}(Y)$

P(S,H)	**s**	**¬ s**	P(H)
h	.02	.01	.03
¬ h	.28	.69	.97
P(S)	.30	.70	

P(S|H)	**s**	**¬ s**
h	.666	.334
¬ h	.29	.71

A. Yes B. No C. It depends D. 42 E. Please make it stop

Lecture Overview

- Recap with Example
- Marginal Independence
- Conditional Independence

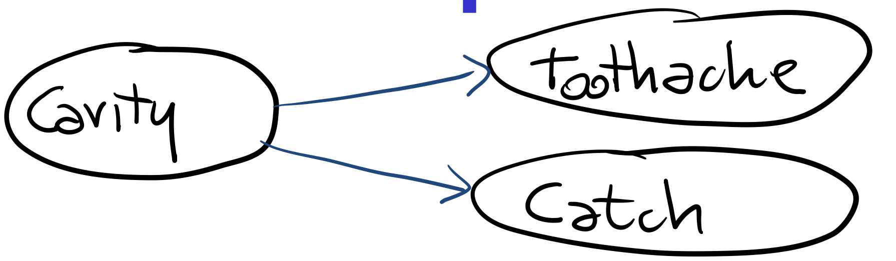
Conditional Independence

- With marginal independence, for *n* independent random vars, $O(d^n) \rightarrow O(n*d)$ space complexity

$$P(x_1, \ldots, x_n) = P(x_1) * \ldots * P(x_n)$$

- Absolute independence is powerful **but** when you model a **particular domain**, it is **rare**
- Dentistry is a large field with hundreds of variables, few of which are independent (e.g.,*Cavity, Heart-disease*).
- What to do?

Look for weaker form of independence



- **P**(*Toothache, Cavity, Catch*)

- Are *Toothache* and *Catch* marginally independent?
  P(*Toothache* | *Catch*) = **P**(*Toothache*) ? NO

- BUT **If I have a cavity**, does the probability that the probe catches depend on whether I have a toothache? **NO**
  (1) P(*catch* | *toothache, cavity*) = **P**(*catch* | *cavity*)

- What if I haven't got a cavity?
  (2) P(*catch* | *toothache,*$\neg$*cavity*) = **P**(*catch* | $\neg$*cavity*)

Each is directly caused by the cavity, but neither has a direct effect on the other

Conditional independence

- In general, *Catch* is conditionally independent of *Toothache* given *Cavity*:

1 $\mathbf{P}(Catch \mid Toothache, Cavity) = \mathbf{P}(Catch \mid Cavity)$

- Equivalent statements:

2 $\mathbf{P}(Toothache \mid Catch, Cavity) = \mathbf{P}(Toothache \mid Cavity)$

3 $\mathbf{P}(Toothache, Catch \mid Cavity) =$
$\mathbf{P}(Toothache \mid Cavity)\, \mathbf{P}(Catch \mid Cavity)$

Proof of equivalent statements

① If $P(X|YZ) = P(X|Z)$ ⇒

$$\Rightarrow \text{(A)}\ \frac{P(X,Y,Z)}{P(Y,Z)} = \frac{P(X,Z)}{P(Z)} \Rightarrow$$

②

$$\Rightarrow \frac{P(X,Y,Z)}{P(X,Z)} = \frac{P(Y,Z)}{P(Z)} \Rightarrow P(Y|X,Z) = P(Y|Z)$$

③

$$P(X,Y|Z) = \frac{P(X,Y,Z)}{P(Z)} \overset{\text{from A}}{\Rightarrow} \frac{P(Y,Z)\,P(X,Z)}{P(Z)} \cdot \frac{1}{P(Z)}$$

$$= \frac{P(Y,Z)}{P(Z)} \cdot \frac{P(X,Z)}{P(Z)} = P(Y|Z) \cdot P(X|Z)$$

Conditional Independence: Formal Def.

Sometimes, two variables might not be marginally independent. However, they *become* independent after we observe some third variable

DEF. Random variable **X** is conditionally independent of random variable **Y** given random variable **Z** if, for all $x_i \in dom(X)$, $y_k \in dom(Y)$, $z_m \in dom(Z)$

$$P(X= x_i \mid Y= y_k , Z= z_m) = P(X= x_i \mid Z= z_m)$$

That is, knowledge of **Y**'s value doesn't affect your belief in the value of **X,** given a value of **Z**

Side note: storing distributions

Joint Probability Distribution (JPD): has $O(d^n)$ values

- but they have to **sum to 1**, so do we need to store all of them?
- how many do we need to store?

Conditional Probability Table (CPT): has $O(d^n)$ values

- but **each row** has to sum to 1, so do we need to store all of them?
- how many do we need to store?

Conditional independence: Use

- Write out full joint distribution using chain rule:

  P(*Cavity, Catch, Toothache*)

  = **P**(*Toothache* | *Catch, Cavity*) **P**(*Catch* | *Cavity*) **P**(*Cavity*)

  = **P**(*Toothache* | Cavity) **P**(*Catch* | *Cavity*) **P**(Cavity)

  how many probabilities?

- The use of conditional independence often reduces the size of the representation of the joint distribution from exponential in *n* to linear in *n*, **where n is the # of variables**

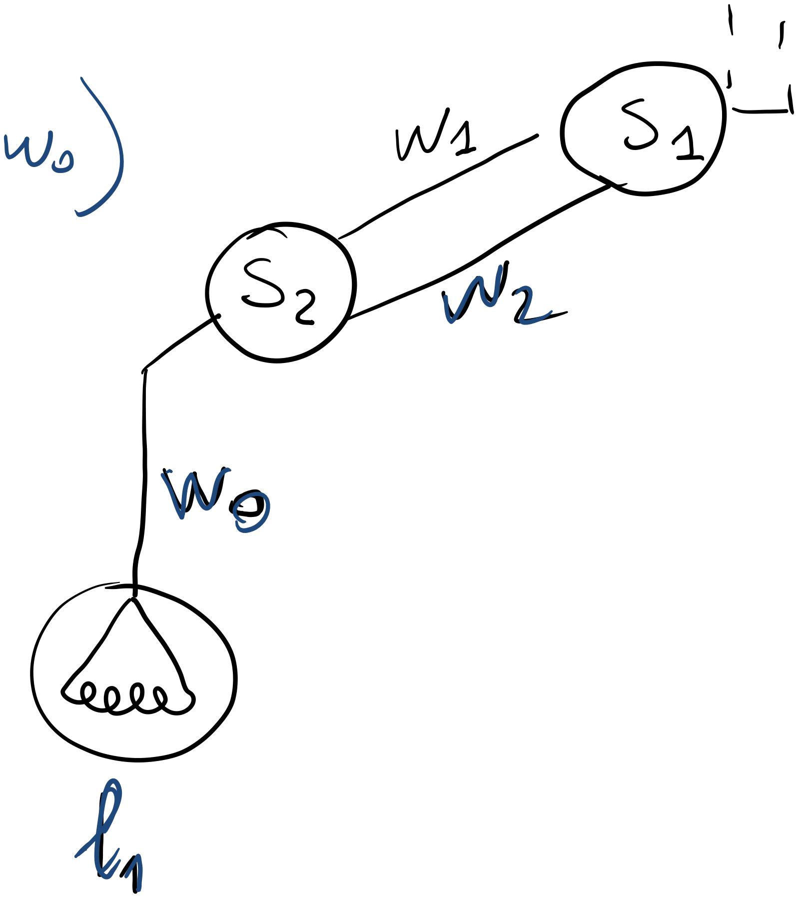
- **Conditional independence** is our **most basic** and **robust** form of **knowledge** about **uncertain environments.**

Conditional Independence Example 2

- Given whether there is/isn't power in wire w0, is whether light l1 is lit or not, independent of the position of switch s2?

$$P(l_1 \mid s_2, w_0) \stackrel{?}{=} P(l_1 \mid w_0)$$

yes!



Conditional Independence Example 3

- Is every other variable in the system independent of whether light l1 is lit, given whether there is power in wire w0 ?

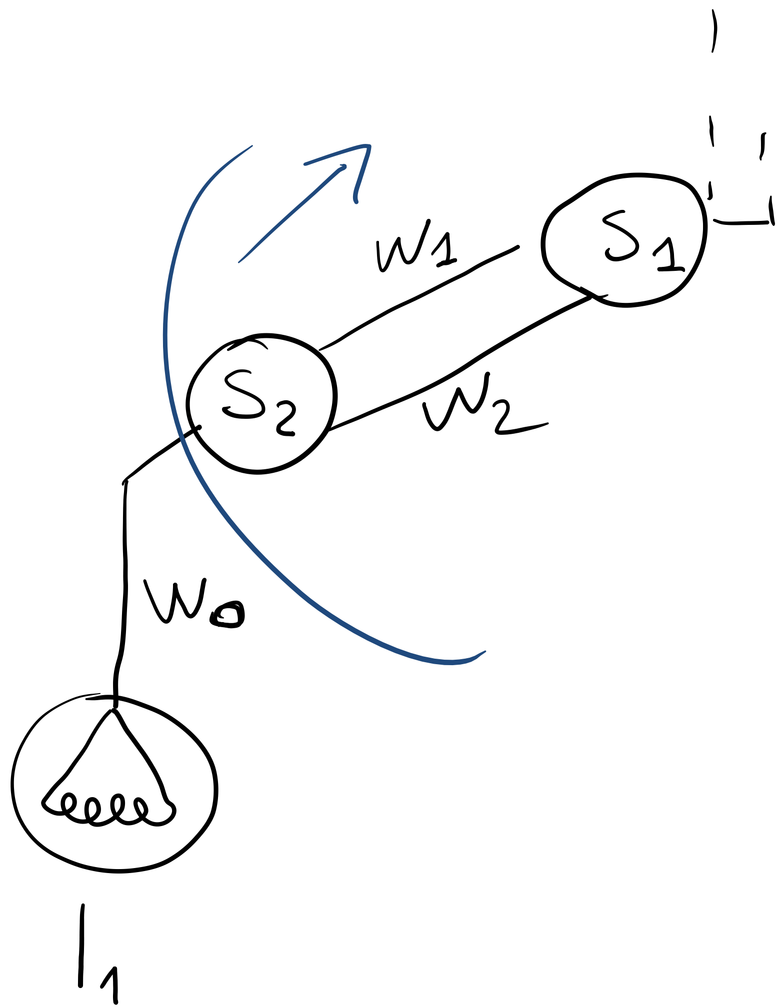
$$P(s_1 \mid l_1, w_0) = P(s_1 \mid w_0)$$

w_1

w_2

⋮

yes!



Where are we? (Summary)

- Probability is a rigorous formalism for uncertain knowledge
- Joint probability distribution specifies probability of every possible world
- Queries can be answered by summing over possible worlds
- For nontrivial domains, we must find a way to reduce the joint distribution size
- **Independence** *(rare)* and **conditional independence** *(frequent)* provide the tools

Next Class

- Bayesian Networks (Textbook 8.3)