Marginal Independence and Conditional Independence

Computer Science cpsc322, Lecture 26

(Textbook Chpt 6.1-2)

Nov, 7, 2012

Lecture Overview

- Recap with Example
  - Marginalization
  - Conditional Probability
  - Chain Rule
- Bayes' Rule
- Marginal Independence
- Conditional Independence

our most basic and robust form of knowledge about uncertain environments.

Recap Joint Distribution

H=True H=False

- 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

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

$\sum 1$

$2^3 - 1$

$2^K - 1$

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	.03
¬ h	.28	.69	.97
	.3	.7	

P(H)?

P(S)?

Recap Conditional Probability

P(H,S)	s	¬ s	P(H)
h	.02	.01	.03
¬ h	.28	.69	$\boxed{.97}$
P(S)	.30	.70	

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

$$P(s \mid \neg h) = \frac{P(s, \neg h)}{\boxed{P(\neg h)}}$$

Two probability distributions for S

$\boxed{\text{P(S\|H)}}$	s	¬ s	
h	.666	.333	$\Sigma\ 1$
¬ h	.29	.71	$\Sigma\ 1$

P(H|S)

do this as an exercise

Recap Conditional Probability (cont.)

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

$P(S \mid H, F)$

$P(X_1 \dots, X_n \mid Y_1 \dots Y_k)$

binary

Two key points we covered in the previous lecture

- We derived this equality from a possible world semantics of probability
- It is not a probability distributions but….. set of prob. distrib.

- One for each configuration of the conditioning var(s)

If conditioned by K binary vars, set 2^K prob. distributions

Recap Chain Rule

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

$$\cancel{P(H)} * \frac{P(S,H)}{\cancel{P(H)}} * \frac{P(F,H,S)}{\cancel{P(H,S)}}$$

Bayes Theorem

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

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

rewrite

$$P(H|S)\,P(S) = P(S,H)$$

Substitute

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

Lecture Overview

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

Do you always need to revise your beliefs?

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

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

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

Marginal Independence: Example

- *X* and *Y* are independent iff: $P(X) = P(X|Y) = \frac{P(X,Y)}{P(Y)}$

  $\boxed{\mathbf{P}(X|Y) = \boxed{\mathbf{P}(X)}}$ or $\boxed{\mathbf{P}(Y|X) = \mathbf{P}(Y)}$ or $\boxed{\mathbf{P}(X, Y) = \mathbf{P}(X)\,\mathbf{P}(Y)}$
- That is new evidence Y(or X) does not affect current belief in X (or Y)
- Ex: $\mathbf{P}(\underline{Toothache, Catch, Cavity}, \underline{Weather})$ $|dom| = 4$ → Sunny Cloudy rainy snowy

  $= \boxed{\mathbf{P}(Toothache, Catch, Cavity)}\ \boxed{P(weather)}$
- JPD requiring 32 entries is reduced to two smaller ones (8 and 4)

Joint prob. distribution

In our example are Smoking and Heart Disease marginally Independent ?

What our probabilities are telling us….?

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

$P(S|H) \overset{?}{=} P(S)$

No!

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

$\overset{?}{=}$ $\overset{?}{=}$

Lecture Overview

- Recap with Example
- Marginal Independence
- Conditional Independence

Conditional Independence

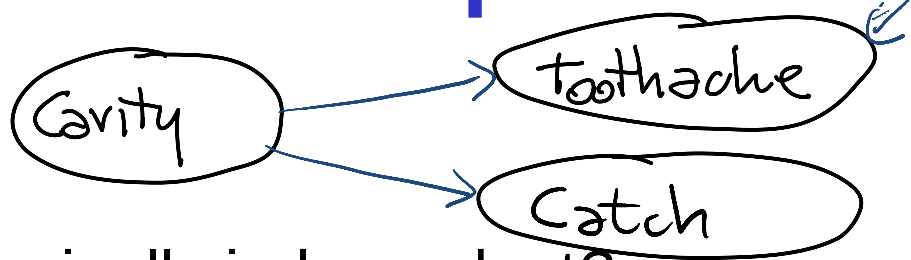
- With marg. Independence, for *n* independent random vars, $O(2^n) \rightarrow O(n)$

  $$P(X_1 \ldots X_n) = P(X_1) \times \ldots \times P(X_n)$$

- 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(|) = 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,*¬*cavity*) = P(catch | ¬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 $\mathbf{P}(Toothache, Catch \mid Cavity) = \mathbf{P}(Toothache \mid Cavity)\, \mathbf{P}(Catch \mid Cavity)$

$$P(X, Y \mid Z) = P(X \mid Z)\, P(Y \mid Z)$$

Proof of equivalent statements

1 If $\boxed{P(X|YZ) = P(X|Z)} \Rightarrow$

$$\Rightarrow \text{A} \quad \frac{P(X,Y,Z)}{P(Y,Z)} = \frac{P(X,Z)}{P(Z)} \Rightarrow \text{2}$$

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

$$\text{3} \quad P(X,Y|Z) = \frac{P(X,Y,Z)}{P(Z)} \overset{\text{from A}}{\Rightarrow} \frac{P(Y,Z)\,P(X,Z)}{P(Z)} \cdot \frac{1}{P(Z)}$$

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

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**

Conditional independence: Use

- Write out full joint distribution using chain rule:

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

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

  $= \underbrace{\mathbf{P}(Toothache \mid Cavity)}_{2}\, \underbrace{\mathbf{P}(Catch \mid Cavity)}_{2}\, \underbrace{\mathbf{P}(Cavity)}_{1}$

  how many probabilities? $2^3 - 1 = 7$

  $2 + 2 + 1 = 5$

- The use of conditional independence often reduces the size of the representation of the joint distribution from exponential in *n* to linear in *n*. **What is n?** # of vars

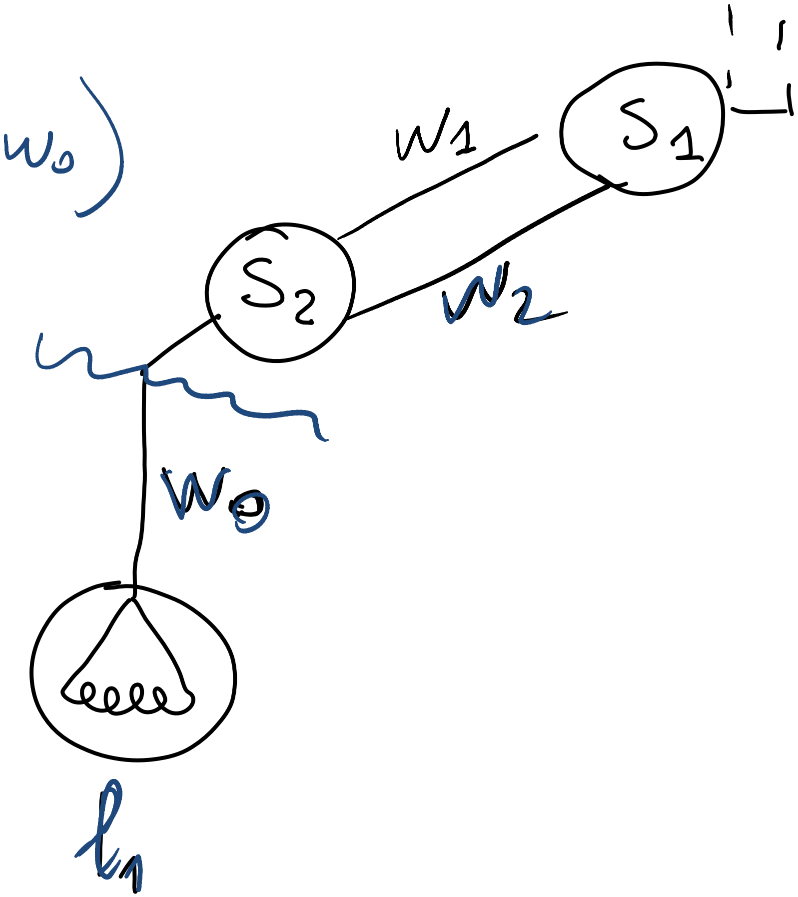
- **Conditional independence** is our **most basic** and **robust** form of **knowledge** about **uncertain environments**.

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?

$$P(\ell_1 \mid s_2, w_0) \stackrel{?}{=} P(\ell_1 \mid w_0)$$

yes!



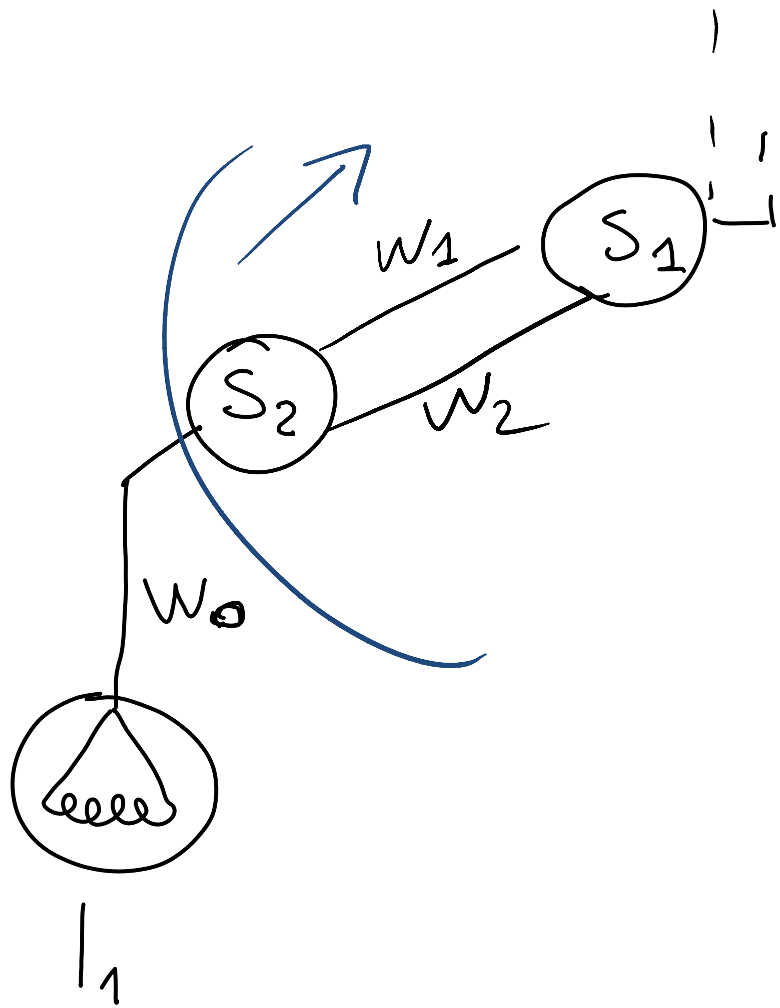
Conditional Independence Example 3

- 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
$\vdots$

yes!



Learning Goals for today's class

- **You can:**
- Derive **the Bayes Rule**
- Define and use **Marginal Independence**
- Define and use **Conditional Independence**

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 (Chpt 6.3)

Start working on assignments3 !