

Uncertainty

CS 580

Intro to Artificial Intelligence

Sensor uncertainty

We've been assuming we can tell what state we're currently in with perfect accuracy. This is often not true!

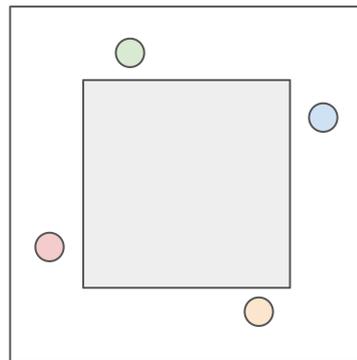
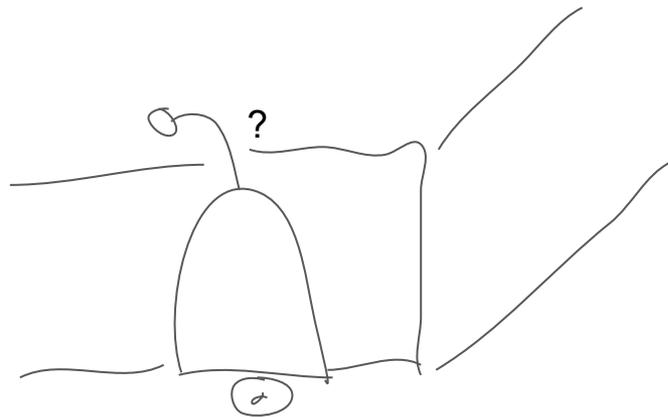
- Is it going to rain in the next hour?
- Do I have allergies or a cold?
- How much time do I have before my battery runs out?

Probabilities to the rescue!

$S = \{\text{'loc':'A', 'A-clean':True, 'B-clean':False}\}$ vs

$S = \{\text{'loc-is-A':0.5, 'A-clean':0.9, 'B-clean':0.2}\}$

We can ask questions like: "What's the probability that both A and B are clean?"



Probabilistic Inference

Represent components of state as **Random Variables**:

$$p(X_i = \text{TRUE}) = p, \quad 0 \leq p \leq 1$$

Random Variables can be **discrete** or continuous

Raining $\in \{\text{TRUE}, \text{FALSE}\}$

Battery $\in [0, \infty)$

The function that determines the **probability** value is called the “distribution” of that RV

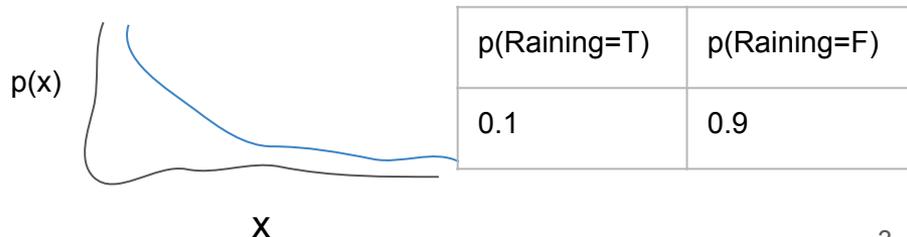
How do we know the distribution? Two ways:

Sampling by observing many times

- Expected utility in RL

Provided as part of the problem description

- $p(\text{Raining}) = \langle 0.1, 0.9 \rangle$
- $p(\text{Battery} > x) = e^{-x}$



Rules of Probability

1. The probability that a RV takes on some value is always between 0 and 1

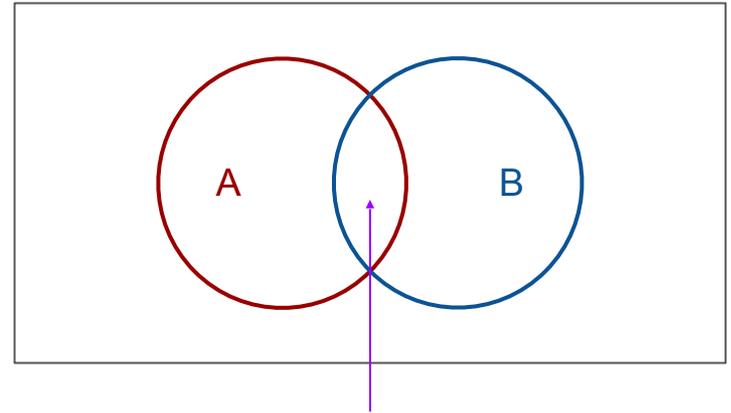
$$0 \leq p(X = x) \leq 1$$

2. Probability of deterministic events

$$p(\text{TRUE}) = 1, \quad p(\text{FALSE}) = 0$$

3. Additivity

$$p(A \text{ or } B) = p(A) + p(B) - p(A \text{ and } B)$$



A and B

$$p(A = \text{TRUE} \vee A = \text{FALSE}) = p(A = \text{TRUE}) + p(A = \text{FALSE}) - p(A = \text{TRUE} \wedge A = \text{FALSE})$$

$$1 = p(A = \text{TRUE}) + p(A = \text{FALSE}) - 0$$

$$1 - p(A = \text{FALSE}) = p(A = \text{TRUE})$$

Interactions between Random Variables

With multiple interacting RVs, the probability distribution that includes **all** of them together is called the **joint distribution**

Example: $p(\text{Cavity}, \text{Toothache}, \text{Catch})$

Notation

$p(X)$: distribution, function or table
 $p(X=x)$: probability, single number

	Toothache		\neg Toothache	
	Catch	\neg Catch	Catch	\neg Catch
Cavity	0.108	0.012	0.072	0.008
\neg Cavity	0.016	0.064	0.144	0.576

Facts: table sums to 1, all combinations of all RVs, one cell per “configuration”

$$p(\text{Toothache}=F, \text{Cavity}=F, \text{Catch}=F) = 0.576$$

Events & Marginalization

- An **event** is a setting of some subset of random variables
- You can use the joint distribution to compute the probability of any event by **adding up** all the table entries that correspond with the given configuration

Marginalization

$$p(\text{Cav}=T) = 0.108 + 0.012 + 0.072 + 0.008 \\ = 0.2$$

$$p(\text{Cav}=T \text{ or } \text{Tooth}=T) = 0.108 + 0.012 + 0.072 + 0.008 + 0.016 + 0.064 = 0.28$$

$$p(\text{Cav}=T \text{ and } \text{Tooth}=T) = 0.108 + 0.012 = 0.12$$

	Tooth		¬Tooth	
	Cat	¬Cat	Cat	¬Cat
Cav	0.108	0.012	0.072	0.008
¬Cav	0.016	0.064	0.144	0.576

Conditional Probability (1)

- When one Random Variable **impacts** the value of another
“If $X=v$, what is the probability of $Y=w$?”

$$p(Y = w \mid X = v) = \frac{p(Y = w \text{ and } X = v)}{p(X = v)}$$

“Conditioned variable”

Notation

$p(Y|X=v)$: distribution, function or table
 $p(Y=w|X=v)$: probability, single number

Conditional Probability (2)

Typically, an agent wants to know conditional probabilities

$$p(S = s \mid E = e)$$



	Tooth		\neg Tooth	
	Cat	\neg Cat	Cat	\neg Cat
Cav	0.108	0.012	0.072	0.008
\neg Cav	0.016	0.064	0.144	0.576

Example:

$$p(\text{Cav}=\text{T} \mid \text{Tooth}=\text{T}) = p(\text{Cav}=\text{T} \text{ and } \text{Tooth}=\text{T})/p(\text{Tooth}=\text{T})$$

$$= (0.108+0.012)/(0.108+0.012+0.016+0.064) = 0.6$$

$$p(\text{Cav}=\text{F} \mid \text{Tooth}=\text{T}) = (0.016+0.064)/(0.108+0.012+0.016+0.064) = 0.4$$

Normalization

$$\begin{aligned} p(\text{Cav}=\text{T} \mid \text{Tooth}=\text{T}) &= 0.6 \\ p(\text{Cav}=\text{F} \mid \text{Tooth}=\text{T}) &= 0.4 \end{aligned}$$

Notice that the two probabilities summed to 1 across Cavity (**not** Toothache)

In general, we don't have to know/compute the denominator, we can "normalize"

$$p(X \mid E = e) = \frac{p(X, E = e)}{p(E = e)} = \alpha \cdot p(X, E = e)$$

$$1 = \alpha \cdot p(X = \text{TRUE}, E = e) + \alpha \cdot p(X = \text{FALSE}, E = e)$$

$$1 = \alpha [p(X = \text{TRUE}, E = e) + p(X = \text{FALSE}, E = e)]$$

$$\alpha = \frac{1}{[p(X = \text{TRUE}, E = e) + p(X = \text{FALSE}, E = e)]}$$

Hidden Variables

Frequently our agent won't have settings for **all** of the random variables.

Solution? Sum them out!

$$p(X \mid E = e) = \alpha \cdot p(X, E = e) = \alpha \sum_{h \in H} p(X, E = e, H = h)$$

In the previous example, **Catch** was a **hidden** variable:

$$p(\text{Cav}=T \mid \text{Tooth}=T) = \alpha p(\text{Cav}=T, \text{Tooth}=T, \text{Cat}=T) + \alpha p(\text{Cav}=T, \text{Tooth}=T, \text{Cat}=F)$$

Example - marginalization and normalization

What's the probability you have a cavity if you don't have a toothache?

$$\begin{aligned} p(\text{Cav}|\text{Tooth}=F) &= \alpha p(\text{Cav}, \text{Tooth}=F) = \alpha \sum_{h=\{T,F\}} p(\text{Cav}, \text{Tooth}=F, \text{Catch}=h) \\ &= \alpha [p(\text{Cav}, \text{Tooth}=F, \text{Cat}=T) + p(\text{Cav}, \text{Tooth}=F, \text{Cat}=F)] \\ &= \alpha [\langle 0.072, 0.144 \rangle + \langle 0.008, 0.576 \rangle] \\ &= \alpha \langle 0.08, 0.72 \rangle \end{aligned}$$

Must sum to 1, so

$$\alpha * (0.08+0.72) = \alpha * (0.8) = 1$$

$$\alpha = 1.25$$

$$p(\text{Cav}|\text{Tooth}=F) = \langle 0.1, 0.9 \rangle$$

	Tooth		¬Tooth	
	Cat	¬Cat	Cat	¬Cat
Cav	0.108	0.012	0.072	0.008
¬Cav	0.016	0.064	0.144	0.576

Conditioning

- Given full joint probability, agent can compute anything about the RVs!
- Usually, agent only has access to some conditional probabilities, not their joint distribution
- Can get around this by marginalizing and leveraging the definition of conditional probability

Ex: How can we get $p(X)$ (marginal) if we don't know $p(X, Y)$, just $p(X|Y)$ and $p(Y)$?

Cond. Prob. def. \longrightarrow $p(X | Y) = \frac{p(X, Y)}{p(Y)}$

$$p(X | Y) \cdot p(Y) = p(X, Y)$$

Marginalize Y \longrightarrow $\sum_y p(X | Y = y)p(Y = y) = \sum_y p(X, Y = y) = p(X)$

Independence

Two random variables are **independent** if and only if their joint probability is **the same** as the product of their marginals

$$\begin{aligned}A \perp B &\iff \\p(A, B) &= p(A)p(B) \\p(A | B) &= p(A)\end{aligned}$$

Similarly, two random variables can be **conditionally independent** given a third

$$\begin{aligned}A \perp B | C &\iff \\p(A, B | C) &= p(A | C)p(B | C) \\p(A | B, C) &= p(A | C)\end{aligned}$$

Independence Example

Are Catch and Toothache independent?

$$p(\text{Catch}=\text{T}, \text{Toothache}=\text{T}) = 0.124$$

$$p(\text{Cat}=\text{T}) * p(\text{Tooth}=\text{T}) = (0.108 + 0.016 + 0.072 + 0.144) * (0.108 + 0.012 + 0.016 + 0.064) \\ = (0.34) * (0.2) = 0.068 \neq 0.124 \text{ (NOT INDEPENDENT)}$$

What about conditioned on Cavity?

$$p(\text{Cat}=\text{T}, \text{Tooth}=\text{T} | \text{Cav}=\text{T}) = 0.54$$

$$p(\text{Cat}=\text{T} | \text{Cav}=\text{T}) * p(\text{Tooth}=\text{T} | \text{Cav}=\text{T}) = 0.9 * 0.6 = 0.54$$

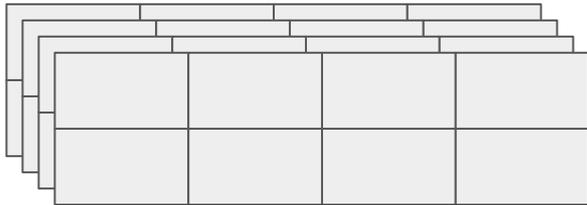
... also checks out for each setting of Cat, Cav, and Tooth
(CONDITIONALLY INDEPENDENT given Cavity)

	Tooth		¬Tooth	
	Cat	¬Cat	Cat	¬Cat
Cav	0.108	0.012	0.072	0.008
¬Cav	0.016	0.064	0.144	0.576

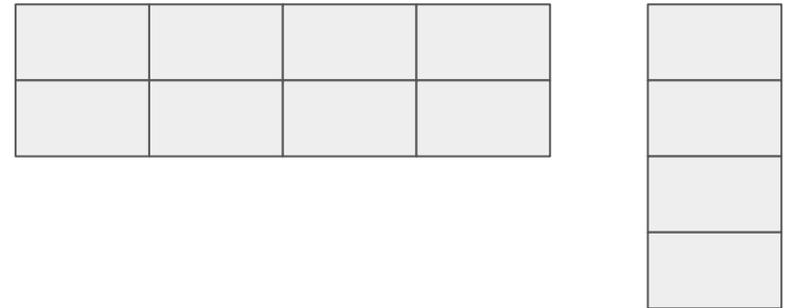
Factoring the joint probability with independence

- Suppose we add Weather as a RV to the Cav,Cat,Tooth model, where whether can take on 4 values: {Sunny, Rainy, Foggy, Snowy}

$$p(\text{Cav,Tooth,Catch,Weather}) = p(\text{Cav,Tooth,Catch}) * p(\text{Weather})$$



$2 \times 2 \times 2 \times 4 = 32$ cells



$2 \times 2 \times 2 + 4 = 12$ cells

Bayes Rule

Additional TRUE FACT about conditional probabilities

$$p(A | B) = \frac{p(B | A) \cdot p(A)}{p(B)}$$

Follows by plugging in definition of conditional probability

$$p(B | A) \cdot p(A) = p(A, B) = p(A | B)p(B)$$

This generalizes to more than two RVs, and gives us the **product rule**

$$p(A, B, C) = p(A, B | C) \cdot p(C) = p(A | B, C) \cdot p(B | C) \cdot p(C)$$

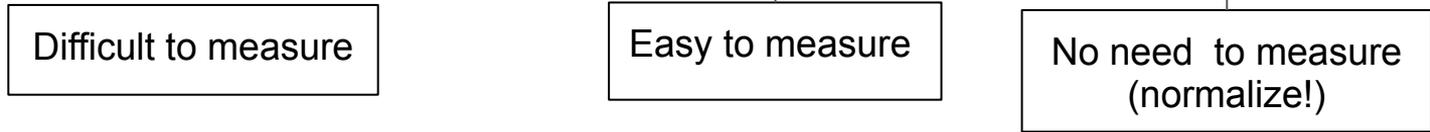
$$p(X_1, X_2, \dots, X_d) = p(X_1 | X_2, \dots, X_d) \cdot p(X_2 | X_3, \dots, X_d) \cdot \dots \cdot p(X_{d-1} | X_d) \cdot p(X_d)$$

Note: this is true for **any** ordering of the X_j !

Using Bayes Rule

We can use Bayes Rule to “flip” the conditioning variable

$$p(\text{CAUSE} \mid \text{EFFECT}) = p(\text{EFFECT} \mid \text{CAUSE}) \frac{p(\text{CAUSE})}{p(\text{EFFECT})}$$



Example: Medical Diagnosis “What’s the probability I have the flu given I have a cough?”

$$p(\text{Flu} \mid \text{Cough}=\text{T}) = \alpha \langle p(\text{Cough}=\text{T} \mid \text{Flu}=\text{T}) p(\text{Flu}=\text{T}), p(\text{Cough}=\text{T} \mid \text{Flu}=\text{F}) p(\text{Flu}=\text{F}) \rangle$$

Incorporating multiple pieces of evidence

“What’s the probability of Cavity given Catch and Toothache?”

Applying bayes rule and normalizing:

$$p(\text{Cav} \mid \text{Cat}, \text{Tooth}) = \alpha p(\text{Tooth}, \text{Cat} \mid \text{Cav}) p(\text{Cav})$$

But since Catch and Tooth are conditionally independent given Cavity

$$p(\text{Cav} \mid \text{Cat}, \text{Tooth}) = \alpha p(\text{Tooth} \mid \text{Cav}) p(\text{Cat} \mid \text{Cav}) p(\text{Cav})$$

IF our **evidence** variables are conditionally independent from one another **given** the **cause** variable, we can significantly simplify things!

Summary and preview

Wrapping up

- To handle uncertain sensors, state now describes **probabilities**
- The joint probability distribution contains **all** the information we need to answer any question about any subset of the random variables it describes
- Marginalization, Normalization, and Bayes Rule are all “probability tricks” we can use to simplify/compute specific probabilities

Next time

- Detailed example: Avoid the wumpus!