Reasoning under Uncertainty: Introduction to Probability

CPSC 322 Lecture 23

Learning Goals for today's class

You can:

- Define and give examples of **random variables**, their domains and **probability distributions**.
- Calculate the **probability of a proposition f** given $\mu(w)$ for the set of possible worlds.
- Define a **joint probability distribution**

Lecture Overview

- Big Transition
- Intro to Probability
- ….

R&R systems we'll cover in this course

		Environment	
Problem		**Deterministic**	**Stochastic**
Static	Constraint Satisfaction	*Variables + Constraints* Search Arc Consistency Local Search	
	Query	*Logics* Search	*Bayesian (Belief) Networks* Variable Elimination
Sequential	Planning	*STRIPS* Search	*Decision Networks* Variable Elimination

Representation
Reasoning Technique

Intro to Probability (Motivation)

- *What will the temperature be tomorrow? In 10 days? In 50 years?*
- *Right now, how many people are in this room? in this building (DMP)? At UBC? ….Yesterday?*
- *The fire alarm just went off. Is there a fire?*

- AI agents (and humans ☹) are not omniscient
- And the problem is not only predicting the future or “remembering” the past

Intro to Probability (Key points)

- Are agents all ignorant/uncertain to the same degree?

- Should an agent act only when it is certain about relevant knowledge?
  - Can it ever be certain?
  - (not acting often has implications)

- So agents need to **represent and reason about their ignorance/ uncertainty**

Probability as a formal measure of uncertainty/ignorance

- We can use **probabilities** to represent **belief**
- Consider a simple dice roll
- Suppose we want to know whether we rolled a **6**
  - What is the probability **P(6)**?
- Suppose I tell you that we rolled an even number
  - This **evidence** forces us to **update our beliefs**
  - What is the new conditional probability **P(6|even)**?
- I saw what we rolled
  - What is my conditional probability **P(6|saw_number)**?

Probability as a formal measure of uncertainty/ignorance

- Belief in a proposition *f* (e.g., *it is raining outside, there are 31 people in this room*) can be measured in terms of a number between 0 and 1 – this is the probability of *f*
  - The probability *f* is 0 means that *f* is believed to be…
  - The probability *f* is 1 means that *f* is believed to be…
  - Using 0 and 1 is purely a **convention** (though a useful one!)

Random Variables

- A random variable is a **variable** (like the ones we have seen in **CSP** and **Planning**), but the agent can be **uncertain about its value**.
- As usual
  - The domain of a random variable *X*, written *dom(X),* is the set of values *X* can take
  - values are mutually exclusive and exhaustive

Examples (Boolean and discrete)

Random Variables (cont')

- A tuple of random variables <X_1 ,….., X_n> is a complex random variable with domain

$$\mathrm{dom}(X_1) \times \ldots \times \mathrm{dom}(X_n)$$

- Assignment $X = x$ means X has value x

$$outsideRaining = T$$

- A proposition is a Boolean formula made from assignments of values to variables

Examples

outsideRaining = T v peopleInRoom = 42

mostUsefulThing = “towel”

Possible Worlds

- A possible world specifies an assignment to each random variable

E.g., if we model only two Boolean variables *Cavity* and *Toothache*, then there are 4 distinct possible worlds:

$Cavity = T \wedge Toothache = T$
$Cavity = T \wedge Toothache = F$
$Cavity = F \wedge Toothache = T$
$Cavity = F \wedge Toothache = F$

cavity	*toothache*
T	T
T	F
F	T
F	F

As usual, possible worlds are mutually exclusive and exhaustive

$w \models X=x$ means variable X is assigned value x in world w

Semantics of Probability

- The belief of being in each possible world *w* can be expressed as a probability *μ(w)*
- For sure, I must be in one of them……so

$$\sum_{w \in W} \mu(w) = 1$$

μ(w) for possible worlds generated by three Boolean variables: ***cavity***, ***toothache***, ***catch*** *(dentist probe catches in the tooth)*

cavity	*toothache*	*catch*	*μ(w)*
T	T	T	.108
T	T	F	.012
T	F	T	.072
T	F	F	.008
F	T	T	.016
F	T	F	.064
F	F	T	.144
F	F	F	.576

Probability of proposition

- What is the **probability of a proposition *f*** ?

cavity	*toothache*	*catch*	$\mu(w)$
T	T	T	.108
T	T	F	.012
T	F	T	.072
T	F	F	.008
F	T	T	.016
F	T	F	.064
F	F	T	.144
F	F	F	.576

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ *cavity*	**.016**	**.064**	**.144**	**.576**

For any *f*, **sum** the prob. of the worlds where f is **true**:

$$P(f) = \Sigma_{w \models f} \mu(w)$$

Ex: P(*toothache* = *T*) =

Probability of proposition

- What is the **probability of a proposition *f*** ?

cavity	*toothache*	*catch*	$\mu(w)$
T	T	T	.108
T	T	F	.012
T	F	T	.072
T	F	F	.008
F	T	T	.016
F	T	F	.064
F	F	T	.144
F	F	F	.576

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ *cavity*	**.016**	**.064**	**.144**	**.576**

For any *f*, sum the prob. of the worlds where it is true:

$$P(f) = \sum_{w \models f} \mu(w)$$

P(cavity=T and toothache=F) =

A. 0.2 B. 0.8 C. 0.000576 D. 0.08 E. 0.42

Probability of proposition

- What is the **probability of a proposition *f*** ?

cavity	*toothache*	*catch*	$\mu(w)$
T	T	T	.108
T	T	F	.012
T	F	T	.072
T	F	F	.008
F	T	T	.016
F	T	F	.064
~~F~~	~~F~~	~~T~~	~~.144~~
~~F~~	~~F~~	~~F~~	~~.576~~

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ *cavity*	**.016**	**.064**	~~**.144**~~	~~**.576**~~

For any *f*, sum the prob. of the worlds where it is **true**:

$$P(f)=\sum_{w \models f} \mu(w)$$

P(cavity **or** *toothache*) = .108 + .012 + .016 + .064 +
+ .072 + .08 = **.28**
(= 1- (.144 + .576))

One more example

- *Weather*, with domain {sunny, cloudy)
- *Temperature*, with domain {hot, mild, cold}
- There are now 6 possible worlds:
- What's the probability of it being cloudy **or** cold?

Weather	*Temperature*	*µ(w)*
sunny	hot	0.10
sunny	mild	0.20
sunny	cold	0.10
cloudy	hot	0.05
cloudy	mild	0.35
cloudy	cold	0.20

A. 1 B. 0.3

C. 0.6 D. 0.7 E. 0.42

- Remember
  - The probability of proposition f is defined by: $P(f) = \Sigma_{w \models f}\ \mu(w)$
  - sum of the probabilities of the worlds w in which *f* is true

Probability Distributions

- A probability distribution **P** on a random variable **X** is a function $dom(X) \to [0,1]$ such that

$$x \to P(X=x)$$

cavity? *P(cavity=T)=*

P(cavity=F)=

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	.108	.012	.072	.008
¬ cavity	.016	.064	.144	.576

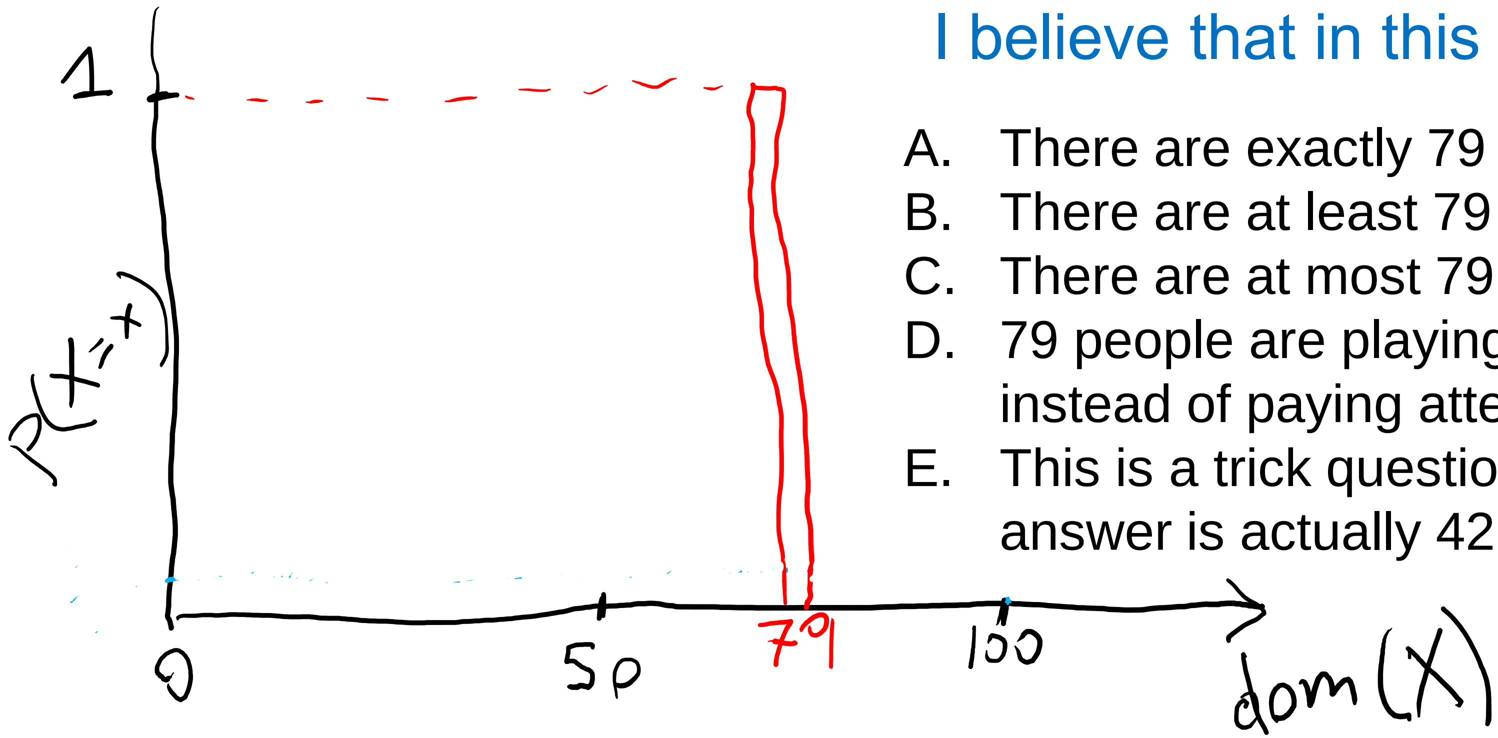
cavity	*toothache*	*catch*	$\mu(w)$
T	T	T	.108
T	T	F	.012
T	F	T	.072
T	F	F	.008
F	T	T	.016
F	T	F	.064
F	F	T	.144
F	F	F	.576

Probability distribution (non binary)

- A probability distribution P on a random variable X is a function *dom(X)* - > *[0,1]* such that

$$x \rightarrow P(X=x)$$

- Number of people in this room at this time



I believe that in this room….

A. There are exactly 79 people
B. There are at least 79 people
C. There are at most 79 people
D. 79 people are playing Hearthstone instead of paying attention
E. This is a trick question, and the answer is actually 42

Probability distribution (non binary)

- A probability distribution P on a random variable X is a function $dom(X) \to [0,1]$ such that

  $$x \to P(X=x)$$

- Number of people in this room at this time
- *Let's represent 3 different kinds of belief:*
  - *complete certainty*
  - *perfect uncertainty*
  - *some reasonable guess*

Joint Probability Distributions

- When we have **multiple random variables**, their joint distribution is a probability distribution over the variables' Cartesian product
  - E.g., $P(<X_1 , \ldots., X_n>)$
  - You can think of a joint distribution over n variables as an n-dimensional table
  - Each entry, indexed by $X_1 = x_1, \ldots., X_n = x_n$ corresponds to $P(X_1 = x_1 \wedge \ldots. \wedge X_n = x_n)$
  - The sum of entries across the whole table is 1

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	.108	.012	.072	.008
¬ cavity	.016	.064	.144	.576

Questions

Suppose you have the joint probability distribution of n variables.

- Can you compute the probability distribution for each variable?
- Can you compute the probability distribution for any combination of variables?
- Can you update these probabilities if you know something about some of the variables?
- Is there a downside to the joint probability distribution?

Next Class

More probability theory

- Marginalization
- Conditional Probability
- Chain Rule
- Bayes' Rule
- Independence