Last Time:

- ☐ values & Q-values
- ☐ hands-on exercise

Lecture 4

HRI, Fall '24

Andrea Bajcsy

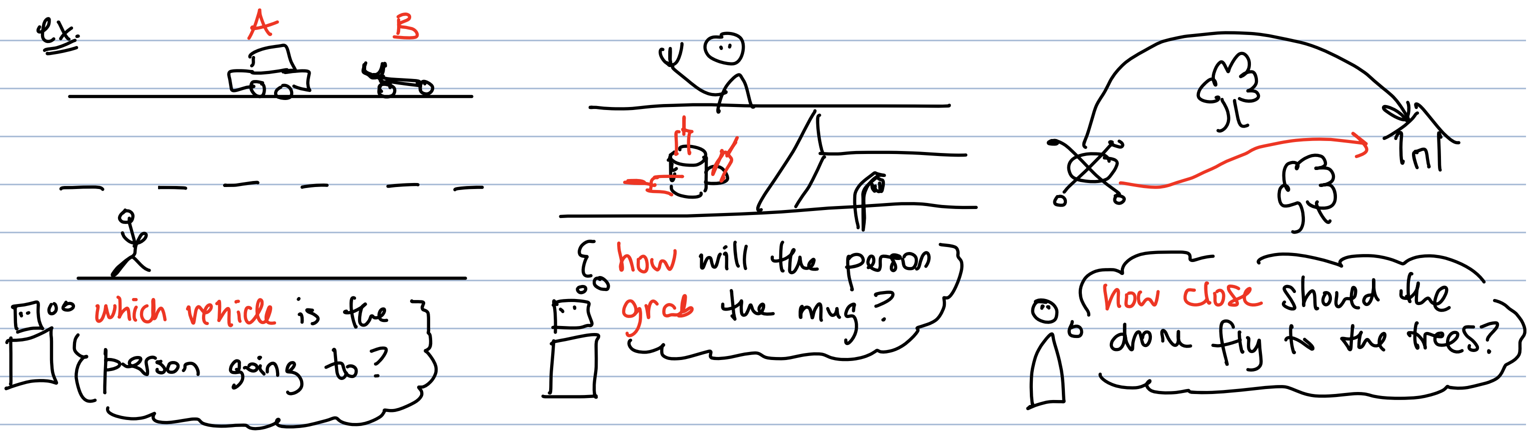
Today:

- ☐ Probability
- ☐ Bayes' Rule

Uncertainty

So far, we have assumed that we are <u>certain</u> about – for example – the reward function that an agent is optimizing. In reality, when we interact with other people, or the world, we may not know what their reward is
goals, intent

ex.



lets model this type of uncertainty + recap some basic probability along the way. In probability theory, we want to make probabilistic assertions about possible "worlds".

NOTATION

$P(x)$ is probability of random variable x

e.g. $P(\text{cloudy})$, $\text{cloudy} \in \{\text{TRUE}, \text{FALSE}\}$

$P(x, y)$ is probability of x and y. Also called <u>joint distribution</u>

e.g. $P(\text{cloudy}, \text{rain})$

$P(x \mid y)$ is probability of x given y.

e.g. $P(\text{cloudy} \mid \text{rain})$

USEFUL RULES

- $P(x,y) = P(x|y)P(y)$ is called <u>product rule</u> or <u>chain rule</u>

  <u>e.g.</u> for it to be cloudy and rain (i.e. $P(\text{cloudy}, \text{rain})$) we need it to rain (i.e. $P(\text{rain})$) and we need it to be cloudy given rain (i.e. $P(\text{cloudy} | \text{rain})$)

$$\begin{aligned} P(A, \underbrace{B, C, D}) &= P(A|B,C,D)\,P(B,C,D) \\ &= P(A|B,C,D)\cdot P(B|C,D)\,P(C,D) \\ &= P(A|B,C,D)\cdot P(B|C,D)\,P(C|D)\,P(D) \end{aligned}$$

- $P(x) = \sum_{y \in Y} P(x,y)$ is <u>marginalization</u> (or "summing out") sums up probabilities of each possible value of a target variable (e.g. y), taking it out of the equation

$P(\text{cloudy}, \text{rain})$

	rain = T	rain = F
cloudy = T	0.5	0.05
cloudy = F	0.15	0.30

$\longrightarrow$ $P(\text{cloudy}) = \sum_{\text{rain} \in \{T,F\}} P(\text{cloudy}, \text{rain})$

$P(\text{cloudy})$

cloudy = T	0.55
cloudy = F	0.45

- <u>independence</u> is when one random variable's probability doesn't depend on the other.

$$\begin{aligned} P(x,y) &= P(x|y)P(y) \\ &= P(x)P(y) \end{aligned}$$ if x & y independent!

  <u>eg.</u> $P(\text{cloudy}, \text{heads}) = P(\text{cloudy})\,P(\text{heads})$

- <u>conditional independence</u>: if x and y are conditionally independent given z.

$$P(x,y|z) = P(x|z)\,P(y|z)$$ ← if x & y are C.I. given z.

<u>e.g.</u> $P(\text{keys}, \text{wallet} | \text{drive}) = P(\text{keys} | \text{drive})\,P(\text{wallet} | \text{drive})$

BAYES' THEOREM

What if we know the probability of x given y ($P(x|y)$), but, we want to know probability of y given x (i.e. $P(y|x)$)...

goal; e.g. action; e.g. goal; action

Bayes Theorem lets us establish this relationship! It underlies a lot of the mathematical human models we will see in class.

goal action action goal

$$\boxed{P(y|x) = \frac{P(x|y)P(y)}{P(x)}}$$

Q: can you derive this with the rules we have seen so far?

HINT: start from $P(x,y)$ ....

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

$\Downarrow$ divide by $P(x)$

$$\frac{P(x,y)}{P(x)} = P(y|x) \quad \xRightarrow[\text{apply chain rule to numerator}]{} \quad \frac{P(x|y)P(y)}{P(x)} = P(y|x)$$

What we will see is that each component of Bayes' Theorem has a name:

(observation model)
likelihood function

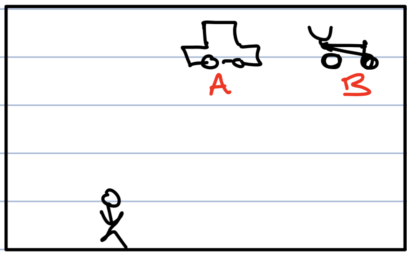
Prior

$$\underbrace{P(y|x)}_{\text{posterior}} = \frac{P(x|y)P(y)}{P(x)}$$

posterior

This rule helps us perform automatic inference over what humans want!

ex.



we want to infer which vehicle the person wants to move to (A or B). Let's use Bayes' Theorem to help us out!

$\mathcal{A} = \{\uparrow, \rightarrow\}$

$P(\text{goal})$ — A, B

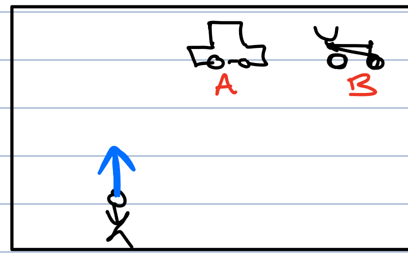
goal = A	goal = B
0.5	0.5

Now, I also give you an observation model, which tells you "how likely is a person's action, given a goal"?

$P(a \mid \text{goal})$

$P(a=\uparrow \mid \text{goal}=A) = 0.8$
$P(a=\rightarrow \mid \text{goal}=A) = 0.2$
$P(a=\uparrow \mid \text{goal}=B) = 0.4$
$P(a=\rightarrow \mid \text{goal}=B) = 0.6$

Now you OBSERVE $a = \uparrow$



Q $P(\text{goal} \mid a=\uparrow) = \dfrac{P(a=\uparrow \mid \text{goal})\,P(\text{goal})}{P(a=\uparrow)}$ // Bayes' Rule

$$= \frac{P(a=\uparrow \mid \text{goal})\,P(\text{goal})}{\sum_{\text{goal}\in\{A,B\}} P(a=\uparrow, \text{goal})}$$ // marginalize over goals

$$= \frac{P(a=\uparrow \mid \text{goal})\,P(\text{goal})}{\sum_{\text{goal}\in\{A,B\}} P(a=\uparrow \mid \text{goal})\,P(\text{goal})}$$ // apply product rule

$P(\overset{A,\,B}{goal})$ prior

goal = A	goal = B
0.5	0.5

$P(goal \mid a = \uparrow)$ posterior

goal = A	goal = B
$0.6\overline{6}$	$0.3\overline{3}$

consider goal = A

$$P(goal = A \mid a = \uparrow) = \frac{\overbrace{P(a=\uparrow \mid goal = A)}^{0.8}\ \overbrace{P(goal = A)}^{0.5}}{\underbrace{P(a=\uparrow \mid goal = A)}_{0.8}\ \underbrace{P(goal = A)}_{0.5} + \underbrace{P(a=\uparrow \mid goal = B)}_{0.4}\ \underbrace{P(goal = B)}_{0.5}}$$

$$= 0.6\overline{6}$$

$\boxed{\text{Boltzmann Rational Model}}$ → [Luce, 1959, 1977; Ziebart, 2010]

Assume people optimize some reward function (e.g. goal location).

This model says that people behave approximately optimally (or, approximately rationally) in pursuit of their goals.

action $(\uparrow, \rightarrow, \downarrow, \leftarrow)$

"parameterized by"

$$P(a \mid s \,;\, goal) = \frac{e^{Q(s,a;g)}}{\sum_{\bar{a} \in \mathcal{A}} e^{Q(s,\bar{a};g)}}$$

e.g. in grid world