ARTIFICIAL INTELLIGENCE

Quantifying Uncertainty

NOTICE

Presentation slides will be available for download along with the recording of this review session.

You are free to take pictures regardless if you'd like.

DISCLAIMER

I have tried to pull all information from the slides, the textbook, and other authorized resources, but I cannot guarantee the veracity of any information in the slides hereafter.

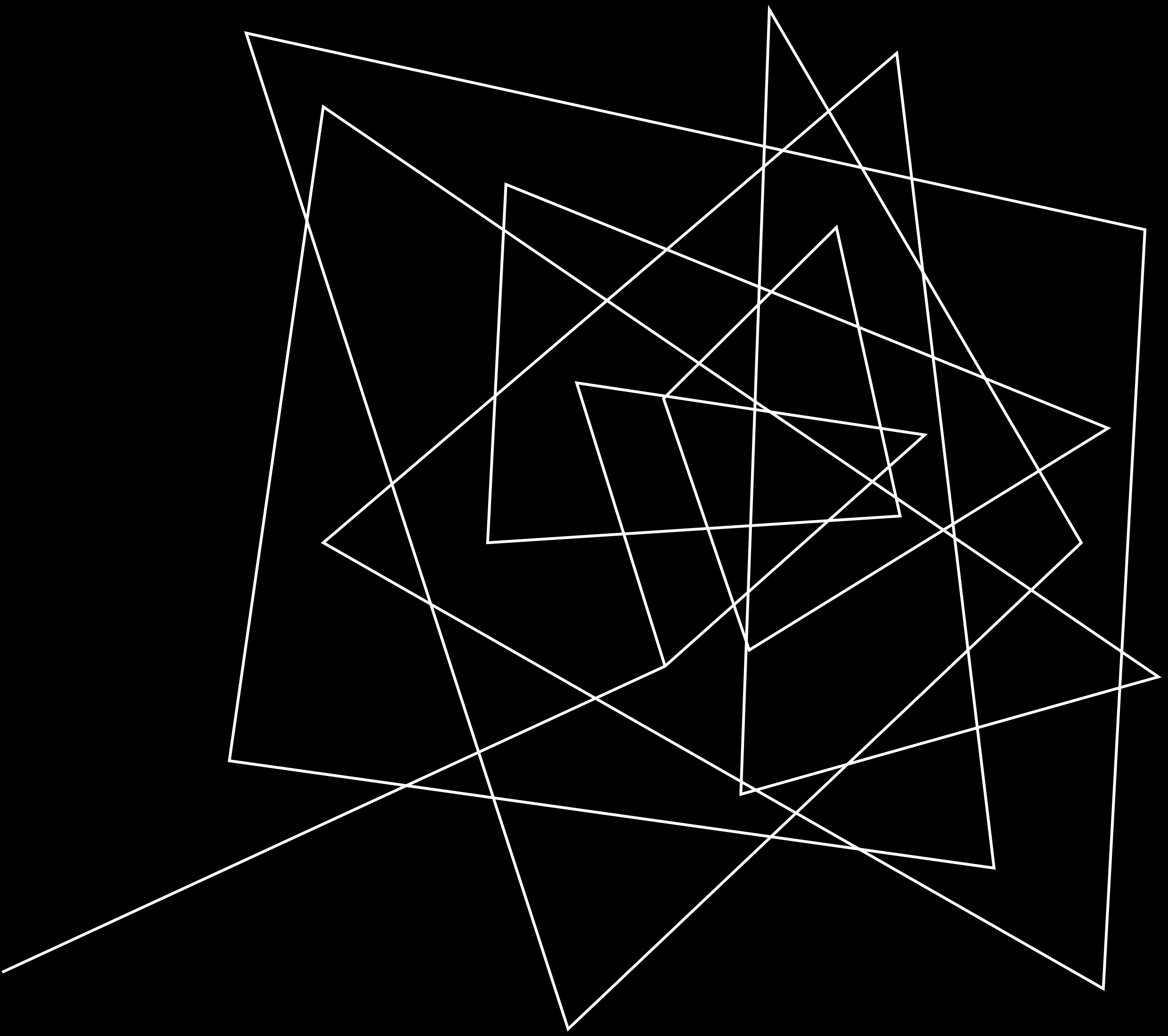
See presentation notes for sources

EDITION DIFFERENCES

- Chapter 13 in 3rd edition is now chapter 12 in 4th edition.
- I will continue to use 4th edition pgs.
- The chapter we're covering is "Quantifying Uncertainty."

OUTLINE

- Uncertainty
- Probability
- Syntax and Semantics
- Inference



UNCERTAINTY

When we don’t know everything

UNCERTAINTY

- **Uncertainty** in driving to school:
  - Traffic

UNCERTAINTY

- **Uncertainty** in driving to school:
  - Traffic
  - Flat/deflated tire

UNCERTAINTY

- **Uncertainty** in driving to school:
  - Traffic
  - Flat/deflated tire
  - FIT Parking

UNCERTAINTY

- **Uncertainty** in driving to school:
  - Traffic
  - Flat/deflated tire
  - FIT Parking
  - When your car makes a weird noise and you're pretty sure it's fine but you let it run for a while to make sure there's no popping or burning sounds, but then the car shudders, so like what is that supposed to mean? I'm a CS major, not a mechanic for heaven's sake! I don't

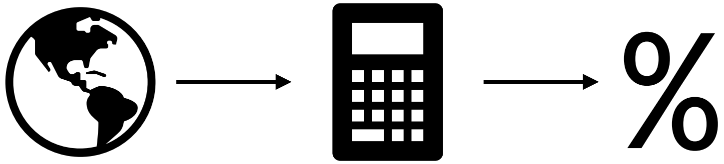
INTO THE UNKNOWN (ELSA WOULD BE PROUD)

Default/Nonmonotonic Logic

Fudge Factors

$$Sprinkler \longmapsto_{0.99} WetGrass$$
$$WetGrass \longmapsto_{0.7} Rain$$

Probability



THE UTILITY OF PROBABILITY

- Probabilistic assertions summarize effects of **laziness** and **ignorance**.
- In **Bayesian** or **subjective** probability, probabilities relate propositions to one's own state of knowledge.
  - This means probabilities change with new evidence.

PARKING WITH BAYES

$OpenSpot$ = The event where there's an open spot between the Olin buildings and Clemente/PDH

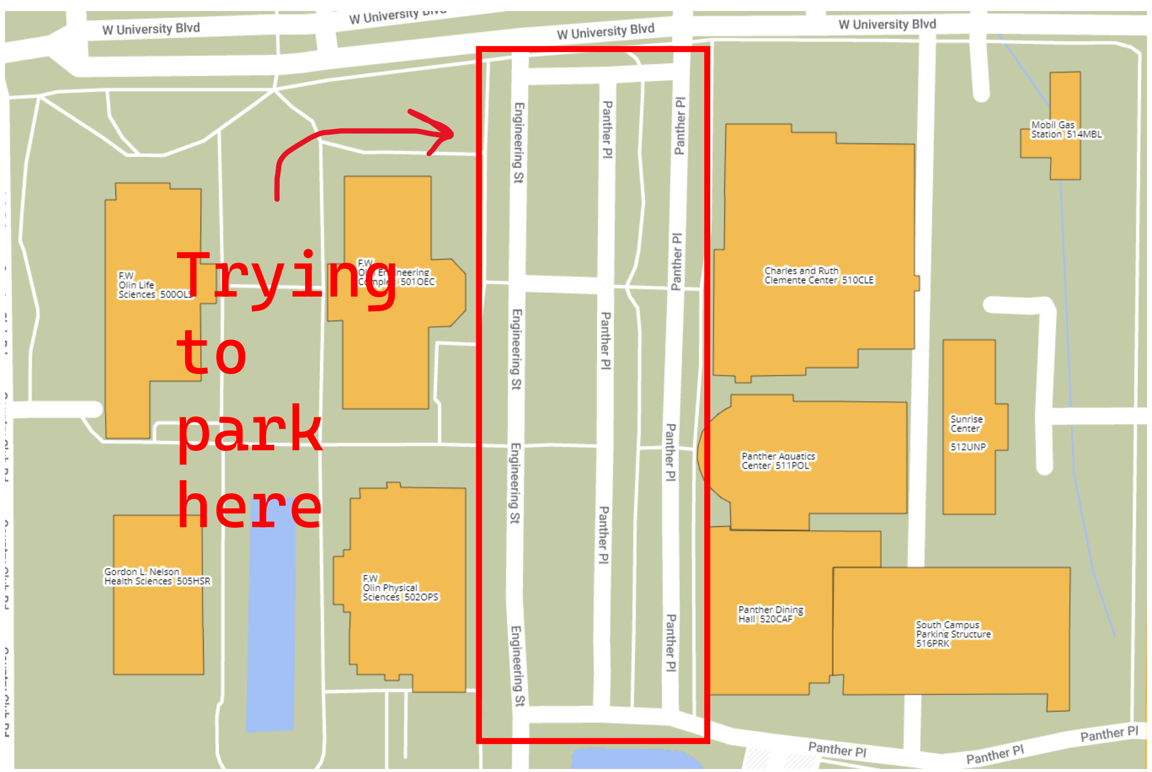
$$P(OpenSpot|\text{no events, 9:50 a.m.}) = 0.05$$

$$\Big\downarrow \text{ Driving into campus}$$

$$P(OpenSpot|\text{career fair/senior design/other event, 9:50 a.m.}) = 10^{-6}$$

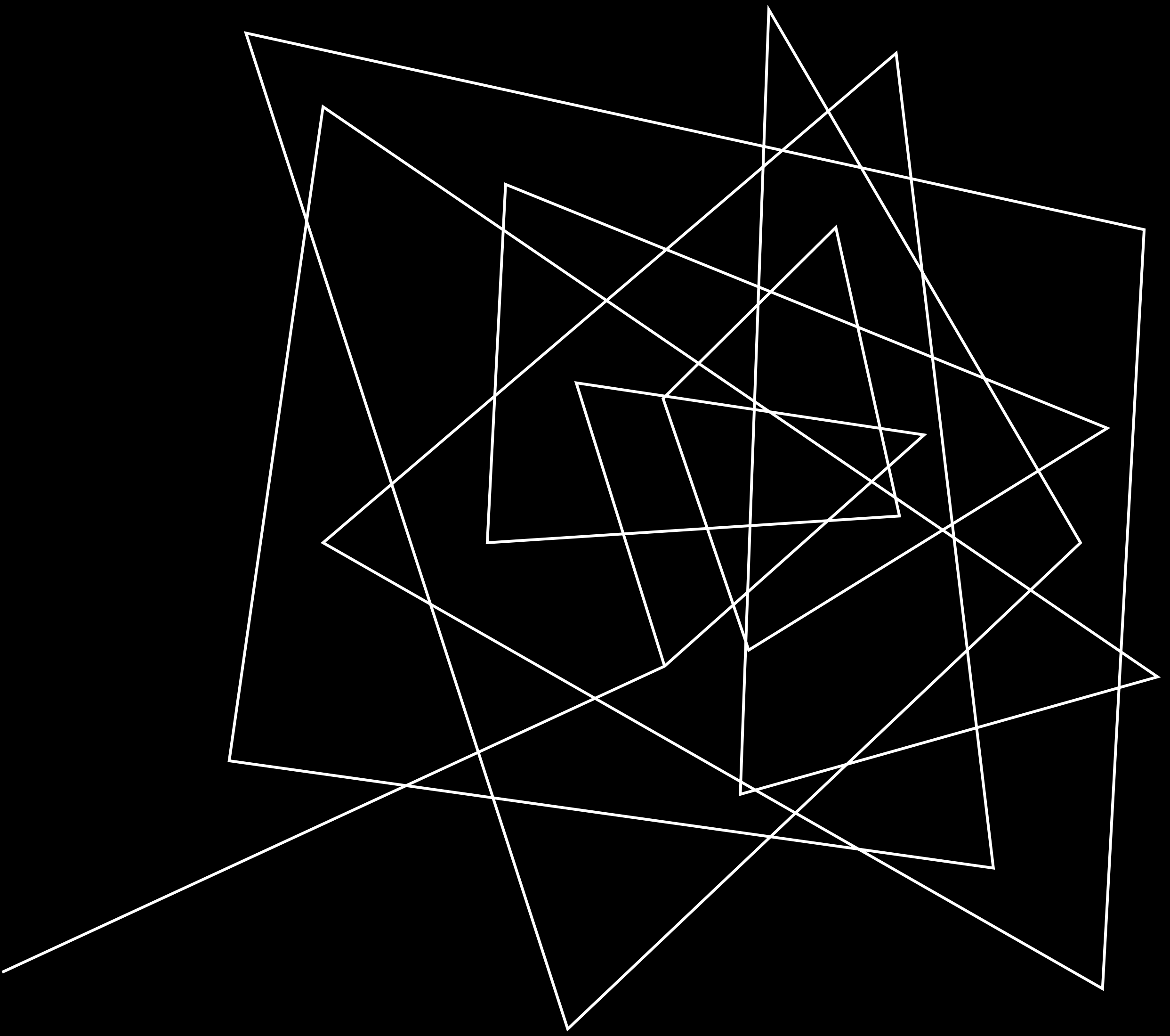
BUT THAT'S JUST A THEORY
(A DECISION THEORY)

BATTLEMAP



Time of Arrival	Probability of Success
9:50	0.04
8:50	0.7
6:00	0.95
2:00	0.9999

Utility Theory = Preferences
Decision Theory = Utility Theory + Probability Theory

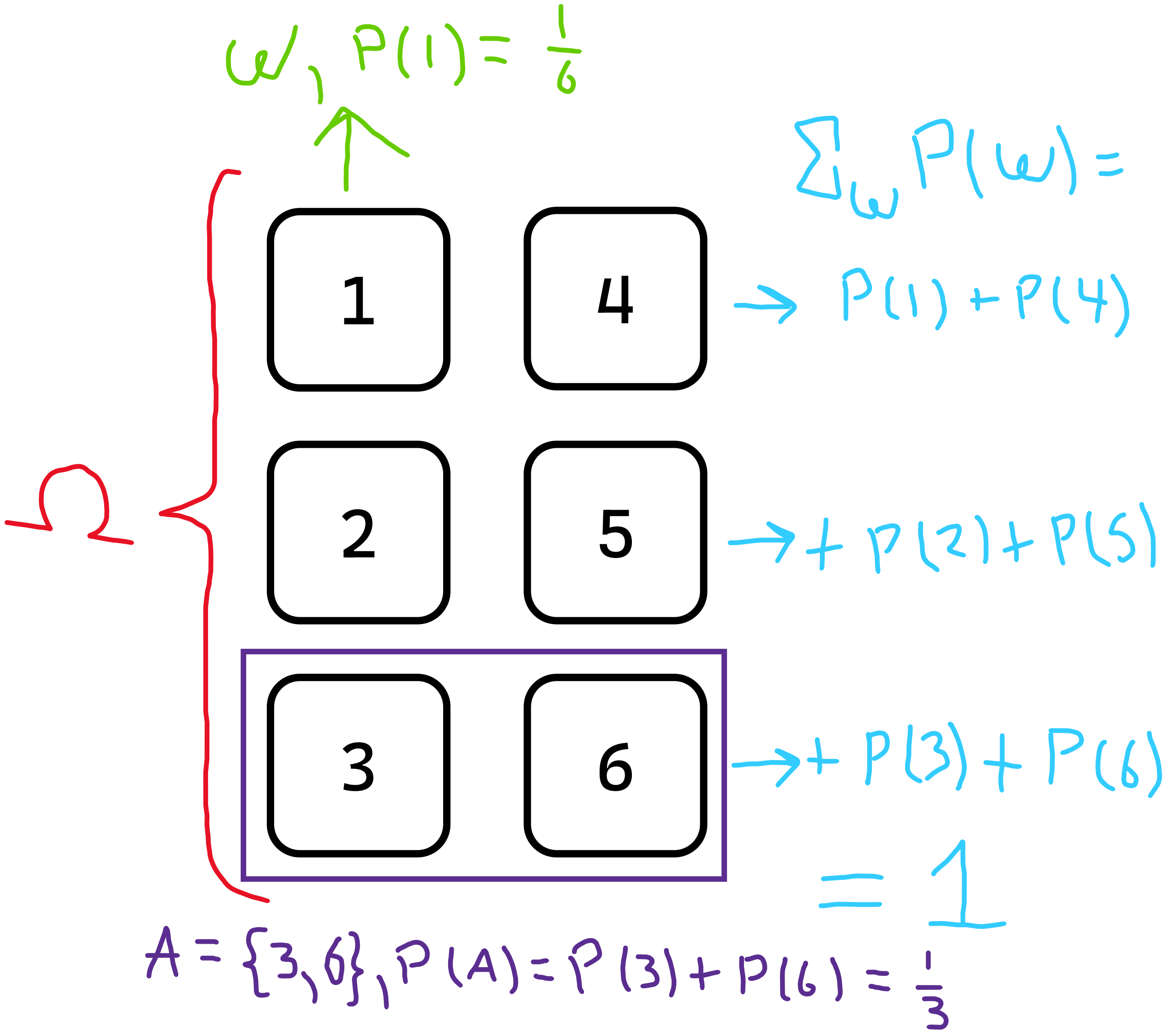


PROBABILITY

They never want you to tell them the odds,
but what about the evens?

INTRO TO ~~SMART GAMBLING~~ PROBABILITY

- **Sample space** = Ω
- **Sample point** = $\omega \in \Omega$
- **Probability space/model** = $P(\omega)$ exists for all $\omega \in \Omega$
  - $0 \leq P(\omega) \leq 1$
  - $\sum_{\omega} P(\omega) = 1$
- An **event** = subset of Ω
  - For event A
  - $P(A) = \sum_{\{\omega \in A\}} P(\omega)$



RANDOM VARIABLES

- **Random variable** = function
  - **Domain:** sample points
  - **Range:** any (reals, booleans, etc.)
- P induces a **probability distribution** for any r.v. X:
  - $P(X = x_i) = \sum_{\{\omega : X(\omega) = x_i\}} P(\omega)$

$$Odd(\boxed{1}) = true$$

$$Odd(\boxed{2}) = false$$

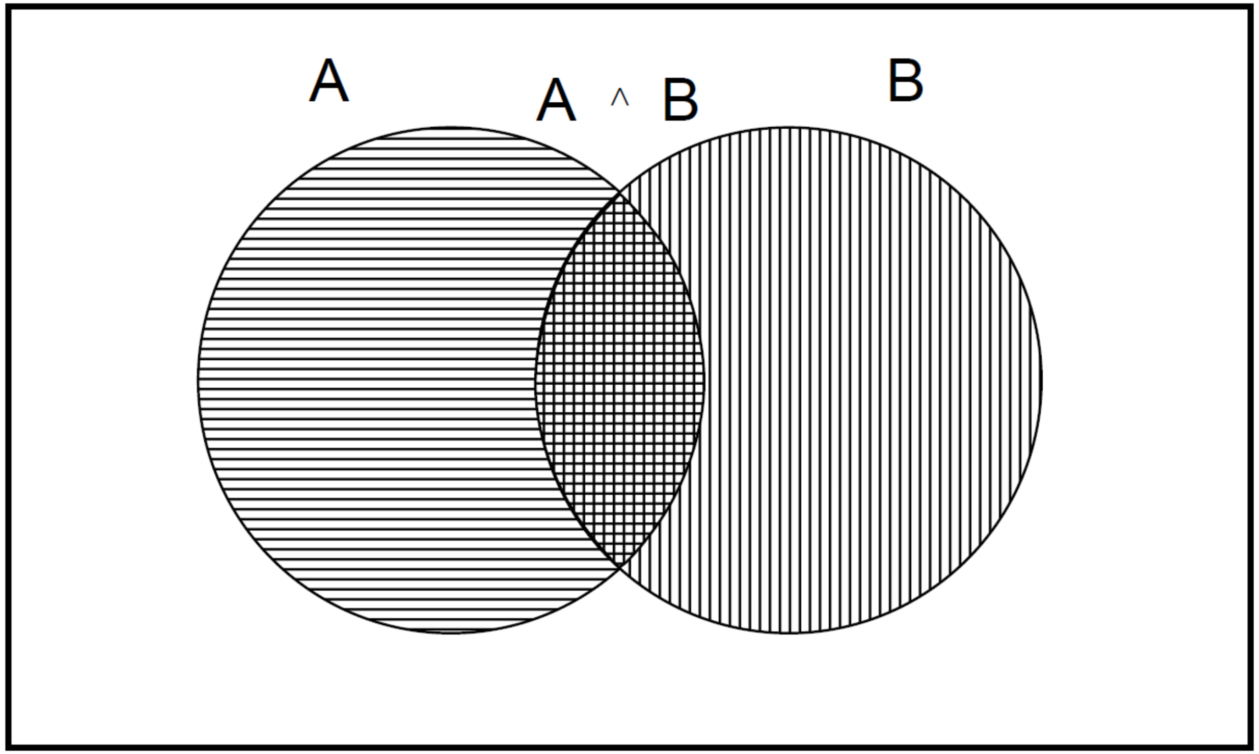
$$P(Odd = true) = P(\boxed{1}) + P(\boxed{3}) + P(\boxed{5}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$

PROPOSITIONS

- Think of a **proposition** as the **event** where the **proposition** is true.
- Given Boolean random variables $GoodSpot$ and $OnTime$:
  - Event $goodspot$ = set of days (sample points) where I parked in a spot I wanted ($GoodSpot(\omega) = true$)
  - Event $\neg goodspot$ = set of points where $GoodSpot(\omega) = false$
  - Event $goodspot \wedge ontime$ = set of all days where I parked in a spot I wanted and was on time for class.
- Sample points **defined** by values of R.V.s
- With Boolean R.V.s, sample point = propositional logic expression
  - Ex. $A = true, B = false$ = $a \wedge \neg b$

A NOTABLE IMPLICATION

- $P(a \vee b) = P(a) + P(b) - P(a \wedge b)$



PROPOSITIONAL CATEGORIZATION
(I JUST LIKE USING BIG WORDS)

Random Variable Type	Range	Example
Propositional/Boolean	$true, false$	$Cavity$ • Range: $true, false$ • Proposition: $Cavity = true$ • Also Written as: $cavity$
Discrete	$< assignment(1), assignment(2), \ldots$ $assignment(n-1), assignment(n) >$	$Weather$ • Range: $< sunny, rain, cloudy, snow >$ • Proposition: $Weather = rain$
Continuous (Bounded)	$[x, y]$	$CreditHours$ • Range: $[0, 18]$ • Proposition: $CreditHours = 13$ • Proposition: $CreditHours \geq 12$
Continuous (Unbounded)	$[-\infty, \infty]$	$Temp$ • Range:$[-\infty, \infty]$ • Proposition: $Temp = 21.6$ • Proposition: $Temp < 22.0$

PRIOR PROBABILITIES PRESENTED

- **Prior** or **unconditional probabilities** of propositions
  - Probabilities **prior** to evidence being **given (|)**.
  - Probabilities without conditions
- Examples:
  - $P(Cavity = true) = 0.1$
  - $P(Weather = sunny) = 0.72$

PROBABILITY DISTRIBUTION

- **Probability distribution,** probability for each assignment of an R.V.
  - Ex. $P(Weather) = < 0.72, 0.1, 0.08, 0.2 >$
  - Ex. $P(Cavity) =< 0.1,0.9 >$
- These probabilities are **normalized** to 1.

NORMALIZATION

- **Normalizing** a set of probabilities = multiplying them to sum to 1
- **Normalization constant** α
  - $\alpha < 0.12, 0.08 > = < 0.6, 0.4 >$
  - $\alpha = \frac{1}{\sum_i probability_i}$
  - $\frac{0.12}{0.12+0.08} = 0.6$
- Normalization is **relative to the set of sample points.**

Your Probabilities

Normalized Probabilities

JOINT PROBABILITY DISTRIBUTION

- **Joint probability distribution** for a set of R.V.s gives every sample point by enumerating through all possible assignments of the R.V.s
- Syntax: $P(Weather, Cavity)$

$Weather =$	$sunny$	$rain$	$cloudy$	$snow$
$Cavity = true$	0.144	0.02	0.016	0.02
$Cavity = false$	0.576	0.08	0.064	0.08

POP QUIZ!

How many values must be known to complete this table?

$Weather =$	$sunny$	$rain$	$cloudy$	$snow$
$Cavity = true$	?	?	?	?
$Cavity = false$	?	?	?	?

PROBABILITY FOR CONTINUOUS VARIABLES

Covered in chapter13.pdf slides 13-14

ADDING CONDITIONS

- We have unconditional probabilities, let's talk **conditional/posterior** probabilities.
  - i.e. The probability of x given some proposition
- Ex. $P(cavity|toothache) = 0.8$
  - **Given (|)** that I have a toothache, update the probability of $cavity$

$$P(cavity) = 0.1 \xrightarrow{\text{New Evidence}} P(cavity|toothache) = 0.8$$

CONDITIONAL DISTRIBUTION

$$P(Cavity|Toothache)$$

	$toothache$	$\neg toothache$
$cavity$	0.08	0.02
$\neg cavity$	0.02	0.88

ON THE TOPIC OF USELESSNESS

$$P(cavity|toothache, cavity) = 1$$

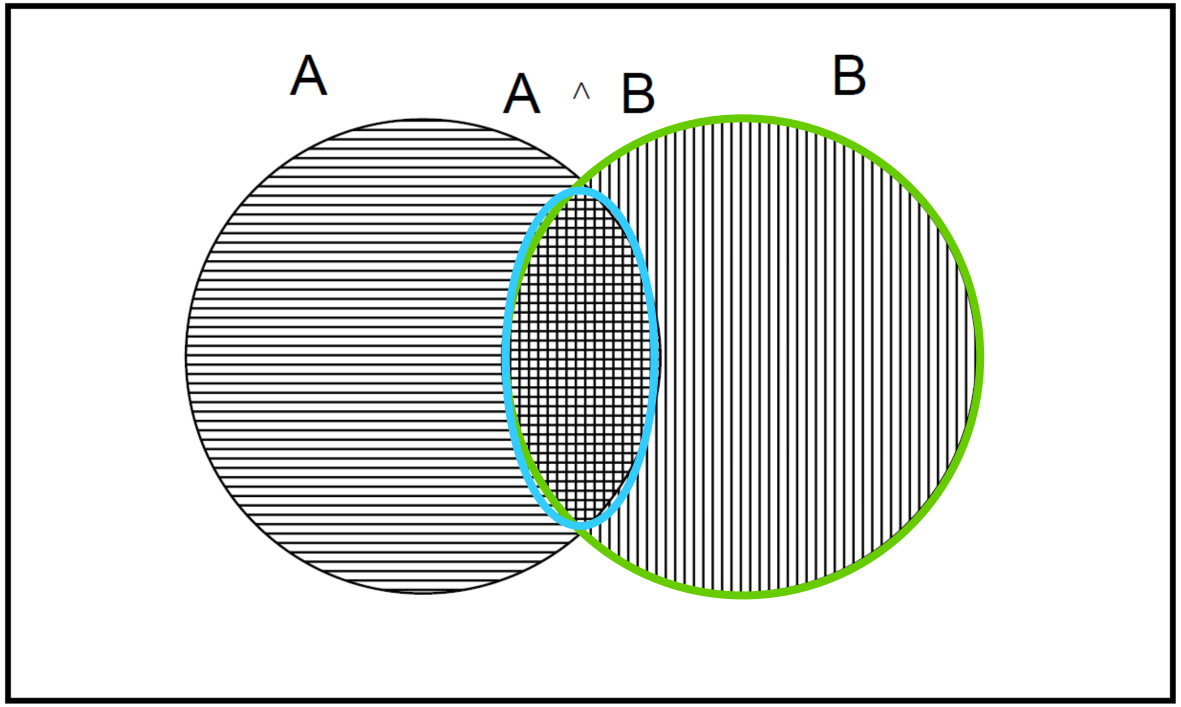
Valid, relevant, but useless

$$P(cavity|toothache, 49ersWin)$$
$$= P(cavity|toothache)$$
$$= 0.8$$

Valid, but irrelevant and useless

UNDER THE HOOD

- Defining conditional probability:
  - $P(a|b) = \dfrac{P(a \wedge b)}{P(b)}$ $if\ P(b) \neq 0$



PRODUCT RULE FOR CONDITIONAL PROBABILITY
(NOT CALCULUS, DON'T CRY YET)

$$P(a|b) = \frac{P(a \wedge b)}{P(b)}, \frac{P(a \wedge b)}{\cancel{P(b)}} * \cancel{P(b)} = P(a \wedge b)$$

Therefore,

$$P(a \wedge b) = P(a|b)P(b) = P(b|a)P(a)$$

CHAIN RULE FOR CONDITIONAL PROBABILITY
(FOR REAL, STILL NOT CALCULUS)

Product Rule holds for whole distributions:

$$P(Weather, Cavity) = P(Weather|Cavity)P(Cavity)$$

Chain rule derived from successive application:

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

$$= \prod_{i=1}^{n} P(X_i|X_1, \ldots, X_{i-1})$$

INFERENCE BY ENUMERATION
(OUR OLD FRIEND)

	$toothache$		$\neg toothache$	
	$catch$	$\neg catch$	$catch$	$\neg catch$
$cavity$	.108	.012	.072	.008
$\neg cavity$	.016	.064	.144	.576

atomic event

atomic event

- For a proposition like $toothache$, $P(toothache)$ is the sum of the probability of every atomic event where it's true.
  - $P(proposition) = \sum_{event:event \models proposition} P(event)$
  - $P(toothache) = 0.108 + 0.012 + 0.016 + 0.064 = 0.2$

INFERENCE BY ENUMERATION
(OUR OLD FRIEND)

	$toothache$		$\neg toothache$	
	$catch$	$\neg catch$	$catch$	$\neg catch$
$cavity$	.108	.012	.072	.008
$\neg cavity$	.016	.064	.144	.576

- $P(toothache \vee cavity) = 0.108 + 0.012 + 0.016 + 0.064 + 0.072 + 0.008 = 0.28$

INFERENCE BY ENUMERATION
(OUR OLD FRIEND)

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

$$P(\neg cavity \mid toothache) = \frac{P(\neg cavity \wedge toothache)}{P(toothache)}$$

$$= \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064} = 0.4$$

NORMALIZATION RETURNS

	$toothache$		$\neg toothache$	
	$catch$	$\neg catch$	$catch$	$\neg catch$
$cavity$	.108	.012	.072	.008
$\neg cavity$	.016	.064	.144	.576

- Denominator like normalization constant α
- $P(Cavity \mid toothache) = \frac{1}{P(toothache)} * P(Cavity \wedge toothache)$

$= \alpha P(Cavity, toothache)$

$= \alpha \sum_{catch} P(Cavity, toothache, Catch = catch)$

$= \alpha[P(Cavity, toothache, catch) + P(Cavity, toothache, \neg catch)]$

$= \alpha[< 0.108, 0.016 > + < 0.012, 0.064 >]$

$= \alpha < 0.12, 0.08 > = < 0.6, 0.4 >$

EXPANDING CONDITIONAL PROBABILITIES

- Ω = all variables
- Q = query variables
- E = evidence variables
- H = hidden variables
  - $H = \Omega - Q - E$

- Start with probability: $P(Q|E = e)$
- Apply product rule:
  - $P(Q|E = e) = \frac{1}{P(E = e)} P(Q \wedge E = e) = \alpha P(Q, E = e)$
- Sum the hidden variables:
  - $\alpha \sum_h P(Q, E = e, H = h)$

THANK YOU

Joshua Sheldon

jsheldon2022@my.fit.edu