CS109: Probability for Computer Scientists

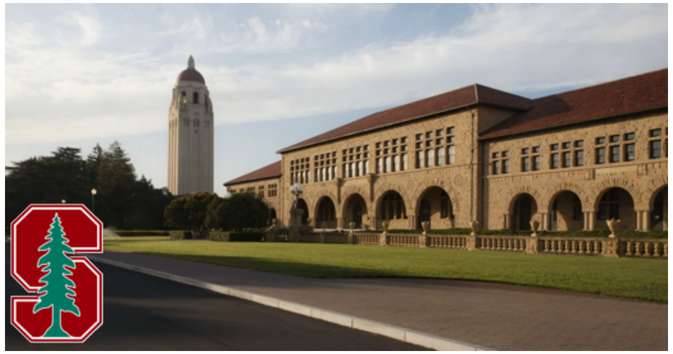
Lisa Yan

September 23, 2019

Lisa Yan

Yes, my undergrad was here…

…But now I'm here!!!



PhD: Tools to understand student learning

My interests over time

Networks, Data Science

Teaching

Education Tools

Why I like probability

- I like data
- I want to help people
- Probability helps me help people with data
- Also Pokemon

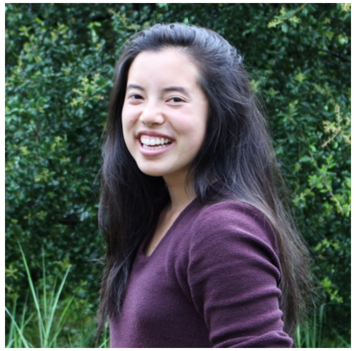
Me, circa 2003

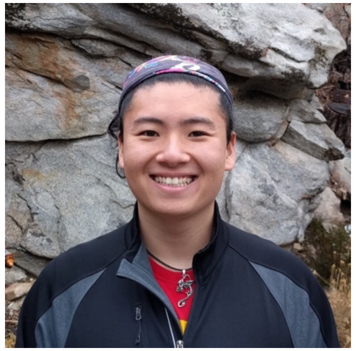
$$a = \frac{(3 \times \mathrm{HP}_{\mathrm{max}} - 2 \times \mathrm{HP}_{\mathrm{current}}) \times \mathrm{rate} \times \mathrm{bonus}_{\mathrm{ball}}}{3 \times \mathrm{HP}_{\mathrm{max}}} \times \mathrm{bonus}_{\mathrm{status}}$$

Teaching team








What about you?

Today's plan

Course Mechanics

Why you should take CS109

Counting!

Course mechanics (light version)

- For more info, read the Administrivia handout
- Course website:

  http://cs109.stanford.edu/

Prerequisites

CS106B/X

Programming
Recursion
Hash tables
Binary trees

MATH 51/CME 100

Multivariate differentiation
Multivariate integration
Basic facility with linear algebra (vectors)

CS103
(co-requisite OK)

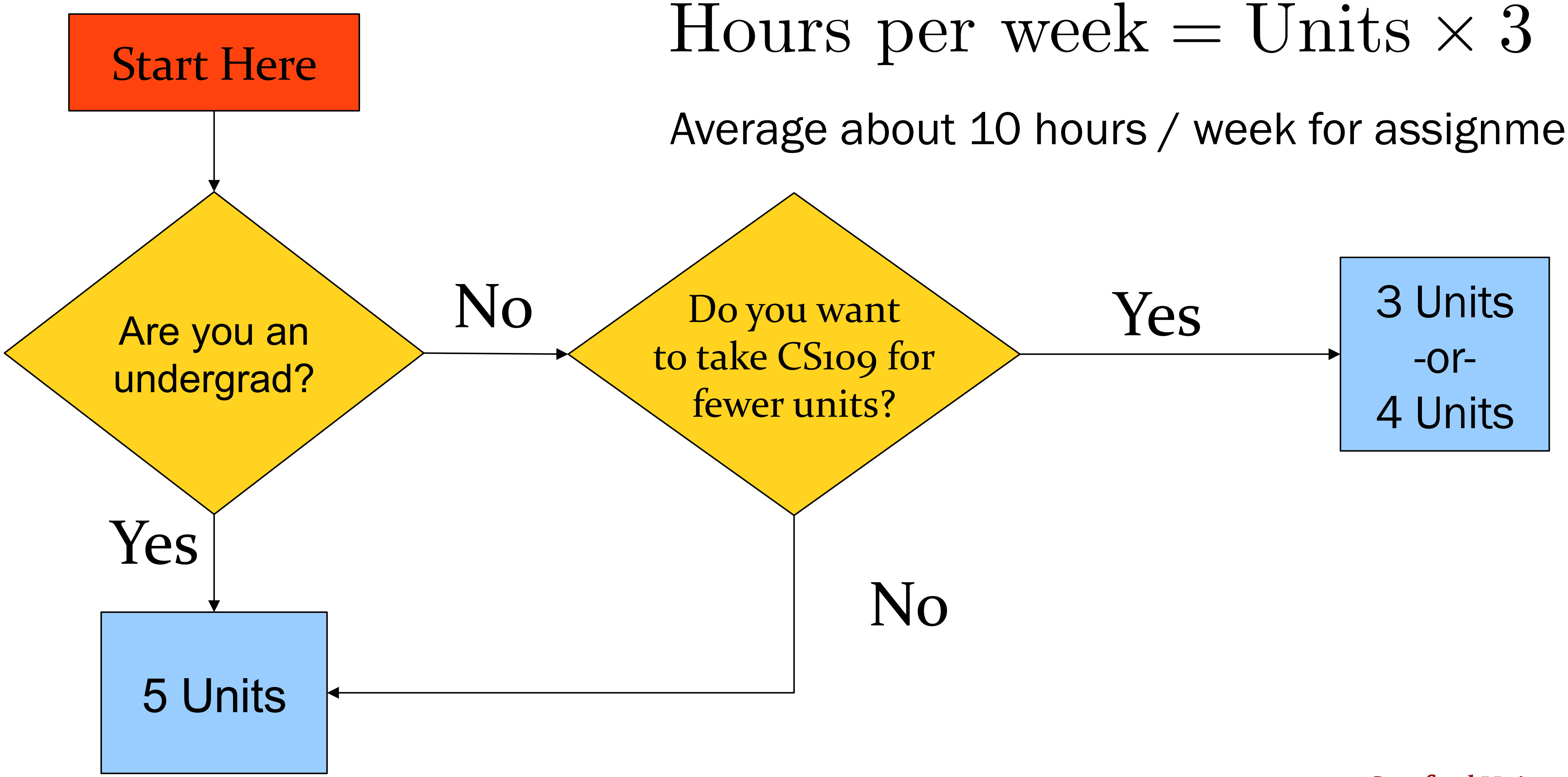
Proofs (induction)
Set theory
Math maturity

Important!

Staff contact

- Piazza
- Email cs109@cs.stanford.edu
- Working office hours
- Contact Lisa for course level issues, extensions, etc.

How many units should I take?



$$\text{Hours per week} = \text{Units} \times 3$$

Average about 10 hours / week for assignments

Where you learn

- Lectures (not videotaped)
- Lecture notes (on website)
- Textbook readings (optional)
- Discussion Section
- Problem Sets

Class breakdown

45%	**6 Problem Sets**	
20%	**Midterm**	Tuesday, October 29th, 7:00–9:00pm
30%	**Final**	Wednesday, December 11th, 3:30–6:30pm
5%	**Participation**	• Weekly concept checks (due Mondays 1pm) • Section participation

Problem Sets

Late Days:

2

(class days)
(for Problem Sets only)

Review session this Friday
(time/location TBA)

Stanford Honor Code

Permitted

- Talk to the course staff
- Talk with classmates
  (cite collaboration)
- Look up general material online

NOT permitted:

- Copy answers:
  - from classmates
  - from former students
  - from previous quarters
- Copy answers from the internet
  - Besides, these are usually incorrect

Questions on logistics?

Today's plan

Course Mechanics

Why you should take CS109

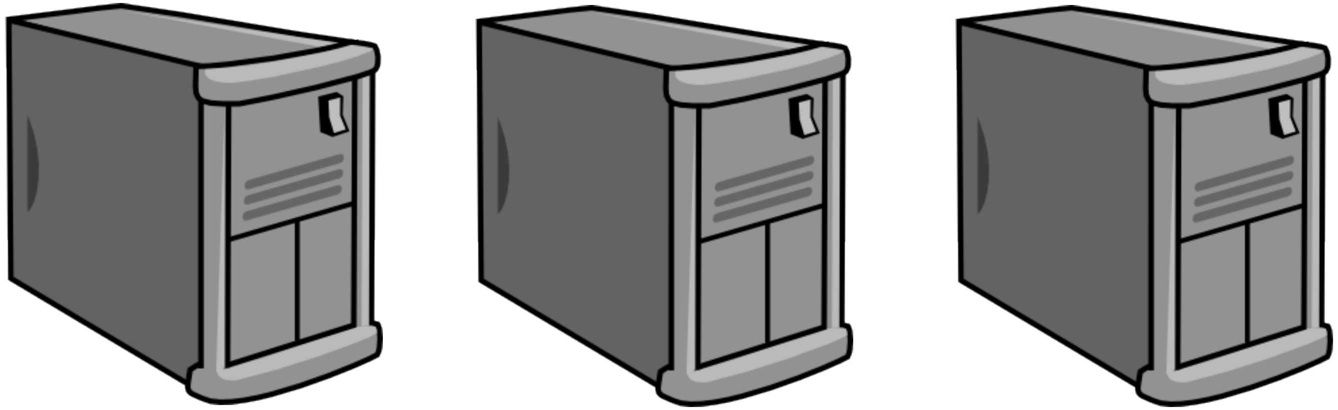
Counting!

Traditional View of Probability

CS view of probability

`http://www.site.com`

But wait… There's MORE!!



Machine Learning

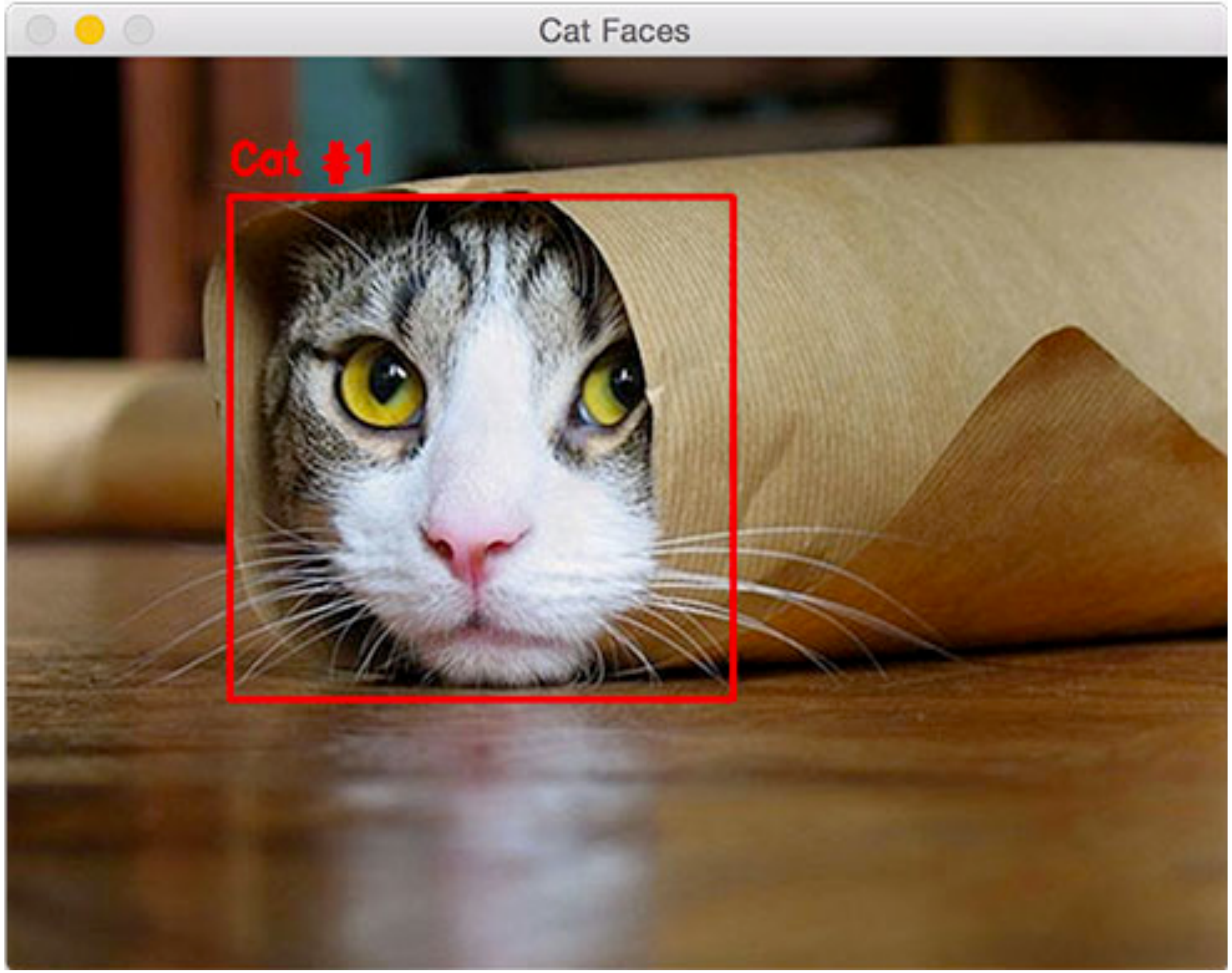
= Machine (compute power)

+ Probability

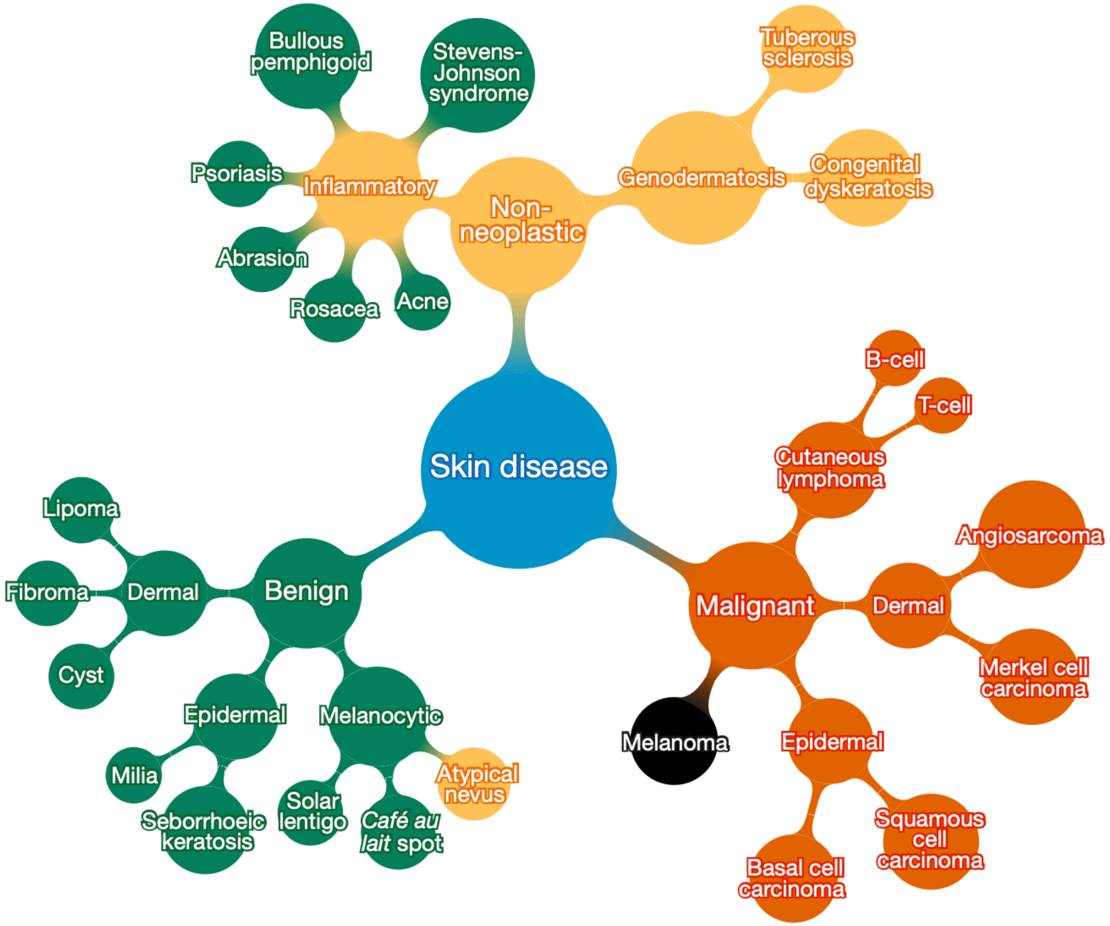
+ Data

Machine Learning Algorithm

Classification



Where is this useful?



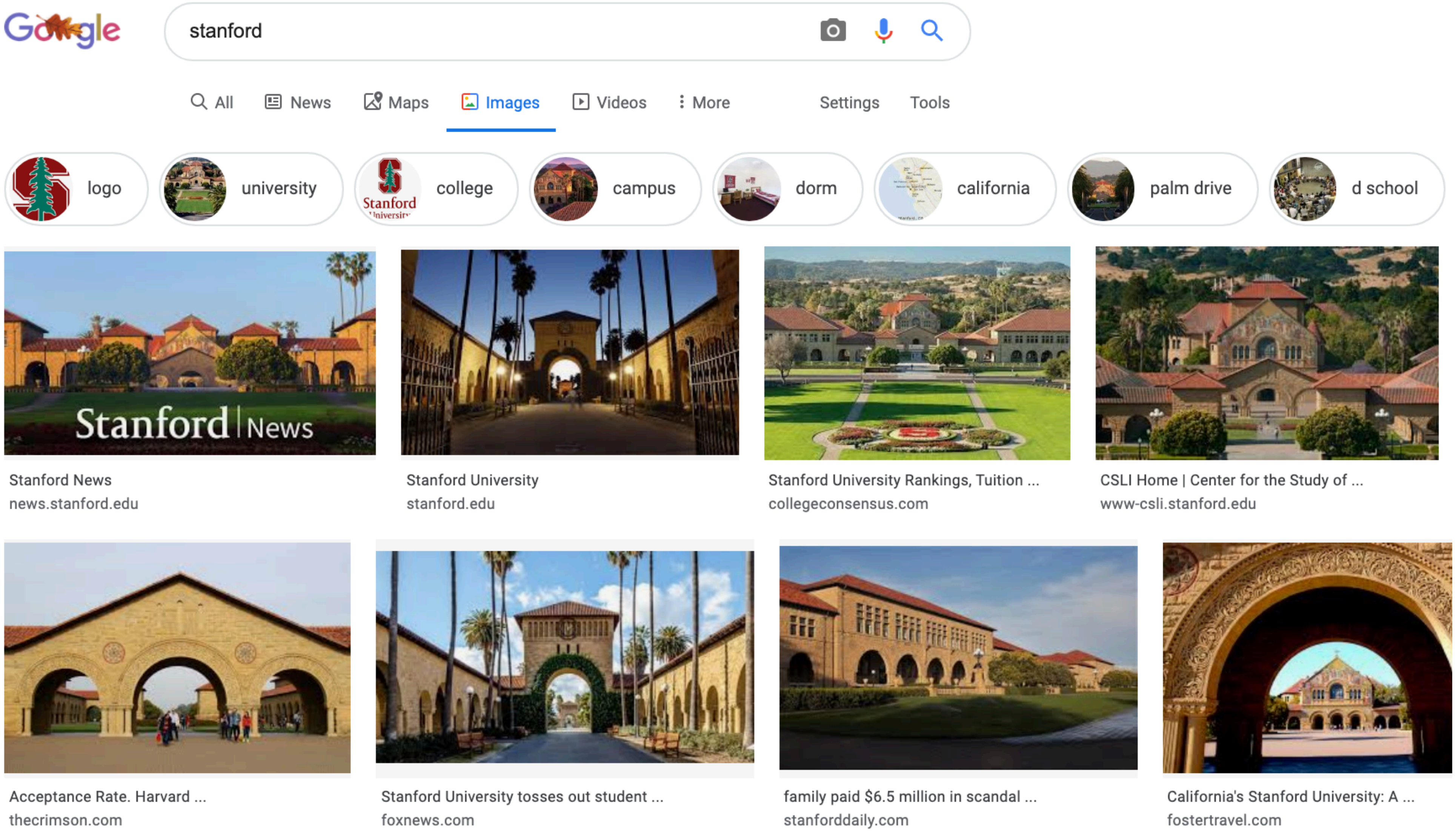
A machine learning algorithm performs **better than** the best dermatologists.

Developed in 2017 at Stanford.

Esteva, Andre, et al. "Dermatologist-level classification of skin cancer with deep neural networks." *Nature* 542.7639 (2017): 115-118.

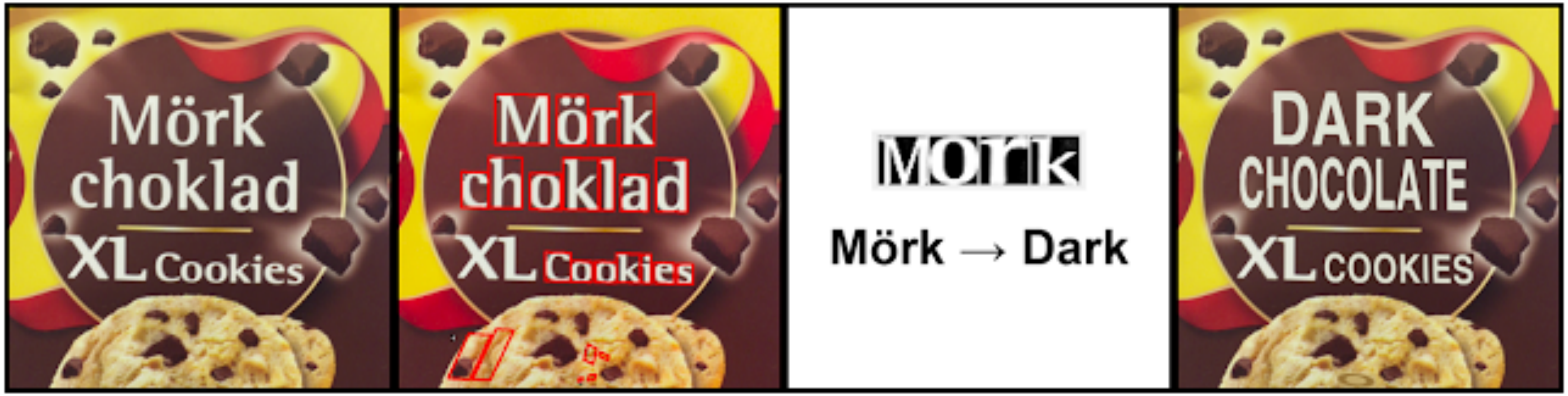
The last remaining board game

Image tagging



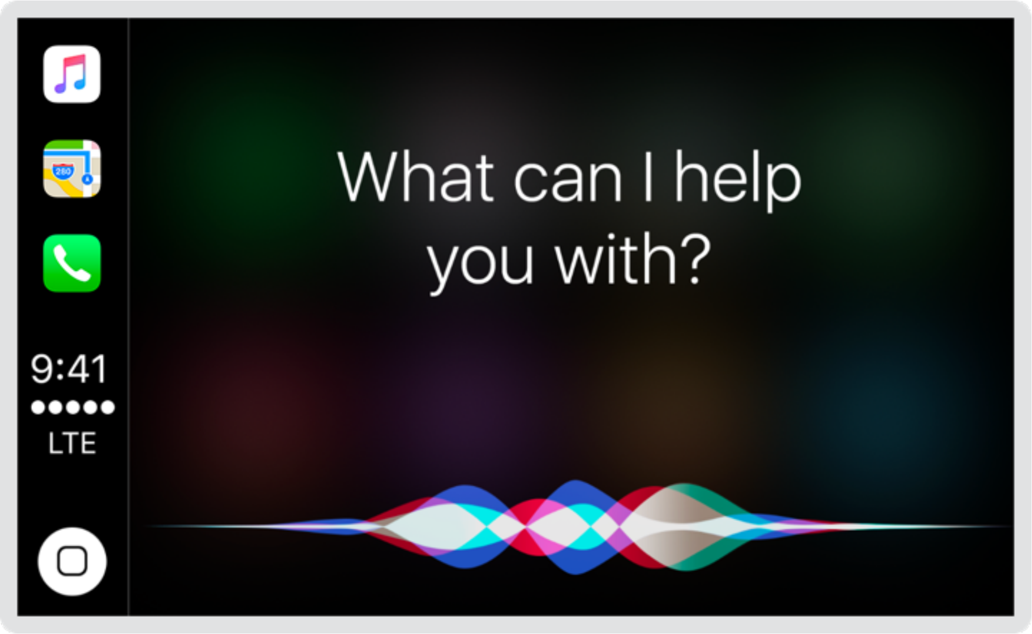
Self-driving cars

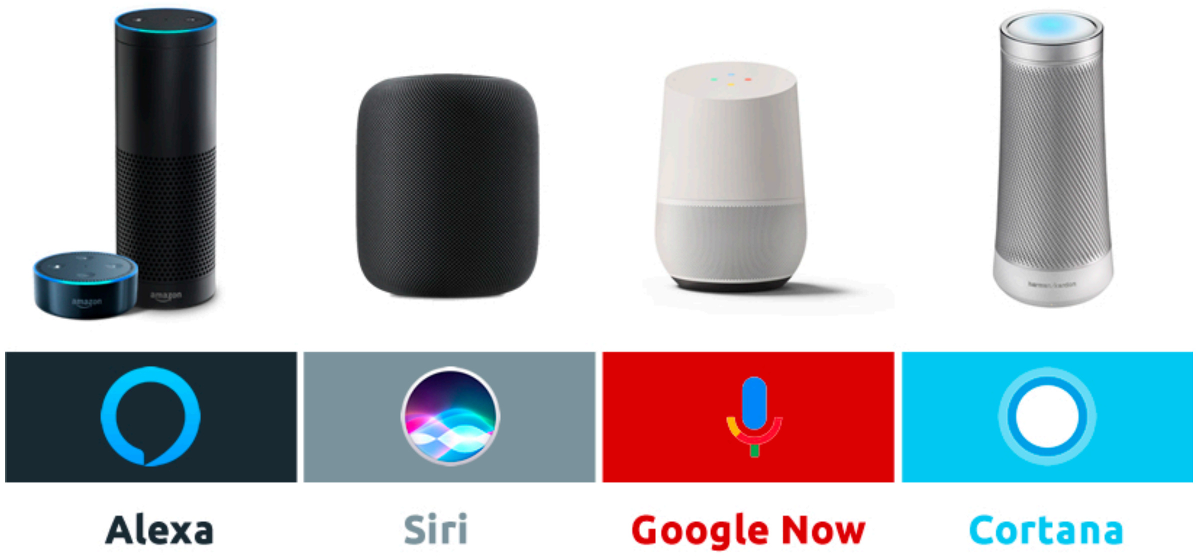
Augmented Reality Machine Translation



Automatic machine translation on Google Translate

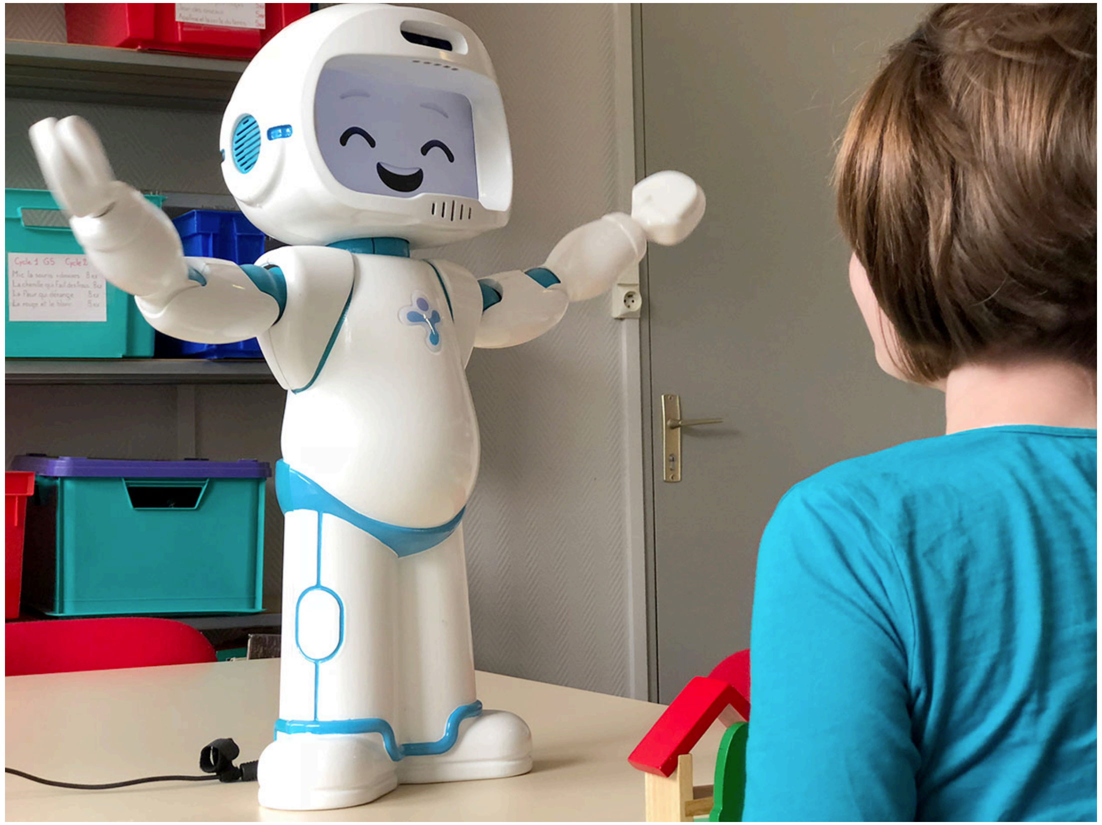
Voice assistants





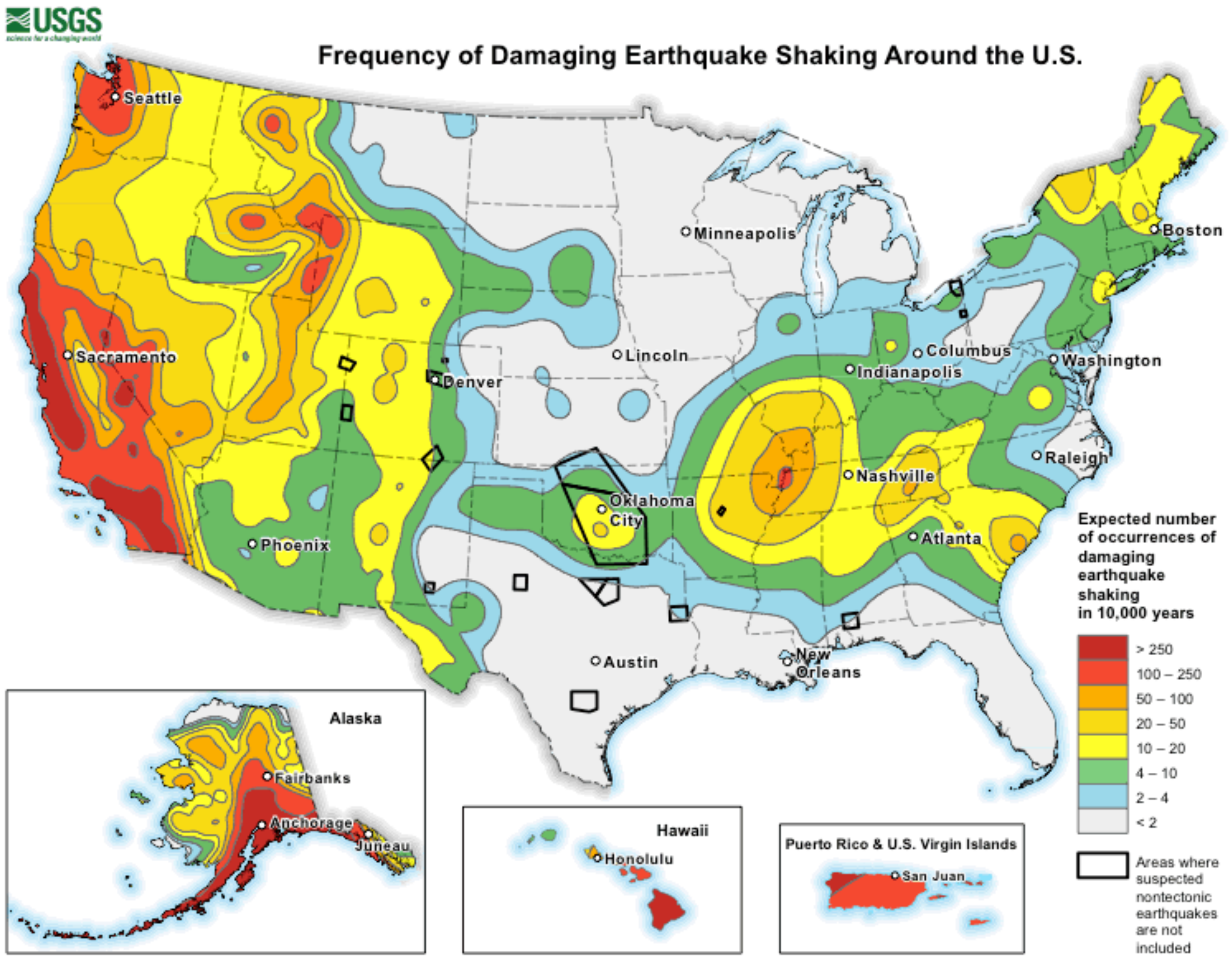
Probability is *more* than just machine learning.

Probability and medicine

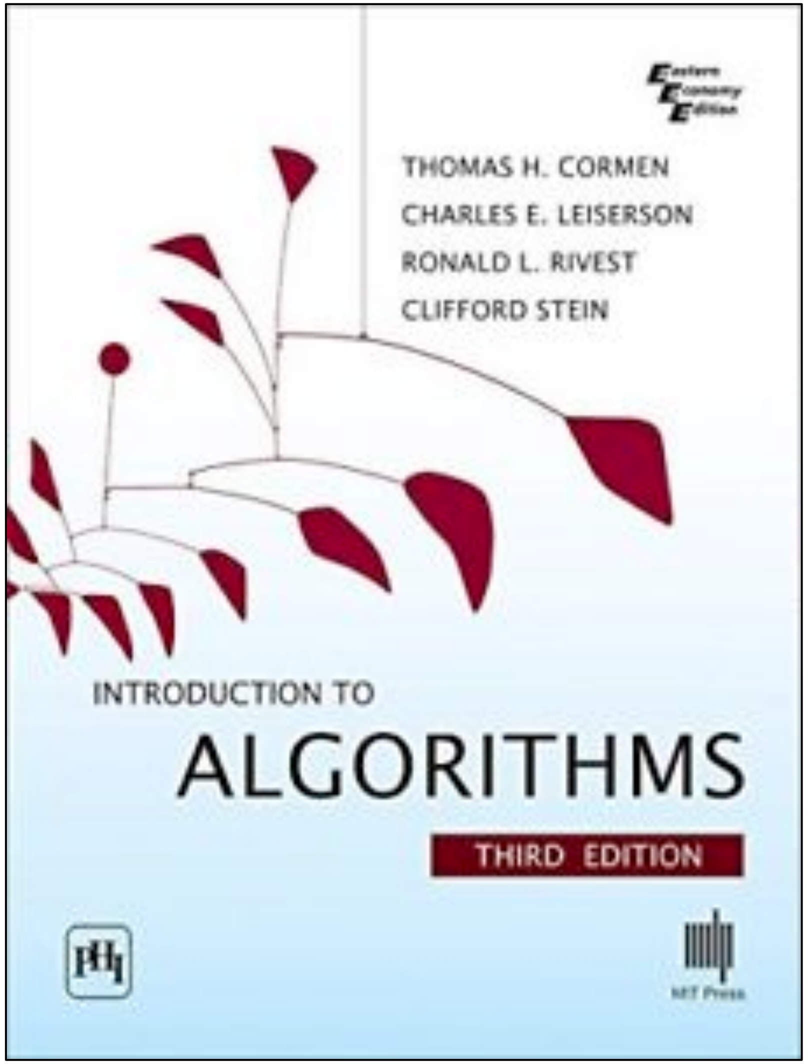


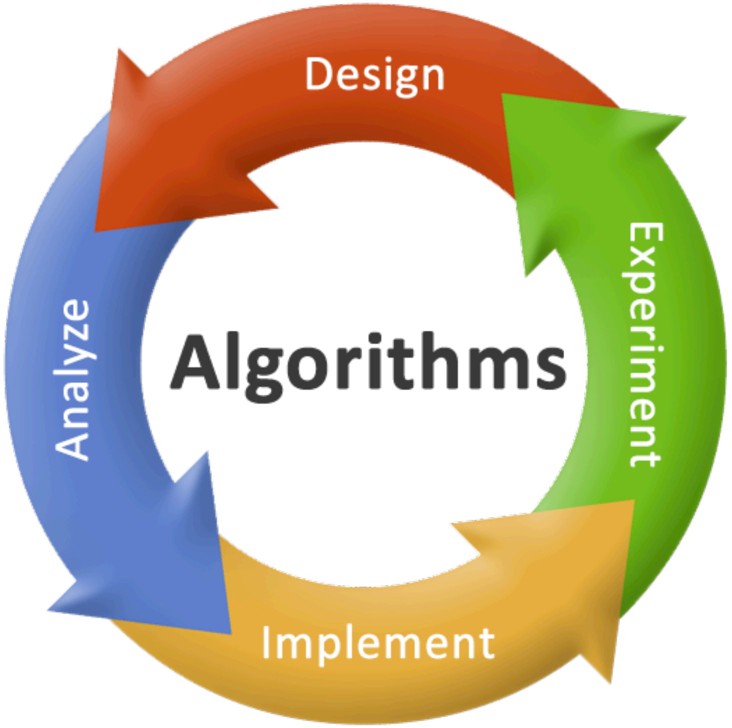
Probability and art

Probability and climate

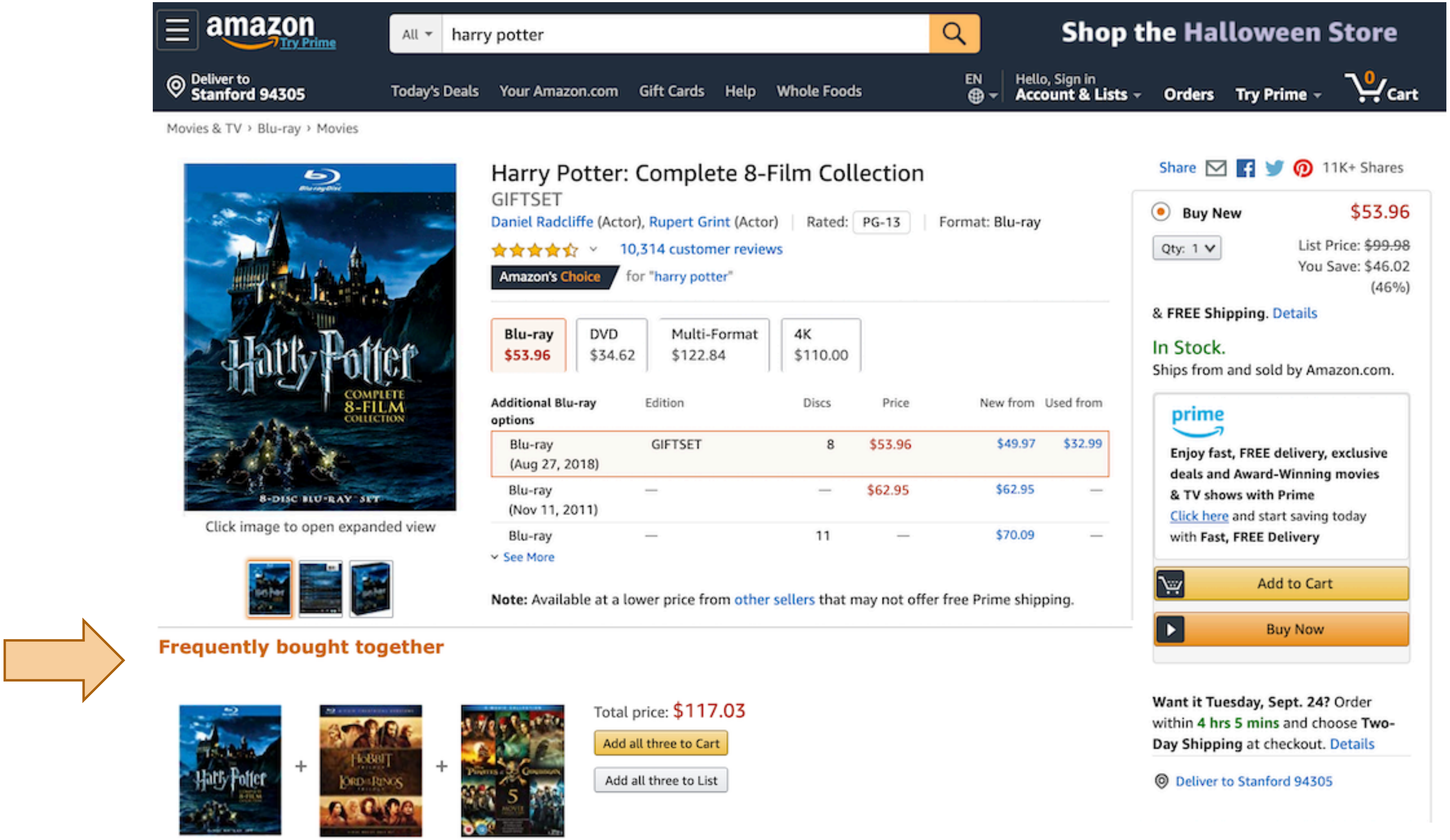


Probabilistic analysis of algorithms

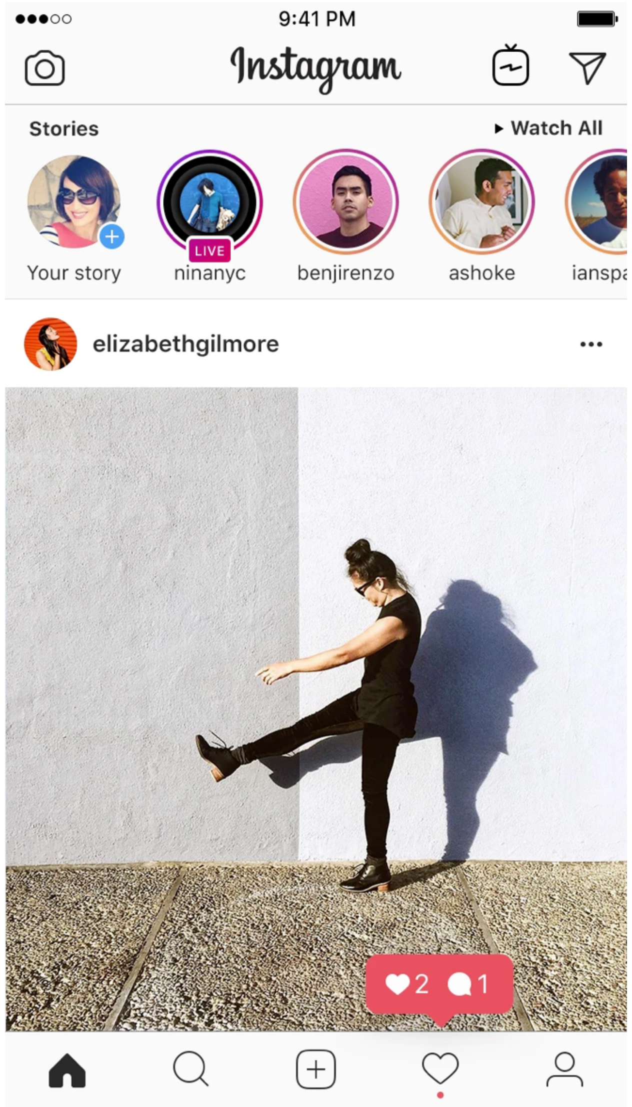




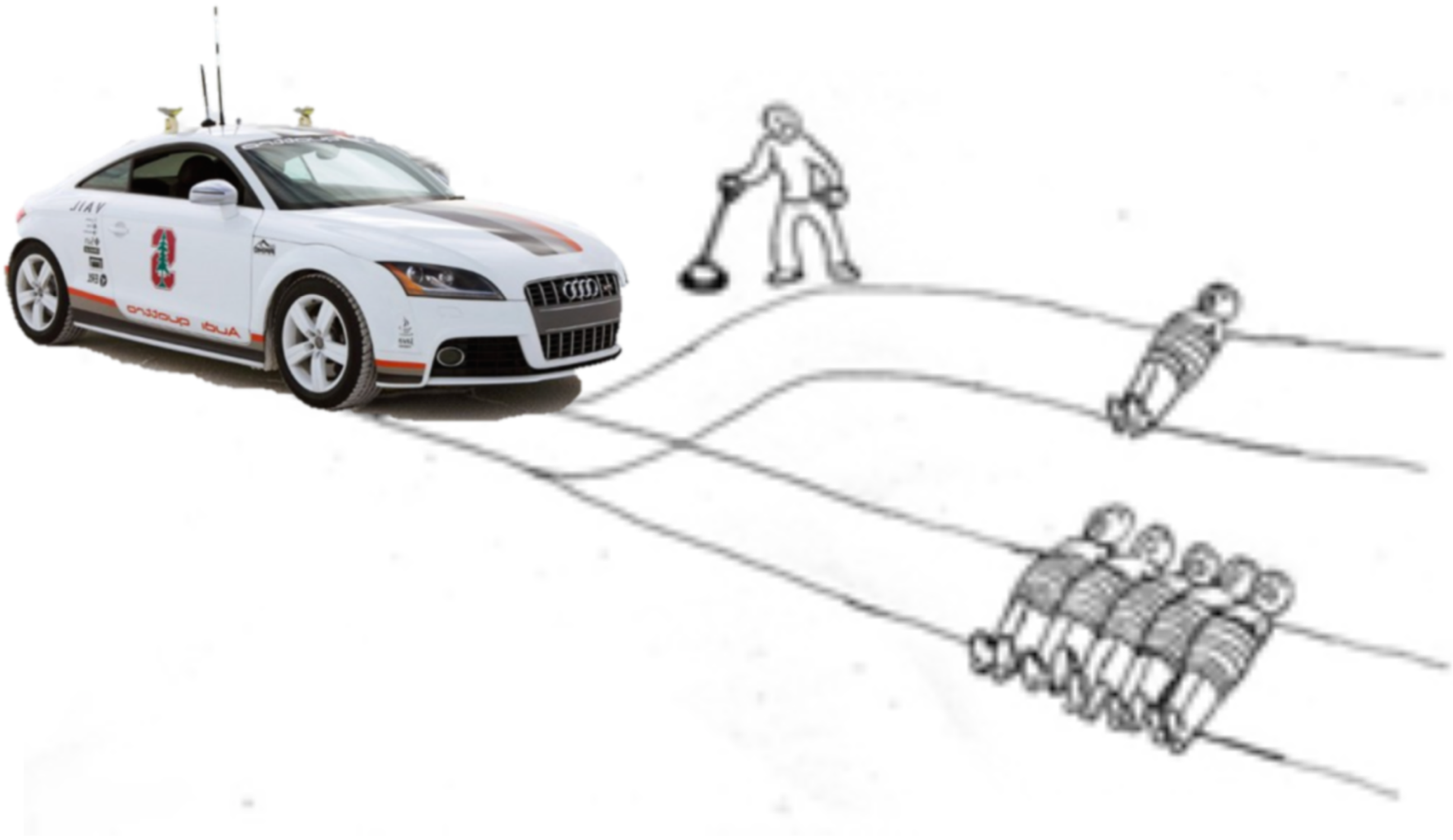
Probability in practice



Probability at your fingertips

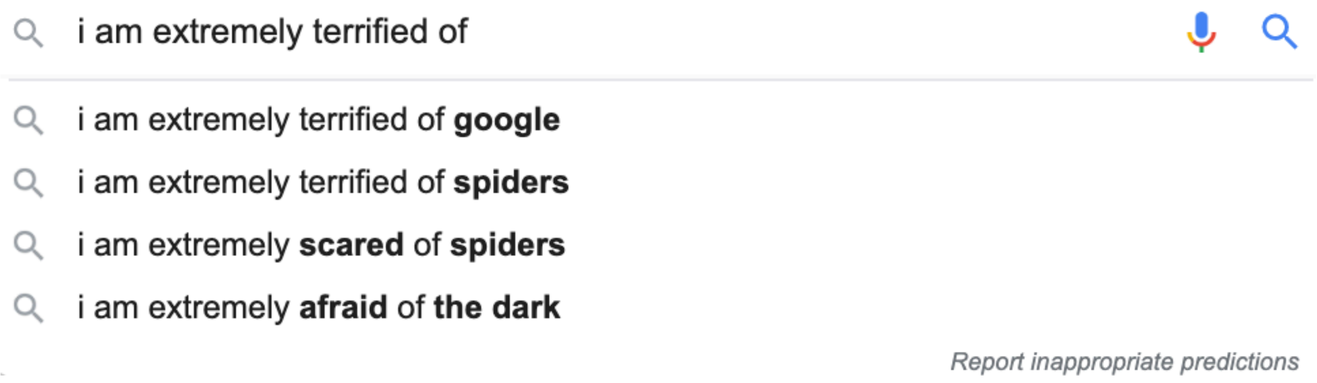


Probability and philosophy



Probability for good

Algorithms of Oppression,
Safiya Umoja Noble. 2018



How do we identify systemic biases in our data and incorporate human judgment into our probabilistic models?

We’ll get there!

Probability is not always intuitive.

Zika test

A patient takes a Zika test that returns positive.

What is the probability that they have the Zika virus?

- 0.8% of people have the virus
- Test has 90% positive rate for people with the virus
- Test has 7% positive rate for people without the virus

Correct answer: 9%

Probability = Important
+ Needs Studying

Today's plan

Course Mechanics

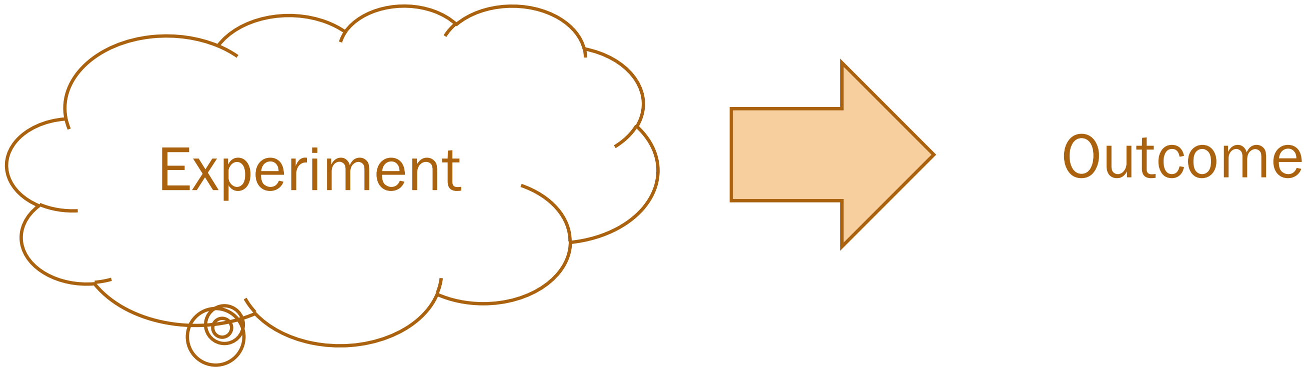
Why you should take CS109

Counting!

01: Counting

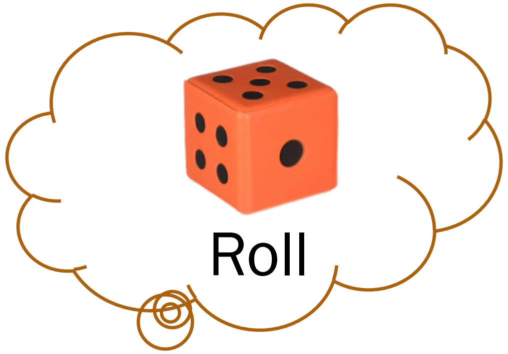
What is Counting?

An experiment in probability:



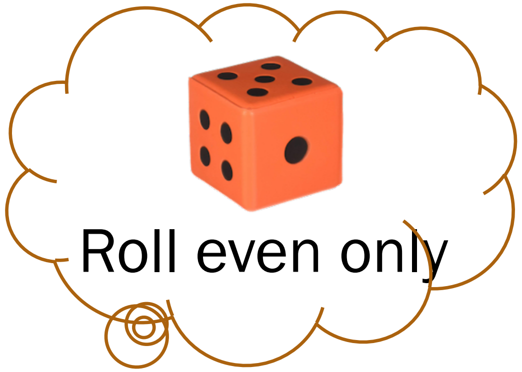
Counting: How many possible outcomes can occur from performing this experiment?

What is Counting?



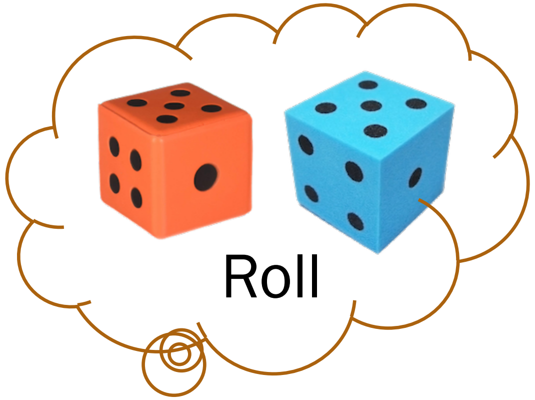
6

{1, 2, 3, 4, 5, 6}



3

{2, 4, 6}



36

{(1, 1) , (1, 2), (1, 3), (1, 4), (1, 5), (1, 6),
(2, 1) , (2, 2), (2, 3), (2, 4), (2, 5), (2, 6),
(3, 1) , (3, 2), (3, 3), (3, 4), (3, 5), (3, 6),
(4, 1) , (4, 2), (4, 3), (4, 4), (4, 5), (4, 6),
(5, 1) , (5, 2), (5, 3), (5, 4), (5, 5), (5, 6),
(6, 1) , (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)}

Sum Rule of Counting

If the outcome of an experiment can be either from

Set A, where $|A| = m$,

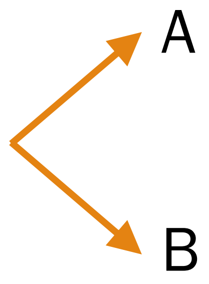
<u>or</u> Set B, where $|B| = n$,

where $A \cap B = \emptyset$,

Then the number of outcomes of the experiment is

$|A| + |B| = m + n.$

One experiment



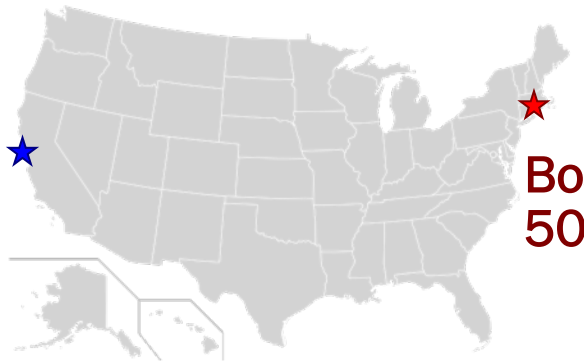
Video streaming application

A $\quad A \cap B = \emptyset$

B $\quad |A| + |B| = m + n$

Your application has distributed servers in 2 locations.

If a server request is sent to the application, how large is the set of servers it can get routed to?



San Jose
100 servers

Boston
50 servers

Goal

Outcome server is in either San Jose or Boston

Define

A : San Jose

B : Boston

Note: $A \cap B = \emptyset$

Solve

$|A| + |B| =$

150 servers

Product Rule of Counting

If an experiment has two parts, where

The first part's outcomes are from Set A, where $|A| = m$, and

The second part's outcomes are from Set B, where $|B| = n$,

Then the number of outcomes of the experiment is

$$|A||B| = mn.$$

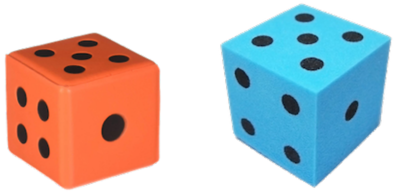
Two-step experiment

→ A → B

Dice

$\longrightarrow$ A $\longrightarrow$ B $\quad |A||B| = mn$

How many possible outcomes are there from rolling two 6-sided dice?



{(1, 1) , (1, 2), (1, 3), (1, 4), (1, 5), (1, 6),
(2, 1) , (2, 2), (2, 3), (2, 4), (2, 5), (2, 6),
(3, 1) , (3, 2), (3, 3), (3, 4), (3, 5), (3, 6),
(4, 1) , (4, 2), (4, 3), (4, 4), (4, 5), (4, 6),
(5, 1) , (5, 2), (5, 3), (5, 4), (5, 5), (5, 6),
(6, 1) , (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)}

Goal

Outcome roll contains an outcome from both die 1 and die 2

Define

A : Die 1 outcomes

B : Die 2 outcomes

Solve

$|A|\times|B| = 36$

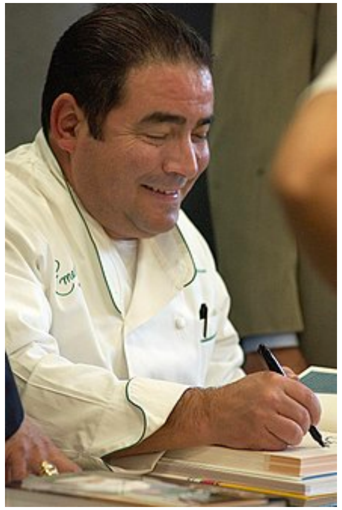
36 outcomes

TOP DEFINITION

kick it up a notch

To make things more intense, exciting, or interesting.

(introduced by **chef Emeril Lagasse** in reference to **spicing** up his recipes)



Inclusion-Exclusion Principle

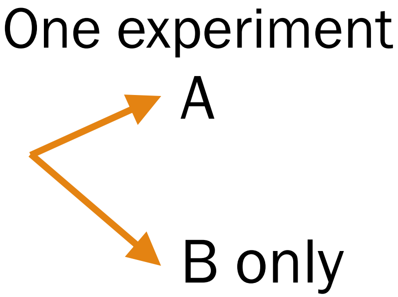
If the outcome of an experiment can be either from

Set A or set B,

where A and B may overlap,

Then the total number of outcomes of the experiment is

$$|A \cup B| = |A| + |B| - |A \cap B|.$$



Sum Rule of Counting:
A special case

Transmitting bytes over a network

A → B only

Inclusion-Exclusion Principle
$|A \cup B| = |A| + |B| - |A \cap B|$

An 8-bit string is sent over a network.

- The receiver only accepts strings that either start with 01 or end with 10.

How many 8-bit strings will the receiver accept?

`01001100`

byte (8 bits)

Define

A : 8-bit strings starting with 01

B : 8-bit strings ending with 10

1. What is $|A|$?
   A. 2^8
   B. 2^6
   C. 2^4
   D. 0

2. What is $|A \cap B|$?
   A. 2^8
   B. 2^6
   C. 2^4
   D. 0

Transmitting bytes over a network

A
B only

Inclusion-Exclusion Principle
$|A \cup B| = |A| + |B| - |A \cap B|$

An 8-bit string is sent over a network.

- The receiver only accepts strings that either start with 01 or end with 10.

How many 8-bit strings will the receiver accept?

01001100

byte (8 bits)

Define
A : 8-bit strings starting with 01
B : 8-bit strings ending with 10

1. What is $|A|$?
   A. 2^8
   B. 2^6
   C. 2^4
   D. 0

2. What is $|A \cap B|$?
   A. 2^8
   B. 2^6
   C. 2^4
   D. 0

Transmitting bytes over a network

A

B only

Inclusion-Exclusion Principle
$|A \cup B| = |A| + |B| - |A \cap B|$

An 8-bit string is sent over a network.

- The receiver only accepts strings that either start with 01 or end with 10.

01001100

byte (8 bits)

How many 8-bit strings will the receiver accept?

Define

A : 8-bit strings starting with 01

B : 8-bit strings ending with 10

1. What is $|A|$?

   B. 2^6

2. What is $|A \cap B|$?

   C. 2^4

Solve

$$|A \cup B| = |A| + |B| - |A \cap B|$$

$$= 2^6 + 2^6 - 2^4 \quad = 112 \text{ outcomes}$$

General Principle of Counting

If an experiment has r steps, such that

Step i has n_i outcomes for all $i = 1, \dots, r$,

Then the number of outcomes of the experiment is

$$n_1 \times n_2 \times \cdots \times n_r = \prod_{i=1}^{r} n_i \,.$$

Multi-step experiment

→ 1 → 2 → …

Product Rule of Counting: A special case

License plates

→ 1 → 2 → …

General Principle of Counting
$n_1 \times n_2 \times \cdots \times n_r = \prod_{i=1}^{r} n_i$

How many CA license plates are possible if…

(pre-1982)

6-part experiment:

A-Z → A-Z → A-Z → digit → digit → digit

$26 \times 26 \times 26 \times 10 \times 10 \times 10$

$= 17{,}576{,}000$

(present day)

2-part experiment:

digit → 6-place license plate experiment

$10 \times 17{,}576{,}000$

$= 175{,}760{,}000$

Floors and ceilings

Floor function

$\lfloor x \rfloor$

The largest integer $\leq x$

Ceiling function

$\lceil x \rceil$

The smallest integer $\geq x$

Check it out:

$\lfloor 1/2 \rfloor = 0$ $\lfloor 2.9 \rfloor = 2$ $\lfloor 8.0 \rfloor = 8$ $\lfloor -1/2 \rfloor = -1$

$\lceil 1/2 \rceil = 1$ $\lceil 2.9 \rceil = 3$ $\lceil 8.0 \rceil = 8$ $\lceil -1/2 \rceil = 0$

Pigeonhole Principle

For positive integers m and n,

if m objects are placed in n buckets,

then at least one bucket must contain at least $\lceil m/n \rceil$ objects.


Pigeons in holes

21st century pigeons

Example:

m objects = 10 pigeons

n buckets = 9 pigeonholes

At least one pigeonhole must contain $\lceil m/n \rceil = 2$ pigeons.

Bounds: an important part of CS109

Balls and urns

≥ 1 bucket must contain at least $\lceil m/n \rceil$ objects

n balls

r urns (buckets)

~~Balls and urns~~ Hash Tables and strings

≥ 1 bucket must contain at least $\lceil m/n \rceil$ objects

Consider a hash table with 100 buckets.

950 strings are hashed and added to the table.

$$n = 100$$
$$m = 950$$

1. Is it guaranteed that at least one bucket contains *at least* 10 entries?
2. Is it guaranteed that at least one bucket contains **at least** 11 entries?
3. Is it possible to have a bucket with ***no entries***?

~~Balls and urns~~ Hash Tables and strings

≥ 1 bucket must contain at least $\lceil m/n \rceil$ objects

Consider a hash table with 100 buckets.

950 strings are hashed and added to the table.

$$n = 100$$
$$m = 950$$

1. Is it guaranteed that at least one bucket contains *at least* 10 entries? — Yes
2. Is it guaranteed that at least one bucket contains **at least** 11 entries? — No
3. Is it possible to have a bucket with ***no entries***? — Sure

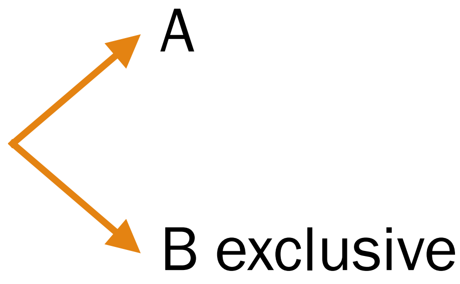
Takeaways from this lecture

Inclusion-Exclusion Principle (generalized Sum Rule)

If the outcome of an experiment can be either from Set A **or** set B, where A and B may overlap, then the total number of outcomes of the experiment is

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

One-step experiment

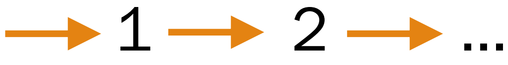


General Principle of Counting (generalized Product Rule)

If an experiment has r **steps**, such that step i has n_i outcomes for all $i = 1, \ldots, r$, then the total number of outcomes of the experiment is

$$n_1 \times n_2 \times \cdots \times n_r = \prod_{i=1}^{r} n_i .$$

Multi-step experiment



Unique 6-digit passcodes

How many unique 6-digit passcodes are possible?

<u>Steps</u>:

1. First digit in passcode — 10 outcomes
2. Second digit in passcode — 10 outcomes

...

6. Sixth digit in passcode — 10 outcomes

$$\begin{aligned}\text{Total} &= n_1 \times n_2 \times \cdots \times n_6 \\ &= 10 \times 10 \times 10 \times 10 \times 10 \times 10 \\ &= 10^6 \text{ passcodes}\end{aligned}$$

Unique 6-digit passcodes with six smudges

How many unique 6-digit passcodes are possible if a phone password uses each of six distinct numbers?

Sort n indistinct objects

Sort n distinct objects

Sort n distinct objects

Ayesha | Tim | Irina | Joey | Waddie

1st | 2nd | 3rd | 4th | 5th

Steps:

1. Choose 1st can — 5 options
2. Choose 2nd can — 4 options

...

5. Choose 5th can — 1 option

$$\text{Total} = 5 \times 4 \times 3 \times 2 \times 1$$
$$= 120$$

Permutations

A permutation is an ordered arrangement of distinct objects.

The number of unique orderings (permutations) of n distinct objects is

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1.$$

Unique 6-digit passcodes with six smudges

How many unique 6-digit passcodes are possible if a phone password uses each of six distinct numbers?

$$\text{Total} = 6!$$
$$= 720 \text{ passcodes}$$

Unique 6-digit passcodes with five smudges

How many unique 6-digit passcodes are possible if a phone password uses each of five distinct numbers?

<u>Steps</u>:

1. Choose digit to repeat — 5 outcomes
2. Create passcode — (permute 4 distinct, 2 indistinct)

$$\text{Total} = 5 \times \frac{6!}{2!}$$

$$= 1{,}800 \text{ passcodes}$$