Here Early?

Here for CSE 312?

Welcome! You're early!

We'll start in ~1 minute

Want a copy of these slides to take notes?

You can download them from the webpage cs.uw.edu/312

Want to be ready for the end of the lecture?

Download the Activity slide from the same place

Go to pollev.com/cse312 and login with your at-uw email (not at-cs!)

You should see something about "presentation hasn't started yet"

Introduction and Counting

CSE 312
Spring 21

Zoom Logistics

We'll always have a TA watching chat – if you have a question, ask it there (either general or direct to the TA – TAs are "co-hosts" of the call).

Don't send direct to me, I won't see it 😬

TA may answer directly, interrupt me, or wait a few minutes and have me answer at a good stopping point.

If you're comfortable (and have the wifi) to turn on your video please do

Nodding/confused looks/glazed over eyes help me know if I said something super confusing.

We will put recordings of (both) lectures on the course webpage.

Staff

Instructor: Robbie Weber

Ph.D. from UW CSE in theory
First year as teaching faculty

Email: rtweber2@cs.washington.edu

TAs

Timothy Akintilo
Joseph Azizeh
Harrison Bay
Anna Goncharenko
Kushal Jhunjhunwalla
Aleks Jovcic
Melissa Lin
Jerome Paliakkara
Alicia Stepin
Alice Wang
Howard Xiao

Logistics

There are two lectures, MWF

A lecture 9:30-10:20, B lecture 1:30-2:20

Both will be recorded, with recordings posted when possible (usually up by the night of the lectures).

Try to attend your officially registered lecture, but you don't need permission to occasionally attend the other.

The lectures should be nearly identical, but there will be small differences (e.g. more/fewer/different questions)

Logistics

Sections meet on Thursdays (starting this week)

Zoom links will be on the pinned post on Ed by Thursday morning.

Please go to your assigned section.

If you can't make your assigned section one week, you can ask the TA(s) in charge of another for permission to join.

Sections will **not** be recorded – we want you to be able to ask questions and give feedback without worrying about being recorded.

We will have separate "walkthrough" videos for those who can't make sections.

Syllabus

When in doubt, it's on the webpage:

it will be

https://courses.cs.washington.edu/courses/cse312/21sp/

Work

Concept Checks (10%)

Short "quiz" for each lecture on gradescope; identify misconceptions right away.

All three for a week due Monday morning; recommend you do them day of lecture.

80% → 100%

→ individual work

Approximately 9 Homeworks (70%)

Mostly written problems, but a few programming questions.

3 "probability in the real world" mini-projects (10%)

A chance to think about applying concepts to real life/some ethical implications of the tools from this class

Final during finals week (10%)

"take home" exam, some collaboration will be allowed.
No proctorio or zoom meetings or anything like that.

Exact dates TBD, you'll have at least 24 hours to spread out the work.

Communication

Ed Discussion board will be our primary means of communication.

Please check frequently.

We'll send announcement emails via Ed.

If you want to contact us:

- Private post on Ed (seen by staff, all TAs)
- Email Robbie
- Anonymous Feedback form on webpage

Collaboration Policy

PLEASE collaborate! Please talk to each other and work with each other.

(subject to the policy – details on webpage)

We're remote – it's going to be harder to find people to work with.

- Ed posts to help find people
- Stay after section Thursday
- Let us know how we can help.

What is This Class?

We're going to learn fundamentals of probability theory.

A **beautiful** and *useful* branch of mathematics.

Applications in:
Machine Learning
Natural Language Processing
Cryptography
Error-Correcting Codes
Data Structures
Data Compression
Complexity Theory
Algorithm Design
…

Content

Combinatorics (*fancy* counting)

Permutations, combinations, inclusion-exclusion, pigeonhole principle

Formal definitions for Probability

Probability space, events, conditional probability, independence, expectation, variance

Common patterns in probability

Equations and inequalities, "zoo" of common random variables, tail bounds

Continuous Probability

pdf, cdf, sample distributions, central limit theorem, estimating probabilities

Applications

Across CS, but with some focus on ML.

Themes

Precise mathematical communication

Both reading and writing dense statements.

Probability in the “real world”

A mix of CS applications

And some actual “real life” ones.

Refine your intuition

Most people have some base level feeling of what the chances of some event are.

We’re going to train you to have better gut feelings.

Counting

Why Counting?

Sometimes useful for algorithm analysis.

- The easiest code to write for "find X" is "try checking every spot where X could be"
- "Given an array, find a set of elements that sum to 0"
- "Given an array, find a set of 2 elements that sum to 0"
- Gut check of "we can 'brute force' this or we can't" is super useful.

A building block toward probability theory

"What are the chances" is usually calculated by

$$= \frac{\text{how many ways can I succeed}}{\text{how many ways can I succeed} + \text{how many ways can I fail}}$$

Remember sets?

A set is an **unordered** list of elements, ignoring repeats.

$\{1,2,3\}$ is a set. It's the same set as $\{2,1,3\}$.

$\{1,1,2,3\}$ is a very confusing way of writing the set $\{1,2,3\}$.

The **cardinality** of a set is the number of elements in it.

$\{1,2,3\}$ has cardinality 3

$|\{1,2,3\}| = 3.$

Counting Rules

How many options do I have for dinner?

I could go to Chili’s where there are 3 meals I choose from, or I could go to Thaiger Room where there are 5 meals I choose from (and none of them are the same between the two).

How many total choices?

$3 + 5 = 8$

Sum Rule: If you are choosing one thing between n options in one group and m in another group with no overlap, the total number of options is: $n + m$.

Counting Rules

I'm still hungry…

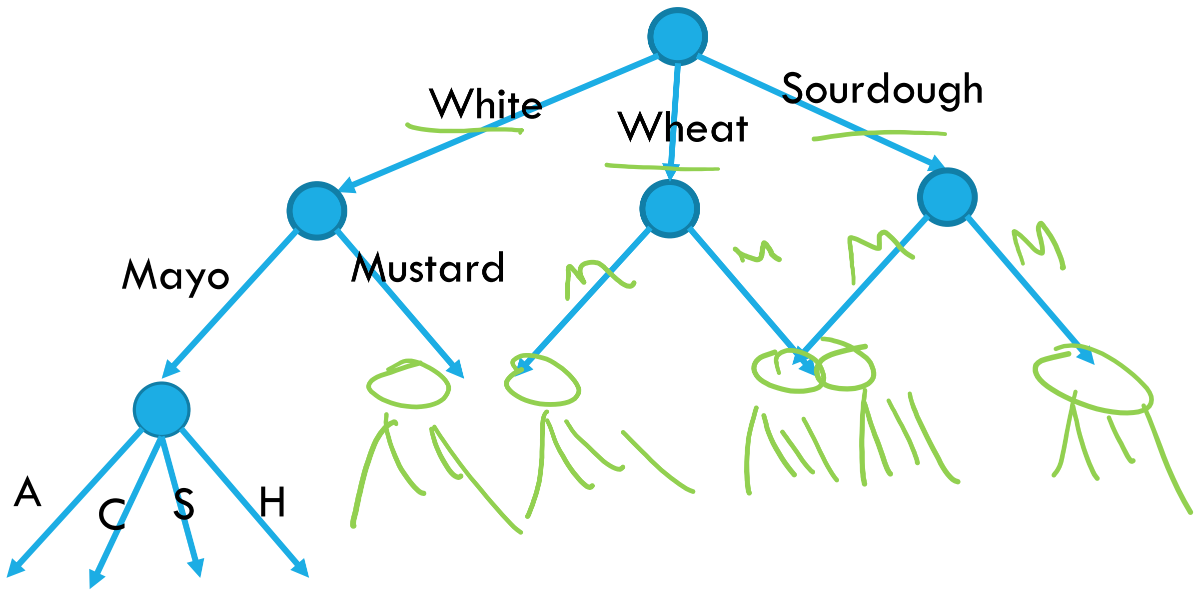
I decide to make a sandwich. My sandwiches are always:

- One of three types of bread (white, wheat, or sourdough).
- One of two spreads (mayo or mustard)
- One of four cheeses (American, cheddar, swiss, or Havarti)

How many sandwiches can I make?

Sandwiches

"Sequential process"



Step 1: choose one of the three breads.

Step 2: regardless of step 1, choose one of the two spreads.

Step 3: regardless of steps 1 and 2, choose one of the four ~~breads~~ cheese.

$$3 \cdot 2 \cdot 4 = 24.$$

Counting Rules

Sum Rule: If you are choosing one thing between n options in one group and m in another group with no overlap, the total number of options is: $n + m$.

Product Rule: If you have a sequential process, where step 1 has $\boldsymbol{n_1}$ options, step 2 has $\boldsymbol{n_2}$ options,…,step $\boldsymbol{k}$ has $\boldsymbol{n_k}$ options, and you choose one from each step, the total number of possibilities is $\boldsymbol{n_1 \cdot n_2 \cdots n_k}$

Applications of the product rule

Remember Cartesian products?

$S \times T = \{(x, y) : x \in S, y \in T\}$

$\{1,2\} \times \{a, b, c\} = \{(1, a), (1, b), (1, c), (2, a), (2, b), (2, c)\}$

How big is $S \times T$? (i.e. what is $|S \times T|$?)

Step 1: choose the element from S.

Step 2: choose the element from T.

Total options: $|S| \cdot |T|$

Power Sets

$\mathcal{P}(S) = \{X : S \subseteq X\}$

$\mathcal{P}(\{1,2,3\}) = \{\, \emptyset,\ \{1\}, \{2\}, \{3\},\ \{1,2\}, \{1,3\}, \{2,3\},\ \{1,2,3\}\,\}$

$S = \{e_1, \ldots, e_{|S|}\}$

How many subsets are there of S, i.e. what is $|\mathcal{P}(S)|$? $2^{|S|}$

Step 1: include e_1 or not 2

Step 2: " e_2 " " 2

⋮

Step $|S|$ " $e_{|S|}$ or not 2

$2^{|S|}$

Power Sets

$\mathcal{P}(S) = \{X: S \subseteq X\}$

$\mathcal{P}(\{1,2,3\}) = \{\ \emptyset,\ \{1\}, \{2\}, \{3\},\ \{1,2\}, \{1,3\}, \{2,3\},\ \{1,2,3\}\ \}$

How many subsets are there of S, i.e. what is $|\mathcal{P}(S)|$?

If $S = \{e_1, e_2, \dots, e_{|S|}\}$

Step 1: is e_1 in the subset?

Step 2: is e_2 in the subset?

…

Step $|S|$: is $e_{|S|}$ in the subset?

$2 \cdot 2 \cdots 2$, |S| times, i.e., $2^{|S|}$.

Baseball Outfits

The Husky baseball team has three hats (purple, black, gray)

Three jerseys (pinstripe, purple, gold)

And three pairs of pants (gray, white, black)

How many outfits are there (consisting of one hat, jersey, and pair of pants) if

the pinstripe jersey cannot be worn with gray pants,
the purple jersey cannot be worn with white pants,
and the gold jersey cannot be worn with black pants.

Baseball Outfits

Step 1: 3 choices for hats.

Step 2: 3 choices for jerseys

Step 3:…

Baseball Outfits

Step 1: 3 choices for hats.

Step 2: 3 choices for jerseys.

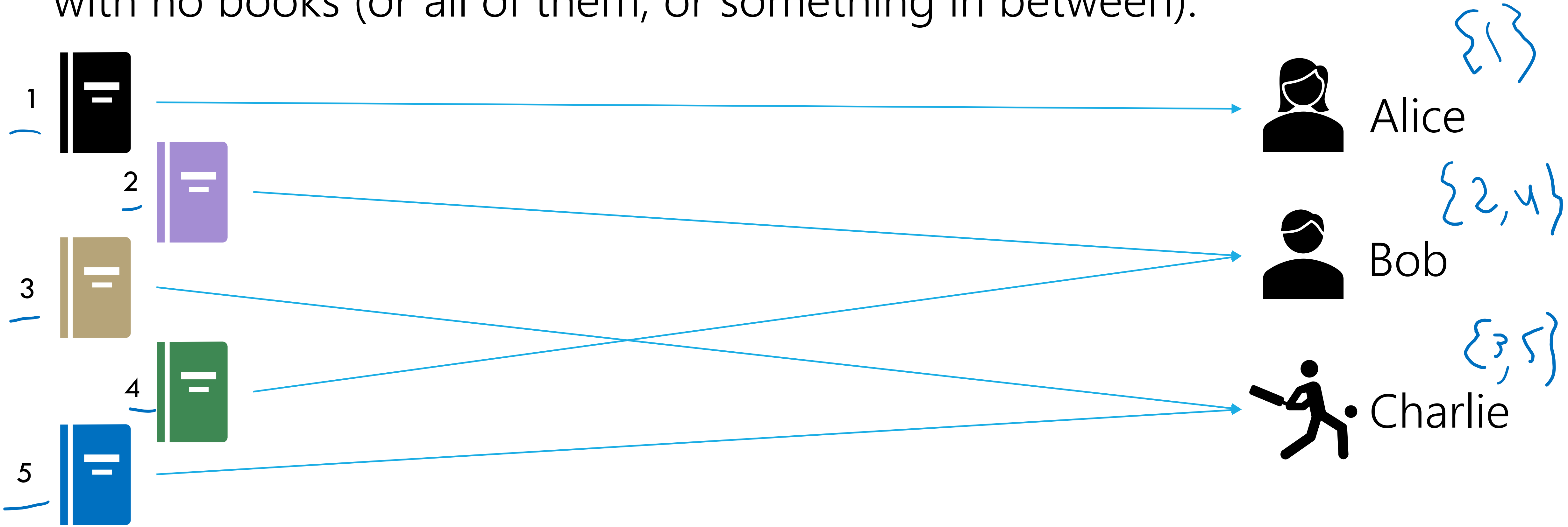
Step 3: Regardless of which jersey we choose, we have 2 options for pants (even though there are three options overall).

$$3 \cdot 3 \cdot 2 = 18.$$

Assigning Books

We have 5 books to split to 3 people (Alice, Bob, and Charlie)

Every book goes to exactly one person, but each person could end up with no books (or all of them, or something in between).



Assigning Books

We have 5 books to split to 3 people (Alice, Bob, and Charlie)

Every book goes to exactly one person, but each person could end up with no books (or all of them, or something in between).

Attempt 1: We're choosing subsets!

Alice could get any of the $2^5 = 32$ subsets of the books.

Bob could get any of the $2^5 = 32$ subsets of the books.

Charlie could get any of the $2^5 = 32$ subsets of the books.

Total is product of those three steps $32 \cdot 32 \cdot 32 = 32768$

Activity

Attempt 1: We're choosing subsets!

Alice could get any of the $2^5 = 32$ subsets of the books.

Bob could get any of the $2^5 = 32$ subsets of the books.

Charlie could get any of the $2^5 = 32$ subsets of the books.

Total is product of those three steps $32 \cdot 32 \cdot 32 = 32768$

Introduce yourselves!

If you can turn your video on, please do.

If you can't, please unmute and say hi.

If you can't do either, say "hi" in chat.

Choose someone to share screen, showing this pdf.

Fill out the poll everywhere so Robbie can adjust his explanation

Go to pollev.com/cse312 and login with your UW identity

Activity

Attempt 1: We're choosing subsets!

Alice could get any of the $2^5 = 32$ subsets of the books.

Bob could get any of the $2^5 = 32$ subsets of the books.

Charlie could get any of the $2^5 = 32$ subsets of the books.

Total is product of those three steps $32 \cdot 32 \cdot 32 = 32768$

We overcounted!

If Alice gets $\{1,2\}$, Bob can't get any subset, he can only get a subset of $\{3,4,5\}$. And Charlie's subset is just whatever is leftover after Alice and Bob get theirs…

Fixing All The Books

You **could**

List out all the options for Alice.

For each of those (separately), list all the possible options for Bob and Charlie.

Use the Summation rule to combine.

~OR~ you could come at the problem from a different angle.

Fixing All the Books

Instead of figuring out which books Alice gets, choose book by book which person they go to.

Step 1: Book 1 has 3 options (Alice, Bob, or Charlie).

Step 2: Book 2 has 3 options (A, B, or C)

…

Step 5: Book 5 has 3 options.

Total: 3^5.

More Practice

Strings

How many strings of length 5 are there over the alphabet $\{A, B, C, \ldots, Z\}$? (repeated characters allowed)

How many binary strings of length n are there?

Strings

How many strings of length 5 are there over the alphabet $\{A, B, C, \ldots, Z\}$? (repeated characters allowed)

26^5

How many binary strings of length n are there?

2^n