Here Early?

Here for CSE 312?

Welcome! You're early!

Want a copy of these slides to take notes?

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

Introduction and Counting

CSE 312 26Wi
Lecture 1

Staff

Instructor: Robbie Weber

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

Email: rtweber2@cs.washington.edu
Office: CSE2 311

TAs

Aadi Anand
Anika Arugunta
Advait Bhargav
Owen Boseley
Isaiah Greenlee
Caleb Hsu
Emma Huang
Heon Jwa
Krisha Khandelwal
Vlad Murad
Lucas Ng
Cindy Ni
Vinay Pritamani
Sam Sullivan
Cleah Winston
Jai Wongcharoensangsiri
Alice Zhu

Logistics

Sections meet on Thursdays (starting this week)

Please go to your assigned section.

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

Section participation is **required**.

- The course is designed to give you high-level ideas here and practice in section.
- Get participation credit by doing problems; either in section or emailed to your TA.
- Details on Syllabus and Ed (including what to do if you can't attend section).

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

Handouts and solutions will be posted.

Concept Checks

Each Lecture will have a "concept check" associated with it.

Idea: Make sure you've understood today's topic before we build on it 2 days later.

Available on gradescope at 2:20 (after B lecture). Due the day of the next lecture at 9:30 (start of A lecture).

- You can submit as many times as you want.
- When you get an answer correct the explanation will appear.

Worth a small amount of your grade, and an 80% average for the quarter is all you need for full credit.

- Details in the syllabus.
- You can (and should!) submit until you get full credit!
- CC1 out this afternoon; due on Friday (in case of late adds/logistics issues)

Exam Dates

We're going to have 4 quizzes in sections(20 minutes each)

Thursdays: Jan 22, Feb 5, Feb 26, Mar 5

One evening midterm: Thursday February 19th (6 PM)

"Second chance quizzes"; optional chances to retake two of your quizzes to improve scores

Quiz second-chance: Evening Thursday Mar 12

Final Exam: **Combined** Mon March 16, 12:30

Syllabus

When in doubt, it's on the webpage:

(or it will be soon ☺)

https://courses.cs.washington.edu/courses/cse312/26wi/

Video coming later this week covering more syllabus information.

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 Rule #1 (problem)

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 Little Thai 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 Rule #2 (problem)

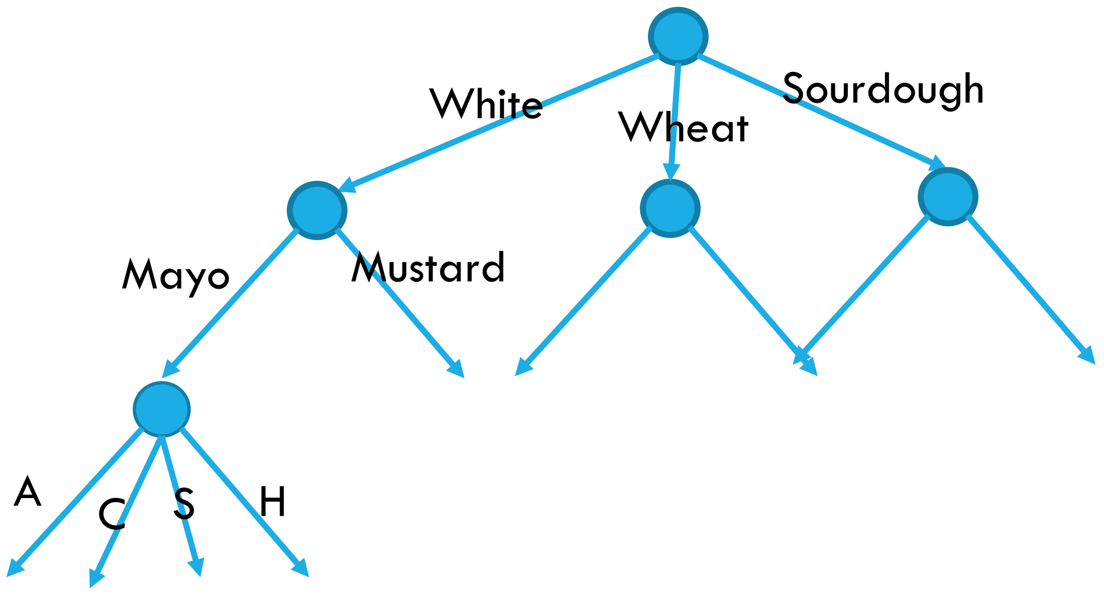
I'm still hungry…

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



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 cheeses.

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

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

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

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

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

Power Sets (answer)

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

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

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 (setup)

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 (attempt 1)

Step 1: 3 choices for hats.

Step 2: 3 choices for jerseys

Step 3:….

Baseball Outfits (answer 1)

Step 1: 3 choices for hats.

Step 2: 3 choices for jerseys.

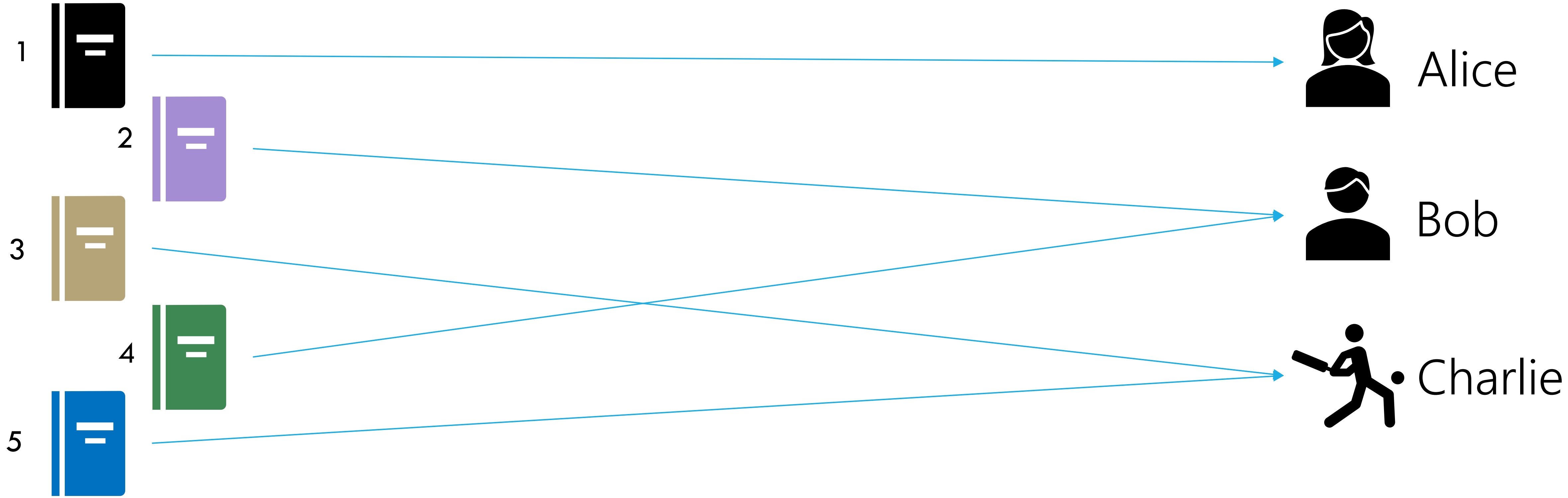
Step 3: Regardless of which jersey we choose, we have 2 options for pants (even though there are three options overall).

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

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 (attempt 1)

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 $32 \cdot 32 \cdot 32 = 32768$

Activity (answer)

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 (attempt 1)

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 (attempt 2)

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 sequence practice

How many length 3 sequences are there consisting of distinct elements of $\{1,2,3\}$.

Pause

Questions in combinatorics and probability are often dense. A single word can totally change the answer. Does order matter or not? Are repeats allowed or not? What makes two things "count the same" or "count as different"?

Let's look for some keywords

How many length 3 sequences are there consisting of distinct elements of $\{1,2,3\}$.

> **Sequences** implies that order matters – $(1,2,3)$ and (2,1,3) are different.
> **Distinct** implies that you can't repeat elements (1,2,1) doesn't count.
>
> $\{1,2,3\}$ is our "universe" – our set of allowed elements.

More sequence practice (result)

How many length 3 sequences are there consisting of distinct elements of $\{1,2,3\}$.

Step 1: 3 options for the first element.

Step 2: 2 (remaining) options for the second element.

Step 3: 1 (remaining) option for the third element.

$3 \cdot 2 \cdot 1$.

Factorial

That formula shows up a lot.

The number of ways to "permute" (i.e. "reorder", i.e. "list without repeats") n elements is "n factorial"

n factorial

$$n! = n \cdot (n-1) \cdot (n-2) \cdots 1$$

We only define $n!$ for natural numbers n.

As a convention, we define: 0!=1.

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 (result)

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