Welcome to CSE 312! 🥳 You're early!

- Grab a **handout** or download from the course website

- Want a **copy of these slides** to take notes? You can download them from the webpage cs.uw.edu/312

- Go to ~~pollev.com/cse312~~ and login with your at-uw email (not at-cs!) There's a poll running to make sure things are working

- Feel free to introduce yourself to the people sitting around you ☺

Introduction and Counting

CSE 312 23Su

Lecture 1

Staff

Instructor: Claris Winston
Email: clarisw@cs.washington.edu
Office hours: MW 1:15-2:30pm

Teaching Assistants (TAs)

✨ Zhi Yang Lim
✨ Hannah Swoffer
✨ Michael Whitmeyer
✨ Cleah Winston

Syllabus and Course Logistics

This Wednesday is a holiday, so we want to get through as much in lecture today as possible!

Syllabus walkthrough video posted on course website: cs.uw.edu/312
Watch this video and read through the syllabus

Immediate Logistics

Sections meet on Thursdays (starting this week)

- Section participation is graded – you can either
    - (1) Go to your assigned section in person – *recommended* ☺
    - (2) Attempt required problems and email to your TA

Concept Checks for each lecture and due before next lecture
- First one released after lecture today, due Friday 12pm

Homework 1 released today afternoon

Office hours
my office hours start today, TA's office hours start Tuesday

Lectures

- Recorded on Panopto
- Attendance it not required but highly recommended (and hopefully fun!)
- Please participate and ask questions!!!
- Lecture notes for each lecture – filled in notes posted on website

Accommodations

If you have something significant happen during the quarter, please reach out to use as soon as possible!

For now:

- Fill out **DRS accommodations** if needed

- Fill out **religious accommodations request form** if needed

- **Email me if you have other accessibility concerns** that are not covered by DRS (!)

Foundations of Computing?

We're going to learn probability and statistics for computer scientists ☺

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
…

Roadmap

- 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.

What should I get out of this class?

Precise mathematical communication
Both reading and writing dense statements.
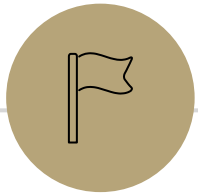
Probability in the "real world"
A mix of CS applications
And some actual "real life" ones.

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
(not as easy as 1-2-3?)

Why Counting?

We can't always just list out possible options and count it up!

Combinatorics: techniques for counting things that are difficult to count

Basics of **algorithm analysis**

Is this problem brute-forceable?

Estimation of how many operations an operation like find() would require

Basics of **probability theory**

Usually, "What are the chances of X" =

$$\frac{\text{how many ways can X occur}}{\text{how many ways can X occur} + \text{how many ways can X not occur}}$$

Remember sets from 311?

Remember sets from 311?

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}.

Remember sets from 311?

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.

Remember sets from 311?

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.

We're going to learn counting techniques to count the size of more complex sets!

Lunchtime 🍕

How many options do I have for lunch?

I could go to **Delfino's where there are 6 pizzas** I choose from, <u>or</u> I could go to **Supreme where there are 4 pizzas** I choose from (and **none of them are the same between the** two).

How many total choices?

Lunchtime 🍕

What are we counting here?

We're counting the size/cardinality of the set of all possible lunches

Lunchtime 🍕

How many options do I have for lunch?

I could go to **Delfino's where there are 6 pizzas** I choose from, <u>or</u> I could go to **Supreme where there are 4 pizzas** I choose from (and **none of them are the same between the** two).

How many total choices?

6 + 4 = 10

Lunchtime 🍕

How many options do I have for lunch?

I could go to **Delfino's where there are 6 pizzas** I choose from, <u>or</u> I could go to **Supreme where there are 4 pizzas** I choose from (and **none of them are the same between the** two).

How many total choices?

$6 + 4 = 10$

Lunchtime 🍕

How many options do I have for lunch?

I could go to **Delfino's where there are 6 pizzas** I choose from, <u>or</u> I could go to **Supreme where there are 4 pizzas** I choose from (and **none of them are the same between the** two).

How many total choices?

$D \sim$ Delfins $S \sim$ Suprme

$|D \cup S| = |D| + |S|$ bc $|D \cap S| \emptyset$

$6 + 4 = 10$

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: $n + m$.

Coffeetime

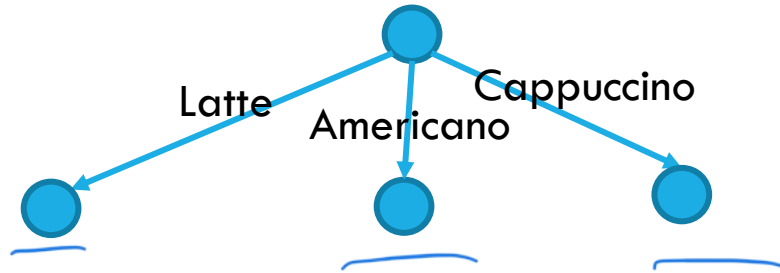Now I'm sleepy!

I decide to get coffee. My coffee is always:
One of three bases (latte americano, cappuccino).
One of two preperations (iced or hot)
One of four syrups (caramel, peppermint, hazelnut, gingerbread)
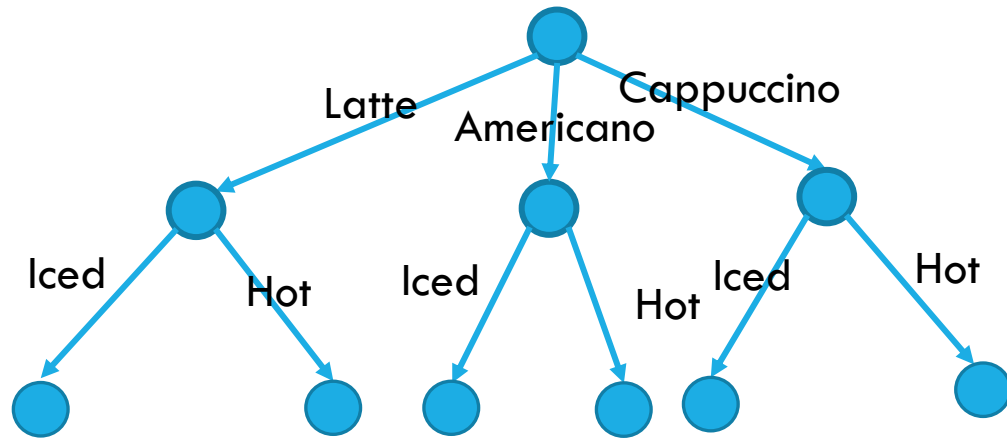
How many possible coffee orders do I have?

Coffeetime

Step 1: choose one of the <u>three</u> bases.
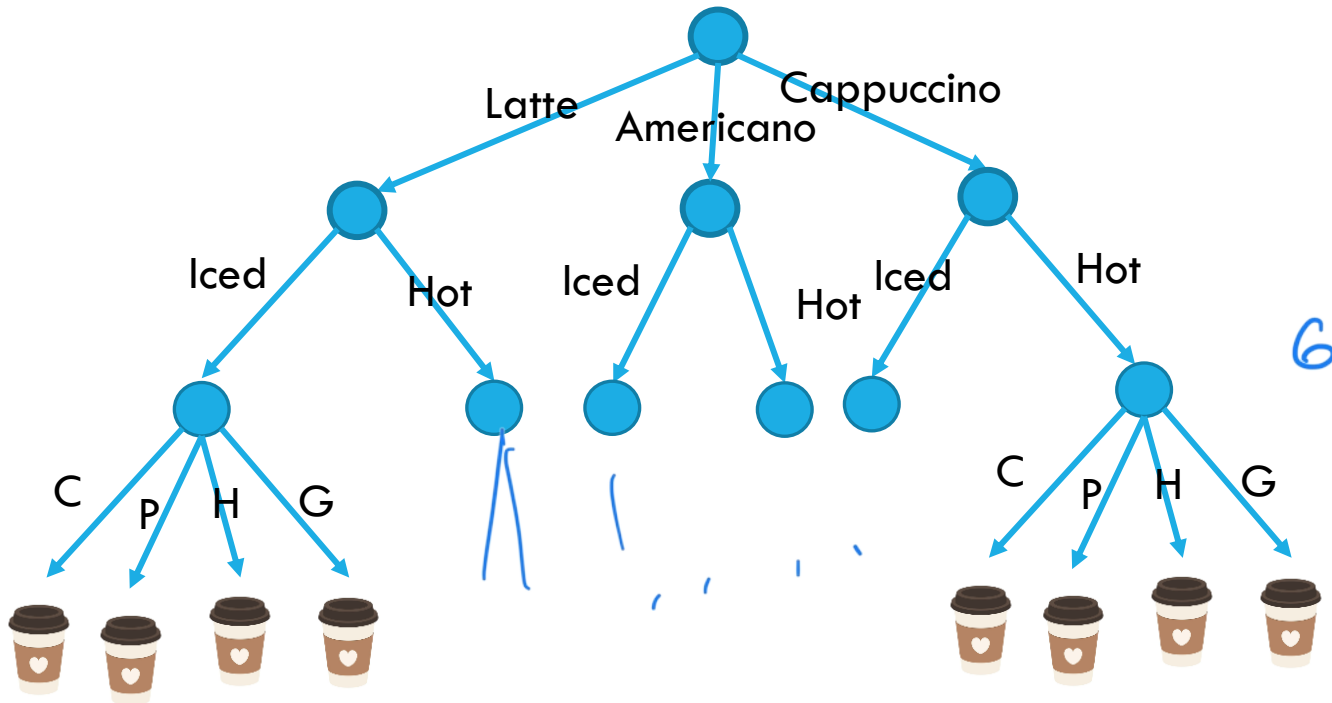
Coffeetime

Step 1: choose one of the <u>three</u> bases.

Step 2: regardless of step 1, choose one of the <u>two</u> preparations.

Coffeetime

Step 1: choose one of the <u>three</u> bases.

Step 2: regardless of step 1, choose one of the <u>two</u> preparations.

Step 3: regardless of steps 1 and 2, choose one of the <u>four</u> syrups.
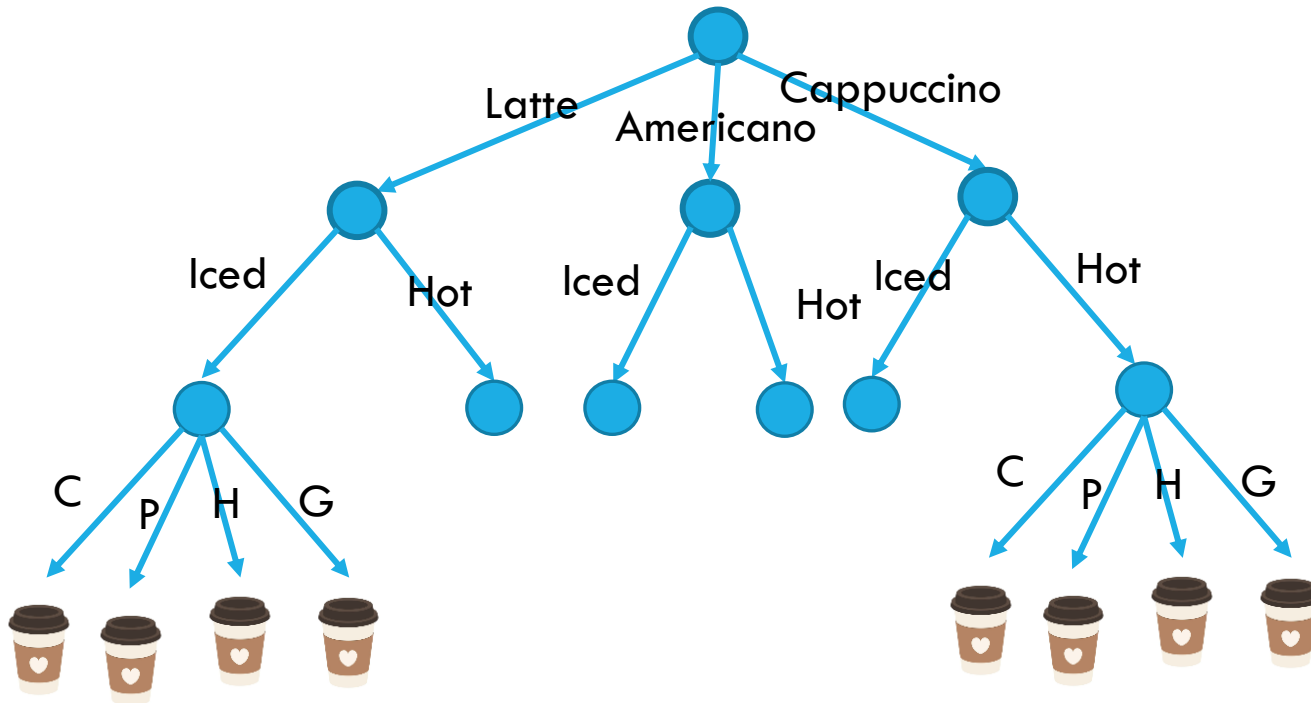
Coffeetime

Step 1: choose one of the <u>three</u> bases.

Step 2: regardless of step 1, choose one of the <u>two</u> preparations.

Step 3: regardless of steps 1 and 2, choose one of the <u>four</u> syrups.

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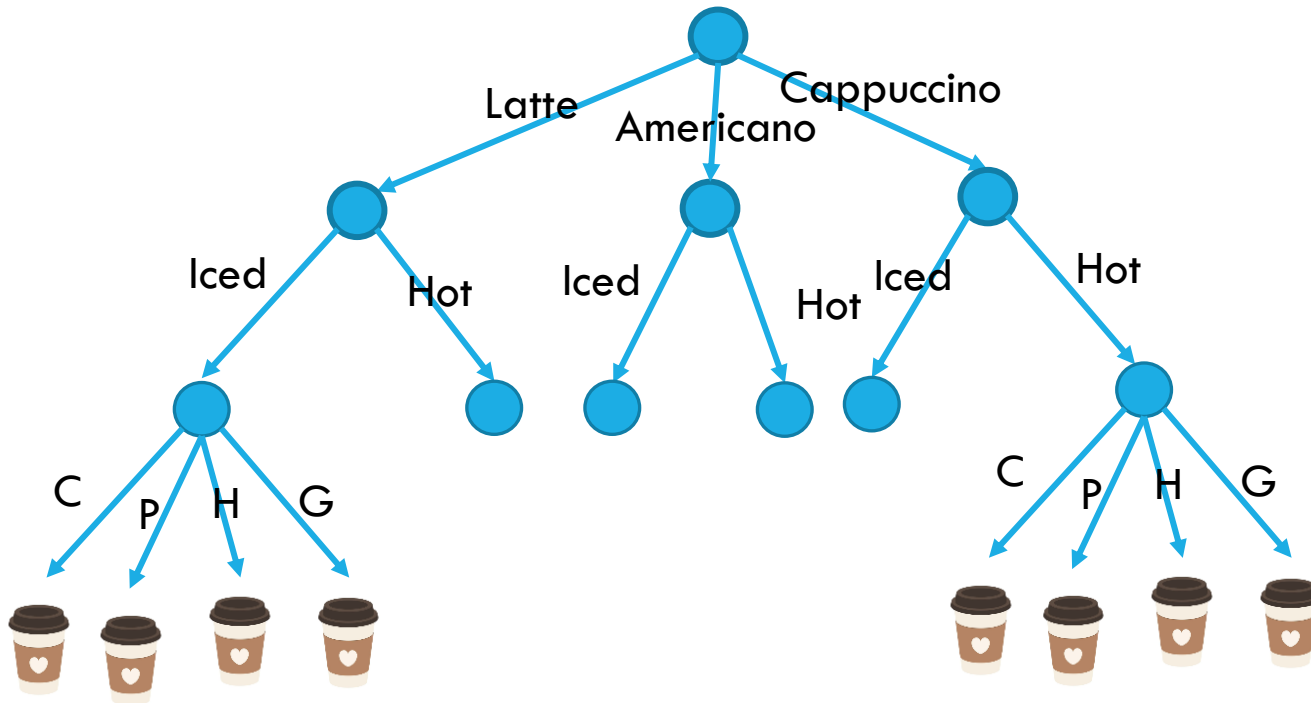
Coffeetime

Step 1: choose one of the <u>three</u> bases.

Step 2: regardless of step 1, choose one of the <u>two</u> preparations.

Step 3: regardless of steps 1 and 2, choose one of the <u>four</u> syrups.

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

Product Rule: If you have a <u>sequential process</u>, where step 1 has n_1 options, step 2 has n_2 options,...,step k has n_k options, and you choose one from each step, the total number of possibilities is $n_1 \cdot n_2 \cdots n_k$

Counting Rules

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

⭐ Used when we're counting something when we have **cases that don't overlap** ⭐

Product Rule: If you have a sequential process, where step 1 has n_1 options, step 2 has n_2 options,...,step k has n_k options, and you choose one from each step, the total number of possibilities is $n_1 \cdot n_2 \cdots n_k$
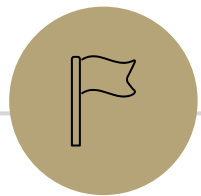
⭐ Used when we're counting something with a **sequential process** ⭐

Application of the Product Rule

Cartesian Products

Remember **Cartesian products**?

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

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

$|S \times T|$ how many pairs w one elem from S and one from t

Cartesian Products

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. $|S|$

Step 2: choose the element from T. $|T|$

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

Power Sets

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 $|\mathcal{P}(S)|$?

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 $S = \{e_1, e_2, \ldots, e_{|S|}\}$

Step 1: is e_1 in the subset? — yes

Step 2: is e_2 in the subset? yes

...   no

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

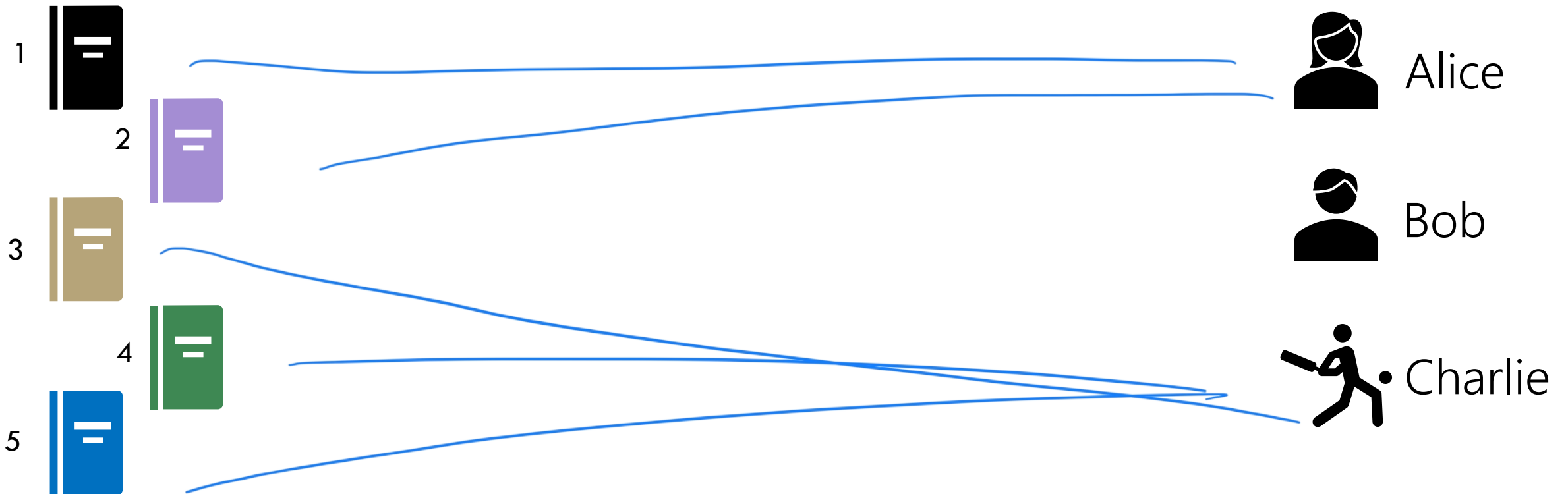$2 \cdot 2 \cdots 2$, $|S|$ times, i.e., $2^{|S|}$.

$\{e_1, e_2\}$

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

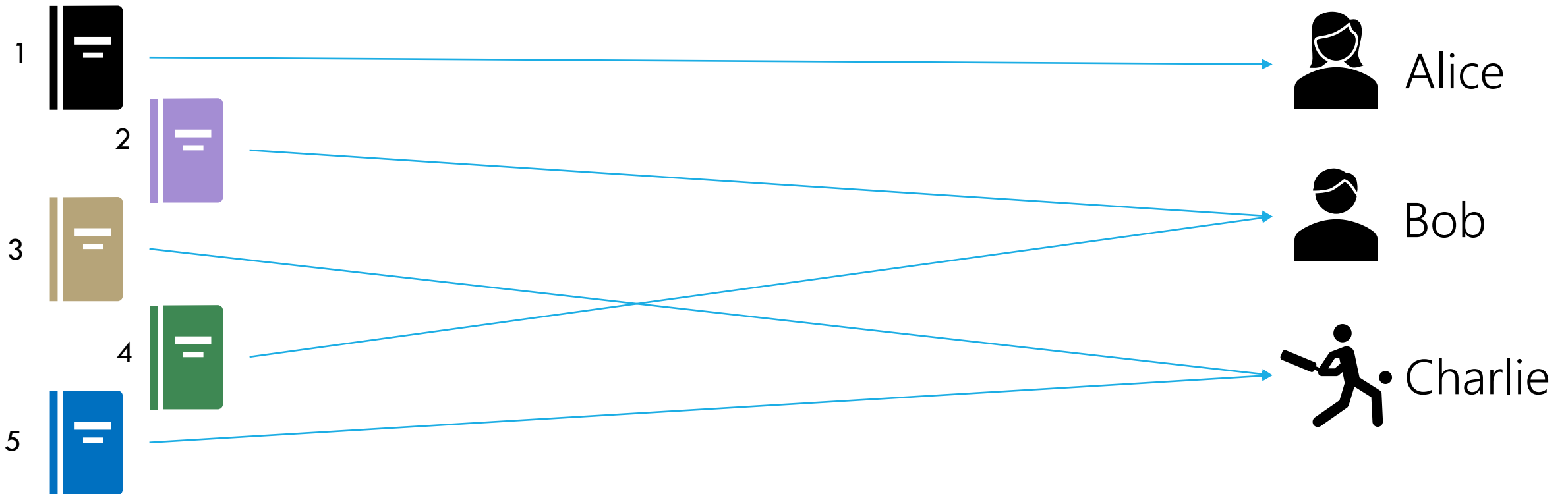We have 5 books to split to 3 people (Alice, Bob, and Charlie)

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

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!
1. Alice could get any of the $2^5 = 32$ subsets of the books. $1,2,3,4,5$
2. Bob could get any of the $2^5 = 32$ subsets of the books. $1,2,34,5$
3. Charlie could get any of the $2^5 = 32$ subsets of the books. $1,2,345$
Total is product of those three steps $32 \cdot 32 \cdot 32 = 32768$

overcount?   undercounts?   correct?

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! **This doesn't account for restriction of 1 book->1 person**

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.

$$3 \cdot 3 \cdot 3 \cdot 3 \cdot 3$$

Total: 3^5.
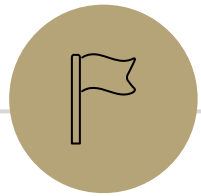
Strings 📝 (and more counting rules!)

Strings (with **repeats** allowed)

How many length 3 sequences are there consisting of elements from $\{1, 2, 3, 4, 5\}$ with <u>repeats allowed</u>?

312   112

$$5 \cdot 5 \cdot 5 = 5^3$$

Strings (with **repeats** allowed)

How many length 3 sequences are there consisting of elements from $\{1, 2, 3, 4, 5\}$ with repeats allowed?

Step 1: Pick for the 1st character – 5 options for this

Step 2: Pick for the 2nd character – 5 options for this

Step 3: Pick for the 3th character – 5 options for this

Applying the product rule: $5 \cdot 5 \cdot 5 = 5^3$ possible strings

Strings (with **repeats** allowed)

How many length 3 sequences are there consisting of elements from $\{1, 2, 3, 4, 5\}$ with repeats allowed?

Step 1: Pick for the 1st character – 5 options for this

Step 2: Pick for the 2nd character – 5 options for this

Step 3: Pick for the 3th character – 5 options for this

Applying the product rule: $5 \cdot 5 \cdot 5 = 5^3$ possible strings

$k-$**sequences** *(an application of the product rule)*

n^k length k sequences from an alphabet of size n, **with repeats allowed**

Strings (with **distinct** elements, use **all** characters)

How many length 5 sequences are there consisting of <u>distinct</u> elements of $\{1, 2, 3, 4, 5\}$. *(in other words, rearranging these 5 elements)*

Step 1: <u>5</u> options for the *first* character.

Step 2: <u>4</u> (remaining) options for the *second* character.

Step 3: <u>3</u> (remaining) options for the *third* character.

Step 4: <u>2</u> (remaining) options for the *fourth* character.

Step 5: <u>1</u> (remaining) options for the *fifth* character.

12345

34521

Strings (with **distinct** elements, use **all** characters)

How many length 5 sequences are there consisting of <u>distinct</u> elements of $\{1, 2, 3, 4, 5\}$. *(in other words, rearranging these 5 elements)*

Step 1: 5 options for the *first* character.

Step 2: 4 (remaining) options for the *second* character.

Step 3: 3 (remaining) options for the *third* character.

Step 4: 2 (remaining) options for the *fourth* character.

Step 5: 1 (remaining) options for the *fifth* character.

Applying the <u>product rule</u>: $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5!$.

Strings (with **distinct** elements, use **all** characters)

How many length 5 sequences are there consisting of <u>distinct</u> elements of $\{1, 2, 3, 4, 5\}$. *(in other words, rearranging these 5 elements)*

Step 1: 5 options for the *first* character.

Step 2: 4 (remaining) options for the *second* character.

Step 3: 3 (remaining) options for the *third* character.

Step 4: 2 (remaining) options for the *fourth* character.

Step 5: 1 (remaining) options for the *fifth* character.

Applying the <u>product rule</u>: $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5!$.

n **factorial** *(another application of the product rule)*

The number of ways to arrange n elements. $n! = n \cdot (n-1) \cdot (n-2) \cdots 1$

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: $0! = 1$.

Other examples of factorials:

- Ordering n people in a line $n!$
- Shuffling a <u>deck</u> <u>of cards</u> $52!$

Strings (with **distinct** elements, but we don't use all the letters)

How many length 3 sequences are there consisting of distinct elements of $\{1, 2, 3, 4, 5\}$.

312   311

Step 1: 5 options for the *first* element.

Step 2: 4 (remaining) options for the *second* element.

Step 3: 3 (remaining) options for the *third* element.

Applying the product rule: $5 \cdot 4 \cdot 3$.

Strings (with **distinct** elements, but we don't use all the letters)

How many length 3 sequences are there consisting of <u>distinct</u> elements of $\{1, 2, 3, 4, 5\}$.

Step 1: 5 options for the *first* element.

Step 2: 4 (remaining) options for the *second* element.

Step 3: 3 (remaining) options for the *third* element.

Applying the <u>product rule</u>: $5 \cdot 4 \cdot 3$.

k permutation *(another application of the product rule)*

Counts number of k-element sequences of distinct symbols from set of n symbols

k-permutation

k-permutation

The number of k-element sequences of distinct symbols from a universe of n symbols is:

$$P(n, k) = n \cdot (n - 1) \cdots (n - k + 1) = \frac{n!}{(n - k)!}$$

k-permutation

k-permutation

The number of k-element sequences of distinct symbols from a universe of n symbols is:

$$P(n, k) = n \cdot (n - 1) \cdots (n - k + 1) = \frac{n!}{(n - k)!}$$

Examples:
- Ordering k people from a group of n people in a line
- Assigning k players from a team of n to k different positions (so order matters)
- Awarding gold, silver, and bronze medal among n participants - $P(n, 3)$

$P(n,k)$

$P(n,k)$

Strings (same as before, but now we don't care about order)

How many subsets of 3 numbers are there consisting of distinct elements of $\{1, 2, 3, 4, 5\}$.

This means that (1, 2, 3), (2, 3, 1), etc. should be counted as 1 outcome

Number of Subsets - k-combination

k-combination

The number of k-element subsets from a set of n distinct symbols is:

$$C(n,k) = \frac{P(n,k)}{k!} = \frac{n!}{k!\,(n-k)!} = \binom{n}{k}$$

Examples:
- Selecting k people from a group of n to form a team
- Choosing k courses from a set of n courses to take in fall quarter
- Select 3 winners among n participants (no ordering between winners) - $C(n,3)$

$\binom{n}{k}$

$\binom{n}{k}$

$\binom{n}{3}$

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

Pause – looking for keywords!

How many length 3 sequences are there consisting of distinct elements of {1, 2,3, 4, 5}?

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.

{1,2,3,4,5} is our "universe" – our set of allowed elements.

Looking for an **ordered sequence** of **distinct elements** from a larger set → **permutation**!

unordered ⟶ *combination*

One more counting technique!

How many length 5 strings are there over $\{a, b, \dots, z\}$ with **at least 1 a**?

One more counting technique!

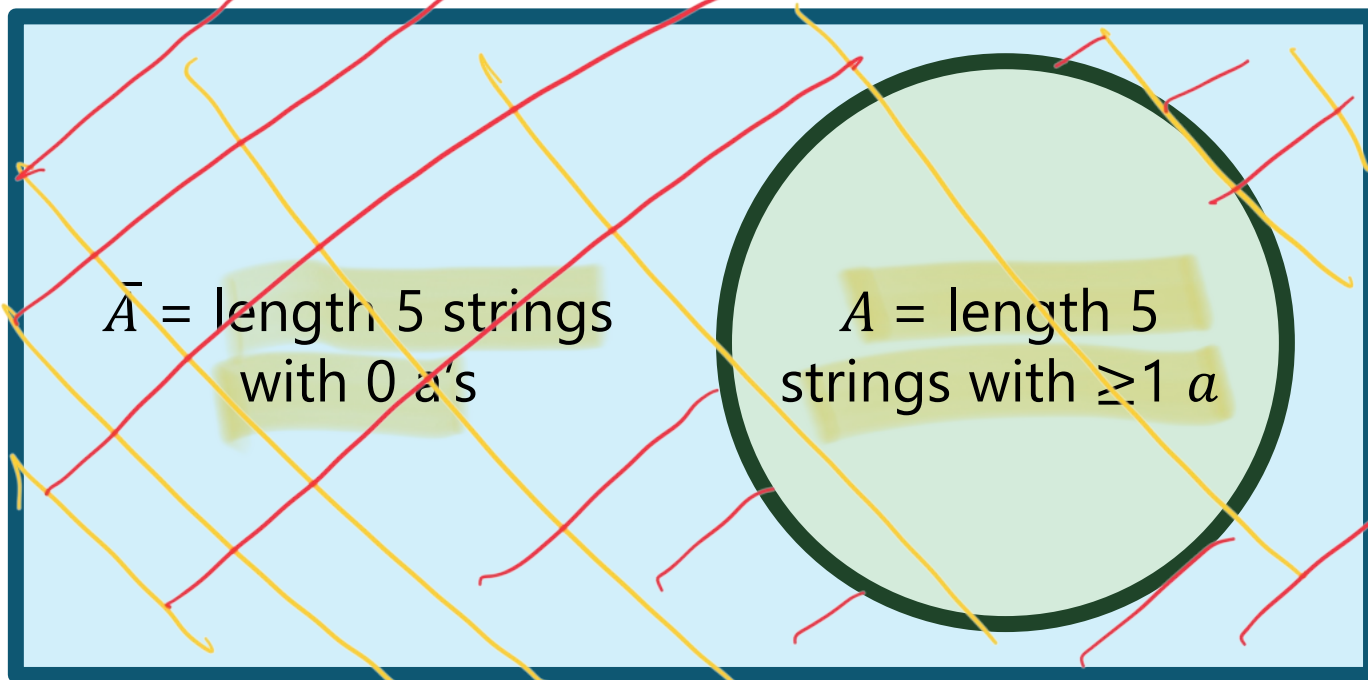How many length 5 strings are there over $\{a, b, \dots, z\}$ with **at least 1** a?

we can rephrase this as: "How many length 5 strings are there that do <u>not</u> have no a's?

One more counting technique!

How many length 5 strings are there over $\{a, b, \dots, z\}$ with **at least 1** a?
we can rephrase this as: "How many length 5 strings are there that do <u>not</u> have no a's?

$\mathcal{U}$ = all possible length 5 strings

$\bar{A}$ = length 5 strings with 0 a's
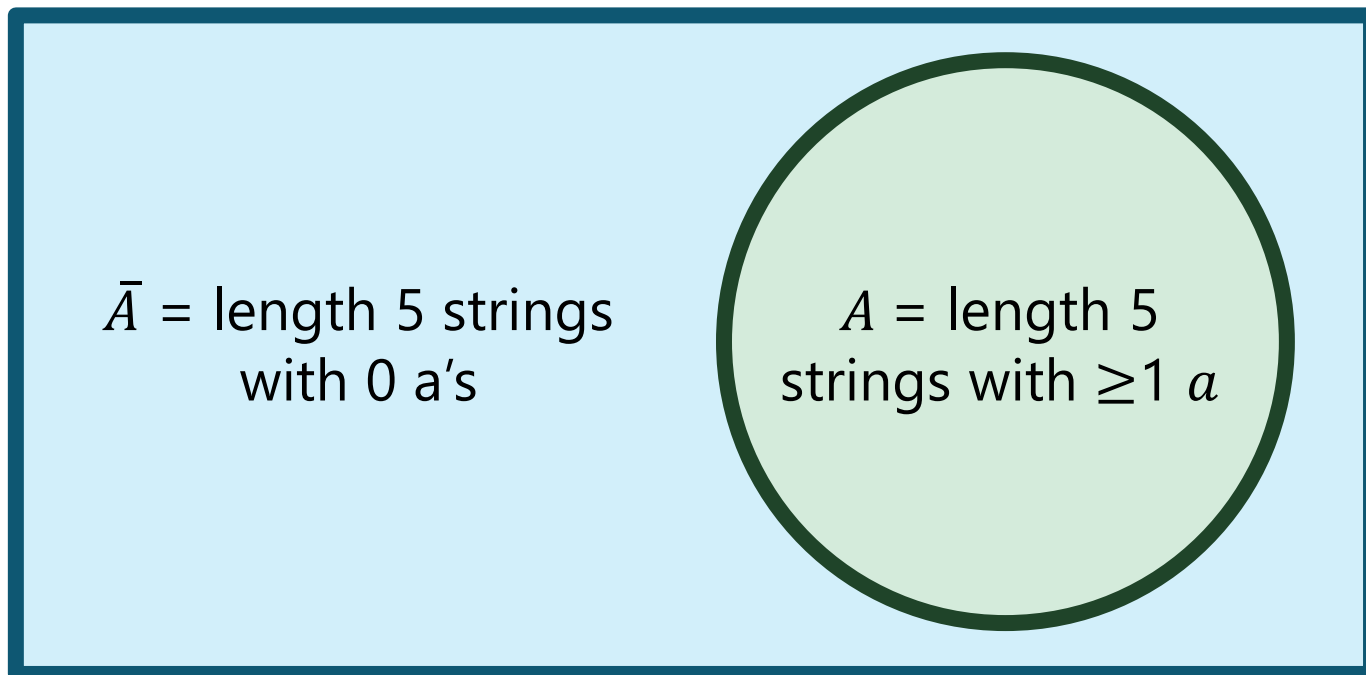
A = length 5 strings with ≥ 1 a

$$|A| = |\mathcal{U}| - |\bar{A}|$$
$$= 26^5 - 25^5$$

One more counting technique!

How many length 5 strings are there over $\{a, b, \ldots, z\}$ with **at least 1 a**?

we can rephrase this as: "How many length 5 strings are there that do <u>not</u> have no a's?

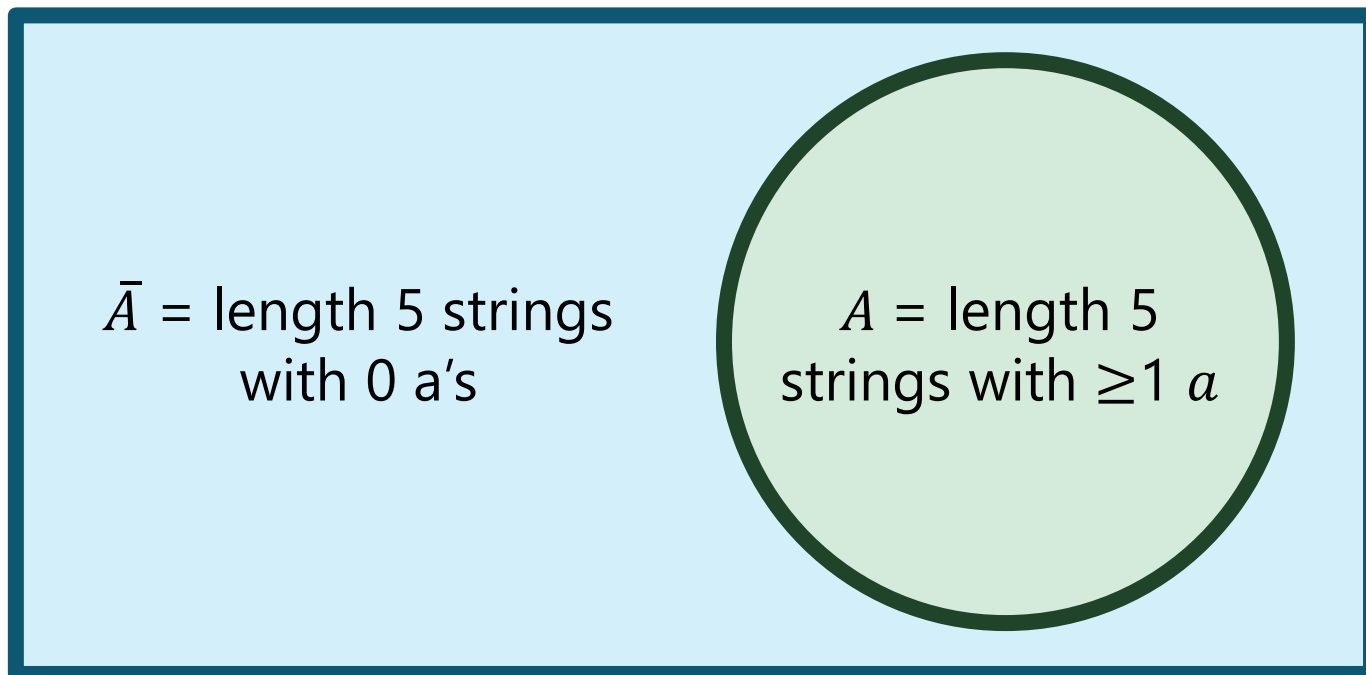$\mathcal{U}$ = all possible length 5 strings

We're interested in $|A|$ -

$$|A| = |\mathcal{U} \backslash \bar{A}| = |\mathcal{U}| - |\bar{A}|$$

$\bar{A}$ = length 5 strings with 0 a's

A = length 5 strings with ≥ 1 a

One more counting technique!

How many length 5 strings are there over $\{a, b, \ldots, z\}$ with **at least 1 a**?
we can rephrase this as: "How many length 5 strings are there that do <u>not</u> have no a's?

$\mathcal{U}$ = all possible length 5 strings

$\bar{A}$ = length 5 strings
with 0 a's

A = length 5
strings with $\geq 1\ a$

We're interested in $|A|$ -

$$|A| = |\mathcal{U} \backslash \bar{A}| = |\mathcal{U}| - |\bar{A}|$$
$$= 26^5 - 25^{5}$$

Complementary Counting

$$|A| = |\mathcal{U} \backslash \bar{A}| = |\mathcal{U}| - |\bar{A}|$$

total options – options for A to not occur = options for A to occur

$\mathcal{U}$ = all possible options

$\bar{A}$ = ways for A to not occur

A = what we're interested in
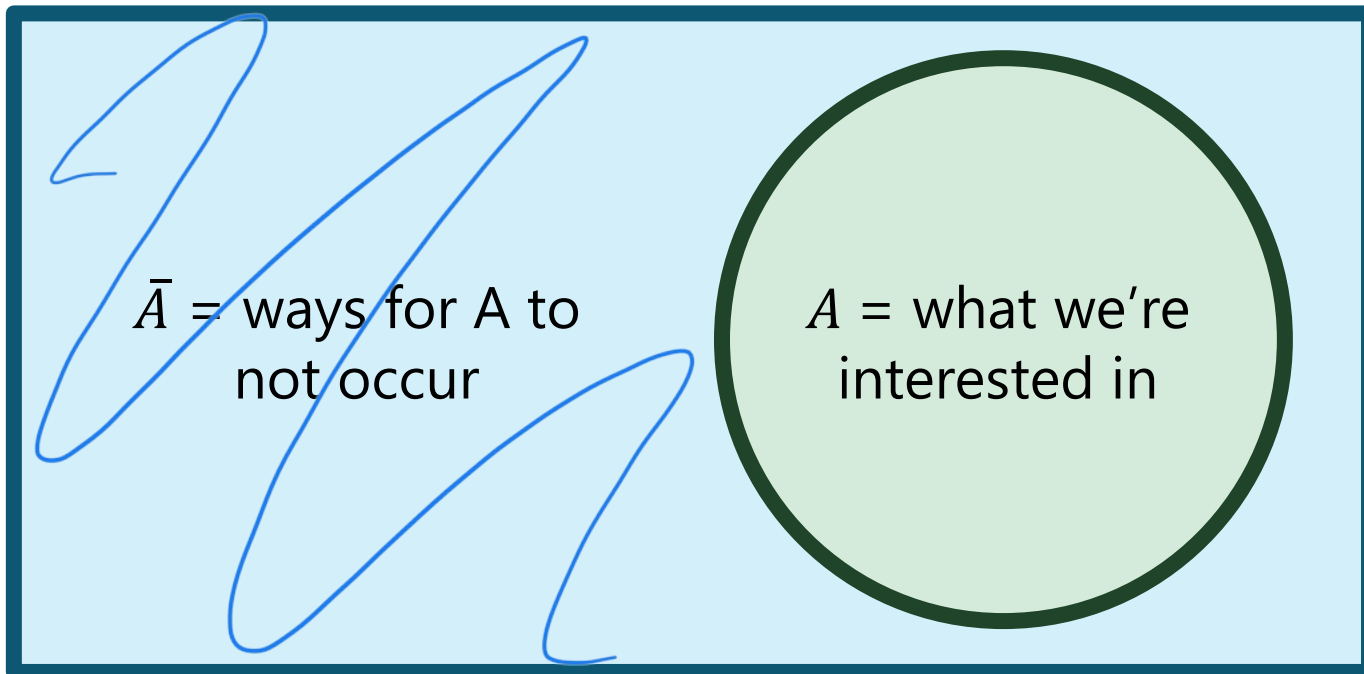
Phrases that *might* indicate complementary counting:
- ..options so that __ does not occur
- ..ways for at least one __ to occur = total – ways for none of _ to occur

That was a lot!!

"High level" counting approaches

- Sum rule – counts options where we have disjoint cases
- Product rule – counts options when we have a sequential process
- Complementary counting – we count the options by finding that total options and subtracting the non-desired options

These may be used together! And may be used with the following --

That was a lot!!

- (direct application of product rule) k-Sequences
  - n^k ways to make a k length ordered sequence from a universe of n elements

- (direct application of product rule) Factorial
  - $n!$ ways to rearrange n objects (e.g., arranging n people in a line for a photo)

- Permutations
  - $P(n, k)$ ways to select k objects from n in total <u>where order matters</u> (e.g, arranging k people from a group of n people in a line for a photo)

- Combinations
  - $C(n, k)$ ways to select k objects from n in total <u>where order does not matter</u> (e.g., selecting k people from a group of n people to be in a photo, but we don't care about the order they stand in)

That was a lot!!

Section this week will be **very** helpful with getting practice with applying these concepts!

TODOs from today

⭐ read through syllabus + watch syllabus walk through

⭐ make sure DRS accommodations are setup

⭐ email me with other accessibility concerns

⭐ first concept check will be released after class and <u>due Friday 12pm</u>

⭐ attend first section on Thursday!

⭐ don't come to class on Wednesday ☺

⭐ PSet1 released today evening

FRIDAY
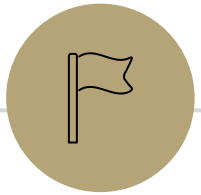
More Practice

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

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.

Step 1: __ choices for hats

Step 2: __ choices for jerseys

Step 3: __ choices for pants

Total options: ____

Baseball Outfits

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.

Step 1: 3 choices for hats

Step 2: 3 choices for jerseys

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

Total options: $3 \cdot 3 \cdot 2 = 18$.