You’re here early! FYI:

These slides are now posted on the course webpage: www.cs.uw.edu/312

We will be using PollEverywhere, with link: www.pollev.com/312 (there’s a pre-course poll running to make sure things are working!)

Intro and Counting

CSE 312 23Su

Lecture 1

Staff

Instructor: Shreya Jayaraman

Email: shreyaj@cs.uw.edu

OH: Monday & Wednesday after class

TAs

Zhi Yang Lim

Alex Liu

Vlad Murad

Sophie Lin Robertson

Davin Tjia

Jolie Zhou

Course webpage

www.cs.uw.edu/312

- **Syllabus**, and other important information is available here
- Homework/section materials/etc. will be posted here

Logistics

Lectures

MWF 12:00pm-1:00pm

Recorded on Panopto

Questions encouraged!

Sections

Th 12:00pm-1:00pm

Attend your registered section

Not recorded for student privacy

Handouts and solutions posted on course webpage

Office Hours

Linked on course webpage, start tomorrow (Thursday)

Logistics

Gradescope

Assignment submission

EdStem

Discussion board

Canvas

Panopto recordings

Assignment Breakdown

Concept Checks (10%)

★ submit empty if you want the answers later & don't plan to submit answers.

Short per-lecture 'quizzes' to keep pace with course and check gaps in understanding – due before the next lecture (11.30am)

~7 Homeworks (50%)

Mostly written (must be typeset), occasional programming question

1 homework per week

Midterm (15%) – July 17th (tentative)

Final (20%) – August 18th

EC: Effort, Participation and Altruism (EPA)

Contacting Us

EdStem

Use discussion board for content-related questions

Check frequently: all announcements will be posted here

For personal/private matters,

EdStem private post: visible to TAs and Shreya

Email Shreya (shreyaj@cs.uw.edu)

Anonymous feedback form on webpage

Collaboration Policy

We encourage **collaboration** – there is a description of academically honest collaboration in the syllabus

Essentially: on homeworks you are **encouraged** to collaborate, but the writeup must be your own.

- If you're working with others, don't take notes from your discussions
- And take a **30 minute break** between discussions and your writeup

The goal of these rules is to make sure you've learned how to do the problems.

TODOs:

- Make sure you have access to course Canvas, EdStem discussion board and Gradescope (linked on website).
- File for DRS accommodations if relevant.

Foundations of Computing II?

Intro to probability and statistics for computer scientists ☺

Applications seen in…

- Machine learning
- Natural language processing
- Cryptography
- Error-correcting codes
- Data structures
- Complexity theory
- Algorithm design
- And more!

Roadmap

- Combinatorics (counting)
  - Permutations, combinations, inclusion-exclusion, pigeonhole principle
- Formal definitions of 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, with some focus on ML

What should you take out of this class?

Precise mathematical communication

- Ability to use and comprehend probabilistic notation

Probability in the ‘real world’

- Applications of probability in CS & the real world

Refine your intuition

- About how likely events are in the real world – and when people are lying to you with statistics!

Foundation of higher-level computation

- Probabilistic data structures, algorithmic analysis, machine learning, etc.

What should you take out of this class?

This class often comes across as teaching many problem-solving **tools**

What comes naturally to us after years of school is… Memorizing **how** to apply a tool

What I hope you take away is more **intuition** about when to apply a tool to a problem – whether you go on to take machine learning classes or not, I think this is the fun part of problem-solving ☺

Counting

Why does this matter?

Combinatorics: techniques for counting things that are larger or more difficult to count with what we usually think of as counting (enumeration)

Basis of probability

Count the number of possible outcomes

$\frac{1}{2}$

"what are the chances of X" = $\frac{\#\ of\ ways\ for\ X\ to\ occur}{\#\ of\ ways\ for\ X\ to\ occur + \#\ of\ ways\ for\ X\ to\ not\ occur}$

Basis of algorithm analysis

Is a problem brute-forceable, given the number of possible outcomes?

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

Counting Rules (Yes, really)

Counting Rules

We go to a boba shop, with a menu consisting of:

- 5 milk tea flavors
- 8 fruit tea flavors

If I order exactly 1 drink, how many possible orders could I make?

$5 + 8 = 13$

Counting Rules

We go to a boba shop, with a menu consisting of:

- 5 milk tea flavors
- 8 fruit tea flavors

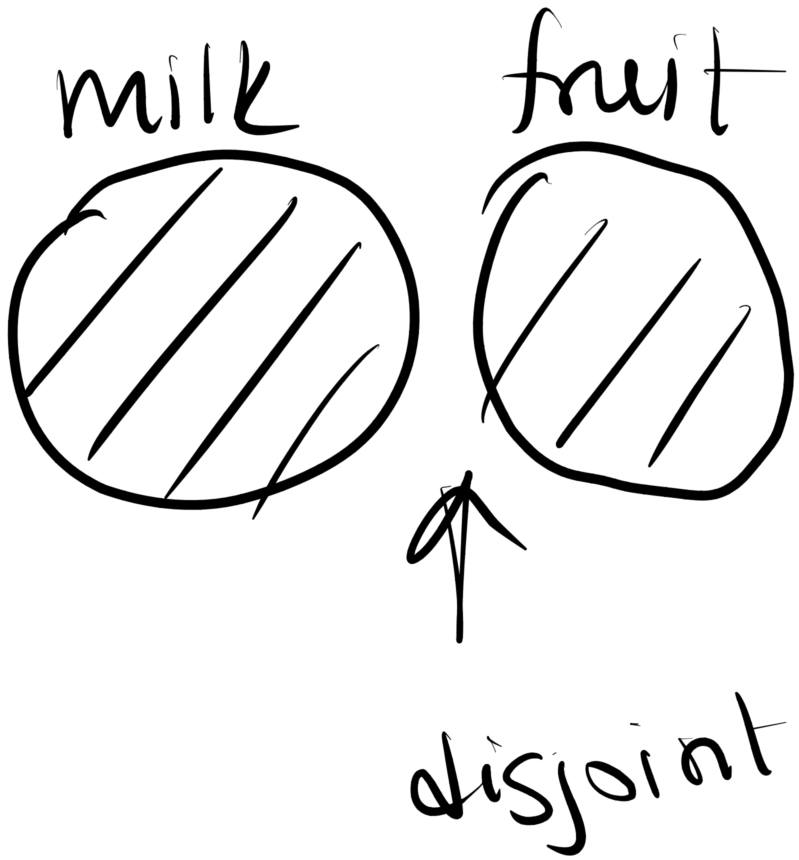
If I order exactly 1 drink, how many possible orders could I make?

5 + 8 = 13 possible orders!

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: $n + m$.

What are we counting?

- The size of a **set** of objects with a given, specific property



Remember sets? (We <3 311)

A set is an **unordered** collection of **unique** elements.

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

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

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

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

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

Counting Rules

I'm still hungry…

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

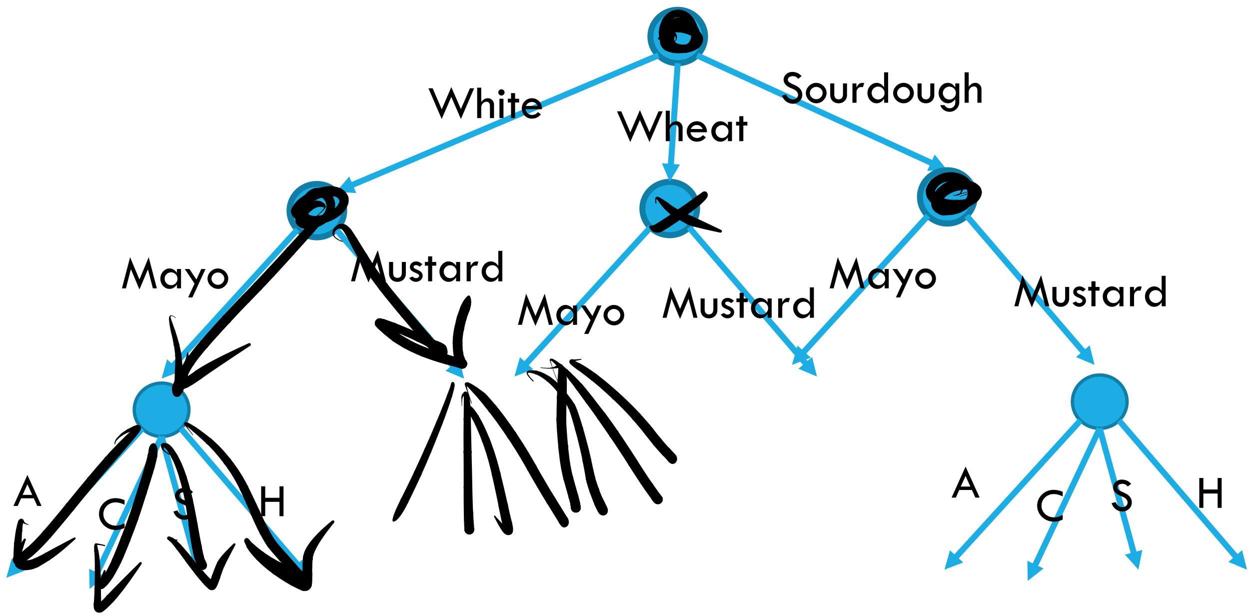
- One of three types of bread (white, wheat, or sourdough). 3
- One of two spreads (mayo or mustard)
- One of four cheeses (American, cheddar, swiss, or Havarti)

How many sandwiches can I make?

☆ sequential process

3

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 breads.

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

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 $\boldsymbol{n_1 \cdot n_2 \cdots n_k}$

Applying counting rules

Strings

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

$$\boxed{26} \cdot \boxed{26} \cdot \boxed{26} \cdot \boxed{26} \cdot \boxed{26} = 26^5$$

How many binary strings of length n are there?

$\{0,1\}$

$$\underset{1}{\boxed{2}} \cdot \underset{2}{\boxed{2}} \cdot \ldots \cdot \underset{n}{\boxed{2}} = 2^n$$

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

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 $|S \times T|$?)

$|S \times T|$

(x, y)

$|S \times T| = |S| \times |T|$

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.

Step 2: choose the element from T.

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

Power Sets

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

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

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

S

T -

$= 2 \times 2 \times 2$

$|\mathcal{P}(S)| = 2^{|T|}$

Power Sets

$X \subseteq S$

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

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

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?

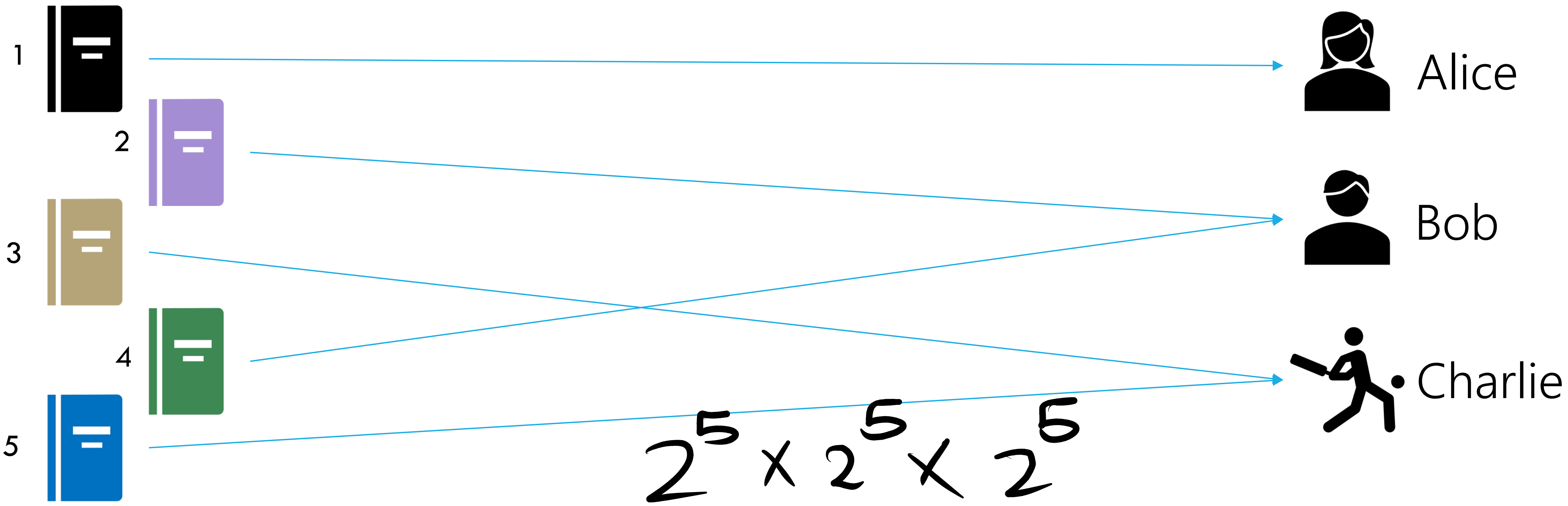
$2 \cdot 2 \cdots 2$, |S| times, i.e., $2^{|S|}$.

Assigning Books

$2^{|S|} = 2^5$ subsets of 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).



$2^5 \times 2^5 \times 2^5$

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. $\{1, 2\}$

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 $32 \cdot 32 \cdot 32 = 32768$

Is this attempt counting correctly?

Overcounting or undercounting?

Not sure?

Fill out the poll everywhere so Shreya can adjust her explanation

Go to pollev.com/312 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 Sum 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.

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.

One More Counting Technique

How many length 5 strings are there over the alphabet {a, b, …, z} with **at least** 1 'a'?

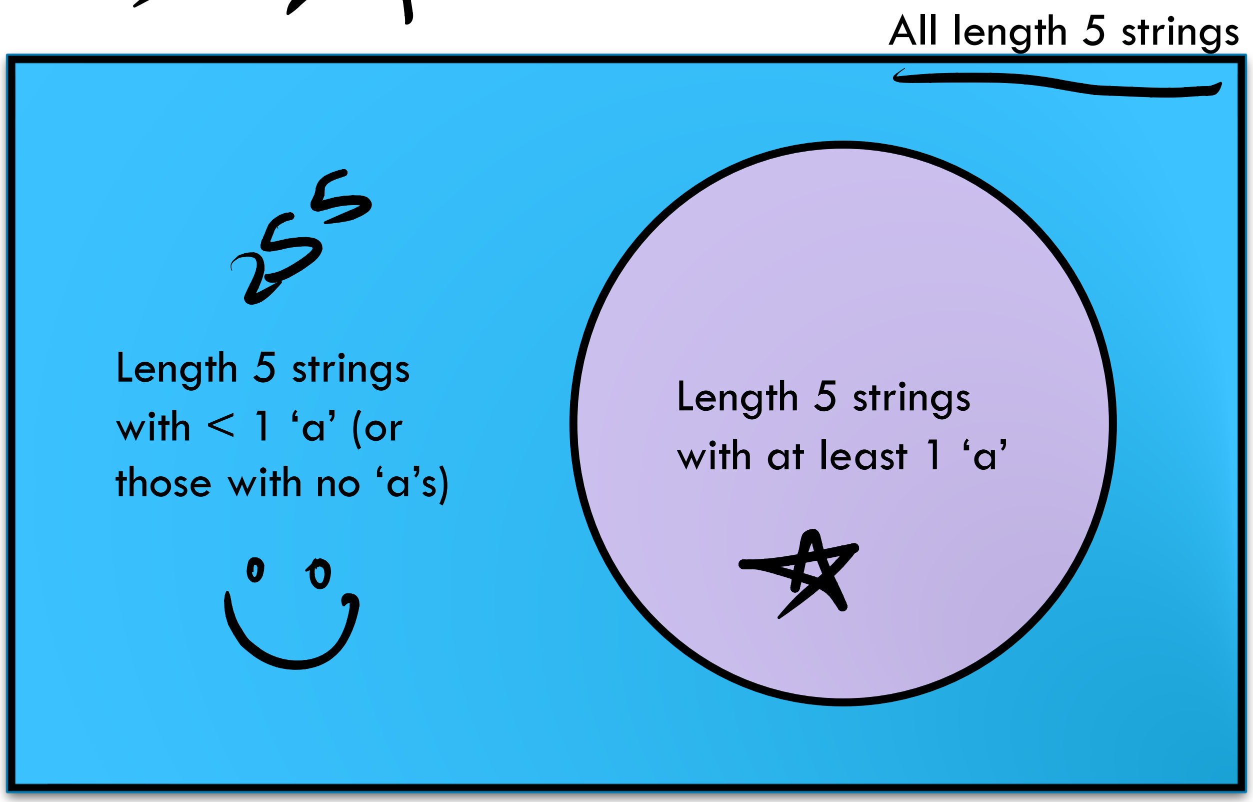
One More Counting Technique

$25 \cdot \_ \cdot \_ \cdot \_ \cdot 25 = 25^5$

How many length 5 strings are there over the alphabet {a, b, …, z} with **at least** 1 ‘a’?

≥ 1

$\underbrace{< 1 \text{ 'a'}}_{0 \text{ 'a'}}$

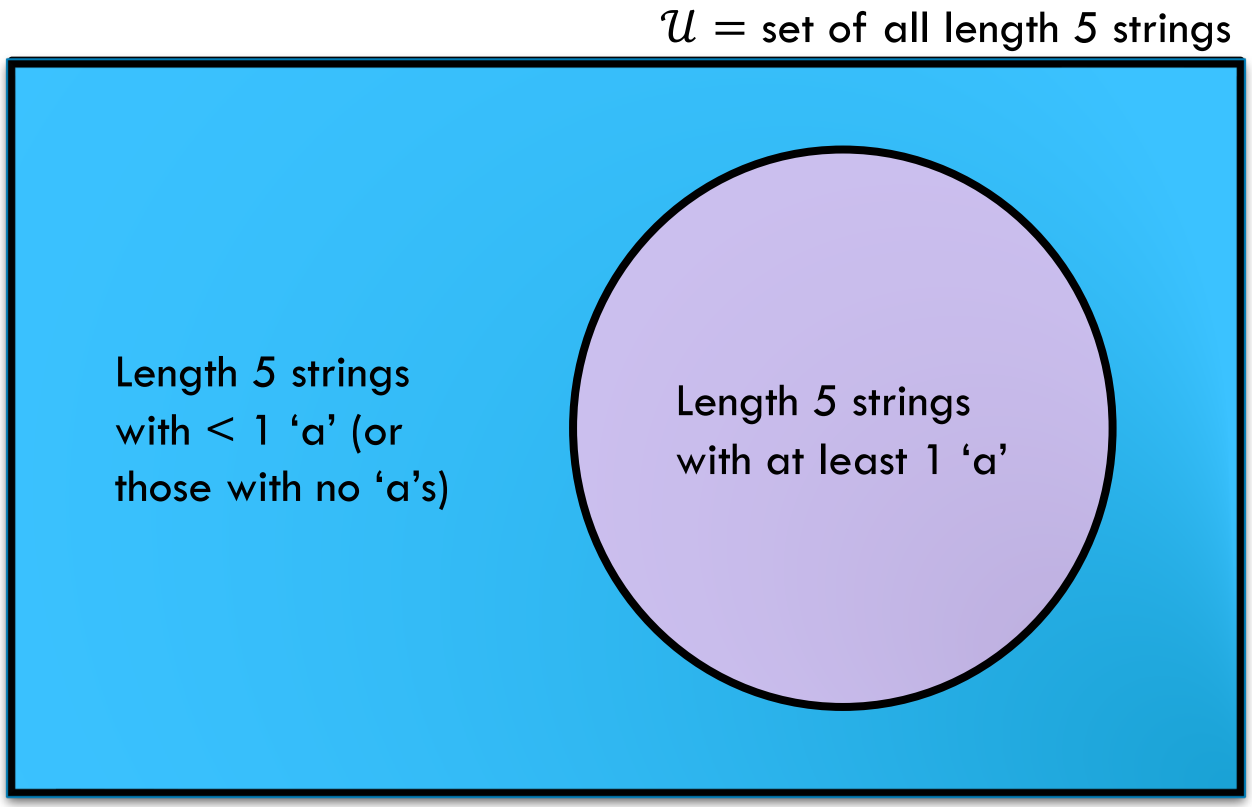


-26^5

$26-1$

One More Counting Technique

How many length 5 strings are there over the alphabet {a, b, …, z} with **at least** 1 ‘a’?



We are interested in |A|.

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

One More Counting Technique

Complementary Counting

Count the complement of the set A you’re interested in.

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

More sequence practice

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

order matters

no repeats

5 $\{1,2,3,4,5\} \rightarrow 5!$

$\underline{3} \times \underline{2} \times \underline{1} = 3!$

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 $\{1,2,3\}$?

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

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., "order n distinct 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: $0! = 1$.

Key Takeaways

- 312 is useful for being a better arguer! And being better at LeetCode! But hopefully more than that…
- Problem-solving technique: sometimes, we need to be creative about fitting a problem into a method that we know how to use (i.e., we need to change our perspective in order to apply our tools!)
- Keywords are important in 312! (This is a lot like a math class)

- sum rule, product rule, complementary counting, permutations (factorials)

Extra Practice

ATM Pins

How many 10-digit pin codes are there with no repeating digits?

- Each digit is one of {0, 1, 2, 3, 4, 5, 6, 7, 8, 9}.

$$10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 10!$$

ATM Pins

How many 10-digit pin codes with **at least one digit repeated once?**

- E.g. 1111111111, 1234567899, 1224567889

$$\begin{aligned} &\mathit{pins\ with\ at\ least\ one\ digit\ repeated\ once} \\ &= \mathit{all\ pins}\ - \mathit{pins\ with\ no\ repeats} \\ &= 10^{10}\ -10! \end{aligned}$$

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).

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