CSE 312

Foundations of Computing II

Anna R. Karlin

Welcome and happy new year!!

https://courses.cs.washington.edu/312

Questions and Discussions

- **Office hours throughout the week (starting Friday)**

- **Ed Discussion**
  - You should have received an invitation (synchronized with the class roster)
  - Use Ed discussion forum as much as possible. You can make private posts that only the staff can view! Email instructors for personal issues.

Lectures and Sections

- **Lectures MWF**
  - 1:30-2:20pm
  - Recorded on Panopto
- **Ask questions during class**
  - I love it when the class is interactive! And trust me, you'll like it more too.
  - If you don't understand something, ask!
  - There will also be an edstem thread for each lecture where you can ask follow-up questions.
  - Try to answer your fellow classmate's questions on edstem!

Lectures and Sections

- **Sections Thu (start tomorrow!)**
  - Not recorded, for privacy of student discussion
  - Important preparation for homework.

To make tomorrow's section useful, I will try to get through a lot today.

Concept checks

- **Concept checks after each lecture**
  - Released 3pm (30 mins after lecture). Must be done by 30 mins before the next lecture.
  - Simple questions to reinforce concepts taught in each class.
  - Keep you engaged throughout the week, so that homework becomes less of a hurdle

- **Important!!!!**
  - If you don't submit by the due date, you won't be able to see the solutions until much later. **So submit, even if you don't answer any of the questions!! You can resubmit as many times as you want before the deadline.**

Grading etc.

I will discuss this in more detail on Friday.

For now, check out the syllabus for policies on late submission, grading breakdown, etc.

Bring questions about syllabus and administrivia to class Friday.

Homework

- HW1 is already out and is due 11:59pm next Wednesday.
- In general, every HW is due one week after it's released.

Foundations of Computing II

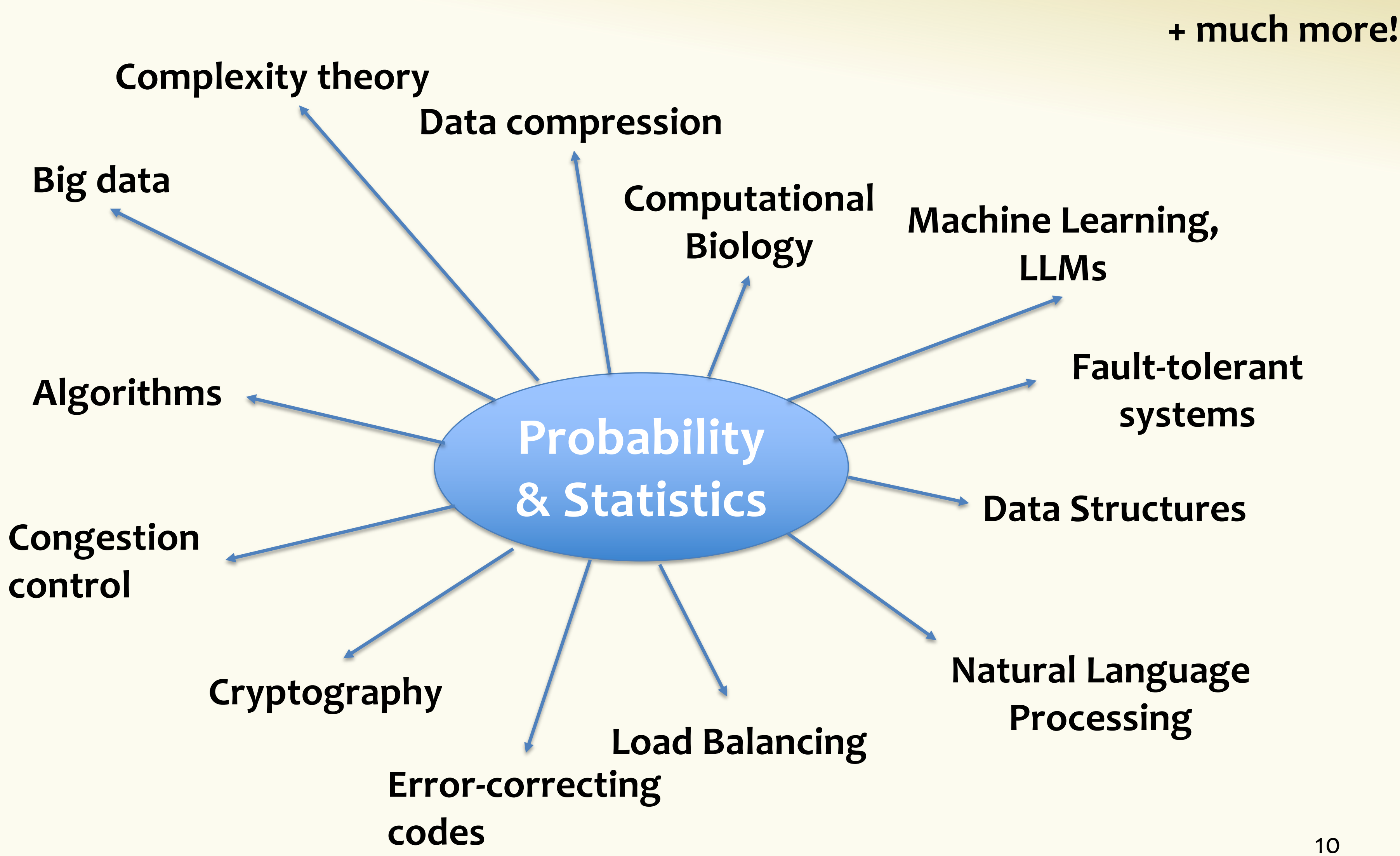
=

Introduction to Probability & Statistics

for computer scientists

What is probability??

Why probability?!


+ much more!
Complexity theory
Data compression
Big data
Computational Biology
Machine Learning, LLMs
Algorithms
Probability & Statistics
Fault-tolerant systems
Data Structures
Congestion control
Cryptography
Natural Language Processing
Load Balancing
Error-correcting codes

Content

- Counting (basis of discrete probability)
  - Counting, Permutation, Combination, inclusion-exclusion, Pigeonhole Principle
- What is probability
  - Probability space, events, basic properties of probabilities, conditional probability, independence, expectation, variance
- Properties of probability
  - Various inequalities, Zoo of discrete random variables, Concentration, Tail bounds
- Continuous Probability & Statistics
  - Probability Density Functions, Cumulative Density Functions, Uniform, Exponential, Normal distributions, Central Limit Theorem, Estimation
- Applications
  - A sample of randomized algorithms, maching learning, differential privacy, …

CSE 312

Foundations of Computing II

Lecture 1: Counting

PAUL G. ALLEN SCHOOL OF COMPUTER SCIENCE & ENGINEERING

Today: A fast introduction to counting, so you will have enough to work on in section tomorrow.

We are interested in counting the number of objects with a certain given property.

"How many ways are there to assign 7 TAs to 5 sections, such that each section is assigned to two TAs, and no TA is assigned to more than two sections?"

"How many integer solutions $(x, y, z) \in \mathbb{Z}^3$ *does the equation* $x^3 + y^3 = z^3$ *have?"*

Generally: Question boils down to computing cardinality $|S|$ of some given set S.

(Discrete) Probability and Counting are Twin Brothers

"What is the probability that a random student from CSE312 has black hair?"

$$= \frac{\#\ students\ with\ black\ hair}{\#students}$$

Today – Two basic rules

- Sum rule
- Product rule

Sum Rule

If you can choose from

- Either one of n options,
- OR one of m options with NO overlap with the previous n,

then the number of possible outcomes is

$$n + m$$

Counting "lunches"

If your starter can be **either** one soup (6 choices) **or** one salad (8 choices), how many possible starters?

Product Rule: If each outcome is constructed by a sequential process where there are

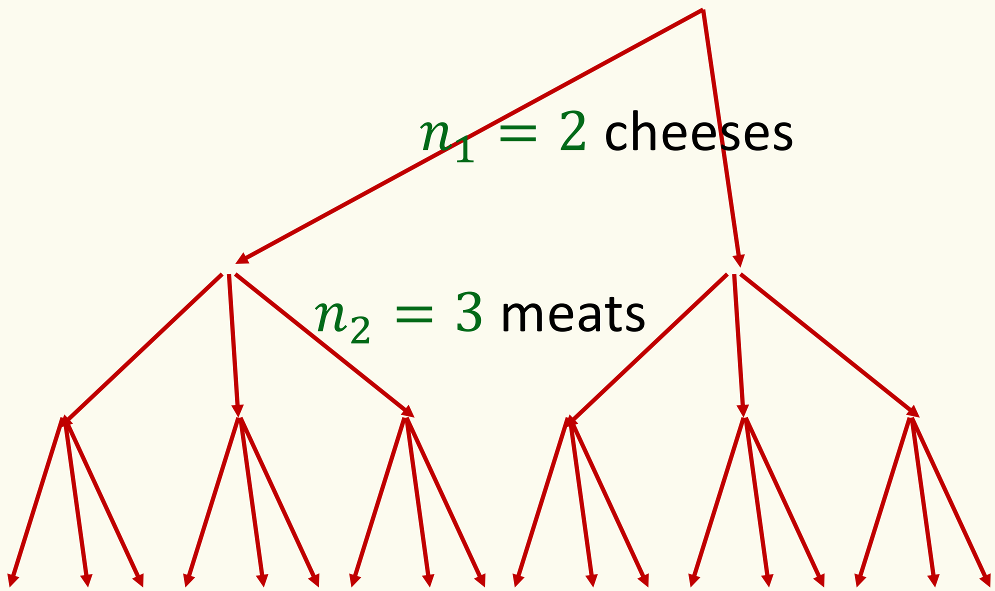
- n_1 choices for the first step,
- n_2 choices for the second step (given the first choice), … , and
- n_k choices for the k^{th} step (given the previous choices),

then the total number of outcomes is $n_1 \times n_2 \times \cdots \times n_k$

Product Rule: In a sequential process, if there are

- n_1 choices for the first step,
- n_2 choices for the second step (given the first choice), … , and
- n_k choices for the k^{th} step (given the previous choices),

then the total number of outcomes is $n_1 \times n_2 \times \cdots \times n_k$



Example: "How many subways?"

$\square \times \square \times \square = \square$

$n_3 = 3$ veggies

Example – Strings

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

- E.g., AZURE, BINGO, TANGO, STEVE, SARAH, …

$$\square \times \square \times \square \times \square \times \square = \square$$

Example – Strings

How many binary strings of length n over the alphabet $\{0,1\}$?

- E.g., $0 \cdots 0,\ 1 \cdots 1,\ 0 \cdots 01$, …

$$\square \times \square \times \square \times \square \times \square = \square$$

Example – Power set

Definition. The **power set** of S is

$$2^S \stackrel{\text{def}}{=} \{X : X \subseteq S\}$$

Example.

$S = \{\bigstar, \spadesuit\}$ $\quad 2^{\{\bigstar, \spadesuit\}} = \{\emptyset, \{\bigstar\}, \{\spadesuit\}, \{\bigstar, \spadesuit\}\}$

$S = \emptyset \quad 2^{\emptyset} = \{\emptyset\}$

...

How many different subsets of S are there if $|S| = n$?

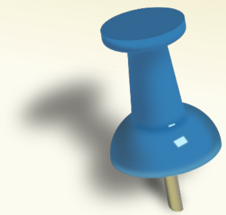
Example – Power set – number of subsets of S

$$S = \{e_1, e_2, e_3, \cdots, e_n\}$$

What is the number of subsets of S, i.e., $|2^S|$?

$\square \times \square \times \square \times \ldots \times \square = \square$

Example – ATMs and Pin codes



- How many 4 –digit pin codes are there?
- Each digit one of {0, 1, 2,… , 9}

☐ x ☐ x ☐ x ☐ = ☐

# possible **first** digits	# possible **second** digits	# possible **third** digits	# possible **fourth** digits	# possible **pins**

Example – ATMs and Pin codes – Stronger Pins

- How many **10–digit** pin codes are there with no repeating digit?
- Each digit one of {0, 1, 2,... , 9}; must use each digit ***exactly once***

☐ x ☐ x ☐ x ... x ☐ x ☐ = ☐

# possible **first** digits	# possible **second** digits	# possible **third** digits	...			# possible **pins**

Permutations

"How many ways to order n distinct objects?"

Answer = $n\times(n-1)\times(n-2)\times\cdots\times2\times1$

Definition. The **factorial function** is

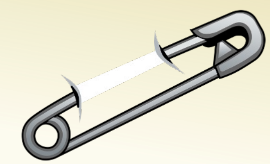
$$n! = n\times(n-1)\times\cdots\times2\times1$$

Read as ``n factorial"

Note: $0! = 1$

Huge: Grows exponentially in n

Example – ATMs and Pin codes – Tricky Pins

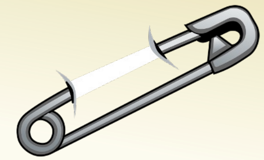


- How many 10–digit pin codes with **at least one digit repeated once**?
- Examples: 1111111111, 1234567889, 1353483595

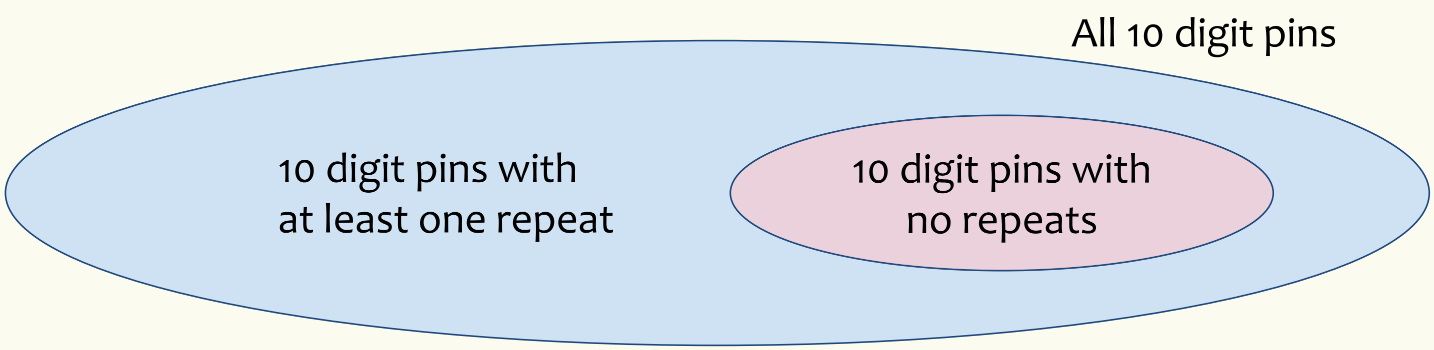
[ ] x [ ] x [ ] x … x [ ] x [ ] = [ ]

possible **first** digits | # possible **second** digits | # possible **third** digits | … | # possible **pins**

Example – ATMs and Pin codes – Tricky Pins



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



Complementary Counting

Let U be a set and S a subset of interest

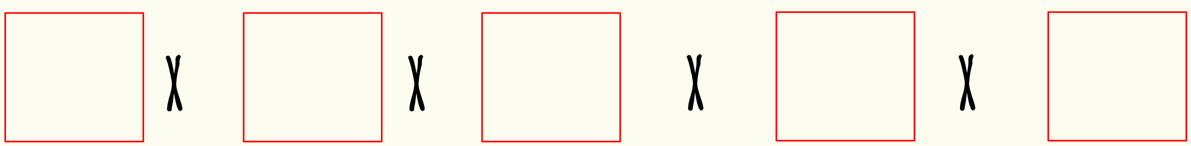
Let U \S denote the set difference (the part of U that is not in S)

Then **|U\S| = |U| - |S|**

Distinct Letters

"How many sequences of 5 *distinct alphabet letters from* $\{A, B, \ldots, Z\}$*?"*

E.g., AZURE, BINGO, TANGO. But not: STEVE, SARAH



Distinct Letters

"How many sequences of 5 *distinct alphabet letters from* $\{A, B, \ldots, Z\}$*?"*

E.g., AZURE, BINGO, TANGO. But not: STEVE, SARAH

Answer: $26 \times 25 \times 24 \times 23 \times 22 = 7893600$

In general

Aka: k-permutations

Fact. # of ways to arrange k out of n distinct objects in a sequence.

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

We sometimes say ``n pick k''

Number of Subsets

*"How many size-5 **subsets** of $\{A, B, \ldots, Z\}$?"*

E.g., {A,Z,U,R,E}, {B,I,N,G,O}, {T,A,N,G,O}. But not: {S,T,E,V}, {S,A,R,H},…

Number of Subsets

"How many size-5 ***subsets*** *of* $\{A, B, \dots, Z\}$*?"*

E.g., {A,Z,U,R,E}, {B,I,N,G,O}, {T,A,N,G,O}. But not: {S,T,E,V}, {S,A,R,H},…

Different from k-permutations: **NO ORDER**

Different sequences: TANGO, OGNAT, ATNGO, NATGO, ONATG …

Same set: {T,A,N,G,O}, {O,G,N,A,T}, {A,T,N,G,O}, {N,A,T,G,O}, {O,N,A,T,G}… …

How to count number of 5 element subsets of $\{A, B, \dots, Z\}$?

Consider the following process:

1. Choose an **unordered** subset $S \subseteq \{A, B, \dots, Z\}$ of size $|S| = 5$
   e.g. $S = \{A, G, N, O, T\}$
2. Choose a permutation of letters in S
   e.g., *TANGO, AGNOT, NAGOT, GOTAN, GOATN, NGOAT, …*

Outcome: An **ordered** sequence of 5 distinct letters from $\{A, B, \dots, Z\}$

x =

Number of Subsets – Idea for how to count

Consider the following process:

1. Choose an **unordered** subset $S \subseteq \{A, B, \dots, Z\}$ of size $|S| = 5$.   e.g. $S = \{A, G, N, O, T\}$
1. Choose a permutation of letters in S
   e.g., *TANGO, AGNOT, NAGOT, GOTAN, GOATN, NGOAT, …*

Outcome: An **ordered** sequence of 5 distinct letters from $\{A, B, \dots, Z\}$

$$??? =$$

? ? ?

×

=

Number of Subsets – Idea for how to count

Consider the following process:

1. Choose an **unordered** subset $S \subseteq \{A, B, \dots, Z\}$ of size $|S| = 5.$ e.g. $S = \{A, G, N, O, T\}$

1. Choose a permutation of letters in S
   e.g., *TANGO, AGNOT, NAGOT, GOTAN, GOATN, NGOAT, …*

$$???\ \times\ 5!\ =\ \frac{26!}{21!}$$

Outcome: An **ordered** sequence of 5 distinct letters from $\{A, B, \dots, Z\}$

$$??? = \frac{26!}{21!\,5!} = 65780$$

Number of Subsets -- don't care about order

Fact. The number of subsets of size k of a set of size n is

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

we say "n choose k"

[also called **combinations** or **binomial coefficients**]

Quick Summary

- Sum Rule

  If you can choose from

  – Either one of n options,

  – OR one of m options with NO overlap with the previous n,

  then the number of possible outcomes of the experiment is $n + m$

- Product Rule

  In a sequential process, if there are

  - n_1 choices for the first step,
  - n_2 choices for the second step (given the first choice), …, and
  - n_k choices for the k^{th} step (given the previous choices),

  then the total number of outcomes is $n_1 \times n_2 \times \cdots \times n_k$

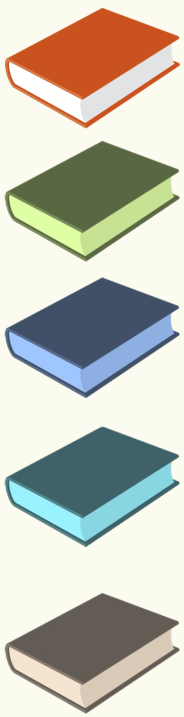
- Complementary Counting

Quick Summary

- K-sequences: How many length k sequences over alphabet of size n? repetition allowed.
  - Product rule ➔ n^K
- K-permutations: How many length k sequences over alphabet of size n, without repetition?
  - Permutation ➔ $\frac{n!}{(n-k)!}$
- K-combinations: How many size k subsets of a set of size n (without repetition and without order)?
  - Combination ➔ $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

Product rule – Another example

5 books



> ***"How many ways are there to distribute 5 books among Alice, Bob, and Charlie?"***

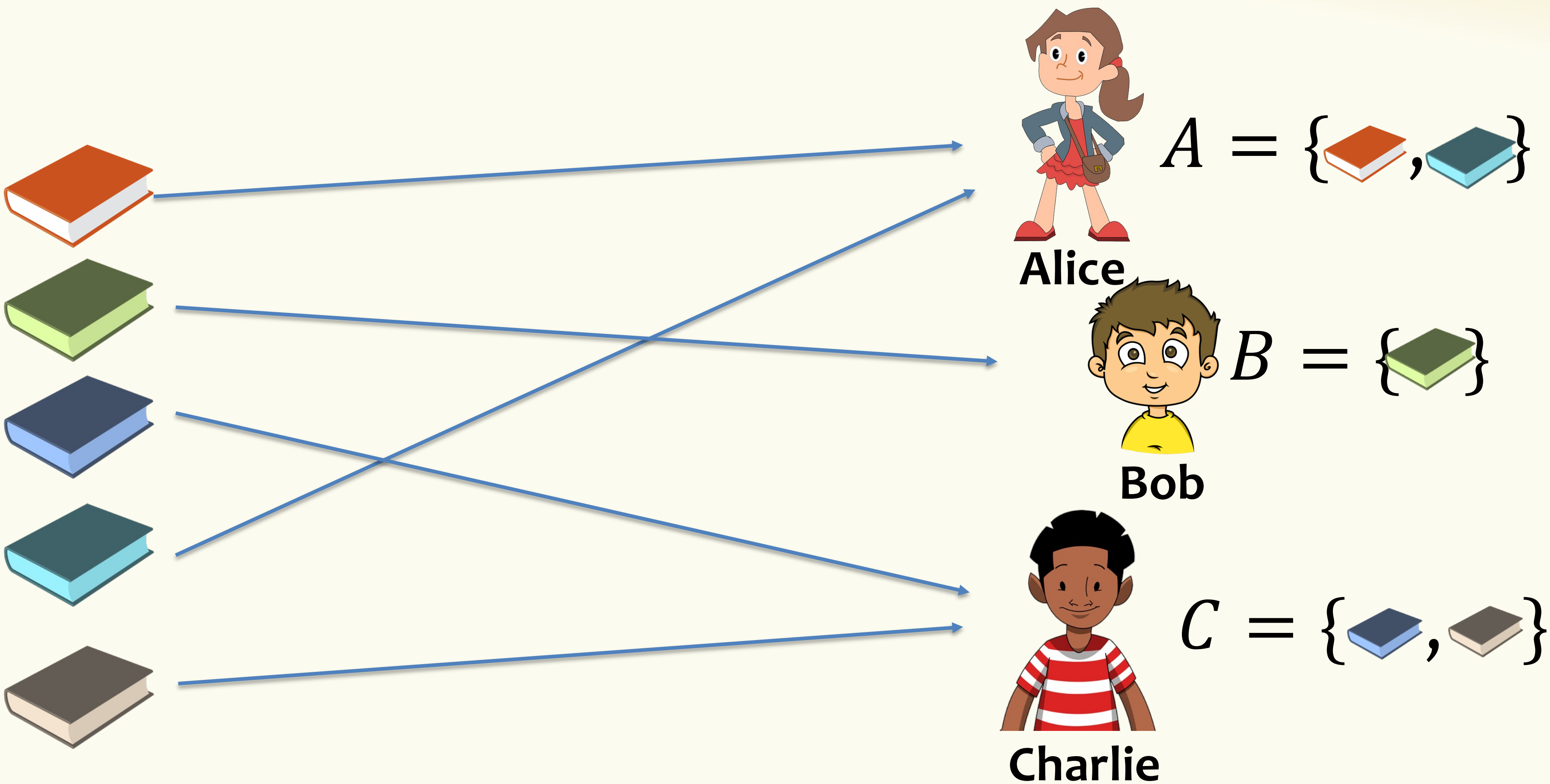
Every book to one person, everyone gets ≥ 0 books.

Alice

Bob

Charlie

Example Book Assignment



$A = \{$, $\}$

$B = \{$ $\}$

$C = \{$, $\}$

Book assignment

$2^5 = 32$ options

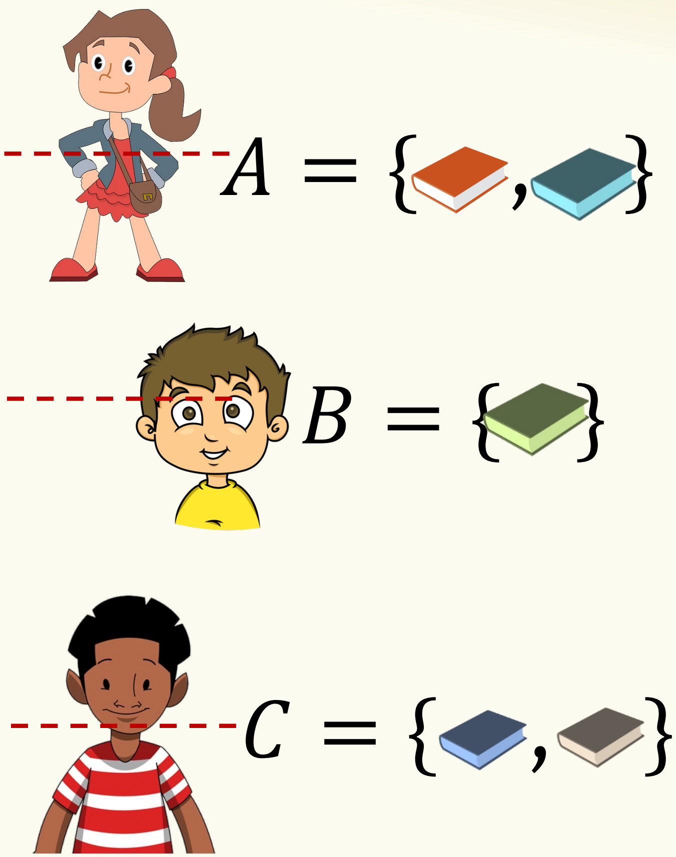
×

$2^5 = 32$ options

×

$2^5 = 32$ options

= 32^3 assignment



Book assignment – Modeling

Correct?

Poll:

A. right

B. Overcount

C. Undercount

D. No idea

https://pollev.com/annakarlin185

$2^5 = 32$ options

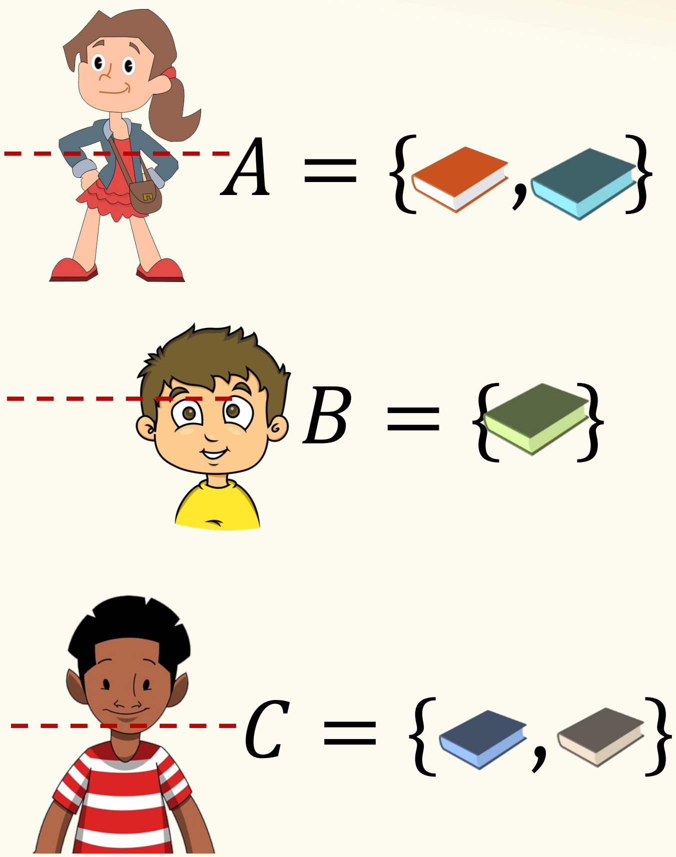
×

$2^5 = 32$ options

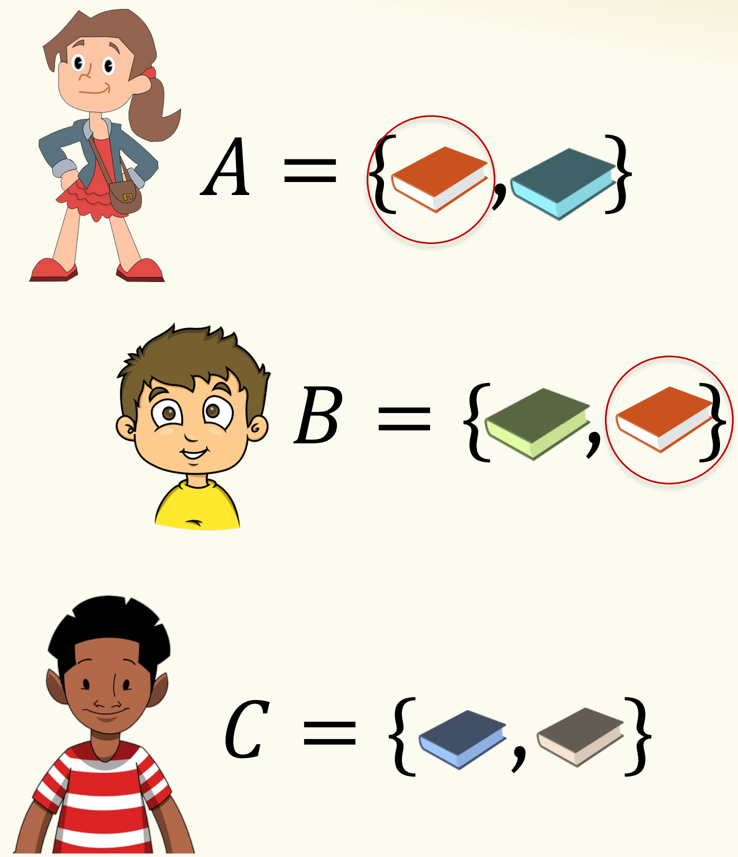
×

$2^5 = 32$ options

= 32^3 assignment



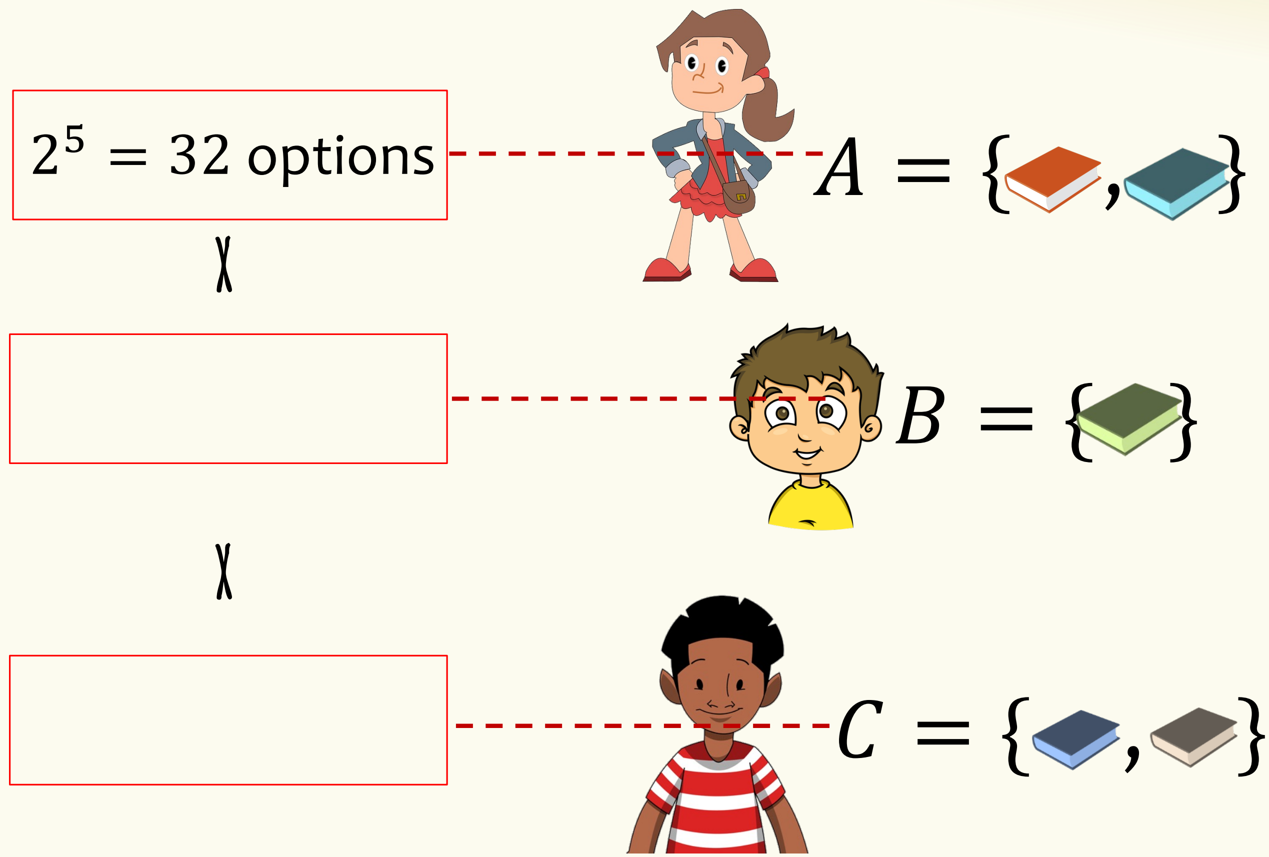
Problem – Overcounting



Problem: We are counting some <u>invalid</u> assignments!!!
→ **overcounting**!

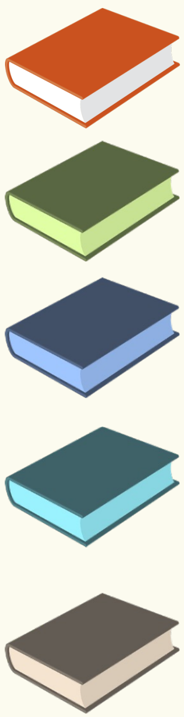
What went wrong in the sequential process?
After assigning set A to Alice, set B is no longer a valid option for Bob

Book assignment – Second try



Product rule – A better way

5 books



> ***"How many ways are there to distribute 5 books among Alice, Bob, and Charlie?"***

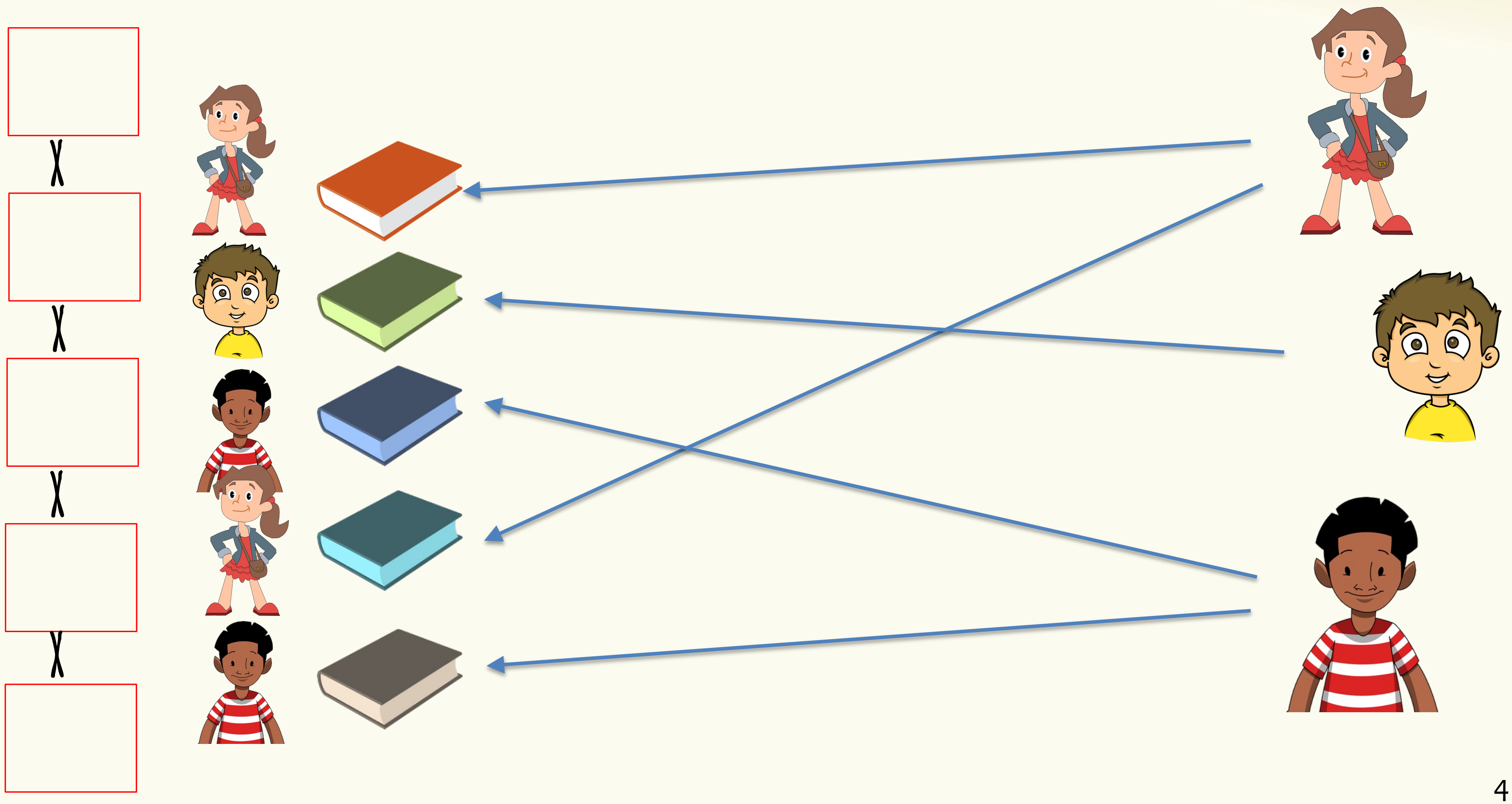
Every book to one person, everyone gets ≥ 0 books.

Alice

Bob

Charlie

Book assignments – Choices tell you who gets each book



Lesson: Representation of what we are counting is very important!

Think about the various possible ways you could make a sequence of choices that leads to an outcome in the set of outcomes you are trying to count.

The first concept check (CC) will be out by 3pm and is due 1:00pm Friday

The concept checks are meant to help you immediately reinforce what is learned in each lecture. (Today's CC also reviews summation and product notation.)

Students from previous quarters found them really useful!