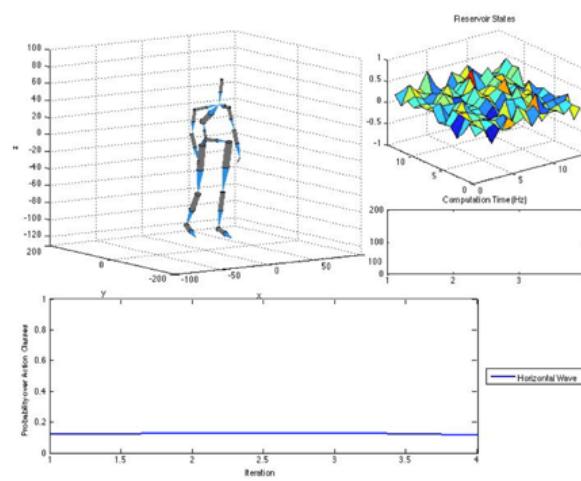
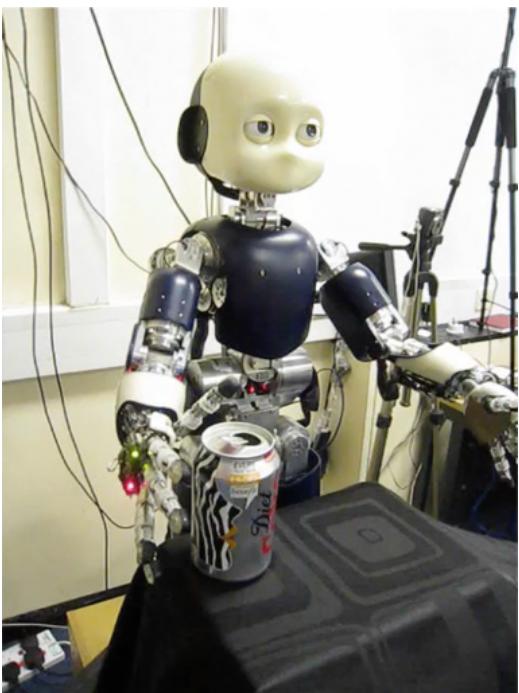


CS2040S: DATA STRUCTURES & ALGORITHMS

Harold Soh
harold@comp.nus.edu.sg



SELF-INTRODUCTION



Harold Soh
Assistant Professor
Dept of Computer Science

Research Interests:
Human-Centered Artificial Intelligence, Machine Learning, Human-Robot Interaction

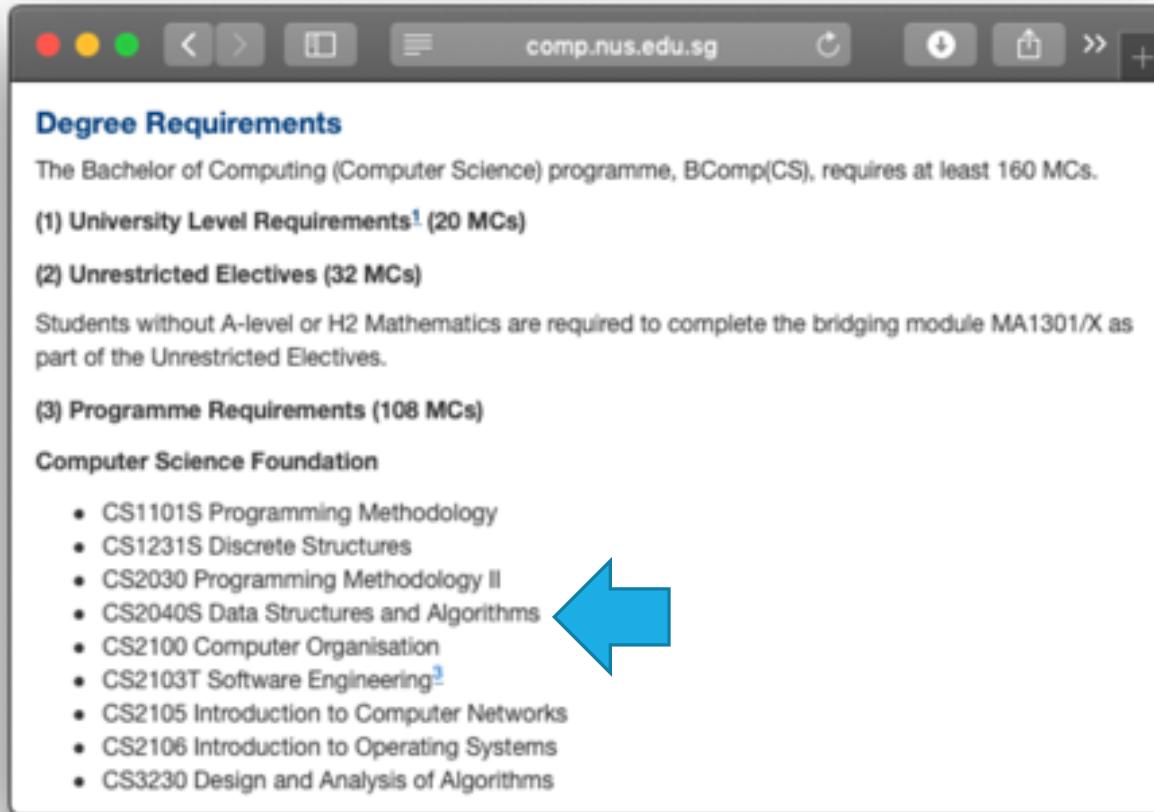
Office: COM2-03-03

Office Hours: By appointment

WHY ARE YOU HERE?



IT'S A CORE MODULE...



The Bachelor of Computing (Computer Science) programme, BComp(CS), requires at least 160 MCs.

(1) University Level Requirements¹ (20 MCs)

(2) Unrestricted Electives (32 MCs)

Students without A-level or H2 Mathematics are required to complete the bridging module MA1301/X as part of the Unrestricted Electives.

(3) Programme Requirements (108 MCs)

Computer Science Foundation

- CS1101S Programming Methodology
- CS1231S Discrete Structures
- CS2030 Programming Methodology II
- CS2040S Data Structures and Algorithms
- CS2100 Computer Organisation
- CS2103T Software Engineering³
- CS2105 Introduction to Computer Networks
- CS2106 Introduction to Operating Systems
- CS3230 Design and Analysis of Algorithms



BUILD USEFUL SYSTEMS FOR PEOPLE



START A TECH COMPANY



CREATE LEARNING MACHINES



COMPUTER SCIENTISTS HAVE CHANGED THE WORLD...



...BY SOLVING PROBLEMS

CS2040S is about **solving problems**.

It prepares you for what lies ahead...

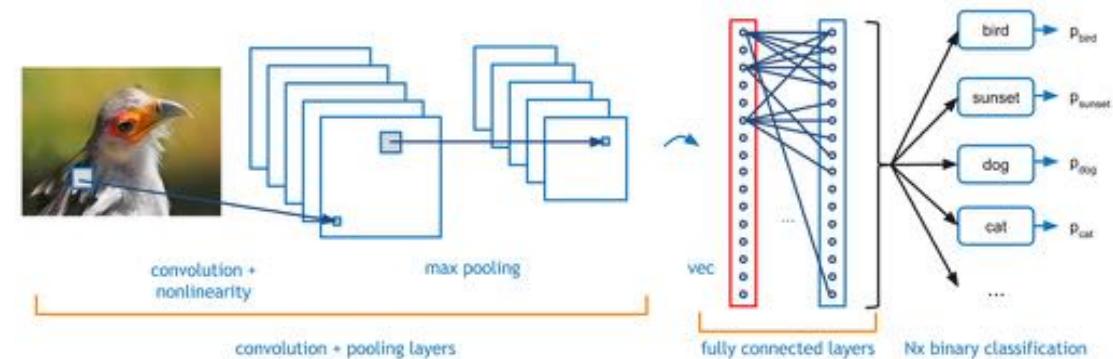
Search



Google Search

I'm Feeling Lucky

Image Recognition





Poll Everywhere

<https://bit.ly/2LvG9bq>



POLL-EVERYWHERE

Via a browser:

[Pollev.com/haroldsohsoo986](https://pollev.com/haroldsohsoo986)

Or SMS:

Text a code to +65 8241 0042

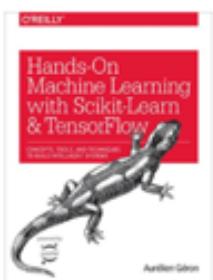
I'm excited about CS2040S:

- A. Yes!
- B. No! ☹
- C. Not sure
- D. What is CS2040S?

The video instructions on using polleverywhere is incorrect for the first lecture
(Fatir was not able to log into my account)
Fatir will provide updated instructions in class.
We will use my polleverywhere account for the 3rd lecture onwards.

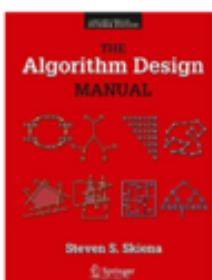
A PROBLEM - RECOMMENDER SYSTEM

Recommended for You in Computers & Technology



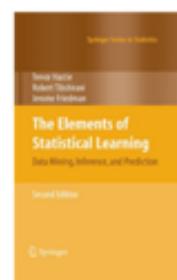
Hands-On Machine Learning with...
Aurélien Géron
★★★★★ 49
Kindle Edition
\$24.99

[Buy now with 1-Click](#)



The Algorithm Design Manual
Steven S Skiena
★★★★★ 135
Kindle Edition
\$53.17

[Buy now with 1-Click](#)



The Elements of Statistical Learning:
Trevor Hastie
★★★★★ 129
Kindle Edition
\$56.67

[Buy now with 1-Click](#)

The screenshot shows a Mac OS X desktop with a Netflix window open. The window title is "netflix.com" and the page title is "Netflix". The main content area displays a "Recently Watched" section with a thumbnail for "HOUSE" featuring Hugh Laurie. Below it is a "Popular on Netflix" section with thumbnails for "Breaking Bad", "ROBOCOP", and "ALWAYS SUNNY IN PHILADELPHIA". At the bottom left is a "Watched by Pennsatucky" section with thumbnails for "jesus camp", "Saved!", and "For The BIBLE Tells Me So". A "Recommend" button is visible at the bottom center. The top right of the window has "Browse" and "Taste Profile" buttons.

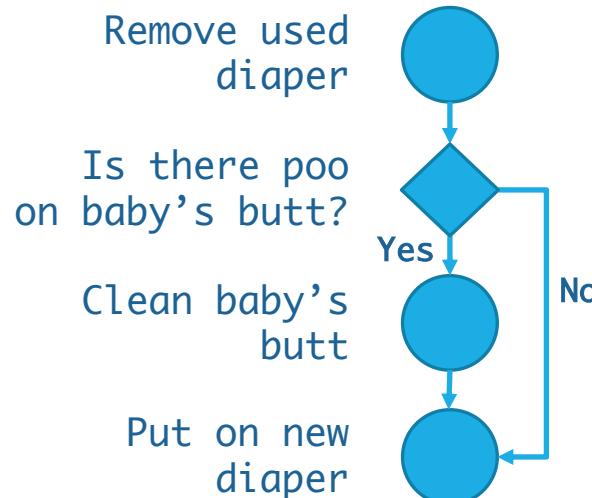
A PROBLEM - RECOMMENDER SYSTEM

Given a set of N items $i = 1, 2, \dots, N$ and
a score s_i for each item i where $s_i \in [0,1]$
find the top- k items with highest scores.

INPUT		OUTPUT	
item	score	item	score
1	0.2	5	0.9
2	0.5	3	0.6
3	0.6	2	0.5
4	0.1		
5	0.9		

DATA STRUCTURES & ALGORITHMS

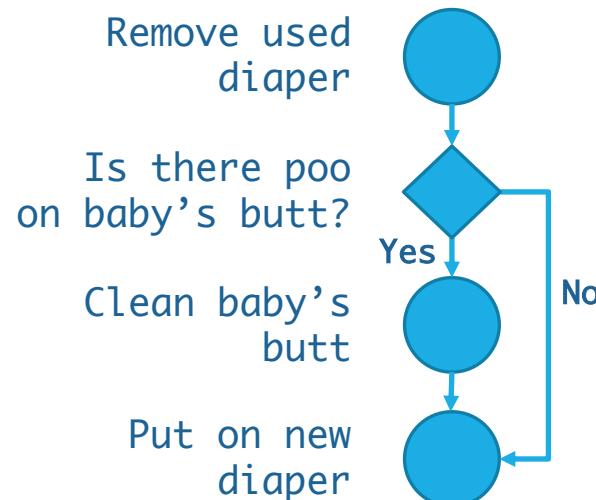
A “**Recipe**”:
List of instructions for
solving a problem



DATA STRUCTURES & ALGORITHMS

Can be expressed in many ways:

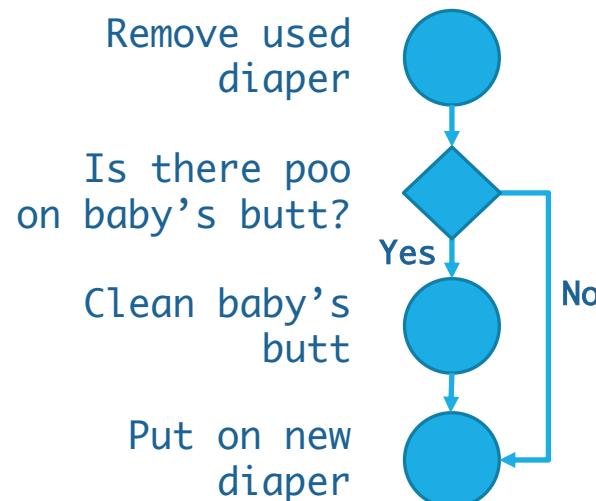
- Pseudocode
- Programming Language (e.g. Java)
- Flowchart
- etc.



DATA STRUCTURES & ALGORITHMS

Can be expressed in many ways:

- Pseudocode
- Programming Language (e.g. Java)
- Flowchart
- etc.



```
butt.removeDiaper()  
if butt.hasPoo() is True  
    butt.clean()  
butt.addDiaper()
```

BACK TO OUR PROBLEM - RECOMMENDER SYSTEM

Given a set of N items $i = 1, 2, \dots, N$ and
a score s_i for each item i where $s_i \in [0,1]$
find the top- k items with highest scores.

INPUT		$N = 5$
item	score	
1	0.2	
2	0.5	
3	0.6	
4	0.1	
5	0.9	

\longrightarrow

$k = 3$

OUTPUT	
item	score
5	0.9
3	0.6
2	0.5

FIND AN ALGORITHM FOR THIS PROBLEM



Given a set of N items $i = 1, 2, \dots, N$ and
a score s_i for each item i where $s_i \in [0,1]$
find the top- k items with highest scores.

INPUT $N = 5$

item	score
1	0.2
2	0.5
3	0.6
4	0.1
5	0.9

$k = 3$

OUTPUT

item	score
5	0.9
3	0.6
2	0.5



WHICH ALGORITHM RUNS FASTER?

Algorithm A:

```
items.sort(descending)  
items.printFirst(k)
```

Algorithm B:

```
for i = 1 to k  
    e = items.findMaxElement()  
    items.print(e)  
    items.remove(e)
```

Which algorithm is faster for large N and k ?

- A. Algorithm A
- B. Algorithm B
- C. Both have the same performance.
- D. It depends.

DATA STRUCTURES & ALGORITHMS

Algorithm A:

```
items.sort(descending)  
items.printFirst(k)
```

Algorithm B:

```
for i = 1 to k  
    e = items.findMaxElement()  
    items.print(e)  
    items.remove(e)
```

What is items?

- how fast can we sort items?
- how fast can we print the first k elements?
- how fast can we find the maximum element?
- how fast can we remove an item?

ALGORITHM A WITH AN **ARRAY**

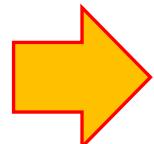
Algorithm A:

```
items.sort(descending)  
items.printFirst(k)
```

If items is an unsorted Array N elements, how many operations are needed to sort items using a Bubble Sort?

$$\sim N^2$$

Week 3!



using a Merge Sort:

$$\sim N \log N$$

ALGORITHM A WITH AN **ARRAY**

Algorithm A:

```
items.sort(descending)  
items.printFirst(k)
```

If **items** is an sorted Array
N elements, how many
operations are needed to
print the first **k** elements?

$\sim k$

ALGORITHM A WITH AN **ARRAY**

Algorithm A:

```
items.sort(descending)  
items.printFirst(k)
```

Total Num. Operations:
 $\sim N \log N + k$

ALGORITHM B WITH AN **ARRAY**

Algorithm B:

```
for i = 1 to k  
    e = items.findMaxElement()  
    items.print(e)  
    items.remove(e)
```

If `items` is an unsorted Array N elements, how many operations are needed to find the maximum element?

$\sim N$

then print and remove it?

~ 2

ALGORITHMS A & B USING ARRAYS

Algorithm A:

```
items.sort(descending)  
items.printFirst(k)
```

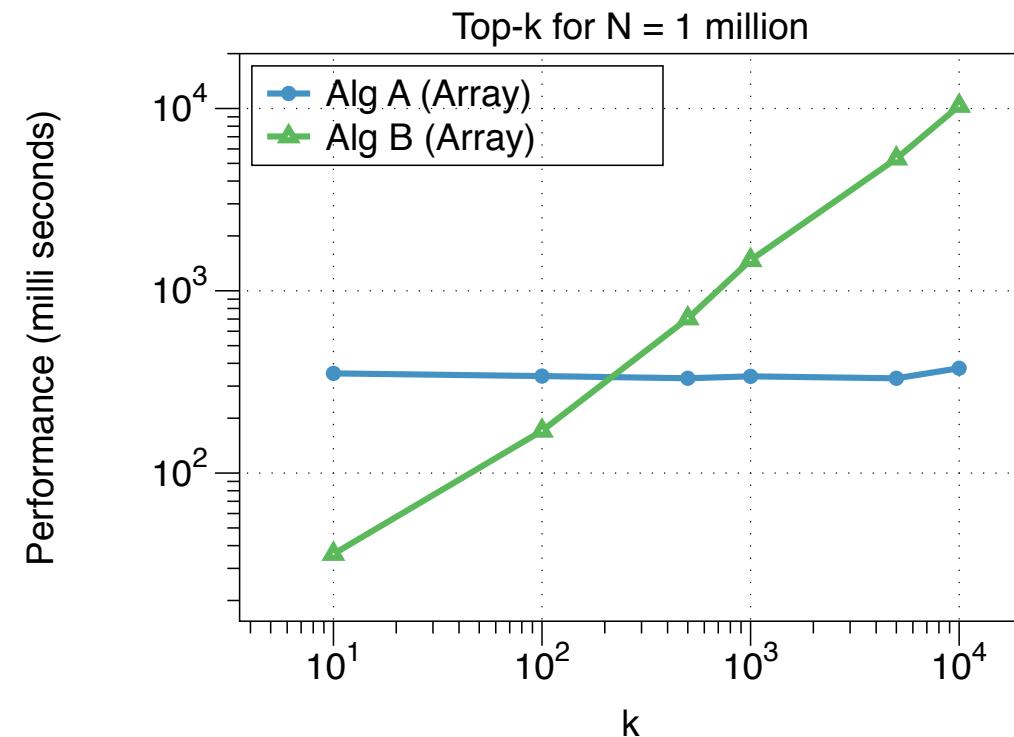
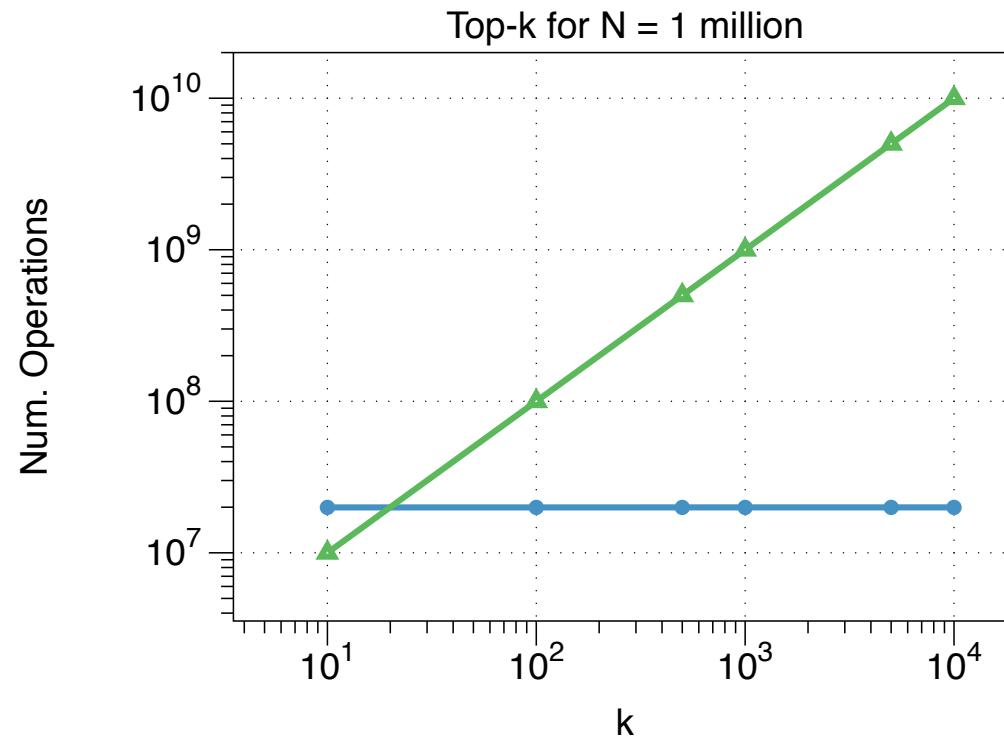
Algorithm B:

```
for i = 1 to k  
    e = items.findMaxElement()  
    items.print(e)  
    items.remove(e)
```

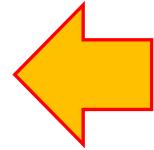
Total Num. Operations:
 $\sim N \log N + k$

Total Num. Operations:
 $\sim k(N + 2) = kN + 2k$

THEORETICAL & REAL-WORLD PERFORMANCE



ALGORITHM B WITH AN MAX HEAP



Week 4!

If items is a Max Heap data structure, how many operations are needed to find the maximum element?

Algorithm B:

```
for i = 1 to k  
    e = items.findMaxElement()  
    items.print(e)  
    items.remove(e)
```

1

then print and remove it?

$1 + \log N$

ALGORITHMS A (ARRAY) & B (MAX HEAP)

Algorithm A:

```
items.sort(descending)  
items.printFirst(k)
```

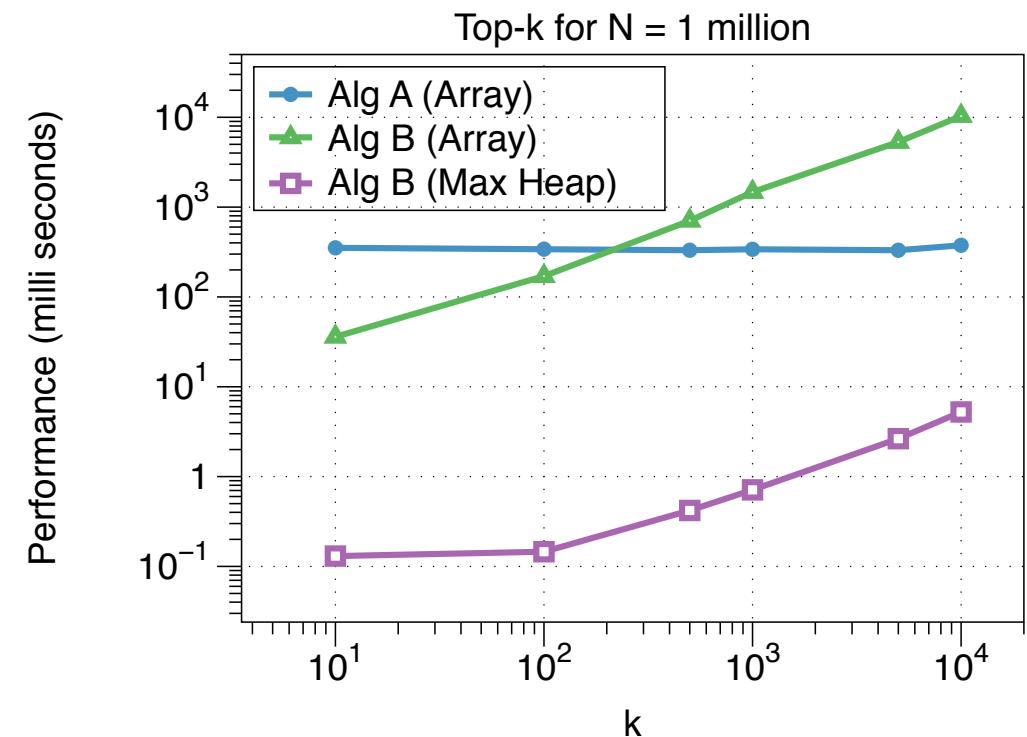
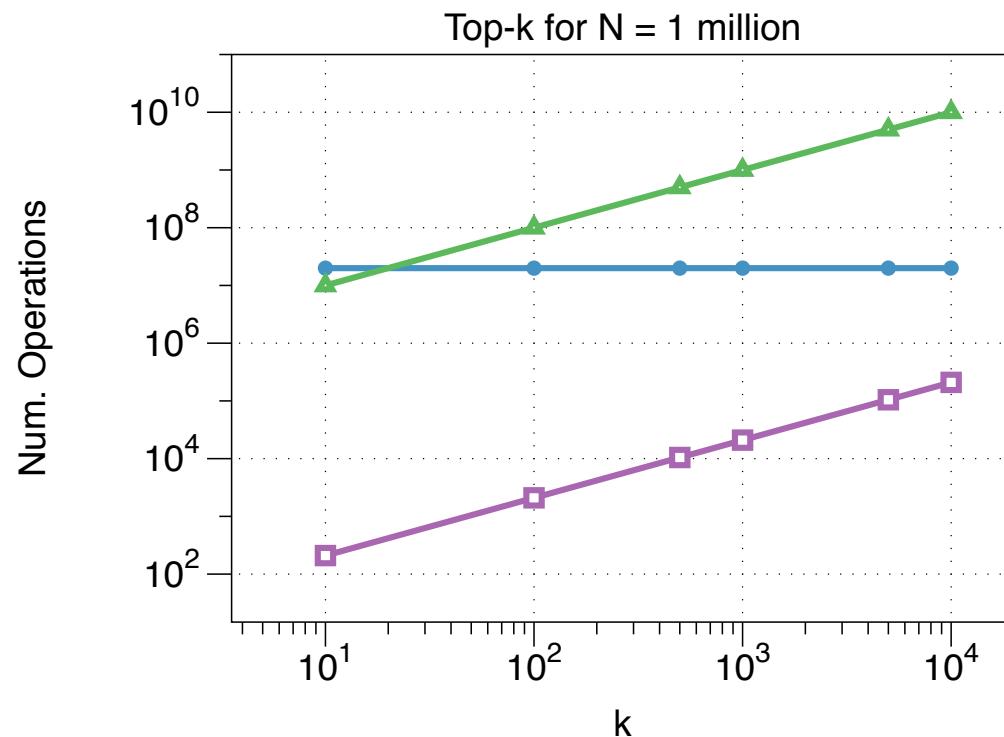
Total Num. Operations:
 $N \log N + k$

Algorithm B:

```
for i = 1 to k  
    e = items.findMaxElement()  
    items.print(e)  
    items.remove(e)
```

Total Num. Operations:
 $k(2 + \log N) = k \log N + 2k$

THEORETICAL & REAL-WORLD PERFORMANCE



"Bad programmers worry about the code. Good programmers worry about *data structures and their relationships*."

- Linus Torvalds*



*Creator of git and the linux kernel (in Linux OS, Chrome OS, Android)

FRAMEWORK: ALGORITHM & DATA STRUCTURE

What problem does it solve?

How does it work?

How to implement it?

What is its asymptotic performance?

What is its real world performance?

What are the trade-offs?



LEARNING OUTCOMES

By the end of this course, you should be able to:

- **Apply algorithmic thinking and techniques** for solving computational problems.
- **Describe the structure and operation** of different **data structures and algorithms** under the standard computational model.
- **Assess the suitability** of different data-structures and algorithms for a specific computational problem.
- **Adapt existing data-structures and algorithms** to solve specific computational problems.

QUESTIONS?



5 MINUTE BREAK





QUESTION: TOP-K ITEMS

How many operations do we need to find the top-k items in unsorted order?

- A. $N^2 + k$
- B. $N \log N + k$
- C. $kN + k \log k$
- D. $N + k \log N$
- E. $N + k \log k$
- F. N

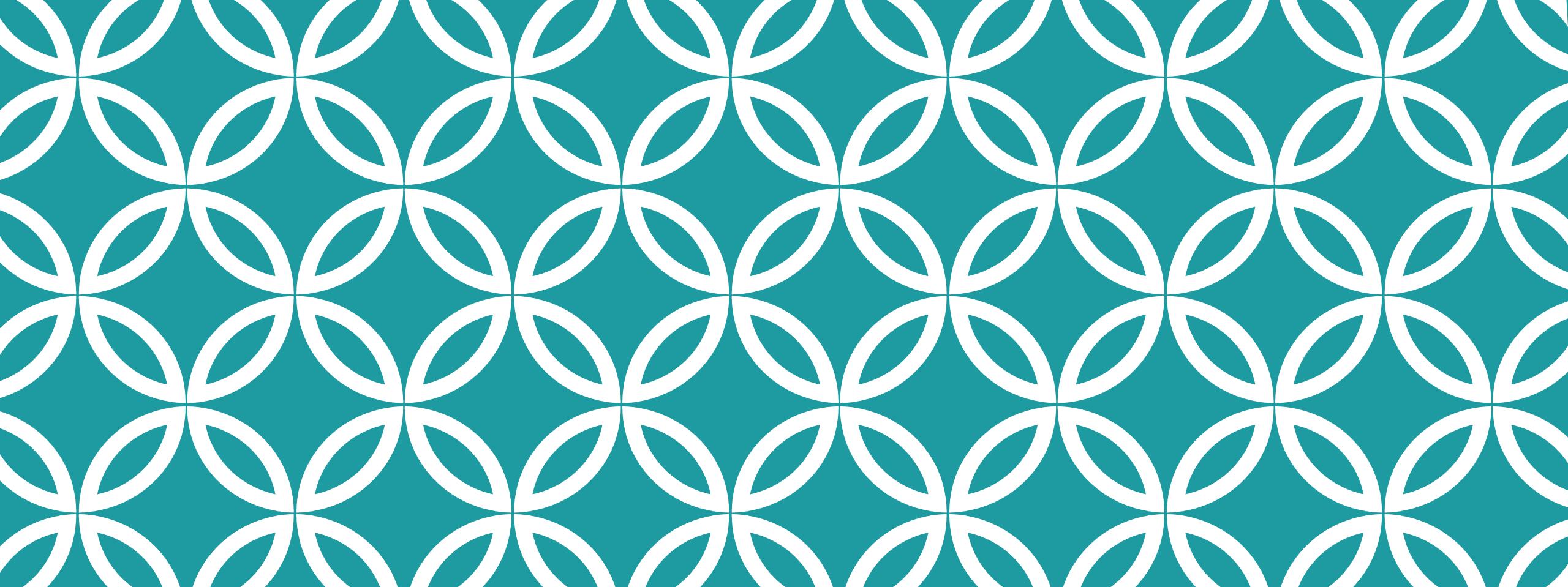
INPUT $N = 5$

item	score
1	0.2
2	0.5
3	0.6
4	0.1
5	0.9

OUTPUT

item	score
2	0.5
5	0.9
3	0.6

$k = 3$



PART 2: ADMINISTRATIVE ISSUES

Pre-reqs, Grading, Schedule,
Quizzes, Exams, Textbooks, etc.

COURSE PREREQUISITES

You should know Java or take CS2030 in parallel

- Our focus is not to teach Java
- Java Quickstart (tomorrow!)

CS1010 or its equivalent

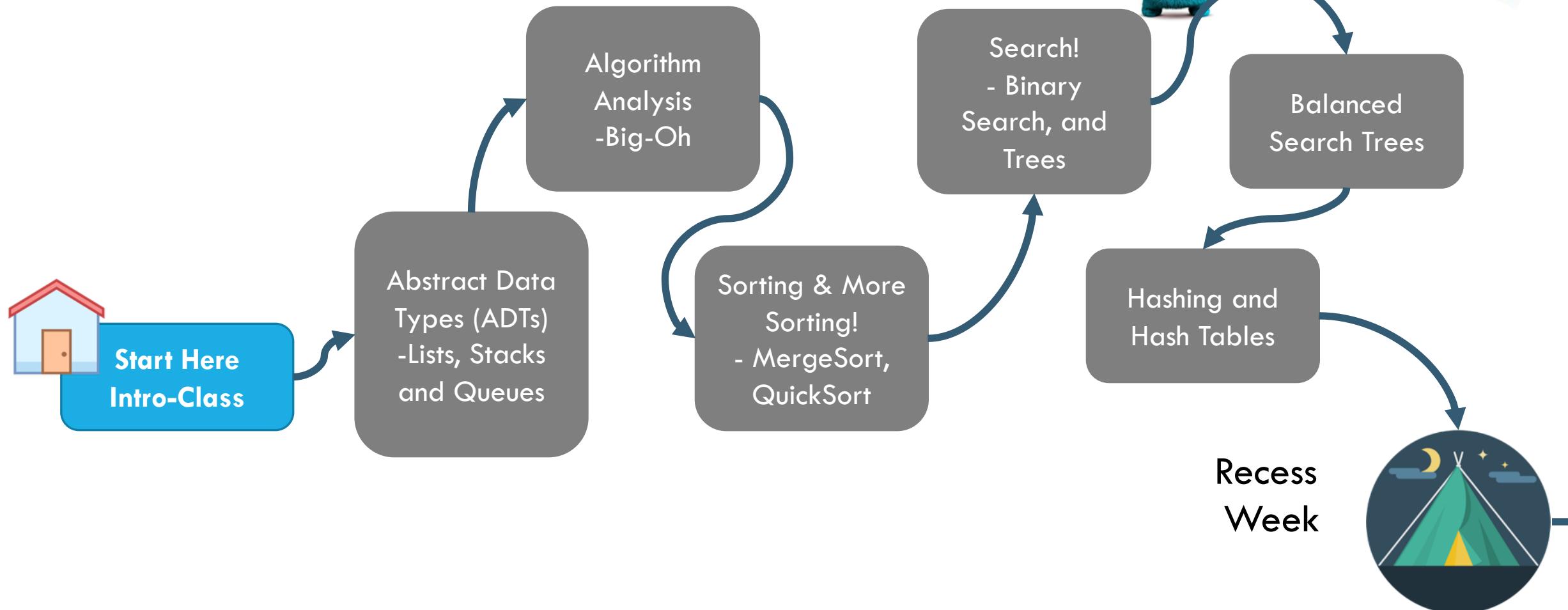
Preclusions: CS1020, CS1020E, CS2020, CS2010, CS2040

- If you've already taken these, it's time to pick another class.



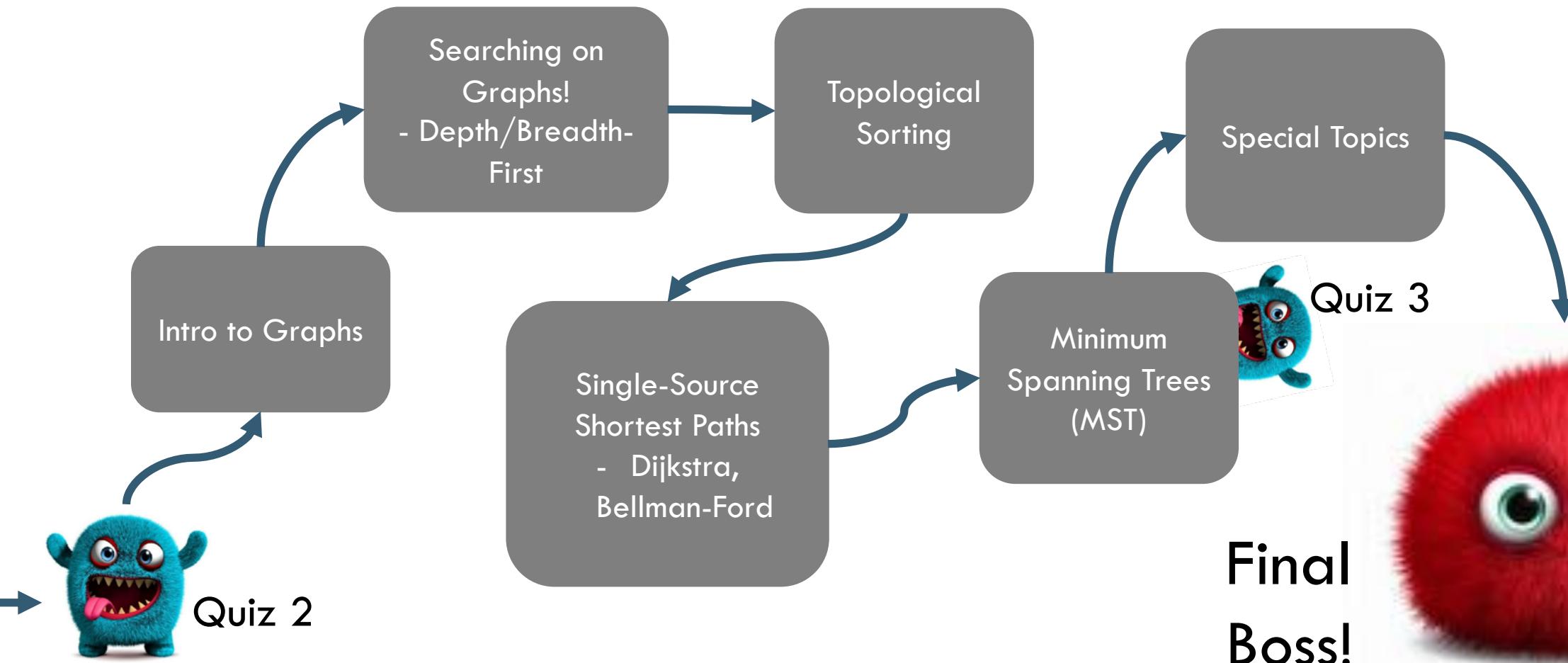
PATH TO MASTERY / COURSE STRUCTURE

DRAFT



PATH TO MASTERY / COURSE STRUCTURE

DRAFT



WEEKLY SCHEDULE

2x Lectures @ LT19

- Tuesday: 12pm-1:30pm
- Wednesday: 12pm-1:30pm

1x Tutorials

- 10 slots (Fridays)
- Starts in Week of 26th August

1x Discussion Groups (Labs)

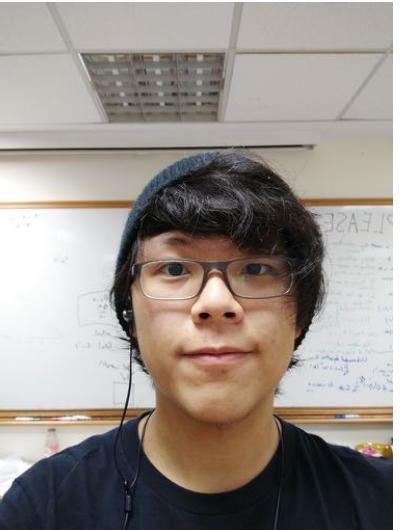
- 9 slots (Thursdays)
- Starts in Week of 26th August

SIGN UP!

ModReg

YOUR GUIDES



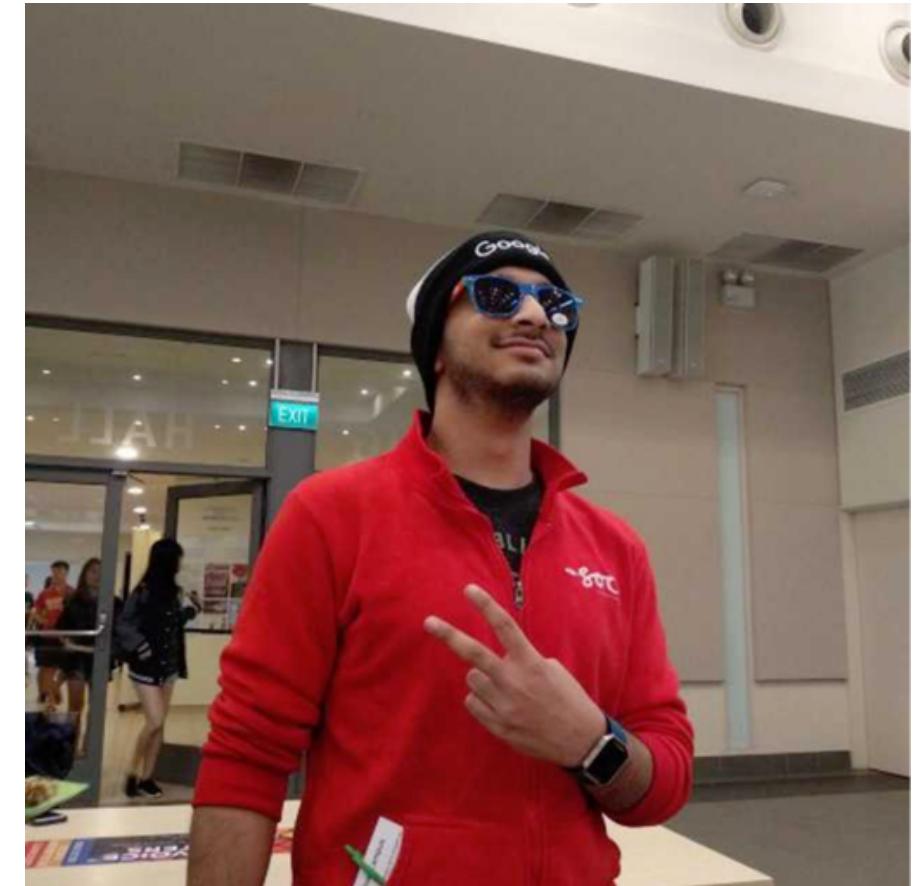


TUTORS

Ayush Chatteraj

Personal Quote:

“Nobody:
Me: Yes, I'm a CS major, why do you ask?”



TUTORS

Abdul Fatir

Personal Quote:

"I waste way too much time watching cat videos while pretending to wait for my model to train."

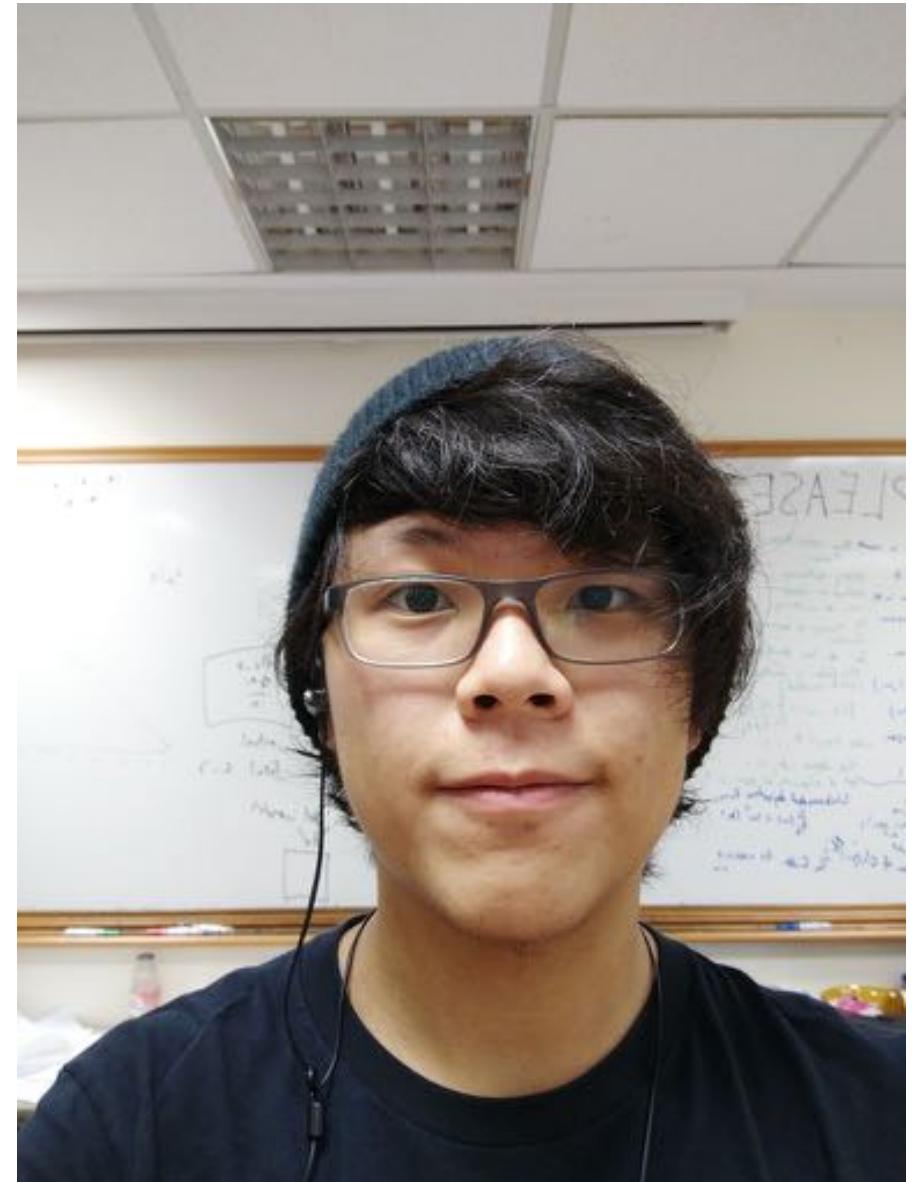


TUTORS

Eldon Chung

Personal Quote:

"As for the quote, err I'll get back to you"



TUTORS

Enzio Kam

Personal Quote:

"I used to like Markov Chains. Now I like to Cuckoo Search."



TUTORS

Esther Chong

Personal Quote:

*"It is said that there are three kind
of lies: lies, damned lies
and ~~statistics~~ machine learning
models created without validation"*



TUTORS

Irham Rasyidi bin Zainal

Personal Quote:

*“What do you mean I can't pass an
`int` by reference in Java?”*



TUTORS

Govind Venugopalan

Personal Quote:

*“If your code runs slower than me,
you are in trouble.”*



TUTORS

Travis Ching

Personal Quote:

“I don't care if your glass is half empty or full. Is there coffee in it?”



TUTORS

Ryan Chew

Personal Quote:

“as seen on TV!”



TUTORS

Ang Wei Neng

Personal Quote: “ ”

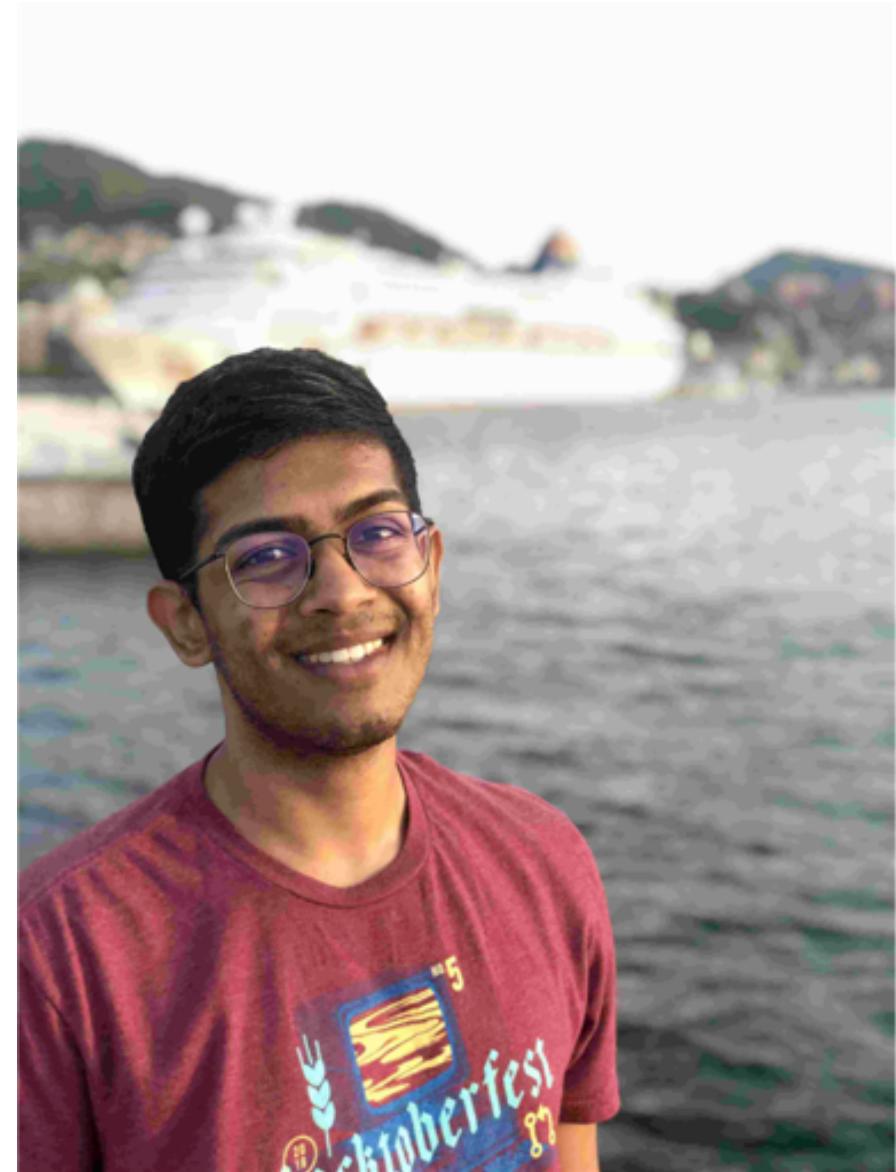


TUTORS

Vignesh Shankar

Personal Quote:

*“Anything can be solved using the
bLocKcHaIn”*

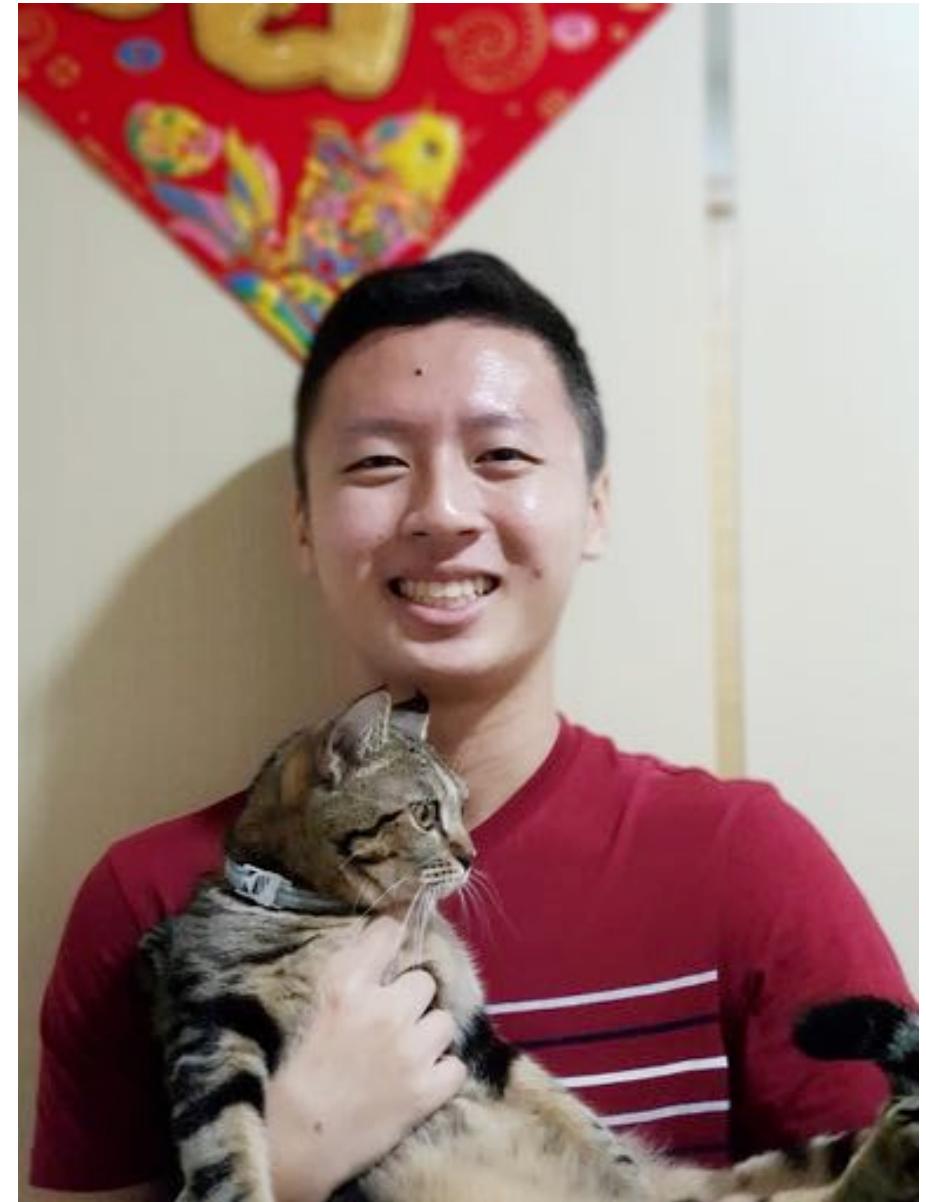


TUTORS

Sim Yu Jie

Personal Quote:

*“I think traits associated with me
are often : Fast, Ambitious,
Talented, Individualistic, and Neat.”*



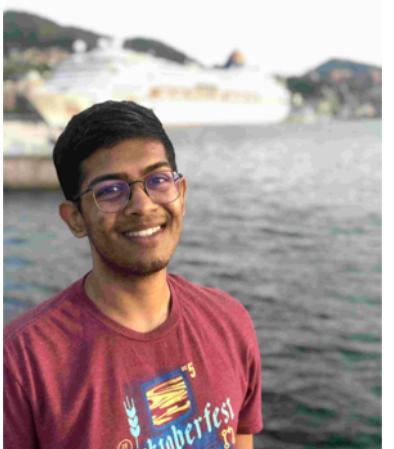
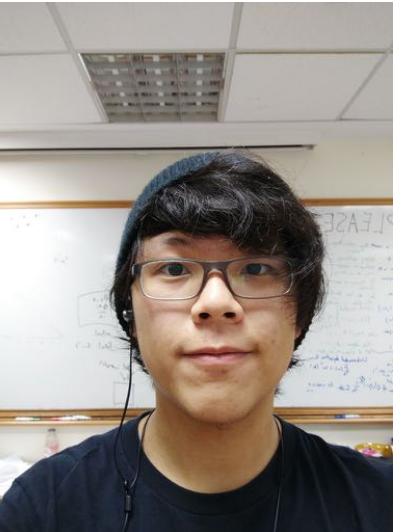
TUTORS

Zhi Jian

Personal Quote:

“Think more, debug less!”





GRADE BREAKDOWN

Component	%
Problem Sets (x5)	30
Quiz (Best 2/3)	20
Participation	10
Final Exam	40

NUS.KATTIS.COM (in progress...)

Access Problem Sets Online

Submit Solutions

- Get (near) instant test feedback (Auto-grading!)

Will provide link

Tomorrow: Java Quickstart + Kattis info



PROBLEM SETS: 30%

Grading based on correctness.

- Grades *not* based on speed ranks.
- Time limit exceeded (TLE) means you have *not* solved the problem

Late Submissions

- 24 hours: 20% penalty
- 2 weeks: 40% penalty
- Last Second of Last Day of class: 60% penalty

Submission and Problems on Kattis



PROBLEM SETS: 30%

Can I collaborate with my friends?

- **Yes! And you should!**

Policy:

- Free to collaboratively discuss ideas and derive solutions.
- Do not be a skiver!
- Must write/code solutions on your own (no copy-pasting nonsense!).
- Don't cheat! Cheating/plagiarism will be dealt with harshly.



WHERE TO ASK QUESTIONS?



Face-to-Face: Lectures / Discussion Groups / Tutorials

Online: Piazza <http://piazza.com/nus.edu.sg/fall2019/cs2040s>

A screenshot of the Piazza website homepage. The page features a dark header with the Piazza logo and navigation links for Product, In Professors' Words, Support, About Us, Companies, Sign Up, and Login. Below the header, a main heading reads "The incredibly easy, completely free Q&A platform" followed by the subtext "Save time and help students learn using the power of community". To the left, there's a small thumbnail image of a Piazza interface showing a video player and some text. To the right, a bulleted list highlights features: "Wiki style format enables collaboration in a single space", "Features LaTeX editor, highlighted syntax and code blocking", "Questions and posts needing immediate action are highlighted", "Instructors endorse answers to keep the class on track", and "Anonymous posting encourages every student to participate".

The incredibly easy, completely free Q&A platform
Save time and help students learn using the power of community

- Wiki style format enables collaboration in a single space
- Features LaTeX editor, highlighted syntax and code blocking
- Questions and posts needing immediate action are highlighted
- Instructors endorse answers to keep the class on track
- Anonymous posting encourages every student to participate

QUIZZES: 20%



Quiz 1 : Week 4 (during Lecture Slot)

Quiz 2 : Week 7 (during Lecture Slot)

Quiz 3 : Week 11 (during Lecture Slot)

All Dates To be Confirmed (TBC)

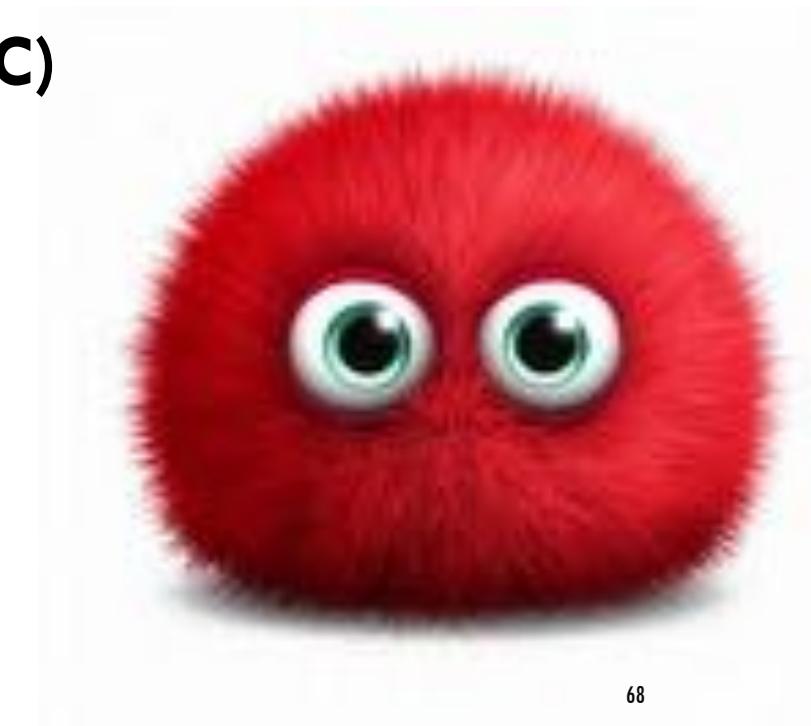
We will take the best 2/3 (10% each)



FINAL EXAM: 40%

4 DEC 2018 (Morning, 9-11 am)

Date To be Confirmed (TBC)



PARTICIPATION: 10%

Discussion Groups

Tutorials

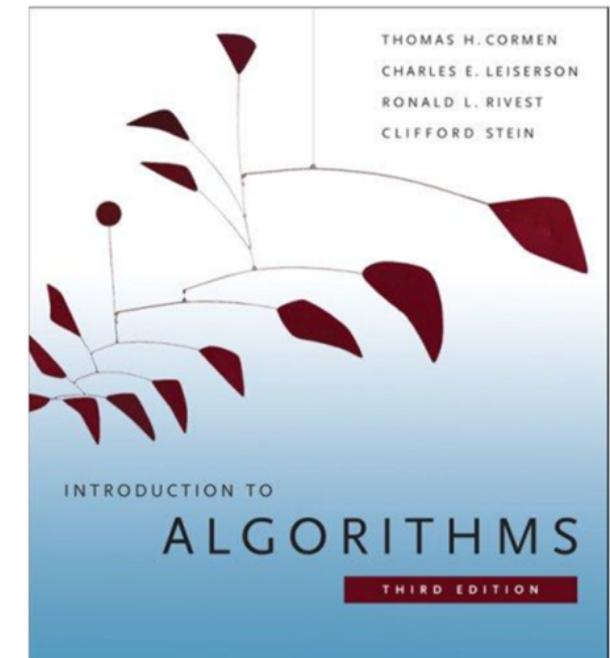
Piazza Forums



TEXTBOOK

No Official Textbook.

- *Algorithms* by Sedgewick and Wayne
- *Introduction to Algorithms* by Cormen, Leiserson, Rivest, Stein



VISUALGO.NET

The screenshot shows the VISUALGO.NET homepage. At the top, there is a navigation bar with links for "Training", "Translation", and "Login". Below the navigation bar, the title "VISUALGO .NET/EN" is displayed, followed by the subtitle "visualising data structures and algorithms through animation". A search bar is located below the title. To the right, there is a logo for "CP3" and a "Buy Now" button for Lulu.

Do You Know? Search the term 'algorithm visualization' in your favorite Search Engine, do you see VisuAlgo in the first page of the search result? Next level: Search that term again, but in your native language (if it is not English). Is VisuAlgo still listed in the first page? And get ready to be surprised: Search the name of your favorite data structure or algorithm without mentioning the keyword 'animation' or 'visualization'. Is VisuAlgo still listed in the first page? :)

Sorting Training

array algorithm bubble select

Bitmask Training

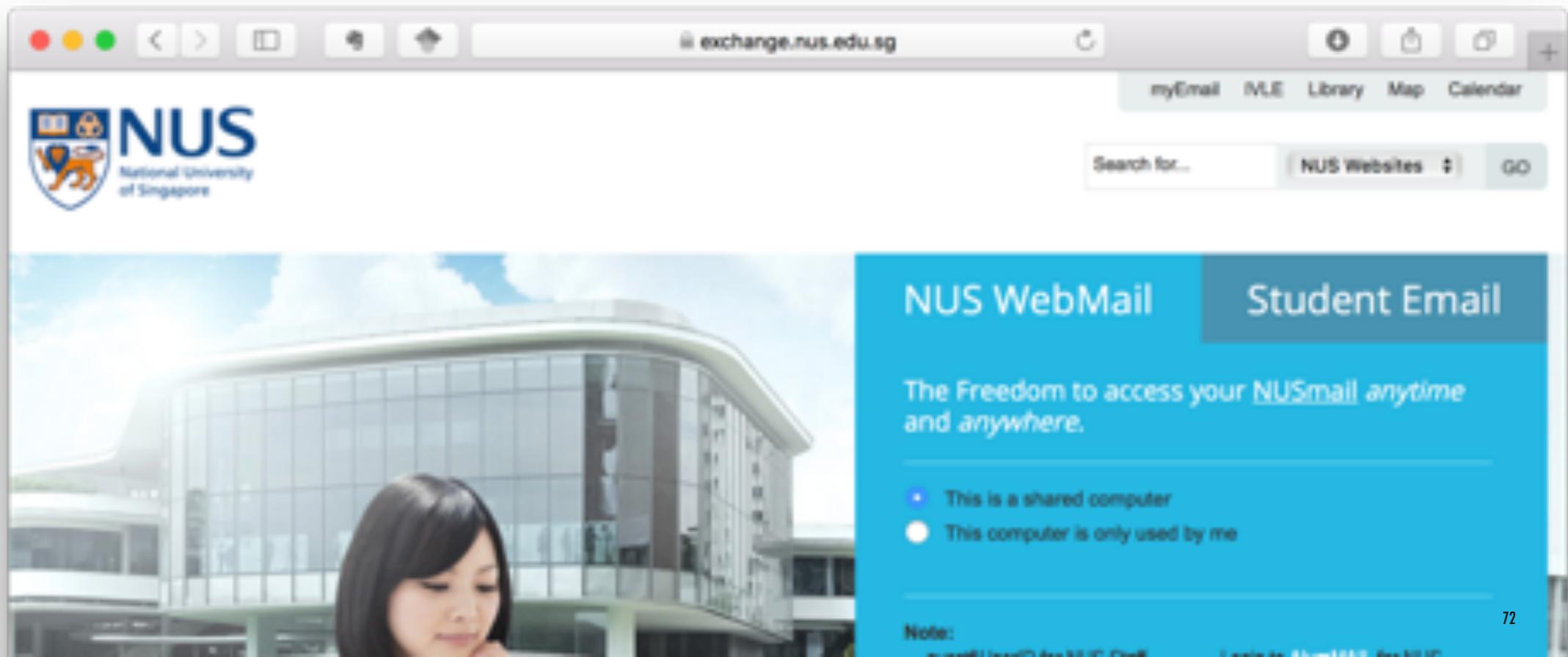
bit manipulation set cs3233

Linked List Training

stack queue doubly deque

ANNOUNCEMENTS ON LUMINUS

And sent to your NUS email (check regularly)!



LUMINUS WILL ALSO HAVE:

Course Description

Weblinks

Lesson Plan

Slides

Grades

and more...

DROP DATES

Refer to Modreg:

<http://www.nus.edu.sg/ModReg/>

Drop with:

- 'W' grade: **26/08/2019 00:00** onwards.
- 'F' grade: **30/09/2019 00:00** onwards.



BEFORE LECTURE TOMORROW

Java Quickstart

Sign up on Kattis:

https://nus.kattis.com/courses/CS2040S/CS2040S_AY_19-20_Sem_1

Bring along your laptop if you want to follow along.

Make sure Java JDK is already installed.

- At least version 8

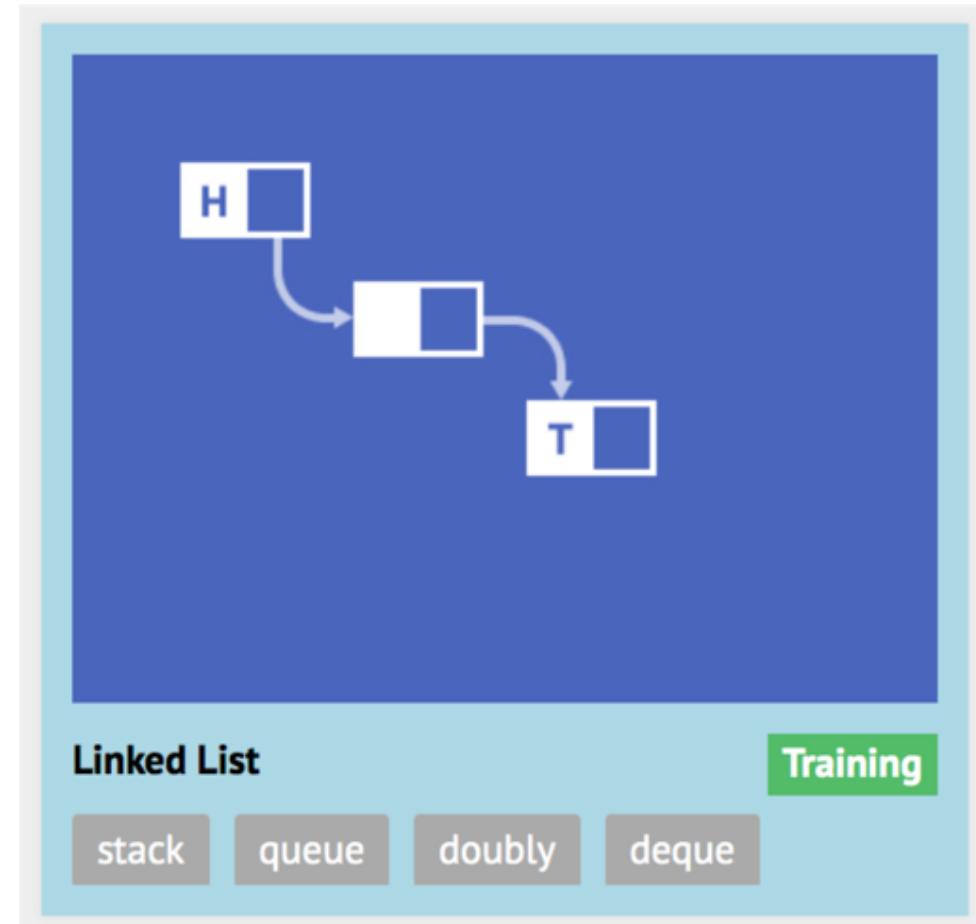
BEFORE LECTURE NEXT WEEK

Go to Visualgo.net and do the
Linked List Module:

<https://visualgo.net/en/ll>

Required: Sections 1 - 5 (Queue)

Optional: 6 onwards.



QUESTIONS?



Email me at harold@comp.nus.edu.sg

BEFORE LECTURE NEXT WEEK

Go to Visualgo.net and do the Linked List Module:

<https://visualgo.net/en/ll>

Required: Sections 1 - 5 (Queue)

Optional: 6 onwards.

My email:

harold@comp.nus.edu.sg

