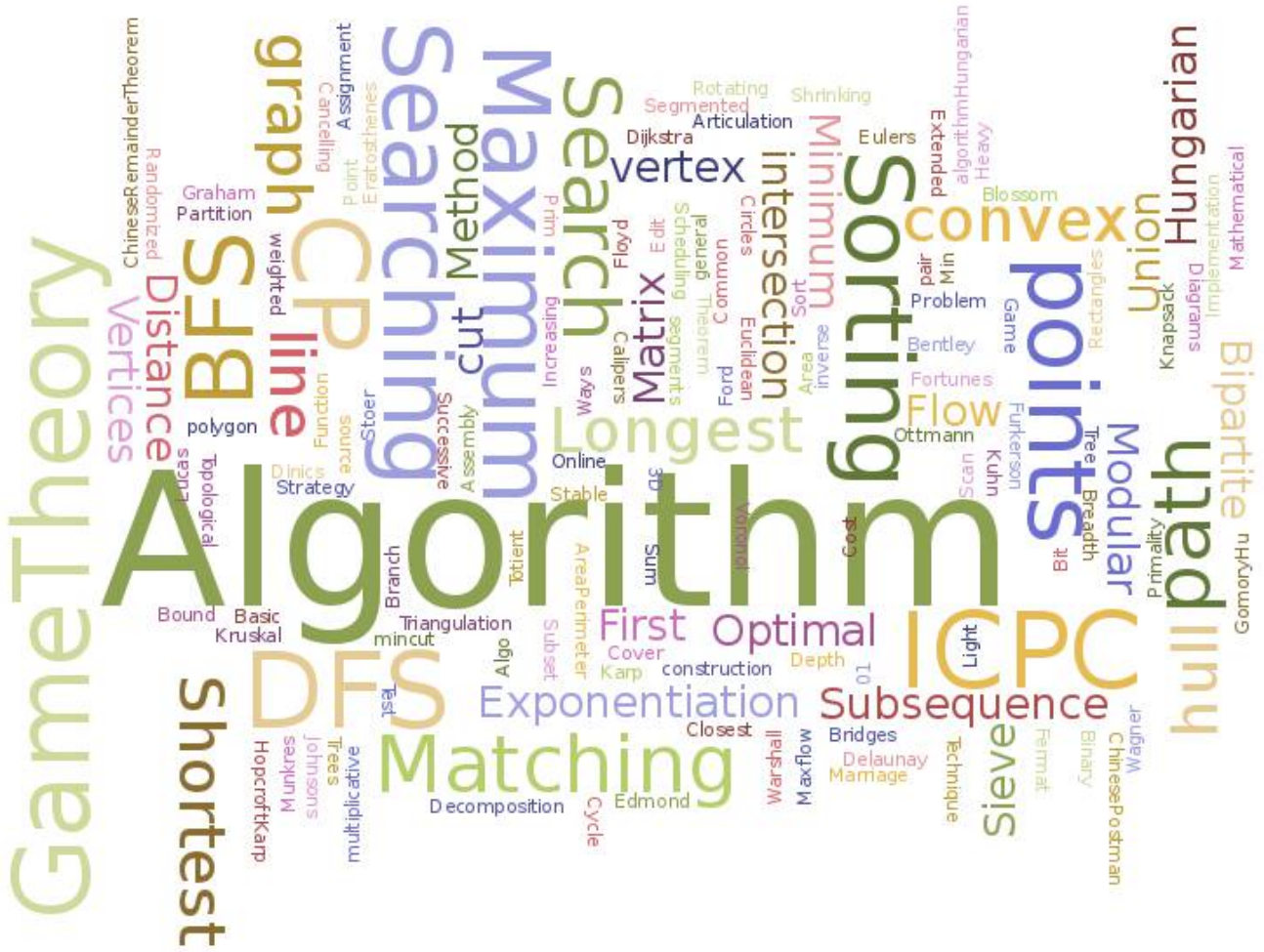
CS2040C Data Structures and Algorithms



Today

- Introduction
- Course Information
- From C to C++

Lecturer

- CHENG, Ho Lun, Alan, 鄭浩璘
  - Hong Kong
  - UIUC, Duke (USA)
- Hobbies:
  - Teaching
  - Research Interests: Graphics, Geometry, Games
- Email:
  - hcheng@comp.nus.edu
  - (alan@comp.nus.edu)
  - DO NOT send to dcschl
  - DO NOT send codes to me
  - Office: AS6 #05-03
- Ext. 68732

Just in case, the less cute one is me

About Me

- *"Those who know the truth are not up to those who love it; those who love the truth are not up to those who delight in it."*

  -- Confucius

- I want to make this course fun and interesting, but still fair, meaningful and with deep knowledge

- To me, it's always a challenge to maintain the balance between fun and fair

- My goal: Minimize your revision time

About Me and this Course

- Some feedback from previous[2] years
  - Make the complex and dry theory into interesting and easy to understand examples
  - “Very funny lecturer. Always explains complex concepts in an easy way for understanding
  - “I can not expect anything more from him. He worked so hard to prepare for the test as well as to help students. The way he explain abstract concepts is just amazing. Furthermore, I can always receive the almost instant replies on the forum from him.“
  - “play nice guitar”

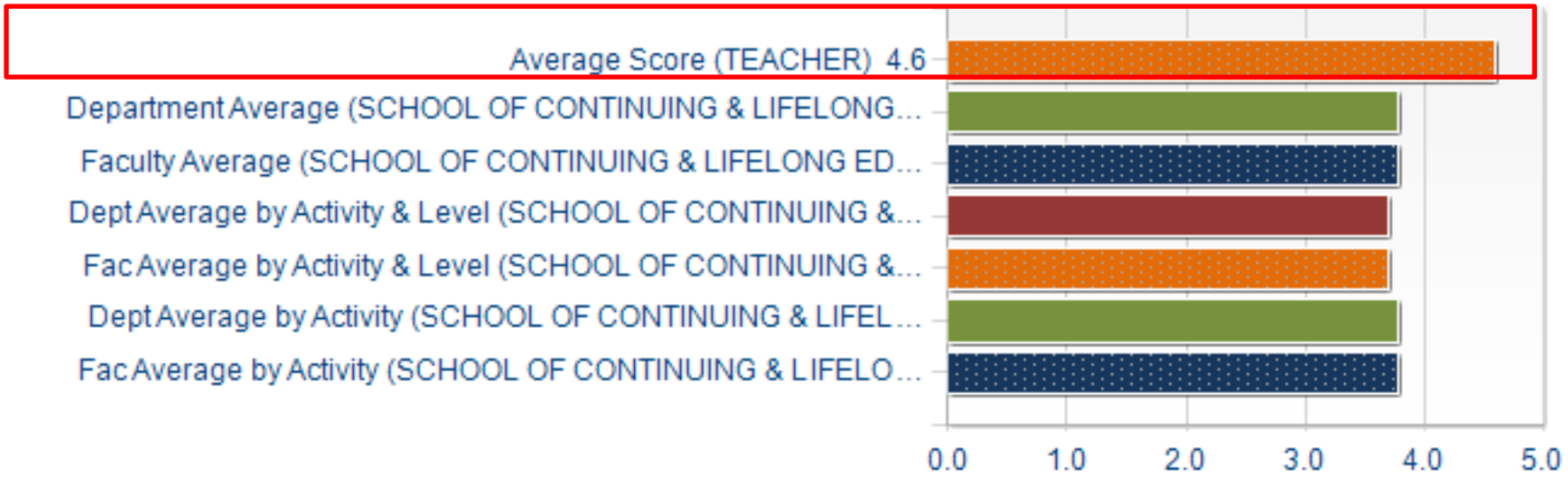
Last Year

Make the complex and dry theory into interesting and easy to understand examples
Weekly quizzes are good to encourage students to study weekly and earn marks
Interesting and attractive teaching style. And gives an easier way of understanding the difficult content. Good arrangement and well balanced between lecture and tutorial. And he will guide TA what to go through in the tutorial. This is the most important thing.
his coaching still very interesting and we are very enjoy in his class. He increased my interest in this module.
Have good sense of humour, lecture is not boring.
Patient and able to articulate well
Provide us very interesting practice questions which require us to apply the knowledge that we have learned from the module. It enhanced our critical thinking and problem solving skills. Through the exercise, we have a better understanding of the knowledge that Mr. Alan Cheng thought us.
Engaging, improvises teaching to accommodate situation.
His funny way of teaching and analogy make us easier to remember and learn new things
his lectures are interesting and able to provide layman examples to relate to the teaching topic which is important as students are new to the topic. especially during tutorial, he recapped and explained further with more examples which enhance the understanding of the topic
Clear concept and written practise to enhance students learning experience beside lecture.
Caring, compassionate, understanding, humorous, understanding, teaches complex concepts using simple examples.
Creatively using images. Use quizzes to encourage students to stay consistent studying. Interactive lessons. Provide lots of resources to suit different type of students study preferences

Last Last Year

Comments
Make the complex and dry theory into interesting and easy to understand examples
Weekly quizzes are good to encourage students to study weekly and earn marks
Interesting and attractive teaching style. And gives an easier way of understanding the difficult content. Good arrangement and well balanced between lecture and tutorial. And he will guide TA what to go through in the tutorial. This is the most important thing.
his coaching still very interesting and we are very enjoy in his class. He increased my interest in this module.
Have good sense of humour, lecture is not boring.
Patient and able to articulate well
Provide us very interesting practice questions which require us to apply the knowledge that we have learned from the module. It enhanced our critical thinking and problem solving skills. Through the exercise, we have a better understanding of the knowledge that Mr. Alan Cheng thought us.
Engaging, improvises teaching to accommodate situation.
His funny way of teaching and analogy make us easier to remember and learn new things
his lectures are interesting and able to provide layman examples to relate to the teaching topic which is important as students are new to the topic. especially during tutorial, he recapped and explained further with more examples which enhance the understanding of the topic
Clear concept and written practise to enhance students learning experience beside lecture.
Caring, compassionate, understanding, humorous, understanding, teaches complex concepts using simple examples.
Creatively using images. Use quizzes to encourage students to stay consistent studying. Interactive lessons. Provide lots of resources to suit different type of students study preferences

Overall, the teacher is effective



Question	Average Score (TEACHER)		Department Average (SCHOOL OF CONTINUING & LIFELONG EDN)		Faculty Average (SCHOOL OF CONTINUING & LIFELONG EDN)	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
The teacher has enhanced my thinking ability.	4.7	0.5	3.9	0.8	3.9	0.8
The teacher provided timely and useful feedback.	4.5	0.6	3.8	0.8	3.8	0.8
The teacher has increased my interest in the subject.	4.6	0.6	3.8	0.9	3.8	0.9
Average of Q1-Q3	4.6	0.6	3.8	-	3.8	-

CS2040C = CS2020 x TIC2001

- CS2020 Data Structures and Algorithms (Accelerated)

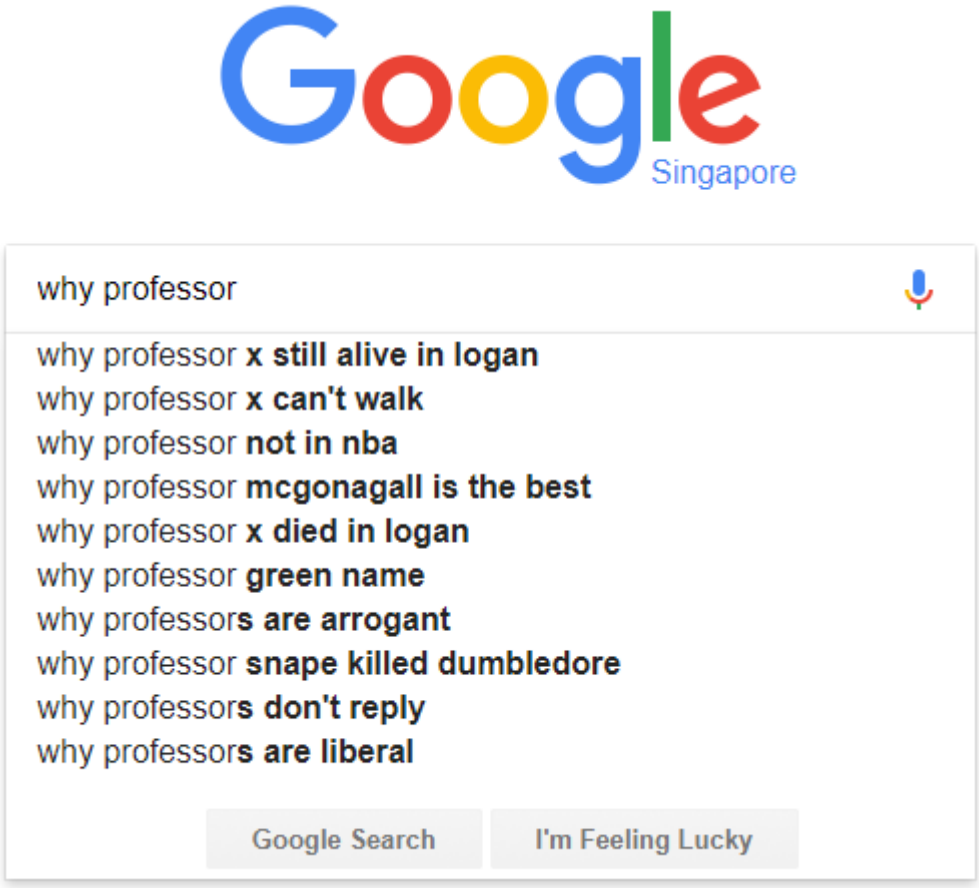
- TIC2001 Data Structures and Algorithms (SCALE)

About This Module

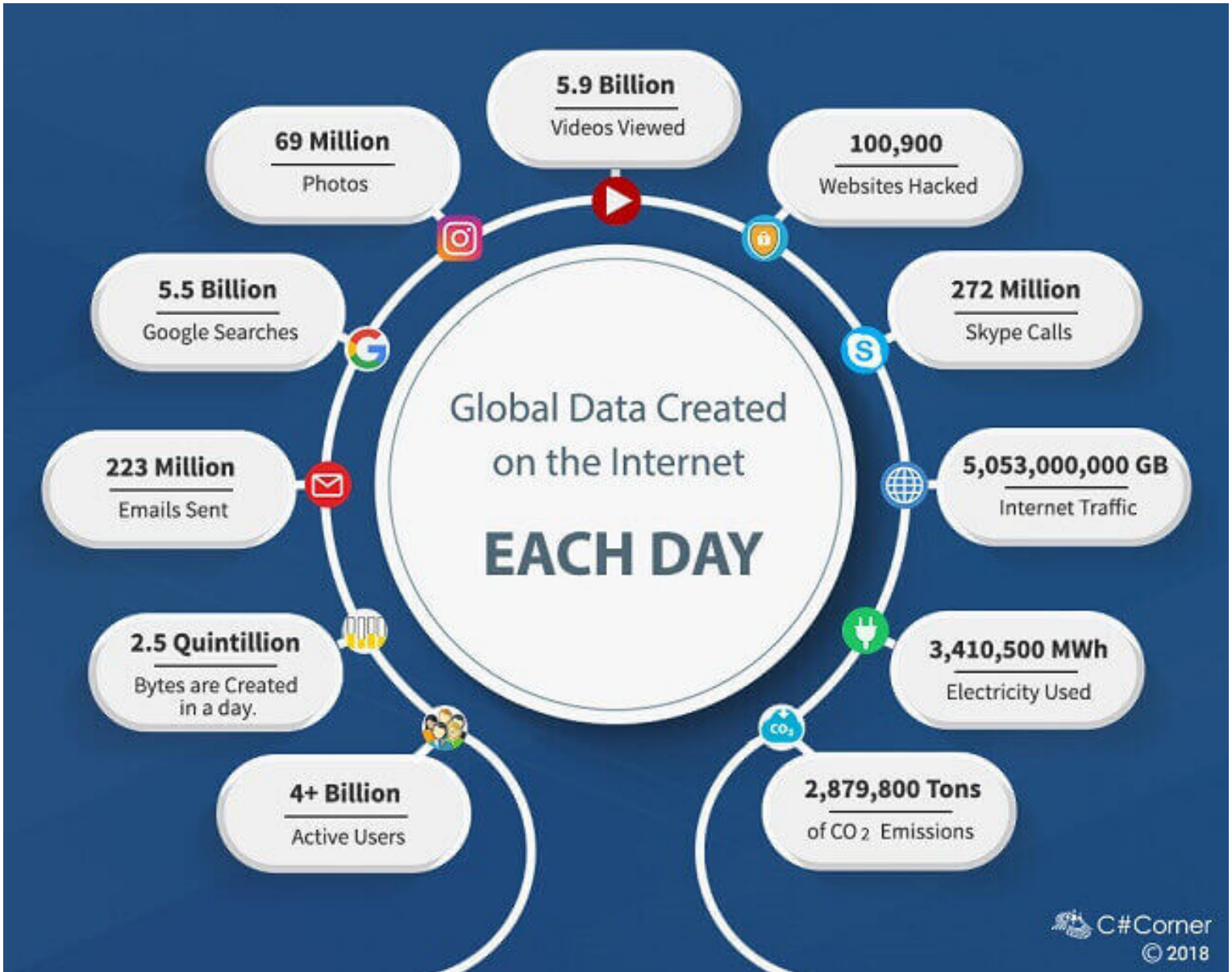
Why This Module?

- How much data did you handle in CS2040C?
- Think BIG!

Google?


Google
Singapore
why professor
why professor x still alive in logan
why professor x can't walk
why professor not in nba
why professor mcgonagall is the best
why professor x died in logan
why professor green name
why professors are arrogant
why professor snape killed dumbledore
why professors don't reply
why professors are liberal
Google Search
I'm Feeling Lucky

How much data does Internet have?



According to MarTech

- According to MarTech, the total data size of the Internet is 2.7 Zettabytes until 2017.

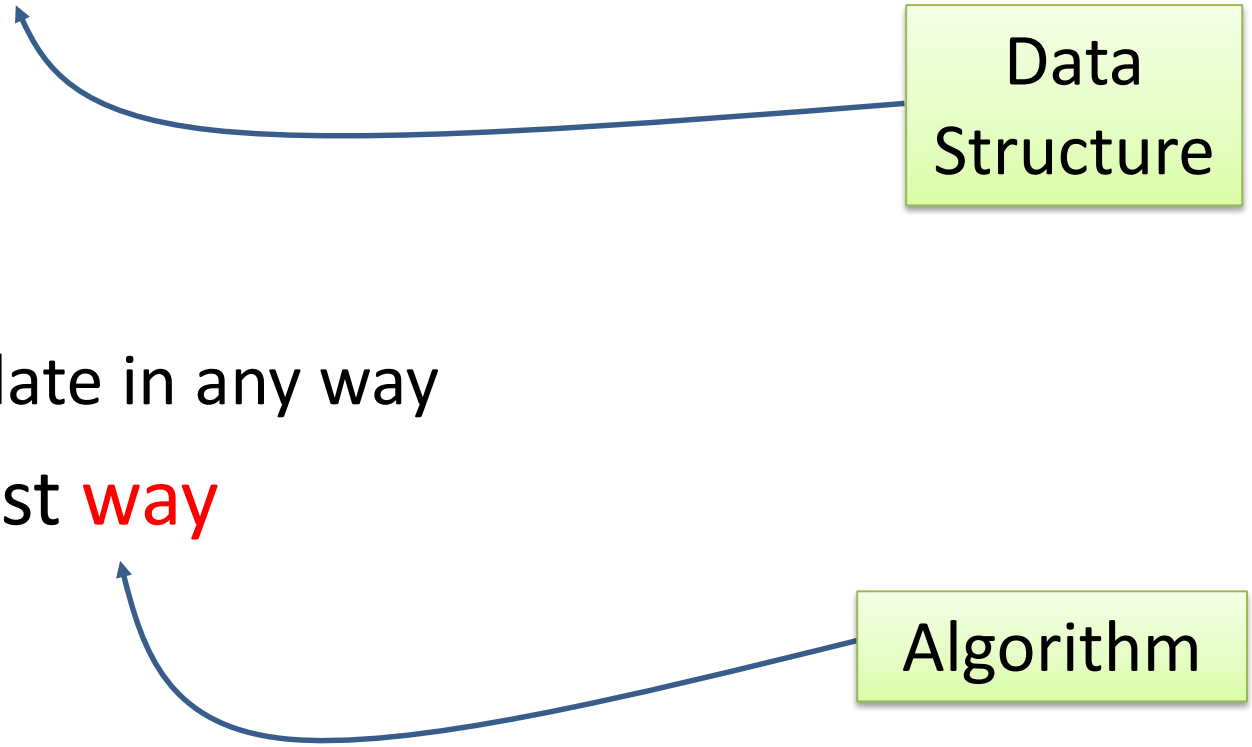
Unit	Decimal Value	Binary Value	Size (in bytes)
Bit (b)	0 or 1	0 or 1	$1/8^{th}$
Byte (B)	8 bits	8 bits	1
Kilobyte (KB)	1000^1 bytes	1024^1 bytes	1,000
Megabyte (MB)	1000^2 bytes	1024^2 bytes	1,000,000
Gigabyte (GB)	1000^3 bytes	1024^3 bytes	1,000,000,000
Terabyte (TB)	1000^4 bytes	1024^4 bytes	1,000,000,000,000
Petabyte (PT)	1000^5 bytes	1024^5 bytes	1,000,000,000,000,000
Exabyte (EB)	1000^6 bytes	1024^6 bytes	1,000,000,000,000,000,000
Zettabyte (ZB)	1000^7 bytes	1024^7 bytes	1,000,000,000,000,000,000,000
Yottabyte (YB)	1000^8 bytes	1024^8 bytes	1,000,000,000,000,000,000,000,000

How much data does Google handle?

- About 10 to 15 Exabyte of data
  - 1 Exabyte(EB)= 1024 Petabyte(PB)
  - 1 Petabyte(PB) = 1024 Terabytes(TB)
  - 1 Terabyte(PB) = 1024 Gigabytes(TB)
    - = 4 X 256GB iPhone
- So Google is handling about 60 millions of iPhones

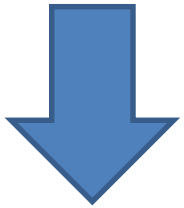
Data Structures

- Always deal with large volume of data
- How to organize data so that we can
  - Modify
  - Search
  - Analyze
  - or manipulate in any way
- in the fastest way



Find a number

```
[257, 186, 63, 231, 321, 312, 210, 249, 149, 339, 420, 299, 252, 34, 128, 281,
313, 370, 109, 120, 454, 133, 496, 460, 302, 419, 9, 466, 412, 326, 493, 248, 3
1, 222, 201, 193, 23, 214, 165, 157, 253, 206, 223, 268, 45, 217, 156, 447, 367
, 392, 86, 402, 85, 380, 294, 65, 131, 292, 277, 172, 192, 340, 459, 492, 140,
341, 336, 27, 346, 208, 282, 278, 71, 124, 254, 243, 47, 96, 436, 189, 418, 371
, 242, 389, 94, 87, 481, 19, 444, 26, 499, 52, 245, 62, 461, 205, 169, 14, 143,
163]
```



```
[9, 14, 19, 23, 26, 27, 31, 34, 45, 47, 52, 62, 63, 65, 71, 85, 86, 87, 94, 96,
109, 120, 124, 128, 131, 133, 140, 143, 149, 156, 157, 163, 165, 169, 172, 186,
189, 192, 193, 201, 205, 206, 208, 210, 214, 217, 222, 223, 231, 242, 243, 245,
248, 249, 252, 253, 254, 257, 268, 277, 278, 281, 282, 292, 294, 299, 302, 312,
313, 321, 326, 336, 339, 340, 341, 346, 367, 370, 371, 380, 389, 392, 402, 412,
418, 419, 420, 436, 444, 447, 454, 459, 460, 461, 466, 481, 492, 493, 496, 499]
```

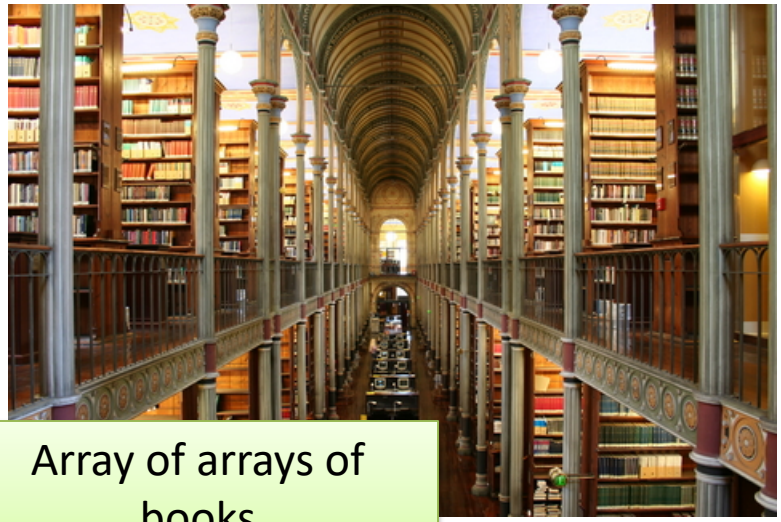
Data Structure and Algorithm

- Data Structure
  - Store the numbers in a sorted way
- Algorithm
  - Searching for a number x:
    - Look for the “middle” number, m
    - If hit, found
    - Search the left or the right part depending if x > m or not

```
[9, 14, 19, 23, 26, 27, 31, 34, 45, 47, 52, 62, 63, 65, 71, 85, 86, 87, 94, 96,
109, 120, 124, 128, 131, 133, 140, 143, 149, 156, 157, 163, 165, 169, 172, 186,
189, 192, 193, 201, 205, 206, 208, 210, 214, 217, 222, 223, 231, 242, 243, 245,
248, 249, 252, 253, 254, 257, 268, 277, 278, 281, 282, 292, 294, 299, 302, 312,
313, 321, 326, 336, 339, 340, 341, 346, 367, 370, 371, 380, 389, 392, 402, 412,
418, 419, 420, 436, 444, 447, 454, 459, 460, 461, 466, 481, 492, 493, 496, 499]
```

Data Structures

- How to organize a lot of data



Array of arrays of books

- with some sophisticated ways

Array of arrays of baskets of books

Algorithm

- What is an algorithm?
- Set of instructions for solving a problem
  1. “wash the tomatoes.”
  2. “peel and cut the carrots.”
  3. “mix the olive oil and vinegar.”
  4. “combine everything in a bowl.”
- Finite sequence of steps
- English, Chinese, pseudocode, Java, etc.
- Unambiguous

Ambiguous Human Language

Algorithms

- History
  - Named for al-Khwārizmī (780-850)
  - Persian mathematician

- Many ancient algorithms
  - Multiplication: Rhind Papyrus
    - Babylon and Egypt: ~1800BC
  - Euclidean Algorithm: Elements
    - Greece: ~300BC
  - Sieve of Eratosthenes
    - Greece: ~200BC

"If you need your software to run twice as fast, hire better programmers.

But if you need your software to run more than twice as fast, use a better **algorithm.**"

-- *Software Lead at Microsoft*

Language Does Not Matter

Algorithms are more important:

Fact: C can be 20x as fast as Python!

Algorithm	Language	Time	10,000 elements
Fast (MergeSort)	**Slow** (Python)	2n log(n) µs	0.266s
Slow (InsertionSort)	**Fast** (C)	$0.01n^2$ µs	1s

(Source: MIT 6.006, 2008)

Course Information

Programming Language

- C++
- We assume you know C well already
  - Including pointers
- IDE will be MSVC
  - Not command line
  - Not MS Visual Studio Code
- We will only provide a crash course for C++
  - More OOP practices will be introduced in CS2030

Syllabus

- Linked List
- Sorting
- Balanced Binary Trees (AVL)
- Hash Tables
- Binary Heaps
- Graphs
  - SSSP
  - MST
  - And more
- Extra topics, e.g. Computational Geometry

Module Information

CA	
Visualgo Quizzes	5%
Assignments (x5)	20%
Quizzes (x2)	20%
PE (F2F)	15%
Final Assessment:	40%

May subject to changes , ±5%

Make-up PEs, Quizzes

- You are qualified for make-ups for
  - Sick leave with a valid MCs from doctors
  - NS (for boys)
  - Represent NUS for international oversea events e.g. Olympics
    - Excluding trainings
- Any thing other than the above are not qualified, some unqualified examples:
  - Oversea Travel (except the above)
  - Representing halls or clubs
    - Including with letters from hall officials
- There will be no more make-up for make-ups.
- By NUS rule, once you started a PE or assessment, you cannot do any make-ups. Same goes for other assessments.


WARNING

Plagiarism

The act of presenting another's work or idea that as your own.

NUS students punished for cheating on take-home exam

PUBLISHED JUN 16, 2020, 5:00 AM SGT

Wong Yang

A handful of National University of Singapore (NUS) students who cheated on a practical examination they were allowed to take home in March were punished last Tuesday, when results for the module were released.

On top of receiving zero marks for the exam, they have been barred from exercising the satisfactory/unsatisfactory option for the module.

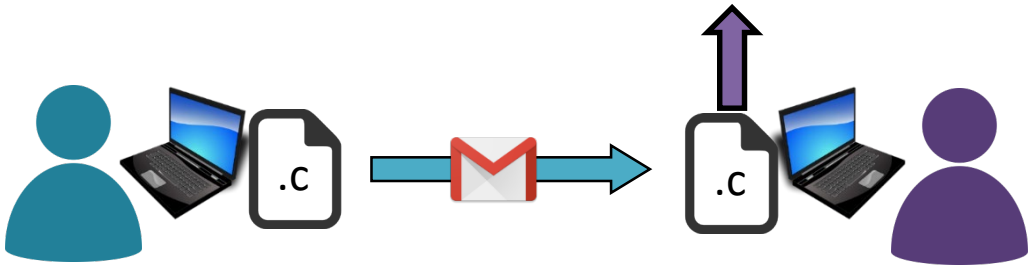
The option to write off grades for up to 10 modular credits for any module taken last semester was made available to NUS students in the light of the disruption to learning caused by the pandemic.

News

- NUS students punished for cheating on take-home exam
  - “On top of receiving zero marks for the exam, they have been barred from exercising the satisfactory/unsatisfactory option for the module.”
  - "I talked to some seniors and they said not to confess because it was just a scare tactic. But I was scared and worried that if I didn't confess, it would be worse later on.
    - "I'll never cheat again, that is for sure."
  - “The students have been reprimanded, and the plagiarism offence **will be included in their formal educational records at the university**”

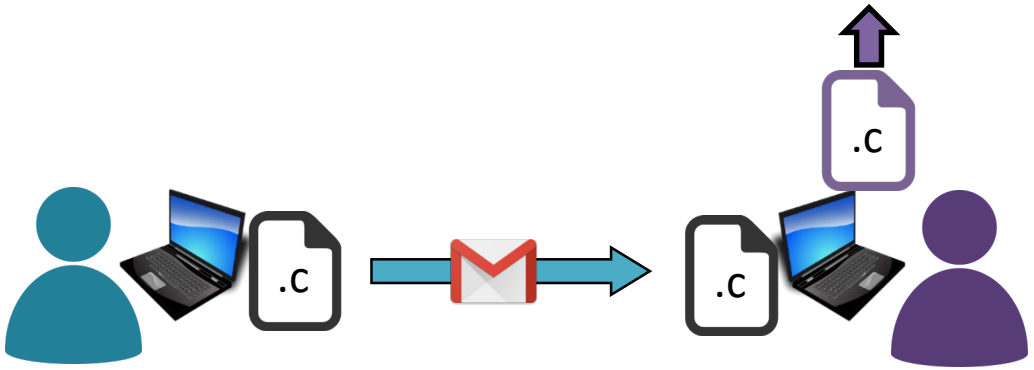
Plagiarism #1

Direct copying



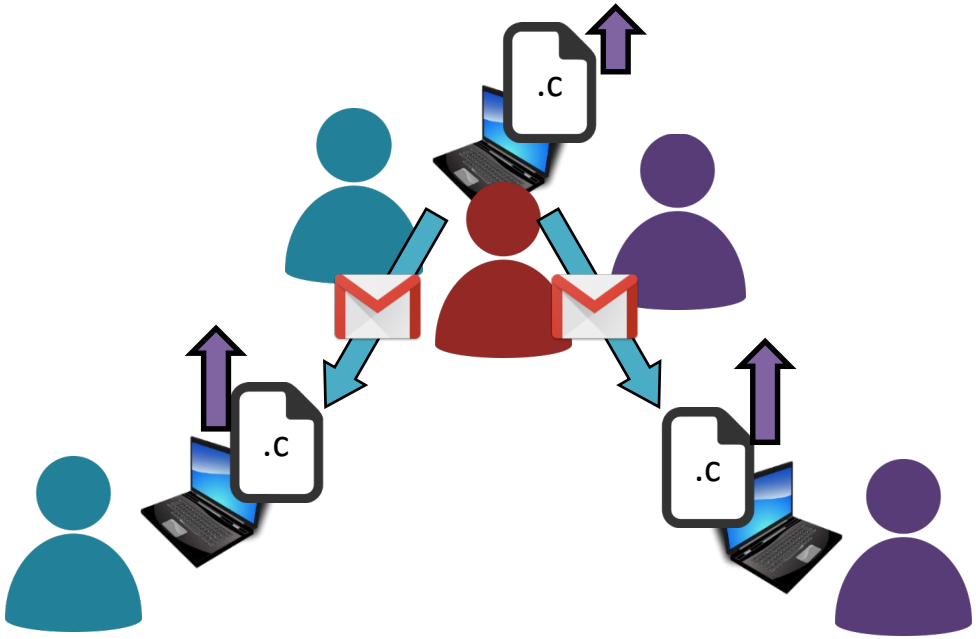
Plagiarism #2

"Referencing"



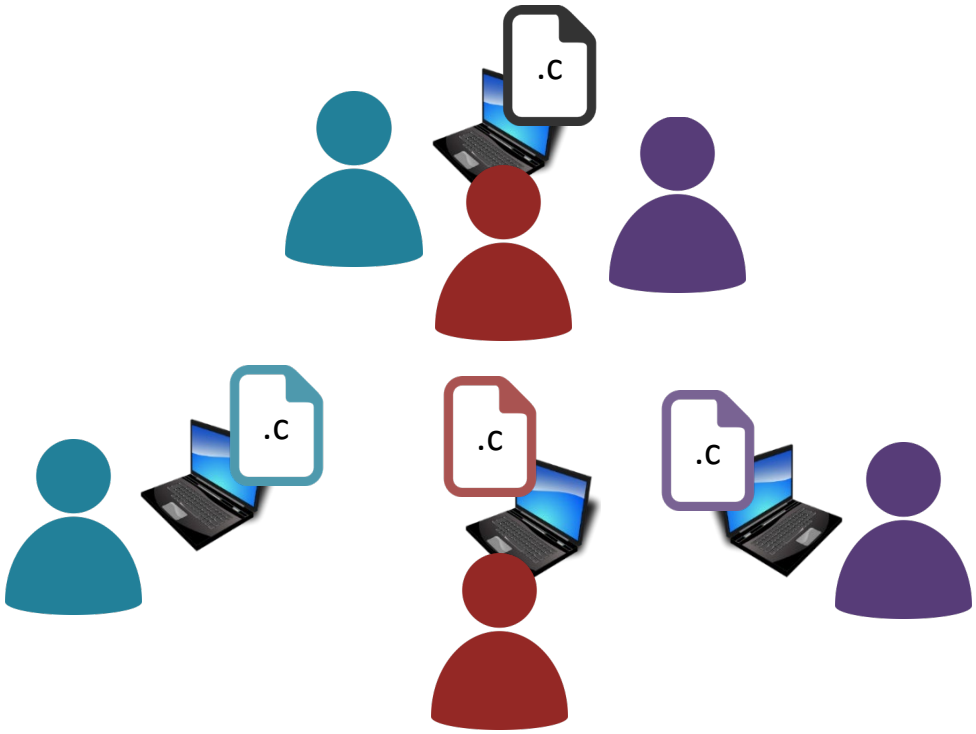
Plagiarism #3

"Discussion"



Not Plagiarism

Discussion (the proper way)



Not Plagiarism

Assistance

Some Plagiarism Techniques (PE)

- Level 0
  - Just copy
- Level 1
  - Add comments
  - Rename variables
  - Add useless lines/statements
  - Shuffle the orders of statements/functions
- Level 2
  - Copy/extract some part of the function to make another function
  - etc.
- Level 3
  - Etc.., etc.

Other Cheating Methods

- Dual screens/keyboards
- Online/local shared storage
  - Dropbox, Google Drive, etc
- Searching/"collaborating" on internet for answers
- Etc. etc.

Conclusion

The knowledge you need to know how to cheat and escape from the plagiarism checker > CS2040C Course materials

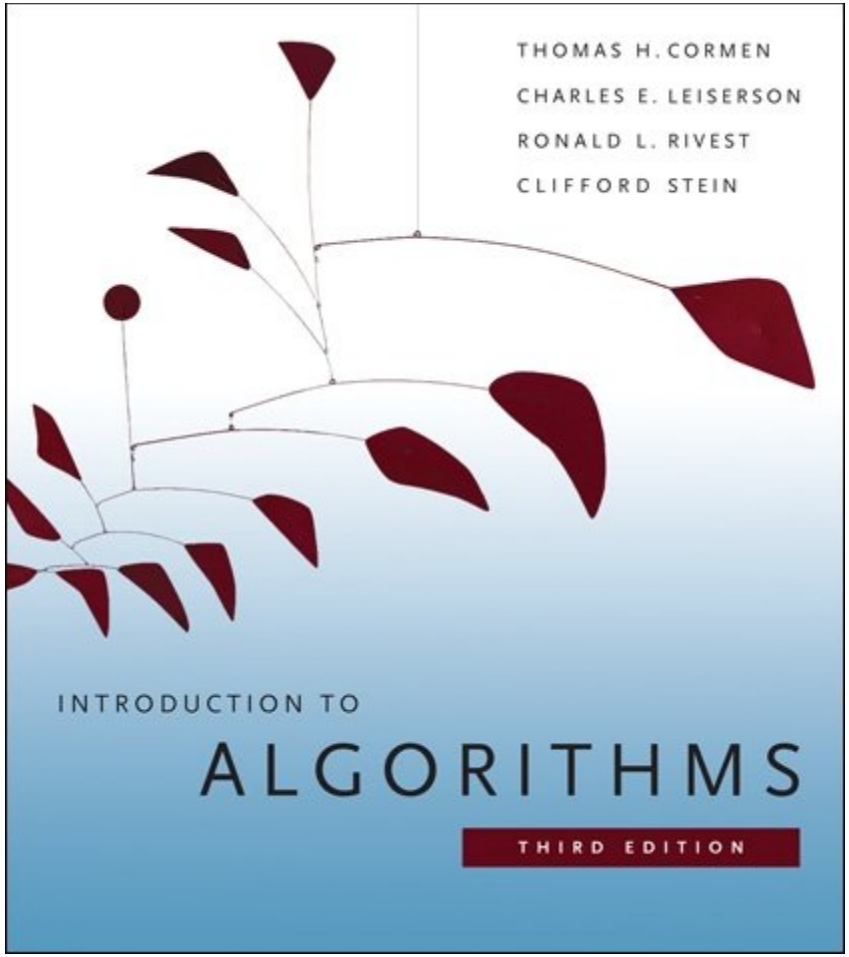
https://youtu.be/hMloyp6NI4E?t=69

Problem Sets: Collaboration Policy

- You can “work” together with your classmates
- You must write/code your problems sets alone.
- You must list on your submission the name of everyone you worked with, and all sources used.
- Cheating / plagiarism will be dealt with harshly.
- Please do not post any code in forum or any public domain

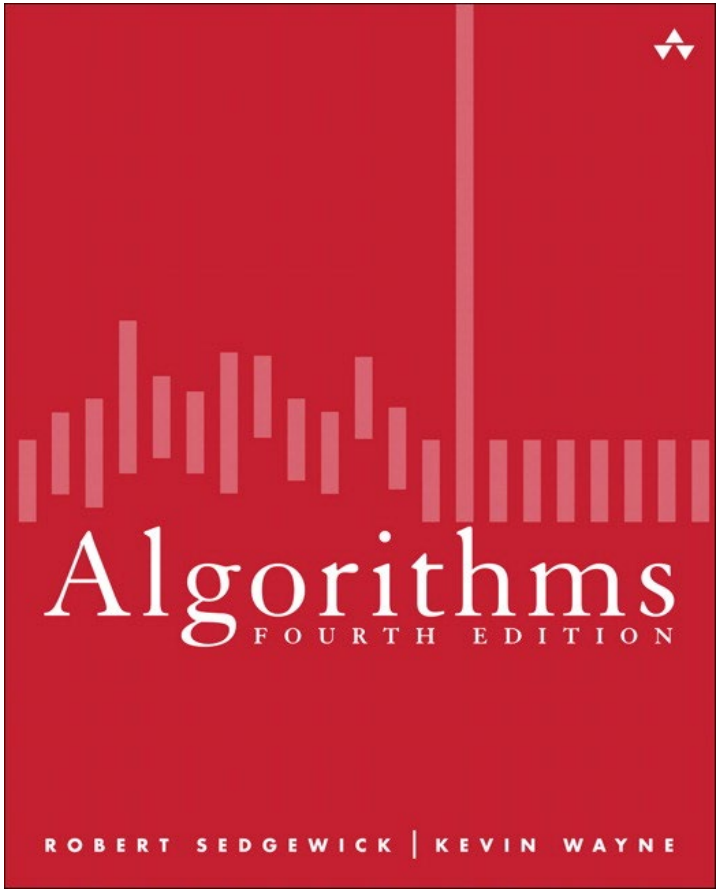
Textbooks
(Not Necessary but Good to have)

- Introduction to Algorithms, Third Edition
  - Cormen, Leiserson, Rivest, Stein



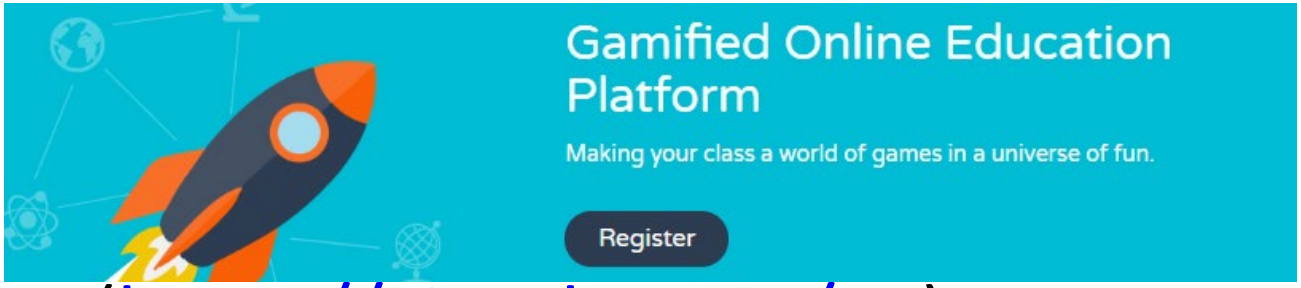
Textbooks
(Not Necessary but Good to have)

- Algorithms
  - Robert Sedgewick and Kevin Wayne

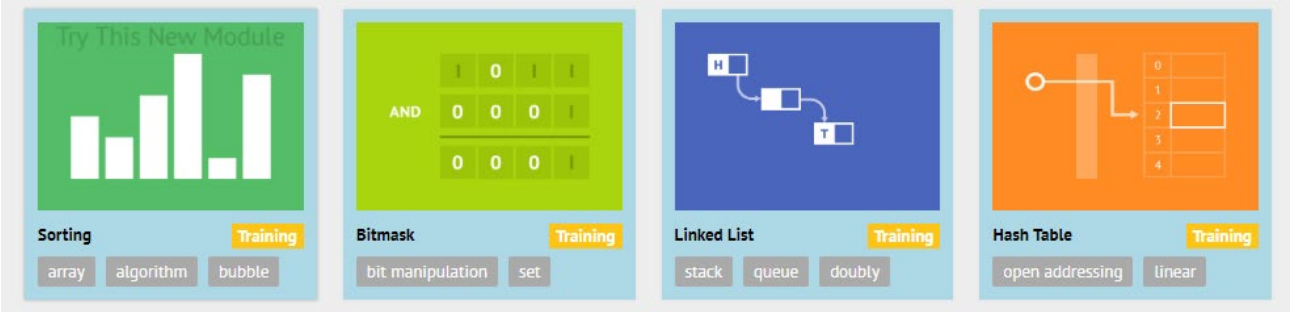


Online Teaching Tools

- Coursmology (https://coursemology.org/)



- Visualgo (https://visualgo.net/en)



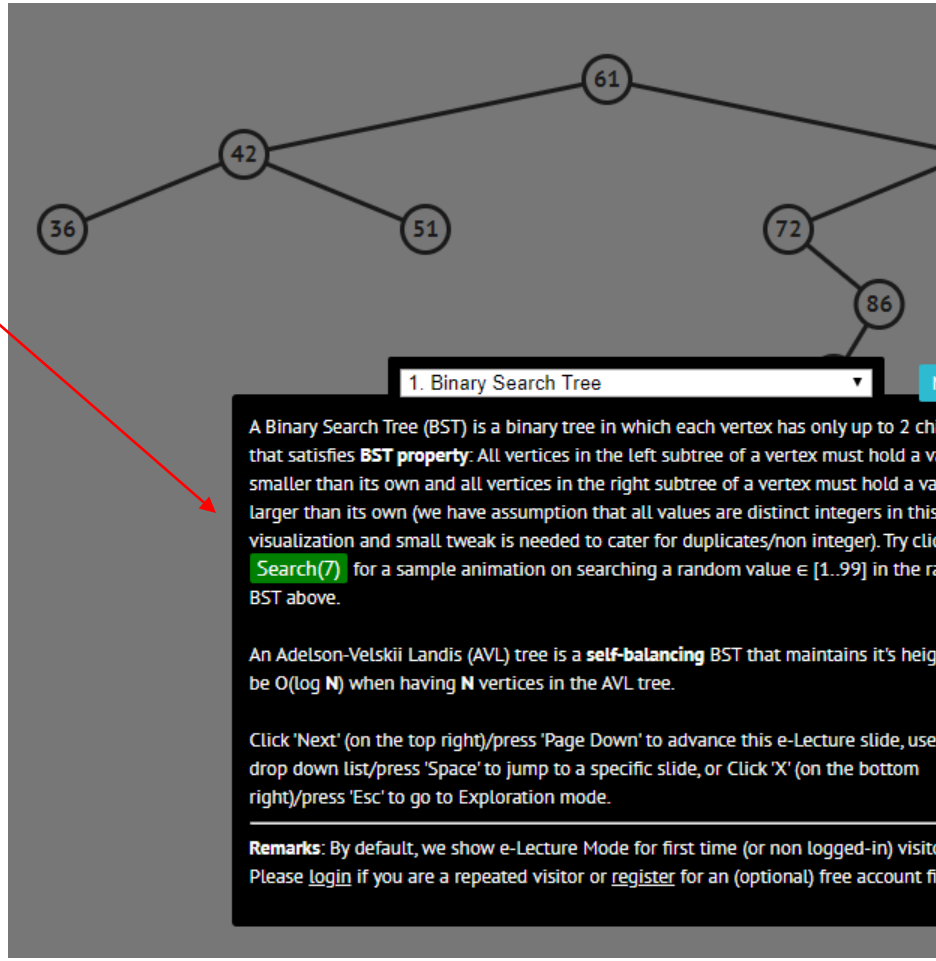
- Archipelago (https://archipelago.comp.nus.edu.sg/)

Coursemology

- Our main platform
  - All materials
  - Homework submission
  - Trainings
  - PE
  - Etc
- We only use Luminus for grade verification at the end of the semester

Visualgo

- Interactive demo
- Flip classroom lecture
- Visualgo Quiz



Programming Tools

- Microsoft Visual Studio 20XX Community Version
  - https://visualstudio.microsoft.com/vs/older-downloads/
  - Not MS VS Code
  - You may download other versions, but may have problem
  - ***Remember to register***
- Mac users can use XCode
- However, in our PE, we will stick on Windows platform with MSVS C++

C++ is so Difficult?

How to be a good programmer

- What is the different between old and new fighting movies?

Grit

https://www.ted.com/talks/angela_lee_duckworth_grit_the_power_of_passion_and_perseverance?language=en