CSIEB0100
Data Structures

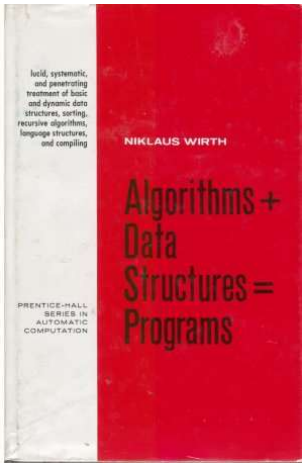
Shiow-yang Wu 吳秀陽
Department of Computer Science and Information Engineering
National Dong Hwa University

Happy Moon Festival !!

- Also known as Mid-Autumn Festival or Mooncake Festival
- One of the most important cultural festival in the Greater China region.

What is Data Structures

- A data structure is a way of organizing and storing data so that it can be processed effectively and efficiently by a computer program.
- This course is therefore about the organization, storage and effective processing of data for computer programs.



Learning Objectives

- Understand the concept of abstract data types(ADT) for data modeling.
- Study different types of **data structures** and the **algorithms** that operate on them.
- Learn how to choose appropriate data structures and algorithms for problem solving.
- Learn to evaluate the performance and cost of data structures and algorithms.
- Learn how to design new data structures and algorithms if necessary.

Lecture and Lab

- This is a lecture-oriented course with associated lab course CSIE@0700.
- It is strongly recommended that you take both courses at the same semester.
- The sample code will be presented in **C++**.
- It is a prerequisite of this class to be familiar with the C++ programming language.
- We will use the free Code::Blocks IDE in class.
- You may choose any C++ compiler you like.

Topics 1

- What are data structures
- Abstract data types (ADTs)
- C++ review and basic algorithms
- Arrays and strings
- Stacks and queues
- Linked lists (singly/doubly linked)
- Trees (basic concepts, binary trees, search, heap)
- Graphs (basic concepts, representations, shortest paths, spanning trees)

Topics 2

- Internal sorting (insertion sort, quick sort, merge sort, heap sort, radix sort)
- External sorting
- Hashing and maps (dictionary structures)
- Priority queues**
- Efficient search structures**
- Advanced data structures**

- **: will be covered if time permits

Special Topic: Data Structures and AI

- Relationship between DSA(Data Structures and Algorithms) and AIML(Artificial Intelligence and Machine Learning)
- Commonly used DSA for AIML
- AIML improves DSA
- It's all about data!

- (Will be covered if time permits)

Administrative Information 1

- **Course Title**: Data Structures
- **Course Number**: CSIEB0100
- **Lecture Time**: Thu 09:10~12:00
- **Classroom**: Science and Engineering Building II A329 (理工二館 A329)
- Office Hours: Thu 17:00~18:00
- Grading Policy: (105%!!)
  - Assignments 35%
  - Midterm Exam 35%
  - Final Exam 35%

Administrative Information 2

- **Course Homepage**: (not on 東華e學苑!)
  http://web.csie.ndhu.edu.tw/showyang/DS2023f/index.html
- **Instructor's Homepage**:
  http://web.csie.ndhu.edu.tw/showyang/index.html
- **TA leader**: 楊舜博(include7212@gmail.com)

Lecture and Lab are graded separately!!

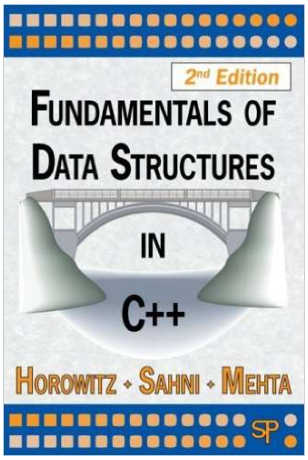
Should take both!!

Online Course Links

- **Online Class Link** (if needed): https://teams.microsoft.com/l/meetup-join/19%3a6YyTrVR5AUYDTXhwe4pJnn-yDyfLudv_IYYKNgockU81%40thread.tacv2/1658041678814?context=%7b%22Tid%22%3a%22edba3211-8174-4411-b089-357c588fa127%22%2c%22Oid%22%3a%22e83708da-2e73-4b78-a037-e2bbca1f4d94%22%7d
- **Join by ID and code**:
  - Meeting ID: 480 724 033 709
    Passcode: f9WC2X

Textbook

- Ellis Horowitz, Sartaj Sahni and Dinesh Mehta. ***Fundamentals of Data Structures in C++, 2nd Edition***, Silicon Press, Summit, New Jersey, 2007.
  (http://www.silicon-press.com/books/isbn.0-929306-37-6/index.html)
  (Code: https://inside.mines.edu/~dmehta/FDS_CPP/)



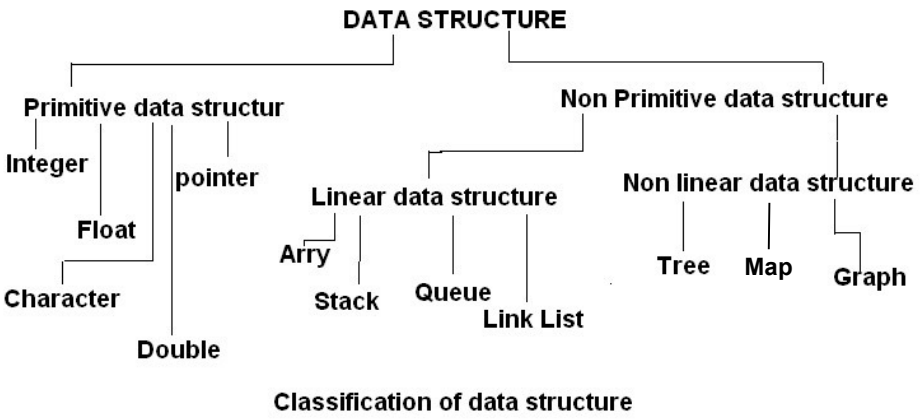
References

- Steven S. Skiena. *The Algorithm Design Manual, 3rd Edition*. Springer, 2020.
- Narasimha Karumanchi. *Data Structures And Algorithms Made Easy, 5th Edition*. CareerMonk Publications, 2016.
- Aditya Bhargava. *Grokking Algorithms: An Illustrated Guide for Programmers and Other Curious People*. Manning, 2016.
- George T. Heineman, Gary Pollice and Stanley Selkow. *Algorithms in a Nutshell, 2nd Edition*. O'Reilly Media, Inc., 2015.
- Clifford A. Shaffer. *Data Structures and Algorithm Analysis, Edition 3.2.0.10*. March 28, 2013.
- Robert Sedgewick and Kevin Wayne. *Algorithms, 4th Edition*. Addison-Wesley Professional, 2011.
- Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest and Clifford Stein. *Introduction to Algorithms, 3rd Edition*. The MIT Press, 2009.

Online References

- Wikibooks. *Fundamental Data Structures*, (http://en.wikipedia.org/wiki/Book:Fundamental_Data_Structures).
- Open Content. *Open Data Structures*, (https://opendatastructures.org/).

Classification of Data Structures



Classification of data structure

Course Outline

- Introduction
- **Part 1: Basic Concepts and Abstract Data Types**
  - What are data structures?
  - Why do we study data structures?
  - What is an abstract data type(ADT)?
  - Relationship between data structures and ADTs
  - Classification of data structures
  - Algorithms and performance analysis
  - Complexity and asymptotic notations

Course Outline

- **Part 2: OO and C++ Review**
  - Object orientation
  - Object-oriented design(OOD)
  - Object-oriented programming(OOP)
  - C++ review
  - C++ templates
  - Data structures in C++
  - OOP with C++

Course Outline

- **Part 3: Linear Data Structures**
  - Arrays and strings
  - Stacks and queues
  - Linked lists
- **Part 4: Non-linear Data Structures**
  - Trees
  - Graphs

Course Outline

- **Part 5: Sorting**
  - Internal sorting (insertion sort, quick sort, merge sort, heap sort, radix sort)
  - External sorting (to handle massive amounts of data)
- **Part 6: Hashing and Maps**
  - Associative arrays
  - Hash functions
  - Hash tables and maps(dictionaries)

Course Outline

- **Part 7: Priority Queues****
  - Priority queues
  - Double-ended priority queues
- **Part 8: Search Structures****
  - AVL trees
  - Red-black trees
  - B-Trees
  - B+-Trees
  - Digital search structures

Course Outline

- **Part 9: Advanced Data Structures****
  - Advanced lists
  - Trie
  - DSA for spatiotemporal data
  - DSA for big data and data science
  - DSA for streaming data
  - DSA for social network analysis(SNA)

(**: will be covered if time permits)

Course Outline

- **Part 10: Data Structures and AI****
  - How does ChatGPT(and other similar LLMs, Large Language Models) work?
  - What data structures and algorithms(DSA) are used in ChatGPT (and other AI&ML frameworks)?
  - AI can help in learning and using DSA.
  - AI can help in designing new DSA.

(**: will be covered if time permits)

Data Structures at Different Levels

- Data structures can be studied at 4 levels:
  - Conceptual level: Understand the definitions, meaning, structures & algorithms
  - Operational level: Know how to actually work on sample cases (with pencil and paper).
  - Programming level: Can actually implement the data structures with programming languages.
  - Problem solving level: Can use the data structures to solve real world problems.
- We will cover the first three levels. You are to master the 4th level by applying them in real world problem solving.

Problem Solving with Data Structures

- CSIE is about computer problem solving.
- Computers need programs to operate.
- Programs = Data Structures + Algorithms
- Data structures are the means and tools to reach our goal.
- It is a good idea to always keep in mind that the theme of the course is

Problem solving with data structures and algorithms

Homework and Programming Assignments

- There will be several homework and programming assignments to give you hands on experience on computer problem solving with data structures.
- Should not be confused with the Lab assignments.
- Late submission will only get partial credit.
- We will use C++ and Code::Blocks.
- You may choose your own IDE.

Use ChatGPT in Coding

- Can I use ChatGPT (or other AI tools) in assignments?
- YES, of course !!
- ChatGPT coding (uses ChatGPT to assist in writing code) is quite popular lately due to:
  - Faster coding
  - Improved accuracy
  - Better productivity
- No free lunch, though. You still need to:
  - Provide the right prompts.
  - **Verify** that the generated code to make sure that it is correct and effective enough.
  - **Improve** the code to satisfy the requirements.

ChatGPT(AI in general) Policy

- With such a useful tool like ChatGPT, there is no reason to ban it.
- ChatGPT coding (coding with AI in general) is becoming an important practice in the industry.
- However, with electronic calculator, you still need to learn arithmetic.
- With ChatGPT, you still need to learn DSA and coding.
- It is OK to use it in assignments.
- You **CANNOT** use it in the exams !!

Exams

- We will have both midterm and final exams
  - To evaluate student performance
  - To help improve the course
- Both are written exams during the official exam weeks.
- Both are open-book exams (open books, notes, codes, …, any paper-based documents).
- **No electronic devices allowed.**
- Obviously, you cannot rely on ChatGPT to assist you in answering exam questions.

Grading and Recommendations

- Grading policy revisited: (105%!!)
  - Assignments (35%)
  - Midterm exam (35%)
  - Final exam (35%)
- Learning Recommendations
  - Read the book(s)/article(s) before the class, participate in the discussion, ask questions!
  - Learning by doing.
  - Practice makes perfect.

Resources

- Wikipedia page on Data Structure: http://en.wikipedia.org/wiki/Data_structure
- Wikipedia, List of data structures https://en.wikipedia.org/wiki/List_of_data_structures
- GeeksforGeeks: https://www.geeksforgeeks.org/data-structures/?ref=shm
- VISUALGO: https://visualgo.net/en
- Tutorialpoint: https://www.tutorialspoint.com/data_structures_algorithms/index.htm
- MIT OpenCourseWare, Introduction to Algorithms https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-006-introduction-to-algorithms-fall-2011/