Cheat Sheet CSE13s
Nelly Sin

Passing by value v. Passing by reference

- "Pass by value": duplicates passed values onto the stack
- "Passing by reference": duplicates a pointer onto the stack
    - It allows "returning" multiple values
    - Allows passing large amounts of data quickly (not copying data but telling it where to store)

Pointer arithmetic

- Adding and subtracting an integer to a pointer = make sense
- Subtracting 2 pointers = make sense (getting the distance between the two are calc.)
- Adding/div/mult 2 pointers = no sense

Comparative sorting algorithms:

- $O(n^2)$ sorting algorithms:
    - Bubble Sort
    - Insertion sort
    - Selection sort
    - Quick sort (worst case)
- $O(n^{5/3})$ sorting algorithm:
    - Shell sort
- O(n log n) sorting algorithms:
    - Merge sort
    - Heap sort
    - Quick sort (average case)

Big O notation:

- To describe how much space an algorithm takes in
- To describe how many steps an algorithm takes

Redex sort:

- Runs in time proportional to the number of digits in the key times the number of records
- It was invented for the mechanical sorting of punched cards

Static v. Dynamic Analyzers

- Static analyzers, *infer*, operate by analyzing the source code for a program before it's run
    - Code is compared against a set (or multiple sets) for coding rules for bugs
    - Only surface level; can't check if a function behaves completely as it's supposed to when it's executed.
    - *Infer* isn't perfect, reported memory leaks are later freed on the code in other functions
- Dynamic analyzers, *valgrind*, operate by tracking down errors that occur during program execution
    - It's good for checking if the program executes as it's supposed to
    - And can only analyze what happens during execution,
    - Things that don't occur during execution aren't analyzed

Recursions:

Recursion algorithms are not inherently inefficient

- A function call requires creating a stack frame, and takes time and space
- All tail recursion functions can be rewritten as iteration (often by a compiler)

Binary search:

- O(Log(n))

If the list is empty then it is not there, if the key is smaller look on the left, if the key is larger than look in the right

Stacks:

Stack operates using the LIFO (last in first out) principle and queue operates using the FIFO principle.

- Operations that can be performed on a queue

- empty , full, enqueue, dequeue

Depth-first-search (DFS) uses a stack and breadth-first-search (BFS) used a queue

Data Compression, Makefiles, and Files

Pointers and Dynamic Memory:

- A pointer is variable that contains a memory address (just a value)
  - Pointers can point to functions as well
- Pointer that don't contain a valid address should point NULL
- The address of a variable can be obtained with the address operator (&)
- The object a pointer points to can be accessed by dereferencing the pointer (*)
- Arbitrary levels of pointing: pointers can point pointers which point to pointers, which point to pointers.. Etc.
- Useful for passing around large data structures

Graphs:

Graphs pervade computer science

- Shortest path finding
- Graph coloring
- Network flow
- Dependency ordering
- Etc.

Come in undirected and directed forms

Used generally to indicate relationships between entities

Can be represented using either an adjacency matrix or using adjacency lists

Bit vectors and sets:

Logical shift:

- Left:
  - Zeroes are shifted in on the right
- Right
  - Zeros are shifted in on the left

Arithmetic shift

- Left
  - Zeros are shifted in on the right
- Right
  - Sign bits are shifted in on the left

Getting high-order nibble

1. A high-order nibble in a byte = most significant 4 bits
2. bit -shift right 4x so that the high-order nibble takes the place of low-order nibble
3. AND with 0x0F

Data Compressions:

- Entropy: a measure of uncertainty of the occurrence of an event
  - $H = -\sum_{i=1}^{n} p_i \log_2(p_i)$

File Access:

- Sequential access
  - Read all bytes/records from the beginning
  - Cannot jump around, may rewind or back up
  - Convenient when medium was magnetic tape
- Random access
  - Bytes/records read in any order
  - Essential for database systems

    - Read can be:
    - Move file marker (seek), then read or…
    - Read then move file marker

Makefile:

- "*" performs expansion and "%" serves as a placeholder
- Useful for when various programs need to use a common set of variable definitions
- Generating a prerequisite by including its Makefile and building it where needed
- Makefile overrides variables and targets with the same name.

Linked Lists:

- Linked data structure:
- Singly: each node contains a data field and pointer to the next node in the list
- Doubly: linked list, each node contains a data field and pointers to the next and previous nodes in the list.
- The last node in the list points to a terminator, usually a null pointer

Regular Expression:

- Series of characters that defines a search pattern
- Used by data bases, websites, vim, and command line programs such as grep and awk
- Lexical analysis in a compiler

Hashing:

- Expected time in arrays and linked lists:
    - Linear search: look-ups in unsorted arrays take O(n)
    - Binary search: lookups in sorted arrays take O(log(n))
    - Linked list look-ups take O(n)
    - Random access arrays: searching in O(1) time
- Hash functions:
    - Map data of arbitrary size to fixed sized values
    - Aim to avoid collisions
- Hash table:
    - A data structure that maps keys to values
    - Provides O(1) access time on average
    - Index into a hashtable is calculated by a hash function
- Bloom filters
    - Probabilistic data structure that tests whether an element is a member of a set
    - The element is either not in the set or it may be in the set.

Cryptography:

- Public key cryptography is usually *slow*
    - Used for:
        - Encrypting small amounts of data or establish a shared key for symmetric encryption algorithms
        - Currently, the most popular method, RSA, involves prime rings and exponentiation
        - Security relies on the difficulty of factoring large composites
        - Other methods involve discrete logarithms and elliptic curves