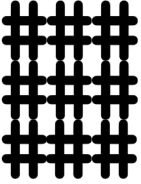


CS5112 Preliminary Exam

October 25, 2018

This test is closed book and closed notes. You have the entire class period, 75 minutes. There is an additional blank page at the end that you can use if you run out of space for your answers.

You are welcome to use the backs of pages for scratch. However, we will not grade anything you write on the back of any page, including the blank one at the end.

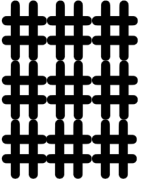
We also will not look at anything written in the top left corner over the hash mark. Be sure that your answer is not ambiguous – if you write multiple solutions, there are no promises which one we will grade.

Point values are given at the top of each page, and in [square brackets] at the start of each question. Take care to answer all the questions. We have used the symbol $\Longrightarrow$ to indicate that an answer is required, and questions that should be true or false are marked T/F.

WRITE YOUR NAME BELOW. WRITE YOUR NETID AT TOP RIGHT OF EACH PAGE (not your name, e.g. "rz13" not "Ramin Zabih")

_____

NAME



Question 1: data structures [10 points]

1. [2] Which of the following data structures is traditionally used in the implementation of Breadth-First Search? ⟹
    a. Hash table
    b. Stack
    c. Queue
    d. Tree

2. [2] Which of the following data structures is traditionally used in the implementation of Depth-First Search? ⟹
    a. Queue
    b. Hash table
    c. Blockchain
    d. Stack

3. [2] Which of the following data structures can be used to improve the running time of Dijkstra's Algorithm? ⟹
    a. Doubly Linked List
    b. Stack
    c. Priority Queue
    d. Binary Search Tree

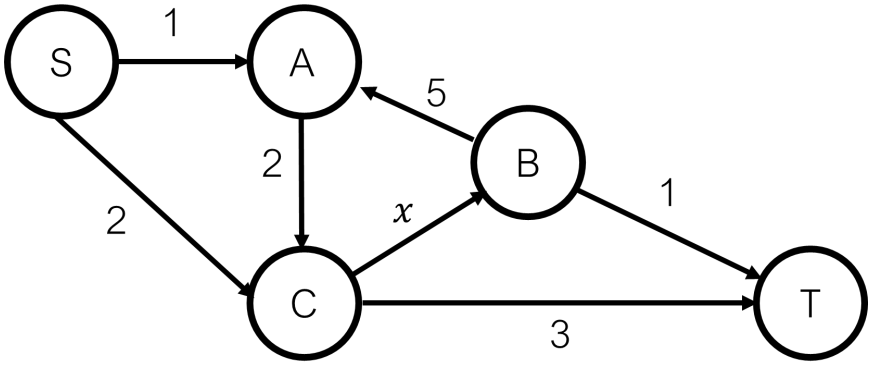
4. [2] Which of the following data structures can be used in the implementation of a hash table to handle collisions (e.g. chaining)? ⟹
    a. Linked List
    b. Binary Search Tree
    c. A second Hash Table
    d. All of the above
    e. a & b
    f. a & c
    g. b & c

5. [2] Suppose the following data structures were being used to store (key, value) pairs. Which of them would allow for the fastest runtime of a `lookup` function which returns the value associated with a given key? ⟹
    a. Linked List
    b. Stack
    c. Paxos
    d. Balanced Binary Search Tree

Netid: _____

Question 2: shortest paths [14 points]

The questions below refer to the following graph, where the weight of one edge is a variable x.



a) [2] What is the shortest path from S to T when $x = 1$? Write down the vertices in order. $\Longrightarrow$

b) [2] When $x = 6$? $\Longrightarrow$

c) [4] When $x = -1$? $\Longrightarrow$

d) [6] Give a value of x for which there is no shortest path. $\Longrightarrow$

Question 3: hashing [29 points]

Consider the following three hash functions that take as input a non-negative integer and produce a single digit output:

$h_z(x)$ = sum of digits modulo z, recall that modulo computes the remainder after division, e.g. 7 modulo 3 = 1, so $h_3(142) = 1$,
$g(x)$ = 2nd to last digit, zero if there is none, so $g(142) = 4$, $g(3) = 0$,
$f(x)$ = table lookup in T below on 3rd to last digit, zero if there is none, so $f(100)=2$, $f(3)=0$

y	0	1	2	3	4	5	6	7	8	9
$T(y)$	4	2	1	3	5	6	8	4	7	0

Given this collection of hash functions we ask the following:

a) [4] Suppose the inputs are {5112, 2110, 312, 6001}. Is there a pair of inputs that collides under the hash function h_4? If so, give the pair. $\Longrightarrow$

b) [4] Give a set of 3 numbers for which f is a perfect hash function. $\Longrightarrow$

c) Suppose we make a Bloom filter $\mathbf{B}$ of length 10 bits, that uses the two hash functions h_5, g. Recall that the output of a Bloom filter is either "the filter has not seen this value" or "the filter may have seen this value".

1. [5] If $\mathbf{B}$ has only seen 5112, 2110 how many of the 10 bits are set to 1? $\Longrightarrow$

2. [3] Suppose that $\mathbf{B}$ has only seen 5112, 2110, 312. The output of $\mathbf{B}$ is that it has not seen 6045. T/F $\Longrightarrow$

3. [5] Give an example input X such that after $\mathbf{B}$ has only seen 5112, 2110, 312 it will output that it may have seen X. $\Longrightarrow$

d) [8] Consider a 10-bit Bloom filter that uses a single hash function. We want it to be the case that if the Bloom filter has only seen 5112, 2110, 312, 6001 it will say that it has not seen 8413. Give an example of such a hash function. $\Longrightarrow$

Question 4: Boyer-Moore majority algorithm [28 points]

Consider an input stream that contains no symbols other than a, b or c. We run the Boyer-Moore majority algorithm, and it produces the output c with count 4.

(a) [3] We can conclude that c is the majority letter T/F $\Longrightarrow$

(b) [3] We can conclude that if there is a majority letter it is c T/F $\Longrightarrow$

(c) In addition, assume that the input contained a, b and c in equal numbers.

c.1 [4] Input contained exactly 12 characters T/F $\Longrightarrow$

c.2 [4] Input contained at least 12 characters T/F $\Longrightarrow$

c.3 [4] There was a c in the last 4 characters T/F $\Longrightarrow$

(d) [10] Give a string of length 12 with an equal number of a, b, c that produces the output c and count 4. To help you we have provided 12 boxes below. $\Longrightarrow$

Question 5: Dynamic programming [19 points]

You would like to open new ice-cream shops at exits along a highway. There are M exits along the highway, spaced a mile apart, where you could place a shop. At each exit m $(0 \leq m < M)$, if you open an ice-cream shop there, you will make a profit $P_m \geq 0$. However, you only have enough money to open (at most) S shops, and because of zoning laws, they must be at least D miles apart.

Example: with M = 9 exits, $D = 3$ miles between shops, a max of $S = 3$ shops, and expected profits:

$$P_0 = 3, P_1 = 4, P_2 = 7, P_3 = 2, P_4 = 5, P_5 = 1, P_6 = 8, P_7 = 0, P_8 = 2$$

The highest possible profit is to open 2 shops, at positions 2 and 6 (4 miles apart), for a profit of 15.

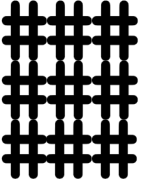
We will write a dynamic program for this problem, using a dynamic programming table T (m, s), using this subproblem: at $T(m, s), 0 \leq m < M, 0 \leq s \leq S$ we will store the maximum possible profit using only exits up to and including m and at most s shops.

Here is a partially-filled-in recursive definition for T ; write the expressions that should go in place of "(Part A)" and "(Part B)" below. Each of your two answers should reference at most 2 cells of $T()$ and exactly 1 of the profits P_m.

$$T(m,s) = \begin{cases} 0 & \text{if} \quad s = 0 \\ P_0 & \text{if} \quad s > 0 \quad \text{and} \quad m = 0 \\ \text{(Part A)} & \text{if} \quad s > 0 \quad \text{and} \quad 1 < m < D \\ \text{(Part B)} & \text{if} \quad s > 0 \quad \text{and} \quad m \geq D \end{cases}$$

Part A: [9] $\Longrightarrow$

Part B: [10] $\Longrightarrow$



This page intentionally left blank. Except for the first sentence, and this one explaining it.