COL 106 Data-Structures Major Test (A) 23rd Nov., 2014

Name _____ Entry Number _____ Gp No.____

Answer all the questions in the space provided for each question.

1. For each of the problems below, describe how you will solve this problem and how much time it will take (in the size of the input).

a. [8 marks] Suppose you are given a dictionary in a foreign language (where the words appear in alphabetical order) and are asked to find the order of characters that appear in the language. For English, say our dictionary has "accord, apple, canary, corner, dice" in that order. Then your output must be "a, c, d, e, i, l, n, o, p, r, y" (for example, the fact that "accord" appears before "canary" means that "a" appears before "c", and so on).

3 marks for identifying the graph
3 marks for edges and topological sort

Build a directed graph G, where the vertices are the letters. For every pair of words, w_1, w_2, (w_1 appears before w_2), find the first position where they differ, put a directed edge between the corresponding letters.
Run topological sort.

2 marks:

n words, m letters

$n^2 L + m + n^2$

$= O(m + n^2 L)$, L: max. length of a word.

b. [8 marks] Snake and ladder problem: You are given an n-by-n snake and ladder board. The squares in the board are labeled with numbers from 1 to n^2 in a continuous manner. At some of the squares, you may find a ladder which goes to a square labeled with a higher number, whereas a snake brings

Modelg: 8 pts (graph & vertices)
Edges: 3 pts
BFS: 1 pt

Build a graph, vertices are the squares. For every square which is not a start point of ladder/snake, 6 edges to the next 6 squares. For a square which is a starting point, an edge to the end of the ladder/snake.

you down to a square with a smaller number. Assume that you have a dice which can show a number between 1 and 6. Whenever you throw the dice, you move ahead by these many steps – if you end-up at a square which has a ladder or a snake, you get transferred to the corresponding square. You are asked to find the minimum possible number of dice throws required to reach the last cell from the first cell.

~~Running time~~ :

Run BFS on this graph.

Running time : $O(\#\text{squares}) + O(\#\text{edges})$: degree of every vertex is ≤ 6

$= O(n^2)$.

$\Rightarrow \#\text{edges} = O(\#\text{vertices})$

2 marks

(if they say Dijkstra, ok with the correct running time of $O(n^2 \log n)$)

$|a - b| \leq 1$

2. [2 marks] Consider two vertices x and y that are simultaneously on the queue during execution of BFS from vertex s in an undirected graph. Which of the following is/are true?

I. The number of edges on the shortest path from s to x is at most one more than the number of edges on the shortest path from s to y.

II. The number of edges on the shortest path from s to x is at least one less than the number of edges on the shortest path from s to y.

III. There is a path from x to y.

(a) Only one of these statements is true. (b) Items I. and II. are true.
(c) Items I. and III. are true. (d) Items II. and III. are true.
(e) All three statements are true.

(c) → 2

X

0 for all else

3. [12 marks] You are given a directed acyclic graph (DAG), represented using an adjacency list, with non-negative edge weights and a fixed source vertex s. Give a O(|V| + |E|) time algorithm to find the shortest path distances from source vertex to all other vertices.

4 mark — Topological sort.

8 mark

can assume

$D[s] = 0$, $D[v] = \infty$ (s has indegree 0)

for $i = 1$ to n do

  for each nbr j of i

    update

      $D[j] \leftarrow D[i] + d(i,j)$ if

      $D[j] > D[i] + d(i,j)$.

Runtime : $O(|V| + |E|)$.

Dijkstra : 4 if correctly written.

4. [12 marks] A bipartite graph is an undirected graph whose vertex set can be partitioned into two sets A and B such that each edge in the graph goes between a vertex in A and a vertex in B (no edges exist between vertices in the same set).
Give an O(|V|+|E|) time algorithm that takes an input graph represented as an adjacency list, and decides if the graph is bipartite.

Run BFS. All cross-edges should be between different levels

6 marks 6 mks.

if the number's starts from 0

i children $di+1$, $d(i+1)$

i parent $\lfloor \frac{i}{d} \rfloor - 1$

5. [10 marks] A d-ary heap is like a binary heap but instead of two children, nodes have d children.

$i \rightarrow d(i-1)+2$ to di

(a) How would you represent a d-ary heap in an array?

Number the vertices starting with 1 and level-wise.

child of a vertex labeled i : ~~di to $di+(d-1)$~~

parent of a vertex i : $\lfloor \frac{i-2}{d} \rfloor + 1$

(b) What is the height of a d-ary heap of n elements in terms of n and d?

$O(\log_d n)$

(c) How much time will deleteMin and insert operations take?

$O(d \log_d n)$, $O(\log_d n)$

6. [6 marks] There is a bag containing 8 white and 8 black balls. You repeat the following experiment till you see a white ball : draw two balls uniformly at random out of the bag. If both are white, stop. Otherwise, put both back in the bag. What is the expected number of times you will need to draw two balls from the bag ?

pr. of 2 white balls $= \frac{{}^8C_2}{{}^{16}C_2} = \frac{8 \cdot 7}{16 \cdot 15} = \frac{7}{30}$.

Expected # times $= 30/7$

$= \sum_i i \cdot (1 - \frac{7}{30})^{i-1} \frac{7}{30}$

7. [2 marks] Suppose we have finished the first iteration of quick sort. What were the possible pivot elements, if the array looks like the following after the first

iteration:
13, 10, 12, 14, 15, 18, 17, 16, 19, 24

14, 15, ~~18~~, 19, 24 (½ pt. for each correct, -½ pt. for each incorrect)

8. [12 marks] You are given an array of n elements with the guarantee that each element is at most K positions from its target position. Devise an algorithm that sorts it in O(n log K) time.

(i) Heapify K elements — 6

6 — repeat [each step: 1 deletemin, insert the next element]

(ii)

6 — Sort each block in $k \lg k$ time: $\frac{N}{K} \times k \log K = N \lg K$

6 — Merge with the next block (starting from the first block)

9. [4 marks] Recall that in the Rabin Karp algorithm, the text subsequence t[i .. i+M-1] is hashed to the number $x(i) = t[i]\, b^{M-1} + t[i+1]b^{M-2} + \ldots + t[i+M-1]$
Suppose we replace the hash function by

$x(i) = t[i]\, R_{M-1} + t[i+1]R_{M-2} + \ldots + t[i+M-1]$ for some fixed numbers $R_{M-1} \ldots R_0$.

Is this a good choice for the hash function? Justify your answer.

No Cannot reuse computation of previous hash fn. for the next one.

1 (

3.

10. [4 marks] Consider the following program for finding a number x in an array A of size n. Assuming that the number x appears exactly twice in the array A, what is the probability that the program will output true ?

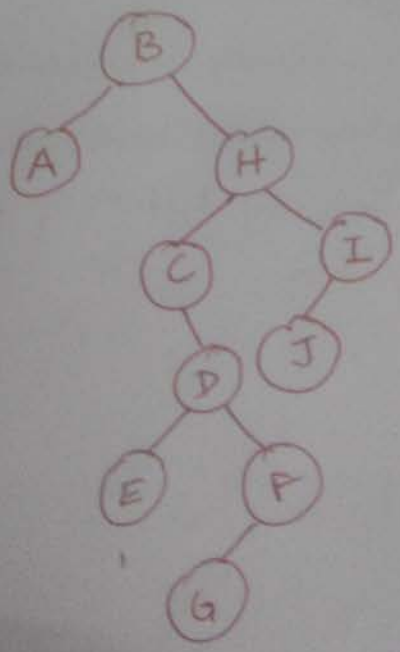
```
function find(A, x) {
    for (i=1; i <= n; i++) {
        pick a random index j between 0 and n-1.
        If (A[i]==x)
                output true;
    }
    output false;
}
```

prob. that we find in one iteration $= \frac{2}{n}$) 1

prob. that we never find it $= \left(1-\frac{2}{n}\right)^n$ 3.

answer $= 1 - \left(1-\frac{2}{n}\right)^n$.

11. [4 marks] Suppose you have a binary tree whose data fields are single characters. The inorder travesal of the tree gives ABCDEFGHIJ, and the preorder traversal gives BAHCEDGFIJ. Draw the binary tree showing the data in each node and the pointers between nodes.



pre: B A H C E D G F I J

in: A B C D E F G H I J

12. [6 marks] You are given the following points in 2-dimensional space :

(100,100), (2,1), (2,2), (3,5), (5,4), (-2, -3), (-5,5)

Draw the corresponding quad-tree and K-D tree to represent the above data.

K-D Tree — 3 marks
Quad Tree — 3 marks.