July 31st 2024

CS 240 Final Help Session

Question 1

First Principles

a) Show that $n \log(n) \in \omega(14n + 22)$ from first principles

Question 2

Algorithm Analysis

Find a tight Θ bound on the runtime of the following algorithm:

```
def algo(n):
  i = 0
  j = n
  while i < j^2:
    i += 1
    j -= 1
```

Question 3

Priority Queues

What would be the best time complexity achievable for `insert` and `delete-max` in a priority-queue implemented with the following data structures

a) dynamic-array
b) binary heap
c) AVL tree
d) Skip list
e) Hash table implementing Cuckoo Hashing

Question 4

Tries

a) True or false, the height of a compressed trie is always in $O(\log n)$
b) Why do we need to store indices in the internal nodes of a compressed trie, but not a pruned trie

c) Given a string S and an array A of forbidden substrings, find the length of the longest substring of S which does not contain any element of A. It may be that $|S| = n$ and $|A|$ are very large, but $|A[i]|$ will be small for each i and can be treated as a constant. (Leetcode Hard)

Question 5

Hashing

a) True or false, if the load factor (α) is 1 and we're using hashing with linear probing then the next insertion is guaranteed to fail
b) True or false, if using Cuckoo hashing then insertion may take an arbitrarily long time
c) Design a data structure that has amortized expected $O(1)$ insert, search and delete and amortized $O(\log n)$ insert, search and delete

Question 6

Range Search

Given an array A containing n unique values, design a data structure which can, given 2 values x and y and two indices i and j return all indices k such that $i \leq k \leq j$ and $x \leq A[k] \leq y$ in

a) $O(n \log n)$ construction time, $o(n) + s$ wost case time per query and $O(n)$ space
b) $O\left(n(\log n)^2\right)$ construction time, $o(\sqrt{n}) + s$ worst case time per query allowing potentially for $\omega(n)$ space.

Question 7

Karp Rabin

Consider using Karp Rabin with the hash function

$$h(p) = (\# \text{ vowels in p})$$

Find a generalizable example of a pattern P and text T which requires $\Theta(nm)$ character comparisons

Question 8

Boyer Moore

Consider the following modifications to the Boyer Moore algorithm. State whether they are still guaranteed to correctly determine whether the pattern P is in the text T or not. If the modification still works, explain why and if not give a concrete example where it fails.

a) Replace the last-occurrence helper array with a first-occurrence helper array
b) When checking a pattern shift, compare characters at the start of the pattern and move forward instead of scanning backwards

Question 9

Huffman Trees

Assume we have some text T on an alphabet of size $|\Sigma| = n \geq 2$. Furthermore assume that the frequency of the i^{th} character in T is the i^{th} Fibonacci number. Show that the height of the Huffman Tree associated with the text is $\geq n - 1$. You may use the fact that $\sum_{i=1}^{n} F(i) = F(n + 2) - 1$

Question 10

LZW Encoding

A secret message was intercepted as follows.

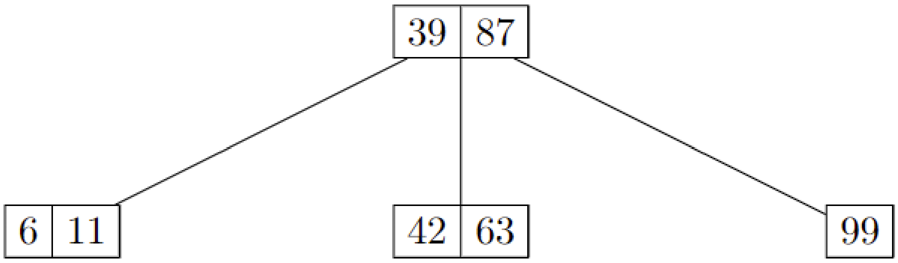
00011001 00100001 01000000 00101000 00100001 00011000 00100001 01000000 00001011 00001011

Intelligence from other sources has indicated that the message was encoded using LZW encoding with 8 bit codewords where the initial dictionary had size 64 although the original dictionary is unknown. Agent Mark did not know the original message or the original dictionary but was immediately able to tell that their intelligence was false. How was he able to determine this?

Question 11

2-4 Trees

Insert 30, 75, 24, 56 into the following into the 2-4 tree below



39
87
6
11
42
63
99