CS 477: Problem Set 1

Section: MW 2-3:50 pm

Total: 185pts Due: 03/04/2020

Instructions:
1. I leave plenty of space on each page for your computation. If you need more sheet, please attach your work right behind the corresponding problem. If your answer is incorrect but you show the computation process, then partial credits will be given. Please staple your solution and use the space wisely.

2. You are allowed to work on the homework in a group (you can stick to your presentation group). No late assignment is accepted. Identical solutions (same wording, paragraph, code), turned in by different groups (persons), will be considered cheating.

3. Full credit will be given only to the correct solution which is described clearly. Convoluted and obtuse descriptions might receive low marks, even when they are correct. Also, aim for concise solutions, as it will save you time spent on write-ups, and also help you conceptualize the key idea of the problem.

4. CLRS refers to the Introduction to Algorithms (3rd edition) textbook while SW refers to the Algorithms (4th edition) textbook specified in the syllabus.

First Name:

Last Name:

Group ID:

Score: **/ 185**

Problem 1 Data Compression: 20 + 20pts

In the SW book, we discussed couple compression mechanisms. One of which is the run-length encoding. We have an open disucssion regarding if it is possible to have a mapping mechansim to map the image such that the resulting image is always of size less than the original. The worst case we showed is the bitstreams all have alternating bits and the run-length might run into problems.

(a) Design a program that can take a 32x48 bits (see example in BinaryDump on the textbook site) and can produce the most efficient reduction you can come up. Please write in either C++ or python.

(b) If it is impossible to design such a compression using run-length approach, can you write a program to analyze the performance of your algorithm on a simple 4 by 4 binary matrix (i.e. only 16 bits, and you have 2^{16} possibl 0-1 matrices), i.e. $sum_{i=1}^{2^{16}} \frac{c_i}{2^{16}}$ where c_i is the compression rate your algorithm has for matrix i. Can your algorithm outperform the stand run-length algorithm?

Problem 2 Data Compression: 15 pts

Please show that Huffman compression is optimal.

Problem 3 Data Compression: 10 pts

Consider the four variable-length codes shown as below: code 1 = (0, 100, 10, 11), code 2 =(0, 1, 00, 11), code 3 =(1, 01, 001, 0001), code 4 =(1, 01, 001, 000) for letters A, B, C and D. Which coding schemes are prefix-free?

Problem 4 Data Compression: 10 pts

Consider the four variable-length codes shown as below: code 1 = (0, 100, 10, 11), code 2 =(0, 1, 00, 11), code 3 =(1, 01, 001, 0001), code 4 =(1, 01, 001, 000) for letters A, B, C and D. Which coding schemes are prefix-free?

Problem 5 Complexity: 10pts

Use a recursion tree to determine a good asymptotic upper bound on the recurrence $T(n) = T(n-a) + T(a) + cn$ where $a \geq 1$ and $c > 0$ are constants.

Problem 6 Quicksort: 20pts

(A) What value of q does PARTITION return when all elements in the array $A[p \cdots r]$ hav the same value? Modify PARTITION so that $= floor(p + r/2)$ when allelements in the array $A[p \cdots r]$ have the same value.

(B) How would you modify QUICKSORT to sort into nonincreasing order?

Problem 7 Simple Random Variables: 10 + 30 points

(A) Let X be a random variable that is equalt to the nubmer of heads in two flips of a biased coin (60 percent Head, 40 percent Tail; Head is 1, Tail is 0). What is $E[X^2]$ and what is $E^2[X]$

(B) Please do the Average sorting problem 8.5 in the CLRS textbook for (a)(b)(c)

Problem 8 Heapsort: 10*4 points

A d -ary heap is like a binary heap, but (with one possible exception) non-leaf nodes have d children instead of 2 children.
(a) How would you represent a d -ary heap in an array?

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

(c) Give an efficient pseudo code EXTRACT-M AX in a d -ary max-heap. Analyze its running time in terms of d and n.

(d). Give an pseduocode of INSERT in a d -ary max-heap. Analyze itsrunning time in terms of d and$n$.