SSN College of Engineering

Department of Computer Science and Engineering

CS1403 — Design and Analysis of Algorithms

2019 – 2020

Session — 06

January 29, 2020

- This homework is due by 4pm on January 29, 2020
- Grace period may be given up to midnight of January 30, 2020
- You can upload only one ZIP file
- The naming convention is "<Your first name (first letter capital and all the other letters small)>-CS1403-S06.zip"
- The questions marked as "OPTIONAL" are, as the name implies, optional! Complete your core assignment first and attempt the optional problem only if you have sufficient time.

---

1. Given probabilities of occurrences of symbols from an alphabet Σ, such as the instance shown below,

Symbol	Probability
a	0.24
b	0.16
c	0.11
d	0.09
e	0.40

find optimal binary code, may be of varying lengths, with prefix property (code for a symbol should not be a prefix for code for another symbol).

To meet the optimality requirement, you should find a code with minimum weighted average which is computed as:

$$\sum_{s \in \Sigma} p_s \times len(code_s)$$

(a) Implement the minheap data structure using a linear array like structure. Your implementation should provide standard heap operations such as: buildHeap, insert, deleteMin, decreaseKey, and increaseKey

(b) Implement Huffman coding using a greedy strategy with time complexity not exceeding $O(n \log n)$. Use your minheap implementation to work with trees maintained in a "forest" (realized using a array like "nodespace").

(c) Write a function to read out the code for a symbol from your final tree.

(d) (OPTIONAL) Assume that the input symbols are sorted according to their probability of occurrence. Rewrite your code to find the Huffman codes in linear time.

(e) (OPTIONAL) Read about adaptive Huffman coding. Learn about the Faller-Gallegar-Knuth Algorithm and Vitter's Algorithm for adaptive Huffman coding.