CS140 - Assignment 7

Due: Sunday, Mar. 24 at 11:59pm

http://www.smbc-comics.com/index.php?db=comics&id=2217

For this assignment, you may (and are encouraged to) work with a partner.

1. **[5 points]** Huffman codes

   Suppose the symbols a, b, c, d, e occur with frequencies 30, 20, 15, 40, 60 respectively, in a file.

   (a) What is the Huffman encoding of the alphabet? Note, follow the algorithm in class, where the smallest value is the left branch and the next smallest is the right branch. For encoding, the left branch will be 0 and the right branch 1.

   (b) If this encoding is applied to the file, what is the length of the encoded file in bits?

Greedy or DP

One of the problems below can be solved more efficiently using a greedy approach and the other cannot (i.e. you must use dynamic programming). For each problem clearly describe your algorithm and state the run-time. For the dynamic programming problem, make sure to describe the table and how it is filled in. For the greedy problem, argue (a formal proof isn't necessary) that your solution is optimal.

2. **[10 points]** You're going on a road trip with friends. Unfortunately, your headlights are broken, so you can only drive in the daytime. Therefore, on any given day you can drive no more than d miles. You have a map with n different hotels and the distances from your start point to each hotel $x_1 < x_2 < ... < x_n$. Your final destination is the last hotel. Describe an algorithm that determines which hotels you should stay in if you want to minimize the number of days it takes you to get to your destination.

3. **[10 points]** Same setup as above, however, you also want to do some sightseeing along the way. To make sure you don't spend too little or too much time in any one place, you decide to add a penalty for having too much free time. If you travel x miles in a day, then the penalty for that day is $(d - x)^2$. Describe an algorithm that determines the hotel sequence that minimizes the total penalty, that is the sum of the daily penalties over all travel days.

More fun

4. **[10 points]** Given a set of points $x_1, x_2, ..., x_n$ on the real line, describe a greedy algorithm that determines the smallest set of unit-length (i.e., length=1) closed intervals that contains all of the given points. State the worst case running time and prove that your algorithm is correct. You do not need write pseudo-code, but make your description clear.

5. **[7 points]** Hashtable review

   (a) **[3 points]** Show the result of inserting 5, 28, 19, 15, 20, 10, 33, 12, 17 into a hashtable with collision resolution by chaining. The table should have 9 slots and use $h(k) = k \bmod 9$ for the hash function.

(b) **[3 points]** Show the result of inserting the first 6 of these into another hashtable using open addressing and linear probing.

(c) **[1 points]** For the these insertions, what was the largest number of collisions you had before finding an open slot and what key was it?