

Problem Set 2: Greedy Algorithms

Handed out Monday, February 3. Due at the start of class Wednesday, February 12.

Homework Information: Some of the problems are probably too long to attempt the night before the due date, so plan accordingly. No late homework will be accepted. You must cite any sources used outside of the course materials and textbook, including anyone you discussed the problems with.

- (4 pts) What is the approximation ratio of the BetterChange algorithm (below), for the set of denominations $\{25, 20, 10, 5, 1\}$? Explain your answer. What is the approximation ratio of the BetterChange algorithm for a set of arbitrary denominations $\{c_1, c_2, \dots, c_d\}$ where $c_1 > c_2 > \dots > c_d$? Explain your answer.

Input: An amount of money M , and an array of d denominations $c = (c_1, c_2, \dots, c_d)$ in decreasing order of value ($c_1 > c_2 > \dots > c_d$).

Output: A list of d integers i_1, i_2, \dots, i_d such that $c_1 i_1 + c_2 i_2 + \dots + c_d i_d = M$ and $i_1 + i_2 + \dots + i_d$ is as small as possible.

```

BetterChange (M, c, d)
    r = M
    for k in 1 to d:
         $i_k = r / c_k$ 
         $r = r - c_k * i_k$ 
    return  $(i_1, i_2, \dots, i_d)$ 

```

- (8 pts) We want to obtain the DNA sequence of a new genome. Because whole genomes are too large to sequence directly, it is common practice to randomly shear multiple copies of its fragments, creating a total of m (possibly overlapping) fragments $f_1 \dots f_m$. Assume that the genome's length is known to be N . Careful alignment of each fragment back to a related genome reveals that fragment f_i runs from position b_i to position e_i . Our goal is to find a minimal tiling path of fragments, that is, the smallest subset of fragments (minimize number of fragments) such that at least one fragment covers every base of the genome. Provide pseudocode for a greedy algorithm to find a minimal tiling path. In particular, tell me what metric the algorithm uses to select a fragment. Is your greedy algorithm correct? Why? (Hint: Prove it is correct or give a counterexample)
- (6 pts) Find a permutation with no decreasing strips for which there exists a reversal that reduces the number of breakpoints. How many permutations are there of 6 elements? How many have a single breakpoint? How many permutations of 6 elements have exactly 2 breakpoints?
- (3 pts) Let b_k be the number of permutations on n elements with exactly k breakpoints. Does b_k always increase as k increases? Explain your answer.
- (5 pts) Write a Python program that takes in a coding DNA sequence and outputs the protein amino acid sequence. In other words, transcribe the DNA into RNA and then translate the RNA into proteins. A codon table is available in Lecture 1, slide 12. The 1-letter abbreviations for each amino acid are also in Lecture1, slides 9 and 10. Include in your homework write-up, your amino acid sequence for the `CodingsequenceofthehumanCFTRgene.txt` file and a print-out of your code file.