Data Structures and Algorithms
Homework Assignment 6

Given: February 22, 2018 **Due:** February 28, 2018

Note: The homework is due **electronically on Gradescope and Canvas** on Wednesday, February 28 by 11:59 pm EST. For late submissions, please refer to the Late Submission Policy on the course webpage. You may use a maximum of 2 late days on this homework.

A. **Gradescope**: You must select the appropriate pages on Gradescope. Gradescope makes this easy for you: before you submit, it asks you to associate pages with the homework questions. Failing to do so will get you points off, which cannot be argued against after the fact. Gradescope may prompt you with a warning to select your cover page, please ignore this warning.

B. **LaTeX**: You must use the `hw121.cls` Latex template provided on the course website, or a harsh penalty will be incurred. Handwritten solutions or solutions not typeset in Latex will not be accepted.

C. **Solutions**: Please write concise and clear solutions; you will get only a partial credit for correct solutions that are either unnecessarily long or not clear. Please refer to the Written Homework Guidelines for all the requirements.

D. **Algorithms**: Whenever you present an algorithm, your answer must include 3 separate sections:

1. A precise description of your algorithm in English. No pseudocode, no code.
2. Proof of correctness of your algorithm
3. Analysis of the running time complexity of your algorithm

E. **Collaboration**: You are allowed to discuss **ideas** for solving homework problems in groups of up to 3 people but *you must write your solutions independently*. Also, you must write on your homework the names of the people with whom you discussed. For a clarification on the collaboration policy, please see Piazza @547

F. **Outside Resources**: Finally, you are not allowed to use *any* material outside of the class notes and the textbook. Any violation of this policy may seriously affect your grade in the class. If you're unsure if something violates our policy, please ask.

---

1. [15 pts - Huffman Codeword Length]

A. Using Huffman encoding scheme on a set S of n symbols with frequencies $f_1, f_2, \ldots, f_n$, what is the longest a codeword could possibly be? Give an example set of frequencies that would produce this case.

B. Prove that if some character occurs with a frequency more than $2/5$, then there is guaranteed to be a codeword of length 1.

C. If all characters occur with frequency less than $1/3$, then there is guaranteed to be no codeword of length 1.

2. [20pts - Ternary Huffman Encoding] In class we studied *binary* Huffman coding of the letters in the alphabet S. Generalize Huffman's algorithm to create a ternary encoding of the letters in S.