COT 6417
Algorithms on Strings and Sequences
Fall 2020
Homework Assignment 2

Please complete this assignment individually (and not in a group). Please email me a single pdf document with the solutions (*only pdf is accepted*). This homework is due on ***Thursday, October 22, 2020***.

1. Let S be a string of length m and let SA denote the suffix array for S. Provide an $O(m)$ time algorithm to construct the LCP array for SA. (Recall that the LCP array stores the length of the longest common prefixes between consecutive pair of suffixes in SA.)

For each of the questions below, assume that string $S = aabaaabbaacda\$$.

2. Apply the $O(m \log m)$ bucket sorting based suffix array construction algorithm of Manber and Myers (which we discussed in class) to string S. For this exercise, you only need to provide the contents of the suffix array at the end of each stage (i.e. contents of SA_1, SA_2, SA_4,. *etc*.). (Note: m is the length of the string).

3. Recall the definition of the successor function Ψ used to construct a compressed suffix array. Compute the successor function Ψ values for string S from its suffix array SA.

4. Compute the Burrows Wheeler Transform for string S using the cyclic shift approach. What are the F and L arrays for string S? For the Backward_Search algorithm discussed in class, show the $<sp,ep>$ values at the end of each iteration when searching for the pattern $P = aaab$ in string S.

5. Generate a wavelet tree for string S.