

**Homework 5****Due: Monday, 12/06/2021, 11:59 PM US EDT****Instructions:**

Type your answers to the following questions and submit them in a single PDF file on Gradescope by the due date and time listed above. Put each question on a separate page and assign pages accordingly when you submit. **You should give justification for all your answers, even if the question does not explicitly say so.**

**Important:**

- From the syllabus: ***"Figures, diagrams, and complex mathematical notations can be handwritten and included in the PDF as an image, but illegible solutions will receive no credit. Note that the definitions of "complex" and "illegible" are at the grader's discretion, so it is in your best interest to type your solutions whenever possible."*** All images and diagrams must be placed in a vertical layout and with a clean background. A picture with smudges, blots, or lazy handwriting will receive no credit. We will not accept regrade requests if an image or diagram does not have a bare minimum of professional presentation.
- Homework assignments in Gradescope without correct page selection will be graded with 0 even if the answers are correct. We will not give partial credits in such a case either.
- Individual homework assignments reschedule due to a valid reason (e.g., medical, military, grief) should be requested through the Office of the Dean of Students. You need to send your request with anticipation. We will not consider requests made after the respective deadline. **We will not grant extensions beyond 11/10/2021 due to academic calendar.**
- Check the syllabus for additional information on late days, extensions, grading, and regrade requests.

**On academic dishonesty:**

Although you are allowed to discuss these problems with other people, your work must be entirely your own. It is considered academic dishonesty to read anyone else's solution to any of these problems, share your solution with another student, or post questions/answers on the Internet. We will report all academic dishonesty incidents to the Office of Student Rights and Responsibilities without exception. More information on this matter is in the syllabus.

**Problem 1 (10 points):**

Show the Brute-Force pattern matching algorithm on the following two patterns for the same text and give the total number of compares that are necessary in each case.

T = where\_there\_is\_a\_will\_there\_is\_a\_way

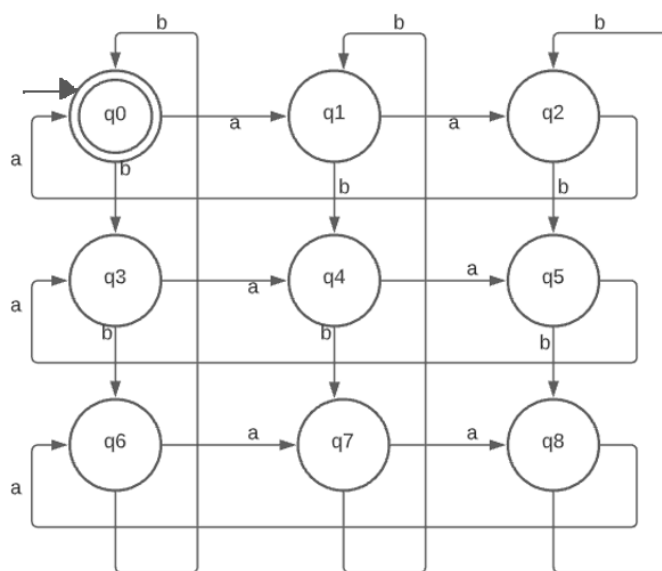
P = a\_way

**Problem 2 (10 points):**

Show the KMP pattern matching algorithm on the following pattern and text and give the total number of compares that are necessary. Be sure to show the failure function.

T = sqraaaqrrrsqsqsrsq

P = sqrsq

**Problem 3 (20 points):**

a) Is the above FA Deterministic?

b) Does the FA accept or reject the following strings?

- aabba
- ababab
- bbbba
- babaa
- baaabb
- bababa
- abababbbaabbbabababa
- ab
- bbb
- bbaaab

c) What are the two smallest strings which can be accepted by the FA?

- d) From part (b), and the state transition diagram of the FA, generalize the type of input string recognized by the FA.

**Problem 4 (10 points):**

Kate wants to send a message to Mike. Since the message is confidential, she wants to encode the message and then send the encoded message to Mike. Mike also receives a “key” (code) from Kate so he can decode the received encoded-message. Assume that the message Kate sent has the following characters appearing with their corresponding frequencies, as shown in the table. Construct a Huffman Tree for the given data in the table and show the final Huffman code for each character. Note- when assigning codes in the tree, make sure to assign the left child as 0 and right child as 1. Also, calculate the average number of bits needed to encode a character (use appropriate units).

A	0.4
B	0.2
C	0.05
D	0.15
–	0.2

**Problem 5 (20 points):**

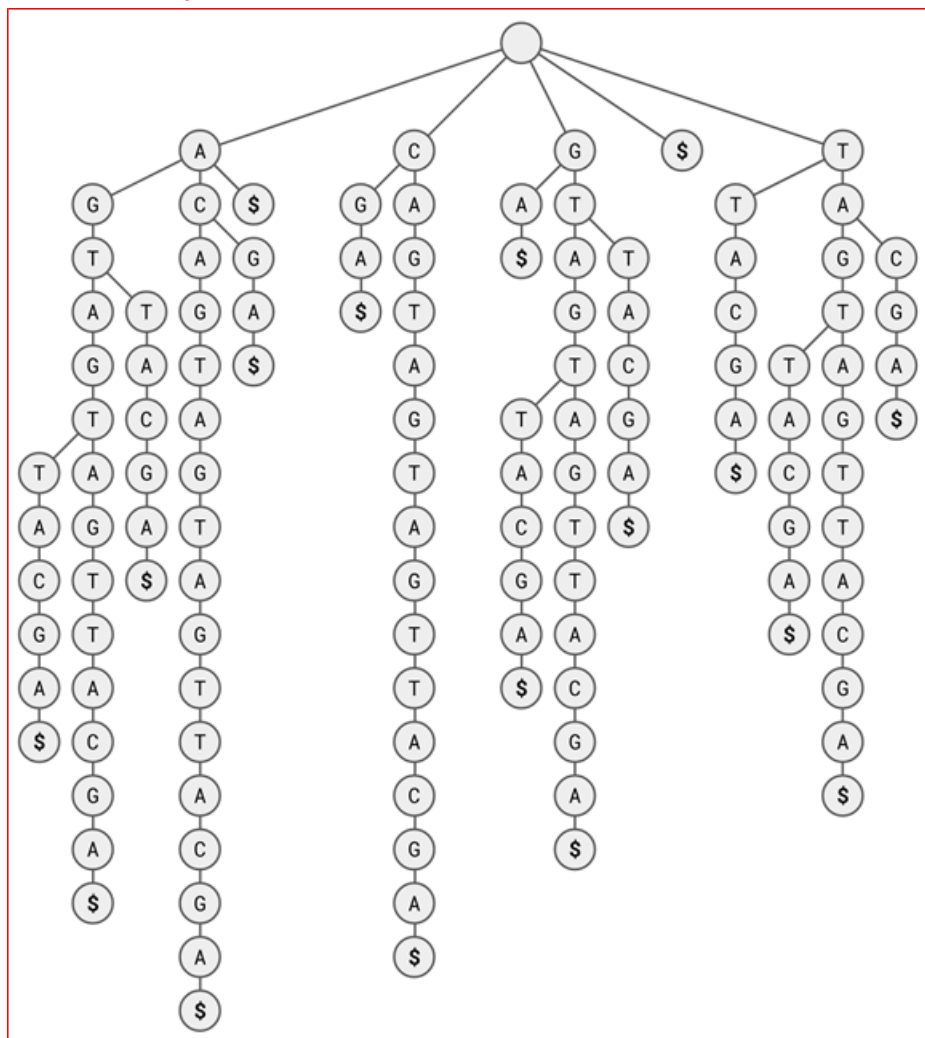
Consider the following regular expressions (space characters added for readability only):

- $(ab)^* (ba)^* \mid (aa)^*$
- $((ab) \mid (aab))^* a^*$
- $((a^* b^* a^*)^* b)^*$
- $((ba) \mid b)^* \mid ((bb) \mid a)^*$

- Give a non-trivial string (i.e., the empty string  $\epsilon$ ) accepted by all four regular expressions. If not possible, explain formally why it is not possible.
- For each regular expression, give a string only recognized by such regular expression and not by the other three. If not possible, explain formally why it is not possible.

**Problem 6 (15 points):**

- a) For each of the following sets of strings, construct a Standard Trie. Add the dollar sign at the end of each word to allow word to be prefixes of other words in the same set.
- {feet, feat, fiction, feature, farther, fashion, fathom, father, feast, feather}
  - {let, lettuce, letter, lexicon, recur, radius, reflex}
  - {sneak, pause, cover, react, taunt}
- b) Convert each of the Standard Tries you constructed on the previous question into a Compressed Trie.

**Problem 7 (15 points):**

- a) Construct a Suffix Trie from the following string: bibbidibobbidi\$

b) Given the following Suffix Trie, state whether each string below occurs in the source string, and if so, how many times it occurs.

- TAG
- AGT
- CG
- A
- GTAC

c) What is the longest substring in the Suffix Trie above that appears more than once?