

CS345 Assignment 2

Q1) Suppose there are two Pattern strings P_1 and P_2 and our goal is to find all occurrence of both the patterns strings in the text string. Suppose the pattern strings are on symbol set Σ . Suppose using KMP technique FSAs, M_1 and M_2 , are designed for each of the patterns where s_1, F_1 are the start state, set of final-states for M_1 and its transition function is d_1 . Similarly s_2, F_2, d_2 are for M_2 . Design a new automaton that can detect all occurrences of P_1 and P_2 .

Can this idea be extended for any number of patterns? What is the disadvantage of combining a large number of patterns in a single FSA?

Q2) Suppose a pattern string is specified as follows. It is constituted by the symbols of Σ and uses two special symbols '?' and '*'. Symbol '?' matches with any one symbol of Σ and '*' matches with any substring. Each Σ symbol only matched with itself. Describe how to find the first occurrence of such a pattern in the text string starting from the first symbol a_1 in the text string?