COMP9319 Exercises

Solution : Please come to the consultations if you have questions with the answers below.

Question 1

Suppose that the BWT encoded string BWT(T) = **arbbrraa$**

where **$** is the last character of T.

Derive the number of matches for the search pattern **ar** using backward search.

Ans: 2 matches

Question 2

Suppose that the BWT encoded string BWT(T) = **acb$cccbaabbcab**

where **$** is the last character of T.

Derive the number of matches for the search pattern **abc** using backward search.

Ans: 2 matches

Question 3

Suppose that the BWT encoded string BWT(T) = **n$rsoocimpse**

Derive the S, B, and B' arrays after applying RLFM index on T.

Ans: S=n$rsocimpse B=111110111111 B'=111111101111

Question 4

Suppose that the RLFM encoded string of text T is **cgc$agagatc** where **$** is the last character of T. Its corresponding bit array B is **1101011101110011**.

Derive its B'.

Ans: B'=1111001101101011

Question 5

Suppose that the RLFM encoded string of text T is **cgc$agagatc** where **$** is the last character of T. Its corresponding bit array B is **1101011101110011**.

Derive the number of matches for the search pattern **cag** using backward search.

Ans: 2 matches

Question 6

Suppose that the RLFM encoded string of text T is **cgc$agagatc** where **$** is the last character of T. Its corresponding bit array B is **1101011101110011**.

Derive the last 4 characters of T.

Ans: agcagcagactg**gac$**