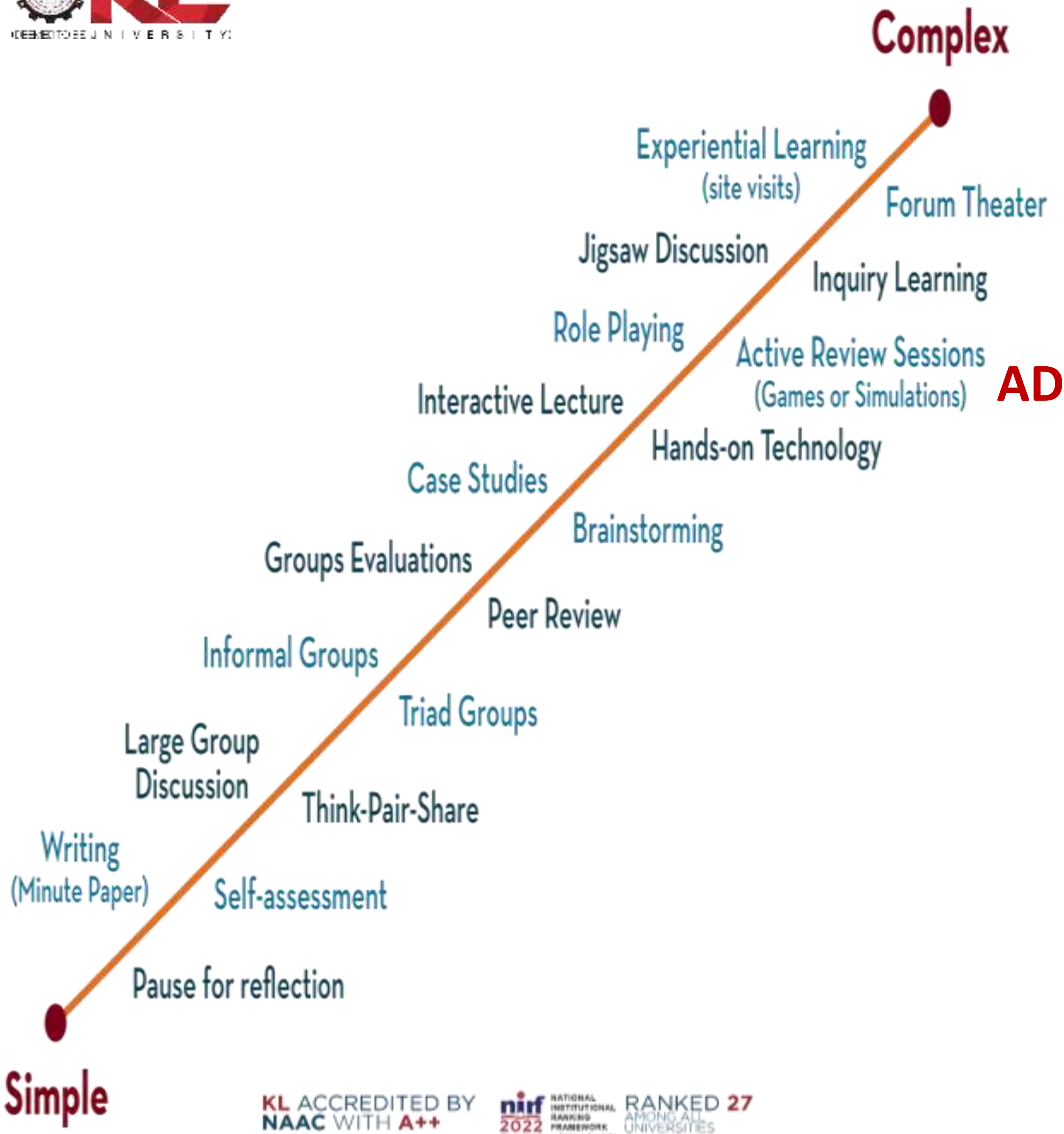
Advanced Algorithms & Data Structures



Department of CSE

ADVANCED ALGORITHMS AND DATA STRUCTURES
23CS03HF

Topic:

Boyer-Moore Algorithnm

Session - 32

AIM OF THE SESSION

To familiarize students with the concept of Boyer Moore algorithm.

INSTRUCTIONAL OBJECTIVES

This Session is designed to:

1. Demonstrate :- Boyer Moore algorithm.
2. Describe :- Sequence of steps in Boyer Moore algorithm

LEARNING OUTCOMES

At the end of this session, you should be able to:

1. Define :- Boyer Moore algorithm.
2. Describe :- Sequence of steps in Boyer Moore algorithm
3. Summarize:- Identification of patterns using Boyer Moore algorithm

Boyer Moore Algorithm

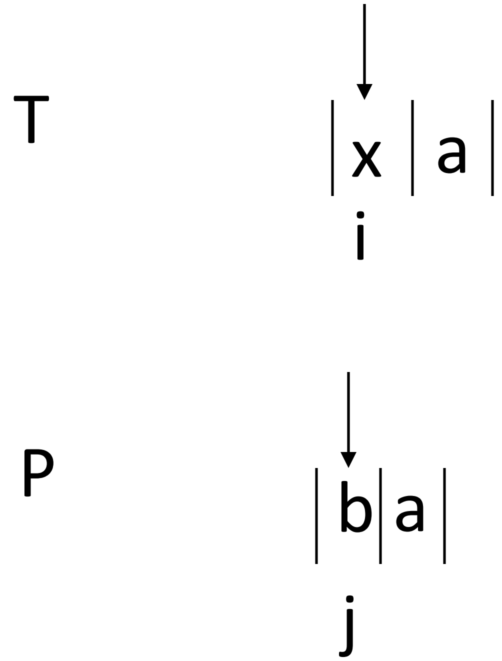
- The Boyer-Moore pattern matching algorithm is based on two techniques.

- 1. The *looking-glass* technique
  - find P in T by moving *backwards* through P, starting at its end

Boyer Moore Algorithm

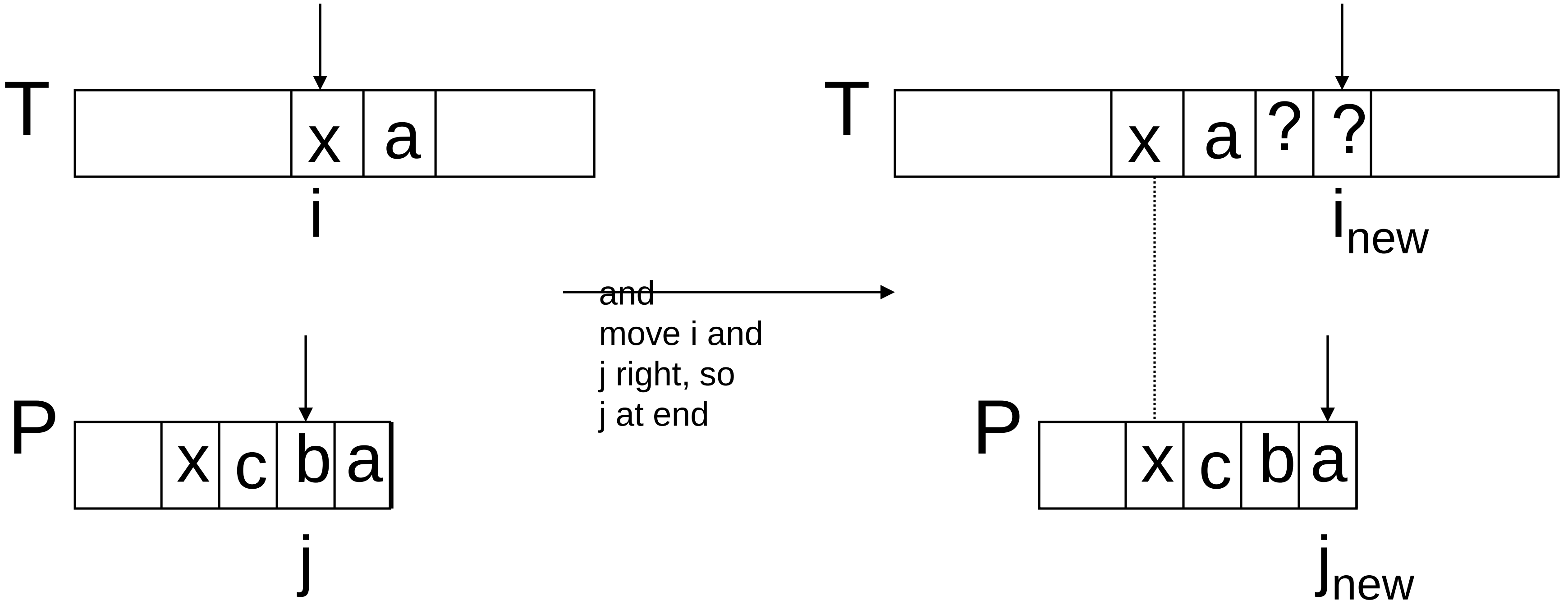
- 2. The *character-jump* technique
  - when a mismatch occurs at T[i] == x
  - the character in pattern P[j] is not the same as T[i]

- There are 3 possible cases, tried in order.



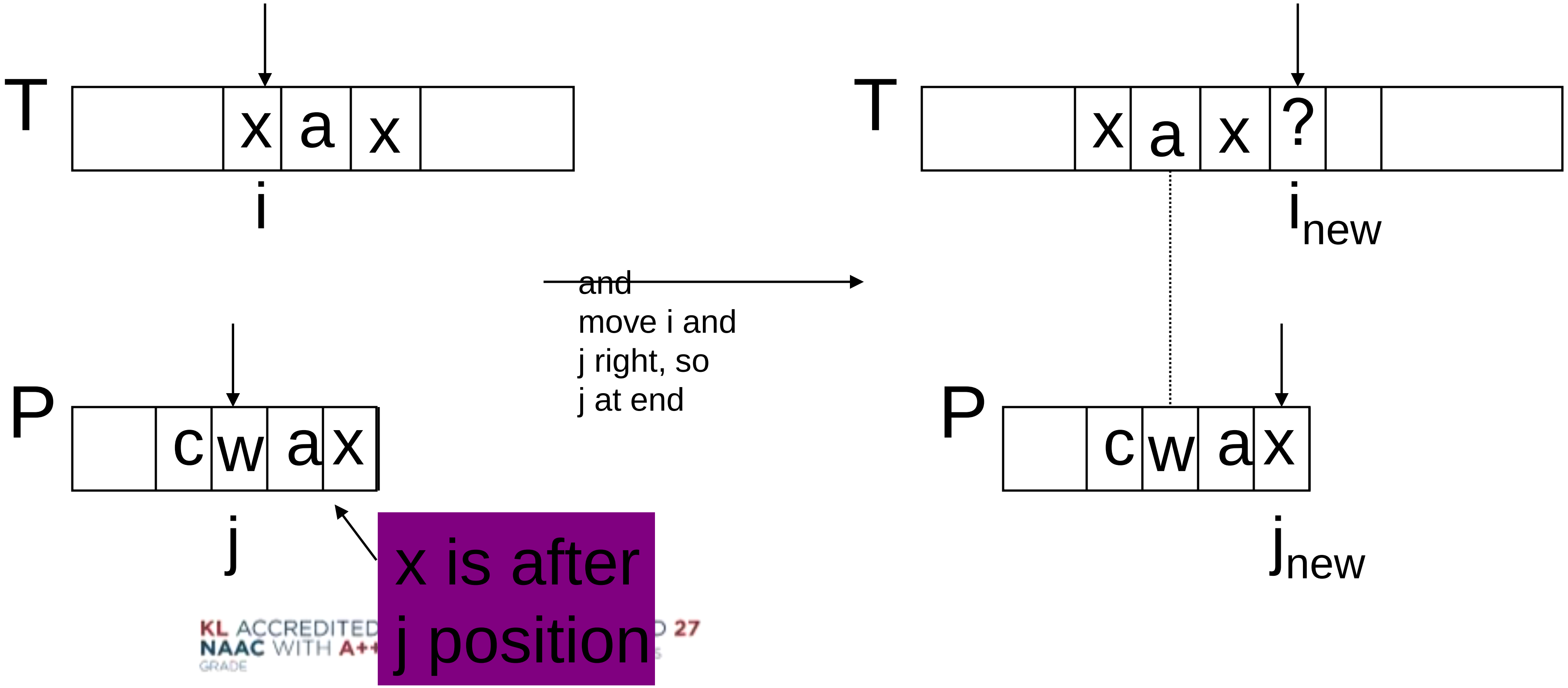
Case 1

- If P contains x somewhere, then try to *shift P* right to align the last occurrence
  of x in P with T[i].



Case 2

- If P contains x somewhere, but a shift right to the last occurrence is *not* possible, then *shift P* right by 1 character to T[i+1].

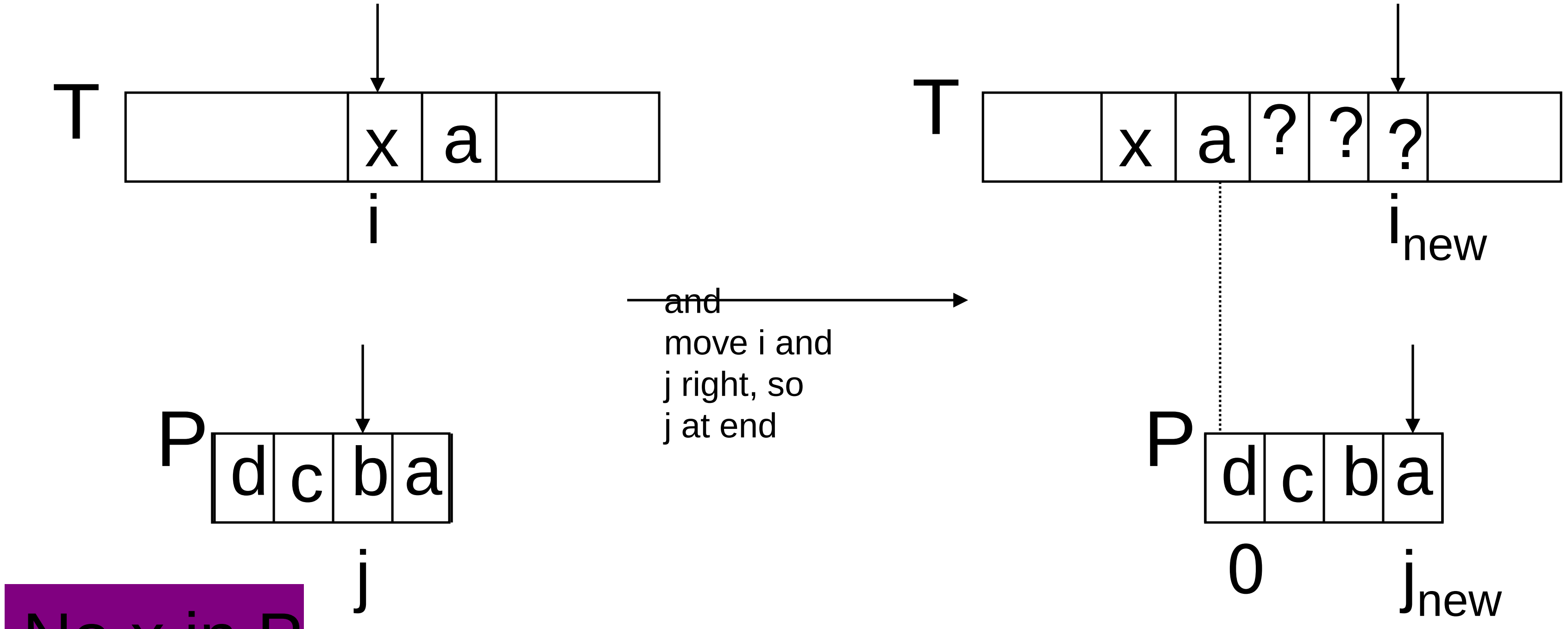


and
move i and
j right, so
j at end

x is after
j position

Case 3

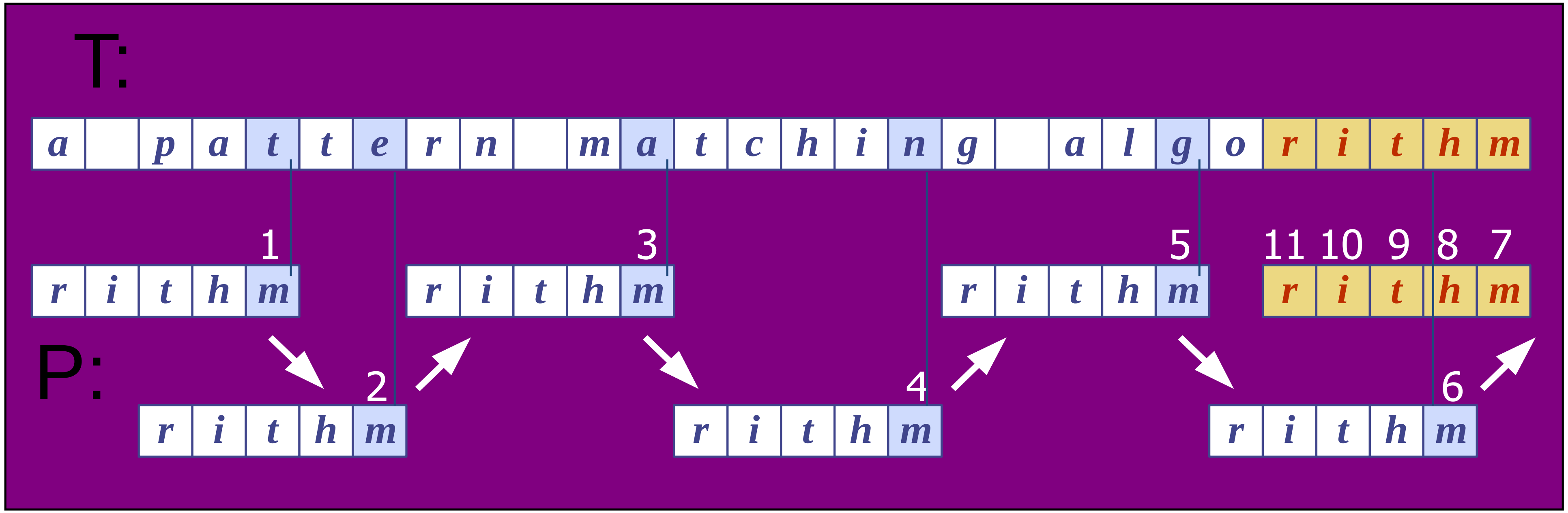
- If cases 1 and 2 do not apply, then *shift* P to align P[0] with T[i+1].



and
move i and
j right, so
j at end

No x in P

Boyer Moore Algorithm



SUMMARY

Boyer-Moore is known for its **sublinear time complexity** in average cases. It avoids unnecessary comparisons by skipping over portions of the text, making it faster than other algorithms like **Naive String Matching**.

SELF-ASSESSMENT QUESTIONS

The Boyer-Moore algorithm performs well when the alphabet size is:

(a) Small

(b) Medium

(c) Large

(d) Any size

In the Boyer-Moore algorithm, which character is compared first when matching the pattern to the text?

(a) The leftmost character of the pattern

(b) The rightmost character of the pattern

(c) The middle character of the pattern

(d) Any character in the pattern

TERMINAL QUESTIONS

1. What are the two main heuristics used by the Boyer-Moore algorithm? Explain each heuristic in detail with an example.
2. Discuss the advantages and disadvantages of the Boyer-Moore algorithm. When would you choose it over other string matching algorithms?

REFERENCES FOR FURTHER LEARNING OF THE SESSION

Reference Books :

1. Introduction to Algorithms, Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein., 3rd, 2009, The MIT Press.
2 Algorithm Design Manual, Steven S. Skiena., 2nd, 2008, Springer.
3 Data Structures and Algorithms in Python, Michael T. Goodrich, Roberto Tamassia, and Michael H. Goldwasser., 2nd, 2013, Wiley.
4 The Art of Computer Programming, Donald E. Knuth, 3rd, 1997, Addison-Wesley Professiona.

MOOCS :

1. https://www.coursera.org/specializations/algorithms?=
2.https://www.coursera.org/learn/dynamic-programming-greedy-algorithms#modules

THANK YOU