

# Advanced Kmp Pattern Matching Algorithm

Narendar Kumboji

Hyderabad Institute of Technology and Management, Telangana State, India

Department: Computer Science of Engineering

Email:narendarkumboji@gmail.com

**Abstract:** Pattern Matching is the process of finding a sequence of characters often known as patterns in the text provided. Pattern matching plays an important role in various fields like intrusion detection, text processing, information retrieval, artificial intelligence etc. In the wide spectrum of pattern matching subject various algorithms were already proposed. Advanced Knuth Morris Pratt algorithm is the one among them and improvement of Knuth Morris Pratt algorithm.

**Index terms:** Pattern, text, KMP algorithm.

## 1. INTRODUCTION:

In computer science, pattern matching is the act of checking input sequence of tokens for the presence of the constituents of some pattern. In contrast to the match usually has to be same. The patterns generally have the form of either sequences or structures. Uses of pattern matching include outputting the locations of a pattern within a token sequence or text, to output some component of the matched pattern.

**Knuth Morris Pratt algorithm:** The scientists Knuth, Morris and Pratt discovered first linear time string-matching which checks the characters from left to right. When a pattern has a sub-pattern appears more than once, it uses that property to improve the time complexity, also for in the worst case. By avoiding this waste of information, it achieves a running time of  $O(m + n)$  using failure function. The implementation of Knuth-Morris-Pratt algorithm is efficient because it reduces the total number of comparisons of the pattern against the input string. The KMP matching algorithm uses degenerating Property. The pattern having same sub-patterns appearing more than once in the pattern.

**TEXT: aabdaaabdaabab**

**PATTERN: aabdaab**

FIGURE 1: KMP COMPUTATION

In the above example the pattern contains same Sub-pattern (aab). It will make to iterate forward easily with checking of common prefix and suffix. It will take 14 comparisons to find out

The pattern from the text.

## 2. ADVANCED KMP ALGORITHM

The algorithm is similar to Knuth Morris Pattern Algorithm which checks the characters from left to right. It is efficient because it reduces the Total number of comparisons of

The pattern against the input string like KMP.

It also uses failure function table. Advanced KMP algorithm also uses degenerating Property. The pattern having same sub-patterns appearing more than once in the pattern. The difference between KMP matching Algorithm and advanced KMP algorithm is that when mismatch in suffix positions. It begins the iteration from prefix where value Return by failure function of prefix position itself. The other difference is that when mismatch Occur at Index equal to failure function Value at that point then begin iteration of pattern from beginning. The time complexity is  $O(m + n)$ .

**TEXT: aabdaaabdaabab**

**PATTERN: aabdaab**

In above example the pattern contains pattern (aab). It will make to iterate forward with checking of common prefix and suffix. In above example it will take 13 comparisons to find out the Pattern from the text.

## 3. ADVANCED KMP APPLICATIONS

The Algorithm may have similar applications like other string pattern matching algorithms Like spell checkers, intrusion detection system Spam filters, search engines, plagiarism detection, DNA sequencing, text processing, information Retrieval, Digital forensics, Artificial intelligence etc.

## 4. CONCLUSION:

String matching algorithms has greatly influenced the field of computer science and

Will play an important role in various fields. Efficient new string matching algorithms will

Be introduced as the Usage of String matching algorithms is increasing, Advanced Knuth Morris Pattern Algorithm is one such type of algorithm. There are many string matching algorithms has been suggested still now, which are used in many areas in which string matching plays an important

role. Knuth Morris Pattern algorithm while processing it does not perform back tracking of text while iterating. It performs back tracking of pattern while iterating using Failure function table. Advanced Knuth Morris Pattern algorithm while processing it also does not perform back tracking of text while iterating. It performs back tracking of pattern while iterating using failure function table which is more efficient than Knuth Morris Pattern algorithm. The search time of both algorithm is same when a pattern doesn't contains common substrings. The search time of advanced KMP algorithm is less than KMP, when a pattern contains common substrings.

**AUTHOR PROFILE:**

Narendar Kumboji has received his Bachelor of Technology in Computer Science & Engineering from CMR College of Engineering and Technology, Telangana, India. He is pursuing his Master in Engineering in Computer Science & Engineering from Hyderabad institute of engineering and Technology, Telangana, India. His main research interests include: String matching algorithms.