

Sunday, 24 April 16

Crux

Text Processing

Manisha Khattar



String Class in Java

1. Immutable
2. Substring
3. Concat
4. indexOf
5. startsWith
6. endsWith

Mutable String?

StringBuffer

1. Mutable String
2. Append
3. Insert
4. Reverse
5. `setCharAt`
6. `.toString`

How to choose between String and StringBuffer?

Text Processing

1. Pattern Matching
 - a. Brute Force
 - b. Boyer Moore
 - c. Tries
2. Text Compression
3. Text Similarity
 - a. LCS
 - b. Edit Distance

Pattern Matching Algorithms

Lets implement the Brute Force Algorithm

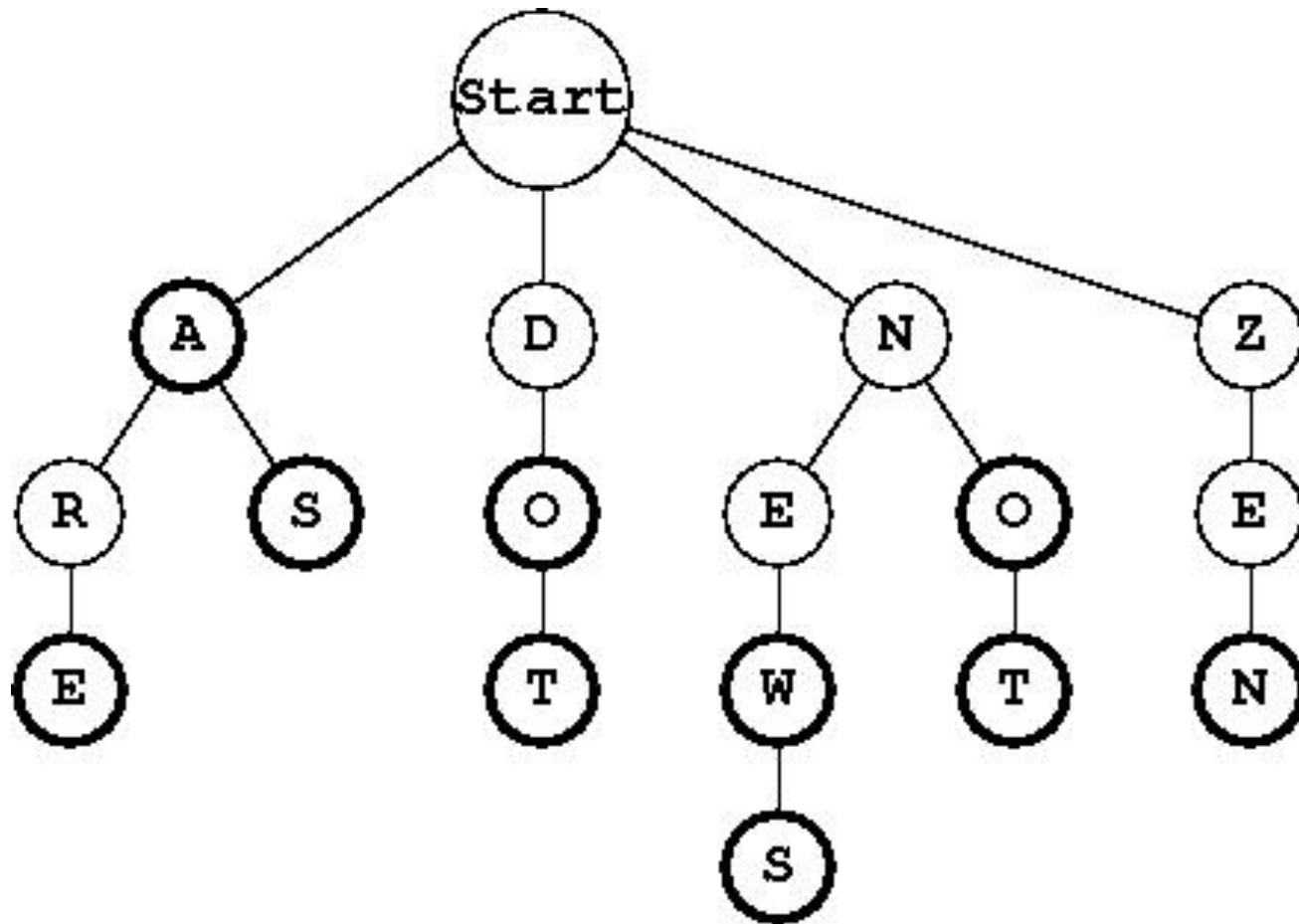
Boyer Moore Algorithm

Lets implement !!

What if we get multiple queries for a fixed String?

Trie

Trie

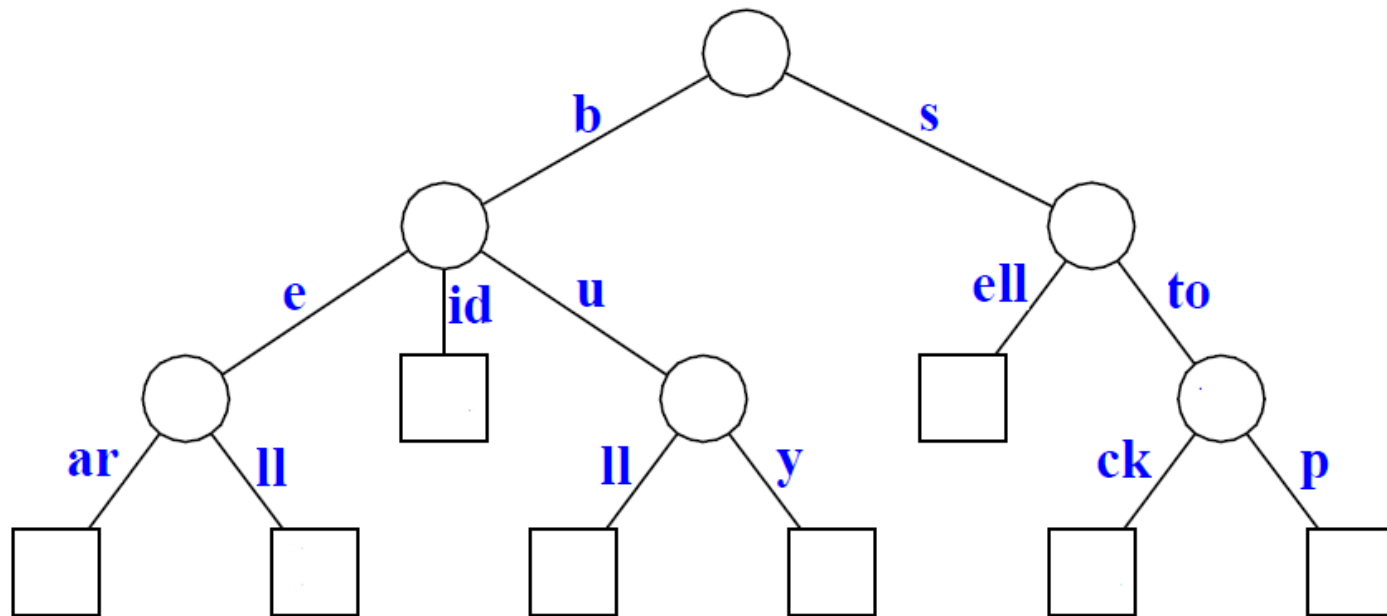


Lets Implement Trie Class



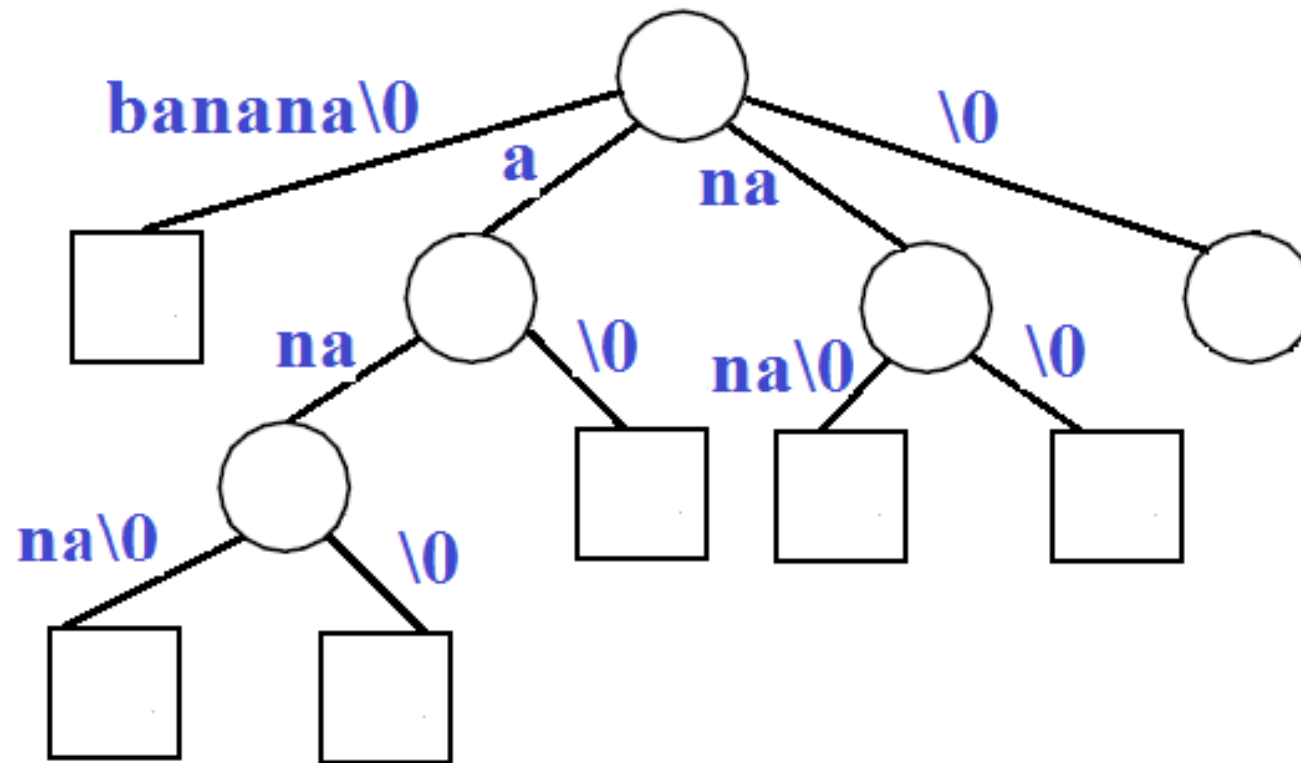
Compressed Tries

Trie



Suffix Tries

Trie



Text Compression

Huffman Coding

Text Similarity

Longest Common Subsequence

Edit Distance



Thank You !! 😊

Manisha Khattar
manisha@codingblocks.com