

Lecture:- Rabin karp

Qul Given 2 strings A and B,  $|A|=n$  and  $|B|=m$ ,  $N \leq m$   
 [lower case]  
 Count no. of substrings in B which are permutations of A

Eg:

A = "xyz"

B = "xyyzyzyzy"      ans = 3.

Brute force:

Go to all substrings of B of len = A.length() —  $O(n^2)$

cmp with string B. —  $O(n)$

if (equal) — count++

TC:  $O(n)$

SC:  $O(1)$

Approach 2:

A = a b a c      { a:2    b:1    c:1 }

B = b b c a a c b b c b a c a b  
 w1 w2 w3 w4 w5

s	e	freqB	Is match?
0	3	b: 2 c: 1 a: 1	No
1	4	b: <del>2</del> 1 c: 1 a: <del>1</del> 2	Yes — cnt++
2	5	b: <del>1</del> 0 c: <del>1</del> 2 a: 2	No
⋮	⋮	⋮	⋮

Code:

```
int cntPermutations(string A, string B) {
```

```
    n = A.length();
```

```
    m = B.length();
```

```
    freqA[26], freqB[26];
```

$O(n)$  ——— { for ( $i=0$ ;  $i < n$ ;  $i++$ ) {  
                    freqA[A.charAt(i) - 'a'] ++;  
                    }

$O(m)$  ——— for ( $i=0$ ;  $i < m$ ;  $i++$ ) {  
                    freqB[B.charAt(i) - 'a'] ++;  
                    if ( $i \geq A.length()$ ) {  
                        freqB[B.charAt(i - A.length()) - 'a'] --;  
                    }

$O(1) \Rightarrow O(26)$  ——— if (cmp(freqA, freqB)) {

```
    cnt ++;
```

```
    }
```

```
}
```

```
return cnt;
```

```
}
```

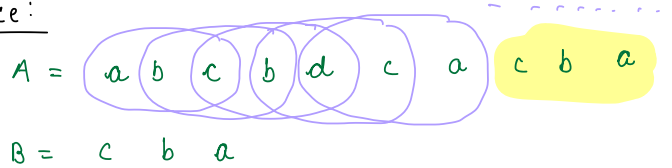
TC:  $O(n+m)$

SC:  $O(1)$

Q2: Given a string A and string B.  $|A| = n$   $|B| = m$   
 find count of substrings of A equal to B.

Eg: A = a b c b d c a c b a      ans = 1.  
 B = c b a

Brute force:



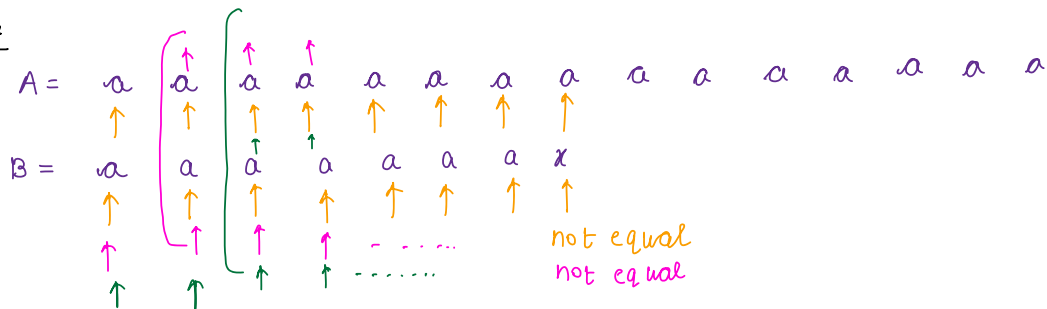
✓ substrings of len (m) of string A.

check if equal to B —

if (true) — cnt++.

TC:  $(n-m+1) * m = O(n-m+1) * O(m) \approx O(n*m)$   
 SC:  $O(1)$

Issue:



cmp two integers =  $O(1)$

" " strings =  $O(n)$

↑  
 idea of Rabin karp

## Robin Karp

Idea 1:

$$h(abc) = \text{sum of all characters} \\ = 1 + 2 + 3 = 6$$

$$h(adc) = 1 + 4 + 3 = 8$$

$$A = \begin{array}{cccc} & \underline{6} & & \\ a & b & c & a & b & a & c \end{array}$$

$$B = \begin{array}{ccc} a & b & a \\ \underline{4} & & \end{array}$$

$$\text{if } [h(s1) \neq h(s2)] \rightarrow s1 \neq s2. \quad [100\%]$$

$$\text{if } [h(s1) == h(s2)] \rightarrow s1 == s2 \quad [100\% \text{ sure}] \times \text{collision.}$$

$$\text{Eg: } h(aba) = 4 \\ h(aab) = 4.$$

a:1  
b:2  
c:3  
d:4  
e:5  
:  
:  
:  
f:26

Idea 2:

$$326 = 3 * 10^2 + 2 * 10^1 + 6 * 10^0$$

$$23 = 2 * 10^1 + 3 * 10^0$$

$$326 = 3 + 2 + 6 = 11$$

$$263 = 2 + 6 + 3 = 11$$

$$h(abc) = 1 * p^2 + 2 * p^1 + \underset{(3)}{c} * p^0$$

$$h(bca) = 2 * \underbrace{p^2}_{\text{base}} + 3 * p^1 + 1 * p^0$$

base [Generally  $p = 29$ ]

for class,  $p = 10$

$$A = \begin{array}{ccccccc} & & w1 & & & & \\ & & \underline{a \quad b \quad c} & & a & b & a \quad c \end{array}$$

$$B = \begin{array}{ccccccc} & & & w2 & & & \\ & & & \underline{a \quad b \quad a} & & & \\ & & & & w3 & & \end{array}$$

$$h(aba) = 1 * 10^2 + 2 * 10^1 + 1 * 10^0 \\ = 100 + 20 + 1 = 121.$$

$$w1[abc] = 1 * 10^2 + 2 * 10^1 + 3 * 10^0 = 123.$$

$$w2[bca] = 2 * 10^2 + 3 * 10^1 + 1 * 10^0 = 231$$

derive value of  $w2$  from  $w1$ .

$$w1 = 123.$$

$$w1 = a \quad b \quad c$$

$$w2 = b \quad c \quad a$$

$$\longrightarrow \left[ \underset{\substack{\text{char removed} \\ \uparrow \\ \text{pos}}}{123 - 1 * 10^2} \right] * 10 + \underset{\substack{\text{new char added} \\ \uparrow \\ \text{pos}}}{1 * 10^0} \quad \xrightarrow{\text{h/w.}}$$

$$\left. \begin{array}{l} \text{if } (h(s1) \neq h(s2)) \longrightarrow s1 \neq s2 \\ \text{if } (h(s1) == h(s2)) \longrightarrow s1 == s2 \end{array} \right\} 100\% \text{ sure.}$$

If len of string is large  $\rightarrow 10^5$ .

$$str = a b c \dots \dots \dots$$

$$h(str) = 1 * P^{10^5} \longrightarrow \text{overflow.}$$

$$h(str) = \left( \_ * p^- + \_ * p^- + \_ * p^- \right) \% \text{ mod.}$$

↑  
collisions

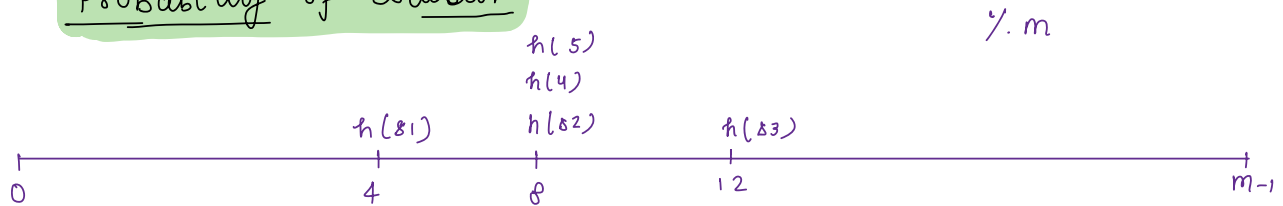
$$\text{generally } \text{mod} = 10^9 + 7$$

## Rolling hash function

$$h(\text{str}) = \left( \sum_{i=0}^{n-1} \text{str}[i] * p^{n-i-1} \right) \% m.$$

$$h(\overset{2}{a}\overset{1}{b}\overset{0}{c}) = a * p^2 + b * p^1 + c * p^0$$

## Probability of collision



Strings

1

2

3

4

⋮

n.

Prob of collision

0

$\frac{1}{m.}$

$\frac{2}{m.}$

$\frac{3}{m.}$

⋮

$\frac{n-1}{m.}$

Generally

$$n = 10^5$$

$$m = 10^9 + 7$$

$$\frac{10^5 - 1}{10^9 + 7} \approx 0.0001$$

Break: 8:45 AM

Best TC:

A = a b c d e f g h i j k l m n o p  
          <sup>h2</sup>       <sup>h4</sup>  
          <sup>h3</sup>  
B = x y z u v  
          <sup>h1</sup>

TC:  $O(n-m+1) = O(n) \rightarrow$  No matches.

Worst TC:

a = x x x x x x x x x x x x x x x x x  
          <sup>h2</sup>  
b = x x x x x x x x  
          <sup>h1</sup>

if ( $h1 == h2$ ) —  $s1 == s2$  [100% sure] X  
          ↑  
          char by char comparison.

TC:  $O(n \times m)$

SC:  $O(1)$



Ques Given stream of characters, check if palindrome.

Eg: str: a b c b a op: a  
 abcba

←  $h_2$   
 m a d a m  
 →  $h_1$  ∴  $h_1 = h_2$

Eg: <sup>0 1 2 3 4</sup>  
 a b c b a

↖ forward hash

$$1. \quad a \quad \therefore \quad \left. \begin{array}{l} fh = a * p^0 \\ bh = a * p^0 \end{array} \right\}$$

$$fh = bh$$

↑  
 99.9999% sure of palindrome  
 Do char by char comparisons

$$2. \quad \begin{array}{l} \underline{ab} \rightarrow fh = a * p^1 + b * p^0 \\ \underline{ba} \rightarrow bh = b * p^1 + a * p^0 \end{array} \quad [fh \neq bh]$$

$$3. \quad abc \quad fh = [a * p^1 + b * p^0] * p + \text{new char}$$

$$\underline{\text{final ans:}} \quad a * p^2 + b * p^1 + c * p^0$$

$$\underline{cba} \rightarrow bh = b * p^1 + a * p^0 + \text{newchar} * p^{\text{index}}$$

$$\underline{\text{final ans:}} \quad c * p^2 + b * p^1 + a * p^0$$

4.  $abcba$  :-

$$bh = cp^2 + bp^1 + ap^0 + \text{new char} * p^{\text{index}}$$

$$bcba :- b * p^3 + c * p^2 + b * p^1 + a * p^0$$

```
void printPalindromes (String str) {  
    base = 29;  
    fh = 0    → forward-hash  
    bh = 0    → backward-hash  
    power = 1;  
    String ans = "";  
    mod = 109 + 7;  
    for (i = 0; i < n; i++) {  
        char ch = str.charAt(i);  
        ans = ans + ch;  
        fh = (fh * base + ch) % mod;  
        bh = bh + ch * power  
        power = power * 29;  
        if (fh == bh) {  
            if (isPalindrome (str.substring(0, i+1))) {  
                print (ans);  
            }  
        }  
    }  
}
```

TC:  $O(n * m)$

SC:  $O(1)$

Thankyou ☺