# Rabin-Karp algorithm in Java → The number of equal substring pairs

**5** users solved this problem. Latest completion was **22** days ago.

Hard (9 43 minutes (2) Share:

#### You've skipped this problem 1 day ago.

Maybe now is a good time to try it out again?

Code Challenge — Write a program

Given a string s and a set of pairs, where each element represents the starting and the ending position of a substring of s. Write a program that counts the number of pairs of equal substrings.

**Input:** the first line contains a string s. The second line contains an integer t. Each of the following t lines contains 4 integers i, j, k and m separated by space, such that  $0 \le i < j \le |s|, 0 \le k < m \le |s|$ .

**Output:** the number of pairs, such that s[i:j] is equal to s[k:m].

In this problem for a polynomial hash with constants a=53 and  $m=10^9+9$  it is guaranteed that if hash values for two substrings are equal, then the strings are equal as well.

**Hint 1:** calculate hash values of all prefixes of s and then use them to find a hash value for any substring of s.

**Hint 2:** for a substring s[i:j] a polynomial hash is calculated as follows:

$$hash(s[i:j]) = \left(\sum_{k=i}^{j} s[k] \cdot a^{k-i}
ight) \; mod \; m.$$

Multiplying the equality by  $a^i$  we will get the following:

$$hash(s[i:j]) \cdot a^i = \left(\sum_{k=i}^j s[k] \cdot a^k
ight) \; mod \, m = \left(\left(\sum_{k=0}^j s[k] \cdot a^k
ight) - \left(\sum_{k=0}^{i-1} s[k] \cdot a^k
ight)
ight) \; mod \, m.$$

That is,

$$hash(s[i:j])\cdot a^i=(hash(s[0:j])-hash(s[0:i-1])) \,\,mod\,m.$$

So, if we know hash values for all prefixes of s, we can calculate a hash value for any substring of s. The only problem is that we get a hash value multiplied by  $a^i$ . Then, if we want to compare hash values of s[i:i'] and s[j:j'] (assuming  $i \leq j$ ) we need to multiply the first one by  $a^{j-i}$ .

#### Sample Input 1:

# Sample Output 1:

1

## Sample Input 2:

seamlessness 3 4 8 8 12 5 8 9 12 0 1 11 12

# Sample Output 2:

2

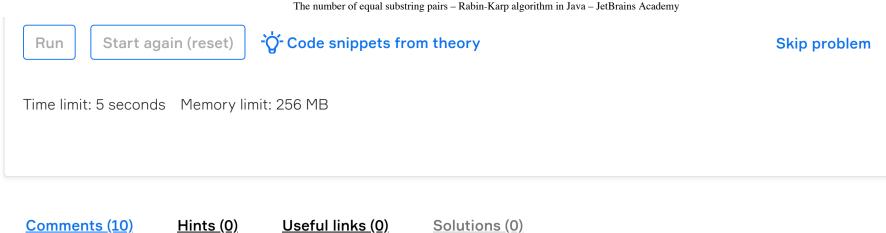
# Code Editor

<u>IDE</u>

Java

public class Main {
 // your code here
}

https://hyperskill.org/learn/step/5723



Share something, Sergey Kubatko

Sort by: Last posted ▼

#### Sergey Kubatko 1 day ago

spent two days, algo is right, cannot pass test#6 as Scanner reads only 986895 chars instead of 2M, pls give a clue how it can be solved

0 Reply

#### AN Adrian Nachev 4 months ago Report

Would be nice to say in the downloaded tests what is the expected result and what was the actual result. Otherwise a test with a massive string is completely pointless.

O Reply

#### AS **Andrey Shinkaryov** 4 months ago Report

16mb string) Got correct answer on computer, but runtime error here.

O Reply

#### JoyceS 5 months ago Report

1. we need clarification on including/excluding the indexes

in all equations of the hints, substring s[i...j] means from index i to index j inclusive. however, in the example input, the right index is given as exclusive.

2. the modulo operations are more complicated than simple multiplying ref: https://en.wikipedia.org/wiki/Modulo\_operation#Equivalences  $(a*b) \mod m = ((a \mod m)*(b \mod m)) \mod m$ 

3. test case #6 is a HUGE one, maybe we need to figure out a smarter way for all the calculations and storage? would be much more helpful if we can have an expected output

○ 0 Reply

# LA **LAURENT APICELLA** 5 months ago Report

In all the equalities with a sigma sum sign, it is j-1 instead of j, and in the last equality:

 $hash(s[i..j]) \cdot a i = (hash(s[0..j]) - hash(s[0..i-1])) \mod m$ .

it is i instead of i-1

○ 0 Reply

#### andioz 5 months ago Report

Ok, I think I am struggling with this: you corrected the definition with Timurs comment, now j means exclusive index, right? But for me the formulas still include j! Am i correct?

0 Reply

#### andioz 5 months ago Report

I don't understand the last formula, maybe a bug? "hash(s[i..i]) ai = (hash(s[0..i]) - hash(s[0..i-1])) mod m" - but "i" can start at 0, so "i - 1" -> -1 sound wrong to me?

○ 0 Reply

andioz 5 months ago (Fixed)

https://hyperskill.org/learn/step/5723 2/3 k has two meanings in explanation. use different letters.

 $\bigcirc$  0

Timur Ishakov 5 months ago Fixed

\* do not

 $\bigcirc$  0

Timur Ishakov 5 months ago Fixed

Samples do match input definition in task. Task states that  $0 \le i < j < |s|, 0 \le k < m < |s|$ , but in sample 1 - |s| = 8, and m is 8 in first

O Show all

https://hyperskill.org/learn/step/5723