

Strings

Algoritmo de Rabin-Karp

Prof. Edson Alves - UnB/FGA

2019

1. Algoritmo de Rabin-Karp

Algoritmo de Rabin-Karp

Definição

- O algoritmo de Rabin-Karp é um algoritmo que contabiliza o número de ocorrências da string P , de tamanho m , na string S , de tamanho n
- Ele foi proposto por Michael O. Rabin e Richard M. Karp em 1987
- A ideia principal do algoritmo é computar o *hash* $h_P = h(P)$ e compará-lo com todas as substrings $h_{ij} = S[i..j]$ de S de tamanho m
- Caso $h_P \neq h_{ij}$, segue que $P \neq S[i..j]$ e o algoritmo pode prosseguir
- Se $h_P = h_{ij}$, as strings ou são iguais ou houve uma colisão
- Esta dúvida pode ser sanada através da comparação direta, enquanto strings, entre $S[i..j]$ e P
- O algoritmo tem complexidade $O(mn)$ no pior caso, por conta do custo do cálculo dos *hashes* e das possíveis comparações diretas entre as strings

Pseudocódigo do algoritmo de Rabin-Karp

Algoritmo 1 Algoritmo de Rabin-Karp – Naive

Input: Duas strings P e S

Output: O número de ocorrências occ de P em S

```
1: function RABINKARP( $P, S$ )
2:    $m \leftarrow |P|$ 
3:    $n \leftarrow |S|$ 
4:    $occ \leftarrow 0$ 
5:    $h_P \leftarrow h(P)$ 
6:   for  $i \leftarrow 1$  to  $n - m + 1$  do
7:      $h_S \leftarrow h(S[i..(i + m - 1)])$ 
8:     if  $h_S = h_P$  then
9:       if  $S[i..(i + m - 1)] = P$  then
10:         $occ \leftarrow occ + 1$ 
11:   return  $occ$ 
```

1. CP-Algorithms. [String Hashing](#), acesso em 06/08/2019.
2. **CROCHEMORE**, Maxime; **RYTTER**, Wojciech. *Jewels of Stringology: Text Algorithms*, WSPC, 2002.
3. **HALIM**, Steve; **HALIM**, Felix. *Competitive Programming 3*, Lulu, 2013.
4. Wikipédia. [Rabin-Karp algorithm](#), acesso em 08/08/2019.