

Exercício 5.2.2 - Unique substrings of length L (Algs4)

Nome: Gabriel de Russo e Carmo

N USP: 9298041

Data: 05/06/2016

Relatório sobre o exercício

Estudo dos dígitos de pi: Como sugerido, analisei os 10^6 primeiros dígitos de pi, assim como os 10^7 primeiros. Os maiores valores de L encontrados foram 4 e 5, respectivamente. No menor teste, o programa rodou em aproximadamente 2.5s. No maior, em aproximadamente 28.3s. Para 10^9 eu não consegui realizar o teste pois a quantidade de substrings únicas começa a ficar muito grande e meu computador não tem memória suficiente para guardar tanta informação junto com o input. Além disso, o tempo de execução é muito alto.

Maior valor de N calculado: O maior valor de N que eu consegui calcular está na ordem de 10^8 . Para este teste, o resultado obtido foi L = 6. Segue uma tabela com os valores de N testados (testado num i5 @ 1.4ghz).

10^x	2	3	4	5	6	7	8
L/tempo	1/0.18s	2/0.16s	3/0.19s	4/0.36s	4/1.84s	5/28.5s	6/7:57m

Comparação com outras estruturas: Para realizar a comparação com outras estruturas, optei por rodar o programa num teste que gera uma sequência aleatória para N = 10^8 (testado num i7 @ 3.7 ghz). Seguem os resultados (sempre os mesmos) e a média dos tempos:

HashMap	L = 6. Tempo = 3:45m
TST	L = 6. Tempo = 1:51m
10-way Trie	L = 6. Tempo = 1:29m
ARNE	L = ?. Tempo > 10 min.

OBS: A RedBlack esquerdista se mostrou completamente inviável, demorando mais de 10 minutos para realizar o primeiro teste.

Dessa forma, podemos concluir que uma 10-way se mostrou mais eficiente do que uma TST. Mesmo assim, a TST ainda é muito melhor do que um Hash ou uma ARNE.