

MAC 5710 - Estruturas de dados e suas aplicações
Primeiro semestre de 2009

Quarto Exercício-Programa – Devolução: 1 de julho

Hashing × skip lists

O objetivo deste exercício-programa é fazer um estudo comparativo parcial de hashing universal e de skip lists.

Seu programa deve funcionar em dois modos: o modo *teste* e o modo *avaliação*. O modo teste tem por objetivo a certificação de que suas implementações estão corretas. O modo avaliação deve gerar uma batelada de entradas aleatórias e medir o tempo que cada uma das duas implementações consome para processar tais entradas, imprimindo o tempo total médio de processamento de cada entrada pelas duas implementações, o espaço alocado médio, o tempo médio por operação, o tempo máximo, mínimo, etc.

Baseando-se nos dados obtidos pelo seu programa no modo avaliação, você deve escrever um relatório descrevendo o que você pode notar, e entregá-lo junto com o seu EP. Faça testes suficientes para que você possa comparar as duas estruturas em situações variadas. Seja sistemático nos testes, para que as suas conclusões sejam bem abalizadas.

MODO TESTE

Implemente no seu programa uma rotina `IMPRIME_HASHING`, que recebe uma tabela de hashing e imprime todos os elementos do conjunto armazenado nesta tabela. Similarmente, implemente uma rotina `IMPRIME_SKIPLIST`, que recebe uma skip list e imprime todos os elementos do conjunto armazenados na skip list.

No modo teste, o programa deve ler os dados de um arquivo de entrada com o formato descrito a seguir e processá-los. Um arquivo conterá vários conjuntos de teste.

Considere que o universo das chaves é o conjunto de inteiros entre de 0 e $1.073.676.289 = 32767^2$ (`INT_MAX`). Um conjunto de teste consiste de um subconjunto S do conjunto universo de chaves, dado pelo número n de elementos em S seguido pelos n elementos de S . Um valor de $n = 0$ indica o fim do arquivo. Veja abaixo um exemplo de arquivo de entrada.

```
10
4 17 3 90 105 13 60 73 36 44
20
100 83 48 91 40 3 62 88 104 200 5 128 21 39 16 95 77 69 42 51
0
```

Para cada conjunto de teste dado, seu programa deve escolher adequadamente o tamanho da tabela de hashing e o valor do $\lg N_{\max}$ para a skip list. Deve então criar uma tabela de hashing deste tamanho e uma skip list, ambas inicialmente vazias, e deve proceder inserindo, na ordem dada no arquivo, um a um cada elemento do conjunto nas duas EDs. Ao final, o seu programa deve imprimir o conjunto armazenado na sua tabela de hashing e o conjunto armazenado na skip list. A ordem dos elementos do conjunto é irrelevante, porém você deve imprimi-los no mesmo formato do arquivo de entrada (tamanho do conjunto seguido de uma lista de seus elementos separados por brancos). Essas impressões devem ser feitas em um arquivo, de modo que o monitor possa usar este arquivo para verificar se sua implementação está correta.

MODO AVALIAÇÃO

No modo de avaliação, o seu programa deve testar vários tamanhos de conjunto. Para cada tamanho, deve gerar aleatoriamente vários subconjuntos do conjunto universo daquele tamanho. Para cada um destes, deve inserir os seus elementos um a um numa tabela de hashing (de tamanho adequado) inicialmente vazia, e numa skip list inicialmente vazia, cronometrando o tempo para cada uma das duas estruturas, para poder, ao final imprimir as estatísticas que usará na confecção do seu relatório.

Lembre-se que o universo das chaves é o conjunto de inteiros entre de 0 e $1.073.676.289$. Use adequadamente o gerador de números aleatórios para gerar inteiros neste intervalo com probabilidade uniforme.

RELATÓRIO

Faça testes variados e sistemáticos para que a comparação descrita no seu relatório seja significativa e rica. Considere conjuntos grandes comparados com o tamanho do universo, e conjuntos pequenos. Leve em consideração na sua avaliação também a quantidade de espaço das duas EDs.