

Lista de Exercícios AEDII – 02 – Profº Helton

BoaNoche Development Group™

1. Defina colisão e dê exemplos e cite duas estratégias comumente utilizadas para lidar com colisão.

Quando duas chaves A e B recebem uma mesma saída para uma função hash. Por exemplo, na função hash $h = k \% 10$, as chaves 10 e 20 terão a mesma saída, o zero. Duas estratégias comumente utilizadas para lidar com a colisão são o uso de buckets, que armazena mais de uma chave em um nó, ou o método “predictor”, que diz quantos rehashs serão necessários para inserir a chave em uma posição válida.

2. O que é uma função de hash? Que propriedades deve possuir uma boa função de hash?

Uma função hash é uma função que recebe uma chave, faz um processamento com a mesma e retorna uma posição da tabela em que a dada chave deve ser inserida. Uma boa função hash deve depender de todos os bits da chave e ser sensível a permutações.

3. O que é agrupamento primário? E secundário? Por que ambos devem ser evitados?

Agrupamento primário é o aumento da probabilidade de uma posição na tabela ser preenchida. E o secundário, é quando duas chaves colidem numa mesma posição sigam colidindo por um “caminho” de rehash até serem inseridas. Ambos devem ser evitados, pois diminuem a eficácia da busca utilizando tabelas hash.

4. Descreva as técnicas de resolução de colisões por rehashing linear e por rehashing duplo.

A técnica de Hash Linear consiste em associar o número de vezes que a chave foi avaliada, ou seja $reHash(i,j) = (i+j) \% TS$, onde i refere-se a chave gerada anteriormente e j ao número de vezes que foi utilizada a função rehash.

A técnica de HashDuplo, trata-se de uma técnica que utiliza duas funções hash, $hash1()$ e $hash2()$, onde o rehash é dado por: $rehash(j) = (j + hash2(c)) \% TS$, onde “c” vem do $hash1$,

$hash1(c) = i \% TS$

5. Implemente o método de dobra(incluindo a transformação da chave em número binário) para o cálculo de uma função hash. Assuma que o vetor tem 1024 posições.

```
int Dobra(int key){
/*
    Créditos ao Henrique Leme
*/
int n = (int)(log((double)%TS)/log(2.0))+1;
int p = pow(2,n)-1;
int d = 0;
int s = sizeof(int)*8;
while (s>0) {
    d=d^(key&p);
    key=key>>n;
    s=s-n;
}
```

```
    return ((int)d)%TS;
}
```

6. Suponha tem-se um conjunto de chaves onde cada chave possui seis dígitos. Implemente uma função de hash que consiste em extrair o 1o, 3o e 4o dígitos de cada chave. Comente a eficiência desta função.

```
int transforma(int k){
    char test[100];
    sprintf(test,"%d",k);

    int result = 0;
    char x1 = test[0];
    char x2 = test[1];
    char x3 = test[2];

    result += test[0] -48;
    result += test[2] -48;
    result += test[3] -48;
    return result;
}
```

7. Utilizando o método de divisão inteira e o rehashing linear para resolver colisões, armazene as chaves: 673686, 412368, 643686, 420435, 643728, 486435, 433401, 598935, 703602 em um vetor de 19 posições. Quantas colisões ocorreram?

Ocorreram 2 colisões

8. Combine o método de extração de dígitos da questão 5 com a divisão inteira, extraíndo os dígitos 1 , 3 e 4 de cada chave e em seguida considerando o resto da divisão por 19.

Use as mesmas chaves do exercício 7 e compare o número de colisões.

Ocorreram 2 colisões

9. Explique o conceito de Bucket (cesto) e diga onde ele é usado.

É uma maneira de implementação de hash que utiliza de cestos para armazenar mais de uma chave em uma posição da tabela.

10. Diferencie as técnicas de hashing extensível e hashing linear.

Ambos utilizam dos buckets, a diferença está na duplicação do diretório. Na técnica de hashing extensível o diretório é duplicado em um único passo, já no hashing linear os buckets vão sendo duplicados gradualmente. Outra diferença é que no extensível a função de hash não varia. Já no linear, há variação da função hash.

11. Para que serve a área extra quando se usa hashing linear (também conhecida como área de overflow)?

A área de overflow é utilizada quando uma chave é inserida em um bucket cheio, e não é a vez dele se duplicar. Dessa maneira, um novo bucket é ligado por um ponteiro a esse bucket já cheio, e a chave é inserida nesse novo cesto, que é a tal área extra.