

Estruturas de Dados II
Arquivos Indexados por Chaves Secundárias

Mário Henrique de Paiva Perché

25 de Novembro de 2010

1. Quais as principais diferenças entre índices primário e secundário?
2. Por que podemos ter no máximo um índice primário, mas diversos índices secundários para um arquivo de dados?
3. Considere um disco com tamanho de bloco $B = 512$ bytes. Um ponteiro de bloco é do tamanho $P = 6$ bytes, e um ponteiro de registro é do tamanho $Pr = 7$ bytes. Um arquivo b possui $r = 30.000$ registros de EMPREGADO de tamanho fixo. Cada registro possui os seguintes campos: NOME (30 bytes), SSN (9 bytes), CODIGODEPARTAMENTO (9 bytes), ENDEREÇO (40 bytes), TELEFONE (9 bytes), DATANASC (8 bytes), SEXO (1 bytes), CODIGOCARGO(4 bytes), SALARIO (4 bytes, número real). Um byte adicional é usado como marcador de exclusão.
 - (a) Calcule o tamanho do registro R em bytes.
 - (b) Quantos registros “cabem” em um bloco?
 - (c) Calcule o número de blocos do arquivo b .
4. Considere-se um arquivo tal como o que se segue, e que se deseja utilizar como entrada para obtenção de um arquivo multilista com chaves de acesso ENDEREÇO e PERÍODO. Pede-se esboçar a aplicação do algoritmo de Lefkowitz, sendo NÚMERO a chave primária.

Exemplo de Multilista

A1

CP	INC	CA1	CA2
NÚMERO	NOME	ENDEREÇO	PERÍODO
15	João	MG	1
18	José	SP	3
25	Maria	RJ	3
37	Gérson	MG	3
41	Sandra	SP	1
43	Miguel	MG	1
45	Ana	SP	1

Para o arquivo acima a seguinte notação é adotada:

- ED = endereço de disco. Numeração dos registros em relação ao início do arquivo
- CP = chave primária dos registros do arquivo.
- CA = chave de acesso. Atributo chave secundária que criou o multilista.
- CL = comprimento de lista. Número de registros com o mesmo valor de chave de acesso.
- EC = endereço de cabeça. ED do primeiro registro com um dado valor de chave de acesso.
- INC = itens não chave. Atributos de registro que não influem no processamento de multilista.
- P = ponteiro para o sucessor na multilista da respectiva chave de acesso