



1 Introdução e objetivos

Este conjunto de atividades visa continuar a trabalhar o conceito de indexação e a compreensão de sua dinâmica.

2 Preparação

A preparação para esta atividade envolve as leituras indicadas:

- Folk e Zoellick (1992), Cap. 6, Seção 6.7
- Moreira (2011), Unidade 5, a partir da seção 5.3.3

3 Atividades

Cada atividade proposta pode ser resolvida individualmente ou em grupos.

Exercício 1.

Considere os dados da Figura 1 e apresente como um índice secundário com listas invertidas pode ser estruturado. Use a *sigla do curso* como chave. Note também que o RA pode ser usado como chave primária para o arquivo.

Exercício 2.

Usando o arquivo e a indexação elaborada no Exercício 1, execute as ações indicadas, na ordem apresentada. Mostre, mostre as modificações no arquivo e na indexação.

1. Faça a inserção do registro (984784, Maria Isis Gois Xisco, EAgr)
2. Faça a remoção (lógica) dos registros:
 - RA 984454 (João Andrade Guterres)
 - RA 984839 (Arthur Henrique Batista)
3. Faça a inserção do registro (985855, Bianca Gomes Botelho, EnfP)

Questão 1.

Defina o termo *bloco*. Há dúvidas sobre seu papel no armazenamento de arquivos?

Exercício 3.

Um arquivo tem 244.880 bytes de tamanho. Sabendo-se que ele está armazenado em um disco rígido cujos blocos têm tamanho 4KiB, qual o espaço em disco que tal arquivo efetivamente ocupa?

Figura 1: Arquivo de dados com campos e registros de comprimentos fixos. Os dados são, nesta ordem, *registro acadêmico* (6 bytes), *nome* (32 bytes) e *sigla de curso* (6 bytes). A coluna à esquerda indica o deslocamento em relação ao início do arquivo e as sublinhas indicam espaços.

0000	9 8 3 9 8 2	N o a h _ C a e t a n o _ V i l
0022	e l a _ _ _ _ _ _ _ _ _ _	E A g - A r
0044	9 8 3 6 2 2	M a y a _ G o d i n h o _ R a m
0066	o s _ _ _ _ _ _ _ _ _ _	P e d L N _
0088	9 8 3 7 4 0	J o a n a _ G o d i n h o _ _ _
0110	_ _ _ _ _ _ _ _ _ _ _ _	P e d L N _
0132	9 8 4 4 5 4	J o a o _ A n d r a d e _ G u t
0154	e r r e s _ _ _ _ _ _ _ _	F i l B _ _
0176	9 8 3 1 7 3	M a r i n a _ R i b e i r o _ N
0198	o v a i s _ _ _ _ _ _ _ _	F i l B _ _
0220	9 8 5 8 4 6	A l i c e _ B r i t e s _ G u e
0242	d e s _ _ _ _ _ _ _ _ _ _	A d m - L S
0264	9 8 2 4 1 0	J o a o _ G u i l h e r m e _ M
0286	e n e z e s _ C a p e l o _ _ _	A d m - L S
0308	9 8 1 9 8 5	B i a n c a _ G a l v a o _ D r
0330	u m o n d _ _ _ _ _ _ _ _	E A g - A r
0352	9 8 5 9 2 4	M a r i a _ L a u r a _ H e r n
0374	a n d e s _ O l i v e i r a _ _ _	E A g - A r
0396	9 8 4 5 5 8	A n a _ S o p h i a _ V i l e l
0418	a _ A l e n c a s t r o _ _ _ _	F i l B _ _
0440	9 8 3 2 2 1	C a r l o s _ E d u a r d o _ G
0462	o u l a r t _ _ _ _ _ _ _ _	A d m - L S
0484	9 8 3 9 8 0	C e c i l i a _ C a v a l c a n
0506	t e _ _ _ _ _ _ _ _ _ _	P e d L N _
0528	9 8 4 6 7 0	A n a _ B e a t r i z _ C a m a
0550	c h o _ M o d e s t o _ _ _ _	A d m - L S
0572	9 8 4 8 3 9	A r t h u r _ H e n r i q u e _
0594	B a p t i s t a _ _ _ _ _ _	E n f _ _ _
0616	9 8 5 6 1 1	M a r i a _ V i t o r i a _ G o
0638	n z a g a _ Q u a r t e i r a _	P e d L _ _
0660	9 8 2 0 7 3	A m a n d a _ A g u i a r _ M a
0682	l d o n a d o _ _ _ _ _ _ _	P e d L _ _

Informações

Considere as seguintes informações:

- Os blocos são de 1024 bytes;
- Um registro de dados tem comprimento fixo de 146 bytes;
- Uma entrada de índice (chave e ponteiro) possuem 18 bytes;
- Há controle do uso dos blocos, de forma que **registros de dados** e **entradas de índice** não sejam “quebrados” entre blocos;
- Existem no arquivo de dados 430.000 registros.

Exercício 4.

Considere que o arquivo de dados esteja ordenado por uma dada chave e que haja um índice primário esparsa para esse arquivo. Pergunta-se:

- Qual o tamanho da fragmentação mínima de cada bloco no arquivo de dados?
- Quantos blocos ocupa o arquivo que contém o índice?
- Qual o espaço em disco que ele ocupa?

Exercício 5.

Considere que o arquivo de dados não possua nenhum tipo de ordenação e que haja um índice secundário usando uma chave primária. Pergunta-se:

- Quantos blocos ocupa o arquivo que contém o índice?

Informações

Considere um arquivo ordenado por uma dada chave e que tenha registros de tamanho fixo. Para a realização de uma inserção ou de uma remoção de registro é preciso, em princípio, fazer o deslocamento de todos os registros com chave maior que a dele uma posição em direção ao fim ou ao início do arquivo. O custo dessa operação é $O(n)$, uma vez que, em média, seriam deslocados $\frac{n}{2}$ registros.

Suponha que esse arquivo esteja organizado usando os blocos para acolher os registros de forma organizada.

Questão 2.

Discuta como uma remoção neste arquivo pode ser apenas lógica, eliminando a necessidade de deslocamentos.

Questão 3.

Discuta como uma inserção neste arquivo pode usar o espaço de uma remoção lógica caso haja espaço ocioso no bloco onde deve ocorrer a inserção.

Questão 4.

Discuta possíveis alternativas para o cenário da Questão 3 quando não há espaço no bloco. Em particular, veja as consequências dos deslocamentos e o uso de um **bloco de estouro** (*overflow block*).

4 Encerramento

Use o fórum de dúvidas do AVA para postar dúvidas ou comentários que tiver sobre esta atividade.

Referências

FOLK, M.; ZOELLICK, B. *File structures*. USA: Addison-Wesley Publishing Company, Inc., 1992.

MOREIRA, J. *Armazenamento e recuperação da informação*. São Carlos: UFScar, 2011. Disponível em: http://audiovisual.uab.ufscar.br/impresso/2016/SI/SI_Jander_RecuperacaoInformacao.pdf.