



1 Introdução e objetivos

Este conjunto de atividades visa 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-Indexing
 - Até seção 6.5.
- Guliato (2004)
 - Até seção 6.3.1.
- Moreira (2011), Unidade 5-Organização de dados em memória secundária
 - Seção 5.3.4.

3 Atividades

Cada atividade proposta pode ser resolvida individualmente ou em grupos. Os arquivos das Figuras 1 e 2 contêm uma versão editada da lista dos mais ricos do mundo, segundo Época Negócios¹.

Exercício 1.

Construa um índice para o arquivo da Figura 1, usando o nome como chave. Assuma que nome seja uma chave primária neste caso.

Exercício 2.

Construa um índice para o arquivo da Figura 2, usando o nome como chave. Assuma que nome seja uma chave primária neste caso.

Questão 1.

Para ambos os Exercícios 1 e 2, poderiam ser usados o RNN (*relative record number*) ou o deslocamento (*offset*) para indicar a posição. Concorda?

Exercício 3.

Construa um índice esparsa para o arquivo da Figura 1, usando o nome como chave. Assuma que nome seja uma chave primária neste caso e que caibam cinco registros por bloco.

Exercício 4.

Descreva a busca no índice gerado no Exercício 3 para a chave nome com valor “Dietrich Mateschitz”.

Questão 2.

O índice do Exercício 3 é agrupado? Justifique.

Exercício 5.

Construa um índice para o arquivo da Figura 3 usando nome como chave. Considere que nome seja uma chave primária.

Descreva passo a passo as alterações feitas no arquivo e no índice para as seguintes operações, nesta ordem:

- Localização do registro “Bill Gates”
- Remoção do registro “Bill Gates”
- Localização do registro “Charles Koch”
- Remoção do registro “Charles Koch”
- Inserção de um novo registro^a (Francois Pinaut, 27, Kering)

^aAdote uma opção de reuso arbitrária.

Exercício 6.

Para o Exercício 5, calcule as porcentagens da fragmentação interna da externa depois das manipulações solicitadas.

Exercício 7.

Construa um índice para o arquivo da Figura 3 usando como chave a *atuação*. Use repetições de entradas no índice para acomodar a chave secundária.

Questão 3.

Segundo a definição de índice primário adotada na disciplina, é possível criar esse tipo de índice para o arquivo da Figura 3? Estabeleça seus argumentos.

Figura 1: Arquivo de dados com campos e registros de comprimentos fixos. Os dados são, nesta ordem, *nome* (21 bytes), *fortuna* em bilhões de dólares (5 bytes) e *atuação* (16 bytes). A coluna à esquerda indica o deslocamento em relação ao início do arquivo e as sublinhas indicam espaços.

```
0000 A l i c e _ W a l t o n _ _ _ _ _
0021 4 6 _ _ _ W a l - M a r t _ _ _ _ _
0042 A m a n c i o _ O r t e g a _ _ _ _ _
0063 7 0 _ _ _ Z a r a _ _ _ _ _
0084 B e r n a r d _ A r n a u l t _ _ _ _ _
0105 7 2 _ _ _ L V M H _ _ _ _ _
0126 C a r l o s _ S l i m _ H e l u _ _ _ _ _
0147 6 7 . 1 _ t e l e c o m u n i c a c o e s
0168 C h a r l e s _ K o c h _ _ _ _ _
0189 6 0 _ _ _ d i v e r s i f i c a d o _ _ _
0210 D a v i d _ K o c h _ _ _ _ _
0231 6 0 _ _ _ d i v e r s i f i c a d o _ _ _
0252 D a v i d _ T h o m s o n _ _ _ _ _
0273 2 5 _ _ _ T h o m s o n _ R e u t e r s _
0294 D i e t r i c h _ M a t e s c h i t z _ _
0315 2 3 _ _ _ R e d _ B u l l _ _ _ _ _
0336 F r a n c o i s e _ B e t t e n c o u r t
0357 4 2 . 4 _ L ' O r e a l _ _ _ _ _
0378 F r a n c o i s _ P i n a u l t _ _ _ _ _
0399 2 7 _ _ _ K e r i n g _ _ _ _ _
0420 G i o v a n n i _ F e r r e r o _ _ _ _ _
0441 2 3 _ _ _ d o c e s _ _ _ _ _
0462 H u i _ K a _ Y a n _ _ _ _ _
0483 3 0 . 3 _ C h i n a _ E v e r g r a n d e
0504 J a c q u e l i n e _ M a r s _ _ _ _ _
0525 2 3 . 6 _ d o c e s _ _ _ _ _
0546 J e f f _ B e z o s _ _ _ _ _
0567 1 1 2 _ _ A m a z o n _ _ _ _ _
0588 J o h n _ M a r s _ _ _ _ _
0609 2 3 . 6 _ d o c e s _ _ _ _ _
0630 J o s e p h _ S a f r a _ _ _ _ _
0651 2 3 . 5 _ b a n c o _ _ _ _ _
0672 L a r r y _ E l l i s o n _ _ _ _ _
0693 5 8 . 5 _ s o f t w a r e _ _ _ _ _
0714 M a _ H u a t e n g _ _ _ _ _
0735 4 5 . 3 _ T e n c e n t _ _ _ _ _
0756 M a r k _ Z u c k e r b e r g _ _ _ _ _
0777 7 1 _ _ _ F a c e b o o k _ _ _ _ _
0798 M i c h a e l _ B l o o m b e r g _ _ _
0819 5 0 _ _ _ B l o o m b e r g _ L P _ _ _
0840 M i c h a e l _ D e l l _ _ _ _ _
0861 2 2 . 7 _ D e l l _ _ _ _ _
0882 P h i l _ K n i g h t _ _ _ _ _
0903 2 9 . 6 _ N i k e _ _ _ _ _
0924 R o b s o n _ W a l t o n _ _ _ _ _
0945 4 6 . 2 _ W a l - M a r t _ _ _ _ _
0966 S c h a e f f l e r _ _ _ _ _
0987 2 5 . 3 _ S c h a e f f l e r _ _ _ _ _
1008 S e r g e y _ B r i n _ _ _ _ _
1029 4 7 . 5 _ G o o g l e _ _ _ _ _
1050 S t e v e _ B a l l m e r _ _ _ _ _
1071 3 8 . 4 _ M i c r o s o f t _ _ _ _ _
1092 S u s a n n e _ K l a t t e n _ _ _ _ _
1113 2 5 _ _ _ B M W _ _ _ _ _
1134 W a n g _ J i a n l i n _ _ _ _ _
1155 3 0 _ _ _ i m o v e i s _ _ _ _ _
```

Figura 2: Arquivo de dados com campos e registros de tamanhos variáveis, usando-se | para terminar campos e # para terminar registros. Os dados são, nesta ordem, *nome*, *fortuna* em bilhões de dólares e *atuação*. A coluna à esquerda indica o deslocamento em relação ao início do arquivo e as sublinhas indicam espaços.

```
0000 A l i c e _ W a l t o n | 4 6 | W a l -
0020 M a r t | # A m a n c i o _ O r t e g a
0040 | 7 0 | Z a r a | # B e r n a r d _ A r
0060 n a u l t | 7 2 | L V M H | # C a r l o
0080 s _ S l i m _ H e l u | 6 7 . 1 | t e l
0100 e c o m u n i c a c o e s | # C h a r l
0120 e s _ K o c h | 6 0 | d i v e r s i f i
0140 c a d o | # D a v i d _ K o c h | 6 0 |
0160 d i v e r s i f i c a d o | # D a v i d
0180 _ T h o m s o n | 2 5 | T h o m s o n _
0200 R e u t e r s | # D i e t r i c h _ M a
0220 t e s c h i t z | 2 3 | R e d _ B u l l
0240 | # F r a n c o i s e _ B e t t e n c o
0260 u r t | 4 2 . 4 | L ' O r e a l | # F r
0280 a n c o i s _ P i n a u l t | 2 7 | K e
0300 r i n g | # G i o v a n n i _ F e r r e
0320 r o | 2 3 | d o c e s | # H u i _ K a _
0340 Y a n | 3 0 . 3 | C h i n a _ E v e r g
0360 r a n d e | # J a c q u e l i n e _ M a
0380 r s | 2 3 . 6 | d o c e s | # J e f f _
0400 B e z o s | 1 1 2 | A m a z o n | # J o
0420 h n _ M a r s | 2 3 . 6 | d o c e s | #
0440 J o s e p h _ S a f r a | 2 3 . 5 | b a
0460 n c o | # L a r r y _ E l l i s o n | 5
0480 8 . 5 | s o f t w a r e | # M a _ H u a
0500 t e n g | 4 5 . 3 | T e n c e n t | # M
0520 a r k _ Z u c k e r b e r g | 7 1 | F a
0540 c e b o o k | # M i c h a e l _ B l o o
0560 m b e r g | 5 0 | B l o o m b e r g _ L
0580 P | # M i c h a e l _ D e l l | 2 2 . 7
0600 | D e l l | # P h i l _ K n i g h t | 2
0620 9 . 6 | N i k e | # R o b s o n _ W a l
0640 t o n | 4 6 . 2 | W a l - M a r t | # S
0660 c h a e f f l e r | 2 5 . 3 | S c h a e
0680 f f l e r | # S e r g e y _ B r i n | 4
0700 7 . 5 | G o o g l e | # S t e v e _ B a
0720 l l m e r | 3 8 . 4 | M i c r o s o f t
0740 | # S u s a n n e _ K l a t t e n | 2 5
0760 | B M W | # W a n g _ J i a n l i n | 3
0780 0 | i m o v e i s | #
```

*Moreira, J. – Universidade Federal de São Carlos – Departamento de Computação – Rodovia Washington Luis, km 235 – 13565-905 – São Carlos/SP – Brasil – jander@dc.ufscar.br

¹Fonte: <https://epocanegocios.globo.com/empreendedorismo/noticia/2018/03/dono-da-amazon-jeff-bezos-lidera-lista-de-mais-ricos-do-mundo-da-forbes.html>, visitado em 17/7/2018.

Figura 3: Arquivo de dados com campos e registros de tamanhos variáveis, usando-se | para terminar campos e # para terminar registros. Os dados são, nesta ordem, *nome*, *fortuna* em bilhões de dólares e *atuação*. A coluna à esquerda indica o deslocamento em relação ao início do arquivo e as sublinhas indicam espaços.

```

0000 L e e _ S h a u _ K e e | 3 0 . 3 | i m
0020 o v e i s | # C h a r l e s _ K o c h |
0040 6 0 | d i v e r s i f i c a d o | # A l
0060 i c e _ W a l t o n | 4 6 | W a l - M a
0080 r t | # J i m _ W a l t o n | 4 6 . 4 |
0100 W a l - M a r t | # B i l l _ G a t e s
0120 | 9 0 | M i c r o s o f t | # D a v i d
0140 _ K o c h | 6 0 | d i v e r s i f i c a
0160 d o | # S t e v e _ B a l l m e r | 3 8
0180 . 4 | M i c r o s o f t | # M a r k _ Z
0200 u c k e r b e r g | 7 1 | F a c e b o o
0220 k | # W a n g _ J i a n l i n | 3 0 | i
0240 m o v e i s | # J o h n _ M a r s | 2 3
0260 . 6 | d o c e s | # J a c q u e l i n e
0280 _ M a r s | 2 3 . 6 | d o c e s | # R o
0300 b s o n _ W a l t o n | 4 6 . 2 | W a l
0320 - M a r t | # G i o v a n n i _ F e r r
0340 e r o | 2 3 | d o c e s | #

```

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.

GULIATO, D. *Estruturas de índice*. 2004. Material de apoio - Gerenciamento de Banco de Dados 2 - UFU. Disponível em: <http://www.facom.ufu.br/~guliato/disciplinas/gbd2/estruturas%20de%20indice.doc>.

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.