

# UNIVERSIDADE FEDERAL DA PARAÍBA

## CENTRO DE INFORMÁTICA

### DEPARTAMENTO DE INFORMÁTICA

<b>DISCIPLINA</b>	Ordenação e Recuperação de Dados	<b>TURMA</b>	01
<b>HORÁRIO</b>	T15502, I15502	<b>SALA</b>	T07
<b>Nº DE CRÉDITOS</b>	4	<b>CARGA HORÁRIA</b>	60 horas
<b>PROFESSOR</b>	Ulysses	<b>PERÍODO</b>	2017.2

## PLANO DE CURSO

### 1 Objetivos do Curso

#### 1.1 Objetivos Gerais

Apresentar conceitos, implementações e avaliações de estruturas de dados e algoritmos devotados a busca e ordenação de dados.

#### 1.2 Objetivos Específicos

Ao final do curso, o aluno deverá ser capaz de:

- Entender os conceitos de hierarquias de memória, caching e localidade de referência, bem como saber aplicar esses conceitos na implementação de algoritmos.
- Conhecer o conceito de tabela de busca e suas operações básicas.
- Saber descrever, implementar e avaliar tabelas de busca usando as seguintes estruturas de dados:
  - ◊ Lista indexada ordenada e sem ordenação
  - ◊ Lista encadeada simples
  - ◊ Lista encadeada com salto
  - ◊ Árvore binária de busca
  - ◊ Árvore multidirecional em memórias principal e secundária
  - ◊ Trie
- Entender a diferença entre implementação de estruturas de dados em memória principal e em memória secundária.
- Saber implementar pelo menos três algoritmos de casamento de strings.
- Conhecer o conceito de dispersão e saber implementar e avaliar funções de dispersão utilizando diversos métodos.
- Entender o que é colisão e saber resolvê-la utilizando encadeamento e endereçamento aberto.
- Entender dispersão extensível em memória secundária e saber implementá-la.
- Entender e saber implementar o conceito de heap binário usando array.

## 2 | Plano de Curso – Ordenação e Recuperação de Dados

- Conhecer o conceito de fila de prioridade e suas operações fundamentais, bem como saber implementar esse conceito de diversas maneiras.
- Entender o funcionamento dos principais algoritmos de ordenação de dados.

## 2 Pré-requisitos

Para bom acompanhamento desta disciplina, é desejável que o aluno apresente bom conhecimento dos seguintes assuntos:

- Programação em C
- Matemática do ensino médio, notadamente sobre somatórios, polinômios e logaritmos
- Lógica matemática, especialmente sobre conjectura e indução finita

## 3 Programa

UNIDADE	TÓPICO	NÚMERO DE AULAS
0	Revisão de Análise de Algoritmos	4
1	Organização de Memória	4
2	Busca Linear em Memória Principal	8
3	Busca Hierárquica em Memória Principal	12
4	Análise Amortizada de Algoritmos	2
5	Busca em Memória Secundária	10
6	Strings e Texto	8
7	Dispersão	4
8	Filas de Prioridade e Heaps	2
9	Ordenação de Dados em Memória Principal	4
10	Ordenação de Dados em Memória Secundária	2

**TABELA 1: PROGRAMA DA DISCIPLINA**

## 4 Metodologia

Aulas expositivas utilizando giz e quadro negro.

## 5 Método de Avaliação

Serão realizadas três avaliações periódicas (**Exercícios Escolares**), uma **Prova de Reposição** e o **Exame Final** de acordo com a programação mostrada na **Tabela 2**.

AVALIAÇÃO	CONTEÚDO	DATA
1º Exercício Escolar	Unidades 0 a 4	10/04/2018
2º Exercício Escolar	Unidades 5 e 6	17/05/2018

**TABELA 2: AVALIAÇÕES**

3º Exercício Escolar	Unidades 7 a 10	14/06/2018
Prova de Reposição	De acordo com a resolução vigente	19/06/2018
Exame Final	Toda a matéria ministrada	21/06/2018

TABELA 2: AVALIAÇÕES

## 6 Bibliografia

O acrônimo *ED* refere-se à disciplina Estruturas de Dados, enquanto que *ORD* refere-se a Ordenação e Recuperação de Dados.

### 6.1 Bibliografia Essencial

Com exceção da primeira referência, que se encontra em ordem de importância, as demais referências estão em ordem alfabética do primeiro autor.

- [1] Oliveira, U. de, *Estruturas de Dados Usando a Linguagem C – Volume II: Busca e Ordenação*, 2017 (em estágio de confecção).

Esta é a principal leitura sugerida, pois contém todo o material que será explorado em extensão e profundidade, além de outros tópicos que não serão considerados.

- [2] Dale, N. B., *C++ Plus Data Structures*, Jones & Bartlett Publishers; 3rd Edition, 2003.

É o livro sobre estruturas de dados mais didático que conheço. Mas essa virtude também é seu maior defeito, pois ele evita temas que não podem ser trivializados e, às vezes, exagera na didática. (A analogia entre notação ó e o elefante com o peixinho é, no mínimo, bizarra.) Além disso, contém muito pouco sobre busca, de modo que é indicado para ED, mas esse não é o caso para ORD (a não ser que você tenha dificuldade em entender ordenação).

- [3] Goodrich, M. T. e Tamassia, R., *Algorithm Design and Applications*, John Wiley & Sons, 2015.

Este livro é muito bom, mas, algumas vezes, difícil de entender, especialmente porque ele não usa terminologia padrão da área. Cobre praticamente todo o assunto de ED e ORD, mas a distribuição do material é um tanto não convencional, de modo que você terá que frequentemente utilizar o índice remissivo para encontrar o que procura. Este livro ainda serve como complemento de uma disciplina sobre grafos e algoritmos.

- [4] Goodrich, M. T., Tamassia, R. e Mount, D. M., *Data Structures & Algorithms in C++*, Second Edition, John Wiley & Sons, 2011.

Em termos conceituais, o conteúdo deste livro é praticamente o mesmo que o livro anterior. Os dois primeiros capítulos são dedicados a uma revisão de C++, enquanto que os demais incluem implementações em C++ de algoritmos associados às principais estruturas de dados.

- [5] Kruse, R. L. e Ryba, A. J., *Data Structures and Program Design in C++*, Prentice-Hall, 2000.

#### 4 | Plano de Curso – Ordenação e Recuperação de Dados

Apresenta belíssima diagramação e ilustrações um tanto exageradas que o fazem parecer um livro do ensino fundamental. Tem boa didática e cobre praticamente todo o assunto de ED e ORD, mas nem todos os tópicos são cobertos com a devida profundidade.

- [6] Tenenbaum, A. M., Langsam, Y. e Augenstein, M. J. *Data Structures Using C*, Pearson Education, 1990.

Está um pouco desatualizado, mas é um excelente livro.

- [7] Weiss, M. A., *Data Structures and Algorithm Analysis in C++*, 4th Edition, Pearson Education, 2013.

É um dos melhores livros de Estruturas de Dados que conheço, apesar de usar alguns termos estranhos e às vezes ser difícil de entender para o iniciante. Ele cobre virtualmente todos os tópicos de ED e ORD e apresenta implementações em C++. Existe uma versão bem mais antiga deste livro em C, mas que não parece ser tão boa quanto esta versão.

### 6.2 Bibliografia Recomendável

Todas as referências a seguir estão em ordem alfabética do primeiro autor.

- [1] Cormen, T. H., Leiserson, C. E., Rivest, R. L. e Stein, C., *Introduction to Algorithms*, 3rd Edition, The MIT Press, 2009.

Considerado uma bíblia da área, mas é um tanto formal demais. Não contém implementação em nenhuma linguagem.

- [2] Das, V. V., *Principles of Data Structures Using C and C++*, New Age International, 2008.

Consegue milagrosamente resumir quase todo o assunto em apenas 375 páginas. Contém implementações em C e C++, mas não cobre todo o assunto e é resumido demais. Lida bem com alguns poucos tópicos, mas, na avaliação geral não é um bom livro.

- [3] Drozdek, A., *Data Structures and Algorithms in C++*, Second Edition, Brooks/Cole Thomson Learning, 2001.

É um bom livro que cobre quase todo os tópicos de ED e ORD, mas contém poucas implementações e as ilustrações são um tanto confusas.

- [4] Drozdek, A., *Data Structures and Algorithms in Java*, Second Edition, Course Technology, 2005.

É a versão Java do livro anterior com diagramação e ilustrações melhores do que aquelas do livro anterior.

- [5] Harris, S. e Ross, J., *Beginning Algorithms*, Wiley Publishing, 2005.

Este livro é superficial demais e é péssimo em profundidade e conteúdo, mas tem boa didática em alguns tópicos (triviais).

- [6] Loudon, K., *Mastering Algorithms with C*, O'Reilly Media, 1999.

Muito fraco do ponto de vista conceitual e de abrangência, mas contém algumas implementações úteis. Boa parte do livro nada tem a ver com estruturas de dados.

- [7] Sedgewick, R., *Algorithms in C*, 3rd Ed, Addison-Wesley, 1998.
- Este é um livro muito utilizado em todo o mundo, mas ele usa uma terminologia que não é padrão na área e, se você não ler o livro a partir do início, terá dificuldade em entender tópicos isolados. O estilo de programação do autor é simplesmente péssimo e duvido que ele tenha implementado certos programas que ele *ensina* como implementar. Mesmo assim, ele lida com alguns tópicos com muita maestria, como, por exemplo, avaliação experimental de algoritmos e estruturas de dados (i.e., sem usar notação  $\mathcal{O}$ ). O grande mérito de Sedgewick foi ter *colorido* as árvores rubro-negras vários anos atrás.
- [8] Sedgewick, R. e Wayne, K., *Algorithms*, 4th Ed, Pearson Education, 2011.
- Esta é uma versão revista do livro anterior de Sedgewick, mas parte dos comentários apresentados acima continua válida. Este livro apresenta implementações em Java com um estilo de programação infinitamente superior ao do livro anterior. A seção referente a *extendible hashing*, que acredito que Sedgewick não sabia como implementar, foi discretamente omitida!
- [9] Tenenbaum, A. M., Langsam, Y. e Augenstein, M. J. *Estruturas de Dados Usando C*, Makron Books, 1995.
- A versão em Português do livro de Tenenbaum é um lixo! Mesmo assim, eu a uso com relativa frequência (porque não tenho a versão original), mas é difícil para quem não tem experiência saber o que está certo, errado ou é erro de tradução.

### 6.3 Bibliografia Complementar

Todas as referências a seguir estão em ordem alfabética do primeiro autor.

- [1] Bryant, R. E. e O'Hallaron, D. R., *Computer Systems: A Programmer's Perspectives*, Second Edition, Addison-Wesley Publishing, 2010.
- Todo estudante de Ciências da Computação deveria ler este livro. Ele não é essencial para esta disciplina, mas contém uma discussão bem mais aprofundada de caching, paginação e memória virtual do que aquela que será vista na **Unidade 2** desta disciplina.
- [2] Deshpande, P. S. e Kakde, O. G., *C & Data Structures*, Charles River Media Computer Engineering, 2004.
- Contém bons exemplos de programação em C escolhidos de acordo com o gosto do autor. Contém poucos exemplos relacionados com busca, de modo que é mais indicado para ED do que para ORD.
- [3] *Dictionary of Algorithms and Data Structures*, publicado em <https://xlinux.nist.gov/dads/>, acessado em 23/01/2014.
- É uma excelente fonte de referência sobre estruturas de dados.
- [4] *Dr. Dobb's Essential Books on Algorithms and Data Structures, Release 2*, Miller Freeman, 1999.
- É uma boa fonte de referência sobre estruturas de dados, mas a maioria dos livros não apresenta boa didática e está desatualizada.
- [5] Epp, S. S., *Discrete Mathematics with Applications*, Cengage Learning, 2010.

## 6 | Plano de Curso – Ordenação e Recuperação de Dados

É um livro com excelente didática se você pretende entender mais sobre análise de algoritmos e teoria dos números.

- [6] Heineman, G. T., Pollice, G. e Selkow, S., *Algorithms in a Nutshell*, O’Reilly Media, 2009.

É uma boa obra de referência, mas não cobre todo o conteúdo nem em abrangência nem em profundidade.

- [7] Horowitz, E. e Sahni, S., *Fundamentals of Data Structures*, Computer Science Press, 1976.

É um clássico na mesma linha do livro de Cormen. Se você gostar dele, suspeito de sua sanidade mental, mas na minha época de estudante era normal. A tradução em Português deste livro é lastimável! É mais fácil implementar remoção em árvores rubro-negras de olhos vendados do que entender a versão em Português deste livro, que é carinhosamente conhecida como *Horrorosovitz...*

- [8] Karumanchi, N., *Data Structures and Algorithms Made Easy: Data Structure and Algorithmic Puzzles*, CareerMonk Publications, 2011.

Este livro contém resumo da teoria que norteia estruturas de dados e inúmeras questões resolvidas ou com sugestões de como podem ser resolvidas. É um excelente livro de natureza pragmática, mas não pode ser usado isoladamente. Ele serve para ED e para ORD.

- [9] Konheim, A. G. *Hashing in Computer Science: Fifty Years of Slicing and Dicing*, John Wiley & Sons, 2010.

Este livro foi incluído apenas para você perceber que, ao contrário do que muitos pensam, dispersão é um tópico tão complexo que é possível escrever um livro com mais de 400 páginas sobre o assunto! Serve apenas para ORD (mas não é grande coisa!).

- [10] Knuth, D.E., *The Art of Computer Programming, Volume 1: Fundamental Algorithms*, 3rd Edition, Addison-Wesley, 1997.

A obra de Knuth é enciclopédica e é uma inesgotável fonte de referência. Ela é sempre atual em muitos tópicos e, no presente, compreende quatro volumes, sendo que o autor pretende lançar o quinto volume em 2025, aos 87 anos. Se a área de Estruturas de Dados tem pai, esse pai é Donald Knuth.

- [11] Knuth, D.E., *The Art of Computer Programming, Volume 3: Sorting and Searching*, 3rd Edition, Addison-Wesley, 1997.

Idem.

- [12] Oliveira, U. de, *Programando em C: Volume I – Fundamentos*, Editora Ciência Moderna, 2008.

Sou suspeito para comentar, mas este livro contém uma boa introdução à programação de baixo nível em C, que é essencial para bom entendimento de dispersão (**Unidade 6**).

- [13] Oliveira, U. de, *Programando em C: Volume II – A Biblioteca Padrão*, Editora Ciência Moderna, 2009.

Após a homologação do padrão C11, este livro ficou um pouco desatualizado, mas ainda é uma boa fonte de referência para a biblioteca padrão de C.

- [14] Rosen, K. H., *Discrete Mathematics and Its Applications*, 7th Edition, McGraw-Hill, 2012.

Este livro é similar ao livro de Epp analisado acima, mas é um pouco mais didático.

- [15] Skiena, S. S., *The Algorithm Design Manual*, Second Edition, Springer-Verlag, 2008.

Não é bem organizado como um livro de estruturas de dados, mas apresenta algumas boas discussões sobre algoritmos básicos.

- [16] Stallman, R., et alii, *Debugging with GDB – The Gnu Source-Level Debugger*, Ninth Edition, GNU Press, 2002.

Se você implementar algoritmos que não são triviais, como muitos que são apresentados nesta disciplina, certamente seus programas conterão bugs e alguns deles requerem o uso de um depurador para corrigi-los. Este livro sobre GDB é uma mão na roda.

- [17] Wirth, N., *Algorithms + Data Structures = Programs*, Prentice-Hall, 1976.

É um clássico que foi bom em sua (e minha) época, mas ainda contém algumas implementações úteis, desde que o leitor entenda Pascal ou Modula-2, que foram linguagens inventadas pelo autor e que ele usa em suas implementações.

- [18] Wirth, N., *Algorithms and DataStructures*, Prentice Hall, 1985.

Idem.

- [19] Wilf, H. S., *Algorithms and Complexity*, 2nd Edition, A K Peters/CRC Press, 2002.

Excelente fonte de referência em análise de algoritmos.

## 7 Recursos na Internet

A página da disciplina na internet provê todos os recursos necessários para seu bom acompanhamento. O endereço desta página é: <http://www.ulysseso.com/ord>. Não deixe de visitá-la com frequência. O professor pode ser contatado por meio do email: [ulysses@ulysseso.com](mailto:ulysses@ulysseso.com).

Para todos os algoritmos de ordenação apresentados nesta disciplina, existem animações, simulações e até danças folclóricas que facilitam o entendimento desses algoritmos de modo lúdico. Utilize o Google e o YouTube para encontrá-las e aprender mais divertindo-se.

## 8 Cronograma de Aulas

O cronograma de aulas apresentado na **Tabela 3** está sujeito a alterações e ajustes no decorrer do período.

8 | Plano de Curso – Ordenação e Recuperação de Dados

AULA	DATA	CONTEÚDO PREVISTO
1 <sup>a</sup>	20/02	<ul style="list-style-type: none"> <li>• Apresentação do Curso</li> <li>• Revisão de Análise de Algoritmos</li> </ul>
2 <sup>a</sup>	22/02	<ul style="list-style-type: none"> <li>• Revisão de Análise de Algoritmos</li> </ul>
3 <sup>a</sup>	27/02	<ul style="list-style-type: none"> <li>• Organização de Memória</li> <li>• Meios de Armazenamento</li> <li>• Acesso a Dispositivos de Entrada e Saída</li> <li>• Hierarquias de Memória</li> </ul>
4 <sup>a</sup>	01/03	<ul style="list-style-type: none"> <li>• Caching</li> <li>• Localidade de Referência</li> <li>• Análise de Algoritmos e Memória Externa</li> </ul>
5 <sup>a</sup>	06/03	<ul style="list-style-type: none"> <li>• Busca: Definições Fundamentais</li> <li>• Busca Sequencial com Registros em Memória Principal</li> </ul>
6 <sup>a</sup>	08/03	<ul style="list-style-type: none"> <li>• Busca Sequencial com Movimentação de Registros</li> <li>• Busca Sequencial com Registros em Memória Secundária</li> <li>• Busca Linear em Tabela Ordenada</li> </ul>
7 <sup>a</sup>	13/03	<ul style="list-style-type: none"> <li>• Listas com Saltos</li> </ul>
8 <sup>a</sup>	15/03	<ul style="list-style-type: none"> <li>• Busca com Chaves Secundárias</li> <li>• Busca de Piso e Teto de Chaves</li> <li>• Busca de Intervalo</li> <li>• Menor e Maior Chaves</li> </ul>
9 <sup>a</sup>	20/03	<ul style="list-style-type: none"> <li>• Busca Hierárquica em Memória Principal</li> <li>• Revisão de Árvores Binárias</li> </ul>
10 <sup>a</sup>	22/03	<ul style="list-style-type: none"> <li>• Árvores Binárias de Busca</li> </ul>
11 <sup>a</sup>	27/03	<ul style="list-style-type: none"> <li>• Rotações em Árvores Binárias de Busca</li> <li>• Árvores AVL</li> </ul>
12 <sup>a</sup>	29/03	<ul style="list-style-type: none"> <li>• Árvores Binárias Afuniladas</li> </ul>
13 <sup>a</sup>	03/04	<ul style="list-style-type: none"> <li>• Árvores Binárias Afuniladas</li> </ul>
14 <sup>a</sup>	05/04	<ul style="list-style-type: none"> <li>• Análise Amortizada</li> </ul>
15 <sup>a</sup>	10/04	<ul style="list-style-type: none"> <li>• <b>1º Exercício Escolar</b></li> </ul>
16 <sup>a</sup>	12/04	<ul style="list-style-type: none"> <li>• Revisão de Processamento de Arquivos em C</li> <li>• Busca em Memória Secundária</li> <li>• Árvores Multidirecionais de Busca Descendentes</li> <li>• Árvores Multidirecionais de Busca em Memória Principal</li> </ul>
17 <sup>a</sup>	17/04	<ul style="list-style-type: none"> <li>• Árvores Multidirecionais de Busca em Memória Secundária</li> </ul>
18 <sup>a</sup>	19/04	<ul style="list-style-type: none"> <li>• Árvores B</li> </ul>
19 <sup>a</sup>	24/04	<ul style="list-style-type: none"> <li>• Árvores B</li> </ul>
20 <sup>a</sup>	26/04	<ul style="list-style-type: none"> <li>• Árvores B+</li> <li>• Outras Variantes de Árvores B</li> <li>• Comparação de Árvore Multidirecionais</li> </ul>

**TABELA 3: CRONOGRAMA DE AULAS**



AULA	DATA	CONTEÚDO PREVISTO
21 <sup>a</sup>	03/05	<ul style="list-style-type: none"> <li>• Casamento de Strings: Terminologia e Visualização</li> <li>• Casamento de Strings com Retrocesso</li> <li>• Casamento por Força Bruta (FB)</li> </ul>
22 <sup>a</sup>	08/05	<ul style="list-style-type: none"> <li>• Algoritmo de Knuth, Morris e Pratt (KMP)</li> <li>• Algoritmo de Boyer e Moore</li> <li>• Algoritmo de Boyer, Moore e Horspool (BMH)</li> </ul>
23 <sup>a</sup>	10/05	<ul style="list-style-type: none"> <li>• Algoritmo de Karp e Rabin (KR)</li> <li>• Comparação de Algoritmos de Casamento de Strings</li> </ul>
24 <sup>a</sup>	15/05	<ul style="list-style-type: none"> <li>• Tries</li> </ul>
25 <sup>a</sup>	17/05	<ul style="list-style-type: none"> <li>• <b>2º Exercício Escolar</b></li> </ul>
26 <sup>a</sup>	22/05	<ul style="list-style-type: none"> <li>• Dispersão: Conceitos</li> <li>• Funções de Dispersão</li> <li>• Dispersão com Encadeamento</li> </ul>
27 <sup>a</sup>	24/05	<ul style="list-style-type: none"> <li>• Dispersão com Endereçamento Aberto</li> </ul>
28 <sup>a</sup>	29/05	<ul style="list-style-type: none"> <li>• Filas de Prioridade</li> <li>• Heaps Binários</li> </ul>
29 <sup>a</sup>	05/06	<ul style="list-style-type: none"> <li>• Métodos de Ordenação: Ordenação com Custo <math>O(n^2)</math></li> </ul>
30 <sup>a</sup>	07/06	<ul style="list-style-type: none"> <li>• Métodos de Ordenação: Ordenação com Custo <math>O(n \log n)</math></li> </ul>
31 <sup>a</sup>	12/06	<ul style="list-style-type: none"> <li>• Métodos de Ordenação: Ordenação com Custo <math>O(n)</math></li> <li>• Ordenação em Memória Secundária</li> </ul>
32 <sup>a</sup>	14/06	<ul style="list-style-type: none"> <li>• <b>3º Exercício Escolar</b></li> </ul>
33 <sup>a</sup>	19/06	<ul style="list-style-type: none"> <li>• <b>Reposição</b></li> </ul>
34 <sup>a</sup>	21/06	<ul style="list-style-type: none"> <li>• <b>Exame Final</b></li> </ul>

TABELA 3: CRONOGRAMA DE AULAS

Última alteração: 19/02/2018

© Prof. Dr. Ulysses de Oliveira

