

MERGULHANDO NO
PYTHON

**O guia rápido e prático
para dominar o Python**

Mark Pilgrim

Mergulhando no Python

Copyright © 2000, 2001, 2002, 2003, 2004 Mark Pilgrim (mail to: mark@diveintopython.org . This book lives at <http://diveintopython.org/>. If you're reading it somewhere else, you may not have the latest version. Permission is granted to copy, distribute, and/or modify this document under the terms of the GNU Free Documentation License, Version 1.1 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in Appendix G, GNU Documentation License. The example programs in this book are free software; you can redistribute and/or modify them under the terms of the Python license as published by the Python Software Foundation. A copy of the license is included in Appendix H, Python license.

Todo o esforço foi feito para fornecer a mais completa e adequada informação, contudo a editora e o(s) autor(es) não assumem responsabilidade pelos resultados e usos da informação fornecida. Recomendamos aos leitores testar a informação, bem como tomar todos os cuidados necessários (como o backup), antes da efetiva utilização. Este livro não contém CD-ROM, disquete ou qualquer outra mídia.

Erratas e atualizações: Sempre nos esforçamos para entregar a você, leitor, um livro livre de erros técnicos ou de conteúdo; porém, nem sempre isso é conseguido, seja por motivo de alteração de software, interpretação ou mesmo quando alguns deslizes constam na versão original de alguns livros que traduzimos. Sendo assim, criamos em nosso site, www.altabooks.com.br, a seção Erratas, onde relataremos, com a devida correção, qualquer erro encontrado em nossos livros.

Avisos e Renúncia de Direitos: Este livro é vendido como está, sem garantia de qualquer tipo, seja expressa ou implícita.

Marcas Registradas: Todos os termos mencionados e reconhecidos como Marca Registrada e/ou comercial são de responsabilidade de seus proprietários. A Editora informa não estar associada a nenhum produto e/ou fornecedor apresentado no livro. No decorrer da obra, imagens, nomes de produtos e fabricantes podem ter sido utilizados, e desde já a Editora informa que o uso é apenas ilustrativo e/ou educativo, não visando ao lucro, favorecimento ou desmerecimento do produto/fabricante.

Produção Editorial Editora Alta Books

Coordenação Editorial: Fernanda Silveira

Tradução: Marcelo Soares

Revisão: Gustav Schmid

Diagramação: Aldenir Gil de Oliveira

Capa: Fernanda Silveira

Impresso no Brasil

O código de propriedade intelectual de 1º de Julho de 1992 proíbe expressamente o uso coletivo sem autorização dos detentores do direito autoral da obra, bem como a cópia ilegal do original. Esta prática generalizada nos estabelecimentos de ensino, provoca uma brutal baixa nas vendas dos livros a ponto de impossibilitar os autores de criarem novas obras.



Av. Nilo Peçanha, 155, cjs. 1101 a 1106 - Castelo Rio de Janeiro - RJ.

CEP: 20020-100

Tel: 21 2532-6556/ Fax: 2215-0225

www.altabooks.com.br

e-mail: altabooks@altabooks.com.br

Índice

Capítulo 1. Instalando o Python	1
1.1. Qual Python é o mais adequado para você?	1
1.2. Python no Windows	1
1.3. Python no Mac OS X	2
1.4. Python no Mac OS 9	3
1.5. Python no RedHat Linux	3
1.6. Python no Debian GNU/Linux	4
1.7. Instalação do Python a partir do código-fonte	5
1.8. O shell interativo	5
1.9. Resumo	6
Capítulo 2. Seu primeiro programa em Python	7
2.1. Mergulhando	7
2.2. Declarando funções	7
2.3. Funções de documentação	8
2.4. Tudo são objetos	8
2.5. Parágrafos no código	10
2.6. Testando módulos	11
Capítulo 3. Tipos de dados nativos	13
3.1. Apresentando os dicionários	13
3.2. Apresentado as listas	15
3.3. Introduzindo as tuples	20
3.4. Declarando variáveis	22
3.5. Formatando strings	24
3.6. Mapeando listas	26
3.7. Juntando listas e dividindo strings	28
3.8. Resumo	29
Capítulo 4. O poder da introspecção	33
4.1. Mergulhando	33
4.2. Usando argumentos opcionais e nomeados	34
4.3. Usando type, str, dir e outras funções embutidas	34
4.5. Filtrando listas	39
4.6. A peculiar natureza de and e or	40
4.7. Usando funções lambda	42
4.8. Juntando tudo	43
4.9. Resumo	45
Capítulo 5. Objetos e orientação a objetos	47
5.1. Mergulhando	47
5.2. Importando módulos usando from módulo import	49
5.3. Definindo classes	50

5.4.	Instanciando classes	52
5.5.	Explorando UserDict: uma classe Wrapper	53
5.6.	Métodos especiais de classes	55
5.7.	Métodos especiais de classes avançados	57
5.8.	Apresentando os atributos de classes	58
5.9.	Funções privadas	59
5.10.	Resumo	60

Capítulo 6. Tratamento de exceções e de arquivos..... 61

6.1.	Tratando das exceções	61
6.2.	Trabalhando com objetos de arquivos	63
6.3.	Iterando com Loops for	66
6.4.	Usando sys.modules	68
6.5.	Trabalhando com diretórios	70
6.6.	Juntando tudo	73
6.7.	Resumo	74

Capítulo 7. Expressões regulares..... 77

7.1.	Mergulhando	77
7.2.	Estudo de caso: endereços postais	77
7.3.	Estudo de caso: numerais romanos	78
7.4.	Usando a sintaxe {n,m}	81
7.5.	Expressões regulares verbose	83
7.6.	Estudo de caso: fazendo parsing de números de telefone	84
7.7.	Resumo	87

Capítulo 8. Processamento de HTML 89

8.1.	Mergulhando	89
8.2.	Introduzindo sgmlib.py	93
8.3.	Extraindo dados de documentos HTML	94
8.4.	Introduzindo BaseHTMLProcessor.py	96
8.5.	locals e globals	98
8.6.	Formatação de strings baseada em dicionário	100
8.7.	Colocando valores de atributos entre aspas	101
8.8.	Apresentando dialect.py	102
8.9.	Juntando tudo	104
8.10.	Resumo	105

Capítulo 9. Processamento de XML..... 107

9.1.	Mergulhando	107
9.2.	Pacotes	112
9.3.	Fazendo parsing de XML	114
9.4.	Unicode	116
9.5.	Procurando elementos	119
9.6.	Acessando atributos dos elementos	121
9.7.	Finalizando	122

Capítulo 10. Scripts e streams	123
10.1. Abstraindo fontes de input	123
10.2. Input, output e erro padrões	126
10.3. Armazenando buscas de nós em cache	129
10.4. Encontrando descendentes diretos de um nó	130
10.5. Criando handlers separados por tipo de nó	130
10.6. Manipulando argumentos de linha de comando	132
10.7. Juntando tudo	134
10.8. Resumo	136
Capítulo 11. Serviços Web HTTP	137
11.1. Mergulhando	137
11.2. Como não obter dados através do HTTP	139
11.3. Recursos do HTTP	139
11.4. Debugging de serviços web HTTP	141
11.5. Definindo o User-Agent	142
11.6. Manipulando Last-Modified e ETag	143
11.7. Manipulando redirecionamentos	145
11.8. Manipulando dados comprimidos	149
11.9. Juntando tudo	151
11.10. Resumo	153
Capítulo 12. Serviços Web SOAP	155
12.1. Mergulhando	155
12.2. Instalando as bibliotecas SOAP	156
12.3. Primeiros passos com o SOAP	157
12.4. Debugging de serviços Web SOAP	158
12.5. Apresentado a WSDL	159
12.6. Fazendo introspecção de serviços Web SOAP com a WSDL	160
12.7. Buscando no Google	162
12.9. Resumo	167
Capítulo 13. Testando as unidades	169
13.1. Introdução aos numerais romanos	169
13.2. Mergulhando	169
13.3. Apresentando romantest.py	170
13.4. Testando o sucesso	172
13.5. Testando o fracasso	174
13.6. Testando a sanidade	176
Capítulo 14. Testando antes de programar	179
14.1. roman.py, etapa 1	179
14.2. roman.py, etapa 2	182
14.3. roman.py, etapa 3	185
14.4. roman.py, etapa 4	188
14.5. roman.py, etapa 5	190

Capítulo 15. Refatorando	193
15.1. Tratando os bugs	193
15.2. Manipulando as mudanças de requerimentos	194
15.3. Refatorando	200
15.4. Post-scriptum	203
15.5. Resumo	205
Capítulo 16. Programação funcional	207
16.1. Mergulhando	207
16.3. Listas de filtragem revisitadas	210
16.4. Mapeamento de listas revisitado	211
16.5. Programação centrada nos dados	212
16.6. Importando módulos dinamicamente	213
16.7. Juntando tudo	214
16.8. Resumo	216
Capítulo 17. Funções dinâmicas	217
17.1. Mergulhando	217
17.2. plural.py, etapa 1	217
17.3. plural.py, etapa 2	219
17.4. plural.py, etapa 3	220
17.5. plural.py, etapa 4	221
17.6. plural.py, etapa 5	223
17.7. plural.py, etapa 6	224
17.8. Resumo	226
Capítulo 18. Ajuste do desempenho	227
18.1. Mergulhando	227
18.2. Usando o módulo timeit	229
18.3. Otimizando expressões regulares	230
18.4. Otimizando buscas em dicionários	232
18.5. Otimizando operações de listas	235
18.6. Otimizando a manipulação de strings	236
18.7. Resumo	237
Apêndice A. Leituras adicionais	239
Capítulo 1. Instalando o Python	239
Capítulo 2. Seu primeiro programa em Python	239
2.3. Funções de documentação	239
2.5. Parágrafos no código	239
2.6. Testando módulos	239
Capítulo 3. Tipos de dados nativos	239
3.3. Apresentando as tuples	240
3.5. Formatando strings	240

3.6. Mapeando listas	240
3.7. Juntando listas e dividindo strings	240
Capítulo 4. O poder da introspecção	240
4.2. Usando argumentos opcionais e nomeados	240
4.5. Filtrando listas	241
Capítulo 5. Objetos e orientação a objetos	241
5.2. Importando módulos usando from módulo import	241
5.3.2. Sabendo quando usar self e __init__	241
5.5. Explorando UserDict: uma classe wrapper	242
5.7. Métodos especiais de classes avançados	242
5.9. Funções privadas	242
Capítulo 6. Tratamento de exceções e de arquivos	242
6.5. Trabalhando com diretórios	243
Capítulo 7. Expressões regulares	243
7.6. Estudo de caso: fazendo parsing de números de telefone	243
Capítulo 8. Processamento de HTML	243
8.4. Apresentando BaseHTMLProcessor.py	243
8.9. Juntando tudo	243
Capítulo 9. Processamento de XML	243
9.4. Unicode	243
Capítulo 10. Scripts e streams	243
Capítulo 11. Serviços Web HTTP	243
11.1. Mergulhando	243
Capítulo 12. Serviços Web SOAP	243
12.1. Mergulhando	243
12.8. Resolvendo problemas dos serviços Web SOAP	244
Capítulo 13. Testando as unidades	244
13.1. Introdução aos numerais romanos	244
13.3. Apresentando romantest.py	244
Capítulo 14. Testando antes de programar	244
Capítulo 15. Refatorando	244
15.5. Resumo	244
Capítulo 16. Programação funcional	244
Capítulo 17. Funções dinâmicas	244

17.7. plural.py, etapa 6	244
Capítulo 18. Ajuste do desempenho	244
18.1. Mergulhando	244
Apêndice B. Uma revisão de 5 minutos	245
Capítulo 1. Instalando o Python	245
1.1. Qual Python é o mais adequado para você?	245
1.2. Python no Windows	245
1.3. Python no Mac OS X	245
1.4. Python no Mac OS 9	245
1.5. Python no RedHat Linux	245
1.6. Python no Debian GNU/Linux	245
1.7. Instalação do Python a partir do código-fonte	245
1.8. O shell interativo	245
1.9. Resumo	245
Capítulo 2. Seu primeiro programa em Python	245
2.1. Mergulhando	245
2.2. Declarando funções	245
2.3. Funções de documentação	245
2.4. Tudo é um objeto	245
2.5. Parágrafos no código	245
2.6. Testando módulos	246
Capítulo 3. Tipos de dados nativos	246
3.1. Apresentando os dicionários	246
3.2. Apresentando as listas	246
3.3. Apresentando as tuples	246
3.4. Declarando variáveis	246
3.5. Formatando strings	246
3.6. Mapeando listas	246
3.7. Juntando listas e dividindo strings	246
3.8. Resumo	246
Capítulo 4. O poder da introspecção	246
4.1. Mergulhando	246
4.2. Usando argumentos opcionais e nomeados	246
4.3. Usando type, str, dir e outras funções embutidas	246
4.4. Obtendo referências a objetos com getattr	247
4.5. Filtrando listas	247
4.6. A peculiar natureza de and e or	247
4.7. Usando funções lambda	247
4.8. Juntando tudo	247
4.9. Resumo	247
Capítulo 5. Objetos e orientação a objetos	247
5.1. Mergulhando	247

5.2.	Importando módulos usando from módulo import	247
5.3.	Definindo classes	247
5.4.	Instanciando classes	247
5.5.	Explorando UserDict: uma classe wrapper	247
5.6.	Métodos especiais de classes	247
5.7.	Métodos especiais de classes avançados	248
5.8.	Apresentando os atributos de classes	248
5.9.	Funções privadas	248
5.10.	Resumo	248
Capítulo 6. Tratamento de exceções e de arquivos		248
6.1.	Tratando das exceções	248
6.2.	Trabalhando com objetos de arquivos	248
6.3.	Iterando com Loops for	248
6.4.	Usando sys.modules	248
6.5.	Trabalhando com diretórios	248
6.6.	Juntando tudo	248
6.7.	Resumo	248
Capítulo 7. Expressões regulares		248
7.1.	Mergulhando	248
7.2.	Estudo de caso: endereços postais	249
7.3.	Estudo de caso: numerais romanos	249
7.4.	Usando a sintaxe {n,m}	249
7.5.	Expressões regulares verbose	249
7.6.	Estudo de caso: fazendo parsing de números de telefone	249
7.7.	Resumo	249
Capítulo 8. Processamento de HTML		249
8.1.	Mergulhando	249
8.2.	Apresentando sgmlib.py	249
8.3.	Extraindo dados de documentos HTML	249
8.4.	Apresentando BaseHTMLProcessor.py	249
8.5.	locals e globals	250
8.6.	Formatação de strings baseada em dicionário	250
8.7.	Colocando valores de atributos entre aspas	250
8.8.	Apresentando dialect.py	250
8.9.	Juntando tudo	250
8.10.	Resumo	250
Capítulo 9. Processamento de XML		250
9.1.	Mergulhando	250
9.2.	Pacotes	250
9.3.	Fazendo parsing de XML	250
9.4.	Unicode	250
9.5.	Procurando elementos	250
9.6.	Acessando atributos dos elementos	251
9.7.	Finalizando	251

Capítulo 10. Scripts e streams	251
10.1. Abstraindo fontes de input	251
10.2. Input, output e erro padrões	251
10.3. Armazenando buscas de nós em cache	251
10.4. Encontrando descendentes diretos de um nó	251
10.5. Criando handlers separados por tipo de nó	251
10.6. Manipulando argumentos de linha de comando	251
10.7. Juntando tudo	251
10.8. Resumo	251
Capítulo 11. Serviços Web HTTP	252
11.1. Mergulhando	252
11.2. Como não obter dados através do HTTP	252
11.3. Recursos do HTTP	252
11.4. Debugging de serviços web HTTP	252
11.5. Definindo o User-Agent	252
11.6. Manipulando Last-Modified e ETag	252
11.7. Manipulando redirecionamentos	252
11.8. Manipulando dados comprimidos	252
11.9. Juntando tudo	252
11.10. Resumo	252
Capítulo 12. Serviços Web SOAP	252
12.1. Mergulhando	252
12.2. Instalando as bibliotecas SOAP	252
12.3. Primeiros passos com o SOAP	252
12.4. Debugging de serviços web SOAP	253
12.5. Apresentando a WSDL	253
12.6. Fazendo introspecção de serviços web SOAP com a WSDL	253
12.7. Buscando no Google	253
12.8. Resolvendo problemas dos serviços web SOAP	253
12.9. Resumo	253
Capítulo 13. Testando as unidades	253
13.1. Introdução aos numerais romanos	253
13.2. Mergulhando	253
13.3. Apresentando romantest.py	253
13.4. Testando o sucesso	253
13.5. Testando o fracasso	253
13.6. Testando a sanidade	253
Capítulo 14. Testando antes de programar	254
14.1. roman.py, etapa 1	254
14.2. roman.py, etapa 2	254
14.3. roman.py, etapa 3	254
14.4. roman.py, etapa 4	254
14.5. roman.py, etapa 5	254
Capítulo 15. Refatorando	254

15.1. Tratando os bugs	254
15.2. Manipulando as mudanças de requerimentos	254
15.3. Refatorando	254
15.4. Post-scriptum	254
15.5. Resumo	255
Capítulo 16. Programação funcional	255
16.1. Mergulhando	255
16.2. Encontrando o caminho	255
16.3. Listas de filtragem revisitadas	255
16.4. Mapeamento de listas revisitado	255
16.5. Programação centrada nos dados	255
16.6. Importando módulos dinamicamente	255
16.7. Juntando tudo	255
16.8. Resumo	255
Capítulo 17. Funções dinâmicas	255
17.1. Mergulhando	255
17.2. plural.py, etapa 1	255
17.3. plural.py, etapa 2	256
17.4. plural.py, etapa 3	256
17.5. plural.py, etapa 4	256
17.6. plural.py, etapa 5	256
17.7. plural.py, etapa 6	256
17.8. Resumo	256
Capítulo 18. Ajuste do desempenho	256
18.1. Mergulhando	256
18.2. Usando o módulo timeit	256
18.3. Otimizando expressões regulares	256
18.4. Otimizando buscas em dicionários	256
18.5. Otimizando operações de listas	256
18.6. Otimizando a manipulação de strings	256
18.7. Resumo	256
Apêndice C. Dicas e truques	245
Capítulo 1. Instalando o Python	245
Capítulo 2. Seu primeiro programa em Python	245
2.1. Mergulhando	245
2.2. Declarando funções	245
2.3. Funções de documentação	245
2.4. Tudo são objetos	245
2.5. Parágrafos no código	245
2.6. Testando módulos	245
Capítulo 3. Tipos de dados nativos	245
3.1. Apresentando os dicionários	245

3.1.2. Modificando dicionários	245
3.2. Apresentando as listas	246
3.2.3. Buscando em listas	246
3.3. Apresentando as tuples	246
3.4. Declarando variáveis	246
3.5. Formatando strings	246
3.7. Juntando listas e dividindo strings	246
Capítulo 4. O poder da introspecção	246
4.2. Usando argumentos opcionais e nomeados	246
4.3.3. Funções embutidas	246
4.7. Usando funções lambda	246
4.8. Juntando tudo	247
Capítulo 5. Objetos e orientação a objetos	247
5.2. Importando módulos usando from módulo import	247
5.3. Definindo classes	247
5.4. Instanciando classes	247
5.5. Explorando UserDict: uma classe wrapper	247
5.7. Métodos especiais de classes avançados	248
5.8. Apresentando os atributos de classes	248
5.9. Funções privadas	248
Capítulo 6. Tratamento de exceções e de arquivos	248
6.1. Tratando das exceções	248
6.5. Trabalhando com diretórios	248
Capítulo 7. Expressões regulares	248
7.4. Usando a sintaxe {n,m}	248
Capítulo 8. Processamento de HTML	249
8.2. Apresentando sgmlib.py	249
8.4. Apresentando BaseHTMLProcessor.py	249
8.5. locals e globals	249
8.6. Formatação de strings baseada em dicionário	249
Capítulo 9. Processamento de XML	249
9.2. Pacotes	249
9.6. Acessando atributos dos elementos	249
Capítulo 10. Scripts e streams	250
Capítulo 11. Serviços web HTTP	250
11.6. Manipulando Last-Modified e ETag	250
Capítulo 12. Serviços web SOAP	250
Capítulo 13. Testando as unidades	250

13.2. Mergulhando	250
Capítulo 14. Testando antes de programar	250
14.3. roman.py, etapa 3	250
14.5. roman.py, etapa 5	250
Capítulo 15. Refatorando	250
15.3. Refatorando	250
Capítulo 16. Programação funcional	250
16.2. Encontrando o caminho	250
Capítulo 17. Funções dinâmicas	250
Capítulo 18. Ajuste do desempenho	250
18.2. Usando o Módulo timeit	250
Apêndice D. Sobre o livro	261
Apêndice E. Licença de documentação livre GNU	263
E.1. Aplicabilidade e definições	263
E.2. Cópia literal	263
E.3. Copiando em quantidades	264
E.4. Modificações	264
E.5. Combinando documentos	265
E.6. Coleções de documentos	265
E.7. Agregação com trabalhos independentes	265
E.8. Tradução	265
E.9. Invalidação	266
E.10. Revisões futuras desta licença	266
E.11. Como usar esta Licença para os seus documentos	266
Apêndice F. Licença do Python	267
F.A. História do software	267
F.B. Termos e condições para acessar ou usar de qualquer forma o Python	267
F.B.1. Contrato de licença da PSF	267
F.B.2. Contrato de licença de código livre do BeOpen Python, versão 1	267
F.B.3. Contrato de licença CNRI de código livre compatível com a GPL	268
F.B.4. Declaração de permissões e isenção de responsabilidades do CWI	269

