

Symbols

- operator 62, 261
 - operator 261
 - != operator 262
 - \$ 54, 102, 126
 - anchoring matches 64, 68
 - metacharacter 61
 - \$" 102, 103, 109, 126
 - \$& 67
 - \$, 102, 103
 - \$. 100
 - \$/ 55, 126
 - controlling output 102
 - reading data 85, 97
 - record separator 84, 98, 111
 - special values 113
 - \$::RD_AUTOACTION 215
 - \$::RD_HINT 220
 - \$::RD_TRACE 220
 - \$_ 65, 99, 104, 113, 180
 - \$| 105
 - \$' 67
 - \$' 67
 - \$1 63, 65, 66, 67, 71
 - \$a 40, 41
 - \$b 40, 41
 - %_ 180
 - & 263
 - && operator 262
 - * operator 61, 63, 64, 261
 - ** operator 261
 - + metacharacter 61, 63, 64
 - + operator 261
 - ++ operator 261
 - . operator 61, 261
 - / operator 66, 261
 - < operator 262
 - <= operator 262
 - <=> operator 40, 262
 - <> operator 35, 84
 - == operator 262
 - ~= operator 67
 - > operator 262
 - >= operator 262
 - ? metacharacter 61, 63, 64
 - @item array 214
 - \ 61
 - \A 68
 - \B 64
 - \b 64
 - \D 62
 - \d 62
 - \S 62
 - \s 62
 - \W 62
 - \w 62, 64
 - \Z 68
 - ^ metacharacter 61, 64, 68
 - ^ operator 263
 - {n,m} 64
 - | metacharacter 61, 62
 - | operator 263
 - || operator 41, 107, 262
 - ||= 43
 - ~ operator 263
 - A**
 - a command line option 54
 - Aas, Gisle 143
 - access method 30
 - ActivePerl 16, 164
 - ActiveState 16, 164
 - Aho 162
 - alt attribute 158
 - alternate matches 62
 - alternate phrases 61
 - anchoring matches 64
 - and operator 262
 - anonymous array 83
 - Apache 123, 124
 - API 47
 - Apple Macintosh
 - carriage return 88
 - arrays 22, 257
 - (see also hashes)
 - array of arrays 82, 85
 - array of hashes 23, 138
 - array slice 260
 - array variable 103
 - associative arrays 259
 - context 67, 97, 115
 - example 83, 86
 - examples of assignment 257
 - flexibility 22
 - lookup 44
 - reference 269
 - slice 116
 - vs. hashes 82
 - vs. lists 258
- arrow operator 260
 - article 210
 - ASCII 9, 82, 130
 - binary 139
 - character set description 139
 - converting to EBCDIC 87
 - data conversion 87
 - template options 135
 - ascii2ebcdic 87
 - example 88

- assign 58
- attributes 24, 30
 - AutoCommit 236
 - Chopblanks 235
 - DECIMAL_DIGITS 239, 240
 - DECIMAL_FILL 239, 240
 - DECIMAL_POINT 239, 240
 - errstr 234
 - INT_CURR_SYMBOL 239, 240
 - Kids 234
 - KILO_SUFFIX 239, 240
 - LongReadLen 235
 - LongTruncOk 235
 - MEGA_SUFFIX 239, 240
 - MON_DECIMAL_POINT 239, 240
 - MON_THOUSANDS_SEP 239, 240
 - NAME 238
 - NEG_FORMAT 239, 240
 - NULLABLE 238
 - NUM_OF_FIELDS 238
 - PrintError 234
 - RaiseError 234
 - statement handle 238
 - THOUSANDS_SEP 238, 239, 240
 - TYPE 238
 - warn 234
- audit log 36, 38
- audit trail 36, 37
- awk 61

- B**

- backslash 62
- bare block 85
- bash 76
- benchmarking 51, 52, 56, 76, 122
 - Benchmark.pm 55
- Bentley, Jon 38
- binary data 14, 17, 82, 150
 - and \$/ 114
 - binmode 141
 - definition 13
 - file formats 139
 - line endings 88
 - PNG file 140
 - reading data chunks 140
 - reading the signature 140
 - working with, 139–144
- binding operator 67
- binmode 141
- bitwise 263
- bless 188
- block of Perl 40
- body 5, 218, 221
- Bowie, David 6
- Bragg, Billy 6
- Bunce, Tim 47, 55
- business rules 26, 29, 38
 - encapsulation methods 25, 26
 - reason to encapsulate 26

- C**

- C 14, 15, 16, 17
- c command line option 256
- C++ 16
- caching data
 - currency conversion 106
 - Memoize.pm 107
- capturing 63, 66
- caret 64
- carriage return 62, 88
- case transformations 60
- CD example 4, 5
 - adding subrecords 151
 - complex data files 150
 - complex records 111–113
 - creating a template 153
 - data parsing 6
 - data structure 20, 23–25
 - Data::Dumper 49
 - grammar 218, 219
 - hash 21, 22
 - parsing 217
 - printing lists 103–104
 - Schwartzian transform 43
- CD-ROM 9
- sensor 77
- census 4
- chain extension 35
- channels 194
- Chapman, Nigel 17, 255
- character classes 61, 62, 83
- checksum 142
- China 87
- chomp 98, 99, 112
- Christiansen, Tom 17, 55, 255
- chunk data 142
- chunk footer 142
- closelog 37
- cmp 41, 262
- COBOL 128
- code 26
 - reuse 16
- colon 64, 74
 - data separated by 73
- column widths 133
- command line options 53
 - a 54
 - c 256
 - d 256
 - e 53
 - F 54
 - i 54
 - M 53
 - n 53, 54
 - p 54
 - T 256
 - w 256
- command line scripts 53, 56
- comma-separated files (*see* CSV)
- comp.lang.perl.misc 43
- compiled 16
- Compilers: Principles, Techniques and Tools* 162
- complex data files 150
 - g modifier 156
 - HTML 154
 - metadata 151
 - subrecords 151
 - XML 154
- complex data records
 - expanded CD file 152
 - reading, example 152
- complex data structures 268, 272
 - using references 269
- complex sorts 41
- Comprehensive Perl Archive Network 16
- computer science 60
- concatenate 51, 261
- conditional execution 263
- connection string 48
- consonants 62
- context 5
- control character 89
- control of input 84

- Convert::EBCDIC 87
 - example 88
 - Convert::SciEng 91, 92, 94
 - converting number formats 90
 - converting the character set 87
 - Conway, Damian 28, 38, 210, 224
 - cool-talk 76
 - CPAN 16
 - ASCII to EBCDIC module 87
 - binary files 143
 - date modules 120, 126
 - HTML parser 165
 - libwww 164
 - Logfile 124
 - Memoize.pm 107
 - MP3 data 144, 145
 - parser module 174
 - POD 197
 - reformatting numbers 91
 - Text::Bastardize module 76
 - Text::CSV_XS 109
 - top-down parsers 210
 - XML parser 178, 208
 - XML parsers 191
 - CSV 108
 - anatomy 108
 - data records 108
 - generic read/write routine 110
 - Text::CSV_XS 109
 - currency conversion
 - memoize 107
 - customer object 29
 - Customer.pm 29, 30
 - Customer_Rules.pm 27, 28
- D**
-
- d command line option 256
 - d00dz 76
 - dash 62
 - data
 - CSV records 109
 - exchange 176
 - field 4
 - fields 109
 - formats 12–14
 - reading chunks 140
 - transformation 6
 - types of, 256
 - data conversion 87
 - ASCII 87
 - ASCII and EBCDIC 87
 - binary data 88
 - converting the character set 87
 - EBCDIC 87
 - line endings 88
 - multibyte characters 87
 - number format 90
 - reformatting numbers 91
 - sprintf 91
 - data file 9, 129
 - corruption check 151
 - footer 151
 - header 150
 - parts of 150
 - sections 218
 - data filtering 6
 - data munging 4, 9, 17
 - and data structures 20
 - audit trail 37
 - business rules 25
 - command line scripts 53
 - examples 8
 - filter model 33
 - importance 7
 - processes 19, 22
 - producing output 23
 - record-oriented data 98
 - simple hash 21
 - sources of data 12
 - tasks 73
 - things to know 229–231
 - usefulness 228
 - data munging beast 1
 - data parsing 6
 - building parsers 209
 - complex data formats 147
 - HTML 163
 - XML 175
 - data pipes 9, 11
 - data processing 7
 - data recognition 5, 6, 16
 - data record 4
 - complex structure 110
 - data sink 9
 - data source 9, 19
 - data structure
 - building 216, 217
 - Data::Dumper 51
 - designing 20, 25, 38
 - examples of use, 23
 - munging processes 19, 22
 - Parse::RecDescent 218
 - parsing 6
 - read example 74
 - reading 82
 - simple hash 21
 - transformation 84
 - Data::Dumper 49, 56, 222
 - data_munger 33, 34, 35
 - data_reader 33, 34
 - data_writer 33, 34
 - database
 - and CSV 108
 - as a datasource 10, 47
 - auditing 36
 - caching data 106
 - combining data 19
 - communicating with 11
 - connecting 48
 - data source 9
 - design 7
 - read example 49
 - database driver (*see* DBD)
 - database handle 49
 - Database Interface (*see* DBI)
 - database schema 10
 - date
 - Date::Calc 120
 - daylight savings time 115
 - fields 114
 - formatting 118
 - functions 114
 - international formats 120
 - manipulation 117, 120, 121
 - manipulation examples 117, 118, 121
 - previous Saturday 121
 - date modules
 - benchmarking 122
 - choosing 122
 - Date::Calc 126
 - functions 240
 - Date::Manip 121, 126
 - functions 242
 - daylight savings time 115
 - DBD 47, 49, 55
 - DBD::mysql 48
 - DBI 11, 55
 - assigning attributes 234
 - connect function 47
 - functions 48, 233, 235

- DBI (continued)
 - handle 234
 - handle attributes 236
 - prepare function 49
 - sample program 48
 - DBM 10
 - Debug style 181
 - debugging 49
 - decimal places 91, 136
 - decoupling 19, 20, 22
 - defining record structure 133
 - field-end markers 134
 - fixed-width numbers 133
 - delimited data 101
 - delimiters 66
 - Descartes, Alligator 55
 - digit 62
 - DOCTYPE 194
 - Document Object Model
 - (see DOM)
 - Document Type Definitions
 - (see DTD)
 - dollar sign 64
 - DOM 191
 - domain specific constraints 25
 - Dominus, Mark-Jason 107, 228, 229
 - DOS 82
 - binary/text files 141
 - carriage return/line feed 88
 - dot character 62
 - double quotes 104
 - DTD
 - setting handlers 189
 - specialized parsers 193
 - valid documents 178
 - dump 172
 - Dumper 50, 51
- E**
-
- e command line option 53
 - e modifier 70, 72
 - EBCDIC
 - converting to ASCII 87
 - data conversion 87
 - ebcdic2ascii 87
 - Eckstein, Robert 177
 - Effective Perl Programming* 55
 - Elements of Programming with Perl* 17, 255
 - elsif 264
 - emacs 82
 - email header 64
 - encapsulate 38
 - business rules 25
 - encryption 77
 - end 64
 - end-of-record marker 132
 - English 14
 - epoch 114, 116
 - eq operator 262
 - equity research 8
 - escape sequences 62
 - eval 70, 181
 - exchange rates 106
 - Expat 178
 - exponential notation 90
 - Extensible Mark-up Language
 - (see XML)
 - extra data 112
- F**
-
- F command line option 54
 - fatal exception 181
 - fetchrow_array 49
 - field separator 112
 - field-end markers 134
 - fields 103, 112
 - extracting 100, 130
 - web access log 123
 - file handle 100
 - no end-of-record marker 132
 - output 104, 105
 - reading data 97
 - writing fixed-width data 136
 - file input operator 97, 98, 99
 - file open 15
 - file pointer 132
 - file transfer methods 9
 - File Transfer Protocol (see FTP)
 - file transfers 10
 - filename 36
 - filtering 4, 17
 - filter model 11, 32, 35
 - line end conversion 89
 - financial models 8
 - finite state machine 160
 - fixed-width data 101, 128–139
 - data item 128
 - defining record structure 133
 - encode field structures 135
 - end-of-record, example 132
 - end-of-record marker 132
 - extracting fields, example 129, 130
 - multiple record types 131, 133
 - pack 135
 - printf 136
 - reading 128
 - record description 128
 - record padding 128
 - sprintf 136
 - sprintf, example 137
 - template 130, 131
 - writing 135
 - fixing 91
 - floating point numbers 136
 - floating points 90
 - floppy disk 9
 - flushed 105
 - footer 5, 218, 221
 - foreach 113
 - form feed 62
 - format description 136
 - format specifier 137
 - format_bytes
 - example 93
 - format_negative
 - example 93, 94
 - format_number
 - example 93
 - format_picture
 - example 93
 - free beer 15
 - Free Software Foundation 15
 - free speech 15
 - Friedl, Jeffrey 78
 - FTP 9
 - fun 17
 - functions 26
 - Add_Delta_Days 242
 - available_drivers 233
 - bind_col 237
 - bind_columns 238
 - bind_param 236
 - bind_param_inout 236
 - called from DBI 233
 - called via statement
 - handle 236
 - check_date 241

functions (continued)
 chomp 98, 99, 112
 commit 236
 connect 48, 233
 data_sources 233
 Date_Cmp 243
 Date_DayOfWeek 244
 Date_DayOfYear 244
 Date_DaysInMonth 244
 Date_DaysInYear 244
 Date_DaySuffix 244
 Date_GetPrev 243
 Date_LeapYear 244
 Date_to_Text 242
 DateCalc 243
 Day_of_Week 241
 Day_of_Year 241
 Days_in_Month 240
 Days_in_Year 240
 Decode_Month 242
 Delta_Days 241
 Delta_DHMS 241
 do 235
 each 260
 Easter_Sunday 242
 errstr 234
 execute 237
 fetch 237
 fetchall_arrayref 237
 fetchrow_array 237
 fetchrow_arrayref 237
 fetchrow_hashref 237
 finish 237
 get 244
 get_rate 107
 getprint 245
 getstore 245
 gmtime 115
 handler 246
 head 244
 keys 260
 leap_year 241
 localtime 115, 116
 map 261
 memoize 107
 mirror 245
 Monday_of_Week 241
 new 245
 Nth_Weekday_of_Month_Year 241
 open 97
 pack 135

 parse 245
 parse_file 245
 ParseDate 122, 123
 ParseDateString 242
 ParseDelta 243
 ParseRecur 243
 prepare 49
 print 103, 115
 printf 135
 read 132
 ref 269
 return 268
 rollback 236
 seek 132
 select 104
 selectall_arrayref 235
 selectrow_array 235
 sprintf 135, 138
 strftime 118
 strict_comment 245
 strict_names 246
 substr 129, 132
 System_Clock 242
 time 116
 timelocal 116
 trace 234
 UnixDate 243
 unpack 130, 132, 133, 134
 values 260
 Week_Number 241
 xml_mode 246

G

 g modifier 69, 71, 85, 156
 gcc 15
 ge operator 262
 general principles 19
 get 164
 get_next_cust_no 27
 get_rate 107
 Getopt::Std 203
 getprint 164
 getstore 164, 172
 getting Perl 15
 GIF 139, 143
 creating 140
 gmtime 115
 GNU 15
 grammar 211, 214, 215
 debugging 220

 syntax errors 220
 Windows INI 213
 graphics files 139
 Graphics Interchange Format
 (see GIF)
 graphs 7
 greedy, definition 156
 grep 61
 group 63
 group of characters 61
 grouping 63
 gt operator 262
 Guttman, Uri 46
 Guttman-Rosler transform 46

H

 Hall, Joseph 42, 55
 handler 165
 Handlers 189
 hash 22, 86
 hash lookup 44
 hashes 259
 accessing values 260
 converting from a list 261
 hash slice 260
 reference 269
 syntax 260
 header 4, 218
 HTML tags 171
 processing 5
 processing, example of, 153
 rule 221
 hierarchical 17
 hierarchical data 13
 Hietaniemi, Jarkko 55
 Horatio 58
 HTML 12, 13, 173
 converting from XML 197
 entities 167
 extracting <h1> elements 169
 extracting from the web 164
 listing header tags 170
 listing specific links 168
 parsing 154
 parsing example 165, 171,
 172
 parsing links 168
 prebuilt parsers 167
 removing tags 154, 155
 tag 165

- HTML (continued)
 tag attributes 157
 testing the parser 166
 HTML::Element 171, 172
 HTML::LinkExtor 167, 168, 169
 functions 247
 HTML::Parser 154, 161,
 165, 174
 functions 245
 handlers 246
 HTML::Tokenizer 169, 173
 functions 248
 HTML::TreeBuilder 171
 functions 249
 HTTP 12
 extracting data WWW 164
 hyperreductionist English 76
 Hypertext Mark-up Language
 (see HTML)
- I**
-
- i 54
 -i command line option 54
 i modifier 68, 69
 I/O
 chaining 35
 independence 33, 38
 pipes 32
 redirection 31, 32
 IBM
 data conversion 87
 if 14, 263, 264
 Image::Info 143, 145
 index 78
 Information Systems 7
 Information Technology 7
 input 19
 record separator 55, 102
 record separator example
 111, 113
 record separator, reading 97
 routine 20
 integers 90, 136
 international date formats 120
 Internet 12
 interpolate 51
 interpreted 16
 interstices 8, 228
 IP addresses 46
 is_valid_sales_ref 28
- J**
-
- jaded 17
 Japan 87
 Jargon File 4
 Johnson, Andrew 17, 255
 JPG 143
- K**
-
- k3wl 76
 Kernigan, Brian 38
 keyboard 31
- L**
-
- label 54
 lc 60
 lcfirst 60
 le operator 262
 league table 4
 leaning toothpick syndrome 66
Learning Perl 17, 255
 left align 91
 lexer 160
 lexical variable 74
 Library for WWW Programming
 (see LWP)
 line end conversion 89
 line endings 88
 line feed 62, 88
 linguistics 14
 links, HTML 167
 Linux
 line feed 88
 list
 separator variable 103
 vs. arrays 258
 list functions 259
 listing particular users 76
 listing users 75
 load_products 35
 locale 60
 localtime 115, 118, 122
 Logfile 124, 125
 logic 25
 logical operators 263
 bitwise 263
 loops 264
 control of, 265
 for 264
 foreach 265
 last 265
 next 265
 redo 265
 while 265
 Lotus 139
 Loukides, Mike 38
 lower case 60
 lp 75
 lvalue 58
 LWP 172, 174
 getting HTML from the
 WWW 164
 mailing list 174
 LWP::Simple 164, 173
 functions 244
- M**
-
- M command line option 53
 m modifier 65, 66, 68, 69
 m/ 65, 78
 Macdonald, John 55
 make 15
 map 44
*Mastering Algorithms with
 Perl* 55
*Mastering Regular
 Expressions* 78
 match modifiers 65, 68
 match operator 68
 matching alternatives 62
 memoize 107
 Memoize.pm 107
 metacharacters 61
 metadata 133, 135, 150
 definition 151
 methods
 format_bytes 240
 format_negative 240
 format_number 239
 format_picture 240
 format_price 240
 round 239
 unformat_number 240
 Microsoft 139
 Access 10
 Excel 139
 Windows 31
 Word 61
 MIME 9

- modifiers 71
 - g 85
 - m 68
 - o 90
 - s 68
 - x 68
 - modules
 - business rules 26, 27, 38
 - convert::SciEng 91
 - Data::Calc 240
 - Date::Manip 242
 - HTML::Element 249
 - HTML::Element 171
 - HTML::LinkExtor 167, 247
 - HTML::Parser 154, 247
 - HTML::TokeParser 169, 248
 - HTML::TreeBuilder 171, 249
 - Image::Info 143
 - LWP 164
 - LWP::Simple 164, 244
 - MPEG::MP3Info 144
 - Number::Format 91, 238
 - Parse::RecDescent 153, 162
 - Parse::Yapp 162
 - POSIX 118
 - Time::Local 116
 - XML::DOM 191
 - XML::Parser 154, 161, 178, 250
 - monitor 31
 - MP3 files 143
 - ID3 data 144
 - MPEG::MP3Info 144, 145
 - multibyte characters 87
 - multiple record types 131
 - munge 4
 - munging 19
 - My Netscape 193
 - MySQL 48
- N**
-
- n command line option 53, 54
 - Named Pipe 11
 - Nandor, Chris 144, 193
 - natural code 228
 - natural numbers 90
 - ne operator 262
 - Network File System (*see* NFS)
 - newline character 113, 257
 - as field separator 112
 - input control 84
 - metacharacter 62
 - output records 102
 - record-oriented data 97
 - NFS 9
 - node 184
 - noncolon 64
 - nondigit 62
 - nongreedy 156
 - nonspace character 62
 - nonvowel 62
 - nonword character 62
 - normalization 7
 - Notepad 82
 - noun_phrase 210
 - Number::Format 91, 92, 94
 - attributes 238
 - example 93, 94
 - methods 239
 - numerical abbreviations 77
 - numerically 40
- O**
-
- o modifier 70, 90
 - O'Reilly, Tim 38
 - object 26, 38, 210
 - object class 28
 - Object Oriented Perl* 28, 38
 - Object style
 - vs. Tree 188
 - object-oriented programming
 - (*see* OOP)
 - objects 38
 - Objects style 186
 - OOP 28
 - open function 97
 - Open Source 140
 - openlog 37
 - operators 84
 - concatenation 261
 - logical 262
 - mathematical 261
 - string multiplication 261
 - optional 61
 - or operator 262
 - Oracle 10, 47
 - oraperl 47
 - Orcish Manoeuvre 42, 55, 107
 - Orwant, Jon 17, 55
 - output 19, 31
 - buffer 105
 - field separator 102, 103
 - file handle 105
 - list separator 102
 - record separator 54, 102
 - routine 19, 20
- P**
-
- p 54
 - p command line option 54
 - pack 135, 141, 144
 - packed-default sort 46
 - padding 128
 - paragraph mode 84, 113
 - parameter passing 267
 - parameters 58
 - parentheses 63, 65
 - parse tree 172
 - Parse::RecDescent 153, 162, 210, 216
 - \$::RD_AUTOACTION 215
 - \$::RD_HINT 220
 - \$::RD_TRACE 220
 - autotrees 223
 - dynamic rules 224
 - error handling 224
 - example 217
 - incremental parsing 224
 - look-ahead rules 223
 - precompiling parsers 224
 - subrule argument 224
 - Parse::Yapp 162
 - parse_file 165, 172
 - ParseDate 122, 123
 - parsefile 184, 192, 196
 - parser actions 217
 - parsers 154
 - adding actions 220
 - building your own 209
 - checking output 222
 - HTML 154
 - HTML::Element 171
 - HTML::LinkExtor 167
 - HTML::Parser 165
 - HTML::TokeParser 169
 - HTML::TreeBuilder 171
 - Parse::RecDescent 153, 210
 - Parse::Yapp 162
 - XML::DOM 191
 - XML::Parser 178
 - XML::RSS 193

- parsing 4, 17
 - (*see also* data parsing)
 - actions 214
 - bottom-up 160
 - building a data structure 216
 - definition 158
 - example, simple English sentences 210
 - failures 180
 - HTML 157, 158, 165
 - (*see also* HTML)
 - HTML, example 171
 - HTML, prebuilt parsers 167
 - jargon 159
 - LL parser 161
 - LR parser 160
 - parser object, creating 211
 - production 159
 - regular expressions 214
 - RSS file, example 196
 - structured data 157
 - subrules 159
 - terminal 159
 - testing for HTML 166
 - tokens 159
 - top-down 161
 - Windows INI 212
 - XML 178, 187, 189
 - (*see also* XML)
 - password 48
 - pattern matching 58, 60
 - payroll 4, 7
 - Peek, Jerry 38
 - performance 76
 - Perl 8, 14–16, 17
 - arrays 82, 129
 - benchmarking date modules 122
 - building parsers 210
 - command line options 255, 256
 - command line scripts 53
 - community 229
 - complex data structures 268
 - complex record structure 111
 - conditional execution 263
 - data munging advantages 16
 - data structure 16
 - data types 256
 - date and time manipulation 117
 - daylight savings time 115
 - dealing with complex records 110
 - default variable 99
 - end-of-record marker 132
 - fixed-width data 130
 - flow of control 263
 - getting HTML from the WWW 164
 - handling dates 114
 - HTML parsing 157
 - idioms 40
 - input record separator 112
 - international date formats 120
 - interpreter 48, 255, 256
 - line ends 89
 - local time 115
 - loops 264
 - multibyte characters 87
 - opening a file 262
 - operators 261
 - output file handles 105
 - parsers 161, 162
 - record-oriented data 97, 98
 - references 268
 - references and complex data structures 269
 - regular expressions 61, 65
 - running 255
 - sorting 40
 - string handling 58
 - subroutines 266
 - substrings 59
 - variables 256
 - web access logs 124
 - why use 228
 - writing objects 38
 - writing separate records 102
 - XML parsers 178
 - Perl Mongers 229
 - Perl Monks 229
 - Perl: The Programmer's Companion* 17, 255
 - perldata 261
 - perldoc 17, 55
 - perlfunc 126, 141, 144
 - perle 61, 77
 - perlrun 256
 - perlunicode 94
 - perlvar 94, 126
 - perltot 28
 - phrase matching 61
 - phrases 61
 - Pig Latin 76
 - Pike, Rob 38
 - pipe separated files 108
 - plain old documentation
 - (*see* POD)
 - plain text 12, 197
 - PNG files 140, 143
 - IHDR chunk 142
 - reading 140
 - signature 141
 - testing the reader 143
 - POD 197
 - pointers (*see* references)
 - pop 259
 - port number 12
 - Portable Network Graphics
 - (*see* PNG files)
 - positive integers 90
 - positive lookahead 65
 - POSIX 118
 - POSIX::strftime 120, 126
 - postfix 91
 - prebuilt parsers 161
 - print 58, 115
 - lists of items 103
 - record-oriented data 102
 - to different file handles 104
 - printf 135, 144
 - producing different document formats 197
 - production 159
 - program input 31
 - programmers 17
 - Programming Pearls* 38
 - Programming Perl* 17
 - Programming the Perl DBI* 55
 - Programming Web Graphics with Perl & GNU Software* 140
 - pronoun 210
 - pseudocode 19
 - punctuation 62, 66
 - push 259
- Q**
-
- quantifying matches 63
 - Quicksort 42
- R**
-
- RDBMS 10, 11
 - read 132, 144

- reading /etc/passwd 74
 - recognition 4, 17
 - recognizing numbers 90
 - record 101, 103, 110, 153
 - record based 73
 - record-oriented data 13, 17, 96–126
 - caching 105
 - CD data file 97
 - complex records 110
 - CSV 108 (*see also* CSV)
 - current record number 100
 - date fields 114
 - different file handles 104
 - extracting data fields 100
 - print lists of items 102
 - processing simple structures 100
 - reading 97
 - reading a record at a time 98
 - splitting into fields 102
 - web access logs 123
 - writing separate records 102
 - writing simple structures 102
 - redirection 31, 32
 - references 256, 272
 - complex data structures 269
 - creating 268
 - to subroutines 269
 - using 269
 - reformatting numbers
 - CPAN modules 91
 - sprintf 91
 - regular expressions 58, 60–77
 - and `S/` 114
 - complex 65
 - definition 60
 - delimiter 66
 - extracting fixed-width data 129
 - greedy 156
 - limitations 157
 - matching numbers 90
 - metacharacters 61
 - nongreedy 156
 - parsing 214
 - strings within strings 60
 - syntax 61
 - text replacement example 85
 - text transformation 84
 - using 65
 - relational database management system (*see* RDBMS)
 - repeated 61
 - replacement string 70, 71
 - requirements 21
 - rich site summary (*see* RSS)
 - right align 91
 - root 75
 - Rosler, Larry 46
 - rot13 77
 - round 93
 - example 93
 - RSS
 - creating with XML::RSS 195
 - definition 193
 - multiple channels 194
 - parsing 196
 - sample file 193
 - rule matching 215
- S**
-
- s modifier 68, 69
 - s/// 69, 78
 - save_cust_record 27, 28
 - scalars 115, 256
 - array element 260
 - context 97
 - variable 84
 - Schwartz, Randal 17, 43, 55, 255
 - Schwartzian transform 43, 45, 55
 - variation 46
 - scope 85
 - sed 61
 - seek 132
 - select 104
 - sentence 210, 212
 - separated data 101
 - separating parsing 22
 - setHandlers 189
 - Sethi 162
 - setlogmask 37
 - SGML 13
 - shift 185, 259
 - short-circuiting 41, 107, 263
 - SI 91
 - signature 140
 - simple sorts 40
 - sinks 17, 36
 - slash 61
 - Slashdot 193
 - slurp mode 113
 - sorting 40–47
 - and the Schwartzian transform 45
 - code 44
 - Orcish Manoeuvre 42
 - packed-default 46
 - sort function 40
 - sort key 44
 - source data 5, 17, 36
 - space 62, 64
 - SPICE 91
 - splice 185, 259
 - split 83, 101, 112, 113
 - spreadsheets 7, 108
 - sprintf 51, 92, 135, 138, 144
 - SQL 11, 49
 - square brackets 62
 - squares 70
 - Standardized General Mark-up Language (*see* SGML)
 - start 64
 - statement modifiers 264
 - statistics 86
 - STDOUT 104, 105
 - Stream style 181, 182
 - stream-based 182
 - strftime 118
 - strings 58, 59, 60
 - finding within strings 59
 - index 59
 - matching 65
 - replacement 69
 - index 59
 - structured data 17, 157
 - Structured Query Language (*see* SQL)
 - Style 179
 - sub 266
 - subject 210
 - subrecords 151
 - subroutines 40
 - reference to 269
 - subrules 159
 - optional, list of 213
 - repeating, list of 213
 - suffixes 213
 - Subs style 182
 - substitution modifiers 69
 - substr 58, 78, 129, 132
 - substrings 58, 59
 - successor state 160

Sybase 47
 Sybase Adaptive Server
 Anywhere 10
 Sybase Adaptive Server
 Enterprise 10
 syperl 47
 syntactic sugar 99
 synthetic code 228
 Sys::Syslog 37
 syslog 37
 System 242

T

-T command line option 256
 tab 62, 257
 tab character 103
 tables 7
 tab-separated data 101, 108
 tag 168
 tape 9
 TCP/IP 12
 TCP/IP Socket 12
 telephone book 41
 templates 130, 131, 135
 terminals 159, 211
 text 17
 editor 88
 file 141
 formats 4
 matching 78
 replacement 84
 statistics 85, 86
 substitution 70, 78
 transformation 84
 Text::Bastardize 76
 Text::CSV 109, 110
 Text::CSV_XS 109
 Text::Wrap 203
The Art of Compiler Design 162
The Dragon Book 162
 “The man of descent” 224
The Perl Cookbook 55
The Perl Journal 108, 224, 229
The Practice of Programming 38
The UNIX Programming Environment 38
 TIFF 143
 time 116
 (see also date)
 Date::Calc 120
 daylight savings time 115

 formatting 118
 functions 114
 manipulation 117, 120, 121
 Time::Local 116
 timelocal 116
 timethese 51, 52
 toascii 87
 toebcdic 87
 tokens 169
 definition 159
 Torkington, Nathan 55
 transaction-level locking 27
 transferring data 7
 transformation 4, 16, 17
 HTML output 206
 POD output 205
 text output 207
 XML 198
 transforming data 36
 translate 72
 translate_products 35
 translating data 85
 caching 105
 currency rates 106
 English to American 70
 example 72
 Tree style 182
 example 183, 185, 204
 vs. Objects style 186, 188
 two-digit year 116

U

uc 60, 78
 ucfirst 60
 Ullman 162
 unfixing 91
 unformat_number
 example 93
 Unicode 87
 Unicode::Map8 87
 Unicode::String 87
 Unisys 139
 UNIX 11, 14, 16, 33
 ASCII text files 82
 binary/text files 141
 data files 9
 data pipes 11
 databases 10
 epoch 114
 filter model overview 31
 (see also UNIX filter model)

 fortune 111
 I/O redirection character
 strings 32
 line feed 88
 POSIX 118
 system logs 37
 tools 61
 using system logs 37
 UNIX filter model 35, 38
 Schwartzian transform 45
 translations 35, 73
UNIX Power Tools 38
 unless 14, 264
 unpack 144
 by column widths 133
 example 134, 141, 153
 fixed-width data 130, 131
 no end-of-record marker 132
 unshift 259
 unstructured data 12, 17, 82
 upper case 60
 use locale 60
 username 48
 UTF-8 format 87
 utf8 module 87, 94

V

valid vs. well-formed 177
 verb 210
 vertical bar 62
 vi 61, 82
 vowel 62

W

-w command line option 256
 Wall, Larry 14, 17
 Wallace, Shawn P. 140
 war3z 76
 web access log
 date 124
 HTTP request 124
 IP address 123
 response code 124
 time 124
 URL 124
 web server 12, 123
 while 113
 \$_variable 99
 reading a PNG file 141

white space 62, 75, 83, 113
Windows 16
 ASCII text files 82
 binary/text files 141
 carriage return/line feed 88
 data files 9
 databases 10
 I/O redirection 32
 INI file grammar 213
 INI file parsing 212
word boundaries 64
word character 62
World Wide Web Consortium
 (W3C) 191
writing separate records 102

X

x modifier 68, 69
x operator 261
XML 8, 13, 154
 introduction 176
 parsers 178
 parsing failures 180
 parsing input 203
 parsing using handlers 189
 parsing, example 178, 191
 parsing, Objects style 187
 sample file 177
 tag 179
transformation script 198–205
Unicode 87

XML Pocket Reference 177
XML::DOM 191, 192
XML::Parser 154, 161
 CPAN alternatives 191
 Debug style 181
 example 178, 203
 functions 250
 RSS file 196
 Stream style 188
XML::RSS 193
 creating a file 195

Y

Yahoo! 177
Yorick 58