

Contents

Acknowledgments	xi
Preface	xiii
What to Expect from This Book	xiv
How This Book Is Organized	xv
What You'll Find in the Patterns	xvi
Who Should Read This Book	xvi
How to Read This Book	xvii
Languages and Tools Used in This Book	xviii
I Getting Started with Parsing	1
1 Language Applications Cracked Open	3
1.1 The Big Picture	3
1.2 A Tour of the Patterns	5
1.3 Dissecting a Few Applications	9
1.4 Choosing Patterns and Assembling Applications	17
2 Basic Parsing Patterns	21
2.1 Identifying Phrase Structure	22
2.2 Building Recursive-Descent Parsers	24
2.3 Parser Construction Using a Grammar DSL	26
2.4 Tokenizing Sentences	27
P.1. Mapping Grammars to Recursive-Descent Recognizers	29
P.2. <i>LL(1)</i> Recursive-Descent Lexer	33
P.3. <i>LL(1)</i> Recursive-Descent Parser	38
P.4. <i>LL(k)</i> Recursive-Descent Parser	43

3	Enhanced Parsing Patterns	49
3.1	Parsing with Arbitrary Lookahead	50
3.2	Parsing like a Pack Rat	52
3.3	Directing the Parse with Semantic Information	52
	P.5. Backtracking Parser	55
	P.6. Memoizing Parser	62
	P.7. Predicated Parser	68
II	Analyzing Languages	71
4	Building Intermediate Form Trees	73
4.1	Why We Build Trees	75
4.2	Building Abstract Syntax Trees	77
4.3	Quick Introduction to ANTLR	84
4.4	Constructing ASTs with ANTLR Grammars	86
	P.8. Parse Tree	90
	P.9. Homogeneous AST	94
	P.10. Normalized Heterogeneous AST	96
	P.11. Irregular Heterogeneous AST	99
5	Walking and Rewriting Trees	101
5.1	Walking Trees and Visitation Order	102
5.2	Encapsulating Node Visitation Code	105
5.3	Automatically Generating Visitors from Grammars	107
5.4	Decoupling Tree Traversal from Pattern Matching	110
	P.12. Embedded Heterogeneous Tree Walker	113
	P.13. External Tree Visitor	116
	P.14. Tree Grammar	119
	P.15. Tree Pattern Matcher	123
6	Tracking and Identifying Program Symbols	131
6.1	Collecting Information About Program Entities	132
6.2	Grouping Symbols into Scopes	134
6.3	Resolving Symbols	139
	P.16. Symbol Table for Monolithic Scope	141
	P.17. Symbol Table for Nested Scopes	146
7	Managing Symbol Tables for Data Aggregates	155
7.1	Building Scope Trees for Structs	156
7.2	Building Scope Trees for Classes	158
	P.18. Symbol Table for Data Aggregates	161
	P.19. Symbol Table for Classes	167

8	Enforcing Static Typing Rules	181
	P.20. Computing Static Expression Types	184
	P.21. Automatic Type Promotion	193
	P.22. Enforcing Static Type Safety	201
	P.23. Enforcing Polymorphic Type Safety	208
III	Building Interpreters	217
9	Building High-Level Interpreters	219
	9.1 Designing High-Level Interpreter Memory Systems	220
	9.2 Tracking Symbols in High-Level Interpreters	222
	9.3 Processing Instructions	224
	P.24. Syntax-Directed Interpreter	225
	P.25. Tree-Based Interpreter	230
10	Building Bytecode Interpreters	239
	10.1 Programming Bytecode Interpreters	241
	10.2 Defining an Assembly Language Syntax	243
	10.3 Bytecode Machine Architecture	245
	10.4 Where to Go from Here	250
	P.26. Bytecode Assembler	252
	P.27. Stack-Based Bytecode Interpreter	259
	P.28. Register-Based Bytecode Interpreter	267
IV	Translating and Generating Languages	277
11	Translating Computer Languages	279
	11.1 Syntax-Directed Translation	281
	11.2 Rule-Based Translation	282
	11.3 Model-Driven Translation	284
	11.4 Constructing a Nested Output Model	292
	P.29. Syntax-Directed Translator	296
	P.30. Rule-Based Translator	302
	P.31. Target-Specific Generator Classes	308
12	Generating DSLs with Templates	313
	12.1 Getting Started with StringTemplate	314
	12.2 Characterizing StringTemplate	317
	12.3 Generating Templates from a Simple Input Model	318
	12.4 Reusing Templates with a Different Input Model	321

► CONTENTS

12.5	Using a Tree Grammar to Create Templates	324
12.6	Applying Templates to Lists of Data	331
12.7	Building Retargetable Translators	337
13	Putting It All Together	349
13.1	Finding Patterns in Protein Structures	349
13.2	Using a Script to Build 3D Scenes	350
13.3	Processing XML	351
13.4	Reading Generic Configuration Files	353
13.5	Tweaking Source Code	354
13.6	Adding a New Type to Java	355
13.7	Pretty Printing Source Code	356
13.8	Compiling to Machine Code	357
A	Bibliography	359
	Index	361