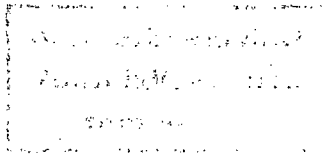




RELAX NG



Eric van der Vlist

O'REILLY®

Beijing • Cambridge • Farnham • Köln • Paris • Sebastopol • Taipei • Tokyo

Table of Contents

Foreword by James Clark	ix
Foreword by Murata Makoto	xi
Preface	xiii

Part I. Tutorial

1. What RELAX NG Offers	3
Diversity	3
Keeping Documents Independent of Applications	3
Validation Has Many Aspects	4
The Best Way to Validate XML Document Structures	4
RELAX NG's Diverse Applications	5
RELAX NG as a Pivot Format	6
Why Use Other Schema Languages?	7
2. Simple Foundations Are Beautiful	8
Documents and Infosets	8
Different Types of Schema Languages	9
A Simple Example	10
A Strong Mathematical Background	12
Patterns, and Only Patterns	13

3. First Schema	14
Getting Started	14
First Patterns	15
Complete Schema	20
4. Introducing the Compact Syntax	24
First Compact Patterns	25
Full Schema	28
XML or Compact?	29
5. Flattening the First Schema	31
Defining Named Patterns	33
Referencing Named Patterns	34
The grammar and start Elements	35
Assembling the Parts	36
Problems That Never Arise	42
Recursive Models	43
Escaping Named Pattern Identifiers in the Compact Syntax	44
6. More Complex Patterns	45
The group Pattern	45
The interleave Pattern	46
The choice Pattern	48
Pattern Compositions	49
Order Variation as a Source of Information	50
Text and Empty Patterns, Whitespace, and Mixed Content	51
Why Is It Called interleave?	54
Mixed Content Models with Order	57
A Restriction Related to interleave	59
A Missing Pattern: Unordered Group	61
7. Constraining Text Values	62
Fixed Values	62
Co-Occurrence Constraints	62
Enumerations	70
Whitespace and RELAX NG Native Datatypes	70
Using String Datatypes in Attribute Values	72
When to Use String Datatypes	73
Using Different Types in Each Value	74

Exclusions	74
Lists	76
Data Versus Text	78
8. Datatype Libraries	80
W3C XML Schema Type Library	80
DTD Compatibility Datatypes	98
Which Library Should Be Used?	101
9. Using Regular Expressions to Specify Simple Datatypes	104
A Swiss Army Knife	104
The Simplest Possible Pattern Facets	105
Quantifying	106
More Atoms	107
Common Patterns	115
10. Creating Building Blocks	120
Using External References	120
Merging Grammars	134
A Real-World Example: XHTML 2.0	149
Other Options	153
11. Namespaces	160
A Ten-Minute Guide to XML Namespaces	160
The Two Challenges of Namespaces	165
Declaring Namespaces in Schemas	166
Accepting Foreign Namespaces	171
Namespaces, Building Blocks, and Chameleon Design	179
12. Writing Extensible Schemas	188
Extensible Schemas	188
The Case for Open Schemas	206
Extensible and Open?	208
13. Annotating Schemas	211
Common Principles for Annotating RELAX NG Schemas	211
Documentation	226
Annotation for Applications	238

14. Generating RELAX NG Schemas	250
Examplotron: Instance Documents as Schemas	251
Literate Programming	258
UML	264
Spreadsheets	270
15. Simplification and Restrictions	275
Simplification	275
Restrictions	295
16. Determinism and Datatype Assignment	305
What Is Ambiguity?	305
The Downsides of Ambiguous and Nondeterministic Content Models	315
Some Ideas to Make Disambiguation Easier	318

Part II. Reference

17. Element Reference	323
18. Compact Syntax Reference	371
19. Datatype Reference	409

Part III. Appendixes

A. DSDL	451
B. The GNU Free Documentation License	458
Glossary	467
Index	473