

---

# **NATURAL LANGUAGE PROCESSING IN LISP**

---

## **An Introduction to Computational Linguistics**

**Gerald Gazdar**

University of Sussex

**Chris Mellish**

University of Edinburgh



**Addison-Wesley Publishing Company**

Wokingham, England · Reading, Massachusetts · Menlo Park, California

New York · Don Mills, Ontario · Amsterdam · Bonn

Sydney · Singapore · Tokyo · Madrid · San Juan

---

# CONTENTS

---

## Preface

v

### 1

#### Introduction

1

- 1.1 The origins of natural language processing 2
- 1.2 The imposition of structure 5
- 1.3 The representation of meaning 8
- 1.4 The role of knowledge 12
- 1.5 The emergence of a new technology 15
- 1.6 Using LISP for natural language processing 17

### 2

#### Finite-state techniques

21

- 2.1 Finite-state transition networks 22
- 2.2 A notation for networks 27
- 2.3 Deterministic FSTNs in LISP 37
- 2.4 Non-deterministic FSTNs in LISP 38
- 2.5 Traversing FSTNs 41
- 2.6 Traversing FSTNs in LISP 44
- 2.7 Finite-state transducers 50
- 2.8 Limitations of finite-state machines 59

### 3

#### Recursive and augmented transition networks

63

- 3.1 Recursive transition networks 64
- 3.2 Modelling recursion in English grammar 68
- 3.3 Traversing RTNs 72
- 3.4 RTNs in LISP 78
- 3.5 Pushdown transducers 82
- 3.6 Advantages and limitations of RTNs 88
- 3.7 Augmented transition networks 89
- 3.8 Developing ATNs 93
- 3.9 Implementation of ATNs 99
- 3.10 Some reflections on ATNs 107

<b>4</b>	<b>Grammars</b>	<b>111</b>
4.1	Grammar as knowledge representation	112
4.2	Words, rules and structures	116
4.3	Grammars in LISP and random generation	122
4.4	Subcategorization and the use of features	127
4.5	Encoding feature specifications in LISP	138
4.6	Classes of grammars and languages	142
<b>5</b>	<b>Parsing, search and ambiguity</b>	<b>149</b>
5.1	A simple parsing problem	150
5.2	Bottom-up parsing	151
5.3	Top-down parsing	161
5.4	Comparing strategies	166
5.5	Breadth-first and depth-first search	168
5.6	Storing intermediate results	170
5.7	Ambiguity	172
5.8	Determinism and lookahead	176
<b>6</b>	<b>Well-formed substring tables and charts</b>	<b>181</b>
6.1	Well-formed substring tables	182
6.2	The active chart	191
6.3	The fundamental rule of chart parsing	195
6.4	Initialization	201
6.5	Rule invocation	203
6.6	Search strategy	205
6.7	Housekeeping	207
6.8	Alternative rule invocation strategies	209
6.9	Efficiency	211
<b>7</b>	<b>Features and the lexicon</b>	<b>215</b>
7.1	Feature-theoretic syntax	216
7.2	Feature structures as graphs	219
7.3	Feature structures in LISP	226
7.4	Subsumption and unification	228
7.5	The status of rules	237
7.6	Representing PATR grammars in LISP	239
7.7	Random generation revisited	244
7.8	Chart parsing with feature-based grammars	249
7.9	Representation of lexical knowledge	260
7.10	Implementing a lexicon in LISP	273
7.11	DAGs versus terms	276

<b>8</b>	<b>Semantics</b>	<b>283</b>
8.1	Compositionality	284
8.2	Meaning as reference	286
8.3	Translation to a meaning representation language	292
8.4	A database query language	294
8.5	Computational semantics as feature instantiation	297
8.6	Transitive verbs and quantification	299
8.7	Ambiguity, preferences and timing	305
8.8	Building semantic checking into the grammar	307
<b>9</b>	<b>Question answering and inference</b>	<b>317</b>
9.1	Question answering	318
9.2	Evaluating DBQ formulae	321
9.3	Standard logical inference	330
9.4	Implementing backwards and forwards inference	338
9.5	The pathological nature of logical inference	344
9.6	Primitives and canonical forms	348
9.7	Inheritance and defaults	351
9.8	A simple semantic network in LISP	357
<b>10</b>	<b>Pragmatics</b>	<b>363</b>
10.1	Ambiguity and levels of language processing	364
10.2	Semantic and pragmatic roles of noun phrases	365
10.3	Given versus new information	368
10.4	Understanding by prediction	376
10.5	More controlled versions of prediction	380
10.6	Problems with prediction	385
10.7	Using discourse structure	386
10.8	Language generation as a goal-oriented process	390
10.9	Language understanding and plan recognition	402
<b>Appendix Code listings</b>		<b>407</b>
<b>Solutions to selected exercises</b>		<b>487</b>
<b>Bibliography</b>		<b>495</b>
<b>Name Index</b>		<b>517</b>
<b>General Index</b>		<b>521</b>