

Inhaltsverzeichnis

Vorwort	v
Einleitung	vii
Symbolverzeichnis	xviii
I Grundlagen	1
1 Was ist Parsing?	3
1.1 Parsing und Compilerbau	5
1.2 Parsing und Kognitive Psychologie	10
1.3 Parsing in der Computerlinguistik	13
1.3.1 Die Rolle von Parsern in natürlichsprachlichen Systemen . . .	13
1.3.2 Parsing natürlicher Sprache	14
1.3.3 Grammatikformalismen	15
2 Begriffe & Repräsentationen	19
2.1 Einige grundlegende Begriffe	19
2.2 Zur Klassifikation von Parsingalgorithmen	22
2.2.1 Verarbeitungsrichtung	22
2.2.2 Analyserichtung	23
2.2.3 Suchstrategien	24
2.3 Komplexität und Effizienz	27
2.3.1 Zur Bewertung von empirischen Tests von Parsingsystemen .	28
2.4 Repräsentation des linguistischen Wissens	30
2.5 Die Beschreibung von Algorithmen und Prozeduren	33
2.5.1 Der Beschreibungskopf	33
2.5.2 Der Spezifikationsteil	34
2.5.3 Der Operationsteil	34

II	Elementare Analysealgorithmen	37
3	Top-down-Parsing	39
3.1	Grundzüge	39
3.2	Kontrollstrukturen	44
3.2.1	Depth-first-Parsing mit Backtracking	46
3.2.2	Breadth-first-Parsing ohne Backtracking	50
3.3	Grenzen des Top-down-Parsings	52
3.4	Implementierung	54
3.4.1	Lisp	54
3.4.2	Prolog	59
4	Bottom-up-Parsing	63
4.1	Grundzüge	63
4.2	Ein deterministischer Shift-reduce-Recognizer	64
4.3	Ein Shift-reduce-Recognizer mit Backtracking	67
4.3.1	Daten und Prozeduren	68
4.3.2	Backtracking	70
4.4	Vom Recognizer zum Parser	74
4.5	Ein Shift-reduce-Parser in Prolog	74
4.5.1	Die Repräsentation der Stacks und Regeln	74
4.5.2	Programmstruktur	76
4.5.3	Das Programm im Überblick	77
4.6	Ein breadth-first arbeitender Shift-reduce-Recognizer	77
4.7	Ein Shift-reduce-Recognizer in Lisp	80
5	Left-corner-Parsing	83
5.1	Grundzüge	83
5.2	Ein Left-corner-Erkennungsalgorithmus	84
5.3	Ein Left-corner-Parsingalgorithmus	89
5.4	Left-corner-Parsing mit Look-ahead	93
5.5	Left-corner-Parsing in Lisp	95
5.5.1	Der Recognizer	95
5.5.2	Der Parser	96
5.6	Das BUP-System	98
5.6.1	Das Kernprogramm als Recognizer	98
5.6.2	Optimierung: Die LINK-Relation	100
III	Chart-Parsing	101
6	Bottom-up-Chart-Parser	107
6.1	Ein genereller Bottom-up-Chart-Parser	107

6.2	Der Cocke-Kasami-Younger-Algorithmus	110
6.2.1	Berechnung einer einfachen Analysematrix	111
6.2.2	Berechnung einer komplexen Analysematrix	114
6.2.3	Zeit- und Speicherbedarf des Cocke-Kasami-Younger-Algorithmus	116
6.3	Implementierung der Algorithmen	116
6.3.1	Prolog-Implementierung des generellen Bottom-up-Chart-Parsers	116
6.3.2	Der Cocke-Kasami-Younger-Algorithmus	118
7	Der Earley-Algorithmus	123
7.1	Unterschiede zwischen dem CKY- und dem Earley-Algorithmus . . .	123
7.2	Darstellung des Erkennungsalgorithmus	124
7.2.1	Version 1: Berechnung der Chart durch Vorwärtsverkettung .	129
7.2.2	Version 2: Zwischenspeicherung neuer Kanten	132
7.3	Erweiterung des Algorithmus	134
7.3.1	Tilgungs- und Kettenregeln	134
7.3.2	Look-ahead	136
7.4	Die Generierung von Strukturbeschreibungen	138
7.4.1	Explizite Repräsentation	138
7.4.2	Implizite Repräsentation	140
7.5	Implementierung der Algorithmen	141
7.5.1	Prolog	142
7.5.2	Lisp	144
8	Insel-Parsing	155
8.1	Motivation für Insel-Parsing	155
8.2	Notwendige Modifikationen des Earley-Algorithmus	157
8.2.1	Datenstrukturen und Operationen	158
8.2.2	Berechnung der Chart durch Vorwärtsverkettung	162
8.2.3	Zwischenspeicherung der neu generierten Kanten	163
8.3	Implementierung	165
8.3.1	Kantenselektoren	165
8.3.2	Identifikation der Inseln im Satz	165
8.3.3	Die Hauptfunktionen	166
9	ID/LP-Syntaxen	169
9.1	Motivation für den ID/LP-Formalismus	169
9.2	Parsen mit ID/LP-Syntaxen	173
9.2.1	Indirektes Parsen	173
9.2.2	Direktes Parsen	174
9.3	Implementierung	178
9.3.1	Lisp	178

IV	Deterministisches Parsen	181
10	LR-Parsing	183
10.1	LR-Syntaxen	183
10.2	LR-Parser	186
10.3	Ein Algorithmus zur Generierung kanonischer LR(1)-Tabellen	188
10.4	Effiziente Repräsentation von LR-Parsingtabellen	191
10.5	Implementierung	193
10.5.1	Lisp	193
10.5.2	Prolog	195
11	Der Tomita-Algorithmus	197
11.1	LR-Parsing beliebiger kontextfreier Syntaxen	197
11.1.1	Stack-Listen	198
11.1.2	Baum-strukturierter Stack	200
11.1.3	Graph-strukturierter Stack	201
11.1.4	Effiziente Strukturverwaltung	202
11.2	Der Algorithmus	204
11.2.1	Der Erkennungsalgorithmus	204
11.2.2	Der Parsingalgorithmus	212
11.3	Komplexität	216
11.4	Implementierung	216
V	Unifikationsbasierte Grammatikformalismen	223
12	Definite Clause Grammatiken	233
12.1	Syntaktische Regeln als logische Ausdrücke	233
12.2	Differenzlisten	234
12.3	Vom Recognizer zum Parser	236
12.4	Nicht-kontextfreie Erweiterungen	236
12.5	Ein earley-basierter Parser für DCGs	239
13	Ein Chart-Parser für unifikationsbasierte Syntaxen	247
13.1	Indirektes Parsen	247
13.2	Direktes Parsen	252
13.2.1	Repräsentation von Regeln und Kategorien	252
13.2.2	Parsen mit merkmalsbasierten Syntaxen	254
13.3	Implementierung	260

Inhaltsverzeichnis	xvii
A Grundbegriffe der Syntaxtheorie	265
B Graphentheoretische Begriffe	269
C Datenstrukturen	271
D Repäsentation von Syntax und Lexikon in LISP	275
E Musterlösungen	277
Literaturverzeichnis	285
Namens- und Sachverzeichnis	299