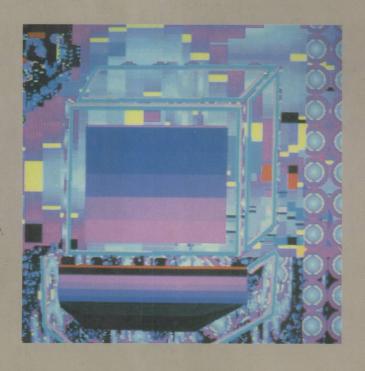


Logic Programming and Knowledge Engineering Tore Amble



Logic Programming and Knowledge Engineering

Tore Amble





E8863951



Wokingham, England · Reading, Massachusetts · Menlo Park, California New York · Don Mills, Ontario · Amsterdam · Bonn · Sydney Singapore · Tokyo · Madrid · Bogota · Santiago · San Juan

- © 1987 Addison-Wesley Publishers Ltd.
- © 1987 Addison-Wesley Publishing Company, Inc.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior written permission of the publisher.

The programs in this book have been included for their instructional value. They have been tested with care but are not guaranteed for any particular purpose. The publisher does not offer any warranties or representations, nor does it accept any liabilities with respect to the programs.

Cover graphic by Laurence M. Gartel. Typeset by Columns of Reading. Printed and bound in Great Britain by T J Press (Padstow) Ltd, Cornwall.

First printed 1987.

British Library Cataloguing in Publication Data

Amble, Tore

Logic programming and knowledge engineering.

- —(International computer science series).
- 1. Expert systems (Computer science)
- 2. Electronic digital computers—

Programming 3. Logic, symbolic and mathematical

I. Title II. Series

005.13'1 OA76.76.E95

ISBN 0-201-18043-X

Library of Congress Cataloging in Publication Data

Amble, Tore, 1945-

Logic programming and knowledge engineering.

(International computer science series)

Bibliography: p.

Includes index.

1. Logic programming. 2. Expert systems (Computer science) I. Title. II. Series.

QA76.6.A464 1987 005.1 87-13556

ISBN 0-201-18043-X

Logic Programming and Knowledge Engineering

子司 心光子主意

INTERNATIONAL COMPUTER SCIENCE SERIES

Consulting editors

A D McGettrick University of Strathclyde

J van Leeuwen University of Utrecht

OTHER TITLES IN THE SERIES

Programming in Ada (2nd Edn.) J G P Barnes

Software Engineering (2nd Edn.) I Sommerville

An Introduction to Numerical Methods with Pascal LV Atkinson and PJ Harley

The UNIX System S R Bourne

Handbook of Algorithms and Data Structures GH Gonnet

Microcomputers in Engineering and Science J F Craine and G R Martin

UNIX for Super-Users E Foxley

Software Specification Techniques N Gehani and A D McGettrick (eds.)

Introduction to Expert Systems P Jackson

Data Communications for Programmers M Purser

Local Area Network Design A Hopper, S Temple and R C Williamson

Modula-2: Discipline & Design A H J Sale

The UNIX System V Environment SR Rourne

Prolog Programming for Artificial Intelligence I Bratko

Prolog F Giannesini, H Kanoui, R Pasero and M van Caneghem

Programming Language Translation: A Practical Approach PD Terry

Data Abstraction in Programming Languages J M Bishop

System Simulation: Programming Styles and Languages W Kreutzer

The Craft of Software Engineering A Macro and J Buxton

UNIX System Programming K F Haviland and B Salama

An Introduction to Programming with Modula-2 PD Terry

Pop-11 Programming for Artificial Intelligence A M Burton and N R Shadbolt

The Specification of Computer Programs W M Turski and T S E Maibaum

Software Development with Ada I Sommerville and R Morrison

UNIXTM is a trademark of AT & T.

Preface

Pioneering research in computer science in the 1970s has led to a revolution in information technology this decade. Two of the most important new directions in software engineering are logic programming and knowledge engineering. They both embody the basic change in computing – a move away from program-controlled computing to computing derived by deduction and reasoning.

The key idea behind logic programming is to replace programming and computation by a logical description of a problem and an automatic proof mechanism for deducing the answers.

The basis of knowledge engineering is to make intelligent, problem-solving expert systems knowledge-based by finding and applying the rules and heuristics that govern experts' problem-solving processes. Knowledge-based systems differ from traditional programs in that they are derived directly from these rules, with no intervening programming.

The combination of these two ideas will change the future of software engineering and replace ideas that have become too formalized, inflexible and problematic. Computer technology will adapt itself to support this development.

This book is written for students and professionals with an interest in engineering, who need a theoretical as well as practical introduction to logic programming and how it can be used to build knowledge-based systems. It is suitable for an undergraduate course at third or fourth year level. For complete understanding, it requires two years of programming experience with some knowledge of Pascal, but parts of the book should be comprehensible to a wider readership.

The book falls roughly into two parts: an introduction to logic programming; and applications of logic programming. The first part introduces predicate logic, resolution and the programming language Prolog. It looks in detail at a subset of the language, list processing in Prolog and advanced programming techniques, with the emphasis on meta-level logic programming. The second part of the book looks at the application of Prolog for formula manipulation, including program verification, and as a sophisticated query language for relational databases. It also describes how to process formal languages, including compiler writing; how to process an interesting and useful subset of natural language; and how to apply knowledge for problem-solving. In addition, it describes an expert system shell in Prolog, with several applications, and gives a short overview of a knowledge engineering project using these ideas.

The appendix contains a collection of useful Prolog predicates.

The language used for the programming examples will be a subset of the Edinburgh Prolog, specifically the C-Prolog VAX version (Pereira, 1984) under the VMS operating systems. The language subset is kept simple, so that the ideas of logic programming are easier to follow. This text, therefore, is not intended to be a reference manual of the fast, user-friendly Prolog system.

Acknowledgements

This book about logic programming and knowledge engineering is a result of 12 years of research and teaching at the University of Trondheim's RUNIT Computing Centre and Division of Computing Science. It is a sequel to an earlier version (Amble, 1984). Many of my colleagues deserve to be mentioned for their help with that version, but I must confine the list to those who have been particularly helpful with the current volume.

First, I am grateful to my colleague Haakon Styri for his encouragement and good advice, based on his impressive knowledge of all aspects of logic programming.

I was fortunate enough to have the opportunity of working on two different projects, both applying logic programming for knowledge engineering. The first was an expert system for diagnosing and teaching television repair; in this field, Roger Eide is an enthusiastic expert and teacher. The other project was to make an expert system for designing welding processes. On this project, I had the pleasure of working with Bjarte H. Nes, who is co-author of the chapter on knowledge engineering. Both projects confirm that the idea of applying logic programming for knowledge engineering works.

Among the people who have scrutinized the manuscripts at various stages, Arild Waaler and Catherine Churchill deserve thanks for their helpful corrections and comments.

A draft version of the book was used for a logic programming course at the University of Trondheim in the fall of 1986. I gained invaluable feedback from the class.

Finally, my bad conscience leads me to express my gratitude towards my wife and children: young children's parents shouldn't really write books.

Figure 13.4 is taken from Hayes-Roth, Waterman and Lenat, *Building Expert Systems*, © 1983, Addison-Wesley Publishing Company Inc., Reading, Massachusetts (p. 389, Figure A.1), and is reprinted with permission.

8863951

Contents

		Preface	v
Chapter	1	Introduction	1
	1.1	Make specifications, not programs	1
	1.2	Symbolic processing language	2
	1.3	Fifth generation computer systems	3
	1.4	, , , ,	4
		1.4.1 Aristotelian logic	4
		1.4.2 Symbolic logic	5
		1.4.3 Logic programming	7
	1.5	Artificial intelligence	8
		1.5.1 The limits of mind	9
		1.5.2 Knowledge-based systems	9
		1.5.3 Expert systems1.5.4 Knowledge engineering	10 11
Chapter	2	Introduction to Logic	12
	2.1	8	12
	2.2	2.2.1 The elimination rule	15
		2.2.2 Clausal form	16
		2.2.3 Refutation proofs	17
	2.3	The state of the s	20
		2.3.1 Predicates and arguments	21
		2.3.2 Quantifier-free notation	22
		2.3.3 Formalizing queries and contradictions	22
		2.3.4 Horn clause resolution	23
		2.3.5 Alternative proof strategies	24
		2.3.6 Functions in predicate logic	25
		2.3.7 Unification of functional terms	26
		2.3.8 Logic programming	26
Chapter	3	Resolution	29
	3.1	Some logical concepts	29
	3.2	Quantifiers	31
		3.2.1 Examples of quantifiers	32
		3.2.2 Second order logic concepts	32
	3.3	First order predicate calculus	33
		3.3.1 Skolem functions	34
			vii

viii CONTENTS

		3.3.2 From predicate logic to clausal form	35
		3.3.3 Clause normalization algorithm	36
		3.3.4 The complete unification algorithm	38
	3.4		39
		3.4.1 Resolution step	39
		3.4.2 Resolution proof	40
		3.4.3 Resolution proof search strategies	41
Chapter	4	Predicate Logic as a Programming Language	44
	4.1	Syntax	44
		4.1.1 Basic syntax	44
		4.1.2 Functions	45
		4.1.3 Clause syntax	45
		4.1.4 Program structure	45
		4.1.5 Describing predicates	46
		4.1.6 Input/output and comments	46
	4.2	The semantics of Prolog	46
	4.3	Search tree	48
	4.4	Recursion	48
	4.5	On variable bindings	50
		4.5.1 Anonymous variables	51
		4.5.2 Renaming variables	51
		4.5.3 Occur check	52
	4.6	J F F F F F F F F F F F F F F F F F F F	53
		4.6.1 Symmetry of sequence of conclusions	53
		4.6.2 Symmetry of sequence of conditions	53
		4.6.3 Test-or-generate symmetry	54
		4.6.4 Input/output parameter symmetry	54
	4.7	8	54
	4.8	0	56
		4.8.1 Negation as failure	56
		4.8.2 Cut unnecessary search	58
		4.8.3 Cut destroys symmetry	58
		4.8.4 Variable conditions	59
		4.8.5 Equality and inequality	60
Chapter	5	Programming in Prolog	62
	5.1	Predicate library	62
	5.2	Interactive Prolog	63
	5.3	Basic input and output	65
	5.4	Built-in operators in Prolog	66
	5.5	Evaluation of expressions	67
	5.6	Query processing	68
	5.7	Manipulating the database	69
		5.7.1 Assert	69
		5.7.2 Retract	70
	5.8	Operator declarations	71

			CONTENTS	ix
		5.8.1 Extralogical features		75
Chapter	6	List Processing		78
	6.1	List processing		78
		S-expression		78
		The empty node		79
	6.4	List notation		80
		6.4.1 Transforming list notation to dot notation		82
		6.4.2 Transforming dot notation to list notation		83
		6.4.3 Extension to Prolog S-expression		83
	6.5	J 1		84
		6.5.1 The cons predicate		84
		6.5.2 The member predicate		85
		6.5.3 The append predicate6.5.4 The delete predicate		86 87
		6.5.5 Naive reverse		88
		6.5.6 Smart reverse		88
	6.6	Lists and sets		89
	0.0	6.6.1 Representing information as lists or as facts	s	91
		6.6.2 The unexpected nature of the built-in setof		92
		6.6.3 Set construction without databases		94
	6.7	D-lists		95
		6.7.1 D-list manipulation		96
		6.7.2 Limitations of Prolog list structures		97
	6.8	11		97
		6.8.1 Mergesort		97
		6.8.2 Quicksort		98
	6.9	Alternative list syntax		99
		6.9.1 Strings		99
		6.9.2 Round lists		100
		6.9.3 List processing with round lists		101
Chapter	7	Logic Programming Techniques		104
	7.1	Constructing recursive programs		105
		7.1.1 A closer look at recursion		105
		7.1.2 Path problems		106
		7.1.3 Finding the path		107
		Constructing iterative programs		108
~	7.3	Possible implications		112
		7.3.1 An application: Mastermind		112
	7.4	The cut operator considered harmful		114
		7.4.1 Examples of hazardous cuts		115
	7.5	7.4.2 Structured use of cut		115
	7.5	Resolution preprocessing		116
	7.6	Inversion		117
	7.7 7.8	Non-Horn logic programming		119
	1.0	Meta-programming		121

x CONTENTS

	7.8.1 Element by element application	122
	7.8.2 Aggregate functions	123
1	7.9 Meta-logic	124
	7.9.1 Explaining facility in meta-level logic	126
Chapter	8 Formula Manipulation	130
8	8.1 Symbolic differentiation	130
	8.2 Manipulation	131
8	8.3 Anatomy of operator expressions	131
8	8.4 Formula evaluation	133
8	8.5 Algebraic simplification	135
	8.5.1 Common subexpressions	137
8	8.6 Integration	138
8	8.7 Program verification	139
	8.7.1 Program verification in Prolog	141
	8.7.2 A verification condition generator	141
Chapter	9 Logic and Databases	146
ç	9.1 Relational databases	146
	9.1.1 A relational example	147
	9.1.2 Binary relations	148
	9.1.3 Composite keys	148
9	9.2 Database retrieval	149
	9.2.1 Efficient retrieval	149
	9.2.2 Virtual tables	150
	9.2.3 Symbolic naming	151
9	9.3 Database updating	153
9	9.4 Data modelling	154
	9.4.1 Normal forms	155
	9.4.2 Relational normal forms	155
	9.5 Beyond the relational model	158
9	9.6 Semantic nets	158
	9.6.1 The class concept	159
9	7.7 The course model	163
	9.7.1 Coupling semantic nets to tables	164
	9.7.2 Typical questions	165
Chapter 10	0 Logic Programming and Compiler Writing	167
10	1.1 Language processing	167
10	2.2 Lexical analysis	167
10		170
	10.3.1 Clause grammar	173
	10.3.2 Table-driven parsing	175
	10.3.3 Constructing a syntax tree	177
	10.3.4 Prettyprinting a syntax tree	177
10.	.4 Semantics and production	179

	CON	NTENTS	X1
10.5	Advanced grammar formalisms		182
	10.5.1 Two-level grammars		182
	10.5.2 Attribute grammars		183
Chapter 11	Natural Language Processing		188
Chapter 11	Natural Language Processing		
11.1	What is natural language?		188
11.2			188
11.3	0 0 ,		189
	11.3.1 Definite clause grammars		190
	11.3.2 Natural language is ambiguous		190
11.4	Soft Systems		194
	11.4.1 The Soft Systems language		194
	11.4.2 Sample questions		196
	11.4.3 The dialogue context		196
	11.4.4 The reference model		197
	11.4.5 Lexical analysis		198
	11.4.6 Syntax analysis		200
	11.4.7 A short attribute grammar for de-verbed		
	language		200
	11.4.8 Semantic analysis with semantic nets		202
	11.4.9 Query processing		203
11.5	Pure natural language		204
	11.5.1 A logic for commonsense knowledge		204
	11.5.2 Why do we do what we do?		205
	11.5.3 A Prolog program for a room situation		206
11.6	Natural language processing in the future		207
Chapter 12	Logic for Problem Solving		209
12.1	What is the problem?		209
12.2			209
12.3			210
12.4	Knowledge for problem solving		211
12. 1	12.4.1 Generate-and-test		213
	12.4.2 Generate-or-test		214
12.5	A meta-problem solver		216
12.6	Robot planning		218
	12.6.1 Kowalski's formulation		218
	12.6.2 Linear planning		222
12.7	Using estimates to guide searches		222
	12.7.1 Stepwise increasing length of solution		223
	12.7.2 Finding short paths		224
	12.7.3 Making a plan before execution		226
Chapter 13	Expert Systems		229
-	•		
13.1	Expert systems		229
13.2	Expert systems in Prolog		230

xii CONTENTS

13.3	Principles of the EXPLAIN expert system shell	231
10.0	13.3.1 Why and how – explanation	231
13.4	•	232
13.5		235
10.0	13.5.1 Important predicates	235
	13.5.2 EXPLAIN program skeleton	236
	13.5.3 Handling of negation	237
	13.5.4 Opening the closed world	237
	13.5.5 Storage versus recomputation	238
13.6	EXPLAIN reference manual	239
10.0	13.6.1 Rules	239
	13.6.2 Base variables and equality	240
	13.6.3 Coupling to relational tables	241
13.7	EXPLAIN user guide	242
2017	13.7.1 Expert system components	242
	13.7.2 Calling EXPLAIN	243
	13.7.3 Summary of commands	243
	13.7.4 The EXPLAIN dialogue explained	244
13.8	Another example of EXPLAIN: pollution detection	245
	13.8.1 The pollution detection knowledge base	245
	13.8.2 Sample dialogue	248
	13.8.3 The structure of the pollution knowledge base	249
13.9	EXPLAIN expert system shell listing	250
13.10	Non-exact reasoning	254
	13.10.1 Multivalued logic	254
	13.10.2 Uncertain logic	255
	13.10.3 Uncertainties in EXPLAIN	257
Chapter 14	Knowledge Engineering	261
14.1	A knowledge engineering example	261
14.1	14.1.1 The problem	261
	14.1.2 The project	263
	14.1.2 The project 14.1.3 Knowledge acquisition	263 264
	14.1.4 Knowledge acquisition 14.1.4 Knowledge base statistics	264
	14.1.5 Performance	266
	14.1.6 User reactions	266
	14.1.7 Sample dialogue for a welding consultation	267
14.2	Comparison with traditional system development	268
12	companion with traditional system development	200
Appendix	Predicate Library	269
	Bibliography	274
	Index	279

Chapter 1 Introduction

Make rules, not programs!

1.1 Make specifications, not programs

The first computers were built to alleviate the burden of numerical computation. The computer could with little effort be programmed to perform any sequence of operations, and such was people's confidence in the machines that some engineers predicted that only a few computers would ever be needed for all the world's data processing.

Every day, computer applications once regarded as an unthinkable misuse of costly laboratory instruments are being developed. In fact, so many new applications are appearing that programming itself is losing its status as an art or handicraft, and is becoming more like traditional engineering. To cope with problems in large scale programming, the computer languages that control the computers have become more and more high level and standardized so that programmers can now exchange programs between systems.

To increase the level of programming, users have to specify what a program should do, but not how it should be done. However, as computing machinery became faster and cheaper, economics allowed more details to be automated by advanced computer software.

It is amusing today to read old arguments explaining why the computer should translate mathematical formulae, as in FORTRAN, instead of letting them be programmed more efficiently by skilled assembly language programmers (Backus, 1958). Few programmers today realize that FORTRAN formulae were once regarded as high-level mathematical specifications of a problem, rather than as programs themselves.

The tendency towards making specifications rather than programs is never ending. The ultimate goal would seem to be natural language specifications, therefore. However, natural language is imprecise, ambiguous and sensitive to context, and precision cannot be compromised. A formalism that approximates to precise natural language descriptions is logic. The ideal specification language of the future may well be some

kind of symbolic logic, leaving natural language for communication through the user interface.

Logic programming is the use of logic to define computer programs. The most notable programming language for logic programming is Prolog, an acronym for PROgramming in LOGic. What makes Prolog not just another programming language is its emphasis on the specifications of a problem. Rather than defining the algorithm for solving the problem, Prolog solves problems by systematically searching for a solution. Thus, Prolog is a significant step towards automatic programming.

Since the conception of Prolog in the early 1970s, the number of Prolog programmers has been doubling every year, which has created a market for fast and intelligent implementations of the language.

1.2 Symbolic processing language

Prolog is a language for processing symbolic information, that is, for processing general dynamic symbolic data structures and creating new structures during run-time. This is also a dominant feature of the programming language LISP (Winston and Horn, 1981). In symbolic processing languages, identifiers represent themselves rather than being names of storage locations as in languages such as Pascal. Prolog is also a declarative language. Programmers describe what they know in a precise manner, and leave the rest to the Prolog interpreter.

For example, a programmer wants to say that Harold is the father of Robin. In Pascal, he or she could imagine a linked list of name-records with pointers representing the father relations:

```
TYPE

person = RECORD

name : PACKED ARRAY [1..12] OF CHAR;
father : person;
next : person;
END

var a,b: person

a^.name := 'HAROLD';
b^.name := 'ROBIN';
b^.father := a;
.
```

In addition, separate programs must be written in Pascal before any information can be obtained.

In Prolog, this symbolic information can be expressed much more directly as

father(harold,robin),

which Prolog accepts as a true statement. If Prolog gets the query

?-father(X,robin),

it will respond with the answer

X=harold;

The query

?-father(harold, Y).

will give the answer

Y=robin;

while the query

?-father(X,Y).

will elicit the answer

X=harold

Y=robin;

1.3 Fifth generation computer systems

Prolog is now a generic name for various languages and dialects around a logic programming paradigm, and has gained great enthusiasm among an increasing number of programmers. This enthusiasm was boosted by the announcement of the Japanese Fifth Generation Computing Systems Programme (Moto-Oka, 1982), which adopted Prolog as the kernel language (Fifth Generation Kernel Language, FGKL) of the new type of computers called knowledge information processing systems (KIPS).

The fifth generation computer project aims to develop a completely new type of computer that can communicate in everyday language, reason intelligently and bring an enormous amount of stored knowledge to help solve any problem that the user may have. The machines will be cheap and reliable, and so widespread in offices, factories and homes that society itself may be changed by the huge fund of expertise made available.

The fifth generation computer will have three main components:

- an intelligent man machine interface (natural language, written or spoken, graphical);
- problem solving (with logic and statistical reasoning);
- a knowledge base facility able to store and retrieve vast amounts of data, and also give judgements and advice.

The programme's detailed plans promise to make incredibly fast Prolog machines for the 1990s; with a reasoning capacity of up to a billion logic inferences per second (one gigalips), and with associated memory storage of hundreds of gigabytes.

By 1986, the Japanese were already producing high performance Prolog workstations as a tool to accelerate their own development. Needless to say, recent achievements in the Japanese high-technology industries have conditioned us to expect a new success.

In light of this, logic programming, and Prolog in particular, is becoming very important.

1.4 History of logic programming

1.4.1 Aristotelian logic

Logic programming is based on logic, so the right place to start is with the origin of logic, which has its roots with the philosopher Aristotle (382–324 BC). He had an enormous influence on scientific thinking, but some of his ideas did not deserve much reverence. For example, Aristotle said that horses had more teeth than men, without counting them, and that the brain was an organ for cooling the blood, which is only true metaphorically. However, he will be remembered forever for his classic work *Organon*, where he summarized the laws of correct systematic thinking. According to Aristotle, correct reasoning proceeds by the application of strict rules of inference called **syllogisms**. A small example is:

Premise 1: All humans are mortal.

Premise 2: All Greeks are humans.

Premise 3: Socrates is Greek.

Conclusion: Socrates is mortal.