

Second Edition

Andries P. Engelbrecht

Computational Intelligence

An Introduction

Companion Website

 WILEY

TP18
E57
v.2

Computational Intelligence

An Introduction

Second Edition

Andries P. Engelbrecht
University of Pretoria
South Africa



E2008000045

John Wiley & Sons, Ltd

Copyright © 2007

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester,
West Sussex PO19 8SQ, England

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on www.wileyeurope.com or www.wiley.com

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, scanning or otherwise, except under the terms of the Copyright, Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1T 4LP, UK, without the permission in writing of the Publisher. Requests to the Publisher should be addressed to the Permissions Department, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, or emailed to permreq@wiley.co.uk, or faxed to (+44) 1243 770620.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The Publisher is not associated with any product or vendor mentioned in this book. All trademarks referred to in the text of this publication are the property of their respective owners.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the Publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 42 McDougall Street, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 6045 Freemont Blvd, Mississauga, Ontario, L5R 4J3, Canada

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Cataloging-in-Publication Data

Engelbrecht, Andries P.

Computational intelligence : an introduction / Andries P. Engelbrecht. – 2nd ed.
p. cm.

Includes bibliographical references.

ISBN 978-0-470-03561-0 (cloth)

1. Computational intelligence. 2. Neural networks (Computer science) 3. Evolutionary programming (Computer science) I. Title.

Q342.E54 2007

006.3-dc22

2007021101

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-470-03561-0 (HB)

Typeset by the author

Printed and bound in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire

This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

Computational Intelligence

*To my parents, Jan and Magriet Engelbrecht,
without whose loving support
this would not have happened.*

Preface to the Second Edition

Man has learned much from studies of natural systems, using what has been learned to develop new algorithmic models to solve complex problems. This book presents an introduction to some of these technological paradigms, under the umbrella of computational intelligence (CI). In this context, the book includes artificial neural networks, evolutionary computation, swarm intelligence, artificial immune systems, and fuzzy systems, which are respectively models of the following natural systems: biological neural networks, evolution, swarm behavior of social organisms, natural immune systems, and human thinking processes.

Why this book on computational intelligence? Need arose from a graduate course, where students did not have a deep background of artificial intelligence and mathematics. Therefore the introductory perspective is essential, both in terms of the CI paradigms and mathematical depth. While the material is introductory in nature, it does not shy away from details, and does present the mathematical foundations to the interested reader. The intention of the book is not to provide thorough attention to all computational intelligence paradigms and algorithms, but to give an overview of the most popular and frequently used models. For these models, detailed overviews of different implementations are given. As such, the book is appropriate for beginners in the CI field. The book is also applicable as prescribed material for a third year undergraduate course.

In addition to providing an overview of CI paradigms, the book provides insights into many new developments on the CI research front to tempt the interested reader. As such, the material is useful to graduate students and researchers who want a broader view of the different CI paradigms, also researchers from other fields who have no knowledge of the power of CI techniques, e.g. bioinformaticians, biochemists, mechanical and chemical engineers, economists, musicians and medical practitioners.

The book is organized in six parts. Part I provides a short introduction to the different CI paradigms and a historical overview. Parts II to VI cover the different paradigms, and can be reviewed in any order.

Part II deals with artificial neural networks (NN), including the following topics: Chapter 2 introduces the artificial neuron as the fundamental part of a neural network, including discussions on different activation functions, neuron geometry and learning rules. Chapter 3 covers supervised learning, with an introduction to different types of supervised networks. These include feedforward NNs, functional link NNs, product unit NNs, cascade NNs, and recurrent NNs. Different supervised learning algorithms are discussed, including gradient descent, conjugate gradient methods, LeapFrog and

Computational Intelligence: An Introduction, Second Edition A.P. Engelbrecht
©2007 John Wiley & Sons, Ltd

particle swarm optimization. Chapter 4 covers unsupervised learning. Different unsupervised NN models are discussed, including the learning vector quantizer and self-organizing feature maps. Chapter 5 discusses radial basis function NNs. Reinforcement learning is dealt with in Chapter 6. Much attention is given to performance issues of supervised networks in Chapter 7. The focus of the chapter is on accuracy measures and ways to improve performance.

Part III introduces several evolutionary computation models. Topics covered include: an overview of the computational evolution process and basic operators in Chapter 8. Chapter 9 covers genetic algorithms, Chapter 10 genetic programming, Chapter 11 evolutionary programming, Chapter 12 evolution strategies, Chapter 13 differential evolution, Chapter 14 cultural algorithms, and Chapter 15 covers coevolution, introducing both competitive and symbiotic coevolution.

Part IV presents an introduction to two types of swarm-based models: Chapter 16 discusses particle swarm optimization, while ant algorithms are discussed in Chapter 17.

Artificial immune systems are covered in Part V, with the natural immune system being discussed in Chapter 18 and a number of artificial immune models in Chapter 19.

Part VI deals with fuzzy systems. Chapter 20 presents an introduction to fuzzy logic with a discussion of membership functions. Fuzzy inferencing systems are explained in Chapter 21, while fuzzy controllers are discussed in Chapter 22. An overview of rough sets is given in Chapter 23.

Throughout the book, assignments are given to highlight certain aspects of the covered material and to stimulate thought. Some example applications are given where they seemed appropriate to better illustrate the theoretical concepts.

The accompanying website of this book, which can be located at <http://ci.cs.up.ac.za>, provides algorithms to implement many of the CI models discussed in this book. These algorithms are implemented in Java, and form part of an opensource library, Cilib, developed by the Computational Intelligence Research Group in the Department of Computer Science, University of Pretoria. Cilib (<http://cilib.sourceforge.net>) is a generic framework for easy implementation of new CI algorithms, and currently contains frameworks for particle swarm optimization, neural networks, and evolutionary computation. Lists with acronyms and symbols used in the book can also be downloaded from the book's website.

As a final remark, it is necessary to thank a number of people who have helped to produce this book. First of all, thanks to my mother, Magriet Engelbrecht, who has helped with typing and proofreading most of the text. Also, thanks to Anri Henning who spent a number of nights proofreading the material. The part on artificial immune systems was written by one of my PhD students, Attie Graaff. Without his help, this book would not have been so complete. Lastly, I thank all of my postgraduate students who have helped with the development of Cilib.

Pretoria, South Africa

Contents

	Page
Figures	xxi
Tables	xxiii
Algorithms	xxvii
Preface	xxix
Part I INTRODUCTION	1
1 Introduction to Computational Intelligence	3
1.1 Computational Intelligence Paradigms	4
1.1.1 Artificial Neural Networks	5
1.1.2 Evolutionary Computation	8
1.1.3 Swarm Intelligence	9
1.1.4 Artificial Immune Systems	9
1.1.5 Fuzzy Systems	10
1.2 Short History	11
1.3 Assignments	13
Part II ARTIFICIAL NEURAL NETWORKS	15
2 The Artificial Neuron	17
2.1 Calculating the Net Input Signal	17
2.2 Activation Functions	18
2.3 Artificial Neuron Geometry	20
2.4 Artificial Neuron Learning	21
2.4.1 Augmented Vectors	23
2.4.2 Gradient Descent Learning Rule	24
2.4.3 Widrow-Hoff Learning Rule	25
2.4.4 Generalized Delta Learning Rule	25
2.4.5 Error-Correction Learning Rule	25
2.5 Assignments	25
3 Supervised Learning Neural Networks	27
3.1 Neural Network Types	27
3.1.1 Feedforward Neural Networks	28
3.1.2 Functional Link Neural Networks	29
3.1.3 Product Unit Neural Networks	30

3.1.4	Simple Recurrent Neural Networks	32
3.1.5	Time-Delay Neural Networks	34
3.1.6	Cascade Networks	35
3.2	Supervised Learning Rules	36
3.2.1	The Supervised Learning Problem	36
3.2.2	Gradient Descent Optimization	38
3.2.3	Scaled Conjugate Gradient	45
3.2.4	LeapFrog Optimization	49
3.2.5	Particle Swarm Optimization	49
3.3	Functioning of Hidden Units	49
3.4	Ensemble Neural Networks	51
3.5	Assignments	54
4	Unsupervised Learning Neural Networks	55
4.1	Background	55
4.2	Hebbian Learning Rule	56
4.3	Principal Component Learning Rule	58
4.4	Learning Vector Quantizer-I	59
4.5	Self-Organizing Feature Maps	62
4.5.1	Stochastic Training Rule	62
4.5.2	Batch Map	65
4.5.3	Growing SOM	65
4.5.4	Improving Convergence Speed	67
4.5.5	Clustering and Visualization	69
4.5.6	Using SOM	71
4.6	Assignments	71
5	Radial Basis Function Networks	73
5.1	Learning Vector Quantizer-II	73
5.2	Radial Basis Function Neural Networks	73
5.2.1	Radial Basis Function Network Architecture	74
5.2.2	Radial Basis Functions	75
5.2.3	Training Algorithms	76
5.2.4	Radial Basis Function Network Variations	80
5.3	Assignments	81
6	Reinforcement Learning	83
6.1	Learning through Awards	83
6.2	Model-Free Reinforcement Learning Model	86
6.2.1	Temporal Difference Learning	86
6.2.2	Q-Learning	86

6.3	Neural Networks and Reinforcement Learning	87
6.3.1	RPROP	87
6.3.2	Gradient Descent Reinforcement Learning	88
6.3.3	Connectionist Q-Learning	89
6.4	Assignments	91
7	Performance Issues (Supervised Learning)	93
7.1	Performance Measures	93
7.1.1	Accuracy	93
7.1.2	Complexity	98
7.1.3	Convergence	98
7.2	Analysis of Performance	98
7.3	Performance Factors	99
7.3.1	Data Preparation	99
7.3.2	Weight Initialization	106
7.3.3	Learning Rate and Momentum	107
7.3.4	Optimization Method	109
7.3.5	Architecture Selection	109
7.3.6	Adaptive Activation Functions	115
7.3.7	Active Learning	116
7.4	Assignments	124
Part III	EVOLUTIONARY COMPUTATION	125
8	Introduction to Evolutionary Computation	127
8.1	Generic Evolutionary Algorithm	128
8.2	Representation – The Chromosome	129
8.3	Initial Population	132
8.4	Fitness Function	133
8.5	Selection	134
8.5.1	Selective Pressure	135
8.5.2	Random Selection	135
8.5.3	Proportional Selection	135
8.5.4	Tournament Selection	137
8.5.5	Rank-Based Selection	137
8.5.6	Boltzmann Selection	138
8.5.7	$(\mu \dagger \lambda)$ -Selection	139
8.5.8	Elitism	139
8.5.9	Hall of Fame	139
8.6	Reproduction Operators	139

8.7	Stopping Conditions	140
8.8	Evolutionary Computation versus Classical Optimization	141
8.9	Assignments	141
9	Genetic Algorithms	143
9.1	Canonical Genetic Algorithm	143
9.2	Crossover	144
9.2.1	Binary Representations	145
9.2.2	Floating-Point Representation	146
9.3	Mutation	153
9.3.1	Binary Representations	154
9.3.2	Floating-Point Representations	155
9.3.3	Macromutation Operator – Headless Chicken	156
9.4	Control Parameters	156
9.5	Genetic Algorithm Variants	157
9.5.1	Generation Gap Methods	158
9.5.2	Messy Genetic Algorithms	159
9.5.3	Interactive Evolution	161
9.5.4	Island Genetic Algorithms	162
9.6	Advanced Topics	164
9.6.1	Niching Genetic Algorithms	165
9.6.2	Constraint Handling	169
9.6.3	Multi-Objective Optimization	170
9.6.4	Dynamic Environments	173
9.7	Applications	174
9.8	Assignments	175
10	Genetic Programming	177
10.1	Tree-Based Representation	177
10.2	Initial Population	179
10.3	Fitness Function	180
10.4	Crossover Operators	180
10.5	Mutation Operators	182
10.6	Building Block Genetic Programming	184
10.7	Applications	184
10.8	Assignments	185
11	Evolutionary Programming	187
11.1	Basic Evolutionary Programming	187
11.2	Evolutionary Programming Operators	189
11.2.1	Mutation Operators	189

11.2.2	Selection Operators	193
11.3	Strategy Parameters	195
11.3.1	Static Strategy Parameters	195
11.3.2	Dynamic Strategies	195
11.3.3	Self-Adaptation	198
11.4	Evolutionary Programming Implementations	200
11.4.1	Classical Evolutionary Programming	200
11.4.2	Fast Evolutionary Programming	201
11.4.3	Exponential Evolutionary Programming	201
11.4.4	Accelerated Evolutionary Programming	201
11.4.5	Momentum Evolutionary Programming	202
11.4.6	Evolutionary Programming with Local Search	203
11.4.7	Evolutionary Programming with Extinction	203
11.4.8	Hybrid with Particle Swarm Optimization	204
11.5	Advanced Topics	206
11.5.1	Constraint Handling Approaches	206
11.5.2	Multi-Objective Optimization and Niching	206
11.5.3	Dynamic Environments	206
11.6	Applications	207
11.6.1	Finite-State Machines	207
11.6.2	Function Optimization	208
11.6.3	Training Neural Networks	209
11.6.4	Real-World Applications	210
11.7	Assignments	210
12	Evolution Strategies	213
12.1	(1 + 1)-ES	213
12.2	Generic Evolution Strategy Algorithm	215
12.3	Strategy Parameters and Self-Adaptation	216
12.3.1	Strategy Parameter Types	216
12.3.2	Strategy Parameter Variants	218
12.3.3	Self-Adaptation Strategies	219
12.4	Evolution Strategy Operators	221
12.4.1	Selection Operators	221
12.4.2	Crossover Operators	222
12.4.3	Mutation Operators	224
12.5	Evolution Strategy Variants	226
12.5.1	Polar Evolution Strategies	226
12.5.2	Evolution Strategies with Directed Variation	227

12.5.3	Incremental Evolution Strategies	228
12.5.4	Surrogate Evolution Strategy	229
12.6	Advanced Topics	229
12.6.1	Constraint Handling Approaches	229
12.6.2	Multi-Objective Optimization	230
12.6.3	Dynamic and Noisy Environments	233
12.6.4	Niching	233
12.7	Applications of Evolution Strategies	235
12.8	Assignments	235
13	Differential Evolution	237
13.1	Basic Differential Evolution	237
13.1.1	Difference Vectors	238
13.1.2	Mutation	239
13.1.3	Crossover	239
13.1.4	Selection	240
13.1.5	General Differential Evolution Algorithm	241
13.1.6	Control Parameters	241
13.1.7	Geometrical Illustration	242
13.2	DE/ $x/y/z$	242
13.3	Variations to Basic Differential Evolution	245
13.3.1	Hybrid Differential Evolution Strategies	245
13.3.2	Population-Based Differential Evolution	250
13.3.3	Self-Adaptive Differential Evolution	250
13.4	Differential Evolution for Discrete-Valued Problems	252
13.4.1	Angle Modulated Differential Evolution	253
13.4.2	Binary Differential Evolution	254
13.5	Advanced Topics	255
13.5.1	Constraint Handling Approaches	256
13.5.2	Multi-Objective Optimization	256
13.5.3	Dynamic Environments	257
13.6	Applications	259
13.7	Assignments	259
14	Cultural Algorithms	261
14.1	Culture and Artificial Culture	261
14.2	Basic Cultural Algorithm	262
14.3	Belief Space	263
14.3.1	Knowledge Components	264
14.3.2	Acceptance Functions	265

14.3.3	Adjusting the Belief Space	266
14.3.4	Influence Functions	267
14.4	Fuzzy Cultural Algorithm	268
14.4.1	Fuzzy Acceptance Function	269
14.4.2	Fuzzified Belief Space	269
14.4.3	Fuzzy Influence Function	270
14.5	Advanced Topics	271
14.5.1	Constraint Handling	271
14.5.2	Multi-Objective Optimization	272
14.5.3	Dynamic Environments	273
14.6	Applications	274
14.7	Assignments	274
15	Coevolution	275
15.1	Coevolution Types	276
15.2	Competitive Coevolution	276
15.2.1	Competitive Fitness	277
15.2.2	Generic Competitive Coevolutionary Algorithm	279
15.2.3	Applications of Competitive Coevolution	280
15.3	Cooperative Coevolution	281
15.4	Assignments	283
Part IV	COMPUTATIONAL SWARM INTELLIGENCE	285
16	Particle Swarm Optimization	289
16.1	Basic Particle Swarm Optimization	289
16.1.1	Global Best PSO	290
16.1.2	Local Best PSO	291
16.1.3	<i>gbest</i> versus <i>lbest</i> PSO	292
16.1.4	Velocity Components	293
16.1.5	Geometric Illustration	294
16.1.6	Algorithm Aspects	296
16.2	Social Network Structures	300
16.3	Basic Variations	303
16.3.1	Velocity Clamping	303
16.3.2	Inertia Weight	306
16.3.3	Constriction Coefficient	309
16.3.4	Synchronous versus Asynchronous Updates	310
16.3.5	Velocity Models	310
16.4	Basic PSO Parameters	312

16.5	Single-Solution Particle Swarm Optimization	314
16.5.1	Guaranteed Convergence PSO	316
16.5.2	Social-Based Particle Swarm Optimization	317
16.5.3	Hybrid Algorithms	321
16.5.4	Sub-Swarm Based PSO	326
16.5.5	Multi-Start PSO Algorithms	333
16.5.6	Repelling Methods	337
16.5.7	Binary PSO	340
16.6	Advanced Topics	342
16.6.1	Constraint Handling Approaches	342
16.6.2	Multi-Objective Optimization	343
16.6.3	Dynamic Environments	346
16.6.4	Niching PSO	350
16.7	Applications	354
16.7.1	Neural Networks	354
16.7.2	Architecture Selection	356
16.7.3	Game Learning	356
16.8	Assignments	357
17	Ant Algorithms	359
17.1	Ant Colony Optimization Meta-Heuristic	360
17.1.1	Foraging Behavior of Ants	360
17.1.2	Stigmergy and Artificial Pheromone	363
17.1.3	Simple Ant Colony Optimization	364
17.1.4	Ant System	368
17.1.5	Ant Colony System	372
17.1.6	Max-Min Ant System	375
17.1.7	Ant-Q	378
17.1.8	Fast Ant System	379
17.1.9	Antabu	380
17.1.10	AS-rank	380
17.1.11	ANTS	381
17.1.12	Parameter Settings	383
17.2	Cemetery Organization and Brood Care	384
17.2.1	Basic Ant Colony Clustering Model	385
17.2.2	Generalized Ant Colony Clustering Model	386
17.2.3	Minimal Model for Ant Clustering	391
17.3	Division of Labor	391
17.3.1	Division of Labor in Insect Colonies	392

17.3.2	Task Allocation Based on Response Thresholds	393
17.3.3	Adaptive Task Allocation and Specialization	395
17.4	Advanced Topics	396
17.4.1	Continuous Ant Colony Optimization	396
17.4.2	Multi-Objective Optimization	398
17.4.3	Dynamic Environments	402
17.5	Applications	405
17.5.1	Traveling Salesman Problem	406
17.5.2	Quadratic Assignment Problem	407
17.5.3	Other Applications	411
17.6	Assignments	411
Part V ARTIFICIAL IMMUNE SYSTEMS		413
18	Natural Immune System	415
18.1	Classical View	415
18.2	Antibodies and Antigens	416
18.3	The White Cells	416
18.3.1	The Lymphocytes	417
18.4	Immunity Types	421
18.5	Learning the Antigen Structure	421
18.6	The Network Theory	422
18.7	The Danger Theory	422
18.8	Assignments	424
19	Artificial Immune Models	425
19.1	Artificial Immune System Algorithm	426
19.2	Classical View Models	428
19.2.1	Negative Selection	428
19.2.2	Evolutionary Approaches	429
19.3	Clonal Selection Theory Models	431
19.3.1	CLONALG	431
19.3.2	Dynamic Clonal Selection	433
19.3.3	Multi-Layered AIS	433
19.4	Network Theory Models	436
19.4.1	Artificial Immune Network	436
19.4.2	Self Stabilizing AIS	438
19.4.3	Enhanced Artificial Immune Network	440
19.4.4	Dynamic Weighted B-Cell AIS	441
19.4.5	Adapted Artificial Immune Network	442

19.4.6 aiNet	442
19.5 Danger Theory Models	445
19.5.1 Mobile Ad-Hoc Networks	445
19.5.2 An Adaptive Mailbox	446
19.5.3 Intrusion Detection	448
19.6 Applications and Other AIS models	448
19.7 Assignments	448
Part VI FUZZY SYSTEMS	451
20 Fuzzy Sets	453
20.1 Formal Definitions	454
20.2 Membership Functions	454
20.3 Fuzzy Operators	457
20.4 Fuzzy Set Characteristics	459
20.5 Fuzziness and Probability	462
20.6 Assignments	463
21 Fuzzy Logic and Reasoning	465
21.1 Fuzzy Logic	465
21.1.1 Linguistics Variables and Hedges	466
21.1.2 Fuzzy Rules	467
21.2 Fuzzy Inferencing	468
21.2.1 Fuzzification	469
21.2.2 Inferencing	470
21.2.3 Defuzzification	471
21.3 Assignments	472
22 Fuzzy Controllers	475
22.1 Components of Fuzzy Controllers	475
22.2 Fuzzy Controller Types	477
22.2.1 Table-Based Controller	477
22.2.2 Mamdani Fuzzy Controller	477
22.2.3 Takagi-Sugeno Controller	478
22.3 Assignments	478
23 Rough Sets	481
23.1 Concept of Discernibility	482
23.2 Vagueness in Rough Sets	483
23.3 Uncertainty in Rough Sets	484
23.4 Assignments	485