

First BMI Press hardcover edition, 2012, the 2nd print
Copyright ©2012 by Brain-Mind Institute

All rights reserved. No part of this publication may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

Printed and bound in the United States of America.

Library of Congress Control Number: 2012952697

ISBN 978-0-9858757-2-5

Natural and Artificial Intelligence

Introduction to Computational Brain-Mind

Juyang Weng

BMI Press
Okemos, Michigan

Contents

Acknowledgments	xv
Preface	xix
1 Agents and Tasks	1
1.1 Agents	1
1.1.1 Sensors and effectors	2
1.1.2 Task environment and beyond	2
1.1.3 Humans as agent constructors	2
1.1.4 Constructors of a human agent	3
1.2 A Paradox	4
1.3 Principles of Task Muddiness	6
1.4 Muddiness Frame Examples	7
1.4.1 Rawness of input	7
1.4.2 Size of input space	8
1.4.3 A two-dimensional muddiness frame	9
1.4.4 Richness of goal	10
1.4.5 A three-dimensional muddiness frame	11
1.5 Muddiness Frames	12
1.5.1 External environment	12
1.5.2 Input	14
1.5.3 Internal environment	15
1.5.4 Output	17
1.5.5 Goal	18
1.5.6 Composite muddiness	18
1.5.7 Questions about the muddiness	20
1.6 Three Task Categories	21
1.6.1 Category 1: Clean tasks	21
1.6.2 Category 2: Muddy tasks	22
1.6.3 Category 3: Very muddy tasks	22
1.7 Intelligence Metrics	23
1.8 Conclusions	25

Bibliographic notes	25
Problems	26
2 Representation and Search	29
2.1 Brain-Mind and AI	29
2.1.1 Basic concepts	29
2.1.2 AMD: Overcome Disciplinary Limitations	32
2.2 Traditional Approaches	34
2.2.1 Knowledge-based	34
2.2.2 Behavior-based	34
2.2.3 Learning-based	35
2.2.4 Evolutional	36
2.3 Developmental Approach: Task Nonspecificity	37
2.3.1 Autonomous Mental Development	37
2.3.2 Task Nonspecificity	37
2.3.3 Conceptual Comparison of Approaches	38
2.3.4 Necessity of Autonomous Development	39
2.4 Symbolic Representations	40
2.4.1 Symbolic Representation	40
2.4.2 Symbolic Monolithic	41
2.4.3 Symbolic Contextual	47
2.5 Emergent Representations	59
2.5.1 Computer Symbols	59
2.5.2 Symbols in Human Communications	59
2.5.3 Emergent Representation	60
2.5.4 Networks that Use Emergent Representations	61
2.5.5 Comments	61
2.6 Conclusions	62
Bibliographic notes	63
Problems	65
3 Autonomous Development	69
3.1 Human Physical Development	69
3.1.1 Biological physical development	69
3.1.2 The brain development	73
3.2 Human Mental Development	74
3.2.1 Biological-maturation perspective	75
3.2.2 Environmental-learning perspective	76
3.2.3 Constructivist perspective	77
3.2.4 Cultural-context perspective	79
3.2.5 Computational perspective	80
3.3 Brain-Mind Research Needs Six Disciplines	81
3.4 Mental Development	81

3.4.1	Natural vs. Artificial Agents	82
3.4.2	Body Development vs. Mental Development	82
3.4.3	Machine Learning vs. Autonomous Development	82
3.5	A New Engineering Paradigm	84
3.5.1	AMD paradigm	84
3.5.2	Comparison of two paradigms	85
3.5.3	The five practical factors	86
3.5.4	Eight Computational Requirements	91
3.5.5	Developmental programs	92
3.6	Interfaces, Interactions and Experiences	93
3.6.1	Biased and unbiased sensors	94
3.6.2	Learning types	95
3.6.3	Autonomous Animal-like Learning	97
3.7	Conclusions	98
	Bibliographic notes	99
	Problems	100
4	Neurons and Features	103
4.1	Neurons	103
4.1.1	Brain and neurons	103
4.1.2	Firing rate model	105
4.1.3	Spiking neuron models	107
4.1.4	Feature detectors	108
4.2	Shift Invariant Features	109
4.2.1	Sampling and discrete images	112
4.2.2	Perfect invariance in the brain?	113
4.3	Types of neuronal learning algorithms	114
4.4	Mean	115
4.4.1	Batch mean	115
4.4.2	Incremental mean	116
4.4.3	Amnesic mean	116
4.4.4	Linewise normalization	117
4.5	Principal Component Analysis	118
4.5.1	Competition among neurons	119
4.5.2	PCA: Most expressive but dense representation	120
4.5.3	Principal components as eigenvectors	122
4.5.4	Making parameters directly observable	124
4.5.5	Mean has the least deviation	124
4.5.6	Incremental PCA	125
4.5.7	Amnesic PCA	126
4.5.8	Higher-order eigenvectors	127
4.5.9	Candid Covariance-free Incremental PCA	127

4.5.10	Computational complexity	129
4.5.11	PCA net	129
4.6	Lobe Component Analysis	130
4.6.1	Concepts	132
4.6.2	Local approximation of high-dimensional density	132
4.6.3	Lobe components	133
4.6.4	Optimal spatial representation: Lobe components	135
4.6.5	Temporal optimality: Automatic step sizes	136
4.6.6	Statistical efficiency	139
4.6.7	Automatic scheduling of optimal step sizes	139
4.6.8	Time-variant distribution	140
4.7	Lobe Component Analysis algorithm	140
4.7.1	CCI LCA Algorithm	140
4.7.2	Time and space complexities	141
4.7.3	LCA examples	142
4.7.4	LCA net	142
4.8	Conclusions	143
	Bibliographic notes	143
	Problems	145
5	Properties of Representation	149
5.1	Feature Extraction	149
5.1.1	Properties of feature extraction	150
5.1.2	Fourier Transforms	152
5.1.3	Gabor Filters	153
5.1.4	Hopfield nets and Boltzmann machines	154
5.1.5	ART	155
5.1.6	Support vector machines	156
5.1.7	Bayesian classifier	157
5.1.8	Linear discriminant analysis	158
5.1.9	Independent Component Analysis	161
5.2	Efficient Estimator	164
5.2.1	Cramér-Rao lower bound	165
5.2.2	Cramér-Rao lower bound: Vector Form	166
5.2.3	Unbiasedness of amnesic mean	166
5.2.4	Convergence speed of CCI PCA	167
5.2.5	Convergence speed of amnesic mean	168
5.2.6	Responses after PCA are uncorrelated	170
5.2.7	Three stringent operative conditions	170
5.2.8	Correct convergence and speed	172
5.3	LCA Properties	174
5.3.1	Entropy and mutual information	174

5.3.2	Dual optimality of LCA	175
5.3.3	Developmental neuron	177
5.3.4	LCA complexity and in-place development	177
5.3.5	LCA is ICA for super-Gaussians	178
5.3.6	Conceptual comparison for LCA and ICA	178
5.3.7	Convergence speed*	179
5.3.8	Top-k winner	179
5.4	Summary	180
	Bibliographic notes	180
	Problems	181
6	Brain-Mind Architecture	185
6.1	Introduction	185
6.2	The Central Nervous System	187
6.2.1	Spinal cord	188
6.2.2	Brainstem	189
6.2.3	Forebrain	189
6.3	5+1 chunks of the brain-mind puzzle	190
6.4	A Basic Model: DN	191
6.4.1	Developmental Network (DN)	191
6.4.2	Conceptual comparison between SN and DN	192
6.4.3	Algorithm of DN	193
6.4.4	Area function of DN	194
6.4.5	FA as a special case of a DN	195
6.4.6	Discussion of Theorem 6.1	201
6.4.7	GDN optimally performs while frozen	202
6.4.8	GDN optimally performs while learning	202
6.4.9	Two AFA-GDN examples	203
6.5	Brain Architectures	204
6.5.1	Architecture considerations	204
6.5.2	Symbols in the brain?	205
6.5.3	Formulations of the brain	208
6.5.4	Brain connections	209
6.5.5	Why is the brain not a cascade of areas?	210
6.5.6	Multimodal integration	213
6.6	Cortex Scale: Prescreening	216
6.6.1	Laminar cortex	216
6.6.2	Prescreening	216
6.7	Layer Scale: the Dually Optimal LCA	218
6.8	Spatiotemporal Cognition and Behaviors	220
6.8.1	Spatial attention	221
6.8.2	Temporal attention	221

6.8.3	Spatiotemporal attention, cognition and behaviors	222
6.9	Experiments	222
6.10	Conclusions and Discussion	223
	Bibliographic notes	224
	Problems	225
7	Spatial Processing	229
7.1	Introduction	229
7.2	Intent network	231
7.3	Where-What Network	232
7.3.1	Area function	232
7.3.2	Hextuple fields of each neuron	234
7.3.3	Area computation	235
7.3.4	Receptive fields are selective and dynamic	237
7.3.5	Concepts	237
7.3.6	Learning	238
7.3.7	WWN algorithm	240
7.3.8	Concept purity	241
7.3.9	WWN as an Intent Network	242
7.3.10	Goal directed perception	246
7.3.11	Top-down representational effect	248
7.3.12	Y sub-areas	250
7.4	Experimental Examples	250
7.4.1	Multiple objects in complex background	251
7.4.2	25 3-D objects with 360° of viewing variation	253
7.4.3	Outdoor vehicles and non-vehicles	253
7.4.4	Stereo perception	253
7.4.5	Natural languages with each word as an image	254
7.5	Conclusions and Discussion	254
	Bibliographic notes	255
	Problems	257
8	Temporal Processing	261
8.1	Introduction	261
8.2	TCM Algorithm	263
8.3	Temporal Representations	264
8.3.1	Handcrafted representations	264
8.3.2	Emergent representations	265
8.3.3	Emergent representation with states	265
8.4	Temporal Mechanisms	266
8.4.1	FA: Symbolic	266
8.4.2	Temporal window: symbolic in time	266
8.4.3	Local feedbacks in emergent representation	267

8.4.4	Motor feedbacks in emergent representation	267
8.5	Architecture of TCM	269
8.5.1	Sensory modality	269
8.5.2	Intractable sensory space	269
8.5.3	Recursive abstraction for actions	270
8.5.4	Recursion on state: open temporal length	271
8.5.5	Using internal representations	272
8.5.6	Temporal attention	273
8.5.7	Observations	273
8.6	Bridge Representation	274
8.6.1	Internal neurons as soft AND of X and Z	274
8.6.2	X and Z neurons as soft OR of Y cases	274
8.6.3	Prediction for image and motor	274
8.6.4	Almost no local minima	275
8.6.5	Internal sensing and actions	275
8.7	Properties	275
8.7.1	For any FA there is a TCM	275
8.7.2	Context dependent attention	276
8.7.3	Active time warping	276
8.7.4	Temporal attention	277
8.7.5	Time duration	278
8.7.6	Generalization for new sentences	279
8.7.7	Complexity	280
8.8	Experimental Examples	281
8.8.1	Video processing as temporal processing	282
8.8.2	Text processing as temporal processing	284
8.9	Conclusions and Discussions	290
	Bibliographic notes	291
	Problems	293
9	Modulation	295
9.1	Modulatory System	295
9.1.1	Why motivation?	296
9.1.2	Types of motivational systems	297
9.1.3	Psychological and neuroscience studies	297
9.1.4	Four types of neural transmitters	298
9.1.5	Ions, neurotransmitters and synapses	300
9.1.6	Punishment, reward, uncertainty, and novelty	301
9.2	Models of serotonin and dopamine systems	302
9.2.1	Architecture	304
9.2.2	Operation and Learning	305
9.2.3	Glutamate, serotonin and dopamine integration	306

9.2.4	How the glutamate subsystem works	307
9.2.5	How serotonin and dopamine subsystems work	308
9.2.6	A degenerate case	310
9.2.7	Delayed rewards	310
9.2.8	Teaching schedule	310
9.2.9	A hypothetical task setting	312
9.2.10	Experimental examples	313
9.2.11	Effects of serotonin and dopamine on Y	315
9.3	Models of Acetylcholine and Norepinephrine Systems	316
9.3.1	Acetylcholine system	316
9.3.2	Norepinephrine system	317
9.3.3	Synapse trimming	318
9.3.4	Change of the maintenance is bi-directional	319
9.3.5	Cross-section synapse maintenance	319
9.3.6	Experimental examples	320
9.4	Conclusions	322
	Bibliographic notes	323
	Problems	325
10	Generalization	327
10.1	General Task Learning and Execution	327
10.1.1	Task conditions	328
10.1.2	Teaching schedule	329
10.1.3	Task success	329
10.2	Attention	330
10.2.1	Spatial attention	331
10.2.2	Multiple attended concepts	333
10.2.3	Multiple image patches	333
10.2.4	Temporal attention	334
10.2.5	Spatiotemporal events and behaviors	335
10.2.6	Environmental calibration	336
10.2.7	Generalization using distributed (vector) representation	336
10.3	Transfer	337
10.3.1	Specificity	337
10.3.2	Transfer	338
10.3.3	State-based transfer in FA	339
10.3.4	Aspect 1: Transfer due to state abstraction	340
10.3.5	Aspect 2: Transfer due to sensory attention	340
10.3.6	Aspect 3: Transfer due to motor attention	341
10.3.7	Aspect 4: Transfer due to off-task processes	343
10.4	Sample Tasks	346
10.4.1	Classical conditioning	346

10.4.2	Instrumental conditioning	349
10.4.3	Language acquisition	350
10.4.4	Autonomous navigation	352
10.5	Beyond Sample Tasks	353
10.5.1	Non-deterministic FA	354
10.5.2	Symbols inside the brain skull?	355
10.5.3	Emergence of brain areas	357
10.5.4	Completeness of AFAs	358
10.6	Experimental Results Using DN	361
10.7	Autonomous Thinking	362
10.7.1	Cognition observed as behaviors	363
10.7.2	Internal and external contexts	364
10.7.3	Autonomous reasoning	365
10.7.4	Autonomous planning	367
10.8	Conclusions	368
	Bibliographic notes	368
	Problems	369
11	Group Intelligence	373
11.1	Beyond traditional views	373
11.2	Groups	374
11.2.1	Autonomous development	374
11.2.2	Five factors of development	375
11.2.3	Group developmental program	376
11.3	The World	377
11.4	Conflicts	380
11.4.1	Biological basis of conflicts	380
11.4.2	Groupish minds	381
11.4.3	US international policies	384
11.4.4	US domestic policies	385
11.4.5	Other types of group	386
11.5	Conclusions	386
	Bibliographic notes	387
	Problems	388
A	Mathematical Background	421
A.1	Vector	421
A.2	Matrix	421
A.3	Random Variables	422
A.4	Random Vectors and Matrices	423
A.5	Eigenvectors and Eigenvalues	423

Acknowledgments

Our work for autonomous intelligence development has gone through several systems: Cresceptron (1991 - 1995), SHSOLIF (1993 - 2000), SAIL (1996 - 2010), Dav (1999 - present), MILN (2004-2008), WWN (WWN-1 through WWN-7, 2007-present). Many colleagues and friends have provided insights, encouragements and suggestions during this period. I would like to thank my colleague Prof. Anil K. Jain in the Department of Computer Science and Engineering at Michigan State University (MSU) who encouraged and supported my early-stage ideas. I would like to thank my colleague Prof. George Stockman in the same department whose critical comments are so useful for a less-guarded mind. My early understanding about animal learning benefited from Prof. Mark E. Rilling in the Department of Psychology at MSU who kindly accepted me as one of his students in his animal learning course in the fall of 1997. My knowledge about neuroscience was greatly enriched by attending MIT 9.011 Principles of Neuroscience, a course in the Department of Brain and Cognitive Sciences, in the fall of 1999, while I was on sabbatical at MIT, hosted by Prof. Sandy Pentland. This six-credit-hour per week intensive course for graduate students was co-taught by Profs. Ann Graybiel, Guosong Liu, Richard Wurtman, Peter Chiller, Chris Brown, Emilio Bizzi, Mriganka Sur, Earl Miller, and Suzanne Corkin, from whom I have learned a lot and with whom I had a great time discussing issues, inside and outside the classroom. I had many extensive discussions with Ida Stockman, a well known, extremely nice developmental psychologist and expert in special education, who taught me many of her perspectives about development. I especially like to mention late Esther Thelen, a well know developmental psychologist, who suggested the term “autonomous” before “mental development” which now has become a well known acronym as AMD. During Fall 2006 and Spring 2007, Prof. Mriganka Sur at MIT hosted me as a visiting professor at the Department of Brain and Cognitive Sciences. This was an invaluable period where I learned about many latest results in neuroscience from faculty, post-doctoral associates and graduate students and got some rich experience. During this full academic year, I was also fortunate to have the opportunity to learn from the courses MIT 7.012 Introductory Biology taught by Eric Lander and Robert Weinberg, MIT 9.036 the Visual System taught by Peter H. Schiller, and MIT 9.285J Neural Coding and Perception of Sound taught by Bertrand A.R. Delgutte. During the summer of 2012, I took and passed BMI 811 Biology for Brain-Mind Research, taught by Dr. Jon Stoltzfus, and BMI 821 Neuroscience for Brain-Mind Research, taught by Prof. Lynwood G. Clemens. The Brain-Mind Institute gave me an opportunity to formally

take these two courses, as the oldest student in each course.

I would also like to extend thanks to my colleagues in related fields at several institutions, for collaborations and idea exchanges. They are: Prof. Narendra Ahuja with UIUC, Prof. Ronald Arkin with Georgia Tech., Prof. Dean Aslam with MSU, Prof. Dana Ballard with Rochester U., Prof. Neil E. Berthier with U. Mass, Prof. Rodney A. Brooks with MIT, Prof. Mingmin Chi with Fudan U., Prof. Chris Brown with Rochester U., Prof. Susan Carey with New York U. and later Harvard U., Prof. Rama Chellappa with U. Maryland, Prof. Rachel Clifton with U. Mass, Prof. Gedeon Deak with UCSD, Prof. Gerald M. Edelman with The Scripps Research Institute, Prof. John R. Deller with MSU, Prof. Jeff Elman with UCSD, Prof. Jianfeng Feng with U. of Warwick, Prof. Stan Franklin with U. of Memphis, Prof. Stephen Grossberg with Boston U., Prof. John Henderson with MSU, Prof. Patrick Dickson with MSU, Prof. Michael Gabriel with UIUC, Prof. Tony Jebara with Columbia U., Prof. Susan A. Gelman with U. Michigan, Prof. Roderic Grupen with U. Mass, Prof. Thomas S. Huang with UIUC, Prof. Xuanjing Huang with Fudan U., Prof. Takeo Kanade with CMU, Prof. Stephen Levinson with UIUC, Prof. Taosheng Liu with MSU, Prof. Hong Lu with Fudan U., Prof. Sridhar Mahadevan with U. Mass, Prof. Maja Mataric with USC, Prof. James McClelland with Stanford U., Dr. Jun Miao with the Institute of Computing Technology of the Chinese Academy of Sciences, Dr. Javier R. Movellan with UCSD, Prof. Ranjan Mukherjee with MSU, Prof. Michael M. Merzenich with UCSF, Prof. Alex Pentland with MIT, Prof. Jianguo Qi with MSU, Prof. Brian Scassellati with Yale U., Prof. Tomaso Poggio with MIT, Prof. Fathi Salam with MSU, Prof. Klaus J. Schulten with UIUC, Prof. Terrence J. Sejnowski with Salk Institute, Prof. Tiande Shou with Fudan U., Prof. Olaf Sporns with Indiana U., Prof. David A. Stewart with MSU, Prof. Ida Stockman with MSU, Prof. Mriganka Sur with MIT, Prof. Lal Tammala with MSU, Prof. Esther Thelen with Indiana U., Prof. Eric Torng with MSU, Prof. Xiaofeng Wu with Fudan U., Prof. Ning Xi with MSU, Prof. Xiangyang Xue with Fudan U., Prof. Hongbo Yu with Fudan University, Prof. Jun Zhang with U. Michigan, Prof. Liming Zhang with Fudan U., Prof. Yang Wang with MSU, Prof. Wenqiang Zhang with Fudan U., Dr. Paul J. Werbos with NSF, Drs. Cha Zhang and Dr. Zhengyou Zhang with Microsoft Research, and many others.

The related projects would not be possible without contributions from my research advisees who worked on various components. They are: Gil Abramovich, Hamid Alavi, Leif E. Alton, Micky L. Badgero, Laura A. Blackwood, Greg J. Bloy, David A. Bordoley, Kevin A. Brown, Yi Chen, David M. Cherba, Paul Cornwell, Matthew P. Ebrom, Shaoyun Chen, Yuntao Cui, Colin H. Evans, Raja Ganjikutta, Laura Grabowski, Kamen Guentchev, Xiao Huang, Jianda Han, Mao Hu, Sally Howden, Wey-Shiuan Hwang, Zhengping Ji, Ameet V. Joshi, Jeffrey A. Kohler, Gongjun Li, Yong-Beom Lee, Quo L. Liu, Matthew D. Luciw, Tianyu Luwang, Jason L. Massey, Kenneth Marable, Kajal Miyan, Rossen Nedeltchev, Lisa Ossian, Stephen Paslaski, Rebecca Smith, Mojtaba Solgi, Xiaoying Song, Jason Sperber, Benjamin Syzek, Daniel L. Swets, Keng Yew Tham, Wei Tong, Nikita N. Wagle, Jingliang Wang, Yuekai Wang, Jamal A. Wills, Changjiang Yang, Shuqing Zeng, Nan Zhang, Yilu Zhang, Zejia Zheng, and Yang Zhou. They were with MSU or with Fudan U. where I served

as a visiting professor.

I would also like to extend my special thanks to the students in the MSU CSE902 Special Topics course in Mental Development and Online Learning, MSU CSE941 Special Topics in Autonomous Mental Development, MSU CSE841 Artificial Intelligence, and MIT 9.915/6.887 Computational Cognitive and Neural Development that I was fortunate to teach using earlier versions of the volume as the textbook. They all provided valuable written and verbal comments and suggestions, which significantly helped me in improving the writing and reducing typos.

I would also like to acknowledge NSF and DARPA for their support of my new line of work. I would especially like to thank Drs. Mike Lesk, Howard Moraff, and Larry Reeker at NSF and Drs. Mark Swinson and Doug Gage at DARPA for their encouragement and support. I would also like to thank Jing Xiao at NSF and Mark Swinson at DARPA for their support of the “Workshop on Development and Learning” (WDL), held at MSU, April 5-7, 2000, which has grown into a regularly scheduled conference series called the International Conference on Development and Learning (ICDL) which later led to the *AMD Newsletter* and the *IEEE Transactions on Autonomous Mental Development* with much support from the interested IEEE societies and the growing research community. The NSF/DARPA support also enabled us to create the web based *Mental Development Repository* at www.mentaldev.org. I am very grateful to the financial supports from Microsoft Research, General Motors, and Toyota Research Institute North America. Such supports were indeed visionary and I wish that scientific rules could be put in place for our peer reviews so that our peer reviewers can be visionary too.

The great participation and support of the research community and professional societies to the *Brain-Mind Institute* and the *Brain-Mind Magazine* are also greatly appreciated, as we hope that such educational and research activities can bridge the wide communication gaps between disciplines and between the research community and general public.

I extend thanks to Shelly Villarreal, Kimberly Thompson and Stephanie Beal who read the manuscript for typos and grammar. However, all of the remaining typos are mine and not theirs. Please report errors to me at weng@cse.msu.edu.

Last, but not the least, I’m grateful to my wife, Min Guo, for her supports, encouragements, and assuming many of my chores to give me time for this endeavor, as well as to my two lovely boys, Colin and Rodney, some of whose early learning behaviors were invaluable to me while I was fascinated by development.

Juyang Weng

October, 2012
East Lansing, USA

Preface

The mind is what the brain does. This volume tries to map a mind model to the corresponding brain so as to not only deepen our understanding of both the brain and the mind, but also unveil computational underpinnings. That is why the words “Brain-Mind” are hyphenated in the title.

This volume strives to unify natural intelligence with artificial intelligence. It approaches intelligence through not only what intelligence *is* but also how intelligence *arises*.

In terms of natural intelligence, instead of modeling what a brain *does*, this volume explains how the brain *develops* — how the brain circuits and mind functions emerge from interactions between the DN and the external environments.

In terms of artificial intelligence, instead of modeling an intelligent agent as a probabilistic version of a *static*, task-specific, and handcrafted finite automaton (FA), this volume explains how a genome-like program (Developmental Program, DP) autonomously grows and adapts a brain-like network (Developmental Network, DN).

In contrast with each FA, the DP is task-nonspecific, general purpose, and capable of growing a wide variety of intelligent agents (DNs). Each DN is an attentive, dynamically changing, and probabilistically optimal brain, natural or artificial, that is drastically smaller (e.g., a large constant) than its static FA-equivalent in terms of the number of states (i.e., exponential in the number of concepts to model).

In contrast with traditional artificial intelligence (AI) approaches, humans are not in the loop of handcrafting an FA. Instead, they are part of the external environments, interacting with the DNs as they interact with human children. Different training paths lead to different careers of those DN brains, natural or artificial.

Fundamentally different from traditional artificial neural networks, each DN can abstract, at least in principle, as well as any symbolic AI system.

Like all other books on this highly challenging subject, this volume has not resolved all the details of the brain-mind problem. It is only an introduction to a new departure. Much of the material in this volume seems to be ahead of time, so I beg the reader to tolerate the facts, approaches, methods, analyses, and views expressed in this volume.

Constrained by my limited exposure to the related disciplines, examples of fundamental discipline questions discussed or implied in this volume include:

Biology: How could autonomous and individual cells interact to give rise to animal behaviors, and what cellular roles could the genome likely play?

Neuroscience: From an overarching perspective, how could a brain self-wire, perform top-down attention, and develop its functions?

Psychology: How does an integrated brain architecture accomplish multiple psychological learning models and develop behaviors?

Computer Science: How does a brain-like network compute, adapt, reason, and generalize, and how is the automaton theory related to the brain-like network?

Electrical Engineering: How does a brain-like network perform general-purpose, nonlinear, feedback sensing-and-control, beyond traditional nonlinear control?

Mathematics: How does a brain-like network perform general-purpose, nonlinear optimization, and how does a brain realize emergent functionals?

Physics: How do meanings arise from physics, and how does a brain-like network treat space and time in a unified way, reminiscent of relativity?

Social sciences: How do computational principles of human brains provide insight into possible solutions to a variety of social and political problems?

A basic reason for such a concise volume to be able to discuss the above wide variety of discipline problems is that nature is governed by basic laws that are fewer and more abstract than the number of observable concrete natural phenomena, analogous to how Newton's laws of motion explain rich phenomena of natural motions.

As both a research monograph and a textbook, the problems at the end of each chapter are meant for senior undergraduate students, graduate students, researchers and other interested readers to practice. The problems marked with asterisk "*" are relatively more challenging. The required mathematical background has been reviewed in the Appendix. For those readers who like to see more concrete DN examples, I have cited the related publications. A less analytical reader should be able to get the basic ideas if he pays attention to rich meanings of the mathematical formulations.

Juyang Weng

October, 2012
East Lansing, USA

Chapter 1

Agents and Tasks

An intelligent system is for performing tasks. However, there are many different tasks in all walks of life, from playing chess, to driving a car, to performing scientific research. The subjects of artificial intelligence (AI) include almost all other disciplines. Suppose that one wants to build an intelligent system for a task T in discipline D ; then we may expect that he must first learn knowledge in discipline D so that he knows how to construct an intelligent machine for task T . If we treat AI in this task-specific way, we would probably never be able to finish this course, since we would have to learn almost all disciplines. Therefore, we need to approach AI in a systematic way. The first basic concept of this systematic way is agents. We discuss agents and their environments. From these basic agent-related concepts, we will study various properties of a task, which fall into five categories: external environment, input, internal environment, output and goal.

1.1 Agents

An *agent* is anything that senses and acts, as shown in Fig. 1.1. Thus, any task executor is an agent, regardless if it is natural or artificial. A cat or a human is a natural agent. A computer program or a robot is an artificial agent. A mixed agent is also possible, such as a human whom is controlled by a computer program or a robot remotely controlled by a human.

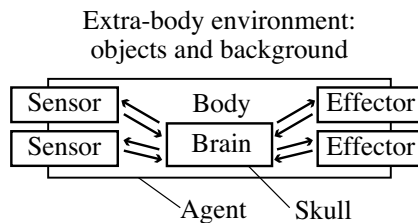


Figure 1.1: The abstract model of an agent