

Tetsuya Hoya

Artificial Mind System – Kernel Memory Approach

Studies in Computational Intelligence, Volume 1

Editor-in-chief

Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw
Poland
E-mail: kacprzyk@ibspan.waw.pl

Further volumes of this series
can be found on our homepage:
springeronline.com

Vol. 1. Tetsuya Hoya
*Artificial Mind System – Kernel Memory
Approach*, 2005
ISBN 3-540-26072-2

Tetsuya Hoya

Artificial Mind System

Kernel Memory Approach

 Springer

Dr. Tetsuya Hoya
RIKEN Brain Science Institute
Laboratory for Advanced
Brain Signal Processing
2-1 Hirosawa, Wako-Shi
Saitama, 351-0198
Japan
E-mail: hoya@brain.riken.jp

Library of Congress Control Number: 2005926346

ISSN print edition: 1860-949X
ISSN electronic edition: 1860-9503
ISBN-10 3-540-26072-2 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-26072-1 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable for prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media
springeronline.com
© Springer-Verlag Berlin Heidelberg 2005
Printed in The Netherlands

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: by the authors and TechBooks using a Springer L^AT_EX macro package

Printed on acid-free paper SPIN: 10997444 89/TechBooks 5 4 3 2 1 0

To my colleagues, educators, and my family

Preface

This book was written from an engineer's perspective of mind. So far, although quite a large amount of literature on the topic of the mind has appeared from various disciplines; in this research monograph, I have tried to draw a picture of the holistic model of an artificial mind system and its behaviour, as concretely as possible, within a unified context, which could eventually lead to practical realisation in terms of hardware or software. With a view that "*mind is a system always evolving*", ideas inspired/motivated from many branches of studies related to brain science are integrated within the text, i.e. artificial intelligence, cognitive science/psychology, connectionism, consciousness studies, general neuroscience, linguistics, pattern recognition/data clustering, robotics, and signal processing. The intention is then to expose the reader to a broad spectrum of interesting areas in general brain science/mind-oriented studies.

I decided to write this monograph partly because now I think is the right time to reflect at what stage we currently are and then where we should go towards the development of "brain-style" computers, which is counted as one of the major directions conducted by the group of "creating the brain" within the brain science institute, RIKEN.

Although I have done my best, I admit that for some parts of the holistic model only the frameworks are given and the descriptions may be deemed to be insufficient. However, I am inclined to say that such parts must be heavily dependent upon specific purposes and should be developed with careful consideration during the domain-related design process (see also the Statements to be given next), which is likely to require material outside of the scope of this book.

Moreover, it is sometimes a matter of dispute whether a proposed approach/model is biologically plausible or not. However, my stance, as an engineer, is that, although it may be sometimes useful to understand the underlying principles and then exploit them for the development of the "artificial" mind system, only digging into such a dispute will not be so beneficial for the development, once we set our ultimate goal to construct the mechanisms

functioning akin to the brain/mind. (Imagine how fruitless it is to argue, for instance, only about the biological plausibility of an airplane; an artificial object that can fly, but not like a bird.) Hence, the primary objective of this monograph is not to seek such a plausible model but rather to provide a basis for imitating the functionalities.

On the other hand, it seems that the current trend in general connectionism rather focuses upon more and more sophisticated learning mechanisms or their highly-mathematical justifications without showing a clear direction/evidence of how these are related to imitating such functionalities of brain/mind, which many times brought me a simple question, “*Do we really need to rely on such highly complex tools, for the pursuit of creating the virtual brain/mind?*” This was also a good reason to decide writing the book.

Nevertheless, I hope that the reader enjoys reading it and believe that this monograph will give some new research opportunities, ideas, and further insights in the study of artificial intelligence, connectionism, and the mind. Then, I believe that the book will provide a ground for the scientific communications amongst various relevant disciplines.

Acknowledgment

First of all, I am deeply indebted to Professor Andrzej Cichocki, Head of the Laboratory for Advanced Brain Signal Processing, Brain Science Institute (BSI), the Institute of Physical and Chemical Research (RIKEN), who is on leave from Warsaw Institute of Technology and gave me a wonderful opportunity to work with the colleagues at BSI. He is one of the mentors as well as the supervisors of my research activities, since I joined the laboratory in Oct. 2000, and kindly allowed me to spend time writing this monograph. Without his continuous encouragement and support, this work would never have been completed. The book is moreover the outcome of the incessant excitement and stimulation gained over the last few years from the congenial atmosphere within the laboratory at BSI-RIKEN. Therefore, my sincere gratitude goes to Professor Shun-Ichi Amari, the director, and Professor Masao Ito, the former director of BSI-RIKEN whose international standing and profound knowledge gained from various brain science-oriented studies have coalized at BSI-RIKEN, where exciting research activities have been conducted by maximally exploiting the centre’s marvelous facilities since its foundation in 1997. I am much indebted to Professor Jonathon Chambers, Cardiff Professorial Fellow of Digital Signal Processing, Cardiff School of Engineering, Cardiff University, who was my former supervisor during my post-doc period from Sept. 1997 to Aug. 2000, at the Department of Electrical and Electronic Engineering, Imperial College of Science, Technology, and Medicine, University of London, for undertaking the laborious proofreading of the entire book written by a non-native English speaker. Remembering the exciting days in London, I would like to express my gratitude to Professor Anthony G.

Constantinides of Imperial College London, who was the supervisor for my Ph.D. thesis and gave me excellent direction and inspiration. Many thanks also go to my colleagues in BSI, collaborators, and many visitors to the ABSP laboratory, especially Dr. Danilo P. Mandic at Imperial College London, who has continuously encouraged me in various ways for this monograph writing, Professor Hajime Asama, the University of Tokyo, Professor Michio Sugeno, the former Head of the Laboratory for Language-Based Intelligent Systems, BSI-RIKEN, Dr. Chie Nakatani and Professor Cees V. Leeuwen of the Laboratory for Perceptual Dynamics, BSI-RIKEN, Professor Jianting Cao of the Saitama Institute of Technology, Dr. Shuxue Ding, at the University of Aizu, Professor Allan K. Barros, at the University of Maranhão (UFMA), and the students within the group headed by Professor Yoshihisa Ishida, who was my former supervisor during my master's period, at the Department of Electronics and Communication, School of Science and Engineering, Meiji University, for their advice, fruitful discussions, inspirations, and useful comments.

Finally, I must acknowledge the continuous and invaluable help and encouragement of my family and many of my friends during the monograph writing.

BSI-RIKEN, Saitama
April 2005

Tetsuya Hoya

Statements

Before moving ahead to the contents of the research monograph, there is one thing to always bear in our mind and then we need to ask ourselves from time to time, “What if we successfully developed artificial intelligence (AI) or humanoids that behaves as real mind/humans? Is it really beneficial to human-kind and also to other species?” In the middle of the last century, the country Japan unfortunately became a single (and hopefully the last) country in the world history that actually experienced the aftermath of nuclear bombs. Then, only a few years later into the new millennium (2000), we are frequently made aware of the peril of bio-hazard, resulting from the advancement in biology and genetics, as well as the world-wide environmental problems. The same could potentially happen if we succeeded the development and thereby exploited recklessly the intelligent mechanisms functioning quite akin to creatures/humans and eventually may lead to our existence being endangered in the long run. In 1951, the cartoonist Osamu Tezuka gave birth to the astro-boy named “Atom” in his works. Now, his cartoons do not remain as a mere fiction but are like to become reality in the near future. Then, they warn us how our life can be dramatically changed by having such intelligent robots within our society; as a summary, in the future we may face to the relevant issues as raised by Russell and Norvig (2003):

- People might lose their jobs to automation;
- People might have too much (or too little) leisure time;
- People might lose their sense of being unique;
- People might lose some of their privacy rights;
- The use of AI systems might result in a loss of accountability;
- The success of AI might mean the end of the human race.

In a similar context, the well-known novel “Frankenstein” (1818) by Mary Shelley also predicted such a day to come. These works, therefore, strongly suggest that it is high time we really needed to start contemplating the (near)

XII Statements

future, where AIs or robots are ubiquitous in the surrounding environment, what we humans are in such a situation, and what sort of actions are necessary to be taken by us. I thus hope that the reader also takes these emerging issues very seriously and proceeds to the contents of the book.

Contents

1	Introduction	1
1.1	Mind, Brain, and Artificial Interpretation	1
1.2	Multi-Disciplinary Nature of the Research	2
1.3	The Stance to Conquest the Intellectual Giant	3
1.4	The Artificial Mind System Based Upon Kernel Memory Concept	4
1.5	The Organisation of the Book	6

Part I The Neural Foundations

2	From Classical Connectionist Models to Probabilistic/Generalised Regression Neural Networks (PNNs/GRNNs)	11
2.1	Perspective	11
2.2	Classical Connectionist/Artificial Neural Network Models	12
2.2.1	Multi-Layered Perceptron/Radial Basis Function Neural Networks, and Self-Organising Feature Maps	12
2.2.2	Associative Memory/Hopfield's Recurrent Neural Networks	12
2.2.3	Variants of RBF-NN Models	13
2.3	PNNs and GRNNs	13
2.3.1	Network Configuration of PNNs/GRNNs	15
2.3.2	Example of PNN/GRNN – the Celebrated Exclusive OR Problem	17
2.3.3	Capability in Accommodating New Classes within PNNs/GRNNs (Hoya, 2003a)	19
2.3.4	Necessity of Re-accessing the Stored Data	20
2.3.5	Simulation Example	20
2.4	Comparison Between Commonly Used Connectionist Models and PNNs/GRNNs	25

2.5	Chapter Summary	29
3	The Kernel Memory Concept – A Paradigm Shift from Conventional Connectionism	31
3.1	Perspective	31
3.2	The Kernel Memory	31
3.2.1	Definition of the Kernel Unit	32
3.2.2	An Alternative Representation of a Kernel Unit	36
3.2.3	Reformation of a PNN/GRNN	37
3.2.4	Representing the Final Network Outputs by Kernel Memory	39
3.3	Topological Variations in Terms of Kernel Memory	41
3.3.1	Kernel Memory Representations for Multi-Domain Data Processing	41
3.3.2	Kernel Memory Representations for Temporal Data Processing	47
3.3.3	Further Modification of the Final Kernel Memory Network Outputs	49
3.3.4	Representation of the Kernel Unit Activated by a Specific Directional Flow	52
3.4	Chapter Summary	57
4	The Self-Organising Kernel Memory (SOKM)	59
4.1	Perspective	59
4.2	The Link Weight Update Algorithm (Hoya, 2004a)	60
4.2.1	An Algorithm for Updating Link Weights Between the Kernels	60
4.2.2	Introduction of Decay Factors	61
4.2.3	Updating Link Weights Between (Regular) Kernel Units and Symbolic Nodes	62
4.2.4	Construction/Testing Phase of the SOKM	63
4.3	The Celebrated XOR Problem (Revisited)	65
4.4	Simulation Example 1 – Single-Domain Pattern Classification .	67
4.4.1	Parameter Settings	67
4.4.2	Simulation Results	68
4.4.3	Impact of the Selection σ Upon the Performance	69
4.4.4	Generalisation Capability of SOKM	71
4.4.5	Varying the Pattern Presentation Order	72
4.5	Simulation Example 2 – Simultaneous Dual-Domain Pattern Classification	73
4.5.1	Parameter Settings	74
4.5.2	Simulation Results	74
4.5.3	Presentation of the Class IDs to SOKM	74
4.5.4	Constraints on Formation of the Link Weights	75
4.5.5	A Note on Autonomous Formation of a New Category .	76

4.6	Some Considerations for the Kernel Memory in Terms of Cognitive/Neurophysiological Context	77
4.7	Chapter Summary	79

Part II Artificial Mind System

5	The Artificial Mind System (AMS), Modules, and Their Interactions	83
5.1	Perspective	83
5.2	The Artificial Mind System – A Global Picture.....	84
5.2.1	Classification of the Modules Functioning With/Without Consciousness	86
5.2.2	A Descriptive Example	87
5.3	Chapter Summary	93
6	Sensation and Perception Modules.....	95
6.1	Perspective	95
6.2	Sensory Inputs (Sensation)	96
6.2.1	The Sensation Module – Given as a Cascade of Pre-processing Units.....	97
6.2.2	An Example of Pre-processing Mechanism – Noise Reduction for Stereophonic Speech Signals (Hoya et al., 2003b; Hoya et al., 2005, 2004c)	98
6.2.3	Simulation Examples	105
6.2.4	Other Studies Related to Stereophonic Noise Reduction	113
6.3	Perception – Defined as the Secondary Output of the AMS ...	114
6.3.1	Perception and Pattern Recognition	114
6.4	Chapter Summary	115
7	Learning in the AMS Context	117
7.1	Perspective	117
7.2	The Principle of Learning	117
7.3	A Descriptive Example of Learning	119
7.4	Supervised and Unsupervised Learning in Conventional ANNs	121
7.5	Target Responses Given as the Result from Reinforcement	122
7.6	An Example of a Combined Self-Evolutionary Feature Extraction and Pattern Recognition Using Self-Organising Kernel Memory	123
7.6.1	The Feature Extraction Part: Units 1)-3)	124
7.6.2	The Pattern Recognition and Reinforcement Parts: Units 4) and 5)	125
7.6.3	The Unit for Performing the Reinforcement Learning: Unit 5).....	126
7.6.4	Competitive Learning of the Sub-Systems	126

7.6.5	Initialisation of the Parameters for Human Auditory Pattern Recognition System	128
7.6.6	Consideration of the Manner in Varying the Parameters i)-v)	129
7.6.7	Kernel Representation of Units 2)-4)	130
7.7	Chapter Summary	131
8	Memory Modules and the Innate Structure	135
8.1	Perspective	135
8.2	Dichotomy Between Short-Term (STM) and Long-Term Memory (LTM) Modules	135
8.3	Short-Term/Working Memory Module	136
8.3.1	Interpretation of Baddeley & Hitch's Working Memory Concept in Terms of the AMS	137
8.3.2	The Interactive Data Processing: the STM/Working Memory \longleftrightarrow LTM Modules	139
8.3.3	Perception of the Incoming Sensory Data in Terms of AMS	140
8.3.4	Representation of the STM/Working Memory Module in Terms of Kernel Memory	141
8.3.5	Representation of the Interactive Data Processing Between the STM/Working Memory and Associated Modules	143
8.3.6	Connections Between the Kernel Units within the STM/Working Memory, Explicit LTM, and Implicit LTM Modules	144
8.3.7	Duration of the Existence of the Kernel Units within the STM/Working Memory Module	145
8.4	Long-Term Memory Modules	146
8.4.1	Division Between Explicit and Implicit LTM	146
8.4.2	Implicit (Nondeclarative) LTM Module	147
8.4.3	Explicit (Declarative) LTM Module	148
8.4.4	Semantic Networks/Lexicon Module	149
8.4.5	Relationship Between the Explicit LTM, Implicit LTM, and Semantic Networks/Lexicon Modules in Terms of the Kernel Memory	149
8.4.6	The Notion of Instinct: Innate Structure, Defined as A Built-in/Preset LTM Module	151
8.4.7	The Relationship Between the Instinct: Innate Structure and Sensation Module	152
8.4.8	Hierarchical Representation of the LTM in Terms of Kernel Memory	153
8.5	Embodiment of Both the Sensation and LTM Modules – Speech Extraction System Based Upon a Combined Blind Signal Processing and Neural Memory Approach	155

8.5.1	Speech Extraction Based Upon a Combined Subband ICA and Neural Memory (Hoya et al., 2003c)	156
8.5.2	Extension to Convolutional Mixtures (Ding et al., 2004) ..	164
8.5.3	A Further Consideration of the Blind Speech Extraction Model	167
8.6	Chapter Summary	168
9	Language and Thinking Modules	169
9.1	Perspective	169
9.2	Language Module	170
9.2.1	An Example of Kernel Memory Representation – the Lemma and Lexeme Levels of the Semantic Networks/Lexicon Module	171
9.2.2	Concept Formation	175
9.2.3	Syntax Representation in Terms of Kernel Memory	176
9.2.4	Formation of the Kernel Units Representing a Concept ..	179
9.3	The Principle of Thinking – Preparation for Making Actions ..	183
9.3.1	An Example of Semantic Analysis Performed via the Thinking Module	185
9.3.2	The Notion of Nonverbal Thinking	186
9.3.3	Making Actions – As a Cause of the Thinking Process ..	186
9.4	Chapter Summary	186
10	Modelling Abstract Notions Relevant to the Mind and the Associated Modules	189
10.1	Perspective	189
10.2	Modelling Attention	189
10.2.1	The Mutual Data Processing: Attention \longleftrightarrow STM/Working Memory Module	190
10.2.2	A Consideration into the Construction of the Mental Lexicon with the Attention Module	192
10.3	Interpretation of Emotion	194
10.3.1	Notion of Emotion within the AMS Context	195
10.3.2	Categorisation of the Emotional States	195
10.3.3	Relationship Between the Emotion, Intention, and STM/Working Memory Modules	198
10.3.4	Implicit Emotional Learning Interpreted within the AMS Context	199
10.3.5	Explicit Emotional Learning	200
10.3.6	Functionality of the Emotion Module	201
10.3.7	Stabilisation of the Internal States	202
10.3.8	Thinking Process to Seek the Solution to Unknown Problems	202
10.4	Dealing with Intention	203

XVIII Contents

10.4.1 The Mutual Data Processing: Attention \longleftrightarrow Intention Module	204
10.5 Interpretation of Intuition	205
10.6 Embodiment of the Four Modules: Attention, Intuition, LTM, and STM/Working Memory Module, Designed for Pattern Recognition Tasks	206
10.6.1 The Hierarchically Arranged Generalised Regression Neural Network (HA-GRNN) – A Practical Model of Exploiting the Four Modules: Attention, Intuition, LTM, and STM, for Pattern Recognition Systems (Hoya, 2001b, 2004b)	207
10.6.2 Architectures of the STM/LTM Networks	208
10.6.3 Evolution of the HA-GRNN	209
10.6.4 Mechanism of the STM Network	214
10.6.5 A Model of Intuition by an HA-GRNN	215
10.6.6 Interpreting the Notion of Attention by an HA-GRNN ..	217
10.6.7 Simulation Example	219
10.7 An Extension to the HA-GRNN Model – Implemented with Both the Emotion and Procedural Memory within the Implicit LTM Modules	226
10.7.1 The STM and LTM Parts	227
10.7.2 The Procedural Memory Part	230
10.7.3 The Emotion Module and Attentive Kernel Units	230
10.7.4 Learning Strategy of the Emotional State Variables	232
10.8 Chapter Summary	234
11 Epilogue – Towards Developing A Realistic Sense of Artificial Intelligence	237
11.1 Perspective	237
11.2 Summary of the Modules and Their Mutual Relationships within the AMS	237
11.3 A Consideration into the Issues Relevant to Consciousness	240
11.4 A Note on the Brain Mechanism for Intelligent Robots	242
References	245
Index	261

List of Abbreviations

ADF	ADaptive Filter
AI	Artificial Intelligence
ALCOVE	Attention Learning COVEring map
ALE	Adaptive Line Enhancer
AMS	Artificial Mind System
ANN	Artificial Neural Network
ARTMAP	Adative Resonance Theory MAP
ASE	Adaptive Signal Enhancer
BP	Back-Propagation
BSE	Blind Signal Extraction
BSP	Blind Signal Processing
BSS	Blind Source Separation
CMOS	Complimentary Metal-Oxide Semiconductor
CR	Conditioned Response
CS	Conditioned Stimuli
DASE	Dual Adaptive Signal Enhancer
DFT	Discrete Fourier Transform
DOA	Direction Of Arrival
ECG	ElectroCardioGraphy
EEG	ElectroEncephaloGraphy
EGO	Emotionally GrOunded
EMG	ElectroMyoGraphy
EVD	EigenValue Decomposition
FIR	Finite Impulse Response
FFT	Fast Fourier Transform
fMRI	functional Magnetic Resonance Imaging
GCM	Generalised Context Model
GMM	Gaussian Mixture Model
GRNN	Generalised Regression Neural Network
HA-GRNN	Hierarchically Arranged Generalised Regression
HMM	Hidden Markov Model

XX List of Abbreviations

HRNN	Hopfield-type Recurrent Neural Network
ICA	Independent Component Analysis
i.i.d.	Independent Identically Distributed
KF	Kernel Function
KM	Kernel Memory
K-Line	Knowledge-Line
LAD	Language Acquisition Device
LIFO	Last-In-Fast-Out
LMS	Least Mean Square
LPC	Linear Predictive Coding
LTD	Long Term Depression
LTM	Long Term Memory
MDIMO	Multi-Domain Input Multi-Output
MEG	MagnetoEncephaloGraphy
MIMO	Multi-Input Multi-Output
MLP-NN	Multi-Layered Perceptron Neural Network
MORSEL	Multiple Object Recognition and Attentional Selection
M-SSP	Multi-stage Sliding Subspace Projection
NLMS	Normalised Least Mean Square
NM	Neural Memory
NN	Neural Network
NR	Noise Reduction
NSS	Nonlinear Spectral Subtraction
PET	Positron Emission Tomography
PNN	Probabilistic Neural Network
PRS	Perceptual Representation System
PSD	Power Spectral Density
QMF	Quadrature Mirror Filter
RBF	Radial Basis Function
SAD	Sound Activity Detection
SAIM	Selective Attention for Identification Model
SDIMO	Single-Domain-Input Single Output
SE	Signal Separation
SFS	Speech Filing System
SIMO	Single-Input Single Output
SLAM	SeLective Attention Model
SNR	Signal-to-Noise Ratio
SOBI	Second-Order Blind Identification
SOFM	Self-Organising Feature Map
SOKM	Self-Organising Kernel Memory
SPECT	Single-Photon Emission Computed Tomography
SRN	Simple Recurrent Network
SS	Signal Separation
SSP	Sliding Subspace Projection
STM	Short Term Memory

SVD	Singular Value Decomposition
SVM	Support Vector Machine
TDNN	Time Delay Neural Network
UR	Unconditioned Response
US	Unconditioned Stimuli
XOR	eXclusive OR