

AutoUni – Schriftenreihe

AutoUni 

Florian Neukart

Reverse Engineering the Mind

Consciously Acting Machines
and Accelerated Evolution

AutoUni – Schriftenreihe

Band 94

Herausgegeben von/Edited by
Volkswagen Aktiengesellschaft
AutoUni

Die Volkswagen AutoUni bietet Wissenschaftlern und Promovierenden des Volkswagen Konzerns die Möglichkeit, ihre Forschungsergebnisse in Form von Monographien und Dissertationen im Rahmen der „AutoUni Schriftenreihe“ kostenfrei zu veröffentlichen. Die AutoUni ist eine international tätige wissenschaftliche Einrichtung des Konzerns, die durch Forschung und Lehre aktuelles mobilitätsbezogenes Wissen auf Hochschulniveau erzeugt und vermittelt.

Die neun Institute der AutoUni decken das Fachwissen der unterschiedlichen Geschäftsbereiche ab, welches für den Erfolg des Volkswagen Konzerns unabdingbar ist. Im Fokus steht dabei die Schaffung und Verankerung von neuem Wissen und die Förderung des Wissensaustausches. Zusätzlich zu der fachlichen Weiterbildung und Vertiefung von Kompetenzen der Konzernangehörigen, fördert und unterstützt die AutoUni als Partner die Doktorandinnen und Doktoranden von Volkswagen auf ihrem Weg zu einer erfolgreichen Promotion durch vielfältige Angebote – die Veröffentlichung der Dissertationen ist eines davon. Über die Veröffentlichung in der AutoUni Schriftenreihe werden die Resultate nicht nur für alle Konzernangehörigen, sondern auch für die Öffentlichkeit zugänglich.

The Volkswagen AutoUni offers scientists and PhD students of the Volkswagen Group the opportunity to publish their scientific results as monographs or doctor's theses within the "AutoUni Schriftenreihe" free of cost. The AutoUni is an international scientific educational institution of the Volkswagen Group Academy, which produces and disseminates current mobility-related knowledge through its research and tailor-made further education courses. The AutoUni's nine institutes cover the expertise of the different business units, which is indispensable for the success of the Volkswagen Group. The focus lies on the creation, anchorage and transfer of new knowledge.

In addition to the professional expert training and the development of specialized skills and knowledge of the Volkswagen Group members, the AutoUni supports and accompanies the PhD students on their way to successful graduation through a variety of offerings. The publication of the doctor's theses is one of such offers. The publication within the AutoUni Schriftenreihe makes the results accessible to all Volkswagen Group members as well as to the public.

Herausgegeben von/Edited by

Volkswagen Aktiengesellschaft

AutoUni

Brieffach 1231

D-38436 Wolfsburg

<http://www.autouni.de>

Florian Neukart

Reverse Engineering the Mind

Consciously Acting Machines
and Accelerated Evolution

 Springer

Florian Neukart
Wolfsburg, Germany

Any results, opinions and conclusions expressed in the AutoUni Schriftenreihe are solely those of the author(s).

AutoUni – Schriftenreihe
ISBN 978-3-658-16175-0 ISBN 978-3-658-16176-7 (eBook)
DOI 10.1007/978-3-658-16176-7

Library of Congress Control Number: 2016955691

© Springer Fachmedien Wiesbaden GmbH 2017

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, 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.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Fachmedien Wiesbaden GmbH
The registered company address is: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

*The only way to discover the limits of the possible
is to go beyond them into the impossible.*

Sir Arthur C. Clarke

Preface

What we perceive as consciousness seems to be an anomaly, and so is intelligence. Earth features a biodiversity of around $8.7 * 10^6 \pm 1.3 * 10^6$ SE (between $\sim 7,400,000$ and $\sim 10,000,000$) organisms,¹ from those just a few show rudimentary forms of both consciousness and intelligence, and only one seems to be aware and discuss what it means to exist. Although more than 1,000 exoplanets have been detected thitherto (latest estimations predict around $1.7 * 10^9$ only in our galaxy), and 12 of them are probably habitable, there has not been any indication that intelligent life apart from earth has emerged elsewhere. This may be interpreted in many ways, whereby the most appealing ones for me are:

- Intelligence is an anomaly.
This assumption states that the probability for the evolution of human-like intelligence in a species is infinitely small. No other species on earth seems to have developed it, which can be considered as strong indication for the hypothesis that if extraterrestrial species exist, they may not have developed it either. A similar thought experiment is about the physical laws governing our universe. If universe parameters, such as the speed of light or the strength of gravity, would have been allowed to take any value from 0 below infinity, then the occurrence of the current set of parameters and parameter values governing the very existence of the universe (and life) as we know it is infinitely low. Although I personally do not hope and guess that this hypothesis is true, it has not yet been disproved by scientific evidence.
- We are the first ones searching for intelligent life in our galaxy.
Considering the size of our galaxy and the time the universe already exists, it is very likely that if intelligent life except from ours exists in whatever form, we will detect it sooner or later, but: if we are not the only intelligent species in the universe it is, given the almost 14 billion years that elapsed since the big bang, very unlikely that we are the first ones having evolved intelligence, and in consequence the technology to communicate and travel through space. More than half a century ago, even Enrico Fermi has faced the paradox that considering the age of our universe, it should be very likely that intelligent life has already emerged elsewhere. But where is it, then? This became widely known as the Fermi-paradox.
- Other intelligent species have already discovered us, but do not (or cannot) show up for some reason.
- No other species has solved the problem of how to produce exotic energy, which is, as to current knowledge, required for travelling faster than light (to be exact, it is not travelling faster than light, but bridging of distances by warping space-time). As no matter can travel faster than light, it must be space-time that is moved through space-time. This can theoretically be done by creating a warp-field around an area of space-time (and an object like a spaceship within this area), which is bordered by a singularity in front of the object and controlled expansion of space-time behind it. In theory this is possible, as has already been shown with a special solution of Einstein's field equations in general relativity. This solution states that the creation of an energy-impulse-tensor modifying the space-time around a spacecraft in the sense that the distance between start- and endpoint can be

1 Tittensor Derek P. et al. (2011): How Many Species Are There on Earth and in the Ocean? [2013-10-24]; URL: <http://www.plosbiology.org/article/info%3Adoi%2F10.1371%2Fjournal.pbio.1001127>

reduced is possible. However, for creating such a warp-field exotic matter, thus matter not only not consisting of protons, neutrons or electrons, but additionally being of negative energy density must be available² – this may be a problem, as we have not even been able to proof that it exists.

From a solely scientific point of view all of these and a lot more arguments are valid options, but I will focus on the first one here. Mankind has always dreamed of immortality, which has, amongst others, been one of the major reasons for why our species devised the concept of religion. Our brain provides us with the abilities required to understand the universe, and the more we understand about our role in the universe and life itself, the more transhumanistic our views become. It may be not easy to accept that we will most likely take care of our immortal soul ourselves in the not-so-distant future, but this is exactly what will happen. Problems like starvation or over-population may be solved by leaving our bodily existence behind, and hundreds of years-lasting journeys to exoplanets for colonization will not pose a challenge for human minds transferred into the computers of spaceships. This is just a first impression of what the future of mankind could be like. The ideas are countless, and history taught us that advancement of science not always bears only good. Anyway, this work marks the beginning of a journey – the journey towards consciously acting machines and artificially accelerated human evolution.

Florian Neukart³

2 Alcubierre Miguel (1994): The warp drive: hyper-fast travel within general relativity; *Classical and Quantum Gravity* 11:L73-L77, 1994

3 I am especially eager to note any copyrights of multimedia elements such as images and texts used, and if possible to use graphics and text of my own. However, most of the time scientific work is based on the work already done by others, which becomes apparent by having a look at the number of publications cited within this elaboration. Every publishing scientist knows that one challenge is that the more one reads and studies about a topic, the easier the boundaries of their own ideas get blurred with those of others' ideas. Thus, if in this elaboration unmarked, but by third party copyright protected images or a text are found, it was not possible for me to detect the related copyright. In case of such an unintentional copyright violation I will remove the corresponding picture or text element or will mark it with an appropriate copyright notice/citation of sources indicated in the next version of the publication after a short notification.

Table of Contents

Preface	VII
List of Figures	XVII
List of Tables	XXI
List of Algorithms	XXIII
List of Abbreviations	XXV
Introduction	1
Structure	3
1 Evolution’s most extraordinary achievement.....	7
1.1 Anatomy of the human brain.....	7
1.1.1 Truncus cerebri.....	9
1.1.1.1 Cerebellum.....	9
1.1.1.2 Mesencephalon.....	10
1.1.1.3 Pons	11
1.1.1.4 Medulla oblongata.....	12
1.1.2 Paleomammalian	12
1.1.2.1 Corpus amygdaloideum.....	13
1.1.2.2 Hippocampus	14
1.1.2.3 Diencephalon	14
1.1.2.3.1 Hypothalamus.....	15
1.1.2.3.2 Subthalamus	16
1.1.2.3.3 Thalamus dorsalis	16
1.1.2.3.4 Pineal gland and Epithalamus	16
1.1.2.4 Cingulate gyrus	16
1.1.3 Cortex and neocortex.....	17
1.1.3.1 Frontal lobe	18
1.1.3.2 Parietal lobe	18
1.1.3.3 Temporal lobe	19
1.1.3.4 Occipital lobe	20
1.2 Neural information transfer.....	21
1.3 Summary.....	24
2 Pillars of artificial intelligence.....	25
2.1 Machine learning	26
2.1.1 Supervised learning algorithms	26
2.1.2 Unsupervised Learning Algorithms.....	26
2.2 Computer Vision	28
2.3 Logic and reasoning.....	30
2.4 Language and communication	32
2.5 Agents and actions.....	34

2.5.1 Principles of the new agent-centered approach.....	34
2.5.2 Multi-agent behavior	35
2.5.3 Multi-agent learning	36
2.6 Summary	36
3 An outline of artificial neural networks.....	39
3.1 Definition	39
3.2 Paradigms of computational intelligence.....	41
3.3 Neural networks	42
3.3.1 Artificial neural networks	42
3.3.1.1 Suitable problems	44
3.3.1.2 Basic knowledge	45
3.3.1.2.1 Structure	45
3.3.1.2.2 Bias	45
3.3.1.2.3 Gradient descent	45
3.3.1.3 Activation functions	46
3.3.1.3.1 Linear activation function	46
3.3.1.3.2 Sigmoid activation function	47
3.3.1.3.3 Hyperbolic tangent activation function.....	47
3.3.1.3.4 Rectifier linear unit	48
3.3.1.3.5 Gaussian activation function	49
3.3.1.4 Regularization	50
3.3.2 Types of artificial neural networks.....	51
3.3.2.1 Supervised and unsupervised learning	51
3.3.2.2 Feed-forward artificial neural network.....	51
3.3.2.3 Feed-forward artificial neural network with feedback connections.....	51
3.3.2.4 Fully connected artificial neural network.....	52
3.3.2.5 Basic artificial neural network structure.....	52
3.3.2.6 Perceptron	53
3.3.2.6.1 Single layer perceptron	53
3.3.2.6.2 Multi layer perceptron.....	56
3.3.2.6.3 Spiking artificial neural networks.....	58
3.3.2.7 Radial basis artificial neural network.....	59
3.3.2.8 Recurrent artificial neural network	59
3.3.2.8.1 Elman recurrent artificial neural network	60
3.3.2.8.2 Jordan recurrent artificial neural network	61
3.3.2.9 Fully connected artificial neural network	62
3.3.2.9.1 Hopfield artificial neural network.....	64
3.3.2.9.2 Boltzmann Machine	66
3.3.2.9.3 Support vector machine.....	71
3.3.2.9.4 Self-organizing feature map	72
3.3.2.9.5 Committee machines.....	74
3.3.3 Training and learning.....	77
3.3.3.1 Supervised and unsupervised training	77
3.3.3.2 (Root) mean squared error	78
3.3.3.3 Estimators	79
3.3.3.4 Hebb's learning rule.....	79

3.3.3.5	Delta rule.....	80
3.3.3.6	Propagation learning.....	81
3.3.3.6.1	Back propagation training.....	82
3.3.3.6.2	Manhattan update rule training.....	86
3.3.3.6.3	Resilient propagation.....	87
3.3.3.7	Genetic learning (NeuroEvolution).....	90
3.3.3.7.1	Evolutionary search of connection weights.....	93
3.3.3.7.2	Evolutionary search of architectures.....	93
3.3.3.7.3	Evolutionary search of learning rules.....	94
3.3.3.8	Simulated annealing.....	94
3.3.3.9	NeuroEvolution of augmenting topologies (NEAT).....	96
3.3.4	Stability-plasticity dilemma.....	99
3.4	Summary.....	100
4	Advanced artificial perception and pattern recognition.....	103
4.1	Convolutional artificial neural networks.....	104
4.1.1	Data representation.....	105
4.1.2	Structure.....	106
4.1.2.1	Convolutional layers.....	107
4.1.2.2	Different ways of perception and processing.....	111
4.1.2.3	Maxpooling/ downsampling layers.....	112
4.1.2.4	Feature maps.....	112
4.1.2.5	Fully connected layers.....	114
4.1.2.6	Number of neurons.....	114
4.1.3	Training.....	114
4.2	Deep belief artificial neural network.....	118
4.2.1	Stacking together RBMs.....	122
4.2.2	Training.....	122
4.3	Cortical artificial neural network.....	123
4.3.1	Structure.....	125
4.3.1.1	Cortices.....	127
4.3.1.2	Number of neurons.....	127
4.3.1.3	Synapses.....	128
4.3.2	A generic cortical artificial neural network.....	128
4.3.3	Purpose.....	130
4.3.4	Evolution and weight initialization.....	131
4.4	SHOCID recurrent artificial neural network.....	134
4.4.1	Structure.....	134
4.4.1.1	Recurrent layer one.....	135
4.4.1.2	Recurrent layer two.....	136
4.4.1.3	Number of neurons.....	137
4.4.1.4	Synapses.....	138
4.4.2	Purpose.....	138
4.4.3	Evolution and weight initialization.....	139
4.5	Summary.....	141

5	Advanced nature-inspired evolution and learning strategies	143
5.1	Transgenetic NeuroEvolution	143
5.1.1	Fundamentals	144
5.1.2	Host genetic material	145
5.1.3	Endosymbiont.....	146
5.1.4	Algorithm.....	146
5.1.5	Horizontal (endosymbiotic) gene (sequence) transfer.....	147
5.1.5.1	Weight plasmid	148
5.1.5.2	Structure plasmid.....	148
5.1.6	Transposon mutation	149
5.1.6.1	Jump and swap transposon	149
5.1.6.2	Erase and jump transposon	150
5.1.7	Usage	150
5.2	Artificial immune system-inspired NeuroEvolution	150
5.2.1	Fundamentals	151
5.2.2	Clonal selection and somatic hypermutation	152
5.2.3	Danger theory, virus attack and hyperrecombination.....	155
5.2.4	Negative selection	158
5.2.5	Overall algorithm.....	160
5.2.6	Causality	161
5.2.7	Usage	162
5.3	Structural evolution	162
5.3.1	Fundamentals	162
5.3.2	Algorithm.....	163
5.3.3	Generic determination of artificial neural network quality.....	165
5.3.4	Parameterization.....	167
5.3.5	Usage	167
5.4	Summary	167
6	Autonomously acting cars and predicting market behaviour: some application scenarios for ANNs	169
6.1	Analysis and knowledge	169
6.1.1	Supervised and unsupervised functions.....	171
6.1.2	Classification.....	171
6.1.3	Regression.....	172
6.1.4	Clustering.....	172
6.1.5	Attribute importance.....	173
6.1.6	Association.....	173
6.1.7	Interesting knowledge.....	173
6.1.8	Accurate knowledge	174
6.1.9	Interpretable knowledge.....	174
6.1.10	Intelligent processing.....	174
6.1.11	Efficient processing	174
6.2	Autonomously acting cars.....	175
6.2.1	V2X-communication	176
6.2.2	Massively equip car with processing power and AI-algorithms.....	176

6.2.3 Artificial intelligence and environment sensing.....	176
6.2.3.1 Cameras and how AI is applied to related data.....	177
6.2.3.2 RADAR and how AI is applied to related data.....	177
6.2.3.3 LiDAR and how AI is applied to related data.....	178
6.2.3.4 Additional sensors and how AI is applied to related data	178
6.2.3.5 GPS and how AI is applied to related data	179
6.2.3.6 Microphones and how AI is applied to related data.....	179
6.2.3.7 Autonomously acting car's brain – the domain controller	179
6.3 Summary.....	180
7 An outline of quantum mechanics.....	181
7.1 Quantum systems in general	181
7.1.1 Quantum theory.....	182
7.1.1.1 Quantum states.....	182
7.1.1.2 Observables.....	183
7.1.1.3 Quantum measurements	184
7.1.1.4 Quantum dynamics.....	185
7.1.2 Quantum operators	186
7.1.3 Quantum physical effects.....	190
7.1.3.1 Quantum interference	190
7.1.3.2 Quantum linear superposition.....	192
7.1.3.3 Quantum entanglement.....	193
7.2 The unitary evolution U.....	195
7.3 The state vector reduction R	198
7.4 Summary.....	219
8 Quantum physics and the biological brain	221
8.1 Difficulties with U in the macroscopic world.....	222
8.2 The Hameroff-Penrose model of orchestrated objective reduction.....	224
8.2.1 The idea.....	224
8.2.2 Microtubules	225
8.3 Further models.....	228
8.4 Summary.....	229
9 Matter and consciousness.....	231
9.1 Qualia.....	231
9.2 Materialism	232
9.2.1 Eliminative materialism.....	233
9.2.2 Noneliminative materialism.....	233
9.3 Functionalism.....	233
9.3.1 The problem of absent or inverted qualia	234
9.3.2 The Chinese Room argument.....	234
9.3.3 The knowledge argument.....	235

9.4 The Identity Theory	235
9.5 Summary	236
10 Reverse engineering the mind	237
10.1 Theory of mind	237
10.2 Quantum linear superposition in artificial brains	238
10.3 Self-organization	241
10.3.1 Structure and system	242
10.3.1.1 Conservative structure	242
10.3.1.2 Dissipative structure	243
10.3.2 Self-organization in computational intelligence	243
10.3.2.1 Self-organized learning	246
10.3.2.2 Learning with respect to self-organization	246
10.3.2.2.1 Competitive learning	248
10.3.2.2.2 Competitive learning in artificial neural networks	249
10.3.2.3 Adaptive Resonance Theory	251
10.3.3 The transition to the human brain	254
10.3.3.1 Laterally interconnected synergetically self-organizing maps	257
10.3.3.2 The pruning neocortex	259
10.3.3.2.1 Incremental pruning	260
10.3.3.2.2 Selective pruning	260
10.3.3.2.3 Pruning and quantum artificial neural networks	260
10.3.4 Arguments for self-organization in artificial neural systems	261
10.4 Mechanisms apart from self-organization	261
10.4.1 Leader	262
10.4.2 Blueprint	262
10.4.3 Recipe	262
10.4.4 Template	262
10.5 Quantum physics and the artificial brain	263
10.5.1 Quantum artificial neural network	263
10.5.1.1 Structure	264
10.5.1.2 Quantum bits	266
10.5.1.3 Superposition	266
10.5.1.3.1 Superposition of dendrites	266
10.5.1.3.2 Superposition of neurons	267
10.5.1.3.3 Superposition of the quantum artificial neural network	267
10.5.1.4 Entanglement	268
10.5.1.5 Interference	275
10.5.1.6 Processing	277
10.5.1.6.1 Entanglement	278
10.5.1.6.2 Quantum parallelism	279
10.5.1.6.3 From basic operators to the quantum transfer function	279
10.5.1.6.4 Reduction of and information about the quantum perceptron equations	284
10.5.1.6.5 Normalization	289
10.5.1.7 Measurement	290

10.5.1.7.1	Quantum artificial neural network configuration search function	290
10.5.1.7.2	Example processing	291
10.5.1.8	Envisaged implementations of a quantum artificial neural network.....	295
10.5.1.8.1	Adiabatic quantum annealing	298
10.5.1.8.2	Nuclear magnetic resonance.....	298
10.5.1.8.3	Others.....	299
10.6	The artificial neocortex.....	299
10.6.1	Knowledge and data	301
10.6.1.1	Knowledge representation	302
10.6.1.2	Declarative knowledge representation	302
10.6.1.2.1	Semantic networks.....	303
10.6.1.2.2	Object-attribute-value-triplet.....	304
10.6.1.2.3	Frames.....	305
10.6.2	Context recognition and hierarchical learning	309
10.6.2.1	Definition of context-sensitive information.....	310
10.6.2.2	Information Clustering	311
10.6.2.3	Context analysis	312
10.6.2.4	Hierarchical learning	313
10.6.2.5	Interpreting the context.....	315
10.6.2.6	Hidden Markov models and conceptual hierarchies in the neocortex.....	317
10.6.3	Implementation.....	325
10.6.3.1	Acquisition of basic knowledge	328
10.6.3.2	Encoding the acquired knowledge into pattern recognizers	329
10.6.3.3	Access to knowledge and how search engines are similar to the brain.....	333
10.6.3.4	Language processing and understanding	337
10.6.3.5	Quantum pattern recognizers	343
10.6.3.6	Real world input and new experiences.....	345
10.6.3.7	Automatic information interconnection.....	346
10.6.4	A superior goal	346
10.7	A distributed mind	347
10.7.1	Non-invasive transducers.....	349
10.7.2	Semi-invasive, invasive transducers and the neural grid.....	350
10.7.3	Signal processing.....	350
10.7.3.1	<i>Pre-processing</i>	351
10.7.3.2	Feature extraction.....	351
10.7.3.3	Detection and classification	351
10.7.4	BCI requirements for the distributed mind	352
10.8	Summary.....	353
11	Conclusion.....	355
	Glossary – computational intelligence	357
	Glossary – quantum physics.....	365
	Bibliography.....	371

List of Figures

Figure 1 - Human brain.....	8
Figure 2 - Triune brain.....	8
Figure 3 – Truncus cerebri.....	9
Figure 4 - Cerebellum.....	10
Figure 5 - Mesencephalon.....	11
Figure 6 - Pons.....	11
Figure 7 - Medulla oblongata.....	12
Figure 8 - Paleomammalian.....	13
Figure 9 - Corpus amygdaloideum.....	13
Figure 10 - Hippocampus.....	14
Figure 11 - Diencephalon.....	15
Figure 12 - Hypothalamus.....	16
Figure 13 - Frontal lobe.....	18
Figure 14 - Parietal lobe.....	19
Figure 15 – Temporal lobe.....	20
Figure 16 - Occipital lobe.....	21
Figure 17 - Neuron.....	21
Figure 18 - Neuron types.....	22
Figure 19 - Action potential.....	23
Figure 20 - McCulloch & Pitts neuron model.....	43
Figure 21 - Representative processing model.....	43
Figure 22 - Linear activation.....	46
Figure 23 - Sigmoid activation.....	47
Figure 24 - Tangens hyperbolicus activation.....	48
Figure 25 - ReLu activation.....	48
Figure 26 - Gauss activation.....	49
Figure 27 - Simple artificial neural network structure.....	52
Figure 28 – Single layer perceptron.....	54
Figure 29 – OR operator.....	56
Figure 30 – AND operator.....	56
Figure 31 – XOR operator.....	56
Figure 32 – Multi layer perceptron.....	57
Figure 33 - Elman artificial neural network.....	61
Figure 34 - Jordan artificial neural network.....	62

Figure 35 - Hopfield artificial neural network	65
Figure 36 - Boltzmann machine	67
Figure 37 - Restricted Boltzmann machine	68
Figure 38 – Self organizing feature map	73
Figure 39 - Committee machine.....	75
Figure 40 - Hebb's rule	80
Figure 41 - Delta rule.....	81
Figure 42 – NeuroEvolution of augmenting topologies mutation	97
Figure 43 – NeuroEvolution of augmenting topologies recombination of different topologies	98
Figure 44 - 4-dimensional tensor	105
Figure 45 – CNN sparse interconnectivity	106
Figure 46 - CNN architecture.....	107
Figure 47 – Convolution	107
Figure 48 – $\mathbf{x}(\mathbf{t})$	108
Figure 49 – $\mathbf{x}(\boldsymbol{\tau})$	109
Figure 50 – $\mathbf{h}(\mathbf{t})$	109
Figure 51 – $\mathbf{h}\boldsymbol{\tau}$ reflected	110
Figure 52 – $\mathbf{h}\boldsymbol{\tau}$ reflected and shifted	110
Figure 53 – Convolution of $\mathbf{x}\mathbf{t}$ and $\mathbf{h}\mathbf{t} - \boldsymbol{\tau}$	110
Figure 54 – Activation map	112
Figure 55 – Downsampled activation map	112
Figure 56 – Convolution calculation	113
Figure 57 - Featuremap generation	115
Figure 58 - Simple ANN.....	119
Figure 59 – Causal brain influence types	124
Figure 60 – Pre-cortical artificial neural network structure.....	125
Figure 61 – Cortical artificial neural network structure	126
Figure 62 – Cortical artificial neural network structure	129
Figure 63 – SHOCID recurrent artificial neural network single hidden layer.....	135
Figure 64 – SHOCID recurrent artificial neural network multi hidden layer.....	137
Figure 65 - Transgenic NeuroEvolution	145
Figure 66 - Bloch sphere.....	196
Figure 67 - Euler's formula	198
Figure 68 - Photon and half-silvered mirror	202
Figure 69 - Photon, half-silvered and fully-silvered mirrors	202

Figure 70 - Infinite potential step	204
Figure 71 - Finite potential step	207
Figure 72 - Particle lacks energy	211
Figure 73 - Potential barrier	211
Figure 74 - Potential barrier - forces	212
Figure 75 - ART1 Structure	252
Figure 76 - Monkey striate cortex recording,	255
Figure 77 - Quantum artificial neural network	265
Figure 78 - Quantum teleportation network unit	270
Figure 79 - cNOT from H and V	281
Figure 80 - Toffoli gate with controlled V	281
Figure 81 - Toffoli-gate with complex conjugate transpose V	282
Figure 82 - Quantum addition	283
Figure 83 - Quantum artificial neural network calculations	286
Figure 84 - Quantum single layer perceptron diagram	287
Figure 85 - Quantum multi layer perceptron diagram	288
Figure 86 - Reverse the calculation bits	288
Figure 87 - Rotation towards $ xd\rangle$	293
Figure 88 - Quantum Hopfield artificial neural network	296
Figure 89 - Quantum Boltzmann machine	297
Figure 90 - Semantic network	303
Figure 91 - Representation of n-digit predicates	304
Figure 92 - O-A-V-triplet	305
Figure 93 - Markov chain weather prediction	319
Figure 94 - Markov chain interpretation of speech signals	319
Figure 95 - Markov chain	320
Figure 96 - Hidden Markov model	321
Figure 97 - Hidden Markov model weather observations	322
Figure 98 - Hierarchically hidden Markov model	324
Figure 99 - Pattern presented to multiple pattern recognizers	326
Figure 100 - Hierarchical pattern processing	327
Figure 101 - Binary decision tree	331
Figure 102 - Bottom-up ANN tree	332
Figure 103 - Cumulative activation	333
Figure 104 - Viterbi example	338
Figure 105 - Variable explanation	340

List of Tables

Table 1 -	Quality determination	166
Table 2 –	(Quantum) artificial neural network feature comparison	265
Table 3 –	O-A-V-triplet.....	305
Table 4 –	Car-frame	306
Table 5 –	Engine-frame.....	307

List of Algorithms

Algorithm 1 - Basic perceptron learning	55
Algorithm 2 – RBM learning.....	71
Algorithm 3 – SOM learning.....	74
Algorithm 4 - Committee of SA feed-forward ANNs.....	77
Algorithm 5 – Back propagation algorithm.....	85
Algorithm 6 - Manhattan update rule	87
Algorithm 7 - Resilient propagation.....	89
Algorithm 8 - Genetic algorithm.....	91
Algorithm 9 - Simulated annealing algorithm	96
Algorithm 10 – CNN back propagation algorithm	118
Algorithm 11 – DBN training.....	123
Algorithm 12 - Evolution of cortical ANN.....	133
Algorithm 13 - Evolution of SRANN.....	140
Algorithm 14 - Transgenetic NeuroEvolution	147
Algorithm 15 – Clonal selection algorithm.....	153
Algorithm 16 - Clonal selection and hypermutation	154
Algorithm 17 – Danger theory algorithm.....	156
Algorithm 18 - Danger theory, virus attack and hyperrecombination.....	157
Algorithm 19 – Negative selection algorithm	158
Algorithm 20 - Negative selection	160
Algorithm 21 – Immune system-inspired NeuroEvolution.....	161
Algorithm 22 - Structural evolution	164
Algorithm 23 – Quality determination	166
Algorithm 24 – Competitive training.....	250
Algorithm 25 – Adaptive resonance theory.....	253
Algorithm 26 - Lifting weights into superposition.....	267
Algorithm 27 – Quantum teleportation	269
Algorithm 28 – Quantum teleportation artificial neural network	271
Algorithm 29 – Quantum teleportation artificial neural network processing.....	274
Algorithm 30 – Quantum input normalization	289
Algorithm 31 – Word comparison	311
Algorithm 32 – Growing SOFM.....	314
Algorithm 33 – Context interpretation	316
Algorithm 34 – Creation of the sequence of symbols.....	321

Algorithm 35 – Viterbi algorithm	339
Algorithm 36 – Forward recursion.....	341
Algorithm 37 – Backward recursion.....	342

List of Abbreviations

AI	Artificial Intelligence
AIS	Artificial Immune System
ANN	Artificial Neural Network
AOD	Agent Oriented Development
AOP	Agent Oriented Programming
APC	Antigen-Presenting Cell
BP	Back propagation
CI	Computational Intelligence
CSV	Comma separated values
DBN	Deep Belief Network
DM	Data Mining
FFANN	Feed-forward Artificial Neural Network
ETL	Extract, Transform, Load
GA	Genetic Algorithm
MLP	Multi Layer Perceptron
NDS	Neural Data Set
OOP	Object Oriented Programming
PAMP	Pathogen-Associated Molecular Patterns
QANN	Quantum ANN
RANN	Recurrent ANN
RBM	Restricted Boltzmann Machine
RP	Resilient-propagation
(R)MSE	(Root) Mean Squared Error
SA	Simulated annealing
SHOCID	System applying High Order Computational Intelligence in Data Mining
SRANN	SHOCID RANN

Introduction

This elaboration is, in some sense, the first version of a manual describing how to implement an artificial conscious entity and how to extend the very human existence beyond biological limitations. Subsequent versions depend on future research not only conducted by the author, but by numerous scientists from various fields.

Since there is the research field of artificial intelligence (AI), one of the biggest hurdles has always been the creation of conscious experiences in machines. Not only lots of different definitions of what exactly consciousness is exist from a philosophical point of view; it is yet also not completely understood on a neuroscientific level how our brain creates conscious content. Within this elaboration, I will provide a foundation for understanding how conscious experiences emerge by approaching the topic by means of actual and future technology. In my opinion, one of the most important foundations of consciousness is, amongst self-awareness, the ability to understand concepts, or more generally, the understanding of 'things'. I define these things to comprise everything that exists, be it a single atom or a complex lifeform. When it comes to understanding things, learning is an important aspect, and for being able to measure levels of understanding, it is first required to define and quantify when something has been understood. I want to emphasize this challenge at the very beginning of this elaboration, as words such as 'learning', 'understanding' and 'consciousness' are suitcase-words from psychology helping us to discuss complex subjects science has not yet entirely understood. Suitcase-words such as consciousness enable us to include yet unknown processes and (changes of) states associated with the human brain in our everyday-language with ease, not only without being able to describe what accounts for a conscious experience on neuronal or (sub-) atomic layers, but also without being able to explain what consciousness is on a more abstract layer. Lots of scientists from different fields have been working on disclosing the secret of consciousness, and numerous different explanations have been published and discussed controversially. Most of these theories feature a common denominator – the inclusion of a feature set, which is associated with the perception of consciousness. It would be counterproductive to reject such approaches, as only the detailed description and combination of single features will allow us to reproduce conscious behavior in artificial entities. We will deal with this and a lot more in this book, and also define these features crisply, as only then we will be able to create hard- and software capable of not only processing information in the way the human brain does, but also capable of reproducing the conditions that are required for creating conscious experiences.

Not only philosophers, but scientists from numerous different fields have long tried to understand what it is that creates such experiences; from what we can see today, some of the attempts have already been crowned by success. The strength of current AI is not only justified in the fact that today's computers can do things better at which machines have used to be better than humans since the first successful implementations of paradigms of AI. Back in 1996, as IBM's Deep Blue beat Garri Kasparov in chess, one could have argued that this had nothing to do with real intelligence, but resulted from the fact that Deep Blue could calculate $2 * 10^6$ chess moves per second. However, another of IBM's masterpieces, Watson, showed that by the combination of linguistic pre-processors, expert systems, search engines, machine learning, logic, natural language understanding, and by accessing data sources of various kinds intelligently, understanding can be emulated, which makes it more difficult for opponents of AI to argue against it. Both examples are very impressive not only from scientific and engineering points of view, however although these systems are artificially

intelligent, they cannot be considered conscious beings (and, as of now, it has never been the aim of making them conscious).

For what it can mean to create conscious machines becomes even more interesting by taking into account the research of Stuart Hameroff and Roger Penrose, who state that quantum physical phenomena take place on cell-basis, or even on the basis of tubulin dimers, elements occurring within nerve (and all other eukaryotic) cells' cytoskeletons: if quantum coherence, which may constitute a requirement for consciousness, may possibly be maintained at this level, this would result in a dramatic increase of the number of operations per second a brain is able to accomplish. Even if quantum physical phenomena have not yet been linked to the creation of conscious content by proof, it is nevertheless useful taking such into account for the implementation of artificial entities that should emulate or experience conscious content themselves. Thus, within this work the implementation of quantum artificial neural networks (by the leveraging the power of quantum computers) is juxtaposed in opposition to classical artificial neural networks in terms of their probable capabilities and usefulness for the implementation of an artificial mind.

The creation of consciousness in an artificial entity brings up lots of new questions, particularly the one after human immortality; this may not be obvious at a first glance, but if artificial entities are capable of producing conscious content and may theoretically live forever, can we humans then benefit thereof in the sense of transferring our minds, dreams and desires into such vessels, one may ask. But then there is the continuous consciousness-problem: if we copy our mind from our biological brain into an artificial vessel, it is just that – a copy, and in the worst case two of us exist at the same time, one doomed to die and the other blessed to live forever. I will also propose a solution to this problem.

Finally, I ask the reader to pardon me for the technical style of writing in which I sometimes slide. I am admittedly used to creating scientific reports/ papers, or technical and functional documentations, but not so much to writing bedtime lecture.

Last, but not least, I sometimes refer to some computational intelligence paradigms with the prefix 'SHOCID'. This is, because the ideas thereto emerged during the specification phase of the data mining system SHOCID (System Applying High Order Computational Intelligence in Data Mining),⁴ which I developed some years ago. This gives the elaboration at hand a practical touch – most of the discussed and introduced computational intelligence-paradigms have been implemented in SHOCID.

4 Neukart Florian (2013): System Applying High Order Computational Intelligence in Data Mining and Quantum Computational Considerations Concerning the Future of Artificial Intelligence; Brasov: Transilvania University of Brasov

Structure

Chapter 1

In this chapter basic brain functionality as well as the rough anatomy of the human brain is discussed. This is mainly because as an artificial intelligence researcher one has to deal with natural intelligence and how the brain works sooner or later. A subfield of artificial intelligence, computational intelligence, comprises many nature-inspired approaches such as swarm intelligence, genetic algorithms or artificial neural networks. Especially some kinds of artificial neural networks, by means of which we are concerned with creating more or less simple imitations of their biological counterparts, have achieved information processing similar to what we currently understand happens within biological brains.

Chapter 2

The major goal of this elaboration is to work out how specific aspects of the human mind, namely those responsible for higher cognitive functions, function, and to figure out which hardware is required to process an artificial mind with the same, similar or superior capabilities. For this, I will focus on any approach that allows us to reproduce cognitive capabilities, but not necessarily achieve this target by the same means as evolution did. It makes sense, at this point, to provide an introduction to artificial intelligence in order to understand which areas of the neocortex are subject to AI research and development. This chapter notabene provides only a brief overview of the pillars of AI, as even the detailed elaboration of just one sub-area of each of those would suffice to fill books. In the later chapters I will mostly focus on artificial neural networks, logic, knowledge representation, and speech in order to illustrate how I think the human thought processes can be rebuilt artificially. I will also give a brief introduction to one of my research fields, autonomously acting cars, which should help to understand what it takes to create intelligent, autonomous, social and adaptive agents; rebuilding collective and intelligent behavior in artificial systems not only allows us to understand a significant amount of brain evolution, but also how intelligence makes us a highly complex species in terms of thought processes.

Chapters 3, 4 and 5

In these chapters the fundamental concepts and standards of artificial neural networks are particularized under the consideration of actual knowledge and research conducted. The field of computational intelligence, to which artificial neural networks belong, comprises nature-inspired approaches for solving complex problem statements. Thus, this is where we will begin the search for paradigms that seem to be suitable for engineering artificial conscious entities.

In chapters 4 and 5, some sophisticated artificial neural network structures and learning approaches, some of which have been developed and published by the author, are particularized. One reason for explaining various algorithms is to show how far we can go with artificial neural networks and classical computers, how sophisticated learning algorithms can become, and how such structures can be manipulated. The more important reason for going into some detail with several algorithms is that I want to free the reader's mind, as for achieving brain-like capabilities we do not necessarily need to copy the inner workings of our brain but rely on other approaches – in this special case not the journey is its own reward. I

will, amongst others, discuss how artificial neural networks can be efficiently trained by simulating the survival behavior of bacteria in hazardous environments.

The major question to be answered in this chapter is whether these complex approaches are already powerful enough for generating conscious experiences in an artificial entity. Furthermore, some simple application scenarios for the introduced approaches are explained, creating a basic understanding of how such approaches may be applied practically.

Chapter 6

Deep artificial neural networks are, in some aspects, the most human-like artificial way of processing information we have thitherto. Very impressive and important achievements have been made due to complex network structures and sophisticated training algorithms, and what can already be achieved is not that far from the capabilities of what parts of the human brain, such as the visual cortex, can achieve. However until recently, the typical research in that area concerned problems that are perfectly suitable for being processed on a computer, like the prediction of numerical values or clustering of unstructured data. Furthermore, once trained to data of a sub-domain, an ANN can only be used in that sub-domain. What I want to say is that individual approaches are simply not potent enough for approaching the full stack of capabilities of a human brain. But a combination of several techniques may... With actual training algorithms it would be impossible to train a structure consisting of billions of neurons and thousands of neuron layers. This chapter serves the purpose to understand what ANNs are often used for, and for this we start with simple data analysis and evolve towards more complex information processing in the latter chapters.

Chapters 7 and 8

Quantum mechanics is of utmost importance within this elaboration, as, thanks to some of humankind's greatest scientists, within the last century the knowledge within this field has grown extensively and allows more than just an outlook on future developments in artificial intelligence. Some theories of consciousness make use of quantum physical phenomena within the brain for explaining awareness, or conscious experiences. Both chapters are used to explain the most important fundamentals of quantum mechanics and how they may find application within a biological brain. However, as the sole reduction of the brain by means of a quantum computer may not solve all problems, some objections against such hypotheses are raised, contributing to the groundwork for the main chapter of this elaboration.

Chapter 9

Before being able to reverse engineer a human mind, one has to understand some of the most commonly accepted philosophical views of what a mind actually is. Trying to implement algorithms without having thought about what the mind actually is, is impossible. Only philosophy can deliver some answers to this question and I, for my part, chose eliminative materialism to be the most suitable approach to pursue, when grappling with implementation scenarios for an aware, artificial entity. Some of the readers may wonder about the statement that I 'chose' one of many philosophical approaches as useful means for what achieving what is discussed in this book; this is because I encounter consciousness from an engineering point of view, and engineers usually solve problems by the means that are to their disposal. Thus, I consider philosophical views as a means of implementation only in the context of this book.

Chapter 10

The objective of writing this book was the introduction of paradigms and ideas that may serve the purpose of reverse engineering the mind in the near future. Amongst others, in this chapter self-organization in both human and artificial brains, which I consider to be another important aspect for the development of our mental abilities, is discussed. So are some more concepts such as search, information retrieval, or data representation, which form, together with what has been elucidated before, the foundation for giving an artificial conscious entity the ability to understand and interpret the world. Furthermore, I am of the opinion that the implementation of artificial neural networks on quantum systems represents one of the most valuable approaches for achieving the objectives of this research, as only such may deliver the computational power for processing what we understand to be our mind. Of course, everything is connected in order to form a big picture – the picture of how to reverse-engineer the mind.

Finally, the question on how we humans may benefit from such developments in the sense of transferring our minds into artificial vessels is addressed. How can the continuous consciousness-problem be bypassed, thus the required information in and structure of our brain be extracted instead of just being copied? There seems to be a solution for the dilemma in the sense of a distributed mind, which is discussed likewise.

1 Evolution's most extraordinary achievement

When doing research in the field of artificial intelligence, sooner or later one is required to deal with what I consider the most fascinating thing Nature has equipped us with – the human brain. This extraordinary organ does not only allow us to understand the universe, but additionally provides us with feelings, conscious perception or the ability to control our bodies in highest precision. I have always tried not to shelve myself into a specific field of research; however, above all I am a computer scientist and from a computer scientist's point of view it is, at first, interesting how the brain is capable of processing, storing or recalling information. Inevitably, when starting to deal with the matter this yields a lot more questions – questions that cannot be answered as easily for the brain as for hard- and software we are used to work with every day, even if one is equipped with knowledge about the field of artificial intelligence. Later in this work we will see that today's artificial intelligence is something that strongly differs from what we consider to be biological or human intelligence. Thus, I consider a rough explanation of the human brain's anatomy as well as the known and studied workings of its parts to be a good starting point.

1.1 Anatomy of the human brain

Most of the readers will remember the lessons in school where the human brain has been explained according to its evolutionary development: this will also be the way the matter will be dealt with in this chapter. Some of the following explanations have been inspired by the ones given by Bruce F. Katz⁵ and at dasgehirn.info,⁶ as the former perfectly well outline the functional principles of the human brain from an engineering point of view, and the latter from a medical point of view. From the abstractive Figure 1 - Human brain below we can see that the brain consists of four big lobes, the brain stem and the cerebellum, the latter looking like a ball of wool attached to the undersurface of the temporal lobe. However, there is a far more abstractive scheme describing the human brain, called triune brain model, which has been introduced by Paul Maclean (Figure 2 - Triune brain).

The triune brain model shows three differently colored areas, stacked according to their occurrence in the brain evolution. The brightest area in Figure 2 - Triune brain represents what is considered to be the most ancient and basic part of the nerve system, the reptilian brain, which has evolved about $3 * 10^8$ years ago. The reptilian brain is responsible for regulating autonomous functions like the heart rate or breathing. The medial part in the figure represents the limbic brain, or paleomammalian, which has evolved about $3 * 10^8$ years ago and which is partially responsible for recalling information, thus memorizing things, and emotions. The neocortex, or neomammalian, which makes up most of the visible part of the human brain (once it has been made visible in some way) has evolved only $1.2 * 10^5$ years ago, and merges not only with the structure of the paleomammalian, but also in its functionality.

5 Katz Bruce F. (2011): *Neuroengineering the future - Virtual minds and the creation of immortality*; Massachusetts: Infinity Science Press LLC, p. 16 ff.

6 dasgehirn.info: [2013-06-22]; dasgehirn.info; URL: <http://dasgehirn.info/>

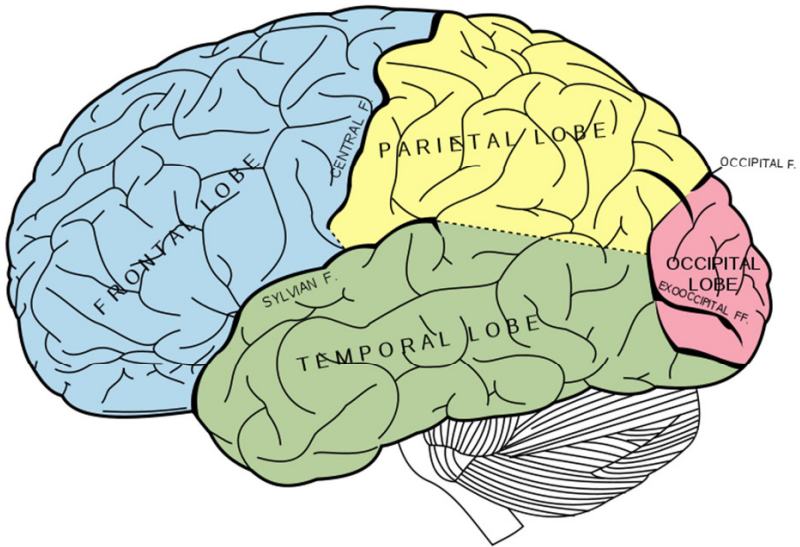


Figure 1 - Human brain⁷

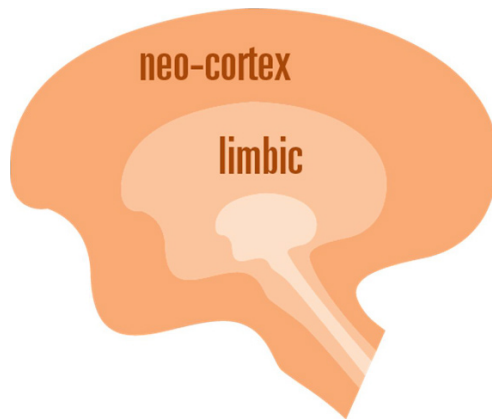


Figure 2 - Triune brain⁸

7 Wikipedia: Human brain [2013-06-19]; Wikipedia; URL: http://en.wikipedia.org/wiki/Human_brain

8 Steven White: Triune brain [2013-06-20]; Steven White; URL: <http://blog.stevenwhite.com/>

1.1.1 *Truncus cerebri*

The truncus cerebri (Figure 3 – Truncus cerebri), or brain stem, is the oldest part of the human brain and only as large as a finger, but nevertheless an essential part of our brain, as it regulates circulatory, breathing or sleeping. It consists of four major parts, whereby each of those may be subdivided into further areas. However, the description of the most basic parts does suffice for our purposes here.

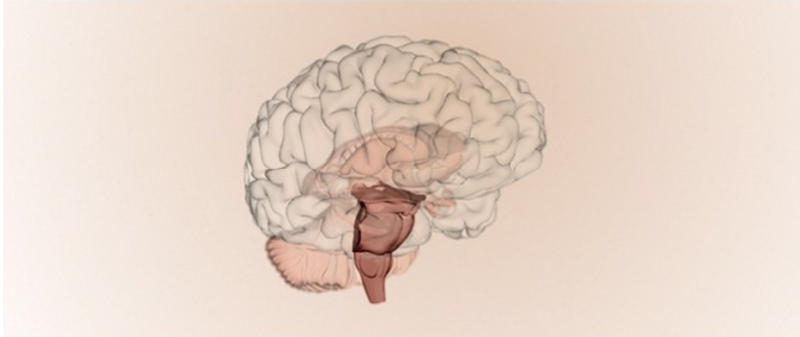


Figure 3 – Truncus cerebri⁹

1.1.1.1 Cerebellum

The cerebellum (Figure 4 - Cerebellum) is known to be responsible for coordination, equilibrium, motor movement and muscle tone. It consists of white matter in the inside and is surrounded by a very tightly folded outer layer of gray matter, called the cerebellar cortex. The number of neurons in the cerebellum makes up one half of all neurons in the brain (~500,000,000), which are used for relaying information between itself and the cerebral cortex' (which we must distinguish from the cerebellar cortex) areas involved in motor controls. An injury of the cerebellum results in disturbances of bodily movements, like a loss of coordination, impossibility of judging distances, the inability for performing fast movements, etc.

⁹ dasgehirn.info: Der Hirnstamm [2013-06-22]; dasgehirn.info; URL: <http://dasgehirn.info/entdecken/anatomic/der-hirnstamm/> (reprinted with permission from dasgehirn.info)