Igor Aizenberg

Complex-Valued Neural Networks with Multi-Valued Neurons

Studies in Computational Intelligence, Volume 353

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

Vol. 331. Athena Vakali and Lakhmi C. Jain (Eds.)
New Directions in Web Data Management 1, 2011
ISBN 978-3-642-17550-3

Vol. 332. Jianguo Zhang, Ling Shao, Lei Zhang, and Graeme A. Jones (Eds.)
Intelligent Video Event Analysis and Understanding, 2011
ISBN 978-3-642-17553-4

Vol. 333. Fedja Hadzic, Henry Tan, and Tharam S. Dillon
Mining of Data with Complex Structures, 2011
ISBN 978-3-642-17556-5

Vol. 334. Álvaro Herrero and Emilio Corchado (Eds.)
Mobile Hybrid Intrusion Detection, 2011
ISBN 978-3-642-18298-3

Vol. 335. Radomir S. Stankovic and Radomir S. Stankovic
From Boolean Logic to Switching Circuits and Automata, 2011
ISBN 978-3-642-11681-0

Vol. 336. Paolo Remagnino, Dorothy N. Monekosso, and Lakhmi C. Jain (Eds.)
Innovations in Defence Support Systems – 3, 2011
ISBN 978-3-642-18277-8

Vol. 337. Sheryl Brahnam and Lakhmi C. Jain (Eds.)
Advanced Computational Intelligence Paradigms in Healthcare 6, 2011
ISBN 978-3-642-17823-8

Vol. 338. Lakhmi C. Jain, Eugene V. Aidman, and Canicious Abeynayake (Eds.)
Innovations in Defence Support Systems – 2, 2011
ISBN 978-3-642-17763-7

Vol. 339. Halina Kwasnicka, Lakhmi C. Jain (Eds.)
Innovations in Intelligent Image Analysis, 2010
ISBN 978-3-642-17933-4

Vol. 340. Heinrich Hussmann, Gerrit Meixner, and Detlef Zuehlke (Eds.)
Model-Driven Development of Advanced User Interfaces, 2011
ISBN 978-3-642-14561-2

Vol. 341. Stéphane Doncieux, Nicolas Bredeche, and Jean-Baptiste Mouret(Eds.)
New Horizons in Evolutionary Robotics, 2011
ISBN 978-3-642-18271-6

Vol. 342. Federico Montesino Pouzols, Diego R. Lopez, and Angel Barriga Barros
Mining and Control of Network Traffic by Computational Intelligence, 2011
ISBN 978-3-642-18083-5

Vol. 343. Kurosh Madani, António Dourado Correia, Agostinho Rosa, and Joaquim Filipe (Eds.)
Computational Intelligence, 2011
ISBN 978-3-642-20205-6

Vol. 344. Atilla Elçi, Mamadou Tadiou Koné, and Mehmet A. Orgun (Eds.)
Semantic Agent Systems, 2011
ISBN 978-3-642-18307-2

Vol. 345. Shi Yu, Léon-Charles Tranchevent, Bart De Moor, and Yves Moreau
Kernel-based Data Fusion for Machine Learning, 2011
ISBN 978-3-642-19405-4

Vol. 346. Weisi Lin, Dacheng Tao, Janusz Kacprzyk, Zhu Li, Ebroul Izquierdo, and Haohong Wang (Eds.)
Multimedia Analysis, Processing and Communications, 2011
ISBN 978-3-642-19550-1

Vol. 347. Sven Helmer, Alexandra Poulovassilis, and Fatos Xhafa
Reasoning in Event-Based Distributed Systems, 2011
ISBN 978-3-642-19723-9

Vol. 348. Beniamino Murgante, Giuseppe Borruso, and Alessandra Lapucci (Eds.)
Geocomputation, Sustainability and Environmental Planning, 2011
ISBN 978-3-642-19732-1

Vol. 349. Vitor R. Carvalho
Modeling Intention in Email, 2011
ISBN 978-3-642-19955-4

Vol. 350. Thanasis Daradoumis, Santi Caballé, Angel A. Juan, and Fatos Xhafa (Eds.)
Technology-Enhanced Systems and Tools for Collaborative Learning Scaffolding, 2011
ISBN 978-3-642-19813-7

Vol. 351. Ngoc Thanh Nguyen, Bogdan Trawiński, and Jason J. Jung (Eds.)
New Challenges for Intelligent Information and Database Systems, 2011
ISBN 978-3-642-19952-3

Vol. 352. Nik Bessis and Fatos Xhafa (Eds.)
Next Generation Data Technologies for Collective Computational Intelligence, 2011
ISBN 978-3-642-20343-5

Vol. 353. Igor Aizenberg
Complex-Valued Neural Networks with Multi-Valued Neurons, 2011
ISBN 978-3-642-20352-7

Igor Aizenberg

Complex-Valued Neural Networks with Multi-Valued Neurons

Springer

Igor Aizenberg
Texas A&m University-Texarkana
7101 University Avenue,
Texarkana, TX, 75503
USA
E-mail: igor.aizenberg@tamut.edu

ISBN 978-3-642-20352-7 e-ISBN 978-3-642-20353-4

DOI 10.1007/978-3-642-20353-4

Studies in Computational Intelligence ISSN 1860-949X

Library of Congress Control Number: 2011926533

Typeset & *Cover Design:* Scientific Publishing Services Pvt. Ltd., Chennai, India.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

In memoriam of my father Naum Aizenberg,

founder of complex-valued neural networks

Acknowledgements

The author would like to express his appreciation and gratitude to many research colleagues from international scientific community who have inspired his thinking and research in the field of complex-valued neural networks.

First of all, the author expresses his great appreciation to the Editor of the Springer book series "Studies in Computational Intelligence" Professor Janusz Kacprzyk (Systems Research Institute of the Polish Academy of Sciences) who encouraged the author to write this book.

The author had a great pleasure and honor and was very happy to collaborate in his research on multi-valued neurons and neural networks based on them with great scientists and great personalities Professor Claudio Moraga (European Centre for Soft Computing, Mieres, Spain and Dortmund University of Technology, Dortmund, Germany), and Professor Jacek M. Zurada (University of Louisville, Kentucky, USA).

The author also appreciates a great opportunity to collaborate with other great scientists Professor Jaakko Astola (Tampere University of Technology, Tampere, Finland) and Professor Joos Vandewalle (Catholic University of Leuven – KU Leuven, Leuven, Belgium). The author also kindly appreciates those great opportunities he had working in KU Leuven (1996-1998) on the invitation of Professor Joos Vandewalle, Tampere University of Technology (1997, 2002, 2005-2006) on the invitation of Professor Jaakko Astola, and Dortmund University of Technology (2003-2005) on the invitation of Professors Claudio Moraga and Brend Reusch.

The author sincerely appreciates great assistance he got from his former and current students and junior colleagues. Particularly, the author would like to express his special thanks to his former students, co-authors in a number of publications, and great programmers whose wonderful contributions to software simulation of complex-valued neural networks with multi-valued neurons is of crucial importance. They are Dr. Dmitriy Paliy (Nokia Research, Helsinki, Finland), Taras Bregin (Jabil Ukraine Ltd, Uzhgorod, Ukraine), and Dr. Constantine Butakoff (University of Pompeu Fabra, Barcelona, Spain).

The author expresses his appreciation and gratitude to his family, to his wife Ella, to his children Zhenia and Alena, and to his mother Svetlana for their continuous support. The author also wants to say many special thanks to his son Zhenia for his kind assistance in preparing some illustrations for the book and in proofing the manuscript.

Some results presented in this book (Sections 3.3, 4.3, 5.2, and 5.3) are based upon work supported by the National Science Foundation under Grant No. 0925080.

Preface

The use of complex numbers in neural networks is as natural as their use in other engineering areas and in mathematics.

The history of complex numbers shows that although it took a long time for them to be accepted (almost 300 years from the first reference to "imaginary numbers" by Girolamo Cardano in 1545[1] to Leonard Euler's[2] and Carl Friedrich Gauss'[3] works published in 1748 and 1831, respectively), they have become an integral part of mathematics and engineering. It is difficult to imagine today how signal processing, aerodynamics, hydrodynamics, energy science, quantum mechanics, circuit analysis, and many other areas of engineering and science could develop without complex numbers. It is a fundamental mathematical fact that complex numbers are a necessary and absolutely natural part of numerical world. Their necessity clearly follows from the Fundamental Theorem of Algebra, which states that every non-constant single-variable polynomial of degree *n* with complex coefficients has exactly *n* complex roots, if each root is counted up to its multiplicity.

Answering a question frequently asked by some "conservative" researches, what one can get using complex-valued neural networks (typical objections are: they have "twice more" parameters, require more computations, etc.), we may say that one may get the same as using the Fourier transform, but not just the Walsh transform in signal processing. There are many engineering problems in the modern world where complex-valued signals and functions of complex variables are involved and where they are unavoidable. Thus, to employ neural networks for their analysis, approximation, etc., the use of complex-valued neural networks is natural. However, even in the analysis of real-valued signals (for example, images or audio signals) one of the most frequently used approaches is frequency domain analysis, which immediately leads us to the complex domain. In fact, analyzing signal properties in the frequency domain, we see that each signal is characterized

[1] G. Cardano's work "Arts Magna" ("Great Art or on Algebraic Rules" was published in 1545. For the first time he introduced a notion of "imaginary numbers", however he considered these numbers useless.

[2] L. Euler proved and published in 1744 the relationship between the trigonometric functions and complex exponential function ($e^{i\varphi} = \cos\varphi + i\sin\varphi$). He also suggested to use symbol *i* for an imaginary unity (the first letter of Latin word *imaginarius*).

[3] C.-F. Gauss gave to complex numbers their commonly used name "complex" and comprehensively described them in his "Memoir to the Royal Society of Göttingen" in 1831. Gauss has also obtained the first mathematically exact proof of algebraic closure of the field of complex numbers in 1799 (this fact was first hypothetically formulated by J. d'Alembert in 1747 and L. Euler in 1751).

by magnitude and phase that carry different information about the signal. To use this information properly, the most appropriate solution is movement to the complex domain because there is no other way to treat properly the phase information. Hence, *one of the most important characteristics of Complex-Valued Neural Networks is the proper treatment of the phase information*. It is important to mention that this phenomenon is important not only in engineering, but also in simulation of biological neurons. In fact, biological neurons when firing generate sequences of spikes (spike trains). The information transmitted by biological neurons to each other is encoded by the frequency of the corresponding spikes while their magnitude is a constant. Since, it is well known that the frequency can be easily transformed to the phase and vice versa, then it should be natural to simulate these processes using a complex-valued neuron.

Complex-Valued Neural Networks (CVNN) is a rapidly growing area. There are different specific types of complex-valued neurons and complex-valued activation functions. But it is important to mention that all Complex-Valued Neurons and Complex-Valued Neural Networks have a couple of very important advantages over their real-valued counterparts. The first one is that they have *much higher functionality*. The second one is their *better plasticity* and *flexibility*: they learn faster and generalize better. The higher functionality means first of all *the ability of a single neuron to learn those input/output mappings that are non-linearly separable in the real domain*. This means the ability to learn them in the initial space without creating higher degree inputs and without moving to the higher dimensional space, respectively. As it will be shown below, such classical non-linearly separable problems as XOR and Parity *n* are about the simplest that can be learned by a single complex-valued neuron.

It is important to mention that *the first historically known complex-valued activation function was proposed in 1971* (!) by Naum Aizenberg and his co-authors[4]. It was 40 years ago, before the invention of backpropagation by Paul Werbos (1974), before its re-invention and development of the feedforward neural network by David Rumelhart (1986), before the Hopfield neural network was proposed by John Hopfield in 1982. Unfortunately, published only in Russian (although in the most prestigious journal of the former Soviet Union), this seminal paper by Naum Aizenberg and his colleagues and a series of their subsequent publications were not available to the international research community for many years. A problem was that in the former Soviet Union it was strictly prohibited to submit scientific materials abroad and therefore there was no way for Soviet scientists to publish their results in international journals. May be, being wider known, those seminal ideas on complex-valued neurons could stimulate other colleagues to join research in this area much earlier than it really happened (only in 1990s and 2000s). May be, this could help, for example, to widely use neural networks for solving not only binary, but multi-class classification problems as far back as more than 30 years ago… However, the history is as it is, we cannot go to the past and change something there. Let us better concentrate on what we have today.

[4] N.N. Aizenberg, Yu. L. Ivaskiv, and D.A. Pospelov, "About one generalization of the threshold function" *Doklady Akademii Nauk SSSR* (*The Reports of the Academy of Sciences of the USSR*), vol. 196, No 6, 1971, pp. 1287-1290 (in Russian).

So what is this book about? First of all, it is not an overview of all known CVNNs. It is devoted to comprehensive observation of one representative of the complex-valued neurons family – the Multi-Valued Neuron (MVN) (and its variation – the Universal Binary Neuron (UBN)) and MVN-based neural networks. The Multi-Valued Neuron operates with complex-valued weights. Its inputs and output are located on the unit circle and therefore its activation function is a function only of argument (phase) of the weighted sum. It does not depend on the weighted sum magnitude. MVN has important advantages over other neurons: its functionality is higher and its learning is simpler because it is *derivative-free* and it is based on the error-correction rule. A single MVN with a periodic activation function can easily learn those input/output mappings that are non-linearly separable in the real domain (of course, including the most popular examples of them, XOR and Parity *n*). These advantages of MVN become even more important when this neuron is used as a basic one in a feedforward neural network. The Multilayer Neural Network based on Multi-Valued Neurons (MLMVN) is an MVN-based feedforward neural network. Its original backpropagation learning algorithm significantly differs from the one for a multilayer feedforward neural network (MLF)[5]. It is derivative-free and it is based on the error-correction rule as it is for a single MVN. MLMVN significantly outperforms many other techniques (including MLF and many kernel-based and neuro-fuzzy techniques) in terms of learning speed, network complexity and generalization capability.

However, when the reader starts reading this book or when the reader even consider whether to read it, it is also important to understand that this book is not the 2nd edition of the first monograph devoted to multi-valued neurons[6]. Since that monograph was published 11 years ago, and many new results were obtained during this time by the author of this book, his collaborators and other researchers, this book, on the one hand, contains a comprehensive observation of the latest accomplishments and, on the other hand, it also deeply observes a theoretical background behind MVN. This observation is based on the today's view on the MVN place and role in neural networks. It is important that today's understanding is much deeper and comprehensive than it was when the first book was published. Thus, the overlap of this book with the first one is minimal and it is reduced to some basic necessarily definitions, which is just about 5-6% of the content. The most significant part of the book is based on the results obtained by the author independently and in co-authorship with other colleagues and his students. However, contributions made by other research colleagues to MVN-based neural networks are also observed.

---

[5] Often this network based on sigmoidal neurons is also referred to as the multilayer perceptron (MLP) or a "standard backpropagation network". We will use a term MLF throughout this book reserving a term "perceptron" for its initial assignment given by Frank Rosenblatt in his seminal paper F. Rosenblatt, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, Cornell Aeronautical Laboratory", *Psychological Review*, v65, No. 6, 1958 pp. 386-408.

[6] I. Aizenberg, N. Aizenberg, and J. Vandewalle, *Multi-Valued and Universal Binary Neurons: Theory, Learning, Applications*, Kluwer Academic Publishers, Boston/Dordrecht/London, 2000.

This book is addressed to all people who work in the fascinating field of neural networks. The author believes that it can be especially interesting for those who use neural networks for solving challenging multi-class classification and prediction problems and for those who develop new fundamental theoretical solutions in neural networks. It should be very suitable for Ph.D. and graduate students pursuing their degrees in computational intelligence. It should also be very helpful for those readers who want to extend their view on the whole area of computational intelligence.

The reader is not expected to have some special knowledge to read the book. All readers with basic knowledge of algebra and calculus, and just very basic knowledge of neural networks (or even without having special knowledge in neural networks area) including students can easily understand it.

We avoid using here too deep mathematical considerations (may be except proofs of convergence of the learning algorithms and analysis of that specific separation of an n-dimensional space, which is determined by the MVN activation function). However, those readers who do not want or do not need to go to those mathematical details may skip over the corresponding proofs.

In this book, we cover the MVN and MVN-based neural networks theory and consider many of their applications. The most important topics related to multi-valued neurons are covered. Chapter 1 should help the reader to understand why Complex-Valued Neural Networks were introduced. It presents a brief observation of neurons, neural networks, and learning techniques. Since the functionality of real-valued neurons and neural networks is limited, it is natural to consider the complex-valued ones whose functionality is much higher. We also observe in Chapter 1 CVNNs presenting the state of the art in this area. In Chapter 2, the multi-valued neuron is considered in detail along with the basic fundamentals of multiple-valued logic over the field of complex numbers, which is a main theoretical background behind MVN. In Chapter 3, MVN learning algorithms are presented. Chapter 4 is devoted to the multi-valued neural network based on multi-valued neurons, its original derivative-free backpropagation learning algorithm and its applications. In Chapter 5, MVN with a periodic activation function and its binary version, the universal binary neuron, are presented, and it is shown how it is possible to solve non-linearly separable problems using a single neuron without the extension of the feature space. In Chapter 6, some other applications of MVN are considered (solving classification and prediction problems, associative memories). The book is illustrated by many examples of applications.

The author sincerely hopes that this book will provide its readers with new interesting knowledge and will encourage many of them to use MVN and MVN-based neural networks for solving new challenging applied problems. The author will be glad and consider his work successful if more researches will be involved through this book in the really magic world of neural networks and particularly its complex-valued part.

The author also hopes that by this book he can pay a tribute to the founder of Complex-Valued Neural Networks, his teacher, colleague, father and a great personality Naum Aizenberg.

Contents