Chapter 2
Are Memristors the Future of AI?
A Review of Recent Progress and Future Perspectives

Robert Kozma, Robinson E. Pino, and Giovanni E. Pazienza

Abstract We review the state-of-the-art of neuromorphic memristor science and technology. We cover principles of memristors and neuromorphic systems, computational models of memristors, and hardware implementations. Potential applications of memristors are also described, including supercomputing, image processing, computer vision, intelligent control, and robotics. This review is based on the chapters of the present volume, which extend the materials of the invited and plenary talks given at the series of events on memristors in 2011. We elaborate on challenges and future perspectives of this promising new research field.

2.1 Introduction

In 2008 scientists from Hewlett-Packard discovered a nano-scale device called the memristor, a hypothetical circuit element predicted in 1971 by Leon Chua, UC Berkeley [1–3]. This has generated unprecedented worldwide interests because, among many applications, memristors can be used as super-dense non-volatile memories for building instant turn-on computers. There are suggestions from many researchers that memristor-based analog memory can be used to build brain-like learning machines with nano-scale memristive synapses. In the past years, the field of memristors has expanded rapidly, with numerous publications and meeting in this area. The present volume is the result of such recent activities [4–5]. Memristor technology has the potential to revolutionize computing and scientific research in the coming decades. There are, however, significant challenges, which may hamper broad proliferation of memristors, and which may ultimately prevent the full realization of its extraordinary potentials. We can learn from past mistakes in scientific discoveries to mitigate the consequences of inevitable pitfalls along the way. This volume can help to establish a roadmap towards materializing the full potential of memristors.

R. Kozma (✉) · G. E. Pazienza
FedEx Institute of Technology, Department of Mathematical Sciences, University of Memphis, Memphis, TN, USA
e-mail: rkozma@memphis.edu

R. E. Pino
Advanced Computing, Air Force Research Laboratory, Rome, NY, USA

R. Kozma et al. (eds.), *Advances in Neuromorphic Memristor Science and Applications,* Springer Series in Cognitive and Neural Systems,
DOI 10.1007/978-94-007-4491-2_2,

There is a need to systematically elaborate the theoretical foundations of memristors, following Chua's visionary work. The invention of the first materials with memristive properties using TiO_2 compounds was quickly followed by the identification of more and more additional materials which exhibit memristive properties and can serve as basis of memristive devices. Clearly, some specific features of various memristive compounds may be beneficial in some context and disadvantageous in other cases. Thus there is a need to conduct significant materials science studies to identify optimal memristors for various technical requirements. Then there is the issue of constructing electrical circuitries using memristive components. The pinched hysteresis feature underlying the operation of memristors indicate the potential of using memristors in a continuous operational mode as part of the analogous computational paradigm. However, the presently dominant digital computers may in fact imply that all efforts should be concentrated on digital manifestations of memristors, in order to reap the benefits of the digital computing paradigm to the fullest extent. As far as applications are concerned, memristors are likely to be implemented in the near future as powerful flash memories. Another application area is the design of electronic circuits with inherent learning capabilities, when data processing and memory functions are not separated; rather those functions are completed on a unified hardware device using memristors.

This leads us to the concept of neuromorphic hardware. Brain tissues are generally assumed to be complex networks of neurons, which change their functional properties as they process sensory data in various cortical areas. This process is called learning and the learned knowledge is stored in the cortex through modified synaptic connectivity. There are various models of learning in the neural tissues, including Hebbian correlational learning, reinforcement, habituation, homeostatis, spike time dependent plasticity, and others. It is conceivable that all these learning effects are in fact manifestations of memristive behavior in the neural tissue. The present volume emphasizes the potential of memristors in neuromorphic designs.

The volume starts with Snyder's thought-provoking review on the potential of memristive nanodevices for building low-power, intelligent machines. He suggests using memristors as discrete memory elements in a digital platform, rather than analog implementations. He contends that this is a technically less risky and economically far more viable path to achieving adaptive machine intelligence. To illustrate his point, he introduces two pairs of paradigms, which he coins the *digital versus analog* and *algebraversus analysis* complementary pairs. He asks the questions: Can intelligence be implemented using digital electronics, or is it necessary to replicate the analog processes found in biological brains? Are nonlinear differential equations the preferred (or necessary) mathematical foundation for building intelligence, or is another mathematical approach possible? He puts forward convincing arguments for answering the first question affirmatively, i.e., machine intelligence is possible and in fact preferable to be implemented on digital hardware. As for the second question, he argues that mathematical description of neural activity using differential equations cannot be a requirement of modeling intelligent behavior, and alternative approaches using matrix algebra and other methods are fully justified, if not preferred.

Part I of the book outlines fundamental concepts of memristors and neuromorphic systems. Di Ventra and Pershin focuses in Chap. 3 on the broad category of memory circuit elements which includes, besides the memristor, the memcapacitor and the meminductor. Their study shows that simple memristive, memcapacitive, and meminductive systems can model a variety of biological processes such as the adaptive behavior of unicellular organisms. Furthermore, it is discussed how networks of memory circuit can be used for solving efficiently graph theory optimization problems and it is proved that a network of memristive devices solves the maze problem faster than any existing algorithm. This example is just an illustration of the general concept; however, it paves the way for numerous practical applications. In view of the recent advances of nanotechnology, it is now feasible to assemble ultra-dense networks of memory circuit elements. As shown in this work, they can be used to create a novel generation of 'smart' electronic circuits whose potentialities are yet to be explored. In Chap. 4, Ames et al. sets forth a systematic analysis of the relevant literature on neuromorphic hardware. They include a thorough discussion concerning the key features, which they consider crucial in future *intelligent* machine embodiments. Based upon this discussion, the authors identify the need of a powerful software platform serving as the *glue* between the neural model and the hardware on which it is implemented. This is the main motivation for a software framework called *Cog* whose purpose is to abstract away the details of the underlying hardware to the software designers and hence allow the seamless integration of new hardware. Currently, *Cog* runs on single core CPU s and GPUs (or a cluster of both processors), but it is specifically designed to reduce communication energy in neuromorphic computing by leveraging the introduction of dense memristive memories close to computing cores. *Cog* has been used in a complex system called MoNETA (Modular Neural Exploring Traveling Agent) that generates complex behaviors based on functionalities that include perception, motivation, decision-making, and navigation. This architecture has been tested in a virtual environment simulating intelligent and autonomous behaviors in robotics applications.

Part I concludes with Chap. 5 by Werbos which explores the question how can memristors be more than just memory devices. Werbos provides an excellent summary of the milestones of neuroengineering and well as cognitive optimization and prediction, which would lead to the emergence of intelligent behavior, ultimately at the level produced by human brains. The author, who has been a key player in these areas since their inception, describes the main events—especially those occurred within the framework of the National Science Foundation—that have led to the better understanding of the mechanisms according to which the brains learn. In the near future, the memristor is expected to allow us to achieve a density of functional features (or devices) on-board chips that is far greater than what we have so far. The author contends that brain-like principles should be exploited on the soon to be available memristive hardware platforms. The author identifies several key applications domains for the new memristor systems; namely, complex large-scale systems such as the management of complex infrastructures over time, and the last generations of electric grids. The chapter poses numerous open questions which will surely encourage the reader to further explore these topics.

Part II describes computational models of memristors and various modeling challenges. In Chap. 6, Pino describes a basic self-reconfigurable neuromorphic computing architecture leveraging the non-linear memristive behavior of ion-conductor chalcogenide-based memristor devices. The chapter begins by describing what can be described as the simplest possible basic building block for a memristor-based neuromorphic computing architecture. Such basic building block seems representative of a self-contained unit-cell that exhibits the property of self-reconfiguration by allowing the user to train the circuit via perception and stimulation with electronic pulses or spikes. Then, the author describes a methodology for the configuration of self-reconfigurable neuromorphic distributed networks. In Chap. 7, Kudithipudi and Merkel present a CMOS/memristor hybrid technology, where memristive devices are integrated within a 3D memristive crossbar architecture. The chapter discusses different models which can be used for implementing memristors as memory, sensing, logic, and neuromorphic units. The authors content that a prime advantage of such computing technology is its ability to offer tera-bit densities with ultra-low power and long data retention times. The main idea is that such distributed computational memristor fabrics can dynamically transform over time to perform heterogeneous computing, with the technological objective of superseding classical CMOS architectures. In Chap. 8, Li and Pino discuss TiO_2 and spintronic memristor-based synaptic designs with a training scheme and explore implications of statistical device variation. The authors discuss the impacts of geometry variation on the electrical properties of these two different types of memristors by analytical modeling and Monte-Carlo simulations. A simple algorithm, based on the latest characterization method of LER (line edge roughness) and thickness fluctuations, is proposed to generate a large volume of geometry variation-aware three-dimensional device structures for Monte-Carlo simulations, and a process-variation aware device model is proposed. These results make it possible for scientists and engineers to map virtual neural networks to physical hardware designs with the corresponding training circuit to mimic a biological system.

In Chap. 9, Verzace et al. analyzes power requirements of arithmetic computational processes implemented on memristive devices. Storing and updating synaptic weights based on synaptic plasticity rules is a computationally very demanding operation in biologically-inspired neural networks using basic operations of addition and multiplication. Memristive hardware holds the promise of greatly reduced power requirements by increasing synaptic memory storage capacity and decreasing wiring length between memory storage and computational modules. The chapter reviews power requirements of various computational algorithms and introduces a novel computational tool based on fuzzy inference for adaptive resonance theory (ART) networks. Fuzzy inference significantly reduces the computational complexity of the memristive hardware; while it is able to learn synaptic weights with the required precisions. In conclusion, fuzzy arithmetic completes the required classification tasks correctly and more efficiently than other approaches. Chapter 10 by Suri and Desalvo reviews basic concepts related to Phase Change Memory (PCM) technology and its potential connection to memristors. They describe a hybrid CMOS/memristive system, in which a PCM device is sandwiched between two spiking CMOS circuits, emulating the synapse connecting a pre- and a post- synaptic neuron. They present basic principles of PCM devices and the current state of the technology. The use of PCM devices for synapses has the potential advantages of high scalability, CMOS compatibility, low programming-current, strong endurance and technological maturity. They demonstrate the use of PCM devices for emulating specific functions of a biological synapse, such as synaptic potentiation, synaptic depression and spike-time dependent plasticity (STDP). They discuss recent research encompassing the study of PCM devices and chalcogenide materials for neuromorphic applications, which are important for realizing PCM-based large-scale neuromorphic systems.

Part III describes various hardware embodiments with memristive properties and potential application areas. Kang and Shin reviews in Chap. 11 recent technology trends in memristive analog and digital electronics, with particular attention to programmable analog and digital circuits, ultra-dense nonvolatile resistive memory architectures, and zero leakage nonvolatile logic gates. A reconfigurable nonvolatile computing platform that harnesses memristor properties is used to deploy massive local nonvolatile memories and advance computing capabilities with much lowered energy consumption compared to conventional charge-based VLSI systems. They describe applications of memristive devices for nonvolatile memories, programmable interconnects, logic gates, and nonvolatile latches with high integration density and CMOS compatibility. They point out that combining memristors with the prevailing CMOS technology would lead to the extension of Moore's Law beyond the hitherto observed technological limitations. Chapter 12 by Taha presents a review of existing memristor modeling techniques and provides simulations that compare several existing models with published memristor characterization data. A discussion of existing models is presented that explains how the equations of each relate to physical device behaviors. The simulations were completed in SPICE and compare the output of the different models to current–voltage relationships of physical devices. Sinusoidal and triangular pulse inputs were used throughout the simulations to test the capabilities of each model. The chapter is concluded by recommending a generalized memristor model that can be accurately matched to several published device characterizations. This generalized model has the potential for more accurate circuit simulation for a wide range of device structures and voltage inputs.

Chapter 13 by Corinto et al. presents a rigorous mathematical study concerning the dynamics of different memristor models with a special emphasis on the influence of initial and boundary conditions of the system. This is the first thorough systematic work in this area and it may have a deep impact on the future literature. The analytical results linking the initial condition of a memristor with its current-voltage characteristic can be used to devise a novel pattern recognition system based on the synchronization of nonlinear dynamical systems. As suggested by the authors, such a pattern recognition system may be used to interpret complex neurophysiological phenomena, such as the binding problem. This contribution can be viewed as a cornerstone of a new class of computing machines, the so-called *Memristor Oscillatory Neurocomputers*, combining memristor synapse circuitry with oscillatory neurocomputers and performing spatial-temporal pattern recognition. The breakthrough provided by such systems may affect numerous disciplines, including intelligent adaptive control and intelligent user interfaces.

Chapter 14 by Pino and Moore explores the implementation of neurophysiological and psychological constructs to develop a hyper-parallel computing platform, termed a neuromorphic computing. The authors describe a model of the primary visual cortex (V1) and simulation results from a high performance computing facility. Their columnar V1 model includes binocular disparity and motion perception, and the V2 model thick and pale stripes were added to produce a V1/V2 stereomotion forming a perception system. Both the V1 and V2 models were based upon structures approximating neocortical minicolumns and functional columns. The neuromorphic strategies included columnar organization, integrate-and-fire neurons, temporal coding, point attraction recurrent networks, Reichardt detectors and "confabulation" networks. The authors seek to find the underlying architecture of hyper-parallel machines, and understand computational methods akin to how a brain deals with sensation, perception, memory, attention, and decision-making. Chapter 15 by Nawrocki et al. is a systematic analysis of Organic Bistable Devices (OBDs) which offer several potential advantages also for memristors. The potential advantages include lower cost and simpler fabrication as compared with their inorganic counterparts. OBDs are now ready for commercial implementations. The chapter provides numerous illustrations to guide the readers throughout the description of this vibrant technology, which is soon expected to find application in various revolutionary neuromorphic memristive systems. The authors mention a few active leading projects in intelligent systems and robotics, in the US and Europe, in which their technology can be particularly successful.

We are pleased to present this volume and hope it will facilitate further progress in the rapidly developing field of memristors.

Acknowledgments R.K. acknowledges partial support by AFOSR Lab Task by the Mathematics and Cognition Program (Dr. J. Myung, PM), and by DARPA Physical Intelligence Program through HRL subcontract (Dr. N. Srinivasa PM).

References

1. Chua LO (1971) Memristor-the missing circuit element. IEEE Trans Circuit Theory 18:507–519
2. Snider G (2007) Self-organized computation with unreliable, memristive nanodevices. Nanotechnology, IOP Press 18(36):365202
3. Strukov DB et al (2008) The missing memristor found. Nature 453:80–83
4. Are memristors the future of AI? Panel session organized by Kozma R, Pazienza G at IEEE/INNS International joint conference on neural networks IJCNN2011, July 30–August 5, 2011, San Jose, CA
5. Future perspectives of neuromorphic memristor science and technology, Workshop organized by Kozma R, Pino R, Pazienza G at IEEE/INNS International joint conference on neural networks IJCNN2011, July 30–August 5, 2011, San Jose, CA

http://www.springer.com/978-94-007-4490-5

Advances in Neuromorphic Memristor Science and Applications
Kozma, R.; Pino, R.E.; Pazienza, G.E. (Eds.)
2012, VIII, 320 p., Hardcover
ISBN: 978-94-007-4490-5