Emergent Criticality in Complex Turing B-Type Atomic Switch Networks

Adam Z. Stieg,* Audrius V. Avizienis, Henry O. Sillin, Cristina Martin-Olmos, Masakazu Aono, and James K. Gimzewski

Recent advances in the neuromorphic operation of atomic switches as individual synapse-like devices demonstrate the ability to process information with both short-term and long-term memorization in a single two terminal junction. Here it is shown that atomic switches can be self-assembled within a highly interconnected network of silver nanowires similar in structure to Turing’s “B-Type unorganized machine”, originally proposed as a randomly connected network of NAND logic gates. In these experimental embodiments, complex networks of coupled atomic switches exhibit emergent criticality similar in nature to previously reported electrical activity of biological brains and neuron assemblies. Rapid fluctuations in electrical conductance display metastability and power law scaling of temporal correlation lengths that are attributed to dynamic reorganization of the interconnected electro-ionic network resulting from induced non-equilibrium thermodynamic instabilities. These collective properties indicate a potential utility for real-time, multi-input processing of distributed sensory data through reservoir computing. We propose these highly coupled, nonlinear electronic networks as an implementable hardware-based platform toward the creation of physically intelligent machines.

1. Introduction

Modern state-of-the-art computers are the product of over half a century spent refining implementations of Turing’s automatic machine (TAM)[1] using Von Neumann’s computational architecture.[2] The TAM is the principal theoretical framework for computation using sequential logical operations on single-purpose hardware consisting of an infinite tape of symbols, a read/write head, and a control mechanism that acts based on a transition table or instruction sheet. Von Neumann’s introduction of the concept of memory into the computer architecture provided a blueprint for the physical realization of multifunctional TAM machines that utilize multiple stored programs via two main functional units – processors and memory. This flexible control mechanism made the TAM truly universal in its capacity to complete any algorithmically defined task.

The von Neumann architecture has the principle advantage of clarity from the engineering perspective. Reduction in the physical size and increased areal density of electronic components directly scales up performance in terms of increased bytes of storage and processor cycles per second. The extension of this trend toward biologically inspired or artificially intelligent computation has resulted in attempts to simulate every neuron in the mammalian cortex and to outperform human experts in games of strategy.[3,4] These achievements, while impressive, are not readily scalable due to the basic constraints of the CMOS architecture, its associated methods of fabrication, and the limits of its operational mechanism.[5] Further, the requisite passage of program instructions and data between processor and memory has evolved as a speed-limiting step known as the “von Neumann bottleneck” (vNB)[6] (Figure 1a), which results in idle processor cycles and power dissipation as information is simply being transferred, not processed. In combination, these factors generate a computational architecture that consumes orders of magnitude more space and energy than intelligent biological systems.

While current state of the art approaches to computation represent tremendous progress in performance and efficiency versus their historical counterparts, computer scientists have drawn inspiration from biology in an effort to develop computational strategies that are able to match the capabilities of biological neural networks (BNN). Remarkably, such concepts were proposed over sixty years ago as Turing’s “B-Type unorganized machine (TBTu)”,[7] and have been subsequently popularized by Rosenblatt’s perceptron, recurrent neural networks, and reservoir computing.[8–15] These bio-inspired designs are generally associated with the notion of “connectionism”. Connectionist

Theories are based on complex networks composed of simple units, which, as a whole, produce emergent behavior not found or associated with any particular unit. What constitutes a “complex system” is difficult to define precisely. However, extensive studies of complex, real-world networks have revealed the importance of both structural topology and internal dynamics. Various models of connectivity and interaction have been utilized to construct a densely interconnected network that connects between electronic components in a cost-efficient manner, especially in designs utilizing unconventional topologies. Recent advances in nanoscale science and technology have enabled the direct self-assembly and integration of functional circuit elements within the wiring scheme of nanoscale devices with the unique architectures. Here, we utilize these concepts to construct a densely interconnected network of synapse-like memristive atomic switches using bottom up self-assembly. We find that this system demonstrates some of the emergent behaviors commonly observed in biological neural networks. These complex atomic switch networks provide as a promising new direction for the development of functional TBTu-inspired neuromorphic computing devices, with specific implications toward physically implementable reservoir computation.

2. Computational Models

Building upon decades of inspired research based in the TAM/von Neumann computational paradigm, modern processors routinely include multiple cores and large memory caches to maximize efficiency by parallelizing computations and reducing memory access times. In addition to physical limitations on component size and the vNB, leakage currents through gate dielectrics, programming challenges in parallel processing, and intolerance to faulty elements have begun to impact performance. These obstacles provide strong motivation to develop and implement alternative computational strategies. To this end, numerous theories and proposals have been put forth toward biologically inspired, neuromorphic computing devices.

Biological neural networks utilize self-configuring, hardware-based architectures capable of dynamic topological alteration and function without the need for pre-programming or an underlying software algorithm. These intrinsically nonlinear, complex systems demonstrate extraordinarily efficient transmission of information and emergent behaviors commonly associated with intelligence such as associative memory, learning, and predictive capacity in non-deterministic environments. One related theoretical construct, the TBTu, was conceived of as a randomly interconnected network of nothing more than modifiable NAND logic gates. Since NAND gates may be combined to perform any other logic function, Turing hypothesized that a sufficiently large network could serve as a usable computer, capable of any TAM operation. Moreover, he showed that its connections and operations could be trained over time to alter its behavior, in a similar fashion to that of a biological brain.

This concept has been applied in the fields of systems neuroscience and artificial intelligence to form the basis of contemporary research into artificial neural networks (ANN). These ANNs are typically implemented as software running on conventional TAM systems, mimicking information processing in natural systems. The earliest ANNs, commonly known as the “perceptron”, utilized a feed-forward design in which artificial neurons are connected by modifiable synaptic weights and can ‘learn’ to map input-output relationships according to any (mathematical) function. The development of recurrent neural networks (RNN) enabled the inclusion of adaptive capacities through feedback
strategies.[9] The existence of cyclical connections makes the RNN a dynamical system, capable of sustaining internal activity in the absence of additional signals, not merely mapping input to output. However, basic RNN training strategies still involve the direct modification of internal synaptic weights implemented abstractly using algorithms inspired by biological neural networks. In addition, ANNs are generally designed and optimized to perform specific computational tasks, occasionally utilizing purpose-built hardware for increased functionality.[8]

This enhanced performance comes at the expense of flexibility, adaptability, and the capacity to synthesize multiple time-varying input signals or to operate in a non-deterministic fashion—all hallmarks of biological neural systems.

Reservoir computation (RC) is a promising extension of RNNs towards more accurately modeling biological neural networks that has been successfully implemented in various engineering applications.[12–15] Instead of tracking and modifying individual synaptic weights, the complex network of artificial neurons is treated as a kind of “black box” which is dynamically modified by the input and retains some memory of previous input signals. The complex network reservoir (CNR) acts to map these lower-dimensional input signals into a higher-dimensional space, represented by patterns in the state of the system and contains temporal information through integration of the input history. Poised between simply periodic and wildly unpredictable oscillations, the CNR operates at the edge of chaos.[16]

This approach overcomes the challenge of training individual synaptic weights inside RNNs by not explicitly modifying them at all. Instead, a separate readout/output function is trained to examine the response of the reservoir, interpreting the spatio-temporal patterns formed by the collective effect of the input signals and transforming this higher-dimensional information into the desired output. Through appropriate training, RC methods are capable of simulating any Turing-type computational machine. Since the reservoir functions autonomously, multiple linear readout functions can be used simultaneously, thereby allowing the system to carry out multiple computational tasks on the same input stream in real time.[12,11]

While simulation and modeling efforts implemented on traditional computational architectures remain the general, near-term focus of reservoir approaches, calls for the development of hardware-based CNR systems continue to form the basis for inquiry into a new paradigm of computational methods. Achieving these goals requires the development of physical systems whose properties mimic those of artificial, simulated reservoirs as well as a means to harness the power of information-rich output patterns they generate. We propose that the former can be achieved by applying the concept of Turing’s connectionist networks to the fabrication of complex device architectures consisting of highly interconnected, nonlinear electronic elements. A near-infinite set of internal system states capable of receiving/storing information from parallel input streams is necessary to combine complex, dynamic signals into a single, higher-dimensional output.

This property is characteristic of systems operating in a critical state, a hallmark of complex networks of nonlinear elements, where the divergence of the system correlation length in both space and time provides all these requisite characteristics.[16,30]

3. Complex Device Architectures

The structure and activity of the biological brain is intrinsically complex, comprised of billions of neurons interacting recurrently through trillions of synaptic interfaces by utilizing a range of signaling chemicals to produce excitatory and inhibitory changes in electro-ionic conductivity. This dynamic, evolving system produces emergent phenomena with which we are intimately familiar such as consciousness, intelligence, learning, and prediction. The realization of hardware-based neuromorphic networks requires the ability to fabricate highly interconnected, complex wiring architectures with integrated circuit elements whose nonlinear properties emulate those of biological neurons and synapses.

Fabrication of micro- and nanoscale devices with complex architectures, especially those with some degree of random structural topology, is difficult using solely lithographic methods due to challenges in forming robust intra- and inter-device connections in a cost-efficient manner. However, combining directed and self-assembly of nanoscale building blocks into functional device components offers a promising route to creating intricate patterns of nanoscale components. To create operable devices based on nanoscale architectures, two basic issues must be addressed: which materials to use and how to pattern them into networks that have some degree of randomness without negatively affecting their functional characteristics.

Simple metals continue to be the material of choice for wires and interconnects in the fabrication of electronic devices. The power-law relationship known as Rent’s Rule formalized the trend between the number of connections in integrated circuit designs and the number of internal components, such as logic gates, and how these are strongly related to both logical capacity and complexity of the interconnect architecture. This relationship infers that the limits on synthetic complex architectures lies in the cost of fabrication, with specific focus on interconnect and wiring strategies.[31,32] Research has shown that biological neural systems also obey this relationship.[13] Whereas biological networks realize a balance of cost and complexity through structural self-similarity and hierarchical modularity, ANN implementations based on TAM/von Neumann architectures remain at the mercy of this “cost of wiring”. While motivating the creation of bio-inspired devices, Rent’s Rule further underscores the fact that new methods, differing not only in scale but also in kind, must be developed to meet these challenges.

Solution phase electrochemistry offers an intriguing approach to the unconventional fabrication of complex metallic structures. In particular, the electroless deposition of various metals through the spontaneous reduction of soluble metal cations is a mature technology that has been employed extensively in macroscopic plating applications and the manufacture of printed circuit boards (PCBs). In contrast to plating applications, dendritic (fractal) growth processes have been studied extensively for various reasons.[14–16] Unwanted, spontaneous growth of dendritic metal protrusions through insulating layers has posed an engineering challenge as the resulting electrical shorts lead to device failures. In a more positive light, interest in these intricate structures generated insightful mechanistic models, such as diffusion-limited aggregation (DLA), that were
catalogue of nonlinear, solid-state electronic elements for use in integrated circuits and solid-state memory. A class of these, known as hysteretic resistive (memristive) switches, has received substantial attention as a synapse-like element for use in next generation neuromorphic computers. Resistive switches (RS) are two-terminal circuit elements that are distinguished from simple resistors by nonlinearities in the relationship between current and voltage across their terminals. These nonlinearities, generically referred to as memristance, can take various functional forms, from a smooth dependence on the time integral of current passed through the device, to discontinuous jumps at some threshold value, or combinations thereof. The resultant nonlinear dynamics can produce behaviors typically associated with biological neural networks, including long-term potentiation, long-term depression, spike timing dependent plasticity, and associativity. The basic RS is a nanoscale device composed of a metal-insulator-metal (MIM) junction that can be fabricated using a variety of materials.

An exciting subset of electro-ionic RS known as atomic switches exhibit common RS characteristics including pinched I–V hysteresis and large ON/OFF switching ratios as well as more exotic behaviors such as multistate switching in quantized increments of conductance. The distinguishing feature of the atomic switch as compared to other memristive systems is its operational mechanism: atomic switches utilize metal filament formation/annihilation and a concurrent bias-catalyzed phase transition within a solid-state electrolyte metal-insulator-metal (MIM) interface. One prevalent atomic switch configuration employs MIM interfaces of silver and silver sulfide (Ag$_2$S). This chalcogenide undergoes a temperature-dependent and bias-catalyzed transition from the monoclinic, semiconducting α-Ag$_2$S phase (acanthite, 2.5 × 10$^{-3}$ Ω$^{-1}$cm$^{-1}$) to a body-centered...
cubic, metallic β-Ag₂S phase (argentine, 1.6 × 10⁹ Ω⁻¹ cm⁻¹). The argentine phase has a remarkably high diffusion coefficient for silver, approximately equal to that of gaseous silver atoms at an equivalent temperature and density. Under applied external bias, this formulation operates via redox coupled ion migration of silver ions within the metallic argentine phase. While some RS are strictly non-volatile, the Ag-Ag₂S atomic switch exhibits nonlinear, time-dependent conductance that has lead to the observation of a number of fascinating synapse-like properties including short- (volatile) and long- (non-volatile) term memory. Robust operation of these devices at rates up to 1 MHz over 10⁵ cycles further enhances their potential applicability as a synthetic synaptic element.

To date, atomic switches have been primarily fabricated through advanced lithographic methods in regular, crossbar-type architectures that are promising candidates for nanoscale memory applications when operated in isolated, single device configurations. However, their operational characteristics are less well understood when connected in series, parallel, or directly coupled through their ionically conductive active layer, as would be required to implement computation in the TBTu/CNR paradigm. Inspired by the exciting synaptic properties of the AgAg₂S|Ag atomic switch configuration and its material compatibility with our scheme for fabricating complex nanowire networks, we have characterized the properties of interconnected atomic switches as a means to examine their potential applicability as physical implementations of TBTu/CNR-based computation.

5. Critical Atom Switch Networks

Complex networks of coupled nonlinear elements commonly manifest non-trivial evolution through dynamic system reconfigurations, which enable enhanced maintenance of spatiotemporal correlations and maximally efficient signal propagation. These features are associated with systems in critical states, and are crucial to the proposed implementation of hardware-based TBTu/CNR-inspired machines. We have fabricated and examined the operational characteristics of an electroionic device composed of a highly interconnected network of interfacial atomic switches wired through electroless self-assembly. Formation of the complex atomic switch network entailed conversion of as-prepared metallic nanogap and crossbar-like interfaces into metal-insulator-metal (MIM) junctions (AgAg₂S|Ag) through gas phase sulfurization. Due to the nature of the electroless deposition process and resulting random network topology, a thorough survey of sulfurization conditions was carried out to optimize the fabrication protocol.

Progressing from isolated, individual synaptic assemblages to an assemblage of electro-ionically coupled units introduces an extensive set of collective interactions capable of producing emergent behaviors. Spatially distributed atomic switch junctions interact through local variations in ionic concentration and electrochemical potential that depend on the combined electrical resistance of the entire network and the memory-dependent state of all other electro-ionically interconnected switches. Dynamical complexity is expected given that atomic switch synapses are volatile memristive systems that exhibit a conductance decay time constant dependent on their operational history.

To examine these properties, atomic switch networks were investigated by I–V spectroscopy. In common with isolated crossbar-type devices, as-fabricated atomic switch networks required an initial forming step during which a sustained, high (~6 V) bias would bring about a large but temporary drop in resistance. While parameters of the forming step varied from device to device, this requisite step indicates the successful preparation of MIM interfaces within the network. After forming, slow voltage sweeps (1 V·s⁻¹) resulted in pinched hysteresis curves (Figure 3a) with an ON/OFF ratio of 10³, further validating the formation of a functional atomic switch network with behavior analogous to that of a two-terminal RS device. Repeatable switching was observed over 10⁴ cycles, and was successfully operated up to a 1 kHz switching rate. Conditions of no applied bias resulted in a return to the OFF state, as expected from the operational mechanism of this particular AgAg₂S|Ag configuration. Un-sulfurized control devices comprised of a purely metallic network demonstrated ohmic I–V characteristics at intermediate voltages (±3 V) followed by irreversible breakdown at high bias.

To rule out the possibility that network activity was simply the result of conductance localization along a dominant pathway, creating in essence a single large, serial atomic switch, the device was characterized using ultrasonic IR imaging at room temperature (Figure 3b). These results revealed thermal emission from Joule heating throughout the network, indicating distributed and dynamic power dissipation during operation. Further, the application of spatially-defined voltage stimulation enabled controlled activation/deactivation of local regions within the network while enhanced overtones in the device frequency response were also observed as predicted by recently reported modeling of current flows in random memristor networks stimulated with a sinusoidal voltage. These results collectively indicate the successful formation of an interconnected network of nonlinear elements, in this case atomic switches.

Emergent behavior was observed during pulsed voltage stimulation, in analogy to methods employed in neuroscience to probe cortical cultures. Under typical conditions (2 V, 10 ms pulses, 10% duty cycle), the current response fluctuated through a wide range of metastable conductance states associated with discrete network configurations (Figure 3c–f), as classified by residence times in a given state ranging from milliseconds (within a single stimulation pulse) to several seconds (across hundreds of pulses). Specifically, all conductance states whose persistence time exceeded that of the measurement bandwidth (10 kHz) were designated as temporally metastable. Observation of both increased and decreased conductivity during stimulation can be attributed to internal network dynamics, as conductance of isolated atomic switches only increases in response to sequences of identical stimulation pulses.

Previously unreported current fluctuations of this kind are ascribed to dynamic redistribution of network connectivity caused by actions of both individual switches as well as electroionic coupling throughout the shared active layer. Specifically, formation of a conducting filament results in localized depletion of silver cations within the solid electrolyte and thereby inhibits the formation of filaments at nearby MIM interfaces.
Due to the high diffusion constant of Ag⁺ in the β-Ag₂S, this non-stoichiometric region may extend relatively large distances and induce weak electro-ionic coupling even between distant switches. Furthermore, concurrent formation and annihilation of conductive filaments will redistribute current flow, thereby modifying local electrical potentials across the network. These local variations sum to produce the observed fluctuations in global network conductance. While direct mechanistic confirmation of the observed conductance fluctuations would be useful, the inferred mechanism proposed here provides a rationale for future optimization of the network architecture.

Critical dynamics are of ultimate importance for applications of TBTu/CNR-based computation. Indicators of criticality typically include power-law scaling of 1/f fluctuations and temporal metastability. Analysis of the power spectral density of network conductivity in the activated state revealed 1/f power law scaling over five orders of magnitude. Electro-ionic coupling within the atomic switch network generated metastable conductance states, which were analyzed for temporal correlations. Comparing the probability of state duration with its likelihood indicated a power law distribution (Figure 3c–f), indicating a diverging temporal correlation length. Observations of both spatially distributed electro-ionic activity within the network and the long-term persistence of metastable state residence times alongside short-term, rapid fluctuations between many available conductance states are strong indicators of critical system dynamics during intermittent pulse operation. These metastable conductance states represent unique configurations of the network and infer behavior similar to those of spatiotemporal states associated with neural dynamics and those required by reservoir computation models.

6. Outlook and Perspectives

The value of exploring new paradigms in computation cannot be overstated, as the challenges of moving “beyond CMOS” undoubtedly provide inspiration and motivation for the next generation of scientists and engineers. Likewise, elucidating the fundamental nature of intelligence remains a question for the ages in fields spanning all of human endeavor. Drawing on a historical perspective of seminal developments in computer science, complex systems theory and neuroscience, we have set out to propose a hardware-based approach to neuro-morphic computation that aims to harness the power of highly coupled, nonlinear systems. We feel that the perspectives and results described herein represent a potentially important link between the requirements for real-time, multi-sensory computation and ongoing advances in neuroscience through a readily addressable physical system with collective behaviors.
analagous to those currently observed in biological neural networks. Research into applications of artificial neural networks toward biologically inspired computation has been greatly facilitated by modern developments in neuroscience. Recent findings have shown biological neural networks to operate in a persistent critical state, a feature commonly associated with the critical point of a second-order phase transition and power law scaling of internal system dynamics. Under such circumstances, the system correlation length diverges in both space and time, indicating that the influence of past events decays slowly and physically distinct points within the system are coupled regardless of the magnitude of separation. Spatiotemporal correlations of this type have been shown to maximize memory, transmission of information, and adaptability within complex networks, such that each part of the system is communicating with every other part of the brain, for every time of its history. A class of critical systems emerge from coupled networks of nonlinear elements governed by threshold dynamics that relax quickly compared to a slower external driving force, an arrangement that allows these systems to settle into a range of correlated metastable states. This model is more than superficially reminiscent of our current understanding of neural dynamics, and has been employed in recent forms of advanced neural network research including, but not limited, to reservoir computation of multi-domain data systems such as those used in autonomous locomotion, proximity sensing and global positioning as well as a wide variety of sensing applications. Technological trends toward the growth of multi-domain and distributed sensing systems represent the seminal challenge for new forms of emerging computation in the center of Turing’s birth.

Acknowledgements

A.Z.S. and A.V.A. contributed equally to this work. The authors gratefully acknowledge Dante Chialvo, Kang Wang, Bob Schwartz, Igor Ovchinnikov and Brian Shieh for their input and assistance. This work was partially supported by the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) World Premier International (WPI) Research Center for Materials Nanoarchitectonics (MANA) and the Defense Advanced Research Projects Agency (DARPA) - Physical Intelligence Program (BAA-09-63), US Department of Defense.

Published online: October 31, 2011