ASSOCIATIVE LEARNING

Hava T. Siegelmann
University of Massachusetts Amherst, Amherst, MA, 01003, USA

Robert Kozma
Tennessee University Professor of Mathematics, The University of Memphis, USA

Keywords: Neural networks, recall, consolidation, reconsolidation, adaptation, Turing computation, Super-Turing computation.

Contents

Summary

Memory is a reflection of who we are, how we understand our environment and how we react to it. The processes of loading new memories, retrieving and adapting them - are foundational to learning, and are done by way of association rather than by scanning memories serially. New input leads to the recall of corresponding memories; similar inputs are associated with each other and with similar memories.

Associative learning concerns association between input and extant memory traces, and the modification of these existing traces. Computational modeling of memory allows researchers to replicate processes that are only partially known to attempt to fill in gaps in understanding. Such models hold the potential to increase our comprehension of disease, develop better education and training methods**,** and create more human-like artificial intelligence**.** Neural networks have, to date, proven the most applicable computational models in this area. The field of Neural Networks places particular emphasis on associative learning to explain how the human brain may work.

This chapter summarizes prominent models describing how this learning occurs, typically using neural network modeling. By way of background, we start with historical works of Turing, and Hebb, culminating with current models.

1. Introduction

The field of neural networks involves mathematical and computational network models, where nodes represent neural cells and their weighted connections represent the synapses connecting them. Updating the connections is considered learning; and, learning algorithms differentiate the various neural network models. The goal of neural networks is to provide functional models describing intelligence; and they are frequently used as the intelligent component of larger, binary computational systems for engineering. Ultimately, "Super-Turing" neural networks, those that are analog and recurrent, are like other neural nets - non-symbolic, semi-parametric and learn from their environment. A binary, Turing system can associate symbols, but is unable to make associations at a sub-symbolic level. Super-Turing neural nets are capable of sub-symbolic associations giving them a greater ability to achieve associative learning and are thus much closer to the way our brains may work.

British mathematician Alan Turing (1912-1954) pioneered the field of computational intelligence. In 1936 as an undergraduate student, he created a model describing how humans calculate basic mathematical operations by following a series of small steps, that was later termed an ``algorithm." It was Turing's innovation to separate the problem itself from the specific input data - focusing instead on the steps designed to solve it. Another of Turing's significant innovations was the use of external memory to hold information for calculations. Over the years, Turing's model, known as the, "Turing machine," has become the foundation for virtually all computers.

Turing suggested that intelligence requires the ability to learn (and be creative) rather than mechanically follow commands; Turing himself refers to the Turing machine as a, "non-intelligent" model. Turing believed that in the future, machines, like people, would have the power to learn, be able to adapt to their environment and to be intelligent. His thoughts correlate closely with the Super-Turing computational model, which mathematically encompasses all adaptive and analog machines (Siegelmann, 2013, 1995).

Turing had a particular interest in mathematically describing the brain. In 1948, he introduced a mathematical structure, he called "unorganized machines" – the earliest conception of an adaptive neural network: The model consisted of a general graph of binary nodes (representing neurons) and interconnecting edges that followed a converging training procedure. Turing, presaging some of today's latest understandings of brain processes and learning, suggested that a baby starts with semi-random brain connectivity that becomes increasingly organized with experience and learning. His pioneering adaptive neural networks, like all subsequent neural networks, are included in the Super-Turing paradigm. Super-Turing computation was named in Turing's honor as a way of advancing computational intelligence models from static (the program is supplied) and binary (memory stores 0-1 bits only) - as in the Turing machine, to adaptive (program can learn and adapt to changing conditions) and continuous (memory can store continuous values like brightness and depth) computational intelligence as in the Super-Turing model.

In 1949, Donald Hebb (1904-1985), a Canadian psychologist, suggested a computational rule of associative learning based on the organization of the nervous system (Hebb,1949). Hebb's postulate, still frequently used in neurobiological modeling, proposes that neurons firing repeatedly in close temporal proximity have causal or semantic connections, and that the synaptic connection between them is strengthened by biological processes as a response to the simultaneous firing. This rule has been stated as "fire together, wire together."

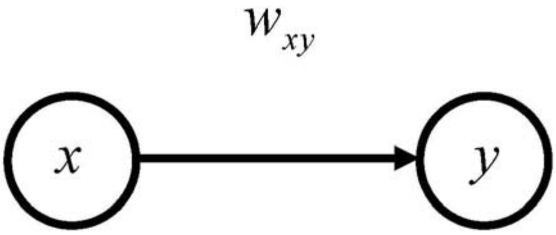

Figure 1. A mathematical way to write the Hebb rule is as follows: If x and y are two binary neurons (either in 0 or 1) with a joint synapse with strength w_{xy} then the synaptic strength can change based on the joint firing of the two neurons, according to the difference: $\Delta w_{xy}(t+1) = \mu x(t)\,y(t)$ where μ is a learning rate parameter and the value $\mu x(t)\,y(t)$ can be positive only when both x and y fire at time t.

The Hebb postulate is the foundation of the modern computational learning rule used most frequently in modeling biological neural networks: The so called "STDP learning rule" (Spike Timing Dependent Plasticity) describes complex biological processes that strengthen connections between neurons that fire at approximately the same time. While the STDP learning rule is very similar to the Hebb rule, the main difference is that the STDP puts emphasis on which neuron fired earlier and which fired later, and the connection is strengthened only in this direction. While the details known today and used in describing the STDP rule are far more sophisticated than those known in the 1940's, the basic principles are very similar. Current science considers STDP part of the explanation of brain function. The influential work explaining this model appears in (Markram et al, 1997).

Following, we will discuss influential models of associative learning, which connect aspects of math, biology, and psychology.

2. Memory as an Attractor System

The Theory of Dynamical Systems is a mathematical field that presents equations that govern the temporal evolution of values of an interacting set of variables. This field has been used to explain various biological and social phenomena, e.g., AIDS, epidemiology, heart arrhythmia, schizophrenia, social obesity; and it has been given rise to well known scientific subfields such as chaos theory, complex systems, and network theory. Neural networks can also be viewed as particular types of dynamical systems where the interacting variables are the neurons.

Dynamical systems are called "dissipative" when their temporal evolution goes to particular repeating areas in their phase space. These areas, which act much like lakes that collect the flow from surrounding rivers, are called attractors.

While the processes of encoding, recalling and adapting memories are not fully understood, the memory system, due to its ability to store and then retrieve different memory traces, is frequently modeled by dissipative neural networks, where the attractors represent different memory traces: When input stimuli is presented, the brain reaches a constant temporal-spatial firing, which represents a particular associated memory trace. Similarly, in an attractor based dynamical system, the inputs belong to different basins of attraction, and the system will flow to the attractor representing the input's basin. Following are prominent models of attractor memory that allow for associative learning.

2.1 The Hopfield Model and Basic Generalizations:

John Hopfield introduced what came to be known as the Hopfield network in 1982 (Hopfield, 1982). The Hopfield net is an approach to modeling memory – a content addressable system with binary threshold nodes. The network consists of binary units connected in symmetric connections: Thus, if neuron <a> is connected to <b> with weight strength <w>, also neuron <b> will be connected to neuron <a> with the same strength <w>. The original use of the Hopfield network was as an "auto-associative memory" where inputs (ideally) retrieve from memory the one item that most resembles them. To accomplish this, the system works in two modes (also called stages): First, loading of memory (consolidation): Memory items are presented to the neurons, connections among the neurons are set so the system will remember them and they will become its attractors**.** In the second mode, partial or noisy inputs that resemble already loaded memory items are presented to neurons in the expectation of recalling the full memory based on the partial information. During this process, neural values (not the values of their connections) change until they reach convergence in firing. The Hopfield network is not a perfect description of the brain processes it seeks to model for a number of reasons: First, the number of memories that can be stored is very small and dependent on the input dimension; second, biologically - the brain is not a symmetric network; third, the Hopfield network saturates quickly and then may converge to other memory items, which are similar but were not intentionally loaded in the first stage.

Since its inception, many improvements have been incorporated to make the Hopfield network a bit more biologically correct or better as an associative memory system. One such work was suggested in (Kosko, 1988), where memory is improved from being only auto-associative (remember the self) to have hetero-association where inputs and outputs can be different.

2.2. The Grossberg Network

Grossberg and colleagues introduced a family of networks, which are based on continuous time updating dynamical systems (Cohen and Grossberg 1983). They assert that their networks are a super-set of the Hopfield network in terms of presenting other dynamical flows in addition to going to attractors such as oscillatory behavior and bi-directional memory. These networks were applied by Grossberg and colleagues to explain some principles appearing in top-down attention processes and other more complex paths for the retrieval of memories.

2.3. Localist Attractor Network (LAN)

In 2001, Zemel and Mozer introduced (Zemel and Mozer, 2001) a layered network model where the attractors themselves are neurons in a higher layer above the dynamic part of the neurons where convergence occurs. This way, spurious attractors do not exist and the system can load as many attractors as needed. The resultant model is a more reliable memory with higher capacity.

2.4 Chaos Based Models

Another line of work suggested that convergence in neurons and the neural system is not necessarily to a fixed point attractor or attractors of a repeating pattern ("limit cycle"), but rather will have chaotic flow or go to a chaotic attractor. While functionally, it is not necessarily the case that chaotic memory systems are richer in abilities than ones that flow to simpler attractors, this sub-field is interesting from a mathematical point of view, and possible connections to biology, most prominently the limbic system, were suggested as well (Kaneko, 1990; Aihara, 1994; Kozma and Freeman, 2001; Kaneko and Tsuda, 2003).

2.5 Kernel Associative Memory (KAM)

KAM is a network model introduced by Nowicki and Siegelmann in 2010 (Nowicki and Siegelmann, 2010). The model combines attractor dynamics, but it is far more practical than the Hopfield and other models described above. The model includes an input space composed of continuous valued vectors, rather than only binary ones, providing the potential of representation of real world, analog information like colors and brightness. The number of attractors in the model is independent of the input dimension as in the LAN (and is thus practically infinite unlike the very bounded number of attractors in the Hopfield and similar networks).

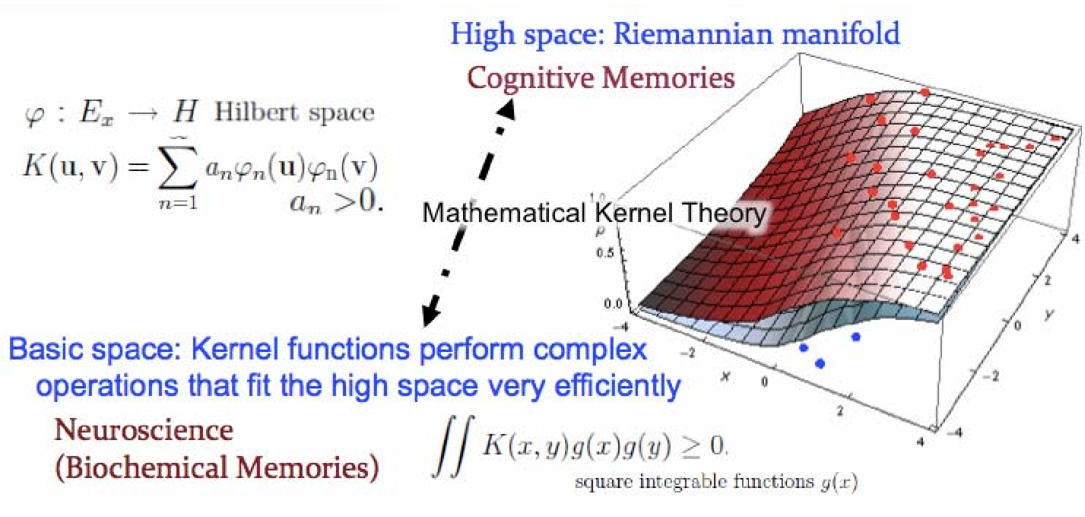

Figure 2. KAM transfers from observable (real) space to storage in internal (neural) high-dimensional space via kernel functions: both loading and retrieval are simple.

-
-

Bibliography

[1] Siegelmann, HT (April 2013). "Turing on Super-Turing and Adaptivity". *Journal Progress in Biophysics & Molecular Biology*. pii: S0079-6107(13)00027-8. [A review paper that describes the Super-Turing model of computation and its possible effect in Artificial Intelligence and Biology].

[2] Siegelmann, HT (April 1995) "Computation Beyond the Turing Limit," *Science* 238(28): 632-637. [A research paper introducing a particular dynamical system that has chaotic properties, and that computationally cannot be described by the Turing machine model and instead can be fully explained by the Super-Turing model].

[3] Hebb, D (1949). *The Organization of Behavior*. New York: Wiley & Sons [Hebb's seminal work describing his postulate of associative learning at the neural level].

[4] Markram H, Lübke J, Frotscher M, Sakmann B (January 1997). "Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs". *Science* 275 (5297): 213–5. [A research paper that building on results of previous researchers, suggests how synapses change plastically by the Spike Timing Dependence associative learning process].

[5] Hopfield JJ (April 1982). "Neural networks and physical systems with emergent collective computational abilities". *Proceedings of the National Academy of Sciences of the USA*, 79(8): 2554–2558. [Hopfield introduces a physics based symmetric network, that can keep memory traces in its attractors; An energy function is defined in which its minima are the attractors.].

[6] Kosko, B (Jan/Feb 1988). "Bidirectional Associative Memories". *IEEE transactions of systems, man and cybernetics*: 49-60. [A generalization of the Hopfield attractor networks to allow for input-output functions].

[7] Cohen MA and Grossberg S (1983). "Absolute stability of global pattern formation and parallel memory storage by competitive neural networks". *IEEE Transactions on Systems, Man, and Cybernetics*, 13: 815-826. [Grossberg's group early research - somewhat unique and separate from the main stream - includes other attractor networks of the time].

[8] Zemel R and Mozer, M (2001). "Localist attractor networks". *Neural Computation* 13: 1045–1064. [A paper that introduced a particular attractor network with three levels: Input, neural, and attractors: The dynamic flow is in the middle layer, while each neuron in the third level stands for a memory trace. This allows for great separation among attractors, a known problem in all Hopfield-like networks].

[9] Kaneko K (1990), "Clustering, coding, switching, hierarchical ordering, and control in network of chaotic elements". *Physica* D 41: 137-172. [Introduces chaos theory in neural networks].

[10] Aihara K (1994). "Chaos in neural response and dynamical neural network models: toward a new generation of analog computing". in *Towards the harnessing of chaos*, edited by M. Yamaguti, Elsevier, Science Publishers B.V., Amsterdam: 83–98. [Introduces analog computation and chaotic behavior in networks].

[11] Kozma R and Freeman WJ (2001). "Chaotic resonance: Methods and applications for robust classification of noisy and variable patterns". *International Journal of Bifurcation and Chaos* 10: 2307-2322. [Practical usages of chaotic dynamic in neural networks].

[12] Kaneko K and Tsuda I eds (2003). Focus issue on chaotic itinerancy. *Chaos* 13: 926-1164. [Special dynamical behavior in networks and memory].

[13] Nowicki D and Siegelmann HT (june 2010). "*Flexible Kernel Memory*". PLOS One 5: e10955. http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0010955 [Memory model that includes unbounded number of memory traces and allows inputs of variation in shades – much more practical than previous model].

[14] Ben-Hur A, Horn D, Siegelmann HT and Vapnik V (2001). "Support vector clustering". *Journal of Machine Learning Research* 2: 125-137. [A very successful clustering method based on kernel functions].

[15] Sara SJ (2000). "Retreival and reconsolidation: Toward a neurobiology of remembering". *Learning and Memory*, 7(2): 73-84. [Coins the term "reconsolidation" as the process of re-inserting information after recall].

[16] Blumenfeld B, Preminger S, Sagi D and Tsodyks M (2006). "Dynamics of memory representations in networks with novelty-facilitated synaptic plasticity". *Neuron* 52(2): 383-394. [A version of Hopfield network that allows the update of synaptic strength over time – rather than only during original loading].

[17] Siegelmann HT (2008). "Analog-Symbolic Memory that Tracks via Reconsolidation". *Physica* D: Nonlinear Phenomena 237 (9): 1207-1214. [A generalization of the localist attractor network to acknowledge the process of reconsolidation].

[18] Osan R, Tort A and Amaral O (2011). "A Mismatch-Based Model for Memory Reconsolidation and Extinction in Attractor Networks". *PLoS ONE*, 6(8):e23113. [A small update to the Hopfield model that allows it to remove memories].

[19] Nowicki D, Verga P, Siegelmann HT (2013). "Modeling Reconsolidation in Kernel Associative Memory". *PLoS ONE* 8(8): e68189. doi:10.1371/journal.pone.0068189 [Richest model of Reconsolidation in computational terms].

[20] Kohonen T and Honkela T (2007). "Kohonen Network". *Scholarpedia*. [Recent paper summarizing self-organizing neural network by Teuvo Kohonen and follows updates and applications].

[21] Tal A and Siegelmann HT (June 2013). "Conscience mechanism in neocortical competitive learning," ICCN2013, Sweden. [Recent work that utilizes a biological version of the SOM to explain recently found biological connection in layer 5].

[22] Roth F, Siegelmann HT and Douglas RJ (2007). "The Self-Construction and -Repair of a Foraging Organism by Explicitly Specified Development from a Single Cell," *Artificial Life* 13(4: 347-368. [Artificial life model that uses the concept of self organization – joining chemistry and physics]

Biographical Sketches

Dr. Siegelmann's research focuses on mathematical and computational studies of the brain, memory, biological computation, and systemic disorders. Her research into biologically inspired cognitive and neural processes has led to theoretical modeling and original algorithms of machine systems capable of superior computation and learning. A unifying theme underlying her research is the study of dynamical and complex systems. Siegelmann's seminal Turing machine equivalence of recurrent neural networks theorem and Super-Turing theory introduced a new subfield of computer science. Her modeling of geometric neural clusters resulted in the highly utile and widely used Support Vector Clustering algorithm with Vladimir Vapnik and colleagues, specializing in the analysis of high-dimensional, complex data. Her work is often interdisciplinary, combining complexity science, dynamical and computational theories with biological sciences, neurology, physics and medicine. Recent contributions include computational and dynamical system studies of reconsolidation, the circadian system, addiction, cancer, and genetic networks; applications in intelligent robotics and advanced human-robot interfaces for military and health applications are currently funded. Her engineering background has been central to her NSF funded development of analog hardware for brain-like intelligence. She remains active in supporting young researchers and encouraging minorities and women to enter and advance in STEM. She has years of experience consulting with industry, creating educational programs including interdisciplinary and international programs, fund raising, and in educational administration and organization.

Dr. Kozma's current research interests include spatio-temporal dynamics of neural processes, random graph approaches to large-scale networks, such as neural networks, computational intelligence methods for knowledge acquisition and autonomous decision making in biological and artificial systems; he also published in related fields including signal processing; and design, analysis, and control of intelligent systems. He has served since 2009 as a Professor of Computer Science, University of Memphis, Memphis, Tennessee, and Professor of Mathematical Sciences, University of Memphis, Memphis, Tennessee. He has also been the director of Computational Neurodynamics Laboratory, presently CLION, FedEx Institute of Technology of the University of Memphis, Memphis since 2001. Kozma serves on the AdCom of IEEE Computational Intelligence Society CIS (2009–2012) and on the Governing Board of the International Neural Network Society INNS (2004–2012). He is Chair of the Distinguished Lecturer Program, IEEE CIS. He has been a Technical Committee Member of IEEE Computational Intelligence Society since 1996, and IEEE Senior Member. He also served in leading positions at over 20 international conferences, including General Chair of IEEE/INNS International Joint Conference on Neural Networks IJCNN09 at in Atlanta; Program Co-Chair of International Joint Conference on Neural Networks IJCNN08/WCCI08 in Hong Kong; Program Co-Chair of IJCNN04, Budapest, Hungary; Chair for Finances of IEEE WCCI06, Vancouver, Canada. He is Associate Editor of ‘Neural Networks (Elsevier),’ ‘IEEE Transactions on Neural Networks,’ ‘Neurocomputing’ (Elsevier), ‘Journal of Cognitive Neurodynamics’ (Springer), Area Editor of ‘New Mathematics and Natural Computation’ (World Scientific), and ‘Cognitive Systems Research.’