Neural-Network Models for Associative Memory Based on Multidimensional Neurons

Roseli S. Wedemann
Universidade do Estado do Rio de Janeiro (UERJ)
Angel R. Plastino
Universidad Nacional del Noroeste de
la Provincia de Buenos Aires (UNNOBA)

Statistical Mechanics for Complexity – 80th Birthday of C. Tsallis, CBPF, Rio 2023

Motivation

- The Hopfield associative memory model has been used to represent an approximation of human memory functioning, and also as an artificial storage device. Neural networks (NNs) used to model associative memory consist of dissipative units (neurons).

- Units interact in a way that the network admits a global energy or Liapunov function.

- The network's global dynamics is such that the system evolves *downhill* in the energy landscape.

- In most models for associative memory, individual neurons are described as one-dimensional, dynamical systems.

- The capacity of human memory to store and retrieve information is central to many mental processes, be they normal, pathological, conscious or unconscious.

- These have been widely studied by psychiatry, psychoanalysis, neuroscience and computational science.

- In recent years, we have developed schematic simulation models based on associative memory that represent aspects of some mental processes as described by psychoanalysis: neurosis, creativity, delusions and the interaction between consciousness and unconsciousness.

- We used Hopfield-like networks and generalizations such as the Boltzmann Machine and Generalized Simulated Annealing (GSA).

1. M. Siddiqui, R. S. Wedemann, and H. J. Jensen, Physica A, 490, 2018.
2. R. S. Wedemann, R. Donangelo, and L. A. V. Carvalho, Chaos 19, 2009.
3. R. S. Wedemann, L. A. V. Carvalho, and R. Donangelo, Neurocomputing, 71, 2008.
4. C. Tsallis and D. A. Stariolo, Physica A, 233, 395—406, (1996).

- Most of the fundamental models in theoretical biology exhibit a dynamics that is dissipative.

- Important examples: Lotka-Volterra models in biological population dynamics; continuous, Hopfield NN models[1]; Cohen-Grossberg network models[1], and various mathematical models for biological, evolutionary processes.

- All biologically inspired models supporting universal computation are nonconservative or, in the case of discrete models, nonreversible.

- They exhibit a modular structure consisting of a set of interacting units, each one characterized by an intrinsic dissipative dynamics.

- In order to represent the complexity of real biological neurons, one may need to describe them as dynamical systems with two or more dimensions. For instance, effects of time-delay lead to multidimensional dynamics.

1. Cohen and Grossberg, IEEE Trans. on Systems, Man, and Cybernetics, 1983 and Hopfield, Proc. Natl. Acad. Sci. **81**, 1984.

- We thus extend the structural description of associative memory NNs to more general scenarios, and formulate a family of NN models, with interacting, dissipative, multi-dimensional neurons (units), encompassing other biological models with nonconservative dynamics.

- These models generate dynamical features akin to those required when modeling associative memory.

1. Nakamura, Y., Torii, K., Munakata, T.: Neural-network model composed of multidimensional spin neurons. Phys. Rev. E, 51 1538-1546, (1995)
2. Stewart, I., Golubitsky, M.: Symmetric networks with geometric constraints as models of visual illusions. Symmetry 11(6), 799 (2019)
3. Solazzi, M., Uncini, A.: Adaptive multidimensional spline neural network for digital equalization, Neural Netw. Signal Process. X. Proc. 2000 IEEE Signal Process. Soc. Workshop (Cat. No. 00TH8501) vol.2, 729–735 (2000)
4. Manzhos, S., Carrington, T., Ihara, M.: Orders of coupling representations as a versatile framework for machine learning from sparse data in high-dimensional spaces. Artificial Intelligence Chemistry 1(2), 100008 (2023)

A Continuous Network with Multidimensional Neurons

We advance a continuous, associative memory, NN model with N multidimensional neurons.

- The i-th neuron $(i = 1, \ldots, N)$ is modeled as an n_i-dimensional, dissipative, dynamical system. Different neurons may have different dimensions.

- States of a neuron, are given by an n_i-dimensional vector $\boldsymbol{X}_i$.

- Two scalar functions $W_i(\boldsymbol{X}_i)$ and $a_i(\boldsymbol{X}_i) > 0$, and two n_i-dimensional vector functions $\boldsymbol{B}_i(\boldsymbol{X}_i)$ and $\boldsymbol{D}_i(\boldsymbol{X}_i)$ are associated with each neuron. These are related to the intrinsic dynamics of i, and with the output signals through which i affects the dynamics of other neurons.

- There is an $n_i \times n_i$ positive-definite, symmetric square matrix $\boldsymbol{M}_i(\boldsymbol{X}_i)$.

- Independent functions that characterize the dynamics of i are a_i, W_i and $\boldsymbol{D}_i$.

- $\boldsymbol{A}^T$ denotes the transpose of a matrix $\boldsymbol{A}$.

$$\boldsymbol{B}_i = -\boldsymbol{M}_i \boldsymbol{\nabla} W_i\,, \tag{1}$$

and

$$\left(\boldsymbol{M}_i^{-1}\right)^T = \frac{\partial \boldsymbol{D}_i}{\partial \boldsymbol{X}_i}\,. \tag{2}$$

- The $n_i \times n_i$ square matrix $\frac{\partial \boldsymbol{D}_i}{\partial \boldsymbol{X}_i}$ is the Jacobian matrix associated with $\boldsymbol{D}_i(\boldsymbol{X}_i)$,

$$\frac{\partial \boldsymbol{D}_i}{\partial \boldsymbol{X}_i} = \begin{bmatrix} \frac{\partial \boldsymbol{D}_{i;1}}{\partial \boldsymbol{X}_{i;1}} & \cdots & \frac{\partial \boldsymbol{D}_{i;1}}{\partial \boldsymbol{X}_{i;n_i}} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \frac{\partial \boldsymbol{D}_{i;n_i}}{\partial \boldsymbol{X}_{i;1}} & \cdots & \frac{\partial \boldsymbol{D}_{i;n_i}}{\partial \boldsymbol{X}_{i;n_i}} \end{bmatrix} \quad . \tag{3}$$

  $\boldsymbol{X}_{i;j}$ and $\boldsymbol{D}_{i;j}$ are the j-th components of vectors $\boldsymbol{X}_i$ and $\boldsymbol{D}_i$.
- The positive-definite character of matrix $\boldsymbol{M}_i$ means that, for any vector $\boldsymbol{V} \in \Re^{n_i}$, one has

$$\boldsymbol{V}^T \boldsymbol{M}_i \boldsymbol{V} > 0 \, . \tag{4}$$

- In the above and similar equations, we consider $\boldsymbol{V}$ as an n_i-dimensional column vector, and $\boldsymbol{V}^T$ as a row vector.
- Interactions between the neurons are codified in an $N \times N$ matrix array $\boldsymbol{\mathcal{C}}$, with elements $\boldsymbol{C}_{ik}, i, k = 1, \ldots, N$. Each $\boldsymbol{C}_{ik}$, which describes the interaction between neuron i and neuron k, is itself an $n_i \times n_k$ rectangular matrix whose elements are real numbers. The elements of the matrices $\boldsymbol{C}_{ik}$ are constants, not depending on the neurons' states. We assume that[1]

$$\boldsymbol{C}_{ki} = \boldsymbol{C}_{ik}{}^T . \tag{5}$$

1. See R. S. Wedemann, A.R. Plastino, C. Tsallis, Phys. Rev. E 94, 2016 for a discussion on symmetry of connections

Dynamics of a single neuron i is governed by the equations of motion

$$\frac{d\boldsymbol{X}_i}{dt} = a_i \boldsymbol{B}_i = -a_i \boldsymbol{M}_i \boldsymbol{\nabla} W_i \,, \tag{6}$$

where $\boldsymbol{\nabla}$ represents the gradient, $\boldsymbol{\nabla} W_i = (\partial W_i/\partial \boldsymbol{X}_{i;1}, \ldots, \partial W_i/\partial X_{i;n_i})^T$.

W_i plays the role of a potential energy (or Liapunov) function. Indeed, let us consider the time derivative of W_i,

$$\frac{dW_i}{dt} = (\boldsymbol{\nabla} W_i)^T \cdot \frac{d\boldsymbol{X}_i}{dt} = -a_i \, (\boldsymbol{\nabla} W_i)^T \, \boldsymbol{M}_i \, (\boldsymbol{\nabla} W_i) < 0 \,, \tag{7}$$

where we used the fact that $\boldsymbol{M}_i$ is a positive-definite matrix.

Therefore, an isolated neuron always evolves *downhill* in the W_i, multi-dimensional landscape. Vector $d\boldsymbol{X}_i/dt$ does not, in general, point in the direction opposite to $\boldsymbol{\nabla} W_i$, but always has a negative projection on the direction of $\boldsymbol{\nabla} W_i$.

We now consider a network of N interacting neurons, whose equations of motion are

$$\frac{d\boldsymbol{X}_i}{dt} = a_i \left[\boldsymbol{B}_i - \sum_{k=1}^{N} \boldsymbol{C}_{ik} \boldsymbol{D}_k \right], \quad i = 1, \ldots, N \,. \tag{8}$$

The first term within the bracket corresponds to the intrinsic dynamics of the i-th neuron, while the second term describes the effects of the other neurons on the dynamics of i.

Strengths of the interactions between neurons are given by the matrices $\boldsymbol{C}_{ik}$.

The effects of the k on i are given by the elements of the matrix $\boldsymbol{C}_{ik}$, and not by a single coefficient, as in a standard Hopfield neural network.

Liapunov Function

We now prove that our NN model admits a Liapunov, or energy function. Let us consider the function

$$\Omega\left(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_N\right) = \left[\sum_{i=1}^{N} W_i(\boldsymbol{X}_i)\right] + \frac{1}{2}\left[\sum_{i,k=1}^{N} \boldsymbol{D}_i{}^T \boldsymbol{C}_{ik} \boldsymbol{D}_k\right]. \tag{9}$$

Ω is a function of the total state of the network, given by the set of vectors $(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_N)$.

We compute the time derivative of Ω,

$$\frac{d\Omega}{dt} = \left\{\sum_{i=1}^{N} (\boldsymbol{\nabla} W_i)^T \cdot \frac{d\boldsymbol{X}_i}{dt}\right\} + \frac{1}{2}\sum_{i,k=1}^{N}\left[\frac{d\boldsymbol{D}_i{}^T}{dt}\boldsymbol{C}_{ik}\boldsymbol{D}_k + \boldsymbol{D}_i{}^T \boldsymbol{C}_{ik}\frac{d\boldsymbol{D}_k}{dt}\right]. \tag{10}$$

Substituting $\frac{d\boldsymbol{X}_i}{dt}$ by the right hand sides of the equations of motion (8), yields

$$\begin{aligned}\frac{d\Omega}{dt} = &\left\{\sum_{i=1}^{N} a_i \left(\boldsymbol{\nabla} W_i\right)^T \cdot \left[\boldsymbol{B}_i - \sum_{k=1}^{N} \boldsymbol{C}_{ik}\boldsymbol{D}_k\right]\right\} \\ &+ \frac{1}{2}\sum_{i,k=1}^{N}\left[\frac{d\boldsymbol{D}_i{}^T}{dt}\boldsymbol{C}_{ik}\boldsymbol{D}_k + \boldsymbol{D}_i{}^T \boldsymbol{C}_{ik}\frac{d\boldsymbol{D}_k}{dt}\right].\end{aligned} \tag{11}$$

Using the expression (1) for the $\boldsymbol{B}_i$, and the symmetry property (5) of the $\boldsymbol{C}_{ik}$,

$$\begin{aligned}\frac{d\Omega}{dt} \quad = \quad & -\left[\sum_{i=1}^{N} a_i \boldsymbol{B}_i{}^T \cdot \left(\boldsymbol{M}_i{}^{-1}\right)^T \boldsymbol{B}_i\right] + \left\{\sum_{i=1}^{N} a_i \boldsymbol{B}_i{}^T \cdot \left(\boldsymbol{M}_i{}^{-1}\right)^T \left[\sum_{k=1}^{N} \boldsymbol{C}_{ik}\boldsymbol{D}_k\right]\right\} \\ & + \left[\sum_{i,k=1}^{N} \frac{d\boldsymbol{D}_i{}^T}{dt} \boldsymbol{C}_{ik}\boldsymbol{D}_k\right] . \end{aligned} \tag{12}$$

We want to prove that the time derivative $d\Omega/dt$ is always non-positive.

We need first to consider the time derivative of the vectors $\boldsymbol{D}_i(\boldsymbol{X}_i)$. We have

$$\frac{d\boldsymbol{D}_i}{dt} = \frac{\partial \boldsymbol{D}_i}{\partial \boldsymbol{X}_i}\frac{d\boldsymbol{X}_i}{dt} = \frac{\partial \boldsymbol{D}_i}{\partial \boldsymbol{X}_i} a_i \left[\boldsymbol{B}_i - \sum_{j=1}^{N} \boldsymbol{C}_{ij}\boldsymbol{D}_j\right], \tag{13}$$

implying that

$$\frac{d\boldsymbol{D}_i{}^T}{dt} = a_i \left[\boldsymbol{B}_i{}^T - \sum_{j=1}^{N} \boldsymbol{D}_j{}^T \boldsymbol{C}_{ji}\right] \left(\frac{\partial \boldsymbol{D}_i}{\partial \boldsymbol{X}_i}\right)^T, \tag{14}$$

where $\frac{\partial \boldsymbol{D}_i}{\partial \boldsymbol{X}_i}$ is the Jacobian matrix of the vector-valued function $\boldsymbol{D}_i(\boldsymbol{X}_i)$ (eq. (3)).

Using relation (2), one obtains

$$\frac{d\boldsymbol{D}_i{}^T}{dt} = a_i \left[\boldsymbol{B}_i{}^T - \sum_{j=1}^{N} \boldsymbol{D}_j{}^T \boldsymbol{C}_{ji} \right] M_i^{-1}. \tag{15}$$

Substituting $\frac{d\boldsymbol{D}_i{}^T}{dt}$ in (12) by the right hand side of (15), one gets

$$\begin{aligned} \frac{d\Omega}{dt} \quad = \quad & - \left[\sum_{i=1}^{N} a_i \boldsymbol{B}_i{}^T \cdot \left(\boldsymbol{M}_i{}^{-1} \right)^T \boldsymbol{B}_i \right] + \left[\sum_{i=1}^{N} a_i \boldsymbol{B}_i{}^T \cdot \left(\boldsymbol{M}_i{}^{-1} \right)^T \left(\sum_{k=1}^{N} \boldsymbol{C}_{ik} \boldsymbol{D}_k \right) \right] \\ & + \left[\sum_{i=1}^{N} a_i \left(\sum_{k=1}^{N} \boldsymbol{C}_{ik} \boldsymbol{D}_k \right)^T \left(M_i^{-1} \right)^T \boldsymbol{B}_i \right] \\ & - \left[\sum_{i=1}^{N} a_i \left(\sum_{k=1}^{N} \boldsymbol{C}_{ik} \boldsymbol{D}_k \right)^T \left(M_i^{-1} \right)^T \left(\sum_{j=1}^{N} \boldsymbol{C}_{ij} \boldsymbol{D}_j \right) \right]. \end{aligned} \tag{16}$$

It follows from (16), and that $a_i(\boldsymbol{X}_i) > 0$ and the matrices $\boldsymbol{M}_i$ are positive-definite that

$$\frac{d\Omega}{dt} = -\sum_{i=1}^{N} a_i \, \boldsymbol{V}_i^T \left(\boldsymbol{M}_i^{-i}\right)^T \boldsymbol{V}_i \leq 0\,, \tag{17}$$

where

$$\boldsymbol{V}_i = \boldsymbol{B}_i - \left(\sum_{j=1}^{N} \boldsymbol{C}_{ij} \boldsymbol{D}_j\right). \tag{18}$$

Our network model thus exhibits a dynamics admitting an energy, Liapunov function Ω. Ω can be decomposed as the sum of two terms: one inherited from the gradient-like character of the units' intrinsic, dissipative dynamics, and one arising from the interactions between the neurons.

The system always evolves downhill in the energy landscape, tending to the landscape's local minima, complying with the basic behavior typical of an associative memory NN.

Connection with the Cohen-Grossberg NN Model

The NN models proposed by Cohen and Grossberg[1] constitute a particular instance of the general models described by the equations of motion (8).

Consider the particular instance of the dynamical system (8), where the behavior of neurons are determined by one-dimensional, dynamical systems, *i.e.* the case where $n_i = 1,\ i = 1, \ldots, N$.

Then the state of each neuron is described by a single number x_i, the functions $a_i(\boldsymbol{X}_i)$, $\boldsymbol{B}_i(\boldsymbol{X}_i)$ and $\boldsymbol{D}_i(\boldsymbol{X}_i)$ become the single-variable, real-valued functions $a_i(x_i)$, $b_i(x_i)$ and $d_i(x_i)$, and each matrice $\boldsymbol{C}_{i,j}$ becomes a single numerical coefficient c_{ij} (the weights).

The NN's equations of motion (8) then reduce to those of the neural network model proposed by Cohen and Grossberg[1], which are

$$\frac{dx_i}{dt} = a_i(x_i)\left[b_i(x_i) - \sum_{k=1}^{N} c_{ik} d_k(x_k)\right], \quad i = 1, \ldots, N\,, \tag{19}$$

where the functions $a_i(x_i)$ and $d_i(x_i)$ comply with $a_i(x_i) > 0$ and $d_i'(x_i) > 0$.

1. Cohen and Grossberg, IEEE Trans. on Systems, Man, and Cybernetics, 1983 and Hopfield, Proc. Natl. Acad. Sci. **81**, 1984.

Condition (5) reduces to the requirement of symmetric weights, $c_{ij} = c_{ji}$.

The energy (Liapunov) function is then

$$\Omega = \sum_{i=1}^{N} W_i(x_i) + \frac{1}{2} \sum_{i,k=1}^{N} c_{ik} d_i(x_i) d_k(x_k)\,, \tag{20}$$

where

$$W_i(x_i) = -\int_0^{x_i} b_i(z) d_i'(z) dz. \tag{21}$$

This energy function coincides with the one derived by Cohen-Grossberg.

The equations of motion (19) can be recast in terms of the partial derivative of the Liapunov function Ω, as

$$\frac{dx_i}{dt} = -\left[\frac{a_i(x_i)}{d_i'(x_i)}\right] \frac{\partial \Omega}{\partial x_i}\,. \tag{22}$$

These equations govern the dynamics of the network's state, which at each instant is represented by the N phase-space variables $\{x_1, x_2, \cdots, x_N\}$.

The Hopfield model for continuous NNs constitutes a particular realization of the Cohen-Grossberg model.

When $a_i(x_i) = -1/\tau_i$, $b_i(x_i) = x_i$, and $d_i(x_i) = g(x_i)$, where all the τ_i's are constant parameters, the equations of motion (19) become

$$\tau_i \frac{dx_i}{dt} = -x_i + \sum_{j=1}^{N} c_{ij}\, g(x_j)\,, \tag{23}$$

which have the same form as the equations governing the continuous Hopfield model.

Concluding Remarks

1. We investigated an extension of the structural scheme of associative memory, NNs to more general settings, where the neurons are described by multi-dimensional, dissipative, dynamical systems.

2. We advanced a coupling scheme for dissipative, multi-dimensional units, that leads to dynamical features akin to those required when modeling associative memory.

3. The family of NNs that we have proposed admits an energy (Liapunov) function, such that the network always evolves downhill in the energy landscape.

4. Our general scheme includes, as particular instances, the continuous NNs proposed by Cohen and Grossberg, as well as the continuous version of the Hopfield model.

5. It would be interesting to explore the implications of our model, with respect to the tension between conservative and nonconservative dynamical models in biology.

6. It would also be worth to investigate possible relations between our approach to networks of multi-dimensional neurons, and intriguing recent developments on the theory of quantum mechanical NNs.

Thank You

We acknowledge financial support from the Brazilian National Research Council (CNPq), the Rio de Janeiro State Research Foundation (FAPERJ) and the Brazilian agency which funds graduate studies (CAPES).

“As Biologists come to focus more of their efforts on the brain-mind, most of them have become convinced that the mind will be to the biology of the twenty-first century what the gene has been to the biology of the twentieth century. Thus, Francois Jacob (1998) writes, ‘the century that is ending has been preoccupied with nucleic acids and proteins. The next one will concentrate on memory and desire. Will it be able to answer the questions they pose?’ ”

Eric Kandel (Nobel prize in physiology or medicine, 2000), “Biology and the Future of psychoanalysis, a new intellectual framework for psychiatry revisited”, 1999, in *Psychiatry, Psychoanalysis and the New Biology of Mind.*