Dynamical Networks with Topological Self-Organization

MICHAIL ZAK

Ultracomputing Group
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109
USA

And

Center for Engineering Science Advanced Research, Oak Ridge National Laboratory

Abstract: - Coupled evolution of state and topology of dynamical networks is introduced. Due to well organized tensor structure, the governing equations are presented in a canonical form, and required attractors as well as their basins can be easily implanted and controlled. This new class of dynamical networks can represent phenomenological models for self-organization of biomolecular sequences or of the genome in order to understand enormous complexity in their behavior

Key-Words:
Tensor structure, attractors, self-organization

Recent advances in nonlinear dynamics opened up a new direction in information processing based upon special properties of solutions to dynamical systems. In this new role, the dynamical system is not derived from the Lagrange or Hamilton principles, but it is rather created to simulate behavior of an observed object whose law of motion is not well understood. For instance, even the simplest living systems interact in a non-Newtonian way via flows of information which are produced and processed by a signaling system whose complexity on a bio-chemical level is enormous. In order to incorporate this kind of a phenomena into the process of self-organization and pattern formations on the physical level of description, one has to find a dynamical equivalent which would capture the phenomenology of the observed behavior. Such an equivalent can be associated with the concept of an attractor which is the most powerful modelling tool for synthesis of complex patterns of behavior. Attractor is a stable dissipative structure which does not depend (at least, within a certain basin) upon the initial conditions. Due to this property, the whole history of evolution prior to attraction becomes irrelevant, and that represents a great advantage for information processing, and in particular, for associative memory and pattern recognition. The central problem of the synthesis of dynamical networks for the purpose of information processing is to place a prescribed type of attractors (with the corresponding basins) at prescribed locations and explicitely control these locations subject to changes of the objective of the performance. Partly this problem is solved by recurrent neural networks which represent the most effective analog tool of information processing. We will start with a brief review of this dynamical model in order to outline some of its limitations.

The standard form of recurrent neural networks (NN) is[1]

$$\dot{x}_i = -x_i + \sigma\left(\sum_i w_{ij} x_j\right), \quad w_{ij} = Const \qquad (1)$$

where x_i are state variables, w_{ij} are control parameters, or interconnection weights (associated with the NN topology), and $\sigma(\cdot)$ is the sigmoid function.

The system (1) is nonlinear and dissipative (i.e., div $\dot{x} = 0$) due to the sigmoid function. The nonlinearity and dissipativity are necessary (but not sufficient) conditions that the system (1) has attractors. The locations of the attractors and their basins in phase (or configuration) space are prescribed by an appropriate choice of the synaptic interconnections w_{ij} which can be found by solving an inverse problem (followed by the stability analysis), or by learning which is a dynamical relaxation procedure based upon iterative adjustments of w_{ij} as a result of comparison of the net output with known correct answers. In both cases, w_{ij} are constant, and that is the first limitation of recurrent NN. Indeed, although the NN architecture (1) is perfectly suitable for such tasks as optimization, pattern recognition, associative memory, i.e., when fixed topology is an advantage, it cannot be exploited for simulation of a complex dynamical behavior which is presumably comprised of a chain of self-organizing patterns (like, for instance, in genome) since for that kind of tasks, variable topology is essential. The latest review of NN with evolutionary topology as well as a numerical proof of concept can be found in [2] However, there is no general analytical approach to the synthesis of such NN. And now we are coming to the second limitation of NN (1): their architecture does not have a tensor structure. Indeed, the state variables and the interconnections w_{ij} cannot be considered as a vector and a tensor, respectively since their invariants are not preserved under linear transformations of the state variables. Obviously, the cause of that is the nonlinearity in the form of the sigmoid function. That is why the dynamical system (1) (even with a fixed topology) cannot be decoupled and written in a canonical form; as a result of that, the main mathematical tools for NN synthesis are based upon numerical runs.

In this paper we propose a new architecture for dynamical networks which reconciles nonlinearities and the tensor structure. It will be shown that these networks can be written in a decoupled form, and that will allow one to implant any number of attractors with prescribed basins whose types and locations can be explicitely changed in real time.

We will start with the following trivial equation:

$$\dot{x} = wx, \tag{2}$$

and assume that the weight coefficient w is coupled with the state variable x by the simplest linear law:

$$w = a - x, \quad a = Const \tag{3}$$

Then, combining (2), and (3) one obtains:

$$\dot{x} = (a - x)x, \tag{4}$$

The solution to Eq. (4) has two equilibrium points:

$$\bar{x} = 0 \text{ and } \bar{x} = \alpha \tag{5}$$

which represent a repeller and a static attractor, respectively, if $a > 0$, and a static attractor and a repeller if $a < 0$.

The basins of attractions are $x > a/2$ and $x < a/2$, respectively.Hence, by changing the parameter a, one can explicitly control the location of the attractor as well as its basin.

Let us introduce a set of equations similar to (2):

$$\dot{x}_i^o = w_{ii}^o x_i^o, \quad w_{ii}^o = a_{ii}^o - \tilde{x}_{ii}^o, \quad i = 1,2,\ldots n \tag{6}$$

assuming that w_{ii}^o, a_{ii}^o, and $\tilde{x}_{ii}^o$ are the principal values of co-axial tensors

$$W^o = \begin{pmatrix} w_{ii}^o & & 0 \\ & \text{O} & \\ 0 & & w_{nn}^o \end{pmatrix}, \quad A^o = \begin{pmatrix} a_{ii}^o & & 0 \\ & \text{O} & \\ 0 & & a_{nn}^o \end{pmatrix},$$
$$X^o = \begin{pmatrix} \tilde{x}_{ii}^o & & 0 \\ & \text{O} & \\ 0 & & \tilde{x}_{nn}^o \end{pmatrix}, \tag{7}$$

and $X^o = \{x_1^o, \ldots x_n^o\}$ is the state variable vector whose components x_i^o are referred to the principal directions of the tensors (7). Here $\tilde{x}_{ii}^o$ are new auxiliary variables which later will be linearly related to the state variables x_i^o.Then Eqs. (6) can be rewritten in an invariant form:

$$\dot{X} = WX, \quad W = A - \tilde{X} \tag{8}$$

and projected onto an arbitrary orthogonal system of coordinates defined by the corresponding orthogonal matrix $\|B\|$ with the components b_{ij}:

$$x_i = b_{ij}x_i^o, \quad w_{ij} = b_{ik}b_{kj}w_{kk}^o, \quad \tilde{x}_{ij} = b_{ik}b_{kj}\tilde{x}_{kk}^o, \quad a_{ij} = b_{ik}b_{kj}a_{kk}^o \tag{9}$$

Let us now define the auxiliary variable $\tilde{x}$ such that

$$\tilde{x}_{ii}^o = x_i^o \tag{10}$$

i.e., the principal coordinates of the state variable vector are equal to the eigenvalues of the tensor X.

After the coordinate transformation (9), the relationship (10) between $\tilde{x}_{ii}$ and x_i will be more sophisticated since the components of the tensor $\tilde{x}_{ii}^o$ and of the vector x_i^o are changed by different rules (see Eq. (9)). Combining Eqs. (9) and (10) yields:

$$\tilde{x}_{ij} = b_{ik}b_{kj}\tilde{x}_{kk}^o = b_{ik}b_{kj}x_k^o = b_{ik}b_{kj}b_{kq}x_q, \text{i.e.}, \tilde{x}_{ij} = b_{ik}b_{kj}b_{kq}x_q \tag{11}$$

Hence, in the new coordinates, the system (6) supplemented by the condition (10) takes the following form:

$$\dot{x}_i = w_{ik}x_k, \; i = 1,2,\ldots n \tag{12}$$

$$w_{ij} = a_{ij} - b_{ik}b_{kj}b_{kq}x_q; \; i,j = 1,2,\ldots n \tag{13}$$

Eq. (12) forms a fully coupled and fully interconnected dynamical network since each state variable x_i directly connected to each of the rest variables. Eq. (13) describes the joint evolution of the state x_i and topology. However, despite the complexity of the system, it has a well organized tensor structure: its in variant properties such as dissipasivity, the number and the type of attractors, their basins, the distances between them, etc., are preserved under the linear transformations, and therefore, they can be found from Eqs. (6) and (10) which represent the same physical object in special (principal) coordinates. Since all the equations in (6) are decoupled, and each of its equations is identical to Eq. (2) and (3), one can conclude that it has a static attractor whose principal coordinates are:

$$\bar{x}_i^o = \begin{cases} a_{ii}^o \text{ if } a_{ii}^o > 0 \\ 0 \text{ if } a_{ii}^o < 0 \end{cases} \quad (14)$$

and the coordinates of the same attractor for the system (12), (13) are found from Eqs. (9):

$$\bar{x}_i = b_{ij}\bar{x}_i^o \quad (15)$$

The coordinates of its boundaries are

$$x_i^b = \tfrac{1}{2} b_{ij} a_{ii}^o \quad (16)$$

Thus, the dynamical network (12), (13) has a static attractor (15) whose position and basin can be explicitly controlled in real time via the corresponding changes of the parameters a_{ii}^o.

In order to place m equilibrium points, the system (8) should be generalized as follows

$$X = \sum_{q=1}^{m} W_q X, \; W_q = A_q - \tilde{X}, \; i = 1,2 \ldots n; \; q = 1,2 \ldots m \quad (17)$$

Then Eqs. (12) and (13) take the form

$$\dot{x}_i = w_{ik_1}^{(1)} w_{k_1k_2}^{(2)} \ldots w_{k_mj}^{(q)} x_j, \quad w_{ij}^{(q)} = a_{ij}^{(q)} - b_{ik} b_{kj} b_{kp} x_p \quad (18)$$

The principal coordinates of the equilibrium points

$$\bar{x}_{i(q)}^o = a_{ii(q)}^o, \quad i = 1,2 \ldots n; q = 1,2 \ldots m \quad (19)$$

and their coordinates in the system (18)

$$\bar{x}_{i(q)} = b_{ij} a_{jj}^o \quad (20)$$

About half of the equilibrium points will be attractors (they will alternate with repellers when q runs over 1,2, n).

So far the only static attractors were discussed. In order to place periodic or chaotic attractors, we will move from differential to difference equations and replace Eqs. (2) and (3) by the following:

$$x^{(t+1)} = w^{(t)} x^{(t)}, \; w^{(t)} = a\left(1 - x^{(t)}\right), \; c = Const \quad (21)$$

Then instead of Eq. (4) one obtains the logistic map:

$$x^{(t+1)} = a\left(1 - x^{(t)}\right)x^{(t)} \quad (22)$$

which has static attractors $(\text{for } a < 2)$, periodic attractors $(\text{for } 2 < a < 3.6)$ and chaotic attractors $(a > 3.6)$[3].

Repeating the same line of argumentation as those described by Eqs. (6) — (13), one arrives at the following:

$$\overset{o}{x}_i^{(t+1)} = \overset{o}{w}_{ii}^{(t)} \overset{o}{x}_i^{(t)}, \quad \overset{o}{w}_{ii}^{(t)} = \overset{o}{a}_{ii}\left(1 - \overset{o}{x}^{(t)}\right) \quad (23)$$

$$X^{(t+1)} = W^{(t)} X^{(t)}, \quad W^{(t)} = A - A\tilde{X}^{(t)} \quad (24)$$

$$x_i^{(t+1)} = w_{ik}^{(t)} x_k^{(t)}, \quad w_{ij}^{(t)} = a_{ij} - b_{ik} b_{kj} b_{kq} x_q^{(t)} \quad (25)$$

The last two equations present the dynamical network in the tensor and the coordinate forms, respectively.

Let us start with placing a static attractor. As follows from Eq. (22), the condition for that is

$$x' = a\left(1 - x'\right)x', \; \text{i.e.,} \quad (26)$$

$$x' = 1 - \tfrac{1}{a} \quad (27)$$

while the other equilibrium point $x' = 0$ is a repeller.

Then the principal and the transformed coordinates of the static attractor for Eqs. (23) are, respectively:

$$\bar{x}_i^o = 1 - \tfrac{1}{a_{ii}}, \quad \bar{x}_i = b_{ij}\left(1 - \tfrac{1}{a_{jj}}\right) \quad (28)$$

If a periodic $(2 < a < 3.6)$ or a chaotic $(a > 3.6)$ attractors of the logistic map (22) in the principal coordinates are approximated by functions

$$\hat{x}_j^o = f_i(t, a_{ii}) \quad (29)$$

then the asymptotic behavior of the solution to the dynamical network (25) in the basin of attraction is expressed as follows:

$$\hat{x}_i = b_{ij} f_j\left(t, a_{jj}\right) \tag{30}$$

Let us now summarize the basic characteristics of the proposed architecture of dynamical networks (12), (13) and (25). Firstly, the nonlinearity and disipativity of the network is created by joint evolution of state and topology variable rather than by a sigmoid function. Secondly, both the state variables and the interconnection weights are represented by the components of a vector and a second rank tensor, respectively, so that the invariant properties of the network can be exposed in a special (principal) system of coordinates. In these coordinates, the network is reduced to a system of simple uncoupled equations which are suitable for an explicit and systematic synthesis of networks exhibiting complex behaviors, by placing and real-time control attractors with prescribed basins. The main area of applications of such networks is inverse problems in dynamics: given solution-find the model. For instance, there is a broad class of behaviors of enormous complexity performed by living systems [4] and in particular, by the human genome. On the level of dynamical description, they can be presented as a chain of different self-organizing patterns whose changes are triggered by external conditions. Since there is no accurate physical model for such systems, one of the most effective ways to study them is a black-box approach when the input-output relationships provided by observation data exploited to develop a phenomenological model. The proposed dynamical network is structured to provide such a model. The black-box capacity of this model is defined by the number of free controllable parameters. These parameters can be divided in two groups. The first group consists of the weights $a_{ii}^{(q)}(i = 1,2,\ldots n;\ q = 1,2,\ldots m)$. They control invariant properties of the dynamics: the number and types of attractors, their basins and relative positions. It should be noticed that smooth changes of these parameters may cause abrupt qualitative changes in the solutions behavior (for instance, when $a_{ii}^{(q)}$ changes its sign, the attractors and repellers exchange their positions). The number of qualitatively different invariant features of the synthesized motion is, loosely speaking, of the order of $n!$ if n is the dimensionality of the network.

The second group of the parameters consists of the components b_{ij} of the coordinate transformation. These parameters are responsible for geometrical representation of the dynamical characteristics in the frame of reference to which the observation data are referred. All the parameters $a_{ii}^{(q)}$ and b_{ij} are supposed to be found by the best fit of the simulated motions into observation data.

Thus the proposed dynamical network can be viewed as a simple and effective tool for reconstruction of models for complex dynamical behaviors based upon observation data.

ACKNOWLEDGMENTS

The research described in this paper was performed by the Jet Propulsion Laboratory, California Institute of Technology. This research was performed in association with the Center for Engineering Science Advanced Research (CESAR), Oak Ridge National laboratory. Funding was provided by the Engineering Research Program of the DOE Office of Science, under contract DE-AC05-00OR22725 with UT-Battelle LLC.

REFERENCES

Bornholdt, S; T. Rohlf, Phys. Rev. Vol. 84, No. 26, June 26, 2000.

Hertz, J; A. Krugh, R. Palmer, Introduction to the theory of Neural Computations, Addison-Wesley, 1991, pp 13-23

Percival, I; D. Richards, Introduction to Dynamics, pp. 200-202, Cambridge Univ. Press, 1982

Zak, M; Phys. Lett. A, 255 (1990) pp 110-118.