ESANN'1999 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 21-23 April 1999, D-Facto public., ISBN 2-600049-9-X, pp. 153-158

Tackling the stability/plasticity dilemma with double loop dynamic systems

Christophe Lecerf
ECArt/LRIA, Université Paris8 Vincennes St-Denis
15, rue Catulienne, F-93526 St-Denis. France
E-mail: clecerf@univ-paris8.fr

Abstract: Open and organized systems such as living organisms regulate their exchanges in order to maintain adaptation to their environment. When one reduces a biological organism to its central nervous system (CNS), adaptation comes up as an information flow exchange between the CNS and its environment. Though, the main mechanism used so far to explain learning is derived from the Hebb's hypothesis and it relies on structural modifications of the network through changing weights on connections. The double loop concept proposed here is the core of a structural and dynamic model tackling with incremental learning in large neural networks. A computer simulation of this concept is briefly described, then is given an equivalent mathematical dynamic system that is related to Thomas' biological feedback theory. Due to the double loop architecture, the observed dynamics shows that the model gives a built-in functional answer to the stability/plasticity dilemma.

1. Introduction

Since the insight of connectionism made by D. Hebb, a large number of models have been proposed (review in Rumelhart & al. [10]). Compared to each other, all of these models differ by their architecture, algorithm, type and necessary features of data, and type of learning. While they have their own advantages and their specific features that make them more suitable to some kind of application, they all rely on the hebbian/anti-hebbian paradigm.

Many biological findings have corroborated the Hebb's hypothesis since it has been proposed, putting more emphasis on the structural dimension of learning in artificial neural networks (ANN). As a reverse proof of the biological hebbian mechanism, epigenesis [2] proved that unused connections could even disappear, at least in the peripheral nervous system. At the lowest-scale level, analytical descriptions of ionic flows in axons and large dendrites give nowadays a perfect explanation of spike propagation in biological cells. Lastly, the NMDA receptor seems to have the necessary features (voltage and ligand dependency) for realizing the hebbian paradigm.

On the other hand, when using a large-scale analysis, living organisms are described as complex open systems relying on their exchanges with their environment for survival. In this case, emphasis is put on exchange flows. The contrast is strong between the approach used in ANN, focused on low-level structural mechanisms, and the key feature for system analysis, which is exchange flow.

A model that would meet both these requirements, e.g. in which both the structural and flow angles would receive an expression, could help to understand how incre-

ESANN'1999 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 21-23 April 1999, D-Facto public., ISBN 2-600049-9-X, pp. 153-158

mental learning takes place in large networks. This paper presents such a model that gives room for learning, data and flow modeling, and that offers a built-in dynamic functional answer to the stability/plasticity dilemma.

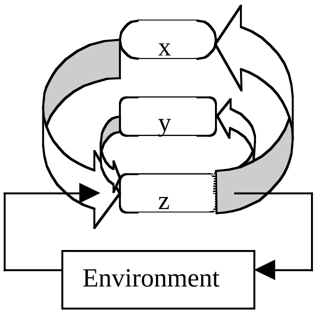
2. The double loop concept

Adaptive and learning systems use feedback signals to adjust the command signal that achieves in the end the correct result. More precisely, nested loops are encountered and we have shown in previous papers [6, 8] that such regular architectures are well fit for incremental learning purposes.

2.1 The mental object

The proposed model uses the double loop mental object as a corner stone. A mental object is defined as a dynamic flow running across a connectionist structure forming a double feedback loop (scheme below). The mental object is a concept sharing features from both connectionist and dynamic systems angles. Basically, every path of the loop is a kind of synfire chain, as introduced by Abeles [1], but the use of two nested feedback loops gives the mental object some special dynamic system properties. Actually, this model explicitly relies on a functional coupling between the connectionist structure and the dynamic flow.

Considered as a dynamic system at steady state, the mental object evokes two coupled and stabilized resonators oscillating on their own specific signal as long as the internal flow is compatible with it. Considered as a connectionist structure producing a regular flow of data (in the steady state case), the mental object is a set of stabilized weights specifying connections in the network. This set is the structural description of the resonator's specific signal. Nevertheless, as the structure supports the flow, which reinforces the structure, the connectionist and the dynamic angles of a double loop mental object cannot be split into pieces.



Indeed, the Hebb's law corollary in the double loop model is the equivalence between learning and adaptation to environment through an established steady state. Local application of Hebb's law at the cell level induces mutual reinforcement between the structure and the flow over it. In order to reach a steady state, the connectionist structure and the dynamic flow have to reinforce each other.

3. Computer simulation

3.1 Cells

In our computer implementation of the model, each cell sets up a path for a unidirectional flow going from its dendritic end to its axonal end. Excitatory and inhibitory effects are properties of cells, so all efferent connections of an excitatory (inhibitory) cell are positive (negative) values. Every input connection is weighed, this weight being modified over time with a hebbian rule, adapted to an epigenetic mechanism.

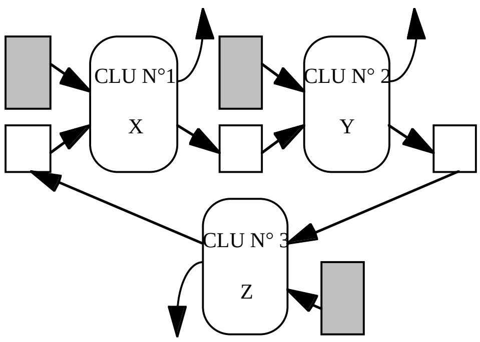
ESANN'1999 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 21-23 April 1999, D-Facto public., ISBN 2-600049-9-X, pp. 153-158

Epigenesis is a biological phenomenon that illustrates the competitive momentum at work between connections during the psychomotor system maturing phase [2]. In the double loop model, taking inspiration from this mechanism, the number of completely efficient connections on each dendrite is limited by a stabilizing factor that is an essential resource for any connections to improve its weight during learning. This epigenetic competition is completed in our model by a process distributing information over a subset of connections in order to achieve a diverging process in the network.

Therefore, the set of weights underlying the connectionist structure is naturally unstable: its natural trend drives toward complete disconnection by decreasing all weights to zero when no regular flows exist. This tendency is counter-balanced by learning according to Hebb's law when a permanent and regular signal flow exists. Here are the pieces of the typical double loops mental objects dynamic equilibrium: the natural trend to forget is balanced by the stabilizing move of learning adaptation through exchanges with the environment.

3.2 Network Architecture

For computational simulations of the network, we use two kinds of elements. The first one gather the learning and computational parts of cells features together, and they are called Computation and Learning Units (CLU). They are autonomous UNIX processes and simulate the activity of a few dozens cells. The second element is used for communication and it relies on a UNIX inter-process facility called Shared Memory Segments (SHM). The SHMs carry the values produced by the cells in each CLU. Both SHMs and CLUs may be manipulated independently, thus providing full facilities for building different architectures with the model. The number of SHMs used in input and output of a CLU is locally configured for each process, and a unique SHM may be used by many CLUs, forming for instance a loop architecture as shown in the above scheme.



4. The double loop mathematical model

The mathematical model we propose here reorganize around the double loop concept many data coming from different fields. On the one hand, we have found great support for trying to use dynamic systems in the results of W. Freeman [3] and the theory of R. Thomas [13] about biological feedback. The set of equations described in this paper is one chosen and modified from his attempt to show that one can build a chaotic dynamic system with a unique feedback loop [12].

On the other hand, there are biological data supporting our conception. M. Abeles [1], as soon as '82, pointed out that time synchrony and phase lags may be the important parameter in information coding. In the recent years, G. Buzsaki and his group [5, 9,

ESANN'1999 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 21-23 April 1999, D-Facto public., ISBN 2-600049-9-X, pp. 153-158

11] and A. Villa [14] confirmed that information is encoded by cell-assemblies, rather than by single cells, in spatio-temporal patterns. According to Buzsaki's group results on interneuron activity, we aggregate a cell assembly as a unique functional unit [11], and we consider its "effective value", i.e. the sum of w_{ij}*c_i, as a unique variable.

Therefore, the model presented here considers the reciprocal influence of small cell-assemblies, forming multiple layers and organized in double loops, rather than some specific weights between cells. As the different coefficients in our system reflect these relative influences and their variations, the system dynamic therefore illustrates the behavior of a large network with structural coupling between the layered assemblies.

The proposed dynamic system for a double loop with 3 layers (x, y, z), or 3 CLUs of the computer model, is the following:

$$\partial x/dt = K(1/x+y+z)x + \sin((y/x+z)y + (z/x+y)z)$$
$$\partial y/dt = K(1/x+y+z)y + \sin((x/y+z)x + (z/x+y)z)$$
$$\partial z/dt = K(1/x+y+z)z + \sin((x/y+z)x + (y/x+z)y)$$

where x, y and z, represents the global output of the three cell assemblies. These values should be understood as the global product of all the cells values by all the weights on their output connections.

where K*1/(x+y+z) is the relative effect of each assembly on itself, e.g. the global result of internal interactions inside the layer of the loop. K is a constant controlling this effect and, according to Thomas' results, it is supposed to be globally inhibitory, which means negative and small ($-1<K<0$). This hypothesis seems to be in accordance with some Buzsaki's group results [11].

where the [y/(x+z)] terms reflects the effect of one cell assembly, here y, relatively to the others, here x and z. This specific term reflects the distribution mechanism we used in our computer model to induce a diverging factor that relies on the output changes of a layer over time.

This system describes a network with nested loops because all layered assemblies feed each other. The structure of each equation directly reflects the structure of the network: each term in the sin() function reflects a directional link between two layers.

The general formulation of a new cell assembly added to the system, which would be represented by a new variable and a new equation, would therefore be the following, in which the αy, βz, γm and … terms reflect the structural coupling between layers:

$$\partial x/dt = E_i x + \sin(\alpha y + \beta z + \gamma m + ...)$$

Assuming the m term represents an external constraint, m would no more be driven by an equation homogenous to the ones of cell assemblies in the system. In such a case, m would be a constant or a linear function of the other variables of the system.

Such a system may receive a numerical resolution using some classical mathematical toolkit. To exhibit the effects of the different parameters, some resolution were made in a 3 layers double loop system with fixed values on the α and β parameters. The

ESANN'1999 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 21-23 April 1999, D-Facto public., ISBN 2-600049-9-X, pp. 153-158

comparison of the resulting curves shows that the global dynamic is driven by the relative influence of the different layers on itself compared to each other, i.e. the Ei/(α+β+...) ratio. All the types of dynamics, classical in such systems, are observed: fixed point and limit cycling, quasi-periodic dynamics and chaos.

In the double loop model, these types of dynamic have a particular meaning. Firstly because of the equivalence between the steady state in the exchange flow with environment and the adaptation/learning process. Secondly because of the signal flow carrier part any double loop plays in the network (see Lecerf [7] for details). Let's consider a subsystem made of a few layers inside a large network, i.e. a subset of equations with limited connections to other variables. A subsystem converging on a fixed point testifies that a regular exchange flow has been established, which means that the considered subsystem has a structure compatible with this exchange flow. On the contrary, all other types of dynamic correspond to irregular exchange flows, thus activating over time outside parts of the network fed by this subsystem because of the changing inputs (i.e. α, β,... factors) in these layers. In the fixed point convergence case, one can say that the changing input coming from environment is carried out by the subsystem's existing structure and that there is no need of the other parts of the network to maintain stability. In the irregular exchange flows case, the changing input is not carried out by the subsystem's existing structure. Because of the irregular output flow changing the Ei/(α+β+...) ratio over time in the layers outside from the subsystem, other parts of the network are activated: architectural plasticity is automatically obtained by recruiting new resources in the network.

Moreover, when forcing the Ei and (α+β+...) factors to fixed values, the relative influence of cell assemblies on each other seems to roughly divide the Ei/(α/β) factor space into two zones (figure 1), therefore confirming that this ratio governs the global dynamic of the system. In the first zone, converging dynamics are systematically observed whatever be the initial conditions sets used, although in the second one, only pseudo-periodic dynamics that do not qualify stability in the double loop model are observed. This suggests a cooperative effect between these parameters.

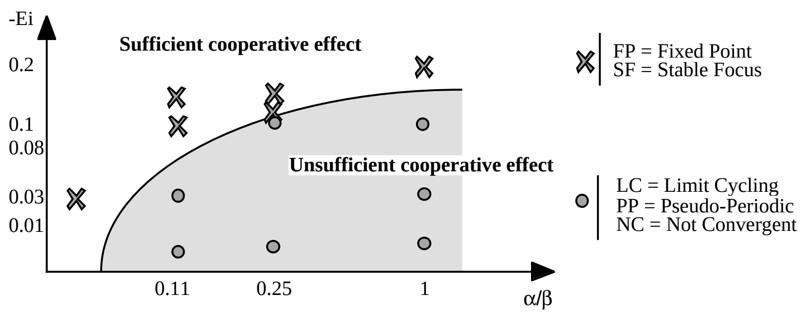


Figure 1: The cooperative effect of the Ei/(α/β) factor

5. Conclusion

We have presented the double loop concept, which associates connectionist and dynamic features, its computer implementation and a large scaled mathematical model.

ESANN'1999 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 21-23 April 1999, D-Facto public., ISBN 2-600049-9-X, pp. 153-158

In this model, a layered cell assembly activity is represented by a variable aggregating value cells and their weights in an "effective value". The coupling between assemblies are modulated by their relative influence. Different network architectures, even some involving identified subsystems, might be described through these parameters. This formal expression of a network by means of a set of equations gives opportunity to qualitatively describe the behavior of large networks. The learning paradigm, associated to the model, relies on adaptation to the environment, which is qualitatively defined by a steady state in the exchange flows. In the double loop model, the stability/plasticity dilemma finds a solution embodied in the relative effect of each cell assembly that is represented by changes in the $Ei/(\alpha+\beta+...)$ factors over time.

6. References

1. M. Abeles, Local cortical circuits, Studies of brain function 6, Springer-Verlag NYC (1982)
2. A. Danchin, P. Courrege, J.P. Changeux, A theory of epigenesis of neural networks by selective stabilization of sysnapses, Proceedings of National Academy of Sciences of USA, No 70, pp. 2974-2978, (1973)
3. W.J. Freeman: Simulation of chaotic EEG patterns with a dynamic model of the olfactory system. Biological Cybernetics 56, pp 139-150, (1987).
4. S. Grossberg, J.W.L. Merrill: the hippocampus and cerebellum in adaptively timed learning, recognition, and movement, Journal of Cognitive Neuroscience 8, pp 257-277, (1996).
5. A. Kamondi, L. Acsady, G. Buzsaki, Dendritic spikes are enhanced by cooperative network activity in the intact hippocampus. Journal of Neuroscience 18, pp 3919-3928, (1998).
6. C. Lecerf, Contribution à la modélisation des réseaux neuronaux et à l'apprentissage automatique. PhD thesis. Université Paris 8 Vincennes Saint-Denis, (1987).
7. C. Lecerf, "Double loop flows and bidirectional Hebb's law in neural networks", VI-DYNN'98 Proceedings, June 22-26, Stockholm, in press, (1998)
8. C. Lecerf, The double loop as a model of a learning neural system, in SCI'98 Proceedings, N. Callaos, O. Omolayole, L. Wang (eds), vol. 1, pp 587-594, July 12-16, Orlando FL, (1998)
9. Z. Nadasdy, A. Bragin, G. Buzsaki: Relationship between recurring spatio-temporal spike patterns and local field oscillations in the hippocampus. Soc. Neurosci. Abst. 23, pp 482, (1997)
10. D.E. Rumelhart, J.L. Mclelland (eds): Parallel distributed Processing, MIT Press, Cambridge MA, (1986).
11. A. Sik, M. Penttonen, G. Buzsaki, Interneurons in the hippocampal dendate gyrus: an in vivo intracellular study. European Journal of Neuroscience 9, pp 573-588, (1997).
12. R. Thomas: Analyse et synthèse de systèmes à dynamique chaotique en terme de circuits de rétroaction, Académie Royale de Belgique, VII, pp 1-6, (1996)
13. R. Thomas, Richard D'Ari: Biological Feedback, CRC Press, Boca Raton FL, (1990)
14. A.E.P. Villa, M. Abeles, Evidence for spatio-temporal firing patterns within the auditory thalamus of the cat, Brain Research 509: pp 325-327, (1990)