

Asymptotic Behavior of an Artificial Neural Network Defined on Multipartite Directed Graph

N.C. Majee and A.B. Roy
Department of Mathematics, Jadavpur University, Calcutta-700032, India

Abstract: Problem statement: Artificial Neural Network (ANN) are simple models to mimic some essential features of the complex central nervous system. ANN models are realistic due to their inherent stochastic nature of neural computation and strong synchronicity. Different ANN models are associated with directed and signed graphs. The present study proceeded by relaxing certain simplifying assumptions in the ANN model. **Approach:** It was assumed that the connected graph associated with the ANN is a multipartite directed graph whose connection comprising of four blocks and two blocks are either both symmetric or both anti symmetric. The convergence of such network was studied in the present research with the help of Lyapunov functional. **Results:** Attractors (fixed points) of such ANN and also limit cycles of different orders are investigated. Bounds of transient length of the neural network were also calculated. Numerical simulation in support of the results was also depicted. **Conclusion:** It was shown that under synchronous updating rule such networks converge to a fixed point or to a limit cycle of period 2 or 4. The bound of transient length was discussed. Conclusions were drawn from the simulation studies carried out in support of the results.

Keywords: Artificial neural network, asymptotic behavior, attractors, fixed points, limit cycles, transient length

INTRODUCTION

Artificial Neural Networks (ANN) are simple models to capture some essential features of the central nervous system. The basic processing unit is (McCulloch and Pitts, 1943) binary units (or linear threshold gate) which is activated only when the weighted sum of its input is larger than some threshold:

$$\sigma = \text{Sgn} \left(\sum_i w_i x_i - \theta \right) \quad (1)$$

where the vector $w = (w_1, \dots, w_n)$ is a vector of weights representing the strength of n pinging synapses from n other neurons with a activation $x = (x_1, \dots, x_n) :: x_i \in \{0,1\}^n$. θ is a threshold value. The output of the network is $\sigma = 1$ if the formal neuron fires a high frequency train of spikes or $\sigma = 0$ if it remains inactivated. Geometrically (1) says that every elementary processor of an ANN corresponds to n dimensional hyperplane:

$$w \cdot x = \theta$$

It is also assumed that a neuron has not fired within the absolute refractory period, the basic cycle-time of

the network and the accumulated Post Synaptic Potential (PSP) decays instantly to zero and the build up of the PSP starts all over again. So the collection of signals, making up the input to a neuron, does not really represent a well defined network state in the dynamical evolution of the network.

It is known that on the average every neuron attempts to fire an action potential in every unit in total cycle-time, independently of all other neurons. This implies a mean updating rate is the inverse of the basic cycle-time. It is difficult to capture such complicated dynamics and instead two simplified versions of network dynamics, namely, synchronous (or parallel) and sequential dynamics. In parallel dynamics, the PSP on each neuron at time $t = n$ is determined by the activation of all other neurons in the time interval $n-1 < t < n$. At the beginning of each period the neuron starts from a zero PSP, that is, after every unit of time all return to these resting membrane potential. This type of dynamics has already been introduced (Caianiello, 1961; Amari, 1972; Little, 1974). Such dynamics is also very favorite in the study of cellular automata (Hoffman, 1987). But ANN models are realistic due to their inherent stochastic nature of neural computation and strong synchronicity.

Corresponding Author: N.C. Majee, Department of Mathematics, Jadavpur University, Calcutta-700032, India

Different forms of the synaptic matrix give rise to different asymptotic dynamics of an ANN under different updating rule. The computational performance of a network, namely, memory, recall is associated with the asymptotic dynamics of the network. There are three types of basic asymptotic dynamics in an ANN namely, chaotic, limit cycle, fixed points. ANN models may never mimic the richness and novelty of human cognition under normal or pathological condition. The ANN, however, may be considered as a metaphor for high brain function which can lead to manifestation of schizophrenia and mania in such disorder (Little and Shaw, 1978).

In ANN models (Hoffman, 1987) many simplifying assumptions were made from neurobiological perspective. This can be summarized as follows:

- (i) Each formal neuron has as many excitatory and inhibitory synapses emanating from axon
- (ii) Each of the neuron can receive input from every neuron and can send output to it and each neuron does not connect directly to itself
- (iii) Symmetric connectivity
- (iv) Updating rule is synchronous/asynchronous

Hopfield (1982) work consists of modeling of Associative Memory by neural network. Given a set of Boolean vector to be memorized, he defines an ANN whose symmetric weights are defined by the Hebbian rule between patterns. Using the analog with spin glass problem he proved that such pattern is a fixed point of the network. He further showed that the sequential network dynamics is driven by a Lyapunov operator.

In the present study we have relaxed slightly the assumptions (ii) and (iii). It is assumed that the connected graph is associated with the ANN is multipartite directed graph and efficacy of the synapses communicating the output neuron j to i is not always equal to the efficacy of the synapse communicating the output of neuron i as input to neuron j . More precisely we have considered connection matrix comprising four blocks out of which two blocks along with the left hand diagonal are null matrices and remaining two blocks are either (i) both symmetric or (ii) both antisymmetric. The convergence of such network is studied in the present study with the help of Lyapunov functional. It is shown that under synchronous updating rule such networks converge to a fixed point or to a limit cycle of period 2 or 4. The bound of transient length is also discussed. Some simulation results in support of our results are also included.

MATERIALS AND METHODS

The model: Let $\pi: \{0,1\}^n \rightarrow \{0,1\}^n$ be a mapping whose components π_1, \dots, π_n are threshold functions:

$$\forall i(1,2,\dots,n), \forall x \in \{0,1\}^n$$

$$\pi(x) = \begin{cases} 1 & \text{if } \sum_i a_{ij}x_j - b_i \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where, $A = (a_{ij})$ is a $(n \times n)$ real matrix known as the connection matrix, $a_{ij} > 0$ means excitatory synaptic weight and $a_{ij} < 0$ an inhibitory weight and (b_1, b_2, \dots, b_n) is a real threshold vector. The synchronous iteration scheme on π is given by:

$$x_i(t+1) = \pi\left(\sum_{j=1}^n a_{ij}x_j(t) - b_i\right) \quad (3)$$

$$i \in I = \{1, 2, \dots, n\}, t_0$$

where, $\pi(x)$ is the threshold function.

Here we have considered the finite neural network with connection matrix A of the form:

$$A = \begin{pmatrix} O_1 & P \\ Q & O_2 \end{pmatrix}$$

where the blocks O_1 and O_2 are null matrices and the blocks P and Q are non null matrices:

Let:

$$I_1 = \{1, 2, \dots, p\}; I_2 = \{p+1, p+2, \dots, n\}$$

$$J_1 = \{1, 2, \dots, n-p\}; J_2 = \{n-p+1, \dots, n\}$$

Such that:

$$I_1 \cup I_2 = I; I_1 \cap I_2 = \emptyset; J_1 \cup J_2 = I; J_1 \cap J_2 = \emptyset$$

We also suppose $b_1 = b_i; i \in J_2$ such that $b_1 \cup b_2 = b$ and $b_1 \cap b_2 = \emptyset$.

Lyapunov functional: We introduce here the Lyapunov functional driving the network dynamics. These operators were first introduced by (Hopfield, 1982) and subsequently by Goles, (1990); Hoffman, (1987) to analyze the fixed point behavior of random sequential iteration of associative network. As for that applications the study of synchronous updating rule and of memory updating were defined and developed in (Hoffman, 1987).

RESULTS

Case 1: We first consider the case when the non null blocks P and Q of the connection matrix A are both symmetric in nature. We can write the blocks P and Q in matrix form as follows:

$$P = (p_{ij})_{n-p \times n-p}: i \in I_2 \text{ and } j \in J_1$$

$$Q = (q_{ij})_{p \times p}: i \in I_1 \text{ and } j \in J_2$$

The synchronous dynamics given by (3) with the strictness condition:

$$\sum_{j \in I} a_{ij} x_j(t) - b_i \neq 0 \quad \forall i \in I \tag{4}$$

We define the Lyapunov functional as follows:

$$E(x(t)) = -\sum_{i \in I} (2x_i(t) - 1) \sum_{j \in I} a_{ij} (2x_j(t-1) - 1) + \sum_{i \in I} (2b_i - \sum_{j \in I} a_{ij}) (2x_i(t) - 1 + 2x_i(t-1) - 1) \tag{5}$$

Now since $a_{ij} = 0 \quad \forall i \in I_1, j \in J_1$ or $\forall i \in I_2, j \in J_2$. We have:

$$E_s(x(t)) = -\sum_{i \in I_1} (2x_i(t) - 1) \sum_{j \in J_2} a_{ij} (2x_j(t-1) - 1) - \sum_{i \in I_2} (2x_i(t) - 1) \sum_{j \in J_1} a_{ij} (2x_j(t-1) - 1) + \sum_{i \in I_1} (2b_i - \sum_{j \in J_2} a_{ij}) (2x_i(t) - 1 + 2x_i(t-1) - 1) + \sum_{i \in I_2} (2b_i - \sum_{j \in J_1} a_{ij}) (2x_i(t) - 1 + 2x_i(t-1) - 1)$$

Lemma 1: If A be the connection matrix of a neural network then $E_s(x(t))$ is a Lyapunov functional for the synchronous iteration given by (3) with (4).

Proof: Let:

$$\begin{aligned} \Delta_t E_s &= E_s(x(t)) - E_s(x(t-1)) - 2 \sum_{i \in I_1} (x_i(t) - x_i(t-2)) \sum_{j \in J_2} a_{ij} (2x_j(t-1) - 1) \\ &\quad - 2 \sum_{i \in I_2} (x_i(t) - x_i(t-2)) \sum_{j \in J_1} a_{ij} (2x_j(t-1) - 1) \\ &\quad + 2 \sum_{i \in I_1} (x_i(t) - x_i(t-2)) (2b_i - \sum_{j \in J_2} a_{ij}) \\ &\quad + 2 \sum_{i \in I_2} (x_i(t) - x_i(t-2)) (2b_i - \sum_{j \in J_1} a_{ij}) \\ &= -4 \sum_{i \in I_1} (x_i(t) - x_i(t-2)) \left(\sum_{j \in J_2} a_{ij} (2x_j(t-1) - 1) - b_i \right) \end{aligned}$$

So we conclude $\Delta_t E_s \leq 0$ and $\Delta_t E_s < 0$ iff $x_i(t) \neq x_i(t-2)$. Therefore $E_s(x(t))$ is a Lyapunov functional for the given dynamics (3) with (4).

Since the network evolves in a finite set $\{0,1\}^n$, the synchronous iteration scheme defined by (3) and (4) converges for any initial configuration in a finite number of steps to a steady state, that is, to a finite cycle. Therefore there exists number $p(x)$ called the period of the cycle of the network, such that:

$$x(t+p(x)) = x(t)$$

Theorem 1: If A be the connection matrix consisting of the non null blocks P and Q of symmetric in nature the orbit of the synchronous iteration are only fixed point and /or cycle of length two.

Proof: Let $(x(0), \dots, x(T-1))$ is a cycle of period T. Then $E_s(x(t))$ is necessarily constant for $t = 0, \dots, T-1$. If $T > 2$ we have $x(2) \neq x(0)$, then from the definition of the Lyapunov functional $E_s(x(2)) < E_s(x(1))$, which is a contradiction. Hence the theorem.

Case 2: Next we consider the case when the blocks P and Q of the connection matrix A are both antisymmetric in nature and the updation rule is given by (3) with (4) with the self-dual condition given by:

$$b_i = \frac{1}{2} \sum_{j \in J_1} p_{ij} + \frac{1}{2} \sum_{j \in J_2} q_{ij} \tag{6}$$

Here also we consider the Lyapunov functional defined by (5). Now using the self dual condition defined by (6) the Lyapunov functional takes the form:

$$\begin{aligned} E_a(x(t)) &= -\sum_{i \in I} (2x_i(t) - 1) \sum_{j \in J} a_{ij} (2x_j(t-1) - 1) \\ &= -\sum_{i \in I_1} (2x_i(t) - 1) \sum_{j \in J_1} a_{ij} (2x_j(t-1) - 1) \\ &\quad - \sum_{i \in I_1} (2x_i(t) - 1) \sum_{j \in J_2} a_{ij} (2x_j(t-1) - 1) \\ &\quad - \sum_{i \in I_2} (2x_i(t) - 1) \sum_{j \in J_1} a_{ij} (2x_j(t-1) - 1) \\ &\quad - \sum_{i \in I_2} (2x_i(t) - 1) \sum_{j \in J_2} a_{ij} (2x_j(t-1) - 1) \end{aligned}$$

Since $a_{ij} = 0 \quad \forall i \in I_1, j \in J_1$ or $\forall i \in I_2, j \in J_2$.

We have:

$$\begin{aligned} E_a(x(t)) &= -\sum_{i \in I_1} (2x_i(t) - 1) \sum_{j \in J_2} a_{ij} (2x_j(t-1) - 1) \\ &\quad - \sum_{i \in I_2} (2x_i(t) - 1) \sum_{j \in J_1} a_{ij} (2x_j(t-1) - 1) \end{aligned}$$

Now we show that $E_a(x(t))$ is Lyapunov functional when the connection matrix A consists of both anisymmetric blocks with the synchronous updating rule:

$$\begin{aligned} \Delta_t E_a &= E_a(x(t)) - E_a(x(t-1)) \\ &= -2 \sum_{i \in I_1} (x_i(t) + x_i(t-2) - 1) \\ &\quad \sum_{j \in J_2} a_{ij} (2x_j(t-1) - 1) - 2 \sum_{i \in I_2} (x_i(t) \\ &\quad + x_i(t-2) - 1) \sum_{j \in J_1} a_{ij} (2x_j(t-1) - 1) \end{aligned}$$

Since $\sum_{j \in J} a_{ij} (x_j(t) - b_i) \neq 0 \forall i \in I$.

We have:

$$\Delta_t E_a = 0 \text{ iff } x_i(t) \neq x_i(t-2)$$

and

$$\Delta_t E_a < 0 \text{ iff } x_i(t) = x_i(t-2)$$

Hence we obtained the result.

Theorem 2: If $E_a(x(t))$ be the Lyapunov functional, then the orbit have a period $T = 4$.

Proof: Let $(x(0), x(1), \dots, x(T-1))$ is an orbit of period T of the given dynamics then $E_a(x(t)) = E_0$ is constant for $t = 0, \dots, T-1$. Hence $x_i(t+4) \neq x_i(t)$ for all t . Hence the orbit have a period $T = 4$.

Bound to transient length: Since $\{0, 1\}^n$ is a finite set, we can bound the Lyapunov functional E_s in the following way.

Let $\{x(t)/_{t=0, \dots, q-1}\}$ be a transient trajectory, that is:

$$x(0) \rightarrow x(1) \rightarrow \dots \rightarrow x(q-1) \rightarrow x(q) \rightarrow \dots$$

where, $q(x)$ is the first vector belonging to the steady state associated to the initial configuration $x(0)$. The transient length of the neural network is defined as the greatest of such lengths. Since in the transient phase, E_s is strictly decreasing there exists:

$$e = 1 \leq t \leq q-1 | \Delta_t E_s | > 0$$

and since:

$$\sum_{j \in J} a_{ij} x_j - b_i \neq 0 \text{ for any } x \in \{0, 1\}^n$$

We can define positive quantity:

$$e_i = \min | \sum_j a_{ij} x_j - b_i |; x \in \{0, 1\}^n \quad i = 1, 2, \dots, n$$

$$\|A\| = \sum_i \sum_j |a_{ij}|$$

Now:

$$E_s(x(t)) = \sum_{i \in I} (E_s(x(t)))_i = \sum_{i \in I_1} (E_s(x(t)))_i + \sum_{i \in I_2} (E_s(x(t)))_i$$

Hence we can found the lower bound as follows:

$$\begin{aligned} (E_s(x(t)))_i &\geq \sum_{j \in J_2} |p_{ij}| - 2|2b_{i_1} - \sum_{j \in J_2} p_{ij}| \\ &+ \sum_{j \in J_1} |q_{ij}| - 2|2b_{i_2} - \sum_{j \in J_1} q_{ij}| \end{aligned}$$

Where:

$$i_1 = \begin{cases} i: & i \in J_1 \\ 0 & \text{otherwise} \end{cases}$$

$$i_2 = \begin{cases} i: & i \in J_2 \\ 0 & \text{otherwise} \end{cases}$$

$$E_s(x(t)) \geq \|P\| - 2\|2b_1 - P\hat{1}\| - \|Q\| - 2\|2b_2 - Q\hat{1}\|$$

Where:

$$\|P\| = \sum_i \sum_j |p_{ij}|$$

$\hat{1} = (1, 1, \dots, 1)$ is a constant vector

To get the upper bound as write $(E_s(x(t)))_i$ as follows:

$$(E_s(x(t)))_i \leq -2e_i + |2b_{i_1} - \sum_{j \in J_2} p_{ij}| + |2b_{i_2} - \sum_{j \in J_1} q_{ij}|$$

Where:

$$i_1 = \begin{cases} i: & i \in J_1 \\ 0 & \text{otherwise} \end{cases}$$

$$i_2 = \begin{cases} i: & i \in J_2 \\ 0 & \text{otherwise} \end{cases}$$

$$E_s(x(t)) \leq -2 \sum_{i \in I} e_i + \|2b_1 - P\hat{1}\| + \|2b_2 - Q\hat{1}\|$$

Since in the transient phase the Lyapunov functional E_s is decreasing:

$$E_s(x(t)) \leq E_s(x(t-1)) - e \leq E_s(x(0)) - te$$

Hence for $t = q$:

$$\begin{aligned}
 E_s(x(q)) &\leq E_s(x(0)) - qe \\
 \Rightarrow \|P\| - 2\|2b_1 - P\hat{1}\| &= \|Q\| - 2\|2b_2 - Q\hat{1}\| \\
 &\leq E_s(x(0)) - qe \\
 \Rightarrow qe &\leq -2\sum e_i + 3(\|2b_1 - P\hat{1}\| + \|2b_2 - Q\hat{1}\|) \\
 &+ \|P\| \|Q\| \\
 \therefore q &\leq (1/e)(-2\sum e_i + 3(\|2b_1 - P\hat{1}\| + \|2b_2 - Q\hat{1}\|) \\
 &+ \|P\| + \|Q\|)
 \end{aligned}$$

DISCUSSION

Let:

The connection matrix $A = \begin{bmatrix} 0 & 0 & 0 & 3 & 4 \\ 0 & 0 & 3 & 0 & 1 \\ 0 & 0 & 4 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \end{bmatrix}$

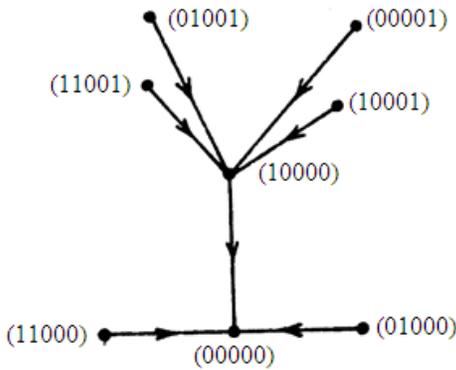


Fig. 1a: Convergent dynamics (00000) as fixed point when both P and Q are symmetric in nature

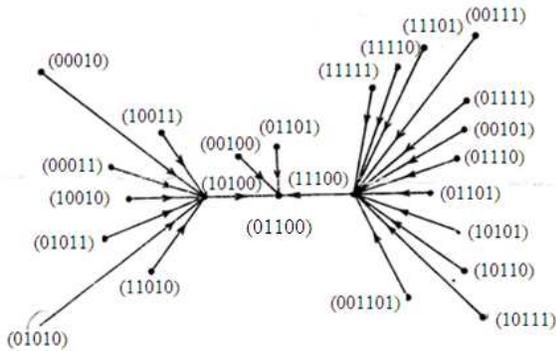


Fig. 1b: Convergent dynamics (01100) as fixed point when both P and Q are symmetric in nature

with threshold vector $b = (2,3,1,4,1,5,3)$. The attractor set consist of two fixed points namely (00000) and (011001) (Fig. 1a and b):

Let:

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & -1 & 0 & 3 \\ 0 & 0 & -2 & -3 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

with threshold vector $b = (1.5,0.5,-2.5,0.5,-0.5)$.

The attractor here is unique limit cycle of length four for every initial configurations (Fig. 2). The limit cycle is (00001) → (11101) → (11110) → (00010).

If condition (6) is violated the nature of attractor is also changed (Fig. 3).

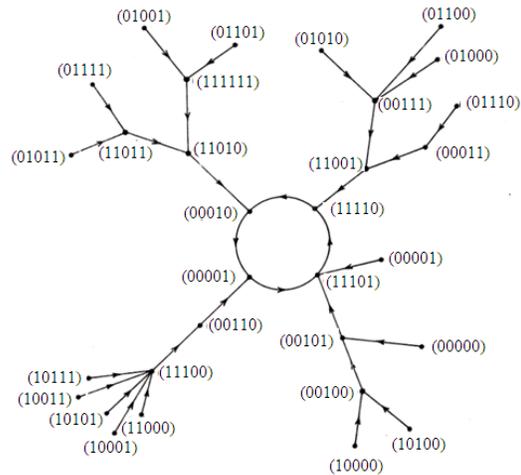


Fig. 2: Limit cycle with period 4 when condition 6 is satisfied

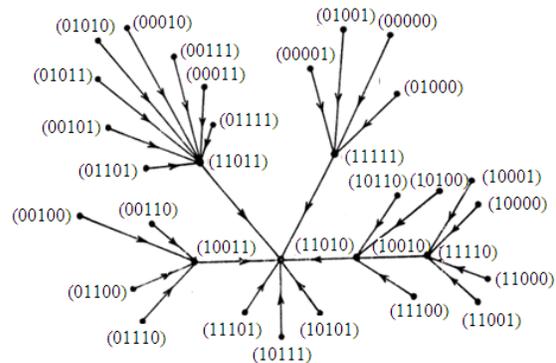


Fig. 3: Fixed point attractor the condition 6 is violated

CONCLUSION

It was shown that under synchronous updating rule such networks converge to a fixed point or to a limit cycle of period 2 or 4. The bound of transient length was discussed. Conclusions were drawn from the simulation studies carried out in support of the results.

REFERENCES

- Amari, S.I., 1972. Learning patterns and pattern sequence by self organizing nets out of threshold elements. *IEEE Trans. Comput.*, 21: 1197-1206. DOI: 10.1109/T-C.1972.223477
- Caianiello, E.R., 1961. Outline of a theory of thought-process and thinking machines. *J. Theor. Biol.*, 1: 204-235.
<http://www.ncbi.nlm.nih.gov/pubmed/13689819>
- Goles, E. and S. Matizer, 1990. *Neural and Automata Networks: Dynamical Behavior and Application*. Kluwer Academic Press, ISBN: 0-7923-0632-5.
- Hoffman, R.E., 1987. Computer simulation of neural information processing and the schizophrenia-mania dichotomy. *Arch. Gen. Psychiat.*, 44: 178-188.
<http://www.ncbi.nlm.nih.gov/pubmed/3813815>
- Hopfield, J.J., 1982. Neural networks and Physical System with emergent collective computational abilities. *Proc. Natl. Acad. Sci., USA.*, 79: 2554-2558.
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC346238/pdf/pnas00447-0135.pdf>
- Little, W.A., 1974. Existence of persistence states in brain. *Math. Biosci.*, 19: 101-120.
<http://www.ams.org/tran/1998-350-12/S0002-9947-98-02083-2/S0002-9947-98-02083-2.pdf>
- Little, W.A. and G.L. Shaw, 1978. Analytical study of the memory storage capacity of a neural network. *Math. Biosci.*, 39: 281-290.
<http://www.ams.org/tran/1998-350-12/S0002-9947-98-02083-2/S0002-9947-98-02083-2.pdf>
- McCulloch, W.S. and W. Pitts, 1943. A logical calculus of the ideas immanent in nervous activity. *Bull. Math Biol.*, 52: 99-115.
<http://www.springerlink.com/content/xw16676748262521/>