Proceedings of the 1999 IEEE
International Symposium on Computer Aided Control System Design
Kohala Coast-Island of Hawai'i, Hawai'i, USA • August 22-27, 1999

WA1-6 11:40

Lotka-Volterra Machine for a General Model of Complex Biological Systems

Masayuki Hirafuji, Kei Tanaka, Scott Hagan
Computational Modeling Laboratory, National Agriculture Research Center
Tsukuba 305-8666 Japan

Abstract

It has been proven that the generalized Lotka-Volterra equations can express arbitrary continuous functions if there are sufficient hidden nodes that are not assigned by observational data. We show that the generalized Lotka-Volterra equations encompass a description of the same structure as do artificial neural networks. Since the equations can be used as a universal modeling framework, we named the computational models of the equations Lotka-Volterra Machines (LVMs). We developed a software tool (Java applet) to construct LVMs and identified unknown parameters in the LVM using a genetic algorithm. It is demonstrated that the LVM can model the growth of organs in a plant and the tool can identify the parameters of the LVM although it may have more than 20 unknown parameters, and that the resulting LVM can correctly forecast the dynamics of organs. Originally the Lotka-Volterra equation was proposed to model ecosystems. We show that the LVM can be transformed into the equations for a hypercycle, a model of chemical processes essential to life. Therefore the LVM has the ability to represent the dynamics at almost all levels of hierarchical structure in complex biological systems. We developed a large LVM simulator (Java applet) to simulate a much larger LVM. Large LVMs on the order of hundreds of nodes show a different character from that of small LVMs. Specific aspects of the character of huge LVMs are discussed from examples of simulations. We propose a self-organizing mechanism according to which a sufficient number of nodes, interconnected by diverse weights in a huge LVM, can maintain their order against fluctuations coming from all other nodes.

1. Introduction

To model physical phenomena, there exist fundamental equations like the Navier-Stokes equation, Maxwell's equations, the Schrödinger equation and so on. Theoretically any physical system can be simulated only by solving such equations. On the other hand simple equations of this sort are not available for biological systems. Artificial neural networks (ANNs) however might be able to provide a framework for such general equations, since some ANNs are able to approximate arbitrary functions in order to represent patterns in complex phenomena. The next phase of research could thus evolve toward algorithms for learning and technologies for neuro-chips.

Similarly, it is most important that general models of complex biological systems, like those governing the dynamics of ecosystems and the growth of individuals, should be able to approximate arbitrary functions. General models that encompass the functioning of neural networks are desirable.

Statistical models have in fact been employed as simple models of biological systems. However statistical modeling is too general as a tool to probe the functioning of complex biological systems. We therefore propose generalized Lotka-Volterra equations as a modeling method more universal and realistic than either neural networks or statistical methods. This method can approximate arbitrary positive continuous functions, and

0-7803-5500-8/99 $10.00

the models comprise the characters of both ecosystems and artificial neural networks. Moreover it has the advantage that hardware implementation (analog circuit) is readily designed. We dubbed it the Lotka-Volterra Machine (LVM). We applied the LVM to the construction of plant growth models and a model of general complex biological systems.

2. LVM: Lotka-Volterra Machine

2.1 Generalized LVM

The generalized Lotka-Volterra equation is

$$\frac{dg_i}{dt} = g_i\left(r_i - \sum_{j=1}^{n} \mu_{ij} g_j\right) \quad (i = 1,2,\cdots,n) \qquad (1)$$

where g_i is abundance such as population or biomes present, μ_{ij} is an interference coefficient between node i and node j, r_i is the intrinsic reproductive rate and n is the number of nodes. Although this equation is nonlinear, it has remarkable advantages that derive from its simplicity and symmetry as well as its universality with respect to software/hardware implementation, a feature that will be demonstrated later.

Eq. (1) subsumes the Lotka-Volterra equation for the population dynamics of two individuals[1], given by

$$\begin{cases} \dfrac{dg_1}{dt} = g_1(r_1 - w_{12}g_2) \\ \dfrac{dg_2}{dt} = g_2(r_2 - w_{21}g_1) \end{cases} \qquad (2)$$

Assume that the parameters in Eq. (2) are

$$\begin{cases} r_1 = w, \quad r_2 = -w \\ w_{12} = w, \quad w_{21} = -w \end{cases} \qquad (3)$$

and that $g_1 g_2$ is small. Then we can derive the approximate solutions,

$$\begin{cases} g_1 = 1 + \alpha\cos(wt + \beta) \\ g_2 = 1 + \alpha\sin(wt + \beta) \end{cases} \qquad (4)$$

where α and β are constants.

If the LVM has n coupled nodes whose parameters, w, are $\varpi, 2\varpi, \cdots, n\varpi$, the LVM synthesizes $2n$ weighted cyclic signals from the outputs of the nodes. This synthesis in the LVM has the same character as a Fourier series (Fig. 1). This means that the LVM can express arbitrary continuous functions if there is a sufficient number of hidden nodes unassigned by the observational data (observational data fixes the input/output nodes). If we can identify the unknown parameters in an LVM comprising several hidden nodes with the observed data of a given complex system, the LVM will constitute a quantitative model of that complex system.

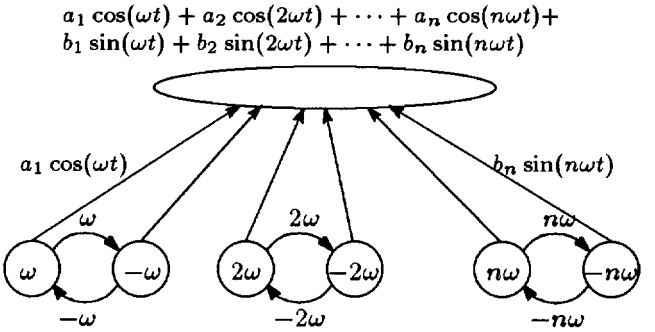

Fig. 1 LVM with oscillating coupled nodes

2.2 Simplified LVM

A simplified LVM encompasses the operation of artificial neural networks like the multi-layered perceptron (e.g. back-propagation model[2]) or the interconnected network.(e.g. Hopfield model[3]).

It has been proven[4] that there exists a Lyapunov function and a global equilibrium point for Eq. (1) under

$$\begin{cases} \mu_{ii} > 0 \\ \mu_{ii} > \mu_{ij} \quad (i \neq j) \end{cases} \qquad (5)$$

We assume the additional and stronger restriction,

$$|\mu_{ii}| >> |\mu_{ij}| \quad (i \neq j) \qquad (6)$$

and rewrite Eq. (1) as

$$\begin{cases} \dfrac{dg_i}{dt} = g_i\left(r_i - z_i - \mu_{ii} g_i\right) \\ z_i = \displaystyle\sum_{j(\neq i)} \mu_{ij} g_j \end{cases} \tag{7}$$

Assuming that z_i is approximately constant over short times, the integrated solution of Eq. (7) is

$$\begin{cases} g_i = \dfrac{1}{1+\exp\left(-p_i t + \theta_i\right)} \\ p_i = r_i - z_i \end{cases} \tag{8}$$

where θ_i is an integration constant. Unknown parameters can be estimated by recurrent neural network learning algorithms[5,6] (t must be constant) or, with empirical data, a special algorithm[7] adapted to the simplified LVM (t may be variable). We developed a soybean growth model using this simplified LVM in which the estimated μ_{ij} were correlated with nutrient flow within the plant body[8]. Both the simplified and the generalized LVM can model neural networks, however the scope of application for the simplified LVM is limited by the restrictions, Eq. (5), (6).

2.3 Identification algorithms for generalized LVM

We employed a genetic algorithm (GA) as a means to estimate the unknown parameters in a generalized LVM. The evaluation function was the distance between estimated values and the observed data. A LVM consisting of n nodes has $n^2 + n$ unknown parameters. If we employ l bits to represent a parameter, the gene size of the LVM is $l(n^2 + n)$, however some parameters can be eliminated by *a priori* restrictions (knowledge).

Since there are so many options, a GUI tool to identify the parameters of a LVM is indispensable. We developed a program (Java applet) for this purpose and can construct a LVM using this tool by interactively choosing options (coding methods: binary code / Gray code, bit size for coding, mutation rate, crossover, data set and so on).

2.4 Applications of a generalized LVM

Plant growth depends on the weather, soil, varieties and so

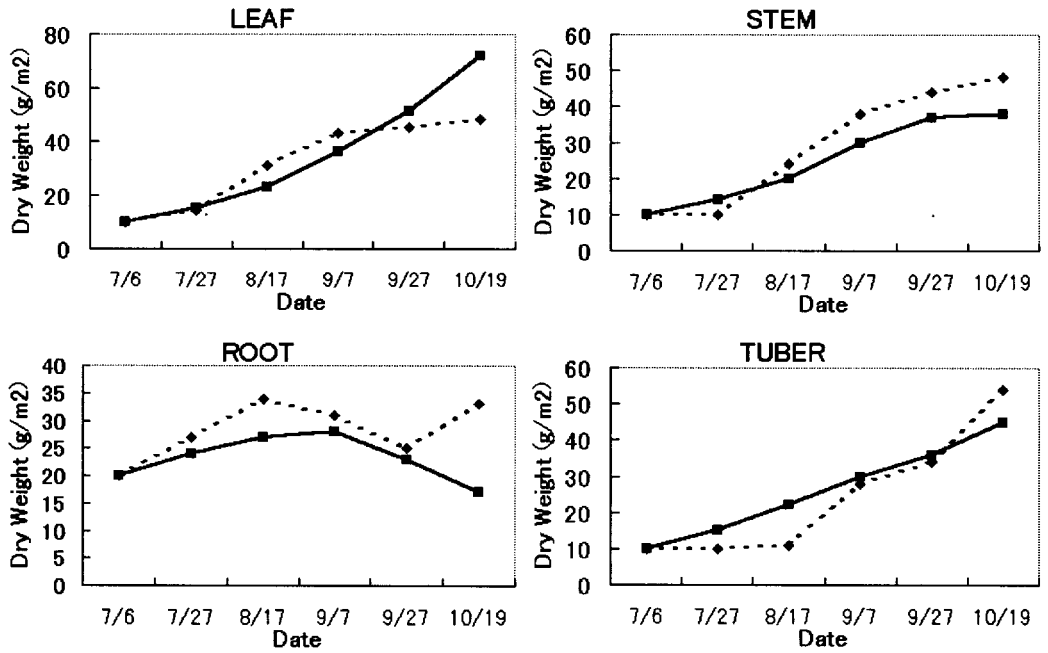

Fig. 2 A comparison of the actual (dotted lines) organ dynamics of sweet potatoes observed in 1982 with the predictions (solid lines) of the model using weather data from the same year. Parameters of the model were estimated using meteorological and observed dry weight data from the period between 1970 and 1980.

on. Farmers and consultants need models specific to their situations and each model must be revised year by year according to new observational data. We constructed 13 types of plant growth models (soybean:6, wheat:4, sweet potato:3) using the LVM construction tool. The predictions of these models are consistent with observed data as shown in Fig. 2.

The plant growth model has m nodes for organs and e nodes for environmental factors (Fig. 3). Environmental nodes are assigned to observed weather data so there need be no connection weights from the organ nodes to environment nodes. The plant growth model contains $m^2 + m + e$ unknown parameters in total. Specifically for the sweet potato model shown in Fig. 3, m was 4 (leaf, stem, root, tuber), e was 2 (temperature, solar radiation) and l was 8 bits. In total, 178 bits were identified in the model.

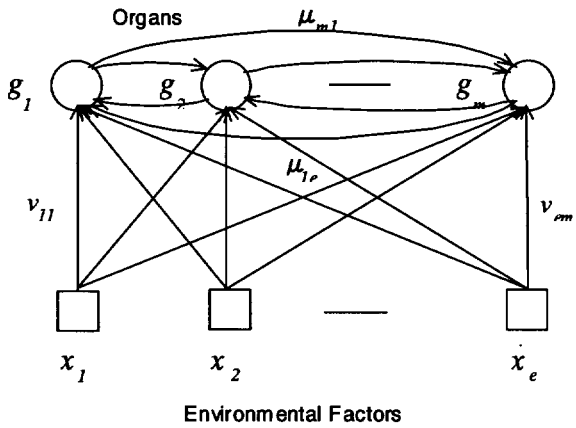

Fig. 3 A plant growth model using the LVM. Organs grow competitively under restrictions imposed on resources such as nutrients, water and space.

3. General model of complex biological systems

The Lotka-Volterra equations were originally conceived as a model for ecosystems and individual organisms such as bacteria and cells. They can also be applied to model organs as has been shown.

One of the fundamental features of complex biological systems is a chemical process, discussed by Eigen[9], known as a hypercycle A hypercycle evolves by autocatalytic reproduction, represented by

$$\frac{dy_i}{dt} = k_i\, y_i\, y_{i-1} \qquad (i = 1,2,\cdots) \tag{9}$$

where y_i is the concentration of polynucleotides and k_i is constant. This equation can be interpreted as a special form of Eq. (1), when

$$\begin{cases} r_i = 0 \\ \mu_{ij} = k_i \ \ (j = i-1) \\ \mu_{ij} = 0 \ \ (j \neq i-1) \end{cases} \tag{10}$$

After all, we can construct models for the dynamics of phenomena from chemical processes to ecosystems, and we can combine them into an all-encompassing model in the framework of the LVM. From the standpoint of system dynamics modeling, the LVM constitutes a general model for complex biological systems (Fig. 4).

For really general models, we should in fact take into account phenomena occurring at the next smallest scale, that is, at the level of giant molecular systems. At this scale we cannot ignore quantum mechanical effects. Although this is an important issue, we will not touch on it here since the LVM cannot naturally represent the Schrödinger equation. We have investigated such quantum-macroscopic complex systems in the context of a simplified two-layer model[10].

Even restricted to the domain of classical mechanics, the generalized model may exhibit interesting phenomena such as self-organization. The model is heterogeneous and the ranges of its parameters extremely broad. Traditionally, complex systems have been modeled by stochastic dynamics or chaotic systems. In such modeling, assumptions about the random nature of node dynamics or restrictions to a small number of degrees of freedom often apply. The LVM for the general model requires an enormously large number of degrees of freedom, and the

structure (network of connections) must be extremely heterogeneous, since real biological systems form a hierarchy which occasionally makes connections between structures at different scales. Are there any specific characters that might be recognized in the heterogeneous structure of vast complex systems? In order to utilize the universal nature of the LVM towards practical applications, we need to investigate the character of such systems.

We have developed a huge LVM simulator as a Java applet to investigate these problems. The only function of this LVM simulator is to simulate Eq. (1). We identified several interesting aspects of the character of this large LVM (Fig. 5). Case (c) in Fig.3 indicates that uniform (or fully random) weak connections offset interference from connected nodes in large complex systems, and if several nodes have diverse interconnections amongst themselves and uniform connections with all other nodes, the nodes

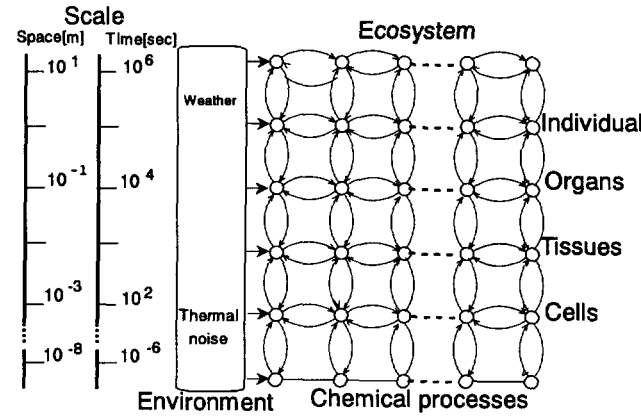

Fig4. A general model of complex biological systems

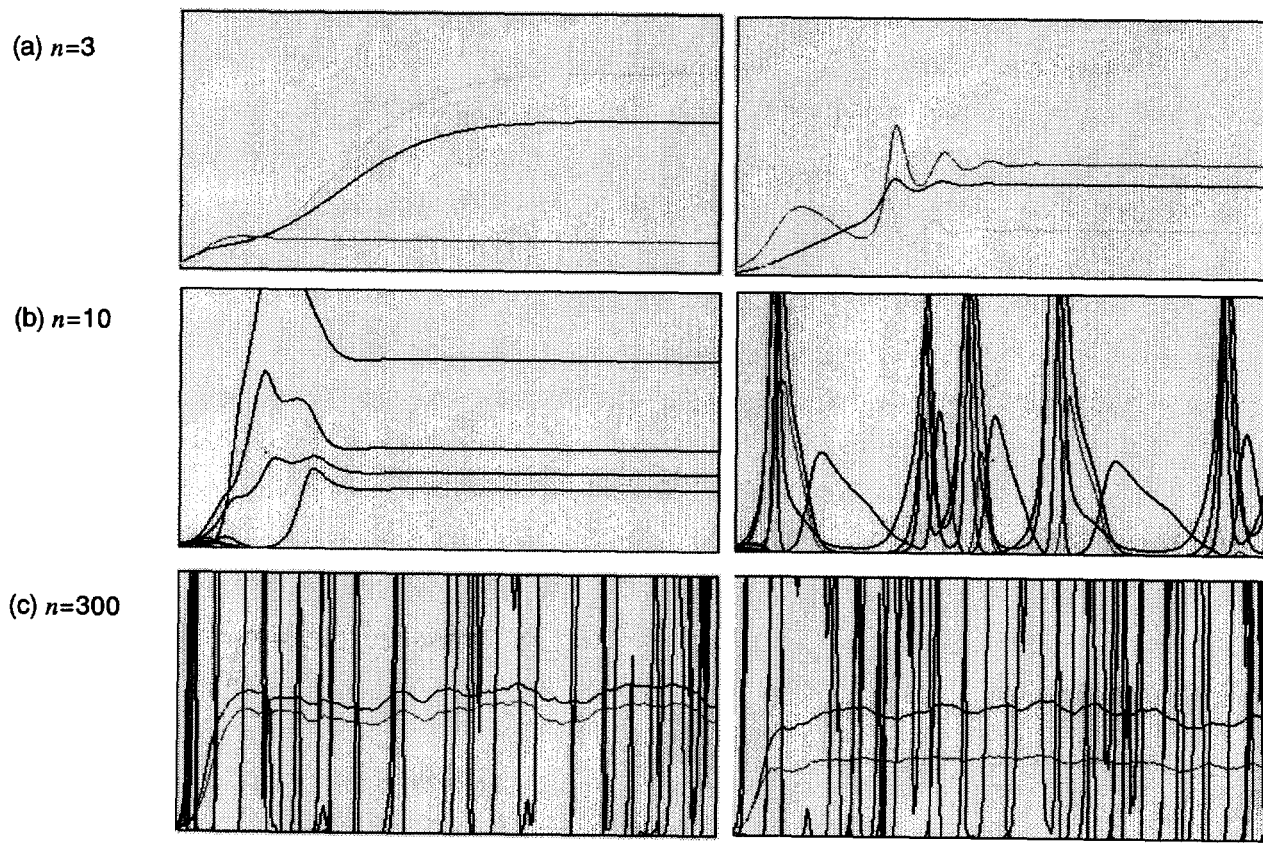

Fig. 5 Typical complex output patterns for three types of LVM. n is the number of nodes in the LVM. The 300-node LVM shown in (c) has three special nodes which have diverse interconnections amongst themselves and uniform (negative) connections with all other nodes. The three nodes are frequently able to maintain their trajectories (stable states) under fluctuations.

can maintain a stable trajectory against fluctuations. This ordering mechanism would be more conspicuous in larger LVMs. When fluctuations are sufficiently weak, the response to fluctuations changes to an enhancement of transitions in nonlinear systems, a phenomenon known as stochastic resonance. Therefore we think that there exists a critical size of the LVM at which the hidden order with respect to fluctuation effects will become evident. This ordering mechanism may account for the relevance of simple legacy models for complex biological systems. For example, legacy yield prediction models and growth models are too simple nowadays. However if these models caught the order of complex systems, their predictions can be more accurate than our expecting accuracy.

4. Discussion -Toward implementation of a huge LVM

The LVM simulator that we have developed can construct thousands of node models. The limit depends on RAM capacity and the CPU's execution speed. By the above discussions, we predict that much larger LVMs with diverse connections will show interesting behaviors. If we try to combine, in the framework of the LVM, quantitative biological models of hierarchically arranged phenomena across levels such as the ecosystem, individuals, organs, cells and self-catalytic reproduction, we will need much larger LVMs. Ideally we want to construct a LVM with billions of nodes to realistically model complex biological systems.

We can propose two ways to construct a billion-node LVM. The first strategy is meta-computing and utilizes the power of PCs connected to the internet by delivering partial threads (tasks) of the LVM simulator. This is the reason we developed the LVM simulator as a Java applet. The second way is a hardware implementation using analog circuits. Fortunately Eq. (1) can be implemented as an analog circuit using only standard analog devices: analog multipliers, amplifiers and low-pass filters. However the number of interconnecting wires in the circuit increases as $O(n^2)$. If we tried to develop a hardware implementation of the LVM with 10^9 nodes, it would need 10^{18} connections. This wiring problem can be solved by multiple frequency coding.

We are attempting both strategies to realize a huge LVM.

References

1) Lotka, A. J., 1925, *Elements of Physical Biology*, Williams & Wilkins, Baltimore.
2) Rumelhart, D.E., Hinton, G.E. and Williams, R.J., 1986, Learning Internal Representations by Error Propagation, in *Parallel Distributed Processing* (Eds. by McClelland,J.L, Rumelhart, D.E. and the PDP Research Group), Vol.1, 354-361, MIT Press.
3) Hopfield, J.J., Tank, D.W., 1986, Computing with Neural Circuits: A Model, *Science*, Vol.233, pp. 625-633.
4) Shigesada, N., Kawasaki, K. and Teramoto, E., 1984, The effects of interference competition on stability, structure and invasion of a multi-species system, *Journal of Mathematical Biology,* **21**, pp. 97-113.
5) Pineda, F.J., 1987, Generalization of Back-Propagation to Recurrent Neural Networks. *Physical Review Letters*, **59**, pp. 2229-2232.
6) Pearlmutter, B.A., 1989, Learning State Space Trajectories in Recurrent Neural Networks, *Proceedings of IJCNN 1989* in Washington D.C., **II**, pp. 365-372.
7) Hirafuji, M., 1991, A Time-continuous Quantitative Plant Growth Model Based on Neural Networks and Lotka-Volterra Equations, *Proc. IFAC/ISHS 1st Workshop on Mathematical and Control Applications in Agriculture and Horticulture*, IFAC Workshop series, Vol.1, Pergamon Press, Oxford, pp. 289-294.
8) Hirafuji, M., 1992, A Plant Growth Model Based on Interference between Organs and the Method of Identification, *J. of Agricultural Meteorology*, **8**, 3, pp. 285-293.
9) Eigen, M. and Schuster, P., 1979, *The hypercycle: A principle of natural self-organization*, Berlin-Heidelberg, Springer.
10) Hirafuji, M. and Hagan, S., 1999, Quantum Tunneling Evolution: A Model of Life as a Global Optimization Process, *Proc. of the Fourth International Symposium on ARTIFICIAL LIFE AND ROBOTICS*, pp. 476-479.