Dynamics of Sequential Decision Making

Mikhail I. Rabinovich* and Ramón Huerta†
Institute for Nonlinear Science, University of California, San Diego, 9500 Gilman Drive 0402, La Jolla, California 92093, USA‡

Valentin Afraimovich
Instituto de Investigación en Comunicación Óptica, UASLP, Avenida Obregón 64, 78000 San Luis de Potosí, SLP, Mexico
(Received 4 May 2006; published 1 November 2006)

We suggest a new paradigm for intelligent decision-making suitable for dynamical sequential activity of animals or artificial autonomous devices that depends on the characteristics of the internal and external world. To do it we introduce a new class of dynamical models that are described by ordinary differential equations with a finite number of possibilities at the decision points, and also include rules solving this uncertainty. Our approach is based on the competition between possible cognitive states using their stable transient dynamics. The model controls the order of choosing successive steps of a sequential activity according to the environment and decision-making criteria. Two strategies (high-risk and risk-aversion conditions) that move the system out of an erratic environment are analyzed.

DOI: 10.1103/PhysRevLett.97.188103 PACS numbers: 87.18.Sn, 05.45.−a, 87.18.Bb

Introduction.—Information-dependent transient activity is the most typical behavior of animals and autonomous intelligent systems [1]. Even in a stationary environment such behavior may not be unique and the brain or cognitive-state machine (CSM) has to make choices; i.e., the behavior is a series of switching or decision-making (DM) procedures [2] (see Fig. 1). It is evident that intelligent decisions in a sequential behavior have to be stable against noise and reproducible to allow memorization and reuse of successful decision sequences in the future. On the other hand, it also has to be sensitive to new information from the environment. These requirements are fundamentally contradictory, and existing approaches [3,4] are not sufficient to explain DM for sequential activity. Here, we formulate a new class of models suitable for analyzing sequential DM based on a generalized winnerless competition (WLC) principle [5].

Model equations.—Decision-making systems consist of subsystems that (i) formulate a goal, (ii) create a decision function, (iii) control the parameters of a CSM, and (iv) are responsible for the generation of spatiotemporal patterns of cognitive states that control the behavior according to incoming information I and DM rules. Here we focused on a dynamical model of the CSM and its parameters that are controlled by a decision function and information about the world, i.e., on items (iii) and (iv) above.

Let us consider a system aiming to realize the maximum possible number of sequential decisions. We assume that life courses can be coded as sequences of events [6], i.e., decision-making events in our case. For the purposes of this Letter we will refer to the number of decisions taken throughout a sequence as the "length of life." The decision-making functions are defined algorithmically (see below). Our model includes ordinary differential equations for the dynamics of cognitive states and equations for control parameters given by DM rules:

$$\dot{a}_i = a_i\left[\sigma_i(I,t) - \left(a_i + \sum_{j\neq i}^{N} \rho_{ij} a_j\right)\right] + \eta_i(t), \qquad (1)$$

$$\tau\dot{\sigma}_i = -\frac{\partial U_i(\sigma_i, I)}{\partial \sigma_i}. \qquad (2)$$

Equation (1) is of the Lotka-Volterra type and models the competitive dynamics of cognitive states $a_i(t)$ (it can be brain modes or competing controllers; see, for example, [7]), where ρ_{ij} denotes the strength of the competitive interaction from the state j to state i, which is based on genetic and memorized information, η_i is external noise, and N is the number of possible cognitive states. The working regime of the CSM is a stimulus dependent competition without winner until the system reaches the "end of life" (a stable equilibrium). Before the CSM reaches this

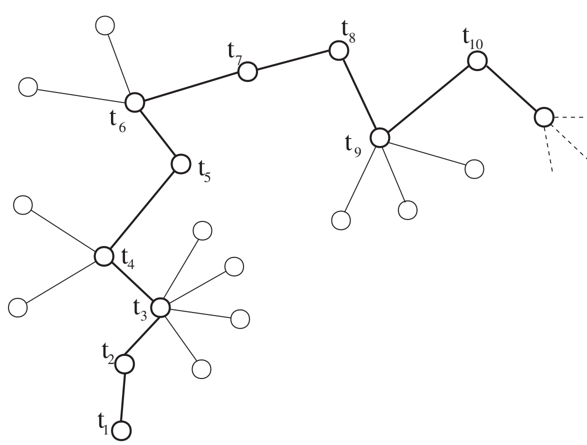

FIG. 1. A sequence of cognitive states: thin lines, possible paths; thick line, realized sequence chosen by the DM according to information fro the environment; t_3, t_4, t_6, t_9 are instants of choices.

 © 2006 The American Physical Society

point the different cognitive states become "winners" just for a short time. This is the WLC principle for the reproducible transient dynamics of neural systems that is ensured by nonsymmetric inhibition (see [5]). The parameters $\sigma_i(t, I)$ control the cognitive-state dynamics and are governed by the independent gradient system (2). Given i, the potential function $U_i(I, \sigma_i)$ has m_k minima

$$\bar{\sigma}_i(I) = \sigma_i^0 + A_i^s(I), \qquad s \in \{1, \ldots, m_k\}, \tag{3}$$

The initial value $\bar{\sigma}_i(t_0)$ contains no memory of the previous history and is determined by the decision-making rule in such a way that it chooses the basin of the only minimum. $A_i^s(I)$ represent the repertoire of actions determined by external stimulation, and σ_i^0 is a constant term common to all stimulus responses. In the following we suppose that the characteristic time τ is so small that the dynamics for σ_i can be neglected; i.e., σ_i take only stable equilibrium values, $\bar{\sigma}_i(I)$. Even in a continuously changing environment, i.e., if I is continuously changing, the σ_i can be changed in a discrete way because perception can be discrete (see, e.g., [6]). Hence, we may assume that the stimulus I acts in such a way that at the instants of choice $t_k = 1, 2, \ldots$, the parameters $\bar{\sigma}_i$ are not unique and may take several values from (3). The times t_k are defined as the instants of arrival of the system to a neighborhood of the fixed points of the system (1), and the number of possibilities m_k and the values of $\bar{\sigma}_i$ depend on the stimulus I. Between the instants of choice, the system (1) evolves according to the values of $\bar{\sigma}_i$ chosen from (3) at time t_k.

The mathematical image of the transient dynamics of cognitive states $a_i(t)$ is a stable heteroclinic sequence (SHS) with steps that are to be chosen according to the DM rule. What happens after the instant of decision? As shown in [8], the system (1) has nontrivial equilibria $S_i = (0, \ldots, 0, \bar{\sigma}_i, 0, \ldots, 0)$ in the absence of noise and $\sigma_i = \bar{\sigma}_i$. The eigenvalues of the linearized at S_i system are $\lambda_{ji} = \bar{\sigma}_j - \rho_{ji}\bar{\sigma}_i$, $j = 1, \ldots, N$, $j \neq i$. Depending on the values of λ_{ji} we can find the following possibilities: (i) If all $\lambda_{ji} < 0$, S_i is a stable fixed point; we say the system reaches its end of life. (ii) If there are at least two values of j, say, j_1 and j_2, such that $\lambda_{j_1 i} > 0$ and $\lambda_{j_2 i} > 0$, we call S_i the "panic state." The system at this point has an infinite numbers of choices (for heteroclinic orbits in such a situation, see [9]). (iii) If there is only one value $j = j_0$ such that $\lambda_{j_0 i} > 0$ and all other eigenvalues are negative, the saddle S_i has an one-dimensional unstable manifold. We will consider only dissipative saddles. Dissipative saddles satisfy the following assumption. Let $\lambda_i^- = \max\{\max_{j \neq j_0, i}\{\lambda_{ji}\}; -\bar{\sigma}_i\}$; then the number $\nu_i = -\lambda_i^-/\lambda_{j_0 i}$ is called the saddle value [10]. The saddle is dissipative if $\nu_i > 1$. If so, we call S_i the "transient state," and then life continues.

A SHS, say, Γ, is a collection of saddles $\{S_{i_k}\}$, $k = 1, \ldots, K$, with a one-dimensional unstable manifold together with a collection of heteroclinic orbits $\{\Gamma_{k,k+1}\}$ such that $\Gamma_{k,k+1}$ connects S_{i_k} and $S_{i_{k+1}}$: $\Gamma = \cup_{k=1}^{K-1}(S_{i_k} \cup \Gamma_{k,k+1})$. Accordingly, Γ must be stable: if an initial point of an orbit lies in a neighborhood of S_{i_1}, then all points of the piece of the orbit will belong to a vicinity of Γ until the arrival time to a neighborhood of S_{i_K}.

For the existence and stability of a SHS it is sufficient to fulfill the following conditions: (i) Every saddle $\{S_{i_k}\}$ is a transient state; we denote the corresponding positive eigenvalue by $\lambda_{i_{k+1}, i_k} = \bar{\sigma}_{i_k} - \rho_{i_{k+1} i_k}\bar{\sigma}_{i_k}$. (ii) The inequalities

$$1 - \rho_{i_{k-1} i_k} \cdot \rho_{i_k i_{k-1}} \neq 0, \tag{4}$$

$$-\bar{\sigma}_{i_k} < \bar{\sigma}_{i_{k-1}} - \rho_{i_{k-1} i_k}\bar{\sigma}_{i_k}, \tag{5}$$

$$\bar{\sigma}_i - \rho_{i i_k}\bar{\sigma}_{i_k} < \bar{\sigma}_{i_{k-1}} - \rho_{i_{k-1} i_k}\bar{\sigma}_{i_k}, \qquad i \notin \{i_{k-1}, i_k, i_{k+1}\} \tag{6}$$

are satisfied (see [8]).

Suppose that an initial condition of the system (1) is placed in the vicinity of the a_{i_1} axis, and assume that there is an integer $m_1 > 0$ with possible values of vectors $\bar{\sigma} = (\bar{\sigma}_1, \ldots, \bar{\sigma}_N)$: $\bar{\sigma}^1, \ldots, \bar{\sigma}^{m_1}$. Before making a decision among them, the system eliminates the following cases based on the intrinsic stimuli: (a) For each value $s \in \{1, \ldots, m_1\}$, the corresponding point S_i is a stable fixed point. (b) For each value $s \in \{1, \ldots, m_1\}$, the point S_i is the panic state. Some trajectories that are passing near such a saddle may be relevant for the formulated goal. However, if the CSM is trying to memorize this specific behavior for future use, it encounters the problem that the unstable trajectories in the vicinity of the panic state are divergent, and as a result, the system moves in different directions in repeated trials (nonreproducibility). Thus a CSM that uses panic states behaves erratically and has a low probability of survival compared with a CSM based on the transient states. (c) There are values of $s \in \{1, \ldots, m_1\}$ for which S_i is a saddle with a one-dimensional unstable manifold, but all of these saddles are nondissipative, i.e., the saddle value $\nu \leq 1$. Case (c) is excluded because it can lead to instability of the sequential behavior, and the dynamics cannot be reproducible [as in case (b)] [11].

Now we assume that there is at least one value of $s \in \{1, \ldots, m_1\}$, say, $s = s'$, such that the corresponding point $S_{i_1}^{s'}$ is a transient state. If such a value is unique, we choose $\bar{\sigma} = \bar{\sigma}^{s'}$, substitute it into (1), and allow the system to evolve. Since the initial point is close to the a_{i_1} axis, the point on the corresponding trajectory comes to a small neighborhood of $S_{i_1}^{s'}$, and because $S_{i_1}^{s'}$ is dissipative, it will follow the heteroclinic orbit joining $S_{i_1}^{s'}$ and $S_{j_0} = (0, \ldots, 0, \bar{\sigma}_{j_0}^{s'}, 0, \ldots, 0)$ on the plane $\cap_{k \neq i_0, j_0}\{a_k = 0\}$ [8]. Now we consider the main case where there are several saddles $S_{i_1}^{s(q)}$, $q = 1, \ldots, p$.

Decision-making functions.—DM evidently depends on the goal. Let us focus on the goal formulated above with two extreme strategies often used by animals to survive [12]. It can be, for example, a risk-aversion DM (stability requirement) or a high-risk DM (minimal time to reach the next decision point).

High-risk DM.—Every saddle $S_{i_1}^{s(q)}$ has only one positive increment $\lambda_{j_0 i_1} = \bar{\sigma}_{j_0} - \rho_{j_0 i_1}\bar{\sigma}_{i_1}$, $j_0 = j_0(q)$, $q = 1, \ldots, p$. We choose q_0 in such a way that

$$0 < \lambda_{j_0(q)i_1} < \lambda_{j_0(q_0)i_1}, \qquad q \neq q_0. \tag{7}$$

In other words, we choose the maximal increment which corresponds to the fastest motion away from saddle S_{i_1}, and therefore to the shortest time for reaching the next saddle in the SHS.

Risk-aversion DM.—Another possibility to make a choice is based on stability considerations. For every $q = 1, \ldots, p$, the corresponding saddle value $\nu_{i_1}^q$ is well defined. We choose q_0 in such a way that

$$1 < \nu_{i_1}^q < \nu_{i_1}^{q_0}, \qquad q \neq q_0. \tag{8}$$

After making a decision the system replaces the corresponding value of $\bar{\sigma} = \bar{\sigma}^{s(q_0)}$ in Eq. (1) and evolves on an orbit close to a heteroclinic until it reaches a neighborhood of the saddle $(0, \ldots, 0, \bar{\sigma}_{j_0(q_0)}, 0, \ldots, 0)$. If this point is fixed as the initial point for the next stage of the process, we denote $j_0(q_0)$ by i_2, taking into account a number m_2 of different vectors $\bar{\sigma}$ and their values. If the saddle S_{i_2} is a transient state, the procedure is repeated again.

Simulation method and parameter values.—The model parameters are chosen as in [8], where $\bar{\sigma}_i^0$ are chosen randomly from the range $\bar{\sigma}_i^0 \in [5, 10]$ according to a uniform probability distribution. Without loss of generality we set the sequence order from 0 to N in the connectivity matrix so that $\rho_{i-1i} = \bar{\sigma}_{i-1}^0/\bar{\sigma}_i^0 + 0.51$ for $i = 2, \ldots, N$, $\rho_{i+1i} = \bar{\sigma}_{i+1^0}/\bar{\sigma}_i^0 - 0.5$ for $i = 1, \ldots, N-1$, and $\rho_{ij} = \rho_{j-1j} + (\bar{\sigma}_i^0 - \bar{\sigma}_{j-1}^0)/\bar{\sigma}_j^0 + 2$ for $i \notin \{j-1, j, j+1\}$. Finally, each of the A_i^s is selected randomly from the range $A_i^s \in [-4, 9]$ according to a uniform distribution. Then, the possible decisions $\underline{A}^s$ are statistically independent.

The system (1) is integrated using a Runge-Kutta method for additive noise [13]. When the trajectory reaches a saddle, S_i within a ball of radius 0.1, the decision-making function is applied. We assume that the number of choices at instants t_k is $m_k = M$.

Results of the modeling.—We calculated the evolution of $a_i(t, t_k)$ by using two antagonistic DM rules, i.e., high-risk and risk-aversion DM. Each DM produces different typical behaviors. A small amount of noise introduces a rich variety of behavior. The noise added to the system is never larger than $\langle \eta(t)\eta(t')\rangle = 10^{-2}\delta(t-t')$.

We calculated the median of the length of life L for different complexity levels of the cognitive states N and number of possible choices M. We first analyzed the high-risk DM function; see Figs. 2 and 3. As shown in Fig. 2 the system can opt to end the sequence soon or wander around until it reaches the last stable fixed point, S_N. It is surprising to find that the system undergoes something like a phase transition for a given number of choices M when its size is large enough (see Fig. 3). Before the transition all the simulations reach fixed points, and after the transition the system either repeats parts of the sequence in a random fashion or enters limit cycles. For $N = 10$ this "phase transition" where the system starts wandering around is not present and it always reaches a stable fixed point. It is also interesting to note that the phase-transition point is not

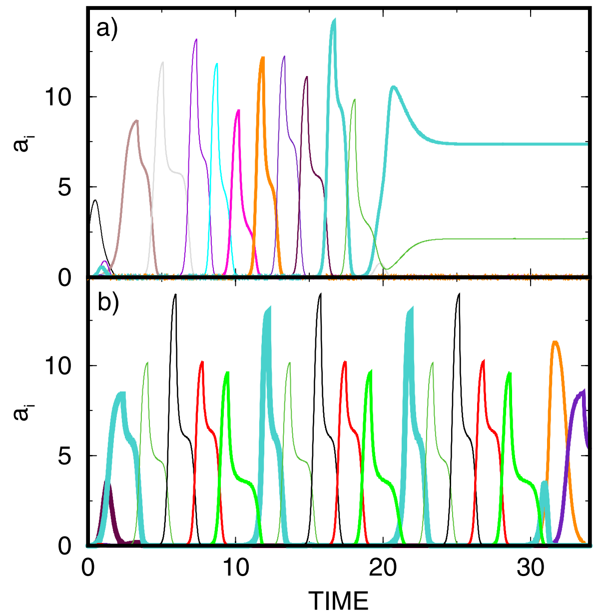

FIG. 2 (color online). High-risk DM dynamics of a system with $N = 20$ and $M = 5$: (a) Most commonly observed DM behavior. (b) An example of the repetitive decisions that can be found in a system with repetitive environment and small percentage of uncertainty. Different tones represent different a_i.

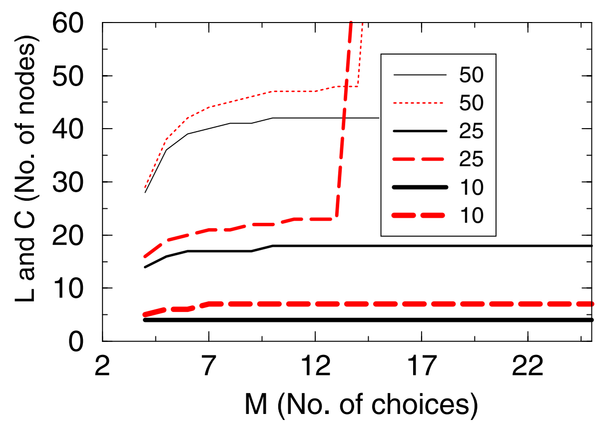

FIG. 3 (color online). Phase transition for the high-risk DM. Median of the length of life L (dashed lines) and median of the number of nodes involved in the sequence (solid lines) versus the number of choices M for the CSM with numbers cognitive states: $N = 10, 25, 50$.

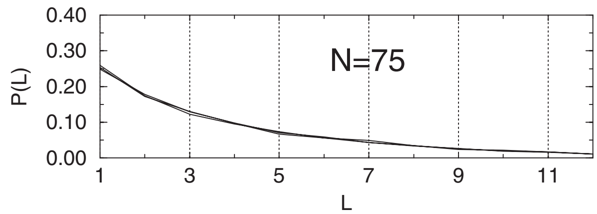

FIG. 4. Probability of finding a sequence of length L as a function of the size of the system $N = 75$ for the *risk-aversion* DM function. $M = 5$, 10, 15 are used in this figure. All of the curves for different M are on top of each other. We do not display the median here because there is no phase transition unlike in Fig. 3.

strongly dependent on the number of choices or the number of cognitive states. These simulations were obtained with 10 000 runs for each N and M. High-risk strategies with sufficiently large numbers of choices last longer.

The risk-aversion DM rule generates completely different results. As our calculations show there is no phase transition in this case. We find that the most important conclusions are the following: First, the system behavior does not depend on the number of available choices M, and second, the length of the sequence decays exponentially (as shown in Fig. 4).

Discussion.—Decision making is a very diverse function of cognitive-state machines, and it may need different approaches for modeling it in different cases. Here we introduced a class of models that describe an uncertainty of the stimulus dependent sequential behavior as a multivariance of the parameters that control the generation of the spatiotemporal patterns responsible for the behavior. To illustrate the potential ability of such models we compared the resulting behaviors of two antagonistic decision functions, i.e., a high-risk and a risk-aversion rule. We showed that high-risk decisions are more effective in increasing the longevity of a behavioral sequence. Despite considering a simple strategy this result is supported by recent psychophysical experiments. In particular, macaque monkeys consistently show faster reactions for larger rewards [14] and good investors, who are not carried away by emotions, avoid risk-aversion strategies [15].

We thank Allen Selverston, Thomas Nowotny, and Pablo Varona for a very productive discussion. This work was supported by the National Institute of Neurological Disorders and Stroke (Grant No. 7R01-NS-38022) and the National Science Foundation (Grant No. EIA-0130708). V. A. was supported by the MEXUS Faculty (Grant No. FE-05-39) and by the CONACyT (Grant No. SEP-2003/C02-42765).

*Electronic address: mrabinovich@ucsd.edu
†Also at Grupo de Neurocomputación Biológica, Escuela Politecnica Superior, Universidad Autónoma de Madrid, 28049 Madrid, Spain.
‡Electronic address: http://inls.ucsd.edu

[1] B. Webb, Nature (London) **417**, 359 (2002).
[2] J. Tanji, Annu. Rev. Neurosci. **24**, 631 (2001); B. B. Averbeck *et al.*, Nat. Neurosci. **9**, 276 (2006).
[3] E. Brown, J. Gao, P. Holmes, R. Bogacz, M. Gilzenrat, and J. D. Cohen, Int. J. Bifurcation Chaos Appl. Sci. Eng. **15**, 803 (2005).
[4] K. L. Briggman, H. D. Abarbanel, and W. B. Kristan, Science **307**, 896 (2005); M. Loh and G. Deco, Eur. J. Neurosci. **22**, 2927 (2005); K.-F. Wong and X.-J. Wang, J. Neurosci. **26**, 1314 (2006).
[5] M. I. Rabinovich, P. Varona, A. I. Selverston, and H. D. I. Abarbanel, Rev. Mod. Phys. (to be published); M. I. Rabinovich, A. Volkovskii, P. Lecanda, R. Huerta, H. D. Abarbanel, and G. Laurent, Phys. Rev. Lett. **87**, 068102 (2001); R. Huerta and M. I. Rabinovich, Phys. Rev. Lett. **93**, 238104 (2004).
[6] R. VanRullen and C. Koch, Trends Cogn. Sci. **7**, 207 (2003); A. Abbott and A. Tsay, Sociological Methods Res. **29**, 3 (2000); A. Abbott, Ann. Rev. Sociology **21**, 93 (1995).
[7] M. D. Fox, M. Corbetta, A. Z. Snyder, J. L. Vincent, and M. E. Raichle, Proc. Natl. Acad. Sci. U.S.A. **102**, 9673 (2005); R. A. Poldrack and M. G. Packard, Neuropsychologia **41**, 245 (2003).
[8] V. S. Afraimovich, V. P. Zhigulin, and M. I. Rabinovich, Chaos **14**, 1123 (2004).
[9] P. Ashwin and J. Borresen, Phys. Lett. A **347**, 208 (2005).
[10] V. S. Afraimovich and S.-B. Hsu, *AMS/IP Studies in Advanced Mathematics*, Lectures on Chaotic Dynamical Systems (International Press, Somerville, MA, 2003).
[11] In principle, it is possible that in spite of the nondissipativity of some saddles in a heteroclinic sequence, the complete sequence nevertheless remains stable (some $\nu_i \leq 1$ but $\prod \nu_i > 1$). This case can be studied by using methods in [8]. We do not consider it here because of the bulkiness of the formulas.
[12] G. Gigerenzer and P. M. Todd, *Simple Heuristics that Make Us Smart* (Oxford University Press, New York, 1999); M. R. Forster, Mind and Machines **9**, 543 (1999).
[13] M. San Miguel and R. Toral, in *Instabilities and Nonequilibrium Structures VI*, edited by Enrique Tirapegui, Javier Martinez, and Rolando Tiemann (Kluwer, Dordrecht, 2000), ISBN 0-7923612-9-6).
[14] M. R. Rowesch and C. R. Olson, J. Neurophysiol. **94**, 1469 (2005).
[15] B. Shiv, G. Loewenstein, A. Bechara, H. Damasio, and A. R. Damasio, Psychol. Sci. **16**, 435 (2005).