

顺序决策的动力学

Dynamics of Sequential Decision Making

Mikhail I. Rabinovich* and Ramo'n Huerta†

Institute for Nonlinear Science, University of California, San Diego, 9500 Gilman Drive 0402, La Jolla, California 92093, USA‡

Valentin Afraimovich

Instituto de Investigacio'n en Comunicacio'n O'ptica, UASLP, Avenida Obreg'o'n 64, 78000 San Luis de Potosi', SLP, Mexico

(Received 4 May 2006; published 1 November 2006)

(SongJian translate)

摘要: 我们提出了一种新的智能决策范式,它适用于动物或人工自主装置的动态顺序活动,该活动取决于内部和外部世界的特性。为此,我们引入了一类新的动力学模型,该模型由常微分方程描述,在决策点上有有限数量的可能性,并且还包括解决这种不确定性的规则。我们的方法是基于可能的认知状态之间的竞争,使用它们稳定的瞬态动力学。该模型根据环境和决策准则控制顺序活动的后续步骤的选择顺序。分析了使系统摆脱不稳定环境的两种策略(高风险和风险规避条件)。

引言。“信息依赖性”瞬态活动是动物和自主智能系统最典型的行为[1]。即使在固定的环境中,这种行为也不是唯一的,大脑或认知状态机制(Cognitive-State Machine, CSM)也必须做出选择;即,这种行为是一系列的转换或决策(Decision-Making, DM)过程[2](见图1)。很明显,顺序行为中的智能决策必须保持稳定、抗噪的,并具有可重复性,以便在将来能够记住和重用成功的决策序列。另一方面,它还必须对来自环境的新信息敏感。这些要求从根本上是矛盾的,现有的方法[3,4]不足以解释顺序活动的DM。在这里,我们基于广义“无赢竞争(Winnerless competition, WLC)”原理[5],建立了一类新的适合于顺序DM分析的模型。

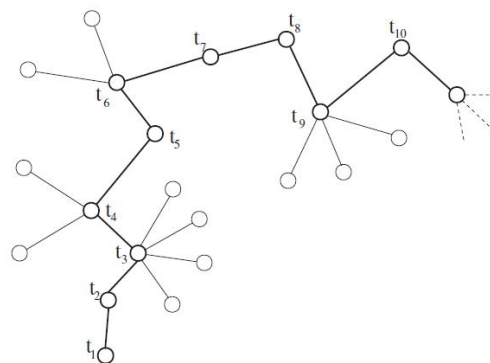


图1: 一系列的认知状态:细线,可能的路径;粗线,根据环境信息通过DM实现选择序列; t_3 , t_4 , t_6 , t_9 是选择的瞬间。

决策系统由以下子系统组成：（i）制定目标；（ii）创建决策函数；（iii）控制 CSM 的参数，和（iv）负责根据输入信息 I 和 DM 规则控制行为的认知状态的时空模式的生成。在这里，我们重点讨论了 CSM 的一个动力学模型及其由决策函数和关于世界的信息控制的参数，即上面的（iii）和（iv）项。

让我们考虑一个系统，其目标是实现顺序决策的最大可能数目。我们假设生命过程可以被编码为事件序列[6]，即我们的案例中的决策事件。在本文中，我们将把整个序列中所做的决定的数量称为“寿命”。决策函数是通过算法定义的（见下文）。我们的模型包括认知状态动力学的常微分方程和由 DM 规则给出的控制参数方程：

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

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

方程（1）是 Lotka-Volterra 类型的，它模拟了认知状态 $a_i(t)$ 的竞争动力学（它可以是大脑模式或竞争控制器；参见，如[7]），其中 ρ_{ij} 表示从状态 j 到状态 i 的竞争交互作用的强度，这是基于遗传和记忆信息， η_i 是外部噪声， N 是可能的认知状态数。CSM 的工作机制是一种依赖刺激的竞争，在系统达到生命终点（一个稳定的平衡点）之前没有赢家。在 CSM 到达这一点之前，不同的认知状态在短时间内成为“赢家”。这是通过非对称抑制（见[5]）确保神经系统可再现瞬态动力学的 WLC 原理。参数 $\sigma_i(t, I)$ 控制认知状态动力学且受到独立的梯度系统（2）控制。在给定 i ，电势函数 $U_i(I, \sigma_i)$ 存在最小值 m_k ，且满足

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

初始值 $\bar{\sigma}_i(t_0)$ 不包含以前历史的记忆，并且由决策规则确定，其方式是选择最小的流域。 $A_i^s(I)$ 代表由外部刺激所决定的动作的全部项目，而 σ_i^0 是所有刺激反应共同的常数项。在下文中，我们假设特征时间 τ 很小，因此可以忽略 σ_i 的动力学；即， σ_i 仅取稳定平衡值 $\bar{\sigma}_i(I)$ 。即使在一个不断变化的环境中，也就是说，如果 I 在不断变化， σ_i 也可以以一种离散的方式变化，因为感知可以是离散的（参见，如[6]）。因此，我们可以假设刺激 I 的作用方式是，在选择 $t_k = 1, 2, \dots$ ，的瞬间，

参数 $\bar{\sigma}_i$ 不是唯一的，可以从 (3) 中取多个值。时间 t_k 被定义为系统到达系统 (1) 不动点附近的瞬间，可能性 m_k 的数目和 $\bar{\sigma}_i$ 的值取决于刺激 I 。在选择的瞬间当中，系统 (1) 根据在时间 t_k 从 (3) 中选择的 $\bar{\sigma}_i$ 的值来演化。

认知状态 $a_i(t)$ 的瞬态动力学的数学图像是一个稳定的异宿序列 (Stable Heteroclinic Sequence, SHS)，且序列步骤根据 DM 规则来选择。在瞬间决策之后会发生什么？如在 [8] 中所示，系统 (1) 在没有噪声且 $\sigma_i = \bar{\sigma}_i$ 时有非平凡平衡点 $S_i = (0, \dots, \bar{\sigma}_i, 0, \dots, 0)$ 。在 S_i 系统上线性化后的特征值是 $\lambda_{ji} = \bar{\sigma}_j - \rho_{ji} \bar{\sigma}_i$ ， $j=1, \dots, N$ ， $j \neq i$ 。根据 λ_{ji} 的值，我们可以找到以下可能性：

(i) 如果所有 $\lambda_{ji} < 0$ ， S_i 是一个稳定的不动点；我们说系统达到了它的生命终点。

(ii) 如果至少存在两个 j 值，比如 j_1 和 j_2 ，使得 $\lambda_{j_1 i} > 0$ 且 $\lambda_{j_2 i} > 0$ ，我们称 S_i 为“恐慌状态 (Panic state)”。此时的系统有无限的选择（对于这种情况下的异宿轨道，见 [9]）。

(iii) 如果只有一个值 $j=j_0$ 使得 $\lambda_{j_0 i} > 0$ ，而且所有其他特征值均为负，则鞍点 S_i 具有一维不稳定流形。我们只考虑“耗散鞍点 (Dissipative saddles)”。耗散鞍座满足以下假设。设 $\lambda_i^- = \max \{ \max_{j \neq j_0, i} \{ \lambda_{ji} \} - \bar{\sigma}_i \}$ ；则 $v_i = -\lambda_i^- / \lambda_{j_0 i}$ 称为鞍点值 [10]。如果 $v_i > 1$ ，鞍点是耗散的。如果是这样，我们称之为“瞬态 (Transient state)”，并且生命继续。

一个 SHS，称为 Γ ，是一个鞍点 $\{S_{i_k}\}$ 集合， $k=1, \dots, K$ ，该集合中有一维不稳定流形以及异宿轨集合 $\{\Gamma_{k,k+1}\}$ ，这个异宿轨集合中元素 $\Gamma_{k,k+1}$ 表示 S_{i_k} 和 $S_{i_{k+1}}$ 之间相连接： $\Gamma = \bigcup_{k=1}^{K-1} (S_{i_k} \cup \Gamma_{k,k+1})$ 。据此可知， Γ 必须是稳定的：如果一条轨线的初始点位于 S_{i_1} 的附近，则这段轨线上的所有点在到达 S_{i_K} 的附近之前都属于 Γ 的邻域。

对于 SHS 的存在性和稳定性，满足以下条件是不够的：(i) 每个鞍点集合 $\{S_{i_k}\}$ 都是一个瞬态；我们用 $\lambda_{i_{k+1}, i_k} = \bar{\sigma}_{i_k} - \rho_{i_{k+1} i_k} \bar{\sigma}_{i_k}$ 表示相应的正特征值；(ii) 满足下列不等式（参见 [8]）

$$1 - \rho_{i_{k-1}i_k} \bullet \rho_{i_k i_{k-1}} \neq 0 \quad (4)$$

$$-\bar{\sigma}_{i_k} < \bar{\sigma}_{i_{k-1}} - \rho_{i_{k-1}i_k} \bar{\sigma}_{i_k} \quad (5)$$

$$\bar{\sigma}_i - \rho_{ii_k} \bar{\sigma}_{i_k} < \bar{\sigma}_{i_{k-1}} - \rho_{i_{k-1}i_k} \bar{\sigma}_{i_k}, \quad i \notin \{i_{k-1}, i_k, i_{k+1}\} \quad (6)$$

假设系统 (1) 的初始条件位于 a_n 轴附近, 并且假设存在一个整数 $m_l > 0$, 满足向量 $\bar{\sigma}$ 的可能值为 $\bar{\sigma} = (\bar{\sigma}_1, \dots, \bar{\sigma}_N): \bar{\sigma}^1, \dots, \bar{\sigma}^{m_l}$ 。在他们之中作出决策之前, 系统根据内在刺激消除下列情况: (a) 对于每个值 $s \in \{1, \dots, m_l\}$, 对应的点 S_i 是一个稳定的不动点。(b) 对于每个值 $s \in \{1, \dots, m_l\}$, 对应的点 S_i 是一个“恐慌状态”。在这样一个鞍点附近经过的一些轨线可能与所制定的目标有关。然而, 如果 CSM 试图记住这一特定行为以备将来使用, 它会遇到恐慌状态附近不稳定的轨线是发散的问题, 因此, 系统在重复试验中朝不同方向移动 (不可产生)。因此, 与基于瞬态的 CSM 相比, 使用恐慌状态的 CSM 行为不稳定, 生存概率较低。(c) 存在 $s \in \{1, \dots, m_l\}$, 其中 S_i 是具有一维不稳定流形的鞍点, 但所有这些鞍点都是非耗散的, 即鞍值 $v \leq 1$ 。情况 (c) 被排除在外, 因为它可能导致顺序行为的不稳定性, 并且动力学不能再现 [如情况 (b)] [11]。

现在我们假设至少存在一个值 $s \in \{1, \dots, m_l\}$, 比方说, $s = s'$, 使得对应的点 $S_{i_1}^{s'}$ 是瞬态。如果这样的值是唯一的, 我们选择 $\bar{\sigma} = \bar{\sigma}^{s'}$, 将其替换到 (1) 式, 并允许系统演化。由于初始点靠近 a_n 轴, 相应轨线上的点到达 $S_{i_1}^{s'}$ 的一个小邻域, 并且由于 $S_{i_1}^{s'}$ 是耗散的, 它将遵循在平面 $\cap_{k \neq i_0, j_0} \{a_k = 0\}$ 上 [8] 的 $S_{i_1}^{s'}$ 和 $S_{j_0} = (0, \dots, 0, \bar{\sigma}_{j_0}^{s'}, 0, \dots, 0)$ 连接的异宿轨线。现在我们主要考虑存在几个鞍点 ($S_{i_1}^{s(q)}$, $q = 1, \dots, p$) 地情况。

决策函数 (Decision-making functions)。 DM 显然取决于目标。让我们把注意力集中在上面用两种极端策略制定的目标上, 这两种极端策略通常被动物用来生存 [12]。例如, 它可以是风险规避 DM (稳定性需求) 或高风险 DM (到达下一个决策点的最短时间)。

高风险决策 (High-risk DM)。 每一个鞍点 $S_{i_1}^{s(q)}$ 仅有一个正的增量 $\lambda_{j_0 i_1} = \bar{\sigma}_{j_0} - \rho_{j_0 i_1} \bar{\sigma}_{i_1}$, $j_0 = j_0(q)$, $q = 1, \dots, p$ 。我们用以下方式选取 q_0 :

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

换句话说，我们选择的最大增量对应于离开鞍点 s_{i_1} 的最快运动，因此也对应于到达 SHS 中的下一个鞍点的最短时间。

风险规避决策 (Risk-aversion DM)。做出选择的另一种可能性是基于稳定性考虑。对于每一个 $q=1, \dots, p$ ，相应的鞍值 $v_{i_1}^q$ 被定义好的。我们选择 q_0 的方式是如下

$$1 < v_{i_1}^q < v_{i_1}^{q_0}, \quad q \neq q_0 \quad (8)$$

在做出一个决策之后，系统会把方程 (1) 中的对应的值替换： $\bar{\sigma} = \bar{\sigma}^{s(q_0)}$ ，并且在接近异宿的轨线上演化，直到到达鞍点 $(0, \dots, 0, \bar{\sigma}_{j_0(q_0)}, 0, \dots, 0)$ 附近。如果此点被固定为流程下一阶段的初始点，我们利用 i_2 定义 $j_0(q_0)$ ，并考虑不同向量 $\bar{\sigma}$ 的数 m_2 和他们的值。如果鞍点 s_{i_2} 是一个瞬态，这个过程会再重复一遍。

仿真方法和参数值。模型参数地选择采用自文献[8]，其中 $\bar{\sigma}_i^0$ 是根据均匀概率分布从范围[5, 10]中随机选取。在不失一般性的前提下，我们将连通矩阵中的序列顺序设置为 0 到 N ，从而

$$\rho_{i-1i} = \bar{\sigma}_{i-1}^0 / \bar{\sigma}_i^0 + 0.51, \quad i=2, \dots, N;$$

$$\rho_{i+1i} = \bar{\sigma}_{i+1}^0 / \bar{\sigma}_i^0 + 0.5, \quad i=1, \dots, N-1$$

$$\rho_{ij} = \rho_{j-1j} + (\bar{\sigma}_i^0 - \bar{\sigma}_{j-1}^0) / \bar{\sigma}_j^0 + 2, \quad i \notin \{j-1, j, j+1\}。$$

最后，每个 A_i 是根据均匀概率分布从范围[-4, 9]中随机选取。则可能的决策 A^s 是统计独立的。

系统(1)采用龙格-库塔方法对加性噪声[13]进行积分。当轨线到达鞍点时， S_i 在半径为 0.1 的球内，决策函数被应用。我们假设 t_k 时刻的选项数 $m_k=M$ 。

模拟结果。我们利用两种对抗性 DM 规则（即“高风险 DM”和“风险规避 DM”）计算了 $a_i(t, t_k)$ 的演化。每个 DM 都会产生不同的典型行为。少量的噪声会导致各种各样的行为。添加到系统中的噪声永远不会大于 $\langle \eta(t) \eta(t') \rangle = 10^{-2} \delta(t - t')$ 。

我们计算了不同复杂程度的认知状态 N 和可能选择数 M 的寿命 L 的中位数。

我们首先分析了“高风险 DM”函数；见图 2 和 3。如图 2 所示，系统可以选择很快结束序列或四处漫游，直到到达最后一个稳定的不动点 $\mathbf{S}_N$ 。令人惊讶的是，当系统的尺寸足够大时，对于给定的选择数 M ，系统经历类似于相变的情况（参见图 3）。在转换之前，所有的模拟都到达不动点，在转换之后，系统要么以随机的方式重复序列的一部分，要么进入极限环。对于 $N=10$ ，系统开始四处游荡的“相变”不存在，并且总是达到一个稳定的不动点。值得注意的是，相变点并不强烈依赖于选择数量或认知状态的数量。这些模拟是在每个 N 和 M 的 10000 次运行中获得的。具有足够多选择的“高风险”策略持续时间更长。

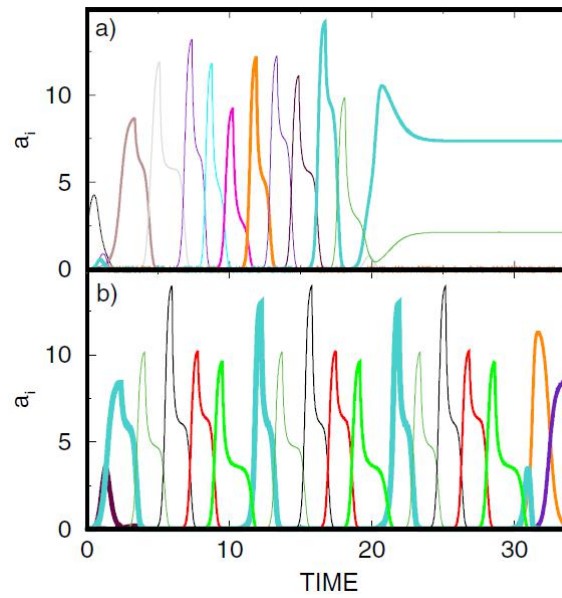


图 2: (彩色线)。具有 $N=20$ 和 $M=5$ 的高风险 DM 动态系统: (a)最常见的 DM 行为。(b)在重复环境和不确定性百分比很小的系统中可以找到重复决策的例子。不同的颜色代表不同的 a_i 。

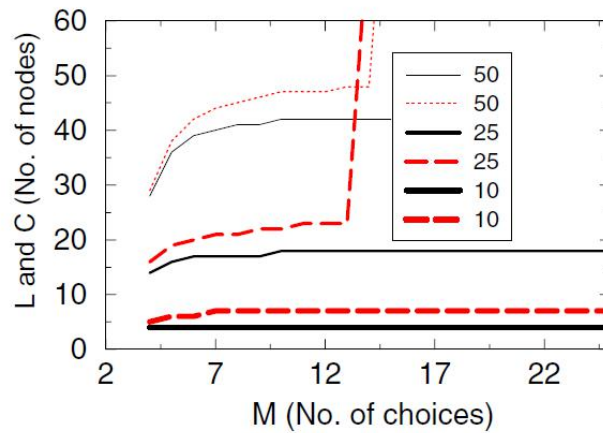


图 3: (彩色线)。高风险 DM 的相变。生命长度的中位数 L (虚线)和参与序列的节点数的中位数(实线)与具有数字认知状态的 CSM 的选择数 M 的中位数: $N=10$ 、25、50。

“风险规避 DM”规则产生完全不同的结果。我们的计算表明在这种情况下没有相变。我们发现最重要的结论如下：第一，系统行为不依赖于可用选项的数量 M ；第二，序列的长度呈指数衰减（如图 4 所示）。

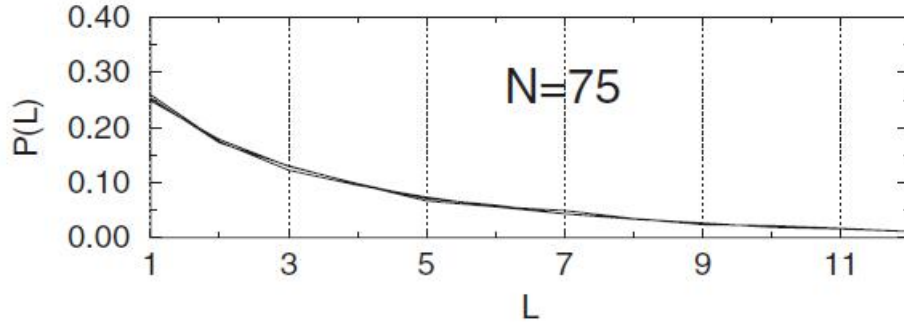


图 4：找到长度为 L 的序列作为系统规模的函数的概率为 $N=75$ 的风险规避 DM 函数。图中使用了 $M=5, 10, 15$ 。不同 M 的所有曲线都是相互重叠的。我们没有在这里显示中值，因为没有像图 3 中那样的相变。

讨论。决策是认知状态机制的一个非常多样化的功能，在不同的情况下，它可能需要不同的建模方法。在这里，我们引入了一类模型，该模型将依赖刺激的顺序行为的不确定性描述为控制该行为的时空模式生成的参数的多变量。为了说明这种模型的潜在能力，我们比较了两个对抗性决策函数的结果行为，即“高风险”和“风险规避”规则。我们发现高风险决策在延长行为序列的寿命方面更有效。尽管考虑了一个简单的策略，这个结果还是得到了最近心理物理实验的支持。特别是，猕猴对更高的回报[14]表现出更快的反应，而好的投资者不会被情绪所左右，会避免风险规避策略[15]。

我们感谢 Allen Selverston、Thomas Nowotny 和 Pablo Varona 的富有成效的讨论。这项工作得到了国家神经疾病和中风研究所（批准号：7R01-NS38022）和国家科学基金会（批准号：EIA—0130708）的资助。V.A.得到了 Mexus 学院（批准号：FE-05-39）和 Conacyt（批准号：SEP-2003/C02-42765）的支持。

参考文献

- [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. 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. Rowes 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).