



## Dynamical Systems

## Robustness in biological regulatory networks I: Mathematical approach

*Robustesse dans les réseaux de régulation biologique I : Approche mathématique*Jacques Demongeot<sup>a</sup>, Jules Waku<sup>a,b</sup><sup>a</sup> AGIM CNRS/UJF 3405, université J. Fourier Grenoble I, faculté de médecine, 38700 La Tronche, France<sup>b</sup> LIRIMA-UMMISCO, université de Yaoundé, faculté des sciences, BP 812, Yaoundé, Cameroon

## ARTICLE INFO

*Article history:*

Received 18 November 2011

Accepted 4 January 2012

Available online 9 February 2012

Presented by the Editorial Board

## ABSTRACT

Numerous indices of complexity are used in biological regulatory networks like the number of their components, their connectance (or connectivity), or the number of the strong connected components of their interaction graph. Concerning the stability of a biological network, it corresponds to its ability to recover from dynamical or parametric disturbance. Complexity is here quantified by the evolutionary entropy, which describes the way the asymptotic presence distribution of the corresponding dynamical system is spread over the state space and the stability (or robustness) is characterized by the rate at which the system returns to this equilibrium distribution after a perturbation. This article shows the mathematical relationships between entropy and stability rate in the general framework of a Markov chain.

© 2012 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

## R É S U M É

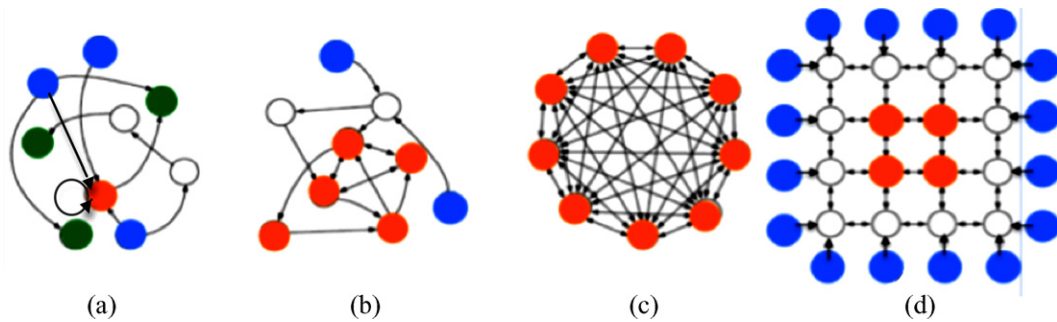
De nombreux indices ont été proposés pour quantifier la complexité des réseaux biologiques de régulation, comme le nombre de leurs composants, leur connectivité, ou le nombre des composantes fortement connexes de leur graphe d'interaction. Quant à la stabilité de ces réseaux biologiques, elle correspond à leur capacité à absorber les changements dynamiques ou paramétriques. La complexité est ici mesurée par l'entropie évolutionnaire, qui décrit la manière dont la probabilité de présence asymptotique du système dynamique correspondant est distribuée dans l'espace d'état, et la stabilité est caractérisée par la vitesse de retour à l'équilibre de cette distribution, après perturbation. Cet article montre les relations mathématiques existant entre entropie et vitesse de retour, de manière générale dans le cadre des chaînes de Markov.

© 2012 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

## 1. Introduction

The complexity–stability relationship in a biological network described in [4] is based both on a mathematical characterization of stability which has its origin in the theory of large deviations [5,6,12] and on a mathematical index of complexity, called evolutionary entropy, introduced by L. Demetrius [2,3] as complementary of other indices of complexity used in biological networks, like the number of their components, their connectance (or connectivity, *i.e.*, the ratio between the number of interactions and the number of components), or the number of strong connected components of their interaction graph.

E-mail address: [Jacques.Demongeot@agim.eu](mailto:Jacques.Demongeot@agim.eu) (J. Demongeot).



**Fig. 1.** Various networks whose interaction graph  $G$  has a unique strong connected component (scc) possessing central points (red), as well as sources (blue) and sinks (green). The connectance of the network increases from 7/10 (a), and 13/9 (b) to 72/9 (c). (a) and (b) are realistic genetic network, (c) is a 9-switch used in plant growth modelling and (d) represents a nearest neighbour ferromagnetic solid model.

The notion of stability is quantified by the rate at which the system returns to its steady state conditions after exogenous and/or endogenous perturbations. We will in this paper review main concepts underlying relationships between complexity and stability in biological networks.

## 2. Evolutionary entropy and stability rate

### 2.1. Biological networks as graphs

A biological network is described as a directed graph  $G$  (digraph), defined in terms of nodes and arcs between the nodes. This graph  $G$  can have different shapes, very regular, but not planar as in 2D crystals (cf. Fig. 1(a)), very regular but not planar, in case of fully connected networks (cf. Fig. 1(c)) in plant growth control metabolic networks (cf. Fig. 1(b)), or very sparse and irregular as those met commonly in genetic regulatory networks (cf. Fig. 1(a)).

The dynamical process defining the evolution of the biological network states invokes a configuration space  $\Omega$  and a potential function  $\varphi : \Omega \rightarrow \mathbb{R}$ . The configuration space  $\Omega$  is the set of all doubly infinite paths generated by the graph  $G$ . Let  $\mu$  denote the Gibbs measure associated with the potential  $\varphi$ : we will represent the biological network by the mathematical object  $(\Omega, \mu, \varphi)$  and consider systems whose Gibbs measure  $\mu$  is invariant under the shift  $\tau$  on  $\Omega$ . Two characteristics are associated with this dynamical system: the Kolmogorov–Sinai invariant  $H_\mu(\tau)$ , and the stability rate  $R$ . We call  $H_\mu(\tau)$  (denoted by  $H$  in the following) evolutionary entropy [4].

### 2.2. Evolutionary entropy

The evolutionary entropy is a measure of the degree of connectance of the network, this property deriving from the Shannon–McMillan–Breiman theorem [3], which says that, for arbitrarily small  $\varepsilon$  and  $\delta > 0$ , and for sufficiently large  $t$ , the set of the partial trajectories  $x$  of length  $t$  (i.e., the collection of states  $x_k$  of the system at time  $k$  observed between times  $k = 0$  and  $k = t - 1$ ) generated by the network dynamics, can be divided into 2 classes  $S_1$  and  $S_2$  such that:

- (a) for every  $x$  in  $S_1$ , we have:  $|\log \mu(x)/t + H| < \varepsilon$ ;
- (b) the sum of the probabilities of the partial trajectories  $x$  belonging to  $S_2$  is less than  $\delta$ .

All partial trajectories of the class  $S_1$  have approximately the same probability, namely  $e^{-tH}$ . This means that the number of partial trajectories in  $S_1$  is approximately  $e^{tH}$ . This number specifies the number of “typical” partial trajectories generated by an individual in a population and it is positively correlated with the connectance of the network.

### 2.3. Complexity and stability in Markov chains

The complexity will be measured by the evolutionary entropy equal to:

$$H = - \sum_{x, y \in \Omega} \mu_x M_{xy} \log M_{xy},$$

where  $M$  is the transition matrix of the Markov chain. The concept of stability (or robustness) pertains to the capacity of a network to maintain its functionality in the face of random perturbations in internal organization and external factors [10]. We proposed in [4] that the stability rate  $R$  could be quantified by  $R = -\log |\lambda_1|$ , where  $\lambda_1$  is the subdominant eigenvalue of  $M$ . We will now give a series of propositions concerning the estimation of  $H$  in the case of Markov chains.

**Proposition 1.**  $H = H_\mu - d_K(P_{x,y}, \mu_x \otimes \mu_y)$ , where  $H_\mu = -\sum_{y \in \Omega} \mu_y \log \mu_y$  is the entropy of the invariant measure  $\mu$  and  $d_K(P_{x,y}, \mu_x \otimes \mu_y)$  is the Kullback divergence between the joined measure  $P_{x,y}$  defined by the transition matrix  $M_{xy}$  and the product measure  $\mu_x \otimes \mu_y$ .

**Proof.** We have:

$$\begin{aligned} H &= - \sum_{x,y \in \Omega} \mu_x M_{xy} \log M_{xy} = - \sum_{x,y \in \Omega} \mu_x \mu_y [P_{x,y}/(\mu_x \mu_y)] [\log(P_{x,y}/(\mu_x \mu_y)) + \log \mu_y] \\ &= - \sum_{x,y \in \Omega} P_{x,y} [\log P_{x,y} - \log(\mu_x \mu_y)] + \sum_{y \in \Omega} \mu_y \log \mu_y = H_\mu - d_K(P_{x,y}, \mu_x \otimes \mu_y). \quad \square \end{aligned}$$

We can consider that  $d_K(P_{x,y}, \mu_x \otimes \mu_y)$  is an index of dependence between the successive random variables  $X_t$  and  $X_{t+1}$  constituting the Markov chain.

We will now propose the following remarks:

- (1) When the Markov is a Bernoulli shift, then all the lines of the transition matrix  $(M_{xy})_{y \in \Omega}$  are the same and  $X_t$  and  $X_{t+1}$  are independent. In this particular case, we have:  $H = H_\mu$ . It is for example the case when the Markov process represents the backward succession of the mother ages in the genealogies of a population dynamics formalized by a Leslie model [1].
- (2) Another circumstance in which  $H = H_\mu$  is observed in getBrens, when  $T$  is tending to infinity, because both  $\mu$  and  $(M_{xy})_{y \in \Omega}$  are the uniform distribution on  $\Omega$ .
- (3) On the opposite,  $H_\mu \approx d_K(P_{x,y}, \mu_x \otimes \mu_y)$  when the asymptotic measure  $\mu$  is uniform over the attractors, with a weak recurrence rate  $\varepsilon$  for the not fixed configurations. Let us suppose that the genetic network is made of 2 genes, i.e.,  $\Omega = \{0, 1\}^2$ , and that the network has two attractors, the fixed point 01, with the attraction basin  $\{00, 01\}$ , and the cycle  $(10, 11)$ . Then, if  $\varepsilon$  is sufficiently small, we have:

$$M = \begin{pmatrix} \varepsilon & 1-\varepsilon & 0 & 0 \\ \varepsilon & 1-\varepsilon & 0 & 0 \\ 0 & 0 & \varepsilon & 1-\varepsilon \\ 0 & 0 & 1-\varepsilon & \varepsilon \end{pmatrix}, \quad \mu = (\varepsilon/3, (1-\varepsilon)/3, 1/3, 1/3) \quad \text{and} \quad H \approx -\varepsilon \log \varepsilon.$$

Let us consider now the symmetric Kullback–Leibler divergence between the measures  $\mu$  and  $\nu$ , defined by:  $d_{KL}(\mu, \nu) = \sum_{x \in \Omega} \mu_x [\log \mu_x - \log \nu_x] + \sum_{x \in \Omega} \nu_x [\log \nu_x - \log \mu_x]$ . Then, we have:

**Proposition 2.** Let suppose that  $\mu$  and  $\nu$  are close:  $\mu_x = \nu_x(1 + \varepsilon_x)$ , where  $\varepsilon_x$  is small. Then, if  $\mu$  is uniform on a subset  $C$  of cardinal  $C$  of  $\Omega$ , we have:  $d_{KL}(\mu, \nu) \approx C \|\mu - \nu\|^2$  where  $\|\cdot\|$  denotes Euclidean norm.

**Proof.** Let suppose that:  $\mu_x = \nu_x(1 + \varepsilon_x)$ , where  $\varepsilon_x$  is small. Then, we have:

$$d_{KL}(\mu, \nu) = \sum_{x \in \Omega} \mu_x [\log \mu_x - \log \nu_x] + \sum_{x \in \Omega} \nu_x [\log \nu_x - \log \mu_x] = \sum_{x \in \Omega} \nu_x \varepsilon_x \log(\mu_x/\nu_x) \approx \sum_{x \in \Omega} \nu_x \varepsilon_x^2.$$

But  $\sum_{x \in \Omega} \nu_x \varepsilon_x = 0$ , hence  $(\sum_{x \in \Omega} \nu_x \varepsilon_x)^2 = 0$  and  $\|\mu - \nu\|^2 = \sum_{x \in \Omega} \nu_x^2 \varepsilon_x^2 = -2 \sum_{x,y \in \Omega, x \neq y} \nu_x \nu_y \varepsilon_x \varepsilon_y$  and  $(\sum_{x \in \Omega} \nu_x \varepsilon_x)(\sum_{y \in \Omega} \varepsilon_y) = 0$ , hence  $d_{KL}(\mu, \nu) = \sum_{x \in \Omega} \nu_x \varepsilon_x^2 = -2 \sum_{x,y \in \Omega, x \neq y} \nu_x \varepsilon_x \varepsilon_y$ . If  $|C| = C$ , we get:  $\|\mu - \nu\|^2 = -2(\sum_{x,y \in \Omega} \varepsilon_x \varepsilon_y)/C^2 \approx d_{KL}(\mu, \nu)/C$ .  $\square$

**Proposition 3.** We suppose that at each iteration of the Markov chain, we have probability 1/2 to stay in a not fixed configuration, 1/2 to pass to the following configuration, and 1 to stay in fixed configurations. If one of the  $m$  attractor is periodic of period  $T$ , each of its  $T$  configurations has a probability 1/2 to stay and 1/2 to pass to the following configuration. Then, we have, for any initial probability measure  $\nu_0$  on  $\Omega$ :

$$\nu_k = M^k \nu_0 \quad \text{and} \quad d_{KL}(\mu, \nu_k) \leq K e^{2k \log |\lambda_1|} \approx L e^{-2kH},$$

where  $\lambda_1$  is the subdominant eigenvalue of the transition matrix  $M$ .

**Proof.** We have:  $\forall k = 1, m$ , the transition matrix  $M$  has on each of its lines only coefficients  $M_{xy}$  equal to 1/2, except for the  $p$  fixed configurations  $z$ , for which  $M_{zz} = 1$ , and for periodic configurations of any limit-cycle like  $(z(1), \dots, z(T))$ , for which  $M_{z(k)z(k+1 \bmod T)} = 1$ . Then, we have:  $\text{Det}(M - \lambda I) = (1/2 - \lambda)^{\sum_{k=1,p} |B(A_k)|} (1 - \lambda)^p (1/2 - \lambda)^{\sum_{k=p+1,m} |B(A_k)|} \prod_{k=p+1,m} [(1/2 - \lambda)^{T_k} + (-1)^{T_{k+1}} (1/2)^{T_k}]$ , where  $|B(A_k)|$  is the cardinal of the basin of the attractor  $A_k$ . Then, the subdominant eigenvalue  $\lambda_1$  of the transition matrix  $M$  verifies:  $\lambda_1 = \frac{1}{2}$  and  $\mu$  has all of its components equal to 0 except those equal to  $1/q$ , which correspond to the  $q = p + \sum_{k=p+1,m} T_k$  attracting configurations, i.e., the fixed and periodic configurations. Hence, we have:

$$H = -\left[1/2 \log(1/2) + 1/2 \log(1/2)\right] = -\log(1/2).$$

We have also for  $k$  sufficiently large [1]:  $\|\mu - \nu_k\|^2 \leq K e^{2k \log|\lambda_1|}$  and the end of the proof comes from Proposition 2, where  $\mathcal{C}$  is the union of the attractors and  $L = KC$ .  $\square$

Results similar to Propositions 2 and 3 can be found in the literature [7–9,11]. They show a direct explicit relationship between a stability ( $R = -\log|\lambda_1|$ ) and a complexity ( $H$ ) index.

### 3. Conclusion

We have presented in this article some theoretical results concerning the correlated indices of complexity and stability for a biological regulatory network. These indices are respectively the evolution entropy  $H$  and the stability index  $R$ . Then, we have developed explicit relationships between  $H$  and  $R$  in the case of Markov chains.

### References

- [1] L. Abbas, J. Demongeot, N. Glade, Synchrony in reaction–diffusion models of morphogenesis: applications to curvature-dependent proliferation and zero-diffusion front waves, *Philos. Trans. R. Soc. Lond. Ser. A* 367 (2009) 4829–4862.
- [2] L. Demetrius, M. Gundlach, G. Ochs, Complexity and demographic stability in population models, *Theor. Pop. Biol.* 65 (2004) 211–225.
- [3] L. Demetrius, M. Ziehe, Darwinian fitness, *Theor. Pop. Biol.* 72 (2007) 323–345.
- [4] J. Demongeot, L. Demetrius, Natural selection and demographic drift: an empirical study of France (1860–1965), *Population* 2 (1989) 231–248.
- [5] M.D. Donsker, S.R.S. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time. I, *Comm. Pure Appl. Math.* 28 (1975) 1–47.
- [6] M.I. Freidlin, A.D. Wentzell, *Random Perturbations of Dynamical Systems*, Springer, New York, 1984.
- [7] S. Goldstein, Entropy increase in dynamical systems, *Israel J. Math.* 38 (1981) 241–256.
- [8] S. Goldstein, O. Penrose, A non-equilibrium entropy for dynamical systems, *J. Stat. Phys.* 24 (1981) 325–343.
- [9] J.H. Jensen, D.P.W. Ellis, M.G. Christensen, S.H. Jensen, Evaluation of distance measures between Gaussian mixture models of MFCCs, in: *Proceedings ISMIR 2007*, Austrian Computer Society, Wien, 2007, pp. 107–108.
- [10] A. Lesne, Robustness: confronting lessons from physics and biology, *Biol. Rev. Camb. Philos. Soc.* 83 (2008) 509–532.
- [11] S. Tuljapurkar, Entropy and convergence in dynamics and demography, *J. Math. Biol.* 31 (1993) 253–271.
- [12] A.D. Wentzell, M.I. Freidlin, On small random perturbations of dynamical systems, *Russian Math. Surveys* 25 (1970) 1–55.