Robustness in biological regulatory networks II: Application to genetic threshold Boolean random regulatory networks (getBren)

Jacques Demongeot\textsuperscript{a}, Jules Waku\textsuperscript{a, b}

\textsuperscript{a} AGIM CNRS/UJF 3405, université J. Fourier Grenoble I, faculté de médecine, 38700 La Tronche, France
\textsuperscript{b} LIRIMA-UMMISCO, université de Yaoundé, faculté des sciences, BP 812, Yaoundé, Cameroon

1. Introduction

Since the recent discoveries on the role of small RNAs (like microRNAs), the problem of evolution of robustness of genetic networks caused by endogenous or exogenous parameter variations has been posed. We will solve partly the question about the parameter sensitivity of the evolutionary entropy \[10\] of the network by relating it to the network frustration.

2. Applications to genetic threshold Boolean random regulatory networks (getBren)

We will now apply the results proved in \[10\] to the biological regulatory networks, using the formalism of the genetic threshold Boolean regulatory networks (getBren), which is the translation, in the genetic language, of the Hopfield networks \[11,12\].
2.1. Genetic regulatory networks

The expression of a gene, i.e., the production of a protein this gene encodes, is regulated by the activity of other proteins. We now describe the classical Boolean formalism [1–13] used for modelling genetic networks, before applying it to particular networks.

2.2. Genetic threshold Boolean random regulatory networks (getBren) definition

Any random automaton $i$ of $N$ owns at time $t$ a state $x_i(t)$ valued in $\{0, 1\}$, $0$ (resp. $1$) meaning that gene $i$ is inactivated (resp. activated). The global state of the getBren at time $t$, called configuration in the sequel, is then defined by: $x(t) = (x_i(t))_{i \in \{1, n\} \in \Omega = \{0, 1\}^n}$.

A getBren $N$ of size $n$ is a triplet $(W, \Theta, P)$ where:

- $W$ is a matrix of order $n$, where the coefficient $w_{ij} \in \mathbb{R}$ represents the interaction weight gene $j$ has on gene $i$. $\text{Sign}(W) = (\text{sign}(w_{ij}))$ is the adjacency matrix of the interaction graph $G$.
- $\Theta$ is an activation threshold vector of dimension $n$, its component $\theta_i$ being the activation threshold attributed to automaton $i$.
- $M : P(\Omega) \rightarrow [0, 1]^{m \times m}$ (where $P(\Omega)$ is the set of subsets of $\Omega$ and $m = 2^N$) is a Markov transition matrix, built from local probability transitions $P_i$ giving the new state of gene $i$ at time $t+1$ according to $W, \Theta$, and configuration $x(t)$ of $N$ at time $t$ such that: $\forall \gamma \in \{0, 1\}, \beta \in \Omega$.

$$P_i^{\beta}(\{x_i(t+1) = \gamma \mid x(t) = \beta\}) = \exp \left[ \gamma \left( \sum_{j \in N_i} w_{ij} \beta_j - \theta_i \right) / T \right] \left[ 1 + \exp \left( \sum_{j \in N_i} w_{ij} \beta_j - \theta_i \right) / T \right].$$

where $N_i$ is the neighbourhood of gene $i$ in the getBren $N$, i.e., the set of genes $j$ (including eventually $i$) such that $w_{ij} \neq 0$, and $P_i^{\beta}$ is the probability for gene $i$ of passing to state $\gamma$, knowing its state $\beta$ on $N_i$. $P$ denotes the transition matrix built from the $P_i^{\beta}$'s and depends on the updating mode. For extreme values of the randomness parameter $T$, we have:

(i) if $T = 0$, the getBren becomes deterministic and the transition can be written as: $x_i(t+1) = 1$, if $\sum_{j \in N_i} w_{ij} x_j(t) - \theta_i > 0; = 0$, otherwise;

(ii) when $T$ tends to infinity, then $P_i^{\beta} = \frac{1}{2}$ and each line of $M$ is the uniform distribution on $\Omega$.

2.3. Entropy and robustness in getBrens

Let $M = (M_{xy})$ denote the Markov transition matrix between the configurations $x$ and $y$ of $\Omega$, defined by an updating mode, and $\mu = (\mu_x) = (\mu(x))_{x \in \Omega}$ its stationary distribution on $\Omega$. The evolutionary entropy $H$ serving as robustness index can be calculated for a getBren: $H = -\sum_{x,y \in \Omega} \mu_x M_{xy} \log M_{xy}$. In sequential updating mode, where the updating order of
the nodes are their index order integer, we have, by denoting $I = \{1, \ldots, i-1\}$, $N \setminus I = \{i, \ldots, n\}$ and identifying $x$ with the set of indices $i$ such that $x_i = 1$:

$$M_{xy} = \prod_{i=1}^{n} \left[ p_{x,i}^{(N(i))} \right]_{i \in y} + p_{x,0}^{(N(i))} \]_{i \notin y}$$

(1)

and $\mu$ is classically the Gibbs measure [12] defined by: $\forall x \in \Omega$, $\mu_x = \exp((\sum_{i \in y} w_{ij} x_j - \theta_i)/T)$, where $Z = \sum_{y \in \Omega} \exp((\sum_{i \in y, j \in N_j} w_{ik} y_j - \theta_i)/T)$. When $T = 0$, $\mu$ is concentrated on the attractors of the deterministic dynamics and $H = 0$; when $T = +\infty$, $\mu$ is scattered uniformly over $\Omega$ and $H = n \log 2$. Let define in this case two characteristics of the attraction basin $B(A)$ of an attractor $A$ (i.e., the set of all initial conditions having $A$ as asymptotic behaviour), its Attraction Basin Relative Size $ABRS(A) = \sum_{x \in B(A) \cup A} \mu_x = \mu(A)$ and its Average Diameter (AD), i.e., the average number of transitions needed to reach $A$ from an initial configuration lying in $B(A)$. We can estimate $H$ from attractor entropy $E_{\text{attractor}}$: $E_{\text{attractor}} = -\sum_{k=1, m \leq 2^n} \mu(A_k) \log \mu(A_k) = -\sum_{k=1, m \in 2^n} ABRS(A_k) \log ABRS(A_k)$. In the next Proposition 1, $E_{\text{attractor}}$ plays the same role as $\Delta_k(P_{xy}, \mu_x \otimes \mu_y)$ in Proposition 1 of [10].

**Proposition 1.** Let suppose that the invariant measure $\mu$ is uniform over all the $m \leq 2^n$ attractors of the deterministic dynamics on the configuration space $\Omega$. Then, if for any configuration $i$ of an attractor $A$, $P_{xy}^i$ is defined for any $\beta \in A$ with a local temperature $T_A$ sufficiently large such as $P_{xy}^\beta$ is scattered uniformly over $B(A) \cup A$ in $\Omega$, such as $ABRS(A_k) = |B(A_k) \cup A_k|/2^n$, we have:

$$H \approx \sum_{k=1, m} ABRS(A_k) \log(2^n ABRS(A_k)) = n \log 2 - E_{\text{attractor}} = H_\mu - E_{\text{attractor}}.$$  

(2)

**Proof.** For any $k = 1, m$, all transition matrix coefficients $M_{xy}$ equal to $1/|B(A_k) \cup A_k|$ on lines corresponding to configurations $x$ belonging to $B(A_k) \cup A_k$ and $\mu$ is uniform over $\Omega$. Then we can write: $H = -\sum_{k=1, m \leq 2^n} ABRS(A_k) \log ABRS(A_k) = n \log 2 - E_{\text{attractor}}$. □

We will now calculate in the framework of a getBren the sensitivity of the entropy $H_\mu$ to a variation of the parameter weights $w_{ij}$. Let us suppose for the sake of simplicity that $w_{ij} = 0$ and $\theta_i = 0$ for all the genes $i$ and that $w_{ij} = \alpha_{ij} w$, where $\alpha_{ij}$ is the sign of the interaction between genes $i$ and $j$, equal to $0, -1$ or $+1$, and $w$ denotes the supposed same common absolute value of all weights. Random function $U$ is defined on $\Omega$ by: $U(x) = \sum_{i \in [1, n]} \alpha_{ij} x_j = Q_+(N) - F(x)$, where $Q_+(N)$ is the number of positive edges in interaction graph $G$ of the network $N$ and $F(x)$ the global self-frustration of $x$, i.e., the number of pairs $(i, j)$ where the values of $x_i$ and $x_j$ are contradictory with the sign $\alpha_{ij}$ of the interaction between genes $i$ and $j$: $F(x) = \sum_{j \in [1, n]} F_i(x)$, where $F_i(x)$ is the local self-frustration [9] of the pair $(i, j)$ defined by: $F_i(x) = 1$, if $\alpha_{ij} = 1$, $x_ix_j = 0$ or $\alpha_{ij} = -1$, $x_ix_j = 1$, and $F_i(x) = 0$, elsewhere. Then, we have:

**Proposition 2.** In sequential updating mode, we have: $\partial H_\mu/\partial w = -w \text{Var} U = -w \text{Var} F$ and $\partial H_\mu/\partial T = \text{Var} F/T^2$, implying that the robustness increases when $\partial H_\mu/\partial T$ decreases, then when the temperature increases, $\mu$ tending in these circumstances to the uniform distribution over $\Omega$.

**Proof.** We have: $H_\mu = -\sum_{x \in \Omega} \mu_x \log \mu_x$, then $\partial H_\mu/\partial w = -\sum_{x \in \Omega} \partial \mu_x/\partial w \log \mu_x = -\sum_{x \in \Omega} \partial \mu_x/\partial w \log \mu_x$ where $\partial \mu_x/\partial w = \partial \exp((\sum_{i \in x, j \in N_j} w_{ij} x_j - \theta_i)/T)/\partial w$. Then $Z = \sum_{y \in \Omega} \exp((\sum_{i \in y, j \in N_j} w_{ik} y_j - \theta_i)/T))$ and $\partial \mu_x/\partial w = \sum_{y \in \Omega} \exp((\sum_{i \in y, j \in N_j} w_{ik} y_j - \theta_i)/T)) Z \partial \mu_x/\partial w$, then:

$$\partial \mu_x/\partial w = \left[ \partial \exp((\sum_{i \in x, j \in N_j} w_{ij} x_j - \theta_i)/T))/\partial w \right] / Z - \exp((\sum_{i \in x, j \in N_j} w_{ij} x_j - \theta_i)/T)) (\partial Z/\partial w)/Z^2$$

and

$$\partial \mu_x/\partial w \log \mu_x = \left( \sum_{i \in x, j \in N_j} \alpha_{ij} x_j/T \right) \mu_x - \sum_{y \in \Omega} \left( \sum_{i \in y, j \in N_j} \alpha_{ij} y_j/T \right) \mu_y \mu_x = \mu_x \partial \log \mu_x/\partial w$$

and

$$\partial \mu_x/\partial w \log \mu_x = \left( \sum_{i \in x, j \in N_j} \alpha_{ij} x_j/T \right)^2 \mu_x - \sum_{y \in \Omega} \left( \sum_{i \in y, j \in N_j} \alpha_{ij} y_j/T \right) \left( \sum_{i \in x, j \in N_j} \alpha_{ij} x_j/T \right) \mu_y \mu_x$$

$$- \left( \sum_{i \in x, j \in N_j} \alpha_{ij} x_j/T \right) \mu_x \log Z + \sum_{y \in \Omega} \left( \sum_{i \in y, j \in N_j} \alpha_{ij} y_j/T \right) \mu_y \mu_x \log Z.$$

Hence: $\partial H_\mu/\partial w = -\sum_{x \in \Omega} \partial \mu_x/\partial w \log \mu_x - \sum_{x \in \Omega} \mu_x \partial \log \mu_x/\partial w = -w E(U^2) + w(E(U))^2 - \sum_{x \in \Omega} \partial \mu_x/\partial w$, but $\sum_{x \in \Omega} \partial \mu_x/\partial w = \partial (\sum_{x \in \Omega} \mu_x)/\partial w = 0$, therefore $\partial H_\mu/\partial w = -\sum_{x \in \Omega} \partial \mu_x/\partial w \log \mu_x = -w \text{Var}(U)$. □
By defining now the local cross-frustration \( G_{ij}(x, y) = 1 \), if \( \alpha_{ij} = 1 \), \( x_i y_j = 0 \) or \( \alpha_{ij} = -1 \), \( x_i y_j = 1 \), and \( G_{ij}(x) = 0 \), elsewhere, the global cross-frustration is \( G(x, y) = \sum_{i,j \in \{1, \ldots, n\}} G_{ij}(x, y) \).

**Proposition 3.** In parallel updating mode, we have: \( \partial H_x / \partial w = -w \text{Var}_x G \), where \( \text{Var}_x \) is taken for the conditional measure \( M_{xy} = \exp(\sum_{i \in x, j \in y} w_{ij} x_i y_j - \theta_i)/T \) \( / \sum_{y \in \Omega} \exp(\sum_{i \in x, j \in y} w_{ij} x_i y_j - \theta_i)/T \).

**Proof.** The proof is the same as proof of Proposition 2, by replacing \( F \) by \( G \). □

**Proposition 4.** In parallel updating mode, we have: \( \partial H / \partial w = -w \text{Var} G + \text{Cov}(H, G) \), where \( \text{Var} \) and \( \text{Cov} \) are taken for random variables \( G(x, y) \) and \( H(x, y) \). \( \text{Var}(x) \) and \( \text{Cov}(x,y) \) are respectively the variance and covariance measures defined by \( M_{xy} \mu_x \), where: \( \mu_x = \sum_{y \in \Omega} \exp(\sum_{i \in x, j \in y} w_{ij} x_i y_j - \theta_i)/T \) \( / \sum_{y \in \Omega} \exp(\sum_{i \in x, j \in y} w_{ij} x_i y_j - \theta_i)/T \).

**Proof.** We have: \( \partial H / \partial w = \sum_{x \in \Omega} (\mu_x \partial H_x / \partial w + H_x \partial \mu_x / \partial w) \). Hence, from Proposition 3, we get: \( \partial H / \partial w = -w \text{Var} G + \sum_{x \in \Omega} H_x \partial \mu_x / \partial w \), where:

\[
\partial \mu_x / \partial w = \sum_{y \in \Omega} \left[ \partial \left( \exp(\sum_{i \in x, j \in y} w_{ij} x_i y_j - \theta_i)/T \right) / \partial w \right] / Z
- \left( \sum_{y \in \Omega} \exp(\sum_{i \in x, j \in y} w_{ij} x_i y_j - \theta_i)/T \right)/Z^2,
\]

where \( Z = \sum_{x \in \Omega, y \in \Omega} \exp(\sum_{i \in x, j \in y} w_{ij} x_i y_j - \theta_i)/T \) and \( \partial Z / \partial w = \sum_{x \in \Omega, y \in \Omega} \alpha_{ij} x_i y_j / T \exp(\sum_{i \in x, j \in y} w_{ij} x_i y_j - \theta_i)/T \), then we have:

\[
\partial \mu_x / \partial w = \sum_{y \in \Omega} \left( \sum_{i \in x, j \in y} \alpha_{ij} x_i y_j / T \right) \exp(\sum_{i \in x, j \in y} w_{ij} x_i y_j - \theta_i)/T / Z
- \mu_x \sum_{x \in \Omega, y \in \Omega} \left( \sum_{i \in x, j \in y} \alpha_{ij} x_i y_j / T \right) \exp(\sum_{i \in x, j \in y} w_{ij} x_i y_j - \theta_i)/T / Z = \mu_x (E_x(G) - E(G)).
\]

Finally, we get: \( \partial H / \partial w = -w \text{Var} G + \sum_{x \in \Omega} H_x \partial \mu_x / \partial w = -w \text{Var} G + \text{Cov}(H, G) \). □

We have shown, in Propositions 2, 3 and 4, a direct link between the sensitivity to \( w \) of \( H_\mu \) and \( H \), and the variability of the frustration of the network.

3. Conclusion

We have in this article developed explicit relationships between the complexity index \( H \) (evolutionary entropy) and the stability rate \( R \) in the particular case of the Markov chain defining the dynamics of a generic threshold Boolean random regulatory network (getBren). Eventually, we have given an explicit expression depending on the network frustration for the sensitivity of the evolutionary entropy \( H \) to the variations of the parameters \( w \) (weight or interaction parameter) and \( T \) (temperature or stochasticity parameter).

References