Contrasting Fisher and Shannon Information in Random Boolean Networks

X. Rosalind Wang[1], Joseph T. Lizier[1,2] and Mikhail Prokopenko[1]

[1]CSIRO Information and Communications Technology Centre, PO Box 76, Epping, NSW 1710, Australia
[2]School of Information Technologies, The University of Sydney, NSW 2006, Australia
Rosalind.Wang@csiro.au

Introduction

Random Boolean Networks (RBNs) (Kauffman, 1993; Gershenson, 2004) have typically been used by Artificial Life researchers as discrete dynamical network models with a large sample space available. In particular, RBNs exhibit a well-known phase transition from ordered to chaotic dynamics, with respect to average connectivity or activity level.

Recently, there have been several attempts to study the order-chaos phase transitions of RBNs using information theory (Ribeiro et al., 2008; Rämö et al., 2007; Lizier et al., 2008). In this study we analyse a phase diagram of RBN dynamics in information-theoretic terms, using Fisher information (Fisher, 1922) which measures the amount of information that an observable random variable carries about an unknown parameter. One could expect this quantity to be maximised near the critical point where system dynamics are most sensitive to control parameters. Furthermore, since some studies of Fisher information discuss its connections to (derivatives of) Shannon information, we intend to clarify the relationship between Shannon and Fisher information, using RBNs.

Random Boolean Networks

An RBN consists of N nodes in a directed *network*. The nodes take *boolean* state values, and update their state values in time as a function of the state values of the nodes from which they have incoming links. The network topology is determined at *random*, subject to whether the in-degree for each node is constant or stochastically determined given an average in-degree $\overline{K}$. Given the topology, the deterministic boolean function or lookup table by which each node computes its next state from its neighbours is also decided at *random* for each node, subject to a probability r of producing outputs of "1" (the *bias*). Note that r close to 1 or 0 gives low activity, whereas r close to 0.5 gives the highest activity for any $\overline{K}$.

RBNs are known to exhibit three distinct phases of dynamics, depending on their parameters: ordered, chaotic and critical. At relatively low connectivity (i.e., low degree K) or activity (i.e., r close to 0 or 1), the network is in an ordered phase, characterised by high regularity of states and strong convergence of similar global states in state space. Alternatively, at relatively high connectivity and activity, the network is in a chaotic phase, characterised by low regularity of states and divergence of similar global states. In the critical phase (the *edge of chaos* (Langton, 1990)), there is percolation in nodes remaining static or updating their values, and uncertainty in the convergence or divergence of similar macro states.

Fisher Information

Fisher information (Fisher, 1922) is a way of measuring the amount of information that an observable random variable X has about an unknown parameter θ, upon which the likelihood function of θ depends. Let $p(x|\theta)$ be the likelihood function of θ given the observations x. Then, Fisher information can be written as:

$$F(\theta) = \int_x \left(\frac{\partial \ln(p(x|\theta))}{\partial \theta} \right)^2 p(x|\theta) dx, \tag{1}$$

where $\ln(p(x|\theta))$ is the log-likelihood of θ given x. Thus, Fisher information is not a function of a particular observation, since the random variable X has been averaged out.

The discrete form of Fisher information is:

$$F(\theta) = \sum_{x_j} p_{x_j} \left(\frac{\Delta \ln(p_{x_j})}{\Delta \theta} \right)^2, \tag{2}$$

where $\Delta \ln(p_{x_j}) = \ln(p'_{x_j}) - \ln(p_{x_j})$ and $p_{x_j} = p(x_j|\theta)$, $p'_{x_j} = p(x_j|\theta+\Delta\theta)$. In this case, $p(x)$ is a discrete probability distribution function, such that $x \in \{x_1, \ldots, x_D\}$, where D is the number of states for the variable X. For example, for a boolean network, $x \in \{0, 1\}$.

We aim to study Fisher information $F(r)$ in RBNs as a function of the probability r:

$$F(r)_{RBN} \triangleq \langle F_i(r) \rangle \tag{3}$$

where $F_i(r)$ is the Fisher information of the i-th node of the RBN calculated using Equation 2.

Shannon Information

When dealing with outcomes of imperfect probabilistic processes, it is useful to define the information content of an outcome x, which has the probability $P(x)$, as $\log_2 \frac{1}{P(x)}$. Crucially, improbable outcomes convey more information than probable outcomes. Given a probability distribution P over the outcomes $x \in \mathcal{X}$ of a discrete random variable, the average Shannon information content of an outcome (Shannon, 1948) is determined by $H(X) = -\sum_{x\in\mathcal{X}} P(x)\log_2 P(x)$. We note the information is measured in bits, and henceforth omit the logarithm base 2. This quantity is known as *(information) entropy*, and may be contrasted with Fisher information in Equation 2.

In this paper we consider the entropy defined in terms of the probability distribution of the states of each node with respect to some parameter θ. Here the probabilities $p(x_j^i|\theta)$ are defined for each possible state x_j^i for each node i (given θ), and Shannon entropy

$$H(X^i|\theta) = -\sum_{x_j} p(x_j^i|\theta)\log p(x_j^i|\theta)$$

is subsequently also defined for each node i given θ, measuring the diversity of system's states. Then this quantity is averaged across the network given θ,

$$H(\theta)_{RBN} \triangleq \langle H(X^i|\theta)\rangle_i. \tag{4}$$

Results and Discussions

We model the RBNs using enhancements to Gershenson's RBNLab software (http://rbn.sourceforge.net). Each simulation run starts from an initial randomised state, ignores a short initial transient of 30 steps to allow the network to settle into the main phase of the computation, and then stops after 400 time steps. 250 repeat runs from random initial states were used for each network.

We calculate $p(x^i|r)$ of each node i in a given RBN over all the repeat runs. This likelihood is used to calculate the Fisher information at node i, thus giving us the average Fisher information of the network, $F(r) = \langle F(r)_{RBN}\rangle$. Similarly, we averaged the entropy measurements $H(r) = \langle H(r)_{RBN}\rangle$ over the network realisations for each r.

We focus on RBNs with $N = 250$ nodes and average connectivity of $\overline{K} = 4.0$, while altering the bias in the network r. $\overline{K} = 4.0$ was chosen because, with it held constant, RBNs at low and high values of r exhibit ordered behaviour and RBNs at mid-range values of r exhibit chaotic behaviour.

Figure 1 shows that the average $F(r)$ has two peaks almost mirrored about $r = 0.5$. These peaks occur approximately at the phase transition between the chaotic and ordered phases. This indicates that close to the phase transition, there is a large increase in the information in the state distribution of the nodes about the parameter r. On the other hand, deep inside the ordered and chaotic phases, the state distribution of the nodes indicates little about r, other than that the network is in one of these phases.

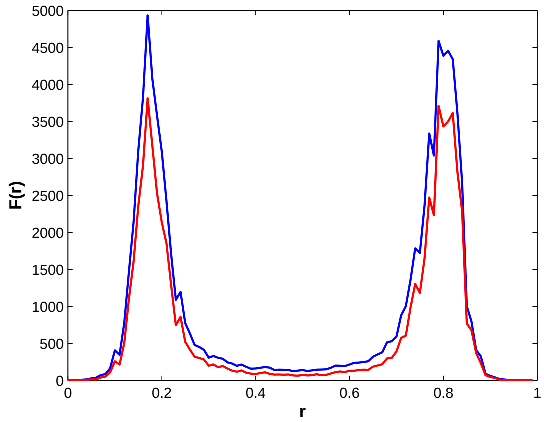

Figure 1: Average Fisher information $F(r)$ versus the bias of the network, r, for networks of size $N = 250$ and average connectivity $\overline{K} = 4.0$. The blue curve shows the Fisher information if we take into account all the nodes in the network, the red curve shows the Fisher information if we ignore those nodes whose logic table has changed due to the change in parameter r, from $p(x|r)$ to $p(x|r+\Delta r)$.

Recently, Frank (2009) proposed that Fisher information is equivalent to the acceleration of Shannon information, i.e. the second derivative of $H(X|\theta)$ with respect to θ. This was shown under certain assumptions, e.g., the assumption that the outer (or averaging) term $p(x|\theta)$ holds constant while differentiating $H(X|\theta)$, thus differentiating $\log p(x|\theta)$ only. However, these assumptions may be too strong, and here we report on similarity between Fisher information and *first* derivative of Shannon information.

Figure 2 shows the derivatives of Shannon information $H(r)$ versus network bias r for RBNs with average connectivity of $\overline{K} = 4.0$: the square of the first derivative of Shannon information, $(\frac{d}{dr}H)^2$, and the second derivative, $\frac{d^2}{dr^2}H$. A comparison with Figure 1 reveals that Fisher information for RBNs is more qualitatively similar to the square of rate of change of Shannon information than the acceleration of Shannon information. Specifically, the peaks for $(\frac{dH}{d\theta})^2$ occur at $r = 0.21$ and $r = 0.79$, coinciding with the Fisher information peaks shown in Figure 1. However, there is a difference in their orders of magnitude — this is because in finding $F(\theta)$, we first differentiate and then square and average the values, while for $(\frac{dH}{d\theta})^2$ we average and then differentiate and square the values.

Let r_{max} denote the maximum Fisher information that occurs with respect to r for fixed $\overline{K}$. Formally, r_{max} for every $\overline{K}$ is set to the global maxima of $F(r)$ in two regions: $0 \le r \le 0.5$ and $0.5 \le r \le 1$. For example, r_{max} correspond to the peaks shown in Figure 1. Figure 3 shows the plot of r_{max} as a function of $\overline{K}$, contrasting it with the

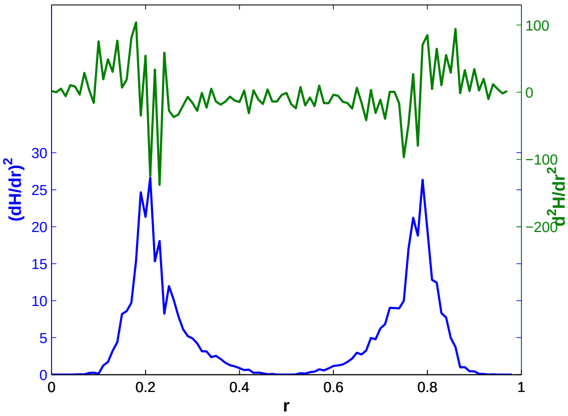

Figure 2: Derivatives of Shannon Information, $H(r)$, for the same networks as Figure 1 ($\overline{K} = 4.0$). (blue) First derivative of H squared. (green) Second derivative of H.

theoretical critical phase (edge of chaos) of the RBNs. We can see from the figure that the phase diagram obtained by maximising Fisher information generally follows the same shape, but is bounded by the theoretical curve for critical K_c versus r (Kauffman, 1993; Gershenson, 2004).

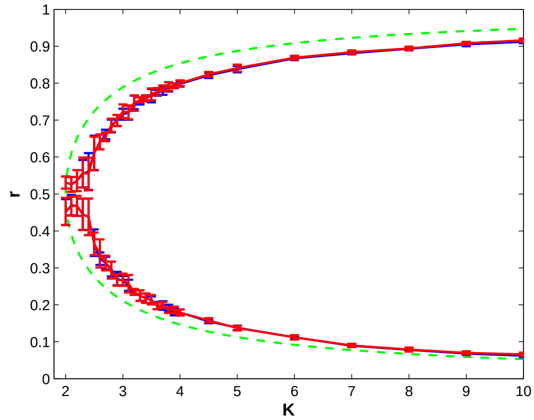

Figure 3: Phase diagram of r_{max} where the maximum Fisher information, $F(r)$, occurs with respect to r for fixed $\overline{K}$, as a function of $\overline{K}$. Blue: when all the nodes in the network were taken into account; Red: when those nodes whose logic table has changed due to the change in parameter r were ignored. The error bars on the curves show the standard deviation of r_{max}. The green dashed line is the theoretical curve for critical K_c versus r.

Conclusion

In this paper, we contrasted Fisher information and Shannon information in the context of Random Boolean Networks (RBNs). RBNs are known to exhibit three distinct phases of dynamics, depending on their parameters: ordered, chaotic and critical, and we analysed the phase diagram of RBN dynamics interpreted in information-theoretic terms.

Both the activity level r and average connectivity K play the role of control parameters, and the phase diagram is obtained by plotting (K, r) points that separate the ordered and chaotic phases. Fisher information about the control parameters was observed to have maxima at the critical (K, r) points. This is because $F(r)$ measures (locally) the amount of information that RBN dynamics carry about the parameter r, and these dynamics are most sensitive to the control parameter near the critical point.

Our analysis showed that an information-theoretic interpretation of the phase diagram (K with respect to r) reveals expected phases (ordered, chaotic and critical) as well as symmetry breaking (slightly obscured by finite-size effects). In addition, the comparison between Fisher information $F(r)$ and a square of a first derivative of Shannon information $H(r)$ uncovered their strong qualitative similarity, albeit separated by an order of magnitude. The analysis shed more light on connections between Fisher information and (derivatives of) Shannon information, and provided a means for further rigorous information-theoretic studies of phase transitions in complex networks.

References

Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Phil. Trans. of the Royal Soc., A*, 220:309–368.

Frank, S. A. (2009). Natural selection maximizes Fisher information. *Journal of Evolutionary Biology*, 22:231–244.

Gershenson, C. (2004). Introduction to random boolean networks. In Bedau, M., Husbands, P., Hutton, T., Kumar, S., and Suzuki, H., editors, *Proceedings of the Workshops and Tutorials of 9 International Conf. on the Simulation and Synthesis of Living Systems (ALife IX), Boston, USA*, pages 160–173.

Kauffman, S. A. (1993). *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford Univ. Press, New York.

Langton, C. G. (1990). Computation at the edge of chaos: phase transitions and emergent computation. *Physica D*, 42(1-3):12–37.

Lizier, J. T., Prokopenko, M., and Zomaya, A. Y. (2008). The information dynamics of phase transitions in random boolean networks. In Bullock, S., Noble, J., Watson, R., and Bedau, M. A., editors, *Proceedings of the 11 International Conf. on the Simulation and Synthesis of Living Systems (ALife XI), Winchester, UK*, pages 374–381, Cambridge, MA. MIT Press.

Rämö, P., Kauffman, S., Kesseli, J., and Yli-Harja, O. (2007). Measures for information propagation in boolean networks. *Physica D*, 227(1):100–104.

Ribeiro, A. S., Kauffman, S. A., Lloyd-Price, J., Samuelsson, B., and Socolar, J. E. S. (2008). Mutual information in random boolean models of regulatory networks. *Physical Review E*, 77(1):011901–10.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 623–656, July, October.