Teoria da Informação e Dinâmica Coletiva em Redes Complexas

Jean Martins,
jean@icmc.usp.br

Danilo Vargas,
dvargas@icmc.usp.br

Universidade de São Paulo
Instituto de Ciências Matemáticas e Computação

São Carlos/SP, 9 de junho de 2011.

Sumário I

RC – Heterogeneidade e Correlação

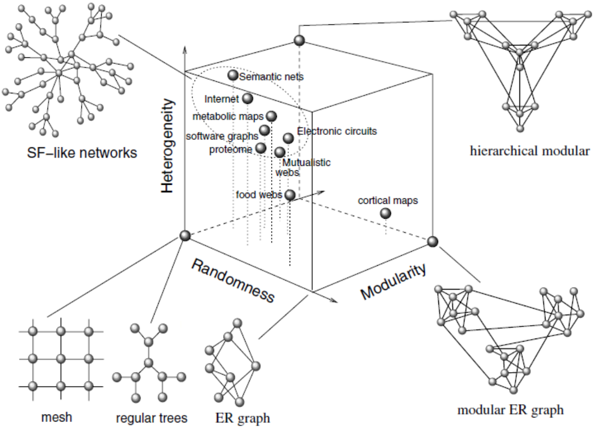


Figura: Figura originalmente publicada em [Solé and Valverde, 2004].

RC – Heterogeneidade e Correlação

1. *Heterogeneidade*:
    - medida pela diversidade de grau dos nós,
2. *Correlação* entre graus dos nós:
    - relação causa/efeito quanto ao grau dos nós e a ocorrências de arestas entre eles.
    - quanto maior o grau de um nó, maior a probabilidade de que esse receba novas conexões.
3. **Como medir tais características do ponto de vista da Teoria da Informação?**

TI – Entropia

Entropia:

1. Medida da incerteza média de uma variável aleatória,
2. O número médio de bits (logaritmo base 2) necessário para descrever tal variável.

Definição (Entropia [Cover et al., 1991])

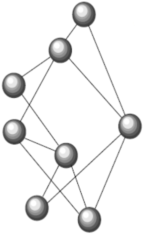
Seja X uma variável aleatória discreta em um alfabeto ξ, função de probabilidade $p(X = x)$, com $x \in \xi$. A entropia $H(X)$ é definida como:

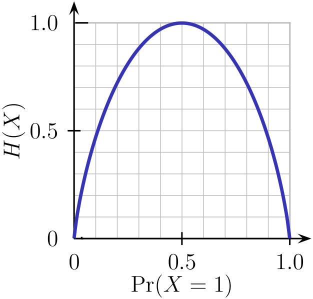
$$H(X) = -\sum_{x \in \xi} p(x) \log p(x) \tag{1}$$

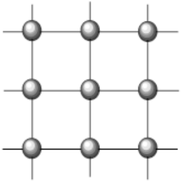
O que pode-se concluir?

TI – Entropia

Conclusão: A entropia pode ser utilizada como medida para a heterogeneidade de uma rede.







TI – *Mutual Information*

Mutual Information:

1. Mutual Information se refere à quantidade de informação que uma variável aleatória possui sobre outra variável.
2. É a redução de incerteza de uma variável aleatória, dado o conhecimento sobre outra variável aleatória.

Definição (Mutual Information [Cover et al., 1991])

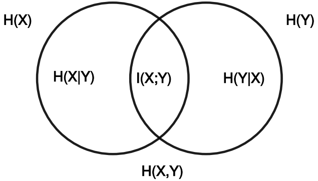
Considere duas variáveis aleatórias X e Y com uma função de probabilidade conjunta $p(X = x, Y = y)$, e probabilidades marginais $p(X = x)$ e $p(Y = y)$. A Mutual Information é definida como:

$$I(X, Y) = \sum_{x \in \xi} \sum_{y \in \xi} p(x)p(y) \log \frac{p(x, y)}{p(x)p(y)} \tag{2}$$

TI – *Mutual Information*

O que pode-se concluir?
A mutual information pode ser utilizada como medida de correlação em redes complexas.

$$I(X,Y) = H(X) - H(X|Y)$$



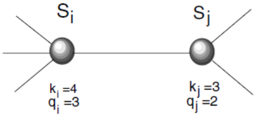
Resumo RC & TI

Características da rede que serão avaliadas:

- Heterogeneidade da rede, medida pela **Entropia**.
- Quantidade de correlação entre os nós da rede: **Mutual Information**,

Teoria da Inf. & Redes Complexas [Solé and Valverde, 2004]

- Foco na **distribuição de grau "restante"** (*remaining degree distribution*),



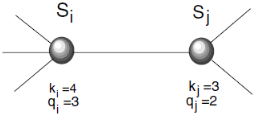
- Que é definida como o número de arestas deixando o vértice i excluindo-se a aresta pela qual chegou-se a tal vértice:

$$q(k) = \frac{(k+1)P_{k+1}}{\langle k \rangle} \tag{3}$$

Teoria da Inf. & Redes Complexas

- Tem-se também a **distribuição conjunta de grau "restante"** $q(k, k')$, a qual assume o valor $q(k)q(k')$ em redes com baixa correlação.
- Caso contrário, tem-se a Equação (4), onde $\pi(k|k')$ é a probabilidade de se observar um vértice com grau k, dado que o outro vértice pertencendo à mesma aresta possua grau k'.

$$q(k, k') = \frac{\pi(k|k')}{q(k')} \qquad (4)$$



TI & RC – Entropia

Utilizando a distribuição $Q = (q(1), \ldots, q(i), \ldots, q(n))$, podemos descrever a entropia e a mutual information de uma rede.

$$H(Q) = -\sum_{k=1}^{n} q(k) \log q(k). \tag{5}$$

$$I(Q) = \sum_{k=1}^{n} \sum_{k'=1}^{n} q(k,k') \log \frac{q(k,k')}{q(k)q(k')}. \tag{6}$$

Aplicação

Como aplicar as funções previamente definidas para a classificação em redes complexas?

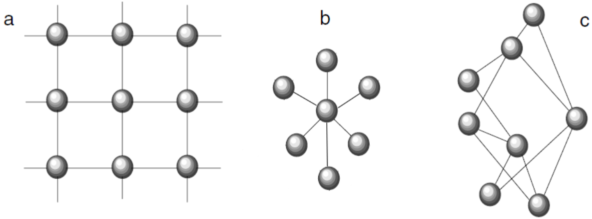


Figura: (a) Redes em Grade, (b) Rede Estrela, (c) Rede Aleatórias.

Redes em Grade

Em uma **grade**, todos os vértices possuem o mesmo grau, portanto:
Observando que $\delta_{k,z} = 1$ se $k = z$ e 0 caso contrário.

$$P_k = \delta_{k,z} \tag{7}$$

1. Desta forma $q(k)$ pode ser simplificado:

$$q(k) = \frac{(k+1)\ P_{k+1}}{\langle k \rangle}$$
$$q(k) \equiv \delta_{k,(z-1)}$$
$$Q = \{0,0,0,0,1,0,\dots,0\}$$
$$H(Q) = -\sum_{k=1}^{n} q(k) \log q(k) = 0$$

Redes em Grade

1. Todos vértices possuem o mesmo grau, portanto:

$$q(k,k') = q(k)\ q(k') = \delta_{k,(z-1)}\ \delta_{k',(z-1)}$$

2. Desta forma a *Mutual Information* é zero, o que implica que não há correlação.

$$I(Q) = \sum_{k=1}^{n}\sum_{k'=1}^{n} q(k,k') \log \frac{q(k,k')}{q(k)q(k')} = \sum_{k=1}^{n}\sum_{k'=1}^{n} \delta_{k,(z-1)}\ \delta_{k',(z-1)} \log \frac{\delta_{k,(z-1)}\ \delta_{k',(z-1)}}{\delta_{k,(z-1)}\delta_{k',(z-1)}}$$

$$I(Q) = 0$$

Rede Estrela

Padrão comumente encontrado em redes *scale-free*.

1. Um vértice de grau $n-1$, demais vértices grau 1:

$$P_k = \frac{n-1}{n}\delta_{k,1} + \frac{1}{n}\delta_{k,(n-1)}$$

2. Portanto, grau médio da rede é dado por:

$$\langle k \rangle = \sum_{k}^{n} k P_k = 2\frac{n-1}{n}$$

3. Substituindo tais valores na equação do grau "restante" $q(k)$, tem-se.

$$q(k) = \frac{1}{2}(\delta_{k,0} + \delta_{k,(n-2)})$$
$$Q = \{0.5, 0, 0, \ldots, 0.5, 0, 0\}$$
$$H(Q) = -2(\frac{1}{2}\log\frac{1}{2} = -2(\frac{1}{2}(\log 1 - \log 2)) = 1$$

Rede Estrela

1. A distribuição de grau "restante" obtida através da equação original é:

$$\begin{aligned} q(k,k') &= q(k')\pi(k|k') \\ &= \frac{1}{2}\left[\delta_{k,0} + \delta_{k,(n-2)}\right]\left[\delta_{k',0} + \delta_{k',(n-2)}\right] \\ q(k,k') &= \frac{1}{2}(\delta_{k,0}\delta_{k',(n-2)} + \delta_{k,(n-2)}\delta_{k',0}) \end{aligned}$$

2. Usando os resultados anteriores, calculamos uma medida da correlação de uma rede estrela.

$$\begin{aligned} I(Q) &= \sum_{k=1}^{n}\sum_{k'=1}^{n} q(k,k')\log\frac{q(k,k')}{q(k)q(k')} \\ &= 2\left[\frac{1}{2}\log\left(\frac{2\delta_{k,0}\delta_{k',(n-2)}}{\delta_{k,0}+\delta_{k,(n-2)}}\right)\right] \\ I(Q) &= \log 2 = 1 \end{aligned} \tag{8}$$

Sumário I

Motivation

- The use of networks of dynamical systems are widespread. But what can be said of the collective dynamics of such systems when they are coupled together?
- Answer: little[Strogatz, 2001].
- It may unravel new ways of doing things. Bioinspiration: The brain uses it, so should not we?.

Sumário I

Master Stability Function

The master stability function (MSF) is a framework developed to study the synchronization state complex of network topologies independent of the peculiarities of the oscillators. It is based on the premise that all nodes are identical dynamical units, described by the following equation:

$$\frac{\partial x_i}{\partial t} = F(x_i), \tag{9}$$

where x_i is a m-dimensional vector and $F(x_i)$ is its evolution function (which is the same for every node). The output of the system is described by $H(x_i)$, which is also identical for all N dynamical units and is a coupling map of the nodes. For example, in y-coupled oscillators $H(x_i) = (0, y, 0)$, where they are only communicating with each other through the y component. The complex network is defined by the node's adjacencies a_{ij} and their respective weights w_{ij} (for unweighted matrix consider $w_{ij} = 1$) expressed by the Equation 10.

$$\frac{\partial x_i}{\partial t} = F(x_i) + \sigma \sum_{i=1}^{N} a_{ij} w_{ij} (H(x_j) - H(x_i)) \quad (10)$$

$$= F(x_i) + -\sigma \sum_{j=1}^{N} G_{ij}(H(x_j)), \quad (11)$$

where σ is the uniform coupling strength ($\sigma > 0$ for diffusive coupling) and G is a coupling matrix defined by next Equation.

$$G_{ij} = \begin{cases} -a_{ij} w_{ij} & i \neq j \\ \sum_{j=1}^{N} a_{ij} w_{ij} & i = j \end{cases}, \quad (12)$$

it follows that G has zero row-sum. Then the synchronization occurs when for all nodes the coupling term vanishes. When this occurs, the nodes will be based solely on their internal dynamics which happen to be the same for all N nodes. Resulting in a state called synchronization manifold, which happens when:

$$x_1(t) = x_2(t) = ... = x_N(t) = x_s(t). \quad (13)$$

The synchronization manifold occurs as a consequence of Equation 13 because the matrix is a zero row-sum and the function $H(x)$ is the same for any node. The stability of synchronization is secured when the system remains entirely inside the synchronization manifold. However, in the real world system are susceptible to perturbations, in this scenario the following equation holds:

$$\frac{\partial \tilde{x}_i}{\partial t} = [JF(x_s) - \sigma\lambda_i JH(x_s)]\tilde{x}_i, \tag{14}$$

where $\tilde{x}_i$ is the deviation from the $x_s(t)$ for $x_i(t)$, J is the Jacobian operator and λ is the eigenvalue of the respective m conditional Lyapunov exponents. Conditional Lyapunov exponents for each value of $\sigma\lambda_i$ can now be acquired.

Lyapunov exponent

The Lyapunov exponent is defined as the rate of separation of two orbits. It is a measure of sensitivity of a given dynamic system from its initial conditions, exactly calculated by:

$$\lambda = \lim_{t\to\infty} \lim_{\delta x_0 \to 0} \frac{1}{t} ln \frac{|\delta x(t)|}{|\delta x_0|}, \quad (15)$$

where δx_0 is the initial difference between the two orbits and $\delta x(t)$ is the difference between the two orbits in instant t. A dynamics system with m-dimensional phase space will have m exponents, forming the Lyapunov spectrum $\lambda_1, \lambda_2, ..., \lambda_m$. The largest value of the Lyapunov spectrum is known as the maximum Lyapunov exponent (MLE) or simply λ_{max}.

- Lyapunov exponents are used in synchronization to verify the stability of a system under small perturbations of its orbit.
- If $\lambda_{max} < 0$ the system is stable because all the remaining Lyapunov is also smaller than λ_{max} which is negative.
- On the other hand, if $\lambda_{max} \geq 0$ the system is said to be unstable, because one exponent is already positive and the system will have by definition an exponential divergence of orbits.
- Not a sufficient condition! - It is important to note that although necessary the Lyapunov exponents do not offer a sufficient condition for the stability of the system. As it will be seen later in the unusual behavior of chaotic systems.

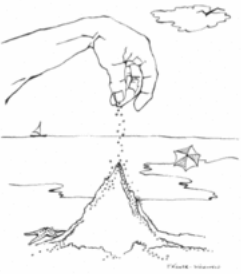
Extensions of the master stability function are still under research, recent results are the MSF near identical systems [Sun et al., 2009].

Identical Oscillators

- **Synchronize or form network dependent pattern** When the set of identical oscillators are coupled by smooth interactions. They often synchronize or form patterns dependent of the network symmetry [Collins and Stewart, 1993].
- **Integrate-and-fire oscillator will fire in unison, independent of network** When oscillators communicate by sudden impulses as commonly found in biology such as neuron spikes. It was proved that N identical integrate-and-fire oscillators (IFO) connected in a all-to-all network will end up firing in unison independent of their initial state [Peskin, 1975], [Mirollo and Strogatz, 1990].
- **Synchrony or self-organized criticality** With such simple models it is yet possible to observe systems which result in synchronization or self-organized criticality [Corral et al., 1995].

Self-Organized Criticality Example

Sandpile Example [Funch, 2008]



Fire Model

Self-Organized Criticality

- **Critical point as attractor** The behavior of a system to self-organize around a critical point is called self-organized criticality (SOC) [Bak et al., 1987],[Turcotte, 1999]. The self-organize capacity is defined as a inner dynamic property of the system. Which independent of its parameters or interferences would drive the system toward a given state.
- **Varied Definitions** Critical point has a varied number of definitions depending on the context it is applied. In mathematics it is the point where either the derivative is 0 or it is non differentiable [Stewart, 2008]. In physics it is where a phase boundary (such as the vapor-liquid point) ceases to exist [Cengel and Boles, 2006].
- **Skepticism** Sometimes, a broad definition definition is used defining a point where a system properties change suddenly to be a critical point. This concept was developed in 1987 but is still view with skepticism.

Sumário I

Kuramoto Model

The Kuramoto model is composed of N oscillators in a complete graph (all-to-all connections, illustrated in Figure 3) and each one has the following dynamics [Kuramoto, 2003]:

$$\frac{\partial \theta_i}{\partial t} = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i), \qquad i = 1 \ldots N, \qquad (16)$$

where ω_i is the natural frequency of the ith oscillator, θ_i is its phase and K is the coupling strength (identical for all edges). The ω_i were drawn from a Lorentzian distribution defined by:

$$f(\omega;\omega_0,\gamma) = \frac{1}{\pi}\left[\frac{\gamma}{(\omega-\omega_0)^2+\gamma^2}\right], \qquad (17)$$

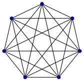


Figura: Diagram of a complete graph illustrating the scenario where the Kuramoto model is applied.

To solve this model analytically when $N \to \infty$, Kuramoto applied the following transformation:

$$re^{i\psi} = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j}, \tag{18}$$

where ψ is the average phase and r measures the coherence of the oscillators. Resulting in the following equation:

$$\frac{\partial \theta_i}{\partial t} = \omega_i + Kr\sin(\psi - \theta_i) \tag{19}$$

In the limit (as $N \to \infty$ and $t \to \infty$), Kuramoto found that the following behavior occurs:

$$r = \begin{cases} 0, & K < K_c \\ \sqrt{1 - \frac{K_c}{K}}, & K \geq K_c \end{cases}, \tag{20}$$

where $K_c = 2\gamma$.

- $K < K_c$ the oscillators are not synchronized.
- $K \geq K_c$ part of the oscillators synchronize, or more specifically lock their phases ($\frac{\partial \theta_i}{\partial t} = 0$) and part are rotating out of synchrony.
- When coupling increases to the limit as $K \to \infty$, the oscillators become totally synchronized approximately to their average phase $\theta_i \approx \psi$ at the same time that $r \to 1$.
- Extensions to the Kuramoto model are still in research [Martens et al., 2009], for the complete proof or more information on variations of the Kuramoto model, such as general frequency distributions, refer to [Acebrón et al., 2005].

Sumário I

Bursting

In y-coupled Rössler systems, after reaching synchronous state above a coupling threshold (i.e., after the longest-wavelength mode becomes stable), they still present a difference between the average x and its component x shown in Figure 4. This difference is expected to be close to 0 in synchronized systems [Ashwin et al., 1994],[Pecora et al., 1997]. This may appear because of unstable periodic orbits (UPO).

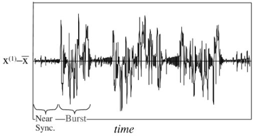


Figura: The difference from the average observed in the x variable of Rössler systems. Image from [Pecora et al., 1997].

Short-length Bifurcation

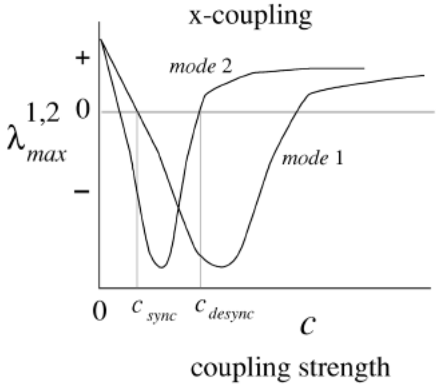


Figura: Stability diagram showing the short-length bifurcation of x-coupled Rössler systems. Image from [Pecora et al., 1997].

Size Effect

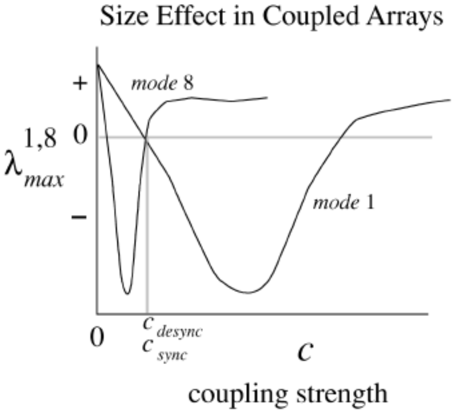


Figura: Stability diagram showing the size effect of 16 x-coupled Rössler systems. Image from [Pecora et al., 1997].

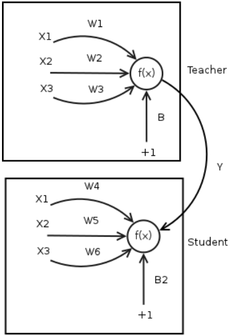
Robustness of Scale-free Networks and its Synchronization Consequences

- Random removal of nodes in a scale-free network will not affect greatly.
- Reason: there are exponentially more nodes with lower degree than higher degree nodes.
- However, scale-free networks are more susceptible to intentional attacks than random networks [Cohen et al., 2001].
- This robustness (or vulnerability) also affects the collective dynamics of scale-free networks, since synchronization inside scale-free networks can sustain random removal of 5% of its nodes with minor affects in behavior. While 1% of wisely selected nodes might destroy completely the synchronization [Lü et al., 2004], [Wang and Chen, 2002].

Sumário I

Teacher and Student

The situation where a student node learn from a teacher node. Which is exactly the interaction defined in supervised learning [Kotsiantis et al., 2007]. The unique difference is that in the supervised learning setting the teacher is not specifically defined as a network node. Dynamics in this simple situation are well defined and studied, an example using two perceptrons is shown on Figure 7.



Self-Interacting

Suppose now the situation where an already trained node is interacting with itself [Bornholdt et al., 2003]. The node can be trained in any arbitrary sequence. This time, however, the node is learning the opposite of its own prediction. Which leads to a situation where its prediction error is 100% and the sequence produced by the node is close to random [Metzler et al., 2001]. When a second Boolean perceptron is added to predict the sequence, it has 78% of prediction error. Which is somewhat better than the self-interacting node, but still worse than a 50% random guessing [Zhu and Kinzel, 1998].

The Figure 8 illustrates the example with two perceptrons, note that this example can be extended to other learning machines.

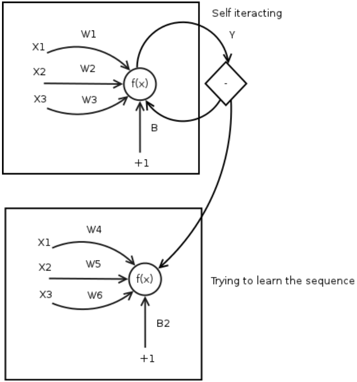


Figura: Example of dynamics a self teaching node and another perceptron trying to learn the output. Where W variables are weights, X are input, B are the bias, $f(x)$ is the activation function and Y is the result from the teacher. The rectangles isolate the nodes of the two units.

Referências I

Acebrón, J., Bonilla, L., Vicente, C., Ritort, F., and Spigler, R. (2005).
The kuramoto model: A simple paradigm for synchronization phenomena.
Reviews of modern physics, 77(1):137.

Ashwin, P., Buescu, J., and Stewart, I. (1994).
Bubbling of attractors and synchronisation of chaotic oscillators.
Physics Letters A, 193(2):126–139.

Bak, P., Tang, C., and Wiesenfeld, K. (1987).
Self-organized criticality: An explanation of the 1/f noise.
Physical Review Letters, 59(4):381–384.

Referências II

Bornholdt, S., Schuster, H., and Wiley, J. (2003).
Handbook of graphs and networks.
Wiley Online Library.

Cengel, Y. and Boles, M. (2006).
Thermodynamics: an engineering approach, volume 988.
McGraw-Hill Higher Education.

Cohen, R., Erez, K., Ben-Avraham, D., and Havlin, S. (2001).
Breakdown of the internet under intentional attack.
Physical Review Letters, 86(16):3682–3685.

Collins, J. and Stewart, I. (1993).
Coupled nonlinear oscillators and the symmetries of animal gaits.
Journal of Nonlinear Science, 3(1):349–392.

Referências III

Corral, Á., Pérez, C., Diaz-Guilera, A., and Arenas, A. (1995).
Self-organized criticality and synchronization in a lattice model of integrate-and-fire oscillators.
Physical review letters, 74(1):118–121.

Cover, T., Thomas, J., Wiley, J., et al. (1991).
Elements of information theory, volume 306.
Wiley Online Library.

Engel, A. and Broeck, C. (2001).
Statistical mechanics of learning.
Cambridge Univ Pr.

Funch, F. (2008).
Self-organized criticality.

Referências IV

Kotsiantis, S., Zaharakis, I., and Pintelas, P. (2007).
Supervised machine learning: A review of classification techniques.
Emerging artificial intelligence applications in computer engineering: real word AI systems with applications in eHealth, HCI, information retrieval and pervasive technologies, 160:3.

Kuramoto, Y. (2003).
Chemical oscillations, waves, and turbulence.
Dover Pubns.

Lü, J., Chen, G., and Cheng, D. (2004).
A new chaotic system and beyond: the generalized lorenz-like system.
Int. J. Bifurc. Chaos, 14(5):1507–1537.

Referências V

- Martens, E., Barreto, E., Strogatz, S., Ott, E., So, P., and Antonsen, T. (2009).
  Exact results for the kuramoto model with a bimodal frequency distribution.
  Physical Review E, 79(2):026204.
- Metzler, R., Kinzel, W., Ein-Dor, L., and Kanter, I. (2001).
  Generation of unpredictable time series by a neural network.
  Physical Review E, 63(5):056126.
- Mirollo, R. and Strogatz, S. (1990).
  Synchronization of pulse-coupled biological oscillators.
  SIAM Journal on Applied Mathematics, 50(6):1645–1662.

Referências VI

- Pecora, L., Carroll, T., Johnson, G., Mar, D., and Heagy, J. (1997).
  Fundamentals of synchronization in chaotic systems, concepts, and applications.
  Chaos: An Interdisciplinary Journal of Nonlinear Science, 7:520.

- Peskin, C. (1975).
  Mathematical aspects of heart physiology.
  Courant Institute of Mathematical Sciences, New York University.

Referências VII

Solé, R. and Valverde, S. (2004).
Information theory of complex networks: On evolution and architectural constraints.
In Ben-Naim, E., Frauenfelder, H., and Toroczkai, Z., editors, *Complex Networks*, volume 650 of *Lecture Notes in Physics*, pages 189–207. Springer Berlin / Heidelberg. 10.1007/978-3-540-44485-5_9.

Stewart, J. (2008).
Calculus: early transcendentals.
Cengage Learning EMEA.

Strogatz, S. (2001).
Exploring complex networks.
Nature, 410(6825):268–276.

Referências VIII

Sun, J., Bollt, E., and Nishikawa, T. (2009).
Master stability functions for coupled nearly identical dynamical systems.
EPL (Europhysics Letters), 85:60011.

Turcotte, D. (1999).
Self-organized criticality.
Reports on progress in physics, 62:1377.

Wang, X. and Chen, G. (2002).
Synchronization in scale-free dynamical networks: robustness and fragility.
Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on, 49(1):54–62.

Referências IX

Zhu, H. and Kinzel, W. (1998).
Antipredictable sequences: harder to predict than random sequences.
Neural computation, 10(8):2219–2230.