

Caracterização Topológica de Dinâmica de Redes Biológicas Modeladas por Redes Booleanas

Aluno: Miguel Pereira Ostrowski

Orientador: Prof. Dr. Ronaldo Fumio Hashimoto

Resumo

As interações entre genes são as responsáveis pela coordenação da execução de funções biológicas, como a divisão celular. Tais interações produzem uma dinâmica que pode ser representada por um diagrama de estados. Tal diagrama pode possuir diversas propriedades, com a finalidade de deixar o sistema mais estável e robusto a pequenas perturbações provindas do ambiente. Uma dessas propriedades é o tamanho da bacia de atração. Neste trabalho pretende-se verificar quais características das interações gênicas possuem a propriedade de gerar redes mais robustas contra pequenas interferências, utilizando Redes *Booleanas* como modelagem.

1 Introdução

Dentro da área de Biologia de Sistemas, muitos modelos matemáticos e computacionais estão sendo desenvolvidos para explicar interações entre seus componentes e existe um número considerável de tentativas para modelar redes biológicas. Um modelo que tem chamado muita atenção é o modelo de redes *Booleanas*, ou *BNs* (*Boolean Networks*), onde utiliza-se grafos direcionados como modelagem a fim de representar uma interação entre genes. Neste grafo, chamado de grafo de interação gênica, cada vértice (ou nó) representa um gene; e um arco liga dois nós se, e somente se, existir uma interação (de ativação ou inibição) do primeiro gene sobre o segundo (ou, analogamente, o segundo gene sofre uma interação do primeiro). Tal grafo, com n genes, pode ser representado por uma matriz de interação gênica $A_{n \times n}$, tal que:

$$a_{i,j} = \begin{cases} 1, & \text{se } x_j \text{ é ativado por } x_i \\ -1, & \text{se } x_j \text{ é inibido por } x_i \\ 0, & \text{se não há interação entre } x_i \text{ e } x_j \end{cases} \quad (1)$$

Um exemplo de um grafo de interações e sua matriz correspondente podem ser vistos na Figura 1a e 1b, respectivamente.

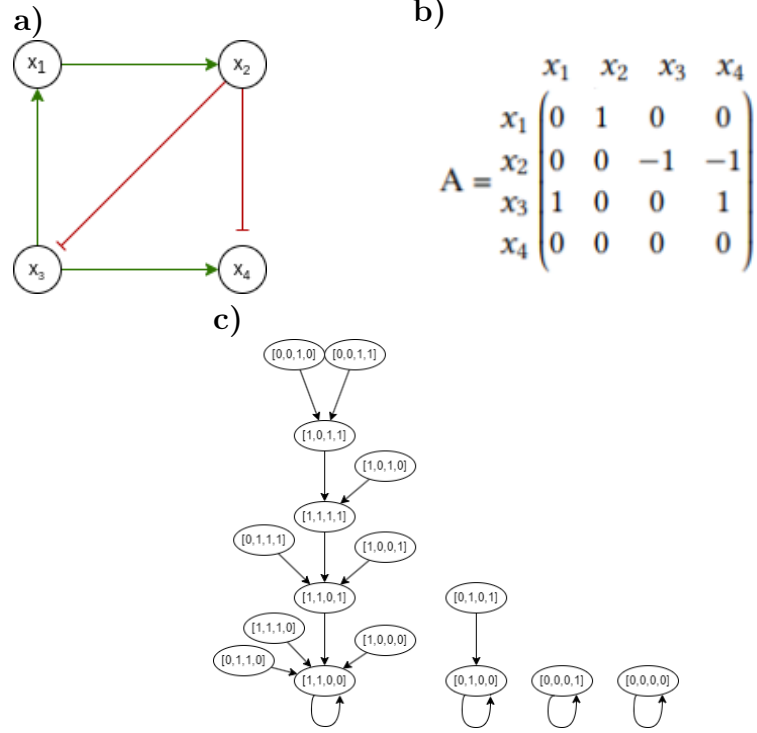


Figura 1: a) Grafo da interação gênica; b) Matriz de interação gênica; c) Diagrama de estados gerado.

Considerando que cada gene x_j em um determinado tempo t pode assumir valor 1 ou 0 (indicando se está ativado ou desativado, respectivamente), podemos, juntamente com a rede *Booleana*, definir a função $\phi_j(t)$, chamada de sinal de entrada no gene x_i no tempo t , a seguir:

$$\phi_j(t) = \sum_{j=1}^n a_{i,j} \cdot x_j(t) \quad (2)$$

Com base na equação podemos calcular no tempo seguinte se um gene é ativado, desativado ou mantém seu estado atual, por meio da seguinte equação:

$$x_j(t+1) = \begin{cases} 1, & \text{se } \phi_j(t) > 0 \\ 0, & \text{se } \phi_j(t) < 0 \\ x_j(t), & \text{se } \phi_j(t) = 0 \end{cases} \quad (3)$$

Como cada gene $x_i(t)$ no tempo t pode assumir somente 2 valores, o vetor

$\mathbf{s}(\mathbf{t})=(x_1(t), \dots, x_n(t))$ formado pelos valores binários das n variáveis da BN forma uma n -*upla* pertencente ao conjunto $\{0, 1\}^n$. Definimos, dessa forma, o estado de uma BN com n variáveis no tempo t como sendo um vetor binário $\mathbf{s}(\mathbf{t}) \in \{0, 1\}^n$. É fácil de ver que uma BN com n genes possui 2^n estados possíveis. Dado um estado no tempo t , a atualização de todos os genes no tempo $t + 1$ leva a rede para o próximo estado $\mathbf{s}(\mathbf{t} + \mathbf{1})$. Podemos representar estas transições de estado através de um diagrama de transição de estados. Veja um exemplo na Figura 1c.

Uma vez que o número de estados de uma BN é finito e que a rede sempre transita de um estado para outro, é fácil ver que depois do sistema iterar por um determinado tempo, certos estados serão revisitados de uma maneira cíclica. Tais estados formam o que chamamos de atratores. Um atrator é um ciclo dirigido (ou um ponto fixo) no diagrama de transição estados da BN . Dado um atrator, o conjunto de todos os estados que conduzem a este atrator (inclusive os próprios estados dentro do atrator) é chamado de bacia de atração. A dinâmica de uma BN pode ser caracterizada pela quantidade de atratores, pelo tamanho médio dos ciclos atratores, pelo tamanho das bacias de atração, entre outras medidas do diagrama de estados.

2 Redes Biológicas modeladas por Redes Booleanas

O ciclo celular da levedura é composto por quatro fases distintas: G_1 (crescimento da célula), S (replicação de DNA), G_2 (preparação para mitose) e M (divisão celular). Neste projeto, nós vamos considerar que o sinal para o início do ciclo celular é o tamanho da célula, indicada pela variável T . As três primeiras fases são conhecidas como intérfase e a última é a mitose. Este ciclo é altamente regulado no sentido de que uma fase não pode ser iniciada sem que a fase anterior tenha sido concluída. Este mecanismo propõe um certo nível de robustez pois as fases devem ser realizadas em sequência para que o ciclo seja bem sucedido. Uma vantagem de se trabalhar com o ciclo celular da levedura é que este processo é relativamente conhecido e bem estudado devido a simplicidade do organismo.

Após uma exaustiva investigação na literatura sobre redes de regulação gênica da levedura, Li et al. [1] modelaram o sistema de controle do seu ciclo celular sob o modelo de redes Booleanas e estudaram suas propriedades dinâmicas e estruturais. A rede de regulação gênica proposta por Li et al. é composta de 11 genes/proteínas representados por x_1, x_2, \dots, x_{11} de forma que $x_i = 1$ e $x_i = 0$ representam o estado ativo e inativo do gene/proteína i , respectivamente. Assim, um estado da rede de Li et al. [1] é uma configuração do espaço $S = \{0, 1\}^{11}$. Temos então 2^{11} estados, rotulados por inteiros do conjunto $\{0, 1, \dots, 2047\}$.

A rede de interação proposta por Li et al. [1] é mostrada na Figura 2. A partir desta figura, é possível construir uma matriz de regulação $A_{11 \times 11}$ da forma descrita

na Eq. 1: (i) uma flecha de um gene i para um gene j representa “ativação” do gene j pelo gene i , e neste caso, $a_{ji} = 1$; (ii) uma linha de um gene i para um gene j com uma barra no final representa “repressão” do gene j pelo gene i , e neste caso $a_{ji} = -1$; (iii) laços pontilhados representam auto-degradação, e neste caso, $a_{ii} = -1$; (iv) para os casos restantes, $a_{ji} = 0$.

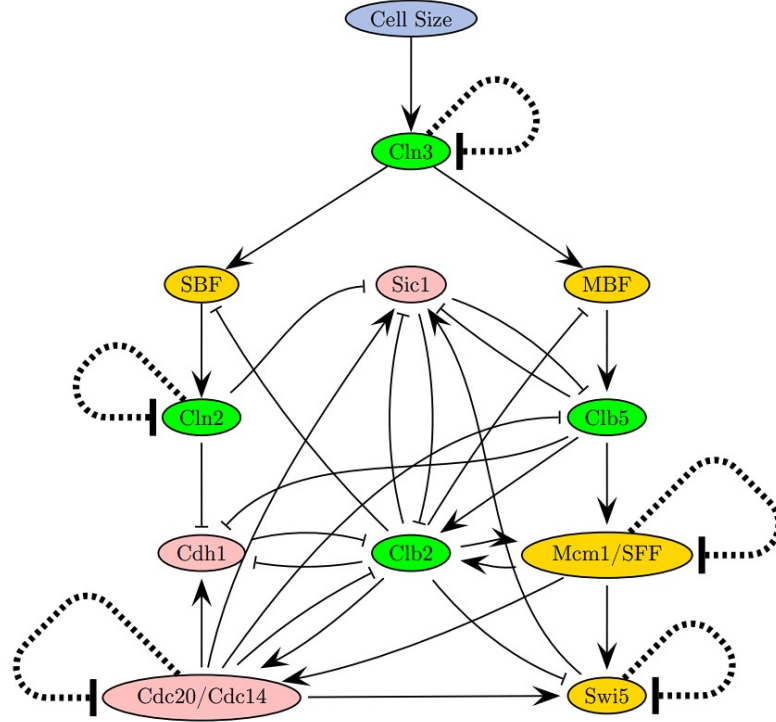


Figura 2: Rede de interações proposta por Li et al [1].

A dinâmica da rede do Li et al. [1] tem 7 atratores (todos pontos fixos) e seus tamanhos estão apresentados na tabela abaixo:

Basin Size	Cln3	MBF	SBF	Cln1,2	Cdh1	Swi5	Cdc20	Cln5	Sic1	Cln2	Mcm1
1,764	0	0	0	0	1	0	0	0	1	0	0
151	0	0	1	1	0	0	0	0	0	0	0
109	0	1	0	0	1	0	0	0	1	0	0
9	0	0	0	0	0	0	0	0	1	0	0
7	0	1	0	0	0	0	0	0	1	0	0
7	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	1	0	0	0	0	0	0

Figura 3: Tamanho de cada bacia de atração (primeira coluna) com seu respectivo atrator.

A bacia de atração de maior tamanho tem aproximadamente 80% dos estados possíveis e está representada na Figura 6. Nesta bacia há uma trajetória com 13 estados (marcada pelos arcos em azul) que representa o ciclo celular da levedura

(modela um fenômeno biológico). Esta trajetória de estados é mostrada na tabela abaixo:

State	Cln3	MBF	SBF	Cln2	Cdh1	Swi5	Cdc20	Clb5	Sic1	Clb2	Mcm1	Phase
s_1	1	0	0	0	1	0	0	0	1	0	0	Excited G_1
s_2	0	1	1	0	1	0	0	0	1	0	0	G_1
s_3	0	1	1	1	1	0	0	0	1	0	0	G_1
s_4	0	1	1	1	0	0	0	0	0	0	0	G_1
s_5	0	1	1	1	0	0	0	1	0	0	0	S
s_6	0	1	1	1	0	0	0	1	0	1	1	G_2
s_7	0	0	0	1	0	0	1	1	0	1	1	M
s_8	0	0	0	0	0	1	1	0	0	1	1	M
s_9	0	0	0	0	0	1	1	0	1	1	1	M
s_{10}	0	0	0	0	0	1	1	0	1	0	1	M
s_{11}	0	0	0	0	1	1	1	0	1	0	0	M
s_{12}	0	0	0	0	1	1	0	0	1	0	0	G_1
s_{13}	0	0	0	0	1	0	0	0	1	0	0	Stationary G_1

Figura 4: Estado do caminho biológico com sua respectiva fase da divisão celular.

Observe que cada estado na trajetória corresponde a alguma fase da divisão celular da levedura.

3 Redes funcionais

Em certas análises é interessante inferir matrizes de regulação que produzam o mesmo caminho biológico no diagrama de estados. Estas redes são chamadas de redes funcionais. Com esse objetivo, as redes funcionais se tornam úteis, uma vez que podem preservar um caminho biológico do diagrama de estados.

Redes funcionais são matrizes de interação geralmente geradas a partir de restrições obtidas a partir de uma trajetória biológica do diagrama de estados $C = [s(1), s(2), \dots, s(r)]$, onde cada $s(t)$ é um estado. No caso da trajetória biológica da levedura, é possível obter aproximadamente 5×10^{34} redes funcionais (caminho que está em azul na Figura 6). Duas redes funcionais podem ser verificados na Figura 5.

4 Objetivos

Na natureza é de suma importância que os sistemas tenham todas suas funções biológicas funcionais e resistentes a perturbações. Entretanto, é extremamente comum que o ambiente interfira no funcionamento do sistema biológico, como a radiação, a temperatura, o pH, a toxicidade, entre outros fatores. Portanto, é necessário que haja mecanismos nas redes biológicas que deixem o sistema mais robusto e resiliente a interferência de fenômenos externos.

Por isso, este trabalho tem a finalidade de investigar quais propriedades das interações entre genes podem deixar a função biológica mais resistente a perturbações, utilizando Redes *Booleanas* como modelagem.

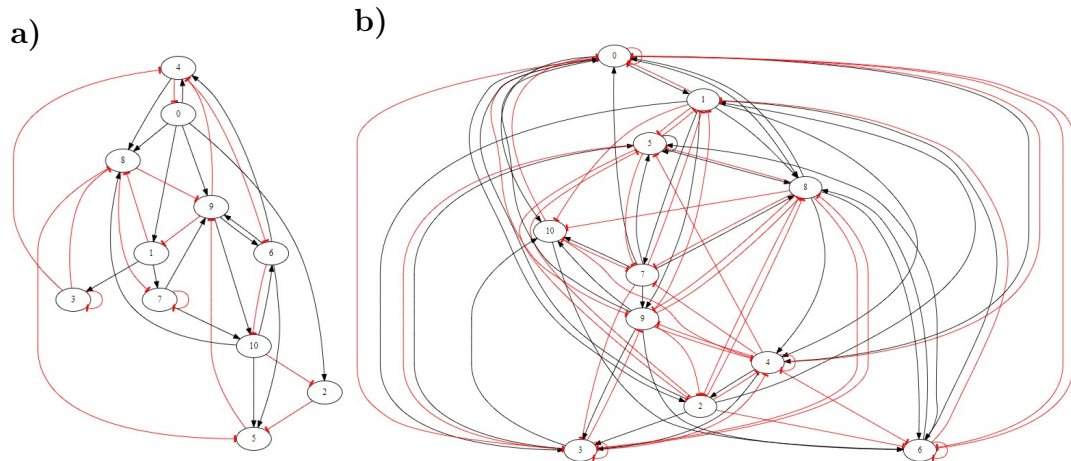


Figura 5: Redes funcionais que preservam o caminho biológico da Figura 6. Arcos pretos representam “ativação” de um gene sobre outro, enquanto os arcos vermelhos representam sua “inibição”.

5 Metodologia

Na literatura são estudadas propriedades das interações gênicas que afetam a dinâmica da rede. Como em [2] afirma, arcos de auto degradação são importantes, em modelos estocásticos, para evitar oscilações na dinâmica. Além disso, em [3] é verificado que, em modelos *Booleanos*, a quantidade de ciclos negativos possui uma correlação com a quantidade de pontos fixos. Adiante, segundo [1], a levedura estudada possui a propriedade de que o caminho biológico está na maior bacia de atração do diagrama de estados, e por isso pequenas perturbações nos estados do caminho biológico podem ter uma alta chance de voltar ao caminho biológico, assim continuando seu processo biológico normalmente.

Dessa forma, pretende-se estudar quais propriedades da matriz de regulação podem influenciar o diagrama de estados e gerar uma bacia de atração com grande quantidade de estados. Um exemplo é estudar a influência de ciclos nas interações gênicas (onde um ciclo é um caminho que começa em um gene e retorna até o mesmo gene seguindo respeitando os arcos, por exemplo o ciclo $[x_1, x_2, x_3, x_1]$ pode ser visto na Figura 1), e se há uma relação direta com o tamanho das bacias de atração.

5.1 Caminho Central

Muitas das análises da dinâmica da rede são feitas pela caracterização de bacias de atração (seus tamanhos, por exemplo) e de seus atratores (quantidade de pontos fixos, por exemplo). Apesar dessas análises, outras características poderiam ser

exploradas. Uma delas, é o caminho central. O caminho central é uma sequência de estados na qual todos os estados convergem. Um caminho central poderia ser o caminho biológico mostrado na Figura 6. Entretanto, em algumas situações podem ser mais difíceis de definir qual o caminho central, como o exemplo da Figura 7.

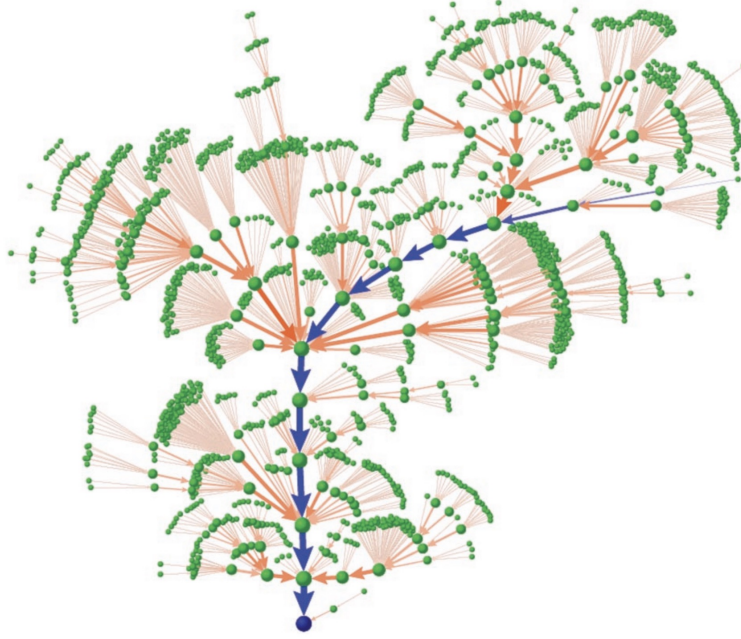


Figura 6: As trajetórias dos 1.764 estados (nós verdes) fluindo para o ponto fixo (nó azul). Os arcos entre os estados indicam a direção de um estado para outro. A sequência indicada pelos arcos em azul é o caminho biológico. O tamanho de um nó e a espessura de um arco são proporcionais ao logaritmo do número de trajetórias que passa por eles. Figura extraída de [1].

Dado isso, uma medida topológica, utilizada em [1], que tenta captar a convergência de cada estado até seu atrator é o valor W descrito a seguir:

$$W = \sum_{i=1}^N w_i/N, \quad (4)$$

onde $N = 2^n$ é a quantidade de estados na rede e w_i é calculado usando a fórmula:

$$w_i = \sum_{k=1}^{L_i} T_{k-1,k}/L_i, \quad (5)$$

sendo que L_i é a distância do estado i até seu atrator e $T_{j,k}$ é o total de fluxo que passa do estado j para o estado k . Esse fluxo é definido como sendo o número de

estados que é possível chegar ao estado k . Veja um exemplo na Figura 7.

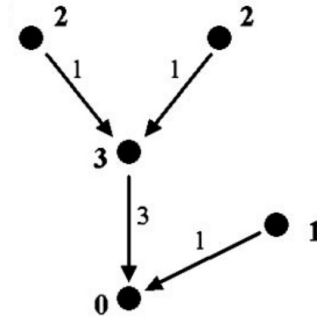


Figura 7: Diagrama de estado com representação do fluxo entre os arcos. O peso em cima de um arco é dado pela quantidade de nós que é possível passar por esse arco ($T_{j,k}$). O peso em cada nó i é a medida de convergência de sua trajetória ao atrator. Assim, a média de convergência é $W = 11/7$. Figura extraída de [1]

Em seguida, é necessário definir as formas com que são obtidos as matrizes de regulação para as análises. A primeira é utilizar matrizes aleatórias que possuem, em média, um valor W alto. E a segunda é utilizar Redes Funcionais, a partir do caminho biológico da levedura de [1].

6 Cronograma Previsto

Abaixo é apresentado o cronograma de execução, dividido em 12 meses deste e projeto.

Descrição das atividades	1°	2°	3°	4°	5°	6°	7°	8°	9°	10°	11°	12°
Estudar medidas para diagramas de estado	X	X										
Estudar medidas topológicas conhecidas			X	X	X							
Estudar medidas que influenciam o tamanho das bacias de atração				X	X							
Propor características que afetam o tamanho da bacia de atração					X	X	X					
Implementação			X	X	X	X	X	X	X			
Teste e análise dos resultados					X	X	X	X	X	X		
Escrever uma monografia							X	X	X	X	X	
Apresentar pôster												X

Referências

- [1] F. Li, T. Long, Y. Lu, Q. Ouyang, and C. Tang, “The yeast cell-cycle network is robustly designed,” *Proceedings of the National Academy of Sciences*, vol. 101, no. 14, pp. 4781–4786, 2004.
- [2] S. Bornholdt, “Boolean network models of cellular regulation: prospects and limitations,” *Journal of The Royal Society Interface*, vol. 5, no. suppl_1, pp. S85–S94, 2008.
- [3] D. Thieffry, “Dynamical roles of biological regulatory circuits1,” *Briefings in Bioinformatics*, vol. 8, pp. 220–225, 07 2007.