

Desenvolvimento de um Framework para Modelagem e Simulação de Redes Regulatórias Genéticas usando Sistemas Multiagente *


Nilzair B. Agostinho¹, Adriano V. Werhli¹, Diana F. Adamatti¹

¹Centro de Ciências Computacionais – Universidade Federal de Rio Grande (FURG)
Av. Itália, km 8, bairro Carreiros, Rio Grande, RS, Brasil.

{nilzairmb,werhli,dianaada}@gmail.com

Abstract. *Biological systems are highly complex and they are separated into individual parts to facilitate their study. The representation of biological systems as Genetic Regulatory Networks, that form a map of the interactions between the molecules in an organism, is a standard way of representing such biological complexity. As Genetic Regulatory Networks are composed of genes that are translated into transcription factors, which in turn regulate other genes. Scientists have worked on the inference and representation of Genetic Regulatory Networks. For simulation and inference purposes, many different mathematical and algorithmic models have been adopted to represent the Genetic Regulatory Networks. Among these methods, we hypothesize that Multiagent Systems are somewhat neglected. In this paper we present the first efforts to develop a simulator using the Multiagent System to model generic Genetic Regulatory Networks. To accomplish this, we develop a Multiagent System that is composed of agents that mimic the biochemical processes of gene regulation.*

Resumo. *Os sistemas biológicos são altamente complexos e a sua separação em partes individuais facilita o estudo. A representação de sistemas biológicos como Redes Regulatórias Genéticas que formam um mapa das interações entre as moléculas num organismo é uma maneira padrão de representar essa complexidade biológica. As Redes Regulatórias Genéticas são compostas de genes que são traduzidos em fatores de transcrição, que por sua vez regulam outros genes. Os cientistas trabalharam na inferência e representação de Redes Regulatórias Genéticas. Para fins de simulação e inferência, muitos modelos matemáticos e algorítmicos diferentes foram adotados para representar as Redes Regulatórias Genéticas nos últimos anos. Entre esses métodos, acreditamos que os Sistemas Multiagentes sejam um pouco negligenciados. Neste trabalho, apresentam-se os primeiros esforços para desenvolver um simulador usando o Sistemas Multiagente para modelar Redes Regulatórias Genéticas genéricos. Neste sentido, está sendo desenvolvido um Sistema Multiagente que é composto por agentes que imitam os processos bioquímicos de regulação de genes.*

*  O trabalho Desenvolvimento de um Framework para Modelagem e Simulação de Redes Regulatórias Genéticas usando Sistemas Multiagente de Nilzair B. Agostinho, Adriano V. Werhli, Diana F. Adamatti está licenciado com uma Licença Creative Commons - Atribuição 4.0 Internacional. <http://creativecommons.org/licenses/by/4.0/>

1. Introdução

Embora o dogma central da biologia afirme que a informação flui através de macro-moléculas, do DNA para o RNA e do RNA para as proteínas (Figura 1), através de Redes Regulatórias Genéticas (Genetic Regulatory Networks - GRN) que por fim funcionam através de reações bioquímicas, a vida não existiria apenas com as macro-moléculas. Então, para que o dogma central seja altamente descritivo, ele deve incluir pequenas moléculas que são os genes [Schreiber 2005].

Esses genes (pequenas moléculas) se traduzem em proteínas que podem ativar ou inibir a transcrição de outros genes. Portanto, identificar os padrões de GRN é de grande valia para uma melhor compreensão dos processos que ocorrem em nível molecular. Então, no cenário atual da área de Sistemas Biológicos, trabalha-se para o desenvolvimento de ferramentas que possam auxiliar na obtenção de novos conhecimentos sobre as interações moleculares. Por este motivo, tem sido desenvolvidas pesquisas com o objetivo principal de obter um detalhamento da organização funcional de Sistemas Biológicos [Gibas and Jambeck 2002, Lopes 2011].

Na Figura 1b pode-se visualizar os sítios de ligação e os Fatores de Transcrição (Transcription Factor - TF) que entram nos sítios de ligação e assim, de acordo com a sua quantidade e as regras de interação existentes, podem produzir ou inibir uma proteína, que é o produto final.

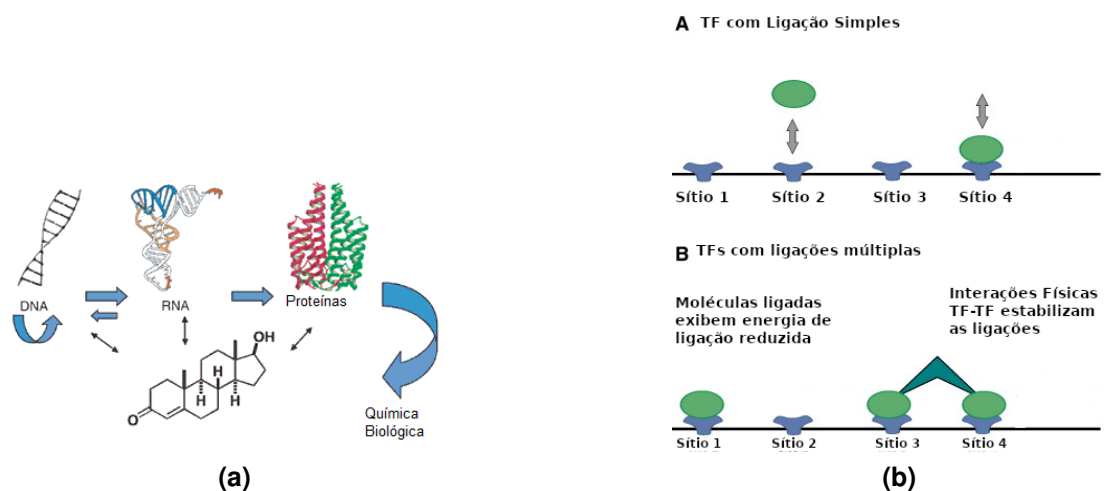


Figura 1. (a) Dogma Central da Biologia: replicação, transcrição e, finalmente o produto final (proteína) - adaptado de [Schreiber 2005]. (b) Fatores de Transcrição e Sítios de Ligação: os Fatores de Transcrição são proteínas que ajudam a ativar ou desativar genes específicos, ligando-se ao DNA próximo; os Fatores de Transcrição, que são ativadores, aumentam a transcrição de um gene, e os que são repressores diminuem a transcrição. Adaptado de [Ezer 2014].

Neste trabalho, o principal objetivo é mostrar o desenvolvimento de uma versão de um Sistema Multiagente (Multiagent Systems - MAS), como um *framework* de modelagem para representação de GRN. Para o desenvolvimento deste MAS, foi realizado um processo de Revisão Sistemática detalhado em [Agostinho et al. 2018]. A Revisão Sistemática foi realizada de acordo com [Mariano et al. 2017], e resul-

tou nos seguintes artigos: [Liu et al. 2016], [Khan et al. 2003], [Haydarlou et al. 2016], [Ghazikhani et al. 2011], [Sanfilippo et al. 2012], [Pham 2008], [Yang and Sun 2011].

Este artigo apresenta a versão inicial entre a interação de dois TF que interagem num Sítio de Ligação específico, e são apresentados alguns resultados preliminares, comparados com o processo de Michaelis-Menten [Johnson and Goody 2011].

O artigo está organizado da seguinte maneira: Operação de uma Rede Regulatória Genética, que explica a fundamentação de uma GRN; Modelo Proposto, que explica a fundamentação de MAS e como o modelo foi desenvolvido; Resultados e Conclusões.

2. Operação de uma Rede Regulatória Genética

Genes são ativados ou inibidos por proteínas chamadas TF, que são genes produtos. Uma GRN descreve a interação entre os TF e os genes que são regulados. Esse processo permite que células produzam as proteínas necessárias e apropriadas em termos de tempo e quantidade [Lopes 2011]. O mecanismo básico de controle regulatório é realizado através da ligação dos TF em um local específico, os Sítios de Ligação, que estão localizados numa região promotora de um gene como é descrito em [Schreiber 2005]. Um TF pode interagir num Sítio de Ligação e ativar ou inibir o gene associado, e, então, incrementar ou decrementar a produção de proteínas, numa taxa específica.

O processo de Michaelis-Menten foi definido há mais de 100 anos e mostra que a taxa de reação catalizada para uma enzima é proporcional à concentração do complexo substrato-enzima, conforme Equação 1.

A taxa de reação varia linearmente com a concentração de substrato [S] (cinética de primeira ordem). No entanto, em níveis mais altos [S] com $[S] \gg K_M$, a reação se torna independente de [S] (cinética de ordem zero) e se aproxima assintoticamente de sua taxa máxima V_{max} [Johnson and Goody 2011]. O processo de regulação entre os genes emerge de suas interações no Sítio de Ligação e pode ser modelado como um processo de cinética enzimática, também conhecido como cinética de Michaelis-Menten [Johnson and Goody 2011]. A taxa de reação varia linearmente com a concentração do substrato [S] (cinética de primeira-ordem). O comportamento da curva de Michaelis-Menten (Figura 2) é usado para comparações com os resultados do modelo proposto.

$$v = \frac{d[P]}{dt} = \frac{V_{max}[S]}{K_M + [S]} \quad (1)$$

3. Modelo Proposto

O modelo proposto é uma abstração do processo bioquímico subjacente e simula a interação entre TF e um gene alvo. As ferramentas MAS são muito flexíveis e são capazes de representar cada indivíduo no sistema, simplificando a avaliação de novas hipóteses sobre o modelo.

O simulador é implementado no NetLogo¹, um software orientado a agentes utilizando a linguagem Logo. Dois arquivos de configuração são definidos para configurar o modelo: definições dos agentes e definições das restrições (Tabela 1a e 1b). Importante lembrar que no NetLogo, cada etapa do tempo é representada por um “ tick ”.

¹<https://ccl.northwestern.edu/netlogo/>

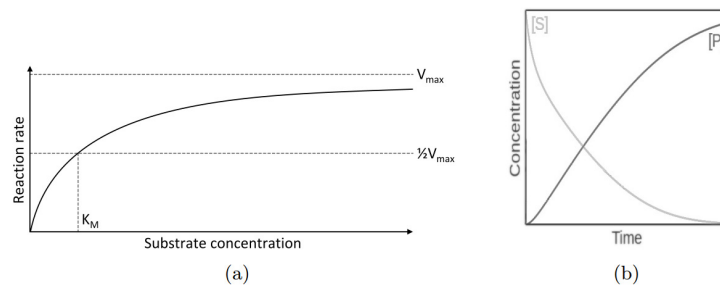


Figura 2. (a) Curva da taxa de reação de Michaelis-Menten. (b) Variação típica na concentração do substrato.

Tabela 1. Arquivos de configuração

(a)					(b)			
Id	Tam	Cor	Qtd	Gene	Regulador	Ação	Regulado	Qtd
0	1	branco	0	"X"	"Y"	1	"X"	3
1	1	vermelho	30	"Y"				

Na Tabela 1a apresenta o arquivo de entrada com as definições dos agentes, onde tem-se 0 gene "X" e 30 genes "Y". É definido cada gene em termos de tamanho (em plano bi-dimensional), cor, quantidade e nome. Na Tabela apresenta o arquivo de entrada para as restrições entre as ações dos agentes, onde a primeira coluna define o nome do gene regulador e a terceira define o gene regulado. A segunda coluna é a definição da ação. No exemplo, a ação de "Y" em "X" é ativação (Ação 1) e 3 agentes de "Y" são requeridos para produzir cada novo gene do tipo "X". O ambiente do simulador é apresentado na Figura 3, onde "X" e "Y" são representados, respectivamente, com círculos brancos e vermelhos. Existe um quadrado verde no ambiente que representa o Sítio de Ligação para "Y".

4. Resultados

No atual estado de desenvolvimento do simulador, os genes (agentes) interagem em Sítios de Ligação, de acordo com as configurações definidas e o comportamento apresentado na Figura 3. No NetLogo cada passo é representado por um "tick". A Figura 3a apresenta a simulação com os genes e o Sítio de Ligação depois de alguns "ticks" de simulação. Na Figura 3b o ambiente de simulação é apresentado após 1684 "ticks". O painel a direita da Figura 3c apresenta a concentração de genes "X" e "Y" ao final da simulação (1684 "ticks"). Salientando que o valor 1684 é o valor no qual consegue-se observar a estabilidade no sistema, ou seja, os TF param de movimentar-se no ambiente.

Salienta-se aqui que o gráfico resultante desta simulação é representado pela concentração de TF em X *versus* "ticks" e concentração de produto em Y *versus* "ticks".

5. Conclusões

Segundo nosso conhecimento, esta é a primeira tentativa de simular GRN usando um *framework* Multiagente. No presente estudo um MAS é definido com poucas regras e parâmetros. Depois de executar a simulação, a concentração das séries temporais dos

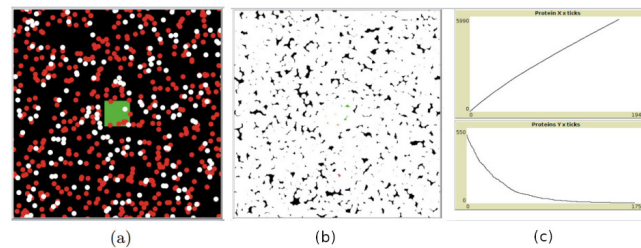


Figura 3. (a) Ambiente de simulação. Os Sítios de Ligação são representados pelos quadrados verdes e os agentes como círculos brancos e vermelhos. (b) Estado do ambiente de simulação após 1684 ticks. (c) Gráficos de concentração temporal dos agentes (genes).

genes foram apresentadas na Figura 3c. Essas curvas apresentam comportamento similar a curva de Michaelis-Menten apresentada na Figura 2b. Apesar de esses serem resultados preliminares, eles indicam que existe um bom local para a pesquisa.

Pretende-se validar esta abordagem comparando-a com o trabalho de [Alon 2007], cujos resultados são apresentados na Figura 4. A Figura 4 mostra o resultado do TF em Y e do produto em Z resultante da interação proveniente do circuito representado na mesma figura. Além disso, pretende-se comparar os resultados deste trabalho com os estudos realizados com o modelo do ciclo circadiano da planta *Arabidopsis thaliana* [Pokhilko et al. 2012] e com o Bio-PEPA [Haydarlou et al. 2016].

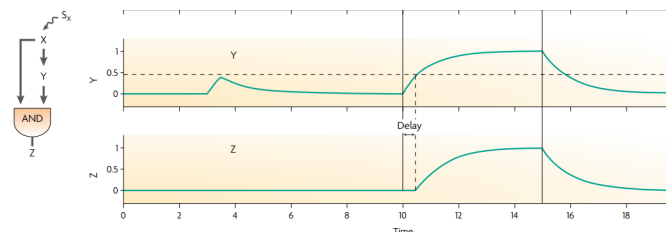


Figura 4. Circuito regulatório e concentrações resultantes de seus componentes [Alon 2007].

6. Agradecimentos

Gostaríamos de agradecer à CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pelo apoio financeiro pela bolsa de estudos.

Referências

- Agostinho, N., Werhli, A., and Adamatti, D. (2018). Revisão sistemática para o desenvolvimento de um ambiente de simulação multiagente para vias regulatórias. *RETEC. REVISTA DE TECNOLOGIAS (OURINHOS)*, 11.
- Alon, U. (2007). Network motifs: theory and experimental approaches. *Nature Reviews Genetics*.
- Ezer (2014). Homotypic clusters of transcription factor binding sites: A model system for understanding the physical mechanics of gene expression. *Comput Struct Biotechnol J*.

- Ghazikhani, A., Akbarzadeh, T., and Monsefi, R. (2011). Genetic regulatory network inference using recurrent neural networks trained by a multi agent system. *International eConference on Computer and Knowledge Engineering (ICCCKE)*.
- Gibas, C. and Jambeck, P. (2002). Developing bioinformatics computer skills. *Yale Journal of Biology and Medicine*.
- Haydarlou, R., Jacobsen, A., Bonzanni, N., Feenstra, K. A., Abeln, S., and Heringa, J. (2016). Bioasf: a framework for automatically generating executable pathway models specified in biopax. *Bioinformatics*, 32:i60–i69.
- Johnson, K. and Goody, R. (2011). The original michaelis constant: Translation of the 1913 michaelis-menten paper. *Biochemistry*.
- Khan, S., Makkena, R., McGeary, F., Decker, K., Gillis, W., and Schmidt, C. (2003). A multi-agent system for the quantitative simulation of biological networks. *AAMAS*.
- Liu, J., Chi, Y. Z. C., and Zhu, C. (2016). A dynamic multiagent genetic algorithm for gene regulatory network reconstruction based on fuzzy cognitive maps. *IEEE TRANSACTIONS ON FUZZY SYSTEMS*, 24.
- Lopes, F. M. (2011). Redes complexas de expressao genica: sintese, identificacao, analise e aplicacoes. Master's thesis, Universidade de Sao Paulo.
- Mariano, D., Leite, C., Santos, L., Rocha, R., and Melo-Minardi, R. (2017). A guide to performing systematic literature reviews in bioinformatics. Technical Report RT.DCC.002/2017, Universidade Federal de Minas Gerais.
- Pham, D. (2008). Multi-agent based simulation of large random boolean network. *MSC-LES*.
- Pokhilko, A., Fernandez, A., Edwards, K., Southern, M., Halliday, K., and Millar, A. (2012). The clock gene circuit in arabidopsis includes a repressilator with additional feedback loops. *molecular systems biology*. *Molecular System Biology*, 8.
- Sanfilippo, A., Haack, J., McDermott, J., Stevens, S., and Stenzel-Poore, M. (2012). Modeling emergence in neuroprotective regulatory networks. *International Conference on Complex Sciences - springer*, pages 291–302.
- Schreiber, S. (2005). Small molecules: the missing link in the central dogma. *Nature Chemical Biology*.
- Yang, T. and Sun, Y. (2011). The reconstruction of gene regulatory network based on multi-agent system by fusing multiple data sources. *2011 IEEE International Conference on Computer Science and Automation Engineering*, 4.