Speeding up Reinforcement Learning for Inference and Control of Gene Regulatory Networks

Anonymous Author(s)
Affiliation
Address
email

Abstract

Motivated by the desire to understand genomic functions through interactions between genes and gene products, the research in the area of gene regulatory networks has become a very important object of study in recent years. Probabilistic Boolean Networks (PBN), which are rules-based dynamic systems, are some of the most studied mathematical models to represent networks and their regulations. However, frequently their representation, regulation, and interactions between genes are overly complex to learn and control, requiring a complex model. Reinforcement Learning (RL) is an interesting technique to deal with this problem because it can learn solutions without the need of a model. This approach is used to train autonomous agents who can find solutions to complex problems, including those of regulation and relationships between genes. But its classical approaches are slow when having to learn tasks with many states especially when these tasks have multiple goals and agents. Besides that, learning bad solutions can make the learning process even more difficult and slow. Therefore, some RL approaches and techniques need to be tested in order to verify the best way to flexibilize, adapt and improve them to intervene and control the gene networks.

1 Introduction

In most living organisms, the genome is concerned with the information that governs life, death and reproduction. Coordinated interactions between genes (both protein coding DNA sequences and non-coded DNA regulatory sequences), RNA, and proteins, form a GRN. If it is possible to build good GRN models and to apply intervention techniques to control genes, it might be possible to develop better treatments for diseases resulting from aberrant gene regulation, like cancer [18]. However, how to model efficient intervention techniques is an open question.

An important class of network models is the PBN model [12] in which each node (gene) can have two possible values, on (1) or off (0), and the way genes interact with each other is formulated by logic functions. Although this model is conceptually simple, it captures some fundamental characteristics of genetic regulation, encapsulating physical and biological information flow by means of rule-based structures. In addition, boolean models can be physically implemented by electronic circuits, and demonstrate rich dynamics that can be studied using mathematical theory, signal processing, and machine learning techniques, e.g., Markov chains [5] and Markov Decision Processes (MDP) [10].

RL [15] is an extensively used technique for learning how to solve MDPs through experimentation. First an action that affects the environment is chosen, then the agent observes how much that action collaborated to the task completion through a reward function. An agent can learn how to optimally solve tasks by executing this procedure multiple times and RL techniques have been used to solve many challenging tasks. The main question to be explored throughout this work (in progress) is how RL algorithms can be adapted to enable learning intervention policies, maintaining a biological system in one desired state, even more so for systems with multiple objectives [11, 17]. The main challenge to be solved is that RL techniques require many interactions with the actual system, which is usually unfeasible because incorrect interventions could harm the biological system. However, like in the human learning process, previous knowledge can greatly accelerate the learning of a new task and might help designing RL techniques applicable to gene-intervention learning [7]. Another question to be answered is how to model GRN inference with RL. The reward function would include conflicting objectives [17, 3], which could further hamper RL traning [4].

Many solutions were developed to address those two issues in general-purpose RL techniques, such as Transfer Learning (TL) and options-based ones. In the work [1] for example, it was proposed a framework which provides a way to generalize and reuse knowledge between tasks by encapsulating states and sequences of actions performed by agents, and we believe it can be easily incorporated into a variety of different RL algorithms and tasks, accelerating their learning processes.

These problems may consist in using RL to improve some algorithms or applying some different ways of solving problems. Such issues may be related to discovering or control the features of the GRN [13]; revealing attractor basins [8] in which RL was used sometimes in order to find subgoals in solutions [3]; allowing humans to intervene in the GRNs while the agent learns through RL algorithm [14], where different other algorithms can be tested; and approximation functions together with RL to regulate networks [7].

2 Research Goal and Expected Contributions

The main objective of this work is to **develop RL algorithms to successfully learn how to infer and intervene in GRNs to maintain a biological system in a desirable state**. Solving this challenge includes (i) finding the best way of modeling the problem as an MDP; (ii) coping with the curse of dimensionality inherent from the domain; and (iii) proposing ways to safely evaluate the proposed methods without harming experiment subjects.

3 Preliminary Results

Previous works mostly focused on improving general-purpose RL methods, which we now plan to apply in this challenging domain. In order to define a representation which allows knowledge generalization, it was proposed an extension of the Options Framework [16], called Multiobjective Options (MOOpt) for Multiobjective Reinforcement Learning (MORL) problems [2, 3] , in which the algorithm *PolicyBlocks* [9] was extended to multiobjective domains and the knowledge obtained was transferred to new tasks. This proposed approach aimed at discovering single-objective options, evaluating them in a different state space, expecting to accelerate learning in MORL problems. In the papers [2, 3] they were only evaluated in the standard MDPs, and an evaluation to verify the benefits from using options with a more robust representation for RL problems (OOMDP) is presented in [1], as well as a probabilistic reuse of learned solutions. Another work that compares the use of this representation (OOMDP) with MOOpt against the MOOpt alone is currently under development. Those works might be a starting point in the investigation of how to use RL and to transfer useful knowledge to be applied to GRN inference and control problems.

4 Next Steps

MO-Opt option-reusing algorithm was shown how to accelerate learning in MORL tasks. The option-based method helped the machine to give a solution faster while providing solutions according to multiple human preferences.

Now, the next step is to model the GRN inference and/or intervention as a sequential decision-making problem, analyzing how RL can be used for this and studying other previous related works in the area, like [6] for instance. After that, we will try to find ways to scale and specialize general-purpose RL methods for our domain. The first steps were performed by the aforementioned works, and we intend to evaluate them in the context of GRN inference and control first, to at least have insights on what must be improved.

References

[1] Rodrigo Cesar Bonini, Felipe Leno da Silva, Ruben Glatt, Edison Spina, and Anna Helena Reali Costa. A framework to discover and reuse object-oriented options in reinforcement learning. In *BRACIS*, 2018.

[2] Rodrigo Cesar Bonini, Felipe Leno da Silva, and Anna Helena Reali Costa. Learning options in multiobjective reinforcement learning. In *AAAI-17 Student Paper*, pages (4708–4709), 2017.

[3] Rodrigo Cesar Bonini, Felipe Leno da Silva, Edison Spina, and Anna Helena Reali Costa. Using options to accelerate learning of new tasks according to human preferences. In *AAAI-17 Workshop Human-Machine Collaborative Learning*, pages (1–8), 2017.

[4] Ivan Brugere, Brian Gallagher, and Tanya Y Berger-Wolf. Network structure inference, a survey: Motivations, methods, and applications. *ACM Computing Surveys (CSUR)*, 51(2):24, 2018.

[5] Aniruddha Datta, Ashish Choudhary, Michael L Bittner, and Edward R Dougherty. External control in markovian genetic regulatory networks. *Machine learning*, 52(1-2):169–191, 2003.

[6] Ricardo De Souza Jacomini, David Correa Martins-Jr, Felipe Leno Da Silva, and Anna Helena Reali Costa. Genice: A novel framework for gene network inference by clustering, exhaustive search, and multivariate analysis. *Journal of Computational Biology*, 24(8):809–830, 2017.

[7] B Faryabi, A Datta, and ER Dougherty. On approximate stochastic control in genetic regulatory networks. *IET Systems Biology*, 1(6):361–368, 2007.

[8] Russell Golman and Scott E Page. Basins of attraction and equilibrium selection under different learning rules. *Journal of evolutionary economics*, 20(1):49, 2010.

[9] Marc Pickett and Andrew G Barto. Policyblocks: An algorithm for creating useful macro-actions in reinforcement learning. In *ICML*, pages 506–513, 2002.

[10] Martin L Puterman. *Markov Decision Processes.: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 2014.

[11] Diederik Marijn Roijers, Peter Vamplew, Shimon Whiteson, and Richard Dazeley. A survey of multi-objective sequential decision-making. *CoRR*, 2014.

[12] Ilya Shmulevich, Edward R Dougherty, Seungchan Kim, and Wei Zhang. Probabilistic boolean networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics*, 18(2):261–274, 2002.

[13] Utku Sirin, Faruk Polat, and Reda Alhajj. Employing batch reinforcement learning to control gene regulation without explicitly constructing gene regulatory networks. In *Twenty-Third International Joint Conference on Artificial Intelligence*, 2013.

[14] Aivar Sootla, Natalja Strelkowa, Damien Ernst, Mauricio Barahona, and Guy-Bart Stan. Toggling a genetic switch using reinforcement learning. *arXiv preprint arXiv:1303.3183*, 2013.

[15] Richard S. Sutton and Andrew G. Barto. *Reinforcement learning: An introduction*. MIT Press, Cambridge, MA, USA, 1st edition, 1998.

[16] Richard S Sutton, Doina Precup, and Satinder Singh. Between mdps and semi-mdps: A framework for temporal abstraction in reinforcement learning. *Artificial intelligence*, pages 181–211, 1999.

[17] Kristof Van Moffaert, Madalina M Drugan, and Ann Nowé. Scalarized multi-objective reinforcement learning: Novel design techniques. In *2013 in ADPRL*, pages 191–199. IEEE, 2013.

[18] Yufei Xiao. A tutorial on analysis and simulation of boolean gene regulatory network models. *Current genomics*, 10(7):511–525, 2009.