GAME THEORETIC MODEL FOR CONTROL OF GENE REGULATORY NETWORKS

Liming Wang and Dan Schonfeld

University of Illinois at Chicago Department of Electrical and Computer Engineering lwang37@uic.edu, dans@uic.edu

ABSTRACT

The intervention in gene regulatory networks has been modelled as the Markov decision process problem. However, this approach only allows one external control, which is inadequate in many situations such as drug and gene therapies. In this paper, we propose the non-cooperative stochastic game model for intervention in the genetic regulatory networks as the generalization of the Markov decision process approach and formulate the intervention problem into solving the Nash equilibrium. The definition of equilibrium has been proposed and the existences for both infinite and finite horizon cases have been proven. We provide the numerical example for using non-cooperative stochastic game model on the mammalian cell cycle network. We also compare the results under the Nash equilibrium and independent Markov decision process approach.

Index Terms— Non-cooperative stochastic game, Gene regulatory networks, Markov decision process.

1. INTRODUCTION

The genetic regulatory modeling is aimed to describe the dynamics of Gene networks, which is useful to identify potential drug targets or alter the system evolutions in a desired manner. It is well known that many diseases such as cancer and tumor are due to the facts of cells improper proliferation or misfunction of certain genes.

The probabilistic Boolean network (PBN) [1] is one the widely used model for genetic regulatory modeling. The purpose of intervention of the PBN is to reduce the probability to visit the unwanted states. The PBN can be modeled as a Markov chain model under certain assumptions [2]. The perturbed gene can be seen as the external control of the Markov process. The Markov decision process (MDP) has been successfully applied to the intervention of PBN [3] and the optimal control policy can be solved by dynamic programming [4].

In these approaches of intervention in PBN, the perturbed the genes are modeled as the control of the MDP and the optimal control policy is solved by maximizing the reward function. However, in many situations, multiple controls may be desired to obtain various optimalities for different purposes. For example, in drug or gene therapy, the patient may need to be treated for various diseases simultaneously. In these situations, a single control is not adequate to describe the interventions. In multiple controls situation, the independent and blind maximization of reward function for each control will not achieve the original aim since the dynamics of the system are bound together. Instead of solving a single optimization problem as in MDP, proper concept of solution in these problems i.e. the Nash equilibrium is required to be defined. As we will see later in the paper, the classical MDP falls into a special case of the non-cooperative stochastic game (NCSG).

In this paper, we propose the NCSG model for control of the genetic regulatory networks and formulate the intervention problem into solving the Nash equilibrium. In section 2. we first provide the necessary background of the PBN and introduce the NCSG model. In section 3, the MDP is shown to be a special case of NCSG and the solving methods are provided. In section 4, the definition of Nash equilibrium has been proposed and the existences for both infinite and finite horizon cases have been proven. In section 5, we provide the numerical example for our NCSG model on the mammalian cell cycle network. We also compare the results under the Nash equilibrium and independent MDP (IMDP) method. Finally, we provide a brief summary and discussion of our results in Section 6.

2. THE MODELS

The PBN [1] consists of n nodes $N = \{x_i\}_{i=1}^n$, where $x_i \in \{0, 1\}$, a sequence of vector valued *predictor func*tions $\{\mathbf{f}_j\}_{j=1}^r$. The x_i represents the expression level of gene with "0" meaning OFF and "1" meaning ON. The predictor functions $f_j = (f_{j1}, \ldots, f_{jn})$ is a predictor of genes 1 through n, when the network j is selected. Note that the Boolean network is deterministic if the predictor functions are given. We have the *switching probability* q to describe the randomness of the network. If the switch is on, the predictor functions are chosen randomly according to probability measure $\{p_l\}_{l=1}^q$. The vector $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))$ can be seen as the binary expansion of $\{0, \ldots, 2^n - 1\}$, which is called *Gene-activity profile* (GAP). For the intervention, we assume there are K controls. In the context of Game theory, we refer them to be K players. For each player k, it controls genes $\mathbf{a}_k = \{g_{k1}, \ldots, g_{km_k}\}$. Therefore the action set for player k is $A_k = \{0, 1, \ldots, 2^{m_k} - 1\}$. We denote $\mathbf{a} = \prod_{k=1}^{K} \mathbf{a}_k$ and $\mathbf{A} = \prod_{k=1}^{K} A_k$. We also assume the genes each player controls are disjoint. At every epoch t the network updates and every player make decision at the same time, i.e. each player only knows the his own decision.

Let $\mathbf{s}_t = (x_1(t), \dots, x_n(t))$ denote the network state. The system can be modelled as the controlled Markov process. The transition probability is defined as,

$$\mathbb{P}\{s'|s,\mathbf{a}\} := \mathbb{P}\{\mathbf{s}_{t+1} = s'|\mathbf{s}_t = s, \mathbf{a}_1 = a^{(1)}, \dots, \mathbf{a}_k = a^{(k)}\}$$
(1)

If the initial state \mathbf{s}_0 is given, then it fixed a stochastic process. Let $\mathbf{r}_t^k(\mathbf{s}_t, \mathbf{a}_1, \dots, \mathbf{a}_k)$ be the immediate reward function for player k at epoch t. So we have the expected reward function $\mathbf{R}_t^k = \mathbb{E}_{s_0}(\mathbf{r}_t^k)$. In this paper, we mainly focus on the finite and infinite horizon discount average reward functions:

$$V^{k}(L) := \sum_{n=1}^{L} \beta_{k}^{n-1} \mathbf{R}_{n}^{k}$$
⁽²⁾

$$V^k := \sum_{n=1}^{\infty} \beta_k^{n-1} \mathbf{R}_n^k \tag{3}$$

The policy $\pi_k = {\pi_{k1}, \pi_{k2}, ...}$ for player k is a sequence of probability distributions such that π_{ki} is the decision distribution on the action set \mathbf{a}_k . A policy is *pure*, if all the decisions in a policy are deterministic. If π_{ki} is independent of time *i*, then the policy is *stationary*.

3. THE SINGLE PLAYER CASE

If we only have one player, i.e. K = 1, then the NCSG model reduces to the Markov decision problem. Assume the cardinality of the states is 2^n and the cardinality of the action set is 2^m . The finite horizon discount problem can be solved by the following dynamic programming:

$$V(n) = \max_{A} \left\{ \mathbf{r}(s, a) + \sum_{s'=0}^{2^{m}-1} \mathbb{P}\{s'|s, a\} V(n-1) \right\}$$
(4)

In the infinite horizon case, the problem can be solved by the following linear programming [7].

$$\max \sum_{s} \sum_{a=0}^{2^{m}-1} \mathbf{r}(s, a) z(s, a)$$

subject to:
$$\sum_{s} \sum_{a=1}^{2^{m}-1} \{\delta(s, s') - \beta \mathbb{P}\{s'|s, a\}\} z(s, a) = \frac{1}{2^{n}}$$

$$z(s, a) \ge 0$$
 (5)

The optimal stationary policy is given by

$$\pi = \frac{z^{\star}}{\sum_{a} z^{\star}} \tag{6}$$

where z^* is the solution of the linear programming (5).

4. THE NON-COOPERATIVE MULTIPLE PLAYERS CASE

The NCSG is the generalization of MDP. In multi-player case, although the goal for each player is to maximize his average reward function, however, the independent IMDP solution for each player is not adequate, i.e. each control works as if it were the only control and treats the problem as the MDP. Because the reward function and transition probability are coupled together by the decisions of all players.

Unlike the case in MDP, where one reward function is to be maximized. In NCSG, all players are interested in maximizing their individual average reward functions. Therefore a proper solution concept is needed to be defined. The concept of Nash equilibrium in static game can be extended as the equilibrium in NCSG.

Definition 1. A policy $\pi^* = (\pi_1, \dots, \pi_K)$ is the infinite horizon discount Nash equilibrium (NE) point if the following inequality holds,

$$V_k(\pi^\star) \ge V_k(\mu_k, \pi_{-k}^\star) \tag{7}$$

for any $\mu_k \in \pi_k$ and any $k \in K$. The π_{-k} denotes the policies of all players except player k.

Similarly, we also have the definition for finite horizon discount Nash equilibrium point.

Definition 2. A policy $\pi^* = (\pi_1, \ldots, \pi_K)$ is the L-step finite horizon discount Nash equilibrium (NE) point if the following inequality holds,

$$V_k(L)(\pi^*) \ge V_k(L)(\mu_k, \pi_{-k}^*) \tag{8}$$

for any $\mu_k \in \pi_k$ and any $k \in K$. The π_{-k} denotes the policies of all players except player k.

The existence of infinite horizon discount equilibrium point has been proven by A. Fink [5]. By using the Nash theorem [6], we can also show the existence in finite horizon case.

Lemma 1. Any policy π_k is a convex combination of the pure policies in finite horizon discount game.

Proof. We show the scheme to construct such convex combinations for L = 2. First consider the π_{k1} , it is straightforward to see

$$\pi_{k1} = \sum_{i=1}^{M} c_i^1 \mu_{k1}^i \tag{9}$$

where $\sum_i c_i^1 = 1$. μ_{k1}^i denotes the pure policy for player k at epoch 1, which has probability 1 on *i*-th gene. Therefore consider the policy

$$\pi^{(j)} = \sum_{i=1}^{M} c_i^1(\mu_{k1}^i, \mu_{k1}^j) \tag{10}$$

. The first action in this policy $\pi^{(j)}$ is the same as π_{k1} , the second action is the pure policy which has probability 1 on the *j*-th gene. So

$$(\pi_{k1}, \pi_{k2}) = \sum_{j=1}^{M} c_j^2 \pi^{(j)}$$
(11)

where $\sum_{i} c_i^2 = 1$. Therefore

$$(\pi_{k1}, \pi_{k2}) = \sum_{i,j=1}^{M} c_i^2 c_i^1(\mu_{k1}^i, \mu_{k1}^j)$$
(12)

We see that the policy is convex combination of pure policies. It follows by mathematical induction for arbitrary stage L.

Theorem 1. The finite horizon discount game has a Nash equilibrium point.

Proof. For any given stage L and initial state s_0 , since the states and actions sets are finite, there are finitely many pure policies. By lemma 1, any policies is convex combination of pure policies. Therefore the set of all policies is a closed, bounded, and convex polyhedron. By Nash theorem [6], the equilibrium exists for finite horizon discount game.

Under certain additional conditions, finding Nash equilibrium could be solved by linear programming or non-linear programming method. For example, if we have a two-player zero-sum game, i.e. K = 2 and $\mathbf{r}^1 = -\mathbf{r}^2$. This game can be interpreted as the matrix game and can be solved by Newton's method [7].

5. EXPERIMENTS

In this section, we conduct an experiment based on the mammalian cell cycle with a mutated phenotype. The cycle regulation is proposed in [8]. We order these genes as $s = \{CycD, Rb, E2F, CycE, CycA, Cdc20, Cdh1, Ubc, CycB\}$. The state sets $\{s\}$ can be interpreted as the binary expansion of $\{0, 1, \ldots, 511\}$. Depending on the value of input CycD. We have two constituent Boolean networks of the PBN. We assume the two constituent networks have the same probability and the probability of switching is 0.01. In Table 1, we show the logic relation of all these genes.

Gene	Predictors
CycD	Input
Rb	$(\overline{CycD} \land \overline{CycE} \land \overline{CycA} \land \overline{CycB})$
E2F	$(\overline{Rb} \land \overline{CycA} \land \overline{CycB})$
CycE	$(E2F \wedge \overline{Rb})$
CycA	$(E2F \land \overline{Rb} \land \overline{Cdc20} \land (\overline{Cdh1 \land Ubc})) \lor$
	$(CycA \land \overline{Rb} \land \overline{Cdc20} \land (\overline{Cdh1 \land Ubc}))$
Cdc20	CycB
Cdh1	$((\overline{CycA} \land \overline{CycB}) \lor \overline{Cdc20})$
Ubc	$(\overline{Cdh1}) \lor (Cdh1 \land Ubc \land Cdc20 \lor CycA \lor$
	CycB)
CycB	$(\overline{Cdc20}\wedge\overline{Cdh1})$

 Table 1. Boolean functions of mammalian cell cycle.

We have the Rb and CycA genes as the two controls a_1 and a_2 of the PBN respectively. We also assume the following reward functions.

$$\mathbf{r}^{1} = \begin{cases} 10 & \text{if } \mathbf{a}_{1} = 0 \text{ and } (CycD, Rb) \neq (0, 0) \\ 0 & \text{if } \mathbf{a}_{1} = 1 \text{ and } (CycD, Rb) = (0, 0) \\ 2 & \text{otherwise} \end{cases}$$
(13)

$$\mathbf{r}^{2} = \begin{cases} 9 & \text{if } \mathbf{a}_{2} = 0 \text{ and } (CycD, CycA) \neq (0,0) \\ 1 & \text{if } \mathbf{a}_{2} = 1 \text{ and } (CycD, CycA) = (0,0) \\ 3 & \text{otherwise} \end{cases}$$
(14)

In Fig. 1. we show the average rewards under two scenarios. The first one is the PBN with only one control on gene Rb by using the linear programming method in section 3. The second scenario is the PBN has two control, but each control employs IMDP method. We can find in the illustration that the average reward under the second scenario is lower than the first one. The reason is that under the second scenario, the transition probability is coupled with the two controls together. In general the IMDP method with negligence of other controls will not achieve the optimality. The Fig. 2. shows the same situation for control on gene CycA.

In Fig. 3. and Fig. 4, we compare the average rewards under the IMDP and the equilibrium. We can find in the illustrations, the rewards under the equilibrium are generally better than IMDP method case. Since the effects of other controls have been taken into account and no player can profitably deviate by any of his own actions, which serves as a solution concept under multiple players situation.

6. CONCLUSION

In this paper, we propose the NCSG model for intervention of the genetic regulatory networks which extends previous MDP model into multiple controls case. We formulate the intervention problem into solving the Nash equilibrium. The definition of equilibrium has been proposed and the existences for



Fig. 1. Illustrations of the average reward of the optimal intervention for control a_1 on gene Rb under the case only one control and the average reward of the optimal intervention for control on Rb using IMDP method with the other control a_2 on CycA ignored.



Fig. 2. Illustrations of the average reward of the optimal intervention for control a_2 on gene CycA under the case only one control and the average reward of the optimal intervention for control on CycA using IMDP method with the other control a_1 on Rb ignored.



Fig. 3. Illustrations of the average reward of the optimal intervention for control a_1 on gene Rb under the two controls case using IMDP method and the average reward of the optimal intervention for control on CycA under equilibrium

both infinite and finite horizon cases have been provided. We also provide the numerical example for using our proposed



Fig. 4. Illustrations of the average reward of the optimal intervention for control a_2 on CycA under the two controls case using IMDP method and the average reward of the optimal intervention for control on Rb under equilibrium

NCSG model on the mammalian cell cycle network.

7. REFERENCES

- I. Shmulevich, E.R. Dougherty, S. Kim, and W. Zhang, "Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks," *Bioinformatics*, vol. 18, no. 2, pp. 261–274, 2002.
- [2] M. Brun, E.R. Dougherty, and I. Shmulevich, "Steadystate probabilities for attractors in probabilistic Boolean networks," *Signal Processing*, vol. 85, no. 10, pp. 1993– 2013, 2005.
- [3] B. Faryabi, G. Vahedi, J.F. Chamberland, A. Datta, and E.R. Dougherty, "Optimal Constrained Stationary Intervention in Gene Regulatory Networks," *EURASIP Journal on Bioinformatics and Systems Biology*, vol. 2008, 2008.
- [4] E. Altman, *Constrained Markov decision processes*, Chapman & Hall, 1999.
- [5] A.M Fink, "Equilibrium in a stochastic n-person game," *Journal of Science in Hiroshima University, Series AI*, vol. 28, pp. 89–93, 1964.
- [6] J.F. Nash, "Equilibrium points in n-person games," Proceedings of the National Academy of Sciences of the United States of America, pp. 48–49, 1950.
- [7] J.A. Filar, K. Vrieze, and OJ Vrieze, Competitive Markov decision processes, Springer Verlag, 1997.
- [8] A. Faure, A. Naldi, C. Chaouiya, and D. Thieffry, "Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle," *Bioinformatics*, vol. 22, no. 14, 2006.