SER-SIT Stochastic Games and Vertical Linear Complementarity Problem [1]

G. S. R. Murthy
Indian Statistical Institute
Street No.8, Habshiguda
Hyderabad - 500017
INDIA

T. Parthasarthy
Indian Statistical Institute
New Delhi - 110 016
INDIA

D. Sampangi Raman
Indian Statistical Institute
110, Nelson Manikham Road
Chennai - 600029
INDIA

Abstract

In this note, we consider a two-person nonzerosum finite stochastic game with SER-SIT properties, that is, with separable rewards and state independent transitions. We formulating the given problem as a vertical linear complementarity problem for the discounted and the undiscounted stochastic games. Our result generalizes an earlier result obtained by Parthasarthy, Tijs and Vrieze.

Key words: Stochastic games, ordered field property, vertical linear complementarity problem.

1 Introduction

The theory of stochastic games began in 1953 with a seminal paper of Shapley [18]. He showed that a zero-sum discounted stochastic game has a value and the two players have optimal stationary strategies. Theory of zero-sum undiscounted stochastic games started with a paper by Gillette [5], while nonzero-sum stochastic games were considered by Fink [4], Takahashi [20] and Sobel [19]. Shapley also remarked that the value of a stochastic game may not lie in the rational field even if the data defining the stochastic game comes from the rational field. Parthasarthy and Raghavan [15] introduced a subclass of the

[1]Dedicated to Stef Tijs for his 65^{th} birthday

stochastic games in which one player controls the transitions and they have shown such games possess the ordered filed property. That is a stochastic game is said to possess the ordered filed property if the value and the components of an optimal (or equilibrium) strategy lie in the rational filed if the data defining the stochastic game comes from the rational field. If a stochastic game has the ordered field property, there is always a hope to find a solution with a finite step algorithm. Many researchers have found interesting subclasses of stochastic games which possess ordered field property and in most of these instances, algorithms were found using results from Linear Programming or Linear Complementarity Problems. Readers are urged to refer to the monograph by Filar and Vrieze [3]. See also Mohan *et.al.* [12] and Zamir [22] for the most recent results.

2 Generalization of SER-SIT Stochastic Games :

First, we describe a generalization of SER-SIT stochastic games. We then prove the existence of equilibrium points for SER-SIT games.

Let $\mathcal{M} = [M_1, M_2, \ldots, M_k]$ where each $M_i \in \mathcal{R}^{m\times n}$. By a row (column) representation $\mathcal{R}(C)$ from $\mathcal{M}$, we denote a matrix of order $m \times n$ whose i^{th} row (j^{th} column) is chosen as the i^{th} row (j^{th} column) of some $M \in \mathcal{M}$.

Example 2.1 Let $\mathcal{M} = \left\{ \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \end{bmatrix}, \begin{bmatrix} -1 & 0 & 0 \\ 10 & 5 & 1 \end{bmatrix}, \begin{bmatrix} 20 & -1 & 5 \\ 0 & 0 & 2 \end{bmatrix} \right\}$.
Then the following are examples of row representative matrices from $\mathcal{M}$:

$$\begin{bmatrix} 1 & 2 & 3 \\ 10 & 5 & 1 \end{bmatrix}, \begin{bmatrix} 20 & -1 & 5 \\ 2 & 4 & 5 \end{bmatrix}.$$

Similarly, the following are some column representative matrices from $\mathcal{M}$:

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 5 & 1 \end{bmatrix}, \begin{bmatrix} 20 & -1 & 3 \\ 0 & 0 & 5 \end{bmatrix}.$$

A generalized two-person stochastic game is denoted by $\Gamma = \{S, \mathcal{A}, \mathcal{B}, r_k, q\}$ where the finite set S stands for states of a system, $\mathcal{A}$ and $\mathcal{B}$ stand for a finite collection of real matrices of order $m \times n$, r_k stands for the reward to the k^{th} player, $k = 1, 2$ and q stands for the transition probability.

We now give a description of the game. Observing the (initial) state s, player 1 selects (x, A) where $x \in \mathcal{R}^m$ is a probability vector and A a row representative matrix from $\mathcal{A}$ and player 2 selects (y, B) where $y \in \mathcal{R}^n$ is a probability vector and B a column representative matrix from $\mathcal{B}$. As a result two things happen:

(i) Player 1 receives an immediate reward $r_1(s, x, y) = c_1(s) + x^t A y$ and player 2 receives an immediate reward $r_2(s, x, y) = c_2(s) + x^t B y$. Here, c_1 and c_2 depend only s.

(ii) The system moves to a new state s' according to the transition law $q(s'|s, x, y)$. Players play the game observing s' as before.

With respect to the evaluation of the pay-off stream, we will consider two cases namely the discounted case and the undiscounted case. In the β-discounted case, where $\beta \in [0, 1)$, we suppose that player $k, (k = 1, 2)$ wants to maximize $\sum\limits_{n=1}^{\infty} \beta^{n-1} r_k^{(n)}$ where, $r_k^{(n)}$ stands for the direct reward on the n^{th} day for player k. In the discounted case, player k wants to maximize $\liminf\limits_{n \to \infty} \left(\frac{1}{n} \sum\limits_{j-1}^{n} r_k^{(j)} \right)$.

In this paper, we are interested in games, satisfying the following additional conditions.

Condition SIT (State Independent Transitions): A transition law defined by q depends only on the actions chosen by the two players. That is,

$$q(s'|s, i, j) = q(s'|i, j) \; \forall x \in S \quad \text{and} \quad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n$$

Define $q(s'|s, x, y) + \sum_j \sum_i q(s'|s, i, j) \xi_i \eta_j$ where $x = (\xi_1, \xi_2, \ldots, \xi_m), y = (\eta_1, \eta_2, \ldots, \eta_n)$.

By our SIT condition, $q(s'|s, x, y) = q(s'|, x, y)$ for all $s \in S$ and clearly $q(s'|, x, y)$ is continuous in (x, y).

Condition SER (Separable Rewards): This condition means,

$$\begin{aligned} r_1(s,x,y) &= c_1(s) + x^t A y \quad \text{for every} \quad s, x, A \quad \text{and} \quad y \\ r_2(s,x,y) &= c_2(s) + x^t B y \quad \text{for every} \quad s, x, B \quad \text{and} \quad y. \end{aligned}$$

A game, with the above two conditions, is called as a SER-SIT game.

One of the main advantages in the present set up is that each player is allowed partially to change his/her immediate reward at each stage using the finite collection of pay-off matrices available to him/her. This is a significant departure from the classical theory. For example, if the state is s, player choose strategies x, y then in the classical version, immediate reward to player 1 will be $r_1(s,x,y)$, whereas under our present set up, it will be one of the r_1's, where $r_1(s,x,y) = c_1(s) + x'Ay$ where A is a row representative matrix from $\mathcal{A}$. This gives some freedom in the choice of immediate rewards to the players. In our opinion, it seems to be very reasonable, especially, when the games are played over the infinite future at discrete points of time.

We are now ready to state and prove a result on SER-SIT stochastic games.

Theorem 2.1 Let $\Gamma\{S, \mathcal{A}, \mathcal{B}, r_k, q, \beta\}$ stand for a generalized two-person SER-SIT finite stochastic game with the discount factor β, where S is a finite set and $\mathcal{A}$ and $\mathcal{B}$ are finite sets in $R^{m\times n}$. Then, there exists a row representative matrix M and a strategy $\bar{x}$ for player 1 and a column representative matrix N and a strategy $\bar{y}$ for player 2, which form a pair of equilibrium point for the given stochastic game.

Before giving a proof, we would like to make the following remarks.

Remark 2.1 Note that the strategy $\bar{x}$ (as well as the $\bar{y}$) given in the above theorem are probability vectors with $m(n)$ coordinates, which will be independent of the state of the system. That is, at every point of time, the players will choose $(\bar{x}, M)$ and $(\bar{y}, N)$.

Remark 2.2. Here, we give the proof of Theorem 2.1 using a lemma due to Gowda and Sznajder [7], which is the key to proving Theorem 2.1. In the next section, we present a finite step algorithm by formulating the given problem as a VLCP.

Lemma 2.1 Consider the generalized bimatrix game of Gowda and Sznajder. Let $\mathcal{A}$

and $\mathcal{B}$ are finite sets in $R^{m\times n}$. Then, there exists a pair $(\bar{x}, \bar{y})$ of probability vectors and a row representative matrix M and a column representative matrix N with the following property :

$$\bar{x}M\bar{y} \geq xR\bar{y} \quad \text{and} \quad \bar{x}N\bar{y} \geq \bar{x}Cy$$

for all probability vectors x and y, for all row representative matrices R of $\mathcal{A}$ and for all column representative matrices C of $\mathcal{B}$.

For a proof of Lemma 2.1 see [7].

Proof of Theorem 2.1: Let Γ be the game described in the statement of Theorem 2.1 and $\beta \in [0, 1)$. Now consider the generalized bimatrix game:

$$\left[a(i,j) + \beta \sum_{t\in S} q(t|i,j)c_1(t), b(i,j) + \beta \sum_{t\in S} q(t|i,j)c_2(t)\right]$$

where, $A = [a(i,j)]$ is a row representative matrix $\mathcal{A}$ and $\mathcal{B} = [b(i,j)]$ is a column representative matrix from $\mathcal{B}$.

Now invoke Lemma 2.1 to conclude the existence of $\bar{x}, \bar{y}, M$ and N satisfying

$$\begin{aligned} \bar{x}^t(M + \beta Q_1)\bar{y} &\geq x^t(R + \beta Q_1)\bar{y} \quad \text{and} \\ \bar{x}^t(N + \beta Q_2)\bar{y} &\geq \bar{x}^t(C + \beta Q_2)y \end{aligned}$$

for all probability vectors x, y and for all row representative matrices R and column representative matrices C, where Q_1, Q_2 are matrices whose ij^{th} elements are given by $\sum_{t\in S} q(t|i,j)c_1(t)$ and $\sum_{t\in S} q(t|i,j)c_2(t)$. We will show that $(\bar{x}.\bar{y})$ form an equilibrium for our stochastic game Γ.

To make the notation less cumbersome, we will write $R(x, y)$ instead of x^tRy, $Q_l(x, y)$ instead of $\sum_{t\in S} q(t|i,j)c_1(t)(= x^tQ_{ly})$ for $l = 1, 2$.

We have $(\bar{x}, \bar{y})$ satisfying $M(\bar{x}, \bar{y}) + \beta Q_1(\bar{x}, \bar{y}) \geq R(x, \bar{y}) + \beta Q_1(x, \bar{y})$ and $N(\bar{x}, \bar{y}) + \beta Q_2(\bar{x}, \bar{y}) \geq C(\bar{x}, y) + \beta Q_2(\bar{x}, y)$ folr all x, y, R and C. Let $I_l(\bar{x}, \bar{y}, \beta)(s)$ be the total expected discounted reward for player l when the game starts at an initial state s and if $(\bar{x}, \bar{y})$ is used at every point of time, the game is played. Also, M and N are chosen at every point of time. Then,

$$I_l(\bar{x}, \bar{y}, \beta)(s) \quad = \quad c_1(s) + M(\bar{x}, \bar{y}) + \beta \sum_{t'} \left(c_1(t') + M(\bar{x}, \bar{y})\right) q(t'|\bar{x}, \bar{y}) +$$

$$\beta^2 \sum_{t''} \sum_{t'} (c_l(t'') + M(\bar{x}, \bar{y}))q(t''|\bar{x}, \bar{y})q(t'|\bar{x}, \bar{y}) + \dots$$
$$= \quad c_1(s) + (M(\bar{x}, \bar{y}) + \beta Q_l(\bar{x}, \bar{y})) + \beta(M(\bar{x}, \bar{y}) + \beta Q_l(\bar{x}, \bar{y})) + \dots$$

Since, $M(\bar{x}, \bar{y}) + \beta Q_l(\bar{x}, \bar{y}) \geq R(x, \bar{y}) + \beta Q_l(x, \bar{y}) \; \forall \, x$ and R, it follows that,

$$I_1(\bar{x}, \bar{y}, \beta)(s) \geq I_1(f, \bar{y}, \beta)(s) \; \forall \, f$$

where, f is any arbitrary strategy for player 1 in the stochastic game. (We may assume $f :\to \mathcal{R} \times X$ where, $\mathcal{R}$ = row representative matrices formed from $\mathcal{A}$ and X = space of probability vectors in R^m). Similarly we can show that,

$$I_2(\bar{x}, \bar{y}, \beta)(s) \geq I_2(\bar{x}, g, \beta)(s) \; \forall \, g$$

where g is any arbitrary strategy for player 2, that is, g is a map from $S \to C \times Y$ where C = the class of column representative matrices from $\mathcal{B}$ and Y = the space of probability vectors in R^n. Thus $(\bar{x}, \bar{y})$ turns out to be an equilibrium point for the given stochastic game. This concludes the proof of Theorem 2.1

We now state a theorem for undiscounted generalized SER-SIT stochastic games.

Theorem 2.2 Let $\Gamma = \{S, \mathcal{A}, \mathcal{B}, r_k, g\}$ stand for a generalized two-person SER-SIT finite stochastic game where, $S, \mathcal{A}, \mathcal{B}$ are finite sets. Then there exists a row representative matrix M with a probability vector x for player 1 and a column representative matrix N with a probability vector y for player 2 which form a pair of equilibrium point for the undiscounted stochastic game.

Proof. As in the proof of Theorem 2.1, we consider the following generalized bimatrix game :

$$\left[a(i,j) + \sum_t q(t|i,j)c_1(t), b(i,j) + \sum_t 1(t|ij)c_2(t) \right]$$

where, $A = [a(i,j)]$ is a row representative matrix from $\mathcal{A}$ and $B = [b(i,j)]$ is a column representative matrix from $\mathcal{B}$. It can be shown as before the existence of $(\bar{x}, M)$ $(\bar{y}, N)$ satisfying

$$\begin{aligned} M(\bar{x}, \bar{y}) + Q_1(\bar{x}, \bar{y}) &\geq R(x, \bar{y}) + Q(x, \bar{y}) \quad \text{and} \\ N(\bar{x}, \bar{y}) + Q_2(\bar{x}, \bar{y}) &\geq C(\bar{x}, y) + Q_2(\bar{x}, y) \end{aligned}$$

for all x, R, y, C. Indeed, one can show following the proof of Theorem 4.2 in [16] that $(\bar{x}, M)$, $(\bar{y}N)$ will form an equilibrium point for the undiscounted stochastic game Γ. This concludes the proof of Theorem 2.2.

3 VLCP and ordered field property for the generalized SER-SIT Stochastic Games

We start with the description of VLCP, a generalization of Linear Complementarity Problem due to Cottle and Dantzig [1]. Let $A \in R^{m\times k}$. Let A be partitioned into k blocks. That is, $A = [A61, A62, \ldots, A^k]^T$. Suppose A^j is of order $m_j \times k$, $j = 1, 2, \ldots, k$. Clearly, $\sum_{j-1}^{k} m_j = m$. Then, we say A is a vertical block of type $(m_1, m_2, \ldots, m_k)$. Given a vertical block matrix $A \in R^{m\times k}(m \geq k)$ of type $(m_1, m_2, \ldots, m_k)$ and $q \in R^m$, where $\sum_{j-1}^{k} m_j = m$, the generalized linear complementarity problem is to find $w \in R^m$ and $z \in R^k$, such that,

$$\begin{aligned} w - Az &= q, w \geq 0, z \geq 0 \text{ and} \\ z_j \Pi_{i-1}^{m_j} w_i^j &= 0 \text{ for } j = 1, 2, \ldots, k \end{aligned}$$

where, w_i^j denotes the i^{th} row in the j^{th} block matrix. This generalization is also known as Vertical Linear Complementarity Problem and is denoted by VLCP (q, A) in the literature [11,6]. Recently Mohan, Sridhar [11] had formulated certain classes of stochastic games and gave an algorithm to solve such games. The generalized SER-SIT stochastic game can also be formulated as a VLCP, borrowing ideas from [11]. Also the generalized SER-SIT stochastic game can be solved using the algorithm given in [11].

We illustrate the algorithm with an example.

Example 3.1 Let $\mathcal{A} = \left\{ \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix} \right\}, \mathcal{B} = \left\{ \begin{bmatrix} 5 & 1 \\ 1 & 3 \end{bmatrix}, \begin{bmatrix} 2 & 3 \\ 4 & 2 \end{bmatrix} \right\}$ where the entries of the above matrices have been obtained after considering the rewards and the transition probabilities. We have not given explicitly the rewards, transition probabilities etc. for simplicity. The VLCP formulation is, following [11]. The matrix which appears below is written as following. First two rows are obtained taking the first row from the two matrices in $\mathcal{A}$ and 5^{th} 6^{th} rows are obtained taking the first column from the two matrices in $\mathcal{B}$. Similarly other rows are written.

$$\begin{pmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 4 \\ 5 & 1 & 0 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 3 & 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix} + \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ w_6 \\ w_7 \\ w_8 \end{pmatrix}$$

We can see that $(1, 0, 0, 1)$ is a solution to the above VLCP. Here z_1, z_2 corresponds to player 1 and z_3, z_4 corresponds player 2. Since $z_2 = 0$, player 1 can choose any second row of $N \in \mathcal{A}$ as the second row in his choice of row representative matrix. Since $z_1 > 0$ and $w_1 > 0$, $w_2 > 0$ player 1 will choose (3,1) as first row in his row representative matrix. In other words, player 1 can have either $\begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$ and $\begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix}$ as his choice of row representative matrix. Similarly, player 2 can have either $\begin{pmatrix} 5 & 1 \\ 1 & 3 \end{pmatrix}$ or $\begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}$ as column representative matrix. Thus, $\left\{(1,0), \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}\right\}, \left\{(0,1), \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}\right\}$ form an equilibrium pair.

Theorem 3.1 Assumptions as in Theorem 2.1. Then, the discounted generalized SER-SIT stochastic game possess the ordered field property.

Theorem 3.2 Assumptions as in Theorem 2.2. Then, the undiscounted generalized SER-SIT stochastic game possess the ordered field property.

Remark 3.1 When $\mathcal{A}$ (and as well $\mathcal{B}$) contain just one matrix, then our result specializes to an earlier result on SER-SIT stochastic games due to Parthasarthy, Tijs and Vrieze

[16].

4 Further remarks and open problems

It is clear that Theorem 2.1 (and Theorem 2.2) is true as long as the number of matrices in $\mathcal{A}$ (and $\mathcal{B}$) is finite.

The main advantage of enlarging the class is that we are able to give an algorithm by formulating the problem as a VLCP. This we may not be able to achieve but perhaps there are other ways to circumvent this difficulty. At the moment we do not have an algorithm like Cottle-Dantzig to get a generalized equilibrium point if we do not enlarge the class.

Throughout the paper, we considered only two-person generalized SER-SIT games. One can perhaps extend the notion to more than two players. We are not sure whether we will be able to give algorithm as in the two-player game.

5 Acknowledgement

We thank Drs. S. R. Moha,, S. K. Neogy and G. Ravindran for some useful discussions regarding the contents of the paper. We are grateful to Dr. Seetharama Gowda for his interest in this work - in fact, we wanted him to be one of the co-authors of this article but he wisely declined it.

References

1. Cottle R. W, G. B. Dantzig (1970) *A generalization of the linear complementarity problem,* Journal of Combinatorial Theory 8: 70-90.

2. Cottle R. W. J. S. Pang, R. E. Stone (1992) *The Linear Complementarity Problem,* Academic Press, New York

3. Filar J, K. Vrieze (1997) *Competitive Markov Decision Process.* Springer-Verlag, New York

4. Fink A. M. (1964) *Equilibrium in a Stochastic n-person Game,* Journal of Science of Hiroshima University Series A-I 28; 89-93.

5. Gillette D (1957) *Stochastic Games with Zero Stop Probabilities. In: Tucker A. W., M Dresher, P. Wolfe (eds.)* Contributions to the Theory of Games, Vol.III (Annals of Mathematics Studies, No.39, Princeton University Press, Princeton), pp. 179-187.

6. Gowda M. S., R. Sznajder (1994) *The generalized Order Linear Complementarity Problems.* SIAM Journal on Matrix Analysis and Applications 15; 779-795.

7. Gowda M. S., R. Sznajder (1996) *A Generalization of Nash Equilibrium Theorem on Bimatrix Games.* International Journal of Game Theory 25; 1-12.

8. Lemke C. E. (1965) *Bimatrix Equilibrium points and mathematical programming* Management Science 11; 681-689.

9. Lemke C. E. (1970) *Recent results on complementarity problems* In: Rosen J. B., O.L.Mangasarian, K. Ritter (eds.) Nonlinear Programming, Academic Press, New York, pp. 349-384.

10. Lemke C. E., J. T. Howson (1964) *Equilibrium points of bimatrix games.* SIAM Journal on Applied Mathematics 12; 413-423.

11. Mohan S. R., S. K. Neogy, R. Sridhar (1996) *The generalized linear complementarity problem revisited.* Mathematical Programming Series A 74; 197-218.

12. Mohan S. R., S. K. Neogy, T. Parthasarthy, S. Sinha (1999) *Vertical linear complementarity problem and discounted zero-sum stochastic games with ARAT structure.* Mathematical Programming Series A 86; 637-648.

13. Nash J. F. (1951) *Noncooperative Games.* Annals of mathematics 54; 286-295.

14. Parthasarthy T, T. E. S. Raghvan (1971) *Some topics in two-person games* American Elsevier Publishing Company, New York.

15. Parthasarthy T. T. E. S. Raghavan (1981) *An ordered field property for stochastic games When One Player Controls Transition Probabilities.* Journal of optimization theory and its applications 33; 375-392.

16. Parthasarthy T, S. H. Tijs, O. J. Vrieze (1984) *Stochastic games with State Independent Transition and Separable Rewards* In: Hammer G. D. Pallashke (eds.) Lecture notes in Economics and Mathematical Systems NO.226, Springer-Verlag, pp. 262-271.

17. Raghavan T. E. S., J. A. Filar (1991) *Algorithms for Stochastic Games - A Survey* Zeitschrift fúr Operations Research 35; 437-472.

18. Shapley L. S. (1953) *Stochastic Games,* Proceedings of the National Academy of Sciences, USA 39; 1095-1100.

19. Sobel M. J. (1971) *Non-cooperative Stochastic Games.* Annals of Mathematical Statistics 42; 1930-1935.

20. Takahashi M. (1964) *Equilibrium points of Stochastic Non-cooperative n-person Game.* Journal of Science of Hiroshima University Press A-I 28; 95-99.

21. Vandenberghe L., B. De Moor, J. Vandewalle (1989) *The Generalized Linear Complementarity Problem Applied to the Complete Analysis of Resistive Piece-wise Circuits.* IEEE Transactions on Circuits and Systems 36; 1382-1392.

22. Zamir U. S. (1999) *Algorithms for Stochastic Games and Related Topics.* Ph.D Thesis submitted to the University of Illinois at Chicago, Chicago.