Orderfield Property of Stochastic Games via Dependency Graphs

(Extended Abstract)

Nagarajan Krishnamurthy[1, 2], T Parthasarathy[3, 4], G Ravindran[3]

1 Introduction

Some classes of stochastic games have been shown to possess the orderfield property (e.g. [3]) which means, given the input data - payoffs, transition probabilities and a discount factor (in case of discounted stochastic games) from an ordered field, such games have at least one solution, that is, an optimal value (for zero-sum games) or a Nash equilibrium value (for non-zero-sum games) and corresponding strategies (probabilities with which the actions are played) in the same ordered field. In this paper, we shall consider rational inputs to classes of stochastic games with finite number of states and finite number of actions.

We propose the concept of *dependency graphs* of stochastic games (and mixtures of classes of stochastic games) and use it to derive conditions which are sufficient for a class of stochastic games to possess the orderfield property. Given a stochastic game or class, these conditions can be checked in polynomial time. As far as we know, previous game theoretic work involving graphs have primarily been on Win-Lose Games [6] which are games with 0-1 payoffs and a graph with the payoff matrices as adjacency matrices is constructed, Graphical Games [7] where the graph represents interacting players (neighbors) in a multi-player setup, or Simple Stochastic Games [1]. Moreover, we found no previous work on computational aspects of the orderfield property – that is, given a stochastic game, is there an efficient way to check if it has the orderfield property or not?

We briefly discuss these concepts in the following sections.

2 Stochastic Games

A *two-player finite state, finite action space stochastic game* consists of

1. A finite, non-empty state space $S = \{1, 2, ..., s\}$.
2. Finite, non-empty sets $A_1 = \{1, 2, ..., m\}$ and $A_2 = \{1, 2, ..., n\}$ of actions for players 1 and 2 respectively.
3. Immediate rewards $r_1(t, i, j)$ for player 1 and $r_2(t, i, j)$ for player 2, where $t \in S$, $i \in A_1$, $j \in A_2$, when the game is in state *t* and the players choose actions i and j respectively. If $r_2 = -r_1$ then, we have a zero-sum game. We will use *r* in place of r_1 in case of zero-sum games. We will denote the matrix of immediate rewards by R_1 and R_2 for players 1 and 2 respectively (and by R in case of zero-sum games).
4. A transition law $q = (q(t' \mid t, i, j): (t, t') \in S \times S, i \in A_1, j \in A_2)$, where $q(t' \mid t, i, j)$ denotes the probability of transition from state *t* to state *t′* given that player 1 and player 2 choose actions $i \in A_1$, $j \in A_2$ respectively. We will use $q(t, i, j)$ to denote the corresponding probability distribution.

Given a starting state $t_0 \in S$, the players simultaneously choose actions $i_0 \in A_1$ and $j_0 \in A_2$, resulting in the payoffs $r_1(t_0, i_0, j_0)$ and $r_2(t_0, i_0, j_0)$ to players 1 and 2 respectively. In case of zero-sum games, player 2 pays a sum $r(t_0, i_0, j_0)$ to player 1. (In the case of non-zero-sum games the players 1 and 2 get payoffs $r_1(t_0, i_0, j_0)$ and $r_2(t_0, i_0, j_0)$, respectively). The game moves to a new state t_1 according to $q(t_0, i_0, j_0)$ and the game continues infinitely.

In general, strategies can depend on complete histories of the game until the current stage. Such strategies are called *behavioral strategies*. We shall look at the simpler class of *stationary strategies* which depend only on the current state *t* and not on how *t* was reached.

For player 1, a *mixed stationary strategy*, *f*, is a function from the state space *S* to the set of probability distributions P_{A1} on the action set A_1, for each state. Similarly for player 2, a *mixed stationary strategy*, g, is a function from the state space, *S*, to the set of probability distributions, P_{A2}, on the action set, A_2. *Pure stationary strategies* are simply a set of actions, one per state.

Given strategies *f*, *g* (pure or mixed) and an initial state t_0, we define the β-discounted payoff

$$[I_\beta(f,g)](t_0) = \sum_{t=0}^{\infty}(\beta^t r_t(t_0, f, g)), \beta \in (0,1) \tag{1}$$

---

1 Email id: naga@cmi.ac.in, naga.research@gmail.com
2 Computer Science, Chennai Mathematical Institute, Chennai, India.
3 SQC & OR Unit, Indian Statistical Institute, Chennai Center, India.
4 The author would like to thank the Indian National Science Academy (INSA) for financial support.

and the undiscounted (or limiting average) payoff

$$[\phi(f,g)](t_0)=\liminf_{T\uparrow\infty}[1/(T+1)]\sum_{t=0}^{T}(r_t(t_o,f,g)) \tag{2}$$

Here, $r_t(t_0, f, g)$ is the expected reward (and not the immediate reward) to player 1 at the t^{th} stage.

In the undiscounted zero-sum game, a pair of strategies (f^*, g^*) is optimal, if $\forall\ t \in S$

$$[\phi(f,g^*)](t) \leq [\phi(f^*,g^*)](t) \leq [\phi(f^*,g)](t) \tag{3}$$

for any pair of strategies (f, g) of players 1 and 2.

The definition for discounted case is similar. For the non-zero sum case, we define Nash equilibria as for bimatrix games. The extension of two-player stochastic games (zero/ non-zero sum, discounted/ undiscounted), as defined above, to stochastic games with 3 or more players is natural.

3 Classes and Mixtures of Classes of Stochastic Games with the Orderfield Property

Examples of 2-player discounted stochastic games with the orderfield property are zero-sum and non-zero-sum single controller stochastic games, SER-SIT (Separable Reward – State Independent Transition) games, perfect information stochastic games, zero-sum ARAT (Additive Reward – Additive Transition) games etc. We point the reader to Raghavan's survey [4] for a comprehensive list of such classes and relevant pointers. We define these classes below:

- ***Stochastic Games with Perfect Information*** are stochastic games in which in every state, the action space of (at least) one of the players is a singleton.
- ***Single Controller Stochastic Games*** are stochastic games in which only one of the players controls the transitions. When player 2 controls transitions, this means $q(t' \mid t, i, j) = q(t' \mid t, j) \quad \forall\ i \in A_1, j \in A_2, t, t' \in S$.
- ***Switching Control Stochastic Games*** are games where the law of motion is controlled by player 1 alone when the game is played in a certain subset of states and by player 2 alone when the game is played in other states. In other words, a switching control game is a stochastic game in which the set of states are partitioned into sets S_1 and S_2 where the transition probabilities are given by
    - $q(t' \mid t, i, j) = q(t' \mid t, i)$,
      for $t' \in S$, $t \in S_1$, $i \in A_1$ and $\forall\ j \in A_2$
    - $q(t' \mid t, i, j) = q(t' \mid t, j)$,
      for $t' \in S$, $t \in S_2$, $j \in A_2$ and $\forall\ i \in A_1$
- In ***SER-SIT Games*** the rewards are separable, that is, $r(t, i, j) = c(t) + a(i, j)$ for the zero-sum case. For the non-zero sum case,

  $r_1(s, i, j) = c(s) + a(i, j)$ for P_1,
  $r_2(s, i, j) = d(s) + b(i, j)$ for P_2

  The transitions are state independent, that is,

  $q(t' \mid t, i, j) = q(t' \mid i, j)\ \forall i\in A_1, j\in A_2, t, t'\in S$.
- In ***ARAT Games***

  (i) the reward function can be written as the sum of two functions, one depending on player 1 and the other on player 2. For the zero-sum case, $r(t, i, j) = r_1(t, i) + r_2(t, j)\ \forall\ i \in A_1, j \in A_2, t \in S$, and

  (ii) the transition probabilities can be written as a sum of two functions, one depending on player 1 and the other on player 2. For the zero-sum case, $q(t' \mid t, i, j) = q_1(t' \mid t, i) + q_2(t' \mid t, j)$, $i \in A_1$, $j \in A_2$, $(t, t') \in S\times S$.
  The transition probabilities can be written in a similar way.

Mixtures of some of these classes have been shown to have the orderfield property as well. For example, Neogy et al. ([8]) show that a mixture of Single Controller and ARAT Games has the orderfield property. On the other hand, it is also known that given classes of stochastic games having the orderfield property, mixtures of these classes may not have the orderfield property. For example, Sinha [9] discusses the following example (fig. 1.) where a mixture of two (zero-sum) SER-SIT games does not possess the orderfield property. Sinha [9] shows that the value of this game is irrational (though the inputs are rational), by actually computing the value.

State 1

0 (1, 0, 0, 0)	0 (0, 0, ½ , ½)
0 (0, 1, 0, 0)	0 (1, 0, 0, 0)

State 2

-1 (1, 0, 0, 0)	-1 (0, 0, ½ , ½)
-1 (0, 1, 0, 0)	-1 (1, 0, 0, 0)

State 3

-1 (1, 0, 0, 0)	-1 (0, 0, ½ , ½)
0 (0, 1, 0, 0)	0 (1, 0, 0, 0)

State 4

0 (1, 0, 0, 0)	0 (0, 0, ½ , ½)
1 (0, 1, 0, 0)	1 (1, 0, 0, 0)

Fig. 1. (In each cell, the entries on top are the rewards and the numbers in parentheses are the probabilities of transition to states 1, 2, 3, 4 respectively).

4 Dependency Graphs, Sufficient Conditions and Our Algorithm

We define two types of dependency graphs as described below. Given a stochastic game, Γ, we construct the dependency graph G = (V, E) as follows:

i. Type I: The graph has one vertex per state of the game. We draw an edge from vertex u to v if there exists a positive transition probability from u to v for some choice of actions of the players.

ii. Type II: The graph has one vertex per (state, action pair) combination. Though this graph is huge, it aids structural analysis in cases where a pure optimal (or Nash equilibrium) exists. For example, perfect information stochastic games.

We modify the above SER-SIT mixture (fig. 1.) to get the following instance of the SER-SIT mixture class (fig. 2.). Without actually computing the value and by just looking at the transitions, we claim that this new mixture has a rational value. This follows from the fact that SER-SIT games possess the orderfield property and the dependency graph (Type I as described in (i) above) does not have cycles involving vertices from both the (given) SER-SIT games. In this example, states 3 and 4 do not have transitions into states 1 and 2 and hence this SER-SIT game (leaving out states 1 and 2) can be solved independently. We can then plug in the values of states 3 and 4 into the other SER-SIT game and solve for states 1 and 2 as well.

This technique can be used for a mixture of more than two types of classes as well. We can also use this technique for complicated mixtures where one class is a zero-sum game, another is a non-zero sum game, a third class has 3 players (say). (It is important that these constituent classes already possess the orderfield property).

This works even if a "sink" (absorbing vertex in the graph) is an undiscounted game; and does not always work if an internal vertex is undiscounted. For example, "Big Match" [2] is an acyclic (not considering self-loops) mixture of 3 undiscounted games, considering each state as originally absorbing and a separate game. (We have misused notation here and are talking about the graph where each vertex is a class/ game. Class III in fig. 3 is one such undiscounted game which is a sink).

The above technique also works for subclasses where there are cycles involving more than one class. Fig. 3 talks of one such example. We discuss this in detail in the full version of our paper.

State 1

0 (1, 0, 0, 0)	0 (0, 0, ½ , ½)
0 (0, 1, 0, 0)	0 (1, 0, 0, 0)

State 2

-1 (1, 0, 0, 0)	-1 (0, 0, ½ , ½)
-1 (0, 1, 0, 0)	-1 (1, 0, 0, 0)

State 3

-1 (0, 0, 1, 0)	-1 (0, 0, ½ , ½)
0 (0, 0, ¼, ¾)	0 (0, 0, 0, 1)

State 4

0 (0, 0, 1, 0)	0 (0, 0, ½ , ½)
1 (0, 0, ¼, ¾)	1 (0, 0, 0, 1)

Fig. 2. (In each cell, the entries on top are the rewards and the number in parentheses are the probabilities of transition to states 1, 2, 3, 4 respectively).

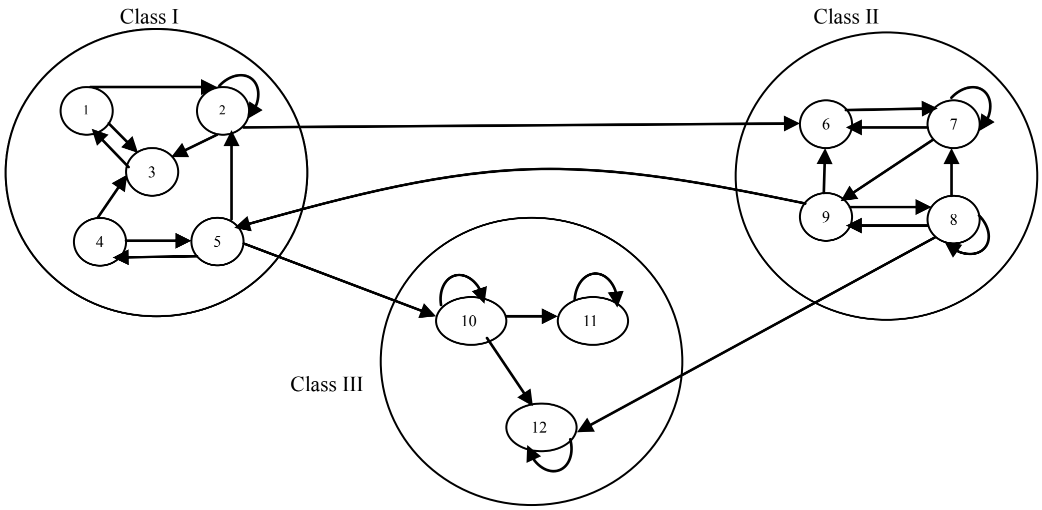

Fig. 3. Example of a subclass of a mixture of 3 stochastic game classes with the orderfield property. An arrow from state i to state j means that there exists a non-zero transition probability from state i to state j for some choice of actions for the players. These classes can be any of the classes that possess the orderfield property. For example, let Class I comprising of states 1, 2, 3, 4 and 5 be a 3-player discounted zero-sum perfect information game, let Class II comprising states 6, 7, 8 and 9 be a 2-player discounted zero-sum ARAT game and let Class III be a 2-player undiscounted non-zero-sum SER-SIT game. Class III is a "sink" as there are no edges going out of this class into the other classes. Though there are cycles through states from both Class I and Class II ($2\to6\to7\to9\to5\to2$ is one such cycle), this game has a rational value (given rational inputs) due to the following reasons. There is an edge from state 2 (in Class I) to state 6 (in Class II) and from state 9 (in Class II) to state 5 (in Class I) and these are the edges responsible for the cycles involving both Classes I and II. Now, as Class I has no cycle with states 2 and 5, Class II has no cycle with states 7 and 9, our claim follows. In this example, Class III does not depend on the other classes and we can solve it independently. Then we can solve the game with states 6 and 7, solve the game with states 1, 2 and 3 next, followed by the game with states 4 and 5 and finally, the game with states 8 and 9.

The following theorems are formal statements of the above "sufficient conditions". We have informally outlined the proof above. Formal proofs can be found in the full version of the paper.

THEOREM 1: *(Sufficient Condition for a Discounted Stochastic Game to Possess the Orderfield Property)*:

Let Γ be a 2-player, discounted stochastic game. Let the set of states, S, of Γ be partitioned into S_1, S_2 and S_3, where states in S_1 are controlled by player 1, states in S_2 are controlled by player 2 and there is no restriction on states in S_3. Let G = (V, E) be the dependency graph of Γ. Γ is such that states in S_3 are not a part of any cycle in G. If G has cycles, each cycle has either vertices from S_1 or vertices from S_2. In other words, Γ has only simple non-intersecting cycles of vertices from S_1 or S_2. Let C_Γ be the class of stochastic games consisting of all games Γ as defined above. Then C_Γ possesses the orderfield property.

THEOREM 2: *(Sufficient Condition for a Mixture of Classes of Discounted Stochastic Games to Possess the Orderfield Property)*:

Given classes $C_1, C_2, \ldots, C_n$ of discounted stochastic games that have the orderfield property, a subclass, C, of a mixture of these classes has the orderfield property if any of the following holds:

(1) The dependency graph G of each $\Gamma \in C$ has only simple non-intersecting cycles as described in Theorem 1 above.
(2) The dependency graph G of each $\Gamma \in C$ has no cycle involving vertices from different classes. That is, if $v_i \in C_i$, $v_j \in C_j$, $i \neq j$, no cycle in G contains both v_i and v_j.
(3) The dependency graph G of each $\Gamma \in C$ may have cycles involving vertices from different classes, but those vertices in these classes do not form a cycle amongst themselves.

THEOREM 3: *(Sufficient Condition for a Mixture of Classes of Discounted and Undiscounted Stochastic Games to Possess the Orderfield Property)*:

In addition to classes in theorems 1 and 2, we allow classes of undiscounted stochastic games with no edges going from these classes into other classes.

ALGORITHM: (*To Determine if a Stochastic Game Belongs to One of the Classes as Described in the Above Theorems and hence has a Rational Solution. The Algorithm can Also Be Extended to Take a Class of Stochastic Games as Input and Verify if it is a Subclass of One of the Above Classes*):

Input: A stochastic game Γ – number of players, finite number of states, rational payoffs for each state, rational transition probabilities for each pair of actions, a rational discount factor (for discounted games).

Step 1: Construct G, the dependency graph.
Step 2: Find cycles in G.
Step 3: Go through all states in Γ to determine special structures (such as ARAT, SER-SIT) as defined above.
Step 4: Check if conditions of theorem 1 or 2 or 3 are satisfied. If "yes", then the given game has a rational solution - rational value and rational strategies. (If the input is a Class of Stochastic Games, then we conclude that the Class possesses the orderfield property).

[One way to implement this algorithm is to color the vertices. For example, if there are states satisfying conditions of theorem 1, color all states in S_1 with one color, states in S_2 with a second color and color each state in with 2 different colors. After finding cycles in G, we perform a DFS (Depth First Search) to verify that there are only monochromatic cycles (if there are any cycles). Similarly, we can check conditions of theorems 2 and 3 as well. Details regarding the algorithm and its time complexity can be found in the full version of the paper].

References

1. Anne Condon, 1992, "The Complexity of Stochastic Games", Information and Computation, Vol. 96, pp. 203-224.
2. D. Blackwell and T. S. Ferguson, 1968, "The Big Match", Ann. Math. Stat., 39, pp. 159-168.
3. T. Parthasarathy and T. E. S. Raghavan, 1981, "An Orderfield Property for Stochastic Games when One Player Controls Transition Probabilities", Journal of Optimization Theory and Applications, Vol. 33, No. 3, pp. 375-392.
4. T. E. S. Raghavan, 2003, "Finite-Step Algorithms for Single Controller and Perfect Information Stochastic Games", NATO Science Series: C, Stochastic Games and Applications, Eds. Abraham Neyman and Sylvain Sorin, Kluwer Academic Publishers Group, Chapter 15.
5. F. Thuijsman and T. E. S. Raghavan, 1997, "Perfect Information Stochastic Games and Related Classes", International Journal of Game Theory, Vol. 26, pp. 403-408.
6. Louigi Addario-Berry, Neil Olver and Adrian Vetta, 2007, "A Polynomial Time Algorithm for finding Nash Equilibria in Planar Win-Lose Games", Journal of Graph Algorithms and Applications, Vol.11, pp.309-319.
7. Michael Kearns, Michael L. Littman and Satinder Singh, 2001, "Graphical Models for Game Theory", Proc. of the 17th Conference in Uncertainty in Artificial Intelligence, pp. 253-260.
8. S. K. Neogy, A. K. Das, S Sinha and A. Gupta, 2007, "On a Mixture Class of Stochastic Game with Ordered Field Property", World Scientific Book, pp. 25-53.
9. Sagnik Sinha, 1989, "A Contribution to the Theory of Stochastic Games", Ph.D. thesis, Indian Statistical Institute, New Delhi, India.