Stochastic Games and their Applications

Gurpreet Singh

150259

1. Introduction

Most of the models we have concerned ourselves with in class are simplistic models unrepresentative of the complexities and possibilities of the real-world. In most realistic problems, there are multiple agents acting in an environment with each other, competing against each other, or both. Bowling and Veloso (2000) provides some examples of multiple RL agents existing in the same environment. Exploration and mapping is an example of cooperating multi-agent system in which multiple agents are trying to find the map of an unknown environment. A game of chess between two agents is an example of a competitive environment, whereas a game of soccer between two teams of robots could be considered as an example of an environment in which agents exhibit both cooperative as well as competitive behaviour.

These examples cannot be modeled as simple dynamic or bayesian games, however require to be formulated using more complex models. Stochastic games deliver such models to extricate us out of this situation. The examples discussed above can infact be formulated as stochastic games in which agents perform actions in a state of the environment, thereby transitioning the environment to new states with the transition modeled as a stochastic process. The promise of stochastic games is to simulate an agent or multiple agents in an environment on the notion of utility (rewards) based on a set of game characteristics. Stochastic games find applications in many fields, such as Industrial Organization, Macroeconomics and Political Economy.

A stochastic game is an extensive game played by one or more players in a sequence of stages. The characteristics of the game at each stage are described by a state, a set of player strategies, and transition probabilities. In a general sense, a stochastic game is a collection of normal-form games played by a set of agents (players) repeatedly. The particulars of the game played at any time depend probabilistically on the previous game played and the actions of the agents in that game. Stochastic games, in this sense, extend Markov Decision Processes to involve multiple agents and simulate them simultaneously in a defined environment.

At the beginning of each stage, the game is in some state. The players select actions and each player receives a payoff that depends on the current state and the chosen actions. The game then moves to a new random state whose distribution depends on the previous state and the actions chosen by the players. The procedure is repeated at the new state and play continues for a finite or infinite number of stages. The total payoff to

a player is often taken to be the discounted sum of the stage payoffs or the limit inferior of the averages of the stage payoffs. Stochastic games generalize both Markov decision processes and repeated games.

In this review, I survey some of the methods used in stochastic games and discuss some of the common applications of stochastic games in economics. The rest of the survey is organized as follows: Section 2 provides some background on the formal definition of a stochastic game and discusses some of the preliminaries used in the rest of the survey; Section 3 talks about the early works on developing Algorithms for solving Stochastic games categorized with different conditions; Section 4 is devoted to discussion on applications of stochastic processes to some problems in economics and game theory in general.

2. Stochastic Games

Here I formally define what a stochastic game is and what it consists of.

Definition 1.1. A *stochastic game* is defined by

1. A finite set of states $\mathcal{S} = \{s_1, s_2, \ldots s_m\}$ with a stage game defined for each state

2. A finite set of agents or players $\mathcal{I} = \{1, 2, \ldots, n\}$

3. A set of strategies $\mathcal{A}_i$ for each agent i. Define $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_n$

4. A **transition probability $\boldsymbol{\pi}$** $: \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0,1]$ where $\boldsymbol{\pi}\left(s' \,|\, s, \mathbf{a}\right)$ denotes the probability of transitioning to state s' if the strategy profile $\mathbf{a} = (a_1, \ldots, a_n)$ is used when the environment is at state s

5. A set of payoff functions for each agent i given by $u_i : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$

In a more general setting, we can have different sets of strategies/actions for each agent at different states of the game, however our definition makes the inessential but simplifying assumption that each agent's strategy space is the same in all games.

We make two additional simplifying assumptions: the set of states is finite; all strategies are *Stationary Strategies i.e.* the distribution over actions (mixed strategy) depends only on the current state s and is independent on the stage or the time-step. We can extend the definition by removing such assumptions assuming however for the sake of this review we contrain to this simplified definition of a stochastic game.

Example of a Stochastic Game: The Romeo-Juliet Game

Suppose two players Romeo (R) and Juliet (J) are trying to decide what kind of movie they want to see. They will go for a movie only if both of them choose to watch the same kind of movie.

We attempt to formulate this as a stochastic game. Suppose there are two states for the game $\mathcal{S} = \{\text{Avoid}, \text{Meet}\}$. The Avoid state describes a scenario where Juliet wants to avoid Romeo and Meet state describes a scenario where Juliet wants to meet Romeo. Suppose they can go for either an Action Movie

	Juliet	
	A	**R**
A	a, a	b, c
R	c, b	d, d

Romeo — **State**: Meet

	Juliet	
	A	**R**
A	a, a	b, c
R	c, b	d, d

Romeo — **State**: Avoid

Table 1: Romeo-Juliet Game: Utility Table

	Juliet	
	A	**R**
A	0.5, 0.5	1.0, 0.0
R	0.0, 1.0	0.5, 0.5

Romeo — **State**: Meet

	Juliet	
	A	**R**
A	0.5, 0.5	0.0, 1.0
R	1.0, 0.0	0.5, 0.5

Romeo — **State**: Avoid

Table 2: Romeo-Juliet Game: Transition Table

(AM) or Romantic Movie (RM), *i.e.* $\mathcal{A}_R = \mathcal{A}_J = \{AM, RM\}$. The utility functions are given as matrices in Table 1 for both the states.

We also need to give the transition probabilities from each state and each strategy profile. These are also represented using matrices and are given in Table 2. The tuples represent the transition probabilites for a certain strategy profile and a state in the order (Meet, Avoid), *i.e.* the first value at any cell represents the probability of the game moving to the Meet state.

The intuitive explanation of the transition probabilities is that if they go for a movie, there's a 50% chance that the state will change and a 50% chance that the sate will remain the same. If they do not go for a movie, the state remains the same almost surely (*i.e.* with probability 1).

Since the game would be repeated (possibly) infinite number of times (with different states), the net reward for each agent would be an accumulation of rewards (defined formally later) over each stage or repetition. Therefore, the notion of Nash Equilibrium does not hold directly in this context. However, under certain conditions, we can say something about the equilibria in such games.

2.1. Preliminaries

Reward. Before talking about the equilibria in infinitely repeated games such as any stochastic game, we need to define a notion of reward. Unline finite games, we cannot assign the total reward to be a sum of rewards at all stages as that could potentially be infinite. Instead, we define two different kinds of rewards: average reward and discounted reward.

Definition 1.2. Given an infinite sequence of payoffs $(r_1, r_2 \ldots)$ for agent i, the *average reward* of i is given as

$$\lim_{k \to \infty} \sum_{j=1}^{k} \frac{r_j}{k} \tag{1}$$

Definition 1.3. Given an infinite sequence of payoffs $(r_1, r_2 \ldots)$ for agent i, the *future discounted reward* of i is given as

$$\lim_{k \to \infty} \sum_{j=1}^{k} \delta^j r_j \tag{2}$$

where $\delta \in [\, 0, 1 \,]$ is the discount factor.

An agent with discounted reward cares more about the immediate reward than the reward he will obtain in the future. This is representative of standard human nature and makes intuitive sense, therefore discounted reward is used very often in many literatures.

Markov Perfect Equilibrium.   Markov Perfect Equilibrium is a refinement on the concept of the subgame perfect equilibrium. Maskin and Tirole (2001) define Markov perfect equilibrium to be a set of mixed strategies for each of the players which satisfy the following criteria:

1. The strategies have the Markov property of memorylessness, meaning that each player's mixed strategy can be conditioned only on the state of the game. These strategies are called Markov reaction functions.

2. The state can only encode payoff-relevant information. This rules out strategies that depend on non-substantive moves by the opponent. It excludes strategies that depend on signals, negotiation, or cooperation between the players (e.g. cheap talk or contracts).

3. The strategies form a subgame perfect equilibrium of the game.

For the sake of brevity within this review, I use Equilibrium and Markov Perfect Equilibrium exchangeably.

<++>

Since the literature of Stochastic games itself is quite vast, leaving out methods and applications is inevitable. I restrict the discussion to Two-Player Zero-Sum games for the following sections. The review consists of different algorithms used in previous works to solve Two-Player Zero-Sum Stochastic games under categorized under different conditions/classes.

3.   Two Player Zero-Sum Games

Zero-Sum games are purely competitive in which there are two agents and the payoff of one player is the negative of the other. $u_1 = -u_2$. These are called zero-sum games as the payoff functions sum to zero.

Although zero-sum games are simplistic games not representative of the real-world challenges, they have an extensive theory.

Solving a Zero-Sum Matrix Game   We describe a Zero-Sum Matrix Game as a game between two players with each player possessing a finite number of strategies. The matrix $A = \| a_{ij} \|$ where a_{ij} is the payoff of player 1 if she chooses to play strategy i and player 2 chooses to play strategy j. Since it is a zero-sum game, the payoff of player 2 for the same strategy profile is given by $-a_{ij}$ gives the utility values for player 2.

The general optimality principle in a two-person zero-sum game declares that player 1 tends to choose a strategy i^*0 such that

$$ i^* \quad = \quad \arg\max_i \min_j \ a_{ij} \quad \text{and} \quad \underline{v} = \min_j \ a_{i^*j} \tag{3} $$

whereas player 2 tends to choose a strategy j^*

$$ j^* \quad = \quad \arg\min_j \max_i \ a_{ij} \quad \text{and} \quad \overline{v} = \max_i \ a_{ij^*} \tag{4} $$

If $\underline{v} = \overline{v}$, then the pair (i^*, j^*) is called a *saddle point* of the game and the value $a_{i^*j^*}$ is called the *value* of the game. In case such a saddle point does not exist, then there exists no solution in pure strategies and we always have $\underline{v} < \overline{v}$. We alternately define an optimal profile using mixed strategies. Suppose the two players play mixed strategies with the distribution given by two probability vectors $\mathbf{x}^*$ and $\mathbf{y}^*$ then, for player 1, we have

$$ \mathbf{x}^* \quad = \quad \arg\max_{\mathbf{x}} \min_{\mathbf{y}} \ \mathbf{x}A\mathbf{y}^{\mathrm{T}} \quad \text{and} \quad \underline{v} = \min_{\mathbf{y}} \ \mathbf{x}^*A\mathbf{y}^{\mathrm{T}} \tag{5} $$

and, for player 2,

$$ \mathbf{y}^* \quad = \quad \arg\min_{\mathbf{y}} \max_{\mathbf{x}} \ \mathbf{x}A\mathbf{y}^{\mathrm{T}} \quad \text{and} \quad \overline{v} = \min_{\mathbf{x}} \ \mathbf{x}A\mathbf{y}^{*\mathrm{T}} \tag{6} $$

. The solution to this is given using linear programming and is not discussed in this review.

The minimax principle claims that the values $\underline{v}$ and $\overline{v}$ are equal and defined as the *value* of the matrix game. This is true for all finite zero-sum matrix games. The formal definition of the minimax principle is as given below.

Theorem 1.1 (Minimax Principle). Given a real $n \times m$ matrix $A = \| a_{ij} \|$, there exists a pair of probability vectors $\mathbf{x}^* = (x_1^*, x_2^*, \dots, x_m^*)$ and $\mathbf{y}^* = (y_1^*, y_2^*, \dots, y_m^*)$ such that for a unique constant n

$$ \sum_{i'} a_{i'j} x_{i'}^* \geq v \geq \sum_{j'} a_{ij'} y_{j'}^*, \quad \text{for all } i, j \tag{7} $$

where v is the value of the matrix game as defined earlier.

In the following sections, I discuss multiple algorithms to solve differenct cases of Zero-Sum Stochastic Games that originated early in the literature of Stochastic Games. For this, I borrow content presented by Raghavan

and Filar (1991) and Bowling and Veloso (2000).

3.1. Iterative Algorithms for Discounted Zero-Sum Stochastic Games

Algorithm 1: Shapley

Shapley (1953) showed that there exist an optimal solution for δ-discounted stochastic zero-sum games using stationary strategies. The algorithm for finding the solution to such a stochastic game comes from Shapley's proof for the prior statement (for details, refer (Shapley, 1953)). This algorithm is given in Algorithm 1

Algorithm 1: Shapley's Algorithm for Stochastic Games

---

1. Start with a arbitrary approximation for the true value $\mathbf{v}$ of the stochastic game, say
   $$\mathbf{v}^{(0)} = \left(v^{(0)}(s) \right)_{s \in \mathcal{S}}$$

2. Repeat until convergence

   (a) For each state $s \in \mathcal{S}$ compute the matrix

   $$\mathrm{G}_s^{(t)} \quad = \quad \left[a_{ij}^s + \delta \sum_{s' \in \mathcal{S}} \boldsymbol{\pi} \left(s' \,|\, s, (i,j) \right) v^{(t)}(s') : (i,j) \in \mathcal{A} \right]$$

   (b) For each state $s \in \mathcal{S}$ update

   $$v^{(t+1)}(s) \quad = \quad \mathrm{Value} \left(\mathrm{G}_s^{(t)} \right)$$

   where the Value(G) function gives the value of the matrix game G.

3. Compute the optimal stationary strategies $\mathbf{X}^*, \mathbf{Y}^*$ as

   $$\mathbf{x}_1, \mathbf{y}_1 \quad = \quad \mathrm{Solve} \left(G_s^{(T)} \right)$$

   where the Solve(G) function gives the optimal strategy profile for the matrix game G.

In his algorithm, Shapley attempts to define the complete stochastic game as one zero-sum matrix game by approximating the expected gain/reward if the player starts playing at a particular state. This is the reason the update step in the algorithm given by Shapley (Algorithm 1) is analogous to the *Bellman equation* or the *optimality equation* of dynamic programming. The matrix $G_s^{(t)}$ at any step t approximates the expected gain matrix for each player if he starts playing at that state and as $t \to \infty$, the matrix $G_s^{(t)}$ converges to the actual gain matrix. At this point, solving the gain matrices for each state would afford us the optimal solution to the Stochastic Game.

Shapley proves that given an optimal stationary strategy $\mathbf{X}^*$ of player 1 the expected discounted

reward $r(s, \mathbf{X}^*, \mathbf{Y})$ is at least $v(s)$ against any strategy $\mathbf{y}$ of player 2. Similarly for an optimal stationary strategy $\mathbf{y}^*$ of player 2, the expected discounted reward $r(\mathbf{x}, \mathbf{y}^*)$ is at most $v(s)$ against all strategies $\mathbf{x}$ of player 1.

Algorithm 2: Pollatschek and Avi-Itzhak

Although Shapley's algorithm derives near-optimal stationary strategies when $\mathbf{v}^{(t)} \approx \mathbf{v}$, it fails to utilize the information contained in the optimal strategies of $\mathcal{A}^s\left(\mathbf{v}^{(t)}\right)$ at each iteration. Pollatschek and Avi-Itzhak (1969) proposed another algorithm that is based on utilizing both the approximate value vectors and the associated optimal strategies of the matrix games constructed at the intermediate steps.

Algorithm 2: Pollatschek and Avi-Itzhak's Algorithm for Stochastic Games

1. Start with an arbitrary approximation for the true value $\mathbf{v}$ of the stochastic game, say
$$\mathbf{v}^{(0)} = \left(v^{(0)}(s)\right)_{s \in \mathcal{S}}$$

2. Repeat until convergence

   (a) For each state $s \in \mathcal{S}$ compute the matrix
   $$\mathbf{G}_s^{(t)} \quad = \quad \left[a_{ij}^s + \delta \sum_{s' \in \mathcal{S}} \boldsymbol{\pi}\left(s' \mid s, (i,j)\right) v^{(t)}(s') : (i,j) \in \mathcal{A} \right]$$

   (b) For each state $s \in \mathcal{S}$, solve the matrix games $\mathbf{G}^{(t)}$ to find optimal strategies $\left(\mathbf{x}_s^{(t)}, \mathbf{y}_s^{(t)}\right)$. Define $\mathbf{X}^{(t)} = \left(\mathbf{x}_1^{(t)}, \ldots, \mathbf{x}_m^{(t)}\right)$ and $\mathbf{Y}^{(t)} = \left(\mathbf{y}_1^{(t)}, \ldots, \mathbf{y}_m^{(t)}\right)$. Update
   $$\mathbf{v}^{(t+1)} \quad = \quad \left[\mathbf{I} - \delta Q\left(\mathbf{X}^{(t)}, \mathbf{Y}^{(t)}\right)\right]^{-1} r\left(\mathbf{X}^{(t)}, \mathbf{Y}^{(t)}\right)$$

   where $Q(\mathbf{X}, \mathbf{Y})$ gives an $m \times m$ transition matrix for mixed stationary strategies given by $(\mathbf{X}, \mathbf{Y})$ and $r(\mathbf{X}, \mathbf{Y})$ represents the total expected reward when the actions are dictated by the strategy profiles $\mathbf{X}, \mathbf{Y}$ are chosen at every state.

Although the algorithm given by Pollatschek and Avi-Itzhak has special importance because of its relationship with the classical Newton-Raphson procedure (Raghavan and Filar, 1991), the convergence of the algorithm is proven only under some restrictive conditions.

Algorithm 3: Van der Wal

der Wal J (1977) gives an algorithm (Algorithm 3) which uses the concepts of stopping times and successive approximations. Van der Wal has provided a detailed analysis on the convergence guarantees of his algorithm, however, since the maths is quite extensive, I exclude that analysis from

---

Algorithm 3: Van der Wal's Algorithm for Stochastic Games

---

1. Start with an arbitrary approximation for the true value $\mathbf{v}$ of the stochastic game, say $\mathbf{v}^{(0)} = \left(v^{(0)}(s) \right)_{s \in \mathcal{S}}$ and an integer $k > 0$

2. Repeat until convergence

   (a) For each state $s \in \mathcal{S}$ compute the matrix

   $$
   \mathrm{G}_s^{(t)} \quad = \quad \left[a_{ij}^s + \delta \sum_{s' \in \mathcal{S}} \boldsymbol{\pi} \left(s' \,|\, s, (i, j) \right) \mathrm{V}^{(t)}(s') : (i, j) \in \mathcal{A} \right]
   $$

   (b) For each state $s \in \mathcal{S}$, find the optimal strategy $\mathbf{y}_s^{(t)}$ for player 2 in the matrix game $\mathrm{G}^{(t)}$.

   (c) Since we have an estimate of the optimal strategy of one player, the problem is reduced to a MDP[a]. Since it is possible to solve this MDP using the Bellman Equation (updates such as in Shapley's algorithm, we compute the m-th iterate of the value function, $\mathbf{v}^m$ for the one player MDP and assign $\mathbf{v}^{(t+1)} = \mathbf{v}^m$

---

[a]MDP stands for Markov Decision Process, here and in the following text

this survey.

Note that if we fix $m = 1$, then Van der Wal's algorithm reduces to Shapley's algorithm (Algorithm 1) as finding the m-th iterate of the MDP problem is simply making updates using the Bellman equation.

The problem with the above algorithms is that they are not finite step algorithms. More particularly, Algorithms 1-3 provide only asymptotic convergence guarantees, however cannot guarantee convergence in finite steps. The reason for this, as highlighted by Raghavan and Filar (1991), is the lack of *ordered field properties* in general stochastic games. The ordered field property states that the optimal stationary strategies and the value functions be in the same Archimedian field, however, this is not always true for stochastic games. For this reason, it is not possible to formulate finite step algorithms.

Although the above remarks stand true, there has been extensive research to find classes of stochastic games which afford the ordered field property. Raghavan and Filar (1991) further discuss these classes and certain finite step algorithms to deal with each class.

3.2. Algorithms for Undiscounted Zero-Sum Games

The problem with finding optimal strategies for undiscounted zero-sum games us that the players might not possess optimal stationary strategies D (1957). This raises two questions in the literature -

1. Do undiscounted stochastic games possess the value vector?

2. What characterizes the class of undiscounted stochastic games solvable in optimal stationary strategies?

Mertens and Neyman (1981) proved that we can find value vectors for undiscounted stochastic games, therefore, answering the first question. Following Gillette, there was extensive researching on finding sufficient conditions for the existence of optimal stationary strategies. In this survey, I discuss only two of these cases and look at one algorithm to solve the games in each of these categories.

The Single-Controller Undiscounted Game

In single-controller games, we assume player 2 to be single controller of states, $i.e.$ $\boldsymbol{\pi}(s' \mid s, (a_1, a_2))$ is independent of a_1. The solution to such a game in the undiscounted case follows from the following theorem.

Theorem 1.2. Let $\mathbf{x}_i, \mathbf{y}_j$ represent a pair of pure stationary strategies in an undisounted single-controller game. Construct an alternate matrix game $\mathrm{H}(s)$ for each state s such that the (i, j)-th entry of this matrix is equal to $r(s, \mathbf{x}_i, \mathbf{y}_j)$, where $r(s, \mathbf{x}_i, \mathbf{y}_j)$ gives the total expected reward given that we play $(\mathbf{x}_1, \mathbf{y}_j)$ at that particular state. Then, for every state s, we have $\mathrm{val}\,(\mathrm{H}(s)) = v(s)$, the value of the stochastic game. Moreover, there exists a common optimal strategy for player 1 of the matrix games $\mathrm{H}(s)$ and this common optimal strategy is the same as the optimal strategy for player 1 in the stochastic game.

Based on the above theorem, OJ (1981) and Hordijk and Kallenberg (1981) independently proposed an algorithm to solve the single-controller undiscounted game. The algorithm is given in Algorithm 4. In order to understand this algorithm better, one can refer to (Raghavan and Syed, 2002) which gives a detailed derivation of the algorithm and Raghavan and Filar (1991) present the working of the algorithm using an example of a rather simple stochastic game.

The SER-SIT Undiscounted Game

In the SER-SIT game, we assume that the rewards are separable, $i.e.$ $u(s, i, j) = c(s) + q(i, j)$ for all s, i, j. We further assume that the transitions are state independent, $i.e.$ $\boldsymbol{\pi}(s, i, j, s') = \boldsymbol{\pi}(i, j, s')$. Parthasarathy et al. (1984) showed that the Undiscounted SER-SIT game has a closed form solution and can be solved in a one-step finite algorithm as well. This algorithm is given in 5.

3.3. General-Sum Stochastic Games

Again we consider only the case of two player games. Similar to the zero-sum case, we denote the expected δ-discounted payoff for player k as $r_\delta^k(\mathbf{x}, \mathbf{y})$ when strategy profile $(\mathbf{x}, \mathbf{y})$ is used. Although solving such games is still an open problem, there are certain classes of games for which it is possible to find Nash Equilibria. A rather recent work by Prasad and Bhatnagar (2012) show, however, that it is possible to find sufficient as well as necessary conditions that would categorize the existence of

---

Algorithm 4: Solving a Single-Controller Undiscounted Game

---

1. Let $(\mathbf{v}^*, \mathbf{w}^*, \mathbf{x}^*)$ be any optimal solution to the linear program

$$\max_{\mathbf{v},\mathbf{x},\mathbf{w}} \sum_s v(s)$$

   subject to the following conditions for all s, j

   (a) $v(s) \geq \sum_{s'} \boldsymbol{\pi}(s' \,|\, s, j) v(s')$
   (b) $v(s) + w(s) \geq \sum_i u(s, i, j) x_i(s) + \sum_{s'} \boldsymbol{\pi}(s' \,|\, s, j) w(s')$
   (c) $\sum_i x_i(s) = 1$
   (d) $x_i(s) \geq 0$

2. Compute the dual solution $(\mathbf{f}, \mathbf{g}, \mathbf{h})$, where $\mathbf{f}$, $\mathbf{g}$, and $\mathbf{h}$ are the dual variables (or the Lagrange multipliers) respectively for conditions (a)-(c) given above. We compute the optimal strategy $\mathbf{y}^*$ as

$$y_j^*(s) \quad = \quad \begin{cases} g_j(s)/g_s & \text{for all } j, \text{ if } s \notin S' \\ f_j(s) + g_j(s)/d_s & \text{for all } j, \text{ if } s \in S' \end{cases}$$

   where $g_s = \sum_j g_j g_j(s)$ and $d_s = \sum_j f_j(s) + g_j(s)$ for every s, and set $S' = \left\{ s \,|\, f_s = 0 \right\}$. The optimal stationary strategy profile, therefore, is given as $(\mathbf{x}^*, \mathbf{y}^*)$

Algorithm 5: Solving a SER-SIT Undiscounted Game

---

1. Solve the matrix game $\mathrm{W}(\mathbf{c}) = \left[q(i, j) + \sum_{s'} \boldsymbol{\pi}(i, j, s') c(s') \right]$ and let $\mathbf{f}, \mathbf{g}$ denote the optimal strategy profile for this game. Let u denote the value of the matrix game $\mathrm{U}(\mathbf{c})$.

2. Compute the value vector as $\mathbf{v}^* = u\mathbf{1}$, where $\mathbf{1}$ denotes a vector of ones, $\mathbf{x}^*(s) = \mathbf{f}$ and $\mathbf{y}^*(s)$ for all s for optimal stationary strategies for both players.

Nash Equlibria.

Theorem 1.3. (Raghavan and Filar, 1991) Under the following conditions/cases, a general-sum stochastic game has equilibria in stationary strategies.

1. The transition matrix is irreducible for any set of stationary strategy choice by the players.

2. There is a special state s^* which is visited infinitely often with probability 1.

3. There exists a non-empty set of states A and a positive integer N such that from any state s, the game moves to some state in A with an expected waiting time of at most T.

4. Single-controller stochastic games

5. SER-SIT games

There have been recent works which attempt to find out optimal strategies for the general-sum case as well. One of the recently popular algorithms is the Nash-Q algorithm proposed by Hu and Wellman (2003).

4. Applications of Stochastic Games in Economics

4.1. Features of Economic Applications

There are a variety of reasons Stochastic Games are popular. The motivations behind applying stochastic games to economis are because of the features of the general models used in economics. These features are briefly listed below (Amir, 2003)

1. **Discounted Payoffs** Situations where finite-time events are irrelevant are unnatural in economics. The presence of positive rate of interest is ubiquitous in economic life. For example, if someone offered you Rs. 100 today or Rs. 110 a year from now, the obvious choice is the former. This is because for beings, short-term profits outweight long-term profits.

2. **Pure Strategies** Mixed strategies have limited acceptance due to lack of compelling universal interpretation and their inherent ex-post regret property. They are considered only in cases where pure strategy equilibria do not exist.

3. **Uncountable state and action sets** Most models in economics have infinite state and action states however there are many applications of game theory in economics which use finite sets, for example models following discrete pricing.

4. **Simplicity** Most economic models assume simple action sets to avoid complications and allow closed-form equilibrium strategies.

5. **Predicitive power of models** Most applications are motived by the search of clear-cut conclusions. For this reason, most models are formulated using a highly structured and relatively aggregated models of stochastic games. This is another reason to motivate the study of Stochastic Games in different categorizations.

4.2. Examples of Economic Applications of Stochastic Games

This section briefs on a few applications of Stochastic Games without discussing much on finding the optimal strategies for the said games.

Strategic Resource Extraction or Capital Accumulation

Levhari and Mirman (1980) formulated a game between two agents who noncooperatively exploit a natural resource. Two agents compete to consume a natural resource with stock $z^{(t)}$ at time step t.

We represent Agent i's consumption at time t using $a_i^{(t)}$ and δ_i denoting his discount factor $(i = 1, 2)$. To model this as a stochastic game, we need to find the state and action sets, the payoff function and the state transition function. We give them as follows

1. The state set is given as the possible values of the resource stock at any point of time, *i.e.* $\left\{ z \,\middle|\, z > 0 \right\}$

2. The action set for both the players is the set of possible amounts they can consume. Since these are state dependent, we can say for each agent $i = 1, 2$, we have the action set $\mathcal{A}(z) = \left\{ (a_1, a_2) \,\middle|\, a_1, a_2 \geq 0 \text{ and } a_1 + a_2 \leq z \right\}$

3. Since the transitions are deterministic, the transition probability function assigns zero probability to all states other than the one given by the stock update equation. The stock update equation says $z^{(t+1)} = \left(z^{(t)} - a_1^{(t)} - a_2^{(t)} \right)^{\alpha}$ *i.e.* we have

$$\boldsymbol{\pi}\left(z, a_1, a_2, z'\right) = \begin{cases} 1 & \text{if } z' = (z - a_1 - a_2)^{\alpha} \\ 0 & \text{else} \end{cases}$$

4. The payoff functions for both the players are equal and defined as $u_1(a) = u_2(a) = \log a$. The discounted reward, therefore, is given as $r^i(\mathbf{a}_i) = \sum_t \delta_i^t \log a_i^{(t)}$ for $i = 1, 2$.

Levhari and Mirman (1980) showed, using induction, the existence of Markovian equilibiria with linear consumption strategies and logarithmic value functions. Furthermore, the equilibirum resource stock converges to a unique globally stable steady-state level given by $\overline{z} = \left(\frac{1}{\alpha \delta_1} + \frac{1}{\alpha \delta_2} - 1 \right)^{\frac{\alpha}{\alpha - 1}}$. This problem, solved using Stochastic games, has seen many applications and analysis in the literature, such as in works by Dutta and Sundaram (1992, 1993a,b) or cases in continuous time.

4.3. The Class of Linear-Quadratic Games

General linear-quadratic games are categorized by a quadtratic cost function that the agents want to minimize (or maximize the negative of the cost function). This cost function is given as

$$J_i(\mathbf{a}_1, \mathbf{a}_2) \quad = \quad \sum_{t=1}^{T} \left\{ \mathbf{z}^{(t+1)} \mathbf{Q}_i^{(t+1)} \mathbf{z}^{(t+1)\,\mathrm{T}} + \sum_j \mathbf{a}_j^{(t)} \mathbf{R}_{ij}^{(t)} \mathbf{a}_i^{(t)\,\mathrm{T}} \right\}$$

where the matrices $\mathbf{R}_{ij}^{(t)}$ are positive-definite and the matrices $\mathbf{Q}_i^{(t+1)}$ are symmetric and positive semi-definite. It is also possible to have a continuous time space in which case the cost function would be converted to an integral. The state vector in this case, again, is given as a deterministic transform from the previous state

$$\mathbf{z}^{t+1} = \mathbf{A}^{(t)} \mathbf{z}^{(t)} + \sum_j \mathbf{B}_j^{(t)} \mathbf{a}_j^{(t)}$$

One can formulate the conditions under which a Markov equilibrium can be given in closed form. These

are discussed in detail in a work by Engwerda (2006). Such games have found many applications in Game Theory and Economics, such as The Game of Defending a Target (Li and Cruz, 2008), Resource Extraction (Yang and Gao, 2016), etc.

4.4. Applications to Dynamic Oligopoly

Among models with dynamic strategic competition, there is a certain class which is categorized by the price competition and involvement of the constumers. Rosenthal (1982) considered a duopology competition under complete customer loyalty from where significant research emerged. With one firm's market share as the natural state variable, Rosenthal classified a Markov-stationary equilibrium where prices remained above marginal costs indefinitely.

In a follow-up work on this, Chen and Rosenthal (1996) considered and showed a closed form mixed-strategy equilibrium. Some part of the literature is also concerned with long-run price competition when customers face costs for switching between producers.

Cabral and Riordan (1994) characterize the long-term consequences of reducing production cost with production experience within a Markov-stationary framework with firms' cumulative sales as the natural state variables. Another source of strategic dynamic competition also arises due to technological progressions.

References

Rabah Amir. Stochastic games in economics and related fields: An overview. In Abraham Neyman and Sylvain Sorin, editors, *Stochastic Games and Applications*, pages 455–470, Dordrecht, 2003. Springer Netherlands. ISBN 978-94-010-0189-2.

Michael H. Bowling and Manuela M. Veloso. An analysis of stochastic game theory for multiagent reinforcement learning. 2000.

Luis M. B. Cabral and Michael H. Riordan. The learning curve, market dominance, and predatory pricing. *Econometrica*, 62(5):1115–1140, 1994. ISSN 00129682, 14680262. URL `http://www.jstor.org/stable/2951509`.

Yongmin Chen and Robert W. Rosenthal. Dynamic duopoly with slowly changing customer loyalties. *International Journal of Industrial Organization*, 14(3):269 – 296, 1996. ISSN 0167-7187. doi: https://doi.org/10.1016/0167-7187(95)00487-4. URL `http://www.sciencedirect.com/science/article/pii/0167718795004874`.

Gilette D. Stochastic games with zero stop probabilities. 1957.

Van der Wal J. Discounted markov games: Successive approximation and stopping times. 1977.

Prajit K. Dutta and Rangarajan Sundaram. Markovian equilibrium in a class of stochastic games: existence theorems for discounted and undiscounted models. *Economic Theory*, 2(2):197–214, Jun 1992. ISSN 1432-0479. doi: 10.1007/BF01211440. URL `https://doi.org/10.1007/BF01211440`.

Prajit K. Dutta and Rangarajan K. Sundaram. How different can strategic models be? *Journal of Economic Theory*, 60(1):42 – 61, 1993a. ISSN 0022-0531. doi: https://doi.org/10.1006/jeth.1993.1034. URL http://www.sciencedirect.com/science/article/pii/S0022053183710343.

Prajit K. Dutta and Rangarajan K. Sundaram. The tragedy of the commons? *Economic Theory*, 3(3): 413–426, Sep 1993b. ISSN 1432-0479. doi: 10.1007/BF01209694. URL https://doi.org/10.1007/BF01209694.

J.C. Engwerda. Linear quadratic games: An overview. Workingpaper, Macroeconomics, 2006. Subsequently published in Advances in Dynamic Games and their Applications (book), 2009 Pagination: 32.

Arie Hordijk and Lodewijk Kallenberg. *Linear programming and Markov games II*, pages 307 – 319. 01 1981.

Junling Hu and Michael P. Wellman. Nash q-learning for general-sum stochastic games. *J. Mach. Learn. Res.*, 4:1039–1069, December 2003. ISSN 1532-4435. URL http://dl.acm.org/citation.cfm?id=945365.964288.

David Levhari and Leonard J. Mirman. The great fish war: An example using a dynamic cournot-nash solution. *The Bell Journal of Economics*, 11(1):322–334, 1980. ISSN 0361915X. URL http://www.jstor.org/stable/3003416.

Dongxu Li and Jose B. Cruz. Game of defending a target: A linear quadratic differential game approach. *IFAC Proceedings Volumes*, 41(2):2643 – 2648, 2008. ISSN 1474-6670. doi: https://doi.org/10.3182/20080706-5-KR-1001.00445. URL http://www.sciencedirect.com/science/article/pii/S1474667016393508. 17th IFAC World Congress.

Eric Maskin and Jean Tirole. We shall define a Markov Perfect Equilibrium (MPE) to be a subgame perfect equilibrium in which all players use markov strategies. *Journal of Economic Theory*, 2001.

J. F. Mertens and A. Neyman. Stochastic games. *International Journal of Game Theory*, 10(2):53–66, Jun 1981. ISSN 1432-1270. doi: 10.1007/BF01769259. URL https://doi.org/10.1007/BF01769259.

Vrieze OJ. Linear programming and undiscounted stochastic games. *OR Spectrum*, 1981.

T. Parthasarathy, S. H. Tijs, and O. J. Vrieze. Stochastic games with state independent transitions and separable rewards. In Gerald Hammer and Diethard Pallaschke, editors, *Selected Topics in Operations Research and Mathematical Economics*, pages 262–271, Berlin, Heidelberg, 1984. Springer Berlin Heidelberg. ISBN 978-3-642-45567-4.

M. A. Pollatschek and B. Avi-Itzhak. Algorithms for stochastic games with geometrical interpretation. *Manage. Sci.*, 15(7):399–415, March 1969. ISSN 0025-1909. doi: 10.1287/mnsc.15.7.399. URL http://dx.doi.org/10.1287/mnsc.15.7.399.

H.L. Prasad and S. Bhatnagar. General-sum stochastic games: Verifiability conditions for nash equilibria. *Automatica*, 48(11):2923 – 2930, 2012. ISSN 0005-1098. doi: https://doi.org/10.1016/j.automatica.2012.06.088. URL http://www.sciencedirect.com/science/article/pii/S0005109812003585.

T. E. S. Raghavan and J. A. Filar. Algorithms for stochastic games — a survey. *Zeitschrift für Operations Research*, 35(6):437–472, Nov 1991. ISSN 1432-5217. doi: 10.1007/BF01415989. URL `https://doi.org/10.1007/BF01415989`.

T. E. S. Raghavan and Zamir Syed. Computing stationary nash equilibria of undiscounted single-controller stochastic games. *Mathematics of Operations Research*, 27(2):384–400, 2002. ISSN 0364765X, 15265471. URL `http://www.jstor.org/stable/3690596`.

Robert W Rosenthal. A dynamic model of duopoly with customer loyalties. *Journal of Economic Theory*, 27(1):69 – 76, 1982. ISSN 0022-0531. doi: https://doi.org/10.1016/0022-0531(82)90015-1. URL `http://www.sciencedirect.com/science/article/pii/0022053182900151`.

L. S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences*, 39(10):1095–1100, 1953. ISSN 0027-8424. doi: 10.1073/pnas.39.10.1095. URL `https://www.pnas.org/content/39/10/1095`.

X. Yang and J. Gao. Linearquadratic uncertain differential game with application to resource extraction problem. *IEEE Transactions on Fuzzy Systems*, 24(4):819–826, Aug 2016. ISSN 1063-6706. doi: 10.1109/TFUZZ.2015.2486809.