

lution of binary stars. MODEST (Modeling Dense Stellar systems)<sup>1</sup> is one of such efforts to develop a simulation code, or a loosely coupled collection of codes, to handle complex systems.

## Bibliography

1. Aarseth SJ (1963) MN 126:223
2. Aarseth SJ (1985) In: Blackbill JU, Cohen BI (eds) Multiple Time Scales. Academic Press, New York, pp 377–418
3. Aarseth SJ (2003) Gravitational N-Body Simulations. Cambridge University Press, Cambridge, pp 430
4. Barnes J, Hut P (1986) Nature 324:446
5. Binney J, Tremaine S (2008) Galactic Dynamics, 2nd edn. Princeton University Press, Princeton
6. Cohn H (1980) ApJ 242:765
7. Cohn H, Hut P, Wise M (1989) ApJ 342:814
8. Goodman J (1984) ApJ 280:298
9. Hachisu I, Nakada Y, Nomoto K, Sugimoto D (1978) Prog Theor Phys 60:393
10. Heggie DC (1984) MN 206:179
11. Henon M (1971) Astrophys Space Sci 13:284
12. Hernquist L (1990) J Comp Phys 87:137
13. Hut P, Makino J, McMillan S (1995) ApJL 443:L93
14. Lynden-Bell D, Eggleton PP (1980) MN 191:483
15. Makino J (1990) J Comput Phys 87:148
16. Makino J (1996) Dynamical Evolution of Star Clusters. In: Hut P, Makino J (eds) Proc IAU Symp 174. Kluwer, Amsterdam, pp 151–160
17. Makino J, Hut P (1988) ApJS 68:833
18. Makino J, Hut P, Kaplan M, Saygin H (2006) New Astron 12:124
19. Pfalzner S, Gibbon P (1996) Many-Body Tree Methods in Physics. Cambridge University Press, Cambridge
20. Preto M, Tremaine S (1999) AJ 118:2532
21. Salmon JK, Warren MS (1994) J Comp Phys 111:136
22. Sanz-Serna JM, Calvo MP (1994) Numerical Hamiltonian Problems. Chapman and Hall, London
23. Skeel RD, Biesiadecki JJ (1994) Ann Numer Math 1:191
24. Skeel RD, Gear CW (1992) Physica D 60:311
25. Spitzer Jr L (1987) Dynamical Evolution of Globular Clusters. Princeton University Press, Princeton
26. Sugimoto D, Bettwieser E (1983) MN 204:19P
27. Yoshikawa K, Fukushige T (2005) PASJ 57:849

## Stochastic Games

EILON SOLAN

The School of Mathematical Sciences,  
Tel Aviv University, Tel Aviv, Israel

## Article Outline

Glossary

Definition of the Subject

Strategies, Evaluations and Equilibria

<sup>1</sup><http://www.manybody.org/modest/>

Zero-Sum Games

Multi-Player Games

Correlated Equilibrium

Imperfect Monitoring

Algorithms

Additional and Future Directions

Acknowledgments

Bibliography

## Glossary

**A stochastic game** A repeated interaction between several participants in which the underlying state of the environment changes stochastically, and it depends on the decisions of the participants.

**A strategy** A rule that dictates how a participant in an interaction makes his decisions as a function of the observed behavior of the other participants and of the evolution of the environment.

**Evaluation of stage payoffs** The way that a participant in a repeated interaction evaluates the stream of stage payoffs that he receives (or stage costs that he pays) along the interaction.

**An equilibrium** A collection of strategies, one for each player, such that each player maximizes (or minimizes, in case of stage costs) his evaluation of stage payoffs given the strategies of the other players.

**A correlated equilibrium** An equilibrium in an extended game in which at the outset of the game each player receives a private signal, and the vector of private signals is chosen according to a known joint probability distribution. In the extended game, a strategy of a player depends, in addition to past play, on the signal he received.

## Definition of the Subject

Stochastic games, first introduced by Shapley [60], model dynamic interactions in which the environment changes in response to the behavior of the players. Formally, a stochastic game is a tuple  $G = (N, S, (\mathcal{A}_i, A_i, u_i)_{i \in N}, q)$  where

- $N$  is a set of players.
- $S$  is a state space. If  $S$  is uncountable, it is supplemented with a  $\sigma$ -algebra of measurable sets.
- For every player  $i \in N$ ,  $\mathcal{A}_i$  is a set of actions for that player, and  $A_i: S \rightarrow \mathcal{A}_i$  is a set-valued (measurable) function that assigns to each state  $s \in S$  the set of actions  $A_i(s)$  that are available to player  $i$  in state  $s$ . If  $\mathcal{A}_i$  is uncountable, it is supplemented with a  $\sigma$ -algebra of measurable sets. Denote  $SA = \{(s, a) : s \in S, a =$

$(a_i)_{i \in N}, a_i \in A_i(s) \forall i \in N\}$ . This is the set of all possible *action profiles*.

- For every player  $i \in N$ ,  $u_i: SA \rightarrow \mathbf{R}$  is a (measurable) *stage payoff function* for player  $i$ .
- $q: SA \rightarrow \Delta(S)$  is a (measurable) *transition function*, where  $\Delta(S)$  is the space of probability distributions over  $S$ .

The game starts at an initial state  $s^1$ , and is played as follows. At each stage  $t \in \mathbf{N}$ , each player  $i \in N$  chooses an action  $a_i^t \in A_i(s^t)$ , receives the stage payoff  $u_i(s^t, a^t)$ , where  $a^t = (a_i^t)_{i \in N}$ , and the game moves to a new state  $s^{t+1}$  that is chosen according to the probability distribution  $q(\cdot | s^t, a^t)$ .

A few comments are in order.

1. A stochastic game lasts infinitely many stages. However, the model also captures finite interactions (of length  $t$ ), by assuming the play moves, at stage  $t$ , to an absorbing state with payoff 0 to all players.
2. In particular, by setting  $t = 1$ , we see that stochastic games are a generalization of matrix games (games in normal form/games in normal form), which are played only once.
3. Stochastic games are also a generalization of repeated games, in which the players play the same matrix game over and over again. Indeed, a repeated game is equivalent to a stochastic game with a single state.
4. Stopping games are also a special case of stochastic games. In these games every player has two actions in all states, *continue* and *quit*. as long as all players choose *continue* the stage payoff is 0; once at least one player chooses *quit* the game moves to an absorbing state.
5. Markov decision problems (see, e.g., [49]) are stochastic games with a single player.
6. The transition function  $q$  governs the evolution of the game. It depends on the actions of all players and on the current state, so that all the players influence the evolution of the game.
7. The payoff function  $u_i$  of player  $i$  depends on the current state as well as on the actions chosen by all players. Thus, a player's payoff depends not only on that player's choice, but also on the behavior of the other players.
8. Though we refer to the functions  $(u_i)_{i \in N}$  as "stage payoffs", with the implicit assumption that each player tries to maximize his payoff, in some applications these functions describe a stage *cost*, and then the implicit assumption is that each player tries to minimize his cost.
9. The action of a player at a given stage affects both his stage payoff and the evolution of the state variable, thereby affecting his future payoffs. These two, sometimes contradicting effects make the optimization problem of the players quite intricate, and the analysis of the game challenging.
10. The players receive a stage payoff at each stage. So far we did not mention how the players evaluate the infinite stream of stage payoffs that they receive, nor did we say what is their information at each stage: Do they observe the current state? Do they observe the actions of the other players? These issues will be discussed later.
11. The actions that are available to the players at each stage, the payoff functions, and the transition function, all depend on the current state, and not on past play (that is, past states that the game visited, and past actions that the players chose). This assumption is without loss of generality. Indeed, suppose that the actions available to the players at each stage, the payoff functions, and the transition function, all depend on past play, as well as on the current state. For every  $t \in \mathbf{N}$  let  $H_t$  be the set of all possible *histories* of length  $t$ , that is, all sequences of the form  $(s^1, a^1, s^2, a^2, \dots, s^t)$ , where  $s^k \in S$  for every  $k = 2, 3, \dots, t$ ,  $a^k = (a_i^k)_{i \in N}$  and  $a_i^k$  is an available action to player  $i$  at stage  $k$ , for every  $k = 1, 2, \dots, t - 1$ . Then the game is equivalent to a game with state space  $H := \bigcup_{t \in \mathbf{N}} H_t$ , in which the state variable captures past play, and the state at stage  $t$  lies in  $H_t$ . In the new game, the sets of available actions, the payoff function, and the transition function, depend on the current state rather than on all past play.

The interested reader is referred to [20,42,72] for further reading on stochastic games. We now provide a few applications.

*Example 1 (Capital Accumulation ([7,18,19,34,45]))* Two (or more) agents jointly own a natural resource or a productive asset; at every period they have to decide the amount of the resource to consume. The amount that is not consumed grows by a known (or an unknown) fraction. Such a situation occurs, e.g., in fishery: Fishermen from various countries fish in the same area, and each country sets a quota for its fishermen. Here the state variable is the current amount of resource, the action set is the amount of resource to be exploited in the current period, and the transition is influenced by the decisions of all the players, as well as possibly by the random growth of the resource.

*Example 2 (Taxation ([14,48]))* A government sets a tax rate at every period. Each citizen decides at every period how much to work, and, from the total amount of money he or she has, how much to consume; the rest is saved for the next period, and grows by a known interest rate. Here the state is the amount of savings each citizen has, the stage payoff of a citizen depends on the amount of money that he consumed, on the amount of free time he has, and on the total amount of tax that the government collected. The stage payoff of the government may be the average stage payoff of the citizens, the amount of tax collected, or a mixture of the two.

*Example 3 (Communication Network [58])* A single-cell system with one receiver and multiple uplink transmitters share a single, slotted, synchronous classical collision channel. Assume that all transmitted packets have the same length, and require one time unit, which is equal to one time slot, for transmission. Whenever a collision occurs, the users attempt to retransmit their packets in subsequent slots to resolve collision for reliable communication.

Here a state lists all relevant data for a given stage: e. g., the number of packets waiting at each transmitter, or the length of time each has been waiting to be transmitted. The players are the transmitters, and the action of each transmitter is which packet to transmit, if any. The stage cost may depend on the number of time slots that the transmitted packet waited, on the number of packets that have not been transmitted at that period, and possibly on additional variables. The transition depends on the actions chosen by the players, but it has a stochastic component, which captures the number of new packets that arrive at the various transmitters during every time slot.

*Example 4 (Queues [1])* Individuals that require service have to choose whether to be served by a private slow service provider, or by a powerful public service provider. This situation arises, e. g., when jobs can be executed on either a slow personal computer or a fast mainframe. Here a state lists the current load of the public and private service providers, and the cost is the time to be served.

The importance of stochastic games stems from the wide range of applications they encompass. Many repeated interactions can be recast as stochastic games; the wide range of theoretical results that have been obtained provide insights that can help in analyzing specific situations and suggesting proper behavior to the participants. In certain classes of games algorithms that have been developed may be used to calculate such behavior.

## Strategies, Evaluations and Equilibria

So far we have not described the information that the players have at each stage. In most of the chapter we assume that the players have complete information of past play; that is, at each stage  $t$ , they know the sequence  $s^1, a^1, s^2, a^2, \dots, s^t$  of states that were visited in the past (including the current state) and the actions that were chosen by all players. This assumption is too strong for most applications, and in the sequel we will mention the consequences of its relaxation.

Since the players observe past play, a *pure strategy* for player  $i$  is a (measurable) function  $\sigma_i$  that assigns to every finite history  $(s^1, a^1, s^2, a^2, \dots, s^t)$  an action  $\sigma_i(s^1, a^1, s^2, a^2, \dots, s^t) \in A_i(s^t)$ , with the interpretation that, at stage  $t$ , if the finite history  $(s^1, a^1, s^2, a^2, \dots, s^t)$  occurred, player  $i$  plays the action  $\sigma_i(s^1, a^1, s^2, a^2, \dots, s^t)$ . If the player does not know the complete history, then a strategy for player  $i$  is a function that assigns to every possible information set, an action that is available to the player when the player has this information. A *mixed strategy* for player  $i$  is a probability distribution over the set of his pure strategies. The space of mixed strategies of player  $i$  is denoted by  $\sigma_i$ .

A simple class of strategies is the class of *stationary* strategies; a strategy  $\sigma_i$  for player  $i$  is *stationary* if  $\sigma_i(s^1, a^1, s^2, a^2, \dots, s^t)$  depends only on the current state  $s^t$ , and not on past play  $s^1, a^1, s^2, a^2, \dots, a^{t-1}$ . A stationary strategy of player  $i$  can be identified with an element  $x = (x_s)_{s \in S} \in \times_{s \in S} \Delta(A_i(s))$ , with the interpretation that player  $i$  plays the mixed action  $x_s$  whenever the current state is  $s$ . Denote by  $X_i = \times_{s \in S} \Delta(A_i(s))$  the space of stationary strategies of player  $i$ .

There are three common ways to evaluate the infinite stream of payoffs that the players receive in a stochastic game: The *finite-horizon evaluation*, in which a player considers the average payoff during the first  $T$  stages, the *discounted evaluation*, in which a player considers the discounted sum of his stage payoffs, and the *limsup evaluation*, in which a player considers the limsup of his long-run average payoffs. We now formally define these evaluations.

Every profile  $\sigma = (\sigma_i)_{i \in N}$  of mixed strategies, together with the initial state, induces a probability distribution  $\mathbf{P}_{s_1, \sigma}$  over the space of infinite plays  $H_\infty := SA^N$ . We denote the corresponding expectation operator by  $\mathbf{E}_{s_1, \sigma}$ .

**Definition 5** Let  $\sigma$  be a profile of mixed strategies. For every finite horizon  $T \in \mathbf{N}$ , the *T-stage payoff* under  $\sigma$  for player  $i$  is

$$\gamma_i^T(s_1, \sigma) := \mathbf{E}_{s_1, \sigma} \left[ \frac{1}{T} \sum_{t=1}^T u_i(s^t, a^t) \right].$$

For every discount factor  $\lambda \in (0, 1]$ , the  $\lambda$ -discounted payoff under  $\sigma$  for player  $i$  is

$$\gamma_i^\lambda(s_1, \sigma) := \mathbf{E}_{s_1, \sigma} \left[ \lambda \sum_{t=1}^{\infty} (1 - \lambda)^{t-1} u_i(s^t, a^t) \right].$$

The limsup payoff under  $\sigma$  for player  $i$  is

$$\gamma_i^\infty(s_1, \sigma) := \mathbf{E}_{s_1, \sigma} \left[ \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T u_i(s^t, a^t) \right].$$

The  $T$ -stage payoff captures the situation in which the interaction lasts exactly  $T$  stages. The  $\lambda$ -discounted evaluation captures the situation in which the game lasts “many” stages, and the player discounts stage payoffs – it is better to receive \$1 today than tomorrow. The limsup payoff also captures the situation in which the game lasts “many” stages, but here the player does not discount his payoffs, and the payoff at each given stage is insignificant as compared to the payoff in all other stages. Equivalently, one could consider the liminf payoff in which the player considers the liminf of the long-run average payoffs.

As usual, an equilibrium is a vector of strategies such that no player can profit by a unilateral deviation. For every player  $i$  and every strategy profile  $\sigma = (\sigma_i)_{i \in N}$  we denote the strategy profile of all other players, except player  $i$ , by  $\sigma_{-i} = (\sigma_j)_{j \neq i}$ .

**Definition 6** Let  $\varepsilon \geq 0$ . A profile of strategies  $\sigma$  is a  $T$ -stage  $\varepsilon$ -equilibrium if

$$\gamma_i^T(s_1, \sigma) \geq \gamma_i^T(s_1, \sigma'_i, \sigma_{-i}) - \varepsilon, \\ \forall s_1 \in S, \forall i \in N, \forall \sigma'_i \in \Sigma_i.$$

It is a  $\lambda$ -discounted  $\varepsilon$ -equilibrium if

$$\gamma_i^\lambda(s_1, \sigma) \geq \gamma_i^\lambda(s_1, \sigma'_i, \sigma_{-i}) - \varepsilon, \\ \forall s_1 \in S, \forall i \in N, \forall \sigma'_i \in \Sigma_i.$$

It is a limsup  $\varepsilon$ -equilibrium if

$$\gamma_i^\infty(s_1, \sigma) \geq \gamma_i^\infty(s_1, \sigma'_i, \sigma_{-i}) - \varepsilon, \\ \forall s_1 \in S, \forall i \in N, \forall \sigma'_i \in \Sigma_i.$$

The payoff that corresponds to an  $\varepsilon$ -equilibrium, that is, either one of the quantities  $\gamma^T(s_1, \sigma)$ ,  $\gamma^\lambda(s_1, \sigma)$  and  $\gamma^\infty(s_1, \sigma)$ , is called an  $\varepsilon$ -equilibrium payoff at the initial state  $s_1$ .

As we will see below, when both state and action spaces are finite, a  $T$ -stage and a  $\lambda$ -discounted 0-equilibrium exist.

However, when the state or action spaces are infinite such a 0-equilibrium may fail to exist, yet  $\varepsilon$ -equilibria may exist for every  $\varepsilon > 0$ .

As the length of the game  $T$  varies, or as the discount factor  $\lambda$  varies, the equilibrium strategy profile varies as well. A strategy profile that is an  $\varepsilon$ -equilibrium for every  $T$  sufficiently large and for every  $\lambda$  sufficiently small is called a uniform  $\varepsilon$ -equilibrium.

**Definition 7** Let  $\varepsilon > 0$ . A strategy profile  $\sigma$  is a uniform  $\varepsilon$ -equilibrium if there are  $T_0 \in \mathbf{N}$  and  $\lambda_0 \in (0, 1)$  such that for every  $T \geq T_0$  the strategy profile  $\sigma$  is a  $T$ -stage  $\varepsilon$ -equilibrium, and for every  $\lambda \in (0, \lambda_0)$  it is a  $\lambda$ -discounted  $\varepsilon$ -equilibrium.

If for every  $\varepsilon > 0$  the game has a ( $T$ -stage,  $\lambda$ -discounted, limsup or uniform)  $\varepsilon$ -equilibrium with corresponding payoff  $g_\varepsilon$ , then any accumulation point of  $(g_\varepsilon)_{\varepsilon > 0}$  as  $\varepsilon$  goes to 0 is a ( $T$ -stage,  $\lambda$ -discounted, limsup or uniform) equilibrium payoff.

## Zero-Sum Games

A two-player stochastic game is zero-sum if  $u_1(s, a) + u_2(s, a) = 0$  for every  $(s, a) \in SA$ . As in matrix games, every two-player zero-sum stochastic game admits at most one equilibrium payoff at every initial state  $s_1$ , which is termed the value of the game at  $s_1$ . Each player's strategy which is part of an  $\varepsilon$ -equilibrium is termed  $\varepsilon$ -optimal. The definition of  $\varepsilon$ -equilibrium implies that an  $\varepsilon$ -optimal strategy guarantees the value up to  $\varepsilon$ ; for example, in the  $T$ -stage evaluation, if  $\sigma_1$  is an  $\varepsilon$ -optimal strategy of player 1, then for every strategy of player 2 we have

$$\gamma_1^T(s_1, \sigma_1, \sigma_2) \geq v^T(s_1) - \varepsilon,$$

where  $v^T(s_1)$  is the  $T$ -stage value at  $s_1$ .

In his seminal work, Shapley [60] presented the model of two-player zero-sum stochastic games with finite state and actions spaces, and proved the following.

**Theorem 8 [60]** For every two-player zero-sum stochastic game, the  $\lambda$ -discounted value at every initial state exists. Moreover, both players have  $\lambda$ -discounted 0-optimal stationary strategies.

*Proof* Let  $\mathcal{V}$  be the space of all functions  $v: S \rightarrow \mathbf{R}$ . For every  $v \in \mathcal{V}$  define a zero-sum matrix game  $G_s^\lambda(v)$  as follows:

- The action spaces of the two players are  $A_1(s)$  and  $A_2(s)$  respectively.
- The payoff function (that player 2 pays player 1) is

$$\lambda u_1(s, a) + (1 - \lambda) \sum_{s' \in S} q(s' | s, a) v(s').$$

The game  $G_s^\lambda(v)$  captures the situation in which, after the first stage, the game terminates with a terminal payoff  $v(s')$ , where  $s'$  is the state reached after stage 1. Define an operator  $\varphi: \mathcal{V} \rightarrow \mathcal{V}$  as follows:

$$\varphi_s(v) = \text{val}(G_s^\lambda(v)),$$

where  $\text{val}(G_s^\lambda(v))$  is the value of the matrix game  $G_s^\lambda(v)$ . Since the value operator is non-expansive, it follows that the operator  $\varphi$  is contracting:  $\|\varphi(v) - \varphi(w)\|_\infty \leq (1 - \lambda)\|v - w\|_\infty$ , so that this operator has a unique fixed point  $\widehat{v}^\lambda$ . One can show that the fixed point is the value of the stochastic game, and every strategy  $\sigma_i$  of player  $i$  in which he plays, after each finite history  $(s^1, a^1, s^2, a^2, \dots, s^t)$ , an optimal mixed action in the matrix game  $G_{s^t}^\lambda(\widehat{v}^\lambda)$ , is a  $\lambda$ -discounted 0-optimal strategy in the stochastic game.  $\square$

*Example 9* Consider the following two-player zero-sum game with three states,  $s_0, s_1$  and  $s_2$ ; each entry of the matrix indicates the payoff that player 2 (the column player) pays player 1 (the row player, the payoff is in the middle), and the transitions (which are deterministic, and are denoted at the top-right corner).

	L	R	
T	0 $s^2$	1 $s^1$	
B	1 $s^1$	0 $s^0$	
	State	$s_2$	

	L	
T	1 $s^1$	
	State $s_1$	

	L	
T	0 $s^0$	
	State $s_0$	

The states  $s_0$  and  $s_1$  are absorbing: Once the play reaches one of these states it never leaves it. State  $s_2$  is non-absorbing. Stochastic games with a single non-absorbing state are called *absorbing games*. For every  $v = (v_0, v_1, v_2) \in \mathcal{V} = \mathbf{R}^3$  the game  $G_{s_2}^\lambda(v)$  is the following matrix game:

	L	R	
T	$(1 - \lambda)v_2$	$\lambda + (1 - \lambda)v_1$	
B	$\lambda + (1 - \lambda)v_1$	$(1 - \lambda)v_0$	
	The game $G_{s_2}^\lambda$		

	L	
T	$\lambda + (1 - \lambda)v_1$	
	The game $G_{s_1}^\lambda$	

	L	
T	$(1 - \lambda)v_0$	
	The game $G_{s_0}^\lambda$	

The unique fixed point of the operator  $\text{val}(G^\lambda)$  must satisfy

- $\widehat{v}_0 = \text{val}(G_{s_0}^\lambda(\widehat{v}))$ , so that  $\widehat{v}_{s_0}^\lambda = \widehat{v}_0 = 0$ ;
- $\widehat{v}_1 = \text{val}(G_{s_1}^\lambda(\widehat{v}))$ , so that  $\widehat{v}_{s_1}^\lambda = \widehat{v}_1 = 1$ ;

- $\widehat{v}_2 = \text{val}(G_{s_2}^\lambda(\widehat{v}))$ . By Theorem 8 both players have a stationary  $\lambda$ -discounted 0-optimal strategy. Denote by  $x$  (resp.  $y$ ) a mixed action for player 1 (resp. player 2) that is part of a  $\lambda$ -discounted 0-optimal strategy at the state  $s_2$ . Since we know that in the fixed point  $\widehat{v}_0 = 0$  and  $\widehat{v}_1 = 1, \widehat{v}_2$  must be the unique solution of

$$v_2 = y(1 - \lambda)v_2 + (1 - y) = y,$$

so that  $\widehat{v}_{s_2}^\lambda = \widehat{v}_2 = (1 - \sqrt{\lambda})/(1 - \lambda)$ . The 0-optimal strategy of player 2 at state  $s_2$  is  $y = \widehat{v}_2 = (1 - \sqrt{\lambda})/(1 - \lambda)$ , and the 0-optimal strategy of player 1,  $x = \widehat{v}_2 = (1 - \sqrt{\lambda})/(1 - \lambda)$ , can be found by finding his 0-optimal strategy in  $G_{s_2}^\lambda(\widehat{v})$ .

Bewley and Kohlberg [11] proved that when the state and action spaces are finite, the function  $\lambda \mapsto v_s^\lambda$ , that assigns to every state  $s$  and every discount factor  $\lambda$  the  $\lambda$ -discounted value at the initial state  $s$ , is a *Puiseux function*, that is, it has a representation  $v_s^\lambda = \sum_{k=K}^\infty a_k \lambda^{k/M}$  that is valid for every  $\lambda \in (0, \lambda_0)$  for some  $\lambda_0 > 0$ , where  $M$  is a natural number,  $K$  is a non-negative integer, and  $(a_k)_{k=K}^\infty$  are real numbers. In particular, the function  $\lambda \mapsto v_s^\lambda$  is monotone in a neighborhood of 0, and its limit as  $\lambda$  goes to 0 exists. This result turned out to be crucial in subsequent study on games with finitely many states and actions.

Shapley's work has been extended to general state and action spaces; for a recent survey see [46]. The tools developed in [46], together with a dynamic programming argument, prove that under proper conditions on the payoff function and on the transitions the two-player zero-sum stochastic game has a  $T$ -stage value.

Maitra and Sudderth [35] proved that the limsup value exists in a very general setup. Their proof follows closely that of Martin [36] for the determinacy of Blackwell games.

The study of the uniform value emanated from an example, called the "Big Match", due to Gillette [28], that was solved by Blackwell and Ferguson [13].

*Example 10* Consider the following stochastic game with two absorbing states and one non-absorbing state.

	L	R	
T	0 $s^2$	1 $s^2$	
B	1 $s^1$	0 $s^0$	
	State $s_2$		

	L	
T	1 $s^1$	
	State $s_1$	

	L	
T	0 $s^0$	
	State $s_0$	

Suppose the initial state is  $s_2$ . As long as player 1 plays  $T$  the play remains at  $s_2$ ; once he plays  $B$  the play moves to either  $s_0$  or  $s_1$ , and is effectively terminated. By finding the fixed point of the operator  $\varphi$  one can show

that the discounted value at the initial state  $s_2$  is  $\frac{1}{2}$ , and a  $\lambda$ -discounted stationary 0-optimal strategy for player 2 is<sup>1</sup>  $[\frac{1}{2}(L), \frac{1}{2}(R)]$ . Indeed, if player 1 plays  $T$  then the expected stage payoff is  $\frac{1}{2}$  and play remains at  $s_2$ , while if player 1 plays  $B$  then the game moves to an absorbing state, and the expected stage payoff from that stage onwards is  $\frac{1}{2}$ . In particular, this strategy guarantees  $\frac{1}{2}$  for player 2 both in the limsup evaluation and uniformly. A  $\lambda$ -discounted 0-optimal strategy for player 1 is  $[\frac{1}{1+\lambda}(T), \frac{\lambda}{1+\lambda}(B)]$ .

What can player 1 guarantee in the limsup evaluation and uniformly? If player 1 plays the stationary strategy  $[x(T), (1-x)(B)]$  that plays at each stage the action  $T$  with probability  $x$  and the action  $B$  with probability  $1-x$ , then player 2 has a reply that ensures that the limsup payoff is 0: If  $x = 1$  and player 2 always plays  $L$ , the payoff is 0 at each stage; if  $x < 1$  and player 2 always plays  $R$ , the payoff is 1 until the play moves to  $s_0$ , and then it is 0 forever. Since player 1 plays the action  $B$  with probability  $1-x > 0$  at each stage, the distribution of the stage in which play moves to  $s_0$  is geometric. Therefore, the limsup payoff is 0, and if  $\lambda$  is sufficiently small, the discounted payoff is close to 0.

One can verify that if player 1 uses a bounded-recall strategy, that is, a strategy that uses only the last  $k$  actions that were played, player 2 has a reply that guarantees that the limsup payoff is 0, and the discounted payoff is close to 0, provided  $\lambda$  is close to 0. Thus, in the limsup payoff and uniformly finite memory cannot guarantee more than 0 in this game (see also [27]).

Intuitively, player 1 would like to condition the probability of playing  $T$  on the past behavior of player 2: If in the past player 2 played the action  $L$  more often than the action  $R$ , he would have liked to play  $T$  with higher probability; if in the past player 2 played the action  $R$  more often than the action  $L$ , he would have liked to play  $B$  with higher probability. Blackwell and Ferguson [13] constructed a family of good strategies  $\{\sigma_1^M, M \in \mathbb{N}\}$  for player 1. The parameter  $M$  determines the amount that the strategy guarantees: The strategy  $\sigma_1^M$  guarantees a limsup payoff and a discounted payoff of  $\frac{M}{2M+1}$ , provided the discount factor is sufficiently low. In other words, player 1 cannot guarantee  $\frac{1}{2}$ , but he may guarantee an amount as close to  $\frac{1}{2}$  as he wishes by choosing  $M$  to be sufficiently large. The strategy  $\sigma_1^M$  is defined as follows: At stage  $t$ , play  $B$  with probability  $\frac{1}{(M+l_t-r_t)^2}$ , where  $l_t$  is the number of stages up to stage  $t$  in which player 2 played  $L$ , and  $r_t$  is the number of stages up to stage  $t$  in which player 2 played  $R$ .

Since  $r_t + l_t = t - 1$  one has  $r_t - l_t = 2r_t - (t - 1)$ . The quantity  $r_t$  is the total payoff that player 1 received in the first  $t - 1$  stages if player 1 played  $T$  in those stages (and the game was not absorbed). Thus, this total payoff is a linear function of the difference  $r_t - l_t$ . When presented this way, the strategy  $\sigma_1^M$  depends on that total payoff. Observe that as  $r_t$  increases,  $r_t - l_t$  increases as well, and the probability to play  $B$  decreases.

Mertens and Neyman [38] generalized the idea presented at the end of Example 10 to stochastic games with finite state and action spaces.<sup>2</sup>

**Theorem 11** *If the state and action spaces of a two-player zero-sum stochastic game are finite, the game has a uniform value  $v_s^0$  at every initial state  $s \in S$ . Moreover,  $v_s^0 = \lim_{\lambda \rightarrow 0} v_s^\lambda = \lim_{T \rightarrow \infty} v_s^T$ .*

In their proof, Mertens and Neyman describe a uniform  $\varepsilon$ -optimal strategy. In this strategy the player keeps a parameter,  $\lambda_t$ , which is a fictitious discount factor to use at stage  $t$ . This parameter changes at each stage as a function of the stage payoff; if the stage payoff at stage  $t$  is high then  $\lambda_{t+1} < \lambda_t$ , whereas if the stage payoff at stage  $t$  is low then  $\lambda_{t+1} > \lambda_t$ . The intuition is as follows. As mentioned before, in stochastic games there are two forces that influence the player's behavior: He tries to get high stage payoffs, while keeping future prospects high (by playing in such a way that the next stage that is reached is favorable). When considering the  $\lambda$ -discounted payoff there is a clear comparison between the importance of the two forces: The weight of the stage payoff is  $\lambda$  and the weight of future prospects is  $1 - \lambda$ ; the lower the discount factor, the more weight is given to the future. When considering the uniform value (or the uniform equilibrium) the weight of the stage payoff is 0. However, if the player never attempts to receive a high stage payoff, the overall payoff in the game will not be high. Therefore, the player has a fictitious discount factor; if past payoffs are low and they do not meet the expectation, player 1 increases the weight of the stage payoff by increasing the fictitious discount factor; if past payoffs are high, player 1 increases the weight of the future by lowering this fictitious discount factor.

## Multi-Player Games

Takahashi [75] and Fink [21] extended Shapley's [60] result to discounted equilibria in non-zero-sum games.

<sup>1</sup>That is, at each stage player 2 plays  $L$  with probability  $\frac{1}{2}$  and  $R$  with probability  $\frac{1}{2}$ .

<sup>2</sup>Mertens and Neyman's [38] result actually holds in every stochastic game that satisfies a proper condition, which is always satisfied when the state and action spaces are finite.

**Theorem 12** Every stochastic game with finite state and action spaces has a  $\lambda$ -discounted equilibrium in stationary strategies.

*Proof* The proof utilizes Kakutani's fixed point theorem [31]. Let  $M = \max_{i,s,a} |u_i(s, a)|$  be a bound on the absolute values of the payoffs. Set  $X = \times_{i \in N, s \in S} (\Delta(A_i(s)) \times [-M, M])$ . A point  $x = (x_{i,s}^A, x_{i,s}^V)_{i \in N, s \in S} \in X$  is a collection of one mixed action and one payoff to each player at every state. For every  $v = (v_i)_{i \in N} \in [-M, M]^{N \times S}$  and every  $s \in S$  define a matrix game  $G_s^\lambda(v)$  as follows:

- The action spaces of each player  $i$  is  $A_i(s)$ ;
- The payoff to player  $i$  is

$$\lambda u_i(s, a) + (1 - \lambda) \sum_{s' \in S} q(s' | s, a) v_i(s').$$

We define a set-valued function  $\varphi: X \rightarrow X$  as follows.

- For every  $i \in N$  and every  $s \in S$ ,  $\varphi_{i,s}^A$  is the set of all best responses of player  $i$  to the strategy vector  $x_{-i,s} := (x_{j,s})_{j \neq i}$  in the game  $G_s^\lambda(v)$ . That is,

$$\varphi_{i,s}^A(x, v) := \left\{ \operatorname{argmax}_{y_{i,s} \in \Delta(A_i(s))} \lambda r_i(s, y_{i,s}, x_{-i,s}) + (1 - \lambda) \sum_{s' \in S} q(s' | s, y_{i,s}, x_{-i,s}) v_{i,s'} \right\}.$$

- For every  $i \in N$  and every  $s \in S$ ,  $\varphi_{i,s}^V(x, v)$  is the maximal payoff for player  $i$  in the game  $G_s^\lambda(v)$ , when the other players play  $x_{-i}$ :

$$\varphi_{i,s}^V(x, v) := \max_{y_{i,s} \in \Delta(A_i(s))} \left( \lambda r_i(s, y_{i,s}, x_{-i,s}) + (1 - \lambda) \times \sum_{s' \in S} q(s' | s, y_{i,s}, x_{-i,s}) v_{i,s'} \right).$$

The set-valued function  $\varphi$  has convex and non-empty values and its graph is closed, so that by Kakutani's fixed point theorem it has a fixed point. It turns out that every fixed point of  $\varphi$  defines a  $\lambda$ -discounted equilibrium in stationary strategies.  $\square$

This result has been extended to general state and action spaces by various authors. These results assume a strong continuity on either the payoff function or the transition function, see, e. g., [39,44,62].

As in the case of zero-sum games, a dynamic programming argument shows that under a strong continuity assumption on the payoff function or on the transitions a  $T$ -stage equilibrium exists.

Regarding the existence of the limsup equilibrium and uniform equilibrium little is known. The most significant result in this direction is Vieille [77,78], who proved that every two-player stochastic game with finite state and action spaces has a uniform  $\varepsilon$ -equilibrium for every  $\varepsilon > 0$ . This result has been proved for other classes of stochastic games, see, e. g., [6,24,25,61,63,65,76]. Several influential works in this area are [22,32,71,79]. Most of the papers mentioned above rely on the vanishing discount factor approach, which constructs a uniform  $\varepsilon$ -equilibrium by studying a sequence of  $\lambda$ -discounted equilibria as the discount factor goes to 0.

For games with general state and action spaces, a limsup equilibrium exists under an ergodicity assumption on the transitions, see e. g. Nowak [44], Remark 4 and Jaskiewicz and Nowak [30].

A game has perfect information if there are no simultaneous moves, and both players observe past play. Existence of equilibrium in this case was proven by Mertens [37] in a very general setup.

### Correlated Equilibrium

The notion of correlated equilibrium was introduced by Aumann [8,9], see also Forges [▶ Correlated Equilibria and Communication in Games](#). A *correlated equilibrium* is an equilibrium of an extended game, in which each player receives at the outset of the game a private signal such that the vector of signals is chosen according to a known joint probability distribution. In repeated interactions, such as in stochastic games, there are two natural notions of correlated equilibria: (a) each player receives one signal at the outset of the game (*normal-form correlated equilibrium*); (b) each player receives a signal at each stage (*extensive-form correlated equilibrium*). It follows from Forges [26] that when the state and action sets are finite, the set of all correlated  $T$ -stage equilibrium payoffs (either normal-form or extensive-form) is a polytope.

Nowak and Raghavan [47] proved the existence of an extensive-form correlated discounted equilibrium under weak conditions on the state and action spaces. In their construction, the strategies of the players are stationary, and so is the distribution according to which the signals are chosen after each history; both depend only on the current state, rather than on the whole past play. Roughly, their approach is to apply Kakutani's fixed point theorem to the set-valued function that assigns to each game  $G_s^\lambda(v)$  the set of all correlated equilibrium payoffs in this game, which is convex and compact.

Solan and Vieille [66] proved the existence of an extensive-form correlated uniform equilibrium payoff when the

state and action spaces are finite. Their approach is to let each player play his uniform optimal strategy in a zero-sum game in which all other players try to minimize his payoff. Existence of a normal-form correlated equilibrium was proved for the class of absorbing games [68].

Solan [64] characterized the set of extensive-form correlated equilibrium payoffs for general state and action spaces and a general evaluation on the stage payoffs, and provided a sufficient condition that ensures that the set of normal-form correlated equilibrium payoffs coincides with the set of extensive-form correlated equilibrium payoffs.

### Imperfect Monitoring

So far it has been assumed that at each stage the players know the past play. There are cases in which this assumption is too strong; in some cases players do not know the complete description of the current state (Examples 3 and 4), and in others players do not fully observe the actions of all other players (Examples 2, 3 and 4). For a most general description of stochastic games see Mertens, see Chapter IV in Sorin and Zamir [40] and Coulomb [17].

In the study of the discounted equilibrium, the  $T$ -stage equilibrium or the limsup equilibrium, one may consider the game as a one-shot game: The players simultaneously choose strategies, and the payoff is either the discounted payoff, the  $T$ -stage payoff or the limsup payoff. If the strategy spaces of the players are compact (e. g., if the state and action spaces are finite), and if the payoff is upper-semicontinuous in each player's strategy, keeping the strategies of the other players fixed, then an equilibrium exists. This approach can be used successfully for the discounted equilibrium or the  $T$ -stage equilibrium under weak conditions (see, e. g., [4]), and may be used for the limsup equilibrium under a proper ergodicity condition.

Whenever there exists an equilibrium in stationary strategies (e. g., a discounted equilibrium in games with finitely many states and actions) the only information that players need in order to follow the equilibrium strategies is the current state. In particular, they need not observe past actions of the other players. As we now show, in the "Big Match" (Example 10) the limsup value and the uniform value may fail to exist when each player does not observe the past actions of the other player.

*Example 13 (Example 10: Continued.)* Assume that no player observes the actions of the other player, and assume that the initial state is  $s_2$ . Player 2 can still guarantee  $\frac{1}{2}$  in the limsup evaluation by playing the stationary strategy  $[\frac{1}{2}(L), \frac{1}{2}(R)]$ . One can show that for every strategy of player 2, player 1 has a reply such that the limsup payoff is

at least  $\frac{1}{2}$ . In other words,  $\inf_{\sigma_2} \sup_{\sigma_1} \gamma^\infty(s_2, \sigma_1, \sigma_2) = \frac{1}{2}$ . We now argue that  $\sup_{\sigma_1} \inf_{\sigma_2} \gamma^\infty(s_2, \sigma_1, \sigma_2) = 0$ . Indeed, fix a strategy  $\sigma_1$  for player 1, and  $\varepsilon > 0$ . Let  $\theta$  be sufficiently large such that the probability that under  $\sigma_1$  player 1 plays  $B$  for the first time after stage  $t$  is at most  $\varepsilon$ . Observe that as  $t$  increases, the probability that player 1 plays  $B$  for the first time after stage  $t$  decreases to 0, so that such a  $\theta$  exists. Consider the following strategy  $\sigma_2$  of player 2: Play  $R$  up to stage  $\theta$ , and play  $L$  from stage  $t + 1$  and on. By the definition of  $\theta$ , either player 1 plays  $B$  before or at stage  $\theta$ , and then the game moves to  $s_0$ , and the payoff is 0 at each stage thereafter, or player 1 plays  $T$  at each stage, and then the stage payoff after stage  $\theta$  is 0, or, with probability less than  $\varepsilon$ , player 1 plays  $B$  for the first time after stage  $\theta$ , the play moves to  $s_1$ , and the payoff is 1 thereafter. Thus, the limsup payoff is at most  $\varepsilon$ . A similar analysis shows that

$$\inf_{\sigma_2} \sup_{\sigma_1} \gamma^\lambda(s_2, \sigma_1, \sigma_2) = \frac{1}{2},$$

$$\sup_{\sigma_1} \inf_{\sigma_2} \lim_{\lambda \rightarrow 0} \gamma^\lambda(s_2, \sigma_1, \sigma_2) = 0,$$

so that the uniform value does not exist as well.

This example shows that in general the limsup value and the uniform value need not exist when the players do not observe past play. Though in general the value (and therefore also an equilibrium) need not exist, in many classes of stochastic games the value and an equilibrium do exist, even in the presence of imperfect monitoring.

Rosenberg et al. [55] and Renault [54] showed that the uniform value exists in the one player setup (Markov Decision Problem), in which the player receives partial information regarding the current state. Thus, a single decision maker who faces a dynamic situation and does not fully observe the state of the environment can play in such a way that guarantees high payoff, provided the interaction is sufficiently long or the discount factor is sufficiently low.

Altman et al. [5,6] and Flesch et al. [24] studied stochastic games in which each player has a "private" state, which only he can observe, and the state of the world is composed of the vector of private states. Altman et al. [5,6] studied the situation in which players do not observe the actions of the other players, and Flesch et al. [24] studied the situation in which players do observe each others payoffs. Such games arise naturally in wireless communication (see [5]); take for example several mobiles who periodically send information to a base station. The private state of a mobile may depend, e. g., on its exact physical environment, and it determines the power attenuation between the mobile and the base station. The throughput (the amount of bits per second) that a mobile can send to the base station depends on the power attenuations of



all the mobiles. Finally, the stage payoff is the stage power consumption.

Rosenberg et al. [57] studied the extreme case of two player zero-sum games in which the players observe neither the current state nor the action of the other player, and proved that the uniform value does exist in two classes of games, which capture the behavior of certain communication protocols. Classes of games in which the actions are observed but the state is not observed were studied, e. g., by Sorin [69,70], Sorin and Zamir [74], Krausz and Rieder [33], Flesch et al. [23], Rosenberg et al. [56], Renault [52,53]. For additional results, see [16,73].

### Algorithms

There are two kinds of algorithms: Those that terminate in a finite number of steps, and those that iterate and approximate solutions. Both kinds of algorithms were devised to calculate the value and optimal strategies (or equilibria) in stochastic games.

It is well known that the value of a two-player zero-sum matrix game and optimal strategies for the two players can be calculated efficiently using a linear program. Equilibria in two-player non-zero-sum games can be calculated by the Lemke–Howson algorithm, which is usually efficient, however, its worst running time is exponential in the number of pure strategies of the players [59]. Unfortunately, to date there are no efficient algorithms to calculate either the value in zero-sum stochastic games, or equilibria in non-zero-sum games. Moreover, in Example 9 the discounted value may be irrational for rational discount factors, even though the data of the game (payoffs and transitions) are rational, so it is not clear whether linear programming methods can be used to calculate the value of a stochastic game. Nevertheless, linear programming methods were used to calculate the discounted and uniform value of several classes of stochastic games, see [20,50]. Other methods that were used to calculate the value or equilibria in discounted stochastic games include fictitious play [80], value iterates, policy improvement, and general methods to find the maximum of a function (see [20,51]), a homotopy method [29], and algorithms to solve sentences in formal logic [15,67].

### Additional and Future Directions

The research on stochastic games extends to additional directions than those mentioned in earlier sections. We mention a few here. Approximation of games with infinite state and action spaces by finite games was discussed by Whitt [81], and further developed by Nowak [43]. Stochastic games in continuous time have also been studied, as

well as hybrid models that include both discrete and continuous aspects, see, e. g., [2,10].

Among the many directions of future research in this area, we will mention here but a few. One challenging question is the existence of a uniform equilibrium and a limsup equilibrium in multi-player stochastic games with finite state and action spaces. Another is the development of efficient algorithms that calculate the value of two-player zero-sum games. A third direction concerns the identification of applications that can be recast in the framework of stochastic games, and that can be successfully analyzed using the theoretical tools that the literature developed. Another problem that is of interest is the characterization of approachable and excludable sets in stochastic games with vector payoffs (see [12] for the presentation of matrix games with vector payoffs, and [41] for partial results regarding this problem).

### Acknowledgments

I thank Eitan Altman, János Flesch, Yuval Heller, Jean-Jacques Herings, Ayala Mashiach-Yakovi, Andrzej Nowak, Ronald Peeters, T.E.S. Raghavan, Jérôme Renault, Nahum Shimkin, Robert Simon, Sylvain Sorin, William Sudderth, and Frank Thuijmsman, for their comments on an earlier version of the entry.

### Bibliography

#### Primary Literature

1. Altman E (2005) Applications of dynamic games in queues. *Adv Dyn Games* 7:309–342
2. Altman E, Gaitsgory VA (1995) A hybrid (differential-stochastic) zero-sum game with fast stochastic part. *Ann Int Soc Dyn Games* 3:47–59
3. Altman E, Solan E (2007) Games with constraints with networking applications. Preprint
4. Altman E, Avrachenkov K, Marquez R, Miller G (2005) Zero-sum constrained stochastic games with independent state processes. *Math Methods Oper Res* 62:375–386
5. Altman E, Avrachenkov K, Bonneau N, Debbah M, El-Azouzi R, Sadoc Menasche D (2008) Constrained cost-coupled stochastic games with independent state processes. *Oper Res Lett* 36:160–164
6. Amir R (1996) Continuous stochastic games of capital accumulation with convex transitions. *Games Econ Behav* 15:111–131
7. Aumann RJ (1974) Subjectivity and correlation in randomized strategies. *J Math Econ* 1:67–96
8. Aumann RJ (1987) Correlated equilibrium as an expression of bayesian rationality. *Econometrica* 55:1–18
9. Başar T, Olsder GJ (1995) *Dynamic noncooperative game theory*. Academic Press, New York
10. Bewley T, Kohlberg E (1976) The asymptotic theory of stochastic games. *Math Oper Res* 1:197–208

12. Blackwell D (1956) An analog of the minimax theorem for vector payoffs. *Pac J Math* 6:1–8
13. Blackwell D, Ferguson TS (1968) The big match. *Ann Math Stat* 39:159–163
14. Chari V, Kehoe P (1990) Sustainable plans. *J Political Econ* 98:783–802
15. Chatterjee K, Majumdar R, Henzinger TA (2008) Stochastic limit-average games are in EXPTIME. *Int J Game Theory* 37:219–234
16. Coulomb JM (2003) Absorbing games with a signalling structure. In: Neyman A, Sorin S (eds) *Stochastic games and applications*. NATO Science Series. Kluwer, Dordrecht, pp 335–355
17. Coulomb JM (2003) Games with a recursive structure. In: Neyman A, Sorin S (eds) *Stochastic games and applications*. NATO Science Series. Kluwer, Dordrecht, pp 427–442
18. Dutta P, Sundaram RK (1992) Markovian equilibrium in a class of stochastic games: Existence theorems for discounted and undiscounted models. *Econ Theory* 2:197–214
19. Dutta P, Sundaram RK (1993) The tragedy of the commons? *Econ Theory* 3:413–426
20. Filar JA, Vrieze K (1996) *Competitive Markov decision processes*. Springer
21. Fink AM (1964) Equilibrium in a stochastic  $n$ -person game. *J Sci Hiroshima Univ* 28:89–93
22. Flesch J, Thuijsman F, Vrieze K (1997) Cyclic Markov equilibria in stochastic games. *Int J Game Th* 26:303–314
23. Flesch J, Thuijsman F, Vrieze OJ (2003) Stochastic games with non-observable actions. *Math Meth Oper Res* 58:459–475
24. Flesch J, Schoenmakers G, Vrieze K (2008) Stochastic games on a product state space. *Math Oper Res* 33:403–420
25. Flesch J, Thuijsman F, Vrieze OJ (2007) Stochastic games with additive transitions. *Europ J Oper Res* 179:483–497
26. Forges F (1990) Universal mechanisms. *Econometrica* 58:1341–1364
27. Fortnow L, Kimmel P (1998) Beating a finite automaton in the big match. In: *Proceedings of the 7th conference on theoretical aspects of rationality and knowledge*. Morgan Kaufmann, San Francisco, pp 225–234
28. Gillette D (1957) *Stochastic games with zero stop probabilities, contributions to the theory of games, vol 3*. Princeton University Press, Princeton
29. Herings JJP, Peeters RJAP (2004) Stationary equilibria in stochastic games: Structure, selection, and computation. *J Econ Theory* 118:32–60
30. Jaskiewicz A, Nowak AS (2006) Zero-sum ergodic stochastic games with Feller transition probabilities. *SIAM J Control Optim* 45:773–789
31. Kakutani S (1941) A generalization of Brouwer's fixed point theorem. *Duke Math J* 8:457–459
32. Kohlberg E (1974) Repeated games with absorbing states. *Ann Stat* 2:724–738
33. Krausz A, Rieder U (1997) Markov games with incomplete information. *Math Meth Oper Res* 46:263–279
34. Levhari D, Mirman L (1980) The great fish war: An example using a dynamic Cournot–Nash solution. *Bell J Econ* 11(1):322–334
35. Maitra A, Sudderth W (1998) Finitely additive stochastic games with Borel measurable payoffs. *Int J Game Theory* 27:257–267
36. Martin DA (1998) The determinacy of Blackwell games. *J Symb Logic* 63:1565–1581
37. Mertens JF (1987) Repeated games. In: *Proceedings of the international congress of mathematicians, American Mathematical Society, Berkeley, California*, pp 1528–1577
38. Mertens JF, Neyman A (1981) Stochastic games. *Int J Game Th* 10:53–66
39. Mertens JF, Parthasarathy T (1987) Equilibria for discounted stochastic games, CORE Discussion Paper No. 8750. (Also published in *Stochastic Games and Applications*, Neyman A, Sorin S (eds), NATO Science Series, Kluwer, 131–172)
40. Mertens JF, Sorin S, Zamir S (1994) Repeated games, CORE Discussion Paper 9420-9422
41. Milman E (2006) Approachable sets of vector payoffs in stochastic games. *Games Econ Behav* 56:135–147
42. Neyman A, Sorin S (2003) *Stochastic games and applications*. NATO Science Series. Kluwer
43. Nowak AS (1985) Existence of equilibrium stationary strategies in discounted noncooperative stochastic games with uncountable state space. *J Optim Theory Appl* 45:591–620
44. Nowak AS (2003)  $N$ -person stochastic games: Extensions of the finite state space case and correlation. In: Neyman A, Sorin S (eds) *Stochastic games and applications*. NATO Science Series. Kluwer, Dordrecht, pp 93–106
45. Nowak AS (2003) On a new class of nonzero-sum discounted stochastic games having stationary Nash equilibrium points. *Int J Game Theory* 32:121–132
46. Nowak AS (2003) Zero-sum stochastic games with Borel state spaces. In: Neyman A, Sorin S (eds) *Stochastic games and applications*. NATO Science Series. Kluwer, Dordrecht, pp 77–91
47. Nowak AS, Raghavan TES (1991) Existence of stationary correlated equilibria with symmetric information for discounted stochastic games. *Math Oper Res* 17:519–526
48. Phelan C, Stacchetti E (2001) Sequential equilibria in a Ramsey tax model. *Econometrica* 69:1491–1518
49. Puterman ML (1994) *Markov decision processes: Discrete stochastic dynamic programming*. Wiley, Hoboken
50. Raghavan TES, Syed Z (2002) Computing stationary Nash equilibria of undiscounted single-controller stochastic games. *Math Oper Res* 27:384–400
51. Raghavan TES, Syed Z (2003) A policy improvement type algorithm for solving zero-sum two-person stochastic games of perfect information. *Math Program Ser A* 95:513–532
52. Renault J (2006) The value of Markov chain games with lack of information on one side. *Math Oper Res* 31:490–512
53. Renault J (2007) The value of repeated games with an informed controller. Preprint
54. Renault J (2007) Uniform value in dynamic programming. Preprint
55. Rosenberg D, Solan E, Vieille N (2002) Blackwell optimality in Markov decision processes with partial observation. *Ann Statistics* 30:1178–1193
56. Rosenberg D, Solan E, Vieille N (2004) Stochastic games with a single controller and incomplete information. *SIAM J Control Optim* 43:86–110
57. Rosenberg D, Solan E, Vieille N (2006) Protocol with no acknowledgement. *Oper Res*, forthcoming
58. Sagduyu YE, Ephremides A (2003) Power control and rate adaptation as stochastic games for random access. *Proc 42nd IEEE Conf Decis Control* 4:4202–4207
59. Savani R, von Stengel B (2004) Exponentially many steps for finding a Nash equilibrium in a bimatrix game. *Proc 45th Ann IEEE Symp Found Comput Sci* 2004:258–267

60. Shapley LS (1953) Stochastic games. *Proc Nat Acad Sci USA* 39:1095–1100
61. Simon RS (2003) The structure of non-zero-sum stochastic games. *Adv Appl Math* 38:1–26
62. Solan E (1998) Discounted stochastic games. *Math Oper Res* 23:1010–1021
63. Solan E (1999) Three-person absorbing games. *Math Oper Res* 24:669–698
64. Solan E (2001) Characterization of correlated equilibria in stochastic games. *Int J Game Theory* 30:259–277
65. Solan E, Vieille N (2001) Quitting games. *Math Oper Res* 26:265–285
66. Solan E, Vieille N (2002) Correlated equilibrium in stochastic games. *Games Econ Behav* 38:362–399
67. Solan E, Vieille N (2007) Calculating uniform optimal strategies and equilibria in two-player stochastic games. Preprint
68. Solan E, Vohra R (2002) Correlated equilibrium payoffs and public signalling in absorbing games. *Int J Game Theory* 31:91–122
69. Sorin S (1984) Big match with lack of information on one side (part 1). *Int J Game Theory* 13:201–255
70. Sorin S (1985) Big match with lack of information on one side (part 2). *Int J Game Theory* 14:173–204
71. Sorin S (1986) Asymptotic properties of a non-zero-sum stochastic games. *Int J Game Theory* 15:101–107
72. Sorin S (2002) A first course on zero-sum repeated games. *Mathématiques et Applications*, vol 37. Springer
73. Sorin S (2003) Stochastic games with incomplete information. In: Neyman A, Sorin S (eds) *Stochastic Games and Applications*. NATO Science Series. Kluwer, Berlin, pp 375–395
74. Sorin S, Zamir S (1991) Big match with lack of information on one side (part 3). In: Raghavan TES et al (eds) *Stochastic games and related topics*. Kluwer, pp 101–112
75. Takahashi M (1962) Stochastic games with infinitely many strategies. *J Sci Hiroshima Univ Ser A-I* 26:123–134
76. Thuijsman F, Raghavan TES (1997) Perfect information stochastic games and related classes. *Int J Game Theory* 26:403–408
77. Vieille N (2000) Equilibrium in 2-person stochastic games I: A Reduction. *Israel J Math* 119:55–91
78. Vieille N (2000) Equilibrium in 2-person stochastic games II: The case of recursive games. *Israel J Math* 119:93–126
79. Vrieze OJ, Thuijsman F (1989) On equilibria in repeated games with absorbing states. *Int J Game Theory* 18:293–310
80. Vrieze OJ, Tijs SH (1982) Fictitious play applied to sequences of games and discounted stochastic games. *Int J Game Theory* 12:71–85
81. Whitt W (1980) Representation and approximation of noncooperative sequential games. *SIAM J Control Optim* 18:33–48
- Mertens JF, Sorin S, Zamir S (1994) Repeated games. CORE Discussion Paper 9420-9422
- Raghavan TES, Shapley LS (1991) *Stochastic games and related topics: In honor of Professor L.S. Shapley*. Springer
- Vieille N (2002) Stochastic games: Recent results. In: Aumann RJ, Hart S (eds) *Handbook of game theory with economic applications*, vol 3. Elsevier, pp 1833–1850

### Books and Reviews

- Başar T, Olsder GJ (1995) *Dynamic noncooperative game theory*. Academic
- Filar JA, Vrieze K (1996) *Competitive Markov decision processes*. Springer
- Maitra AP, Sudderth WD (1996) *Discrete gambling and stochastic games*. Springer
- Mertens JF (2002) Stochastic games. In: Aumann RJ, Hart S (eds) *Handbook of game theory with economic applications*, vol 3. Elsevier, pp 1809–1832

## Stochastic Loewner Evolution: Linking Universality, Criticality and Conformal Invariance in Complex Systems

HANS C. FOGEDBY<sup>1,2</sup>

<sup>1</sup> Department of Physics and Astronomy, University of Aarhus, Aarhus, Denmark

<sup>2</sup> Niels Bohr Institute, Copenhagen, Denmark

### Article Outline

Glossary  
 Definition of the Subject  
 Introduction  
 Scaling  
 Conformal Invariance  
 Loewner Evolution  
 Stochastic Loewner Evolution  
 Results and Discussion  
 Future Directions  
 Acknowledgments  
 Bibliography

### Glossary

**Bessel process** The Bessel process operates in  $d$  dimensions and describes the radial distance  $r$  from the origin of a particle performing Brownian motion. The random motion is governed by the Langevin equation  $dr/dt = \kappa(d-1)/2r + \xi$ , where  $\langle \xi \xi \rangle(t) = \kappa \delta(t)$ . For  $d \leq 2$  the motion is recurrent, i. e., returns to the origin; for  $d > 2$  the motion goes off to infinity. In SLE the Bessel process describes the transition between simple curves and self-intersecting curves.

**Brownian motion** Brownian motion,  $B_t$ , is the scaling limit of random walk. Brownian motion is plane-filling, has the fractal dimension  $D = 2$ , and is described by the Langevin equation  $dB_t/dt = \eta_t$ ,  $\langle \eta_t \eta_s \rangle = \delta(t-s)$ .  $B_t$  is characterized by i) the stationarity property,  $B_{t+t'} - B_t$  and  $B_{t'}$  identical in distribution and ii) the independence property,  $B_{\Delta t'}$  and  $B_{\Delta t}$  independent for  $\Delta t \neq \Delta t'$ . The correlations are given by