ARTICLE IN PRESS
YGAME:2638

Games and Economic Behavior ••• (••••) •••–•••

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Acceptable strategy profiles in stochastic games ☆

Eilon Solan

The School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997800, Israel

ARTICLE INFO

Article history:
Received 15 August 2016
Available online xxxx

JEL classification:
C72
C73

Keywords:
Stochastic games
Acceptable strategy profiles
Automata

ABSTRACT

This paper presents a new solution concept for multiplayer stochastic games, namely, acceptable strategy profiles. For each player i and state s in a stochastic game, let $w_i(s)$ be a real number. A strategy profile is *w-acceptable*, where $w = (w_i(s))$, if the discounted payoff to each player i at every initial state s is at least $w_i(s)$, provided the discount factor of the players is sufficiently close to 1. Our goal is to provide simple strategy profiles that are w-acceptable for payoff vectors w in which all coordinates are high.

1. Introduction

Shapley (1953) presented the model of *stochastic games*, which are dynamic games in which the state variable changes from stage to stage as a function of the current state and the actions taken by the players. Shapley (1953) proved that the discounted value exists in two-player zero-sum stochastic games, and provided an equation that the discounted value satisfies.

This seminal work led to an extensive research in several directions (see the surveys by, e.g., Neyman and Sorin, 2003; Mertens et al., 2015; Solan and Vieille, 2015; Solan and Ziliotto, 2016; and Jaśkiewicz and Nowak, 2016, 2017), including the study of the discounted value in games with general state and action sets, the study of discounted equilibria in multiplayer stochastic games, and the study of the robustness of equilibria.

A commonly studied robustness concept is that of uniform equilibrium. A strategy profile is a *uniform ε-equilibrium* for $\varepsilon \geq 0$ if it is an ε-equilibrium in (a) the discounted game, provided the discount factor is sufficiently close to 1, namely, the players are sufficiently patient, and (b) the finite horizon game, provided the horizon is sufficiently long.

Progress in the study of the uniform equilibrium turned out to be slow, existence of such a strategy profile was proven only in special cases (see, e.g., Mertens and Neyman, 1981; Solan, 1999; Vieille, 2000a, 2000b; Solan and Vieille, 2001, Simon, 2007, 2012, 2016; and Flesch et al., 1997, 2008, 2009), and the strategy profiles that are uniform ε-equilibria are usually quite complex.

The present paper proposes a new solution concept for stochastic games that combines simplicity in behavior with relatively high payoffs. Let $w = (w_i(s))$ be a vector, where i ranges over all players and s ranges over all states. A strategy profile in a stochastic game is *w-acceptable* if when the players follow it, for every discount factor sufficiently close to 1, the discounted payoff of each player i is at least $w_i(s)$ when the initial state is s. Thus, when the players follow such a

☆ The author thanks Eitan Altman for helping in identifying relevant references, and acknowledges the support of the Israel Science Foundation, Grant #323/13.

E-mail address: eilons@post.tau.ac.il.

http://dx.doi.org/10.1016/j.geb.2017.01.011
0899-8256/

Please cite this article in press as: Solan, E. Acceptable strategy profiles in stochastic games. Games Econ. Behav. (2017), http://dx.doi.org/10.1016/j.geb.2017.01.011

strategy profile, they forgo the option to profit by deviation in order to guarantee a reasonably high payoff for each player. A strategy profile is *min–max ε-acceptable* (resp. *max–min ε-acceptable*) if it is w-acceptable for the vector $w = (w_i(s))$ that is defined by $w_i(s) := \overline{v}_i^1(s) - \varepsilon$ (resp. $w_i(s) := \underline{v}_i^1(s) - \varepsilon$), where $\overline{v}_i^1(s)$ (resp. $\underline{v}_i^1(s)$) is the uniform min–max value (resp. uniform max–min value) of player i at the initial state s. By Neyman (2003), $\overline{v}_i^1(s)$ is the amount that player i can uniformly guarantee when the other players cooperate to lower his payoff, and $\underline{v}_i^1(s)$ is the same amount when the other players can correlate their actions.

In their study of correlated equilibrium, Solan and Vieille (2002) constructed a min–max ε-acceptable strategy profile in every multiplayer stochastic game and for every $\varepsilon > 0$. Their construction uses the technique of Mertens and Neyman (1981) for designing a uniform ε-optimal strategy in two-player zero-sum stochastic games, and in particular is history dependent.

Our goal in this paper is the construction of simple strategy profiles that are min–max ε-acceptable or max–min ε-acceptable, where simplicity is measured by the size of the automata that are needed to implement the individual strategies of the players.

A naïve suggestion for a stationary min–max ε-acceptable strategy profile is a stationary discounted equilibrium, for some discount factor sufficiently close to 1. As we now explain, this approach is bound to fail. The discounted payoff that corresponds to a stationary strategy profile is the weighted average of the payoffs that are received in the various states, where the weight of a state is equal to the discounted time that the play spends in that state. A discounted equilibrium yields a high discounted payoff to all players, which implies that this weighted average is high. It might happen that while the average payoff of all players is high, some players get high payoff in some states, while other players get high payoff in other states. When we fix a λ-discounted equilibrium and we calculate the payoff according to a discount factor λ' that goes to 1, the weights of the various states change, and there is no guarantee that the weighted average payoff of all players remains high. This phenomenon in fact happens, as can be seen in Example 2.6 below.

We prove the existence of a max–min ε-acceptable strategy profile, in which the strategy of each player can be implemented by an automaton whose number of states is at most the number of states in the stochastic game times the number of players.[1] We also prove the existence of a min–max ε-acceptable stationary correlated strategy, which can be implemented by an automaton whose number of states is the number of states in the stochastic game. The proofs are constructive and identifies (at least) one such strategy profile.

Another view on the concept of w-acceptability stems from the folk theorem. The folk theorem for repeated games states that under proper technical conditions, every feasible and individually rational payoff vector is an equilibrium payoff. Solan (2001) extended this result to stochastic games when considering extensive-form correlated equilibria rather than Nash equilibria. The identification of the set of feasible and individually rational payoffs in multiplayer stochastic games is open. A strategy profile is min–max (resp. max–min) ε-acceptable if it generates a feasible and ε-individually rational payoff vector when punishment is given by the uniform min–max (resp. max–min) value.

Identifying individually rational (when punishment is given by the min–max value) correlated strategy profiles in the discrete-time game is useful for continuous-time stochastic games. Indeed, an ε-individually rational correlated strategy in the discrete-time game can be transformed into an ε-equilibrium in the continuous-time game, see Neyman (2012).

The paper is organized as follows. The model of stochastic games, the concept of acceptable strategy profiles, the main results, a discussion, and open problems appear in Section 2. The proof of the main results and additional discussion appear in Section 3.

2. Model and main results

2.1. The model of stochastic games

A multiplayer *stochastic game* is a vector $\Gamma = (I, S, (A_i)_{i\in I}, (u_i)_{i\in I}, q)$ where

- I is a finite set of players.
- S is a finite set of states.
- A_i is a finite set of actions available to player i at each state.[2] Denote by $A := \times_{i\in I} A_i$ the set of all action profiles.
- $u_i : S \times A \to \mathbf{R}$ is player i's payoff function. We assume w.l.o.g. that the payoffs are bounded between 0 and 1.
- $q : S \times A \to \Delta(S)$ is a transition function, where $\Delta(X)$ is the set of probability distributions over X, for every nonempty finite set X.

The game is played as follows. The initial state $s^0 \in S$ is given. At each stage $n \ge 0$, the current state s^n is announced to the players. Each player i chooses an action $a_i^n \in A_i$; the action profile $a^n = (a_i^n)_{i\in I}$ is publicly announced, the new state s^{n+1} is drawn according to the probability distribution $q(\cdot \mid s^n, a^n)$, and the game proceeds to stage $n+1$.

[1] We consider automata in which the output function depends only on the automaton's state, and not on the input (Moore machine). If the output function can depend both on the automaton's state and on the input (Mealy machine), then the number of required automaton's states is at most the number of players.

[2] We could have assumed that the action set of a player depends on the current state. This would have complicated the definition of an automaton that implements a strategy, hence we prefer to assume that the action set is independent of the state.

Please cite this article in press as: Solan, E. Acceptable strategy profiles in stochastic games. Games Econ. Behav. (2017), http://dx.doi.org/10.1016/j.geb.2017.01.011