Operations Research Letters 43 (2015) 507–513

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Approximated cooperative equilibria for games played over event trees

Elena Parilina [a], Georges Zaccour [b,*]

[a] Saint Petersburg State University, Saint Petersburg, Russia

[b] Chair in Game Theory and Management, GERAD, HEC Montréal, Montréal, Canada

ARTICLE INFO

Article history:
Received 28 May 2015
Received in revised form
20 July 2015
Accepted 20 July 2015
Available online 26 July 2015

Keywords:
Stochastic games
S-adapted strategies
Cooperative solution
Subgame-perfect ε-equilibrium
Trigger strategies

ABSTRACT

We consider the class of stochastic games played over event trees. We suppose that the players agree to cooperate and maximize their joint payoff. To sustain cooperation over the event tree, we use trigger strategies. As we are dealing with a finite horizon, it is known that deviation from cooperation in the last stage cannot be deterred, as there is no possibility for punishing the deviator(s). Consequently, we focus on epsilon equilibria. We prove the existence of an epsilon-perfect equilibrium.

1. Introduction

An important issue in cooperative dynamic games is the sustainability of the agreement over time, i.e., how to ensure that all players stick to their cooperative controls as time goes by. The literature in (state-space) dynamic games has dealt with this issue along essentially two lines, namely, the design of time-consistent mechanisms and cooperative equilibria. In a nutshell, the determination of a time-consistent solution involves a two-step procedure. The first step is the computation of the cooperative solution and selection of an imputation in, e.g., the core or Shapley value. Second, the definition of a payment schedule over time such that: (i) the total stream of payments to a player corresponds to her imputation in the overall cooperative game; (ii) at any intermediate instant of time, the cooperative payoff-to-go dominates its noncooperative counterpart. A time-consistent solution is not an equilibrium, nor is based on unilateral-deviation thinking, that is, either there is an agreement where all parties are on board, or there is no agreement at all. (For a review of time consistency in differential games, see [20].) In the second approach, the idea is to embed the cooperative solution with an equilibrium property, and, hence self-supported. This is achieved by letting the players use non-Markovian (or history-based) strategies that effectively deter cheating on the cooperative agreement.

The objective of this paper is to design a cooperative equilibrium solution for dynamic games played over event trees (DGPET), that is, games where the random process is an act of nature and is not influenced by the players' actions. This class of games, which involves flow (control) and stock (state) variables, is useful to model competition and cooperation between players interacting repeatedly over time in the presence of an accumulation process. As an example, the set of players could be firms belonging to the same industry, where each firm makes an investment (control variable) to increase its production capacity (state variable), and with the price of the product being dependent on all firms' outputs and on some random event (weather, state of the economy, etc.). This class of games was initially introduced in [19, 10] to study noncooperative equilibria in the European natural gas market, involving four suppliers competing over a long-term planning horizon in nine markets described by stochastic demand laws. The solution concept was termed S-adapted equilibrium, where S stands for *sample* of realizations of the random process (see [9] for details). Recently, cooperative DGPET have been considered, with a focus on the sustainability of cooperation over time (see [13,15]). In this paper, our concern is the construction in a finite-horizon setting of an approximated cooperative equilibrium solution for this class of games.

We use a trigger strategy that is based on the following simple rule: if cooperation has prevailed till now, then choose the

* Corresponding author.
E-mail addresses: e.parilina@spbu.ru (E. Parilina), georges.zaccour@gerad.ca (G. Zaccour).

http://dx.doi.org/10.1016/j.orl.2015.07.006
0167-6377/

cooperative control in the current stage; and if a deviation has been observed, then implement a noncooperative (or punishing) control for the rest of the game. The so-called folk theorem about the existence of a subgame perfect equilibrium in trigger strategies for infinitely repeated games was proved long ago (see, e.g., [2]). A similar theorem for stochastic games is proved in [4], see also [12]. In differential games, [18] considered strategies with memory, which are called cooperative, as they were built as behavior strategies incorporating cooperative open-loop controls and feedback strategies used as threats in order to enforce the cooperative agreement.

Folk theorems are for infinite-horizon dynamic games. It is well known that enforcing cooperation in finite-horizon games is more difficult. The reason is that, at the last stage, defection from the agreement is individually rational and this deviation cannot be punished. Using a backward-induction argument, it is easy to show that the unique subgame perfect equilibrium in repeated and multistage games is to implement Nash equilibrium controls at each stage of the finite game. This theoretical result has not always received empirical support. Indeed, experiments show that cooperation may be realized, at least partially, in finite-horizon games (see, e.g. [1]).

The literature has came out with different ways to cope with the difficulties in enforcing cooperation in finite-horizon dynamic games. For instance, in [5] it is proposed to support collusive behavior in finite repeated games by having the players post bonds, which can be forfeited if they detect from cooperative behavior. The idea of ε-equilibrium was proposed and its existence for finitely repeated games is proved in [14]. Other options can be used in the class of finitely repeated games when there exist more than one Nash equilibria in a one-shot game [3].

Recently, the problem of the existence of a subgame-perfect ε-equilibrium in pure strategies has been investigated, see, e.g., [6,17]. The concept of φ-tolerance equilibrium perfect-information games of infinite duration where φ is a function of history was proposed in [7]. A strategy profile is said to be a φ-tolerance equilibrium, if, for any history h, this strategy profile is a $\varphi(h)$-equilibrium in the subgame starting at h. This concept is close to the one we investigate in this paper and to the contemporaneous perfect ε-equilibrium proposed in [11]. Contrary to [6,7,17], we examine finite horizon games with perfect information. Further, different authors retained different measures for payoffs. For instance, in [3,14] it is assumed that the payoff is the average payoff for one stage of the game, whereas in [11] and here, the players' payoffs are given by a stream of discounted payments.

The rest of the paper is organized as follows: Section 2 recalls the main ingredients of DGPET. Section 3 states the problem of strategic support and the main results. We provide an illustrative example in Section 4, and briefly conclude in Section 5.

2. Game over event tree

This section draws heavily on [9,13]. Let $\mathcal{T} = \{0, 1, \ldots, T\}$ be the set of periods. The stochastic process is represented by an event tree, which has a root node n^0 in period 0 and a set of nodes $\mathcal{N}^t = \{n_1^t, \ldots, n_{N_t}^t\}$ in period $t = 1, \ldots, T$. Denote by $a(n_l^t) \in \mathcal{N}^{t-1}$ the unique predecessor of node $n_l^t \in \mathcal{N}^t$ on the event-tree graph for $t = 1, \ldots, T$, and by $\mathcal{S}(n_l^t) \subset \mathcal{N}^{t+1}$ the set of all possible direct successors of node $n_l^t \in \mathcal{N}^t$ for $t = 0, \ldots, T-1$. A path from the root node n^0 to a terminal node n_l^T is called a *scenario*. Each scenario has a probability and the probabilities of all scenarios sum up to 1. We denote by $\pi(n_l^t)$ the probability of passing through node n_l^t, which corresponds to the sum of the probabilities of all scenarios that contain this node. In particular, $\pi(n^0) = 1$ and $\pi(n_l^T)$ is equal to the probability of the single scenario that terminates in (leaf) node n_l^T. Observe that each node $n_l^t \in \mathcal{N}^t$ represents a possible sample value of the history of the stochastic process up to time t. The tree graph structure represents the nesting of information as one time period succeeds the other.

Denote by $M = \{1, \ldots, m\}$ the set of players. For each player $j \in M$, we define a set of decision variables indexed over the set of nodes. Denote by $u_j(n_l^t) \in \mathbb{R}^{m_j}$ the decision variables of player j at node n_l^t, and let $u(n_l^t) = \big(u_1(n_l^t), \ldots, u_m(n_l^t)\big)$. Let $X \subset \mathbb{R}^p$, with $p \in \mathbb{N}_+$, be a state set. For each node $n_l^t \in \mathcal{N}^t$, $t = 0, 1, \ldots, T$, let $U_j^{n_l^t} \subset \mathbb{R}^{\mu_j^{n_l^t}}$, with $\mu_j^{n_l^t} \in \mathbb{N}_+$, be the control set of player j. Denote by $U^{n_l^t} = U_1^{n_l^t} \times \cdots \times U_j^{n_l^t} \times \cdots \times U_m^{n_l^t}$ the product control sets. A transition function $f^{n_l^t}(\cdot, \cdot) : X \times U^{n_l^t} \mapsto X$ is associated with each node n_l^t. The state equations are given by

$$x(n_l^t) = f^{a(n_l^t)}\left(x\left(a\left(n_l^t\right)\right), u\left(a\left(n_l^t\right)\right)\right), \tag{1}$$

$$u\left(a\left(n_l^t\right)\right) \in U^{a(n_l^t)}, \quad n_l^t \in \mathcal{N}^t,\ t = 1, \ldots, T. \tag{2}$$

At each node n_l^t, $t = 0, \ldots, T-1$, the reward to player j is a function of the state and the controls of all players, given by $\phi_j^{n_l^t}(x(n_l^t), u(n_l^t))$. At a terminal node n_l^T, the reward to player j is given by the function $\Phi_j^{n_l^T}(x(n_l^T))$.

We assume that player $j \in M$ maximizes her expected stream of payoffs discounted at rate λ_j $\big(0 < \lambda_j < 1\big)$. The state equations and the reward functions define the following multistage game, where we let

$\mathbf{x} = \{x(n_l^t) : n_l^t \in \mathcal{N}^t, t = 0, \ldots, T\}$,
$\mathbf{u} = \{u(n_l^t) : n_l^t \in \mathcal{N}^t, t = 0, \ldots, T-1\}$,
and $J_j(\mathbf{x}, \mathbf{u})$ be the payoff to player j, that is,

$$J_j(\mathbf{x}, \mathbf{u}) = \sum_{t=0}^{T-1} \lambda_j^t \sum_{n_l^t \in \mathcal{N}^t} \pi(n_l^t) \phi_j^{n_l^t}(x(n_l^t), u(n_l^t)) + \lambda_j^T \sum_{n_l^T \in \mathcal{N}^T} \pi(n_l^T) \Phi_j^{n_l^T}(x(n_l^T)), \quad j \in M, \tag{3}$$

s.t.

$$x(n_l^t) = f^{a(n_l^t)}(x(a(n_l^t)), u(a(n_l^t))), \quad x(n^0) = x^0, \tag{4}$$
$$u(a(n_l^t)) \in U^{a(n_l^t)}, \quad n_l^t \in \mathcal{N}^t,\ t = 1, \ldots, T.$$

Definition 1. An admissible S-adapted strategy of player j is a vector $\mathbf{u}_j = \{u_j(n_l^t) : n_l^t \in \mathcal{N}^t, t = 0, \ldots, T-1\}$, that is, a plan of actions adapted to the history of the random process represented by the event tree.

The S-adapted strategy vector of the m players is $\mathbf{u} = (\mathbf{u}_j : j \in M)$. We can thus define a game in normal form, with payoffs $W_j(\mathbf{u}, x^0) = J_j(\mathbf{x}, \mathbf{u})$, $j \in M$, where $\mathbf{x}$ is obtained from $\mathbf{u}$ as the unique solution of the state equations that emanate from the initial state x^0.

If the game is played noncooperatively, then the players will seek a Nash equilibrium in S-adapted strategies defined as follows:

Definition 2. An S-adapted Nash equilibrium is an admissible S-adapted strategy profile $\mathbf{u}^N$ such that for every player $j \in M$ the following condition holds:

$$W_j(\mathbf{u}^N, x^0) \geq W_j([\mathbf{u}_j, \mathbf{u}_{-j}^N], x^0),$$

where $[\mathbf{u}_j, \mathbf{u}_{-j}^N]$ is the S-adapted strategy profile when all players $i \neq j$, $i \in M$, use their Nash equilibrium policies.

Remark 1. We suppose that the joint-optimization solution and the Nash equilibrium in the whole game and in any subgame are unique. The uniqueness for the joint-optimization solution