Operations Research Letters 45 (2017) 269–274

Contents lists available at ScienceDirect

Operations Research Letters

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

Time consistent cooperative solutions for multistage games with vector payoffs

Denis Kuzyutin [a,b], Mariya Nikitina [b,*]

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

[b] National Research University Higher School of Economics (HSE), Saint Petersburg, Russia

ARTICLE INFO

Article history:
Received 25 October 2016
Received in revised form
3 April 2017
Accepted 10 April 2017
Available online 19 April 2017

Keywords:
Multicriteria game
Dynamic game
Time consistency
Shapley value
Cooperative solution

ABSTRACT

To ensure sustainable cooperation in multistage games with vector payoffs we use the payment schedule based approach. The main dynamic properties of cooperative solutions used in single-criterion multistage games are extended to multicriteria games.

We design two recurrent payment schedules that satisfy such advantageous properties as the efficiency and the time consistency conditions, non-negativity and irrational behavior proofness.

1. Introduction

Multicriteria games (or games with vector payoffs) are used to model interactive decision situations in which every player attempts to optimize more than one objective. For example, in a multiobjective environmental game [5,24] a player (i.e. a country) aims at simultaneously increasing production, obtaining larger quota for emissions, saving health care costs, etc. Multicriteria games can be used when modeling various real-life situations where several objectives have to be taken into account, especially if the players do not have an a priori estimation of the relative importance of the components of their payoff vectors. Starting from the pioneering papers [2,27], much research has been done on non-cooperative multicriteria games (see, e.g., [3,30,29,11,12]). The problem of cooperative behavior in games with vector payoffs was examined in [23,24].

This paper is mainly focused on the dynamic aspects of cooperation in n-person multicriteria games. Namely, we deal with multicriteria multistage games in extensive form with perfect information [15,21,11,12] where the players' vector payoffs $h_i(x_t)$ at each node x_t of the game tree are assumed to be non-negative. The first step of cooperation is to obtain the largest possible total vector payoff. We assume that the players have agreed on a rule γ regarding how to choose a particular Pareto optimal solution (using, for instance, the approaches from [25,8,14,16,12]) and hence, the respective optimal cooperative trajectory $\omega = (x_0, \bar{x}_1, \ldots, \bar{x}_t, \bar{x}_{t+1}, \ldots, \bar{x}_T)$.

The next step that should be done to achieve stable cooperation is to choose an allocation mechanism to divide the total cooperative vector payoff between the players. Such an allocation is available if the cooperative game satisfies a certain kind of transferable utility property, i.e., if the payoff transfer w.r.t. the same criterion k from one player to another is possible. As an example of a single-valued cooperative solution, we consider the Shapley value [26] (an easily computed solution possessing a number of advantageous properties such as, e.g., monotonicity, see [13]), which was extended into multicriteria games in [24].

Finally, the main issue in dynamic cooperative games is the consistency (or sustainability) of the cooperative agreement over time. This problem consists in designing an appropriate payment schedule (or imputation distribution procedure [20,22,21,28]) that satisfies certain advantageous properties such as efficiency, time consistency [19,20,7,21,9,10,12,17,18], irrational behavior proofness [31], non-negativity, etc. It is worth noting that not all of the criteria can be measured in monetary terms. However, we use the term *payment* with respect to each criterion for the sake of simplicity and uniformity.

The above mentioned properties can be briefly described as follows. If the payment schedule is *time consistent* (TC), then at no subgame along the optimal trajectory can any player do better by

* Corresponding author.
E-mail address: maryaniki@gmail.com (M. Nikitina).

http://dx.doi.org/10.1016/j.orl.2017.04.004
0167-6377/

deviating from the original cooperative agreement. The *irrational behavior proof* condition (IBP) ensures that each player has an incentive to cooperate even if he assumes that the cooperation can be destroyed due to the irrational behavior of the other players at some intermediate node before the end of the game. Note that the problem of time consistency and irrational behavior proofness of the Shapley value for multicriteria games in extensive form has not been studied yet. To ensure the sustainability of a cooperative agreement over time we use the payment schedule based approach which was extensively studied for single-criterion differential and multistage games (see, e.g., [20–22,28]).

The main goal of this paper is to design time consistent and irrational behavior proof payment schedules for a multicriteria multistage game while guaranteeing non-negative payments to the players along the optimal cooperative trajectory. In particular, we propose a modification of the TC condition (called the *time consistency inequality*) and construct a *refined payment schedule* that satisfies efficiency and non-negativity properties as well as the time consistency inequality. Finally, we extend the IBP property to multicriteria multistage games to obtain a *generalized payment schedule* which satisfies efficiency and non-negativity conditions, the TC inequality and the strong IBP as well.

The refined payment schedule can be considered as an improvement of the allocation procedure (12), suggested in [20,22,21] for single-criterion games, and also as an extension of known payment schedule into the class of multicriteria extensive games. The generalized payment schedule (meeting both the TC inequality and the IBP) has no analogues for single-criterion extensive games.

The rest of this paper is organized as follows: in Section 2 we introduce the game and define the optimal cooperative trajectory. Section 3 states the problem of time consistency and provides an illustrative example (a three-person bicriteria multistage game) that will be used extensively throughout the paper. In Section 4, the refined payment schedule is formalized. Section 5 is devoted to extending the IBP property to the class of multicriteria games and designing a generalized payment schedule. Finally, Section 6 contains conclusions.

2. Multistage multicriteria game

In this section we consider a multistage r-criteria game with perfect information following [21,11,12]. To start with we define the following notations that will be used in the sequel:

- $N = \{1, \dots, n\}$ is the set of players;
- K is the game tree with the set of nodes P and the root x_0;
- $S(x)$ is the set of all direct successors of the node x in K and $S^{-1}(y)$ is the unique predecessor of the node $y \neq x_0$ such that $y \in S(S^{-1}(y))$;
- $\omega = (x_0, \dots, x_{t-1}, x_t, \dots, x_T)$ is the path (or trajectory) in the game tree, $x_{t-1} = S^{-1}(x_t)$, $1 \leqslant t \leqslant T$; $S(x_T) = \varnothing$;
- P_i is the set of all player i's decision nodes, $P_i \cap P_j = \varnothing$ for $i \neq j$, and P_{n+1} is the set of all terminal nodes, $\cup_{i=1}^{n+1} P_i = P$;
- $h_i(x) = (h_{i/1}(x), \dots, h_{i/r}(x))$ is the ith player's vector payoff at the node $x \in P$.

We assume that

$$h_{i/k}(x) \geq 0; \quad \forall i \in N;\ k = 1, \dots, r;\ x \in P. \tag{1}$$

In the following, we will focus on the games with perfect information where the players use pure strategies (see, e.g., [15,21]). The pure strategy $u_i(\cdot)$ of player i is a function with domain P_i that specifies for every node $x \in P_i$ the next node $u_i(x) \in S(x)$ which the player i should choose at x. Let U_i denote the (finite) set of all ith player's pure strategies, $U = \prod_{i \in N} U_i$. The strategy profile $u = (u_1, \dots, u_n) \in U$ generates the trajectory $\omega = (x_0, \dots, x_t, x_{t+1}, \dots, x_T)$, where $x_{t+1} = u_j(x_t) \in S(x_t)$ if $x_t \in P_j$ and, respectively, a collection of all players' vector payoffs.

Denote by

$$H_i(u) = h_i(\omega) = \sum_{t=0}^{T} h_i(x_t), \tag{2}$$

the value of player i's vector payoff function, given by strategy profile $u = (u_1, \dots, u_n)$.

Let $a, b \in R^m$; we consider the following vector preferences: $a \geqq b$ if $a_k \geqslant b_k$, $k = 1, \dots, m$; $a > b$ if $a_k > b_k$, $k = 1, \dots, m$; $a \geq b$ if $a \geqq b$ and $a \neq b$. The last vector inequality means that b is Pareto dominated by a.

For a given finite set A and a vector-function $f \colon A \to R^m$, denote by $Max_{a \in A} f(a)$ the set of all Pareto optimal (nondominated) elements from A, i.e.:

$$b \in Max_{a \in A} f(a) \quad \text{if} \nexists\, a \in A : f(a) \geq f(b).$$

If the players agree to cooperate, they maximize w.r.t. the binary relation $\geq$ the total vector payoff $\sum_{i=1}^{n} H_i(u)$. Denote by U^c the set of all strategy profiles u^c such that $u^c \in Max_{u \in U} \sum_{i=1}^{n} H_i(u)$, i.e.:

$$\nexists u : \sum_{i=1}^{n} H_i(u) \geq \sum_{i=1}^{n} H_i(u^c). \tag{3}$$

Since the set U (and $\{\sum_{i=1}^{n} H_i(u)\}_{u \in U}$) is finite, the set U^c of all Pareto optimal solutions is known to be nonempty (see, e.g., [25,8]).

There are different approaches on how to choose a particular Pareto optimal solution [25,8,14,16,12]. We will assume henceforth that the players have agreed to use a specific rule γ in order to choose a particular strategy profile $\bar{u} = \gamma(U^c)$ from the set U^c. We will call $\bar{u}$ the optimal cooperative strategy profile and the corresponding path $\bar{\omega} = (\bar{x}_0, \bar{x}_1, \dots, \bar{x}_t, \dots, \bar{x}_T)$ with $\bar{x}_0 = x_0$ will be referred to as the *optimal cooperative trajectory*.

3. Time consistent Shapley value

When designing a cooperative agreement it is important to choose an allocation mechanism to divide the total cooperative vector payoff $\sum_{i=1}^{n} H_i(\bar{u})$ between the players.

Let $\Gamma^{x_0}(N, V^{x_0})$ be a multicriteria TU cooperative game [24], where $S \subseteq N$ is a coalition, $V^{x_0}(S) : 2^N \to R^r$ is a (vector-valued) characteristic function of the game with $V^{x_0}(\varnothing) = (0, \dots, 0)$, and $V^{x_0}(N) = \sum_{i=1}^{n} H_i(\bar{u})$. To construct a characteristic function for the given strategic game one can use different approaches (see, e.g., [1,4,6] for single-criterion games). The main results presented in this paper can be used equally well for any characteristic function. However, we chosen to use the so-called α-characteristic function [1] in Example 1

Definition 1 (*[26,24]*)**.** The Shapley value of $\Gamma^{x_0}(N, V^{x_0})$ denoted by φ^{x_0} is defined for each player $i \in N$ as

$$\varphi_i^{x_0} = \sum_{S \subset N, i \in S} \frac{(n - |S|)!(|S| - 1)!}{n!} (V^{x_0}(S) - V^{x_0}(S \setminus \{i\})). \tag{4}$$

Remark 1 (*[26,24]*)**.** The Shapley value in a multicriteria cooperative game is an imputation, i.e.:

$$\varphi_i^{x_0} \geqq V^{x_0}(\{i\}) \quad \forall i \in N, \tag{5}$$

$$\sum_{i=1}^{n} \varphi_i^{x_0} = V^{x_0}(N) = \sum_{\tau=0}^{T} \sum_{i=1}^{n} h_i(\bar{x}_\tau). \tag{6}$$