Chapter 10
Cooperation Enforcing in Multistage Multicriteria Game: New Algorithm and Its Implementation

Denis Kuzyutin, Ivan Lipko, Yaroslavna Pankratova, and Igor Tantlevskij

Abstract To enforce the long-term cooperation in a multistage multicriteria game we use the imputation distribution procedure (IDP) based approach. We mainly focus on such useful properties of the IDP like "reward immediately after the move" assumption, time consistency inequality, efficiency and non-negativity constraint. To overcome the problem of negative payments along the optimal cooperative trajectory the novel refined A-incremental IDP is designed. We establish the properties of the proposed A-incremental payment schedule and provide an illustrative example to clarify how the algorithm works.

Keywords Dynamic game · Multistage game · Multicriteria game · Cooperative solution · Shapley value · Time consistency · Imputation distribution procedure

10.1 Introduction

The theory of multicriteria games (multiobjective games or the games with vector payoffs) develops at the overlap of classical game theory and multiple criteria decision analysis. It can be used to model various real-world decision-making problems where several objectives (or criteria) have to be taken into account (see, e.g., [1, 2, 14, 26] a player aims at simultaneously increasing production, obtaining large quote for the use of a common resource, saving costs of water purification, saving health care costs, etc. Starting from [29], much research has

D. Kuzyutin (✉)
Saint Petersburg State University, St. Petersburg, Russia

National Research University Higher School of Economics (HSE), St. Petersburg, Russia
e-mail: d.kuzyutin@spbu.ru

I. Lipko · Y. Pankratova · I. Tantlevskij
Saint Petersburg State University, St. Petersburg, Russia
e-mail: st052710@student.spbu.ru; y.pankratova@spbu.ru; tantigor@bk.ru

L. A. Petrosyan et al. (eds.), *Frontiers of Dynamic Games*,
Static & Dynamic Game Theory: Foundations & Applications,
https://doi.org/10.1007/978-3-030-51941-4_10

been done on non-cooperative multicriteria games (see, e.g., [8, 12, 24, 31]). Different cooperative solutions for static and dynamic multicriteria games were examined in [9–11, 13, 22].

This paper is mainly focused on the dynamic aspects of cooperation enforcing in an n-person multistage multicriteria games in extensive form (see, e.g., [6, 7, 20]) with perfect information. In order to achieve and implement a long-term cooperative agreement in a multicriteria dynamic game the players have to solve the following problems. First, when players seek to reach the maximal total vector payoff of the grand coalition, they face the problem of choosing a unique Pareto optimal payoffs vector. In the dynamic setting it is necessary that a specific method the players agreed to accept in order to choose a particular Pareto optimal solution not only takes into account the relative importance of the criteria, but also satisfies time consistency [5, 6, 9, 17–20, 25, 27], i.e., a fragment of the optimal cooperative trajectory in the subgame should remain optimal in this subgame. In the paper, we assume that the players employ the refined leximin (RL) algorithm, introduced in [11], to select a unique Pareto optimal solution for each multicriteria optimization problem they face. This approach allows constructing time consistent cooperative trajectory and vector-valued characteristic function. Another appropriate method—the rule of the minimal sum of relative deviations from the ideal payoffs vector—was suggested in [13].

After choosing the cooperative trajectory it is necessary to construct a vector-valued characteristic function. For instance, when analyzing the Example 10.1, we employ a friendly computable ζ-characteristic function introduced in [3] as well as the RL-algorithm in order to choose a particular Pareto efficient solution for the auxiliary vector optimization problems. To determine the optimal payoff allocation we adopt the vector analogue of the Shapley value [9, 22, 28]. Such an approach is based on the assumption that the payoff can be transferred between the players within the same criterion. It is worth noting that the main measurable criteria used in multicriteria resource management problems usually satisfy this component-wise transferable utility property.

Lastly, to guarantee the sustainability of the achieved long-term cooperative agreement the players are expected to use an appropriate imputation distribution procedure (IDP), i.e. a payoff allocation rule that determines the actual current payments to every player along the optimal cooperative trajectory. The IDP based approach was extensively studied for single-criterion differential and multistage games (see, e.g. [15, 16, 18, 20, 21]) and was extended to multicriteria multistage games in [9, 10]. The detailed review of useful properties the IDP may satisfy for multistage multicriteria games is presented in [9–13].

In particular, two novel properties an acceptable payment schedule for the multistage game should satisfy which take into account the sequence of the players' actions along the optimal cooperative trajectory were suggested in [12]. Firstly, a player which moves at position x according to the cooperative scenario expects to receive some reward for the "correct" move immediately after this move, while the other players (which are inactive at x) should get zero current payments. Furthermore, if the position x is the last player i's node along the cooperative trajectory this player should get the rest of her optimal payoff right after her last move. These properties were formalised in the so-called Reward Immediately after the Move (RIM) assumption (see [12] for details).

In this paper we mainly focus on the RIM assumption, efficiency and non-negativity constraint as well as time consistency property. The first so-called "incremental" IDP was suggested in [18] to ensure time consistency of the solution in differential single-criterion game, then this simple IDP was extended to different classes of dynamic games. The A-incremental IDP that satisfies RIM assumption, efficiency constraint and time consistency for multicriteria multistage game was designed in [12]. However, as it is demonstrated in the paper, the A-incremental IDP, as well as the classical incremental IDP may imply negative current payments to some players at some nodes (see [4, 9, 20] for details). One approach how to overcome this negative feature of the incremental IDP—the refined payment schedule for multicriteria games—was constructed in [9]. Another regularisation method for single-criterion multistage game was proposed in [4]. In this paper we provide a refinement of the A-incremental imputation distribution procedure for multicriteria multistage game. This "refined A-incremental IDP" is proved to satisfy the RIM assumption, non-negativity constraint, efficiency condition and time consistency inequality.

Hence, the main contribution of this paper is twofold:

– we reveal one possible disadvantage of the A-incremental payment schedule, namely that it may imply negative current payments to the players. To overcome this drawback we design the novel A-refined imputation distribution procedure which satisfies a number of usefull properties (in particular, non-negativity).
– we provide the step-by-step algorithm how to implement this novel allocation rule. Then we compare the implementation of the simple A-incremental IDP and the refined A-incremental IDP for given 3-person bicriteria multistage game.

The rest of the paper is organized as follows: The class of r-criteria multistage n-person games in extensive form with perfect information is formalized in Sect. 10.2. The optimal cooperative trajectory and vector-valued characteristic function are constructed in Sect. 10.3 using the refined leximin algorithm. We provide an illustrative example of the 3-person bicriteria multistage game here. Different useful properties of imputation distribution procedure are formulated in Sect. 10.4. In Sect. 10.5, we discuss the implementation of the A-incremental IDP and reveal the problem of negative payments. We provide a refined A-incremental IDP and the algorithm of its implementation in Sect. 10.6 and a brief conclusion in Sect. 10.7.

10.2 Multistage Game with Vector Payoffs

We consider a finite multistage r-criteria game in extensive form with perfect information following [7, 9, 20]. First we define the following notations that will be used throughout the paper:

- $N = \{1, \ldots, n\}$ is the finite set of players;
- K is the game tree with the root x_0 and the set of all nodes P;
- $S(x)$ is the set of all direct successors (descendants) of the node x and $S^{-1}(y)$ is the unique predecessor (parent) of the node $y \neq x_0$ such that $y \in S(S^{-1}(y))$;
- P_i is the set of all player i's decision nodes, $P_i \cap P_j = \varnothing$ for $i \neq j$, and $P_{n+1} = \{y^j\}_{j=1}^m$ is the set of all terminal nodes, $S(y^j) = \varnothing \; \forall y^j \in P_{n+1}$, $\cup_{i=1}^{n+1} P_i = P$;
- $\omega = (x_0, \ldots, x_{t-1}, x_t, \ldots, x_T)$ is the trajectory (or path) in the game tree, $x_{t-1} = S^{-1}(x_t)$, $1 \leqslant t \leqslant T$; $x_T = y^j \in P_{n+1}$, the lower index t in x_t denotes the number of the node within the trajectory ω and can be interpreted as the "time index", T is an ordinal number of the last node of the trajectory ω;
- $h_i(x) = (h_{i/1}(x), \ldots, h_{i/r}(x))$ is the i-th player's vector payoff at the node $x \in P \backslash \{x_0\}$.

We assume that

$$h_{i/k}(x) \geq 0; \;\; \forall i \in N; \;\; k = 1, \ldots, r; \;\; x \in P \backslash \{x_0\}.$$

Let us use $MG^P(n, r)$ to denote the class of all finite multistage n-person r-criteria games in an extensive form with perfect information. Since we will focus on the games with perfect information we restrict ourselves to the class of pure strategies (see, e.g., [7, 20]). 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 i-th player's pure strategies, $U = \prod_{i \in N} U_i$. Every strategy profile $u = (u_1, \ldots, u_n) \in U$ generates the trajectory $\omega(u) = (x_0, \ldots, x_t, x_{t+1}, \ldots, x_T) = (x_0, x_1(u), \ldots, x_t(u), x_{t+1}(u), \ldots, x_T(u))$, where $x_{t+1} = u_j(x_t) \in S(x_t)$ if $x_t \in P_j$, $0 \leq t \leq T-1$, $x_T \in P_{n+1}$, and, respectively, a collection of all players' vector payoffs.

Denote by

$$H_i(u) = (H_{i/1}(u), \ldots, H_{i/r}(u)) = \tilde{h}_i(\omega(u)) = \sum_{\tau=1}^{T} h_i(x_\tau(u)),$$

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

In the multistage multicriteria game Γ^{x_0} defined above every intermediate node $x_t \in P \backslash P_{n+1}$ generates a subgame Γ^{x_t} with the subgame tree K^{x_t} and the subroot x_t as well as a factor-game with the factor-game tree $K^D = \{x_t\} \cup (K \backslash K^{x_t})$ (see, for instance [20]). Decomposition of the original extensive game Γ^{x_0} at node x_t into the subgame Γ^{x_t} and the factor-game Γ^D generates the corresponding decomposition of pure strategies.

Let $P_i^{x_t}(P_i^D)$, $i = 1, \ldots, n$ denote the restriction of P_i on the subtree $K^{x_t}(K^D)$, and $u_i^{x_t}(u_i^D)$, $i = 1, \ldots, n$, denote the restriction of the player i's pure strategy $u_i(\cdot)$ in Γ^{x_0} on $P_i^{x_t}(P_i^D)$. The strategy profile $u^{x_t} = (u_1^{x_t}, \ldots, u_n^{x_t})$ generates the trajectory $\omega^{x_t}(u^{x_t}) = (x_t, x_{t+1}, \ldots, x_T) = (x_t, x_{t+1}(u^{x_t}), \ldots, x_T(u^{x_t}))$ and, respectively, a collection of all player's vector payoffs in the subgame. Denote by

$$H_i^{x_t}(u^{x_t}) = \tilde{h}_i^{x_t}(\omega^{x_t}(u^{x_t})) = \sum_{\tau=t+1}^{T} h_i(x_\tau(u^{x_t})), \tag{10.1}$$

the value of player i's vector payoff function in the subgame Γ^{x_t}, and by $U_i^{x_t}$ the set of all player i's pure strategies in Γ^{x_t}, $U^{x_t} = \prod_{i \in N} U_i^{x_t}$. Note that

$$\begin{aligned} H_i(u) = \tilde{h}_i(\omega(u)) = \sum_{\tau=1}^{T} h_i(x_\tau(u)) = \sum_{\tau=1}^{t} h_i(x_\tau(u)) + \\ \sum_{\tau=t+1}^{T} h_i(x_\tau(u^{x_t})) = \tilde{h}_i(\underline{\omega}^{x_t}(u)) + \tilde{h}_i^{x_t}(\omega^{x_t}(u^{x_t})), \end{aligned} \tag{10.2}$$

where $\underline{\omega}^{x_t}(u) = (x_0, x_1, \ldots, x_{t-1}, x_t)$ denotes a part of trajectory $\omega(u)$ before the subgame Γ^{x_t} starts.

Remark 10.1 Since $P_i = P_i^{x_t} \cup P_i^D$ while $P_i^{x_t} \cap P_i^D = \varnothing$, one can compose the player i's pure strategy $W_i = (u_i^D, v_i^{x_t}) \in U_i$ in the original game Γ^{x_0} from his strategies $v_i^{x_t} \in U_i^{x_t}$ and $u_i^D \in U_i^D$ in the subgame Γ^{x_t} and factor-game Γ^D respectively [20].

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

10.3 Designing a Cooperative Solution

If the players agree to cooperate in multicriteria game Γ^{x_0}, they maximize w.r.t. the binary relation $\geq$ the total vector payoff $\sum_{i=1}^{n} H_i(u)$. Denote by $PO(\Gamma^{x_0})$ the set of all Pareto optimal strategy profiles from U, i.e.:

$$u \in PO(\Gamma^{x_0}) \; if \; \nexists \, v \in U : \sum_{i \in N} H_i(v) \geq \sum_{i \in N} H_i(u)$$

The set $PO(\Gamma^{x_0})$ is known to be nonempty (see, e.g., [23]) and in general it contains multiple strategy profiles. Since the set $PO(\Gamma^{x_0})$ may contain more than one strategy profile, the players face the problem how to select a unique Pareto optimal cooperative strategy profile $\bar{u} \in PO(\Gamma^{x_0})$ and corresponding optimal cooperative trajectory $\bar{\omega} = \bar{\omega}(\bar{u}) = (\bar{x}_0, \bar{x}_1, \ldots, \bar{x}_T)$. In a dynamic game it is essential that a specific method the players agreed to employ in order to choose a particular Pareto optimal solution has to satisfy time consistency, that is, a fragment $\bar{\omega}^{x_t}(\bar{u}^{x_t}) = (\bar{x}_t, \bar{x}_{t+1}, \ldots, \bar{x}_T)$ of the optimal trajectory $\bar{\omega}$ in the subgame $\Gamma^{\bar{x}_t} \in G(\bar{u})$ should remain optimal trajectory for this subgame.

We employ the so-called Refined Leximin algorithm, introduced in [11] to find optimal cooperative trajectory in Example 10.1. This approach looks reasonable for the special case when the criteria have significantly different importance, and all the players rank the criteria in the same order. In other circumstances, the players may employ other appropriate methods to choose a unique Pareto optimal solutions (an example of such methods—the rule of minimal sum of relative deviations from the ideal payoffs vector—was suggested in [13]). Note that the main result of the paper—Proposition 10.1—does not depend on the particular time consistent rule which the players have agreed to use in order to choose a unique Pareto optimal solution.

Let us briefly remind the main idea of the RL algorithm and the notations (the reader could find the comprehensive specification of this algorithm in [11, 12]). Suppose that all the criteria are ordered in accordance with their relative importance for the players, namely let criterion 1 be the most important for every player $i \in N$, the next to be the 2-nd criterion, and so on, and the last criterion r be the least important one. When choosing the optimal cooperative trajectory the players are expected to maximise the total vector payoff primarily on the first criterion, i.e.

$$\max_{u \in U} \sum_{i \in N} H_{i/1}(u) = \sum_{i \in N} H_{i/1}(\bar{u}) = \bar{H}_1.$$

If there exists a unique trajectory $\bar{\omega} = \omega(\bar{u})$ satisfying this condition then this trajectory is called the optimal cooperative trajectory while $\bar{u}$) is the optimal cooperative strategy profile.

If there are several trajectories $\omega(u)$ with $\sum_{i \in N} H_{i/1}(u) = \bar{H}_1$ the players should choose such trajectory from this set $\overline{PO}_1(\Gamma^{x_0})$ that

$$\max_{u \in \overline{PO}_1(\Gamma^{x_0})} \sum_{i \in N} H_{i/2}(u) = \bar{H}_2,$$

and so on. Lastly, if there are several trajectories $\omega \in \{\omega(u), u \in \overline{PO}_r(\Gamma^{x_0})\}$, the players should choose the trajectory from this set with minimal number j of the terminal node y^j.

We will suppose henceforth that the players have agreed to use the RL algorithm in order to choose the *optimal cooperative strategy profile* $\bar{u} \in PO(\Gamma^{x_0})$ and the corresponding *optimal cooperative trajectory* $\bar{\omega} = \omega(\bar{u}) = (\bar{x}_0, \ldots, \bar{x}_T)$.

Let

$$Max^L_{u \in U} \sum_{i \in N} H_i(u) = \sum_{i \in N} H_i(\bar{u}) \tag{10.3}$$

denote the maximal (in the sense of the RL algorithm) total vector payoff. Note that the Pareto optimal cooperative trajectory $\bar{\omega} = \omega(\bar{u}) = (\bar{x}_0, \dots, \bar{x}_T)$ based on the RL algorithm was proved to satisfy time consistency [11].

Let us use the following example to demonstrate how the players choose the cooperative trajectory and then to explore and compare the A-incremental IDP and the refined A-incremental payment schedule.

Example 10.1 (A 3-Player Bicriteria Multistage Game) The game tree K is shown in Fig. 10.1. Let $n = 3$, $r = 2$, $P_1 = \{\overline{x}_0, \overline{x}_2, \overline{x}_4, \overline{x}_6\}$, $P_2 = \{\overline{x}_1, \overline{x}_5\}$, $P_3 = \{\overline{x}_3\}$, $P_{n+1} = \{z_1, \dots, z_9\}$,

$$h(x_t) = \begin{pmatrix} h_{1/1}(x_t) & h_{2/1}(x_t) & h_{3/1}(x_t) \\ h_{1/2}(x_t) & h_{2/2}(x_t) & h_{3/2}(x_t) \end{pmatrix},$$

i.e. the columns correspond to the players while the rows correspond to the criteria. The players' payoffs at all nodes $x \in P \backslash \{x_0\}$ are:

$$h(x_1) = \begin{pmatrix} 0 & 6 & 0 \\ 0 & 0 & 12 \end{pmatrix}, \; h(x_2) = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 0 & 12 \end{pmatrix}, \; h(x_3) = \begin{pmatrix} 0 & 6 & 0 \\ 0 & 0 & 12 \end{pmatrix},$$

$$h(x_4) = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 12 & 0 \end{pmatrix}, \; h(x_5) = \begin{pmatrix} 0 & 6 & 0 \\ 0 & 0 & 12 \end{pmatrix}, \; h(x_6) = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 0 & 12 \end{pmatrix},$$

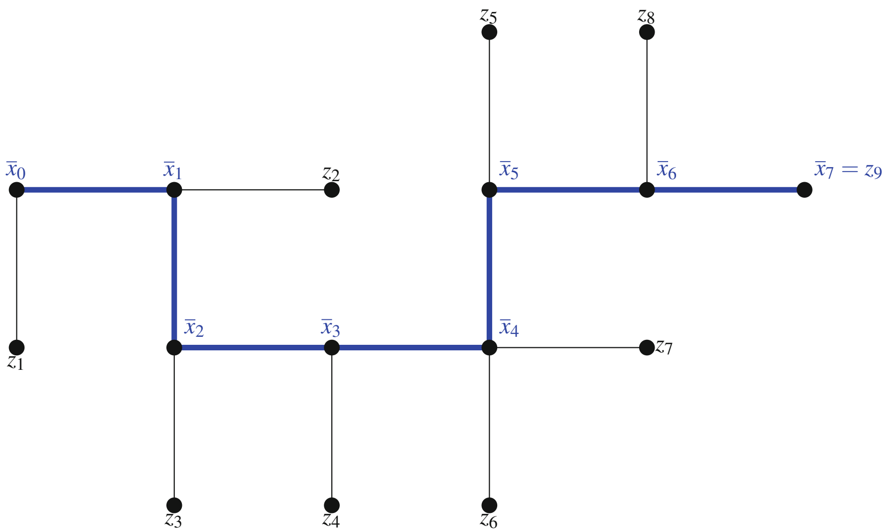

Fig. 10.1 The game tree

$$h(x_7) = \begin{pmatrix} 30\ 30\ 30 \\ 60\ 30\ 30 \end{pmatrix},\ h(z_1) = \begin{pmatrix} 18\ 0\ 0 \\ 18\ 0\ 0 \end{pmatrix},\ h(z_2) = \begin{pmatrix} 0\ 18\ 0 \\ 0\ 18\ 0 \end{pmatrix},$$

$$h(z_3) = \begin{pmatrix} 18\ 0\ 0 \\ 18\ 0\ 0 \end{pmatrix},\ h(z_4) = \begin{pmatrix} 0\ 0\ 18 \\ 0\ 0\ 18 \end{pmatrix},\ h(z_5) = \begin{pmatrix} 0\ \ 18\ 0 \\ 60\ 18\ 0 \end{pmatrix},$$

$$h(z_6) = \begin{pmatrix} 18\ 0\ 0 \\ 18\ 0\ 0 \end{pmatrix},\ h(z_7) = \begin{pmatrix} 90\ \ 6\ 0 \\ 162\ 0\ 6 \end{pmatrix},\ h(z_8) = \begin{pmatrix} 18\ \ 0\ 0 \\ 150\ 0\ 0 \end{pmatrix},$$

There are three pure strategy Pareto optimal strategy profiles in $PO(\Gamma^{x_0})$:

$$\bar{u}_1(x_0) = x_1,\ \bar{u}_2(x_1) = x_2,\ \bar{u}_1(x_2) = x_3,\ \bar{u}_3(x_3) = x_4,\ \bar{u}_1(x_4) = x_5,$$

$$\bar{u}_2(x_5) = x_6,\ \bar{u}_1(x_6) = x_7$$

that generates trajectory $\bar{\omega}(\bar{u}) = (x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7)$,

$$u_1'(x_0) = x_1,\ u_2'(x_1) = x_2,\ u_1'(x_2) = x_3,\ u_3'(x_3) = x_4,\ u_1'(x_4) = z_7,$$

that generates trajectory $\bar{\omega}(u') = (x_0, x_1, x_2, x_3, x_4, z_7)$ and

$$u_1''(x_0) = x_1,\ u_2''(x_1) = x_2,\ u_1''(x_2) = x_3,\ u_3''(x_3) = x_4,\ u_1''(x_4) = x_5,$$

$$u_2'(x_5) = x_6,\ u_1''(x_6) = z_8$$

that generates trajectory $\bar{\omega}(u'') = (x_0, x_1, x_2, x_3, x_4, x_5, z_8)$.

Using RL algorithm the players choose the optimal cooperative strategy profile $\bar{u} = (\bar{u}_1, \bar{u}_2, \bar{u}_3)$ which generates the optimal cooperative trajectory $\bar{\omega} = \omega(\bar{u}) = (x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7) = (\bar{x}_0, \bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4, \bar{x}_5, \bar{x}_6, \bar{x}_7)$.

After selecting a cooperative trajectory it is necessary to construct a vector-valued characteristic function for a multicriteria cooperative game. In Example 10.1 we use a vector-valued analogue of the so-called ζ-characteristic function introduced in [3] and again the RL algorithm (see [11] for details). Namely:

$$V^{x_0}(S) = \begin{cases} 0, & S = \emptyset \\ \underset{u_j, j \in N \setminus S}{Min^L} \sum_{i \in S} H_i(\bar{u}_S, u_{N \setminus S}), & S \subset N, \\ \underset{u \in U}{Max^L} \sum_{i \in N} H_i(u), & S = N \end{cases} \tag{10.4}$$

where

$$\underset{u_j, j \in N \setminus S}{Min^L} \sum_{i \in S} H_i(\bar{u}_S, u_{N \setminus S}) = - \underset{u_j, j \in N \setminus S}{Max^L} (- \sum_{i \in S} H_i(\bar{u}_S, u_{N \setminus S})).$$

Let $\Gamma^{x_0}(N, V^{x_0})$ denote multicriteria game $\Gamma^{x_0} \in MG^P(n, r)$ with characteristic function V^{x_0}. It is worth noting that one can use other approaches to construct characteristic function (CF) for multicriteria game, say classical α-CF or δ-CF [21], but as it was mentioned in [3] the ζ-characteristic function is much more easy to compute (it is essential especially for multicriteria case). Note that the main result of the paper—Proposition 10.1—does not depend on the specific method which the players employ to calculate the vector-valued characteristic function.

We assumed that the players adopt a single-valued cooperative solution φ^{x_0} (for instance, the vector analogue of the Shapley value [9, 28]) for the cooperative game $\Gamma^{x_0}(N, V^{x_0})$ which satisfies the efficiency property

$$\sum_{i=1}^{n} \varphi_i^{x_0} = V^{x_0}(N) = \sum_{\tau=1}^{T} \sum_{i=1}^{n} h_i(\bar{x}_\tau), \tag{10.5}$$

and the individual rationality property

$$\varphi_i^{x_0} \geqq V^{x_0}(\{i\}),\ i = 1, \dots, n. \tag{10.6}$$

Denote by $\Gamma^{\bar{x}_t}(N, V^{\bar{x}_t})$, $\bar{x}_t \in \bar{\omega}(\bar{u})$, $t = 0, \dots, T-1$ a subgame along the optimal cooperative trajectory with the characteristic function $V^{\bar{x}_t}$ which can be computed in the subgame using (10.4). Note that $V^{\bar{x}_t}(N) = \sum_{\tau=t+1}^{T} \sum_{i \in N} h_i(\bar{x}_\tau)$.

In addition, we assume that the same properties (10.5) and (10.6) are valid for the cooperative solutions $\varphi^{\bar{x}_t}$ at each subgame $\Gamma^{\bar{x}_t}(N, V^{\bar{x}_t})$, $t = 0, \dots, T-1$.

10.4 Imputation Distribution Procedure and Its Properties

Let $\beta = \{\beta_{i/k}(\bar{x}_\tau)\}, i = 1, \dots, n;\ k = 1, \dots, r;\ \tau = 1, \dots, T$ denote the Imputation Distribution Procedure—IDP [9, 18, 20, 25] or the payment schedule. The IDP-based approach implies that the players have agreed to accumulate the cooperative vector payoff $\sum_{i \in N} H_i(\bar{u}) = V^{x_0}(N)$, obtained using the initial payoffs $h_i(\bar{x}_\tau)$, and then allocate this summary payoff between the players along the optimal cooperative trajectory $\bar{\omega}(\bar{u})$. Then $\beta_{i/k}(\bar{x}_\tau)$ corresponds to the actual current payment which the player i receives at $\bar{x}_\tau$ w.r.t. criterion k (instead of $h_{i/k}(\bar{x}_\tau)$) according to the IDP β.

From now on we suppose that the IDP β should satisfy the following assumption (see [12] for details):

Assumption RIM (Reward Immediately After the Move) If $\bar{x}_t \in P_i$, $t = 0, \dots, T-1$, then $\beta_j(\bar{x}_{t+1}) = 0$ for all $j \in N \backslash \{i\}$, i.e. the only player who can receive nonzero current payment at node $\bar{x}_{t+1}$ is the player i which moves at the previous node $\bar{x}_t = S^{-1}(\bar{x}_{t+1})$.

For given player $i \in N$ let $(y_1^i, y_2^i, \dots, y_{T(i)}^i)$ denote the ordered set of all the positions from the set $P_i \cap \bar{\omega}$ along the optimal trajectory $\bar{\omega}$, where nodes $\{y_\tau^i\}$ are listed in order of their location in $\bar{\omega}$. Namely,

$$y_1^i = \bar{x}_{t^i(1)},\ y_2^i = \bar{x}_{t^i(2)}, \dots, y_{T(i)}^i = \bar{x}_{t^i(T(i))};$$

and for all $y_\lambda^i = \bar{x}_{t^i(\lambda)}$ and $y_m^i = \bar{x}_{t^i(m)}$, we have $\lambda < m$ if and only if $t(\lambda) < t^i(m)$.

Below, we introduce a number of useful properties an acceptable IDP may satisfy (see [9–11]). Note that we need to modify known definitions of efficiency and time consistency to take assumption RIM into account.

To simplify the notations, henceforth we will omit superscript i in $t^i(\lambda)$, $\lambda = 1, \dots, T(i)$, i.e. we will write $\beta_i(\bar{x}_{t(\lambda)+1})$ instead of $\beta_i(\bar{x}_{t^i(\lambda)+1})$, e.t.c.

Definition 10.1 ([12]) The imputation distribution procedure $\beta = \{\beta_{i/k}(\overline{x}_t)\}$ satisfies the *efficiency condition* if

$$\sum_{t=1}^{T} \beta_i(\bar{x}_t) = \sum_{\lambda=1}^{T(i)} \beta_i(\bar{x}_{t(\lambda)+1}) = \varphi_i^{\bar{x}_0},\ i = 1, \dots, n. \tag{10.7}$$

Indeed, if (10.7) holds then the payment schedule for every player can be considered as a rule for the step-by-step allocation of the player i's optimal payoff.

Definition 10.2 The IDP $\beta = \{\beta_{i/k}(\bar{x}_t)\}$ meets the *time consistency (TC) inequality* if for every player $i \in N$ such that $|T(i)| \geqslant 2$, for all $\tau = 1, \dots, T(i) - 1$ it holds that

$$\sum_{\lambda=1}^{\tau} \beta_i(\bar{x}_{t(\lambda)+1}) + \varphi_i^{\bar{x}_{t(\tau)+1}} \geqq \varphi_i^{\bar{x}_0}. \tag{10.8}$$

The vector inequality (10.8) implies that every player has an incentive to continue cooperation at every subgame along the cooperative trajectory.

Definition 10.3 ([9]) The imputation distribution procedure $\beta = \{\beta_{i/k}(\bar{x}_t)\}$ satisfies the *balance condition* if $\forall t = 0, \dots, T$; $\forall k = 1, \dots, r$ it holds that

$$\sum_{\tau=1}^{t} \sum_{i=1}^{n} \beta_{i/k}(\bar{x}_\tau) \leq \sum_{\tau=1}^{t} \sum_{i=1}^{n} h_{i/k}(\bar{x}_\tau) \tag{10.9}$$

Note that (10.9) is always satisfied for $t = T$ due to the efficiency condition (10.7) and (10.5). If β does not satisfy (10.9) at some intermediate node $\bar{x}_t$, we will suppose that the players may borrow the required amount on account of future earnings. For the sake of simplicity we assume that an interest-free loan is available for the grand coalition N while recognising that in general case the enforcing of a cooperative agreement may require extra costs (see [9]).

Definition 10.4 ([9]) The IDP β satisfies the *non-negativity constraint* if

$$\beta_{i/k}(\bar{x}_t) \geqslant 0,\ i = 1, \ldots, n;\ k = 1, \ldots, r;\ t = 1, \ldots, T.$$

Note that there could be different payment schedules that may or may not satisfy the properties listed above (several IDP for multicriteria games are examined in [9–11]). The A-incremental IDP that satisfies RIM assumption, efficiency constraint and time consistency (equation) for multicriteria multistage game was suggested in [12]:

Definition 10.5 The *A-incremental imputation distribution procedure* $\beta = \{\beta_{i/k}(\bar{x}_t)\}, t = 0, \ldots, T; i \in N$ is formulated as follows

(c1) $\beta_{i/k}(\bar{x}_0) = 0, i = 1, \ldots, n; k = 1, \ldots r;$
(c2) if $\bar{x}_t \in P_i, t = 0, \ldots, T-1$, then $\beta_j(\bar{x}_{t+1}) = 0$ for all $j \in N\backslash\{i\}$;
(c3) if $\bar{x}_t \in P_i$ and $T(i) = 1$, i.e. $\bar{\omega} \cap P_i = (y_1^i) = \{\bar{x}_{t(1)}\}$, then

$$\beta_i(\bar{x}_{t(1)+1}) = \varphi_i^{\bar{x}_0} \tag{10.10}$$

(c4) if $\bar{x}_t \in P_i$ and $T(i) = 2$, i.e. $\bar{\omega} \cap P_i = (y_1^i, y_2^i) = (\bar{x}_{t(1)}, \bar{x}_{t(2)})$, then

$$\beta_i(\bar{x}_{t(1)+1}) = \varphi_i^{\bar{x}_0} - \varphi_i^{\bar{x}_{t(1)+1}}; \quad \beta_i(\bar{x}_{t(2)+1}) = \varphi_i^{\bar{x}_{t(1)+1}} \tag{10.11}$$

(c5) if $\bar{x}_t \in P_i$ and $T(i) \geq 3$, i.e. $\bar{\omega} \cap P_i = (y_1^i, y_2^i, \ldots, y_{T(i)}^i) = (\bar{x}_{t(1)}, \bar{x}_{t(2)}, \ldots, \bar{x}_{t(T(i))})$, then

$$\begin{aligned}
&\beta_i(\bar{x}_{t(1)+1}) = \varphi_i^{\bar{x}_0} - \varphi^{\bar{x}_{t(1)+1}};\\
&\beta_i(\bar{x}_{t(\lambda)+1}) = \varphi_i^{\bar{x}_{t(\lambda-1)+1}} - \varphi_i^{\bar{x}_{t(\lambda)+1}},\ \lambda = 2, \ldots, T(i)-1;\\
&\beta_i(\bar{x}_{t(T(i))+1}) = \varphi_i^{\bar{x}_{t(T(i))-1)+1}}.
\end{aligned} \tag{10.12}$$

10.5 A-Incremental IDP May Imply Negative Current Payments

Let us use the game from Ex. 1 to demonstrate the A-incremental IDP implementation and properties and to reveal one possible disadvantage of this payment schedule. We will adopt the vector analogue of the Shapley value as an optimal cooperative solution when analysing Ex. 1.

Definition 10.6 ([22, 28]) 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

$$\varphi_i^{x_0} = \sum_{S \subset N, i \in S} \frac{(n - |S|)!(|S| - 1)!}{n!} (V^{x_0}(S) - V^{x_0}(S \setminus \{i\})). \tag{10.13}$$

Example 10.1 (Continued) The values of the vector-valued ζ-characteristic function (10.4) for the game Γ^{x_0} are

S	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	N
$V^{x_0}(S)$	0	0	0	18	0	0	126
	0	0	0	0	12	0	192

and the Shapley value for original game Γ^{x_0} is

$$\varphi^{x_0} = \begin{pmatrix} 45 & 45 & 36 \\ 66 & 60 & 66 \end{pmatrix}.$$

The vector-valued ζ-characteristic functions and the respective Shapley values for the subgames along the cooperative trajectory $\bar{\omega}$ can be constructed using the same approach.

The subgame $\Gamma^{x_1}(N, V^{x_1})$:

S	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	N
$V^{x_1}(S)$	0	0	0	12	0	0	120
	0	0	0	0	0	12	180

$$\varphi^{x_1} = \begin{pmatrix} 42 & 42 & 36 \\ 56 & 62 & 62 \end{pmatrix}.$$

The subgame $\Gamma^{x_2}(N, V^{x_2})$:

S	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	N
$V^{x_2}(S)$	0	0	0	6	6	0	114
	0	0	0	0	84	0	168

$$\varphi^{x_2} = \begin{pmatrix} 40 & 37 & 37 \\ 70 & 28 & 70 \end{pmatrix}.$$

The subgame $\Gamma^{x_3}(N, V^{x_3})$:

S	{1}	{2}	{3}	{1, 2}	{1, 3}	{2, 3}	N
$V^{x_3}(S)$	0	0	0	0	6	0	108
	0	0	0	0	72	12	156

$$\varphi^{x_3} = \begin{pmatrix} 37 & 34 & 37 \\ 60 & 30 & 66 \end{pmatrix}.$$

The subgame $\Gamma^{x_4}(N, V^{x_4})$:

S	{1}	{2}	{3}	{1, 2}	{1, 3}	{2, 3}	N
$V^{x_4}(S)$	0	0	0	72	0	0	102
	60	0	0	90	72	0	144

$$\varphi^{x_4} = \begin{pmatrix} 46 & 46 & 10 \\ 95 & 29 & 20 \end{pmatrix}.$$

The subgame $\Gamma^{x_5}(N, V^{x_5})$:

S	{1}	{2}	{3}	{1, 2}	{1, 3}	{2, 3}	N
$V^{x_5}(S)$	0	0	0	66	0	0	96
	60	0	0	90	60	12	132

$$\varphi^{x_5} = \begin{pmatrix} 43 & 43 & 10 \\ 85 & 31 & 16 \end{pmatrix}.$$

The subgame $\Gamma^{x_6}(N, V^{x_5})$:

S	{1}	{2}	{3}	{1, 2}	{1, 3}	{2, 3}	N
$V^{x_6}(S)$	30	0	0	60	60	0	90
	60	0	0	90	90	0	120

$$\varphi^{x_6} = \begin{pmatrix} 60 & 15 & 15 \\ 90 & 15 & 15 \end{pmatrix}.$$

Applying the A-incremental IDP (10.10), (10.12), (10.13) we obtain the following current payments along the optimal cooperative path $\bar{\omega} = (\bar{x}_0 = y_1^1,\ \bar{x}_1 = y_1^2,\ \bar{x}_2 = y_2^1,\ \bar{x}_3 = y_1^3,\ \bar{x}_4 = y_3^1,\ \bar{x}_5 = y_2^2,\ \bar{x}_6 = y_4^1,\ \bar{x}_7) : \beta_1(\bar{x}_1) = \varphi_1^{\bar{x}_0} - \varphi_1^{\bar{x}_1} =$

$\begin{pmatrix} 3 \\ 10 \end{pmatrix}$, $\beta_1(\bar{x}_3) = \varphi_1^{\bar{x}_1} - \varphi_1^{\bar{x}_3} = \begin{pmatrix} 5 \\ -4 \end{pmatrix}$, $\beta_1(\bar{x}_5) = \varphi_1^{\bar{x}_3} - \varphi_1^{\bar{x}_5} = \begin{pmatrix} -6 \\ -25 \end{pmatrix}$, $\beta_1(\bar{x}_7) = \varphi_1^{\bar{x}_5} = \begin{pmatrix} 43 \\ 85 \end{pmatrix}$, since $T(1) = 4$;

$\beta_j(\bar{x}_1) = \beta_j(x_3) = \beta_j(x_5) = \beta_j(x_7) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $j = 2, 3$;

$\beta_2(\bar{x}_2) = \varphi_2^{\bar{x}_0} - \varphi_2^{\bar{x}_2} = \begin{pmatrix} 8 \\ 32 \end{pmatrix}$, $\beta_2(\bar{x}_6) = \varphi_2^{\bar{x}_2} = \begin{pmatrix} 37 \\ 28 \end{pmatrix}$ since $T(2) = 2$;

$\beta_j(\bar{x}_2) = \beta_j(\bar{x}_6) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $j = 1, 3$;

$\beta_3(\bar{x}_4) = \varphi_3^{\bar{x}_0} = \begin{pmatrix} 36 \\ 66 \end{pmatrix}$ since $T(3) = 1$; $\beta_j(\bar{x}_4) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $j = 1, 2$.

The efficiency condition for the player 1 and criterion 2 takes the form:

$$\sum_{t=1}^{7} \beta_{1/2}(\bar{x}_t) = \sum_{\lambda=1}^{4} \beta_{1/2}(\bar{x}_{t(\lambda)+1}) = 10 - 4 - 25 + 85 = 66 = \varphi_1^{\bar{x}_0}.$$

The time consistency equations for the player 1 and criterion 2 take the form:

$$\tau = 1 : \sum_{\lambda=1}^{1} \beta_{1/2}(\bar{x}_{t(\lambda)+1}) + \varphi_1^{\bar{x}_1} = 10 + 56 = 66 = \varphi_1^{\bar{x}_0};$$

$$\tau = 2 : \sum_{\lambda=1}^{2} \beta_{1/2}(\bar{x}_{t(\lambda)+1}) + \varphi_1^{\bar{x}_3} = 10 - 4 + 60 = 66 = \varphi_1^{\bar{x}_0};$$

$$\tau = 3 : \sum_{\lambda=1}^{3} \beta_{1/2}(\bar{x}_{t(\lambda)+1}) + \varphi_1^{\bar{x}_5} = 10 - 4 - 25 + 85 = 66 = \varphi_1^{\bar{x}_0}.$$

As it was mentioned in the Introduction, the A-incremental IDP, as well as the classical incremental IDP may imply negative current payments to some players at some nodes (see [4, 9, 20] for details). Thus, in Example 10.1 the A-incremental IDP implies negative payments to player 1 at $\bar{x}_3$ and $\bar{x}_5$.

10.6 Refined A-Incremental IDP and Its Implementation

Below we introduce a refinement of A-incremental payment schedule that is designed to satisfy assumption RIM, efficiency, time consistency inequality and non-negativity constraint.

We will use an auxiliary integer variable $a_{i/k}(\lambda)$ to denote the number of nodes $\bar{x}_{t(\tau)+1}$ on the optimal cooperative path $\bar{\omega}$ from $\bar{x}_\tau$ to $\bar{x}_{t(\lambda)+1}$ for which $\beta_{i/k}(\bar{x}_{t(\tau)+1}) = 0$ after the last positive current payment (one may call it the payment delay variable). We assume that $t^i(0) = -1$ for any i.

Definition 10.7 The *refined A-incremental imputation distribution procedure* $\hat{\beta} = \{\hat{\beta}_{i/k}(\bar{x}_t)\}$, $t = 0, \ldots, T$; $i \in N$ is formulated as follows

(c1) $\hat{\beta}_{i/k}(\bar{x}_0) = 0, i = 1, \ldots, n; k = 1, \ldots r;$

(c2) if $\bar{x}_t \in P_i, t = 0, \ldots, T-1$, then $\hat{\beta}_j(\bar{x}_{t+1}) = 0$ for all $j \in N \backslash \{i\}$;

(c3) if $\bar{x}_t \in P_i$ and $T(i) = 1$, i.e. $\bar{\omega} \cap P_i = (y_1^i) = \{\bar{x}_{t(1)}\}$, then

$$\hat{\beta}_{i/k}(\bar{x}_{t(1)+1}) = \varphi_{i/k}^{\bar{x}_0} \tag{10.14}$$

(c4) if $\bar{x}_t \in P_i$ and $T(i) = 2$, i.e. $\bar{\omega} \cap P_i = (y_1^i, y_2^i) = (\bar{x}_{t(1)}, \bar{x}_{t(2)})$, then

$$\hat{\beta}_{i/k}(\bar{x}_{t(1)+1}) = \max\{\varphi_{i/k}^{\bar{x}_0} - \varphi_{i/k}^{\bar{x}_{t(1)+1}}, 0\} \tag{10.15}$$

$$\hat{\beta}_{i/k}(\bar{x}_{t(2)+1}) = \varphi_{i/k}^{\bar{x}_0} - \hat{\beta}_{i/k}(\bar{x}_{t(1)+1}) \tag{10.16}$$

(c5) if $\bar{x}_t \in P_i$ and $T(i) \geq 3$, i.e. $\bar{\omega} \cap P_i = (y_1^i, y_2^i, \ldots, y_{T(i)}^i) = (\bar{x}_{t(1)}, \bar{x}_{t(2)}, \ldots, \bar{x}_{t(T(i))})$, then

Step 1 ($\lambda = 1$):

$$\hat{\beta}_{i/k}(\bar{x}_{t(1)+1}) = \max\{\varphi_{i/k}^{\bar{x}_0} - \varphi^{\bar{x}_{t(1)+1}}, 0\}; \tag{10.17}$$

- if $\hat{\beta}_{i/k}(\bar{x}_{t(1)+1}) > 0$, then $a_{i/k}(1) = 0$ (no delay in payment compared to A-incremental IDP);
- if $\hat{\beta}_{i/k}(\bar{x}_{t(1)+1}) = 0$, then $a_{i/k}(1) = 1$ (the delay in payment for a one step).

Step 2 ($\lambda = 2$):

$$\hat{\beta}_{i/k}(\bar{x}_{t(2)+1}) = \max\{\varphi_{i/k}^{\bar{x}_{t(1-a_{i/k}(1))+1}} - \varphi_{i/k}^{\bar{x}_{t(2)+1}}, 0\}, \tag{10.18}$$

- if $\hat{\beta}_{i/k}(\bar{x}_{t(2)+1}) > 0$, then $a_{i/k}(2) = 0$;
- if $\hat{\beta}_{i/k}(\bar{x}_{t(2)+1}) = 0$, then $a_{i/k}(2) = a_{i/k}(1) + 1$.

Step λ ($\lambda = 2, \ldots, T(i) - 1$):

$$\hat{\beta}_{i/k}(\bar{x}_{t(\lambda)+1}) = \max\{\varphi_{i/k}^{\bar{x}_{t(\lambda-1-a_{i/k}(\lambda-1))+1}} - \varphi_{i/k}^{\bar{x}_{t(\lambda)+1}}, 0\}, \tag{10.19}$$

- if $\hat{\beta}_{i/k}(\bar{x}_{t(\lambda)+1}) > 0$, then $a_{i/k}(\lambda) = 0$ (no delay in payment);
- if $\hat{\beta}_{i/k}(\bar{x}_{t(\lambda)+1}) = 0$, then $a_{i/k}(\lambda) = a_{i/k}(\lambda - 1) + 1$ (the delay in payment at $\bar{x}_{t(\lambda)+1}$ for $a_{i/k}(\lambda)$ steps).

Step $\lambda = T(i)$:

$$\hat{\beta}_{i/k}(\bar{x}_{t(T(i))+1}) = \max\{\varphi_{i/k}^{\bar{x}_0} - \sum_{\lambda=1}^{T(i)-1} \hat{\beta}_{i/k}(\bar{x}_{t(\lambda)+1}), 0\} = \varphi_{i/k}^{\bar{x}_{t(T(i)-1-a_{i/k}(T(i)-1))+1}} \tag{10.20}$$

By the construction of this refined payment schedule the following proposition holds.

Proposition 10.1 *Refined A-incremental IDP satisfies assumption RIM, efficiency condition (10.7), non-negativity constraint and time consistency inequality (10.8).*

Now we will apply the refined A-incremental algorithm (10.14)–(10.20) to the game from Example 10.1.

Example 10.1 (Continued) Note that if the A-incremental IDP implies non-negative current payments to the ith player w.r.t. criterion k at all nodes along the cooperative trajectory $\bar{\omega}$, then $\hat{\beta}_{i/k}(\bar{x}_t) = \beta_{i/k}(\bar{x}_t)$, $\bar{x}_t \in \bar{\omega}$. Hence, the current payments to the player 2 and 3 according to the refined A-incremental IDP $\hat{\beta}$ will not change compared to the A-incremental IDP β.

Let us now consider the payments to the player $i = 1$:

$$\begin{aligned}
&\hat{\beta}_{1/1}(\bar{x}_1) = \max\{\varphi_{1/1}^{\bar{x}_0} - \varphi_{1/1}^{\bar{x}_1}; 0\} = 3, a_{1/1}(1) = 0;\\
&\hat{\beta}_{1/2}(\bar{x}_1) = \max\{\varphi_{1/2}^{\bar{x}_0} - \varphi_{1/2}^{\bar{x}_1}; 0\} = 10, a_{1/2}(1) = 0;\\
&\hat{\beta}_{1/1}(\bar{x}_3) = \max\{\varphi_{1/1}^{\bar{x}_1} - \varphi_{1/1}^{\bar{x}_3}; 0\} = \max\{5; 0\} = 5, a_{1/1}(2) = 0;\\
&\hat{\beta}_{1/2}(\bar{x}_3) = \max\{\varphi_{1/2}^{\bar{x}_1} - \varphi_{1/2}^{\bar{x}_3}; 0\} = max\{-4; 0\} = 0, a_{1/2}(2) = 1;\\
&\hat{\beta}_{1/1}(\bar{x}_5) = \max\{\varphi_{1/1}^{\bar{x}_3} - \varphi_{1/1}^{\bar{x}_5}; 0\} = \max\{-6; 0\} = 0, a_{1/1}(3) = 1;\\
&\hat{\beta}_{1/2}(\bar{x}_5) = \max\{\varphi_{1/2}^{\bar{x}_1} - \varphi_{1/2}^{\bar{x}_5}; 0\} = max\{-29; 0\} = 0, a_{1/2}(3) = 2;\\
&\hat{\beta}_{1/1}(\bar{x}_7) = \varphi_{1/1}^{\bar{x}_3} = 37;\\
&\hat{\beta}_{1/2}(\bar{x}_7) = \varphi_{1/2}^{\bar{x}_1} = 56.
\end{aligned}$$

All the payments to player $i = 1$ are non-negative now. Note that the current payments at $\bar{x}_7$ are less than the relevant payments according to the simple A-incremental IDP.

The efficiency condition for the player 1 and criterion 2 now takes the form:

$$\sum_{t=1}^{7} \hat{\beta}_{1/2}(\bar{x}_t) = \sum_{\lambda=1}^{4} \hat{\beta}_{1/2}(\bar{x}_{t(\lambda)+1}) = 10 + 0 + 0 + 56 = 66 = \varphi_1^{\bar{x}_0}.$$

The time consistency inequalities for the player 1 and criterion 2 take the form:

$$\tau = 1 : \sum_{\lambda=1}^{1} \hat{\beta}_{1/2}(\bar{x}_{t(\lambda)+1}) + \varphi_1^{\bar{x}_1} = 10 + 56 \geq 66 = \varphi_1^{\bar{x}_0};$$

$$\tau = 2 : \sum_{\lambda=1}^{2} \hat{\beta}_{1/2}(\bar{x}_{t(\lambda)+1}) + \varphi_1^{\bar{x}_3} = 10 + 0 + 60 \geq 66 = \varphi_1^{\bar{x}_0};$$

$$\tau = 3 : \sum_{\lambda=1}^{3} \hat{\beta}_{1/2}(\bar{x}_{t(\lambda)+1}) + \varphi_1^{\bar{x}_5} = 10 + 0 + 0 + 85 \geq 66 = \varphi_1^{\bar{x}_0}.$$

Note that the refined A-incremental IDP may not necessarily satisfy balance condition (10.9). Let us for instance consider the balance condition in Example 10.1 for $t = 4$ and $k = 2$:

$$\sum_{\tau=1}^{4} \sum_{i=1}^{3} \hat{\beta}_{i/2}(\bar{x}_\tau) = 108 > \sum_{\tau=1}^{4} \sum_{i=1}^{3} h_{i/2}(\bar{x}_\tau) = 48.$$

As it was firstly noted in [9], in general it is impossible to design a time consistent IDP which satisfies both the balance condition and non-negativity constraint.

10.7 Conclusion

When analyzing Example 10.1, we adopt the Shapley value as an optimal imputation and use the RL algorithm for choosing a unique Pareto optimal solution (to find optimal cooperative trajectory and to construct vector-valued characteristic function). It is worth noting that the provided algorithm to calculate the refined A-incremental IDP as well as Proposition 10.1 remains valid if the players employ another optimal imputation, other approach to calculate the characteristic function and other time consistent rule for choosing a particular Pareto optimal solution, for instance, the rule of minimal sum of relative deviations from the ideal payoffs vector [13].

Note that, since the set of active players in extensive form game changes while the game is evolving along the optimal path, multistage game could be considered as an example of the so-called “games with changing conditions”. The RIM assumption and the proposed refined A-incremental payment schedule allows taking into account this specific feature of a n-person multistage game. It is worth noting that similar assumptions could be implied implicitly in some ancient texts—cf., for instance, the so-called “History of King David’s ascent to power” in connection with David’s activity at the beginning of his career (see, e.g.: [30]). The detailed interdisciplinary analysis of the relevant motivation for “optimal” behaviour could be an interesting issue for further research.

Acknowledgments We would like to thank the anonymous Reviewer for the valuable comments. The research of the first and the third author was funded by RFBR under the research project 18-00-00727 (18-00-00725). The research of the fourth author was funded by RFBR under the research project 18-00-00727 (18-00-00628).

References

1. Climaco, J., Romero, C., Ruiz, F.: Preface to the special issue on multiple criteria decision making: current challenges and future trends. Intl. Trans. Oper. Res. **25**, 759–761 (2018) http://dx.doi.org/10.1111/itor.12515
2. Finus, M.: Game Theory and International Environmental Cooperation. Edward Elgar Ed (2001).
3. Gromova, E. V., Petrosyan, L. A.: On an approach to constructing a characteristic function in cooperative differential games. Autom. Remote Control **78**(9), 1680–1692 (2017)
4. Gromova, E.V., Plekhanova, T.M.: On the regularization of a cooperative solution in a multistage game with random time horizon. Discrete Appl. Math. **255**, 40–55 (2019) http://dx.doi.org/10.1016/j.dam.2018.08.008
5. Haurie, A.: A note on nonzero-sum diferential games with bargaining solution. J. Optim. Theory Appl. **18**, 31–39 (1976)
6. Haurie, A., Krawczyk, J. B., Zaccour, G.: Games and Dynamic Games. Scientific World, Singapore (2012)
7. Kuhn, H.: Extensive games and the problem of information. Ann. Math. Stud. **28**, 193–216 (1953)
8. Kuzyutin, D.: On the consistency of weak equilibria in multicriteria extensive games. In: Petosyan, L.A., Zenkevich, N.A. (eds.) Contributions to Game Theory and Management, Vol. V, pp. 168–177 (2012)
9. Kuzyutin, D., Nikitina, M.: Time consistent cooperative solutions for multistage games with vector payoffs. Oper. Res. Lett. **45**(3), 269–274 (2017) http://dx.doi.org/10.1016/j.orl.2017.04.004
10. Kuzyutin, D., Nikitina, M.: An irrational behavior proof condition for multistage multicriteria games. In: Proceedings of the Consrtuctive Nonsmooth Analysis and Related Topics (Dedic. to the Memory of V.F. Demyanov), CNSA 2017, pp. 178–181. IEEE (2017)
11. Kuzyutin, D., Gromova, E., Pankratova, Ya.: Sustainable cooperation in multicriteria multistage games. Oper. Res. Lett. **46**(6), 557–562 (2018). http://dx.doi.org/10.1016/j.orl.2018.09.004
12. Kuzyutin, D., Pankratova, Y., Svetlov, R.: A-subgame concept and the solutions properties for multistage games with vector payoffs. In: Petrosyan, L., Mazalov, V., Zenkevich, N. (eds.) Frontiers of Dynamic Games. Static & Dynamic Game Theory: Foundations & Applications. Birkhäuser, Cham (2019) http://dx.doi.org/10.1007/978-3-030-23699-1_6

13. Kuzyutin, D., Smirnova, N., Gromova, E.: Long-term implementation of the cooperative solution in multistage multicriteria game. Oper. Res. Perspect. **6**, 100107 (2019) http://dx.doi.org/10.1016/j.orp.2019.100107
14. Mendoza, G.A., Martins, H.: Multi-criteria decision analysis in natural resource management: A critical review of methods and new modelling paradigms. For. Ecol. Manage. **230**, 1–22 (2006)
15. Parilina, E., Zaccour, G.: Node-consistent core for games played over event trees. Automatica **55**, 304–311 (2015)
16. Parilina, E., Zaccour, G.: Node-consistent Shapley value for games played over event trees with random terminal time. J. Optim. Theory Appl. **175**(1), 236–254 (2017)
17. Petrosyan, L.: Stable solutions of differential games with many participants. Vestn. Leningr. Univ. **19**, 46–52 (1977) (in Russian)
18. Petrosyan, L., Danilov, N.: Stability of the solutions in nonantagonistic differential games with transferable payoffs. Vestn. Leningr. Univ. **1**, 52–59 (1979) (in Russian)
19. Petrosyan, L.A., Kuzyutin, D.V.: On the stability of E-equilibrium in the class of mixed strategies. Vestn. St.Petersburg Univ. Math. **3**(15), 54–58 (1995) (in Russian)
20. Petrosyan, L., Kuzyutin, D.: Games in Extensive Form: Optimality and Stability. Saint Petersburg University Press (2000) (in Russian)
21. Petrosyan, L., Zaccour, G.: Time-consistent Shapley value allocation of pollution cost reduction. J. Econ. Dyn. Control **27**(3), 381–398 (2003)
22. Pieri, G., Pusillo, L.: Multicriteria partial cooperative games. Appl. Math. **6**, 2125–2131 (2015)
23. Podinovskii, V., Nogin, V.: Pareto-Optimal Solutions of Multicriteria Problems. Nauka (1982) (in Russian)
24. Puerto, J., Perea, F.: On minimax and Pareto optimal security payoffs in multicriteria games. J. Math. Anal. Appl. **457**(2), 1634–1648 (2018). http://dx.doi.org/10.1016/j.jmaa.2017.01.002
25. Reddy, P., Shevkoplyas, E., Zaccour, G.: Time-consistent Shapley value for games played over event trees. Automatica **49**(6), 1521–1527 (2013)
26. Rettieva, A.: Cooperation in dynamic multicriteria games with random horizons. J. Glob. Optim. (2018). http://dx.doi.org/10.1007/s10898-018-0658-6
27. Sedakov, A.: On the strong time consistency of the core. Autom. Remote Control **79**(4), 757–767 (2018)
28. Shapley, L.: A value for n-person games. In: Kuhn, H., Tucker, A.W. (eds.) Contributions to the Theory of Games, II, pp. 307–317. Princeton University Press, Princeton (1953)
29. Shapley, L.: Equilibrium points in games with vector payoffs. Nav. Res. Logist. Q. **6**, 57–61 (1959)
30. Tantlevskij, I.: King David and His Epoch in the Bible and History. 3rd edn. Revised and Enlarged. RChGA Publishing House (2020) (in Russian)
31. Voorneveld, M., Vermeulen, D., Borm, P.: Axiomatizations of Pareto equilibria in multicriteria games. Games Econom. Behav. **28**, 146–154 (1999)