Artificial Intelligence 173 (2009) 180–195

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

Strong mediated equilibrium ☆

Dov Monderer, Moshe Tennenholtz *

Faculty of Industrial Engineering and Management, Technion – Israel Institute of Technology, Haifa 32000, Israel

A R T I C L E I N F O

Article history:
Received 22 November 2007
Received in revised form 5 October 2008
Accepted 7 October 2008
Available online 14 October 2008

Keywords:
Multi-agent systems
Game theory
Mediator
Mediated equilibrium
Strong equilibrium

A B S T R A C T

Stability against potential deviations by sets of agents is a most desired property in the design and analysis of multi-agent systems. However, unfortunately, this property is typically not satisfied. In game-theoretic terms, a strong equilibrium, which is a strategy profile immune to deviations by coalition, rarely exists. This paper suggests the use of mediators in order to enrich the set of situations where we can obtain stability against deviations by coalitions. A mediator is defined to be a reliable entity, which can ask the agents for the right to play on their behalf, and is guaranteed to behave in a pre-specified way based on messages received from the agents. However, a mediator cannot enforce behavior; that is, agents can play in the game directly, without the mediator's help. A mediator generates a new game for the players, the mediated game. We prove some general results about mediators, and mainly concentrate on the notion of strong mediated equilibrium, which is just a strong equilibrium at the mediated game. We show that desired behaviors, which are stable against deviations by coalitions, can be obtained using mediators in several classes of settings.

1. Introduction

When considering a prescribed behavior in a multi-agent system, it makes little sense to assume that an agent will stick to its part of that behavior, if deviating from it can increase its payoff. This leads to much interest in the study of Nash equilibrium in games. When agents are allowed to use mixed strategies, Nash equilibrium always exists. However, Nash equilibrium does not take into account deviations by non-singleton sets of agents. While stability against deviations by subsets of the agents, captured by the notion of strong equilibrium [4], is a most natural requirement, it is well known that obtaining such stability is possible only in rare situations.[1]

In order to tackle this issue we consider in this paper the use of *mediators*. A mediator is a reliable entity that can interact with the players and perform on their behalf actions in a given game. However, a mediator cannot enforce behavior. Indeed, an agent is free to participate in the game without the help of the mediator. This notion is highly natural in a setting in which there exists some form of reliable party or administrator that is ready to serve as a mediator. For example, when Ebay is offering proxy services, it actually acts as a mediator and not only as an organizer. Notice that we assume that the multi-agent interaction formalized as a game is given, and that all the mediator can do is to communicate with the agents

☆ An extended abstract of this paper appears in the proceedings of the Twenty-First National Conference on Artificial Intelligence (AAAI-06). Almost all proofs are missing from the extended abstract. This version of the paper contains all of these missing proofs, and provides additional discussion and results. Furthermore, some of the definitions that appear in the extended abstract have been slightly modified. This work has been partially supported by the Israel Science Foundations (ISF).

* Corresponding author.
E-mail address: moshet@ie.technion.ac.il (M. Tennenholtz).

[1] For example, in the context of congestion games, Holzman and Law-Yone [13] characterized the networks where strong equilibrium always exist. They showed that strong equilibrium is guaranteed only in a very restricted type of networks.

0004-3702/$ – see front matter
doi:10.1016/j.artint.2008.10.005

and perform actions on behalf of the agents that allow it to do so. The mediator's behavior is pre-specified and depends on the messages received from all agents. This natural setting is different from the one discussed in the theory of mechanism design, where a designer designs a new game from scratch in order to yield some desired behavior.

Indeed, many markets employ very powerful forms of mediators like brokers, or routers in communication networks.[2] We find the notion of a mediator as central to the study of multi-agent systems. Indeed, while in economic theory, the dominant theme is that rational agents are to behave independently without any interference of a mediator, the (either explicit or implicit) existence of a party that provides suggestions, protocols, and rules of behavior has always been fundamental in the AI context of multi-agent systems (see [9,25] for some early introductions). As a result, in this paper we develop a rigorous study of mediators, aiming at the study of their use in establishing stability against deviations by coalitions.

A mediator for a given game is defined by sets of messages, one set for each player, and by an action function defined on vectors of messages; when a player sends a message to the mediator she gives the right to play to the mediator who will choose an action on her behalf (possibly by randomization) by applying the action function to the vector of messages sent to him. However, the mediator cannot enforce the players to use his services. The mediator generates a new game for the players, which we call the *mediated game*. In this game every player can either send a message to the mediator or play without the mediator. The outcome generated in the given game by an equilibrium in the mediated game is called a *mediated equilibrium*. An outcome generated by a strong equilibrium at the mediated game is called a *strong mediated equilibrium*. In an extreme case the message space of each player is a singleton. That is, this mediator accepts only one possible message: "I give you the right to play on my behalf". Such a mediator is called a *minimal mediator*. An important mediator is the one already developed in [32], where the set of messages of each player is the set of possible programs in a given programming language; in this case the action function is the one that executes the programs. Hence, program equilibrium is a particular type of mediated equilibrium. We further discuss the connections between mediators and the notion of program equilibrium in Section 10. In this paper we concentrate on the notion of strong mediated equilibrium.

In order to illustrate the power of a reliable mediators as discussed in this paper, consider the following simple example:

	Cooperate	Defect
Cooperate	4,4	0,6
Defect	6,0	1,1

In this classical Prisoners' dilemma game we get that in the unique equilibrium both agents will defect, yielding both of them a payoff of 1. However, this equilibrium, which is also a dominant strategy equilibrium, is inefficient; indeed, if both agents deviate from defection to cooperation then both of them will improve their payoffs. Formally, mutual defection is not a strong equilibrium.

Consider a reliable minimal mediator who offers the agents the following action function: if **both** agents give the mediator the right to play, he will perform cooperate on behalf of both agents. However, if only one agent agrees to give the right to play, the mediator he will perform defect on behalf of that agent. Hence, the mediator generates the following mediated game:

	Mediator	Cooperate	Defect
Mediator	4,4	6,0	1,1
Cooperate	0,6	4,4	0,6
Defect	1,1	6,0	1,1

[2] One interesting type of such markets is that of lottery syndicates. A lottery syndicate coordinates agents' activities in a lottery by trying to optimize the participants' joint actions. Such syndicates are known to be successful in the UK. It seems, however, that they are considered illegal at the US.

The mediated game has a most desirable property: it possesses a strong equilibrium; that is, an equilibrium which is stable against deviations by coalitions. In this equilibrium both agents will give the mediator the right to play, which will lead them to a payoff of 4 each! Hence, cooperation in the Prisoners' Dilemma game is a strong mediated equilibrium. In Sections 3 and 4 we explore general properties of mediators. Given the general concept of a mediator, we prove that mediators can indeed significantly increase the set of multi-agent encounters in which desired outcomes, which are stable against deviations by coalitions, can be obtained. We first prove that every two-person game possesses a strong mediated equilibrium, which also leads to optimal surplus. For general n-person games we prove that every *balanced* symmetric game possesses a strong mediated equilibrium, which also leads to optimal surplus. The precise definition of a balanced game is given in Section 7.1.[3] On an intuitive level, a game is balanced if there exists a profile of strategies yielding a payoff vector with the property that for each coalition of players their aggregate payoffs in this vector is at least as high as the aggregate payoff they can grantee themselves in the game by using a correlated strategy. For example, the Prisoner's Dilemma game discussed above is a balanced game. The profile of strategies (c, c) yields the payoff vector $(4, 4)$; No player can guarantees herself more than 4, and the coalition of the two players cannot guarantee itself more than 8.

In between equilibrium and strong equilibrium one can naturally define k-strong equilibrium as an outcome, which is immune to deviations of coalitions of size at most k. Indeed, if one considers the distributed computing and cryptography literature, it typically requires stability against deviations by up to k (typically faulty or malicious) agents, which can be viewed as a particular form of game-theoretic stability [17]. Similarly, we define the notion of k-strong mediated equilibrium. We show that in every symmetric game with n agents, if $k!$ divides n there exists a k-strong mediated equilibrium, leading to optimal surplus.[4]

We concentrate in this paper on the study of mediators within the classical NTU (non-transferable utility) model; however, we also extend our study, to a restricted case of the TU (transferable utility) model via the concept of an aggregate mediated equilibrium. In such an equilibrium, deviations which include re-distribution of payments, are taken into account. In fact, part of our study of strong aggregate mediated equilibrium serves as a technical tool for establishing the above mentioned results on the existence of strong mediated equilibrium in the NTU setting.

Finally, we want to report about two recent developments in the theory of mediators, which occurred after the publication of an extended abstract of this paper in the Twenty-First National Conference on Artificial Intelligence (AAAI-06). In [27] the concept of mediator has been generalized to allow it to condition its choices on the realizations of the actions of the players who do not use its services. Such a mediator is natural in routing systems, and has much more power than the mediator introduced in this paper. A mediator for games with incomplete information has been defined and analyzed in [3], where as an application the authors construct mediators for position auctions.[5]

We end this introduction with a discussion of a related literature.

A non-strategic model of mediation was introduced in [4] via the concept of c-acceptable strategies. This is an abstract notion that captures the "reasonable outcomes" obtained when subsets of the set of agents can correlate their activities. Our work introduces an explicit model of the mediation activity. In Section 5 we show that the introduction of an explicit mediator makes a difference; the outcomes that can be obtained using strong mediated equilibria constitute a subset of the set of outcomes that can be obtained using c-acceptable strategies.

The simplest form of mediator discussed in the game theory literature is captured by the notion of correlated equilibrium [6]. Indeed, since mediation via correlation device makes perfect sense from the CS/AI perspective, where protocols are typically recommended to the participants, this topic got considerable attention in the CS literature (e.g., [14,23]). This notion was generalized to communication equilibrium by [11,21]. Another, more powerful type of mediators is discussed in [19]. However, in all these settings the mediator can not perform actions on behalf of the agents that allow it to do so.

Situations where mediators can act on behalf of agents were discussed in the literature (see, e.g., [15]). However, these studies concentrated on 2-person games, and the central issue of stability against deviations by coalitions was not discussed in that literature. Mediators were also discussed in some restricted settings, under different titles; one interesting example is the study of bidding rings in auctions (see, e.g., [7,12,18]); here the bidding ring organizer can be viewed as a form of a mediator. A recent paper [31] implicitly discusses mediators via the notion of commitment – a commitment device serves as a mediator. This paper deals only with equilibrium (in contrast to strong equilibrium).

2. Games in strategic form: Strong equilibrium

Some notational preliminaries are needed. Let Y be a nonempty finite set. The set of probability distributions over Y is denoted by $\Delta(Y)$. That is, every $c \in \Delta(Y)$ is a function $c : Y \to [0, 1]$ such that $\sum_{y \in Y} c(y) = 1$. For every $y \in Y$ we denote by δ_y the probability distribution that assigns probability 1 to y. Let I be a nonempty finite set of indices, and let $c_i \in \Delta(Y_i)$, $i \in I$, where Y_i is a nonempty finite set for every $i \in I$. We denote by $\times_{i \in I} c_i$ the product probability distribution on $\mathbf{Y} = \times_{i \in I} Y_i$ that assigns to every $\mathbf{y} \in \mathbf{Y}$ the probability $\prod_{i \in I} c_i(y_i)$.

[3] As explained in Section 7.1 the title "balanced" is inherited from cooperative game theory.

[4] As an anecdote, the parliament in Israel contains $120 = 5!$ members. Hence, every anonymous game played by this Parliament possesses an optimal surplus symmetric 5-strong mediated equilibrium. While no parliament member is able to give the right of voting to a mediator, this right of voting may be replaced in real life by a commitment to follow the mediator's algorithm.

[5] That is, auctions similar to the ones used by Google to sell ads. See e.g., [33].

A game in strategic form is a tuple $\Gamma = \langle N, (X_i)_{i\in N}, (u_i)_{i\in N}\rangle$, where N is a nonempty finite set of players, X_i is the nonempty strategy set of player i, and $u_i : \mathbf{X} \to \mathbb{R}$ is the payoff function for player i, where $\mathbf{X} = \times_{i\in N} X_i$. If $|N| = n$, whenever convenient we assume $N = \{1, \dots, n\}$. Γ is *finite* if the strategy sets are finite. A nonempty subset of the players is also called a *coalition*.

Let $\Gamma = \langle N, (X_i)_{i\in N}, (u_i)_{i\in N}\rangle$ be a finite game. For every coalition S we denote $\mathbf{X}_S = \times_{i\in S} X_i$. Note that $\mathbf{X}_N = \mathbf{X}$ and $\mathbf{X}_{\{i\}} = X_i$. When the set N is clear, we will denote the complement of S in N by S^c or by $-S$. Thus, $\mathbf{X}_{N\setminus S}$ will be also denoted by $\mathbf{X}_{S^c}$ or by $\mathbf{X}_{-S}$, and moreover, $\mathbf{X}_{-\{i\}}$ will be also denoted by $\mathbf{X}_{-i}$. Let $S \subseteq N$ be a coalition. Every $c \in \Delta(\mathbf{X}_S)$ is called a *correlated strategy for S*. A correlated strategy for the set of all players N is also called a *correlated strategy*, and for every i, a correlated strategy for $\{i\}$ is also called a *mixed strategy* for i. The expected payoff of i with respect to a correlated strategy c is denoted by $U_i(c)$. That is,

$$U_i(c) = \sum_{\mathbf{x}\in\mathbf{X}} u_i(\mathbf{x})c(\mathbf{x}).$$

For every S, the set of mixed-strategy profiles is denoted by $\mathbf{Q}_S$. That is, $\mathbf{Q}_S = \times_{i\in S} \Delta(X_i)$. We will use $\mathbf{Q}$ for $\mathbf{Q}_N$, and Q_i for $\mathbf{Q}_{\{i\}}$.

The *mixed extension* of the game Γ is the game $\langle N, (Q_i)_{i\in N}, (w_i)_{i\in N}\rangle$, where for every $\mathbf{q} \in \mathbf{Q}$, $w_i(\mathbf{q}) = U_i(q_1 \times \cdots \times q_n)$.

2.1. Three definitions of strong equilibrium

In general, a strong equilibrium in a game is a profile of strategies with the property that no coalition has a beneficial deviation for its members. The above verbal definition leaves a lot of modeling choices. What is a profile of strategies? (pure, mixed, or correlated), what is a deviation? (pure, mixed or correlated), and what is the meaning of "beneficial for its members"? – is a deviation beneficial only when every member is better off, or is it sufficient that some of the members are better off and the others are not worse off? or maybe it is just required that the aggregate payoffs of the players is better. Below we will define three notions of strong equilibrium, and discus their rationale:

Let $\Gamma = \langle N, (X_i)_{i\in N}, (u_i)_{i\in N}\rangle$ be a game in strategic form.

Definition 1. *Let* $\mathbf{x} \in \mathbf{X}$. $\mathbf{x}$ is a *strong equilibrium of type I* in Γ if the following holds:
For every coalition S, and for every $\mathbf{y}_S \in \mathbf{X}_S$ there exists $j \in S$ such that

$$u_j(\mathbf{y}_S, \mathbf{x}_{-S}) \leqslant u_j(\mathbf{x}).$$

Hence, in a strong equilibrium of type I the suggested profile of behavior and the possible profiles of deviations consist of pure strategies, and a deviation is beneficial to a coalition if it is beneficial to each of its members.

All other definitions require that the game Γ is finite because we defined the notion of mixed strategy and correlated strategy only for finite games.

Definition 2. Let $\mathbf{q} = (q_1, \dots, q_n)$ be a profile of mixed strategies. We say that $\mathbf{q}$ is a *strong equilibrium of type* II in Γ if $\mathbf{q}$ is a strong equilibrium of type I in the mixed extension of Γ.

That is, $\mathbf{q}$ is a strong equilibrium of type II in Γ if for every coalition S, and for every profile of mixed strategies $(p_i)_{i\in S}$ there exists $j \in S$ such that $U_j(\times_{i\in S}\, p_i \times_{i\in N\setminus S} q_i) \leqslant U_j(\times_{i\in N}\, q_i)$. Obviously, if $\mathbf{x} \in \mathbf{X}$, and $(\delta_{x_1}, \dots, \delta_{x_n})$ is a strong equilibrium of type II in Γ then $\mathbf{x}$ is a strong equilibrium of type I, but the converse does not hold. Whenever $(\delta_{x_1}, \dots, \delta_{x_n})$ is a strong equilibrium of type II we abuse notations and allow ourselves to say that $\mathbf{x}$ is a strong equilibrium of type II in Γ.

Definition 3. Let $\mathbf{q} \in \mathbf{Q}$. We say that $\mathbf{q}$ is a *strong equilibrium of type* III, if for every coalition S, and for every correlated strategy for S, $c_S \in \Delta(\mathbf{X}_S)$ there exists $j \in S$ such that

$$U_j\big(c_S \times (\times_{i\in S^c}\, q_i)\big) \leqslant U_j(q_1 \times q_2 \times \cdots \times q_n).$$

Let $\mathbf{q} \in \mathbf{Q}$. Obviously $\mathbf{q}$ is a strong equilibrium of type II if $\mathbf{q}$ is a strong equilibrium of type III, but not vice versa.

The requirement that $\mathbf{q}$ is a strong equilibrium of type II seems to be acceptable in an environment in which the players believe that they and others could not possibly correlate their behavior (e.g., when every player is sitting in a separate room, and there is no communication between the players). However, every player can perform a private randomization. In an environment in which the players do not correlate their strategies, but they may fear/hope that such a correlation is possible, we expect $\mathbf{q}$ to be a strong equilibrium of type III in order to be believed/played by the players. At Section 7 we define a fourth notion of strong equilibrium.

3. Strong mediated equilibrium

We now introduce mediators, a general tool for coordinating and influencing agents' behavior in games. A mediator is always assumed to be reliable. However, mediators are classified according to their abilities to interfere in the game.

In this paper we endow the mediator with the ability to play for the players who give him the right to play for them. However, the mediator cannot enforce the players to use his services.

Definition 4. Let Γ be a finite game in strategic form. A *mediator* for Γ is a tuple $((M_i)_{i\in N}, \mathbf{c} = (\mathbf{c}_S)_{\emptyset\neq S\subseteq N})$, where each M_i is a finite set, $M_i \cap X_i = \emptyset$ for every player i, and for every coalition S, $\mathbf{c}_S : \mathbf{M}_S \to \Delta(\mathbf{X}_S)$.

In the above definition M_i is a set of *messages* that may be sent by agent i to the mediator. Agent i may either participate in the game directly or participate in the game using the mediator's services by sending him any message from M_i. The *action function* of the mediator is $\mathbf{c} = (\mathbf{c}_S)_{\emptyset\neq S\subseteq N}$; If the set of players that send messages to the mediator is S, and the members of S send the vector of messages $\mathbf{m}_S = (m_i)_{i\in S}$, the mediator plays in behalf of the members of S the correlated strategy for S, $\mathbf{c}_S(\mathbf{m}_s)$. That is, the mediator chooses a profile of strategies $\mathbf{x}_S$ according to the probability distribution $\mathbf{c}_S(\mathbf{m}_S)$, and plays x_i on behalf of every player $i \in S$.

Every mediator $\mathcal{M}$ for Γ defines a finite game in strategic form, which we call the *mediated game* and denote by $\Gamma(\mathcal{M})$. In the mediated game, the strategy set of player i is $Z_i = X_i \cup M_i$, and the payoff function of i is defined for every $\mathbf{z} \in \mathbf{Z}$ as follows:

$$u_i^{\mathcal{M}}(\mathbf{z}) = U_i\big(\mathbf{c}_{T_\mathbf{z}}(\mathbf{z}_{T_\mathbf{z}}) \times (\times_{j\in N\setminus T_\mathbf{z}}\delta_{z_j})\big),$$

where $T_\mathbf{z} = \{j \in N \mid z_j \in M_j\}$. That is, $T_\mathbf{z}$ is the set of players who use the service of the mediator.[6]

Definition 5. Let Γ be a game in strategic form. A correlated strategy $c \in \Delta(\mathbf{X})$ is a *strong mediated equilibrium* if there exists a mediator for Γ, $\mathcal{M}$, and a vector of messages $\mathbf{m} \in \mathbf{M} = \times_{i\in N} M_i$, with $\mathbf{c}_N(\mathbf{m}) = c$, such that $\mathbf{m}$ is a strong equilibrium of type III in $\Gamma(\mathcal{M})$; Such a mediator is said to *strongly implement c*.

We now define a type of minimal mediators that will play an important role in our subsequent analysis.

Definition 6. Let Γ be a game in strategic form. A mediator $\mathcal{M} = ((M_i)_{i\in N}, (\mathbf{c}_S)_{\emptyset\neq S\subseteq N})$ is *minimal* if each message space is a singleton.

Consider a minimal mediator in which $M_i = \{r_i\}$ for every player i. Let $\mathbf{r} = (r_1, \ldots, r_n)$. When the players are using this mediator each of them can either give the right to play to the mediator (sending r_i) or play independently. If the coalition that gives the right to play is T, the mediator uses the correlated strategy $c_T = \mathbf{c}_T(\mathbf{r}_T)$ in order to play for T. Hence, every minimal mediator is uniquely defined by a vector of correlated strategies, $(c_S)_{\emptyset\neq S\subseteq N}$, one for each coalition. This minimal mediator strongly implements c if $\mathbf{r}$ is a strong mediated equilibrium at the mediated game and $c_N = c$. As it turns out, restricting our attention to minimal mediators does not cause any loss of strong mediated equilibria:

Lemma 1. *Let Γ be a finite game in strategic form. Every strong mediated equilibrium in Γ can be implemented by a minimal mediator.*

Proof. Let $c \in \Delta(\mathbf{X})$ be a strong mediated equilibrium, and let $\mathcal{M} = ((M_i)_{i\in N}, (\mathbf{c}_S)_{\emptyset\neq S\subseteq N})$ implement c by the profile $\mathbf{m} \in \mathbf{M}$. That is, $\mathbf{m}$ is a strong equilibrium of type III in $\Gamma(\mathcal{M})$, and $\mathbf{c}_N(\mathbf{m}) = c$. Define a minimal mediator, $\mathcal{K}$, in which the set of messages for every i is $K_i = \{m_i\}$. The implementing functions are for every coalition S the restriction of $\mathbf{c}_S$ to $\{\mathbf{m}_S\}$. That is, the minimal mediator is $\mathcal{K} = (\mathbf{c}_S(\mathbf{m}_S))_{\emptyset\neq S\subseteq N}$. Hence, $\Gamma(\mathcal{K})$ is obtained from $\Gamma(\mathcal{M})$ by restricting the strategy set of every player i from $X_i \cup M_i$ to $X_i \cup \{m_i\}$. Therefore, $\mathbf{m}$ remains a strong equilibrium of type III in $\Gamma(\mathcal{K})$. $\square$

4. Properties of strong mediated equilibria

For further analysis it is convenient to notice the following properties of mediators.

Every correlated strategy in $\Gamma(\mathcal{M})$, $\xi \in \Delta(\mathbf{Z})$ induces a correlated strategy c_ξ in Γ: for every $\mathbf{x} \in \mathbf{X}$ we have

$$c_\xi(\mathbf{x}) = \sum_{S\subseteq N} \sum_{\mathbf{m}_{-S}\in\mathbf{M}_{-S}} \xi(\mathbf{x}_S, \mathbf{m}_{-S})\mathbf{c}_{-S}(\mathbf{m}_{-S})(\mathbf{x}_{-S}). \tag{1}$$

[6] If $T_\mathbf{z} = \emptyset$ the symbol $\mathbf{c}_{T_\mathbf{z}}(\mathbf{z}_{T_\mathbf{z}})$ is not defined. However, here and in other cases the meaning of such expressions is obvious, and we skip their details for ease of exposition since they can be easily understood by the reader. Here, of course, if $T_\mathbf{z} = \emptyset$, $u_i^{\mathcal{M}}(\mathbf{z}) = U_i(\times_{j\in N}\delta_{z_j}) = u_i(\mathbf{z})$. Similarly, if $T_\mathbf{z} = N$, $u_i^{\mathcal{M}}(\mathbf{z}) = c_N(\mathbf{z})$.

If M is a minimal mediator, (1) has a simpler form:

$$c_\xi(\mathbf{x}) = \sum_{S \subseteq N} \xi(\mathbf{x}_S, \mathbf{r}_{-S}) c_{-S}(\mathbf{x}_{-S}). \tag{2}$$

Hence, in the mediated game:

$$U_i^{\mathcal{M}}(\xi) = U_i(c_\xi). \tag{3}$$

Similarly, for every coalition T, every $\xi_T \in \Delta(\mathbf{Z}_T)$ define $c_{\xi_T} \in \Delta(\mathbf{X}_T)$ as follows:

$$c_{\xi_T}(\mathbf{x}_T) = \sum_{A \subseteq T} \sum_{\mathbf{m}_{T \setminus A} \in \mathbf{M}_{T \setminus A}} \xi_T(\mathbf{x}_A, \mathbf{m}_{T \setminus A}) \mathbf{c}_{T \setminus A}(\mathbf{m}_{T \setminus A})(\mathbf{x}_{T \setminus A}), \quad \mathbf{x}_T \in \mathbf{X}_T. \tag{4}$$

And for a minimal mediator:

$$c_{\xi_T}(\mathbf{x}_T) = \sum_{A \subseteq T} \xi_T(\mathbf{x}_A, \mathbf{r}_{T \setminus A}) c_{T \setminus A}(\mathbf{x}_{T \setminus A}), \quad \mathbf{x}_T \in \mathbf{X}_T. \tag{5}$$

The next proposition shows that by using mediators we don't lose any outcome that can be obtained in a strong equilibrium of the original game:

Proposition 1. *Let Γ be a finite game in strategic form, and let $\mathbf{q}$ be a profile of mixed strategies, which is a strong equilibrium of type* III *in Γ. Then, $q_1 \times q_2 \times \cdots \times q_n$ is a strong mediated equilibrium in Γ.*

Proof. We define a minimal mediator $\mathcal{M} = (c_S)_{\emptyset \neq S \subseteq N}$ as follows: $c_S = \times_{i \in S} q_i$ for every coalition $S \subseteq N$. Let T be a coalition and let $\xi_T \in \Delta(\mathbf{Z}_T)$. We have to show that ξ_T is not a profitable deviation for its members. Let $\xi = \xi_T \times (\times_{i \in N \setminus T} \delta_{r_i}) \in \Delta(\mathbf{Z})$. We have to show that there exists $i \in T$ for which

$$U_i^{\mathcal{M}}(\xi) \leqslant u_i^{\mathcal{M}}(r_1, r_2, \ldots, r_n).$$

That is,

$$U_i(c_\xi) \leqslant U_i(q_1 \times q_2 \times \cdots \times q_n). \tag{6}$$

We will show that

$$c_\xi = c_{\xi_T} \times c_{-T}. \tag{7}$$

Because $\mathbf{q}$ is a strong equilibrium of type III in Γ and $c_{-T} = \times_{j \in N \setminus T} q_j$, (7) implies that there exists $i \in T$ for which (6) holds.

In order to prove (7) note that for every $\mathbf{z} \in \mathbf{Z}$, $\xi(\mathbf{z}) = 0$ unless $\mathbf{z}_{-T} = r_{-T}$. As the mediator is minimal we get from (2), that for every $\mathbf{x} \in \mathbf{X}$,

$$c_\xi(\mathbf{x}) = \sum_{S \subseteq T} \xi(\mathbf{x}_S, \mathbf{r}_{-S}) c_{-S}(\mathbf{x}_{-S}) = \sum_{S \subseteq T} \xi_T(\mathbf{x}_S, \mathbf{r}_{T \setminus S}) c_{-S}(\mathbf{x}_{-S}).$$

However, for every $S \subseteq T$, $c_{-S}(\mathbf{x}_{-S}) = c_{T \setminus S}(\mathbf{x}_{T \setminus S}) c_{-T}(\mathbf{x}_{-T})$. Therefore, by (5)

$$c_\xi(\mathbf{x}) = \left[\sum_{S \subseteq T} \xi_T(\mathbf{x}_S, \mathbf{r}_{T \setminus S}) c_{T \setminus S}(\mathbf{x}_{T \setminus S}) \right] c_{-T}(\mathbf{x}_{-T}) = c_{\xi_T}(\mathbf{x}_T) c_{-T}(\mathbf{x}_{-T}) = (c_{\xi_T} \times c_{-T})(\mathbf{x}). \qquad \square$$

We next show that a new mediator for the game generated by a mediator cannot add acceptable outcomes. That is, a mediator can not help in a situation where a mediator is already in place, beyond what can be obtained by the existing mediator.

Proposition 2. *Let Γ be a game in strategic form, and let $\mathcal{M} = ((M_i)_{i \in N}, (\mathbf{c}_S)_{\emptyset \neq S \subseteq N})$ be a mediator for Γ. If ξ is a strong mediated equilibrium in $\Gamma(\mathcal{M})$, then c_ξ is a strong mediated equilibrium in Γ.*

Proof. Let $\mathcal{M}'$ be a mediator for $\Gamma(\mathcal{M})$ that implement ξ. By Lemma 1 we can assume that $\mathcal{M}'$ is a minimal mediator, say $\mathcal{M}' = (\xi_S)_{\emptyset \neq S \subseteq N}$, and therefore $\xi_N = \xi$. The action of i of giving the right to play to the mediator $\mathcal{M}'$ is denoted by r_i. We define a minimal mediator for Γ, $\mathcal{M}^* = (c_{\xi_S})_{\emptyset \neq S \subseteq N}$. In $\Gamma(\mathcal{M}^*)$ the action of i of giving the right to play to this mediator is also denoted by r_i. We have to show that $\mathbf{r} = (r_1, r_2, \ldots, r_n)$ is a strong equilibrium of type III in $\Gamma(\mathcal{M}^*)$. Let T be a coalition, and let ψ_T be a correlated strategy of T at the game $\Gamma(\mathcal{M}^*)$. Hence, $\psi_T \in \Delta(\times_{i \in T}(X_i \cup \{r_i\}))$. Define $\hat{\psi}_T \in \Delta(\times_{i \in T}(X_i \cup M_i \cup \{r_i\}))$ as follows: $\hat{\psi}_T$ is concentrated on $\times_{i \in T}(X_i \cup \{r_i\})$, and on this set it is defined as ψ_T. Let $\tilde{\xi}$ be the correlated strategy in $\Gamma(\mathcal{M})$ induced by $\hat{\psi}_T \times (\times_{i \in N \setminus T} \delta_{r_i})$. It is easily verified that:

$$\tilde{\xi}(x_A, \mathbf{m}_A) = \sum_{S \subseteq A \cap T} \psi_T(\mathbf{x}_S, r_{T \setminus S}) \xi_{-S}(\mathbf{x}_{A \setminus S}, \mathbf{m}_{-A}).$$

From this, by (1) it can be easily verified that

$$c_{\tilde{\xi}} = c_{\psi_T \times (\times_{i \in N \setminus T} \delta_{r_i})}. \tag{8}$$

As a deviation to $\hat{\psi}_T$ is not profitable to its members at the game $\Gamma(\mathcal{M})(\mathcal{M}')$, there exists a player $i \in T$ such that $U_i(c_{\tilde{\xi}}) \leqslant U_i(c_\xi)$. Therefore, by equality (8), ψ_T is not a profitable deviation from $\mathbf{r}$ at $\Gamma(\mathcal{M}^*)$. $\square$

When defining strong mediated equilibrium we focus on a particular type of pure strategy strong equilibrium of type III in the mediated game. However, the game generated by the mediator may give rise to other possibilities of forming a strong equilibrium of type III. It is therefore important to know that our seemingly restricted definition does not restrict the possible acceptable outcomes. Indeed, we show:

Proposition 3. *Let $\mathcal{M} = ((M_i)_{i\in N}, (\mathbf{c}_S)_{\emptyset \neq S \subseteq N})$ be a mediator for Γ, and let $\hat{q}$ be a vector of mixed strategies in $\Gamma(\mathcal{M})$, which is a strong equilibrium of type* III *in $\Gamma(\mathcal{M})$. Then $c_{\hat{q}_1 \times \hat{q}_2 \times \cdots \times \hat{q}_n}$ is a strong mediated equilibrium in Γ.*

Proof. Because $\hat{q}$ is a strong equilibrium of type III at $\Gamma(\mathcal{M})$ then by Proposition 1 $\hat{q}_1 \times \hat{q}_2 \times \cdots \times \hat{q}_n$ is a strong mediated equilibrium at $\Gamma(\mathcal{M})$. Therefore, by Proposition 2, $c_{\hat{q}_1 \times \hat{q}_2 \times \cdots \times \hat{q}_n}$ is a strong mediated equilibrium in Γ. $\square$

Furthermore, when there is one mediator, other mediators may show up. Any set of mediators $\mathcal{H} = \{\mathcal{M}_1, \ldots, \mathcal{M}^k\}$ generates a game in strategic form $\Gamma(\mathcal{H})$ in which every player can choose any mediator in $\mathcal{H}$ she wishes, and give this mediator the right to play by sending him a message, or play independently. If ξ is a correlated strategy in $\Gamma(\mathcal{H})$ we denote by c_ξ the correlated strategy in Γ generated by ξ. The next proposition shows that the existence of many mediators that the agents can approach does not help beyond the use of a single mediator.

Proposition 4. *Let $\mathcal{H}$ be a set of mediators, and let $\bar{\mathbf{q}}$ be a strong equilibrium of type* III *in $\Gamma(\mathcal{H})$, then $c_{\bar{q}_1 \times \bar{q}_2 \times \cdots \times \bar{q}_n}$ is a strong mediated equilibrium in Γ.*

Proof. One can naturally define a mediator $\mathcal{M}$ which will mimic the strategic possibilities given by the set of mediators: The message sets in $\mathcal{M}$ are just $M_i = \bigcup_{j=1}^k M_i^j$ for every $i \in N$. Every message $\mathbf{m} \in \mathbf{M}_S$ has the form $\mathbf{m}_S = (\mathbf{m}_{A^1}, \ldots, \mathbf{m}_{A^k})$, where A^j is the set of all players in S that choose a message in $\mathbf{M}^j$. Hence, $S = \bigcup_{j=1}^k A^j$. Define $\mathbf{c}_S$ as follows:

$$\mathbf{c}_S(\mathbf{m}_S) = \mathbf{c}_{A^1}(\mathbf{m}_{A^1}) \times \cdots \times \mathbf{c}_{A^k}(\mathbf{m}_{A^k}).$$

It is obvious that $\Gamma(\mathcal{M})$ is strategically equivalent to $\Gamma(\mathcal{H})$, and therefore $\bar{\mathbf{q}}$ is a strong equilibrium of type III at $\Gamma(\mathcal{M})$. Therefore, by Proposition 3 $c_{\bar{q}_1 \times \bar{q}_2 \times \cdots \times \bar{q}_n}$ is a strong mediated equilibrium in Γ. $\square$

5. C-acceptable correlated strategies and strong mediated equilibrium

As mentioned, the main motivation for this work comes from the desire to establish multi-agent behaviors that are stable against deviations by coalitions, using mediators which can offer their services. Hence, a major point of our study is in showing that such mediators are indeed helpful. However, when establishing such a theory, it is important to understand how our study of mediators fits relevant previous foundational work in game theory. Aumann (see [4]) defined **c-acceptable correlated strategies**, which may seem at first glance to implicitly catch the idea of defining the "reasonable outcomes" that can be obtained when agents can correlate their strategies.

Definition 7. A correlated strategy c is *c-acceptable* if there exists a vector $(c_S)_{\emptyset \neq S \subseteq N}$ with $c_N = c$ such that for every coalition S and for every $d_S \in \Delta(X_S)$, there exists $i \in S$ such that

$$U_i(d_S \times c_{-S}) \leqslant U_i(c).$$

It is easy to show:

Proposition 5. *Every strong mediated equilibrium is c-acceptable.*

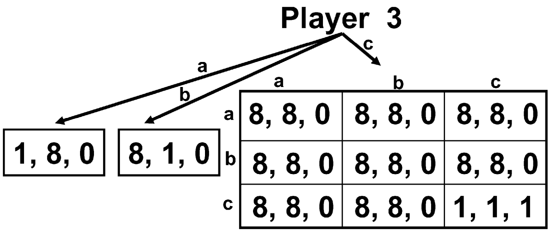
Proof. Let c be a strong mediated equilibrium. Therefore there exists a minimal mediator, $(c_S)_{\emptyset \neq S \subseteq N}$ that implements c. Let T be a coalition. By Definition 5, no deviation of T to $\xi_T \in \Delta(\mathbf{Z}_T)$ is profitable for its members, when T^c uses c_{T^c}. As X_T is a subset of Z_T, it is clear that no deviation of T to $c_T \in \Delta(X_T)$ is profitable for its members. $\square$

However, as we will show, the converse of the above result is not true, for illuminating reasons. One should notice that all that we used in the above proof was the fact that a strong mediated equilibrium is immune against deviations by correlating strategies of the original game. But the concept of strong mediated equilibrium requires also that it is immune against correlation over messages and actions. Hence, the converse to Proposition 5 is not clear. Indeed we are about to prove that the converse is not true. A c-acceptable strategy used by a mediator is not necessary immune against *Trojan horses*: A coalition of players may correlate in such a way that in some realizations a subgroup is pretending to cooperate by sending the right messages. This may be beneficial because these Trojan horses will be part of the punishing group, and they may get a very big payoff for this!

More precisely, the mediator can only see who has chosen its services, and can not distinguish between an agent that just asks for its services to one that asks for its services as part of a sophisticated deviation in which it does so only in particular instances. It turns out that the ability to have such sophisticated deviations, where the mediator can not tell who are the deviators is indeed meaningful, as captured by the following theorem:

Theorem 1. *There exists a game Γ, and a c-acceptable strategy c, which is not a strong mediated equilibrium.*

Proof. Consider the following 3-person game:



The strategy set of each player i is $\{a, b, c\}$. If Player 3 chooses a the resulting payoff vector is $(1, 8, 0)$ regardless of the actions chosen by 1 and 2. Similarly, when 3 chooses b, the resulting payoff is $(8, 1, 0)$. When 3 chooses c, the resulting payoff matrix is the 3×3 matrix in the figure.

We first show that the correlated strategy that is concentrated on (c, c, c) is c-acceptable. We define a vector $(\eta_S)_{S \subseteq \{1,2,3\}}$, where $\eta_{\{1,2,3\}} = \delta_{(c,c,c)}$ that satisfies the conditions for $\delta_{(c,c,c)}$ to be c-acceptable.

As almost all correlated strategies discussed in this example are concentrated on pure strategy profiles, we will abuse notations and for every S and every $\mathbf{x}_S \in \Delta(\mathbf{X}_S)$ we will identify $\mathbf{x}_S$ with the correlated strategy $\delta_{\mathbf{x}_S}$

Indeed, let $\eta_{\{1,2\}} = (a, a)$, $\eta_{\{1,3\}} = (a, b)$, $\eta_{\{2,3\}} = (a, a)$, $\eta_{\{3\}} = a$, $\eta_{\{1\}} = a$, and $\eta_{\{2\}} = a$. It is easily checked that these are punishing strategies that do the job. That is, no deviating group can ensure more than 1 to each of its members. We proceed to show that (c, c, c) is not a strong mediated equilibrium. Assume for contradiction that it is a strong mediated equilibrium, and let $(c_S)_{S \subset N}$ be a minimal mediator that implements it. In particular, $c_N = (c, c, c)$. Consider $c_{\{2,3\}}$. Using this strategy by 2 and 3 must guarantee that player 1 does not get more than 1. Therefore this correlated strategy must assign a probability 1 to player 3 playing a, since otherwise a deviation of player 1 to (say) a would yield him a strict convex combination of 8 and 1. Similarly, in $c_{\{1,3\}}$ player 3 must play b with probability 1. We construct a profitable deviation for $\{1, 2\}$. Let 1 and 2 randomize with equal probability between the two options: "1 plays a and 2 give the right of play to the mediator", and "2 plays a and 1 give the right of play of the mediator". Given the properties of $\eta_{\{1,3\}}, \eta_{\{2,3\}}$ described above, this deviation will give each of them an expected payoff of 4.5. This contradicts the assumption that (c, c, c) is a strong mediated equilibrium. Therefore (c, c, c) is not a strong mediated equilibrium. □

5.1. The β-core and the mediated core

We need the following notations: For every correlated strategy c (a strategy profile $\mathbf{x}$) we denote $\mathbf{U}(c) = (U_1(c), \dots, U_n(c))$ ($\mathbf{u}(\mathbf{x}) = (u_1(\mathbf{x}), \dots, u_n(\mathbf{x}))$). A vector $\mathbf{w} \in \mathbb{R}^n$ will be called a *payoff vector*; $\mathbf{w}$ is called a *feasible payoff vector* for Γ if there exists a correlated strategy d with $\mathbf{U}(d) = \mathbf{w}$. That is, $\mathbf{w}$ is feasible if it belongs to the convex hull of all payoff vectors of the form $\mathbf{u}(\mathbf{x})$, $\mathbf{x} \in \mathbf{X}$. The proof of the following important observation follows directly from the definition of c-acceptable strategies and strong mediated equilibrium, respectively:

Observation 1.

1. *If c is c-acceptable and $\mathbf{U}(d) = \mathbf{U}(c)$ then d is also c-acceptable.*
2. *If c is a strong mediated equilibrium and $\mathbf{U}(d) = \mathbf{U}(c)$ then d is also a strong mediated equilibrium.*

Hence, the question of whether c is a strong mediated equilibrium or c-acceptable depends only on the payoff vector $\mathbf{U}(c)$. Therefore, it is natural to define the concept of "acceptable" payoff vectors. This was done in [5] for the concept of "c-acceptable" via the notion of the β-core of a game in strategic form, as follows:

Definition 8. Let Γ be a game in strategic form. The *β-core* of Γ, $C_\beta(\Gamma)$ is the set of all feasible payoff vectors, $\mathbf{w} = (w_1, \ldots, w_n)$, for which there exists a c-acceptable correlated strategy c such that $\mathbf{U}(c) = \mathbf{w}$. That is, $w_i = U_i(c)$ for every $i \in N$.

Similarly, we define:

Definition 9. The *mediated core* of Γ, $C_m(\Gamma)$, is the set of all feasible payoff vectors, $\mathbf{w}$, for which there exists a strong mediated equilibrium c such that $\mathbf{w} = \mathbf{U}(c)$.

Proposition 5 implies that $C_m(\Gamma) \subseteq C_\beta(\Gamma)$. Considering again the game described in the proof of Theorem 1, as we showed that (c, c, c) is not a strong mediated equilibrium, then by Part 2 in Observation 1, $(1, 1, 1) = \mathbf{u}(c, c, c)$ is not in the mediated core. Therefore we have:

Corollary 1. *For every game in strategic form, Γ, the mediated core is contained in the β-core. That is,*

$$C_m(\Gamma) \subseteq C_\beta(\Gamma).$$

Moreover, there exists a game Γ for which a strict inclusion holds.

For completeness, we end this section with an analysis of the mediated core of the game described in Theorem 1. Actually, it is easy to notice that the correlated strategy, c, which randomizes with equal probabilities between (c, c, c) and (a, a, c) is a strong mediated equilibrium in Γ, in which 1, and 2 get each 4.5, and 3 gets 0.5. Moreover, the mediated core of this game is a singleton, that is $C_m(\Gamma) = \{(4.5, 4.5, 0.5)\}$.

6. Existence: Two-person games

In this section we illustrate the power of mediators, by showing that every two-person game in strategic form possesses a strong mediated equilibrium. Other existence results for general classes of games are presented in Sections 8 and 9.

We need some notation: A correlated strategy c is *Pareto optimal* if for every correlated strategy d there exists $i \in N$ such that $U_i(d) \leqslant U_i(c)$. A feasible payoff vector, $\mathbf{w}$ is *Pareto optimal* if there exists a Pareto optimal strategy c with $\mathbf{U}(c) = \mathbf{w}$.

Obviously Pareto optimality is a necessary condition on c to be a strong mediated equilibrium, as well as to be c-acceptable. It just says that the set of **all** players does not have a profitable deviation for its members. Consequently, every payoff vector in the mediated core or in the β-core is Pareto optimal.

In [5] it was shown that every 2-person game has a c-acceptable strategy. Since the Trojan horse effect, which worked in the proof of Theorem 1, cannot hold with only two players, we have:

Proposition 6. *In a two-person game, every c-acceptable strategy is a strong mediated equilibrium. Consequently, every 2-person game has a strong mediated equilibrium.*

Proof. Let Γ be a two-person game in strategic form, and let c be a c-acceptable correlated strategy in Γ. Denote $c_{\{1,2\}} = c$. By Definition 7 there exist correlated strategies for i, c_i, $i = 1, 2$ such that $(c_1, c_2, c_{\{1,2\}})$ has the following three properties:

$$U_i(d_i) \leqslant U_i(c_i) \quad \text{for every } d_i \in \Delta(X_i),\ i = 1, 2.$$

For every $d \in \Delta(X)$ there exists $i \in \{1, 2\}$ with $U_i(d) \leqslant U_i(c)$.

However, the three properties above are precisely the properties that are required to guarantee that c is a strong mediated equilibrium implemented by the minimal mediator $(c_1, c_2, c_{\{1,2\}})$. Therefore, c is a strong mediated equilibrium. $\square$

7. Aggregate deviations

In this section we develop a theory of another type of implementation by mediators, for settings in which players consider deviations with re-distribution of payments, and discuss its relationships with the previous sections. In the following sections we use the results of this section to prove additional existence theorems for the standard strong mediated equilibrium. However, the theory developed here is interesting for its own.

We begin with a fourth notion of strong equilibrium for games in strategic form. Let Γ be a game in strategic form. For every $\mathbf{x} \in \mathbf{X}$, and for every $S \subseteq N$ let $u_S(\mathbf{x}) = \sum_{i \in S} u_i(\mathbf{x})$. Similarly, for a correlated strategy c we denote

$$U_S(c) = \sum_{i \in S} U_i(c).$$

We say that a correlated strategy c is *surplus optimal* if $\max_{d \in \Delta(\mathbf{X})} U_N(d)$ is attained at c. Obviously, every surplus optimal correlated strategy is also Pareto optimal. Also, one can always find a surplus optimal correlated strategy, which is concentrated on a pure strategy profile.

Definition 10. Let $\mathbf{q} \in \mathbf{Q}$. We say that $\mathbf{q}$ is a *strong equilibrium of type* IV, if for every coalition S, and for every correlated strategy for S, $c_S \in \Delta(\mathbf{X}_S)$,

$$U_S\big(c_S \times (\times_{i \in S^c} q_i)\big) \leqslant U_S(q_1 \times q_2 \times \cdots \times q_n).$$

The requirement that $\mathbf{q}$ is a strong equilibrium of type IV seems to be acceptable in an environment in which the players do not correlate their strategies, and do not redistribute their payoffs, but they may fear/hope that such a correlation and re-distribution is possible. The following is a simple observation:

Observation 2. *Given a game in strategic form, every strong equilibrium of type* IV *in the game is also a strong equilibrium of type* III.

When players consider aggregate deviations as possible, a mediator must make sure that sending him the right messages forms a strong equilibrium of type IV at the mediated game. We therefore define:

Definition 11. Let Γ be a game in strategic form. A correlated strategy $c \in \Delta(\mathbf{X})$ is an *aggregate mediated equilibrium* if there exists a mediator for Γ, $\mathcal{M}$, and a vector of messages $\mathbf{m} \in \mathbf{M}$, with $\mathbf{c}_N(\mathbf{m}) = c$, such that $\mathbf{m}$ is a strong equilibrium of type IV in $\Gamma(\mathcal{M})$. Such a mediator is said to *implement* c with aggregate deviations.

Lemma 1 and Propositions 1, 2, 3, and 4 are true when "strong equilibrium of type III" is replaced with "strong equilibrium of type IV", "strong mediated equilibrium" is replaced with "aggregate mediated equilibrium", and "strongly implements" is replaced with "implements with aggregate deviations". The following is a simple observation:

Observation 3. *Every aggregate mediated equilibrium is a strong mediated equilibrium.*

Observation 1 also holds for the concept of aggregate mediated equilibrium. Hence, we also define the aggregate mediated core as follows:

Definition 12. The *aggregate mediated core* of Γ, $C_{ag-m}(\Gamma)$ is the set of all feasible payoff vectors $\mathbf{w}$, for which there exists an aggregate mediated equilibrium c such that $\mathbf{w} = \mathbf{U}(c)$.

We now define the classical concepts[7] of TU-acceptable strategies, and the TU-core, which will be related to aggregate mediated equilibrium and aggregate mediated core, respectively. However, we define these concepts in the spirit of previous definitions in this paper, rather than in the equivalent classical way, which associates a TU-cooperative game with each game in strategic form. In the next sub-section we will provide also the classical approach, which is useful.

Definition 13. Let Γ be a game in strategic form, and let c be a correlated strategy. We say that c is *TU-acceptable* if there exists a sequence of correlated strategies $(c_S)_{S \subseteq N}$ with $c_S \in \Delta(\mathbf{X}_S)$ and $c_N = c$, such that for every $T \subseteq N$ and for every $d_T \in \Delta(\mathbf{X}_T)$, $U_T(d_T \times c_{-T}) \leqslant U_N(c)$.

Definition 14. A payoff vector $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ is in the *TU-core*, $C_{tu}(\Gamma)$, if there exists a TU-acceptable correlated strategy, c such that $\mathbf{w} = \mathbf{U}(c) = (U_1(c), \ldots, U_n(c))$.

Proposition 7.

1. *Every aggregate mediated equilibrium is TU-acceptable.*
2. *There exists a game Γ, and a TU-acceptable strategy c, which is not an aggregate mediated equilibrium.*
3. *The aggregate mediated core is contained in the TU-core. That is,*

$$C_{ag-m}(\Gamma) \subseteq C_{tu}(\Gamma),$$

and moreover, there exists a game Γ for which a strict inclusion holds.

Proof. 1. The proof mimics the proof of the analogous claim in Proposition 5.

2. Consider the game in strategic form given in Theorem 1, Γ. It can be easily checked that (c, c, c) is a TU-acceptable correlated strategy in this game. By Theorem 1, (c, c, c) is not a strong mediated equilibrium in Γ. Therefore, by Observation 3, (c, c, c) is not an aggregate mediated equilibrium.

3. The proof mimics the proof of Corollary 1. □

[7] See [34].

7.1. The associated TU-cooperative game

A *TU-cooperative game* on the set of players N is a function $v : 2^N \to R$ with $v(\emptyset) = 0$, where 2^N denotes the set of all subsets of N.

Let Γ be a game in strategic form. We define the *associated cooperative game* v_Γ as follows: For every $S \subseteq N$ let

$$v_\Gamma(S) = \min_{c_{-s}} \max_{c_S} U_S(c_S, c_{-S}). \tag{9}$$

Note that:

$$v_\Gamma(N) = \max_{c \in \Delta(X)} U_N(c).$$

A correlated strategy c_S is an *optimal punishing* strategy for S if it ensures that N/S does not obtain more than $v_\Gamma(N/S)$, that is the Min in (9) is attained in c_S. Note that if c is an optimal punishing strategy for N, then it does not punish any one. Actually, such c is a surplus optimal strategy.

An alternative definition of a TU-acceptable strategy is derived from the following observation:

Observation 4. *c is TU-acceptable if and only if for every coalition S,*

$$U_S(c) \geqslant v_\Gamma(S).$$

Consequently, every TU-acceptable strategy c is surplus optimal.

Proof. The proof follows easily from the minimax theorem (see e.g., [34]), according to which:

$$v_\Gamma(S) = \max_{c_S} \min_{c_{-s}} U_S(c_S, c_{-S}). \qquad \square \tag{10}$$

7.2. Symmetric games with nonempty TU-core

In this section we prove that every symmetric game with a nonempty TU-core possesses an aggregate mediated equilibrium that yields an optimal surplus. Needless to say that symmetric games are most popular in the rich literature in the interface between CS/AI and game theory. In particular, much of the extremely rich literature in computer science on congestion games (see e.g. [2,16,26]) deals with symmetric games, in which agents' costs do not depend on their identity and the bundles of resources they can choose from are identical.

We proceed to define symmetric games: A *permutation* of the set of players is a one-to-one function $\pi : N \to N$. We consider games Γ in strategic form, for which all players share the same strategy set, that is $X_i = X_j$ for every $i, j \in N$. Let us call such games *shared-actions games*. Let Γ be a shared-actions game. For every permutation π, and for every strategy profile $\mathbf{x} \in \mathbf{X}$ we denote by $\pi\mathbf{x}$ the permutation of $\mathbf{x}$ by π. That is, $(\pi\mathbf{x})_{\pi i} = x_i$ for every player $i \in N$. Similarly, for a correlated strategy c and a permutation π we denote by πc the correlated strategy defined by $\pi c(\mathbf{x}) = c(\pi\mathbf{x})$ for every $\mathbf{x} \in \mathbf{X}$. The same notations are applied to strategy profiles and correlated strategies in $\Delta(\mathbf{X}_S)$, $S \subset N$.

A shared-actions game Γ is *symmetric* if $u_i(\mathbf{x}) = u_{\pi(i)}(\pi\mathbf{x})$ for every player i, for every action profile $\mathbf{x} \in \mathbf{X}$, and for every permutation π.

Obviously, if Γ is a symmetric game, v_Γ is a symmetric TU-game, that is $v_\Gamma(\pi(S)) = v(S)$ for every $S \subseteq N$ and for every permutation π. Hence $v_\Gamma(S)$ depends only on the number of players in S. Let $f_\Gamma : \{0, 1, \ldots, n\} \to \mathcal{R}$ be a function with $f_\Gamma(0) = 0$ such that $v_\Gamma(S) = f_\Gamma(|S|)$ for every $S \subseteq N$.

We are about to prove that for symmetric games, $C_{tu}(\Gamma) \neq \emptyset$ implies that $C_{ag-m}(\Gamma) \neq \emptyset$. Moreover we show that there exists a mediator that implements a symmetric and surplus optimal correlated strategy c with aggregate deviations, that is, $\mathbf{U}(c) = (\frac{f_\Gamma(n)}{n}, \ldots, \frac{f_\Gamma(n)}{n})$.

Before proving this we need the following lemma that characterizes those symmetric games in strategic form that possess a nonempty TU-core.[8]

Lemma 2. *Let Γ be a symmetric game in strategic form.*

(1) *If $C_{tu}(\Gamma) \neq \emptyset$, then*

$$\left(\frac{f_\Gamma(n)}{n}, \ldots, \frac{f_\Gamma(n)}{n}\right) \in C_{tu}(\Gamma). \tag{11}$$

[8] A sufficient condition for the nonemptiness of the TU-core (not necessarily for symmetric games) is given in [36]. This sufficient condition does not imply our condition. Actually, the TU-core is called there the β-core, but the author rightly explains that this is the β-core in the TU spirit.

Moreover, there exists a symmetric TU-acceptable strategy, c with

$$\mathbf{U}(c) = \left(\frac{f_\Gamma(n)}{n}, \dots, \frac{f_\Gamma(n)}{n}\right).$$

(2) $C_{tu}(\Gamma) \neq \emptyset$ *if and only if*

$$\frac{f_\Gamma(s)}{s} \leqslant \frac{f_\Gamma(n)}{n} \quad \text{for every } 1 \leqslant s < n. \tag{12}$$

Proof. (1) Let $\mathbf{w}$ be in the TU-core of Γ, and Let d be a TU-acceptable strategy with $U_N(d) = \mathbf{w}$. Because v_Γ is a symmetric TU-cooperative game, by Observation 4, πd is also a TU-acceptable strategy for every $\pi \in \Pi$, where Π is the set of all permutations of N. Moreover, $u_N(\pi d) = \pi\mathbf{w}$. Therefore $\pi\mathbf{w} \in C_{tu}(\Gamma)$. As by Observation 4 $C_{tu}(\Gamma)$ is convex set, $\mathbf{w}^\Gamma = \frac{1}{n!}\sum_{\pi\in\Pi} \pi\mathbf{w} \in C_{tu}(\Gamma)$. Moreover,

$$\mathbf{w}^\Gamma = \left(\frac{f_\Gamma(n)}{n}, \dots, \frac{f_\Gamma(n)}{n}\right).$$

Let $c = \frac{1}{n!}\sum_{\pi\in\Pi} \pi d$. Obviously $U_i(c) = \frac{f_\Gamma(n)}{n}$ for every $i \in N$. Therefore, the first assertion is proved.

(2) Assume $C_{tu}(\Gamma) \neq \emptyset$. By the first assertion in this theorem $\mathbf{w}^\Gamma$ is in the TU-core. Let c be a TU-acceptable strategy with $U_N(c) = \mathbf{w}^\Gamma$. Let $1 \leqslant s < n$. Let $S \subseteq N$ be an arbitrary set with s players. Then by Observation 4,

$$s\frac{f_\Gamma(n)}{n} = \sum_{i\in S} w_i^\Gamma \geqslant v_\Gamma(S) = f_\Gamma(s).$$

Therefore, $\frac{f_\Gamma(s)}{s} \leqslant \frac{f_\Gamma(n)}{n}$.

On the other direction, assume condition (12) holds. Let $\mathbf{x}$ be a pure strategy profile with $u_N(\mathbf{x}) = v_\Gamma(N) = f_\Gamma(n)$. Let $c = \frac{1}{n!}\sum_{\pi\in\Pi} \pi\delta_{\mathbf{x}}$. Obviously c is a symmetric surplus optimal correlated strategy, with $u_N(c) = f_\Gamma(n)$ because c is symmetric, $U_i(c) = \frac{f_\Gamma(n)}{n}$ for every $i \in N$. Therefore, by condition (12), c satisfies the conditions to be a TU-acceptable strategy. Hence, $C_{tu}(\Gamma) \neq \emptyset$. □

Definition 15. A symmetric game Γ is called *balanced* if condition (12) holds.

The term “balanced” is inherited from cooperative game theory.[9] By Lemma 2, Γ is balanced if and only if its TU-core is nonempty.

Proposition 8. *Let Γ be a balanced symmetric game. There exists a symmetric aggregate mediated equilibrium c with*

$$\mathbf{U}(c) = \left(\frac{f_\Gamma(n)}{n}, \dots, \frac{f_\Gamma(n)}{n}\right).$$

Proof. For every coalition $S \subset N$, let d_S be an optimal punishing strategy for S. That is, by using d_S the players in S ensure that the members of $N \setminus S$ do not get together more than $v_\Gamma(N \setminus S)$. Obviously, the symmetrization of d_S over all permutations of S is also an optimal punishing strategy for S, which is symmetric in $\Delta(X_S)$. Let us denote this symmetrization by c_S, that is:

$$c_S(\mathbf{x}_s) = \frac{1}{s!}\sum_{\pi_S\in\Pi_S} c_S(\pi_S\mathbf{x}_S), \quad \mathbf{x}_S \in \Delta(\mathbf{X}_S),$$

where s is the number of players in S, and Π_S is the set of permutation of S. Denote $c_N = c$, where c is a symmetric TU-acceptable strategy with $\mathbf{U}(c) = \mathbf{w}^\Gamma = (\frac{f_\Gamma(n)}{n}, \dots, \frac{f_\Gamma(n)}{n})$, whose existence is guaranteed by Lemma 2. We show that the minimal mediator $\mathcal{M} = (c_S)_{\emptyset\neq S\subseteq N}$ implements c with aggregate deviations.

Let $T \subset S$, and let $\xi_T \in \Delta(\mathbf{Z}_T)$ be a potential deviation of the coalition T at the mediated game.

We have to show that ξ_T is not a profitable deviation for its members. Let $\xi = \xi_T \times (\times_{i\in N\setminus T}\delta_{r_i}) \in \Delta(\mathbf{Z})$. We have to show that

$$U_T^{\mathcal{M}}(\xi) \leqslant u_T^{\mathcal{M}}(r_1, r_2, \dots, r_n).$$

That is,

$$U_T(c_\xi) \leqslant U_T(c) = \frac{t}{n} f_\Gamma(n), \tag{13}$$

where $t = |T|$.

[9] The definition of balanced TU-cooperative games is given in [8,29]. By this definition, a symmetric game Γ is balanced if and only if v_Γ is a balanced TU-cooperative game.

Note that for every $\mathbf{z} \in \mathbf{Z}$, $\xi(\mathbf{z}) = 0$ unless $\mathbf{z}_{-T} = r_{-T}$. As the mediator is minimal we get from (2), that for every $\mathbf{x} \in \mathbf{X}$,

$$c_\xi(\mathbf{x}) = \sum_{S \subseteq T} \xi(\mathbf{x}_S, \mathbf{r}_{S^c}) c_{S^c}(\mathbf{x}_{S^c}) = \sum_{S \subseteq T} \xi_T(\mathbf{x}_S, \mathbf{r}_{T \setminus S}) c_{S^c}(\mathbf{x}_{S^c}). \tag{14}$$

Therefore,

$$U_T(c_\xi) = \sum_{S \subseteq T} \left[\sum_{\mathbf{x}_S \in \mathbf{X}_S} \xi_T(\mathbf{x}_S, r_{T \setminus S}) U_T(\mathbf{x}_S, c_{S^c}) \right]. \tag{15}$$

Therefore,

$$U_T(c_\xi) = \sum_{S \subseteq T} \left[\sum_{\mathbf{x}_S \in \mathbf{X}_S} \xi_T(x_S, r_{T \setminus S}) \big(U_S(x_S, c_{S^c}) + U_{T \setminus S}(x_S, c_{S^c}) \big) \right]. \tag{16}$$

Because c_{S^c} is symmetric for the members of S^c,

$$U_{T \setminus S}(x_S, c_{S^c}) = \frac{t-s}{n-s} U_{S^c}(x_S, c_{S^c}) \leqslant \frac{t-s}{n-s} \big(v_\Gamma(N) - U_S(\mathbf{x}_S, c_{S^c}) \big). \tag{17}$$

Therefore,

$$U_S(x_S, c_{S^c}) + U_{T \setminus S}(x_S, c_{S^c}) \leqslant \frac{t-s}{n-s} v_\Gamma(N) + \frac{n-t}{n-s} U_S(x_S, c_{S^c}). \tag{18}$$

Because c_{S^c} is an optimal punishing strategy for S^c,

$$U_S(x_S, c_{S^c}) \leqslant v_\Gamma(S).$$

Therefore, because $v_\Gamma(N) = f_\Gamma(n)$ and $v_\Gamma(S) = f_\Gamma(s)$,

$$U_S(x_S, c_{S^c}) + U_{T \setminus S}(x_S, c_{S^c}) \leqslant \frac{t-s}{n-s} f_\Gamma(n) + \frac{n-t}{n-s} f_\Gamma(s).$$

As the TU-core is not empty, Lemma 2 implies that

$$f_\Gamma(s) \leqslant \frac{s}{n} f_\Gamma(n). \tag{19}$$

Hence,

$$U_S(x_S, c_{S^c}) + U_{T \setminus S}(x_S, c_{S^c}) \leqslant \left[\frac{t-s}{n-s} n + \frac{n-t}{n-s} s \right] \frac{f(n)}{n} = \frac{t}{n} f_\Gamma(n). \tag{20}$$

Plug in (20) in (16), and use the fact that

$$\sum_{S \subseteq T} \left[\sum_{\mathbf{x}_S \in \mathbf{X}_S} \xi_T(x_S, r_{T \setminus S}) \right] = 1$$

to get the desired inequality in (13). □

8. Existence: Balanced symmetric games in strategic form

In this section we prove our next existence result for strong mediated equilibrium.

Theorem 2. *Let Γ be a balanced symmetric game in strategic form. There exists a symmetric mediated equilibrium c with*

$$\mathbf{U}(c) = \left(\frac{f_\Gamma(n)}{n}, \ldots, \frac{f_\Gamma(n)}{n} \right). \tag{21}$$

Proof. By Proposition 8 there exists a symmetric aggregate mediated equilibrium that satisfies (21). By Proposition 3, this aggregate mediated equilibrium is a strong mediated equilibrium. □

If a symmetric game is not balanced, it does not necessarily possess a strong mediated equilibrium. Indeed In [4], Aumann presented a symmetric game that does not have a c-acceptable correlated strategy. Therefore, by Proposition 5 Aumann's game does not have a strong mediated equilibrium.

Consequently, one may conjecture that being a balanced symmetric game is also a necessary condition for the existence of strong mediated equilibrium.

As we show in the next example this conjecture does not hold. In this example, we construct a non-balanced symmetric game, which possesses a strong mediated equilibrium. Moreover, this equilibrium will also be surplus optimal. Hence, the set of symmetric games that possess strong mediated equilibrium strictly contains the set of symmetric balanced games.

Example 3. Consider the following game, Γ. There are 3 players. The strategy set of each player is $\{a_1, a_2\}$,where a_1, a_2 are called locations. A player gets a payoff of 1 if she is alone at one of the locations. Otherwise she gets nothing, i.e., her payoff is 0. Obviously $f_\Gamma(3) = 1$; It is obtained when one player is alone. Also, $f_\Gamma(1) = 0$, because two players can punish the third one by choosing distinct locations. On the other hand, when two players are choosing distinct locations, they ensure a total (sum of) payoffs of 1. Hence, $f_\Gamma(2) = 1$. Therefore, $\frac{f_\Gamma(2)}{2} > \frac{f_\Gamma(3)}{3}$. Hence, Γ is not balanced. We construct a minimal mediator that will strongly implement the correlated strategy c, which is defined as follows: $c = \eta_{\{1,2\}} \times \eta_3$, where $\eta_{\{1,2\}} = (a_1, a_2)$, and η_3 randomizes with equal probabilities between "3 chooses a_1" and "3 chooses a_2". Obviously,

$$\mathbf{U}(c) = \left(\frac{1}{2}, \frac{1}{2}, 0\right).$$

Therefore c is surplus optimal. We continue in constructing the punishing parts in the mediator. If a pair of players come to the mediator, the mediator plays for them some pure strategy in which each of them chooses a different location. If only one player, i goes to the mediator, he plays for him $c_i = (\frac{1}{2}, \frac{1}{2})$. It is easily verified that no coalition T has a profitable deviation at the mediated game.

9. *K*-strong mediated equilibrium

For every "strong" equilibrium concept, and for every $1 \leqslant k \leqslant n$ one can define the corresponding concept of a k-strong equilibrium, in which it is only required that deviation of every subset with at most k players is not profitable. Obviously, a 1-strong equilibrium concept is just a Nash equilibrium, as emphasized in the following definition:

Definition 16. A 1-mediated equilibrium is called a *mediated equilibrium*. That is, a correlated strategy is a mediated equilibrium in a game in strategic form, Γ, if there exists a mediator for Γ, $\mathcal{M}$, and a vector of messages $\mathbf{m} \in \mathbf{M} = \times_{i\in N} M_i$, with $\mathbf{c}_N(\mathbf{m}) = c$, such that $\mathbf{m}$ is pure-strategy equilibrium profile in the mediated game, $\Gamma(\mathcal{M})$; such a mediator is said to *implement c*.

An n-strong equilibrium concept is simply the corresponding strong equilibrium concept. The notion of k-strong equilibrium is very natural; it captures the idea that only a group of a limited size can coordinate a deviation.

Before we prove our main result we need the following lemma:

Lemma 3. *Let Γ be a symmetric game in strategic form. Let $1 \leqslant s < n$ be an integer that divides n. Then,*

$$\frac{f_\Gamma(s)}{s} \leqslant \frac{f_\Gamma(n)}{n}. \tag{22}$$

Proof. One can easily generates a direct proof. However, by [5], v_Γ is super-additive with respect to the grand coalition. That is, If $\emptyset \neq S_j \subseteq N$, $1 \leqslant j \leqslant m$ is a partition of N to nonempty subsets. Then

$$\sum_{j=1}^{m} v_\Gamma(S_j) \leqslant v_\Gamma(N). \tag{23}$$

Because s divides n, one can take such a partition with $m = \frac{n}{s}$, and with $|S_j| = s$ for every $1 \leqslant j \leqslant \frac{n}{s}$. Therefore, (22) follows from (23). □

We are now able to prove:

Theorem 4. *Let Γ be a symmetric game in strategic form. Let $1 \leqslant k \leqslant n$ be an integer. If $k!$ divides n there exists a symmetric k-strong mediated equilibrium, leading to an optimal surplus.*

Proof. Consider the minimal mediator constructed in proof of Proposition 8. Let $1 \leqslant s \leqslant k$, and let S be a coalition with s members. We have to show that S does not have a profitable deviation. One can notice that when proving this result in Proposition 8, the only place in the proof , in which we used the fact that Γ is balanced is in deriving the inequality (19). By Lemma 3 this inequality is satisfied by our game too. Therefore, the result follows. □

Notice that this result implies, for example, that in every symmetric game with even number of agents, a 2-strong mediated equilibrium always exists. Moreover, there is such an equilibrium, in which the agents' sum of payoffs (that is, their social surplus) is maximized. Hence, optimal social surplus can be obtained using a mediator, where deviations by pairs of players are not beneficial.

We believe that this result has quite significant ramifications from the CS/AI perspective. In the recent years much attention has been given to the study of game-theoretic solution concepts in computerized settings (see [10,22] for some recent related overviews). However, many of the results are highly limited due to the fact deviations by coalitions, and even by pairs, are not handled. By introducing the study of mediators, a natural concept in many of the systems discussed in that literature, we are able to obtain real general positive results; for example, if one considers a router in a standard symmetric congestion setting then this router can obtain stability against deviations by singletons and pairs for arbitrary cost functions, when acting as a mediator.

We have already mentioned the example in [4], which we use to show that a non-balanced symmetric game does not necessarily possess a strong mediated equilibrium. This example can be modified to an example in which $k!$ does not divide n, and the game does not have a k-strong mediated equilibrium. On the other hand, Example 3 shows a non-balanced symmetric game, and a k ($k = 2$) such that $k!$ does not divide n ($n = 3$), in which there exists a k-strong symmetric equilibrium.

10. Program equilibrium: A special type of mediators

The theory of mediators discussed in this paper is a very broad one. Indeed, in general, the agents' messages can be arbitrary, and the interpretation of these messages can be arbitrary. One interesting type of messages are those that have the flavor of a standard computer program, using a standard programming language; in this case it is interesting to look at mediators whose role is the mere execution of the programs. As shown in [32] this perspective can be highly productive. We now briefly discuss program equilibrium and its relationships to our setting.

Consider the prisoners dilemma, discussed in the introduction. Denote the possible actions by D (*defect*) and C (*cooperate*). Recall that in the only (dominant strategy) equilibrium of this game both agents will choose D, while mutual cooperation will lead both of the agents to a higher payoff. In [32] this issue has been addressed by considering agents who can use computer programs as their strategies; these computer programs are to run on a single server/machine, and therefore can exploit the famous dual role of computer programs (introduced in [35]): a program can serve both as a set of instructions and as a data file. Consider the program: *IF MY-PROGRAM = YOUR-PROGRAM then C; else D*; The exact syntax and semantics of such programs is discussed in [32], but the reader can easily notice the basic idea: the agent/programmer *instructs* the computer to compare its program to the other program, *as files*, and execute a particular action based on the result of that comparison. There are no circular arguments here, due to the dual role of computer programs. Moreover, this program defines a *program equilibrium*: it is irrational for an agent to deviate from that program assuming the other agent stick to it. As a result, we get cooperation in the one-shot prisoners' dilemma! This result is then extended to a general folk theorem.

An interesting question is whether the role of a mediator can be replaced by a computer program as discussed in [32]. We conjecture:

Conjecture 1. *Given a game in strategic form, the set of outcomes implemented by the program equilibria of the game is equivalent to the set of outcomes implemented by the mediated equilibria of that game.*

11. Discussion

Mediators make perfect sense in the CS/AI setting/literature where the idea of providing agents with protocols and suggestions is typically considered. This is in contrast to most work in economic theory. This paper concentrated on establishing the theory, putting it in the perspective of foundational work in game theory, and proving the usefulness of mediators in establishing behaviors which are stable against group deviations, a significant challenge in multi-agent systems.

Notice that the design of mediators can be viewed as the design of mechanisms for a given game; the mediators do not design new games from scratch, or constrain the agents' behavior in the given game. This is complementary to the literature on mechanism design, which became quite standard in CS/AI (see e.g. [22,24,28]) where games are designed from scratch in order to obtain some desired criteria. The design of mechanisms that lead to desired behavior in a given game is also the subject of work on social laws for artificial agent societies (e.g. [20,30]); however, social laws do constrain the agents' behaviors in the given game.

Finally, an interesting recent line of research deals with the desire to distribute the power of some limited forms of a mediators using alternative cryptographic techniques (see e.g., [1]). It is interesting to note that equipping agents with programs, as discussed in Section 10, can be also viewed as an approach for distributing the power of mediators.

References

[1] I. Abraham, D. Dolev, R. Gonen, J. Halpern, Distributed computing meets game theory: Robust mechanisms for rational secret sharing and multiparty computation, in: Proceedings of the Twenty-Fifth Annual ACM Symposium on Principles of Distributed Computing, 2006, pp. 53–62.

[2] H. Ackermann, P. Goldberg, V. Mirrokni, H. Röglin, B. Vöcking, A unified approach to congestion games and two-sided markets, in: Proceedings of WINE-07, 2007, pp. 30–41.

[3] I. Ashlagi, D. Monderer, M. Tennenholtz, Mediators in position auctions, in: The ACM Conference on Electronic Commerce, EC-2007, 2007.

[4] R. Aumann, Acceptable points in general cooperative n-person games, in: A. Tucker, R. Luce (Eds.), Contribution to the Theory of Games, vol. IV, Annals of Mathematics Studies, vol. 40, 1959, pp. 287–324.

[5] R. Aumann, The core of cooperative games without payments, in: Transactions of the American Mathematical Society, 1961.

[6] R. Aumann, Subjectivity and correlation in randomized strategies, Journal of Mathematical Economics 1 (1974) 67–96.
[7] N. Bhat, K. Leyton-Brown, Y. Shoham, M. Tennenholtz, Bidding Rings Revisited.
[8] O. Bondareva, Some applications of linear programming methods to the theory of cooperative games, Problemi Kibernet 10 (1963) 119–139 (in Russian).
[9] C. Boutilier, Y. Shoham, M. Wellman, Special issue on economic principles of multi-agent systems, Artificial Intelligence 94 (1997).
[10] P. Cramton, Y. Shoham, R. Steinberg, Combinatorial Auctions, MIT Press, 2006.
[11] F.M. Forges, An approach to communication equilibria, Econometrica 54 (6) (1986) 1375–1385.
[12] D. Graham, R. Marshall, Collusive bidder behavior at single-object second-price and English auctions, Journal of Political Economy 95 (1987) 579–599.
[13] R. Holzman, N. Law-Yone, Strong equilibrium in congestion games, Games and Economic Behavior 21 (1997) 85–101.
[14] S. Kakade, M. Kearns, J. Langford, L. Ortiz, Correlated equilibria in graphical games, in: ACM Conference on Electronic Commerce, 2003, pp. 42–47.
[15] E. Kalai, R. Rosenthal, Arbitration of two-party disputes under ignorance, International Journal of Game Theory 7 (1976) 65–72.
[16] E. Koutsoupias, C. Papadimitriou, Worst-case equilibria, in: STACS, 1999.
[17] N. Linial, Game-theoretic aspects of computer science, in: R. Aumann, S. Hart (Eds.), Handbook of Game Theory with Economic Applications, vol. II, North-Holland, 1994, pp. 1340–1395, Chapter 38.
[18] R. McAfee, J. McMillan, Bidding rings, The American Economic Theory 82 (1992) 579–599.
[19] D. Monderer, M. Tennenholtz, K-implementation, Journal of Artificial Intelligence Research (JAIR) 21 (2004) 37–62.
[20] Y. Moses, M. Tennenholtz, Artificial social systems, Computers and Artificial Intelligence 14 (6) (1995) 533–562.
[21] R.B. Myerson, Multistage games with communication, Econometrica 54 (2) (1986) 323–358.
[22] N. Nisan, T. Roughgarden, E. Tardos, V. Vazirani, Algorithmic Game Theory, Cambridge University Press, 2007.
[23] C. Papadimitriou, Computing correlated equilibria in multi-player games, in: Proceedings of the Thirty-Seventh Annual ACM Symposium on Theory of Computing, 2005, pp. 49–56.
[24] D.C. Parkes, Computational mechanism design, in: Lecture Notes of Tutorials at 10th Conf. on Theoretical Aspects of Rationality and Knowledge (TARK-05), Institute of Mathematical Sciences, University of Singapore.
[25] J.S. Rosenschein, G. Zlotkin, Rules of Encounter, MIT Press, 1994.
[26] T. Roughgarden, E. Tardos, How bad in selfish routing? Journal of the ACM 49 (2) (2002) 236–259.
[27] O. Rozenfeld, M. Tennenholtz, Routing mediators, in: Proceedings of the 23rd International Joint Conferences on Artificial Intelligence (IJCAI-07), 2007.
[28] T. Sandholm, Making markets and democracy work: A story of incentives and computing, in: Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI-03), 2003, pp. 1649–1671.
[29] L.S. Shapley, On balanced sets and cores, Naval Research Logistics Quarterly 14 (1967) 453–460.
[30] Y. Shoham, M. Tennenholtz, Social laws for artificial agent societies: Off-line design, Artificial Intelligence 73 (1995).
[31] A. Tauman-Kalai, E. Kalai, E. Lehrer, D. Samet, A commitment folk theorem, Working Paper. Northwestern University, 2007.
[32] M. Tennenholtz, Program equilibrium, Games and Economic Behavior 49 (2004) 363–373.
[33] H. Varian, Position auctions, International Journal of Industrial Organization 25 (2007) 1163–1178.
[34] J. von Neumann, O. Morgenstern, Theory of Games and Economic Behavior, Princeton University Press, Princeton, 1944.
[35] J. von Neumann, First draft of a report on the edvac, contract no. w-670-ord-402 moore school of electrical engineering, Univ. of Penn., Philadelphia, 1945. Reprinted (in part) in: Brian Randell (Ed.), Origins of Digital Computers: Selected Papers, Springer-Verlag, Berlin, Heidelberg, 1982, pp. 383–392.
[36] J. Zhao, A β-core existence result and its application to oligopoly markets, Games and Economic Behavior 27 (1999) 153–168.