

# A Compact Representation Scheme for Coalitional Games in Open Anonymous Environments \*

Naoki Ohta, Atsushi Iwasaki,  
Makoto Yokoo and Kohki Maruono

ISEE, Kyushu University  
6-10-1, Hakozaki, Higashi-ku,  
Fukuoka, 812-8581 JAPAN

(oota@lang./yokoo@iwasaki@maruono@lang.) is.kyushu-u.ac.jp

Vincent Conitzer and Tuomas Sandholm

Computer Science Department  
Carnegie Melon University  
5000 Forbes Avenue  
Pittsburgh, PA 15213 USA

(conitzer/sandholm)@cs.cmu.edu

## Abstract

Coalition formation is an important capability of automated negotiation among self-interested agents. In order for coalitions to be stable, a key question that must be answered is how the gains from cooperation are to be distributed. Recent research has revealed that traditional solution concepts, such as the Shapley value, core, least core, and nucleolus, are vulnerable to various manipulations in open anonymous environments such as the Internet. These manipulations include submitting false names, collusion, and hiding some skills. To address this, a solution concept called the anonymity-proof core, which is robust against such manipulations, was developed. However, the representation size of the outcome function in the anonymity-proof core (and similar concepts) requires space exponential in the number of agents/skills.

This paper proposes a compact representation of the outcome function, given that the characteristic function is represented using a recently introduced compact language that explicitly specifies only coalitions that introduce synergy. This compact representation scheme can successfully express the outcome function in the anonymity-proof core. Furthermore, this paper develops a new solution concept, the anonymity-proof nucleolus, that is also expressible in this compact representation. We show that the anonymity-proof nucleolus always exists, is unique, and is in the anonymity-proof core (if the latter is nonempty), and assigns the same value to symmetric skills.

## Introduction

Coalition formation is an important capability in automated negotiation among self-interested agents. In order for coalitions to be stable, a key question that must be answered is how the gains from cooperation are to be distributed. Coalitional game theory provides a number of solution concepts for this, such as the Shapley value, the core, the least

core, and the nucleolus. Some of these solution concepts have already been adopted in the multi-agent systems literature (Zlotkin & Rosenschein 1994; Shehory & Kraus 1998; Conitzer & Sandholm 2003; 2004).

Besides being of interest to the game-theory and multi-agent systems research communities, the Internet and e-commerce growth have expanded the application areas of coalitional game theory. For example, consider a large number of companies, some subsets of which could form profitable virtual organizations that can respond to larger or more diverse orders than an individual company can. Due to Internet advance, forming such virtual organizations becomes much easier, but the companies must agree on how to divide the profits among themselves.

However, Yokoo *et al.* (2005) have pointed out that existing solution concepts have limitations when applied to open anonymous environments such as the Internet. In such environments, an agent can use multiple identifiers (or *false names*), to pretend to be multiple agents. Also, multiple agents can collude and pretend to be a single agent. Furthermore, an agent might try to hide some of its skills. These manipulations are virtually impossible to detect in open anonymous environments, and have thus become an issue in such environments specifically.

Yokoo *et al.* (2005) have developed a new solution concept called the *anonymity-proof core*, which is robust against these manipulations. However, the anonymity-proof core has one serious limitations, i.e., the representation size of the outcome function requires space exponential in the number of agents/skills.

This paper proposes a compact representation scheme using a synergy coalition group (*SCG*), which only specifies which coalitions introduce new synergy (Conitzer & Sandholm 2003). This compact representation scheme can express the outcome functions of the anonymity-proof core. The size of the representation depends on how much synergy each coalition introduces.

Some prior research has studied compact representation schemes in coalitional games. Deng & Papadimitriou (1994) studied games where the players are nodes of a graph with weights on the edges, and the value of a coalition is determined by the total weight of the edges contained in it. Faigle *et al.* (1997) studied the complexity of testing membership in the core in minimum cost spanning tree games. How-

\*This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (B), 17300049, 2005. Conitzer and Sandholm are supported by the National Science Foundation under ITR grants IIS-0121678 and IIS-0427858, a Sloan Fellowship, and an IBM Ph.D. Fellowship.

Copyright © 2006, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

ever, their results depend heavily on compact game representations specific to the game families under study. In contrast, we study a natural representation that can capture any characteristic function game, similar to (Conitzer & Sandholm 2003; Jeong & Shoham 2005). However, they do not address the various manipulations that are an issue in open anonymous environments.

## Model

Traditionally, value division in coalition formation is studied in *characteristic function games*, where each potential coalition (that is, each subset  $X$  of the agents) has value  $w(X)$  that it can obtain. This assumes that utility is transferable (for example, utility can be transferred using side payments) and that a coalition’s value is independent of what non-members of the coalition do.

The characteristic function by itself does not give sufficient information to assess what manipulations may be performed by agents in an open anonymous environment. Yokoo *et al.* (2005) introduced a more fine-grained representation of what each agent brings to the table. Instead of defining the characteristic function over agents, we define it over the *skills* that the agents possess.<sup>1</sup>

**Definition 1 (skills and agents)** *Let  $T$  be the set of all possible skills. Each agent  $t$  has a subset of skills  $S_t \subset T$ . We assume that the skills are unique:  $\forall t \neq u, S_t \cap S_u = \emptyset$ .*<sup>2</sup>

**Definition 2 (characteristic function defined over skills)** *A characteristic function  $v : 2^T \rightarrow \mathbb{R}$  assigns a value to each set of skills.*

We denote by  $w$  the characteristic function defined over agents and by  $v$  the characteristic function defined over skills. For a given set of agents  $X$ , let  $S_X = \bigcup_{t \in X} S_t$ . Thus,  $w(X) = v(S_X)$ . Typically, both  $v$  and  $w$  are weakly increasing: adding more skills or agents to a coalition never causes harm. We also assume zero-normalized games in the characteristic function, i.e., the value of a single skill is zero. The length of the (naïve) representation of a characteristic function is exponential in the number of skills. We will describe how to represent it compactly later.

We assume that the coalition and the value division (payoffs to agents) are established as follows. There exists a party whom we will call the *mechanism designer*. The mechanism designer knows  $T$ , the set of all possible skills<sup>3</sup>, and  $v$ , the characteristic function over skills. If agent  $t$  is interested in joining a coalition, it declares its skills it has to the mechanism designer. The mechanism designer determines the value division among participants.

<sup>1</sup>The word “skills” should not be interpreted too strictly—while the skills may indeed correspond to the abilities of the agents, they may also correspond to, for example, resources that the agents possess.

<sup>2</sup>This is just for simplifying for the notation; even if there are identical skills, we can give them different names.

<sup>3</sup>We do not require that each skill in  $T$  is actually possessed by some agent; the only thing that is required is that every skill that an agent possesses is indeed in  $T$ . Therefore, the mechanism designer really only needs to know an upper bound on the set of skills possessed by the agents.

For this setting, Yokoo *et al.* (2005) identified the following three types of manipulations by the agents, i.e., *hiding skills*, *using false names*, and *acting in collusion*.

- An agent  $t$  can declare that its skill set is  $S'_t \subseteq S_t$ . It is assumed that an agent cannot claim to have skills that it does not have. Such a lie is detectable because the lie will be exposed if the agent is called on to apply such skills.
- Agent  $t$  can use multiple identifiers and declare that each identifier has subset of the skills  $S_t$ . Because the skills are unique, two different identifiers cannot declare having the same skill.<sup>4</sup> Thus, a false-name manipulation by agent  $t$  corresponds to a partition of  $S_t$  into multiple identifiers.
- Multiple agents can collude, i.e., pretend to be a single agent. They can declare that this agent’s skills are the union of their skills (or a subset of this union, thus combining collusion with skill hiding).

By using the characteristic function defined over skills, the latter two manipulations become ineffective. Thus, we can concentrate our attention on hiding skills to develop a solution concept robust to such manipulations.

## Anonymity-proof core

In this section, we briefly review the recent solution concept called the *anonymity-proof core* (Yokoo *et al.* 2005). As that paper showed, the anonymity-proof core can be characterized by certain axiomatic conditions. We assume that the only knowledge that the mechanism designer has is  $T$  and  $v$ , the set of all possible skills and the characteristic function defined over  $T$ , respectively. He does not know the number of agents, or the skills possessed by each agent. He must define an outcome function  $\pi$  that decides, for all possible skill reports by the agents, how to divide the value generated by these skills.

The outcome function defined in (Yokoo *et al.* 2005) depends on a set of skills that an agent declares and a set of skills that other agents in a coalition declare. However, to simplify the notations, we introduce another formation of an outcome function. We can use this simplified formation without loss of generality as long as the outcome function is anonymity-proof.

More specifically, let  $S$  be the set of skills present in a game. Then, the outcome function  $\pi(s, S)$  takes  $s \in S$  and  $S$  as arguments and returns the payoff to the agent that has  $s$ , when the agent declares its skills as  $s$  and the total set of skills declared by the agents is  $S$ . Notice that, if an agent declares multiple skills, the payoff is the sum of the outcome functions of each of the skills. Before introducing the anonymity-proof core, we formally define an anonymity-proof outcome function.

**Definition 3 (anonymity-proof outcome function)** *An outcome function  $\pi$  is anonymity-proof if*

- $\pi$  achieves Pareto efficiency:  $\forall S, \sum_{s \in S} \pi(s, S) = v(S)$ , and

<sup>4</sup>Alternatively, we can consider a case where agents can declare that they have multiple “copies” of a single skill. We hope to address this model in our future works.

- $\pi$  is robust against hiding skills:  $\forall S, \forall S', \forall S''$ , subject to  $S'' \subset S' \subseteq S$ ,  

$$\sum_{s \in S''} \pi(s, S \setminus (S' \setminus S'')) \leq \sum_{s \in S'} \pi(s, S).$$

Yokoo *et al.* (2005) pointed out that conventional (agent-based) outcome functions are vulnerable to false-name manipulations and collusion, but they also proved that directly applying any solution concept to the skills is robust against false-name manipulations and collusion. Hence, as defined above, anonymity-proof outcome functions are robust against such manipulations.

Now we are ready to introduce the anonymity-proof core. Any anonymity-proof outcome function satisfies Pareto efficiency and provides no incentive for agents to use false-name manipulations, collusion, or hiding of skills. Therefore, to define the anonymity-proof core, we only need to add the *non-blocking* condition described below.

**Definition 4 (anonymity-proof core)** *An anonymity-proof outcome function  $\pi$  is in the anonymity-proof core if it satisfies the non-blocking condition:*

- $\pi$  is never blocked by any coalition  $S'$ , that is,  $\forall S, \forall S' \subseteq S$ ,  $\sum_{s \in S'} \pi(s, S) \geq v(S')$ .

In short, the anonymity-proof core is a combination of core outcomes (each of which satisfies the non-blocking condition) for each possible sets of skills. These core outcomes must be chosen so that the no-hiding condition is satisfied.

**Example 1** *Let there be a set of skills  $T = \{a, b, c, d, e\}$ . Let the characteristic function over skills be:*

- $v(\{a, b, c, d, e\}) = 2$ ,
- $v(\{a, b, d, e\}) = v(\{b, c, d, e\}) = 2$ ,
- $v(\{a, b, c, d\}) = v(\{a, b, c, e\}) = v(\{a, c, d, e\}) = 1$ ,
- $v(\{a, b, c\}) = v(\{a, b, d\}) = v(\{a, b, e\}) = v(\{a, d, e\}) = v(\{b, c, d\}) = v(\{b, c, e\}) = v(\{b, d, e\}) = v(\{c, d, e\}) = 1$ ,
- $v(\{a, b\}) = v(\{b, c\}) = v(\{d, e\}) = 1$ ,
- for any other subset  $S \subset T$ ,  $v(S) = 0$ .

*In this example, if there exist all of five skills, the conventional core gives 1 to  $b$ ,  $p$  to  $d$ ,  $q$  to  $e$ , and 0 to  $a$  and  $c$ , so that  $p \geq 0$ ,  $q \geq 0$ ,  $p + q = 1$ . If  $p = q = 0.5$ , this value division is the nucleolus.*

*Then, the anonymity-proof core gives the following outcome function: for all  $p$  and  $q$ , subject to  $p \geq 0$ ,  $q \geq 0$ ,  $p + q = 1$ ,*

- $\pi(b, \{a, b, \dots\}) = \pi(b, \{b, c, \dots\}) = 1$ ,
- $\pi(d, \{d, e, \dots\}) = p$ ,  $\pi(e, \{d, e, \dots\}) = q$ , and
- for any other skill  $s$  and subset  $S \subset T$ ,  $\pi(s, S) = 0$ .

As Example 1 shows, for the second argument of the outcome functions, all possible combinations of skills should be considered. If the mechanism designer knows the set of skills possessed by agents beforehand, then it is suffice to specify the value division for these skills. However, in this setting, we assume the mechanism designer knows only an upper bound on the set of skills. Thus, the mechanism designer needs to prepare value divisions for all possible subsets of skills. In general, for a game with a set of skills  $S$ ,

where  $|S| = n$ , a traditional solution concept needs to specify the value division only for  $S$ , while an outcome function that is in the anonymity-proof core (and similar concepts) has to specify value divisions for all subsets of  $S$ , the number of which is  $2^{n-1}$ . Thus, the length of the (naïve) representation of the outcome function is exponential in the number of skills.

## Compact representation of outcome functions

To overcome this problem, we develop a compact representation scheme that can express any outcome functions in the anonymity-proof core. We assume that the characteristic function is represented using a recently introduced compact language for this purpose, the *synergy coalition group (SCG)*, which explicitly specifies only coalitions that introduce synergy (Conitzer & Sandholm 2003).

**Definition 5 (synergy coalition group)** *The synergy coalition group SCG is a set of coalitions, each of which has some synergy, i.e.,  $\forall S \in SCG, \forall \{S_1, \dots, S_k\}$ , where all the  $S_j$  are disjoint and  $\bigcup_{1 \leq j \leq k} S_j = S$ ,  $v(S) > \sum_{S_j} v(S_j)$  holds.*

For example, for the coalitional game in Example 1, the SCG is  $[\{a, b\}, \{b, c\}, \{d, e\}]$ .

To describe a characteristic function, it suffices to only specify  $v(S)$  for each element of SCG. The value for coalition  $S$ , which is not an element of SCG, is defined as follows:  $v(S) = \max\{\sum_{1 \leq j \leq k} v(S_j) \mid \bigcup_{1 \leq j \leq k} S_j = S \text{ and all the } S_j \text{ are disjoint}\}$ .

From SCG, we create a group of coalitions called the *generalized synergy coalition group (GSCG)*, which is a superset of the SCG and contains a union of any number of elements of SCG that have a nonempty intersection.

**Definition 6 (generalized synergy coalition group)** *The generalized synergy coalition group GSCG is the smallest group of coalitions that satisfies the following conditions:*

- $\forall S$ , if  $S \in SCG$ , then  $S \in GSCG$ ,
- $\forall S_1 \in GSCG, \forall S_2 \in GSCG$ , if  $S_1 \cap S_2 \neq \emptyset$ , then  $S_1 \cup S_2 \in GSCG$ .

For example, GSCG for the coalitional game in Example 1 is the union of SCG and  $\{a, b, c\}$ .

**Definition 7 (projection onto GSCG)** *For a set of skills  $S$ , projection  $P_S$  of  $S$  onto GSCG is defined as follows:*

$$P_S = \{G \mid G \in GSCG, G \subseteq S, \text{ and } \forall G', \text{ where } G \subset G' \subseteq S, G' \notin GSCG\}.$$

For example, for the five skills game in Example 1, the projection  $P_{\{a, b, c, d, e\}}$  onto GSCG is  $[\{a, b, c\}, \{d, e\}]$ .

Now, we are ready to define a compact representation scheme of an outcome function.

**Definition 8 (compact representation)** *A compact outcome function  $\pi_c$  takes set of skills  $G$ , which is an element of GSCG and a skill  $s$ , which is an element of  $G$ , as arguments, and returns the value division of skill  $s$  when skills  $G$  exist.*

Next, we show how a standard outcome function can be derived from a compact outcome function.

**Definition 9 (compactly expressible outcome function)**

An outcome function  $\pi$  is expressible by a compact outcome function  $\pi_c$  if the following equation holds for all  $s$  and  $S$ , where  $s \in S$ .

$$\pi(s, S) = \begin{cases} \pi_c(s, G) & \text{where } s \in G \text{ and } G \in P_S \\ 0 & \text{if such } G \text{ exists,} \\ & \text{otherwise.} \end{cases}$$

A compact outcome function  $\pi_c$  defines the value division among skills in  $G \in GSCG$ . In Example 1, we describe the value divisions for  $\{a, b\}$ ,  $\{a, c\}$ ,  $\{d, e\}$ , and  $\{a, b, c\}$ .

Next, we examine the condition that  $\pi_c$  should satisfy so that the outcome function  $\pi$  that is expressible by  $\pi_c$  becomes anonymity-proof.

**Definition 10 (no-hiding condition for compact outcome function)**

A compact outcome function  $\pi_c$  satisfies the no-hiding condition if it satisfies the following condition:  $\forall G \in GSCG, \forall S, S', \text{ where } S' \subset S \subseteq G, \text{ let } P_{G \setminus (S \setminus S')}$  be the projection of  $G \setminus (S \setminus S')$  onto  $GSCG$ ,

$$\sum_{P \in P_{G \setminus (S \setminus S')}} \sum_{s \in (S' \cap P)} \pi_c(s, P) \leq \sum_{s \in (G \cap S)} \pi_c(s, G).$$

**Theorem 1** An outcome function  $\pi$ , which is expressible by a compact outcome function  $\pi_c$ , is anonymity-proof if  $\pi_c$  satisfies the no-hiding condition.

**Proof** We derive a contradiction by assuming that hiding some skills is beneficial. We assume there exist three mutually disjoint sets of skills  $S_1, S_2$ , and  $S_3$ , where an agent has  $S_1 \cup S_2$  and other agents have  $S_3$ . We assume for the agent, hiding  $S_2$  is beneficial, i.e.,

$$\sum_{s \in S_1} \pi(s, S_1 \cup S_3) > \sum_{s \in (S_1 \cup S_2)} \pi(s, S_1 \cup S_2 \cup S_3).$$

For simplicity, we assume that the projection of  $S_1 \cup S_2 \cup S_3$  onto  $GSCG$  has a unique element of  $G$ . If the projection has multiple elements, we can derive the same conclusion by applying a similar argument herein to each element.

For each  $s \in (S_1 \cap G)$ , there exists at most one set that contains  $s$  in the projection of  $G \setminus (G \cap S_2)$  onto  $GSCG$ . Let us denote the set as  $P$ . If no such a set exists, we assume  $P$  is an empty set. Also, for each  $s \in (S_1 \cap G)$ , there exists at most one set that contains  $s$  in the projection of  $S_1 \cup S_3$  onto  $GSCG$ . Let us denote the set as  $P'$ . If no such a set exists, we assume  $P'$  is an empty set.

Then, we show that  $P = P'$  holds. It is clear that  $P' \subseteq G$ , otherwise, the projection of  $S_1 \cup S_2 \cup S_3$  onto  $GSCG$  must contain  $G \cup P'$  instead of  $G$ . If  $P$  and  $P'$  are different, then either  $P$  or  $P'$  must be a proper subset of  $P \cup P'$ . However,  $P \cup P'$  is also a member of  $GSCG$  and a subset of  $S_1 \cup S_3$  and  $G$ . This means that  $P \cup P'$  should have been chosen instead of  $P$  or  $P'$ .

From the definition and the no-hiding condition, the following formulae hold:

$$\begin{aligned} \sum_{s \in (S_1 \cup S_2)} \pi(s, S_1 \cup S_2 \cup S_3) &= \sum_{s \in ((S_1 \cup S_2) \cap G)} \pi_c(s, G) \\ &\geq \sum_{s \in (S_1 \cap G)} \pi_c(s, P), \end{aligned}$$

$$\begin{aligned} \sum_{s \in S_1} \pi(s, S_1 \cup S_3) &= \sum_{s \in (S_1 \cap G)} \pi_c(s, P') \\ &= \sum_{s \in (S_1 \cap G)} \pi_c(s, P). \end{aligned}$$

From these formulae, we obtain  $\sum_{s \in S_1} \pi(s, S_1 \cup S_3) \leq \sum_{s \in (S_1 \cup S_2)} \pi(s, S_1 \cup S_2 \cup S_3)$ , but this contradicts the assumption that  $\sum_{s \in S_1} \pi(s, S_1 \cup S_3) > \sum_{s \in (S_1 \cup S_2)} \pi(s, S_1 \cup S_2 \cup S_3)$  holds.  $\square$

Next, we examine the condition that  $\pi_c$  should satisfy so that the outcome function  $\pi$  that is expressible by  $\pi_c$  is in the anonymity-proof core.

**Definition 11 (non-blocking condition for SCG)** A compact outcome function  $\pi_c$  satisfies the non-blocking condition for  $SCG$ , if  $\forall G \in GSCG, \forall S \in SCG, \text{ where } S \subseteq G, \sum_{s \in S} \pi_c(s, G) \geq v(S)$  holds.

**Theorem 2** An outcome function  $\pi$ , which is expressible by a compact outcome function  $\pi_c$ , is in the anonymity-proof core if  $\pi_c$  satisfies the no-hiding condition and the non-blocking condition for  $SCG$ .

**Proof** First, we show that if  $\pi$  is blocked by a coalition, then  $\pi$  is also blocked by an element of  $SCG$  (this characteristic corresponds to Lemma 2 described in Conitzer & Sandholm (2003)). Let us assume  $C$  is a blocking coalition when there exists a set of skills  $S$ , i.e.,  $v(C) > \sum_{P \in P_S} \sum_{s \in (C \cap P)} \pi(s, S)$  holds. Then, let us choose  $\{C_1, \dots, C_k\}$  so that  $C_j \in SCG$ , all of the  $C_j$  are disjoint, and  $v(C) = v(C_1) + \dots + v(C_k)$  holds. From the definition of  $SCG$ , we can always choose such  $\{C_1, \dots, C_k\}$ . Then, the following condition holds:

$$\sum_{1 \leq j \leq k} v(C_j) > \sum_{1 \leq j \leq k} \sum_{P \in P_S} \sum_{s \in (C_j \cap P)} \pi_c(s, S).$$

Thus, for at least one  $C_j$ ,

$$v(C_j) > \sum_{P \in P_S} \sum_{s \in (C_j \cap P)} \pi_c(s, S),$$

which means that  $C_j \in SCG$  is also a blocking coalition.

Now, let us assume that although  $\pi_c$  satisfies the no-hiding condition and the non-blocking condition for each element of  $SCG$ ,  $\pi$  is not in the anonymity-proof core. This means that there exists a blocking coalition. Thus, there must be an element of  $SCG$  that is a blocking coalition. However, this contradicts the assumption that the non-blocking condition for  $SCG$  holds.  $\square$

**Example 2** Consider the skills and the function  $v$  of Example 1. The outcome function in the anonymity-proof core is compactly expressible by the following compact outcome function in the anonymity-proof core. For all  $p$  and  $q$ , subject to  $p \geq 0, q \geq 0, p + q = 1$ ,

- $\pi_c(a, \{a, b, c\}) = 0, \pi_c(b, \{a, b, c\}) = 1, \pi_c(c, \{a, b, c\}) = 0,$
- $\pi_c(a, \{a, b\}) = 0, \pi_c(b, \{a, b\}) = 1,$
- $\pi_c(b, \{b, c\}) = 1, \pi_c(c, \{b, c\}) = 0,$
- $\pi_c(d, \{d, e\}) = p, \pi_c(e, \{d, e\}) = q.$

From this compact representation, we can rederive the original outcome function. For example, let us derive  $\pi(d, \{a, b, c, d, e\})$ . For the five skills game, since  $P_{\{a, b, c, d, e\}}$  consists of  $\{\{a, b, c\}, \{d, e\}\}$ , we have  $\pi(d, \{a, b, c, d, e\}) = \pi_c(d, \{d, e\}) = p$ . Alternatively, when an agent has two skills  $b$  and  $d$ , the outcome function is the sum of  $\pi(b, \{a, b, c, d, e\})$  and  $\pi(d, \{a, b, c, d, e\})$ , which is  $\pi_c(b, \{a, b, c\}) + \pi_c(d, \{d, e\}) = 1 + p$ .

For this game, while the outcome function requires to specify the value for  $\sum_{2 \leq j \leq 5} j \cdot {}_5C_j = 75$  combinations, the compact outcome function needs to specify only 9 combinations. As Example 2 shows, our proposed compact representation scheme significantly reduces the representation size. Furthermore, the following theorem shows that we can find an outcome function that is compactly expressible if the anonymity-proof core is nonempty.

**Theorem 3** *If an outcome function  $\pi$  is in the anonymity-proof core, then there exists an outcome function  $\pi_c$  that is compactly expressible and is in the anonymity-proof core.*

**Proof** *Let us choose a compact outcome function  $\pi_c$  so that  $\forall G \in GSCG, \forall s \in G, \pi_c(s, G) = \pi(s, G)$  holds. From the assumption that  $\pi$  is anonymity-proof, it is clear that  $\pi_c$  satisfies the no-hiding condition and the non-blocking condition. Thus, by Theorem 2, the outcome function which is expressible by  $\pi_c$  is in the anonymity-proof core.  $\square$ .*

### Anonymity-proof nucleolus

In this section, we develop a new solution concept called *anonymity-proof nucleolus*. We extend the traditional nucleolus (Schmeidler 1969) to the anonymity-proof nucleolus using our compact representation.

Before introducing the anonymity-proof nucleolus, we briefly explain the traditional nucleolus. For any outcome (payoff) vector, for any coalition, we can consider the *excess* (or *dissatisfaction*) of that coalition, which is the difference between the value of the characteristic function for that coalition, and the sum of values that agents (skills) in the coalition obtain. Now consider the vector of all the coalitions' excesses, sorted in descending order. The nucleolus chooses the outcome that lexicographically minimizes this vector—that is, it first minimizes the greatest excess, then the second-greatest excess, etc. The nucleolus has some desirable properties: for a transferable utility game, it always exists, is unique (even if the core of the game is empty), is in the core if the core is nonempty, and is symmetric.

The same properties hold for the anonymity-proof nucleolus. In particular, the allocation is uniquely determined, even if the anonymity-proof core is empty, and the anonymity-proof nucleolus belongs to the anonymity-proof core if the anonymity-proof core is nonempty. Before we define the anonymity-proof nucleolus, let us define *excess* and the *compact excess vector*.

**Definition 12 (excess and compact excess vector)** *Given a set of skills  $G$ , which is an element of  $GSCG$ , and a subset of skills  $S \subseteq G$ , let us define the excess of  $S$  as  $v(S) - \sum_{s \in S} \pi_c(s, G)$ . The compact excess vector for a compact outcome function  $\pi_c$  is defined as the vector of*

*excesses for coalitions, where each coalition  $S$  satisfies either that  $S$  is in  $SCG$  or  $S$  consists of exactly one skill, in descending order.*

We can now introduce the anonymity-proof nucleolus.

**Definition 13 (anonymity-proof nucleolus)** *The anonymity-proof nucleolus is the compact outcome function that satisfies the no-hiding condition and gives the (lexicographically) best compact excess vector.*

**Example 3** *Consider the skills and the function  $v$  of Example 1. In this example, if there exist all of five skills, the conventional nucleolus gives 1 to  $b$ , 0.5 to  $d$  and  $e$ , and 0 to  $a$  and  $c$ . On the other hand, the outcome function in the anonymity-proof nucleolus is expressible by the following compact outcome function:*

- $\pi_c(a, \{a, b, c\}) = 0, \pi_c(b, \{a, b, c\}) = 1,$   
 $\pi_c(c, \{a, b, c\}) = 0,$
- $\pi_c(a, \{a, b\}) = 0, \pi_c(b, \{a, b\}) = 1,$
- $\pi_c(b, \{b, c\}) = 1, \pi_c(c, \{b, c\}) = 0,$
- $\pi_c(d, \{d, e\}) = 0.5, \pi_c(e, \{d, e\}) = 0.5.$

We now prove several basic properties of the anonymity-proof nucleolus.

**Theorem 4** *The anonymity-proof nucleolus is unique.*

**Proof** *We derive a contradiction assuming that there exist two distinct compact outcome functions  $\pi'_c$  and  $\pi''_c$  that satisfy the no-hiding condition and give the identical compact excess vector, which is lexicographically best. Let a compact output function  $\pi_c$  be the average of  $\pi'_c$  and  $\pi''_c$ , i.e.,  $\forall s \in G, G \in GSCG, \pi_c(s, G) = (\pi'_c(s, G) + \pi''_c(s, G))/2$ . It is clear that  $\pi_c$  satisfies the no-hiding condition, since both  $\pi'_c$  and  $\pi''_c$  satisfy the no-hiding condition. Also, we can show that  $\pi_c$  gives a lexicographically better compact excess vector than that of  $\pi'_c$  or  $\pi''_c$ , using a similar argument as the proof of Theorem 2 in Schmeidler (1969). This contradicts the assumption that  $\pi'_c$  and  $\pi''_c$  give the lexicographically best compact excess vector.  $\square$*

**Theorem 5** *The anonymity-proof nucleolus always exists.*

**Proof** *It is suffice to show that there exists at least one compact outcome function that satisfies no-hiding condition. Let us consider the following compact outcome function.*

- *Assume a lexicographic order among all skills  $T$  is defined.  $\forall G \subseteq GSCG$ , we set  $\pi_c(s, G)$  as follows:*  
 $\pi_c(s, G) = v(G)$  *if  $s$  is the first skill within  $G$  that appears in the lexicographic order,*  
 $\pi_c(s, G) = 0$  *otherwise.*

*It is clear that this compact output function satisfies the no-hiding condition, i.e., hiding skills is clearly useless.  $\square$ .*

**Theorem 6** *If the anonymity-proof core is nonempty, then the anonymity-proof nucleolus is in the anonymity-proof core.*

**Proof** *Assume  $\pi_c$  is the anonymity-proof nucleolus. From Theorem 3, if the anonymity-proof core is nonempty, there exists a compact outcome function  $\pi'_c$  that gives an outcome function that is in the anonymity-proof core. Each element of*

the compact excess vector of  $\pi'_c$  must be non-positive since  $\pi'_c$  satisfies the non-blocking condition. Since the compact excess vector of  $\pi_c$  is lexicographically better than that of  $\pi'_c$ , each element of the compact excess vector of  $\pi_c$  must be non-positive. This means that the outcome function that is expressible by  $\pi_c$  is also in the anonymity-proof core.  $\square$

**Theorem 7** *If two skills  $s, s'$  are symmetric, i.e.,  $\forall S$ , where  $s \notin S, s' \notin S, v(S \cup \{s\}) = v(S \cup \{s'\})$  holds, then the anonymity-proof nucleolus  $\pi_c$  gives the same value for  $s, s'$ .*

**Proof** *We derive a contradiction by assuming that the anonymity-proof nucleolus  $\pi_c$  gives different values for symmetric skills  $s, s'$ , i.e., one of the following conditions holds:*

$$\pi_c(s, S \cup \{s\}) \neq \pi_c(s', S \cup \{s'\}),$$

$$\pi_c(s, S \cup \{s\} \cup \{s'\}) \neq \pi_c(s', S \cup \{s\} \cup \{s'\}).$$

*Let us choose another anonymity-proof compact outcome function  $\pi'_c$ , which satisfies the following conditions:*

$$\begin{aligned} \pi'_c(s, S \cup \{s\}) &= \pi_c(s', S \cup \{s'\}), \\ \pi'_c(s', S \cup \{s'\}) &= \pi_c(s, S \cup \{s\}), \\ \pi'_c(s, S \cup \{s\} \cup \{s'\}) &= \pi_c(s', S \cup \{s\} \cup \{s'\}), \\ \pi'_c(s', S \cup \{s\} \cup \{s'\}) &= \pi_c(s, S \cup \{s\} \cup \{s'\}), \text{ and} \\ \pi'_c(\cdot, \cdot) &= \pi_c(\cdot, \cdot) \text{ otherwise.} \end{aligned}$$

*It is clear that  $\pi'_c$  gives the same compact excess vector as that of  $\pi_c$ . However, this contradicts the assumption that  $\pi_c$  is the anonymity-proof nucleolus, since from Theorem 4, the anonymity-proof nucleolus must be determined uniquely.  $\square$*

## Discussion

In the worst case, the size of the *GSCG* can be exponential to the size of the *SCG*. For example, assume the following  $k$  elements of the *SCG* exist:  $\{a, a_1\}, \{a, a_2\}, \dots, \{a, a_k\}$ . Then, the size of the *GSCG* becomes  $2^k - 1$ . However, in practice, the size of *GSCG* can be much smaller. To estimate its size, we ran the following simulation: assuming there exists 15 skills, 100 members of *SCG* are drawn iid from a uniform distribution. In this case, the average size of *GSCG* was around 3,400, which is only 10% of  $2^{15}$ .

Note that the anonymity-proof nucleolus does not always minimize the largest excess. Consider the following example with 7 skills:  $v(\{a, b, c\}) = v(\{a, b, d\}) = v(\{a, c, d\}) = v(\{b, c, d\}) = 1, v(\{e, f\}) = v(\{e, g\}) = v(\{f, g\}) = 1$ . In this case, if all 7 skills are declared, the anonymity-proof nucleolus gives  $1/4$  to each of  $\{a, b, c, d\}$  and  $1/3$  to each of  $\{e, f, g\}$ . Hence,  $\{a, b, c, e, f\}$  has an excess of  $1/4 + 1/3 = 7/12$ . Now, let us consider the value division that gives  $(1 + \epsilon)/4$  to each of  $\{a, b, c, d\}$  and  $(1 - \epsilon)/3$  to each of  $\{e, f, g\}$ , that is, we redistribute a little from  $\{e, f, g\}$  to  $\{a, b, c, d\}$ . Now,  $\{a, b, c, e, f\}$  has an excess of  $7/12 - \epsilon/12$ . In fact, in this case, the excess of  $\{a, b, c, e, f\}$  is the largest. Thus the anonymity-proof nucleolus does not minimize the largest excess.

However, the anonymity-proof nucleolus does minimize the largest excess of the members of *SCG*. This appears to be a reasonable compromise given the cost of computing/representing the value division that minimizes the largest excess in general.

## Conclusion

Anonymity-proof solution concepts are developed so that they are robust to various manipulations in open anonymous environments. However, the representation size of the outcome function is exponential in the number of skills that agents declare (while that of the outcome function of conventional solution concepts is linear).

This paper developed a compact representation of the outcome function using a synergy coalition group (*SCG*), which only specifies which coalitions introduce new synergy. We demonstrated that this compact representation scheme can successfully express outcome functions in the anonymity-proof core and can effectively reduce the representation size of the outcome function.

In addition, we introduced a new solution concept, i.e., the anonymity-proof nucleolus. We showed that it always exists, is unique, is symmetric, and belongs to the anonymity-proof core when the latter is nonempty.

## References

- Conitzer, V., and Sandholm, T. 2003. Complexity of determining nonemptiness of the core. In *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI)*, 613–618.
- Conitzer, V., and Sandholm, T. 2004. Computing Shapley values, manipulating value division schemes, and checking core membership in multi-issue domains. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, 219–225.
- Deng, X., and Papadimitriou, C. 1994. On the complexity of cooperative solution concepts. *Mathematics of Operations Research* 19:257–266.
- Faigle, U.; Fekete, S. P.; Hochstättler, W.; and Kern, W. 1997. On the complexity of testing membership in the core of min-cost spanning tree games. *International Journal of Game Theory* 26(3):361–366.
- Ieong, S., and Shoham, Y. 2005. Marginal contribution nets: a compact representation scheme for coalitional games. In *Proceedings of the 6th ACM conference on Electronic commerce (EC '05)*, 193–202.
- Schmeidler, D. 1969. The nucleolus of a characteristic function game. *Society for Industrial and Applied Mathematics Journal of Applied Mathematics* 17:1163–1170.
- Shehory, O., and Kraus, S. 1998. Methods for task allocation via agent coalition formation. *Artificial Intelligence* 101(1–2):165–200.
- Yokoo, M.; Conitzer, V.; Sandholm, T.; Ohta, N.; and Iwasaki, A. 2005. Coalitional games in open anonymous environments. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, 509–515.
- Zlotkin, G., and Rosenschein, J. S. 1994. Coalition, cryptography and stability: Mechanisms for coalition formation in task oriented domains. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, 432–437.