The Knowledge Engineering Review, Vol. 28:4, 381–424. © Cambridge University Press, 2013
doi:10.1017/S0269888913000106
First published online 3 May 2013

Cooperative games and multiagent systems

STÉPHANE AIRIAU

LAMSADE – Université Paris-Dauphine, Place du Maréchal de Lattre de Tassigny, 75775 PARIS Cedex 16, France; e-mail: stephane.airiau@dauphine.fr

Abstract

Forming coalitions is a generic means for cooperation: people, robots, web services, resources, firms; they can all improve their performance by joining forces. The use of coalitions has been successful in domains such as task allocations, sensor networks, and electronic marketplaces. Forming efficient coalitions requires the identification of matching synergies between different entities (finding complementary partners, or similar partners, or partners who add diversity). In addition, the different parties must negotiate a fair repartition of the worth created by the coalition. The first part of this paper is a tutorial on cooperative game theory (also called coalitional games). We then survey the different scenarios and the key issues addressed by the multiagent systems community.

1 Introduction

Coalition formation is an important tool for enabling cooperation in agent societies. Social scientists and economists have studied situations where individuals and businesses benefit from joining forces. The coalition formation problem can be decomposed into two sub-problems. The first problem is the *selection* of the coalition members in the agent population. Then, it is assumed that the members self-organize and achieve their goals, and that the coalition *as a whole* receives a value, that is, the cooperation of the coalition members is rewarded, not the individual agents. The second problem is the *sharing* of this value between its members. These two sub-problems cannot be treated independently (Sandholm *et al.*, 1999): a rational agent will not accept to be a member of a coalition if it can get a higher reward by joining a different coalition. Games that model such cooperation have been extensively studied in the game theory literature in a sub-field called 'Cooperative game theory'. A central point of attention is the stability of the coalition: it is preferable that agents do not have any incentive to leave their current coalition to join a different one. Unfortunately, there is no unique and accepted solution to enforce stability; there are different stability criteria, with their own strengths and weaknesses.

Over the last decade, cooperative game theory has received increased attention in the multiagent systems community as forming dynamic coalitions that may lead to more efficient artificial agent societies. Joining a coalition may be beneficial for an agent: the use of other members' resources may facilitate or enable the solution of a problem. Cooperative game theory has provided a great basis to build coalition formation protocols, but additional issues have risen while trying to apply them. In addition, computer scientists provide insights about computational complexity issues. Owing to the exponential size of the input (when there are n agents, there are 2^n coalitions, and each of them may generate a different value), it may become quickly unfeasible to compute some solution for large values of n.

Multiple scenarios, for example, in the task allocation domains (Shehory & Kraus, 1998) or in the electronic marketplace domain (Tsvetovat *et al.*, 2001) have brought to light many issues and

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

constraints that classical game theory did not address, or some classes of games or representation that allow easier computation of some solution concepts. Some other issues are related with dynamic environments: agents can enter and leave the system at any time, new tasks may appear in the environment, and the environment may be uncertain (uncertainty about the value of the coalitions, about the competence of other agents, etc.). Safety and robustness issues should also be taken into account to guarantee a stable agent society. In addition, researchers must design protocols that are secure to prevent the possibility of manipulation or infiltration by agents or external forces. Another scenario is to consider that the goal of the agent is to maximize utilitarian social welfare. This scenario is not interesting for game theory, as sharing the value between the members is no longer an issue. However, finding the optimal organization is still a hard problem, which can be addressed by AI techniques.

The first part, Section 2, consists of a tutorial on cooperative game theory. We first survey the case in which utility can be transferred between agents (i.e. agents are allowed to make side-payments between them): the transferable utility games (TU games). This is the most important case treated by the game theory literature. We introduce the stability concepts for TU games and provide some results about their complexity. We will also study one special type of TU game that models voting situation, and some extensions of TU games. We then briefly introduce the case where no transfer or comparison of utility are possible between agents: the non-transferable utility games (NTU games) and provide some definitions of stability concepts.

The rest of the paper introduces research from the multiagent systems literature. We first present some applications that have been used to study the formation of coalitions in Section 3. In particular, we discuss the task allocation domain, the electronic marketplace domain, and some variants. We also list some additional domains where coalitions of agents have been used. In Section 4, we survey the cooperative case where the agents' goal is to maximize utilitarian social welfare, that is, the case where the utility of an agent is the total utility of the population. We survey some central algorithms that efficiently search for the optimal partition of agents into coalitions. Finally, in Section 5, we survey some issues raised by the multiagent systems community for which game theory has little (or no) answer so far.

2 Tutorial on cooperative game theory

One branch of game theory studies cooperation between agents, the so-called cooperative game theory (Kahan & Rapoport, 1984; Osborne & Rubinstein, 1994; Peleg & Sudhölter, 2007). The literature is divided into two main models, one in which it is possible to compare utility between two agents and transfer utility (the TU games), and one in which comparison is not possible (the NTU games). In a TU game, a coalition generates a worth, that is, a value achieved through cooperation. The members of a coalition have to share the value of their coalition, and hence they need to compare the utility between them and they must be able to be transfer some utilities between them. In an NTU game, an agent has some preference over the different coalitions, but they cannot provide anything to compensate any agent.

We will introduce TU games in Section 2.1. The solution of a game consists of a partition of the set of agents into coalitions and a payoff distribution to share the value of each coalition. One intuitive and important solution, the core, has an important drawback: for some games, there will not be any solution in the core. Many other solution concepts have been proposed to relax the requirements of the core, and we will study the most important ones (the stable set, the nucleolus, and the kernel). We will also study a solution concept that fosters fairness: the Shapley value.

Then, we will study a particular type of TU games that model voting situations. We will end the study of TU games with few extensions. One extension assumes (in Section 2.8) that the set of agents has already been partitioned into coalitions, or that there exist some affinities between agents. The second extension, called games with externality (in Section 2.9), considers games in which the value of a coalition depends on the other coalitions present in the environment.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

We will then move on to a brief introduction of NTU games in Section 2.10. We will first introduce a subclass of NTU games called hedonic games before moving on the general definition of NTU games. This will conclude the tutorial on cooperative game theory.

2.1 Transferable utility games

In the following, we use a utility-based approach and we assume that 'everything has a price': each agent has a utility function that is expressed in currency units. The use of a common currency enables the agents to directly compare alternative outcomes, and it also enables side-payments.

A TU game involves a set of players N and a characteristic function $v : 2^N \rightarrow \mathbb{R}$ that provides a value for each possible coalition or subset of agents. The characteristic function is common knowledge for the entire population, and the value of a coalition depends only on the players present in its coalition. In the TU games, two questions are asked simultaneously: what coalitions should form (i.e. how to partition the set N into coalitions), and how to share the value of a coalition to each of its members.

In general, it is not always possible to satisfy the interests of all players at the same time. Unfortunately, there is no single criterion for characterizing an acceptable solution. After defining the TU games with more details, we will present some desirable criterion for a solution, and then we will present the main solution concepts.

2.1.1 Notations and types of TU games

We consider a set N of n agents. A *coalition* is a non-empty subset of N. The set N is also known as the *grand coalition*. The set of all coalitions is $\mathscr{C}$ and its cardinality is 2^n. A *coalition structure (CS)* $\mathcal{S} = \{\mathcal{C}_1, \cdots, \mathcal{C}_m\}$ is a partition of N: each set $\mathcal{C}_i$ is a coalition with $\cup_{i=1}^{m} \mathcal{C}_i = N$ and $i \neq j \Rightarrow \mathcal{C}_i \cap \mathcal{C}_j = \emptyset$. The set of all CSs is $\mathscr{S}$ and its size is of the order $O(n^n)$ and $\omega(n^{\frac{n}{2}})$ (Sandholm *et al.*, 1999). The *characteristic function* (or *valuation function*) $v : 2^N \rightarrow \mathbb{R}$ provides the worth or utility of a coalition. For TU games, it is assumed that the valuation of a coalition $\mathcal{C}$ does not depend on the other coalitions present in the population.

DEFINITION 2.1 (TU game). *A transferable utility game (TU game) is defined as a pair (N, v) where N is the set of agents, and $v : 2^N \rightarrow \mathbb{R}$ is a characteristic function.*

We now describe some types of valuation functions.

Additive (or inessential): $\forall \mathcal{C}_1, \mathcal{C}_2 \subseteq N \mid \mathcal{C}_1 \cap \mathcal{C}_2 = \emptyset,\ v(\mathcal{C}_1 \cup \mathcal{C}_2) = v(\mathcal{C}_1) + v(\mathcal{C}_2)$. When a TU game is additive, $v(\mathcal{C}) = \sum_{i \in \mathcal{C}} v(\{i\})$, that is, the worth of each coalition is the same whether its members cooperate or not: there is no gain in cooperation or any synergies between coalitions, which explains the alternative name (inessential) used for such games.

Superadditive: $\forall \mathcal{C}_1, \mathcal{C}_2 \subseteq N \mid \mathcal{C}_1 \cap \mathcal{C}_2 = \emptyset,\ v(\mathcal{C}_1 \cup \mathcal{C}_2) \geq v(\mathcal{C}_1) + v(\mathcal{C}_2)$, in other words, any pair of coalitions is best off by merging into one. In such environments, social welfare is maximized by forming the grand coalition and agents have an incentive to form the grand coalition.

Subadditive: $\forall \mathcal{C}_1, \mathcal{C}_2 \subseteq N \mid \mathcal{C}_1 \cap \mathcal{C}_2 = \emptyset,\ v(\mathcal{C}_1 \cup \mathcal{C}_2) \leq v(\mathcal{C}_1) + v(\mathcal{C}_2)$: the agents are best of when they are on their own, that is, cooperation is not desirable.

Convex games: First, let us call $v(\mathcal{C} \cup \{i\}) - v(\mathcal{C})$ the marginal contribution of a player i to coalition $\mathcal{C}$, that is, it is the increase of value of coalition $\mathcal{C}$ because of the presence of agent i. We call a valuation *convex* if for all $\mathcal{C} \subseteq T$ and $i \notin T\ v(\mathcal{C} \cup \{i\}) - v(\mathcal{C}) \leq v(T \cup \{i\}) - v(T)$. Therefore, a valuation function is convex when the marginal contribution of each player increases with the size of the coalition he joins. Convex valuation functions are superadditive. We will see that such games have some nice properties (e.g. the core of a convex game is non-empty).

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

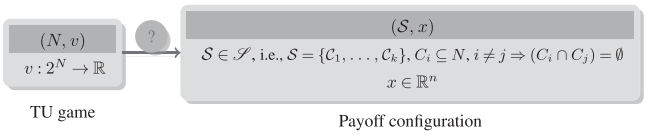


Figure 1 What is solving TU games?

Unconstrained: The valuation function can be superadditive for some coalitions, and subadditive for others: some coalitions should merge when others should remain separated. This is the most difficult and interesting environment.

The valuation function provides a value to a set of agents, not to individual agents. The *payoff distribution* $x = \langle x_1, \cdots, x_n \rangle$ describes how the worth of the coalition is shared between the agents, where x_i is the payoff of agent i. We also use the notation $x(\mathcal{C}) = \sum_{i \in \mathcal{C}} x_i$. A *payoff configuration (PC)* is a pair $(\mathcal{S}, x)$, where $\mathcal{S} \in \mathscr{S}$ is a CS and x is a payoff distribution. $\mathscr{P}$ denotes the set of all PCs.

The solution of a TU game (N, v) is a PC: what are the coalitions that will form and how to distribute the worth of the coalitions (Figure 1). We are now going to present some rationality concepts for PCs, which describe good properties that a solution of the coalition formation should have.

2.1.2 Rationality concepts

We now discuss different rationality concepts for payoff distributions, that is, some properties that link the coalition values to the agents' individual payoff.

Efficiency: $x(N) = v(N)$, the payoff distribution, is an allocation of the whole worth of the grand coalition to all the players. In other words, no utility is lost at the level of the population.

Individual rationality: An agent i will be a member of a coalition only when $x_i \geq v(\{i\})$, that is, to be part of a coalition, a player must be better off than when it is on its own.

Group rationality: $\forall \mathcal{C} \subseteq N,\ x(\mathcal{C}) \geq v(\mathcal{C})$, that is, the sum of the payoffs of a coalition should be at least the value of the coalition (there should not be any loss at the level of a coalition).

Pareto optimal payoff distribution: It may be desirable to have a payoff distribution where no agent can improve its payoff without lowering the payoff of another agent. More formally, a payoff distribution x is Pareto optimal iff

$$\nexists y \in \mathbb{R}^n \mid \exists i \in N \mid \{y_i > x_i \text{ and } \forall j \neq i,\ y_j \geq x_j\}.$$

Two notions will be helpful to discuss some solution concepts. The first is the notion of *imputation*, which is a payoff distribution with the minimal acceptable constraints.

DEFINITION 2.2 (Imputation). *An* imputation *is a payoff distribution that is efficient and individually rational for all agents. The set of all imputations is denoted by* $\mathcal{I}mp$.

An imputation is a solution candidate for a payoff distribution, and can also be used to object a payoff distribution. The second notion is the *excess*, which measures the improvement due to a change of coalition in a CS.

DEFINITION 2.3 (Excess). *The excess related to a coalition* $\mathcal{C}$ *given a payoff distribution* x *is* $e(\mathcal{C}, x) = v(\mathcal{C}) - x(\mathcal{C})$.

We can provide two interpretation of the excess. First, it may measure the total amount that the players would gain or lose if they were to form coalition $\mathcal{C}$. When $e(\mathcal{C}, x) > 0$, it means agents in $\mathcal{C}$ would gain some utility by forming $\mathcal{C}$, and hence they have an incentive in doing so. An agent

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

can use the value of the excess of a coalition $\mathcal{C}$ as a measure of its strength, that is, if it were to form the coalition $\mathcal{C}$, the agent would be able to generate an additional value of $e(\mathcal{C}, x)$. Another interpretation is to view the excess as an amount of complaints: when the excess for a coalition is positive, it means that some utility is lost, which is not acceptable! Some stability concepts (the kernel and the nucleolus, see below) are based on the excess of coalitions. The core can also be defined using the notion of excess.

The solution of the coalition formation problem is a PC $(\mathcal{S}, x)$. The problems of finding the CS (i.e. finding which coalitions are formed) and of finding a payoff distribution (i.e. sharing the value of the coalitions between the members) cannot generally be separated. In the following, we are going to present different solution concepts proposed in the literature. Each has pros and cons, and none is clearly better than all others.

2.2 The stable set

A first idea is to use the concept of dominance between outcomes from non-cooperative game theory and to apply it to the context of cooperative games.

Let x and y be two payoff distributions. We say that x dominates y iff $\exists T \subseteq N$ such that $\forall i \in T,\ x_i > y_i$ and $v(T) \geq x(T)$ and we note $x \succ y$. In other words, there exists a coalition T for which each member prefers the allocation x over y, and they can obtain this utility. This dominance relation may not be complete, that is, two payoff distributions may not be comparable. In addition, dominance may not be transitive. One way to characterize fairness is to ensure, for all agents, that there is no other payoff distribution that dominates the current one. The idea of the stable set is to gather together the payoff distributions that are not comparable between each other, and that dominate some payoff distributions outside the stable set. The formal definition follows:

DEFINITION 2.4 (Stable set). *The* stable set *V is a set of imputations that satisfies the following conditions*:

Internal Stability: $\forall x \in V,\ \nexists y \in V$ such that $y \succ x$

External Stability: $\forall z \notin V,\ \exists y \in V$ such that $y \succ z$.

In other words, internal stability ensures that no payoff distribution in the stable set dominates any other payoff distribution in the stable set. External stability ensures that for any payoff distribution that is not in the stable set, there exists one in the stable set that dominates it. Hence, the stable set represents a set of acceptable payoff distribution from a global point of view, which is akin to the Pareto Optimality concept of non-cooperative game theory: individual player can prefer some distributions over others in the stable set, but not all the players will have the same preferences. Just as in non-cooperative game theory, Pareto Optimality is accepted as a desirable equilibrium criterion: the stable set can be viewed as a desirable property of a solution. Although in many situations, the stable set is guaranteed to be non-empty, it is not always the case.

2.3 The core

Let us assume that we have a TU game (N, v) and that we want to form the grand coalition. The core, which was first introduced by Gillies (1953), is the most attractive and natural way to define stability. A payoff distribution is in the *core* when no set of agents have any incentive to form a different coalition. More formally:

DEFINITION 2.5 (Core). *A payoff distribution $x \in \mathbb{R}^n$ is in the* core *of a game (N, v) iff x is an imputation that is group rational, that is, core*$(N, v) = \{x \in \mathbb{R}^n \mid \sum_{i \in N} x_i = v(N) \wedge \forall \mathcal{C} \subseteq N\ x(\mathcal{C}) \geq v(\mathcal{C})\}$

A payoff distribution is in the core when no group of agents has any interest in rejecting it, that is, no group of agents can gain by forming a different coalition. Note that this condition has to be

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

true for all subsets of N, in particular for all singletons, which ensures individual rationality. The core can thus be defined as a payoff structure that satisfies weak linear inequalities. The core is therefore closed and convex. Another way to define the core is in terms of excess:

DEFINITION 2.6 (Core). *The core is the set of payoff distribution* $x \in \mathbb{R}^n$, *such that* $\forall R \subset N,\ e(R,x) \leq 0$.

In other words, a PC is in the core when there exists no coalition that has a positive excess. This definition is attractive as it shows that no coalition has any complaint: each coalition's demand can be granted.

However, there are multiple concerns associated with using the notion of the core. First, the core can be empty: the conflicts captured by the characteristic function cannot satisfy all the players simultaneously. When the core is empty, at least one player is dissatisfied by the utility allocation and therefore blocks the coalition. Let us consider the following example from Kahan and Rapoport (1984): $v(\{A,B\}) = 90,\ v(\{A,C\}) = 80,\ v(\{B,C\}) = 70$, and $v(N) = 120$. In this case, the core is the PC where the grand coalition forms and the associated payoff distribution is (50, 40, 30). If $v(N)$ is increased, the size of the core also increases. However, if $v(N)$ decreases, the core becomes empty.

The other issue with adopting the core as stability concept concerns computational complexity. Checking whether a payoff distribution is in the core is $\mathcal{NP}$-hard (Conitzer & Sandholm, 2004). In addition, determining the non-emptiness of the core, even for a superadditive game, is $\mathcal{NP}$-hard (Conitzer & Sandholm, 2003), although there exists a transfer scheme to converge to the core (Wu, 1977). In addition, Dieckmann and Schwalbe (2002) introduce a process that leads to a core allocation in non-superadditive games.

However, some classes of games are guaranteed to have a non-empty core. For example, a convex game has a non-empty core. The set of games with non-empty core has been characterized independently by Bondareva (1963) and Shapley (1967), and the result is known as the Bondareva–Shapley theorem. The idea is to use a linear program (LP) to define the core:

$$(LP)\begin{cases}\min\ x(N)\\ \text{subject to } x(\mathcal{C}) \geq v(\mathcal{C}) \text{ for all } \mathcal{C} \subseteq N, S \neq \emptyset\end{cases}$$

The linear constraints correspond to the group rationality conditions. A feasible solution of (LP) is a group rational payoff distribution. For a solution to be member of the core, it has to be efficient as well. The group rationality assumption for the grand coalition guarantees that $x(N) \geq v(N)$. The idea of (LP) is to minimize $x(N)$. Then, when the objective function reaches the value of $v(N)$, it is clear that the core of the game is non-empty. We now introduce some notations needed to state the theorem.

First, we introduce a notation that encodes the members of a coalition in a vector form. The *characteristic vector* $\chi_{\mathcal{C}}$ of a coalition $\mathcal{C} \subseteq N$ is vector of $\mathbb{R}^N$ defined by $\chi^i_{\mathcal{C}} = \begin{cases}1 \text{ if } i \in \mathcal{C}\\ 0 \text{ if } i \in N\backslash\mathcal{C}\end{cases}$.

Next, we introduce a weight function, which is called a map.

DEFINITION 2.7 (Map). *A* map *is a function* $2^N\backslash\emptyset \to \mathbb{R}_+$ *that gives a positive weight to each coalition.*

A special kind of map, the balanced map, can be thought of a percentage of time spent by the agents in each coalition. For each agent i, the sum of a balanced map for all coalitions that contain i must sum up to one. The following definition formalizes this idea. Note that the condition is an equality between vectors of $\mathbb{R}^n$.

DEFINITION 2.8 (Balanced map). *A function* $\lambda : 2^N\backslash\emptyset \to \mathbb{R}_+$ *is a* balanced map *iff* $\sum_{\mathcal{C} \subseteq N} \lambda(\mathcal{C})\chi_{\mathcal{C}} = \chi_N$.

We provide an example of a balanced map for three players in Table 1. Using the definition of balanced map, we are now ready to define a balanced game.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

Table 1 Example of a balanced map for $n = 3$

$$\lambda(\mathcal{C}) = \begin{cases} \frac{1}{2} \text{ if } |\mathcal{C}| = 2 \\ 0 \text{ otherwise} \end{cases}$$

Each of the row sums up to 1:
for all i, we have $\frac{1}{2}\chi^i_{\{1,2\}} + \frac{1}{2}\chi^i_{\{1,3\}} + \frac{1}{2}\chi^i_{\{2,3\}} = \chi^i_{\{1,2,3\}} = 1$,
or using the vector notation $\frac{1}{2}\chi_{\{1,2\}} + \frac{1}{2}\chi_{\{1,3\}} + \frac{1}{2}\chi_{\{2,3\}} = \langle 1, 1, 1\rangle$.

value of $\lambda_{\mathcal{C}}\chi^i_{\mathcal{C}}$ for nonzero $\lambda_{\mathcal{C}}$

$\mathcal{C} \setminus i$	1	2	3
$\{1, 2\}$	$\frac{1}{2}$	$\frac{1}{2}$	0
$\{1, 3\}$	$\frac{1}{2}$	0	$\frac{1}{2}$
$\{2, 3\}$	0	$\frac{1}{2}$	$\frac{1}{2}$

DEFINITION 2.9 (Balanced game). *A game is* balanced *iff for each balanced map* λ *we have* $\sum_{\mathcal{C}\subseteq N, \mathcal{C}\neq\emptyset} \lambda(\mathcal{C})v(\mathcal{C}) \leq v(N)$.

The definition of a balanced game may appear artificial. However, it appears in the dual of (LP), which can be shown to be

$$(DLP) \begin{cases} \max \sum_{\mathcal{C}\subseteq N} y_{\mathcal{C}} v(\mathcal{C}) \\ \text{subject to} \begin{cases} \sum_{\mathcal{C}\subseteq N} y_{\mathcal{C}}\chi_{\mathcal{C}} = \chi_N \text{ and,} \\ y_{\mathcal{C}} \geq 0 \text{ for all } \mathcal{C} \subseteq N, \mathcal{C} \neq \emptyset \end{cases} \end{cases}.$$

The Bondareva–Shapley theorem is a consequence of the duality theorem of linear programming. This theorem completely characterizes the set of games with a non-empty core. One standard application of the Bondareva–Shapley theorem is to prove that a market game has a non-empty core (see Osborne & Rubinstein, 1994, section 13.4).

THEOREM 2.1 (Bondareva–Shapley theorem). *A TU game has a non-empty core iff it is balanced.*

There are few extensions to the concept of the core. As discussed above, one main issue of the core is that it can be empty. In particular, a member of a coalition may block the formation so as to gain a very small payoff. When the cost of building a coalition is considered, it can be argued that it is not worth blocking a coalition for a small utility gain. The *strong* and *weak* ϵ-*core* concepts model this possibility. The constraints defining the strong (respectively the weak) ϵ-core become $\forall T \subseteq N, x(T) \geq v(T) - \epsilon$, (respectively $\forall T \subseteq N, x(T) \geq v(T) - |T| \cdot \epsilon$). In the weak ϵ-core, the minimum amount of utility required to block a coalition is per player, whereas for the strong ϵ-core it is a fixed amount. If one picks ϵ large enough, the strong or weak ϵ-core will be non-empty. When decreasing the value of ϵ, there will be a threshold such that for $\epsilon' < \epsilon$ the ϵ'-core ceases to be non-empty. This special ϵ-core is then called the *least core*.

Another way to relax the requirements of the core is to slightly modify the game. Consider the LP and imagine that a feasible solution does not reach the value of the grand coalition. If one could increase sufficiently the value of the grand coalition, the core of the corresponding game would become non-empty. This is the idea of the *cost of stability* (Bachrach *et al.*, 2009). Given a TU game (N, v) and a value $\Delta \in \mathbb{R}_+$, we consider the game (N, v^Δ) where $\forall \mathcal{C} \subset N$, $v^\Delta(\mathcal{C}) = v(\mathcal{C})$, and $v^\Delta(N) = v(N) + \Delta$. The cost of stability is then defined as the smallest Δ such that the core of (N, v^Δ) is non-empty.

In most traditional work in game theory, the superadditivity of the valuation function is not explicitly stated, but it is implicitly assumed when the core is defined. In particular, this assumption ensures that the grand coalition always emerges. That is one of the reasons why efficiency is defined with respect to the grand coalition. In case of an unconstrained valuation function, the grand coalition may not form, and instead a different CS may emerge. We can define the core for CS, and we borrow the definitions from a study by Chalkiadakis *et al.* (2008), but the definitions are similar to the study by Dieckmann and Schwalbe (2002) and Sandholm and Lesser (1997).

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

A payoff distribution x is *efficient* with respect to a CS $\mathcal{S}$ when $\forall \mathcal{C} \in \mathcal{S}$, $\sum_{i \in \mathcal{C}} x_j = v(\mathcal{C})$. A payoff distribution is an *imputation* when it is efficient (with respect to the current CS) and individually rational (i.e. $\forall i \in N,\ x_i \geq v(\{i\})$). The set of all imputations for a CS $\mathcal{S}$ is denoted by $\mathcal{I}mp(\mathcal{S})$. We can now state the definition of the core:

DEFINITION 2.10 (Core). *The core of a game* (N, v) *is the set of all PCs* $(\mathcal{S}, x)$ *such that* $x \in \mathcal{I}mp(\mathcal{S})$ *and* $\forall \mathcal{C} \subseteq N,\ \sum_{i \in \mathcal{C}} x_j \geq v(\mathcal{C})$.

We now introduce a special type of TU game that will be usefull to define other stability concepts (the nucleolus and the kernel). Aumann and Drèze (1974) propose the definition of a game with CS: in their definition, the CS formed by the agent is fixed (e.g. because of some external constraints such as location). In these types of games, the agents' goal is not to change the CS, but simply to obtain a stable payoff distribution. We will provide more information for these types of games in Section 2.8.1. For now, the definition of the game and the statement of the definition of the core suffice.

DEFINITION 2.11 (Game with coalition structure). *A* game with coalition structure *is a triplet* $(N, v, \mathcal{S})$, *where* (N, v) *is a TU game, and* $\mathcal{S}$ *is a particular CS. In addition, transfer of utility is only permitted within (not between) the coalitions of* $\mathcal{S}$, *that is,* $\forall \mathcal{C} \in \mathcal{S}\ x(\mathcal{C}) \leq v(\mathcal{C})$.

For these types of games the core can be defined as follows:

DEFINITION 2.12 (Core). *The core of a game* $(N, v, \mathcal{S})$ *is the set of all PCs* $(\mathcal{S}, x)$ *such that* $x \in \mathcal{I}mp(\mathcal{S})$ *and* $\forall \mathcal{C} \subseteq N,\ \sum_{i \in \mathcal{C}} x_j \geq v(\mathcal{C})$, that is, $core(N, v, \mathcal{S}) = \{x \in \mathbb{R}^n \mid (\forall \mathcal{C} \in \mathcal{S}, x(\mathcal{C}) = v(\mathcal{C}))\}$.

2.4 The nucleolus

The nucleolus has been introduced by Schmeidler (1969) for games with CS. Let $(N, v, \mathcal{S})$ be a TU game with CS and x be a payoff distribution. Let us start the discussion by recalling the definition of the excess given in Section 2.1.2. The *excess* $e(\mathcal{C}, x)$ of coalition $\mathcal{C}$ at x is the quantity $e(\mathcal{C}, x) = v(\mathcal{C}) - x(\mathcal{C})$. For the nucleolus, we use the interpretation that the excess is a measure of complaints: when $e(\mathcal{C}, x)$ is positive, the members of $\mathcal{C}$ should complain that some utility is lost or not given to them. The vector of excesses over all the coalitions is one way to evaluate the amount of complaints about a payoff distribution. The goal of the nucleolus is to minimize this in a certain way.

Let us consider the game in Table 2 and we want to compare two payoff distributions x and y. *A priori*, it is not clear which payoff should be preferred. One solution is to compare the two vectors of complaints using the lexicographical order (it is the order of a dictionary or a phone book[1]).

$$N = \{1, 2, 3\},\ v(\{i\}) = 0 \text{ for } i \in \{1, 2, 3\}\ v(\{1, 2\}) = 5,\ v(\{1, 3\}) = 6,\ v(\{2, 3\}) = 6,\ v(N) = 8$$

Let us consider two payoff vectors $x = \langle 3, 3, 2\rangle$ and $y = \langle 2, 3, 3\rangle$.

Let l be a sequence of m reals. We denote by $l^{\blacktriangleright}$ the *reordering* of l in *decreasing* order. In the example, $e(x) = \langle -3, -3, -2, -1, 1, 1, 0\rangle$ and then $e(x)^{\blacktriangleright} = \langle 1, 1, 0, -1, -2, -3, -3\rangle$. The first entry of $e(x)^{\blacktriangleright}$ is the maximum excess: the agents involved in the corresponding coalition have the most valid complaint. If the maximum excess of x is larger than the one of y, then y should be preferred. If they have the same maximum complaint, one can check the second entry and iterate the reasoning. Hence, we use the lexicographical ordering for comparing two payoff distributions x and y: we say that x is preferred to y when $e(x)^{\blacktriangleright} \leq_{lex} e(y)^{\blacktriangleright}$: there is a smaller amount of complaints in x than in y.

[1] The formal definition is the following:
DEFINITION 2.13 (Lexicographic ordering). *Let* $(x, y) \in (\mathbb{R}^m)^2$. *We say that x is greater or equal to y in the* lexicographical *ordering, and we note* $x \geq_{lex} y$ *when* $\begin{cases} x = y \text{ or} \\ \exists t \text{ s.t. } 1 \leq t \leq m \text{ s.t. } \forall i \text{ s.t. } 1 \leq i < t\ x_i = y_i \text{ and } x_t > y_t \end{cases}$.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

Table 2 A motivating example for the nucleolus

$x = \langle 3, 3, 2\rangle$		$y = \langle 2, 3, 3\rangle$	
coalition $\mathcal{C}$	$e(\mathcal{C}, x)$	coalition $\mathcal{C}$	$e(\mathcal{C}, y)$
$\{1\}$	-3	$\{1\}$	-2
$\{2\}$	-3	$\{2\}$	-3
$\{3\}$	-2	$\{3\}$	-3
$\{1, 2\}$	-1	$\{1, 2\}$	0
$\{1, 3\}$	1	$\{1, 3\}$	1
$\{2, 3\}$	1	$\{2, 3\}$	0
$\{1, 2, 3\}$	0	$\{1, 2, 3\}$	0

A payoff distribution is in the nucleolus when it yields the 'least problematic' sequence of complaints according to the lexicographical ordering. The nucleolus of the game is a set of payoff distributions such that the corresponding vector of excess $e(x)^{\blacktriangleright}$ is minimal.

DEFINITION 2.14 *Let $\mathcal{I}mp$ be the set of all imputations. The* nucleolus *$Nu(N, v)$ is the set*

$$Nu(N, v) = \left\{x \in \mathcal{I}mp \mid \forall y \in \mathcal{I}mp \;\; e(y)^{\blacktriangleright} \geq_{lex} e(x)^{\blacktriangleright}\right\}.$$

Recall that the core is a payoff distribution that has no complaints: for all coalitions $\mathcal{C}$, $e(\mathcal{C}, x) \leq 0$. The nucleolus relaxes the stability requirements of the core, as it is possible that some coalition has a complaint. This relaxation of the core stability requirements is enough to guarantee that not only the nucleolus is non-empty, but also that it contains at most one element. Hence, agents using the nucleolus are guaranteed to find an agreement, and it is unique.

THEOREM 2.2 *Let (N, v) be a TU game and $\mathcal{I}mp$ is the set of imputations. If $\mathcal{I}mp \neq \emptyset$, then the nucleolus $Nu(N, v)$ is* non-empty.

THEOREM 2.3 *The nucleolus has* at most one *element*.

The next property of the nucleolus shows its relation with the core: when the core of a game is non-empty, the nucleolus is included in the core. Hence, the nucleolus can be seen as a refinement of the core. The intuition behind this result is that the core requires no complaint. For a game with non-empty core, some payoff distributions have no complaints; however, the nucleolus, by picking the one with the lowest complaint, will discriminate between these payoffs in the core.

THEOREM 2.4 *Let (N, v) be a TU game with a non-empty core. Then $Nu(N, v) \subseteq core(N, v)$.*

One drawback of the nucleolus is that it is difficult to compute. It can be computed using a sequence of LPs of decreasing dimensions, but the size of each of these groups is exponential. In some special cases, the nucleolus can be computed in polynomial time (Kuipers *et al.*, 2001; Deng *et al.*, 2006; Elkind & Pasechnik, 2009), but in the general case computing the nucleolus is not guaranteed to be polynomial. Only a few papers in the multiagent systems community have used the nucleolus, for example, Yokoo *et al.* (2006).

2.5 The kernel

The kernel is a stability concept that weakens the stability requirements of the core. It was first introduced by Davis and Maschler (1965) and it is based on the idea that the strength of an agent

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

is measured by the maximum excess that agent can obtain by forming a new coalition (i.e. a different interpretation of the kernel). An agent can consider a payoff distribution to be acceptable by comparing its own 'strength' with the 'strength' of other members of its coalition. When both agents have equal strength, they do not have any incentive to leave the coalition. When their strength differs, the weaker agent can make a payment to the stronger agent to balance out their strength and reach some kind of equilibrium. Although its definition is not as intuitive as the core, the kernel exists and is always non-empty. In addition, there is an algorithm that converges to a payoff distribution in the kernel. The guarantee of finding a payoff distribution and the existence of the algorithm make the kernel an attractive stability concept for applications.

2.5.1 *Definition of the kernel*

We recall that the *excess* related to coalition $\mathcal{C}$ for a payoff distribution x is $e(\mathcal{C}, x) = v(\mathcal{C}) - x(\mathcal{C})$. We saw that a positive excess can be interpreted as an amount of complaint for a coalition. We can also interpret the excess as a potential to generate more utility. Let us consider that the agents are forming a CS $\mathcal{S} = \{\mathcal{C}_1, \ldots, \mathcal{C}_k\}$, and let us consider that the excess of a coalition $\mathcal{C} \notin S$ is positive. Agent $i \in \mathcal{C}$ can view the positive excess as a measure of his strength: if she leaves its current coalition in $\mathcal{S}$ and form coalition $\mathcal{C}$, she has the power to generate some surplus $e(\mathcal{C}, x)$. When two agents want to compare their strength, they can compare the maximum excess of a coalition that contains them, and the kernel is based on this idea.

DEFINITION 2.15 (Maximum surplus). *For a TU game (N, v), the* maximum surplus $s_{k,l}(x)$ *of* agent k over agent l *with respect to a payoff distribution x is*

$$s_{k,l}(x) = \max_{\mathcal{C} \subseteq N \mid k \in \mathcal{C}, l \notin \mathcal{C}} e(\mathcal{C}, x).$$

For two agents k and l, the *maximum surplus* $s_{k,l}(x)$ of agent k over agent l with respect to x is the *maximum excess* from a coalition that *includes* k but does *exclude* l. This maximum surplus can be used by agent k to show its strength over agent l: assuming it is positive and that the agent can claim all of it, agent k can argue that it will be better off without agent l, and hence it should be compensated with more utility for staying in the current coalition. When any two agents in a coalition have the same maximum surplus (except for a special case), the agents are said to be in equilibrium. A payoff distribution is in the kernel when all agents are in equilibrium. The formal definitions follow:

DEFINITION 2.16 (kernel). *Let $(N, v, \mathcal{S})$ be a TU game with coalition structure. The* kernel *is the set of imputations $x \in \mathcal{I}mp(\mathcal{S})$ such that for every coalition $\mathcal{C} \in \mathcal{S}$,* if $(k, l) \in \mathcal{C}^2$, $k \neq l$, *then we have either* $s_{k,l}(x) \geq s_{l,k}(x)$ or $x_k = v(\{k\})$.

$s_{k,l}(x) < s_{l,k}(x)$ calls for a transfer of utility from k to l unless it is prevented by individual rationality, that is, by the fact that $x_k = v(\{k\})$.

The kernel and the nucleolus are linked: the following result shows that the nucleolus is included in the kernel. As a consequence, this guarantees that the kernel is non-empty.

THEOREM 2.5 *The nucleolus is included in the kernel.*

THEOREM 2.6 *When $\mathcal{I}mp \neq \emptyset$, then the kernel is non-empty.*

An approximation of the kernel is the ϵ-kernel, where the equality $s_{k,l}(x) = s_{l,k}(x)$ above is replaced by $|s_{k,l}(x) - s_{l,k}(x)| \leq \epsilon$. One property of the kernel is that agents with the same maximum surplus, that is, symmetric agents, will receive equal payoff. For ensuring fairness, this property is important.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

2.5.2 Computational issues

One method for computing the kernel is the Stearns method (Stearns, 1968). The idea is to build a sequence of side-payments between agents to decrease the difference of surpluses between the agents. At each step of the sequence, the agents with the largest maximum surplus difference exchange utility so as to decrease their surplus: the agent with smaller surplus makes a payment to an agent with higher surplus so as to decrease their surplus difference. After each side-payment, the maximum surplus over all agents decreases. In the limit, the process converges to an element in the kernel. Computing an element in the kernel may require an infinite number of steps as the side-payments can become arbitrarily small, and the use of the ϵ-kernel can alleviate this issue. A criteria to terminate Stearns method is proposed by Shehory and Kraus (1999), and we present the corresponding algorithm in Algorithm 1.

Algorithm 1 Transfer scheme to converge to a ϵ-kernel-stable payoff distribution for the CS S

compute-ϵ-kernel (ϵ, S)
repeat
  for each coalition $\mathcal{C} \in S$ **do**
    for each pair $(i, j) \in \mathcal{C}$, $j \neq i$ **do** {compute the surplus for two members of a coalition in S}
      $s_{ij} \leftarrow \max_{R \subseteq N | (i \in R,\, j \notin R)} v(R) - x(R)$
    end for
    $\delta \leftarrow \max_{(i,j) \in N^2} |s_{ij} - s_{ji}|$
    $(i^\star, j^\star) \leftarrow \text{argmax}_{(i,j) \in N^2}\, s_{ij} - s_{ji}$
  end for
  if $\left(x_{j^\star} - v(\{j^\star\}) < \frac{\delta}{2}\right)$ **then** {payment should be individually rational}
    $d \leftarrow x_{j^\star} - v(\{j^\star\})$
  else
    $d \leftarrow \frac{\delta}{2}$
    $x_{i^\star} \leftarrow x_{i^\star} + d$
    $x_{j^\star} \leftarrow x_{j^\star} - d$
  end if
until $\frac{\delta}{v(S)} \leq \epsilon$

Computing a kernel distribution is of exponential complexity. In Algorithm 1, computing the surpluses is expensive, as we need to search through all coalitions that contains a particular agent and does not contain another agent. Note that when a side-payment is carried out, it is necessary to recompute the maximum surpluses. The derivation of the complexity of the Stearns method to compute a payoff in the ϵ-kernel can be found in studies by Klusch and Shehory (1996b) and Shehory and Kraus (1999), and the complexity for one side-payment is $O(n \cdot 2^n)$. Of course, the number of side-payments depends on the precision ϵ and on the initial payoff distribution. They derive an upper bound for the number of iterations: converging to an element of the ϵ-kernel requires $n \log_2\left(\frac{\delta_0}{\in \cdot v(S)}\right)$, where δ_0 is the maximum surplus difference in the initial payoff distribution. To derive a polynomial algorithm, the number of coalitions must be bounded. The solution used by Klusch and Shehory (1996b) and Shehory and Kraus (1999) is to only consider coalitions whose size is bounded in the interval K_1, K_2. The complexity of the truncated algorithm is $O(n^2 \cdot n_{coalitions})$, where $n_{coalitions}$ is the number of coalitions with a size between K_1 and K_2, which is a polynomial of order K_2.

2.5.3 Fuzzy kernel

To take into account the uncertainty in the knowledge of the utility function, a fuzzy version of stability concept can be used. Blankenburg *et al.* (2003) consider a coalition to be kernel–stable with a degree of certainty. This work also presents a side-payment scheme and shows that the complexity is similar to the crisp kernel, and the idea used by Klusch and Shehory (1996b) can be

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

used for ensuring a polynomial coalition formation algorithm. This approach assumes a linear relationship of the membership and coalition values.

2.6 Shapley value

So far, the solution concepts we introduced focus on the stability of the payoff distribution. The Shapley value focuses on fairness. It was introduced by Shapley (1953), who described the notion of fairness in two different ways. The first one is an axiomatic approach: the Shapley value can be defined by a set of axioms, each of them being a desirable property for fairness. The second approach considers a coalition formation process in which agents enter the coalition one by one and obtain the marginal contribution as payoff. This coalition formation process may be unfair because an agent's payoff depends on the joining order. The Shapley value provides fairness by using an average over all possible joining orders. We present these two views on fairness. Then, we will survey computational issues and present some representations that allows for polynomial computation of the Shapley value.

2.6.1 An axiomatic characterization

We want to define a *value*function ϕ, which assigns an efficient allocation x to a TU game (N, v). What properties should this value function satisfy? We now present few simple axioms. The first axiom uses the definition of a dummy agent: an agent i is a *dummy* when $v(\mathcal{C} \cup i) - v(\mathcal{C}) = v(i)$ for all $\mathcal{C} \subseteq N$ such that $i \notin \mathcal{C}$.

DUM ("Dummy actions"): If agent i is a dummy then $x_i = v(\{i\})$. In other words, if the presence of agent i does not improve the worth of a coalition by more than $v(\{i\})$, the agent does not bring anything to the coalition, and then should obtain only $v(\{i\})$.

SYM ("Symmetry"): When two agents generate the same marginal contributions, they should be rewarded equally: for $i \neq j$ and $\forall \mathcal{C} \subseteq N$ such that $i \notin \mathcal{C}$ and $j \notin S$, if $v(\mathcal{C} \cup \{i\}) = v(\mathcal{C} \cup \{j\})$, then $x_i = x_j$.

ADD ("Additivity"): For any two TU games, (N, v) and (N, w), and corresponding payoff profiles, $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$, the payoff profile should be $x + y$ for the TU game $(N, v + w)$.

Each axiom makes sense and one would want a value function to satisfy them all. Actually, Shapley showed that there is a unique value function that satisfies these three axioms.

THEOREM 2.7 *The Shapley value is the unique value that satisfies axioms DUM, SYM, and ADD.*

The theorem states that these three axioms uniquely define a value function, which is called the *Shapley value*. A proof of this theorem can be found in a study by Osborne and Rubinstein (1994). To prove this result, one needs to show the existence of a value function that satisfies the three axioms, and then prove the unicity of the value function. In addition, one can prove that the axioms are independent. Finally, one can also show that if one of the three axioms is dropped, it is possible to find multiple value functions satisfying the other two axioms.

The axioms SYM and DUM are clearly desirable. The last axiom, ADD, is harder to motivate in some cases. If the valuation function of a TU game is interpreted as an expected payoff, then ADD is desirable. In addition, if we consider cost-sharing games and one TU game corresponds to sharing the cost of one service, then ADD is desirable as the cost for a joint service should be the sum of the cost of the separate services. However, if we do not make any assumptions about the games (N, v) and (N, w), the axiom implies that there is no interaction between the two games. In addition, the game $(N, v + w)$ may induce a behavior that may be unrelated to the behavior induced by either (N, v) or (N, w). Other axiomatizations that do not use the ADD axiom have been proposed by Young (1985) and Myerson (1977). These other axiomatizations reinforce the importance of the Shapley value.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

2.6.2 Ordinal marginal contribution

Another interpretation of the Shapley value is based on the notion of ordered marginal contribution. The marginal contribution of an agent i to a coalition $\mathcal{C} \subseteq N$ is $mc_i(\mathcal{C}) = v(\mathcal{C} \cup \{i\}) - v(\mathcal{C})$. Let us consider that a coalition $\mathcal{C}$ is built incrementally with one agent at a time entering the coalition. Also consider that the payoff of each agent i is its marginal contribution. For example, $\langle mc_1(\emptyset), mc_2(\{1\}), mc_3(\{1,2\})\rangle$ is an efficient payoff distribution for a game $\{1, 2, 3\}, v)$. In this case, the value of each agent depends on the order in which the agents enter the coalition, which may not be fair. For example, consider agents that form a coalition to take advantages of price reduction when buying large quantities of a product. Agents that start the coalition may have to spend large setup costs, and agents that come later benefit from the already large number of agents. To alleviate this issue, the Shapley value averages each agents' payoff over all possible orderings: the value of agent i in coalition $\mathcal{C}$ is the average marginal value over all possible orders in which the agents may join the coalition.

Let π represent a joining order of the grand coalition N: π can also be viewed as a permutation of $\langle 1, \ldots, n\rangle$. We write $mc(\pi)$ the payoff vector where agent i obtains $mc_i(\{\pi(j) \mid j < i\})$. The payoff vector $mc(\pi)$ is called the *marginal vector*. Let us denote the set of all permutations of the sequence $\langle 1, \ldots, n\rangle$ as $\Pi(N)$. The Shapley values can then be defined as

$$Sh(N, v) = \frac{\sum_{\pi \in \Pi(N)} mc(\pi)}{n!}.$$

We provide an example in Table 3 in which we list all the orders in which the agents can enter the grand coalition. The sum is over all joining orders, that is, over $n!$ terms. When computing the Shapley value for one agent, one can avoid some redundancy by summing over all coalitions and noticing that

- Members of $\mathcal{C}$ precede i in $|\mathcal{C}|!$ permutations.
- The remaining members succede i in $|N\backslash(\mathcal{C} \cup \{i\})|!$ permutations, that is, in $(n - |\mathcal{C}| - 1)!$ permutations.

These observations allow us to rewrite the Shapley value for a given agent i as

$$Sh_i(N, v) = \sum_{\mathcal{C} \subseteq N\backslash\{i\}} \frac{|\mathcal{C}|!(n-|\mathcal{C}|-1)!}{n!}(v(\mathcal{C} \cup \{i\}) - v(\mathcal{C})).$$

Note that the example from Table 3 also demonstrates that in general the Shapley value is not in the core or in the nucleolus.

Table 3 Example of a computation of Shapley value

$N = \{1, 2, 3\}$ $v(\{1\}) = 0$ $v(\{2\}) = 0$ $v(\{3\}) = 0$ $v(\{1, 2\}) = 90$ $v(\{1, 3\}) = 80$ $v(\{2, 3\}) = 70$ $v(\{1, 2, 3\}) = 120$

	1	2	3
$1 \leftarrow 2 \leftarrow 3$	0	90	30
$1 \leftarrow 3 \leftarrow 2$	0	40	80
$2 \leftarrow 1 \leftarrow 3$	90	0	30
$2 \leftarrow 3 \leftarrow 1$	50	0	70
$3 \leftarrow 1 \leftarrow 2$	80	40	0
$3 \leftarrow 2 \leftarrow 1$	50	70	0
total	270	240	210
Shapley value Sh	45	40	35

Let $y = \langle 50, 40, 30\rangle$

$\mathcal{C}$	$e(\mathcal{C}, Sh)$	$e(\mathcal{C}, y)$
$\{1\}$	-45	0
$\{2\}$	-40	0
$\{3\}$	-35	0
$\{1, 2\}$	5	0
$\{1, 3\}$	0	0
$\{2, 3\}$	-5	0
$\{1, 2, 3\}$	0	0

This example shows that the Shapley value may not be in the core, and may not be the nucleolus.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

2.6.3 Other properties

The Shapley value always exists and is unique. When the valuation function is superadditive, the Shapley value is individually rational, that is, it is an imputation. When the core is non-empty, the Shapley value may not be in the core. However, when the valuation function is convex, the Shapley value is also group rational, and hence it is in the core.

2.6.4 Computational issues

The nature of the Shapley value is combinatorial, as all possible orderings to form a coalition need to be considered. By using specific representations, it is possible to compute the Shapley value efficiently, and we are surveying few representations.

This computational complexity can sometimes be an advantage as agents cannot benefit from manipulation. For example, a single agent could benefit from using multiple identities to play some games. However, the complexity of determining whether an agent can benefit from such false names has been shown to be $\mathcal{NP}$-complete (Yokoo *et al.*, 2005).

Bilateral shapley value

In order to reduce the combinatorial complexity of the computation of the Shapley value, Ketchpel introduces the Bilateral Shapley value (*BSV*; Ketchpel, 1994a). The idea is to consider the formation of a coalition as a succession of merging between two coalitions. Two disjoint coalitions, $\mathcal{C}_1$ and $\mathcal{C}_2$, with $\mathcal{C}_1 \cap \mathcal{C}_2 = \emptyset$, may merge when $v(\mathcal{C}_1 \cup \mathcal{C}_2) \geq v(\mathcal{C}_1) + v(\mathcal{C}_2)$. When they merge, the two coalitions, called founders of the new coalition $\mathcal{C}_1 \cup \mathcal{C}_2$, share the marginal utility as follows: $BSV(\mathcal{C}_1) = \frac{1}{2}v(\mathcal{C}_1) + \frac{1}{2}(v(\mathcal{C}_1 \cup \mathcal{C}_2) - v(\mathcal{C}_2))$ and $BSV(\mathcal{C}_2) = \frac{1}{2}v(\mathcal{C}_2) + \frac{1}{2}(v(\mathcal{C}_1 \cup \mathcal{C}_2) - v(\mathcal{C}_1))$. This is the expression of the Shapley value in the case of an environment with two agents. In $\mathcal{C}_1 \cup \mathcal{C}_2$, each of the founders gets half of its 'local' contribution, and half of the marginal utility of the other founder. Given this distribution of the marginal utility, it is rational for $\mathcal{C}_1$ and $\mathcal{C}_2$ to merge if $\forall i \in \{1,2\}$, $v(\mathcal{C}_i) \leq BSV(\mathcal{C}_i)$. Note that symmetric founders get equal payoff, that is, for $\mathcal{C}_1$, $\mathcal{C}_2$, $\mathcal{C}$ such that $\mathcal{C}_1 \cap \mathcal{C}_2 = \mathcal{C}_1 \cap \mathcal{C} = \mathcal{C}_2 \cap \mathcal{C} = \emptyset$, $v(\mathcal{C} \cup \mathcal{C}_1) = v(\mathcal{C} \cup \mathcal{C}_2) \Rightarrow BSV(\mathcal{C} \cup \mathcal{C}_1) = BSV(\mathcal{C} \cup \mathcal{C}_2)$. Given a sequence of successive merges from the states where each agent is in a singleton coalition, we can use a backward induction to compute a stable payoff distribution (Klusch & Shehory, 1996a). Although the computation of the Shapley value requires looking at all of the permutations, the value obtained by using backtracking and the BSV only focuses on a particular set of permutations, but the computation is significantly cheaper.

Weigthed graph games

Deng and Papadimitriou (1994) introduce a class of games called weighted graph games: they define a TU game using an undirected weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{W})$ where V is the set of vertices and $W : V \rightarrow V$ is the set of edges' weights. For $(i, j) \in V^2$, w_{ij} is the weight of the edge between the vertices i and j. The coalitional game (N, v) is defined as follows:

- $N = V$, that is, each agent corresponds to one vertex of the graph.
- the value of a coalition $\mathcal{C} \subseteq N$ is the sum of the weights between any pairs of members of $\mathcal{C}$, that is, $v(\mathcal{C}) = \sum_{(i,j) \in \mathcal{C}^2} w_{ij}$.

We provide an example in Figure 2. This representation is succinct as we only need to provide n^2 values to represent the entire game. However, it is not a complete representation as some TU games cannot be represented this way (e.g. it is not possible to represent a majority voting game). If we add some restrictions on the weights, we can further guarantee some properties. For example, when all the weights are non-negative, then the game is convex, and then the game is

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

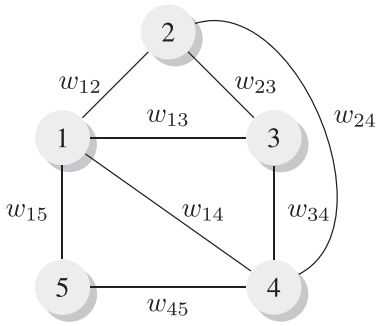


Figure 2 Example of a graph with five agents

guaranteed to have a non-empty core. One other nice property of this representation is that the Shapley value can be computed in quadratic time:

THEOREM 2.8 *Let* (V, W) *be a weighted graph game. The Shapley value of an agent* i *is given by* $Sh_i(N, v) = \sum_{(i,j) \in N^2, i \neq j} w_{ij}$.

This theorem can be proved using the axioms defining the Shapley value.

Multi-issue representation

Conitzer and Sandholm (2004) analyze the case where the agents are concerned with multiple independent issues that a coalition can address. For example, performing a task may require multiple abilities, and a coalition may gather agents that work on the same task but with limited or no interactions between them. A characteristic function v can be decomposed over T issues when it is of the form $v(\mathcal{C}) = \sum_{t=1}^{T} v_t(\mathcal{C})$, in which, for each t, (N, v_t) is a TU game. An agent may have some specific capabilities, and hence may not be able to work on all the issues. For a given issue t, we denote by $\mathcal{I}_t$ the set of agents that are concerned (i.e. that can participate) with the issue t.

The Shapley value for agent i for the characteristic function v is the sum of the Shapley values over the t different issues: $Sh_i(N, v) = \sum_{t=1}^{T} Sh_i(N, v_t)$. When a small number of agents is concerned about an issue, computing the Shapley value for the particular issue can be cheap. For an issue t, the characteristic function v_t concerns only the agents in $\mathcal{I}_t$ when $\forall \mathcal{C}_1 \in \mathscr{C}$, $\mathcal{C}_2 \in \mathscr{C}$ such that $\mathcal{I}_t \cap \mathcal{C}_1 = \mathcal{I}_t \cap \mathcal{C}_2 \Rightarrow v_t(\mathcal{C}_1) = v_t(\mathcal{C}_2)$. When the characteristic function v is decomposed over T issues and when $|\mathcal{I}_t|$ agents are concerned about each issue $t \in [1 \ldots T]$, computing the Shapley value takes $O(\sum_{t=1}^{T} 2^{|\mathcal{I}_t|})$.

Marginal contribution networks

Ieong and Shoham (2005) propose a representation in which the characteristic function is represented by a set of 'rules'. A rule is composed by a pattern and a value: the pattern tells which agent must be present or absent from a coalition so that the value of the coalition is increased by the value of the rule. This representation allows to represent any TU game.

More formally, each player is represented by a boolean variable and the characteristic vector of a coalition is treated as a truth assignment. Each 'rule' associates a pattern ϕ and a weight $w \in \mathbb{R}$. The pattern *phi* is a formula of propositional logic containing variables in N. A positive literal represents the presence of an agent in a coalition, whereas a negative literal represents the absence of an agent in the coalition. The value of a coalition is the sum over the values of all the rules that apply to the coalition.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

When negative literals are allowed or when the weights can be negative, MC-nets can represent any TU game, and hence this representation is complete. When the patterns are limited to conjunctive formula over positive literals and weights are non-negative, MC-nets can represent all and only convex games (in which case they are guaranteed to have a non-empty core).

Using this representation and assuming that the patterns are limited to a conjunction of variables, the Shapley value can be computed in time linear to the size of the input (i.e. the number of rules of the MC-net).

2.7 A special class of TU games: voting games

The formation of coalitions is usual in parliaments or assemblies. Therefore, it is interesting to consider a particular class of coalitional games that models voting in an assembly. For example, we can represent an election between two candidates as a voting game where the winning coalitions are the coalitions of size at least equal to half the number of voters. The formal definition follows:

DEFINITION 2.17 (voting game). *A game* (N, v) *is a* voting game *when*

- *the valuation function takes only two values: 1 for the winning coalitions, 0 otherwise.*
- *v satisfies* unanimity: $v(N) = 1$
- *v satisfies* monotonicity: $S \subseteq T \subseteq N \Rightarrow v(S) \leq v(T)$.

Unanimity and monotonicity are natural assumptions in most cases. Unanimity reflects the fact that all agents agree, and hence the coalition should be winning. Monotonicity tells that the addition of agents in the coalition cannot turn a winning coalition into a losing one, which is reasonable for voting: more supporters should not harm the coalition. A first way to represent a voting game is by listing all winning coalitions. Using the monotonicity property, a more succinct representation is to list only the *minimal winning coalitions*.

DEFINITION 2.18 (Minimal winning coalition). *A coalition* $\mathcal{C} \subseteq N$ *is a minimal winning coalition iff* $v(\mathcal{C}) = 1$ *and* $\forall i \in \mathcal{C}\ v(\mathcal{C} \setminus \{i\}) = 0$.

We can now see how we formalize some common terms in voting: dictatorship, veto player, and blocking coalition.

DEFINITION 2.19 (Dictator). *Let* (N, v) *be a simple game. A player* $i \in N$ *is a* dictator *iff* $\{i\}$ *is a winning coalition.*

Note that with the requirements of simple games, it is possible to have more than one dictator!

DEFINITION 2.20 (Veto Player). *Let* (N, v) *be a simple game. A player* $i \in N$ *is a* veto *player if* $N \setminus \{i\}$ *is a losing coalition. Alternatively,* i *is a* veto *player iff for all winning coalition* $\mathcal{C}$, $i \in \mathcal{C}$.

It also follows that a veto player is member of every minimal winning coalitions. Another concept is the concept of a blocking coalition: it is a coalition that, on its own, cannot win, but the support of all its members is required to win. Put another way, the members of a blocking coalition do not have the power to win, but they have the power to lose.

DEFINITION 2.21 (blocking coalition). *A coalition* $\mathcal{C} \subseteq N$ *is a* blocking coalition *iff* $\mathcal{C}$ *is a losing coalition and* $\forall S \subseteq N \setminus \mathcal{C}$, $S \setminus \mathcal{C}$, *is a losing coalition.*

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

The following theorem characterizes the core of simple games.

THEOREM 2.9 *Let* (N, v) *be a simple game. The core of the game is non-empty iff there exists a veto player and we have* $core(N, v) = \{x \in \mathcal{I}mp \mid x_i = 0$ *for each non−veto player* $i\}$.

A variant of a voting game is a weighted voting game where each agent has a weight and a coalition needs to achieve a threshold or quota to be winning. This is a much more compact representation, as we only use to define a vector of weights and a threshold. However, this is not a complete representation, as some voting games cannot be represented as a weighted voting game. The formal definition follows.

DEFINITION 2.22 (weighted voting game). *A game* (N, v, q, w) *is a* weighted voting game *when*

- $w = (w_1, w_2 \ldots, w_n)$ *is a vector of weights, one for each voter*
- *A coalition* $\mathcal{C}$ *is winning (i.e.* $(v(\mathcal{C}) = 1)$ *iff* $\sum_{i \in C} w_i \geq q$, *it is losing otherwise (i.e.* $(v(\mathcal{C}) = 0)$
- v *satisfies monotonicity*: $S \subseteq T \subseteq N \Rightarrow v(S) \leq v(T)$.

We will note a weighted voting game $(N, w_{i \in N}, q)$ as $[q;\ w_1, \ldots, w_n]$. Note that the weights may not represent the voting power of the player. Let us consider the following weighted voting games:

- [10; 7, 4, 3, 3, 1]: The set of minimal winning coalitions is {{1, 2}{1, 3}{1, 4}{2, 3, 4}}. Player 5, although it has some weight, is a dummy. Player 2 has a higher weight than player 3 and 4, but it is clear that player 2, 3, and 4 have the same influence.
- [51; 49, 49, 2]: The set of winning coalitions is {{1, 2}, {1, 3}, {2, 3}}. It seems that the players have symmetric roles, but it is not reflected in their weights.

The European Union uses a combination of weighted voting games (a decision is accepted when it is supported by 55% of Member States, including at least 15 of them, representing at the same time at least 65% of the Union's population).

The examples raise the subject of measuring the voting power of the agents in a voting game. Multiple indices have been proposed to answer these questions, and we now present few of them. One central notion is the notion of *pivotal player* (also referred to as *swing player*): we say that a voter i is pivotal for a coalition $\mathcal{C}$ when it turns it from a losing to a wining coalition, that is, $v(\mathcal{C}) = 0$ and $v(\mathcal{C} \cup \{i\}) = 1$. Let w be the number of winning coalitions. For a voter i, let η_i be the number of coalitions for which i is pivotal, that is, $\eta_i = \sum_{S \subseteq N \setminus \{i\}} v(S \cup \{i\}) - v(S)$.

Shapley–Shubik index: It is the Shapley value of the voting game, its interpretation in this context is the percentage of the permutations of all players in which i is pivotal. 'For each permutation, the pivotal player gets one more point'. Another interpretation is that the index represent the expected marginal utility assuming all joining orders are equally likely.

$$I_{SS}(N, v, i) = \sum_{\mathcal{C} \subseteq N \setminus \{i\}} \frac{|\mathcal{C}|!(n - |C| - 1)!}{n!}(v(\mathcal{C} \cup \{i\}) - v(\mathcal{C})).$$

One issue is that the voters do not trade the value of the coalition, although the decision that the voters vote about is likely to affect the entire population.

Banzhaff index: For each coalition, we determine which agent is a pivotal agent (more than one agent may be pivotal). The *raw Banzhaff index* of a player i is

$$\beta_i = \frac{\sum_{\mathcal{C} \subseteq N \setminus \{i\}} v(\mathcal{C} \cup \{i\}) - v(\mathcal{C})}{2^{n-1}}.$$

This index corresponds to the expected marginal utility assuming all coalitions are equally likely.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

Table 4 Shapley-Schubik and the Banzhaff indices for the weighted voting game [7; 4, 3, 2, 1]

$\{1, \underline{2}, 3, 4\}$	$\{3, 1, \underline{2}, 4\}$
$\{1, \underline{2}, 4, 3\}$	$\{3, 1, \underline{4}, 2\}$
$\{1, 3, \underline{2}, 4\}$	$\{3, 2, \underline{1}, 4\}$
$\{1, 3, \underline{4}, 2\}$	$\{3, 2, 4, \underline{1}\}$
$\{1, 4, \underline{2}, 3\}$	$\{3, 4, \underline{1}, 2\}$
$\{1, 4, \underline{3}, 2\}$	$\{3, 4, 2, \underline{1}\}$
$\{2, \underline{1}, 3, 4\}$	$\{4, 1, \underline{2}, 3\}$
$\{2, \underline{1}, 4, 3\}$	$\{4, 1, \underline{3}, 2\}$
$\{2, 3, \underline{1}, 4\}$	$\{4, 2, \underline{1}, 3\}$
$\{2, 3, 4, \underline{1}\}$	$\{4, 2, 3, \underline{1}\}$
$\{2, 4, \underline{1}, 3\}$	$\{4, 3, \underline{1}, 2\}$
$\{2, 4, 3, \underline{1}\}$	$\{4, 3, 2, \underline{1}\}$

In red and underlined, the pivotal agent

	1	2	3	4
Sh	$\frac{7}{12}$	$\frac{1}{4}$	$\frac{1}{12}$	$\frac{1}{12}$

winning coalitions:
$\{\underline{1}, \underline{2}\}$
$\{\underline{1}, \underline{2}, 3\}$
$\{\underline{1}, \underline{2}, 4\}$
$\{\underline{1}, \underline{3}, \underline{4}\}$
$\{\underline{1}, 2, 3, 4\}$
In red and underlined, the pivotal agents

	1	2	3	4
β	$\frac{5}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
$I_B(N, v, i)$	$\frac{1}{2}$	$\frac{3}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

For a simple game (N, v), $v(N) = 1$ and $v(\emptyset) = 0$, at least one player i has a power index $\beta_i \neq 0$. Hence, $B = \sum_{j \in N} \beta_j > 0$. The *normalized Banzhaff index* of player i for a simple game (N, v) is defined as

$$I_B(N, v, i) = \frac{\beta_i}{B}.$$

Coleman index: Coleman defines three indices (Coleman, 1970): the power of the collectivity to act $A = \frac{w}{2^n}$ (A is the probability of a winning vote occurring); the power to prevent action $P_i = \frac{\eta_i}{w}$ (it is the ability of a voter to change the outcome from winning to losing by changing its vote); and the power to initiate action $I_i = \frac{\eta_i}{2^n - w}$ (it is the ability of a voter to change the outcome from losing to winning by changing its vote, the numerator is the same as in P_i, but the denominator is the number of losing coalitions, that is, the complement of the one of P_i).

We provide in Table 4 an example of computation of the Shapley–Schubik and Banzhaff indices. This example shows that both indices may be different. There is a slight difference in the probability model between the Banzhaf β_i and Coleman's index P_i: in Banzhaf's, all the voters but i vote randomly, whereas in Coleman's the assumption of random voting also applies to the voter i. Hence, the Banzhaf index can be written as $\beta_i = 2P_i \cdot A = 2I_i \cdot (1 - A)$.

The computational complexity of voting and weighted voting games have been studied in studies by Deng and Papadimitriou (1994) and Elkind *et al.* (2007). For example, the problem of determining whether the core is empty is polynomial. The argument for this result is the following theorem: the core of a weighted voting game is non-empty iff there exists a veto player. When the core is non-empty, the problem of computing the nucleolus is also polynomial, otherwise it is an $\mathcal{NP}$-hard problem. The problem of choosing the weights so that they correspond to a given power index has also been tackled in a study by de Keijzer *et al.* (2010).

2.8 Games with coalition structure and games with a priori unions

We now turn to two special refinements of TU games that are defined using a CS. The first one assumes that a CS has already been formed and the only problem to solve is how to share the value of each coalition. The second one assumes that the grand coalition forms and that members of a coalition in the CS can be represented as one 'meta'-agent.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

2.8.1 Games with coalition structure

In the description of the core and the Shapley value in Sections 2.3 and 2.6, and in most traditional work in game theory, the superadditivity of the valuation function is not explicitly stated, but it is implicitly assumed. When the grand coalition is formed, checking whether the core is empty amounts to checking whether the grand coalition is stable (Wooldridge, 2009). We have already defined games with CS in Section 2.3, and we recall it now.

DEFINITION 2.23 (Game with coalition structure). *A* game with coalition structure *is a triplet (N, v, S), where (N, v) is a TU game, and S is a particular CS. In addition, transfer of utility is only permitted within (not between) the coalitions of S, that is,* $\forall \mathcal{C} \in S,\ x(\mathcal{C}) \leq v(\mathcal{C})$.

Aumann and Drèze (1974) discuss why the coalition formation process may generate a CS that is not the grand coalition. One reason they mention is that the valuation may not be superadditive (and they provide some discussion about why it may be the case). Another reason is that a CS may 'reflect considerations that are excluded from the formal description of the game by necessity (impossibility to measure or communicate) or by choice'. For example, the affinities can be based on location, or trust relations, etc. Another way to understand this definition is to consider that the problems of deciding which coalition forms and how to share the coalition's payoff are decoupled: the choice of the coalition is made first and results in the CS. Only the payoff distribution choice is left open. The agents are allowed to refer to the value of coalition with agents outside of their coalition (i.e. opportunities they would get outside of their coalition) to negotiate a better payoff. Aumann and Drèze use an example of researchers in game theory that want to work in their own country, that is, they want to belong to the coalition of game theorists of their country. They can refer to offers from foreign countries to negotiate their salaries. Note that the agents' goal is not to change the CS, but only to negotiate a better payoff for themselves.

First, we need to define the set of possible payoffs: the payoff distributions such that the sum of the payoff of the members of a coalition in the CS does not exceed the value of that coalition. More formally:

DEFINITION 2.24 (Feasible payoff). *Let* $(N, v, \mathcal{S})$ *be a TU game with CS. The set of* feasible payoff distributions *is* $X_{(N, v, \mathcal{S})} = \{x \in \mathbb{R}^n \mid \forall \mathcal{C} \in \mathcal{S} x(\mathcal{C}) \leq v(\mathcal{C})\}$.

A payoff distribution x is *efficient* with respect to a CS $\mathcal{S}$ when $\forall \mathcal{C} \in \mathcal{S}$, $\sum_{i \in \mathcal{C}} x_j = v(\mathcal{C})$. A payoff distribution is an *imputation* when it is efficient (with respect to the current CS) and individually rational (i.e. $\forall i \in N,\ x_i \geq v(\{i\})$). The set of all imputations for a CS $\mathcal{S}$ is denoted by $\mathcal{I}mp(\mathcal{S})$. We can now state the definition of the core:

DEFINITION 2.25 (Core). *The core of a game* $(N, v, \mathcal{S})$ *is the set of all PCs* $(\mathcal{S}, x)$ *such that* $x \in \mathcal{I}mp(\mathcal{S})$ *and* $\forall \mathcal{C} \subseteq N,\ \sum_{i \in \mathcal{C}} x_j \geq v(\mathcal{C})$, that is, $core(N, v, \mathcal{S}) = \{x \in \mathbb{R}^n \mid \forall \mathcal{C} \in \mathcal{S}, x(\mathcal{C}) = v(\mathcal{C})\}$.

We now provide a theorem by Aumann and Drèze, which shows that the core satisfies desirable properties: if two agents can be substituted, then a core allocation must provide them identical payoffs.

DEFINITION 2.26 (Substitutes). *Let* (N, v) *be a game and* $(i, j) \in N^2$. *Agents i and j are* substitutes *iff* $\forall \mathcal{C} \subseteq N \setminus \{i, j\},\ v(\mathcal{C} \cup \{i\}) = v(\mathcal{C} \cup \{j\})$.

THEOREM 2.10 *Let* $(N, v, \mathcal{S})$ *be a game with coalition structure, let i and j be substitutes, and let* $x \in core(N, v, \mathcal{S})$. *If i and j belong to different members of* $\mathcal{S}$, then $x_i = x_j$.

Aumann and Drèze made a link from a game with CS to a special superadditive game $(N, \hat{v})$ called the superadditive cover (Aumann & Drèze, 1974). First, we need to introduce a notation: for

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

a coalition $\mathcal{C} \subseteq N$, we note $\mathscr{S}_{\mathcal{C}}$ the set of all partitions of $\mathcal{C}$. The valuation function of the superadditive cover $\hat{v}$ is defined as $\hat{v}(\mathcal{C}) = \max_{\mathcal{P} \in \mathscr{S}_{\mathcal{C}}} \{\sum_{T \in \mathcal{P}} v(T)\}$ for all coalitions $\mathcal{C} \subseteq N \backslash \emptyset$, and $\hat{v}(\emptyset) = 0$. In other words, $\hat{v}(\mathcal{C})$ is the maximal value that can be generated by any partition of $\mathcal{C}$. They showed that $core(N, v, \mathcal{S}) \neq \emptyset$ iff $core(N, \hat{v}) \neq \emptyset \wedge \hat{v}(N) = \sum_{\mathcal{C} \in \mathcal{S}} v(\mathcal{C})$ and that when $core(N, v, \mathcal{S}) \neq \emptyset$, then $core(N, v, \mathcal{S}) = core(N, \hat{v})$. This means that a necessary condition for $(N, v, \mathcal{S})$ to have a non-empty core is that $\mathcal{S}$ is an optimal CS.

Aumann and Drèze extend the definition of the core and the Shapley value as well as other stability concepts (nucleolus, Bargaining set, kernel) (Aumann & Drèze, 1974).

2.8.2 Games with a priori unions

So far, a coalition has represented a set of agents that worked on its own. In a CS, the different coalitions are intended to work independently of each other. We can also interpret a coalition to represent a group of agents that is more likely to work together (because of personal or political affinities). The members of a coalition do not mind working with other agents, but they want to be together and negotiate their payoff together, which may improve their bargaining power. This is the idea used in games with a priori unions. Formally, a game with *a priori* unions is similar to a game with CS: it consists of a triplet $(N, v, \mathcal{S})$ where (N, v) is a TU game and S is a CS. However, we assume that the grand coalition forms. The problem is again to define a payoff distribution.

DEFINITION 2.27 (Game with *a priori* unions). *A* game with *a priori* unions *is a triplet* (N, v, S), *where* (N, v) *is a TU game, and* S *is a particular CS. It is assumed that the grand coalition forms.*

Owen (1977) proposes a value that is based on the idea of the Shapley value. The agents form the grand coalition by joining one by one. In the Shapley value, all possible joining orders are allowed. In the Owen value, an agent i may join only when the last agent that joined is a member of i's coalition or when the last agents $(j_1, \ldots, j_k)$ that joined before formed a coalition in S. This is formally captured using the notion of a consistency with a CS:

DEFINITION 2.28 (Consistency with a coalition structure). *A permutation* π *is* consistent *with a CS* S *when, for all* $(i, j) \in \mathcal{C}^2$, $\mathcal{C} \in S$ *and* $l \in N$, $\pi(i) < \pi(l) < \pi(j)$ *implies that* $l \in \mathcal{C}$.

We denote by $\Pi_S(N)$ the set of permutations of N that are consistent with the CS S. The number of such permutations is $m \prod_{\mathcal{C} \in S} |\mathcal{C}|!$ where m is the number of coalitions in S. The Owen value is then defined as follows:

DEFINITION 2.29 (Owen value). *Given a game with a priori union* (N, v, S), *the Owen value* $O_i(N, v, S)$ *of agent* i *is given by*

$$O_i(N, v, S) = \sum_{\pi \in \Pi_S(N)} \frac{mc(\pi)}{|\Pi_S(N)|}.$$

In Table 5, we present the example used for the Shapley value and compute the Owen value. The members of the coalition of two agents improve their payoff by forming a union.

2.9 Games with externalities

A traditional assumption in the literature of coalition formation is that the value of a coalition depends solely on the members of that coalition. In particular, it is independent of non-members' actions. In general, this may not be true: some externalities (positive or negative) can create a dependency between the value of a coalition and the actions of non-members. Sandholm and Lesser (1997) attribute these externalities to the presence of shared resources (if a coalition uses some resource, they will not be available to other coalitions), or when there are conflicting goals: non-members can move the world farther from a coalition's goal state. Ray and Vohra (1999) state

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

Table 5 Example of the computation of an Owen value

$$N = \{1, 2, 3\}$$

$$v(\{1\}) = 0 \quad v(\{2\}) = 0 \quad v(\{3\}) = 0$$
$$v(\{1, 2\}) = 90 \quad v(\{1, 3\}) = 80 \quad v(\{2, 3\}) = 70$$
$$v(\{1, 2, 3\}) = 120$$

$S_1 = \{\{1, 2\}, \{3\}\}$

	1	2	3
$1 \leftarrow 2 \leftarrow 3$	0	90	30
$1 \leftarrow 3 \leftarrow 2$		×	
$2 \leftarrow 1 \leftarrow 3$	90	0	30
$2 \leftarrow 3 \leftarrow 1$		×	
$3 \leftarrow 1 \leftarrow 2$	80	40	0
$3 \leftarrow 2 \leftarrow 1$	50	70	0
total	220	200	60
Owen value $O_i(N, v, S_1)$	55	50	15

$S_2 = \{\{1, 3\}, \{2\}\}$

	1	2	3
$1 \leftarrow 2 \leftarrow 3$		×	
$1 \leftarrow 3 \leftarrow 2$	0	40	80
$2 \leftarrow 1 \leftarrow 3$	90	0	30
$2 \leftarrow 3 \leftarrow 1$	50	0	70
$3 \leftarrow 1 \leftarrow 2$	80	40	0
$3 \leftarrow 2 \leftarrow 1$		×	
total	220	80	180
Owen value $O_i(N, v, S_2)$	55	20	45

that a 'recipe for generating characteristic functions is a minimax argument': the value of a coalition $\mathcal{C}$ is the value $\mathcal{C}$ gets when the non-members respond optimally so as to minimize the payoff of $\mathcal{C}$. This formulation acknowledges that the presence of other coalitions in the population may affect the payoff of the coalition $\mathcal{C}$. As in the studies by Hart and Kurz (1983) and Ray and Vohra (1999), we can study the interactions between different coalitions in the population: decisions about joining forces or splitting a coalition can depend on the way the competitors are organized. For example, when different companies are competing for the same market niche, a small company might survive against a competition of multiple similar individual small companies. However, if some of these small companies form a viable coalition, the competition significantly changes: the other small companies may now decide to form another coalition to be able to successfully compete against the existing coalition. Another such example is a bargaining situation where agents need to negotiate over the same issues: when agents form a coalition, they can have a better bargaining position, as they have more leverage, and because the other party needs to convince all the members of the coalition. If the other parties also form coalition, the bargaining power of the first coalition may decrease.

Two main types of games with externalities are described in the literature, both are represented by a pair (N, v), but the valuation function has a different signature.

Games in partition function form: (Thrall & Lucas, 1963): $v : 2^N \times \mathscr{S} \to \mathbb{R}$. This is an extension of the valuation function of a TU game by providing the value of a coalition, given the current coalition structure (note that $v(\mathcal{C}, \mathcal{S})$ is meaningful when $\mathcal{C} \in \mathcal{S}$).

Games with valuations: $v : N \times \mathscr{S} \to \mathbb{R}$. In these type of games, the valuation function directly assigns a value to an agent given a coalition structure. One possible interpretation is that the problem of sharing the value of a coalition to the members has already been solved.

The definitions of superadditivity, subadditivity, and monotonicity can be adapted to games in partition functions (Bloch, 2003). As an example, we provide the definition for superadditivity.

DEFINITION 2.30 (superadditive games in partition function). *A partition function v is superadditive when, for any CS $\mathcal{S}$ and any coalitions $\mathcal{C}_1$ and $\mathcal{C}_2$ in $\mathcal{S}$, we have* $v(\mathcal{C}_1 \cup \mathcal{C}_2, \mathcal{S} \setminus \{\mathcal{C}_1, \mathcal{C}_2\} \cup \{\mathcal{C}_1 \cup \mathcal{C}_2\}) \geq v(\mathcal{C}_1, \mathcal{S}) + v(\mathcal{C}_2, \mathcal{S})$.

The partition function may also have some regularities when two coalitions merge: either they always have a positive effect on the other coalition, or they always have a negative one.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

More precisely, a partition function exhibits *positive spillovers* when for any CS $\mathcal{S}$ and any coalitions $\mathcal{C}_1$ and $\mathcal{C}_2$ in $\mathcal{S}$, we have $v(\mathcal{C}, \mathcal{S} \setminus \{\mathcal{C}_1, \mathcal{C}_2\} \cup \{\mathcal{C}_1 \cup \mathcal{C}_2\}) \geq v(\mathcal{C}, \mathcal{S})$ for all coalitions $\mathcal{C} \neq \mathcal{C}_1, \mathcal{C}_2$ in $\mathcal{S}$.

These regularities can be exploited when searching for an optimal CS. As shown by Michalak *et al.* (2008), it is possible that the grand coalition is not the CS with maximal social welfare for a superadditive partition function. If the partition function is both superadditive and exhibits a positive spillover, then the grand coalition has maximum social welfare. The similar property holds for subadditive partition function with negative spillovers.

We now turn to considering solution concepts for such games. The issue of extending the Shapley value has a rich literature in game theory. We want the Shapley value to represent an average marginal contribution, but there is a debate over which set of coalition structures. Michalak *et al.* (2010) provide references on different solutions and present three solutions in more details.

Airiau and Sen (2010) considers the issue of the stability of the optimal CS and discusses a possible way to extend the kernel for partition function games. Airiau and Sen (2009) consider coalition formation in the context of games with valuations and propose a solution for myopic agents (an agent will join a coalition only when it is beneficial, without considering long-term effect).

Michalak *et al.* (2009) tackle the problem of representing such games and propose three different representations that depend on the interpretation of the externalities. The first representation considers the value of a coalition in a CS: the value of a coalition can be decomposed into one term that is free of externality and another term that models the sum of the uncertainty because of the formation of the other coalitions. The two other representations consider that the contribution of a coalition in a CS: either by providing the mutual influence of any two coalitions in a CS (outward operational externalities) or by providing the influence of all the other coalitions on a given coalition (inward operational externalities). Michalak *et al.* (2010) and Michalak *et al.* (2010) extend the concept of MC-nets to games with partition function.

2.10 Non-transferable utility games

An underlying assumption behind a TU game is that agents have a common scale to measure the worth of a coalition. Such a scale may not exist in every situation, which leads to the study of games where the utility is non-transferable (NTU games). We start by introducing a particular type of NTU games called *Hedonic games*.

In these games, agents have preferences over coalitions: each agent knows whether it prefers to be in company of some agents rather than others. An agent may enjoy more the company of members of $\mathcal{C}_1$ over members of $\mathcal{C}_2$, but it cannot tell by how much it prefers $\mathcal{C}_2$ over $\mathcal{C}_2$. Consequently, an agent cannot be compensated when it is not part of its favorite coalition. More formally, let N be a set of agents and $\mathcal{N}_i$ be the set of coalitions that contain agent i, that is, $\mathcal{N}_i = \{\mathcal{C} \cup \{i\} \mid \mathcal{C} \subseteq N \setminus \{i\}\}$. For a CS $\mathcal{S}$, we will note $\mathcal{S}(i)$ the coalition in $\mathcal{S}$ containing agent i.

DEFINITION 2.31 (Hedonic games). *A Hedonic game is a tuple* $(N, (\succeq_i)_{i \in N})$ *where*

- N *is the set of agents*
- $\succeq_i \subseteq 2_i^{\mathcal{N}} \times 2_i^{\mathcal{N}}$ *is a complete, reflexive, and transitive preference relation for agent i, with the interpretation that if* $S \succeq_i T$, *agent i prefers coalition T at most as much as coalition S.*

The notion of core can be easily extended for these types of games. Weaker versions of stability can also be defined using the current CS formed, and we now give the definition of stability concepts adapted from a study by Bogomolnaia and Jackson (2002).

Core Stability: A CS $\mathcal{S}$ is *core-stable* iff $\nexists \mathcal{C} \subseteq N \mid \forall i \in \mathcal{C}, \mathcal{C} \succ_i \mathcal{S}(i)$.

Nash Stability: A CS $\mathcal{S}$ is *Nash-stable* iff $(\forall i \in N)\ (\forall \mathcal{C} \in \mathcal{S} \cup \{\emptyset\})\ \mathcal{S}(i) \succsim_i \mathcal{C} \cup \{i\}$. No player would like to join any other coalition in $\mathcal{S}$ assuming the other coalitions did not change.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

Individual Stability: A CS $\mathcal{S}$ is *individually stable* iff $(\nexists i \in N)\ (\nexists \mathcal{C} \in \mathcal{S} \cup \{\emptyset\}) \mid (\mathcal{C} \cup \{i\} \succ_i \mathcal{S}(i))$ and $(\forall j \in \mathcal{C}, \mathcal{C} \cup \{i\} \succsim_j \mathcal{C})$. No player can move to another coalition that it prefers without making some members of that coalition unhappy.

Contractually Individual Stability: A CS $\mathcal{S}$ is *contractually individually stable* iff $(\nexists\, i \in N)\ (\nexists \mathcal{C} \in \mathcal{S} \cup \{\emptyset\}) \mid (\mathcal{C} \cup \{i\} \succ_i \mathcal{S}(i))$ *and* $(\forall j \in \mathcal{C}, \mathcal{C} \cup \{i\} \succsim_j \mathcal{C})$ *and* $(\forall j \in \mathcal{S}(i) \setminus \{i\}, \mathcal{S}(i) \setminus \{i\} \succsim_j \mathcal{S}(i))$. No player can move to a coalition it prefers so that the members of the coalition it leaves and it joins are better off.

If a CS is core-stable, no subset of agents has incentive to leave its respective coalition to form a new one; this is the classic definition of the core in the context of hedonic games. In the other stability solution concepts, the possible deviations feature a single agent i that leaves its current coalition $\mathcal{C}_1$ to join a different existing coalition $\mathcal{C}_2$ or to form a singleton coalition. The difference between the three notions is the behavior of the members of $\mathcal{C}_1$ and $\mathcal{C}_2$. For a Nash stable S, they are not considered: if i prefers to join an existing coalition, it is a valid deviation. This assumes that the agents in $\mathcal{C}_2$ will accept agent i, which is quite optimistic. In individual stability, the deviation is valid if no agent in $\mathcal{C}_2$ is against accepting agent i. Finally, contractual individual stability adds a constraint on the agents from $\mathcal{C}_1$: they have to agree on agent i leaving them. The three stability concepts have the following inclusion: Nash stability is included in individual stability, which is included in contractual individual stability.

The literature in game theory focuses on finding conditions for the existence of the core. In the AI literature, Elkind and Wooldridge have proposed a succinct representation of Hedonic games Elkind and Wooldridge (2009).

We now turn to the most general definition of an NTU game, which uses a set of outcomes that can be achieved by the coalitions. The formal definition is the following:

DEFINITION 2.32 (NTU Game). *A non-transferable utility game (NTU Game)* $(N, X, V, (\succ_i)_{i \in N})$ *is defined by*

- *a set of agents* N;
- *a set of outcomes* X;
- *a function* $V : 2^N \rightarrow 2^X$ *that describes the outcomes* $V(\mathcal{C}) \subseteq X$ *that can be brought about by coalition* $\mathcal{C}$;
- *a preference relation* $\succ_i$ *(transitive and complete) over the set of outcomes for each agent* i.

Intuitively, $V(\mathcal{C})$ is the set of outcomes that $\mathcal{C}$ can bring about by means of its joint action. The agents have a preference relation over the outcomes.

EXAMPLE 1: *Hedonic games as a special class of NTU games. Let* $(N, (\succsim_i^H)_{i \in N})$ *be a hedonic game.*

- For each coalition $\mathcal{C} \subseteq N$, create a unique outcome $x_{\mathcal{C}}$.
- For any two outcomes x_S and x_T corresponding to coalitions S and T that contains agent i, we define $\succsim_i$ as follows: $x_S \succsim_i x_T$ iff $S \succsim_i^H T$.
- For each coalition $\mathcal{C} \subseteq N$, we define $V(\mathcal{C})$ as $V(\mathcal{C}) = \{x_{\mathcal{C}}\}$.

EXAMPLE 2: *A TU game can be viewed as an NTU game. Let* (N, v) *be a TU game.*

- We define X to be the set of all allocations, that is, $X = \mathbb{R}^n$.
- For any two allocations $(x, y) \in X^2$, we define $\succsim_i$ as follows: $x \succsim_i y$ iff $x_i \geq y_i$.
- For each coalition $\mathcal{C} \subseteq N$, we define $V(\mathcal{C})$ as $V(\mathcal{C}) = \{x \in \mathbb{R}^n \mid \sum_{i \in N} x_i \leq v(\mathcal{C})\}$. $V(\mathcal{C})$ lists all the feasible allocation for the coalition $\mathcal{C}$.

First, we can note that the definition of the core can easily be modified in the case of NTU games.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

DEFINITION 2.33 $core(V) = \{x \in V(N) \mid \nexists \mathcal{C} \subset N, \exists y \in V(\mathcal{C}), \forall i \in \mathcal{C}\ y \succ_i x\}$

An outcome $x \in X$ is blocked by a coalition $\mathcal{C}$ when there is another outcome $y \in X$ that is preferred by all the members of $\mathcal{C}$. An outcome is then in the core when it can be achieved by the grand coalition and it is not blocked by any coalition. As is the case for TU game, it is possible that the core of an NTU game is empty.

3 Applications in multiagent systems

In some applications, agents cooperate and share the value of their joint work. In some cases, it is possible to use results from cooperative games to design the agents. In many other cases, some theory was missing to incorporate some aspects of the environment, for example, uncertainty, time constraints, manipulations to name a few. In this section, we start by introducing two generic applications that have been use by the multiagent systems literature to motivate further research in cooperative games. Then, we provide an overview of other applications that have been used in the multiagent system literature. We end the section with some classes of games that have nice computational properties.

3.1 Task allocation problem

A task allocation problem can be easily represented by a coalition formation problem: a coalition of agents is in charge of performing a task (or a subset of tasks). A task may require multiple agents to be performed because of the following reasons:

- Complementary expertise may be required to perform a complex task, and many approaches assume that no agent has all the required expertise to perform a complex task on its own (Shehory & Kraus, 1998; Kraus *et al*., 2003, 2004; Manisterski *et al*., 2006). In the general case, a task can be decomposed into subtasks, and the agents are able to perform a subset of all possible subtasks.
- All the agents have the required ability or expertise to perform a task, but they do not have enough resources on their own to perform the task. For example, robots have the ability to move objects in a plant, but multiple robots are required to move a heavy box (Shehory & Kraus, 1998; Aknine *et al*., 2004).

In addition, the valuation function of a coalitional game has a simple interpretation: it is the benefit of the group of coalitions when the task is performed. The classical stability problem of coalitional games appears as multiple coalitions may be able to perform a complex task, and some coalitions may be better suited to perform a given task. Ideally, an agent should not have any incentive to join a different coalition to work on a different set of tasks.

A generic task allocation problem can be described as follows: a coalition of agents forms to perform a complex task and each agent in the coalition plays a role in the completion of the task (they can all have the same or complementary roles). The completion of a task is rewarded by a payoff. The cost associated with the task completion depends on the coalition members. The value of the coalition is the net benefit (payoff minus cost) of completing the task. Hence, the task allocation problem is well-modeled by a coalition formation problem where the value of a coalition depends only on its members. Note that in the case where agents are not self-interested, the population of agents as a whole may try to maximize the total benefit of completing the tasks. In this case, the agents are trying to optimize utilitarian social welfare and search for the optimal CS (see Section 4.2).

Task allocation problems may be even more complex. First, the tasks may be interdependent. For example, in a study by Shehory and Kraus (1998), there is a partial precedence order between the tasks. This assumption is of particular importance in the transportation domain. The existence of task dependency may promote cooperation between the agents as advocated by

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

Aknine and Shehory (2005): the dependence between the tasks may translate into a certain form of dependence between the agents. If agents realize this fact, they may reciprocally help each other: agent A may help agent B to perform a task needed for the completion of an important task for agent B, and vice versa.

In the generic model, the cost and benefit depends only on the members of the coalition. In environments where the value of a task depends on its completion time, Kraus *et al.* (2003, 2004) suggested that there should be a a cost associated with the time it takes to decide on a coalition. They propose a variant of the task allocation problem where at each round the reward to perform a task is reduced. This forces the agents to decide rapidly whether to form a coalition for taking advantage of the high reward. The first coalition that accepts the contract gets it, and if multiple coalitions agree one coalition is chosen at random. Agents that are only capable of performing a subset of the subtasks must propose or join a coalition. At each round, they can propose a coalition or accept to be part of one. Unlike in a study by Shehory and Kraus (1998) where all the tasks are known in advance, in these works, a coalition is formed incrementally for each task. The order of the tasks may play an important role in the overall payoff to the agents.

If tasks arrive in a pattern, it may be efficient to form similar coalitions for similar tasks. Abdallah and Lesser (2004) assume the existence of a hierarchy of agents. When an agent gets a task for which it does not have the necessary resources, it can ask the agent above it in the hierarchy to take care of the task. If agents placed below it can solve subtasks of the task, the agent can decompose the task and assign it to the agents below in the hierarchy. Learning can be used to choose which agent can perform the task. Abdallah and Lesser show that learning allows for faster and better task assignments.

Another issue that can arise in the task allocation problem is the need to have overlapping coalitions, that is, to have the possibility that agents are members of multiple coalitions. For example, an agent i may have a unique ability that is required to complete two tasks. If i is restricted to be a member of a single coalition, one task cannot be completed, which would be inefficient in the case where i has enough resources to help performing both tasks. An example is in a transportation domain, if each task is to move an item between two points and a coalition is a set of vehicles that carries the item. Overlapping coalitions would model this problem as multiple items could be moved by the same truck (Shehory & Kraus, 1998).

The task allocation problem is in general a computationally hard problem: when agents are limited to perform a single task, the coalition problem resembles the set partitioning problem. When agents are able to perform multiple tasks, the allocation problem gets closer to the set covering problem. In both cases, these problems are $\mathcal{NP}$-complete (Shehory & Kraus, 1998). A taxonomy is proposed to distinguish different complexity classes of the task allocation problems (Lau & Zhang, 2003) on the basis of three factors: (1) Is the same task likely to be offered again? (2) Does the multiagent system have more than enough/just enough/not enough resources to performing a set of tasks? (3) Is the reward intrinsic to the task, or does it only depend on the members performing the task? They show that some combinations of factors lead to polynomial problems, and other combinations have exponential complexity. Shehory and Kraus (1998) restrict the set of possible coalitions by adding a constraint on the size of the coalition. This assumption is motivated by the fact that negotiation with a large number of partners becomes costly, and over a given size, a coalition of agents will not be able to get any benefit. In this case, the size of the set of possible coalitions is a constant, and hence the problem can be solved in polynomial time (possibly a high order polynomial though).

3.2 *Electronic marketplace*

Coalition formation has also been used to model firms or agents in the electronic marketplace (Tsvetovat *et al.*, 2001; Li & Sycara, 2002; Vassileva *et al.*, 2002; Li *et al.*, 2003; Sarne & Kraus, 2003, 2005; Cornforth *et al.*, 2004; Asselin & Chaib-draa, 2006). The field originated from a paper by Tsvetovat *et al.* (2001) where consumer agents can form a coalition (i.e. a buying group) to

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

benefit from the quantity discount provided by sellers. From the point of view of a system designer, the problem is to form a CS, and each coalition is forming a buying group. Desirable property of the CS formed include to be Pareto optimal, that is, no other CS should give more to a consumer without giving less to another one (Asselin & Chaib-draa, 2006) and social welfare maximization, which provides the greatest revenue to the buyers.

First, this problem can be modeled by coalitional games with NTU, as in practice, a buyer may not pay another buyer to join a buying group (it is recognized that side-payments could allow for more efficient outcomes though). It is the model studied by Asselin and Chaib-draa (2006); their goal is to define protocols that find a Pareto optimal solutions, and they only propose a centralized solution.

One variant of the problem is to consider the cost of searching for other coalition members. For example, there is a cost to advertise the possibility to form a buying group, to look for partners, to negotiate price and payment (Sarne & Kraus, 2003). The goal of the agents is to increase the size of the coalition so that the benefit from forming a coalition is worth the effort. The dilemma is about executing the task with the current configuration or starting a costly search to find new partners. The work of Sarne and Kraus analyzes the equilibrium strategies.

Other variants include the problem introduced by Yamamoto and Sycara (2001). Unlike in the original problem in a study by Tsvetovat *et al*. (2001), a buying group does not correspond to a particular item: each buyer agent can have a list of single items or a disjunction of items. In addition, sellers can bid discount prices to sell large volume of items. This allows a formulation that is closer to a combinatorial auction. The proposed solution assumes that each buying group is managed by an agent that has to solve the following two problems: (1) given the requests from the buyer agents, the manager agent chooses the sellers and buys the appropriate items, and (2) the manager agent chooses the price paid by each buyer agent. To address the first problem, the proposed algorithm performs a greedy search. To answer the second problem, Yamamoto and Sycara use a surplus sharing rule that ensures a payoff distribution, which is in the core. In a study by Li and Sycara (2002), the agents can bid in combinatorial auctions: agents bid a reservation value for a bundle of items. This makes the problem even more complex as a winner determination problem has to be solved and a stable payoff distribution must be found. The mechanism design aspect of this problem can be found in a study by Li *et al*. (2003). Li and Sycara (2002) present an algorithm that computes an optimal coalition and a payoff division in the core, but it is not guaranteed to be of polynomial complexity. Hence, they also present an approximation algorithm that is polynomial.

A mathematical model using first-order differential equations is presented in a study by Lerman and Shehory (2000). The model describes the dynamics of the coalitions and allows for the computation of a steady-state equilibrium. The paper shows that a steady-state equilibrium always exists, and that it yields higher utility gain compared with the case where agents are buying on their own, or when leaving a coalition is not allowed. However, the outcome is not guaranteed to be Pareto optimal.

Vassileva *et al*. address long-term coalitions (Breban & Vassileva, 2002; Vassileva *et al*., 2002): in many other papers, a coalition is formed to complete a given task, and the coalition is disbanded when the task is accomplished. In contrast, the goal is to form a coalition of agents that will collaborate for a long period of time. The decision to leave a coalition and join a new one should also be a function of the trust put in another agents, that is, the belief that they will have successful interaction in the future.

Another application of coalition formation in the context of an electronic marketplace is service-oriented computing. A large number of services are offered on the Internet, at different prices and with different quality. Blankenburg *et al*. (2006) propose the use of service Request Agent that can request one (potentially) complex task and a Service Provider Agent that can provide a service. The latter can also, given a task and a set of service advertisements, compose services to form a plan that implements the task. The service requester agents only pay the service provider agent if the task is performed on time. The service provider agents must evaluate the risk

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

involved in accepting a request. In addition, a service provider may be involved in more than one coalition, that is, it can have multiple clients at the same time. Blankenburg *et al.* propose the use of fuzzy coalitions to allow agents to be members of multiple coalitions. The agents use a measure of risk from the finance literature and they accept a proposal if the risk is below a threshold. To distribute the payoff, Blankenburg *et al.* define the kernel for their fuzzy coalition and use Stearns method to converge to a payoff distribution in the kernel.

3.3 Other domains

Coalitions of agents have also been used in many other application domains, and we list some of them in the following. We start with an application for gathering information (Klusch & Shehory, 1996a, 1996b). An agent is associated with a local database, and to answer a query, an agent may require other agents. When the agents form a coalition, all agents in the coalition must cooperate: the members share some of their private data, for example, dependency information. If an agent does not cooperate, it will not have access to some information schema that are available to members of the coalition. The coalition formation process assumes an utilitarian mechanism, and each agent tries to maximize its expected utility. The BSVis used to determine the payoff distribution in a study by Klusch and Shehory (1996a). A kernel-oriented solution is proposed by Klusch and Shehory (1996b) for the same domain.

Coalitions have been used to track a moving target using a sensor network, a problem introduced by Horling *et al.* (2001). The problem is to ensure that at least three agents are sensing the target at the same time to perform triangulation. The problem becomes complex as the target is moving and sensors and communication can be faulty. In a study by Sims *et al.* (2003), the goal is for the agents to self-organize and form an appropriate coalition to track the target. The paper used a variant of the contract net protocol to negotiate a coalition that will be used throughout the tracking. Two valuation functions are studied (local and social utility) and different protocols are empirically tested. Soh *et al.* (2003) also solve a real-time tracking problem. An initiator agent starts the coalition formation process by contacting the neighboring agents that are most suitable for the particular task and engages in negotiation with each of them. Case-based reasoning is used to choose the most promising negotiation protocol. In addition, reinforcement learning is used to estimate the utility of a coalition. The coalition formation process may or may not succeed.

In machine learning, it is possible to combine the results of different classifier to increase the accuracy of the classification. Aknine and Caillou (2004) and Plaza and Ontañón (2006) applied this idea in a coalition formation setting. For example, in the work of Plaza and Ontañón, agents can form committees (i.e. coalitions) to classify a new species of sponge. Each agent has its own expertise, a set of cases, and uses case-based reasoning for the classification problem. In their work, Plaza and Ontañón show how to decide when a committee is needed and how to select the agents to form a committee for a new species of sponge.

Coalitions of agents have been used in the context of distribution and planning of infrastructure for power transportation (Contreras *et al.*, 1997, 1998; Poon *et al.*, 1999). Poon *et al.* (1999) model the trading process between firms that generate, transmit, or distribute power using agents. Agents rank other agents by possible gains and send the BSV of the potential partner when it makes an offer. If both agents send requests to each other, it is beneficial for them to work together and they form a single entity. The process iterates until no further improvement is possible. In the power transmission domain, the problem is to decide whether or not to create a new line or a new plant, and if so, how to share the cost between the different parties involved. Contreras *et al.* (1997, 1998) use similar solutions as Yeung and Poon.

Coalitions of agents have been used in the context of planning and scheduling. For example, Pěchouček *et al.* (2002) tackle the problem of planning humanitarian relief operations and the problem of production planning in a study by Pěchouček *et al.* (2000). For the humanitarian relief operation scenario, different organizations can form coalitions to be more efficient and provide optimal help to the people. However, the different groups that have different capabilities can also have different goals,

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

and hence they might not want to disclose all available information. In that context, the authors propose a formation of alliances: provided some public information, the agents seek to form groups of agents with the same kind of goals. These alliances can be viewed as long-term agreements between agents, and alliances define a partition of the agents. Unlike alliances, coalitions are viewed as short-term agreements to perform a specific task, and to reduce the search space coalitions can form within an alliance. In case of impossibility of forming coalitions within an alliance, agents from different alliances can be used. The authors are interested in the amount of information agents have to disclose: when it sends a request, an agent may reveal private or semi-private information. This can occur when an agent asks an agent of a different alliance to perform a task (revealing that neither it nor its alliance can complete the task). An agent can also decide to disclose private information when it wants to inform other agents, for instance, when they form alliances.

In the context of production planning, instead of using a centralized planning approach, Pěchouček *et al.* want to use local coalition formation to execute tasks in an efficient manner. One requirement is that agents know their possible collaborators well in order to minimize the communication effort, for example, agents have knowledge of the status of surrounding agents, and thus an agent may ask help from another agent if it knows that the agent is not busy. Caillou *et al.* (2002) use the scenario of scheduling classes in a university where a coalition is a schedule. This work considers NTU. Caillou *et al.* propose a protocol where a set of acceptable coalitions is passed from agents to agents, and each time agents can add coalitions or remove coalitions that are not acceptable. The result of the protocol is a Pareto optimal schedule. The protocol also considers reusing existing solutions to compute a solution to a modified problem (e.g. when a class is removed from the schedule, or a professor is coming, previous solutions of the problem can be used to accommodate these changes).

3.4 Some interesting classes of games from the computational point of view

We want to briefly introduce some classes of games that have been studied in the AI literature. Some of these classes of games can be represented more compactly than by using 2^n values, one for each coalition, using an underlying graph structure. In some restricted cases, some solution concepts can be computed efficiently.

Graph games:[2] This class of games was introduced by Deng and Papadimitriou (1994). A game is a pair $\langle V,w\rangle$ where V is the set of agents and is also the set of edges of a weighted undirected graph (V, w) where $w : v^2 \rightarrow \mathbb{R}$ is the weight function, mapping an edge $(i, j) \in V^2$ to a real number. The value of a coalition $\mathcal{C} \subseteq N$ is the sum of the weights of the sub-graph induced by $\mathcal{C}$.

This representation is succinct but not complete (e.g. a majority game cannot be represented using this representation). If all the weights are non-negative, then the game is convex (and consequently, the core is non-empty), and testing the membership in the core can be performed in polynomial time. In addition, the Shapley value can be computed in polynomial time.

Minimum cost spanning tree games: A game is $\langle V, s, w\rangle$ where $\langle V, w\rangle$ is as in a graph game and $s \in V$ is the source node. For a coalition $\mathcal{C}$, we denote by $\Gamma(\mathcal{C})$ the minimum cost spanning tree spanning over the set of edges $\mathcal{C} \cup \{s\}$. The value of a coalition $V \setminus \{s\}$ is given by $\sum_{(i,j) \in \Gamma(\mathcal{C})} w_{i,j}$.

This class of game can model the problem of connecting some agents to a central node played by the source node s. Computing the nucleolus or checking whether the core is non-empty can be done in polynomial time.

Network flow games: A flow network $\langle V, E, c, s, t\rangle$ is composed of a directed graph (V, E) with a capacity on the edge $c : V^2 \rightarrow \mathbb{R}^+$, a source vertex s and a sink vertex t. A network flow is a function $f : E \rightarrow \mathbb{R}^+$ that satisfies the capacity of an edge $(\forall (i, j) \in E,\ f(i, j) \leq c(i, j))$ and that is conserved (except for the source and sink), that is, the total flow arriving at an edge is equal to the total flow leaving that edge $(\forall j \in V,\ \sum_{(i,j) \in E} f(i,j) = \sum_{(j,k) \in E} f(j,k))$. The value of the flow is the amount flowing out of the sink node.

[2] We introduce this class in Section 2.6 about the Shapley value.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

In network flow game (Kalai & Xemel, 1982), $\langle V, E, c, s, t\rangle$, the value of a coalition $\mathcal{C} \subseteq N$ is the maximum value of the flow going through the flow network $\langle \mathcal{C}, E, c, s, t\rangle$.

This class of games can model a situation where some cities share a supply of water or some electricity network. Kalai and Xemel (1982) proved that a network flow game is balanced, and hence it has a non-empty core. Bachrach and Rosenschein (2009) study a threshold version of the game and the complexity of computing power indices.

Affinity games: The class of *affinity games* is a class of hedonic games introduced by Brânzei and Larson (2009a, 2009b). An affinity game is defined using a directed weigthed graph $\langle V, E, w\rangle$ where V is the set of agents, E is the set of directed edges and $w : E \rightarrow \mathbb{R}$ is the weight of the edges. $w(i, j)$ is the value of agent i when it is associated with agent j. The value of agent i for coalition $\mathcal{C}$ is $v_i(\mathcal{C}) = \sum_{j \in \mathcal{C}} w(i, j)$.

Some special classes of affinity games have a non-empty core (e.g. when the weights are all positive or all negative). In these games, there may be a trade-off between stability and efficiency (in the sense of maximizing social welfare) as the ratio between an optimal CS and a stable CS may be infinite.

Skill games: This class of games, introduced by Bachrach and Rosenschein (2008) is represented by a triplet $\langle N, S, \mathrm{T}, u\rangle$ where N is the set of agents, S is the set of skills, T is the set of tasks, and $u : 2^T \rightarrow \mathbb{R}$ provides a value to each set of tasks that is completed. Each agent i has a set of skills $S(i) \subseteq S$, each task t_i requires a set of skills $S(t_i) \subseteq S$. A coalition $\mathcal{C}$ can perform a task t when each skill needed for the task is the skill of at least a member of $\mathcal{C}$ (i.e. $\forall s \in S(t)$, $\exists i \in \mathcal{C}$ such that $S(i) = s$). The value of a coalition $\mathcal{C}$ is $u(T_{\mathcal{C}})$ where $T_{\mathcal{C}}$ is the set of tasks that can be performed by $\mathcal{C}$.

This representation is exponential in the number of agents, but variants of the representation lead to polynomial representation. For example, when the value of a coalition is the number of tasks it can accomplish, or when each task has a weight and the value of a coalition is the sum of the weights of the accomplished tasks. In general, computing the solution concepts with these polynomial representation is hard. However, in some special cases, checking whether the core is empty or computing an element of the core can be carried out in polynomial time. The problem of finding an optimal CS is found in a study by Bachrach *et al.* (2010).

Some more papers are studying the computational complexity of some subclasses of games, for example, Greco *et al.* (2009) and Aziz *et al.* (2010), to name a few. We do not want to provide a full account of complexity problem, which could be the topic of a survey paper on its own.

4 Coalition structure generation problem and related issues

In the previous sections, the focus was on individual agents that are concerned with their individual payoff. In this section, we consider TU games (N, v) in which agents are concerned only about the society's payoff: the agents' goal is to maximize utilitarian social welfare. The actual payoff of the agent or the value of its coalition is not of importance in this setting, only the total value generated by the population matters. This is particularly interesting for multiagent systems designed to maximize some objective functions. In the following, an optimal CS denotes a CS with maximum social welfare. This may model multiagent systems that are designed to optimize an objective function.

More formally, we consider a TU game (N, v), and we recall that a coalition structure (CS) $s = \{\mathcal{S}_1, \cdots, \mathcal{S}_m\}$ is a partition of N, where $\mathcal{S}_i$ is the i^{th} coalition of agents, and $i \neq j \Rightarrow \mathcal{S}_i \cap \mathcal{S}_j = \emptyset$ and $\cup_{i \in [1\ldots m]} \mathcal{S}_i = N$. $\mathscr{S}$ denotes the set of all CSs. The goal of the multiagent system is to locate a CS that maximizes utilitarian social welfare. In other words, the problem is to find an element of $\mathrm{argmax}_{s \in S} \sum_{\mathcal{S} \in s} v(\mathcal{S})$.

The space $\mathscr{S}$ of all CSs can be represented by a lattice, and an example for a population of four agents is provided in Figure 3. The first level of the lattice consists of only the CS corresponding to the grand coalition $N = \{1, 2, 3, 4\}$, and the last level of the lattice contains CS containing singletons only, that is, coalitions containing a single member. Level i contains all the CSs with exactly i coalitions. The number of CSs at level i is $\mathcal{S}(|N|, i)$, where S is the Stirling Number of

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

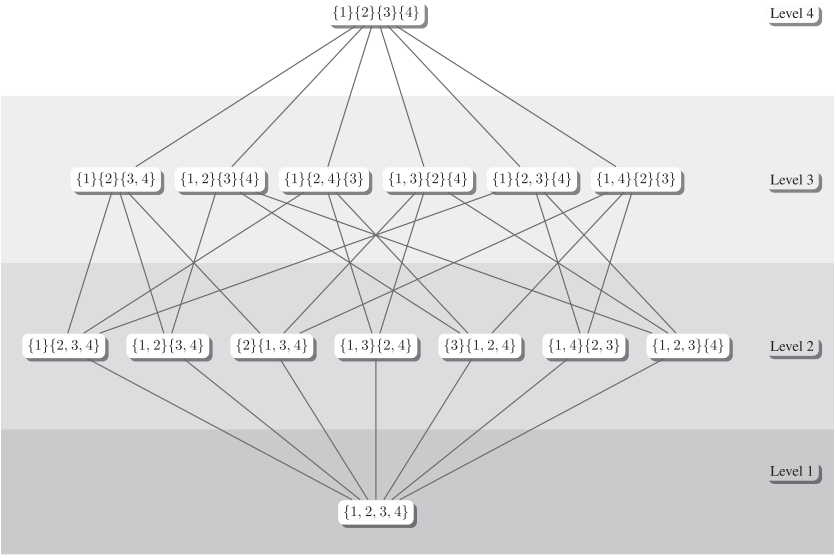


Figure 3 Set of CSs for four agents

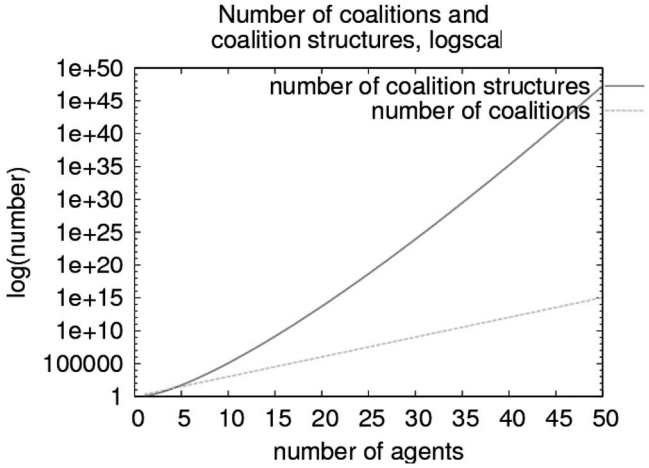


Figure 4 Number of CSs in a population of n agents

the Second Kind[3]. The Bell number, $\mathscr{B}(n)$, represents the total number of CSs with n agents, $\mathscr{B}(n) = \sum_{i=0}^{n} S(n,k)$. This number grows exponentially, as shown in Figure 4, and is $O(n^n)$ and $\omega(n^{\frac{n}{2}})$ (Sandholm *et al.*, 1999). When the number of agents is relatively large, for example, $n \geq 20$, exhaustive enumeration may not be feasible.

The actual issue is the search of the optimal CS. Sandholm *et al.* (1999) show that, given a TU game (N, v), finding the optimal CS is an $\mathcal{NP}$-complete problem. In the following, we will consider centralized search where a single agent is performing the search, as well as the more interesting case of decentralized search where all agents make the search at the same time on different parts of the search space. Before doing so, we review some work where the valuation function v is not known in advance. In a real application, these values need to be computed, and this may be an issue on its own if the computations are hard, as illustrated by an example by Sandholm and Lesser (1997), where the computation of a value requires to solve a traveling salesman problem.

[3] $\mathcal{S}(n,m)$ is the number of ways of partitioning a set of n elements into m non-empty sets.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

4.1 Sharing the computation of the coalition values

Before searching the space of CSs to find an optimal CS, the agents may need to compute the value of each of the coalitions. We are interested in a decentralized algorithm that computes all the coalition values in a minimal amount of time, and that requires minimum communication between the agents.

Shehory and Kraus (1998) were the first to propose an algorithm to share the computation of the coalition values. In their algorithm, the agents negotiate which computation is performed by which agent, which is quite demanding. Rahwan and Jennings proposed an algorithm where agents, once they agree on an identification for each agent participating in the computation, know exactly which coalition values to compute. This algorithm, called DCVC (Rahwan & Jennings, 2007), outperforms the one by Shehory and Kraus. The key observation is that in general it should take longer to compute the value of a large coalition compared with a small coalition (i.e. the computational complexity is likely to increase with the size of the coalition as more agents have to coordinate their activities). Their method improves the balance of the loads by distributing coalitions of the same size to all agents. By knowing the number of agents n participating in the computation and an index number (i.e. an integer in the range $\{0 \ldots n\}$), the agents determine for each coalition size which coalition values to compute. The algorithm can also be adapted when the agents have different known computational speed so as to complete the computation in a minimum amount of time.

4.2 Searching for the optimal coalition structure

The difficulty of searching for the optimal CS lies in the large search space, as recognized by existing algorithms, and this is even more true in the case where there exists externalities (i.e. when the valuation of a coalition depends on the CS). For TU games with no externalities, some algorithms guarantee finding CSs within a bound from the optimum when an incomplete search is performed. Unfortunately, such guarantees are not possible for games with externalities. We shortly discuss these two cases in the following.

4.2.1 Games with no externalities

Sandholm *et al.* (1999) proposed the first algorithm that searches through a lattice as presented in Figure 3. Their algorithm guarantees that the CS found, $\mathcal{S}$, is within a bound from the optimal $\mathcal{S}^*$ when a sufficient portion of the lattice has been visited. The bound considered is $\frac{v(\mathcal{S})}{v(\mathcal{S}^*)} \leq \mathcal{K}$. They prove that, to ensure a bound, it is necessary to visit a least 2^{n-1} CSs (Theorems 1 and 3 in Sandholm *et al.*, 1999), which corresponds to the first two levels of the lattice, that is, the algorithm needs to visit the grand coalition, and all the CSs composed of two coalitions. The bound improves each time a new level is visited. An empirical study of different strategies for visiting the other levels is presented in a study by Larson and Sandholm (2000). Three different algorithms are empirically tested over characteristic functions with different properties: (1) subadditive, (2) superadditive, and (3) picked from a uniform distribution in [0, 1] or in $[0, |\mathcal{C}|]$ (where $|\mathcal{C}|$ is the size of the coalition). The performance of the heuristics differs over the different type of valuation functions, demonstrating the importance of the properties of the characteristic function in the performance of the search algorithm.

The algorithm by Dang and Jennings (2004) improves the one by Sandholm *et al.* (1999) for low bounds from the optimal. For large bounds, both algorithms visit the first two levels of the lattice. Then, when the algorithm by Sandholm *et al.* continues by searching each level of the lattice, the algorithm by Dang and Jennings only searches specific subset of each level to decrease the bound faster. This algorithm is anytime, but its complexity is not polynomial.

These algorithms were based on a lattice as the one presented in Figure 3 where a CS in level i contains exactly i coalitions. The best algorithm to date has been developed by Rahwan *et al.* and uses a different representation called integer-partition (IP) of the search space. It is an anytime algorithm that has been improved over a series of papers: Rahwan *et al.* (2007a, 2007b),

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

Rahwan and Jennings (2008a, 2008b), and Rahwan *et al* (2009). In this representation, the CSs are grouped according to the sizes of the coalitions they contain, which is called a configuration. For example, for a population of four agents, the configuration {1, 3} represents CSs that contain a coalition with a singleton and a coalition with three agents. A smart scan of the input allows to search the CSs with two coalitions: the grand coalition and the CS containing singletons only. In addition, during the scan, the algorithm computes the average and maximum value for each coalition size. The maximum values can be used to prune the search space. When constructing a configuration, the use of the maximum values of a coalition for each size permits the computation of an upper bound of the value of a CS that follows that configuration, and if the value is not greater than the current best CS, it is not necessary to search through the CSs with that configuration, which prunes the search tree. Then, the algorithm searches the remaining configurations, starting with the most promising ones. During the search of a configuration, a branch and bound technique is used. In addition, during the search, the algorithm is designed so that no CS is evaluated twice. Empirical evaluation shows that the algorithm outperforms any other current approach over different distributions used to generate the values of the coalitions.

More recently, Service and Adams (2010, 2011) designed an algorithm that uses dynamic programming and that guarantees a constant factor approximation ratio r in a given time. In particular, the latest algorithm by Service and Adams (2010) guarantees a factor of $\frac{1}{8}$ in $O(2^n)$. Finally, Ueda *et al.* (2010) propose to use a different representation, assuming that the value of a coalition is the optimal solution of a distributed constraint optimization problem (DCOP). The algorithm uses a DCOP solver and guarantees a bound from the optimum. Currently, it is difficult to compare all these different approaches.

4.2.2 Games with externalities

The previous algorithm explicitly uses the fact that the valuation function only depends on the members of the coalition, that is, has no externalities. When this is not the case, that is, when the valuation function depends on the CS, it is still possible to use some algorithms, for example, the one proposed by Larson and Sandholm (2000), but the guarantee of being within a bound from the optimal is no longer valid. Sen and Dutta (2000) use genetic algorithms techniques to carry out the search. The use of such technique only assumes that there exists some underlying patterns in the characteristic function. When such patterns exist, the genetic search makes a much faster improvement in locating higher-valued CS compared with the level-by-level search approach. One downside of the genetic algorithm approach is that there is no optimality guarantee. Empirical evaluation, however, shows that the genetic algorithm does not take much longer to find a solution when the value of a coalition does depend on other coalitions.

More recently, Rahwan *et al.* (2009) and Michalak *et al.* (2008) consider the problem for some class of externalities and modify the IP algorithm for the games with externalities; however, they assume games with negative or positive spillovers. Banerjee and Kraemer (2010) introduce a representation to represent games in partition function games using types: each agent has a single type. They make two assumptions on the nature of the externalities (based on the notions of competition and complementation) and they show that games with negative or positive spillovers are special cases. They provide a branch and bound algorithm for the general setting. They also provide a worst-case initial bound.

5 Issues for applying cooperative games

We now highlight issues that have emerged from the applications presented in Section 3. The protocols and algorithms we cited there present some solutions to these issues. Some additional issues remain unsolved, for example, dealing with agents that can enter and leave the environment at any time in an open, dynamic environment. None of the current protocols can handle these issues without restarting computation, and only few approaches consider how to reuse the already computed solution (Caillou *et al.*, 2002; Belmonte *et al.*, 2004).

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

5.1 Stability and dynamic environments

Real-world scenarios often present dynamic environments. Agents can enter and leave the environment at any time, the characteristics of the agents may change with time, the knowledge of the agents about the other agents may change, etc.

The game-theoretic stability criteria are defined for a fixed population of agents and the introduction of a new agent in the environment requires significant computation to update a stable payoff distribution. For example, for the kernel, all the agents need to check whether any coalition that includes the new agent changes the value of the maximum surplus, which requires reevaluating $O(2^n)$ coalitions. Given the complexity of the stability concept, one challenge that is faced by the multiagent community is to develop stability concepts that can be easily updated when an agent enters or leaves the environment.

In addition, if an agent drops during the negotiation, this may cause problems for the remaining agents. For example, a protocol that guarantees a kernel stable payoff distribution is shown not to be 'safe' when the population of agents is changing: if an agent i leaves the formation process without notifying other agents, the other agents may complete the protocol and find a solution to a situation that does not match the reality. Each time a new agent enters or leaves the population, a new process needs to be restarted (Blankenburg & Klusch, 2004).

In an open environment, manipulations will be impossible to detect: agents may use multiple identifiers (or false names) to pretend to be multiple agents, or the other way around, multiple agents may collude and pretend to be asingle agent, or agents can hide some of their skills. Hence, it is important to propose solution concepts that are robust against such manipulations. We will come back later to some of the solution that have been proposed: the anonymity-proof core (Yokoo *et al.*, 2005) and anonymity-proof Shapley value (Ohta *et al.*, 2009).

5.2 Uncertainty about knowledge and task

In real-world scenarios, agents will be required to handle some uncertainty. Different sources of uncertainty have been considered in the literature:

- The valuation function is an approximation (Sandholm & Lesser, 1997) and agents may not use the same algorithm. Hence, the agents may not know what is the true value.
- Agents may not know some tasks (Blankenburg & Klusch, 2004) or the value of some coalitions. In such cases, the agents play a different coalitional game that may reduce the payoff of some agents compared with the solution of the true game.
- Some information is private, that is, an agent knows some property about itself, but does not know it for other agents. In a study by Kraus *et al.* (2003), it is the cost incurred by other agents to perform a task that is private. In studies by Chalkiadakis *et al.* (2009) and Chalkiadakis and Boutilier (2012), agents have a private type, and the valuation function depends on the types of the coalition's members.
- Uncertainty about the outcome of an action (Chalkiadakis & Boutilier, 2012): when a coalition makes an action, some external factors may influence the outcome of the actions. This can be captured by a probability of an outcome, given the action taken and the type of the members of the coalition.
- There are multiple possible worlds (Ieong & Shoham, 2008), which models the different possible outcomes of the formation of a coalition. Agents know a probability distribution over the different worlds. In addition, an agent may not be able to distinguish some worlds as it lacks information and they know a partition of the worlds (called information sets), each set of the partition represent worlds that appears as indistinguishable.

Some authors also consider that there is uncertainty in the valuation function without modeling a particular source: for example, in a study by Ketchpel (1994b), each agent has an expectation of the valuation function. In studies by Blankenburg *et al.* (2003) and Blankenburg and Klusch (2005),

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

fuzzy sets are used to represent the valuation function. In the first paper, the agents enter bilateral negotiations to negotiate Shapley value, and in the second paper they define a fuzzy version of the kernel.

In the uncertainty model by Ieong and Shoham (2008), the definition of the core depends on the time one reasons about it. They proposed three different definitions of the core that depend on the timing of the evaluation: before the world is drawn or *ex-ante*, not much information can be used; after the world is drawn but before it is known, also called *ex-interim*, an agent knows to which set of its information set the real world belongs, but does not know which one; finally when the world is announced to the agent or *ex-post*, everything is known.

The model by Chalkiadakis and Boutilier (2012) combines uncertainty about the agent types and uncertainty about the outcome of the action taken by the coalition. Each agent has a probabilistic belief about the types of the other agents in the population. Chalkiadakis and Boutilier propose a definition of the core, the Bayesian core (introduced in a study by Chalkiadakis & Boutilier, 2004) in which no agent has the belief that there exists a better coalition to form. As it may be difficult to obtain all the probabilities and reason about them, Chalkiadakis *et al.* (2009) propose to use a 'point' belief: an agent guesses the type of the other agents and reason with these guesses. The paper analyzes the core, simple games (proving that the core of a simple game is non-empty iff the game has a veto player) and some complexity result in this games with belief.

5.3 Safety and robustness

It is also important that the coalition formation process is robust. For instance, communication links may fail during the negotiation phase. Hence, some agents may miss some components of the negotiation stages. This possibility is found in a study by Blankenburg and Klusch (2004) for the KCA protocol (Klusch & Shehory, 1996b): coalition negotiations are not safe when some agents become unavailable (intentionally or otherwise). In particular, the payoff distribution is not guaranteed to be kernel-stable. Belmonte *et al.* (2004) empirically study the robustness of the use of a central algorithm introduced in their paper (Belmonte *et al.*, 2002): the cost to compute a task allocation and payoff distribution in the core is polynomial, but it can still be expensive. In the case of agent failure, the computation needs to be repeated. Belmonte *et al.* propose an alternative payoff division model that avoids such a recomputation, but the solution is no longer guaranteed to be in the core, it is only close to the core. There is a trade-off between computational efficiency and the utility obtained by the agent. They conclude that, when the number of agents is small, the loss of utility compared with the optimal is small; hence, the improvement of the computational efficiency can be justified. However, for a larger number of agents, the loss of utility cannot justify the improvement in computational cost.

5.3.1 Protocol manipulation

When agents send requests to search for members of a coalition or when they accept to form a coalition, the protocol may require disclosure of some private information (Pěchouček *et al.*, 2002). When the agents reveal some of their information, the mechanism must ensure that there is no information asymmetry that can be exploited by some agents (Blankenburg *et al.*, 2005). To protect a private value, some protocol (Blankenburg & Klusch, 2004) may allow the addition of a constant offset to the private value, as long as this addition does not impact the outcome of the negotiation.

Belmonte *et al.* (2004) study the effect of deception and manipulation of their model. They show that some agents can benefit from falsely reporting their cost. In some other approaches (Blankenburg & Klusch, 2004; Conitzer & Sandholm, 2004), even if it is theoretically possible to manipulate the protocol, it is not possible in practice as the computational complexity required to ensure higher outcome to the malevolent agent is too high. For example, Conitzer and Sandholm (2004) show that manipulating marginal contribution-based value division scheme is $\mathcal{NP}$-hard (except when the valuation function has other properties, such as being convex).

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

Other possible protocol manipulations include hiding skills, using false names, colluding, etc. The traditional solution concepts can be vulnerable to false names and to collusion (Yokoo *et al.*, 2005). To address these problems, it is beneficial to define the valuation function in terms of the required skills instead of defining it over the agents: only skills, not agents, should be rewarded by the characteristic function. In that case, the solution concept is robust to false names, collusion, and their combination. But the agents can have incentive to hide skills. A straight, naive decomposition of the skills will increase the size of the characteristic function, and Yokoo *et al.* (2006) propose a compact representation in this case.

5.4 Communication

Whereas one purpose of better negotiation techniques may be to improve the quality of the outcome for the agents, other goals may include decreasing the time and the number of messages required to reach an agreement. For example, learning is used to decrease negotiation time in a study by Soh and Tsatsoulis (2002). The motivation of Lerman's work (Lerman & Shehory, 2000) is to develop a coalition formation mechanism that has low communication and computation cost. In another work, the communication costs are included in the characteristic function Tohmé and Sandholm (1999).

The communication complexity of some protocols has been derived. For instance, the exponential protocol in a study by Tohmé and Sandholm (1999) and the coalition algorithm for forming BSV stable coalition in a study by Klusch and Shehory (1996a) have communication complexity of $O(n^2)$, the negotiation-based protocol in a study by Tohmé and Sandholm (1999) is $O(n^2 2^n)$, and it is $O(n^k)$ for the protocol in a study by Shehory and Kraus (1998) (where k is the maximum size of a coalition). The goal of Procaccia and Rosenschein (2006) is to analyze the communication complexity of computing the payoff of a player with different stability concepts: they find that it is $\Theta(n)$ when either the Shapley value, the nucleolus, or the core are used.

5.5 Scalability

When the population of heterogeneous agents is large, discovering the relevant agents to perform a task may be difficult. In addition, if all agents are involved in the coalition formation process, the cost in time and computation will be large. To alleviate this scalability issue, a hierarchy of agents can be used (Abdallah & Lesser, 2004). When an agent discovers a task that can be addressed by agents below this agent in the hierarchy, the agent picks the best of them to perform the task. If the agents below cannot perform the task, the agent passes the task to the agent above it in the hierarchy and the process repeats. The notion of clans (Griffiths & Luck, 2003) and congregations (Brooks & Durfee, 2003), where agents gather together for a long period, have been proposed to restrict the search space by considering only a subset of the agents (see Section 5.6).

5.6 Long term vs. short term

In general, a coalition is a short-lived entity that is 'formed with a purpose in mind and dissolve when that need no longer exists, the coalition ceases to suit its designed purpose, or critical mass is lost as agents depart' (Horling & Lesser, 2004). It can be beneficial to consider the formation of long-term coalitions, or the process of repeated coalition formation involving the same agents. Vassileva *et al.* (2002) explicitly study the long-term coalitions, and in particular the importance of trust in this content. Brooks and Durfee (2003) refer to a long-term coalition as a congregation. The purpose of a congregation is to reduce the number of candidates for a successful interaction: instead of searching the entire population, agents will only search in the congregation. The goal of a congregation is to gather agents, with similar or complementary expertise to perform well in an environment in the long run, which is not very different from a coalition. The only difference is that group rationality is not expected in a congregation. The notion of congregation is similar to

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

the notion of clans (Griffiths & Luck, 2003): agents gather not for a specific purpose, but for a long-term commitment. The notion of trust is paramount in the clans, and sharing information is seen as another way to improve performance.

5.7 *Fairness*

Stability does not necessarily imply fairness. For example, let us consider two CSs $\mathcal{S}$ and $\mathcal{T}$ with associated kernel–stable payoff distribution x_S and x_T. Agents may have different preferences between the CSs. It may even be the case that there is no CS that is preferred by all agents. If the optimal CS is formed, some agents, especially if they are in a singleton coalition, may suffer from the choice of this CS. Airiau and Sen (2010) propose a modification of the kernel to allow side-payment between coalitions to compensate such agents.

Airiau and Sen (2009) consider partition function games with externalities. They consider a process where, in turns, agents change coalition to improve their immediate payoff. They propose that the agents share the maximal social welfare, and the size of the share is proportional to the expected utility of the process. The payoff obtained is guaranteed to be at least as high as the expected utility. They claim that using the expected utility as a base of the payoff distribution provides some fairness, as the expected utility can be seen as a global metric of an agent performance over the entire set of possible CSs.

5.8 *Overlapping coalitions*

It is typically assumed that an agent belongs to a single coalition; however, there are some applications where agents can be members of multiple coalitions. As explained in the task allocation domain (see Section 3.1), the expertise of an agent may be required by different coalitions at the same time, and the agent can have enough resources to be part of two or more coalitions. In a traditional setting, the use of the same agent i by two coalitions $\mathcal{C}_1$ and $\mathcal{C}_2$ would require a merge of the two coalitions. This larger coalition U is potentially harder to manage, and *a priori*, there would not be much interaction between the agents in $\mathcal{C}_1$ and $\mathcal{C}_2$, except for agent i. Another application that requires the use of overlapping coalition is tracking targets using a sensor networks (Dang *et al.*, 2006). In this work, a coalition is defined for a target, and as agents can track multiple targets at the same time they can be members of different coalitions.

The traditional stability concepts do not consider this issue. One possibility is for the agent to be considered as two different agents, but this representation is not satisfactory as it does not capture the real power of this agent. Shehory and Kraus (1998) propose a setting with overlapping coalition. Each agent has a capacity, and performing a task may use only a fraction of the agent's capacity. Each time an agent commits to a task, the possible coalitions that can perform a given task can change. A mapping to a set covering problem allows to find the coalition. However, the study of the stability is not considered. Another approach is the use of fuzzy coalition (Blankenburg *et al.*, 2006): agents can be members of a coalition with a certain degree that represents the risk associated with being in that coalition. Other work considers that the agents have different degree of membership, and their payoff depends on this degree (Aubin, 1979; Mares, 2001; Nishizaki & Masatoshi, 2001). The protocols in the study by Lau and Zhang (2003) also allow overlapping coalitions.

More recently, Chalkiadakis *et al.* (2010)[4] have studied the notion of the core in overlapping coalition formation. In their model, each agent has one resource and the agent contributes a fraction of that resource to each coalition it participates in. The valuation function v is then $[0,1]^n \rightarrow \mathbb{R}$. A CS is no longer a partition of the agents: a CS S is a finite list of vectors, one for each 'partial' coalition, that is, $S = (r^1, \ldots, r^k)$. The size of S is the number of coalitions, that is, k. The support of $r^{\mathcal{C}} \in S$ (i.e. the set of indices $i \in N$ such that $r_i^{\mathcal{C}} \neq 0$) is the set of agents

[4] An earlier version is Chalkiadakis *et al.* (2008).

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

forming coalition $\mathcal{C}$. For all $i \in N$ and all coalition $\mathcal{C} \in S$, $r_i^{\mathcal{C}} \in [0,1]^n$ represents the fraction of resource that agent i contributes to coalition $\mathcal{C}$; hence, $\sum_{\mathcal{C} \in S} r_i^{\mathcal{C}} \leq 1$ (i.e. agent i cannot contributes more than 100% of its resource). A payoff distribution for a CS S of size k is defined by a k-tuple $x = (x^1, \ldots, x^k)$ where $x^{\mathcal{C}}$ is the payoff distribution that the agents obtain for coalition $\mathcal{C}$. If an agent is not in the coalition, it must not receive any payoff for this coalition, and hence $(r_i^{\mathcal{C}} = 0) \Rightarrow (x_i^{\mathcal{C}} = 0)$. The total payoff of agent i is the sum of its payoffs over all coalitions $p_i(\mathcal{CS}, x) = \sum_{\mathcal{C}=1}^{k} x_i^{\mathcal{C}}$. The efficiency criterion becomes $\forall r^{\mathcal{C}} \in S, \sum_{i \in N} x_i^{\mathcal{C}} = v(r^{\mathcal{C}})$. An imputation is an efficient payoff distribution that is also individually rational. We denote by $I(S)$ the set of all imputations for the CS S.

We are now ready to define the overlapping core. One issue is the kind of permissible deviations: when an agent deviates, she can completely leave some coalitions, reduce her contribution in other coalitions, or contributes to new coalitions. If she stills contribute to a coalition containing non-deviating agents, how should they behave? They may first refuse to give any payoff to the deviating agent, as she is seen as not trustworthy. However, agents that are not affected by the deviation may consider that the deviators agents did not fail them, and consequently they may continue to share payoffs with the deviators. A last case occurs when the deviators are decreasing their implication in a coalition. This coalition may no longer perform the same tasks, but it can still perform some. If there is enough value to maintain the payoff of the non-deviators, the deviators may be allowed to share the surplus generated. Each of these behaviors give rise to different types of deviations, and consequently different definition of a core: the conservative core, the refined core, and the optimistic core. The paper also provides a characterization of conservative core, properties of the different core, including a result showing that convex overlapping coalitional games have a non-empty core.

5.9 *Trust*

The notion of trust can be an important metric to determine whom to interact with. This is particularly important when the coalition is expected to live for a long term. In a study by Blankenburg *et al*. (2005), an agent computes a probability of success of a coalition, on the basis of a notion of trust, which can be used to eliminate some agents from future consideration. This probability is used to estimate the value of different coalitions and help the agent in deciding which coalition to join or form. In a study by Vassileva *et al*. (2002), the decision to leave or join a coalition is function of the trust put in other agents. In this paper, the concept of trust is defined as a belief that agents will have successful interaction in the future; hence, trust is used to consider a subset of the entire population of agents for the formation of future coalitions. Trust is used to compute coalitions, but agents do not compute a payoff distribution. Another work that emphasizes trust is the study by Griffiths and Luck (2003), which introduces the concept of clans. A clan is formed by agents that trust each other with long-term commitments. Given the trust and an estimate of local gain, agents can accept to join a clan. The idea behind this work is that agents that trust each other will be collaborative. Moreover, when an agent needs to form a coalition of agents, it will only search partners in the clan, which reduces the search space. Trust can therefore be very effective for scaling up in large society of agents.

5.10 *Learning*

When agents have to repeatedly form coalitions in the presence of the same set of agents, learning can be used to improve performance of the coalition formation process, both in terms of speed of the process and in terms of better valuation.

A basic model of iteratively playing many coalitional games is presented in a study by Mérida-Campos and Willmott (2004): at each time step, a task is offered to agents that are already organized into coalitions. The task is awarded to the best coalition. The model is made richer in a study by Mérida-Campos and Willmott (2006) where the agents can estimate the value of

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

a coalition and have a richer set of actions: as the agents can fire members from a coalition, join a different coalition, or leave a coalition to replace some agents in a different coalition. However, in both works, the agents are not learning; they have a set of static strategies. Empirical experiments compare the results over populations using either the same strategy or a mix of strategies.

Chalkiadakis and Boutilier (2004, 2008, 2012) also consider a repeated coalition formation problem, sequentially. The setting is a task allocation problem where agents know their own types (i.e. skill to perform some type of tasks), but do not know the ones of other agents in the population. Each time a coalition is formed, the agents will receive a value for that coalition. From the observation of this value, the agents can update a belief about the types of other agents. When an agent is reasoning about which coalition to form, it uses its beliefs to estimate the value of the coalition. This problem can be formulated using a POMPD (Partially Observable Markov Decision Process) where the agents are maximizing the long-term value of their decision over the repetition of the coalition formation process. Solving a POMPD is a difficult task, and the POMPD for the coalition formation problem grows exponentially with the number of agents. In a study by Chalkiadakis and Boutilier (2004), a myopic approach is proposed. More recently, Chalkiadakis & Boutilier (2008) propose additional algorithms to solve that POMPD and empirically compare the solutions.

6 Conclusion

Cooperative game theory has been studied from many decades now, and this survey shows this work relevant to multiagent systems. A TU game is a simple mathematical object: it consists of a set of players N and a valuation that maps any subset of N to a real value. We saw that many solution concepts have been proposed by game theory, and each of them has its pros and cons. The core is perhaps the most intuitive solution promoting stability, but it may be empty. The nucleolus is theoretically quite appealing; however, it is difficult to compute. The Shapley value, also difficult to compute, does not provide stability but provides fairness.

Using a naive representation, the size of the input is exponential, and hence it is computationally expensive to compute a solution concept. One contribution of the multiagent systems community to game theory has been to investigate the computational properties of classes of games for various solution concepts under various representations. We have tried to indicate these results in this paper, but this topic would probably require a survey on its own.

Another contribution of the multiagent system community has been to propose solutions to issues that are less interesting from a game theoretic point of view. For example, searching for an optimal coalition structure or considering issues presented in Section 5, such as uncertainty, the problem of overlapping coalition, learning, etc.Some issues have not been successfully treated, for example, the formation of coalition in open environments.

Acknowledgment

This article originated from the related work of my dissertation, under the supervision of Sandip Sen. I thank him for reviews and comments about earlier versions of this paper. I also thank Krzysztof R. Apt. for sharing his notes. I also thank Ulle Endriss for his comments. An earlier version of this article has been used as lecture notes for the 12th European Agent Systems Summer School held in Saint-Etienne, France.

References

Abdallah, S. & Lesser, V. 2004. Organization-based cooperative coalition formation. In *Proceedings of the IAT 04*, Beijing, China.

Airiau, S. & Sen, S. 2009. A fair payoff distribution for myopic rational agents. In *Proceedings of the Eighth International Conference on Autonomous Agents and Multiagent Systems (AAMAS-09)*, Budapest, Hungary.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

Airiau, S. & Sen, S. 2010. On the stability of an optimal coalition structure. In *Proceedings of the 19th European Conference on Artificial Intelligence (ECAI-2010)*, Lisbon, Portugal, 203–208.

Aknine, S. & Caillou, P. 2004. Agreements without disagreements. In *Proceedings of the 16th Eureopean Conference on Artificial Intelligence, ECAI'2004*, Valencia, Spain, 3–7.

Aknine, S., Pinson, S. & Shakun, M. F. 2004. A multi-agent coalition formation method based on preference models. *Group Decision and Negotiation* **13**(6), 513–538.

Aknine, S. & Shehory, O. 2005. Coalition formation: concessions, task relationships and complexity reduction. *ArXiv Computer Science e-prints*.

Asselin, F. & Chaib-draa, B. 2006. Performance of software agents in non-transferable payoff group buying. *Journal of Experimental and Theoretical Artificial Intelligence* **18**(1), 1–32.

Aubin, J.-P. 1979. *Mathematical Methods of Game and Economic Theory*. North-Holland.

Aumann, R. J. & Drèze, J. H. 1974. Cooperative games with coalition structures. *International Journal of Game Theory* **3**(4), 217–237.

Aziz, H., Brandt, F. & Harrenstein, P. 2010. Monotone cooperative games and their threshold versions. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems – Volume 1*, *AAMAS '10*, van der Hoek, W., Kaminka, G. A., Lespérance, Y., Luck, M. & Sen, S. (eds). International Foundation for Autonomous Agents and Multiagent Systems, 1107–1114.

Bachrach, Y., Elkind, E., Meir, R., Pasechnik, D., Zuckerman, M., Rothe, J. & Rosenschein, J. 2009. The cost of stability in coalitional games. In *SAGT'09: 2nd International Symposium of Algorithmic Game Theory*, Paphos, Cyprus, 122–134.

Bachrach, Y., Meir, R., Jung, K. & Kohli, P. 2010. Coalitional structure generation in skill games. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI-10)*, Atlanta, Georgia, USA, 703–708.

Bachrach, Y. & Rosenschein, J. S. 2008. Coalitional skill games. In *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-08)*, Estoril, Portugal, 1023–1030.

Bachrach, Y. & Rosenschein, J. 2009. Power in threshold network flow games. *Autonomous Agents and Multi-Agent Systems* **18**, 106–132 10.1007/s10458-008-9057-6.

Banerjee, B. & Kraemer, L. 2010. Coalition structure generation in multi-agent systems with mixed externalities. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems – Volume 1*, *AAMAS '10*, van der Hoek, W., Kaminka, G. A., Lespérance, Y., Luck, M. & Sen, S. (eds). International Foundation for Autonomous Agents and Multiagent Systems, 175–182.

Belmonte, M.-V., Conejo, R., de-la Cruz, J.-L. P. & Ruiz, F. T. 2002. A stable and feasible payoff division for coalition formation in a class of task oriented domains. In *ATAL '01: Revised Papers from the 8th International Workshop on Intelligent Agents VIII*, Meyer, J.-J. Ch. & Tambe, M. (eds). Springer-Verlag, 324–334.

Belmonte, M.-V., Conejo, R., de-la Cruz, J.-L. P. & Ruiz, F. T. 2004. A robust deception-free coalition formation model. In *Proceedings of the 2004 ACM Symposium on Applied Computing (SAC '04)*, Haddad, H., Omicini, A., Wainwright, R. L. & Liebrock, L. M. (eds). ACM Press, 469–473.

Blankenburg, B., Dash, R. K., Ramchurn, S. D., Klusch, M. & Jennings, N. R. 2005. Trusted kernel-based coalition formation. In *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems*, Dignum, F., Dignum, V., Koenig, S., Kraus, S., Singh, M. P. & Wooldridge, M. (eds). ACM Press, 989–996.

Blankenburg, B., He, M., Klusch, M. & Jennings, N. R. 2006. Risk-bounded formation of fuzzy coalitions among service agents. In *Proceedings of 10th International Workshop on Cooperative Information Agents*, Edinburgh, UK.

Blankenburg, B. & Klusch, M. 2004. On safe kernel stable coalition forming among agents. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems*, New York, NY, USA.

Blankenburg, B. & Klusch, M. 2005. Bsca-f: efficient fuzzy valued stable coalition forming among agents. In *Proceedings of the IEEE International Conference on Intelligent Agent Technology (IAT)*, Skowron, A., Barthès, J.-P. A., Jain, L. C., Sun, R., Morizet-Mahoudeaux, P., Liu, J. & Zhong, N. (eds). IEEE Computer Society Press, 732–738.

Blankenburg, B., Klusch, M. & Shehory, O. 2003. Fuzzy kernel-stable coalitions between rational agents. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-03)*, Rosenschein, J. S., Wooldridge, M., Sandholm, T. & Yokoo, M. (eds). ACM Press, 9–16.

Bloch, F. 2003. Non-cooperative models of coalition formation in games with spillover. In *The Endogenous Formation of Economic Coalitions*, Carraro, C. (ed.). Edward Elgar, chapter 2, 35–79.

Bogomolnaia, A. & Jackson, M. O. 2002. The stability of hedonic coalition structures. *Games and Economic Behavior* **38**(2), 201–230.

Brânzei, S. & Larson, K. 2009a. Coalitional affinity games. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI-09)*, 1319–1320.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

Brânzei, S. & Larson, K. 2009b. Coalitional affinity games. In *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2009)*, Budapest, Hungary, 79–84.

Breban, S. & Vassileva, J. 2002. A coalition formation mechanism based on inter-agent trust relationships. In *AAMAS '02: Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems*, Gini, M., Ishida, T., Castelfranchi, C. & Lewis Johnson, W. (eds). ACM Press, 306–307.

Brooks, C. H. & Durfee, E. H. 2003. Congregation formation in multiagent systems. *Autonomous Agents and Multi-Agent Systems* **7**(1–2), 145–170.

Caillou, P., Aknine, S. & Pinson, S. 2002. A multi-agent method for forming and dynamic restructuring of pareto optimal coalitions. In *AAMAS '02: Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems*, Gini, M., Ishida, T., Castelfranchi, C. & Lewis Johnson, W. (eds). ACM Press, 1074–1081.

Chalkiadakis, G. & Boutilier, C. 2004. Bayesian reinforcement learning for coalition formation under uncertainty. In *Proceedings of the third International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'04)*, New York, NY, USA.

Chalkiadakis, G. & Boutilier, C. 2008. Sequential decision making in repeated coalition formation under uncertainty. In *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-08)*, Estoril, Portugal.

Chalkiadakis, G. & Boutilier, C. 2012. Sequentially optimal repeated coalition formation under uncertainty. *Autonomous Agents and Multi-Agent Systems* 24(3), 441–484.

Chalkiadakis, G., Elkind, E. & Jennings, N. R. 2009. Simple coalitional games with beliefs. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence*, Boutilier, C. (ed.). Morgan Kaufmann Publishers Inc., 85–90.

Chalkiadakis, G., Elkind, E., Markakis, E. & Jennings, N. R. 2008. Overlapping coalition formation. In *Proceeedings of the 4th International Workshop on Internet and Network Economics (WINE2008)*, Shanghai, China, 307–321.

Chalkiadakis, G., Elkind, E., Markakis, E. & Jennings, N. R. 2010. Cooperative games with overlapping coalitions. *Journal of Artificial Intelligence Research* **39**, 179–216.

Coleman, J. S. 1970. The benefits of coalition. *Public Choice* **8**, 45–61.

Conitzer, V. & Sandholm, T. 2003. Complexity of determining nonemptiness of the core. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, Acapulco, Mexico.

Conitzer, V. & Sandholm, T. 2004. Computing shapley values, manipulating value division schemes, and checking core membership in multi-issue domains. In *Proceedings of the 19th National Conference on Artificial Intelligence (AAAI-04)*, San Jose, CA, USA, 219–225.

Contreras, J., Klusch, M., Shehory, O. & Wu, F. 1997. Coalition formation in a power transmission planning environment. In *Proceedings of the Second International Conference on Practical Applications of Multi-Agent Systems, PAAM97*, London, UK.

Contreras, J., Klusch, M. & Yen, J. 1998. Multi-agent coalition formation in power transmission planning: a bilateral shapley value approach. In *Proceedings of the Fourth International Conference on Artificial Intelligence Planning Systems*, Pittsburgh, PA, USA.

Cornforth, D., Kirley, M. & Bossomaier, T. 2004. Agent heterogeneity and coalition formation: investigating market-based cooperative problem solving. In *Proceedings of the third International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'04)*, New York, NY, USA.

Dang, V. D., Dash, R. K., Rogers, A. & Jennings, N. R. 2006. Overlapping coalition formation for efficient data fusion in multi-sensor networks. In *Proceedings of the Twenty-First Conference on Artificial Intelligence (AAAI-06)*, Boston, USA, 635–640.

Dang, V. D. & Jennings, N. R. 2004. Generating coalition structures with finite bound from the optimal guarantees. In *Proceedings of the third International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'04)*, New York, NY, USA.

Davis, M. & Maschler, M. 1965. The kernel of a cooperative game. *Naval Research Logistics Quarterly* **12**, 223–259.

Deng, X., Fang, Q. & Sun, X. 2006. Finding nucleolus of flow game. In *SODA '06: Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithm*. ACM Press, 124–131.

Deng, X. & Papadimitriou, C. H. 1994. On the complexity of cooperative solution concepts. *Mathematical Operation Research* **19**(2), 257–266.

de Keijzer, B., Klos, T. & Zhang, Y. 2010. Enumeration and exact design of weighted voting games. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2010)*, Toronto, Canada, 391–398.

Dieckmann, T. & Schwalbe, U. 2002. Dynamic coalition formation and the core. *Journal of Economic Behavior & Organization* **49**(3), 363–380.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

Elkind, E., Goldberg, L. A., Goldberg, P. & Wooldridge, M. 2007. Computational complexity of weighted threshold games. In *Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence (AAAI-07)*, Vancouver, Canada, 718–723.

Elkind, E. & Pasechnik, D. 2009. Computing the nucleolus of weighted voting games. In *SODA'09: 20th ACM-SIAM Symposium on Discrete Algorithms*, New York, NY, USA, 327–335.

Elkind, E. & Wooldridge, M. 2009. Hedonic coalition nets. In *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-09)*, Budapest, Hungary, 417–424.

Gillies, D. B. 1953. *Some Theorems on n-Person Games*. PhD thesis, Department of Mathematics, Princeton University.

Greco, G., Malizia, E., Palopoli, L. & Scarcello, F. 2009. On the complexity of compact coalitional games. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence*, Boutilier, C. (ed.). Morgan Kaufmann Publishers Inc., 147–152.

Griffiths, N. & Luck, M. 2003. Coalition formation through motivation and trust. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS '03)*, Rosenschein, J. S., Wooldridge, M., Sandholm, T. & Yokoo, M. (eds). ACM Press.

Hart, S. & Kurz, M. 1983. Endogenous formation of coalitions. *Econometrica* **51**(4), 1047–1064.

Horling, B. & Lesser, V. 2004. A survey of multi-agent organizational paradigms. *The Knowledge Engineering Review* **19**, 281–316.

Horling, B., Vincent, R., Mailler, R., Shen, J., Becker, R., Rawlins, K. & Lesser, V. 2001. Distributed sensor network for real time tracking. In *Proceedings of the Fifth International Conference on Autonomous Agents (Agents'01)*, André, E., Sen, S., Frasson, C. & Müller, J. P. (eds). ACM Press, 417–424.

Ieong, S. & Shoham, Y. 2005. Marginal contribution nets: a compact representation scheme for coalitional games. In *EC '05: Proceedings of the 6th ACM Conference on Electronic Commerce*, Riedl, J., Kearns, M. & Reiter, M. (eds). ACM Press, 193–202.

Ieong, S. & Shoham, Y. 2008. Bayesian coalitional games. In *Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence (AAAI-08)*, Chicago, IL, USA, 95–100.

Kahan, J. P. & Rapoport, A. 1984. *Theories of Coalition Formation*. Lawrence Erlbaum Associates.

Kalai, E. & Xemel, E. 1982. Totally balanced games and games of flow. *Mathematics of Operations Research* **7**, 476–478.

Ketchpel, S. P. 1994a. The formation of coalitions among self-interested agents. In *Proceedings of the Eleventh National Conference on Artificial Intelligence*, Seattle, Washington, 414–419.

Ketchpel, S. P. 1994b. Forming coalitions in the face of uncertain rewards. In *Proceedings of the Eleventh National Conference on Artificial Intelligence*, Seattle, Washington, 414–419.

Klusch, M. & Shehory, O. 1996a. Coalition formation among rational information agents. In *Seventh European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, van Hoe, R. (ed.). Eindhoven, The Netherlands, 204–217.

Klusch, M. & Shehory, O. 1996b. A polynomial kernel-oriented coalition algorithm for rational information agents. In *Proceedings of the Second International Conference on Multi-Agent Systems*, Rosenschein, J. S., Wooldridge, M., Sandholm, T. & Yokoo, M. (eds). AAAI Press, 157–164.

Kraus, S., Shehory, O. & Taase, G. 2003. Coalition formation with uncertain heterogeneous information. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems*, Rosenschein, J. S., Wooldridge, M., Sandholm, T. & Yokoo, M. (eds). ACM Press, 1–8.

Kraus, S., Shehory, O. & Taase, G. 2004. The advantages of compromising in coalition formation with incomplete information. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'04)*, New York, NY, USA.

Kuipers, J., Faigle, U. & Kern, W. 2001. On the computation of the nucleolus of a cooperative game. *International Journal of Game Theory* **30**(1), 79–98.

Larson, K. S. & Sandholm, T. W. 2000. Anytime coalition structure generation: an average case study. *Journal of Experimental & Theoretical Artificial Intelligence* **12**(1), 23–42.

Lau, H. C. & Zhang, L. 2003. Task allocation via multi-agent coalition formation: taxonomy, algorithms and complexity. In *Proceedings of the 15th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2003)*, Sacramento, CA, USA, 346–350.

Lerman, K. & Shehory, O. 2000. Coalition formation for large-scale electronic markets. In *Proceedings of the Fourth International Conference on MultiAgent Systems (ICMAS-2000)*, Durfee, E. (ed.). IEEE Computer Society, 167–174.

Li, C., Rajan, U., Chawla, S. & Sycara, K. 2003. Mechanisms for coalition formation and cost sharing in an electronic marketplace. In *ICEC '03: Proceedings of the 5th International Conference on Electronic Commerce*, Sadeh, N., Dively, M. J., Kauffman, R., Labrou, Y., Shehory, O., Telang, R. & Cranor, L. (eds). ACM Press, 68–77.

Li, C. & Sycara, K. 2002. Algorithm for combinatorial coalition formation and payoff division in an electronic marketplace. In *Proceedings of the First International Joint Conference on Autonomous Agents*

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

and Multiagent Systems (AAMAS '02), Gini, M., Ishida, T., Castelfranchi, C. & Lewis Johnson, W. (eds). ACM Press, 120–127.

Manisterski, E., David, E., Kraus, S. & Jennings, N. R. 2006. Forming efficient agent groups for completing complex tasks. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems AAMAS-06*, Hakodate, Japan.

Mares, M. 2001. Fuzzy cooperative games: cooperation with vague expectations, volume 72 of *Studies in Fuzziness and Soft Computing*. Physica Verlag.

Mérida-Campos, C. & Willmott, S. 2004. Modelling coalition formation over time for iterative coalition games. In *Proceedings of the third International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'04)*, New York, NY, USA.

Mérida-Campos, C. & Willmott, S. 2006. The effect of heterogeneity on coalition formation in iterated request for proposal scenarios. In *Proceedings of the Fourth European Workshop on Multi-Agent Systems (EUMAS 06)*, Lisbon, Portugal.

Michalak, T., Dowell, A., McBurney, P. & Wooldridge, M. 2008. Optimal coalition structure generation in partition function games. In *Proceeding of the 2008 Conference on ECAI 2008*, Ghallab, M., Spyropoulos, C. D., Fakotakis, N. & Avouris, N. (eds). IOS Press, 388–392.

Michalak, T., Rahwan, T., Marciniak, D., Szamotulski, M. & Jennings, N. R. 2010. Computational aspects of extending the shapley value to coalitional games with externalities. In *Proceedings of the 2010 Conference on ECAI 2010: 19th European Conference on Artificial Intelligence*, Coelho, H., Studer, R. & Wooldridge, M. (eds). IOS Press, 197–202.

Michalak, T., Rahwan, T., Marciniak, D., Szamotulski, M., McBurney, P. & Jennings, N. R. 2010. A logic-based representation for coalitional games with externalities. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems – Volume 1*, *AAMAS '10*. International Foundation for Autonomous Agents and Multiagent Systems, 125–132.

Michalak, T., Rahwan, T., Sroka, J., Dowell, A., Wooldridge, M., McBurney, P. & Jennings, N. R. 2009. On representing coalitional games with externalities. In *Proceedings of the 10th ACM Conference on Electronic Commerce 09 (EC'09)*, Stanford, CA, USA.

Myerson, R. B. 1977. Graphs and cooperation in games. *Mathematics of Operations Research* **2**, 225–229.

Nishizaki, I. & Sakawa, M. 2001. Fuzzy and multiobjective games for conflict resolution, volume 64 of *Studies in Fuzziness and Soft Computing*. Physica Verlag.

Ohta, N., Conitzer, V., Satoh, Y., Iwasaki, A. & Yokoo, M. 2009. Anonimity-proof shapley value: extending shapley value for coalitional games in open environments. In *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-09)*, Budapest, Hungary.

Osborne, M. J. & Rubinstein, A. 1994. *A Course in Game Theory*. The MIT Press.

Owen, G. 1977. Values of games with a priori unions. In *Mathematical Economics and Game Theory: Essays in Honor of Oskar Morgenstern*, Hein, O. M. R. (ed.). Springer, 77–88.

Pěchouček, M, Mařík, V. & Bárta, J. 2002. A knowledge-based approach to coalition formation. *IEEE Intelligent Systems* **17**(3), 17–25.

Pěchouček, M., Mařík, V. & čStépánková, O. 2000. Coalition formation in manufacturing multi-agent systems. In *Proceedings of the 11th International Workshop on Database and Expert Systems Applications (DEXA'00)*, London/Greenwich, UK.

Peleg, B. & Sudhölter, P. 2007. *Introduction to the theory of cooperative cooperative games*, 2nd edition Springer.

Plaza, E. & Ontañón, S. 2006. Learning collaboration strategies for committees of learning agents. *Journal of Autonomous Agents and Multi-Agent Systems* **13**(3), 429–461.

Poon, A. S. Y., Yeung, C. S. K. & Wu, F. F. 1999. Game theoretical multi-agent modeling of coalition formation for multilateral trades. *IEEE Transactions on Power Systems* **4**(3), 929–934.

Procaccia, A. D. & Rosenschein, J. S. 2006. The communication complexity of coalition formation among autonomous agents. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems AAMAS-06*, Hakodate, Japan.

Rahwan, T. & Jennings, N. R. 2007. An algorithm for distributing coalitional value calculations among cooperating agents. *Artificial Intelligence* **171**(8–9), 535–567.

Rahwan, T. & Jennings, N. R. 2008a. Coalition structure generation: Dynamic programming meets anytime optimization. In *Proceedings of the 23rd Conference on Artificial Intelligence (AAAI-08)*, Chicago, IL, USA, 156–161.

Rahwan, T. & Jennings, N. R. 2008b. An improved dynamic programming algorithm for coalition structure generation. In *Proceedings of the 7th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-08)*, Estoril, Portugal.

Rahwan, T., Michalak, T., Jennings, N. R., Wooldridge, M. & McBurney, P. 2009. Coalition structure generation in multi-agent systems with positive and negative externalities. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI-09)*, Pasadena, CA, USA.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

Rahwan, T., Ramchurn, S. D., Dang, V. D., Giovannucci, A. & Jennings, N. R. 2007a. Anytime optimal coalition structure generation. In *Proceedings of the Twenty-Second Conference on Artificial Intelligence (AAAI-07)*, Vancouver, Canada, 1184–1190.

Rahwan, T., Ramchurn, S. D., Dang, V. D. & Jennings, N. R. 2007b. Near-optimal anytime coalition structure generation. In *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence (IJCAI'07)*, Hyderabad, India, 2365–2371.

Rahwan, T., Ramchurn, S. D., Jennings, N. R. & Giovannucci, A. 2009. An anytime algorithm for optimal coalition structure generation. *Journal of Artificial Intelligence Research* **34**, 521–567.

Ray, D. & Vohra, R. 1999. A theory of endogenous coalition structures. *Games and Economic Behavior* **26**, 286–336.

Sandholm, T. & Lesser, V. R. 1997. Coalitions among computationally bounded agents. *AI Journal* **94**(1–2), 99–137.

Sandholm, T. W., Larson, K. S., Andersson, M., Shehory, O. & Tohmé, F. 1999. Coalition structure generation with worst case guarantees. *Artificial Intelligence* **111**(1–2), 209–238.

Sarne, D. & Kraus, S. 2003. The search for coalition formation in costly environments. In *Cooperative Information Agents VII, 7th International Workshop, CIA 2003, Helsinki, Finland, August 27–29, 2003, Proceedings*, Lecture Notes in Computer Science **2782**, Springer.

Sarne, D. & Kraus, S. 2005. Cooperative exploration in the electronic marketplace. In *Proceedings of the Twentieth National Conference on Artificial Intelligence*, Veloso, M. & Kambhampati, S. (eds). AAAI Press/The MIT Press, 158–163.

Schmeidler, D. 1969. The nucleolus of a characteristic function game. *SIAM Journal of Applied Mathematics* **17**, 1163–1170.

Sen, S. & Dutta, P. S. 2000. Searching for optimal coalition structures. In *ICMAS '00: Proceedings of the Fourth International Conference on MultiAgent Systems (ICMAS-2000)*, Durfee, E. (ed.). IEEE Computer Society, 287.

Service, T. & Adams, J. 2010. Approximate coalition structure generation. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI-10)*, Atlanta, Georgia, USA, 854–859.

Service, T. C. & Adams, J. A. 2011. Constant factor approximation algorithms for coalition structure generation. *Autonomous Agents and Multi-Agent Systems* **23**(1), 1–17.

Shapley, L. 1953. A value for n-person games. In *Contributions to the Theory of Games*, Kuhn, H. & Tucker, A. (eds). Princeton University Press, vol. 2, 307–317.

Shehory, O. & Kraus, S. 1998. Methods for task allocation via agent coalition formation. *Artificial Intelligence* **101**(1–2), 165–200.

Shehory, O. & Kraus, S. 1999. Feasible formation of coalitions among autonomous agents in nonsuperadditve environments. *Computational Intelligence* **15**, 218–251.

Sims, M., Goldman, C. V. & Lesser, V. 2003. Self-organization through bottom-up coalition formation. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS03)*, Rosenschein, J. S., Wooldridge, M., Sandholm, T. & Yokoo, M. (eds). ACM Press, 867–874.

Soh, L.-K., Tsatsoulis, C. & Sevay, H. 2003. A satisficing, negotiated, and learning coalition formation architecture. In *Distributed Sensor Networks*, Victor, L., Charles O. Jr. & Milind, T. (eds). Springer, 109–138.

Soh, L.-K. & Tsatsoulis, C. 2002. Satisficing coalition formation among agents. In *AAMAS '02: Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems*. ACM Press, 1062–1063.

Stearns, R. E. 1968. Convergent transfer schemes for n-person games. *Transactions of the American Mathematical Society* **134**(3), 449–459.

Thrall, R. M. & Lucas, W. F. 1963. N-person games in partition function form. *Naval Research Logistics Quarterly* **10**(1), 281–298.

Tohmé, F. & Sandholm, T. 1999. Coalition formation processes with belief revision among bounded-rational self-interested agents. *Journal of Logic and Computation* **9**(6), 793–815.

Tsvetovat, M., Sycara, K. P., Chen, Y. & Ying, J. 2001. Customer coalitions in electronic markets. In *Agent-Mediated Electronic Commerce III, Current Issues in Agent-Based Electronic Commerce Systems (includes revised papers from AMEC 2000 Workshop)*, Dignum, F. & Cortès, U. (eds). Springer-Verlag, 121–138.

Ueda, S., Iwasaki, A., Yokoo, M., Silaghi, M. C., Hirayama, K. & Matsui, T. 2010. Coalition structure generation based on distributed constraint optimization. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI-10)*, Atlanta, Georgia, USA, 197–203.

Vassileva, J., Breban, S. & Horsch, M. 2002. Agent reasoning mechanism for making long-term coalitions based on decision making and trust. *Computational Intelligence* **18**(4), 583–595.

Wooldridge, M. 2009. *An Introduction to MultiAgent Systems*, 2nd edition, Wiley.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press

Wu, L.S.-Y. 1977. A dynamic theory for the class of games with nonempty cores. *SIAM Journal on Applied Mathematics* **32**(2), 328–338.

Yamamoto, J. & Sycara, K. 2001. A stable and efficient buyer coalition formation scheme for e-marketplaces. In *AGENTS '01: Proceedings of the Fifth International Conference on Autonomous Agents*, André, E., Sen, S., Frasson, C. & Müller, J. P. (eds). ACM Press, 576–583.

Yokoo, M., Conitzer, V., Sandholm, T., Ohta, N. & Iwasaki, A. 2005. Coalitional games in open anonymous environments. In *Proceedings of the Twentieth National Conference on Artificial Intelligence*, Veloso, M. M. & Kambhampati, S. (eds). AAAI Press AAAI Press/The MIT Press, 509–515.

Yokoo, M., Conitzer, V., Sandholm, T., Ohta, N. & Iwasaki, A. 2006. A compact representation scheme for coalitional games in open anonymous environments. In *Proceedings of the Twenty First National Conference on Artificial Intelligence*, Gil, Y. & Mooney, R. J. (eds). AAAI Press/The MIT Press, 697–702.

Young, H. P. 1985. Monotonic solutions of cooperative games. *International Journal of Game Theory* **14**, 65–72.

https://doi.org/10.1017/S0269888913000106 Published online by Cambridge University Press