Coalition Formation Tools for Achieving Collaboration inside Agent Societies

Gheorghe Cosmin SILAGHI
Faculty of Economics, Babes-Bolyai University, Cluj-Napoca, Romania

This paper reviews the main coalition formation tools for achieving collaboration inside agent societies. We adopted a computational approach for our study, as automated tools are required for a-priori simulation of organizational design. We presented coalition formation as a 3-step process. Algorithms for coalition structure generation are presented. Regarding the division of the payoffs inside coalitions, three kinds of game-theoretical solutions are investigated. We exemplify the usage of the theoretical tools by applying them on an example chosen from virtual organization modelling.
Keywords: *agent-based simulation, cooperative game theory, coalition formation, organizational design, algorithms.*

Acknowledgment: This paper was presented at the 2nd International Conference on Economics and Management of Networks, EMNET-2005, Corvinus University, September 2005. The author acknowledges the fruitful comments received from discussing participants during the conference. This research was carried out as part of the consortium project "Collaborative information systems in global economy" funded by Romanian National University Research Council under contract no. 1/8 from 2005.

Introduction

The proliferation of computer systems led to a new conception about how computer or non-human entities might work together. The concept of "open-systems" groups together a large number of systems of different design that can interact and cooperate in order to accomplish some specific tasks. Given the broad range of tasks that such systems should address, flexible patterns of communication and cooperation are required. Multi-agent systems represent a modern approach for modelling so-called open systems. Although there are situations where an agent can operate usefully by itself, the increasing interconnection and networking is making such situations rare. In the usual state of affairs, agents interact with other agents [3]. When designing multi-agent systems or societies of agents, the society engineer should design the rules (norms) that agents should accomplish in order to behave properly and to be accepted in the society.

These statements are valid also for designing interaction in human organizations. In the last years, human organizations are modelled and simulated more and more with agent systems and new interaction patterns are a-priori evaluated through computer-based simulation. For example, virtual organizations provide a means of bringing together autonomous stakeholders in a dynamic fashion in order to address a specific problem [2]. In this context, we consider that the topic of studying interaction patterns inside agent systems is of great interest.

Our main study concern is related with mechanisms used to enable collaboration inside open systems with heterogeneous agents. One means towards collaboration is to impose strong rules over the agents' behaviour. But, considering such an approach, agents will perform their tasks because they are required, and they could not be happy with their achievements. Therefore, agents will have incentives to leave the society or to oppose resistances to these enforcing rules. We think that such an approach is worth for consideration, but is more suitable for computational cooperation-based decision-making systems, with agents without beliefs and desires.

When dealing with self-interested goal-directed agents, like humans, other approaches need to be considered for achieving collaboration. We described [9] some settings like voting, auctions, bargaining and market mechanisms in order to argue that some beneficial characteristics of the society can be achieved even when imposing weakly enforcing rules over self-interested heterogeneous agents. In this paper, we will review and describe coalition-formation techniques used for studying collaboration inside multi-agent systems. Coalition formation studies how agents can engage voluntarily in coordination, what makes them to keep the coalition. The coalition formation research question can be stated as follows [5]: if each agent of a society has its own goals, how do agents locate other agents with whom they can beneficially collaborate and how they can fairly share the joint gain that accrues to the coalition? With respect to virtual organizations, coalition formation tools can be employed for modelling the operation phase of a virtual organization. Therefore, coalition formation stays at the foundation for establishing which partners in a society will work together for a task or how the operational structure will be re-modelled during the life of the organization.

We will investigate the activities that agents have to perform for establishing coalitions. We will analyze some proposed algorithms for determining possible solutions for the problem of establishing the operational structure of a society of agents. Regarding the payoff division inside a coalition, we will describe, exemplify and evaluate three solution concepts: the core concept [8], the Shapley value [1] and the kernel of the coalition [6]. These concepts are based on the cooperative game theory formalization, and allow one to simulate new designed mechanisms inside computational heterogeneous agent systems.

We will exemplify the usage of those concepts by applying them on a task distribution and realization problem. The main contribution we be to reveal how such formal concepts and computational algorithms can be employed for societal design, in order to a-priori simulate and predict some outcomes of the new rules.

The paper starts by introducing coalition formation as a topic of study inside agent theory. Section 3 will brief concepts from cooperative game theory that stay at the foundation of coalition formation. Section 4 describes in detail activities for coalition formation, with focus on computational tools toward coalition structure generation and payoffs division. Section 5 takes and example and shows the usage of the tools on a numerical setup. Section 6 concludes the paper, drawing out the main results.

2 Coalition formation within agent societies

This section will present the most important properties of agent systems and will argue for the necessity of studying coalition formation within agent societies.

An agent is an entity capable of independent (autonomous) action on its own behalf or the owner of the entity, figuring out what needs to be done to satisfy the design objectives of the society, rather than constantly being told [10]. In the context of computational multi-agent systems, the above-mentioned entities are computer systems, but in the most general case, the entities can be humans or organizations [7]. With respect to the *weak notion of agency* [10], agents have at least the following characteristics: autonomy, reactivity, proactivity and social ability

With respect to our goals of studying means of interaction inside organizations, the social property of agents is of our interest. In economic organizations coalitions of agents are formed on the basis of increased gains of the agents. How the agents negotiate in order to engage in collaboration and how after that, they will decide how to split the new created value is the topic of coalition formation. Coalition formation studies how agents can engage voluntarily in coordination, what makes them to keep the coalition. The question is what algorithms or design alternatives we have for setting up such a society of agents. Formally, we want to partition the set of agents into subsets, each of which is a coalition. The coalition gets a certain utility, which therefore, is divided among the members of the coalition. With this respect, coalition formation resides on the cooperative game theory concepts we will introduce fur-

ther on the paper.
Within our approach, we will consider computational approaches for coalition formation. Computational approaches reside on algorithms for establishing the coalition structure of the agents and the division of the value inside the coalition. With this respect, such algorithms must be [8] [2]

- *local executed*: each agent should be able to execute the coalition formation algorithm locally. Negotiation according to the algorithm should be totally decentralized
- *anytime executed*: after any regular termination of an arbitrary cooperative game in the considered environment, the coalition formation algorithm should output a stable configuration as a solution of the game.

Coalition formation mainly includes 3 activities [8]:

- *Coalition structure generation.* This activity includes all agent interaction performed inside the system such that in the end, coalitions will be formed. Inside the coalitions, agents will coordinate their activities, but agents will not coordinate between coalitions
- *Solving the optimization problem of each coalition.* This activity means pooling the tasks and resources of the agents in the coalition and solving the joint problem. The coalition objective could be to maximize the monetary value, or the overall expected utility. In our problem setup, the objective function of each coalition would be to maximize its utility.
- *Dividing the value* of the generated solution among agents. In the end, each agent will receive a value (money or utility) as a result of participating in the coalition. In some problems, the coalition value the agents have to share is negative, being a shared cost.

We will concentrate our discourse on coalition structure generation and on value division. We will trust the agents will accomplish their tasks well and, therefore, we will not insist on the second phase of the coalition formation process.

3 Cooperative game theory concepts

This section will provide the required game theoretical concepts for describing coalition formation activities. We used as references the reviews of E. Dani [3], M. Klusch [6], the lectures of M. Klusch for the course of Intelligent Information Agents, Free University of Amsterdam, 2001, the teaching materials of Adam Brandenburger for the course of Game Theory and Business Strategy[1], Stern School of Business, New York University.
Given a system of agents (players) $A = \{a_1, a_2, ..., a_n\}$, a coalition is a subset $C \subseteq A$.
A cooperative game $G = \langle A, v \rangle$ consists of two elements:

- the set of players, A
- the characteristic function specifying the value created by different subsets of the players in the game. The characteristic function is a function $v : P(A) \to R_+$ that assigns to each coalition C a value, which measures its utility achievable as a whole by cooperation among its members. $v(\emptyset) = 0$. $v(\{a_i\})$ denotes the self-value of a single agent coalition. $v(C)$ is the worth of (the value created by) the coalition C.

We say that the function v is *super-additive* if for all disjoint coalitions K and L (with $K \cap L = \emptyset$) inequality (3.1) holds:

$$v(K \cup L) \geq v(K) + v(L) \quad (3.1)$$

In this case, the game is said to be super-additive. If the above-mentioned inequality does not hold for all pairs of disjoint coalitions, the game is non-super-additive.
The game is *cohesive* if, for every partition $S_1, S_2, ..., S_k$ of A with $S_i \cap S_j = \emptyset, i \neq j$, inequality (3.2) holds:

$$v(N) \geq \sum_i v(S_i) \quad (3.2)$$

Each partition of A (a set of non-empty disjunctive subsets $C \subseteq A$) is called a *coalition structure* $\aleph$. For a coalition structure it is obvious that $|\aleph| = m \leq n$.
A game is *symmetric* if $\forall a_i, a_j \in A$ symmetric agents, for every $C_k \in \aleph$ not containing

[1] http://pages.stern.nyu.edu/~abranden/

agents a_i, a_j, equation (3.3) holds:

$$v(C \cup \{a_i\}) = v(C \cup \{a_j\}) \qquad (3.3)$$

The *marginal contribution* of player i to a coalition S, denoted by MC_i^S or MC_i if the coalition S is the grand coalition A, is the expression $v(S) - v(S \setminus \{i\})$, where $v(S \setminus \{i\})$ represents the coalition S without player i.

A *payment configuration* $PC = (u, \aleph)$ for the coalition structure $\aleph$ denotes a payoff distribution $u : P(A) \to R$ of coalition values in (A, v) to each agent, $(u_1, u_2, ..., u_n; C_1, C_2, ..., C_m)$.

An allocation $(u_1, u_2, ..., u_n)$ is said to be *individually rational* if inequality (3.4) holds for every player i. The related payment configuration is individual rational.

$$u_i \geq v(\{a_i\}) \qquad (3.4)$$

A payment configuration is *group rational* if $u(A) = v(A)$.

A payment configuration is *coalition rational* if $\forall T \subseteq A, u(T) \geq v(T)$.

A payment configuration is *locally coalition rational* if $\forall T \subseteq C_k \in \aleph, u(T) \geq v(T)$.

A payment configuration is Pareto optimal if does not exist u' such that $\forall i \in \{1,...,n\}$ $u'_i \geq u_i$.

An allocation $(u_1, u_2, ..., u_n)$ is *efficient* with respect to a coalition C if (3.5) holds

$$u(C) \overset{NOT}{=} \sum_{a_i \in C} u_i = v(C) \qquad (3.5)$$

An allocation vector $(u_1, u_2, ..., u_n)$ satisfies the *marginal contribution principle* if inequality (3.6) holds for all players i. It is obvious that an allocation that is individually rational and efficient satisfies the marginal contribution principle.

$$u_i \leq MC_i \qquad (3.6)$$

With regard to cooperative game theory, a payment configuration represents a solution of the game. The solution is stable if the payment configuration is at least individually rational. The core concept, the Shapley value and the kernel are solution concepts for cooperative games. We will analyze these different approaches as part of coalition formation activities.

4 Coalition formation activities

This section will describe some possible approaches for the first and the third phase of coalition formation. Therefore, we will review formally some results regarding coalition structure generation and some algorithms for payoff division inside coalitions. Section 5 of the paper will compare the presented approaches using a classical example from the literature, showing the importance of these tools with respect to the operation phase of organizations.

4.1 Coalition structure generation

Research has focused more on super-additive games [6] [11]. In such games, coalition structure generation is trivial because the agents are best off by forming the grand coalition where all agents operate together. It is argued that almost all games are super-additive because, at worst, the agents in a composite coalition use solutions that they had when they were in separate coalitions [8]. Against superadditivity it can be argued with the costs of coalition formation. In order to generate a coalition, there might be communication costs or antitrust penalties. Solving the optimization problem of the composite coalition may be more complex than solving the problems of the component coalitions. Therefore, under costly computation, component coalitions may be better off staying apart and not forming the composite coalition.

In games that are not super-additive the social welfare maximizing coalition structure varies, and coalition structure generation becomes non-trivial. The goal is to maximize the social welfare of the agents A by finding a coalition structure such that [8]:

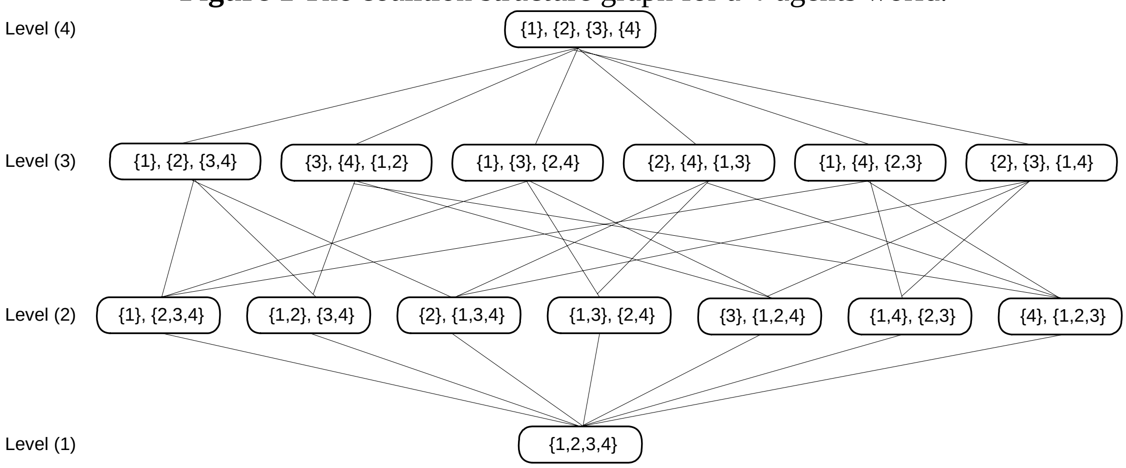
$$CS^* = \arg\max_{CS \in \aleph} \sum_{S \in CS} v_S \qquad (4.1)$$

We can make the following notation: $V(CS) = \sum_{S \in CS} v_S$, where $V(CS)$ represents the global social welfare of coalition structure CS.

The problem is that the number of coalition structures is very large ($O(|A|^{|A|/2})$) and if the total number of agents is not too small, we cannot enumerate all possible permutations of agents. To solve this problem, we should perform a *search* trough a subset of coalition structures and pick the best coalition structure seen so far. As a search problem, the coalition structure generation process can be viewed as a search on the coalition structure graph. The objective of the search, if the graph of evaluated structures has N nodes, would be to guarantee that the selected coalition structure is within a worst case bound from optimal, i.e. $k = \frac{V(CS^*)}{V(CS_n^*)}$ is finite and as small as possible. If we denote by $n_{\min}$ the smallest size of N that allows one to establish such a bound k, it was proved [8] that to bound k it suffices to search the lowest 2 levels of the coalition structure graph. With this search, the bound $k = |A|$ is tight, and the number of nodes searched is $n = 2^{|A|-1}$. No other search algorithm (than the one that searches the bottom two levels) can establish a bound k while searching only $n = 2^{|A|-1}$ nodes. Figure 1 depicts the coalition structure generation graph for a problem with 4 agents. The result of Sandholm [8] guarantees us obtaining a worst case bound from optimum, without searching the whole graph. On the other side, this result shows that exponentially ($n = 2^{|A|-1}$) many coalition structures have to been searched before a bound can be established.

To reduce the bound rapidly, if there is time for more search, the same author [8] proposes to continue with a breadth-first search of the graph starting from the top. Breadth-first search expands the successors of a selected node of the graph, evaluates each expanded node according to the criteria function of the search and stores the nodes on a list of candidates. From the list of candidates, the algorithm selects the most promising node, according with the criteria function value. If the problem domain happens to be super-additive, the algorithm finds the optimal coalition structure immediately. After searching level l, the bound k is given by the equation (4.2).

Figure 1 The coalition structure graph for a 4-agents world.



$$k = \begin{cases} \left\lceil \frac{|A|}{h} \right\rceil, & \text{if } |A| \equiv h-1 \pmod{h} \text{ and } |A| \equiv l \pmod{2} \\ \left\lfloor \frac{|A|}{h} \right\rfloor, & \text{otherwise} \end{cases} \qquad (4.2)$$

where $h = \left\lfloor \frac{|A|-l}{2} \right\rfloor + 2$.

V. D. Dang [2] produced an improved version of the above-presented structure generation algorithm. After searching levels 1, 2 and n of the graph, instead of breadth-first searching one-by-one the levels L_k of the graph, he proposes to search only subsets of those levels. It should search the set of all coalition structures that have at least $k = 3$ coalitions and at least one coalition whose cardinality is not less than $\lceil n(q-1)/q \rceil$, with q running from $\lfloor (n+1)/4 \rfloor$ to 2. The author [2] proved that after one round of searching, the solution is within a finite bound $b = 2q-1$ from optimal. More, the algorithm is an anytime algorithm and is 10^{379} times faster than Sandholm's algorithm [8] for a system containing 1000 agents, with respect to small bounds.

Other algorithms were proposed for generating the coalition structure. Therefore we can have

- *Merging algorithm:* i.e. breadth-first search algorithm from the top of the coalition structure graph. This algorithm cannot establish any bound before it has searched the entire graph. This is because to establish a bound the algorithm needs to see every coalition and the grand coalition only occurs at the bottom node. Visiting first the grand coalition would not help much since at least part of level 2 needs to be searched as well.
- *Splitting algorithm*: i.e. breadth-first search from the bottom of the graph. This algorithm is identical with the algorithm of [8] till the 3rd level is reached. After the 2nd level, splitting was found to reduce the bound much slower than the algorithms we presented above.

We may notice that these algorithms use as criteria value function the global social welfare and they are not concerned if agents indeed will wish to pursue in such a coalition structure. With respect to the behavior of self-interested agents, the payoffs division should validate if a coalition structure is feasible or not.

4.2 Payoffs division

Payoff division strives to divide the value of the chosen coalition structure among the agents in a fair and stable way, so that the agents are motivated to stay with the coalition structure, rather than move out of it. Several ways of dividing the payoff have been proposed. We will analyze the core concept, the kernel and the Shapley value. Payoffs division is of interest only in games with transferable utilities.

4.2.1 The core concept

The core of a characteristic function game with transferable payoffs is a set of payments configurations so that no subgroup is motivated to depart from the coalition structure CS. Formally, given a game $G = \langle A, v \rangle$ and a payment configuration $PC = (u, \aleph)$ then the core is given by equation (4.3).

$$Core(\langle A,v\rangle,\aleph) = \{(u,\aleph) \mid \forall C \subset A, \sum_{a \in C} u(a) \geq v(C), u(A) = v(A)\} \quad (4.3)$$

Only coalition structures that maximize the social welfare can be stable in the sense of core, because for any other coalition structure, the group of agents would prefer to switch to a maximizing social welfare solution.

In order to identify the core of a game we need to solve a set of inequalities, starting from the conditions of (4.3), from the individual rationality restrictions of (3.4) and the efficiency restriction (3.5). Or, we can replace individual rationality and efficiency by considering the marginal contribution principle restrictions of (3.6). E. Dani [3] proposes the usage of linear programming for solving the resulting inequalities system.

Figure 2 Transfer scheme algorithm for core payoff division

Repeat

Choose coalition T

For every agent $a_i \in T$, $u_i^{new} = u_i + \dfrac{v_T - \sum_{a_j \in T} u_j}{|T|}$ and $u_i^{new} = u_i$ for $\forall a_i \notin T$

Maintain feasibility: $\forall a_i \in A$, $u_i = u_i^{new} - \dfrac{\sum_{a_j \in A} u_j^{new} - \sum_{S \in CS} v_S}{|A|}$

The core is the strongest concept in the coalition formation. There are cases when the core is empty. In such games, there is no way to divide the social good such that the coalition structure to become stable. Another problem of the core is that the core may include multiple payoff vectors and the agents must agree which one to select. The often used solution is to pick the *nucleolus*, which is the

payoff vector that is in the center of the set of payoff vectors in the core.
If the number of agents is big, in order to define the core we have to solve and accomplish a big number of constraints from formula (4.3). To reduce the cognitive burden of the agents that try to reach a payoff division in the core, we can use the algorithm of figure 2 (Wu[2] in [8]). This algorithm converges to a solution in the core, starting from any initial payoff division.
The choice of T can be done at random or considering the coalition with the largest $v_T - \sum_{a_j \in T} u_j$ first. But there is no guarantee that a self-interested agent is motivated to follow the transfer scheme truthfully.

4.2.2 The Shapley value

In the core solution, an individual allocation fulfilled the individual rationality criteria and the marginal contribution principle. Therefore, inside a coalition, a player is rewarded with an amount u_i such that $v(\{a_i\}) \le u_i \le MC_i$. Therefore, the marginal contribution of a player restricts the amount she receives inside the coalition.
But, the real contribution of the player to a coalition depends on the position of the player inside the coalition. This position might be the moment (the round) when the player joined the coalition or might be other formalism. Therefore, instead of computing the payoff for a player based on her marginal contribution (which is independent of the coalition formation algorithm), we should let better to reward the player with her *expected contribution*. This idea resides at the foundation of the Shapley value. The Shapley value incorporates the property that gains from participation at the coalition are balanced between participating players.
Shapley value can be characterized axiomatically. An agent a_i is called dummy if $v_{C \cup \{a_i\}} - v_C = v_{\{a_i\}}$ for every coalition C that does not include a_i. Agents a_i and a_j are called interchangeable if $v_{(C \setminus \{a_i\}) \cup \{a_j\}} = v_C$ for every coalition C that includes a_i and does not include a_j. The axioms of the Shapley value are:

- Symmetry: if a_i and a_j are interchangeable, then $u_i = u_j$
- Dummies: If a_i is a dummy, then $u_i = v_{\{a_i\}}$
- Additivity: for any two games with characteristic functions v and w, u_i in $v+w$ equals u_i in v plus u_i in w, where $v+w$ is the game defined by the characteristic function $(v+w)_C = v_C + w_C$

The value of equation (4.4) defines the Shapley value. This value fulfills the 3 axioms stated above [8] [5]. The Shapley value can be interpreted as the marginal contribution of agent a_i to the coalition structure, averaged over all possible joining orders. The joining order matters, since the perceived contribution of agent a_i varies based on which agents have joined the coalition before it. The Shapley value represents the expected contribution of the agent to the game.

$$sv_i = \sum_{C \subseteq A} \frac{(|A| - |C|)!(|C|-1)!}{|A|!}(v_C - v_{C-\{a_i\}}) \quad (4.4)$$

The most remarkable properties of the Shapley value are that it always exists and it is unique, while the core solution does not guarantee either of these proprieties. The Shapley value is Pareto efficient as the entire value of the coalition structure gets distributed among the agents. The Shapley value guarantees that individual agents and the grand coalition are motivated to stay with the coalition structure. The Shapley value is the sole value function that fulfils the balanced contribution principle, accomplishing some ethical requirements for a possible solution.
A weak point is that it does not guarantee that all subgroups of agents are better off in the coalition structure than by breaking off into a coalition of their own. Another problem with the Shapley value is that marginal contributions of each agent have to be computed over all joining orders and there are

---
[2] L.S. Wu, (1977) A dynamic theory for games with non-empty cores, SIAM Journal of Applied Mathematics, vol. 32

$|A|!$ joining orders. One can guarantee each agent an expected payoff equal to the Shapley value by randomizing the joining orders. A trusted third party needs to carry out the randomization since each agent has a strong preference over different joining orders. Zlotkin [11] proposed a non-manipulable protocol for finding a randomized joining order. Within this protocol, every agent constructs a random permutation of the agents, encrypts it and sends it to all other agents. Once an agent has received an encrypted permutation from every other agent, it broadcasts its key. These keys are used to decrypt the permutations. The overall joining order is determined by sequentially permuting the results.

Contreras [1] proposed a method for finding the Shapley values based stable solutions to a CFG. Each agent a has to perform the following steps:

- compute the eigenvalue $v(\{a\})$ and $worth(a,a')$ for each $a' \in A$. Send / receive these values to agents a' in the world.
- compute the coalition values v_C for each coalition C of the actual coalition structure, in the hypothesis that agent a will join that coalition. The computed value will be

$$v_C = \sum_{a,a' \in C} worth(a,a') - (|C|-2)\sum_{a \in C} v_{(\{a\})} \quad (4.5)$$

- compute the own benefit of joining a grand coalition A as Shapley value and own demand for payment ($sv(a) - worth(a,A)$))
- form grand coalition with all other agents in a configuration $(\{A\},(sv(a))_{a\in A})$

$worth(a,a')$ represents the payment agent a receives if it engages in a coalition with agent a'. By $worth(a,C)$ was denoted the payments to agent a for items in coalition C. That is the sum of utilities of self and commissioned productions of a in coalition C. All agents repeat this coalition formation process from step 2 until no more coalitions are possible. If no coalition is possible at a step, the agents look at the second partner, after that at the third one etc. Contreras [1] recommends dividing the accumulated utility in a coalition according with the formula (4.6):

$$u(a) = v(a) + \frac{sv(a) - \sum_{a \neq a', a' \in C} v(a')}{|C|} \quad (4.6)$$

Therefore, it means that in the new coalition each agent will receive its value. The marginal utilities obtained by the coalition from the Shapley values of the agents that joined the coalition in the last round will we divided equally between the members of the coalition. They proved that this payoff dividing scheme is efficient and individual rational in the case of super-additive games. With this algorithm, [1] computed the computation and communication complexity as being $O(2^n n^2)$ and $O(n^2)$.

Regarding the above-mentioned algorithm, we may note that this is suitable for a general CFG environment. The grand coalition will not be necessarily formed. Even if common resources are shared, it does not mean that the agents will have access to them.

A similar approach based on the Shapley values was considered in [5]. Instead of considering coalitions of multiple agents, this work is based on a 2 agents-auction process. That is at the beginning, no coalitions exist. In each round, agents pair each to others, based on the Shapley values. After forming coalitions of size 2, inside the coalition agents decide to name a “chief” of the coalition. The “chief” of the coalition will further negotiate as the coalition is one agent. The following negotiation rounds are similar with the first round, as we can consider each coalition as a simple agent. Within this approach, a new problem arises: how an agent will be selected as a chief of the coalition. And, after some round, when we need to select the new chief, how the agents that are “workers” in the old coalition will participate in the chief selection process.

4.2.3 The Kernel approach

The set of payments configurations of a cooperative game where each pair of agent is in equilibrium is called a *kernel* (Davis and

Maschler[3] in [6]). Formally, a kernel is a set $K = \{(u, \aleph) \mid \forall a_i, a_j \in C \in \aleph, (a_i, a_j) \text{ in equilibrium}\}$, where the following concepts apply:

- *excess* of a coalition $C \notin \aleph$ with respect to an utility distribution $(u, \aleph)$: $e(C,u) = v_C - u_C$
- *surplus* or strength of agent a_k over agent a_l in the same coalition with respect to a payment configuration $(u, \aleph)$: $s_{kl} = \max_{R \notin \aleph, a_k \in R, a_l \notin R} e(R,u)$
- agent a_k ***dominates*** agent a_l in the same coalition with respect to a payment configuration $(u, \aleph)$ if $s_{kl} > s_{lk}$ and $u_l > v(\{a_l\})$
- a pair of agents a_k and a_l in the same coalition $C \in \aleph$ is in *equilibrium* if $s_{kl} = s_{lk}$ or $(s_{kl} > s_{lk}) \wedge (u_l = v(\{a_l\}))$ or $(s_{kl} < s_{lk}) \wedge (u_k = v(\{a_k\}))$

It was proved that the kernel solution has the following desirable properties:

- it exists and is unique for every 3-agents game.
- symmetric agents of some coalition in a given coalition structure get equal payoffs.
- each kernel stable configuration is locally Pareto optimal in the kernel.

In order to build the kernel, usually an algorithmic iterative approach is considered. Such algorithms are based on transferring payoffs (utilities) between agents as they are leaving / entering new coalitions. We will present a basic algorithm for determining the kernel as in the discourse of M. Klusch [6]. A *transfer schema* is a sequence of configurations $(u^1, \aleph), \ldots, (u^i, \aleph) \ldots$ transferring in each step i positive side-payments α among agents a_k and a_l by the assignments of formula (4.7). The computation of the side payments is done according with the principle of stability.

$$(u^{i+1}, \aleph) = \begin{cases} u_k^{i+1} & = u_k^i + \alpha \\ u_l^{i+1} & = u_l^i - \alpha \\ u_j^{i+1}, j \neq k, j \neq l & = u_j^i \end{cases} \quad (4.7)$$

A convergent transfer scheme iteratively computes for a given coalition structure $\aleph$ and an initial payment configuration $(u^0, \aleph)$ another payment configuration $(u\ \ , \aleph)$ which is the kernel of the coalition structure. This convergent transfer scheme is determined by the mutual demand $\alpha = d_{kl}$ for a pair of agents a_k and a_l as an upper bound for the side payments. Formula (4.8) provides means of computing the mutual demand between 2 agents.

$$d_{kl} = \begin{cases} \min((s_{kl} - s_{lk})/2, u_l), \text{ if } & s_{kl} > s_{lk} \\ 0, \quad \text{otherwise} & \end{cases} \quad (4.8)$$

Considering a sufficiently small relative error ε, an algorithm should terminate the iterations in order to compute the transfer schema if condition (4.9) fulfils.

$$re(u) = \frac{\max\{s_{ij} - s_{ji} \mid a_i, a_j \in C \in \aleph\}}{u(A)} \leq \varepsilon \quad (4.9)$$

The algorithm converges after at most $n \log_2 (re(u)/\varepsilon)$ iterations, with the complexity of $O(n2^n)$ for each iteration.

The difference between the kernel approach and the other game theoretical approaches (like the core approach) is that kernel approach provides with this transfer schema iterative algorithm for converging to the payoff division, while the core approach gives (predicts) how the division will look at the end, without giving an explanation of the process of reaching that division. The kernel concept is the first solution concept that possesses the desirable properties of symmetry and desirability (agents want to enter the solution) once at a time.

Unfortunately, the concept has also weak points. The original interpretation of the surplus presumes that agents a and a' compare $s(a,a')$ and $s(a,a')$ to see who could hope for more payoff in the same coalition based on the intensity of the feeling of individual utility. Therefore, it results in an interpersonal comparison of utilities, which, according with microeconomics is not acceptable. Utilities are a personal endowment of agents, and it is recommended to compare only utilities

[3] Davis M., Maschler M, (1965) The kernel of a cooperative game. Naval Research Logistic Quarterly, vol. 12

of the same agent. Granot and Maschler[4] (in [6]) refine this solution interpreting the kernel as a set of payments configurations for which every pair of rational agents are located symmetrically within their bargaining range. While two agents compare their surpluses, the remaining agents are assumed to be content with what they receive.

Klusch [6] proposes an algorithm for coalition formation based on the kernel concept. The algorithm advances the agents from one coalition structure to another in order to increase agents' payoff, thus, increasing rationally motivated cooperation. In each step at least one coalition will make an attempt to improve the payoffs of its members by computing a subsequent payment configuration as a proposal to another, actually most promising coalition. If both agents agree on such a proposal, it will be broadcasted to the other coalitions. Thereby, there are temporally enforced to accept the corresponding payoff distribution for all agents. Since this affects particular agents that were not involved in the bilateral agreement, they may be dissatisfied with their new payoff, and will react with their proposals in the next round. The negotiation continues till all proposals of all coalitions entities of the current valid coalition structure are rejected or a kernel stable payoff solution has been reached in the defined period of time. We can see that this algorithm mixes the two phases of determining of coalition structure and the payoff division.

The algorithm is classified as a negotiation-oriented and decentralized one, and the termination time is polynomial. It results in a kernel stable configuration, and provides the agents with a rational behaviour model for decision making.

5 Applying coalition formation tools

This section will take an example of a coalitional game and will try to identify the solutions of the game according with the tools enumerated in section 4. The numerical example was chosen from [2], who considered the same setup for analyzing creation, operation and maintenance of virtual organizations.

We can imagine a virtual organization where, in a city, an entrepreneur decides to offer full tourist packages for visitors. Therefore, it has to consider hotels where the tourists will be accommodated in the city and local tour organizers who will pick tourists from hotels, will perform a city tour and will bring back the tourists to their hotels. Local tour organizers have 5 car centers (from where they will start the tours) and there are 2 hotels from where they will have to take visitors. This setup models a game with 7 cooperative agents, with the objective of maximizing the entrepreneur profit (or minimizing its cost). Maximizing the profit can be done if local organizers decide in a rational way from which car centre a bus will pick tourists from hotels.

We will denote by C_1, C_2, C_3, C_4, C_5 the car centres and by H_1, H_2 the hotels. A coalition C might contain car centres and hotels. If a coalition C contains only car centres or hotels, its value $V(C)$ is 0, as no work is to be done. The value of a coalition containing both car centres and hotels is the profit the entrepreneur gets from letting the selected car centres to pick up visitors from the selected hotels.

We will consider the values of fig. 3 as being the definition of the characteristic value function.

Considering the structure generation algorithms of section 4.1, we can notice that for our problem, searching the two lower levels of the search graph bring out the most valuable coalition as being $\{\{C_1, C_2, C_3, C_5, H_1\}, \{C_4, H_2\}\}$ with the value 76.

Continuing the search with the approach of Sandholm [8], searching breadth-first the nodes of the graph, when approaching the second upper level of the graph, the most promising coalition is $\{\{C_1\}, \{C_2\}, \{C_3\}, \{C_4, H_2\}, \{C_5\}, \{H_1\}\}$ with the value of 45. The 3rd level of the graph brings to the coalition structure of $\{\{C_1\}, \{C_2, H_1\}, \{C_3\}, \{C_4, H_2\}, \{C_5\}\}$ with the

[4] Granot D., Maschler M., (1997) The Reactive Bargaining Set: Structure, Dynamics and Extensions to NTU Games. in International Journal of Game Theory, vol. 26/2

value of 77. Next level, C_5 can join C_2 and H_1 leading to an increased value of the coalition structure to 86. Then, the algorithm joins C_1 and C_3 in a no-significant coalition, entering a local optima peak of the search space. Returning on the 4th upper level, the algorithm investigates the coalition structure $\{\{C_1\},\{C_2,C_3,H_1\},\{C_4,H_2\},\{C_5\}\}$ and on the 5th level it goes to outcome the solution $\{\{C_1,C_2,C_3,H_1\},\{C_4,H_2\},\{C_5\}\}$ with the value of 87. After some search, the most promising solution is found on iteration 194 on the 3rd upper level of the graph as being $\{\{C_5\},(C_1,C_4,H_1\},\{C_2,C_3,H_2\}\}$ with value 89. We can notice how breadth-first search avoids entering local optima of search space without a chance to recover, as in hill-climbing search algorithms.

The approach of Dang [2], when starting the search from the top of the graph, immediately finds the coalition structure of $\{\{C1,C2,C3,H1\},\{C4,H2\},\{C5\}\}$ with value 87 on the second iteration.

We can notice that the idea of coalition formation algorithms is to quickly find a good approximation of the solution, rather than obtaining the best possible solution in a reasonable time. Analyzing outputs of coalition formation algorithms, we can make an idea about possible final coalition structures at the end of the game.

In order to determine the division of the payments inside the coalition, we need to analyze algorithms of section 4.2. First, we will determine the core, according with the definition of 4.2.1.

If we try to determine the core, related to the grand coalition (all job done by all agents together), we would notice that the core is empty. For example, the marginal utility of the first car centre and the first hotel toward the grand coalition is $MC_{C1,H1} = V_G - V(C_2,C_3,C_4,C_5,H_2) = 49 - 42 = 7$. Therefore, we should have $20 = V(C_1,H_1) \le u_{C1} + u_{H1} \le MC_{C1,H1} = 7$ which is a contradiction. Therefore, in this game, players have no incentive to self-interested cooperate toward the grand coalition. With respect to the social welfare maximizing coalition structure $\{\{C_5\},(C_1,C_4,H_1\},\{C_2,C_3,H_2\}\}$ with value 89, the only useful inference we can make is that the hotels H_1 and H_2 should get a payoff at least 20, otherwise there will be an incentive for them to make a coalition with the 5th car center. This inference is not part of the core restrictions under the final utilities. The division of the payoff inside the coalitions of the structure should therefore (according with the marginal contribution principle) satisfy the following (weak) restrictions: $u_1 \le 23$, $u_4 \le 24$, $u_{H1} \le 45$, $u_2 \le 12$, $u_3 \le 14$, $u_{H2} \le 44$, $u_1 + u_4 \le 44$, $u_1 + u_{H1} \le 44$, $u_4 + u_{H1} \le 44$, $31 \le u_2 + u_{H2} \le 45$, $33 \le u_3 + u_{H2} \le 45$.

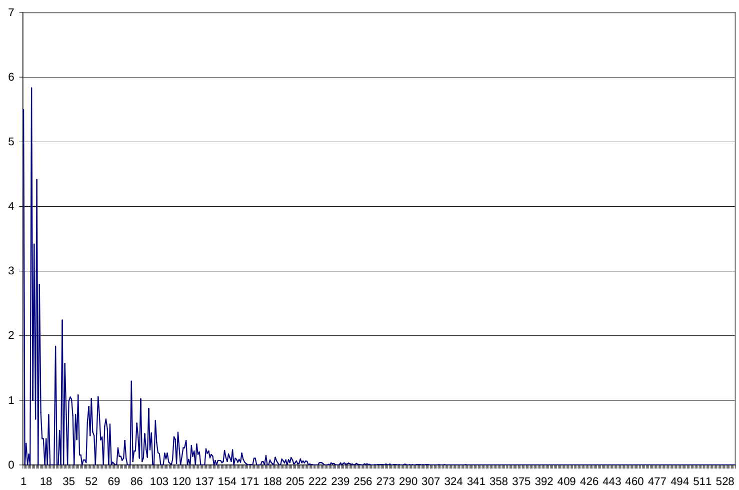
Therefore, the core cannot give us an obvious way of dividing the payoffs inside the welfare maximizing coalition structure. Trying to run the algorithm of fig. 2 in order to determine a solution inside the core of the game, we can notice that after few iterations (8) the algorithms converges to a solution, the one that divides equally the payoffs inside each coalition. Although the transfer scheme is a fair one, it can not assure the stability of the solution, as within this distribution of payoffs, hotels H_1 and H_2 gets less that 20 and, therefore, they have incentive to enter a coalition with C_5.

Regarding the Shapley values of the game, we computed according with formula (4.4) the following numbers: $sv_{C1} = 3.43$, $sv_{C2} = 4.3$, $sv_{C3} = 4.52$, $sv_{C4} = 4.02$, $sv_{C5} = 2.4$, $sv_{H1} = 13.63$, $sv_{H2} = 16.7$. The Shapley values represent a fair division between agents of the created value of making a coalition with regard to the opposite alternative (staying apart the coalition). We can notice that the second hotel has the biggest contribution, as the 5th car centre is the one with the smallest joint contribution to the game. We can easily notice the weak point of this distribution of the created value as (e.g.) car centre C_5 and H_1 have incentive to deviate from the distribution, because if they form alone a coalition they can obtain together a value of 20. The Shapley value solution divides only the new created value, not the whole value of the coalition structure. Shapley value division is not influenced by the coalition structure of the game.

Figure 3 Characteristic function for the organizational game

$V(C_1,H_1)=20 \quad V(C_2,H_1)=32 \quad V(C_3,H_1)=23 \quad V(C_4,H_1)=21 \quad V(C_5,H_1)=20$
$V(C_1,H_2)=22 \quad V(C_2,H_2)=31 \quad V(C_3,H_2)=33 \quad V(C_4,H_2)=45 \quad V(C_5,H_2)=20$
$V(C_1,H_1,H_2)=30 \quad V(C_2,H_1,H_2)=31 \quad V(C_3,H_1,H_2)=28 \quad V(C_4,H_1,H_2)=29$
$V(C_5,H_1,H_2)=26 \quad V(C_1,C_2,H_1)=31 \quad V(C_1,C_3,H_1)=32 \quad V(C_1,C_4,H_1)=44$
$V(C_1,C_5,H_1)=32 \quad V(C_2,C_3,H_1)=33 \quad V(C_2,C_4,H_1)=34 \quad V(C_2,C_5,H_1)=41$
$V(C_3,C_4,H_1)=35 \quad V(C_3,C_5,H_1)=34 \quad V(C_4,C_5,H_1)=31 \quad V(C_1,C_2,H_2)=32$
$V(C_1,C_3,H_2)=43 \quad V(C_1,C_4,H_2)=34 \quad V(C_1,C_5,H_2)=33 \quad V(C_2,C_3,H_2)=45$
$V(C_2,C_4,H_2)=33 \quad V(C_2,C_5,H_2)=31 \quad V(C_3,C_4,H_2)=34 \quad V(C_3,C_5,H_2)=42$
$V(C_4,C_5,H_2)=34 \quad V(C_1,C_2,H_1,H_2)=37 \quad V(C_1,C_3,H_1,H_2)=38$
$V(C_1,C_4,H_1,H_2)=37 \quad V(C_1,C_5,H_1,H_2)=36 \quad V(C_2,C_3,H_1,H_2)=38$
$V(C_2,C_4,H_1,H_2)=38 \quad V(C_2,C_5,H_1,H_2)=37 \quad V(C_3,C_4,H_1,H_2)=37$
$V(C_3,C_5,H_1,H_2)=36 \quad V(C_4,C_5,H_1,H_2)=38 \quad V(C_1,C_2,C_3,H_1)=42$
$V(C_1,C_2,C_4,H_1)=31 \quad V(C_1,C_2,C_5,H_1)=30 \quad V(C_1,C_3,C_4,H_1)=30$
$V(C_1,C_3,C_5,H_1)=31 \quad V(C_1,C_4,C_5,H_1)=29 \quad V(C_2,C_3,C_4,H_1)=31$
$V(C_2,C_3,C_5,H_1)=32 \quad V(C_2,C_4,C_5,H_1)=28 \quad V(C_3,C_4,C_5,H_1)=31$
$V(C_1,C_2,C_3,H_2)=31 \quad V(C_1,C_2,C_4,H_2)=29 \quad V(C_1,C_2,C_5,H_2)=31$
$V(C_1,C_3,C_4,H_2)=30 \quad V(C_1,C_3,C_5,H_2)=30 \quad V(C_1,C_4,C_5,H_2)=29$
$V(C_2,C_3,C_4,H_2)=31 \quad V(C_2,C_3,C_5,H_2)=30 \quad V(C_2,C_4,C_5,H_2)=29$
$V(C_3,C_4,C_5,H_2)=31 \quad V(C_1,C_2,C_3,H_1,H_2)=34 \quad V(C_1,C_2,C_4,H_1,H_2)=35$
$V(C_1,C_2,C_5,H_1,H_2)=31 \quad V(C_1,C_3,C_4,H_1,H_2)=33 \quad V(C_1,C_3,C_5,H_1,H_2)=31$
$V(C_1,C_4,C_5,H_1,H_2)=29 \quad V(C_2,C_3,C_4,H_1,H_2)=34 \quad V(C_2,C_3,C_5,H_1,H_2)=32$
$V(C_2,C_4,C_5,H_1,H_2)=28 \quad V(C_3,C_4,C_5,H_1,H_2)=30 \quad V(C_1,C_2,C_3,C_4,H_1)=32$
$V(C_1,C_2,C_3,C_5,H_1)=31 \quad V(C_1,C_2,C_4,C_5,H_1)=32 \quad V(C_1,C_3,C_4,C_5,H_1)=30$
$V(C_2,C_3,C_4,C_5,H_1)=31 \quad V(C_1,C_2,C_3,C_4,H_2)=43 \quad V(C_1,C_2,C_3,C_5,H_2)=42$
$V(C_1,C_2,C_4,C_5,H_2)=40 \quad V(C_1,C_3,C_4,C_5,H_2)=41 \quad V(C_2,C_3,C_4,C_5,H_2)=42$
$V(C_1,C_2,C_3,C_4,H_1,H_2)=46 \quad V(C_1,C_2,C_3,C_5,H_1,H_2)=44$
$V(C_1,C_2,C_4,C_5,H_1,H_2)=44 \quad V(C_1,C_3,C_4,C_5,H_1,H_2)=45$
$V(C_2,C_3,C_4,C_5,H_1,H_2)=45 \quad V(C_1,C_2,C_3,C_4,C_5,H_1)=30$
$V(C_1,C_2,C_3,C_4,C_5,H_2)=31 \quad V_G=V(C_1,C_2,C_3,C_4,C_5,H_1,H_2)=49$

Figure 4 Convergence of the kernel generation algorithm for payoff divisions. X axis represents the iteration number. Y axis represents the side payments of the transfer scheme



Regarding the kernel approach for dividing the payoffs, if we start the kernel generation algorithm of section 4.2.3 (formulas 4.7 and 4.8) for the best coalition structure $\{\{C_5\},(C_1,C_4,H_1\},\{C_2,C_3,H_2\}\}$ we obtained the following distribution: $u_{C1}=1$, $u_{C2}=3$, $u_{C3}=2$, $u_{C4}=5$, $u_{C5}=0$, $u_{H1}=38$, $u_{H2}=40$. The solution resides inside the core and the most of payoff seems to go to the hotels, as they are key players of the game. The kernel generation algorithm converged in a reasonable time. Fig. 4 depicts the variation of side payments during the iterations of the algorithm. We can notice that after less than 250 iterations we obtained a reasonable division, taking into account that at each iteration time, only one pair of agents exchange value (communicates). The division according with the Kernel approach resides in the core and is stable (no agents have incentive to deviate).

6 Conclusion

This paper investigated methods toward coalition formation inside agent societies. The approach is computational-based, the concern being on the coalition formation process and properties rather than on the subjective reasoning of agents. It was assumed that the designer intends to have tools for simulating what could happen while the organization operates, if every agent fulfils at least some "rational" behaviour.

As coalition formation is a 3-phases process, we insisted only on the determination of the coalition structure and the division of the payoffs.

Determination of coalition structure is represented by a negotiation process in which agents discuss each other and agree or not to enter a coalition. In essence, the coalition structure determination represents a search on the all permutations set of partitions of a set. Therefore, the number of possible coalition structures is very big, being not tractable to run a full search. Searching methods were proposed in the literature. These methods are based on building the search tree of the problem and performing an evaluation of the nodes of the tree. The criteria function employed for the search is the global social welfare. Sandholm [8] and Dang [2] provides with worthwhile algorithms for fast determination of an acceptable coalition structure.

The payoffs division is studied according with game theoretical concepts. Desirable properties are defined and one could impose such properties of the solution of the game.

The core approach represents the strongest solution concept, from the theoretical point of view. It is based on the principle that from the solution of the game, no group of agent has incentive to deviate. The core approach supplies with the final expected payoff division. The core concept is often too strong, as there are cases when, theoretically, no such a solution exists. The core concept can be well applied together with the presented searching methods for the coalition structure. The concept that guides the path through the solution is the maximization of the social welfare.

Other approaches are based on how pairs of agents negotiate or how an agent negotiates with a coalition. The Shapley value is a measure used to quantify the expected contribution of an agent to a coalition. Agents will join coalitions if they can add value to the existing coalition. With respect to the Shapley value, determination of the coalition structure is not any more a search, being more a matter of negotiation. The Shapley value payoff division has the advantage that it exists, is unique, Pareto efficient and ethical, fulfilling the balanced contribution principle. The Shapley value is suitable for agent in bilateral negotiations, when arguing for a division inside a coalition.

Kernel approaches are based on the idea that every two agents in a coalition are in equilibrium, which means that no agent wants to give off the coalition, agreeing with the outcome it obtained. Payment transfer schemas are considered, and convergence algorithms are proposed. Again, the coalition structure determination is a matter of negotiation. Kernel approaches are worth to be applied after the coalition structure is determined, in order to generate payoffs such as agents would have no incentive to deviate. Therefore, if a designer can impose some rules such as the organization to be structured toward the welfare maximizing solution, the kernel approach gives the designer the tool to suitable divide the values inside coalitions.

Each alternative is based on game theoretical principles and supplies with iterative methods for providing with the coalitions. Agents should gain more from the coalition than if they would stay alone. Alone, coalition structure generation algorithms do not verify if this individual rationality criterion is fulfilled. Each method has some drawbacks, as each one proved some to fulfil some desirable properties. It depends on how the given real problem can be formulated which coalition formation method will apply better. In every case, the analyst needs to consider the computational and communicational costs of the solution, as there are expensive searches and costly negotiation. The fact that these algorithms are, by principle, decentralized represents a worth of them.

References

1. Contreras, J, Klusch, M, Shehory, O, Wu, F (1997) Coalition Formation in a Power Transmission Planning Environment. in Proceedings of the 2nd International Conference on Practical Applications of Multi-agent Systems, London, UK

2. Dang, V. D. (2004) *Coalition Formation and Operation in Virtual Organizations*. PhD thesis, University of Southampton, UK

3. Dani E., (1980) *Game Theory*, Babes-Bolyai University, Faculty of Economics, Cluj-Napoca, Romania

4. Huhns, M.H., Stephens, L., (2001) Multi-agent Systems and Societies of Agents. in *Multiagent Systems, a Modern Approach to Distributed Artificial Intelligence*, The MIT Press, London, UK

5. Ketchpel, S., (1993) Coalition Formation Among Autonomous Agents. in *From Reaction to Cognition*, the 5th European Workshop on Modeling Autonomous Agents MAAMAW 1993, Lecture Notes in Computer Science, Vol. 957, Springer-Verlag, Berlin

6. Klusch, M., Shehory, O., (1996) A Polynomial Kernel Oriented Coalition Algorithm for Rational Information Agents. in *Proceedings of the 2nd International Conference on Multi-Agent Systems*, AAAI Press, Kyoto Japan

7. Russell S., Norvig P., (1995) *Artificial Intelligence: A Modern Approach*, 1st ed., Prentice Hall

8. Sandholm, T., (2000) Distributed Rational Decision Making. in *Multiagent Systems, a Modern Approach to Distributed Artificial Intelligence*, The MIT Press, London, UK

9. Silaghi, G.C., (2005) Mechanisms for Collaboration inside Heterogeneous Multi-agent Systems. in *Studia Oeconomica* L(1), Babes-Bolyai University Cluj-Napoca, Romania

10. Wooldridge M., (2002) *An Introduction to Multi-Agent Systems*, John Wiley and Sons, Chichester, UK

11. Zlotkin G., Rosenschein J.S., (1994) Coalition, cryptography and stability: mechanisms for coalition formation in task oriented domains. in *Proceedings of the National Conference on Artificial Intelligence*, Seattle, US