An Abstract Argumentation Framework for Supporting Agreements in Agent Societies

Stella Heras, Vicente Botti, and Vicente Julián

Departamento de Sistemas Informáticos y Computación
Universidad Politécnica de Valencia
Camino de Vera s/n. 46022 Valencia (Spain)
Tel.: (+34) 96 387 73 50 - Fax. (+34) 96 387 73 59

Abstract. In this paper we present an abstract argumentation framework for the support of agreement processes in agent societies. This framework takes into account arguments, attacks between them and the social contex of the agents that put forward arguments. Then, we define the semantics of the framework, providing a mechanism to evaluate arguments in view of other arguments posed in the argumentation process. Finally, the framework is illustrated with an example in a real domain of a water-rights transfer market.

1 Introduction

Research on argumentation is at its peak in the Multi-Agent Systems (MAS) community, since it has proven very successful to implement agents' internal and practical reasoning and to manage multi-agent dialogues [1]. MAS requires agents to have a way of reaching agreements that harmonise the conflicts that come out when they have to collaborate or coordinate their activities. Also, gents in MAS can form societies that link them via dependency relations that emerge from agents' interactions or are predefined by the system. These dependencies define the agents' social context, which has an important influence in the way agents can argue and reach agreements with other agents. Nowadays, argumentation is an active research area in AI and MAS [2].

However, little work has been done to study the effect of the social context of agents in the way that they argue and analyse arguments. Commonly, the term *agent society* is used in the argumentation and AI literature as a synonym for an *agent organisation* [3] or a *group of agents* that play specific roles, follow some interaction patterns and collaborate to reach global objectives [4]. In addition to the dependency relations between agents, we also consider values as an important element of their social context. These values can be individual values that agents want to promote or demote (e.g. solidarity, peace, etc.) or also *social values* inherited from the agents' dependency relations. Thus, we endorse the view of abstract value-based argumentation frameworks [5], which stress the importance of the audience in determining whether an argument is persuasive or not. To our knowledge, no research is done to adapt abstract argumentation

E.S. Corchado Rodriguez et al. (Eds.): HAIS 2010, Part II, LNAI 6077, pp. 177–184, 2010.

frameworks to represent and manage arguments in agent societies taking into account their social context. Nevertheless, this social information plays an important role in the way agents can argue and in the acceptability semantics of arguments. Depending on their social relations with other agents, an agent can accept arguments from a member of its society that it would never accept before acquiring social dependencies with this member.

In this paper we present an abstract argumentation framework for the support of agreement processes in agent societies. This framework takes into account arguments, attacks between them and the social contex of the agents that put forward them. Then, we define the semantics of the framework, providing a mechanism to evaluate arguments in view of other arguments posed in the argumentation process. Finally, an example is provided.

2 Agent Society

In this work, we follow the approach of [6] and [7], who define an *agent society* in terms of a set of *agents* that play a set of *roles*, observe a set of *norms* and a set of *dependency relations* between roles and use a *communication language* to collaborate and reach the global objectives of the *group*. This definition can be adapted to any open MAS where there are norms that regulate the behaviour of agents, roles that agents play, a common language that allow agents to interact defining a set of permitted locutions and a formal semantics for each of these elements. Moreover, the set of norms in open MAS define a *normative context* (covering both the set of norms defined by the system itself as well as the norms derived from agents' interactions)[8].

However, we consider that the values that individual agents or groups want to promote or demote and preference orders over them have also a crucial importance in the definition of an argumentation model for agent societies. These values could explain the reasons that an agent has to give preference to certain beliefs, objectives, actions, etc. Also, dependency relations between roles could imply that an agent must change or violate its value preference order. For instance, an agent of higher hierarchy could impose their values to a subordinate or an agent could have to adopt a certain preference order over values to be accepted in a group. Therefore, we endorse the view of [9], [10] and [5], who stress the importance of the audience in determining whether an argument (e.g. for accepting or rejecting someone else's beliefs, objectives or action proposals) is persuasive or not. Thus, we have included in the above definition of agent society the notion of values and preference orders among them. Next, we provide a formal definition for the model of society that we have adopted:

Definition 1 (Agent Society). *An Agent society in a certain time t is defined as a tuple* $S_t = <$ Ag, Rl, D, G, N, V, Role, $Dependency_{S_t}$, Group, Values, $Valpref_q >$ *where* $Ag = \{ag_1, ag_2, ..., ag_I\}$ *is the set of I agents of the society at the time t,* $Rl = \{rl_1, rl_2, ..., rl_J\}$ *is the set of J roles of the society,* $D = \{d_1, d_2, ..., d_K\}$ *is the set of K possible dependency relations over roles,* $G = \{g_1, g_2, ..., g_L\}$ *is the set of groups that the agents of the society form in t, where* *each* $g_l = \{a_1, a_2, ..., a_M\}, M \leq I$ *consist of a set of agents* $a_i \in A$ *of the society,* N *is the normative context of* S_t *(the defined set of norms that affect the roles that the agents play in the society* S_t*),* $V = \{v_1, v_2, ..., v_P\}$ *is the set of* P *values predefined in the society,* $Role : A \rightarrow 2^R$ *is a function that assigns an agent the role(s) that it plays in the society in this moment,* $Dependency_{S_t} :<_D^{S_t} \subseteq RxR$ *defines a partial order relation over roles* ($\forall r_1, r_2, r_3 \in R, r_1 <_d^{S_t} r_2 <_d^{S_t} r_3$ *implies that* r_3 *has the highest rank with respect to the dependency relation d in the society* S_t*. Also,* $r_1 <_d^{S_t} r_2$ *and* $r_2 <_d^{S_t} r_1$ *implies that* r_1 *and* r_2 *have the same rank with respect to d in the society* S_t*),* $Group : A \rightarrow 2^G$ *is a function that assigns an agent to the groups that it belongs to,* $Values : A \rightarrow 2^V$ *is a function that assigns an agent the set of values that it has and* $Valpref_q \subseteq VxV$*, where* $q = Ag \vee q = G$*, defines a irreflexive, antisymmetric and transitive relation* $<_q^{S_t}$ *over the values.*

3 Framework Formalisation

Most abstract argumentation frameworks (AF) are based on Dung's framework [11], which is defined as a pair $< A, R >$ where A is a set of arguments and $R \subseteq AxA$ is a binary *attack* relation on A. For two arguments A and B, $R(A, B)$ means that the argument A attacks the argument B. AF abstract the structure and meaning of arguments and attacks between them and focus their research efforts on analysing generic properties and *argumentation semantics*. This semantics is the formal definition of the method by which arguments are evaluated in view of other arguments [12]. Semantics can be *extension-based*, which determine the *extensions* or sets of arguments that can be collectively acceptable or *labelling-based*, which label each argument of A with a specific state in a predetermined set of possible states of an argument. Based on Dung's AF, we define an Argumentation Framework for an Agent Society (AFAS) as:

Definition 2 (Argumentation Framework for an Agent Society). *An argumentation framework for an agent society is a tuple:* $AFAS = <A, R, S_t>$ *where A is a set of arguments, R is a abstract attack relation as defined in Dung's framework and* S_t *is a society of agents as defined in Definition 1.*

Then, we specialise AFAS considering them for an specific agent, since each agent of an open MAS can have a different preference order over values. Thus, an audience is defined as a preference order over values. For the definition of our Agent specific Argumentation Framework for Agent Societies we start from the definition of Audience specific Value-based Argumentation Frameworks (AVAF) [5]. This is also based on Dung's and we will extend and adapt it to take into account the social context of agents. An audience-specific value-based argumentation framework is a 5-tuple: $AVAF_a =< A, R, V, val, Valpref_a >$ where A, R , V and val are as defined for a Value-based Argumentation Framework (VAF) [5], $a \in P$ is an audience of the set of audiences P and $Valpref_a \subseteq V \times V$ is a transitive, irreflexive and asymmetric preference relation that reflects the value preferences of the audience a. We extend AVAFs and define our abstract Agent Specific Argumentation Framework in an Agent Society (AAFAS) as follows:

Definition 3 (Agent specific Argumentation Framework in an Agent Society). *An agent specific argumentation framework in an agent society is a tuple:* $AAFAS =< Ag, Rl, D, G, N, A, R, V, Role, Dependency_S, Group, Values, val, Valpref_{ag_i} >$ *where* Ag, Rl, D, G, N, A, R, V, $Role$, $Dependency_{S_t}$, $Group$ *and* $Values$ *are defined as in Definition 1;* $val(ag, a) : Ag \times A \rightarrow 2^V$ *is a function that assigns an agent's argument the value(s) that it promotes and* $Valpref_{ag_i} \subseteq V \times V$*, defines a irreflexive, antisymmetric and transitive relation* $<^{S_t}_{ag_i}$ *over the agent's* ag_i *values in the society* S_t*.*

The aim of AFAS is to determine which agent's argument attacks other agent's argument in an argumentation process performed in a society of agents and in each case, which argument would defeat the other. To do that, we have to consider the values that arguments promote and their preference relation as in AVAFs, but also the dependency relations between agents. These relations could be stronger than value preferences in some cases (depending on the application domain). For the time being, as in [6], we only consider the following dependency relations: a) *Power*: when an agent has to accept a request from other agent because of some pre-defined domination relationship between them (e.g. in a society S that manages the water-rights transfer of a river basin (as explained in the example of section 4), $Farmer <^{S_t}_{Pow} BasinAdministrator$, since farmers must comply with the laws announced by the basin administrator);b) *Authorisation*: when an agent has committed itself to other agent for a certain service and a request from the latter leads to an obligation when the conditions are met (e.g. in the society S_t, $Farmer_i <^{S_t}_{Auth} Farmer_j$, if $Farmer_j$ has contracted a service that offers $Farmer_i$) and *Charity*: when an agent is willing to answer a request from other agent without being obliged to do so (e.g. in the society S_t, by default $Farmer_i <^{S_t}_{Ch} Farmer_j$ and $Farmer_j <^{S_t}_{Ch} Farmer_i$). Thus, we can now define the agent-specific defeat relation of our AAFAS as:

Definition 4 (Defeat). *An agent's* ag_1 *argument* $a_1 \in AAFAS$ *put forward in the context of a society* S_t $defeats_{ag_1}$ *other agent's* $ag_2 \in AAFAS$ *argument* a_2 *iff* $attack(a_1, a_2) \wedge (val(ag_1, a_1) <^{S_t}_{ag_1} val(ag_2, a_2) \notin Valpref_{ag_1}) \wedge (Role(ag_1) <^{S_t}_{Pow} Role(ag_2) \vee Role(ag_1) <^{S_t}_{Auth} Role(ag_2) \notin Dependency_{S_t})$

Therefore, we express that the argument a_1 $defeats_{ag_1}$ from the ag_1 point of view the argument a_2 as $defeats_{ag_1}(a_1, a_2)$ if a_1 attacks a_2, ag_1 prefers the value promoted by a_1 to the value promoted by a_2 and ag_2 does not have a power or authority relation with ag_1. Thus, based on Dung's acceptability semantics, we can define some acceptability concepts. Note that in them we compare arguments of different agents. However, since dependency relations are a partial order relations (reflexive, antisymmetric and transitive), an agent has equal power, authorisation and dependency relations over itself ($ag \leq ag$ (reflexivity) $\rightarrow ag = ag$ (antisymmetry)) and, in that case, the AAFAS would be equivalent to an AVAF and the acceptability criteria of this AVAF would apply. Let $a_i \in A$ be the argument of an agent $ag_i \in Ag$ and $ag \in Ag$ a generic agent.

Definition 5 (Conflict-free). *A set of arguments* $ARG \in A$ *is* $conflict-free_{ag_1}$ *for an agent* ag_1 *in the society* S_t *if* $(\nexists a_1, a_2 \in ARG\ /\ (attacks(a_1, a_2) \vee attacks(a_2, a_1))\ \vee\ ((val(ag_1, a_1) <^{S_t}_{ag_1} val(ag_2, a_2) \notin\ Valpref_{ag_1})\ \wedge\ (Role(ag_1) <^{S_t}_{Pow} Role(ag_2) \vee Role(ag_1) <^{S_t}_{Auth} Role(ag_2) \notin Dependency_{S_t}))$.

That is, if there is no pair of arguments that attack each other and or, otherwise, there is a value preference relation and a dependency relation that invalidates the attack.

Definition 6 (Acceptability). *An argument* $a_1 \in A$ *is* $acceptable_{ag}$ *in a society* S_t *wrt a set of arguments* $ARG \in A$ *iff* $\forall a_2 \in A \wedge defeats_{ag}(a_2, a_1) \rightarrow \exists a_3 \in ARG \wedge defeats_{ag}(a_3, a_2)$.

That is, if the argument is $defeated_{ag}$ by other argument of A, some argument of the subset ARG $defeats_{ag}$ this other argument.

Definition 7 (Admissibility). *A conflict-free set of arguments* $ARG \in A$ *is admissible for an agent* ag *iff* $\forall a \in ARG \rightarrow acceptable_{ag}$.

Definition 8 (Preferred Extension). *A set of arguments* $ARG \in A$ *is a* $preferred-extension_{ag}$ *for an agent* ag *if it is a maximal (wrt set inclusion)* $admissible_{ag}$ *subset of* A.

Then, for any $AAFAS = < Ag, Rl, D, G, N, A, R, V, Role, Dependency_{S_t}, Group, Values, val, Valpref_{ag_i} >$ there is a corresponding $AFAS = <A, R, S_t>$, where $R = defeats_{ag_i}$. Thus, each attack relation of $AFAS$ has a corresponding agent specific $defeat_{ag_i}$ relation in $AAFAS$. These properties are illustrated in the example of the next section.

4 Application of the Framework to the Management of Water-Right Transfer Agreements

To exemplify our framework, let us propose a scenario of an open MAS that represents a water market [13], where agents are users of a river basin, they belong to a society S_t and they can enter or leave the system to buy and sell water-rights. A water-right is a contract with the basin administrator that specifies the volume that can be spent, the water price, the district where the water is settled, etc. Here, suppose that two agents that play the role of farmers F_1 and F_2 in the river basin RB (group) are arguing to decide over a water-right transfer agreement and a basin administrator BA must control the process and make a final decision. The basin has a set of norms N_{RB} and commands a dependency relations of charity (Ch) between two farmers and power (Pow) between a basin administrator and a farmer. In addition, farmers prefer to reach an agreement before taking legal action to avoid the intervention of a jury (J). Also, F_1 prefers economy over solidarity ($SO <^{S_t}_{F_1} J <^{S_t}_{F_1} EC$), F_2 prefers solidarity over economy ($J <^{S_t}_{F_2} EC <^{S_t}_{F_2} SO$) and by default, BA has the value preference order of the basin, which is ($EC <^{S_t}_{BA} SO <^{S_t}_{BA} J$).

In this scenario, F_1 puts forward the argument " I should be the beneficiary of the transfer because my land is adjacent to the owner's land". Here, we suppose that the closer the lands the cheaper the transfers between them and then, this argument could promote economy. However, F_2 replies with the argument "I should be the beneficiary of the transfer because there is a drought and my land is almost dry". In this argument, we assume that crops are lost in dry lands and helping people to avoid losing crops promotes solidarity. In addition, the BA knows that the jury will interfer if the agreement violates the value preferences of the river basin. Then, they can also put forward the following arguments "F_2 should allow me (F_1) to be the beneficiary of the water-right transfer to avoid the intervention of a jury (J)", "F_1 should allow me (F_2) to be the beneficiary of the water-right transfer to avoid the intervention of a jury (J)" and "F_1 should allow F_2 to be the beneficiary of the water-right transfer to avoid the intervention of a jury (J)".

In view of this context, BA could generate a $AFAS = <A,\ R,\ S_t>$ as an extension of abstract argumentation frameworks $AF =< A, R >$. Thus, we have the following arguments in A={A1, A2, A3, A4, A5, A6, A7} (which are all possible solutions for the water-right transfer agreement process): A1 (posed by F_1): F_1 should be the beneficiary of the water transfer ($F_1 w$) to promote economy (EC), A2 (posed by F_2): F_1 should not be the beneficiary of the water transfer ($F_1 nw$) to promote solidarity (SO), A3 (posed by F_2): F_2 should be the beneficiary of the water transfer ($F_1 w$) to promote solidarity (SO), A4 (posed by F_1): F_2 should not be the beneficiary of the water transfer ($F_2 nw$) to promote saving (EC), A5 (posed by F_1): F_2 should allow F_1 to be the beneficiary of the water transfer ($F_1 w \& F_2 nw$) to avoid the intervention of a Jury (J) and A6 (posed by F_2 and BA): F_1 should allow F_2 to be the beneficiary of the water transfer ($F_1 nw \& F_2 w$) to avoid the intervention of a Jury (J).

The BA cannot decide the water tranfer in favour of both water users, so $attacks(A1, A3)$ and vice versa and we assume that it must take a decision favouring at least one part, so $attacks(A2, A4)$ and vice versa. In addition, $attacks(A5, A2)$, $attacks(A5, A3)$ and $attacks(A5, A6)$ and all these arguments attack A5 and $attacks(A6, A1)$, $attacks(A6, A4)$ and $attacks(A6, A5)$ and all these arguments attack A6. Then, R ={*attacks(A1, A3), attacks(A3, A1), attacks(A2, A4), attacks(A4, A2), attacks(A5, A2), attacks(A5, A3), attacks(A5, A6), attacks(A2, A5), attacks(A3, A5), attacks(A6, A5), attacks(A6, A1), attacks(A6, A4), attacks(A1, A6), attacks(A4, A6)*} and $S_t = <Ag,\ Rl,\ D,\ G,\ N,\ V,\ Role,\ Dependency_{S_t},\ Group,\ Values,\ Valpref_q>$ where $Ag = \{F_1,\ F_2,\ BA\}$; $Rl = \{Farmer,\ BasinAdministrator\}$, $D = \{Power,\ Charity\}$; $G = \{RB\}$; $N = N_{RB}$; $V = \{EC,\ SO,\ J\}$, $Role(F_1) = Role(F_2) = Farmer$; $Role(BA) = BasinAdministrator$, $Farmer <^{S_t}_{Pow} BasinAdministrator$, $Farmer <^{S_t}_{Ch} Farmer$, $Group(F_1) = Group(F_2) = Group(BA) = RB$, $Values(F_1) = Values(F_2) = Values(BA) = \{EC,\ SO,\ J\}$, $Valpref_{F_1} = \{SO <^{S_t}_{F_1} J <^{S_t}_{F_1} EC\}$; $Valpref(F_2) = \{EC <^{S_t}_{F_2} J <^{S_t}_{F_2} SO\}$; $Valpref(BA) = \{EC <^{S_t}_{BA} SO <^{S_t}_{BA} J\}$. Therefore, taking into account that F_1 and F_2 have a charity dependency relation between them, the AFAS for this example is shown in the figure 1a.

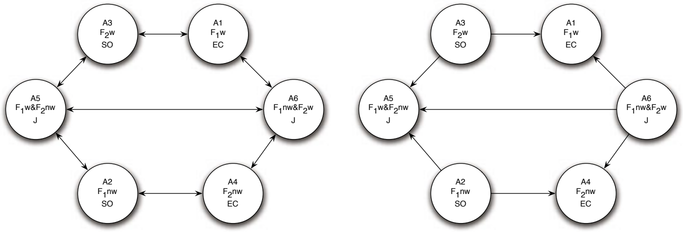

Fig. 1. a) Example $AFAS$; b) Example $AFAS_{F_2}$

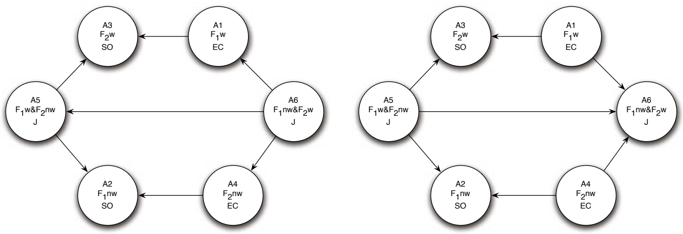

Fig. 2. a) Example of $AFAS_{F_1}$; b) $AFAS_{F_1}$ modified

Now, lets consider what happens with specific agents by creating their AAFAS. For instance, recalling that F_1 prefers economy to other values and gives solidarity the lesser value ($SO <^{S_t}_{F_1} J <^{S_t}_{F_1} EC$) we have that $AAFAS_{F_1}$ is the following: $AAFAS_{F_1} = <$ Ag, Rl, D, G, N, A, R, V, $Role$, $Dependency_{S_t}$, $Group$, $Values$, val, $Valpref_{F_1}$ $>$. Then, eliminating the unsuccessful attacks (due to value preferences of F_1) we have the equivalent $AFAS_{F_1}$ for $AAFAS_{F_1}$ as $AAFAS_{F_1}$ $= <$ A, $\{attacks(A1, A3)$, $attacks(A2, A4)$, $attacks(A5, A2)$, $attacks(A5, A3)$, $attacks(A6, A5)$, $attacks(A6, A1)$, $attacks(A6, A4)\}$, S_t $>$ which is shown in the graph of Figure 2a. This graph has the preferred extension $PE_{F_1} = \{A6\}$, meaning that F_2 should be the beneficiary of the water-right transfer to promote solidarity and the no intervention of a jury. This demonstrates how the power dependency relation of BA prevails over farmers and their arguments. Otherwise, if we change the environment and set a charity dependency relation of basin administrators over farmers $Farmer <^{S_t}_{Ch} BasinAdministrator$, the preferences of F_1 would prevail and the graph would be as the one of Figure 2b. In this case, the preferred extension would be $PE_{F_1} modified = \{A1, A4, A5\}$ that would defend F_1 as the benefitiary of the transfer agreement.

In its turn, F_2 gives the highest value to solidarity, but prefers to avoid a jury over economy ($EC <^{S_t}_{F_2} F <^{S_t}_{F_2} SO$). Therefore, its asssociated $AAFAS_{F_2}$ would be the following: $AAFAS_{F_2} = <$ Ag, Rl, D, G, N, A, R, V, $Role$, $Dependency_{S_t}$, $Group$, $Values$, val, $Valpref_{F_2}$ $>$. Then, eliminating the unsuccessful attacks we have the equivalent $AFAS_{F_2}$ for $AAFAS_{F_2}$ as $AF_{F_2} = <$ A, $\{$*attacks(A3, A1), attacks(A2, A4), attacks(A2, A5), attacks(A3, A5), attacks(A6, A5), attacks(A6, A1), attacks(A6, A4)*$\}$, S_t $>$ which is shown in the graph of the

Figure 1b. This graph has the preferred extension $PE_{F_2} = A2, A3, A6$ that means that F_2 defends its position as beneficiary of the water transfer.

5 Conclusion

In this paper we have presented an abstract argumentation framework to help reaching agreements in agent societies. After defining our concept of agent society, we have provided the formal definition of our argumentation framework. This is an extension of Dung's framework [11] to include agents's values, value preference orders and dependency relations. The framework has been illustrated in a real scenario of a water-rights transfer market.

Acknowledgements. This work is supported by the Spanish government grants CONSOLIDER INGENIO 2010 CSD2007-00022, TIN2008-04446 and TIN2009-13839-C03-01 and by the GVA project PROMETEO 2008/051.

References

1. Rahwan, I.: Argumentation in multi-agent systems. Autonomous Agents and Multiagent Systems, Guest Editorial 11(2), 115–125 (2006)
2. Bench-Capon, T., Dunne, P.: Argumentation in artificial intelligence. Artificial Intelligence 171(10-15), 619–938 (2007)
3. Ferber, J., Gutknecht, O., Michel, F.: From agents to organizations: an organizational view of multi-agent systems. In: Giorgini, P., Müller, J.P., Odell, J.J. (eds.) AOSE 2003. LNCS, vol. 2935, pp. 214–230. Springer, Heidelberg (2004)
4. Oliva, E., McBurney, P., Omicini, A.: Co-argumentation artifact for agent societies. In: Rahwan, I., Moraitis, P. (eds.) Argumentation in Multi-Agent Systems. LNCS (LNAI), vol. 5384. Springer, Heidelberg (2009)
5. Bench-Capon, T., Atkinson, K.: Abstract argumentation and values. In: Argumentation in Artificial Intelligence, pp. 45–64 (2009)
6. Dignum, V.: PhD Dissertation: A model for organizational interaction: based on agents, founded in logic. PhD thesis (2003)
7. Artikis, A., Sergot, M., Pitt, J.: Specifying norm-governed computational societies. ACM Transactions on Computational Logic 10(1) (2009)
8. Criado, N., Argente, E., Botti, V.: A Normative Model For Open Agent Organizations. In: International Conference on Artificial Intelligence, ICAI 2009 (2009)
9. Perelman, C., Olbrechts-Tyteca, L.: The New Rhetoric: A Treatise on Argumentation (1969)
10. Searle, J.R.: Rationality in Action (2001)
11. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n -person games. Artificial Intelligence 77, 321–357 (1995)
12. Baroni, P., Giacomin, M.: Semantics of Abstract Argument Systems. In: Argumentation in Artificial Intelligence, pp. 25–44. Springer, Heidelberg (2009)
13. Botti, V., Garrido, A., Giret, A., Noriega, P.: Managing water demand as a regulated open mas. In: Workshop on Coordination, Organization, Institutions and Norms in agent systems in on-line communities, COIN 2009, vol. 494, pp. 1–10 (2009)