Managing Social Influences through Argumentation-Based Negotiation

Nishan C. Karunatillake[1], Nicholas R. Jennings[1],
Iyad Rahwan[2], Sarvapali D. Ramchurn[1]

[1] School of Electronics and Computer Science, University of Southampton, Southampton, UK.
{nnc02r,nrj,sdr}@ecs.soton.ac.uk
[2] Institute of Informatics, The British University in Dubai, P.O.Box 502216 Dubai, UAE.
(Fellow) School of Informatics, University of Edinburgh, Edinburgh, UK.
irahwan@acm.org

Abstract. Social influences play an important part in the actions that an individual agent may perform within a multi-agent society. However, the incomplete knowledge and the diverse and conflicting influences present within such societies, may stop an agent from abiding by all its social influences. This may, in turn, lead to conflicts that the agents need to identify, manage, and resolve in order for the society to behave in a coherent manner. To this end, we present an empirical study of an argumentation-based negotiation (ABN) approach that allows the agents to detect such conflicts, and then manage and resolve them through the use of argumentative dialogues. To test our theory, we map our ABN model to a multi-agent task allocation scenario. Our results show that using an argumentation approach allows agents to both efficiently and effectively manage their social influences even under high degrees of incompleteness. Finally, we show that allowing agents to argue and resolve such conflicts early in the negotiation encounter increases their efficiency in managing social influences.

1 Introduction

Autonomous agents usually operate as a multi-agent community performing actions within a shared social context to achieve their individual and collective objectives. In such situations, the actions of these individual agents are influenced via two broad forms of motivations. First, the *internal influences* reflect the intrinsic motivations that drive the individual agent to achieve its own internal objectives. Second, as agents reside and operate within a social community, the social context itself influences their actions. For instance, within a structured society an agent may assume certain specific roles or be part of certain relationships. These, in turn, may influence the actions that an agent may perform. Here, we categorise such external forms of motivations as *social influences*.

Now, in many cases, both these forms of influence are present and they may give conflicting motivations to the individual agent. For instance, an agent may be internally motivated to perform a specific action. However, at the same time, it may also be subject to an external social influence (via the role it is enacting or the relationship that it is part of) not to perform it. Also an agent may face situations where different social influences motivate it in a contradictory fashion (one to perform a specific action and the other not

to). Furthermore, in many cases, agents have to carry out their actions in environments in which they are not completely aware of all the roles, relationships, or the ensuing commitments that they and their counterparts enact. Thus, in such instances, an agent may not be aware of the existence of all the social influences that could or indeed should affect its actions and it may also lack the knowledge of certain specific social influences that motivate other agents' actions. Therefore, when agents operate in a society with incomplete information and with such diverse and conflicting influences, they may, in certain instances, lack the knowledge, the motivation and/or the capacity to abide by all their social influences.

However, to function as a coherent society it is important for these agents to have a means to resolve such conflicts, manage their internal and social influences, and to come to a mutual understanding about their actions. To this end, *Argumentation-Based Negotiation* (ABN) has been advocated as a promising means of resolving conflicts within such agent societies [7, 12]. In more detail, ABN allows agents to exchange additional meta-information such as justifications, critics, and other forms of persuasive locutions within their interactions. These, in turn, allow agents to gain a wider understanding of the internal and social influences affecting their counterparts, thereby making it easier to resolve certain conflicts that arise due to incomplete knowledge. Furthermore, the negotiation element within ABN also provides a means for the agents to achieve mutually acceptable agreements to the conflicts of interests that they may have in relation to their different influences.

Against this background, this work advances the state of the art in the following ways. First, our main contribution is to propose a novel ABN approach that allows agents to detect, manage, and resolve conflicts related to their social influences in a distributed manner within a structured agent society. In order to demonstrate the performance benefits of our method, we use our proposed ABN framework to design a number of ABN strategies to manage such conflicts and then use an empirical evaluation to assess their impact. Specifically, we show that allowing agents to argue about their social influences provides them with the capability to not only manage their social influence more effectively, but to do so more efficiently as a society. Furthermore, we show that giving these agents the capability to challenge their counterparts and obtain their reasons for violating social commitments (instead of simply attempting to claim the penalty charges to which they are entitled) allows the agents to manage their social influences even more efficiently. Our second main contribution is to the ABN community. Here, we present a complete ABN framework which allows agents to argue and negotiate and resolve conflicts in the presence of social influences. Furthermore, we demonstrate the versatility of that framework; first, by mapping it to a specific computational problem of a multi-agent task allocation scenario and second, by using it to design a number of ABN strategies to resolve conflicts within a multi-agent society.

To this end, the remainder of the paper is structured as follows. First, Section 2 highlights the theoretical model of our ABN framework. Section 3 then maps this model to a computational context detailing the different representations and algorithms used. Subsequently, Section 4 details the experimental setting, presents our results and an analysis of the key observations. Next, Section 5 discusses the related work and Section 6 concludes.

2 Social Argumentation Model

In this section, we give a brief overview of our formal and computational framework for arguing and negotiating in the presence of social influences. In abstract, our framework consists of four main elements: (i) a *schema* for reasoning about social influence, (ii) a set of *social arguments* that make use of this schema, (iii) a *language and protocol* for facilitating dialogue about social influence, and (iv) a set of *decision functions* that agents may use to generate dialogues within the protocol. In the following sub-sections, we discuss each of these elements in more detail.[1]

2.1 Social Influence Schema

The notion of *social commitment* acts as our basic building block for capturing social influence. First introduced by Castelfranchi [3], it remains simple, yet expressive, and is arguably one of the fundamental approaches for modelling social behaviour among agents in multi-agent systems. In essence, a social commitment $(\mathrm{SC}_{\theta}^{x \rightarrow y})$ is a commitment by one agent x (termed the *debtor*) to another y (termed the *creditor*) to perform a stipulated action θ. As a result of such a social commitment, the debtor is said to attain an *obligation* toward the creditor, to perform the stipulated action. The creditor, in turn, attains certain rights. These include the right to demand or require the performance of the action, the right to question the non-performance of the action, and, in certain instances, the right to demand compensation to make good any losses suffered due to its non-performance. We refer to these as *rights to exert influence*. This notion of social commitment, resulting in an obligation and rights to exert influence, allows us a means to capture social influences between two agents. In particular, obligations reflect the social influences an agent is subjected to, while rights reflect the social influences the agent is capable of exerting on others.

Given this basic building block for modelling social influence between specific pairs of agents, we now proceed to explain how this notion is extended to capture social influences resulting due to factors such as roles and relationships within a wider multi-agent society (i.e., those that rely on the structure of the society, rather than the specific individuals who happen to be committed to one another). Specifically, since most relationships involve the related parties carrying out certain actions for each other, we can view a relationship as an encapsulation of social commitments between the associated roles. For instance, a relationship between the roles *supervisor* and *student* may be associated with a social commitment "to hand over the thesis in a timely manner." This social commitment, in turn, gives the student an obligation toward the supervisor to hand in the thesis, and gives the supervisor the right to exert influence on the student by either demanding that he does so or through questioning his/her non-performance. In a similar manner, the supervisor may be influenced to review and comment on the thesis. This again is another social commitment associated with the relationship. In this instance, it subjects the supervisor to an obligation to review the thesis while the student gains the right to demand its performance. In this manner, social commitment again provides

---

[1] It is important to note that here we only give a basic recap of our model to enable the reader to gain an overall understanding. A comprehensive formal representation of the framework can be found in [8, 9].

an effective means to capture the social influences emanating through roles and relationships of the society (independently of the specific agents who take on the roles). Given this descriptive definition of our model, we now formulate these notions to capture the social influences within multi-agent systems as a schema (refer to Figure 1 and formulae (1) through (6)):

Definition 1: For $n_A, n_R, n_P, n_\Theta \in \mathbb{N}^+$, let:
- $A = \{a_1, \ldots, a_{n_A}\}$ denote a finite set of agents,
- $R = \{r_1, \ldots, r_{n_R}\}$ denote a finite set of roles,
- $P = \{p_1, \ldots, p_{n_P}\}$ denote a finite set of relationships,
- $\Theta = \{\theta_1, \ldots, \theta_{n_\Theta}\}$ denote a finite set of actions,
- Act : $A \times R$ denote the fact that an agent is acting a role,
- RoleOf : $R \times P$ denote the fact that a role is related to a relationship, and
- In : $A \times R \times P$ denote the fact that an agent acting a role is part of a relationship.

If an agent acts a certain role and that role is related to a specific relationship, then that agent acting that role is said to be part of that relationship (as per Cavedon and Sonenberg [4]):

$$\text{Act}(a, r) \land \text{RoleOf}(r, p) \to \text{In}(a, r, p) \qquad \text{(Rel. Rule)}$$

Definition 2: Let SC denote a finite set of social commitments and $\text{SC}_\theta^{x \to y} \in SC$. Thus, as per [3], $\text{SC}_\theta^{x \to y}$ *will result in the debtor attaining an obligation toward the creditor to perform a stipulated action and the creditor, in turn, attaining the right to influence the performance of that action:*

$$\text{SC}_\theta^{x \to y} \to [\text{O}_\theta^{x \to y}]_x^f \land [\text{R}_\theta^{y \to x}]_y, \qquad \text{(S-Com Rule)}$$

where:
- $[\text{O}_\theta^{x \to y}]_x^f$ represents the obligation that x attains that subjects it to an influence of a degree f (refer to [9] for more details) toward y to perform θ and
- $[\text{R}_\theta^{y \to x}]_y$ represents the right that y attains which gives it the ability to demand, question, and require x regarding the performance of θ.

Definition 3: Let:
- DebtorOf : $(R \cup A) \times SC$ denote that a role (or an agent) is the debtor in a social commitment,
- CreditorOf : $(R \cup A) \times SC$ denote that a role (or an agent) is the creditor in a social commitment,
- ActionOf : $\Theta \times SC$ denote that an act is associated with a social commitment, and
- AssocWith : $SC \times P$ denote that a social commitment is associated with a relationship.

If the roles associated with the relationship are both the creditor and the debtor of a particular social commitment, then we declare that social commitment is associated with the relationship (as per Section 2.1).

Applying the Rel. Rule to a society where: $a_i, a_j \in A \land r_i, r_j \in R \land p \in P$ s.t. $\text{Act}(a_i, r_i)$, $\text{Act}(a_j, r_j)$, $\text{RoleOf}(r_i, p)$, $\text{RoleOf}(r_j, p)$ hold true, we obtain:

$$\text{Act}(a_i, r_i) \land \text{RoleOf}(r_i, p) \to \text{In}(a_i, r_i, p) \qquad (1)$$
$$\text{Act}(a_j, r_j) \land \text{RoleOf}(r_j, p) \to \text{In}(a_j, r_j, p). \qquad (2)$$

Now, consider a social commitment $\text{SC}_\theta^{r_i \to r_j}$ associated with the relationship p in this society. Applying this to Definition 3 we obtain:

$$(\text{DebtorOf}(r_i, \text{SC}) \land \text{RoleOf}(r_i, p)) \land (\text{CreditorOf}(r_j, \text{SC}) \land \text{RoleOf}(r_j, p))$$
$$\land \text{ActionOf}(\theta, \text{SC}) \to \text{AssocWith}(\text{SC}_\theta^{r_i \to r_j}, p). \qquad (3)$$

Applying the S-Comm rule to $\text{SC}_\theta^{r_i \to r_j}$ we obtain:

$$\text{SC}_\theta^{r_i \to r_j} \to [\text{O}_\theta^{r_i \to r_j}]_{r_i}^f \land [\text{R}_\theta^{r_j \to r_i}]_{r_j}. \qquad (4)$$

An agent $\boldsymbol{a}_i$ acting the role $\boldsymbol{r}_i$

Leads it to be part of the relationship $\boldsymbol{p}$

With another agent $\boldsymbol{a}_j$ acting the role $\boldsymbol{r}_j$

A social commitment $\boldsymbol{SC}_\theta^{r_i \to r_j}$ associated with $\boldsymbol{p}$
- Leads to $\boldsymbol{a}_i$ attaining an obligation $\boldsymbol{O}$ toward $\boldsymbol{r}_j$,
  Which subjects it to an influence of degree $\boldsymbol{f}$
  To perform the action θ
- And, in turn, leads to $\boldsymbol{a}_j$ attaining the right $\boldsymbol{R}$ toward $\boldsymbol{r}_i$
  To demand, question, and require the performance
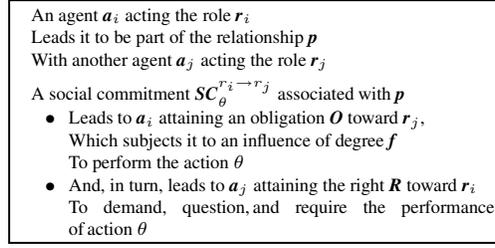 of action θ

Fig. 1. Schema of Social Influence.

Combining (1), (3) and (4) we obtain:

$$\text{In}(a_i, r_i, p) \wedge \text{AssocWith}(\text{SC}_\theta^{r_i \to r_j}, p) \to \left[\text{O}_\theta^{a_i \to r_j} \right]_{a_i}^f . \tag{5}$$

Combining (2), (3) and (4) we obtain:

$$\text{In}(a_j, r_j, p) \wedge \text{AssocWith}(\text{SC}_\theta^{r_i \to r_j}, p) \to \left[\text{R}_\theta^{a_j \to r_i} \right]_{a_j} . \tag{6}$$

2.2 Social Arguments

Having captured the notion of social influence into a schema, we now show how agents can use this schema to systematically identify social arguments to negotiate in the presence of social influences. Specifically, we identify two major ways in which social influence can be used to change decisions and, thereby, resolve conflicts between agents.

Socially Influencing Decisions. One way to affect an agent's decisions is by arguing about the validity of that agent's practical reasoning [2]. Similarly, in a social context, an agent can affect another agent's decisions by arguing about the validity of the other's social reasoning. In more detail, agents' decisions to perform (or not) actions are based on their internal and/or social influences. Thus, these influences formulate the *justification* (or the reason) behind their decisions. Therefore, agents can affect each other's decisions indirectly by affecting the social influences that determine their decisions. Specifically, in the case of actions motivated via social influences through the roles and relationships of a structured society, this justification to act (or not) flows from the social influence schema (see Section 2.1). Given this, we can further classify the ways that agents can socially influence each other's decisions into two broad categories:

1. Undercut the opponent's existing justification to perform (or not) an action by disputing certain premises within the schema that motivates its opposing decision (i.e., dispute a_i is acting role r_i, dispute SC is a social commitment associated with the relationship p, dispute θ is the action associated with the obligation O, etc.).
2. Rebut the opposing decision to act (or not) by,
    i. Pointing out information about an alternative schema that justifies the decision not to act (or act as the case may be) (i.e., point out that a_i is also acting in role r_i, that SC is also a social commitment associated with the relationship p, that θ is the action associated with the obligation O, etc.).
    ii. Pointing out information about conflicts that could or should prevent the opponent from executing its opposing decision (i.e., point out conflicts between two existing *obligations*, *rights*, and *actions*).
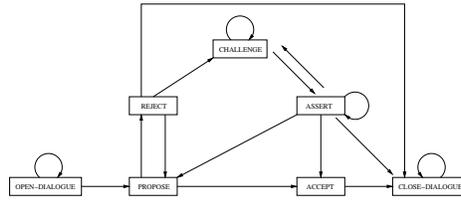
Fig. 2. Dialogue Interaction Diagram.

Negotiating Social Influence. Agents can also use social influences within their negotiations. More specifically, as well as using social argumentation as a tool to affect decisions (as above), agents can also use negotiation as a tool for "trading social influences". In other words, the social influences are incorporated as additional parameters of the negotiation object itself. For instance, an agent can promise to (or threaten not to) undertake one or many future obligations if the other performs (or does not perform) a certain action. It can also promise not to (or threaten to) exercise certain rights to influence one or many existing obligations if the other performs (or does not perform) a certain action. In this manner, the agents can use their obligations, rights, and even the relationship itself as parameters in their negotiations.

2.3 Language and Protocol

To enable agents to express their arguments, we define two complimentary languages: the *domain language* and the *communication language* (see [8] for a complete formal specification). The former allows the agents to express premises about their social context and also the conflicts that they may face while executing actions within such a context. The communication language, on the other hand, enables agents to express premises about the social context in the form of arguments and, thereby, engage in their discourse to resolve conflicts. This consists of seven elocutionary particles (i.e., *OPEN-DIALOGUE*, *PROPOSE*, *ACCEPT*, *REJECT*, *CHALLENGE*, *ASSERT*, and *CLOSE-DIALOGUE*). These locutions can be used together with content expressed in the domain language in order to allow agents to make *utterances* (e.g., assert a particular social premise, challenge a premise, make a specific proposal, and so on).

The protocol, which indicates the legal ordering of communication utterances, has six main stages: (i) *opening*, (ii) *conflict recognition*, (iii) *conflict diagnosis*, (iv) *conflict management*, (v) *agreement*, and (vi) *closing*. The opening and closing stages provide the important synchronisation points for the agents involved in the dialogue, the former indicating its commencement and the latter its termination [11]. The conflict recognition stage, the initial interaction between the agents, brings the conflict to the surface. Subsequently, the diagnosis stage allows the agents to establish the root cause of the conflict and also to decide on how to address it (i.e., whether to avoid the conflict or attempt to manage and resolve it through argumentation and negotiation [7]). Next, the conflict management stage allows the agents to argue and negotiate, thus, addressing the cause of this conflict. Finally, the agreement stage brings the argument to an end, either with the participants agreeing on a mutually acceptable solution or agreeing to disagree due to the lack of such a solution. In operation, it is defined as a dialogue game protocol which gives locutions rules (indicating the moves that are permitted), com-

---

Algorithm 1 Decision making algorithm for *PROPOSE*.

1: **if** $(Capable(do(a_i, \theta_i)) \wedge B^{a_i}_{do(a_j, \theta_j)} > C^{a_i}_{do(a_i, \theta_i)})$ **then**
2: $\quad$ $PROPOSE(do(a_j, \theta_j), do(a_i, \theta_i))$
3: **end if**

Algorithm 2 Decision making algorithm for *ACCEPT* or *REJECT*.

1: **if** $(Capable(do(a_j, \theta_j)) \wedge B^{a_j}_{do(a_i, \theta_i)} > C^{a_j}_{do(a_j, \theta_j)})$ **then**
2: $\quad$ $ACCEPT(do(a_j, \theta_j), do(a_i, \theta_i))$
3: **else**
4: $\quad$ $REJECT(do(a_j, \theta_j), do(a_i, \theta_i))$
5: **end if**

---

mitment rules (defining the commitments each participant incurs with each move), and structural rules (specifying the types of moves available following the previous move). Figure 2 presents these locutions and structural rules in abstract.

2.4 Decision Making Functionality

The protocol described above gives agents a number of different options, at various stages, as to what utterances to make. For instance, after a proposal the receiving agent could either accept or reject it. After a rejection, the agent may choose to challenge this rejection, end the dialogue, or forward an alternative proposal. An agent, therefore, still requires a mechanism for selecting a particular utterance among the available legal options. To this end, for each of the possible dialogue moves, we specify general decision making algorithms to give the agents that capability. Specifically, Algorithms 1 and 2 show two such examples, the former for generating a proposal and the latter for evaluating such a proposal. In abstract, a proposal in our formulation has two aspects; the request and the reward. Thus, when generating a proposal the agent would assess two aspects (i) if it is capable of performing the reward and (ii) the benefit it gains from the request $(B^{a_i}_{do(a_j, \theta_j)})$ is greater than the cost of reward $(C^{a_i}_{do(a_i, \theta_i)})$ (Algorithm 1). On the other hand, when evaluating a proposal, the agent will consider (i) if it is capable of performing the request and (ii) that the benefit of the reward $(B^{a_j}_{do(a_i, \theta_i)})$ is greater than the cost incurred in performing the request $(C^{a_j}_{do(a_j, \theta_j)})$ (Algorithm 2).

3 Argumentation Context

To evaluate how our argumentation model can be used as a means of managing social influences, we require a computational context in which a number of agents interact in the presence of social influences and conflicts arise as a natural consequence of these interactions. To this end, we now proceed to detail how we map our general framework into a specific multi-agent task allocation scenario. We first provide an overview description of the scenario and then proceed to explain how we map the notion of social influence within it. Finally, we detail how the agents can use our ABN model to interact within this social context and manage conflicts related to their social influences.

3.1 The Scenario

The argumentation context is based on a simple multi-agent task allocation scenario (similar to that presented in [7]) where a collection of self-interested agents interact

Table 1. A Sample Scenario

Time	a_0	a_1	a_2
	$c_{(0,0.9)}, c_{(1,0.1)}$	$c_{(0,0.1)}, c_{(1,0.9)}$	$c_{(0,0.4)}, c_{(1,0.5)}$
t_0	$\theta_0 : [c_{(0,0.5)}, 200]$	$\theta_0 : [c_{(1,0.2)}, 500]$	$\theta_0 : [c_{(1,0.5)}, 700]$
t_1	$\theta_1 : [c_{(1,0.3)}, 900]$	$\theta_1 : [c_{(0,0.4)}, 300]$	$\theta_1 : [c_{(1,0.7)}, 100]$
t_2	$\theta_2 : [c_{(1,0.1)}, 400]$	$\theta_2 : [c_{(0,0.8)}, 900]$	
t_3	$\theta_3 : [c_{(0,0.9)}, 600]$		

to obtain services to achieve a given set of actions. In abstract, the context consists of two main elements. On one hand, each agent in the system has a list of *actions* that it is required to achieve. On the other hand, all agents in the system have different *capabilities* to perform these actions. In this context, agents are allowed to interact and negotiate between one another to find capable counterparts that are willing to sell their services to perform their actions. The following introduce these main elements in more detail:

Capability: All agents within the domain have an array of capabilities. Each such capability has two parameters: (i) a type value (x) defining the type of that capability and (ii) a capability level $(d \in [0, 1])$ defining the agent's competence level in that capability (1 indicates total competence, 0 no competence). Given this, we denote a capability as $c_{(x,d)} : [x, d]$.

Action: Each action has four main parameters: (i) the specified time (t_i) the action needs to be performed, (ii) the capability type (x) required to perform it, (iii) the minimum capability level (d_m) required, and (iv) the reward $(r_i$; distributed normally with a mean μ and a standard deviation $\sigma)$ the agent would gain if the action is completed. Given this, we denote an action as $\theta_i : [t_i, c_{(x,d_m)}, r_i]$.

Each agent within the context is seeded with a specified number of such actions. This number varies randomly between agents within a pre-specified range. Table 1 depicts one such sample scenario for a three agent context $(a_0, a_1,$ and $a_2)$ with their respective capabilities and actions.

3.2 Modelling Social Influences

Given our argumentation context, we now describe how social influences are mapped into it. In order to provide the agents with different social influences, we embody a role-relationship structure into the multi-agent society. To do so, first, we define a specific number of roles and randomly link them to create a web of relationships. This defines the role-relationship structure. Figure 3(a) shows an example of such a representation between 3 roles: r_1, r_2, and r_3, where 1 indicates that a relationship exists between the two related roles, and 0 indicates no relationship.

Given this role-relationship structure, we now randomly specify social commitments for each of the active relationship edges (those that are defined as 1 in the mapping). A social commitment in this context is a commitment by one role, to another, to provide a certain type of capability when requested. As per Section 2.1, an important component of our notion of social commitment is its associated degree of influence. Thus, not all social commitments influence the agents in a similar manner (for more

	r_0	r_1	r_2
r_0	0	1	0
r_1	1	0	1
r_2	0	1	0

(a) Rol-Rel mapping.

	r_0	r_1	r_2
r_0	[0:0]	[200:0]	[0:0]
r_1	[400:100]	[0:0]	[200:600]
r_2	[0:0]	[700:200]	[0:0]

(b) Social commitment mapping.

	r_0	r_1	r_2
a_0	1	0	0
a_1	0	1	1
a_2	0	1	0

(c) Ag-Rol mapping.

Fig. 3. Social Influence Model.

details refer to [9]). Here, we map these different degrees of influence by associating each social commitment with a decommitment penalty. Thus, any agent may violate a certain social commitment at any given time. However, it will be liable to pay the specified decommitment value for this violation (this is similar to the notion of levelled commitments introduced in [14]). Since all our agents are self-interested, they prefer not to lose rewards in the form of penalties, so a higher decommitment penalty yields a stronger social commitment (thereby, reflecting a higher social influence). The following represents such a mapping. For instance, in Figure 3(b) the entry [400:100] in row 1, column 2 indicates that the role r_0 is committed to provide capabilities c_0 and c_1 to a holder of the role r_1. If the agent holding the role r_0 chooses not to honour this commitment it will have to pay 400 and 100 (respectively for c_0 and c_1) if asked. Having designed this social structure and the associated social commitments, finally we assign these roles to the actual agents operating within our system as shown in Figure 3(c).

From this representation, we can easily extract the rights and the obligations of each agent within our system. For instance, the agent-role mapping shows the fact that agent a_0 acts the role r_0. Given this, its obligations and rights can be extracted as follows:

- *Obligation to provide*:
  - c_0 to an agent acting r_1; obliged to pay 400 if decommitted.
  - c_1 to an agent acting r_1; obliged to pay 100 if decommitted.
- *Rights to demand*:
  - c_0 from an agent acting r_1; right to demand 200 if decommitted.

Given this global representation of social influence, we will now detail how we seed these agents with this information. Since one of the aims in our experiments is to test how agents use argumentation to manage and resolve conflicts created due to incomplete knowledge about their social influences, we generate a number of settings by varying the level of knowledge seeded to the agents. More specifically, we give only a subset of the agent-role mapping.[2] We achieve this by randomly replacing certain 1s with 0s and give this partial knowledge to the agents during initialisation. Thus, a certain agent may not know all the roles that it or another agent may act. This may, in turn, lead to conflicts within the society, since certain agents may know certain facts about the society that others are unaware of. By controlling this level of change, we

---

[2] Theoretically it is possible to introduce imperfections to all the premises within the schema (i.e., $Act(a_i, r_i)$, $RoleOf(r_i, p)$, $AssocWith(SC^{r_i \leftarrow r_j}, p)$, $InfluenceOf(O, f)$ etc.; see Section 2.1). However, since the objective of our experiments is to prove the concept of how arguments can resolve conflicts, instead of designing an exhaustive implementation with all possible imperfections and arguments, we chose to concentrate on the first two premises. Increasing the imperfections would merely increase the reasons why a conflict may occur, thus, bringing more arguments into play. However, this would have little bearing on the general pattern of the results.

Algorithm 3 The *negotiate()* method.	**Algorithm 4** The *argue()* method.
1: $[p_0, p_1, \ldots, p_{max}] \leftarrow$ *generateProposals()*	1: {Challenge for the opponent's justification}
2: $p \leftarrow p_0$	2: $H_o \leftarrow$ *challenegeJustification()*
3: $isAccepted \leftarrow$ **false**	3: {Generate personal justification}
4:	4: $H_p \leftarrow$ *generateJustification()*
5: {Loop till either the agent agrees or the last proposal fails.}	5:
	6: **if** $(isValid(H_o) =$ **false)** **then**
6: **while** $(isAccepted \neq$ **true** $\| \ p \leq p_{max})$ **do**	7: {Assert invalid premises of H_o}
7: $\quad response \leftarrow PROPOSE(p)$	8: **else**
8: $\quad$ **if** $(response =$ "$accept$") **then**	9: {Adopt premises of H_o into personal knowledge}
9: $\quad\quad isAccepted \leftarrow$ **true**	10: **end if**
10: $\quad$ **else**	11: **if** $(isValid(H_p) =$ **false)** **then**
11: $\quad\quad$ **if** $(p \neq p_{max})$ **then**	12: {Correct invalid premises of H_p within personal knowledge}
12: $\quad\quad\quad p \leftarrow$ *getNextViableProposal()*	
13: $\quad\quad$ **end if**	13: **else**
14: $\quad$ **end if**	14: {Assert H_p}
15: **end while**	15: **end if**
16: **return** $isAccepted$	

generate an array of settings ranging from perfect knowledge (0% missing knowledge) in the society, to the case where agents are completly unaware of their social influences (100% missing knowledge).

To explain this further, consider for instance that when initialising a_0 we seeded it with an incomplete agent-role map by replacing the 1 in column 1, row 1 with a 0. Thus, a_0 is unaware that it is acting the role r_0. As a result, it is not aware of its ensuing obligations and rights highlighted above. Now, when agents interact within the society this may lead to conflicts between them. For example, if a_0 refused to provide c_0 to a_1, it may request that the violation penalty of 400 be paid. However, since a_0 is unaware of its obligation it will not pay the amount. On the other hand, when initialising a_0 if we replace the 1 in column 2, row 3 with a 0, a_0 would now be unaware of its obligations towards agent a_2 since its lacks the information that its counterpart a_2 acts the role r_1. This, in turn, would also lead to conflicts with the society. In these situations, agents can use the argumentation process explained in Section 3.3 to argue and resolve such conflicts.

3.3 Agent Interaction

Having detailed the multi-agent context, we now proceed to discuss how the agents can use our ABN model to interact within this social setting. As mentioned in Section 3.1, agents within the system argue and negotiate with each other to find willing and capable partners to accomplish their actions. In essence, an agent that requires a certain capability will generate and forward *proposals* to another selected agent within the community requesting it to sell its services in exchange for a certain reward (Algorithm 1). If the receiving agent perceives this proposal to be viable and believes it is capable of performing it, then will *accept* it. Otherwise it will *reject* the proposal (Algorithm 2). In case of a reject, the original proposing agent will attempt to forward a modified proposal. The interaction will end either when one of the proposals is accepted or when all valid proposals that the proposing agent can forward are rejected (Algorithm 3). In this context, the two main elements of the negotiation interaction are:

Proposal Generation: When generating a proposal, an agent needs to consider two aspects (Algorithm 1): (i) whether it is capable of carrying out the *reward* and (ii) whether

the *benefit it gains from the request* is greater than the *cost incurred while performing the reward*. To simplify the implementation, we constrain our system to produce proposals with only monetary rewards. Thus, the generic proposal from an agent a_i to an agent a_j takes the form $PROPOSE(do(a_j, \theta_j), do(a_i, m))$ where θ_j is the requested action and m the monetary reward. In this context, calculating the benefit and the cost becomes straight forward. The benefit is the request r_j associated with the action θ_j and the cost of reward is m the monetary reward. Given this, the agent would generate an array of proposals with increasing amounts of monetary rewards, the lowest being 1 and the highest being $(r_j - 1)$.

Proposal Evaluation: When the receiving agent evaluates a proposal it also considers two analogous factors: (i) whether it is capable of performing the *request* and (ii) if the *benefit it gains from the reward* is greater than the *cost of carrying out the request* (Algorithm 2). To evaluate capability, the agent compares its own level with the minimum required to perform the action. In this case, the cost is the current opportunity cost. Here, all agents have a minimum asking price (set to μ the mean reward value, see Section 3.1) if they are not occupied, or, if they are, the cost is the reward plus the decommitment cost of the previously agreed action. The benefit, in the simplest case, is the monetary value of the reward m. However, if the agent has a social commitment to provide that capability type to the requesting agent, then the benefit is the monetary reward plus the decommitment penalty of this social commitment.

Given the negotiation interaction, we will now detail how agents argue (Algorithm 4) to resolve conflicts within the multi-agent society (such as the one highlighted in Section 3.2). Agents first detect conflicts by analysing the decommitment penalties paid by their counterparts for violating their social commitments. In more detail, when an agent with the right to demand a certain capability claims the penalty form another for violating its obligation and the amount paid in response is different from the amount it expects to receive, the agents would detect the existence of a conflict. Once such a conflict is detected agents attempt to resolve it by exchanging their respective justifications. These justifications would take the form of the social influence schema (see Equations 5 and 6 in Section 2.1) and are then analysed to diagnose the cause of the conflict. If there are inconsistencies between them, social arguments (Section 2.2; Type-1) are used to highlight these. If they are both valid, then each agent would point-out alternative justifications via asserting missing knowledge (Section 2.2; Type-2). The defeat-status is computed via a validation heuristic, which simulates a defeasible model such as [1].

4 Managing Social Influences

As mentioned in Section 1, when agents operate within a society with incomplete knowledge and with diverse and conflicting influences, they may, in certain instances, lack the knowledge, the motivation and/or the capacity to enact all their social commitments. In some cases, therefore, an agent may violate specific social commitments in favour of abiding by a more influential internal or external motivation. In other cases it may inadvertently violate such commitments simply due to the lack of knowledge of their existence. However, to function as a coherent society it is important for these agents to have a means to resolve such conflicts and manage their social influences in

Algorithm 5 Claim-Penalty-Non-Argue strategy.	**Algorithm 6** Claim-Penalty-Argue strategy.
1: $isAccepted \leftarrow negotiate()$ 2: **if** $(isAccepted = \textbf{false})$ **then** 3: compensation $\leftarrow$ *demandCompensation*() 4: **end if**	1: $isAccepted \leftarrow negotiate()$ 2: **if** $(isAccepted = \textbf{false})$ **then** 3: compensation $\leftarrow$ *demandCompensation*() 4: **if** (compensation $<$ rightToPenalty) **then** 5: *argue*() 6: **end if** 7: **end if**

a systematic manner. Against this background, we will now investigate a number of different interaction strategies that allow the agents to manage their social influences within a multi-agent context. The underlying motivation for these strategies is our social influence schema (see Section 2.1), which gives the agents different rights; namely the right to demand compensation and the right to challenge non-performance of social commitments. Specifically, in the following we use our ABN model to design both arguing and non-arguing strategies to implement these forms of interactions and assess their relative performance benefits.

The experiments are set within the context described in Section 3 with 20 agents, each having 3 capabilities with different levels of competence (varied randomly). The number of actions each agent has vary between 20 and 30, while their respective rewards are set according to a normal distribution with a mean 1,000 and a standard deviation 500. We use two metrics to evaluate the overall performance of the different strategies (similar to [7, 13]): (i) the *total earnings* of the population as a measure of effectiveness (the higher the value, the more effective the strategy) and (ii) the *total number of messages* used by the population as a measure of efficiency (the lower the value, the more efficient the strategy). Here all reported results are averaged over 40 simulation runs to diminish the impact of random noise, and all observations emphasised are statistically significant at the 95% confidence level.

4.1 Demanding Compensation

If an agent violates a social commitment, one of the ways its counterpart can react is by exercising its right to demand compensation. This formulates our baseline strategy which extends our negotiation algorithm by allowing the agents to demand compensation in cases where negotiation fails (Algorithm 5). Once requested, the agent that violated its social commitment will pay the related penalty.[3] However, in imperfect information settings, a particular agent may violate a social commitment simply because it was not aware of it (i.e., due to the lack of knowledge of its roles or those of its counterparts). In such situations, an agent may pay a decommitment penalty different to what the other believes it should get, which may, in turn, lead to conflicts. In such situations, our second strategy allows agents to use social arguments to argue about their social influences (as per Section 2.2) and, thereby, manage their conflicts (Algorithm 6). Our hypothesis here is that by allowing agents to argue about their social influences we are providing them with a coherent mechanism to manage and resolve their conflicts and,

---

[3] To reduce the complexity, here, we assume that our agents do not attempt to deceive one another. Thus an agent will either honour its obligation or pay the penalty. We could drop this assumption and make it more realistic by incorporating trust and reputation mechanism into the system. However, this is beyond the scope of this paper.

thereby, allowing them to gain a better outcome as a society. To this end, the former strategy acts as our control experiment and the latter as the test experiment. Figures 4 and 5 show our results from which we make the following observations:

Observation 1: *The argumentation strategy allows agents to manage their social influences even at high uncertainty levels.*
If agents are aware of their social influences, they may use them as parameters within their negotiation interactions. Thereby, agents can endorse certain actions which may otherwise get rejected (see Section 2.2). This would, in turn, increase the population earnings as more actions are accomplished. However, if the agents are not aware of their social influences they may not be able to use these influences to endorse such actions. Therefore, we can observe a downward trend in the population earnings for both strategies as the agent's knowledge level about their social influences decrease (0 on the X-axis indicates perfect information, whereas, 100 represents a complete lack of knowledge about the social structure). However, we can observe that the non-argue strategy falls more rapidly than the argue one. This is because the argue method allows agents to manage and resolve conflicts of opinion that they may have about their social influences. For instance, if a certain agent is unaware of a role that another acts, it may correct this through arguing with that agent. Thus, arguing allows agents to correct such gaps in their knowledge and, thereby, resolve any conflicts that may arise as a result. In this manner, ABN allows the agents to manage their social influences even at high uncertainty levels. Thereby, as a society, the agents can accomplish more of their actions and gain a higher total earnings value. The non-arguing approach, which does not allow them to argue about their social influences and manage such conflicts, reduces the population earnings as knowledge imperfections increase within the social system.

Observation 2: *In cases of perfect information and complete uncertainty, both strategies perform equally.*
The reason for both strategies performing equally when there is perfect information (0 level) is because there are no knowledge imperfections. In other words, agents do not need to engage in argumentation to correct conflicts of opinions simply because such conflicts do not exist. On the other hand, the reason for both strategies performing equally when there is a complete lack of knowledge is more interesting. Since, none of the agents within the society are aware of any social influences (even though they exist) they are not able to detect any conflicts or violations. Consequently, agents do not resort to arguing to manage such conflicts (see *conflict recognition* stage in Section 2.3). Thus, when there is a complete lack of knowledge, the strategy that uses the argue strategy performs the same as the non-argue one.

Observation 3: *At all knowledge levels, the argumentation strategy exchanges fewer messages than the non-arguing one.*
Figure 4(b) shows the number of messages used by both strategies under all knowledge levels. Apart from the two end points, where argumentation does not occur (see Observation 2), we can clearly see the non-arguing strategy exchanging more messages (is less efficient) than the argue one. The reason for this is that even though agents use some number of messages to argue and correct their incomplete knowledge, thereafter the agents use their corrected knowledge in subsequent interactions. However, if the agents do not argue to correct their knowledge imperfections, they negotiate more fre-
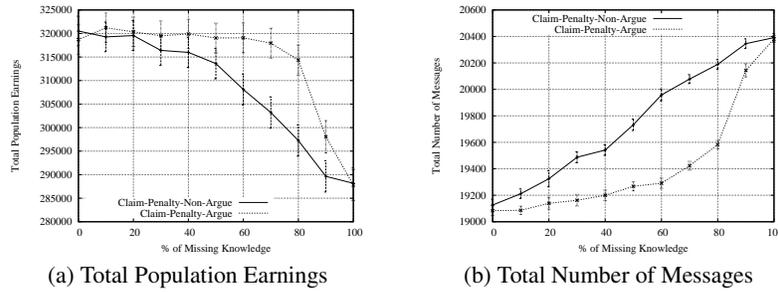
(a) Total Population Earnings   (b) Total Number of Messages

Fig. 4. Efficiency and Effectiveness of the Argue and Non-Argue strategies with 20 Agents and 3 Roles.

quently since they cannot use their social influence. Thus, this one-off increase of argue messages becomes insignificant when compared to the increase in the propose, accept, and reject messages due to the increased number of negotiations.

Observation 4: *When there are more social influences within the system, the performance benefit of arguing is only significant at high levels of knowledge incompleteness.* Figure 4(a) and Figures 5(a) through 5(d) show the effectiveness of both the strategies as the number of roles increases within the society. One of the key observations here is the decline rate of the non-argue strategy. We can see that as the number of roles increase, the rate of decline of the non-argue method becomes less pronounced. Furthermore, the crossover point where the non-argue method starts to be less effective than the argue strategy also shifts increasingly to the right (higher knowledge imperfections). In Figures 5(a) though 5(d) this level is roughly 50%, 70%, 80%, 90%. This again is a very interesting observation. As agents gain a higher number of roles, they aquire an increasing number of social influences. Now, as explained in Observation 1, the agents use these social influences as a resource to endorse their actions. Thus, when an agent has a higher number of social influences, its lack of knowledge about a certain particular influence makes little difference. The agent can easily replace it with another influence (which it is aware of) to convince its counterpart. Therefore, under such conditions, agents arguing about their social influences to correct their lack of knowledge would have little reward since the non-argue method can more simply replace it with another known influence and still achieve the same end. Only when an agent has a near complete lack of knowledge (i.e., 80%, 90%) does the argue strategy yeild significant performance gains. This observation complements our previous emperical study on the worth of argumentation at varying resource levels [7]. There we show that the benefit of arguing is more pronounced at low resource settings and under higher resource conditions the benefit is less.

4.2 Questioning Non-Performance

In the event that a particular social commitment is violated, apart from the right to demand compensation, our social influence schema also gives the agents the right to challenge and demand a justification for this non-performance (see Section 2.1). It is generally argued in ABN theory that allowing agents to exchange such meta-information in the form of justifications gives them the capability to understand each others' rea-
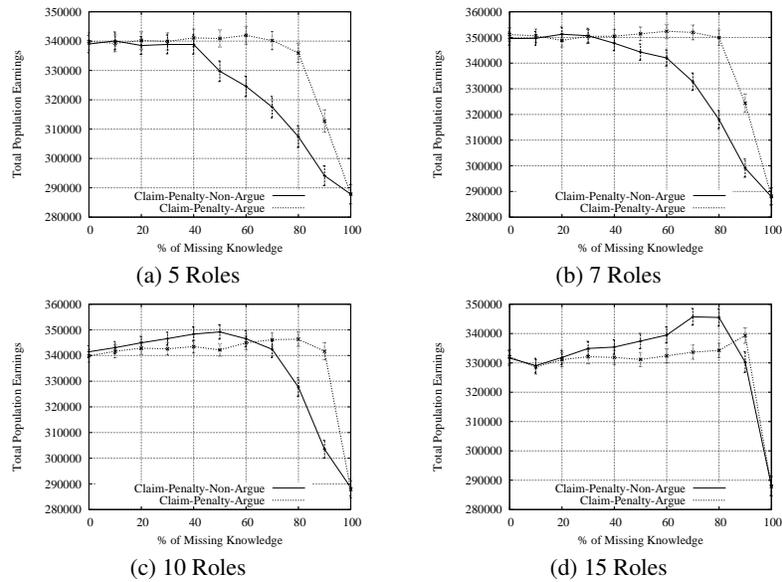
Fig. 5. Total population earnings with 20 agents and a varying number of roles.

sons and, thereby, provides a more efficient method of resolving conflicts under uncertainty [12]. In a similar manner, we believe that providing the agents with the capability to challenge and demand justifications for violating social commitments also allows the agents to gain a wider understanding of the internal and social influences affecting their counterparts, thereby, providing a more efficient method for managing social influences in the presence of incomplete knowledge.

This intuition forms the underlying hypothesis for our next set of experiments. More specifically, we use our previous best strategy *Claim-Penalty-Argue* as the control experiment and design two other strategies (*Argue-In-First-Rejection* and *Argue-In-Last-Rejection*) to experiment with the effect of allowing the agents to challenge non-performance at different stages within the negotiation encounter. The former allows the agent to challenge after the receipt of the first rejection and the latter after the last rejection. Thus, the two differ on when the agent attempts to find the reason (in the first possible instance or after all proposals have been forwarded and rejected).[4] Figures 6(a) and 6(b) show our results and the following highlight our key observations:

Observation 5: *The effectiveness of the various argumentation strategies are broadly similar.*
Figure 6(a) shows no significant difference in the effectiveness of the three ABN strategies. This is due to the fact that all three strategies argue and resolve the conflicts even though they decide to argue at different points within the encounter. Therefore, we do not expect to have any significant differences in number of conflicts resolved. Thus, the effectiveness stays the same.

---

[4] Due to space restrictions we avoid specifying the algorithms for these two strategies here.

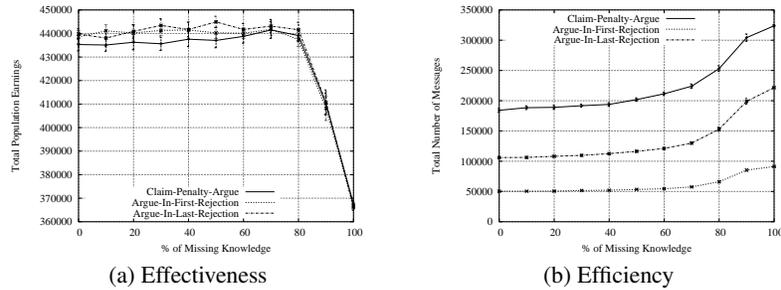(a) Effectiveness   (b) Efficiency

Fig. 6. Efficiency and Effectiveness of the various argumentation strategies.

Observation 6: *Allowing the agents to challenge earlier in the dialogue, significantly increases the efficiency of managing social influences.*

Figure 6(b) shows a significant difference in the number of messages used by the three strategies at all levels of knowledge. In more detail, the number of messages used by the *Argue-In-Last-Rejection* strategy is significantly lower than our original *Claim-Penalty-Argue* one. Moreover, the *Argue-In-First-Rejection* strategy has the lowest number of messages exchanged. The reason for this behaviour is based on how the agents use these reasons exchanged during the argue phase. In the *Claim-Penalty-Argue* strategy the main objective of arguing is to resolve the conflict regarding the penalty value that should be paid. However, it does not attempt to find out the reason for why its counterpart rejected its proposal. For instance, one reason could be the lack of capability. Another could be the reward of the proposal is not high enough to cover the cost. By challenging the reason for the rejection, the latter two strategies gain this meta-information which the agents constructively use in their subsequent interactions. For instance, if the counterpart rejected the proposal due to lack of capability, it can be excluded in future if the agent requires a capability which is equal or greater. In this way such reasons give useful meta-information to the agents for their future negotiations. So these strategies allow the agents to exploit such information and interact more efficiently as a society. Arguing in the first rejection provides this information earlier in the negotiation, which, in turn, gives the agent more capacity to exploit such information (even in the present negotiation) than getting it in the last encounter. Given this, we can conclude that in our context allowing the agents to challenge non-performance earlier in the negotiation allows them to manage their social influences more efficiently as whole.

5   Related Work

As highlighted in Section 1, to function as a coherent society, agents operating within a multi-agent society need the ability to detect, manage, and resolve conflicts in a systematic manner. Here, we will compare our ABN approach with two others suggested in the multi-agent literature. First, we note the work of [5] on electronic institutions where commitments of agents resulting due to social influences are managed through a performative structure. In more detail, they use a central authority to ensure that such commitments are upheld by controlling the type of locutions agents can issue in certain contexts based on the state of their commitments. In a similar vein, [6] provides a mechanism to control, verify, and manipulate commitments through the use of a state

machine. Now, one of the key distinctions of our approach from these is the absence of a central authority. Ours is a decentralised model where agents detect, manage and resolve conflicts about their social influences by arguing between each other. Another key feature in our method is its ability to function under incomplete knowledge. On the other hand, both the above approaches assume complete information within the central entity.

Our ABN framework also extends current ABN research by allowing the agents to argue, negotiate and manage conflicts in a multi-agent society. When compared against the model of Kraus *et al* [10] our framework has two distinct advantages. First, ours expressly takes into account the impact of society by way of social commitments, whereas their main focus is in formulating interactions between two agents. Second, they do not take into account the impact of incomplete information. In contrast, our social arguments captured in Section 2.2 allow agents to argue about their social influences and overcome such conflicts within a society. The work of Sierra *et al*. [15] is an important initial attempt to extend the work of [10] to a social context. Similar to our approach (and unlike [10]) they allow agents to argue in social contexts with imperfect information. However, they only consider authority based relationships, which we believe only capture a specialised form of social context (i.e., institutions or formal organisations). Our work, on the other hand, presents a more generic way of capturing social influences of roles and relationships (i.e., using social commitment with different degrees of influence), thus allowing agents' the ability to argue, negotiate and resolve conflicts under disparate social influences.

6   Conclusions and Future Work

The incomplete knowledge and the diverse conflicting influences present within a multi-agent society may prevent agents from abiding by all their social influences. In such situations, in order to function as a coherent society, agents require a mechanism to manage their social influences in a systematic manner. To this end, this paper develops a novel ABN approach that allows agents to argue, negotiate and, thereby, achieve a consensus, about their social influences. Furthermore, in order to assess the performance benefits of our proposed method, we carry out an empirical analysis by implementing such an ABN approach in a multi-agent task allocation context. Our results can be summarised as three main points. *First*, our method is shown to be both a more efficient and a more effective strategy in managing social influence even at high uncertainty levels when compared to a non-arguing approach. *Second*, we show that our approach can be further enhanced in terms of efficiency by allowing agents to challenge one another earlier in the negotiation encounter and using the meta-information that is gained to guide future negotiation encounters. *Third*, we show that both under complete uncertainty and when there are abundant social influences available in the society, the effectiveness of our approach is not significantly different from a non-arguing one.

In the future, we aim to expand our approach by allowing the agents to explicitly trade social influences in the form of threats and promises (as per Section 2.2) and examine the effect of so doing. At the moment agents only implicitly use these social influences to endorse their proposals. In such a system, we also plan to experiment with the effect of using different argument selection strategies in order to identify if certain strategies allow the agents to argue more efficiently or effectively than others.

7 Acknowledgements

References

1. L. Amgoud and H. Prade. Reaching agreement through argumentation: A possibilistic approach. In D. Dubois, C. A. Welty, and M.-A. Williams, editors, *Proc. of the Ninth International Conference (KR2004)*, pages 175–182, Canada, 2004.

2. K. Atkinson, T. Bench-Capon, and P. McBurney. A dialogue game protocol for multi-agent argument over proposals for action. In *Argumentation in Multi-Agent Systems*, LNAI 3366, pages 149–161, NY, USA, 2004.

3. C. Castelfranchi. Commitments: From individual intentions to groups and organizations. In *Proc. of the first Int. Conf. on Multi-agent Systems (ICMAS'95)*, pages 41–48, San Francisco, CA, 1995.

4. L. Cavedon and L. Sonenberg. On social commitment, roles and preferred goals. In *Proc. of the third Int. Conf. on Multi-Agent Systems (ICMAS'98)*, pages 80–86, 1998.

5. M. Esteva, J. A. Rodríguez, C. Sierra, P. Garcia, and J. L. Arcos. On the formal specifications of electronic institutions. *LNAI*, 1991:126–147, 2001.

6. N. Fornara. *Interaction and Communication among Autonomous Agents in Multiagent Systems*. Phd thesis, Universit della Svizzera italiana, Facolt di Scienze della Comunicazione, 2003.

7. N. C. Karunatillake and N. R. Jennings. Is it worth arguing? In *Argumentation in Multi-Agent Systems*, LNAI 3366, pages 234–250, NY, USA, 2004.

8. N. C. Karunatillake, N. R. Jennings, I. Rahwan, and T. J. Norman. Arguing and negotiating in the presence of social influences. In *Proc. of the fourth Int. Central and Eastern European Conference on Multi-Agent Systems (CEEMAS'05)*, LNAI 3690, pages 223–235, Budapest, Hungary, 2005.

9. N. C. Karunatillake, N. R. Jennings, I. Rahwan, and T. J. Norman. Argument-based negotiation in a social context. In *Proc. of the second Int. Workshop on Argumentation in Multi-Agent Systems (ArgMAS'05)*, pages 74–88, Utrecht, The Netherlands, 2005.

10. S. Kraus, K. Sycara, and A. Evenchik. Reaching agreements through argumentation. *Artificial Intelligence*, 104(1-2):1–69, 1998.

11. P. McBurney, R. van Eijk, S. Parsons, and L. Amgoud. A dialogue-game protocol for agent purchase negotiations. *Autonomous Agents and Multi-Agent Systems*, 7(3):235–273, 2003.

12. I. Rahwan, S. D. Ramchurn, N. R. Jennings, P. McBurney, S. Parsons, and L. Sonenberg. Argumentation-based negotiation. *The Knowledge Engineering Review*, 18(4):343–375, 2003.

13. S. D. Ramchurn, N. R. Jennings, and C. Sierra. Persuasive negotiation for autonomous agents: A rhetorical approach. In *Computational Models of Natural Argument, IJCAI*, pages 9–18, 2003.

14. T. W. Sandholm and V. R. Lesser. Advantages of a leveled commitment contracting protocol. In *Proc. of the 13th Conference on Artificial Intelligence (AAAI'96)*, pages 126–133, OR, USA, 1996.

15. C. Sierra, N. R. Jennings, P. Noriega, and S. Parsons. A framework for argumentation-based negotiation. In *Proc. of fourth Int. Workshop on Agent Theories Architectures and Languages (ATAL'97)*, pages 167–182, 1998.