Arguing and Negotiating in the Presence of Social Influences*

Nishan C. Karunatillake[1], Nicholas R. Jennings[1], Iyad Rahwan[2], and Timothy J. Norman[3]

[1] School of Electronics and Computer Science, University of Southampton, Southampton, UK.
{nnc02r,nrj}@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
[3] Department of Computing Science, University of Aberdeen, Aberdeen, UK.
tnorman@csd.abdn.ac.uk

Abstract. When agents operate in a society with incomplete information and with diverse and conflicting influences, they may, in certain instances, lack the knowledge, the motivation and/or the capacity to enact all their commitments. However, to function as a coherent society it is important for these agents to have a means to resolve such conflicts and to come to a mutual understanding about their actions. To this end, *argumentation-based negotiation* provides agents with an effective means to resolve conflicts within a multi-agent society. However, to engage in such argumentative encounters, agents require four fundamental capabilities; a schema to reason in a social context, a mechanism to identify a suitable set of arguments, a language and a protocol to exchange these arguments, and a decision making functionality to generate such dialogues. This paper presents formulations of all of these capabilities and proposes a coherent framework that allows agents to argue, negotiate, and, thereby, resolve conflicts within a multi-agent society.

Key words: Argumentation-based Negotiation, Conflict Resolution.

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 a social context, their actions 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. Here, we categorise these latter forms as *social influences*. Now, in many cases, both these forms of influence may be present and they may give conflicting motivations to the individual agent. For instance, an agent may be internally motivated to perform a specific action, whereas, at the same time, it may also be subject to an external social influence 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). Moreover, in many cases, agents have to carry out their actions in environments with incomplete information. Thus, for instance, they may not be aware of the existence of all the social influences that could or indeed should affect their actions and they may also lack the knowledge of certain specific internal influences that drive other agents' behaviours. Therefore, when agents operate in a society of incomplete information 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 and to come to a mutual understanding about their actions. To this end,

* The first author is a full time PhD student funded by EPSRC under the project Information Exchange (GR/S03706/01). The authors also extend their gratitude to Pietro Panzarasa, Chris Reed, and Xudong Luo for their thoughts, contributions, and discussions.

Argumentation-Based Negotiation (ABN) has been advocated as a promising means of resolving conflicts within such agent societies [1, 2]. 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.

Now, one of the central features required by an agent to engage in such arguments within a society is the ability to generate valid arguments during the course of the dialogue. We believe this demands four fundamental capabilities: (i) a schema to reason in social settings; (ii) a mechanism to identify a suitable set of arguments; (iii) a language and a protocol to exchange these arguments; and (iv) a decision making functionality to generate such dialogues. This paper builds upon our previous conceptual grounding [3] and formulates a coherent framework that addresses all four of these issues. More specifically, apart from formulating a coherent schema that captures social influence in multi-agent systems (see Section 2.1) and systematically using it, in turn, to identify social arguments to resolve conflicts within an agent society (see Section 2.2), this paper presents three additional contributions. *First*, we construct a language that is capable of expressing such social arguments and, which allows agents to exchange them within their argumentative dialogues (see Section 3.1). *Second*, we define a dialogue game protocol identifying the different guidelines (such as locution rules, structural rules and commitment rules) which will govern these dialogues and guide its participants toward resolving their conflicts. *Finally*, we define the different decision making algorithms required by the agents to engage in such argumentative dialogues to resolve conflicts about their social influences (see Section 3.2).

2 Model for Arguing with Social Influences

Here we outline our ABN model that provides agents with a means to argue, negotiate, and, thereby, resolve their conflicts in relation to social influences. We introduce our model in two stages; first detailing how social influences within a society can be captured into a schema, and second explaining the different ways that agents can use this schema to systematically capture arguments to use within their ABN in a multi-agent community.

2.1 Capturing Social Influence

The notion of *social commitment* acts as our basic building block for capturing social influences. First introduced by Castelfranchi [4], it is 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.[1] Having defined such, Castelfranchi further explains the consequences of a social commitment for both the agents involved. In detail, a social commitment results in the debtor attaining 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 make good any losses suffered due to its non-performance. We refer to these rights the creditor gains as the *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. Thus, when a certain agent is socially committed to another to

[1] In the desire to maintain simplicity within our schema, we avoid incorporating the witness (see [4]) in our model (as Castelfranchi did in his subsequent expositions).

perform a specific action, it subjects itself to the social influences of the other to perform that action. The ensuing obligation, on one hand, allows us to capture how an agent gets subjected to the social influence of another, whereas, the rights to exert influence, on the other hand, model how an agent gains the ability to exert such social influence upon another. Thereby, the notion of social commitment gives an elegant mechanism to capture social influence resulting between two agents.

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. To illustrate this, consider the relationship between the roles supervisor and student. For instance, assume the relationship socially influences the student to produce and hand over his thesis to the supervisor in a timely manner. This influence we can perceive as a social commitment that exists between the roles supervisor and student (the student is socially committed to the supervisor to perform the stipulated action). As a consequence of this social commitment, the student attains an obligation toward the supervisor to carry out this related action. On the other hand, the supervisor gains the right to exert influence on the student by either demanding that he does so or through questioning his 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 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 [5]):

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

Definition 2: Let SC denote a finite set of social commitments and $\text{SC}_\theta^{x \rightarrow y} \in SC$. Thus, as per Castelfranchi, $\text{SC}_\theta^{x \rightarrow y}$ *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:*

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

where:

- $[\text{O}_\theta^{x \rightarrow y}]_x^f$ represents the obligation that x attains that subjects it to an influence of a degree f (refer to [3] for more details) toward y to perform θ and
- $[\text{R}_\theta^{y \rightarrow x}]_y$ 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.

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 $\boldsymbol{SC}_\theta^{r_i \rightarrow r_j}$ 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. Natural Language Representation of the Schema of Social Influence.

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: $a_i, a_j \in A \wedge r_i, r_j \in R \wedge p \in P$ 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:

$$\text{Act}(a_i, r_i) \wedge \text{RoleOf}(r_i, p) \rightarrow \text{In}(a_i, r_i, p) \quad (1)$$

$$\text{Act}(a_j, r_j) \wedge \text{RoleOf}(r_j, p) \rightarrow \text{In}(a_j, r_j, p). \quad (2)$$

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

$$(\text{DebtorOf}(r_i, \text{SC}) \wedge \text{RoleOf}(r_i, p)) \wedge (\text{CreditorOf}(r_j, \text{SC}) \wedge \text{RoleOf}(r_j, p)) \wedge \text{ActionOf}(\theta, \text{SC}) \rightarrow \text{AssocWith}(\text{SC}_\theta^{r_i \rightarrow r_j}, p). \quad (3)$$

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

$$\text{SC}_\theta^{r_i \rightarrow r_j} \rightarrow \left[\text{O}_\theta^{r_i \rightarrow r_j}\right]_{r_i}^{f} \wedge \left[\text{R}_\theta^{r_j \rightarrow r_i}\right]_{r_j}. \quad (4)$$

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

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

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

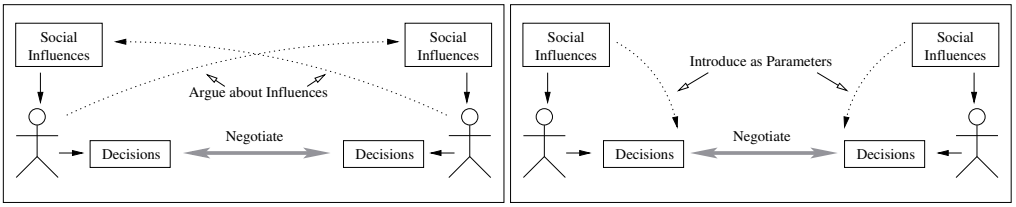
$$\text{In}(a_j, r_j, p) \wedge \text{AssocWith}(\text{SC}_\theta^{r_i \rightarrow r_j}, p) \rightarrow \left[\text{R}_\theta^{a_j \rightarrow r_i}\right]_{a_j}. \quad (6)$$

2.2 Capturing Social Arguments

Having captured the notion of social influence into a schema, here we present how agents can use it to systematically identify arguments to negotiate within a society. We term these arguments *social arguments*, not only to emphasise their ability to resolve conflicts within a society, but also to highlight the fact that they use the social influence present within the system as a core means in changing decisions and outcomes within the society.[2] Specifically, we have identified two major ways in which social influence can be used to change decisions and outcomes and thereby resolve conflicts between agents (see Figure 2).

Socially Influencing Decisions: One way to affect an agent's decisions is by arguing about the validity of that agent's practical reasoning [6, 7]. 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 to perform) 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 (see Figure 2(a)). 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 to act) 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:

[2] Due to space restrictions here we present only a limited subset of social arguments. For a comprehensive list of arguments, together with their formal representation, refer to [3].



(a) Socially Influencing Decisions (b) Negotiating Social Influence

Fig. 2. Interplay of Social Influence and Argumentation-Based Negotiation.

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 a_i is also acting role r_i, point out SC is also a social commitment associated with the relationship p, point out 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*).

Negotiating Social Influence: Agents can also use social influences within their negotiations. More specifically, instead of using social argumentation as a tool to affect decisions (as above), agents can 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 [8] (see Figure 2(b)). 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.

3 The Language, Protocol, and Decision Making Functionality

As mentioned in Section 1, our main objective is to formulate a society of agents that are capable of resolving their conflicts through argumentation-based negotiations. To this end, Section 2 formulated a model that allows the agents to identify such arguments to resolve conflicts in a social context. However, identifying such arguments is merely the first step. Agents also require a means to express such arguments, a mechanism to govern their interactions and guide them to resolve their conflicts, and a functionality to make decisions during the course of such dialogues. To this end, we now present the language, the protocol, and the decision making algorithms of our ABN framework.

3.1 The Language

The language plays an important role in an ABN framework. It not only allows agents to express the content and construct their arguments, but also provides a means to communicate and exchange them within an argumentative dialogue. Highlighting these two distinct functionalities, we define the language in our framework at two levels; namely the *domain language* and the *communication language*. The former allows the agents to specify certain premises about their social context and also the conflicts that they may face while executing actions within such a

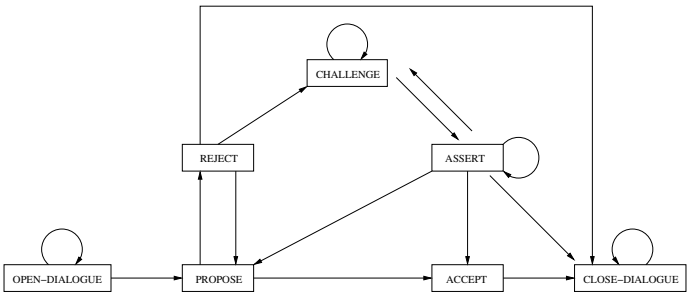


Fig. 3. Dialogue Interaction Diagram.

context. The latter, on the other hand, provides agents with a means to express these arguments and, thereby, engage in their discourse to resolve conflicts. Inspired by the works of Sierra *et al*. [9], this two tier definition not only allows us an elegant way of structuring the language, but also provides a means to easily reuse the communication component within a different context merely by replacing its domain counterpart.

In more detail, our domain language consists of ten elocutionary particles. Of these, eight allow the agents to describe their social context and these flow naturally from our social influence schema (i.e., Act, RoleOf, In, DebtorOf, CreditorOf, ActionOf, InfluenceOf, and AssocWith). Due to space restrictions we avoid repeating these definitions here (see Section 2.1). Furthermore, we define two additional predicates that provide a means to express the conflicts that the agents may face while executing their actions:

Definition 4: Let:

- do: $A \times \Theta$ denote the fact that an agent is performing an action (expressed in the abbreviated form $do(\theta)$ when the agent is unambiguous).
- Conflict: $do(A \times \Theta) \times do(A \times \Theta)$ denote the fact that performing the actions gives rise to a conflict.

On the other hand, our communication language consists of seven elocutionary particles (see Table 1). Mainly inspired form the works of Amgoud *et al*. [10], MacKenzie's system DC [11], and McBurney *et al*. [12], these form the building blocks of our dialogue game protocol explained below (see Section 3.2). Furthermore, these collectively allow the agents to use both of our identified methods of conflict resolution; namely socially influencing decisions and negotiating social influences (see Section 2.2). Due to their integrated nature with our protocol, we will detail their operational functionality and the decision making algorithms associated with each of these locutions alongside the protocol (see Section 3.2).

3.2 The Protocol and the Decision Making Functionality

Given the language component of our ABN framework, we will now proceed to describe both the protocol which governs its interaction and guides the agents to resolve their conflicts, and the various decision making algorithms that would enable the individual agents to participate in such encounters.[3] While the overall structure of our protocol is inspired from the work on computational conflicts by Tessier *et al.* [13], the works on pragma-dialectics proposed by van Eemeren and Grootendorst [14], and that on dialogue games conducted by McBurney *et al*. [12], and Amgoud *et al*. [10] contributed greatly in defining its operational guidelines.

More specifically, our protocol consists of 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

[3] Even though we acknowledge the importance of distinguishing the rules of encounter governed by the protocol from the individual decision mechanisms required by the participants to engage in such dialogues (see [12]), due to space restrictions we choose to describe both these elements in this section.

Table 1. The Protocol

Locution	Effects on CS & IS	Next Valid Moves
OPEN-DIALOGUE	$CS(a_i) \leftarrow$ *OPEN-DIALOGUE* $CS(a_j) \leftarrow$ *OPEN-DIALOGUE*	*OPEN-DIALOGUE* *PROPOSE*(l, m)
PROPOSE(l, m)	$CS(a_i) \leftarrow$ *PROPOSE*(l, m) $CS(a_j) \leftarrow$ *PROPOSE*(l, m) $IS(a_j) \leftarrow$ *Need*$(a_i, l) \wedge$ *Capable*(a_i, m)	*ACCEPT*(l, m) *REJECT*(l, m)
ACCEPT(l, m)	$CS(a_i) \leftarrow$ *ACCEPT*$(l, m) \wedge l \wedge m$ $CS(a_j) \leftarrow$ *ACCEPT*$(l, m) \wedge l \wedge m$ $IS(a_i) \leftarrow$ *Capable*(a_j, l)	*CLOSE-DIALOGUE*
REJECT(l, m)	$CS(a_i) \leftarrow$ *REJECT*(l, m) $CS(a_j) \leftarrow$ *REJECT*(l, m)	*CHALLENGE*(l) *PROPOSE*(l, m') *CLOSE-DIALOGUE*
CHALLENGE(l)	$CS(a_i) \leftarrow$ *CHALLENGE*(l) $CS(a_j) \leftarrow$ *CHALLENGE*(l)	*ASSERT*(l) *CHALLENGE*(l)
ASSERT(l)	$CS(a_i) \leftarrow$ *ASSERT*(l) $CS(a_j) \leftarrow$ *ASSERT*(l)	*PROPOSE*(l, m') *ACCEPT*(l, m) *ASSERT*$(\neg l)$ *CHALLENGE*(l) *CLOSE-DIALOGUE*
CLOSE-DIALOGUE	$CS(a_i) \leftarrow$ *CLOSE-DIALOGUE* $CS(a_j) \leftarrow$ *CLOSE-DIALOGUE*	*CLOSE-DIALOGUE*

in the dialogue, the former indicating its commencement and the latter its termination [12]. The four remaining stages not only adhere to the computational conflict work by Tessier *et al.*, but also comply well with the pragma-dialectics model for critical discussion proposed by van Eemeren and Grootendorst. In more detail, in 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 decide on how to address it (i.e., whether to avoid the conflict or attempt to manage and resolve it through argumentation and negotiation [1]). 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. As mentioned above, these four stages map seamlessly to the four stages in the pragma-dialectics model; namely *confrontation*, rather infelicitously termed *opening*, *argumentation*, and *concluding* respectively.

Given the overall stages of our protocol, we now describe its internal operation. Our protocol follows the tradition of dialogue games [12] where a dialogue is perceived as a game in which each participant make moves (termed dialogue moves) to win or tilt the favour of the game toward itself. In such a context, the protocol defines the different rules for the game such as locutions rules (indicating the moves that are permitted), commitment rules (defining the commitments each participant incurs with each move), and structural rules (that define the types of moves available following the previous move).[4] To this end, Figure 3 depicts the overall structure of our protocol and Table 1 details the different commitment rules and the valid locutions that may follow each move. For ease of reference, here we address the proposing agent as a_i and its responding counterpart as a_j. The commitment rules are shown as effects on the participants' commitment (CS) and information (IS) stores (see [10]) and l and m are propositions constructed in the domain language defined above. The following describes their operation in more detail.

OPEN-DIALOGUE: This indicates the entry point of that agent to the dialogue. As shown in Table 1 this would result in an entry in either agents' commitment stores corresponding to the dialogical commitment [15] of having made the move (i.e., commitment to the fact that a_i has uttered *OPEN-DIALOGUE*). An agent receiving an *OPEN-DIALOGUE* will retort back (if it hasn't already initiated it) by uttering the same. This would put both these agents in the opening stage and their negotiation over actions can commence. For simplicity, we assume that the first

[4] Note, this is not intended to be an exhaustive list of rules, but rather the most important ones in our context. For instance, if the aim of the dialogue governed by the protocol is persuasion, the win-loss rules specifying what counts as a winning or losing position would become a vital component.

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: *PROPOSE*$(do(a_j, \theta_j), do(a_i, \theta_i))$
3: **end if**

Algorithm 2 Decision making algorithm for *ACCEPT* and *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: *ACCEPT*$(do(a_j, \theta_j), do(a_i, \theta_i))$
3: **else**
4: *REJECT*$(do(a_j, \theta_j), do(a_i, \theta_i))$
5: **end if**

agent opening the dialogue is the one attempting to make its counterpart perform (or abstain from performing) an action.

PROPOSE: Each such proposal is composed of two basic elements; the action θ_j that a_i requires a_j to perform and the action θ_i that a_i is willing to perform in return. Thus, in general, a proposal will have the form *PROPOSE*$(do(a_j, \theta_j), do(a_i, \theta_i))$. Here, θ_i could be single atomic action (e.g., I will perform (or will not perform) a certain action in return or I will make a payment of a certain amount) or a composite action (e.g., I will perform action $(\theta_1$ *and* $\theta_2)$ or $(\theta_1$ *or* $\theta_2)$). Therefore, this generic form of proposal allows the agents not only to make simple offers of payment over actions, but also to make simple or composite rewards and/or threats over actions. In this manner, it allows the agents to negotiate and also to use social influences as parameters within their negotiations to resolve conflicts (see Section 2.2). Given this, Algorithm 1 highlights the decision making required to generate such a proposal. In more detail, we assume our agents to be self-interested, thus, the proposals that they generate need to be viable on their behalf (i.e., the cost for a_i in performing the proposed action θ_i (i.e., $C^{a_i}_{do(a_i,\theta_i)}$) should not exceed the benefit it gains from a_j performing the requested action θ_j (i.e., $B^{a_i}_{do(a_j,\theta_j)}$). We also assume our agents do not intentionally attempt to deceive each other with offers that they do not believe feasible on their behalf. Therefore, they will only generate proposals that they believe to have the capability to honour.[5] Once received, as an effect of the proposal, a_j will gain the information that a_i requires θ_j and that a_i has the ability to perform θ_i (see Table 1).

ACCEPT* and *REJECT: Upon receiving a proposal, the agent a_j may choose to either accept or reject it. Now, in order to make this decision, it will need to evaluate the proposal. Similar to above, this evaluation is also based on two factors: a_j needs to have the capability to perform the requested action and the benefit of the proposal should outweigh the cost of performing the suggested action (see Algorithm 2). If both these conditions are satisfied the agent will accept the proposal, otherwise it will reject it. If accepted, both agents will incur commitments to perform their respective actions (see Table 1).

CHALLENGE: Upon rejection of a proposal by its counterpart (a_j), a_i may choose to either forward a modified proposal (i.e., if the reason is apparent such that there can be only one possibility) or challenge a_j's decision in order to identify the underlying reasons for rejection. Apart from this, an agent may use *CHALLENGE* in two other situations (see Figure 3). First, an agent may challenge another's right to challenge (demand or question) its decision (see Section 2.1) if that right is not evident for the agent. This allows an agent to only justify its decisions to others who have the right to challenge its decision. To avoid infinite deepening of challenges, we do not allow such challenges go beyond two levels (i.e., challenge another's right to challenge its own right to challenge). Second, an agent can challenge a certain assertion

[5] First, under these assumptions of self-interest and non-deceit, we believe, viability and feasibility are the two most important factors to consider. Second, even though we choose to specify the algorithms at an abstract level that is independent of any domain, by defining how the agents evaluate these costs and benefits we can easily set these to reflect a given domain. Finally, even though the *PROPOSE* locution defined above has both the elements request and reward explicitly present, either can be null. This allows the agents to express proposals that are mere requests without an explicit reward (such as demands, pleads, and orders) and solitary rewards (such as offers, gifts, and suggestions) that they deem to be viable during their negotiation.

Algorithm 3 Decision algorithm for *CHALLENGE*.

– In case of *REJECT*(l)
  1: **if** (*REJECT*(l) $\in \Delta^{a_i} \wedge$*reason*(*REJECT*(l)) $\notin \Delta^{a_i}$) **then**
  2: *CHALLENGE*(REJECT(l))
  3: **end if**
– In case of *ASSERT*(l)
  1: **if** (*ASSERT*(l) $\in \Delta^{a_i} \wedge$*reason*(*ASSERT*(l)) $\notin \Delta^{a_i}$) **then**
  2: *CHALLENGE*(l)
  3: **end if**
– In case of *CHALLENGE*(l)
  1: **if** (*CHALLENGE*(l) $\in \Delta^{a_i} \wedge R_{\textit{CHALLENGE}(l)} \notin \Delta^{a_i}$) **then**
  2: *CHALLENGE*(CHALLENGE(l))
  3: **end if**

Algorithm 4 Decision algorithm for *ASSERT*.

– In case of *ASSERT*($\neg l$)
  1: **if** ($\neg l \notin \Delta^{a_i} \wedge l \in \Delta^{a_i}$) **then**
  2: *ASSERT*(l)
  3: **end if**
– In case of *CHALLENGE*(l)
  1: **if** (*search-Justification*(l, Δ^{a_i}) $\Rightarrow H$) **then**
  2: *ASSERT*(H)
  3: **end if**

by its counterpart if either that assertion or its contradiction is not within its knowledge (see Algorithm 3 where Δ^{a_i} denotes agent a_i's knowledge-base).

ASSERT: An agent can assert some fact in two possible situations. First, if the agent is challenged for some justification on its decision it can assert that justification. Second, if its counterpart has made an assertion (l), but the agent has justification to believe its contradiction $(\neg l)$, then the agent can assert this to dispute its partner's assertion.[6] This will allow agents to undercut and rebut each others' social reasoning, and, thereby, resolve conflicts (see Section 2.2). Assert can either result in the counterpart generating an alternative proposal (taking into account the reason given) or accepting the proposal (convinced by the persuasion).

CLOSE-DIALOGUE: When either the counterpart has accepted a certain proposal or the proposing agent has no other feasible and worthwhile proposals to forward, an agent will utter *CLOSE-DIALOGUE* (echoed in return by its counterpart) to bring the dialogue to an end.

Having formulated the language, the protocol, and the decision making functions of our ABN system, we now explain how these would interact to provide a means for the agents to resolve their conflicts in a social context. To this end, Figure 4 depicts an illustrative dialogue taking place between Andy, an agent acting the role of a PhD student, and Ben, acting as his supervisor. The case is set within a context where Andy has two distinct obligations, both toward Ben; to finish his thesis θ_1 and to write a journal paper θ_2. However, due to time restrictions, we assume that Andy has decided to perform θ_1 at the expense of θ_2. This choice is in conflict with Ben's own motivations. In this context, Figure 4 illustrates how he can socially influence (see Section 2.2) Andy's decision by undercutting his justification and, thereby, resolve the conflict. More specifically, Figure 4 highlights two specific aspects of our language and protocol. First, it shows how the language component allows the agents to do a straightforward encoding of the natural language locutions into its respective utterances (see locutions L1 to L8 with its corresponding utterances M3 to M10). Second, it also depicts how the dialogue progresses through the six distinct stages of conflict resolution identified above.

4 Conclusions and Future Work

The long term objective of our work is to formulate an agent society that can use argumentative dialogues to resolve their conflicts. To this end, this paper builds upon our previous conceptual grounding on social arguments [3] and formulates a coherent argumentation framework that allows agents to use ABN to resolve conflicts in a multi-agent community. In more detail, we first define a schema that captures social influences in an agent society and then illustrates the different ways that agents can use it to systematically identifying a suitable set of arguments to resolve conflicts in such a social context. Next, we formulate the language, which allows agents

[6] Our current implementation uses a simple arbitration heuristic to resolve such disputes. However, this can be extended by replacing it with either a system based on the strength of justification [10] or a learning heuristic based on commitment (see Section 4).

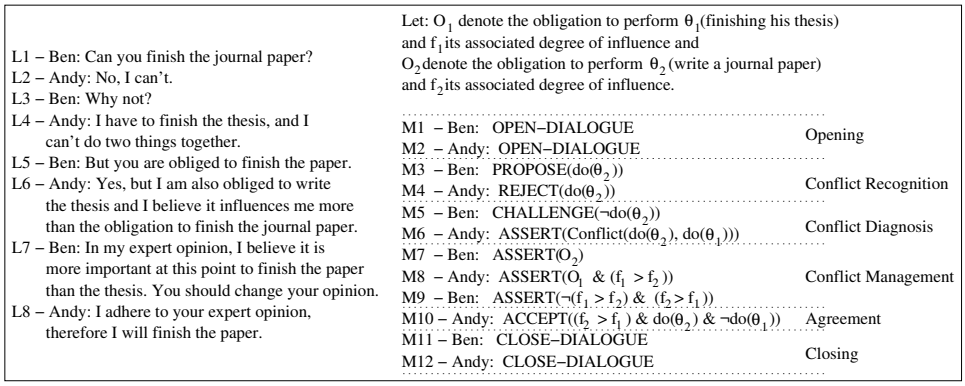


Fig. 4. Resolving Conflicts through Argumentation-based Negotiation.

to construct and express such arguments, and the protocol that would guide the course of the dialogue toward resolving conflicts. Finally, we define the various decision making algorithms that would enable the individual agents to participate in such argumentative encounters. Apart from the models specified in this paper, in our current work we have implemented these in a multi-agent task allocation domain (specified in [1]) in order to empirically test the efficiency and effectiveness of these concepts. In future, we aim to expand upon our current implementation by designing different argument selection strategies, thus, allowing the agents to adopt different tactics in resolving conflicts in an agent society.

References

1. Karunatillake, N.C., Jennings, N.R.: Is it worth arguing? In: Argumentation in Multi-Agent Systems (Proc. of ArgMAS 2004). LNAI 3366, NY, USA, Springer-Verlag (2004) 234–250
2. Rahwan, I., Ramchurn, S.D., Jennings, N.R., McBurney, P., Parsons, S., Sonenberg, L.: Argumentation-based negotiation. The Knowledge Engineering Review **18** (2003) 343–375
3. Karunatillake, N.C., Jennings, N.R., Rahwan, I., Norman, T.J.: Argument-based negotiation in a social context. In: Proc. of the 2^{nd} Int. Workshop on Argumentation in Multi-Agent Systems (ArgMAS'05), Utrecht, The Netherlands (2005) to appear
4. Castelfranchi, C.: Commitments: From individual intentions to groups and organizations. In: Proc. of the 1^{st} Int. Conf. on Multi-agent Systems (ICMAS'95), San Francisco, CA (1995) 41–48
5. Cavedon, L., Sonenberg, L.: On social commitment, roles and preferred goals. In: Proc. of the 3^{rd} Int. Conf. on Multi-Agent Systems (ICMAS'98). (1998) 80–86
6. Atkinson, K., Bench-Capon, T., McBurney, P.: A dialogue game protocol for multi-agent argument over proposals for action. In: Argumentation in Multi-Agent Systems (Proc. of ArgMAS 2004). LNAI 3366, NY, USA, Springer-Verlag (2004) 149–161
7. Walton, D.N.: Argumentation Schemes for Presumptive Reasoning. Erlbaum, Mahwah, NJ (1996)
8. Faratin, P., Sierra, C., Jennings, N.R.: Using similarity criteria to make trade-offs in automated negotiations. Artificial Intelligence **142** (2002) 205–237
9. Sierra, C., Jennings, N.R., Noriega, P., Parsons, S.: A framework for argumentation-based negotiation. In: Proc. of 4^{th} Int. Workshop on Agent Theories Architectures and Languages (ATAL'97), Rhode Island , USA (1998) 167–182
10. Amgoud, L., Parson, S., Maudet, N.: Argument, dialogue and negotiation. In Horn, W., ed.: Proc. of the 14^{th} European Conference on Artificial Intelligence (ECAI'00), Berlin (2000) 338–342
11. MacKenzie, J.: Question-begging in non-cumulative systems. Journal of philosophical logic **8** (1979) 117–133
12. McBurney, P., van Eijk, R., Parsons, S., Amgoud, L.: A dialogue-game protocol for agent purchase negotiations. Autonomous Agents and Multi-Agent Systems **7** (2003) 235–273
13. Tessier, C., Chaudron, L., Müller, H.J., eds.: Agents' Conflicts: New Issues. In: Conflicting Agents Conflict Management in Multi-Agent Systems. Kluwer Academic Publishers (2000) 1–30
14. Eemeren, F.H. van, Grootendorst, R.: Argumentation, Communication, and Fallacies. Lawrence Erlbaum Associates, Inc, Hillsdale NJ (1992)
15. Walton, D.N., Krabbe, E.C.W.: Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning. State Univ. of NY (1995)