Social and Persuasive Argumentation over Organized Actions

Maxime Morge[1]

Abstract. To greater adoption of argumentation technologies, their links with other disciplines need attention. In particular, Sociology provides a pertinent and well-grounded background for analysing the social dimensions of multiagent organisations. In this paper, we explore the social science background which captures the notions of motivation and social power/relationship in order to provide a coordination mechanism for open complex multiagent systems. Moreover, we formalize here these notions and we apply to them a particular argumentation technology for allowing agents to negotiate. Agents argue for persuading each other to collaborate with the help of two different schemes: appeals to common goal and threats. Our framework is exemplified with a simple use case.

1 INTRODUCTION

In the past decade argumentation has become increasingly important in Artificial Intelligence. It has provided a fruitful way of approaching defeasible reasoning, decision support, dialogue, and negotiation [3]. Argumentation has been researched extensively over the last years in application domains such as law, medicine and e-democracy. [13] points out that the links between argumentation and other disciplines need attention to greater adoption of argumentation technologies. For instance, argumentation and social theory require prior theoretical development in order to develop intelligent computer systems to support collaborative work.

Sociology provides a pertinent and well-grounded background for analysing the social dimensions of multiagent organisations. In this paper, we explore the social science background which captures the notions of motivation and social power/relationship in order to provide a coordination mechanism for open complex multiagent systems. Moreover, we formalize here these notions and we apply to them the argumentation technology proposed in [12] for allowing agents to negotiate. Concretely, we formalize the *Sociology of Organized Action* (SOA) [7] through the notion of *Concrete Action Systems* (CAS). Agents argue over it for persuading each other to collaborate with the help of two different schemes: appeals to common goal and threats. Our framework is exemplified with a simple use case.

The paper is organised as follows. Section 2 presents and formalized the social theory that is the basis for our proposal, namely the SOA. The formalization of CAS focus on the major concepts of agent, resource, and goal as well as their relationships. Section 3 introduces a conceptual framework for analyzing the decision problem related to the confident behaviour of agents. Section 4 presents our computational Argumentation Framework (AF) for decision making. Section 5 outlines the social interaction amongst agents. This interaction is illustrated by two persuasion examples: the first one is using an appeal to common goal and the second one is using a threat. Section 6 discusses some related works. Section 7 concludes with some directions for future work.

2 FORMAL CAS

The *Sociology of Organized Action* [7] (SOA), also called *strategic analysis*, studies the interaction amongst agents within an organization with the help of Concrete Action Systems (CAS) which have been formalized in [14]. We simplify and extend here this formalization for our purpose.

In order to study the interaction amongst (human) agents durably engaged in an organization (e.g. a firm, a university, a political institution), the SOA defines *Concrete Action Systems* (CAS) as structured contexts of cooperation among agents constraining their autonomy with respect to the power relationships. The power relationships result from the mastering of one or several resources. Each actor controls some resources and needs some other resources in order to achieve their goals. Therefore, the resources are the media of the power relationships between agents. To summarize, a CAS is an analysis grid of organizations taking into account the *resources*, the *agents*, and their *goals* as well as their relationships: the agents *control* resources, the agents *need* of resources, the resources are *required* for the achievement of goals, the agents *select* goals, the goals *depends* on one another.

The concepts manipulated by a CAS can serve to study (collective) decision-making processes. For this purpose, the selection of a goal by an agent as well as the dependence of goals must be weighted. Priorities, possibly numerical weights, allow goals to be compared to one another in a rational and consistent way as envisaged by multi-criteria decision making [11]. In the same way, the agent's needs (respectively controls) of resources can be weighted. An utility function (respectively payment) allows resources to be compared to one another as envisaged by the game theory [16].

Figure 1 depicts our formalization of CAS with the help of an entity-relation schema. Each resource is mastered by zero or more agents who decide about its availability, and so influence the achievement of the goals of the agents who need it. Each agent masters (needs of) zero or more resources. The agent depending on a resource affects it an *utility* to measure the satisfaction from consumption this good. The agent controlling a resource determines the *payment*, i.e. his reward for allowing its access. The agent selecting a goal affects it a *priority* to measure its importance. Moreover, the dependence relationship over goals is also associated with a priority. Obviously, the relation of requirement can be weight by a measure of necessity.

[1] Universitá di Pisa, Italy, email: morge@di.unipi.it

Since this weight can be calculated with the utilities, the payments, and the priorities, we do not mention it in our framework.

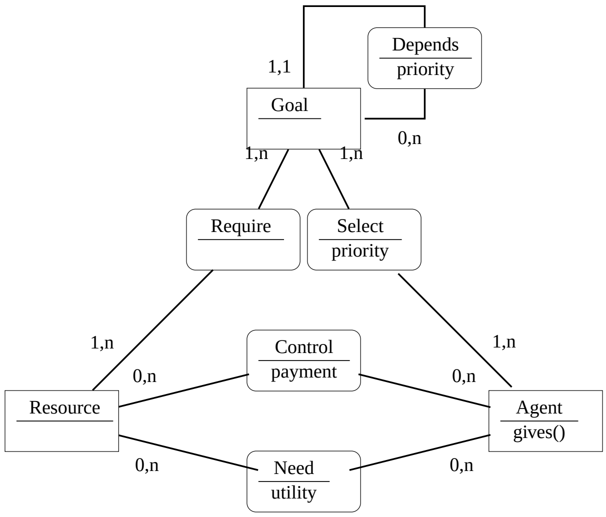

Figure 1. Formalization of a CAS

Using the convention of denoting constants in `typescript` and variables in *italics*, we illustrate the approach with the "hang a picture" [15] example, suitable adapted for illustrating persuasion. We consider here two agents `alice` and `bob`. They share the same goal which consists of seeing a picture hung in their living-room, `hung`. Each agent wants to hang the picture by oneself: `hang`(`alice`) or `hang`(`bob`). Obviously, the fact that the picture is hung depends on one of these two goals. Moreover, `alice` is mad at `bob` and she wants to hit him, `hit`(`alice`). `bob` wants to avoid it, ¬`hit`(`alice`). We consider here two resources: a `hammer` and a `nail`. The agent `alice` (respectively `bob`) controls the `hammer` (respectively the `nail`): `control`(`alice`, `hammer`) (respectively `control`(`bob`, `nail`)). The `hammer` and the `nail` are required for hanging the picture. The `hammer` is required for hitting someone. We can deduce that `alice` (respectively `bob`) needs of the `nail` (respectively the `hammer`). For this purpose, `alice` (respectively `bob`) can give the `hammer` (respectively the `nail`) to `bob` (respectively `alice`): `give`(`alice`, `bob`, `hammer`) (respectively `give`(`bob`, `alice`, `nail`)).

3 DECISION ANALYSIS

Our methodology is to decompose the decisionproblem into elements that can be analyzed and can be brought together to create an overall representation. We use here *influence diagrams* which are simple graphical representations of decision problems [5] including the decisions to make amongst the possible courses of action (called *decision nodes*, represented by squares), the value of the specific outcomes that could result (called *value nodes*, represented by rectangles with rounded corners), and the uncertain events which are relevant information for decision making (called *chance nodes*, represented by ovals). In order to show the relationship amongst these elements, nodes are put together in a graph connected by arrows, called *arcs*. We call a node at the beginning of an arc a *predecessor* and one at the end of an arc a *successor*. The nodes are connected by arcs where predecessors are independent and affect successors. Influence diagrams which are properly constructed have no cycles. In order to capture multi-criteria decision making, it is convenient to include additional nodes (called *abstract value nodes*, represented by double line) that aggregate results from predecessor nodes. While a *concrete value* is specified for every possible combination of decisions and events that feed into this node, an abstract value is specified for every possible combination of values that feed into this node, and so the multiple attributes are represented with a hierarchy of values where the top, abstract values aggregate the lower, concrete values. We assume that influence diagrams are provided by users via a GUI which allows them to communicate user-specific preferences.

We consider here the decision problem of an agent (cf Fig. 2). The fact that the picture is hung (`hung`) depends on its decision, `give`(ag_1, ag_2, res). This top main value is split into two concrete values, the fact that an agent ag_2 hangs the picture `hang`(ag_2) and the fact that an agent ag_2 hits the other one or not (`hit`(ag_2) or ¬`hit`(ag_2)). The evaluation of these criteria depends on the agent knowledge, namely the information about the controls of the resources, `control`(ag_1, res). The agent also provides, through the GUI, her preferences and constraints. For instance, according to `alice`, her own goals (e.g. `hang`(`alice`)) have priority over `bob`'s goals (e.g. `hang`(`bob`)). According to `bob`, the priorities over these goals are similar.

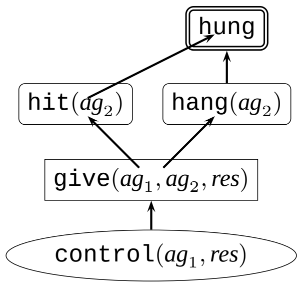

Figure 2. Influence diagram to structure the negotiation

4 ARGUMENTATION FRAMEWORK

According to the approach of defeasible argumentation of [8], arguments are reasons supporting claims which can be defeated[2] by other arguments.

Definition 1 (AF) *An* argumentation framework *is a pair* $AF = \langle \mathcal{A}, \texttt{defeats} \rangle$ *where* $\mathcal{A}$ *is a finite set of arguments and* `defeats` *is a binary relation over* $\mathcal{A}$*. We say that a set* S *of arguments defeats an argument* a *if* a *is defeated by at least one argument in* S*.*

[8] also analysis when a set of arguments is collectively justified.

Definition 2 (Semantics) *A set of arguments* $S \subseteq \mathcal{A}$ *is:*

- conflict-free *iff* $\forall a, b \in S$ *it is not the case that* a *defeats* b*;*
- admissible *iff* S *is conflict-free and* S *defeats every argument* a *such that* a *defeats some arguments in* S*.*
- S *is* preferred *iff* S *is maximally admissible;*

[2] The defeat relation is called attack in [8].

- S *is* complete *iff* S *is admissible and* S *contains all arguments* a *such that* S *attacks all attacks against* a*;*
- S *is* grounded *iff* S *is minimally complete;*

These declarative model-theoretic *semantics* of the AF capture various degrees of justification ranging from very permissive conditions, called *credulous*, to restrictive requirements, called *sceptical*. The semantics of an admissible (or preferred) set of arguments is credulous, in that it sanctions a set of arguments as acceptable if it can successfully dispute every arguments against it, without disputing itself. However, there might be several conflicting admissible sets. That is the reason why various sceptical semantics have been proposed for the AF, notably the grounded semantics and the sceptically preferred semantics, whereby an argument is accepted if it is a member of all maximally admissible sets of arguments. For simplicity, we restrict ourself to admissible semantics.

4.1 Decision framework

Since we want to instantiate our AF for our example, we need to specify a particular framework capturing the decision problem.

Definition 3 (Decision framework) *A* decision framework *is a tuple* $\mathcal{D} = \langle \mathcal{L}, \mathcal{A}sm, \mathcal{I}, \mathcal{T}, \mathcal{P} \rangle$*, where:*

- $\mathcal{L}$ *is the* object language *which captures the statements about the decision problem;*
- $\mathcal{A}sm$*, is a set of sentences in* $\mathcal{L}$ *which are taken for granted, called* assumptions*;*
- $\mathcal{I}$ *is the* incompatibility relation*, i.e. a binary relation over atomic formulas which is asymmetric. It captures the mutual exclusion between the statements;*
- $\mathcal{T}$ *is the* theory *which gathers the statements;*
- $\mathcal{P} \subseteq \mathcal{T} \times \mathcal{T}$ *is a (partial or total) preorder over* $\mathcal{T}$*, called the* priority *relation, which captures the uncertainty of beliefs, the priority amongst goals, and the expected utilities of the decisions.*

In the object language $\mathcal{L}$, we distinguish six disjoint components:

- a set of *abstract goals* (resp. *concrete goals*), i.e. some propositional symbols which capture the abstract values (resp. concrete values) that could result;
- a set of *decisions*, i.e. some predicate symbols which capture the decision nodes;
- a set of *alternatives*, i.e. some constants symbols which capture the mutually exclusive actions for each decision;
- a set of *beliefs*, i.e. some predicate symbols which capture the chance nodes;
- the *names* of rules in $\mathcal{T}$ which are unique.

In $\mathcal{L}$, we consider strong negation (classical negation) and weak negation (negation as failure). A strong literal is an atomic first-order formula, possible preceded by strong negation $\neg$. A weak literal is a literal of the form $\sim L$, where L is a strong literal.

We explicitly distinguish *assumable* (respectively *non-assumable*) literals which can (respectively cannot) be taken for granted, meaning that they can or cannot be assumed to hold as long as there is no evidence to the contrary. Decisions (e.g. $\mathtt{give}(ag_1, ag_2, res) \in \mathcal{A}sm$) as well as some beliefs (e.g. $\mathtt{control}(\mathtt{bob}, \mathtt{nail}) \in \mathcal{A}sm$) can be taken for granted. In this way, $\mathcal{D}$ can capture incomplete knowledge.

The *incompatibility relation* captures the conflicts. We have $L \,\mathcal{I}\, \neg L$, $\neg L \,\mathcal{I}\, L$, and $L \,\mathcal{I} \sim L$ but we do not have $\sim L \,\mathcal{I}\, L$. We say that two sets of sentences Φ_1 and Φ_2 are incompatible ($\Phi_1 \,\mathcal{I}\, \Phi_2$) iff there is at least one sentence ϕ_1 in Φ_1 and one sentence ϕ_2 in Φ_2 such as $\phi_1 \,\mathcal{I}\, \phi_2$.

A *theory* gathers the statements about the decision problem.

Definition 4 (Theory) *A* theory $\mathcal{T}$ *is an extended logic program, i.e a finite set of rules* R*:* $L_0 \leftarrow L_1, \ldots, L_j, \sim L_{j+1}, \ldots, \sim L_n$ *with* $n \geq 0$*, each* L_i *being a strong literal in* $\mathcal{L}$*. The literal* L_0*, called the* head *of the rule, is denoted* $\mathsf{head}(R)$*. The finite set* $\{L_1, \ldots, \sim L_n\}$*, called the* body *of the rule, is denoted* $\mathsf{body}(R)$*. The body of a rule can be empty. In this case, the rule, called a* fact*, is an unconditional statement.* R*, called the unique* name *of the rule, is an atomic formula of* $\mathcal{L}$*. All variables occurring in a rule are implicitly universally quantified over the whole rule. A rule with variables is a scheme standing for all its ground instances.*

For simplicity, we will assume that the names of rules are neither in the body nor in the head of the rules thus avoiding self-reference problems. Considering a decision problem, we distinguish:

- *goal rules* of the form R: $G_0 \leftarrow G_1, \ldots, G_n$ with $n > 0$. Each G_i is a goal literal in $\mathcal{L}$. The head of the rule is an abstract goal (or its strong negation). According to this rule, the abstract goal is promoted (or demoted) by the goal literals in the body;
- *epistemic rules* of the form R: $B_0 \leftarrow B_1, \ldots, B_n$ with $n \geq 0$. Each B_i is a belief literal of $\mathcal{L}$. According to this rule, B_0 is true if the conditions $B_1, \ldots, B_n$ are satisfied;
- *decision rules* of the form R: $G \leftarrow D(a), B_1, \ldots, B_n$ with $n \geq 0$. The head of the rule is a concrete goal (or its strong negation). The body includes a decision literal ($D(a) \in \mathcal{L}$) and a set of belief literals possibly empty. According to this rule, the concrete goal is promoted (or demoted) by the decision $D(a)$, provided that conditions $B_1, \ldots, B_n$ are satisfied.

Due to our representation of decision problems, we assume that the elements in the body of rules are independent, the decisions do not influence the beliefs, and the decisions have no side effects.

In order to evaluate the previous statements, all relevant pieces of information should be taken into account, such as the uncertainty of knowledge, the priority between goals, or the expected utilities of the decisions. In this work, we consider that all rules are potentially defeasible and that the priorities are extra-logical and domain-specific features. We consider that the *priority* $\mathcal{P}$ which is a reflexive and transitive relation considering possible *ex æquo*. $R_1 \mathcal{P} R_2$ can be read "R_1 has priority over R_2". $R_1 \not\mathcal{P} R_2$ can be read "R_1 has no priority over R_2", either because R_1 and R_2 are *ex æquo* or because R_1 and R_2 are not comparable. The priority over concurrent rules depends on the nature of rules. Rules are *concurrent* if their heads are identical or incompatible. We define three priority relations:

- the priority over *goal rules* comes from the *preferences* overs goals. The priority of such rules corresponds to the relative importance of the combination of (sub)goals in the body as far as reaching the goal in the head is concerned;
- the priority over *epistemic rules* comes from the *uncertainty* of knowledge. The prior the rule is, the more likely the rule holds;
- the priority over *decision rules* comes from the *expected utility* of decisions. The priority of such rules corresponds to the expectation of the conditional decision in promoting/demoting the goal literal.

In order to illustrate the previous notions, let us consider the goal rules, the decision rules, and the epistemic rules from alice's viewpoint which are represented in Tab. 1. According to the goal rules, the main goal is reached if: i) either alice hangs the picture and alice does not hit bob (cf r_{01}); ii) or bob hangs the picture and alice hits bob (cf r_{02}). alice prefers to hang the picture by herself, $r_{01}\mathcal{P}r_{02}$. According to the decision rules, the picture is hung by alice ($\mathtt{hang}(me)$) if she controls the hammer and bob gives the nail he controls (cf r_{11}). The picture is hung by bob ($\mathtt{hang}(you)$) if he controls the nail and alice give the hammer she controls (cf r_{12}). alice can hit bob if she controls the hammer (cf $r_{21}(ag)$). Otherwise, alice needs no resource (cf $f_{22}(ag)$). According to the epistemic rules, alice beliefs that she has the hammer, f_2.

$\mathcal{T}$

$$\uparrow \quad r_{01}: \mathtt{hung} \leftarrow \mathtt{hang}(me), \neg\mathtt{hit}(you)$$
$$r_{02}: \mathtt{hung} \leftarrow \mathtt{hang}(you), \mathtt{hit}(you)$$

$\mathcal{T}$

$$r_{11}: \mathtt{hang}(me) \leftarrow \mathtt{give}(you, me, \mathtt{nail}), \mathtt{control}(me, \mathtt{hammer}), \mathtt{control}(you, \mathtt{nail})$$
$$r_{12}: \mathtt{hang}(you) \leftarrow \mathtt{give}(me, you, hammer), \mathtt{control}(you, \mathtt{nail}), \mathtt{control}(me, \mathtt{hammer})$$
$$r_{21}(ag): \mathtt{hit}(ag) \leftarrow \mathtt{control}(ag, \mathtt{hammer})$$
$$f_{22}(ag): \neg\mathtt{hit}(ag) \leftarrow$$

$\mathcal{T}$

$$f_1 : \mathtt{control}(me, \mathtt{hammer}) \leftarrow$$

Table 1. The goal rules, the decision rules, and the epistemic rules.

4.2 Arguments

Since we want that our AF not only suggests some actions but also provides an intelligible explanation of them, we adopt here the tree-like structure for arguments proposed in [17] and we extend it with suppositions on the missing information.

Definition 5 (Argument) *An* argument *built upon $\mathcal{D}$ is composed by a conclusion, a top rule, some premises, some suppositions, and some sentences. These elements are abbreviated by the corresponding prefixes. An argument a can be:*

1. *a* hypothetical argument *built upon an unconditional ground statement. If L is an assumable literal (possibly its negation), then the argument built upon a ground instance of this assumable literal is defined as follows*[3]*: $conc(a) = L$, $top(a) = \theta$, $premise(a) = \emptyset$, $supp(a) = \{L\}$, $sent(a) = \{L\}$.*
   or
2. *a* built argument *built upon a rule such that all the literals in the body are the conclusion of arguments.*

   (a) *If f is a fact in $\mathcal{T}$ (i.e. $body(f) = \emptyset$), then the* trivial *argument a built upon this fact is defined as follows: $conc(a) = head(f)$, $top(a) = f$, $premise(a) = \emptyset$, $supp(a) = \emptyset$, $sent(a) = \{head(f)\}$.*

   (b) *If r is a rule in $\mathcal{T}$ with $body(r) = \{L_1, \ldots, L_j, \sim L_{j+1}, \ldots, \sim L_n\}$ and there is a collection of arguments $\{a_1, \ldots, a_n\}$ such that, for each strong literal $L_i \in body(r)$, $conc(a_i) = L_i$ with $i \leq j$ and for each weak literal $\sim L_i \in body(r)$, $conc(a_i) = \sim L_i$ with $i > j$, we define the* tree *argument a built upon the rule r and the set $\{a_1, \ldots, a_n\}$ of arguments as follows: $conc(a) = head(r)$, $top(a) = r$, $premise(a) = body(r)$, $supp(a) = \cup_{a' \in \{a_1,\ldots,a_n\}} supp(a')$, $sent(a) = \cup_{a' \in \{a_1,\ldots,a_n\}} sent(a') \cup body(r) \cup \{head(r)\}$. The set of arguments $\{a_1, \ldots, a_n\}$ are called the set of* subarguments *of a (denoted $sbarg(a)$).*

The set of arguments built upon $\mathcal{D}$ is denoted $\mathcal{A}(\mathcal{D})$.

[3] θ denotes that no literal is required.

Notice that the subarguments of a tree argument concluding the weak literals in the body of the top rule are hypothetical arguments. Indeed, the conclusion of an hypothetical argument could be a strong or a weak literal while the conclusion of a built argument is a strong literal. As in [17], we consider composite arguments, called *tree* arguments, and atomic arguments, called *trivial* arguments. Contrary to other definitions of arguments (set of assumptions, set of rules), our definition considers that the different premises can be challenged and can be supported by subarguments. In this way, arguments are intelligible explanations. Moreover, we consider *hypothetical* arguments which are built upon missing information or a decision. In this way, our framework allows to reason further by making suppositions related to the unknown beliefs and over possible decisions.

In our example, the argument b (respectively a), concludes that the main goal is reached since bob (respectively alice) hangs the picture and alice hits (respectively does not hit) bob if we suppose that alice (respectively bob) gives the hammer (respectively the nail) and if we suppose that bob controls the nail. The arguments a is depicted in Figure 3. An argument can be represented as a tree where the root is the conclusion (represented by a triangle) directly connected to the premises (represented by losanges) if they exist, and where leafs are either some suppositions (represented by circles) or the empty set. Each plain arrow corresponds to a rule (or a fact) where the head node corresponds to the head of the rule and the tall nodes are in the body of the rule. The tree argument a is composed of one trivial subargument and one tree argument. Neither trivial arguments nor hypothetical arguments contain subarguments.

4.3 Interactions

The interactions amongst arguments may come from their conflicts, from their nature (hypothetical or built), and from the priority of rules. We examine in turn these different sources of interaction.

Since their sentences are conflicting, the arguments interact with one another. For this purpose, we define the following attack relation.

Definition 6 (Attack relation) *Let $a, b \in \mathcal{A}(\mathcal{D})$ be two arguments. a* attacks *b iff $sent(a)\ \mathcal{I}\ sent(b)$.*

This relation encompasses both the direct (often called *rebuttal*) attack due to the incompatibility of the conclusions, and the indirect (often called *undermining*) attack, i.e. directed to a "subconclusion". According to this definition, if an argument attacks a subargument, the whole argument is attacked.

Since arguments are more or less hypothetical, we define the size of their suppositions.

Definition 7 (Supposition size) *Let $a \in \mathcal{A}(\mathcal{D})$ be an argument. The* size of suppositions *for a, denoted $suppsize(a)$, is the number of suppositions of a: $suppsize(a) = |supp(a)|$.*

The size of suppositions for an argument is the number of decision literals and assumable belief literals in the sentences of the argument.

Since arguments have different natures (hypothetical or built) and the top rules of built arguments are more or less strong, we define the strength relation as follows.

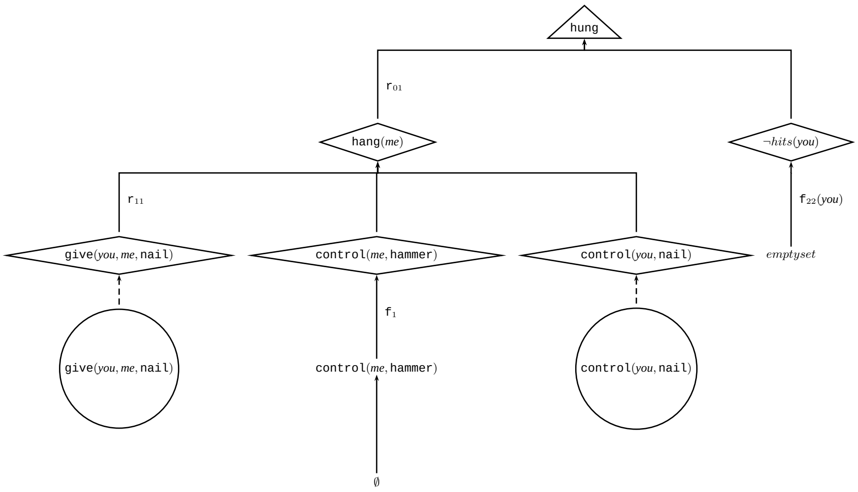

Figure 3. Arguments supporting the fact that bob give the nail

Definition 8 (Strength relation) *Let A_1 be a hypothetical argument, and A_2, A_3 be two built arguments.*

1. *A_2 is stronger than A_1 (denoted $A_2\ \mathcal{P}^{\mathcal{A}}\ A_1$);*
2. *If $(\mathtt{top}(A_2)\mathcal{P}\mathtt{top}(A_3)) \wedge \neg(\mathtt{top}(A_3)\mathcal{P}\mathtt{top}(A_2))$, then $A_2\ \mathcal{P}^{\mathcal{A}}\ A_3$;*
3. *If $(\mathtt{top}(A_2)\not\mathcal{P}\mathtt{top}(\mathrm{A}_3)) \wedge (\mathtt{suppsize}(\mathrm{A}_2) < \mathtt{suppsize}(\mathrm{A}_3))$, then $A_2\ \mathcal{P}^{\mathcal{A}}\ A_3$;*

Since $\mathcal{P}$ is a preorder on $\mathcal{T}$, $\mathcal{P}^{\mathcal{A}}$ is a preorder on $\mathcal{A}(\mathcal{T})$. Since it is preferable to consider fewer suppositions as possible, built arguments are preferred to hypothetical arguments. Moreover, we want to take into account the preferences captured by the priorities. That is the reason why we consider that an argument is stronger than another argument if the top rule of the first argument has a proper higher priority than the top rule of the second argument, or if it is not the case but the number of suppositions made in the first argument is properly smaller than the number of suppositions made in the second argument.

In order to adopt Dung's seminal calculus of opposition, we define the defeat relation.

Definition 9 (Defeat relation) *Let $\mathtt{a}, \mathtt{b} \in \mathcal{A}(\mathcal{D})$ be two arguments. $\mathtt{a}$ defeats $\mathtt{b}$ iff: i) a attacks b; ii) $\neg(\mathtt{b}\ \mathcal{P}^{\mathcal{A}}\ \mathtt{a})$.*

Let us consider our previous example. The arguments a and b attack each other. Since the top rules of a is r_{01} and the top rule of b is r_{02}, a is stronger than b, and so a defeats b. The argument a is in an admissible set, and so can justified the opinion of alice. This arguments is useful for alice to justify its choice in front of bob and to persuade the latter.

5 SOCIAL INTERACTION

The social statements are exchanged during dialogues and notified in the *dialogical commitments*. Our agent drives the interactions by the adherence to protocols.

The negotiation is driven according to the individual/social statements concerning the goals of agents (their own goals and the goals of their interlocutors), the decisions they make, the knowledge, and preferences over them. The social statements are exchanged during dialogues and notified in the *dialogical commitments* which are internal data structures which contain propositional/action social obligations involving the agent, namely with the agent being either the debtor or the creditor. The choice amongst actions is made according to the agent's statements and the preferences over them. The dialogical commitments of alice include commitments involving alice: either alice is the creditor of the commitment, or alice is the debitor of the commitment (see the next section).

A protocol is required to conduct the interaction. For this purpose, the social reasoning uses a boot strap mechanism that initiates the required protocol, the role the agent will play in that protocol, and the other participants. The protocol engine determines the appropriate message to be sent given those parameters. When there is a decision to be made either between the choice of two locutions (e.g. an accept or a reject) to be sent or the instantiation of the content of the locution (e.g. the definition of a proposal), the protocol engine uses a precondition mechanism to prompt the social reasoning. Upon the satisfaction of the precondition, the protocol engine sends the locution. A similar mechanism is used for incoming messages. If it is necessary to update the dialogical commitments of the agent, this can be done with the post condition mechanism which operates in a similar manner.

The agents utter messages to exchange goals, decisions, and knowledge. The syntax of messages is in conformance with a common communication language. We assume that each message: has an identifier, M_k; is uttered by a speaker (S_k); is addressed to a hearer (H_k); responds to a message with identifier R_k; is characterised by a speech act A_k composed of a locution and a content. The locution is one of the following: question, assert, accept, why, withdraw (see Table 2 below for examples). The content is a triple consisting of: a goal G_k, a decision D_k, and a knowledge K_k[4]

Figure 4 depicted our protocol from the initiator viewpoint with the help of a deterministic finite-state automaton. The choice of locutions to send depends on the way the social reasoning fulfills preconditions. For example, the outcome of evaluate decision by the social reasoning will dictate to the protocol engine whether it sends accept, assert or why. The corresponding rule of the protocol engine is as follows:

[4] We will use θ to denote that no goal is given and $\emptyset$ to denote that no knowledge is provided.

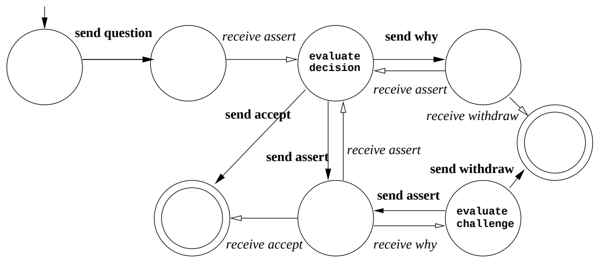

Figure 4. Protocol for the initiator

```
IF receive assert(G,D,K) from interlocutor THEN {
   update commit(interlocutor,[G,D,K]);
   IF evaluate(G,D,K) THEN{
         send accept(G,D,K) to interlocutor;
         commit(me,[G,D,K]);}
   ELSE IF evaluate(G,D2,K2) & D2!=D THEN{
                send assert(G,D2,K2) to interlocutor;
                commit(me,[G,D2,K2]);}
   ELSEIF send why(G,D,K) to interlocutor;}
```

In this rule me denotes the reasoning agent and interlocutor denotes the agent it dialogues with. evaluate(G,D,K) is a predicate which evaluates if the goal G is supported by an admissible argument built upon the decision D and the knowledge K. According to the corresponding rules, the dialogical commitments are updated when a proposal is received. If an admissible proposal have been suggested, then the speech act is an accept. If a new admissible proposal is found, then the speech act is an assert. Otherwise the speech act is a why.

Table 2 depicts the speech acts exchanged between alice and bob playing two different dialogues. They attempt to come to an agreement on the transaction for the resources to reach the common goal hung. The dialogue is initiated by alice. With the message M_1, bob informs alice that it finds out that the action give(alice, bob, hammer) is justified with respect to the common goal (hung). However, alice does not find give(alice, bob, hammer) justified and she proposes give(bob, alice, nail). Since none of these proposals have been jointly accepted, bob attempts to determine the reasons for alice's choice (cf M_3). In the first dialogue (Top of Table 2), alice argues with the goal hang(alice) which is a subgoal hung. In the second dialogue (Top of Table 2), alice argues with the goal hit(alice) demoted by bob. Given alice's response in M_4, bob includes the argument provided by alice. Therefore, it finds Alice's proposal justified whatever the dialogue is. Finally, bob communicates his agreement with the help of an accept (M_5) which closes the dialogue. We can notice that the influence of alice on bob leads this latter to concede the alternative give(bob, alice, nail) which was previously not justified from bob's viewpoint. The agents are able to persuade each other, through argumentation. In the first dialogue, bob includes a new argument concluding a common goal. In the second dialogue, bob includes a new argument concluding a goal he demotes. Whatever the dialogue is, alice persuades bob.

6 RELATED WORKS

We have presented here a model of social and persuasive argumentation. In this perspective, [2] identifies different types of arguments: threats, rewards, appeals to past promise, appeals to prevailing practice, appeals to self-interest, and counter-examples. In our paper, we focus on the threats and we introduce the appeals to common goal. Actually, the concepts of threat and promise are two faces of the same coin. Both of them are based upon the power/dependence relations of agents [9]. The purpose of [2] is to handle these types of arguments within a classical argumentation-based framework. Even if our argumentation-based framework is different, our definitions of arguments are quite similar. However, the strengths of these arguments are different. While the strength of arguments of [2] is based upon the weakest link principle, the strength of our arguments is based upon the last link principle and depends on the hypothetical nature of arguments. Contrary to [2], we adopt the calculus of opposition of Dung [8]. Actually, our work consists of a more refined framework for persuasive negotiation than [15] where the strength of arguments is not only based upon some authority relations.

We have used here the argumentation-based mechanism for decision making proposed in [12]. The framework of [10, 12] incorporates abduction on missing information, while the frameworks of [1, 12] can be applied to a multi-criteria decision making. To the best of our knowledge, the framework of [12] is the only one integrating both of these proprieties required by our application. Moreover, the other existing frameworks do not come with a conceptual framework for creating a model and a representation of decision problems. By relying on [12], the decision problem is firstly analyzed, and so treated.

[14] proposes straight translations of the sociology of organized action into a computer science formalism. We have simplified and extended here this formalism for our purpose. The relation of requirement is weight by a measure of necessity in [14]. Since this weight can be calculated with the utilities, the payments, and the priorities, we do not mention it in our framework. Contrary to [14], we do not distinguish the resources and the stake they represent, i.e. the power relationships amongst the agents which either control the stake or depend on its. Moreover, we have added a relation amongst goals to reflect that they can depend on one another. This extension enrich the coordination model.

According to [6], the four main recurring social dimensions of multiagent organisations are: the social structures of roles and groups, the dialogical structures of the interaction amongst agents, the functional structures of the goals and tasks, and the normative structures incorporating the deontic notions of obligation and permission. With respect to this analysis grid, our framework treats a subset of the social structures captured by the power of agents resulting from the mastering of resources. Our framework considers the functional structures due to the goal decompositions captured by the the dependence relation. Obviously, the proposed negotiation protocol consists of the dialogical structure of our framework. Finally the normative structure of multiagent systems is out of the scope of our framework.

7 CONCLUSIONS

In this paper, we have described a model of autonomous, social, and argumentative agent trying to persuade each other to collaborate. For this purpose we have formalized a social theory that is served to study collective decision-making processes with the help of the concepts of *agent*, *resource*, and *goal*, as well as their relationships. Moreover, we have provided an AF for decision-making to perform the social reasoning which is about how to achieve the individualistic, the social, and the common goals through collaboration. Actually, we have developed a model of internal dialectics between the individual goals and the social goals of agents to capture the interactions which can

M_k	S_k	H_k	A_k	R_k
M_0	alice	bob	question(hung, give(ag_1, ag_2, *res*), ∅)	θ
M_1	bob	alice	assert(hung, give(alice, bob, hammer), ∅)	M_0
M_2	alice	bob	assert(hung, give(bob, alice, nail), ∅)	M_1
M_3	bob	alice	why(hung, give(bob, alice, nail), ∅)	M_2
M_4	alice	bob	assert(hang(alice), give(bob, alice, nail), [control(bob, nail)])	M_3
M_5	bob	alice	accept(hung, give(bob, alice, nail), ∅)	M_1

M_k	S_k	H_k	A_k	R_k
M_0	alice	bob	question(hung, give(ag_1, ag_2, *res*), ∅)	θ
M_1	bob	alice	assert(hung, give(alice, bob, hammer), ∅)	M_0
M_2	alice	bob	assert(hung, give(bob, alice, nail), ∅)	M_1
M_3	bob	alice	why(hung, give(bob, alice, nail), ∅)	M_2
M_4	alice	bob	assert(¬hit(alice), give(bob, alice, nail), [control(bob, nail)])	M_3
M_5	bob	alice	accept(hung, give(bob, alice, nail), ∅)	M_1

Table 2. Dialogue using an appeal to common goal (top) and using a threat (bottom)

exist between the agents interests and their social responsibilities. In order to valid this approach, we use the multiagent platform GOLEM [4] for the deployment of our agents.

The notions of social welfare can be used as a criterion to discriminate amongst arguments for the allocation of resources amongst agents taking into account the utilities/payments (resp. preferences) that agents assign to the resources (respectively goals). Existing results indicate that interaction amongst agents as well as the agents' reasoning can be designed to direct negotiation towards high quality allocations. Future works will aim at building upon the existing results to design and realise effective argumentation-based negotiation mechanisms that can be used to exhibit social-welfare related properties.

ACKNOWLEDGEMENTS

This work is supported by the Sixth Framework IST programme of the EC, under the 035200 ARGUGRID project. The author would like to thank Prof. Christophe Sibertin-Blanc for the precious comments regarding the ideas proposed in this paper.

REFERENCES

[1] Leila Amgoud, 'A general argumentation framework for inference and decision making', in *Proc. of the 21st Conference on Uncertainty in Artificial Intelligence (UAI)*, ed., Tommi Jaakkola Fahiem Bacchus, pp. 26–33, (2005).

[2] Leila Amgoud and Henri Prade, 'Handling threats, rewards, and explanatory arguments in a unified setting', *International journal of intelligent systems*, **20**(12), 1195–1218, (2005).

[3] Trevor Bench-Capon and Paul Dunne, 'Argumentation in artificial intelligence', *Artificial Intelligence*, **171**(10-15), 619–641, (July-October 2007).

[4] Stefano Bromuri and Kostas Stathis, 'Situating cognitive agents in GOLEM', in *Proc. of the Engineering Environment-Mediated Multiagent Systems Conference (EEMMAS)*, eds., Danny Weyns, Sven Brueckner, and Yves Demazeau, pp. 76–93, Leuven (Belgium), (2007). Katholieke Universiteit Leuven.

[5] Robert Taylor Clemen, *Making Hard Decisions*, Duxbury. Press, 1996.

[6] L.R. Coutinho, J.S. Sichman, and O. Boissier, 'Organizational modeling dimensions in multiagent systems', in *Iberagents*, pp. 1–15, Ribeirao Preto, Brazil, (October 2006).

[7] Michel Crozier and Erhard Friedberg, *L'acteur et le système : les contraintes de l'action collective*, Seuil, 1977.

[8] Phan Minh Dung, 'On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games', *Artificial Intelligence*, **77**(2), 321–357, (1995).

[9] Margco Guerini and Cristiano Castelfranchi, 'Promises and threats in persuasion', in *Proc. of the ECAI Workshop on Computational Models of Natu- ral Argument*, pp. 1–8, Riva del Garda, (September 2006).

[10] Antonis Kakas and Pavlos Moraitis, 'Argumentative-based decision-making for autonomous agents', in *Proceedings of the 2nd International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pp. 883–890. ACM Press, (2003).

[11] Ralph Keeney and Howard Raiffa, *Decisions with multiple objectives: Preferences and value tradeoffs*, J. Wiley, 1976.

[12] Maxime Morge, 'The hedgehog and the fox. an argumentation-based decision support system', in *Argumentation in Multi-Agent Systems: Fourth International Workshop ArgMAS, Revised Selected and Invited Papers*, eds., Iyad Rahwan, Simon Parsons, and Chris Reed, volume 4946 of *Lecture Notes in Artificial Intelligence*, pp. 114–131, Honolulu, Hawai, USA, (2008). Springer-Verlag.

[13] Iyad Rahwan and Peter McBurney, 'Argumentation technology', *IEEE Intelligent Systems, Special Issue on Argumentation Technology*, **22**(6), 21–23, (November/December 2007).

[14] C. Sibertin-Blanc, F. Amblard, and M. Mailliard, 'A coordination framework based on the sociology of organized action', in *Proc of AAMAS 2005 International Workshops on Coordination, Organizations, Institutions, and Norms in Multi-Agent Systems*, volume 3913 of *Lecture Notes in Computer Science*, pp. 3–17. Springer Berlin, (2006).

[15] Carles Sierra, Nick R. Jennings, Pablo Noriega, and Simon Parsons, 'A framework for argumentation-based negotiation', in *Proc. of the 4th International Workshop on Agent Theories, Architectures, and Languages (ATAL)*, volume 1365 of *Lecture Notes in Computer Science*, pp. 177–192, Provident, RI, (1997). Springer.

[16] John von Neumann and Oskar Morgenstern, *Theory of games and economic behavior*, Princeton University Press,, Princeton, 1944.

[17] Gerard Vreeswijk, 'Abstract argumentation systems', *Artificial Intelligence*, **90**(1-2), 225–279, (1997).