Argumentation Based Decision Making for Autonomous Agents

Antonis Kakas and Pavlos Mora¨tis
Dept. of Computer Science, University of Cyprus
P.O.Box 20537, CY-1678 Nicosia, Cyprus
antonis@ucy.ac.cy, moraitis@ucy.ac.cy

ABSTRACT

This paper presents an argumentation based framework to support the decision making of an agent within a modular architecture for agents. The proposed argumentation framework is dynamic, with arguments and their strength depending on the particular context that the agent £nds himself, thus allowing the agent to adapt his decisions in a changing environment. In addition, in order to enable the agent to operate within an open environment where the available information may be incomplete we have integrated abduction within this argumentation framework. This is particular useful when the agent £nds himself in a dilemma and hence needs additional information to resolve this. We have also developed, motivated by work in Cognitive Psychology, within the same framework an argumentation based personality theory for agents thus incorporating a dimension of individuality in the decisions of the agent.

Categories and Subject Descriptors

I.2.11 [**Distributed Arti£cial Intelligence**]: Intelligent Agents

General Terms

Theory

Keywords

Agents, Argumentation, Decision Making

1. INTRODUCTION

Automated decision making is an important problem in multi-agent systems. Autonomous and social agents need to take decisions related to their different capabilities, e.g. problem solving, cooperation, communication, etc, under complex preference policies. In most cases these policies need to be dynamic in nature depending on the particular environment in which the agent £nds himself at the time of decision making.

In this work we will adopt a modular agent architecture (see e.g. [21, 14]) where each module is dedicated to one of the capabilities of an agent. The set of these capabilities of an agent determines the behavior of an agent as an individual but also as a social entity in a community. Each module can reason independently on matters of its own knowledge and suggest what is the best course of action for these matters. Then the overall behavior of an agent is the result of the interaction among these different modules and their separate decisions. We therefore have within each module of an agent a deliberation process which is central to the implementation of the associated capability of the agent. The nature of these deliberation processes for each separate module maybe different. However, we can consider that most of them have the common characteristic of decision-making to choose among different possible options, e.g. choice of a goal or plan for a problem solving module, choice of a partner for a cooperation module, etc.

In this paper we propose an argumentative deliberation model (a preliminary version was presented in [13]) as the basis for each separate deliberation process in the different modules of an agent. In this way we are proposing a framework in which the various decision making processes of an agent are treated uniformly thus facilitating the design and implementation of such modular agents.

The proposed framework provides a high level of *adaptability* in the decisions of the agent when his environment changes. In particular, this adaptability can be effected by encompassing the in¤uence that the different relative roles of interacting agents and the context of the particular interaction can have on the deliberation process of each module of the agent. Roles and context de£ne in a natural way dynamic preferences on the decision policies of the agent at two different levels and are represented within the argumentation theory of the agent in two corresponding modular parts. In addition, aiming to provide the agent with a level of *robustness* in the face of incomplete information from the environment we integrate abduction within the argumentation framework. The agent is then able to deliberate on alternative choices and take decisions which are conditional on assumptions about the environment.

In order to give to our agents a dimension of *individuality*, as they operate within a society of agents, we are proposing to enrich the modular architecture of an agent with an additional module of *personality*. This module is based on the same argumentation framework used for the other modules by representing the personality of an agent as a decision policy according to the needs and motivations of the agents. We will adopt the classical model of Maslow [16] in which he sets up a theory of hierarchy of human needs (physiological, safety, af£liation, achievement, self-actualization) corresponding to motivational factors that drive human behavior. Then the mechanism of choosing which need to address next is carried out via a process of argumentative reasoning.

This personality module plays a central role in the decision making of the agent as it can offer an additional judgment on the decision problems of any one of the other modules depending on the

AAMAS'03, July 14–18, 2003, Melbourne, Australia.

various needs that the alternatives in these problems may address. As such the personality module can have an in¤uence on the decisions of the other modules associated with the speci£c capabilities of the agent and hence characterize the overall behaviour of the agent. We study this in¤uence and the possible con¤ict between the results of the personality theory and that of the policies of the other modules of the agent and examine ways to resolve such con¤icts by exploiting the ability of the agent to do hypothetical reasoning via abduction.

The rest of the paper is organized as follows. Section 2 sets up the basic argumentation framework for decision making by an agent and its integration with abduction. In section 3, we develop a framework for an argumentation based personality theory and in section 4 we study its interaction with the other decision policies of the agent. Section 5 discusses brie¤y related and future work.

2. ARGUMENTATIVE DELIBERATION

In this section we present the argumentation framework which will be used as a basis for the decision making processes associated to the different capabilities of an agent. This framework is developed as an extension of the approach to argumentation developed over the last decade through a series of studies [12, 11, 8, 6] on the links of argumentation to non-monotonic reasoning. The particular framework we will use, called Logic Programming without Negation as Failure $(LPwNF)$[1], was proposed originally in [11] and can be seen as a realization of the more abstract frameworks of [8, 3]. The abstract attacking relation, i.e. its notion of argument and counter-argument, is realized through monotonic proofs of contrary conclusions and a priority relation on the sentences of the theory that make up these proofs.

In this paper we extend this framework (a preliminary version was presented in [13]), following the more recent approach of other works [20, 5] to allow this priority relation and thus the attacking relation to be dynamic, making the framework more suitable for the application of agent self deliberation in a dynamically changing environment. In addition, we will integrate with argumentation abduction for dealing with the fact that the agent can be faced with incomplete information in the open environment that he operates.

2.1 Argumentation with Roles and Context

Within this extended framework of $(LPwNF)$ we aim to provide the ability of the agent to adapt its reasoning to the open and changing environment by capturing the basic concepts of agent roles and context of interaction between agents. Agents are always integrated within a (social) environment of interaction. We call this the *context* of interaction. This determines relationships between the possible roles the different agents can have within the environment. We consider, in line with much of the agent literature a *role* as a set of behaviour obligations, rights and privileges determining its interaction with other roles. Generally, the substance of roles is associated to a *default context* that de£nes shared social relations of different forms (e.g. authority, friendship, relationship, etc.) and speci£es the behaviour of roles between each others. Consequently, it implicitly installs a partial order between roles that expresses preferences of behaviour. For instance in the army context an of£cer gives orders that are obeyed by a soldier, or in a everyday context we respond in favour more easily to a friend than to a stranger.

However, a default context that determines the basic roles £lled by the agents is not the only environment where they could interact. For example, two friends can also be colleagues or an of£cer and a soldier can be family friends in civil life. Therefore we consider a second level of context, called *speci£c context*, which can overturn the pre-imposed, by the default context, ordering between roles and establish a different social relation between them. For instance, the authority relationship between an of£cer and a soldier would change under the speci£c context of a social meeting at home or the speci£c context of treason by the of£cer.

In order to accommodate in an agent's argumentative reasoning roles and context and thus make it adaptable to a changing environment of the agent we have extended the framework of $LPwNF$ so that the priority relation of a theory is not simply a static relation but a dynamic relation that captures the non-static preferences associated to roles and context. There is a natural way to do this. Following the same philosophy of approach as in [20], the priority relation can be de£ned as part of the agent's theory with the same argumentation semantics along with the rest of the theory. An extended argumentation theory is then de£ned as follows.

De£nition 1. A **argumentation theory** is a pair of sets of sentences $(\mathcal{T}, \mathcal{P})$ in the **background monotonic logic** $(\mathcal{L}, \vdash)$ of the form $L \leftarrow L_1, \ldots, L_n$, where $L, L_1, \ldots, L_n$ are positive or negative ground literals. For rules in $\mathcal{P}$ the head L refers to an (ir¤exive) higher-priority relation, i.e. L has the general form $L = h\_p(rule1, rule2)$. The derivability relation, $\vdash$, of the background logic is given by the single inference rule of modus ponens.

For simplicity, we will assume that the conditions of any rule in the theory do not refer to the predicate h_p thus avoiding self-reference problems. For any ground atom $h\_p(rule1, rule2)$ its negation is denoted by $h\_p(rule2, rule1)$ and vice-versa.

An **argument** for a literal L in a theory $(\mathcal{T}, \mathcal{P})$ is any subset, T, of this theory that derives L, $T \vdash L$, under the background logic. In general, we can separate out a part of the theory $\mathcal{T}_0 \subset \mathcal{T}$ and consider this as a non-defeasible part from which any argument rule can draw information that it might need. We will call $\mathcal{T}_0$ the **background theory**. The notion of attack between arguments in a theory is based on the possible con¤icts between a literal L and its negation and on the priority relation given by h_p in the theory.

De£nition 2. Let $(\mathcal{T}, \mathcal{P})$ be a theory, $T, T' \subseteq \mathcal{T}$ and $P, P' \subseteq \mathcal{P}$. Then (T', P') **attacks** (T, P) iff there exists a literal L, $T_1 \subseteq T'$, $T_2 \subseteq T$, $P_1 \subseteq P'$ and $P_2 \subseteq P$ s.t.:

(i) $T_1 \cup P_1 \vdash_{min} L$ and $T_2 \cup P_2 \vdash_{min} \neg L$

(ii) $(\exists r' \in T_1 \cup P_1, r \in T_2 \cup P_2$ s.t. $T \cup P \vdash h\_p(r, r')) \Rightarrow (\exists r' \in T_1 \cup P_1, r \in T_2 \cup P_2$ s.t. $T' \cup P' \vdash h\_p(r', r))$.

Here, when L does not refer to h_p, $T \cup P \vdash_{min} L$ means that $T \vdash_{min} L$. This de£nition states that a (composite) argument (T', P') attacks (or is a counter-argument to) another such argument when they derive a contrary conclusion, L, and $(T' \cup P')$ makes the rules of its counter proof at least "as strong" as the rules for the proof by the argument that is under attack. Note that the attack can occur on a contrary conclusion L that refers to the priority between rules.

De£nition 3. Let $(\mathcal{T}, \mathcal{P})$ be a theory, $T \subseteq \mathcal{T}$ and $P \subseteq \mathcal{P}$. Then (T, P) is **admissible** iff $(T \cup P)$ is consistent and for any (T', P') if (T', P') attacks (T, P) then (T, P) attacks (T', P'). Given a ground literal L then L is a **credulous** (respectively **skeptical**) consequence of the theory iff L holds in a (respectively every) maximal (wrt set inclusion) admissible subset of $\mathcal{T}$.

Hence when we have dynamic priorities, for an object-level argument (from $\mathcal{T}$) to be admissible it needs to take along with it priority arguments (from $\mathcal{P}$) to make itself at least "as strong" as the

[1]The historical reasons for the name of this framework are not important for this paper.

opposing counter-arguments. This need for priority rules can repeat itself when the initially chosen ones can themselves be attacked by opposing priority rules and again we would need to make now the priority rules themselves at least as strong as their opposing ones.

We can now de£ne an agent's argumentation theory for describing his policy in an environment with roles and context where the $higher - priority$ rules in $\mathcal{P}$ are separate into two levels.

De£nition 4. An agent's **argumentative policy theory or theory,** T**,** is a triple $T = (\mathcal{T}, \mathcal{P}_R, \mathcal{P}_C)$ where the rules in $\mathcal{T}$ do not refer to h_p, all the rules in $\mathcal{P}_R$ are priority rules with head $h\_p(r_1, r_2)$ s.t. $r_1, r_2 \in \mathcal{T}$ and all rules in $\mathcal{P}_C$ are priority rules with head $h\_p(R_1, R_2)$ s.t. $R_1, R_2 \in \mathcal{P}_R \cup \mathcal{P}_C$.

We therefore have three levels in an agent's theory. In the £rst level we have the rules $\mathcal{T}$ that refer directly to the subject domain of the agent. We call these the **Object-level Decision Rules** of the agent. In the other two levels the rules relate to the policy under which the agent uses his object-level decision rules according to roles and context. We call the rules in $\mathcal{P}_R$ and $\mathcal{P}_C$, **Role (or Default Context) Priorities** and **(Speci£c) Context Priorities** respectively.

As an example, consider the following theory $\mathcal{T}$ representing (part of) the object-level rules of an employee in a company[2].

$r_1(A, Obj, A_1) : give(A, Obj, A_1) \leftarrow requests(A_1, Obj, A)$
$r_2(A, Obj, A_1) : \neg give(A, Obj, A_1) \leftarrow needs(A, Obj)$
$r_3(A, Obj, A_2, A_1) : \neg give(A, Obj, A_2) \leftarrow give(A, Obj, A_1), A_2 \neq A_1.$

In addition, we have a theory $\mathcal{P}_R$ representing the general default behaviour of the code of contact in the company relating to the roles of its employees: a request from a superior is in general stronger than an employee's own need; a request from another employee from a competitor department is in general weaker than its own need. Here and below we will use capitals to name the priority rules but these are not to be read as variables. Also for clarity of presentation we do not write explicitly the full name of a priority rule omitting in the name the parameter terms of the rules.

$R_1 : h\_p(r_1(A, Obj, A_1), r_2(A, Obj, A_1)) \leftarrow higher\_rank(A_1, A)$
$R_2 : h\_p(r_2(A, Obj, A_1), r_1(A, Obj, A_1)) \leftarrow competitor(A, A_1)$
$R_3 : h\_p(r_1(A, Obj, A_1), r_1(A, Obj, A_2)) \leftarrow higher\_rank(A_1, A_2).$

Between the two alternatives to satisfy a request from a superior or from someone from a competing department, the £rst is stronger when these two departments are in the speci£c context of working on a common project. On the other hand, if we are in a case where the employee who has an object needs this urgently then he would prefer to keep it. Such policy is represented at the third level in $\mathcal{P}_C$:

$C_1 : h\_p(R_1(A, Obj, A_1), R_2(A, Obj, A_1)) \leftarrow common(A, Obj, A_1)$
$C_2 : h\_p(R_2(A, Obj, A_1), R_1(A, Obj, A_1)) \leftarrow urgent(A, Obj).$

Note the *modularity* of this representation. For example, if the company decides to change its policy "that employees should generally satisfy the requests of their superiors" to apply only to the direct manager of an employee we would simply replace R_1 by the new rule R_1' without altering any other part of the theory:

$R_1' : h\_p(r_1(A, Obj, A_1), r_2(A, Obj, A_1)) \leftarrow manager(A_1, A).$

Consider now a scenario where we have two agents ag_1 and ag_2 working in competing departments and that ag_2 requests an object from ag_1. This is represented by extra statements in the non-defeasible part, $\mathcal{T}_0$, of the theory, e.g. $competitor(ag_2, ag_1)$, $requests(ag_2, obj, ag_1)$. Should ag_1 give the object to ag_2 or not?

If ag_1 does not need the object then, there are only admissible arguments for giving the object, e.g. $\Delta_1 = (\{r_1(ag_1, obj, ag_2)\}, \{\})$ and supersets of this. This is because this does not have any counter-argument as there are no arguments for not giving the object since $needs(ag_1, obj)$ does not hold. Suppose now that $needs(ag_1, obj)$ does hold. In this case we do have an argument for not giving the object, namely $\Delta_2 = (\{r_2(ag_1, obj, ag_2)\}, \{\})$. This is of the same strength as Δ_1 but the argument Δ_2', formed by replacing in Δ_2 its empty set of rules of priority with $\{R_2(r_2(ag_1, obj, ag_2), r_1(ag_1, obj, ag_2))\}$, attacks Δ_1 and any of its supersets but not vice-versa: R_2 gives higher priority to the rules of Δ_2 and there is no set of priority rules with which we can extend Δ_1 to give its object-level rules equal priority as those of Δ_2. Hence we conclude skeptically that ag_1 will not give the object. This skeptical conclusion was based on the fact that the theory of ag_1 cannot prove that ag_2 is of higher rank than himself. If the agent learns that $higher\_rank(ag_2, ag_1)$ does hold then Δ_2' and Δ_1', obtained by adding to the priority rules of Δ_1 the set $\{R_1(r_1(ag_1, obj, ag_2), r_2(ag_1, obj, ag_2))\}$, attack each other. Each one of these is an admissible argument for not giving or giving the object respectively and so we can draw both conclusions credulously.

Suppose that we also know that the requested object is for a common project of ag_1 and ag_2. The argument Δ_2' is now not admissible since now it has another attack obtained by adding to the priority rule of Δ_1' the extra priority rule $C_1(R_1(ag_1, obj, ag_2), R_2(ag_1, obj, ag_2))$ thus strengthening its derivation of $h\_p(r_1, r_2)$. The attack now is on the contrary conclusion $h\_p(r_1, r_2)$. In other words, the argumentative deliberation of the agent has moved one level up to examine what priority would the different roles have, within the speci£c context of a common project. Δ_2' cannot attack back this attack and no extension of it exists that would strengthen its rules to do so. Hence there are no admissible arguments for not giving and ag_1 draws the skeptical conclusion to give the object.

We have seen in the above example that in several cases the admissibility of an argument depends on whether we have or not some speci£c information in the background theory. For example, ag_1 may not have information on whether their two departments are in competition or not. This means that ag_1 cannot build an admissible argument for not giving the object as he cannot use the priority rule R_2 that he might like to do. But this information maybe just unknown to him and if ag_1 wants to £nd a way to refuse the request he can reason further to £nd *assumptions* related to the unknown information under which he can build an admissible argument. Hence in this example he would build an argument for not giving the object to ag_2 that is *conditional* on the fact that they belong to competing departments. Furthermore, this type of information may itself be dynamic and change while the rest of the theory of the agent remains £xed, e.g. ag_1 may have in his theory that ag_2 belongs to a competing department but he has not yet learned that ag_2 has changed department or that his department is no longer a competing one.

We can formalize this conditional form of argumentative reasoning by de£ning the notion of *supporting information* and extending argumentation with *abduction* on this missing information.

De£nition 5. Let $T = (\mathcal{T}_0, \mathcal{T}, \mathcal{P})$ be an agent theory, and $\mathcal{A}$ a distinguished set of predicates in the language of the theory, called **abducible** predicates[3]. Given a goal G, a set S of ground ab-

[2]Non-ground rules represent their instances in a given Herbrand universe.

[3]Typically, the theory $\mathcal{T}$ does not contain any rules for the ab-

ducible literals consistent with the non-defeasible part $\mathcal{T}_0$ of T, is called a **strong** (respectively **weak) supporting information** for G in T iff G is a skeptical (respectively credulous) consequence of $(\mathcal{T}_0 \cup S, \mathcal{T}, \mathcal{P})$. We say that the agent **deliberates** on a goal G, *deliberate(T,G;S)*, to produce supporting information $S \neq \emptyset$ for G in the theory T.

The supporting information expressed through the abducibles predicates refers to the incomplete and evolving information of the external environment of interaction. Typically, this information relates to the context of the environment, the roles between agents or any other aspect of the environment that is dynamic. Agents can acquire and/or validate such information either through direct observation of the environment or through some interaction with other agents. This ability of the agent to deliberate with the help of abduction will be used to help it compare con¤icting decisions from different policies.

3. AGENT NEEDS AND MOTIVATIONS

In this section, we will study how our argumentation framework can help us model the individual needs and motivations of an agent. In particular, we will examine the argumentative deliberation that an agent has to carry out in order to decide which needs to address at any current situation that he £nds himself. This will then allow us to use the argumentation framework to specify different personalities of agents in a modular way independently from the other architectural elements of an agent.

We will apply the same approach as when we model a preference policy of an agent in a certain module, described in the previous section. We now consider the domain of an agent's needs and motivations where, according to the type or personality of an agent, the agent has a default (partial) preference amongst the different types of needs. Hence now the type of need, or the motivation that this need addresses, plays an analogous role to that of Roles in the previous section determining the basic behaviour of the agent in choosing amongst different goals pertaining to different needs.

We will follow the work of Maslow [16] from Cognitive Psychology (see also [17]) where needs are categorized in £ve broad classes according to the motivation that they address. These are **Physiological, Safety, Af£liation or Social, Achievement or Ego and Self-actualization or Learning**. As the world changes a person is faced with a set of potential goals from which it selects to pursue those that are "most compatible with her/his (current) motivations". We choose to eat if we are hungry, we protect ourselves if we are in danger, we work hard to achieve a promotion etc. The theory states that there is a basic ordering amongst these £ve motivations that we follow in selecting the corresponding goals. But this ordering is only followed in general under the assumption of "other things being equal" and when special circumstances arise it does not apply.

Our task here is then to model and encode such motivating factors and their ordering in a natural way thus giving a computational model for agent behaviour and personality.

Let us assume that an agent has a theory[4], $\mathcal{T}$, describing the knowledge of the agent. Through this, together with his perception inputs, he generates a set of needs that he could possibly address at any particular situation that he £nds himself. We will consider that these needs are associated to goals, G, e.g. to £ll with petrol, to rest, to help someone, to promote himself, to help the community etc. For simplicity of presentation and without loss of generality we will assume that the agent can only carry out one goal at a time and thus any two goals activated by $\mathcal{T}$ oppose each other and a decision is needed to choose one. We also assume for the moment that any one goal G is linked only to one of the £ve motivations above, m_j, and we will thus write G_j, $j = 1, ..., 5$ to indicate this, with $m_1 = Physiological$, $m_2 = Safety$, $m_3 = Affiliation$, $m_4 = Achievement$, $m_5 = Self-actualization$. Given this theory, $\mathcal{T}$, that generates potential goals an agent has a second level theory, $\mathcal{P}_{\mathcal{M}}$, of priority rules on these goals according to their associated motivation. This theory helps the agent to choose amongst the potential goals that it has and forms part of his decision policy for this. It can be de£ned as follows.

De£nition 6. Let Ag be an agent with knowledge theory $\mathcal{T}$. For each motivation, m_j, we denote by S_j the set of conditions, evaluated in $\mathcal{T}$, under which the agent considers that his needs pertaining to motivation m_j are **satis£ed**. Let us also denote by N_j the set of conditions, evaluated in $\mathcal{T}$, under which the agent considers that his needs pertaining to motivation m_j are **critical**. We assume that S_j and N_j are disjoint and hence N_j corresponds to a subset of situations where $\neg S_j$ holds. Then the **default motivation preference theory of** Ag, denoted by $\mathcal{P}_{\mathcal{M}}$, is a set of rules of the form:

$R^1_{ij} : h\_p(G_i, G_j) \leftarrow N_i$

$R^2_{ij} : h\_p(G_i, G_j) \leftarrow \neg S_i, \neg N_j$

where G_i and G_j are any two potential goals, $(i \neq j)$, of the agent associated to motivations m_i and m_j respectively.

The £rst rule refers to situations where we have a critical need to satisfy a goal G_i whereas the second rule refers to situations where the need for G_j is not critical and so G_i can be preferred. Hence when the conditions S_i hold an agent would not pursue goals of needs pertaining to this motivation m_i. In fact, we can assume that $\neg S_i$ holds whenever a goal G_i is activated and is under consideration. On the other side of the spectrum when N_i holds the agent has an urgency to satisfy his needs under m_i and his behaviour may change in order to do so. Situations where $\neg S_j$ and $\neg N_j$ both hold are in between cases where the decision of an agent to pursue a goal G_j will depend more strongly on the other simultaneous needs that he may have.

For example, when a robotic agent has low_energy, that would make it non-functional, the condition N_1 holds and a goal like $G_1 = fill\_up$ has, through the rules R^1_{1j} for $j \neq 1$, higher priority than any other goal. Similarly, when the energy level of the robotic agent is at some middle value, i.e. $\neg S_1$ and $\neg N_1$ hold, then the robot will again consider, through the rules R^2_{1j} for $j \neq 1$, the goal G_1 to £ll up higher than other goals provided also that there is no other goal whose need is critical. However, if in addition the robotic agent is in great danger and hence N_2 holds then rule R^2_{12} does not apply and the robot will choose goal $G_2 = self\_protect$ which gets a higher priority through R^1_{21}. In situations as in this example, the agent has a clear choice of which goal to select. Indeed, we can show that under some suitable conditions the agent can always decide deterministically which goal to choose.

Proposition 1. Let $\mathcal{P}_{\mathcal{M}}$ be a default motivation preference theory for an agent where $N_i \cap N_j = \emptyset$ $(i \neq j)$ and $\neg S_j = N_j$ for each j. Then given any two goals G_i, G_j only one of these goals belongs to an admissible extension of the agents theory.

Similarly, if we have $N_i \cap N_j = \emptyset$ and $\neg S_i \cap \neg S_j = \emptyset$ $(i \neq j)$ then the agent can always make a deterministic choice of which goal to choose to address in any current situation. But these conditions are too strong. There could arise situations where, according to the knowledge of the agent, two needs are not satis£ed and/or

ducible predicates. Also for simplicity of presentation we are not considering in this paper integrity constraints for abducibles.

[4]This could be an argumentation theory as described above.

where they are both urgent/critical. How will the agent decide which one to perform? The agent is in a *dilemma* as its theory will give an admissible argument for each need. For example, a robotic agent may at the same time be low in energy and in danger. Similarly, the robotic agent may be in danger but also need to carry out an urgent task of helping someone.

According to Maslow's theory decisions are then taken following a basic hierarchy amongst needs. For humans this basic hierarchy puts the Physiological needs above all other needs, Safety as the second most important with Af£liation, Achievement and Self-Actualization following in this order. Under this hierarchy a robotic agent would choose to £ll its battery despite the danger or avoid a danger rather than give help. One way to model in $\mathcal{P}_{\mathcal{M}}$ such a hierarchy of needs that helps resolve the dilemmas is as follows. For each pair k, l $s.t.$ $k \neq l$ the theory $\mathcal{P}_{\mathcal{M}}$ contains only one of the rules R^1_{kl} or R^1_{lk} and similarly only of R^2_{kl} or R^2_{lk}. Deciding in this way which priority rules to include in the theory gives a basic personality pro£le to the agent.

But this approach would be too rigid in the sense that the chosen hierarchy in this way can never be overturned under any circumstance. Often we may want a higher degree of ¤exibility in modeling an agent and indeed Maslow's hierarchy of needs applies under the assumption of "other things being equal". In other words, there maybe special circumstances where the basic hierarchy in the pro£le of an agent should not be followed. This extra level of ¤exibility is needed to capture an adaptive dynamic behaviour of an agent. For example, an agent may decide, despite his basic preference to avoid danger rather than help someone, to help when this is a close friend or a child. We can solve these problems by extending the agent theory with a third level analogous to the speci£c context level presented in the previous sections.

De£nition 7. An agent **personality** theory expressing his decision policy on needs is a theory $T = (\mathcal{T}, \mathcal{P}_{\mathcal{M}}, \mathcal{P}_{\mathcal{C}})$ where $\mathcal{T}$ and $\mathcal{P}_{\mathcal{M}}$ are de£ned as above and for each pair $i \neq j$, $\mathcal{P}_{\mathcal{C}}$ contains only one set of the following rules, for each $k = 1, 2$:

$H^k_{ij} : h\_p(R^k_{ij}, R^k_{ji}) \leftarrow true$
$E^k_{ji} : h\_p(R^k_{ji}, R^k_{ij}) \leftarrow sc^k_{ji}$
$C^k_{ji} : h\_p(E^k_{ji}, H^k_{ij}) \leftarrow true$

where sc^k_{ji} are (special) conditions whose truth can be evaluated in $\mathcal{T}$. The rules H^k_{ij} are called the **basic hierarchy** of the theory and the rules E^k_{ji} the **exception policy** of the theory.

Choosing which one of the basic hierarchy rules H^k_{ij} or H^k_{ji} to have in the personality theory determines the default preference of needs G_i over G_j or G_j over G_i respectively (for $k = 1$ in critical situations and for $k = 2$ in non-critical situations). The special conditions sc_{ij} de£ne the speci£c contexts under which this preference is overturned. They are evaluated by the agent in his knowledge theory $\mathcal{T}$. They could have different cases of de£nition that depend on the particular nature of the goals and needs that we are considering in the dilemma.

Each choice of the rules H^k_{ij} to include in the agent personality theory, determining a basic hierarchy of needs, in effect gives a different agent with a different basic pro£le of behaviour. For example, if we have H^k_{34} in $\mathcal{P}_{\mathcal{C}}$ (remember that $m_3 = Affiliation$ and $m_4 = Achievement$) we could say that this is an *altruistic* type of agent, since under normal circumstances (i.e. not exceptional cases de£ned by sc^k_{43}) he would give priority to the af£liation needs over the self-achievement needs. Whereas if we have H^k_{43} we could consider this as a *sel£sh* type of agent. To illustrate this let us consider the speci£c theory $\mathcal{P}_{\mathcal{C}}$ corresponding to Maslow's pro£le for humans. This will contain the following rules to capture the basic hierarchy of Physiological (m_1) over Safety (m_2) and Safety over Af£liation (m_3):

$H^k_{12} : h\_p(R^k_{12}, R^k_{21}) \leftarrow true,\ for\ k = 1, 2$
$H^k_{13} : h\_p(R^k_{13}, R^k_{31}) \leftarrow true,\ for\ k = 1, 2$
$H^k_{23} : h\_p(R^k_{23}, R^k_{32}) \leftarrow true,\ for\ k = 1, 2$
$E^2_{21} : h\_p(R^2_{21}, R^2_{12}) \leftarrow sc^2_{21}$
$C^2_{21} : h\_p(E^2_{21}, H^2_{12}) \leftarrow true$
$E^2_{31} : h\_p(R^2_{31}, R^2_{13}) \leftarrow sc^2_{31}$
$C^2_{31} : h\_p(E^2_{31}, H^2_{13}) \leftarrow true$
$E^2_{32} : h\_p(R^2_{32}, R^2_{23}) \leftarrow sc^2_{32}$
$C^2_{32} : h\_p(E^2_{32}, H^2_{23}) \leftarrow true.$

The conditions sc^2_{21} are exceptional circumstances under which we prefer a safety need over a physiological need, e.g. sc^2_{21} could be true if an alternative supply of energy exists. Similarly for sc^2_{31} and sc^2_{32}. Note that if we are in a situation of critical physiological need (i.e N_1 holds and hence R^1_{12} applies) then this theory has no exceptional circumstances (there is no E^1_{21} rule) where we would not prefer to satisfy this physiological need over a critical safety need. Similarly, this pro£le theory does not allow any af£liation need to be preferred over a critical safety need; it does not allow a "heroic" behaviour of helping. If we want to be more ¤exible on this we would add the following rules in the pro£le:

$E^1_{32} : h\_p(R^1_{32}, R^1_{23}) \leftarrow sc^1_{32}$
$C^1_{32} : h\_p(E^1_{32}, H^1_{23}) \leftarrow true$

where the conditions sc^1_{32} determine the circumstances under which the agent prefers to help despite the risk of becoming non-functional, e.g. when the help is for a child or a close friend in great danger.

Given any personality theory we can show that an agent can always decide which goal to pursue.

Proposition 2. Let $T = (\mathcal{T}, \mathcal{P}_{\mathcal{M}}, \mathcal{P}_{\mathcal{C}})$ be an agent theory according to de£nition 7 and G_i, G_j ($i \neq j$) be any two potential goals addressing different needs. Then given any situation there exists an admissible argument for only one of the two goals.

When goals are associated with more than one need the agent may £nd that he can set up an admissible argument for each one of a set of competing goals depending on which particular needs it considers for each goal. The agent then needs to combine these results in some way in order to reach a conclusion of which goal to choose. In its general form this problem is linked with the problem of multi-criteria reasoning [25]. Here we will adopt a simple qualitative approach, exploiting the result of proposition 2, where the agent considers separately each pair of motivation labels between two competing goals and decides on the goal which is preferred with the label strongest in the basic hierarchy of needs of the agent. This choice is again motivated by Malsow's theory where a human "considers a particular need only when stronger needs are satis£ed".

Hence, given two goals G^a and G^b with sets of labels M^a and M^b respectively we consider each pair of labels $m^a_i \in M^a$ and $m^b_j \in M^b$ and £nd which goal is preferred under these labels alone. We recorded this result as (G^a, m^a_i) or (G^b, m^b_j) depending on the case for each such pair. Note that if the $m^a_i = m^b_j$ then we record both (G^a, m^a_i) and (G^b, m^b_j). Given all these separate decisions the agent then chooses the goal which has recorded the strongest, according to his own basic hierarchy of needs, label m. If such a label exists this goal is then a **skeptical** conclusion of his personality theory. If for both goals the strongest label recorded is the same then the agent cannot decide[5] and we say that both goals are **credulous** conclusions of his personality theory. In this case, the agent

[5]Here we could apply additional methods to compare further such goals, e.g. lexicographic comparison on their recorded labels, but this is beyond the scope of the current paper.

needs more information in order to make a de£nite decision. We will consider this problem in the next section.

4. CAPABILITIES & PERSONALITY

In this section, we study how the personality of an agent can be integrated in his architecture and how this personality can in¤uence the decision making of the agent associated to his different capabilities. We will therefore be examining how the various argumentation based decision processes of the agent can be integrated with that for his personality in order to provide overall decisions that take into account the individual character of the agent.

Let us £rst illustrate through two extensive examples the in¤uence that the personality of an agent can have on his decisions showing how the uniform approach based on argumentation facilitates this. In the £rst example we will consider the task of deciding, within the problem solving module of the agent, which requested task to perform according to his "professional" policy of how to consider such requests. The argumentation theory T_1 below shows a simple part of such a policy.

$r1(A,T1,A1) : perform(A,T1,A1) \leftarrow ask(A1,T1,A)$
$r2(A,T1,T2,A1) : \neg perform(A,T1,A1) \leftarrow perform(A,T2,self)$
$R1 : h\_p(r1(A,T1,A1), r2(A,T1,T2,A1)) \leftarrow higher\_rank(A1,A)$
$R2 : h\_p(r2(A,T1,T2,A1), r1(A,T1,A1)) \leftarrow competitor(A,A1)$
$C1 : h\_p(R1(A,T1,T2,A1), R2(A,T1,T2,A1)) \leftarrow$
$\quad common\_project(A,T1,A1)$
$C2 : h\_p(R2(A,T1,T2,A1), R1(A,T1,T2,A1)) \leftarrow urgent(A,T2)$

Suppose that an agent ag has a request for $task1$ from another agent ag_1 and that also currently he has a separate task, $task2$, to perform for himself, i.e that $perform(ag, task2, self)$ holds. Let us also assume that his background theory contains the following information: $competitor(ag, ag_1)$, $higher\_rank(ag_1, ag)$, $common\_project(ag, task1, ag_1)$. The agent thus has to choose $G^a = perform(ag, task1, ag_1)$ or $G^b = perform(ag, task2, self)$ i.e. between performing $task1$ for ag_1 or performing the $task2$ for himself. According to T_1 and his background theory, the agent will choose skeptically the goal G^a.

As we have presented in the previous section any goal that the agent considers may be associated by the agent to some particular needs or motivations that this goal addresses. Goals are therefore labeled by the agent according to these needs or in other words goals are categorized in one or more of the speci£ed, e.g. by Maslow's theory, categories of motivations. In our present work we consider that this association of the agent's possible goals with the given motivations is part of his background knowledge and it is acquired during the agent's design phase. Hence this association or labeling of goals is computed by the agent via a relatively simple non-defeasible process in the background theory. In a more advanced design of agents this labeling of goals could be the result of a learning process where an agent learns to associate a category m_i to a goal if he observes that when this goal has been achieved in the past it has had a positive repercussion on the given motivation m_i.

Returning to our example let us now assume that the goal G^a is associated (labeled) by the agent's background theory with m_3, i.e the need of the agent to satisfy goals for the society, and G^b with m_4 i.e. his need to satisfy personal needs. We thus write these as G^a_3 and G^b_4. Lets us also consider a personality theory T_2 as follows (shown below partly and in a simpli£ed form) that would rather characterize the agent as sel£sh. The agent prefers goals that refer to his personal achievement (m_4) except when this goal can be dangerous for his company.

$R^2_{43} : h\_p(G_4, G_3) \leftarrow \neg S_4, \neg N_3$
$R^2_{34} : h\_p(G_3, G_4) \leftarrow \neg S_3, \neg N_4$
$H^2_{43} : h\_p(R^2_{43}, R^2_{34}) \leftarrow true$
$E^2_{34} : h\_p(R^2_{34}, R^k_{43}) \leftarrow dangerous\_for\_company(G_4)$
$C^2_{34} : h\_p(E^2_{34}, H^2_{43}) \leftarrow true$

According to this theory T_2 and background knowledge where $\neg S_4, \neg N_3, \neg S_3, \neg N_4$ hold the agent will choose skeptically the goal G^b_4 because he has no information that (performing) $task2$ is dangerous to the company. Therefore, this agent will £nd himself in a dilemma because, reasoning as a professional under T_1, he will have to choose the goal G^a_3 while, reasoning as an individual under T_2, he will choose the G^b_4. We will see below in the next subsection how do deal with such dilemmas.

Suppose now that the personality of the agent is given by the theory T_3 below instead of T_2. This theory would characterize the agent as altruist or collaborative.

$R^2_{34} : h\_p(G_3, G_4) \leftarrow \neg S_3, \neg N_4$
$R^2_{43} : h\_p(G_4, G_3) \leftarrow \neg S_4, \neg N_3$
$H^2_{34} : h\_p(R^2_{34}, R^2_{43}) \leftarrow true$
$E^2_{43} : h\_p(R^2_{43}, R^k_{34}) \leftarrow against\_principle\_reasons(G_3)$
$C^2_{43} : h\_p(E^2_{43}, H^2_{34}) \leftarrow true$

According to T_3, the agent will choose goal G^a_3 because he has no information that choosing to achieve $task1$ could be against some of his personal principles. Therefore, in this case, T_1 and T_3, expressing his way of thinking as a professional and as an individual, respectively, are in agreement and therefore the agent will not have any dilemma to choose the same goal G^a_3.

Let us now consider another analogous example of the interaction between the deliberation process of how an agent chooses partners (i.e his cooperation capability) and his personality. We suppose that the following theory T_4 is part of the agent's policy in his cooperation module under which the agent selects his collaborators for a speci£c task based on purely professional criteria.

$r1 : request\_help(A,T,A1) \leftarrow need\_cooperation(A,T),$
$\qquad relevant(T,A1)$
$r2 : \neg request\_help(A,T,A1) \leftarrow request\_help(A,T,A2), A1 \neq A2$
$R1 : h\_p(r1(A,T,A1), r1(A,T,A2)) \leftarrow management\_task(T),$
$\qquad manager(A1)$
$R2 : h\_p(r1(A,T,A2)), r1(A,T,A1)) \leftarrow technical\_task(T),$
$\qquad expert(A2)$
$C1 : h\_p(R1, R2) \leftarrow market\_share\_increase\_period$
$C2 : h\_p(R2, R1) \leftarrow \neg market\_share\_increase\_period$

This theory, at the object level says that for any speci£c task the agent can have only one collaborator and that collaborators are considered according to their relevant expertise for the task at hand. At the roles level the theory says that if the task contains a management issue he prefers to choose a manager agent while if the task contains technical issues he prefers to choose an expert agent. When both apply (i.e. the task contains both management and technical issues), the theory at the context level expresses a priority according to the current period. If the current period imposes the need of a market share increase, the preference for a manager is stronger, otherwise the preference for an expert is stronger.

Let us assume that the agent has a certain task for which he needs help and that he has two alternatives, either to request help from agent ag_1 (i.e. to choose goal $G^a = request\_help(ag, task, ag_1)$)) or from ag_2 (i.e. to choose goal $G^b = request\_help(ag, task, ag_2)$)). In his background theory the following hold: $management\_task(task)$,

$manager(ag_1), technical\_task(task), expert(ag_2)$, $need\_cooperation(ag, task), relevant(task, ag_1)$, $relevant(task, ag_2), market\_share\_increase\_period$. According to this theory T_4 the agent will choose G^a as this is a skeptical conclusion from this theory.

Let us now also suppose that the two goals G^a and G^b are labeled in the agent's background theory with m_3 and m_5 respectively, i.e. that the agent considers that the choice of ag_1 serves the company while the choice of ag_2 serves his own ambitions, and that this agent has an "ambitious" personality captured by the theory:

$R^1_{53} : h\_p(G_5, G_3) \leftarrow N_5$
$R^1_{35} : h\_p(G_3, G_5) \leftarrow N_3$
$H^1_{53} : h\_p(R^1_{53}, R^1_{35}) \leftarrow true$
$E^1_{35} : h\_p(R^1_{35}, R^k_{53}) \leftarrow job\_loss(self, G_5)$
$C^1_{35} : h\_p(E^1_{35}, H^1_{53}) \leftarrow true$

According to this theory (and assuming that N_3 and N_5 hold) the agent will choose the achievement of the goal G^b_5, which corresponds to the choice of the agent ag_2. This can be overturned only if G^b_5 will lead (according to his background theory) to the agent loosing his job, but the agent does not have such information. Hence again the agent is in a dilemma between his personality based choice for G^b_5 and the professional choice for G^a_3.

4.1 Resolving Policy & Personality Con¤icts

We have seen above that it is possible for the "professional" policy of an agent's module to be in con¤ict with his personality. Such con¤icts need to be resolved in order for the agent to overcome his dilemma and decide on a de£nite action. This problem, again as at the end of section 3, can be addressed by several different approaches drawing from work on multi-criteria decision theory. Here we will present a method of con¤ict resolution that exploits the ability of the agent to synthesize argumentation and abductive reasoning and is based on the assumption that these con¤icts occur due to lack of information at the time of reasoning.

Given two opposing goals G_1 and G_2 there are three possible cases of such a con¤ict. These are:

Case1 G_1 and G_2 are skeptical conclusions of the professional theory of a module and the personality theory respectively,

Case2 G_1 is a skeptical conclusion of the professional theory of a module and G_2 is a credulous conclusion (and hence so is G_1) of the personality theory (or vice-versa),

Case3 G_1 and G_2 are credulous conclusions of the professional theory of a module and the personality theory respectively.

A mechanism for resolving these con¤icts is as follows: (i) Suspend Decision, (ii) Deliberate on these goals to £nd supporting information that would strengthen or weaken the conclusions of the separate theories, (iii) evaluate if possible (some of) this supporting information in the external environment. Then if this results in Case 2 the agent decides for the goal which is skeptically true. Otherwise, if time is out and a decision needs to be made then the agent in Case 1 chooses one of the goals according to a simple preference for or against the personality choice, given to the agent by the designer speci£cally for such contingencies. In case 3 the agent if indeed it is necessary for him to choose can select one of the two goals at random. Note that in the second step of deliberation there may exist several different assumptions that can strengthen or weaken the conclusions and the agent may need to evaluate these with respect to each other. The details of this evaluation are again beyond the scope of this paper.

Let us analyze a little further the second step in this process. In order to weaken a skeptical decision G of either theory T, professional or personality, the agent deliberates on its negation (or an opposing goal), i.e considers $deliberate(T, \neg G; S)$ as in de£nition 5, to £nd a weak supporting information S for this. Thus his argumentation reasoning together with abduction for possible missing information produces information that if true would give an admissible argument for $\neg G$ and hence G would no longer be skeptically true. In the £rst example above the skeptical conclusion of the professional theory, namely $G^a = perform(ag, task1, ag_1)$, can be weakened by deliberating on its negation and hence on $G^b = perform(ag, task2, self)$. This can be supported by the abducible assumption $urgent(self, task2)$. The agent can then evaluate/check whether this is now true in his world and hence be sure whether he can weaken the conclusion G^a. If it is not possible to do this check the agent can still decide to weaken this conclusion under this assumption that his own task was urgent which he can offer as an explanation for his decision if needed. Similarly, to strengthen a credulous conclusion for G to a skeptical conclusion we need to deliberate on this goal to £nd strong supporting information for this. If this information were true then the goal would be a skeptical conclusion of the theory.

This process of strengthening or weakening a conclusion is the same in either type of theory, professional or personality, with the additional possibility in the case of the personality theory of £nding new supporting information that gives new motivation labels to a goal. In fact, here the agent may make directly assumptions on extra needs that a given goal addresses, i.e. we can consider the label as an abducible predicate, which are not testable in the world and thus do not need justi£cation from the external world. Under these "personal assumptions" of the agent he can change the strength of a conclusion from his personality theory and hence decide on which goal to choose justi£ed by these assumptions. In the £rst example above, the agent can weaken the skeptical conclusion of his personality theory for $G^b_4 = perform(ag, task2, self)$ by assuming that the opposing goal $G^a_3 = perform(ag, task1, ag_1)$ also serves the self achievement motivation m_4. With this extra label for G^a now both goals become credulous conclusions of his personality theory T_2. Hence the agent under this "personal assumption" that G^a also serves m_4, is able to resolve the dilemma (by case 2) he had between his professional theory for G^b and his personality theory for G^a in favour of G^b, which is now the only skeptical conclusion.

5. RELATED WORK AND CONCLUSIONS

We have proposed an argumentation based framework for decision making by an autonomous agent that combines properties of adaptability, robustness and individuality for the agent. The framework can embody in a direct way the various decision policies and accompanied knowledge of the agent. This allows changes to the agent's argumentation theories to be localized and easily accommodated. In turn this facilitates the implementation of a modular architecture for agents. The framework and its integration with abduction has been implemented in a general system for argumentative deliberation and is available at http://www.cs.ac.cy/∼nkd/gorgias.

Earlier approaches on agent argumentation have been proposed in [23, 22, 18, 1, 2] and used for various purposes such as modeling dialogues, negotiation and persuasion. In comparison, our framework allows a high degree of ¤exibility in the adaptation of the agent's argumentative reasoning to a changing environment. The dynamic preferences that our framework supports and its encapsulation of roles and context result in the transparent integration of changes in the environment within the argumentative deliberation

process itself. Furthermore, the integration of argumentation with abduction provides a useful means to address in a natural way the dilemmas that can be created due to lack of information.

More generally, our approach to argumentation draws from similar work by [20, 19] where also dynamic priorities are used. Our emphasis though is on the use of argumentation for autonomous agent decision making whereas this work draws its application and motivation from legal reasoning. The integration in our framework of abduction provides additional ¤exibility in the reasoning in the face of incomplete information. Another recent logic programming framework that captures a dynamic form of reasoning is that of [15].

Our approach to decision making, can also be considered in relation to the general £eld of qualitative decision theory (see e.g. [4, 7, 24]). Our work shows explicitly how such a theory can form the basis of decision making for an autonomous agent. Indeed, our approach builds on the most important characteristics for a such theory (see [9]), namely preference revision (i.e. preference statements that can be revised in light of more speci£c information) and defeasible reasoning.

Following the work of Maslow's hierarchy of needs [16], we have used our argumentative deliberation framework to model an agent's needs corresponding to motivational factors. This allows the expression of different personality pro£les of an agent in a modular and ¤exible way. In the agent literature [17] have already used Maslow's theory for guiding the behaviour of deliberative and reactive agents in various unpredictable environments. However, to our knowledge, this is £rst time that an argumentative deliberation framework is used to encode and model these motivation factors, in such a way that, we believe, allows a more natural expression of several behaviours.

In the future we plan to study further the problem of con¤ict resolution between policies using ideas from multi-criteria decision theory [25]. More importantly, we need to understand the mechanisms of goal generation in relation to the needs and motivations that these generated goals address according to the agent's knowledge and personality theory. Also a deeper study is needed to explore the ¤exibility of the framework in modelling different agent personalities. Here we can again draw from work in cognitive science (see e.g. [10]) on the characteristics of human personalities. It is also important to study how these different personalities play a role in the interaction among agents especially in relation to the problem of forming heterogeneous communities of agents, where the deliberation process of an agent may need to take into account (his knowledge of) the personality of the other agents.

Other future work concerns the application of our argumentation framework to different forms of interaction between agents, e.g. negotiation and conversation. We also aim to incorporate a mechanism of updating the knowledge of an agent from changes of the environment and study how our argumentative agent will be affected by this.

6. ACKNOWLEDGMENTS

This work is partially funded by the Information Society Technologies programme of the European Commission, Future and Emerging Technologies under the IST-2001-32250 project. We thank the rest of the partners for their helpful comments and Neophytos Demetriou for his help with the implementation of the framework.

7. REFERENCES

[1] L. Amgoud, N. Maudet, and S. Parsons. Modelling dialogues using argumentation. In *ICMAS-00, pp. 31-38*, 2000.

[2] L. Amgoud and S. Parsons. Agent dialogues with con¤icting preferences. In *ATAL01*, 2001.

[3] A. Bondarenko, P. M. Dung, R. A. Kowalski, and F. Toni. An abstract, argumentation-theoretic framework for default reasoning. *Arti£cial Inelligence*, 93(1-2):63–101, 1997.

[4] C. Boutilier. Toward a logic for qualitative decision theory. In *KR94*, 1994.

[5] G. Brewka. Dynamic argument systems: a formal model of argumentation process based on situation calculus. In *Journal of Logic and Computation, 11(2), pp. 257-282*, 2001.

[6] Y. Dimopoulos and A. C. Kakas. Logic programming without negation as failure. *ILPS'95, pp. 369-384*, 1995.

[7] J. Doyle and M. Wellman. Representing preferences as ceteris paribus comparatives. In *Working Notes of the AAAI Spring Symposium on Decision-Theoretic Planning*, 1994.

[8] P.M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. In *Arti£cial Intelligence, 77, pp. 321-357 (also in IJCAI'93)*, 1995.

[9] V. Ha. Preference Logics for Automated Decision Making. http://www.cs.uwm.edu/public/vu/papers/qdtsurvey.pdf.

[10] Great Ideas in Personality. Five-Factor Model. www.personalityresearch.org/big£ve.html, 2002.

[11] A. C. Kakas, P. Mancarella, and P.M. Dung. The acceptability semantics for logic programs. In *Proc. ICLP'94, pp. 504-519*, 1994.

[12] A.C. Kakas, R.A. Kowalski and F. Toni. Abductive logic programming. In *Journal of Logic and Computation, 2(6), pp. 719-770*, 1992.

[13] A. C. Kakas and P. Moraitis. Argumentative Agent Deliberation, Roles and Context. In Computational Logic in Multi-Agent Systems (CLIMA02), 2002.

[14] N. Karacapilidis and P. Moraitis. Engineering issues in inter-agent dialogues. In *ECAI02, Lyon, France*, 2002.

[15] J.A. Leite. Evolving Knowledge Bases. IOS Press, 2003.

[16] A. Maslow. Motivation and Personality. Harper and Row, New York, 1954.

[17] P. Morignot and B. Hayes-Roth. Adaptable motivational pro£les for autonomous agents. Knowledge Systems Laboratory, TR KSL 95-01, Stanford University, 1995.

[18] S. Parsons, C. Sierra, and N.R. Jennings. Agents that reason and negotiate by arguying. In *Logic and Computation 8 (3), 261-292*, 1998.

[19] H. Prakken. Logical Tools for Modeling Legal Reasoning: A study of Defeasible Reasoning in Law, Kluwer, 1997.

[20] H. Prakken and G. Sartor. A dialectical model of assessing con¤icting arguments in legal reasoning. In *Art£cial Intelligence and Law Vol 4, pp. 331-368*, 1996.

[21] J. Sabater, C. Sierra, S. Parsons, and N.R. Jennings. Engineering executable agents using multi-context systems. Journal of Logic and Computation, 12, 2002.

[22] C. Sierra, N.R. Jennings, P. Noriega, and S. Parsons. A framework for argumentation-based negotiation. In *ATAL-97, pp. 167-182*, 1997.

[23] K. Sycara. Argumentation: Planning other agents' plans. In *IJCAI-89, pp. 517-523*, 1989.

[24] S. Tan and J. Pearl. Qualitative decision theory. In *AAAI-94, pp. 928-932*, 1994.

[25] P. Vincke Multi-criteria Decision Aid. John Wiley, 1992.