A Game Theoretic Approach to Norms and Agents

Guido Boella and Leonardo Lesmo
Dipartimento di Informatica and Centro di Scienza Cognitiva
Università di Torino
e-mail: {guido,lesmo}@di.unito.it

September 6, 2001

Abstract

In recent years, multi-agent systems have become one of the most promising approaches for organizing software. However, the demand for capabilities of autonomous decision making poses some requirements on agent architectures. It is necessary, in fact, that the behavior of an agent be suitably constrained to make it socially compatible with the community of agents. In this paper, we propose a model for obligations and rules which is inspired to E. Goffman's work in sociology, and which can be integrated in existing BDI (Belief, Desire, Intention) agent architectures. We show that *Decision Theory* and *Anticipatory Coordination* can be the basic building blocks for an agent model where obligations are associated with sanctions (and with another agent, the *normative agent*, who takes care of making the obligation respected). Anticipatory Coordination is required to foresee the normative agent reactions, while Decision Theory enables an agent to choose in a rational way the most promising line of action.

1 Introduction

A *Multi-Agent System* is composed of a set of software programs called *agents*; each agent must be able to interact with other agents and to choose autonomously the next action to execute. Usually, this choice is based on the agent's goals and preferences, and it aims at maximizing the agent's satisfaction (or the satisfaction of the human owner of the agent). But it can happen that the preferences of an agent conflict with the preferences of other agents in the community. In such a contexts, the importance of the concepts of norm and obligation has become apparent [27], [23]. Norms have been proposed as a solution for improving agent performances [46] and also for regulating e-commerce [27] and electronic marketplaces [26]. In turn, the conceptual tools developed by agent theories may be useful for a better understanding of the concepts of norms and obligations in AI and social sciences as shown in [23].

In principle, an obligation is something an agent is obliged to do. In other words, given an initial situation, in any course of events produced by the action(s) chosen by the agent, the obligation must be fulfilled. However, this need not be the most rational way for an agent to act. There can be situations where different obligations contrast with each other, or situations where an obligation cannot be reconciled with the agent's personal desires or goals. In these cases, the agent must evaluate the situation carefully and must decide if the obligation (or, which obligation) must be pursued, and in which way.

As we will see in this paper, the game theoretic approach to norms will allow to develop *deliberate normative agents* ([20]), which have an explicit representation of norms and can reason to decide whether fulfilling them or not: they will be able to find a trade off between their goals and the norms they have accepted; fulfilling an obligation will be just one possibility for them, and they will choose the course of action which maximizes their advantage.

The definition of norm we propose in this paper, and the related reasoning framework for taking deliberations dealing with norms, are inspired by the work of E. Goffman ([33]) in the sociological field.

According to Goffman [32], "a norm is that kind of guide to action that is supported by social sanctions".[1] And A. Giddens adds: "a sanction is defined as a reaction of others to the behavior of an individual or a group, a reaction having the goal to enforce the respect of a given norm".[2]
Hence, from the point of view of sociology, norms come always together with sanctions[3]; since sanctions are actions, they presuppose, in turn, someone to perform them.

This assertion could be debatable if Goffman did not project it into a wider framework about multi-agent situations. For Goffman, "individuals, like other objects in this world, affect the surrounding environment in a manner congruent with their own actions and properties. Their mere presence produces signs and marks. Individuals, in brief, exude expressions".[4] On the other hand, "all organisms after their fashion make use of information collected from the immediate environment so as to respond effectively to what is going on around them and to what is likely to occur".[5]

Restating this in multi-agent theory terminology, we have that an agent who acts in an environment, affects it in such a way that other agents may recognize his intentions. Since resources are generally scarce, and, hence, agents have conflicting objectives, it is dangerous for an agent to act in order to achieve his own goals, without considering what other agents know about him. In fact:

> Aware that his actions, expressions, and words will provide information to the observer, the subject incorporates into the initial phases of this activity a consideration of the informing aspects of its later phases, so

[1] [32], p.62

[2] [29], p.120

[3] Actually, written law includes norms without sanction. We believe that they constitute an anomaly (even if not rare): some of them are due to the difficulty of integrating a possible sanction with the rest of the rules, others to the desire of achieving extra-legal (i.e. political) goals.

[4] [33], p.4.

[5] [33], p.10.

> that the definition of the situation he eventually provides for the observer hopefully will be one he feels from the beginning would be profitable to evoke.[6]

In this way, agents engage in "expression games", resorting to their empathy and ability to "take the attitude"[7] of the other observer, as G.H. Mead has remarked: "the agent takes the viewpoint of the observer, but he does not 'identify' his interests with it". He does so "only insofar as the observer is engaged in observing him and ready to make decisions on this basis, and only long enough and deep enough to learn from this perspective what might be the best way to control the response of the person who will make it".

So, when an agent considers which course of action to follow, before he takes a decision, he depicts in his mind the consequences of his action for the other involved agents, their likely reaction, and the influence of this reaction on his own welfare. He will adapt his actions to the other agents' reaction before it can even happen. Goffman calls this form of reasoning "strategic interaction".

As an analytical tool for modeling deliberation in situations of strategic interaction, Goffman proposes *Game Theory* ([51]). In fact, Game Theory enables one to base a decision not only on the *expected payoff* of an action, but also to model in a structured way the situation of other agents, to predict their rational reaction, and finally, to choose what to do on the basis of the predicted possible final outcomes and their utility for the agent.

As Goffman has noticed, an agent who has to follow some norms can be considered as a player in a game, where the payoffs of his actions depend on the subsequent actions of another social agent, a second player who has the role of making the norms respected [8].

In this paper, we focus our attention on the way the first player simulates the reasoning of the second player, in order to anticipate her possible moves. In particular, we assume that both players are intelligent deliberative agents: the *bearer* of the obligation, who must respect the norm, and the *normative* agent (an *authority*, in formal situations), which has posed the obligation, wants that the *bearers* of the norm fulfill it, and (possibly) will sanction the violators.[9] So, as "it is generally acknowledged that norms and normative action emphasize autonomy on the side of *decision*" [10], we

---

[6][33], p.12.

[7][33], p.13.

[8]The game theoretic perspective offers various ways for modeling deliberation, i.e. the techniques developed in the area of decision theoretic planning (i.e., non classical planning based on decision theory and/or game theory, see [13], [34], [30]). It is useful to remember that decision and game theoretic planning have an important role in modeling multi-agent situations, as, for example, [12] and [41] have recently argued.

[9]In [23]'s terminology, we assume that the *sovereign* (the agent who issued the norm) is a *defender* too (i.e., an agent who watches over the norm). This is a rather strong assumption, and it can be argued that it is usually wrong. However, the origin of the norm, as well as the source of the power of the authority who issued it, are complex problems, which are outside the scope of this paper. In other words, we aim at modelling how the norms affect the behavior of agents, and not where they come from.

[10][23], p.100.

can also pay due attention to the fact that norms and obligations are enforced by the *normative* agent, who is an autonomous agent, too.

The first consequence of the approach proposed here is that the *bearer* of the obligation has to consider explicitly the disadvantage of facing the sanction when he considers whether to fulfill the obligation. However, the sanction is not a granted exogenous event, but it is the result of an autonomous activity of the *normative* agent. She has the goal of checking the fulfillment of the norm and has a plan for doing so and eventually posing the sanction. But she also has other goals, preferences and obligations as any other agent.

The *bearer* of the obligation has to take into account all these facts when he considers the advantage of fulfilling or not the obligation: i.e., he has to model (recursively) also the *normative* agent as an agent. The recursive modeling of agents is a trend in agent theories which has been developing in the last years [30], [40], [54] and [6]: these approaches are motivated by the fact that every action of an agent has an impact on the choices of other agents who can react to it. On the basis of the available information about the status of the other agents (i.e., their beliefs, goals, obligations, preferences and available plans), an agent can try to predict what they will do depending on what he decides to do [5].

A consequence of the overall approach is that norms need not be represented by a specific primitive propositional attitude with a distinct ontological status, but as a combination of beliefs and goals of the agent subjected to the norm, and of the agent who has to enforce the respect of the norm: in particular, the goal of avoiding sanctions, the goal of not violating the norm and the belief that there is another agent who has the goal of sanctioning violations. This is the basis of the integration between norms and the other pro-attitudes. In this way, the sophisticated models developed for agents can be exploited for modeling deontic reasoning and, at the same time, agent models can be endowed with normative concepts.

On the other hand, the recursive modeling of the *normative* agent opens the way to another opportunity for the *bearer* besides a better evaluation of the resulting final state. The *bearer* agent can reason about how the *normative* agent will (decide to) check the fulfillment and how and when she will apply the sanction if he discovers a violation. This knowledge can allow an agent to predict when the *normative* agent possibly fails to become aware of a violation and how to induce him to this failure by means of some action. In fact, the *normative* agent has to check whether the obligation has been fulfilled before applying the sanction. But, checking the fulfillment and applying the sanction have a cost for him, so he may decide not to do anything.[11] Finally, the action of checking the fulfillment may fail with a certain probability. In this case, the decrease in the final utility due to the sanction must be weighed according to the probability of success of the *normative* agent (if she fails to discover the violations, she cannot apply the sanction).
So, there are various motivations for an agent to decide not to fulfill an obligation Ω:

1. The agent wants to respect Ω, but he does not know how.

[11] For a public administration, checking fiscal evasion has sometimes a cost which does not cover the returns gained from the payment of monetary sanctions.

2. All possible plans which lead to the fulfillment of Ω achieve a lower utility than some other plan. In particular, this may happen if some of the actions do not ensure that the *normative* agent becomes aware of the fulfillment so that she will probably apply the sanction anyway.
3. There is some plan which does not fulfill Ω but which induces the *normative* agent to believe otherwise.
4. There is some plan which does not fulfill Ω but which make the sanction impossible to apply.
5. The *bearer* may choose a plan which misleads the *normative* agent so that she selects a sanction which she believes can be applied, whereas, as a matter of fact, it cannot be applied.
6. The *bearer* of the obligation can bribe (or menace) the *normative* agent so that she does not apply the sanction.

Obligations have been discussed in the field of multi-agent systems mostly in order to build agents that respect a certain behavior. Hence, the analysis of the possible deviations from the norm seems at first sight inappropriate. Instead, there are a number of reasons for the present work. First of all, obligations must be distinguished from other propositional attitudes as goals and intentions. If the only possible deviation were of kind 1, an obligation would be similar to an intention (as happens, e.g., in [28]): that is, it is given up only when it is impossible to achieve.

But, more importantly, possible deviations should be analysed in order to let agents reason about the behavior of other agents (and human users), which are not necessarily built to respect obligations. In particular, in some domains, agents must be able to judge whether the other ones are *trusted* and maintain obligations concerning *security* and *privacy* (see [21]). In other words, we claim that an agent is able to understand the behavior of another agent just in case he is able to build a model of his behavior in terms of possible plans and goals. So, no deceiving behavior can be understood (i.e. recognized) unless the agent has some knowledge about deceiving: any 'honest' normative agent needs some knowledge about dishonest behavior if it is deemed to detect such a behavior.

If agents who respect (if they can) obligations are needed, there are some ways to enforce the fulfillment of norms:

- The content of a certain obligation Ω can occur also as a preference of the agent: in this way, when adopted, it becomes similar to an intention (reinforced by the possible sanction): the agent directly achieves an utility from the satisfaction of the obliged state (the content of the obligation is a *value* for the agent).[12]
- The agent may have the preference not to mislead the *normative* agent: the former agent does not do anything to induce false beliefs in the *normative* agent, e.g., that the obligation is fulfilled when it is not the case. In this way, the agent cannot exploit the possibilities described at point 3 and 4 above.

[12] Emotions can play an important role, as well, in the decision process leading to norm fulfillment. [31] propose a decision-theoretic framework for modeling emotions. In this perspective, for instance, the set of alternative actions in a decision problem can be affected by the emotional state of the agent.

- The agent has some *social goal* which makes him disprefer situations where he is liable (for example because he does not want that other agents decrease the trust they have on him).

Moreover, as [42] as noticed, the motivations that lead an agent to consider the norm as a value for him can be further analysed. In particular, he claims that in game theoretic formulation of norms, an agent can have preferences towards them if he adopts a "plural perspective" rather than an individualistic one.

In this paper, we do not address generic moral assertions as "there should be no war" or technical assertions as "in order to print a file, you should use the 'lpr' command".[13] Our proposal is directed towards those obligations which are *personal* (i.e., they concern certain individuals), and such that there is some entity which sanctions who violates them.

This assumption restricts the scope of the paper. But it can be noticed that the basic approach is only partially affected by this limitations. Our model, as shown in [7], covers also informal norms such as those involved in natural language communication: as [29] notices, there is a continuum between formal norms issued by some entity and informal ones, whose respect is enforced in a spontaneous way by the members of a group. Also for moral norms, it holds that an agent breaking the norm can be sanctioned in the sense that he can reach a state of negative desirability: perhaps not because of a re-action of a normative agent, but because it enters a negative mood (e.g. shame or fear), or because the entire community he lives within comes to play the role of the normative agent, making him an outcast.

The structure of this paper is the following: in the next section, the agent model is described and a formal definition of obligation is presented; in Sections 3 and 3.1, a detailed example is examined; Discussion, comparison with related work and conclusions end the paper.

2 The Definition of Obligation

According to the observations made in the introduction, we propose that an obligation be defined as follows:

an obligation holds when there is an agent A, the normative *agent, who has a goal that another (or more than one) agent B, the* bearer *agent, satisfy a goal G and who, in case he knows that the agent B has not adopted the goal G, can decide to perform an action Act which (negatively) affects some aspect of the world which (presumably) interests B. Both agents know these facts.*[14]

[13] Sometimes, works on deontic logic use the notion of obligation for modeling both kinds of assertions.

[14] This generalization has an obvious limitation: it is implicit that the normative agent has chosen not to enforce G by means other than the sanction. More importantly, it hides the source of the goal of the normative agent (that the bearer satisfy G); in particular, this goal could come from another obligation, i.e. the obligation the normative agent has towards the sovereign (in case the sovereign and the normative agent are distinct).

Unlike what appears at first sight, this definition covers not only 'institutional' cases, but also other situations like obligations in dialog (see [48] and [8]), which share the characteristic that new goals are acquired as a consequence of social inputs. Moreover, also inner rewards and punishments deriving from moral obligations can be considered.

Since this definition of norms does not introduce further propositional attitudes like *being obliged*, it allows current models of BDI (Beliefs, Desires and Intentions) agents to deal with obligations, since they are able to manage intentions, to take into account the goals of other agents and their behavior, to devise plans for satisfying goals, and to compare the alternative plans according to their preferences. We assume that BDI agents are able to carry out, in a reactive way, their intentions, unless they are already satisfied or are impossible to achieve or have become not relevant; they decide which intention to pursue by means of a plan, starting from a set of goals; they are able to take this decision by comparing different alternatives to achieve the goals on the basis of their *preferences*.

Furthermore, a BDI agent is able to take into account the goals of other agents; this, according to [18], is one of the key capabilities for an agent to be social: social agents must be capable to consider the goals of other agents, and to have attitudes towards those goals, that is, to *adopt* those goals (i.e., "having a state of affairs as a goal *because* another agent has the same state as a goal");[15] moreover, *goal adoption* is at the basis of the definition of cooperation among agents in [6] and [5], and of dialogical interaction in [3].

As it is obvious, the role of goals is to direct the planning process; but the fact that a goal is provided as an input to the planner does not assure that a plan for achieving it will be chosen: in fact, the planner builds all the plans which possibly enable the agent to achieve one of his many goals; among them, just one plan will be chosen for driving the next action of the agent (this chosen plan will be called *intention* in the following). This choice is made on the basis of the agent's preferences. This means that many of the agent's goals usually remain unsatisfied (waiting perhaps for future actions). So, it is possible that among these unsatisfied goals there are also the ones which represent needs or desires of other agents, and in particular the ones (originated by norms and obligations) proper of the normative agents.

It could be observed that the term *goal* does not seem to suit the application of the norm, since the norm usually involves something which is not *per se* advantageous for the bearer agent. However, in general, goals which are *instrumental* for an agent also exist. These are goals that are associated to intermediate steps, in the process of satisfying the agent's preferences, such as, for instance, the subgoals required for achieving the preconditions of an action which satisfies the main goals of the agent. Goals associated with norms are just a special case of these instrumental goals.

In fact, in case of obligations, the *normative* agent A wants [16] that the *bearer* of the obligation adopt the goal G concerning the obligation. Moreover, the *bearer* B

[15] As explained below, adopting a goal does not mean that it will be satisfied or even become an intention of the adopter: he just considers the adopted goal as a possible further objective besides his own goals.

[16] Or, in the more general case the *sovereign* agent wants it.

knows the reaction of agent A if he does not adopt G; the resulting state can be less useful for B, so the goal G is really an instrumental goal for B (if he wants to achieve a state of affairs desirable to him).

So, the bearer agent must be able to foresee the reaction of the *normative* agent (both in case he fulfills the obligation and in case he doesn't). This ability of *anticipatory coordination* is another fundamental feature of social agents, according to [18]. In the field of BDI agents some proposals for this form of reasoning are already present. [40] introduced the notion of "anticipation feedback loop", [30] the "recursive modeling" of agents and [6] a planning framework for anticipatory coordination.

Since we would like to let the agent free not to fulfill an obligation, we need some mechanism for enabling him to choose among the various alternatives. As in all the proposals mentioned above, we exploit the notion of *utility* developed in *Decision Theory*, in order to enable the agent to choose the best alternative for him (the best, depending on the reaction of the partner). Utility is the formalization of the notion of *preference* of persons: therefore, it is possible to express the fact that the reaction of the *normative* agent leads to a less preferred state for the agent, together with the fact that the agent achieves some utility by satisfying his own goals.

2.1 The Agent Model

In this section, we introduce the definition of agent. When in the following we refer to *states*, we mean sets of attributes which are used to describe the world, together with a probability distribution over the values of the attributes at a certain moment. Probabilities enable us to represent the uncertainty of an agent about the current situation. Analogously, probability distributions will be associated to the effects of actions in order to express the fact that an agent is not sure about how the world changes after the execution of an action (either by himself or by someone else).

An agent C is a 6-tuple $\{IB,\ CG,\ f,\ L,\ KP,\ CI\}$ where:

- IB are the agent beliefs (including beliefs about other agents and beliefs about the current state of the world);
- CG is the set of current goals of C;
- f is the utility function of C (a function from states to real numbers); it is used to evaluate the outcomes of C's actions. f applies to states expressed as sets of attributes;[17]
- L is a set of tuples representing the obligations known by C of which he is the *bearer* (they will be described in the next Section);
- KP is the set of actions which C knows (*action recipes* [15]);
- CI is the Current Intention to execute a plan (a newly planned plan or the remaining part of the plan C is currently executing). CI is selected within the set CP of candidate plans produced by the planner.

[17] f is based on the theory of *multi-attribute utility functions* developed by [37]. For a discussion on the usefulness and restrictions of multi-attribute functions see [34].

f is based on a set of attributes, each of which is associated with a utility value, and, by means of a combination function, produces the overall desirability of a state on the basis of its description; so, it is clear that just the actions including effects that involve some attributes appearing in f can affect the evaluation of the state resulting from the execution of the action. The action may affect the utility of the resulting state in a positive or negative way; in particular, a decrease in the utility values is used for representing the costs of executing the action in terms of time and resource consumption.

The process of intention formation slightly modifies a major property of intentions (according to [22]), i.e., their $persistency$. In fact, from a computational point of view, the planner takes always into account the current intention (i.e., the previously chosen plan), by continuously evaluating what remains to be done, unless new information makes it believe that the intended goal has already been achieved, or it is not worth being achieved any more (the basic definition of persistence). But the presence of utilities can lead the planner to believe that a different (totally new) plan can enable the agent to reach higher utility. So, the previous plan can be discarded or modified.

Since the planning framework has been described elsewhere ([6], [5]), in this paper we focus on the criteria which determine the action to execute among the ones which may be executed in the current situation. In the same way we do not present here the agent architecture for the reactive execution of plans described in [6] and [5]: in [6], a hierarchical decision theoretic planner is employed which is inspired to the DRIPS planner [34].

The planner takes as input goals consisting in states or actions (*state goals* and *action goals* in [18]'s terminology): in case the goal is a state, it is considered as a state to be achieved, so that the planner must find all actions which can contribute to achieving the state; in case the goal is an action, the planner assumes that it is a complex action which has to be executed, so that its (easier) task is to find all possible decompositions of (i.e., ways to carry out) the task and to choose the best of them. The latter activity is called *refinement* of the action.

The set CP is produced by the planner starting from the initial state S, and inspecting the KP to find all the action recipes which have among their effects a goal in CG and the recipes which refer to (expand) an action in CG. Then, on the basis of the utility function f, the possible alternatives are examined and the best one (P) is chosen: it becomes the current intention of the agent. The best plan is the one which maximizes the utility:

$$P = argmax_{\{P_i \in CP\}} f(P_i(S))$$

where $P_i(S)$ is the state resulting from the execution of the plan P_i in the state S.

In case of probabilistic effects and uncertain states, $P_i(S)$ is a set of pairs $(S_{i,j}, p_{i,j})$, where $p_{i,j}$ is the probability that $S_{i,j}$ is reached after the execution of P_i. In this case, the best plan is the one which maximises the *expected* utility, as it is usual in Decision Theory:

$$P = argmax_{\{P_i \in CP\}} \sum_{(S_{i,j}, p_{i,j}) \in P_i(S)} p_{i,j} * f(S_{i,j})$$

As discussed above, we have shown [6] that in the context of cooperation among agents it is not sufficient to take into account the resulting state $P_i(S)$, but it is also necessary to consider the possible subsequent behavior of the other agents starting from $P_i(S)$. For instance, in a cooperative setting, it may happen that a state very positive for the agent endangers the activity of the partners, so that the overall (group) goal is harder to achieve. Our solution has been to base the evaluation not on $P_i(S)$, but on the states achievable from the partners starting from $P_i(S)$ (a kind of one-level lookahead in the spirit of min-max search).

So, in presence of anticipatory coordination, C will select the plan P in the set of candidate plans CP such that:

$$P = argmax_{\{P_i \in CP\}} f(P_{P_i}^{best_N}(P_i(S)))$$

where $P_{P_i}^{best_N}$ is the plan that (according to C's beliefs) will be selected by N when C executes P_i in state S.

Also in presence of anticipatory coordination, we must take into account actions which have non-deterministic outcomes. In this case, $P_i(S)$ is a set of states with associated probabilities.
When N plans her reaction, she will be in a specific state of $P_i(S)$ (since C will have already executed the action he chose). Therefore, C has to simulate N's reaction in each of these states. In this situation, $P_{P_i}^{best_N}$ will be a set of (state, probability, plan) tuples (the probability is the one of the state in $P_i(S)$ from which the associated plan has been planned); the above formula must be modified in:[18]

$$P = argmax_{\{P_i \in CP\}} \sum_{(S_{i,j}, p_{i,j}, P_{i,j}^N) \in P_{P_i}^{best_N}} p_{i,j} * f(P_{i,j}^N(S_{i,j}))$$

2.2 Formal Definition of Obligations

According to the model outlined above, and to the last two formulae, C should foresee the possible reactions of N. However, there must occur some sensing action performed by N to detect the violation. If C assumes that N will choose to execute this action and that the action succeeds - so that N knows that a violation occurred - he must try to imagine which action N will do next. Although it is possible that N sanctions him, C should take into account that N has to balance this possibility against other alternatives. So, C must reason about the motivations of N for executing the sanction. In fact, C considers N an *autonomous* agent, whose behavior is driven by her beliefs, goals and intentions.

Since N is an agent (i.e., a tuple $\{IB^N,\ G^N,\ f^N,\ L^N,\ KP^N,\ CI^N\}$), her model includes a utility function. So, if the attributes appearing in the effects of the sanction have the greatest utility value for N, compared to doing nothing or achieving some other goal, then she can select the sanction as her preferred action. This would mean that N must gain an advantage if the violation of the norm is sanctioned.

[18] For the details of the planning algorithm see [5].

But it is also possible that the sanction does not provide N with any direct personal utility. For instance, there is no utility for a policeman to sanction the breaking of a norm. In this case, the execution of the sanction R by N may be due to the existence of another norm, where the policeman acts as the *bearer* and the local administration acts as the *normative* agent. In other words, it is a duty of the policeman to sanction a driver who parked outside the allowed areas: this is a duty established by the administration for which the policeman works and a sanction should be applied to the policeman in case he does not respect the norm.

This opens the possibility of a *regressio ad infinitum*: who monitors the *normative* agent? Also in this case, our model does not pose any restriction for preventing this possibility. As a matter of fact, also complex normative systems as constitutional states do not succeed in ensuring a complete control on the chain of *normative* agents who have to monitor and sanction other *normative* agents, as corruption scandals at the higher level of the legal hierarchy prove.

The need of having some knowledge about the *normative* agent's *operative code*, that is the utility function and goals is a strong requirement. However, some defaults can be applied. So that a set of definitions for the 'standard policeman', or the 'standard web agent' can be used. But in some cases, more detailed user models can be available as the 'policeman I meet everyday in front of my office', the 'notoriously corruptable policeman', or the 'softbot controlling an overworked web site'[19].

If L (the set of obligations) is not empty, then the planning phase and the selection process of the best alternative must be modified, since C should examine whether to satisfy (or respect) the obligations in L (for the sake of brevity, in the following description, only one obligation will be assumed). So, the planning phase must be given as input an extended set of goals, CG', which consists in the union of CG (including the current intention) together with the goal(s) appearing in the obligation(s) in L.

As we said in Section 2.1, obligations can be both action and states, and in fact the planner takes as input both state and action goals. The planner we adopted for implementing the model is DRIPS, which is based on action expansion and refinement. So, in case of state-goals, the higher level actions which can achieve a goal are found and then are passed as the real input to the planner; on the other hand, in case of action goals, these are simply fed to the planner as the real input.

In case the content of the obligation is the negation of an action or, equivalently, a state which must not be achieved (e.g., 'the fuel tank must not be empty'), a different strategy is applied: if an action forbidden by the obligation (or which makes true the state forbidden by the obligation) is inserted in a plan during the planning phase, it will be detected during the simulation of the *normative* agent behavior, and used to

[19] A possible alternative, suggested by [30], is to model the *normative* agent by assigning a probability to the sanction which she would face if she disregards her duty, instead of modeling explicitly this sanction as the action of an autonomous agent (i.e., the administration). For instance, a car driver does not need to wonder why a policeman sanctions him when the car speed is above the maximum allowed; it is enough that he believe that when a policemen notices a car running at a speed just five miles above the maximum, there is a 0.3 probability that he sanctions the driver, while if the speed is 30 miles above the maximum, the probability is 0.95. Of course, this simplifies the matter, but does not enable the agent to exploit in full the foreseen behavior of the *normative* agent, as our model does.

predict the possible application of the sanction.

Formally, in the L component of an agent, an obligation Ω is represented as a 4-tuple $\{B,\ N,\ O,\ R\}$ where:

- B is an agent who is called the *bearer* of the obligation,
- N is an agent called the *normative* agent,
- O is the *content* of the obligation, i.e., the state or action goal which N wants to be adopted and satisfied by B,
- R is an action (called *sanction*) which N will presumably bring about in case he detects a violation of the obligation.

When O directly provides N with some advantage, she will be called the *counter-party*, which benefits from what the obligation is about, as in the beneficiary theory of rights ([36]).

As stated above, the content of the obligation Ω, i.e. O, is not necessarily a state (e.g., "the tank must be full"), but it can be also an action where C is the agent (e.g., "the agent should send the credit card number") or not (e.g., "The owner of the agent must submit his personal certificate"). But it can also be the prescription of not executing an action ("you should not send huge files by mail") or, almost equivalently, the prescription not to achieve a state ("the tank must not be empty"). Notice that not achieving a state means that the state is currently false (the tank is not empty now); so, this is equivalent to maintaining true a state that is currently true (maintenance goals: "the door must stay closed").[20]

2.2.1 Sanctions

As the sociologist Giddens notices, sanctions can be *positive* (a reward offered when the obligation is respected) or *negative* (a punishment for a behavior in contrast with the obligation)[29]. That is, the effect of the sanction can be to affect the world in a way that increases or decreases the utility of the *bearer* of the obligation, in case he, respectively, respects or violates the norm. In the rest of the paper we will refer only to the case of *negative* sanctions; but our model can be applied also to the case of *positive* sanctions, with the difference that the *normative* agents will 'sanction' the *bearer* when the norm is respected. But a sanction can also have side effects which affect the *normative* agent's utility, since they are actions to be executed. These effects can be negative (the consumption of resources, but also worsening the social

[20] While the difference between requiring to execute or to avoid to execute an action seems to be clear enough, the difference between a positive and a negative state is more dubious. In fact, 'open' and 'not closed', are clearly equivalent. We assume that the difference depends on the predicates describing the effects of the actions. A predicate appearing as an effect is considered 'positive', while a predicate not appearing in any effect is considered 'negative'. This is useful to keep apart predicates which can act as goals from predicates which cannot. The former can be used as input to the planner, while the latter cannot. So, if an action 'to-open' has, as its effect 'open', then both 'open' and 'not closed' are taken as positive state specifications (assuming that suitable meaning postulates link 'open' and 'closed').

relationship with the violator) as well as positive: some time ago, in Italy, policemen were rewarded for every fine they assigned to a driver.[21]

As we have seen, a *negative* sanction is executed only in case the *normative* agent believes that there is a violation: in order to establish whether a violation occurred, the *normative* agent has to perform a monitoring action. For this reason, in our model, the sanction R is represented by a complex action consisting of the monitoring action followed by the very action of affecting the world in a way that is relevant (in one way or the other) for the *bearer* agent.[22]

We have chosen to insert the monitoring action in the definition of a norm although it would have been sufficient that the sanctioning act had as a precondition the belief of the normative agent that a violation has possibly occurred. But it must be observed that in formal contexts, say laws, not all ways for knowing that a violation occurred are acceptable. For example, the italian police is not allowed to tape phone calls without the authorisation of an attorney. Illegal recordings do not count as evidence in trials to determine that a violation actually occurred. Therefore, the specification of a norm includes the specification of the possible means for checking violations. In case every means is allowed, the monitoring action will consists of a general action which subsumes all the possible checking actions for a given type of violation.

Finally, it must be noted that the sanction itself can establish another obligation for the violating agent. For example, a policeman can put the obligation of paying a fine for having parked in a no parking area, where the *normative* agent's role is taken by some administration. In turn, the failure to respect the deadline for paying the fine results in another obligation of being brought to trial before another kind of *normative* agent, a judge. And so on in a *crescendo* of more and more negative payoffs for the agent.[23]

However, there is no risk of a regression to infinitum, since this series of obligations usually ends in a sanction which does not pose a new obligation, as Goffman notices:

> still steeper penalties should their judgement be rejected, and still deeper penalties should this, in turn, be rejected and so forth, eventually culminating presumably in physically coerced rulings.[24]

2.2.2 Determining Violations and Applying Sanctions

In our multi-agent framework the behavior of the partners of an agent C is influenced by the actions of C, just insofar the effects of his actions can be noticed by the partners.

[21] With the obvious social consequences this fact implies, since many policemen started sanctioning any dubious violation.

[22] In general, the checking action need not follow the breaking of the norm. It can be an action that the normative agent executes periodically, independently of the actions that the bearer executes. This does not affect the basic model.

[23] The consequences of chaining of obligations, even if not infinite, is clear, at least in Italy: since the deadline for respecting each obligation can be long it is possible that at some point, for example after a long appeal, the violation can not be punished anymore since the sanction must be applied within a certain time interval (prescription).

[24] [33], p.117.

In other words, any action can have a side-effect on the partners' behavior just in case they are able to detect that something relevant for them has happened. This means that when C tries to imagine how another agent N will react to his own action, he must reason not on $P(S)^C$ (the resulting state, how C sees it), but on $P(S)^N_C$, i.e. on the state that (according to C's knowledge) N will see.

This is particularly relevant in the case of obligations. In fact, the N counterparty of an agent C who is the *bearer* of an obligation cannot be assumed to become immediately acquainted with the (possible) violation of the obligation. According to C's knowledge, there is some probability that this happens: in fact, C is assumed to know that N has some actions available to check the fulfillment of O, that these actions may fail, and that just in case of their success, N will consider (not necessarily decide) to apply the sanction.

Finally, any agent knows that actions may fail; also the action of applying the sanction may fail. So, even if the violation has been detected, and N has decided to apply the sanction (which he would not do, in case the cost of applying it is greater than the gained utility), the sanctioning action may fail. C must (or, at least, we claim that rational agents do) weigh all of these possibilities when he chooses the best way of acting.

2.3 The Behavior of the Normative Agent

As stated above, the monitor and sanction action leads to an actual punishment only if the 'monitor' part enables the *normative* agent to *know* (or at least to believe) that there is a violation. In the general case, the *normative* agent is (or may be) unaware of the violation [25] and she has to decide whether to monitor and (possibly) sanction him or to do other things, according to other goals. In order to make her choice, she has to compare the different utilities of her alternatives. Also sanctioning actual violations may provide her with an utility: but why should she choose the monitor and sanction action when she is not aware of the violation? That is, if the effect of this action depends on the actual state of the world which is not known by the agent, how can she evaluate the utility? What is important for this discussion is the role of the action during planning and not during execution by the *normative* agent. During execution, the world can be in one and only one state, and the monitoring action will reflect it. But during the planning process, the *normative* agent has no access to the actual world. The evaluation of the result of an action and of its utility must be made entirely from his belief space. And the *normative* agent may either not know anything about the violation or she may have some *a priori* idea of what happened. We have to find a framework for representing both situations.

In order to deal with this problem, we resort to the probabilistic framework exploited in [34]. [34] distinguishes two kinds of nondeterminism: first, a world state (as well as an action effect) may not be known for sure, but an agent may know a probabilistic distribution of the values of the attributes which describe the world. E.g.,

[25] The monitoring action is a *sensing* action, that is an action which makes the agent's beliefs correspond (at least with a certain probability) to the current situation of the world; sensing involves complex forms of reasoning that go beyond the scope of this paper (see, e.g., [39], [25]).

while at work, from our cubicle, we do not know whether it is raining or not, but the weather forecast said that there is a .3 probability of a sunny day. Second, an agent can have no idea of the probability distribution on those values (this does not mean that he believes they are equally likely): the agent is *uncertain* about the actual state of the world. Coming back to the example, if we haven't listened to the weather forecast we have no idea at all if it is raining or it is a sunny day.

Our approach follows the Theory of Evidence ([43]), where the certainty about the world is represented by the so-called Basic Probability Assignment (BPA). In a BPA, a probability is assigned to each subset of the universe of possible values of a random variable. As usual, the sum of the BPA values is 1. The assignment of the total mass of probability (1) to the whole universe is intended to mean total ignorance[26]. Although this is just a special case of BPA, it is useful to obtain a concise representation of the amount of knowledge an agent has about the current situation. In fact, we assume that just three situation may arise: total ignorance (BPA assigns 1 to the whole universe), knowledge about a probability distribution (BPA assigns values different from 0 just to singletons), certainty about the value of a given attribute (just one singleton has the BPA value 1). In this way, since what is of interest is just the utility of the outcomes, it is possible to focus the attention on the worst and best cases. So, the utility of the different states is evaluated, and the situation of ignorance is represented as a pair whose first element is the state for which the utility is worst and the second element is the state for which the utility is best.[27] The utility of an uncertain state is represented as a pair as well, with the utility of the worst and best case scenario as first and second element. It must be observed that the utility pair can be interpreted as an interval, since the actual utility value necessarily falls within the bounds appearing in the pair.

3 A Sample Scenario

Consider (somewhere in the near future) two (BDI) agents working in internet with different roles. The first agent is an electronic spider that, for the sake of his owner, has the goal to gather information in a short time by accessing different sites, and retrieving documents (possibly paying for them). He can also have preferences about social goals, such as the relationships with other agents.

The second agent controls a given website: her task is to prevent some agents from accessing the site, to assure that all agents pay for the provided data, and to avoid possible *denial of service* inconvenients, which can degrade the interaction with other users. Her concerns are that a spider agent can access the site sending requests for documents at a high rate lowering the response time of the server with other users; or he can use an expired credit card; or he can send to the website a huge file so that the performance of the website is endangered.

[26] This must be contrasted to an assignment of values different from 0 just to singletons, which corresponds to a standard probability assignment, and which is considered as expressing some precise knowledge about the world, i.e. how probable the different outcomes are.

[27] This solution relies on a number of assumptions on the structure of preferences of agents, as shown in [34] or in classical works as [37].

So far as the access rate to the site is concerned, since the site agent does not have the computational resources to check each request and she has other tasks to perform (as clearing payments), she poses some obligations to the (artificial) agents accessing the site. However, posing an obligation is not enough, since this might be violated. Hence, the spider agent must be provided with an *incentive* for respecting the norm, even if it is not advantageous for him to do so. Therefore, the site agent states that if any customer overcomes a given access rate, she will sanction him: subsequent requests coming from that IP address will be rejected. The site agent can check the violations by sampling the log file of the server. But the checking action has a statistical character and it may fail: for example, it can discover only 30% of the violations; hence the violating agent may be neither identified nor sanctioned.

The spider agent, when becomes aware of the obligation, has to take it into account in his decision process[28]. When he selects the access rate, he has to predict the reaction of the site agent: the higher the rate, the greater the utility for the spider agent; but in case the access rate is beyond the norm, the resulting state can be worse because of the reaction of the site agent: a prohibition of further access may greatly decrease the utility of such a state. So, the spider agent has to find a trade-off between the utility of gathering information rapidly and the possibility of future access.

Even worse, the spider agent can act in order to increase the probability of not being discovered. For example, he can send part of his requests via one (or more) proxy server: this is a lengthy process but it allows the agent to send, in a certain amount of time, more requests than it is permitted to do, without disclosing the information that they come from the same IP address; in fact, even if we assume that he and the proxy server do not violate the norm, the spider agent can send at least twice as many requests as the allowed amount.

We represent this scenario by means of a number of attributes. In particular:

- `info = x`; the amount of (valuable) information gathered by the spider agent.
- `ip = y`; the IP address used by spider agent: if it sends requests via a proxy server, this address will be different from the spider's one.
- `over(y)`. This predicate is true iff the agent with IP address y has accessed too much information in a given time, so that he has violated the norm on fair use of the site. The spider agent obviously knows if `over(y)` is true; in contrast, the site agent has to examine the log files to find that a violation occurred by checking the timestamps of the requests from the IP address.
- `t = z` specifies that the current time is z; time will be considered also as a resource so that the longer the time, the lower the advantage for the agents.
- `working`. This is a predicate which is true iff the site is functioning (it did not suffer from a denial of service attack).
- `sanct(y)`. True in case the agent whose IP address is y has been sanctioned; for simplicity, we assume that the value of this attribute is known by both agents without any effort.

[28] Note that, accepting the norm depends just on the belief of the spider agent that the site agent has really the power of applying the sanction. See [19] for a general discussion.

The situation of uncertainty of the *normative* agent can be represented in this way: if the site agent does not know whether an agent `y1` has used a forbidden access rate to the site, then she does not have any belief about the binary predicate `over(y1)`, so its two possible values (T and F) will appear in the pair as the limits of the interval.

The spider agent has at its disposal in KP some actions:

1. `HighRateRequest`: the spider agent sends HTTP requests to the web site at the highest possible rate. The effect of this action is to gather a wide amount of information (say `info=x1`) in a short time (t_x units of time); however, it violates the norm contained in L, so that it has among its effects `over(y1)`. Recall that this condition concerns C's beliefs and the real world but not the normative agent's beliefs space, which is not affected by C's action.
2. `LowRateRequest`: the spider agent sends HTTP requests to the web site at a rate which respects the condition specified in L. It allows him to gather a limited amount of information (`info=x2, x2<<x1`) in the same time as the above action, but it does not affect the `over(y1)` attribute (which remains false).
3. `ViaProxyRequest`: the spider agent sends HTTP requests at almost the forbidden rate directly to the web site, but it sends the same amount of requests via a proxy agent: in this way, in the log file of the web site half C's requests appears as bearing another IP address than C's. C gathers almost the same amount of information as with `HighRateRequests`, but in a little longer time (due to the proxy delay). This action does not violate the letter of the norm beared by C.
4. `Attack`: the spider agent sends HTTP requests at a high rate directly to the web site, and, afterwards, he tries a denial of service attack which makes the site not work anymore (`not working`: assume this attribute is immediately known to both agents).

The site agent is N and we assume that each visitor is informed about the obligation to keep the access rate low. N has at his disposal three actions:

1. `Monitor&sanction`, that we discuss below.
2. `Repair`: in case the site is not working, the effect is to repair it; while repairing it (and afterwards), it is not possible to check the violations (e.g. the log files are deleted); hence, `not sanct(Y)` and `working`.
3. `DoNothing`: doing nothing and preserving resources.

As an example of the full representation of an action, in Figure 1, we have depicted the sanctioning action to be performed by the *normative* agent[29]; it is composed of a sequence of two elementary actions, the monitoring action and the (conditional) action of sanctioning the violator. Besides the actions, their (conditional) effects are depicted in a programming language style. The `bel(X, over(Y))` notation is used to represent the fact that the attribute concerns the view that the agent has about the predicate `over(Y)`; this must be contrasted with `over(Y)`, which refers to the actual world

[29] The value 1 corresponds to logical truth, and 0 to falsity.

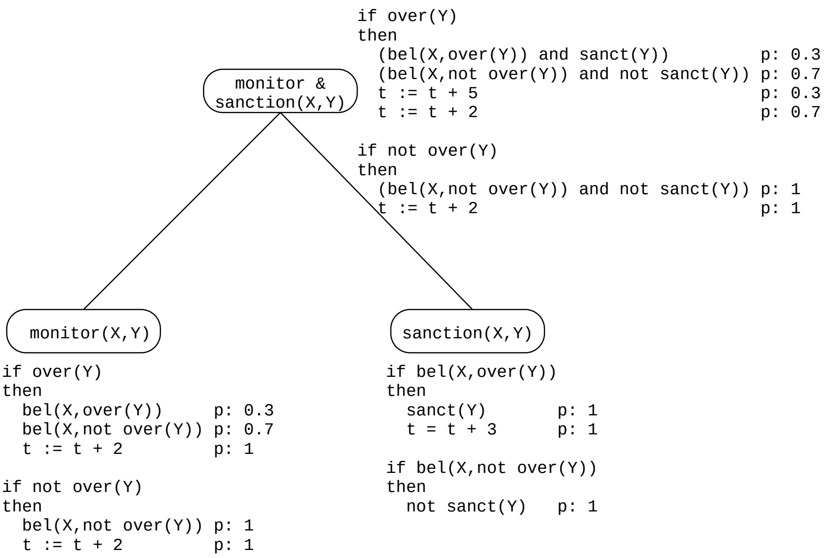

Figure 1: Definition of the `Monitor&Sanction` action.

(or, during the simulation, the beliefs of the *bearer* agent). The monitoring action says that the agent can *sense* the world and discover whether a violation of the access rate has occurred with a .3 probability; in contrast, if no violation happened, he knows this fact for sure. The second action, `sanction`, has a deterministic effect conditioned to the fact that the agent knows that a violation happened. The effects besides the top level complex action `Monitor&Sanction` summarize the effects of its decomposition via a sequential abstraction (see [35] for details). In particular, these effects say that, if there is a violation, there is a .3 chance that the agent becomes aware of this fact and the violator is sanctioned. The 'monitor' action takes 2 time units, and the possible sanctioning takes 3 time units.

This rather complex representation must be combined with the fact that this action, during the planning phase, must be evaluated starting from a world which can contain uncertainty and probability from the *normative* agent's point of view.

As a starting point, we show in Figure 2(a), the notation we adopt to describe the (probabilistic) effect of the execution of the action `Monitor&Sanction`. The hypothesis is that a violation actually occurred: in such a case, it is sufficient to apply the definition appearing in the upper right corner of Figure 1, using the `over(Y)` branch. In Figure 2(b), the action is evaluated in a situation where the *normative* agent has some idea about the probability of a violation. The .8 probability of a (suspected) violation makes the *normative* agent eager to find out and sanction it. A different distribution could lead to optimistic agents who do not suspect of violations and decide not to monitor for them. In Figure 2(c), it is depicted the situation where the *normative* agent has no idea whether a violation occurred: no precise probability can be associated with the possible outcomes. In fact the arcs are labelled with pairs: the probability of finding a violation ranges from 0 (in case no actual violation occurred) to .3 (in case there was a violation, this is the expected probability of finding it); the probability of not finding it, on the contrary, ranges from 0 to 1. Note that the first element of the pair (0 in both cases) is associated with the first element of the pair characterizing the ignorance about the actual state (F, i.e. no violation) and the same for the second element. So, the uncertainty about the initial situation is propagated to the predicted outcomes of the action.

3.1 Obligations: An Extended Scenario

The scenario depicted above is extended by including in the L component of the agent a norm specifying that the access rate must be below some threshold. This condition is expressed via the predicate `over(Y)`: it is false if the norm is respected; it is true otherwise.

Sanctioning violations is a goal for N, which gains an utility from this fact, even if the time consumed in acting has a cost for him; moreover, N has an advantage in maintaining the site working, a goal which conflicts with the first one, since, as stated above, it is not possible to make both of them true.

The tuple representing the norm in L^C is:

$\{C$,N,`not over(`C`)`, `Monitor&Sanction(`N`)`$\}$

`not over(`C`)` is the negation of a predicate and is currently true; so, C need not care about finding actions that make it true, but any plan he devises which makes `over(`C`)` true will make C subject to the sanction. This will happen just in case N realizes that a violation has occurred.

In the initial situation $S0^C$ (from C's point of view), C has gathered no information, he has the IP address $y1$, he has not overcome the allowed access rate and the time is t_0 (see Figure 3). The four alternatives available to C lead to the states $S1^C$, $S2^C$, $S3^C$, and $S4^C$. Now, in order to make the right choice, C must imagine the possible reactions by N. A distinctive feature of this example is that, after three of the four possible actions by C, N will have no perception of a difference in the state. Indeed, the Monitor part of `Monitor&Sanction` has the purpose of enabling N to ascertain which state she is in.

So, inside the large box labelled as *N's point of view*, the four states seen by C produce just two states SX^N_C, and $S4^N_C$. Reasoning on the first of them, N must choose between `Monitor&Sanction` and `DoNothing`. We assume that N has no knowledge about the attribute `over(Y)`, which is expressed by the pair (T F)[30]. Let's suppose that the utility function of N is defined as follows:

$f(t)$ = (-2) * t
f(`sanct(X)`) = 100
f(`not working`) = -200

[30] Alternatively, N could know an *a priori* probability distribution over the two values of the `over` attribute, as discussed in Section 2.3.

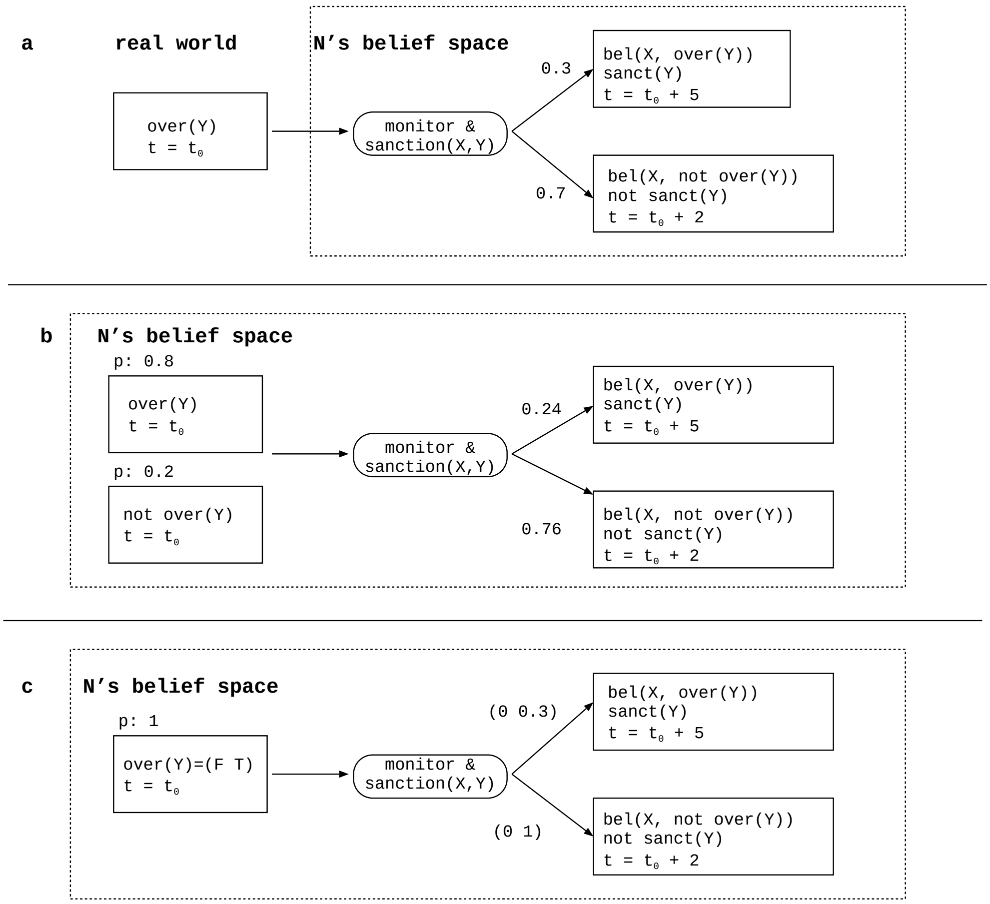

Figure 2: Effects of the execution of the `Monitor&Sanction` action in different situations.

so, the utility of states achievable by N is obtained as the sum of the three different components[31]:

$f(SX1'^N_C)$: ((-2) * t_0 - 30) + 100 = 70
$f(SX1''^N_C)$: ((-2) * t_0 - 30) = - 30
$f(SX2^N_C)$: ((-2) * t_0 - 20) = - 20

on the contrary, if N sees state $S4^N_C$, its only possible action is `repair` which can lead in one of the two states:

$f(S43'^N_C)$: ((-2) * t_0 - 50) = - 50

[31] In order to simplify the expressions, in the following formulas we assume t_0 = 0; this does not affect in any way the various choices. Since `sanct(X)` and `not working` are binary predicates, when they are false, the associated utility is assumed to be zero

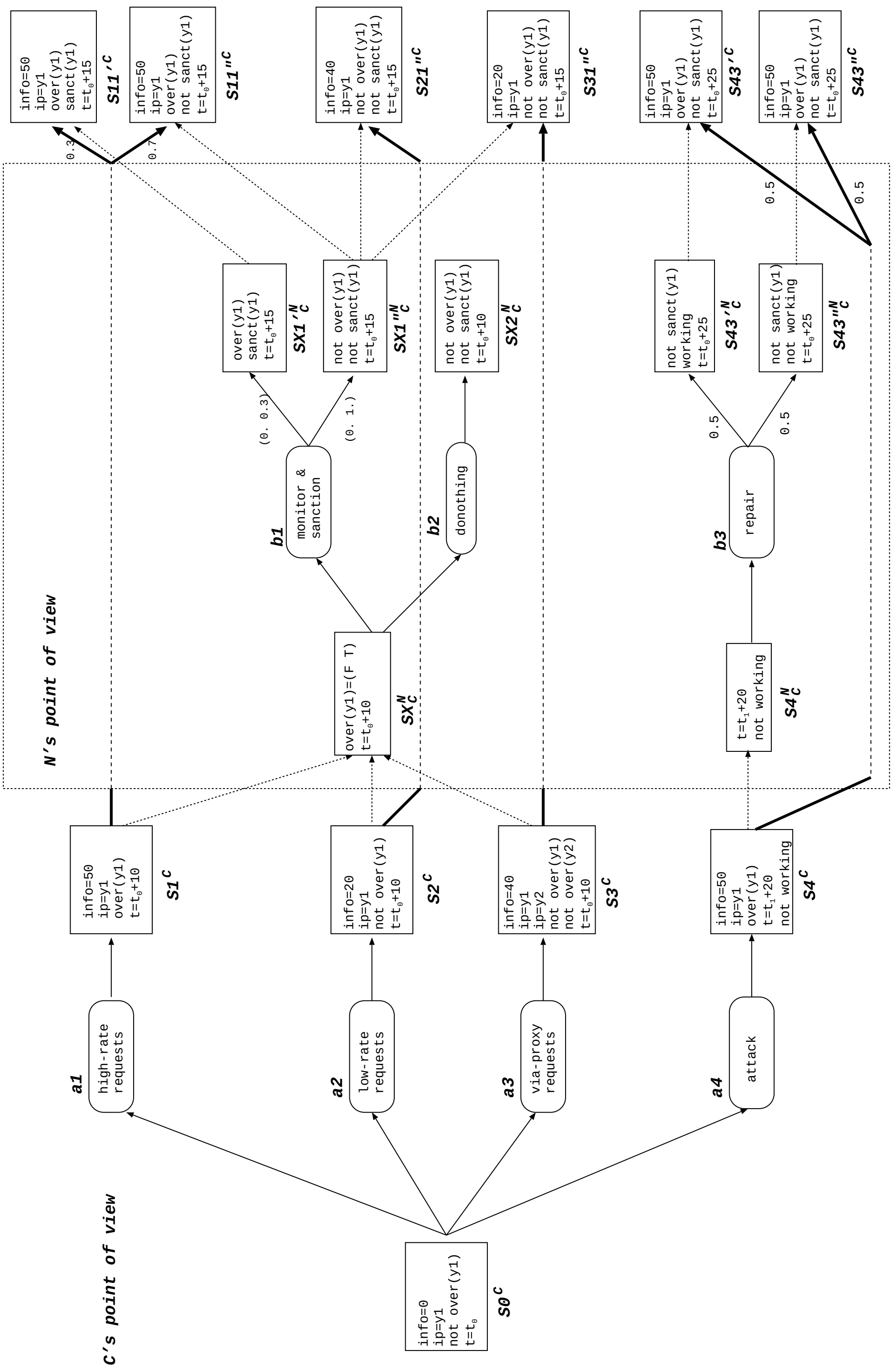

Figure 3: Modelling of Anticipatory Coordination

$f(S43"^N_C)$: ((-2) * t_0 - 50) - 200 = - 250

According to the reasoning mechanism based on the theory of evidence, N, in case she finds herself in state SX^N_C, will associate with the action **b1** an expected utility in the interval [-30, 70], and with the action **b2** an utility of -20. Although N can have no certainty that the first action is better, notwithstanding she can choose by averaging the extreme values of the first interval. Since this gives for **b1** a value of 20, this is the action that N (according to C's simulation) will choose.

On the contrary if C executes `Attack`, no choice is available to N.

Now, C must reason on the actual states (according to his knowledge) resulting from N's chosen actions. Since it has been established that in absence of an attack, N will execute `Monitor&Sanction`, C has to apply `Monitor&Sanction` to his view of state SX^N_C, i.e. to the three different states $S1^C$, $S2^C$, and $S3^C$. Starting from $S1^C$ (where there has been a violation), C will conclude that the violation will be detected with a .3 probability and it will not be detected with a probability of .7, thus leading to states $S11'^C$ or $S11"C$. In state $S2^C$, no violation occurred, so that nothing can be detected (state $S21"^C$). With `ViaProxyRequest` no formal violation occurred, so the resulting state is $S31"^C$. On the other hand, the consequences of an `Attack` can be the states $S43'^C$ or $S43'^C$, according to the success of the `Repair` action.

Now, C must evaluate the utility of these states (taking into account the various probabilities, according to the last formula in section 2.1). We assume the following values for the utility function of C:

f(`info=K`) = K
$f(t)$ = (-2) * t
f(`sanct(X)`) = - 100
f(`not working`) = - 40

The utility of states achievable by C is[32]:

$f(S11'^C)$: ((-2) * t_0 - 20) + 50 - 100 = - 70
$f(S11"^C)$: ((-2) * t_0 - 20) + 50 = 30
$f(S21"^C)$: ((-2) * t_0 - 20) + 20 = 0
$f(S31"^C)$: ((-2) * t_0 - 20) + 40 = 20
$f(S43'^C)$: ((-2) * t_0 - 40) + 50 = 10
$f(S43"^C)$: ((-2) * t_0 - 40) + 50 - 40 = - 30

Finally, C can evaluate the expected utility associated with each of its four possible actions:

f(**a1**) = 0.3 * (- 70) + 0.7 * 30 = 0
f(**a2**) = 0
f(**a3**) = 20
f(**a4**) = 0.5 * 10 + 0.5 (- 30) = - 10

So, C will eventually choose `ViaProxyRequest` as the action to execute.

[32] Note that here, the value of t to be used is the one in the states resulting from C's actions, not the one in the states reached after N's moves; in fact, `info` is available to C before N's reaction.

4 Discussion

In this paper, we have described a computational model of reasoning about norms based on decision theory and anticipatory coordination. This description could have given the reader the impression that we claim that people (and autonomous agents) always base their behavior on this kind of reasoning. This is not so. Our claim is that *sometimes* reasoning on sanctions and on sanctioning agents is necessary and is actually done. But we agree that in most cases this degree of sophistication is not required. The idea is that a model which does not include the ability to reason in this way is fundamentally limited, since there are some behaviors it cannot account for. On the contrary, a complex model, as the one presented here, is *downward scalable*: this means that this same model can pair with the full reasoning mechanism also shortcuts, that may be used in more standardised situations. For instance, as observed in footnote 19, nothing prevents from associating with the normative agent a 'probability of sanction', a move that in standard situation can cut off the need for imagining what she will do. The same model can account for even more simplified situations, where the normative agent does not appear at all: if I drive too fast in the U.S.A., the probability of being sanctioned is higher than in Italy. But it is not required that I understand why it is so; I may simply choose my behavior according to this information. The model accounts for this very easily: it is sufficient to associate a probabilistic outcome with the action *drive fast*, without any need to complicate the model.

Another relevant objection to the approach presented here concerns the relationship between sanctions and moral behavior. This problem can be faced at two different levels: the indirect *social* negative utility of being sanctioned, and the blind acceptance of norms as a cultural attitude. The first level corresponds to the observation made in [19] that "incentives" such as sanctions or rewardings could not be enough for making agents respect norms. Although we agree on this, we must quote Goffman again:

> the meaning of rewards and sanctions is not in their intrinsic and substantial value, but in what they state on the moral status of the actor. Norms are almost always expressed in general terms, as if they applied to a specific event as a member of a class to which the rule must be applied. So, in any situation where the rule is supposed to apply, every deviation can give the impression that the subject is deviant with respect to the whole class of events[33]

Since the social image is (possibly) part of the utility of agents, sanctions have a double role as incentives: they directly affect the 'welfare' of the agent and, indirectly, his position in the society. This argument explains why humans are afraid of violations: the impact on the social image of the violator is amplified by the fact that even a single violation is interpreted by the 'society' as a sign of the agent's character. We are not able to capture this intuition in our model in a general way; but, in [2], we modeled this social phenomenon in a particular setting where obligations are

[33] [32], p.64.

involved: that is, in dialog. In particular, following [14], we modeled the choice of indirect polite speech acts, a very pervasive phenomenon in communication, as an effect of the goal of preventing this form of generalizations; in the case of speech acts the violation correspond to the use of a direct and unpolite speech act: the generalization is that the hearer interpret a single direct interaction as a sign of the general attitude of the speaker towards him.

We do not have very much to say about the second point, i.e. the attitude to respect norms. In a very speculative way, we can suppose that education brings individuals to associate with norm-abiding behaviors a positive utility, but this would deserve more in-depth studies. However, the model presented here should be amended, in order to associate with all breakings of norms a special parameter, to which a (more or less) high negative utility can be attached.

It is worth mentioning a second argument put forward in [19] about the weakness of incentives and sanctions, i.e the fact that they are not always applied: for example, burglars are punished only for 5% of the robberies. Given this very low probability why people do not all start robbing? Evidently, the not-to-rob norm is not respected only for the fear of the sanction.
Apart from the individual attitude towards the respect of norms mentioned above, we must observe that our model allows to show that the argument on the probability of sanction is not entirely correct. In fact, this argument is grounded on the idea that the probability of the sanction is the same for all violators. This idea depends, in turn, on a paradigm according to which an agent takes deliberations only based on a decision theoretic paradigm and not game-theoretic one: that is, the agent considers only the probability of the effects of his action, and the sanction is only one effect which is unlikely to happen and negative. Instead, we claim that our game theoretic view of norm abiding behavior enables us to correlate the probability of the sanction with the agent who violates the norm and his planning ability. In fact, the sanction is not a direct effect of a violation, but a possible effect of the reaction of a player of the game, i.e. the normative agent.
Hence, the probability of the sanction is related to the ability of the agent to violate the norm in order to achieve his own goal, while leaving the normative agent unaware of his behavior. In turn, as stated in Section 2.2, the evaluation of the possibility of avoiding sanction is related to the model of the normative agent we have: if we do not know exactly the monitoring process that leads to sanctions, we are not sure about the risk we are taking in violating the norms in a smart way. For example, robbery requires (presumably) skills, support, knowledge about how to sell stolen goods: this, in turn, implies the risk of blackmail by the receiver, etc. Not everyone has this know-how about robbery and acquiring a sufficient skill is too risky and too costly for most people.[34] If we look at other activities that require less know-how, like shoplifting in department stores, well, ask the managers for the statistics.

In contrast, another argument raised in [19] becomes even more striking in a game theoretic approach to normative reasoning; [19] notices that, in the acquisition of

[34] Our argument would be stronger if we knew that the sanctioned 5% of burglar would be composed of rookies.

norms it is important that an agent believes what the normative agent says about the sanction: nothing prevents, in fact, the normative agents from cheating and overstressing the probability and effect of a sanction. And, actually, not only parents but also the states with their laws resort to cheating for ensuring norms.[35] Since in our model the normative agent is considered a player of the game and his behavior is simulated by the bearer agent, the information that the bearer has about the normative agent are of paramount importance. It is relevant not only that she does not lie about the sanction and the monitoring process, but also that she intends to monitor and sanction violations and, most importantly, that she can do that. This set of beliefs necessary for the bearer to predict the sanction and to respect the norm are really similar to that proposed by [21] under the name of *trust*. The trust is a fundamental quality of the normative agent for ensuring that norms are respected by agents pursuing goals which (possibly) conflict with the norms. People would like trusted normative agents only if they will not have goals contrasting with norms and will believe that the respect of the norms by the community will improve the welfare.

Anyway, we have seen that our model, on the line of [30]'s one is scalable in both directions. We can account for the strategic reasoning of the *normative* agent, but we can also model *bearers* who disregard the strategic reasoning and model sanctions as side effects of their actions, independently of the autonomy of the *normative* agent.

5 Related Work

Since [52], *deontic logic* has been proposed as a formalism for reasoning about obligations, normative concepts and what "should be" (or happen) in the world. The main assumption in most proposals is that verbs as "ought", "should", etc., can be modeled in the same way as other modalities as necessity or belief by means of a possible world framework. Modal operators as O have been introduced in order to express formulas as Op which are true in a world w if the proposition p is true in in all the 'ideal' (possible) worlds which are accessible from w. The ideal worlds represent how the reality should be according to some normative system or preferences.

However, the aim of deontic logic is different from the way obligations are used in agent theories: the main goal of the former is to examine how obligations follow from each other and where paradoxes can arise (see [50]). In contrast, agent theories aim more at examining the relationship between intentions and obligations, i.e. how the agents decide or not to fulfill an obligation. In deontic logic terms, we do not believe that the T axiom ($Op \supset p$), which holds for the *necessity* and *knowledge* operators is fully appropriate here.

[35] For politicians, the strategy of overstressing the possibility of sanctioning and above all the willingness has also a second goal; in fact, they can let the real information about the impossibility or unwillingness of sanctioning be known to a subset of the bearers of the norm: in this way, their 'friends' will be able to violate the norm, if it is advantageous to them, being conscious that the risk is low, while other people will be compelled to follow the norm by their wrong beliefs about the risk of sanctioning. Probably, the latter class of people will appreciate also the effort of the politician to follow the greater good by strictly enforcing sanctions.

While deontic logic has devoted attention to the possibility of violating norms, less attention has been paid to the role of sanctions, even if in one of the first works about obligations, [1], they are reduced to the alethic modality of necessity via the idea of the occurrence of a sanction s:[36] $Op \equiv NEC(\neg p \supset s)$. After a long period where sanctions have received little attention, in recent work their importance is explicitly recognized:

> The threat of punishment might be taken into account when the agent designer considers building into his agent the capability of adhering [to obligations]. [...] When a rule is violated, and the violation is detected, a sanctioning act (or an act of recovery) is effectuated.[37]

In [27], deontic logic is applied in an agent framework for dealing with norms and conventions. This work explicitly models sanctions consequent to violations and relates the fulfillment of obligations to preferences. Moreover, recent works as [16] have addressed interesting issues in a deontic logic framework, as the management of norms in case of collective agency.

Finally, non modal approaches to deontic logic have been proposed, for example, by [45], where the temporal aspects of norms are considered.

In current deontic theories, the center of attention has only been the bearer of the norm. The remarkable exception [27] explicitly deals with sanctions, but does not model the agent who is in charge of monitoring violations and applying the sanction. In [49] the concept of norm is associated with that of a sanction occurring in case of violations, even if the sanction is still viewed as an event and not an action of an *autonomous* agent. In works from the field of economics, as [4], norms require that the violations are "often punished" with a sanction imposed as the result of the decision process of another agent. Finally, also in the area of AI and Law, norms require sanctions [47], [10].

For what concerns agent theories, the notion of obligation has been exploited for the goal of directing the behavior of agents; as an example, in [44] (as well as in similar approaches) there is a different view of obligations, as [38] has noticed: in [44] obligations are used for *regimenting* agents, that is, for assuring that they will behave in a certain way. Because of this goal, the actions of the agent repertoire are constrained by the norms, the axiom T is adopted for modeling obligations, and obligations are constrained to be consistent. In a similar way, [11] proposes to constrain the evaluation module for enforcing norms.

On the contrary, our approach leans towards another view of obligations, which is inspired to [24], where obligations can be violated, *normative* agents can be deceived in order to avoid sanctions, and the fulfillment is motivated by some instrumental relation with some goal or preference. The main difference with [24] is in the role given to the recursive modeling of the *normative* agent, a difference which is more apparent in [20] where an implementation with the DESIRE agent architecture is proposed. In our work too, obligations lead to goal adoption; but here, those goals becomes intentions only after the evaluation of the effects of the agent's alternatives, obtained via

[36] s should be better defined as *liability* since a sanction does not necessarily occur, as noticed in [53].

[37] [38], p. 163.

the recursive modeling of the reaction of the normative agent.
Moreover, we agree with [17] when he says that the limit of *deontic* logic is in the anti-mentalistic approach to obligations: “obligations and permission are just relative to actions, i.e. to the behavior of a cognitive agent (a behavior based on beliefs and directed by goals).

[4] proposes to use a game theoretic framework for modeling norms, too. In particular, our model is similar to Axelrod’s framework in that the sanction is not a necessary event but the result of the decision and action of another agent. However, there are two important differences: first, [4]’s agent are simple ones with limited rationality and not BDI agents: they decide to sanction on the basis of a so called *vengeance* attitude; second, and most important, the decision of the sanctioning agent includes also the deliberation about whether it is worth to monitor violations; in contrast, in [4], there is no monitoring action, but only a probability distribution on the events that a normative agent comes to know about a violation.

6 Conclusion

Our proposal constitutes a step forward in the understanding of deontic reasoning in that we include in the decision process the prediction of the *normative* agent’s autonomous behavior. This is the basis not only for discovering when it does not worth to fulfill an obligation, as well as for enabling agents to reason about how to deceive the *normative* agent. Predicting the possible failures and deceits of obligations is fundamental for building agent communities regulated by norms.

Finally, we used the reasoning process involving the prediction of the behavior of other agents for modeling cooperation among agents ([6]) and for modeling dialog ([8]); this form of reasoning is becoming a widespread methodology in the multi-agent field, as works like [30] witness.

In [9], the details and limitations of the planning process underlying this framework are discussed, while the phenomenon of deceits for avoiding the fulfillment of obligation is the topic of the ongoing work. Finally, the approach described in the paper has been applied for modelling the Hohfeldian concept of Legal Relations [10].

References

[1] A. Anderson. The logic of norms. *Logic et analyse*, 2, 1958.

[2] L. Ardissono, G. Boella, and L. Lesmo. The role of social goals in planning polite speech acts. In *Workshop on Attitude, Personality and Emotions in User-Adapted Interaction at UM’99 Conference*, pages 41–55, Banff, 1999.

[3] L. Ardissono, G. Boella, and L. Lesmo. Plan based agent architecture for interpreting natural language dialogue. *International Journal of Human-Computer Studies*, (52):583–636, 2000.

[4] R. Axelrod. An evolutionary approach to norms. *The American Political Science Review*, 80(4):1095–1111, 1986.

[5] G. Boella. *Cooperation among economically rational agents*. PhD thesis, Università di Torino, 2000.

[6] G. Boella, R. Damiano, and L. Lesmo. Cooperation and group utility. In N. Jennings and Y. Lespérance, editors, *Intelligent Agents VI — Proceedings of the Sixth International Workshop on Agent Theories, Architectures, and Languages (ATAL-99, Orlando FL)*, pages 319–333. Springer-Verlag, Berlin, 2000.

[7] G. Boella, R. Damiano, and L. Lesmo. Social goals in conversational cooperation. In *Proc. of Sigdial Workshop*, Honk Kong, 2000.

[8] G. Boella, R. Damiano, L. Lesmo, and L. Ardissono. Conversational cooperation: the leading role of intentions. In *Amstelogue'99 Workshop on Dialogue*, Amsterdam, 1999.

[9] G. Boella and L. Lesmo. Deliberate normative agents. In R. Conte and C. Dellarocas, editors, *Social order in MAS*. Kluwer, 2001.

[10] G. Boella, L. Lesmo, and L. Favali. The definition of legal relations in a bdi multiagent framework. In *Proc. of AI*IA*. Springer Verlag, Berlin, 2001.

[11] M. Boman. Norms as constraints on real-time autonomous agent action. In *Proc. of MAAMAW'97*, pages 36–44, Berlin, 1997. Springer Verlag.

[12] C. Boutilier. Multiagent systems: Challenges and opportunities for decision-theoretic planning. *AI Magazine*, 20(4):35–43, 1999.

[13] C. Boutilier, T. Dean, and S. Hanks. Decision theoretic planning: Structural assumptions and computational leverage. *Journal of Artificial Intelligence Research*, 1, 1999.

[14] P. Brown and S. C. Levinson. *Politeness: some universals on language usage*. Cambridge University Press, Cambridge, 1987.

[15] S. Carberry. *Plan Recognition in Natural Language Dialogue*. MIT Press, 1990.

[16] J. Carmo and O. Pacheco. Deontic and action logics for collective agency and roles. In R. Demolombe and R. Hilpinen, editors, *Proc. Fifth International Workshop on Deontic Logic in Computer Science (DEON'00)*, pages 93–124, ONERA-DGA, 2000.

[17] C. Castelfranchi. Practical permission. In *Proc. of Practical Reasoning and Rationality Workshop PRR'97*, 1997.

[18] C. Castelfranchi. Modeling social action for AI agents. *Artificial Intelligence*, 103:157–182, 1998.

[19] C. Castelfranchi and R. Conte. Commitment: from intentions to groups and organizations. In *Proc. of Autonomous Agents Workshop on Norms and Institutions in Multi-Agent Systems*, Barcelona, 2000.

[20] C. Castelfranchi, F. Dignum, C. M. Jonker, and J. Treur. Deliberate normative agents: Principles and architecture. In N. Jennings and Y. Lespérance, editors, *Intelligent Agents VI — Proceedings of the Sixth International Workshop on Agent Theories, Architectures, and Languages (ATAL-99)*, Lecture Notes in Artificial Intelligence. Springer-Verlag, Berlin, 2000.

[21] C. Castelfranchi and R. Falcone. Principles of trust for mas: Cognitive anatomy, social importance, and quantification. In *Proc. of ICMAS 98*. 1998.

[22] P. Cohen and H. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42:213–261, 1990.

[23] R. Conte and C. Castelfranchi. *Cognitive and Social Action*. UCL Press, 1995.

[24] R. Conte, C. Castelfranchi, and F. Dignum. Autonomous norm-acceptance. In J. P. Mueller, M. Singh, and A. Rao, editors, *Intelligent Agents V — Proc. of 5th Int. Workshop on Agent Theories, Architectures, and Languages (ATAL-98)*. Springer Verlag, Berlin, 1998.

[25] G. DeGiacomo and H. Levesque. Projection using regression and sensors. In *Proc. IJCAI*, pages 160–165, 1999.

[26] C. Dellarocas. Negotiated shared context and social control in open multi-agent systems. In R. Conte and C. Dellarocas, editors, *Social Order in MAS*. Kluwer, 2001.

[27] F. Dignum. Autonomous agents and social norms. In *ICMAS'96 Workshop on Norms, Obligations and Conventions*, 1996.

[28] F. Dignum, J.-J. Meyer, R. Wieringa, and R. Kuiper. A modal approach to intentions, commitments and obligations: intention plus commitment yields obligation. In *Proc. of DEON'96*, 1996.

[29] A. Giddens. *Sociology*. Polity Press, Cambridge, 1989.

[30] P. J. Gmytrasiewicz and E. H. Durfee. Formalization of recursive modeling. In *Proc. of first ICMAS-95*, 1995.

[31] P. J. Gmytrasiewicz and C. Lisetti. Using decision theory to formalize emotion for MAS. In *Proc. of ICMAS-00*, 2000.

[32] E. Goffman. *Interaction Ritual*. Penguin, Harmondsworth, 1967.

[33] E. Goffman. *Strategic Interaction*. Basil Blackwell, Oxford, 1970.

[34] P. Haddawy and S. Hanks. Utility models for goal-directed, decision-theoretic planners. *Computational Intelligence*, 14:392–429, 1998.

[35] P. Haddawy and M. Suwandi. Decision-theoretic refinement planning using inheritance abstraction. In *Proc. of 2nd Int. Conference on Artificial Intelligence Planning Systems*, pages 266–271, Menlo Park, CA, 1994.

[36] H. Herrestad and C. Krogh. Obligations directed from bearers to counterparties. In *Proc. of 5th. Conf. on AI and Law*. College Park, Md, 1995.

[37] R. L. Keeney and H. Raiffa. *Decision with multiple Objectives: preferences and value tradeoff*. Cambrige University Press, Cambridge (UK), 1976.

[38] C. Krogh. *Normative Structures in Natural and Artificial Systems*. Complex, TANO, Oslo, 1997.

[39] H. J. Levesque, R. Reiter, I. Lesperance, F. Lin, and R. Scherl. Golog: A logic programming language for dynamic domains. *Journal of Logic Programming*, 31(1-3):59–83, 1997.

[40] A. Ndiaye and A. Jameson. Predictive role taking in dialog: global anticipation feedback based on transmutability. In *Proc. 5th Int. Conf. on User Modeling*, pages 137–144, Kailua-Kona, Hawaii, 1996.

[41] M. Pollack and J. Horty. There's more to life than making plans. *AI Magazine*, 20(4):71–84, 1999.

[42] G. Sartor. Why agents comply with norms and why they should. In R. Conte and C. Dellarocas, editors, *Social order in MAS*. Kluwer, 2001.

[43] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[44] Y. Shoham. Agent-oriented programming. *Artificial Intelligence*, 60(1):51–92, 1993.

[45] T. Stratulat, F. Clerin-Debart, and P. Enjalbert. Norms and time in agent-based systems. In *Proc. 8th Conf. on AI and Law*, pages 178–184, St Louis (MO), 2001.

[46] M. Tennenholtz. On stable social laws and qualitative equilibria. *Artificial Intelligence*, 102(1):1–20, 1998.

[47] D. Tiscornia and F. Turchi. Formalization of legislative documents based on a functional model. In *Proc. 6th International Conference on AI and Law (ICAIL)*, pages 63–71, Melbourne, 1997.

[48] D. Traum and J. Allen. Discourse obligations in dialogue processing. In *Proc. 32nd Annual Meeting of ACL*, pages 1–8, Las Cruces, New Mexico, 1994.

[49] R. Tuomela and W. Balzer. Social institutions, norms and practices. In R. Conte and C. Dellarocas, editors, *Social order in MAS*. Kluwer, 2001.

[50] L. van der Torre and Y. Tan. Contrary-to-duty reasoning with preference-based dyadic obligations. *Annals of Mathematics and Artificial Intelligence*, 2000.

[51] J. Von Neumann and O. Morgenstern. *The Theory of Games and Economic Behavior*. Princeton University Press, Princeton, 1947.

[52] G. von Wright. Deontic logic. *Mind*, 60:1–15, 1950.

[53] G. von Wright. A new system of deontic logic. In R. Hilpinen, editor, *Deontic Logic*, pages 105–120. D. Reidel, Dordrecht-Holland, 1971.

[54] P. Xuan and V. R. Lesser. Incorporating uncertainty in agent commitments. In N. Jennings and Y. Lespérance, editors, *Intelligent Agents VI — Proceedings of the Sixth International Workshop on Agent Theories, Architectures, and Languages (ATAL-99)*, Lecture Notes in Artificial Intelligence. Springer-Verlag, Berlin, 2000.