Delegation of Power in Normative Multiagent Systems

Guido Boella[1] and Leendert van der Torre[2]

[1]Dipartimento di Informatica - Università di Torino - Italy. email: guido@di.unito.it
[2]University of Luxembourg. e-mail: leendert@vandertorre.com

Abstract. In this paper we reconsider the definition of counts-as relations in normative multiagent systems: counts-as relations do not always provide directly an abstract interpretation of brute facts in terms of institutional facts. We argue that in many cases the inference of institutional facts from brute facts is the result of actions of agents acting on behalf of the normative systems and who are in charge of recognizing which institutional facts follow from brute facts. We call this relation delegation of power: it is composed of a counts-as relation specifying that the effect of an action of an agent is an institutional fact and by a goal of the normative system that the fact is considered as an institutional fact. This relation is more complex than institutional empowerment, where an action of an agent counts-as an action of the normative system but no goal is involved, and than delegation of goals, where a goal is delegated to an agent without giving it any power. With two case studies we show the importance of the delegation of power. Finally, we show how the new definition can be related with existing ones by using different levels of abstraction.

1 Introduction

It is well known that normative systems include not only regulative norms like obligations, prohibitions and permissions, but also constitutive norms stating what counts as institutional facts in a normative system.

In this paper we introduce a new notion, called *delegation of power* beside constitutive and regulative norms. Thus, the research questions of this paper are:

– What is delegation of power in a normative multiagent system?
– How does it relate to counts-as conditionals?
– How does it relates to regulative norms?

The notion of counts-as introduced by Searle [1] has been interpreted in deontic logic in different ways and it seems to refer to different albeit related phenomena [2]. For example, Jones and Sergot [3] consider counts-as from the constitutive point of view. According to Jones and Sergot , the fact that A counts-as B in context C is read as a statement to the effect that A represents conditions for guaranteeing the applicability of particular classificatory categories. The counts-as guarantees the soundness of that inference, and enables "new" classifications which would otherwise not hold.

An alternative view of the counts-as relation is proposed by Grossi *et al.* [4]: according to the classificatory perspective A counts-as B in context C is interpreted as: A is classified as B in context C. In other words, the occurrence of A is a sufficient

condition, in context C, for the occurrence of B. Via counts-as statements, normative systems can establish the ontology they use in order to distribute obligations, rights, prohibitions, permissions, *etc.*

In [5, 6] we propose a different view of counts-as which focuses on the fact that counts-as often provides an abstraction mechanism in terms of institutional facts, allowing the regulative rules to refer to legal notions which abstract from details.

None of the above analyses, however, considers the motivational aspects behind constitutive norms required by agent theory. Constitutive norms are modelled as counts-as conditionals which allow to infer which institutional facts follow from brute facts and from existing institutional facts. E.g., a car counts as a vehicle for the traffic law. The inference from facts to institutional facts is considered as automatic, i.e., it is assumed not to need any agent or resource to perform it. Agent theory, instead, considers also the resources needed to perform inferences. Calculating the consequences following from some premises has a cost which must be traded off against the benefit of making the inferences. Thus in Georgeff and Ingrand [7] inferences are considered as actions which are planned and subject to decision processes as any other action: there must be an agent taking the decision and executing them.

According to resource bounded reasoning, the assumption made above on constitutive rules, even if useful in some circumstances, is not realistic. In many circumstances facts which in principle should be considered as institutional facts are not recognized as such. In such circumstances, the interpretation of a fact as an institutional fact may depend on the action of some agent who acts to achieve a goal of the normative system that a brute fact is interpreted as an institutional fact: we say that this agent has been *delegated the power* to interpret the fact as an institutional fact. In the next section two such examples are considered.

To answer the research questions of this paper we use our normative multiagent system framework [5, 8] which explicitly takes into account the activity of agents in the definition of sanction based obligations. The basic assumptions of our model are that beliefs, goals and desires of an agent are represented by conditional rules, and that, when an agent takes a decision, it recursively models the other agents interfering with it in order to predict their reaction to its decision as in a game. Most importantly, the normative system itself can be conceptualized as a socially constructed agent with whom it is possible to play games to understand what will be its reaction to the agent's decision: to consider its behavior as a violation and to sanction it. These actions are carried out by agents playing roles in the normative system, like judges and policemen. In the model presented in [5, 6], regulative norms are represented by the goals of the normative system and constitutive norms as its beliefs.

We relate the notion of delegation of power with our previous work on norms providing the definition of counts-as and of obligations at different levels of abstraction depending on whether agents are considered or not as acting for the normative system.

The paper is organized as follows. In Section 2 we show two case studies motivating the paper. In Section 3 we introduce delegation of power and in Section 4 we distinguish three different levels of abstraction. In Section 5 we introduce the formal model explained by examples. Comparison with related work and conclusion end the paper.

2　Motivating examples

To illustrate the necessity of the notion of delegation of power we resort to two motivating case studies. The first one is a case which happened to one of the authors. The second case concerns how the law deals with the enforcement of obligations.

The first example is about traffic norms. Due to increased levels of pollution, on certain days, only ecological vehicles are allowed in major towns. In particular, cars with a catalytic converter count as ecological vehicles. One author of this paper bought many years ago one of the first catalytic cars. So he felt permitted to go around by car in the days when non-catalytic cars were prohibited. Instead, he was stopped by the police and fined. Why? The car was bought before the local law recognized catalytic cars as ecological vehicles. The police agreed that the car had a catalytic converter: they could see it, the car worked only with unleaded fuel, both the manual and the licence of the car said it has a catalytic converter. However, there was a missing rubber stamp by the office declaring that the car counts as an ecological vehicle. The problem is not simply that only catalytic cars bought after a certain date are considered as ecological. Rather, a catalytic car is not ecological unless an agent officialy recognizes it as such. The police has no power to consider the car as ecological, the evidence notwithstanding.

The moral is that even if a brute fact is present and could allow the recognition of an institutional fact, the institutional fact is the result of the action of some agent who is empowered to do that.

The second example concerns how obligations are dealt with by the law. Obligations represent the social order the law aims to achieve. However, specifying this order is not sufficient to achieve it. Thus, obligations are associated with other instrumental norms - to use Hart [9]'s terminology: the lack of the wanted state of affairs is considered as a violation and the violator is sanctioned. These tasks are distributed to distinct agents, like judges and policemen, who have to decide whether and how to fulfill them.

There is, however, an asymmetry between considering something as a violation and sanctioning. The latter can be a physical action like putting into jail, while the former has always an institutional character. So, while the sanction can be directly performed by a policeman, the recognition of a violation can only be performed indirectly by means of some action which counts as the recognition of a violation, e.g., a trial by a judge.

The two examples have some similarities and differences. Both in case of ecological cars and in case of violations an action of an agent is necessary to create the institutional fact. These cases can be modelled by a counts-as relation between the action of an agent (putting a stamp on the car licence or recognizing a violation) and the institutional fact (being an ecological vehicle or having violated an obligation), rather than by a direct counts-as relation between the brute facts and the institutional facts. But at first sight the two cases also have a difference: the recognition of a violation is wanted by the normative system to achieve its social order. In this case besides the counts-as rule between the action and the recognition as a violation there is also the goal of the normative system that this recognition contributes to the social order. In the next section we argue that indeed both cases should be modelled by means of a goal of the normative system and a counts-as relation between actions of agents acting for the normative system and institutional facts: they are both examples of delegation of power.

3   Goal delegation and the delegation of power

In this section we show how both examples of Section 2 can be modelled in the same way, starting from the analysis of two apparently unrelated phenomena: goal delegation and institutional empowerment.

According to Castelfranchi [10], goal delegation is relying on another agent for the achievement of one own's goal: "in delegation an agent A needs or likes an action of another agent B and includes it in its own plan. In other words, A is trying to achieve some of its goals through B's behaviours or actions; thus A has the goal that B performs a given action/behaviour." This is not an institutional phenomenon but a basic capability of agents which enables them to interact with each other.

Institutional empowerment, instead, is by nature an institutional phenomenon which is based on the counts-as relation: an agent is empowered to perform an institutional action - a kind of institutional fact - if some of its actions counts as the institutional action. For example, a director can commit by means of his signature his institution to purchase some goods. Thus it is essentially related to counts-as rules, albeit restricted to actions of agents. Consider as a paradigmatic case the work by Jones and Sergot [3]

Bottazzi and Ferrario [11] argue that the two phenomena are related, as in cases like the two examples of Section 2: an agent which is institutionally empowered, is also delegated the goal of the institution of making true an institutional fact by exercising its power in the specified situations.

The connection between goal delegation and institutional empowerment is not a necessary one. For example, the agent in charge for sanctioning an obligation is delegated the goal of sanctioning, but there is no need of institutional powers in case of physical sanctions. Viceversa, the law institutionally empowers agents to stipulate private contracts which have the force of law, without being delegated by the law to do so, since contracting agents act for their own sake [5].

This connection, which we call *delegation of power*, can be used to explain the two examples above. In the case of cars, for the normative system, catalytic cars have to be considered as ecological vehicles. There are three possibilities: first, recognizing all catalytic cars as ecological vehicles by means of a constitutive norm. This solution, however, does not consider the actual performance of the inference and the possible costs related to it. Second, the normative system can rely on some agent to recognize catalytic cars as ecological vehicles. As said above, this can be done by means of a counts-as relation between an action of an agent and its effects. This solution, however, fails to account for the motivations that the agent should have to perform the action of recognizing ecological vehicles as such. Third, also a goal of the normative system is added to motivate its action: there is an agent who has the institutional power to recognize cars as ecological vehicles and the normative system has delegated it the goal that it does so in order to motivate it.

In the case of obligations, beside the counts-as relation between an action of judge and the institutional fact that a violation is recognized, we have the goal that the specified behavior is considered as a violation. The goal is an instrumental goal associated with the definition of obligation which aims at regulating how violations are prosecuted.

Thus to model delegation of power we need a model where:

– Both constitutive and regulative norms are modelled.
– Since we want to model goal delegation, mental attitudes are attributed to the normative system.
– Agents act inside the normative system.

All these features are present in our model of normative multiagent systems [5]. Our model is based on the so called agent metaphor: social entities like normative systems can be conceptualized as agents by attributing them mental attitudes like beliefs and goals. The cognitive motivations of the agent metaphor underlying our framework are discussed in [12].

Beliefs model the constitutive rules of the normative system, while goals model regulative rules. Thus, in the normative system the interaction between constitutive and regulative rules is the same as the interaction of beliefs and goals in an agent.

However, differently from a real agent, the normative system is a socially constructed agent. It exists only because of the collective acceptance by all the agents and, thus, it cannot act in the world. Its actions are carried out by agents playing roles in the normative system, like legislators, judges and policemen. It is a social construction used to coordinate the behavior of the agents.

Our model of roles, which allows to structure organizations in sub-organizations and roles to make a multiagent system modular and thus manage its complexity, is described in [13]. For space reason, we do not introduce roles explicitly in this paper.

In our model obligations are not only modelled as goals of the normative system, but they are also associated with the instrumental goals that the behavior of the addressee of the norms is considered as a violation and that the violation is sanctioned. Considering something as a violation and sanctioning are actions which can be executed by the normative system itself, or, at a more concrete level of detail, by agents playing roles in the normative system.

The counts-as relation in our model is modelled as a conditional belief of the normative system to provide an abstraction of reality in terms of institutional facts. Regulative norms can refer to this level, thus decoupling them from the details of reality. For example, it must be distinguished the institutional fact that traffic lights are red from the brute fact that red light bulbs in the traffic lights are on: in the extreme case the institutional fact can be true even if all the red bulbs are broken. As a consequence, as discussed in [6, 14], we do not accept the identity rule for counts-as.

In this paper, we consider how counts-as can be used to define delegation of power. Counts-as relations are not used in this case to directly connect brute facts to institutional facts, but only to express the (institutional) effects of actions of agents empowered by the normative system (in the same sense as the act of signing of the director counts as the commitment of the institution he belongs to).

In our model [5], constitutive rules have also other roles, in particular, they specify how the normative system can change. In this sense a normative multiagent system is a system in the sense that it specifies itself how it can change. Since it is a socially constructed agent, it cannot directly change itself. Rather it relies on the actions of agents playing roles in it, like legislators, which count as changes of the system.

4   Level of abstractions in the definition of norms

Obligations and counts-as relations can be defined at different levels of abstraction. We identify three different levels of abstraction, which are suited for different applications. The abstraction dimension is the detail at which we consider agents acting for the normative system: at the higher abstraction level agents have no role, at the middle abstraction level only actions of the normative system are considered but agents are not considered; only at the more concrete level, where agents are in charge of the actual functioning of the normative system concerning regulative and constitutive rules, delegation of power enters the picture:

1. The higher level abstracts from the fact that violations, sanctions and institutional facts are the result of the action that an agent decides to perform. Thus obligations are defined as in Anderson's reduction [15]: the recognition of the violation and the sanction logically follow from the violation. This abstraction level for regulative rules is adopted also by [16, 17] and we use it in [18]. For constitutive rules, up to our knowledge, this is the only level considered.
2. The middle level abstracts from the role of agents in the normative system, but the normative system is in some way personified and is assumed to act in the world: thus the recognition of violations and sanctions are considered as the actions of the normative system itself. We adopt this level of representation for regulative norms in [5, 8]. Analogously, institutional facts follow from actions of the normative system: they are not anymore logical consequences of facts, but consequences of decisions of the normative systems which are traded-off against other decisions. They, thus, do not follow automatically, since the normative system can take a different decision due to conflicts with other goals or to lack of resources.
3. The lower level takes into account the actions of the agents in the normative system. Concerning regulative norms, some agents are delegated the goal to sanction violations and the goal and power of recognizing violations. I.e., they are delegated the power to do so. Concerning constitutive norms, the agents are delegated the goal to recognize some facts as institutional facts and the power to do so by means of their actions. I.e., they are delegated the power to do so. The problem of agents recognizing violations has been partially addressed in [19], but the recognition action was considered as a physical action like the sanction. In this paper we add the counts-as relation to the recognition of violations.

At the lower two levels it becomes possible to answer the question whether constitutive norms can be violated like it happens for regulative ones. A constitutive norm can be violated in the sense that the normative system or the agent who is delegated the goal to recognize the institutional fact and empowered to do so fails to achieve the delegated goal. In our first example the office could fail or refuse to recognize the car as an ecological vehicle. The reason can be the inability to perform the necessary actions, laziness, bribing, *etc.*, like it happens for regulative norms. Moreover, constitutive rules can be abused, in the sense that the delegated agent can exercise its power without being delegated to do so in the given circumstances. This possibility assumes that the institutional power can be exercised beyond the conditions under which it has been delegated the goal to exercise it.

5 The formal model

The definition of the agents is inspired by the rule based BOID architecture [20], though in our theory, and in contrast to BOID, obligations are not taken as primitive concepts. Beliefs, desires and goals are represented by conditional rules rather than in a modal framework. Intentions have been introduced as a form of bounded rationality: since an agent has not enough resources to make the optimal decision at each moment, he maintains its previous choices. In this paper we consider only one decision, so we do not need to introduce intentions to model decisions which persist over time.

5.1 Input/output logic

To represent conditional mental attitudes we take a simplified version of the input/output logics introduced in [21, 22]. A rule set is a set of ordered pairs $P \rightarrow q$. For each such pair, the body P is thought of as an input, representing some condition or situation, and the head q is thought of as an output, representing what the rule tells us to be believed, desirable, obligatory or whatever in that situation. In this paper, to keep the formal exposition simple, input and output are respectively a set of literals and a literal.

The development of input/output logic has been motivated by conditional norms, which do not have a truth value. For that reason, the semantics of input/output logic given by Makinson and van der Torre [21] is an operational semantics, which characterizes the output as a function of the input and the set of norms. However, it is easy to define a classical semantics for conditional norms too. Makinson and van der Torre illustrate how to recapture input/output logic in modal logic, and thus give it a classical possible worlds semantics. More elegantly, as illustrated by Bochman [23], the operational semantics of input/output logic can be rephrased as a bimodel semantics, in which a model of a set of conditionals is a pair of partial models from the base logic (in this paper, propositional logic).

Though the development of input/output logic has been motivated by the logic of norms, the same logic can be used for other conditionals like conditional beliefs and conditional goals – which explains the more general name of the formal system. Moreover, Bochman [23] also illustrates how the same logic is used for causal reasoning and various non-monotonic reasoning formalisms.

Definition 1 (Input/output logic).
Let X be a set of propositional variables, the set of literals built from X, written as $Lit(X)$, is $X \cup \{\neg x \mid x \in X\}$, and the set of rules built from X, written as $Rul(X) = 2^{Lit(X)} \times Lit(X)$, is the set of pairs of a set of literals built from X and a literal built from X, written as $\{l_1, \ldots, l_n\} \rightarrow l$. We also write $l_1 \wedge \ldots \wedge l_n \rightarrow l$ and when $n = 0$ we write $\top \rightarrow l$. For $x \in X$ we write $\sim x$ for $\neg x$ and $\sim(\neg x)$ for x. Moreover, let Q be a set of pointers to rules and $RD : Q \rightarrow Rul(X)$ is a total function from the pointers to the set of rules built from X.

Let $S = RD(Q)$ be a set of rules $\{P_1 \rightarrow q_1, \ldots, P_n \rightarrow q_n\}$, and consider the following proof rules strengthening of the input (SI), disjunction of the input (OR), cumulative transitivity (CT) and Identity (Id) defined as follows:

$$\frac{p \rightarrow r}{p \wedge q \rightarrow r} SI \qquad \frac{p \wedge q \rightarrow r, p \wedge \neg q \rightarrow r}{p \rightarrow r} OR \qquad \frac{p \rightarrow q, p \wedge q \rightarrow r}{p \rightarrow r} CT \qquad \frac{}{p \rightarrow p} Id$$

The following output operators are defined as closure operators on the set S using the rules above.

out_1: SI (simple-minded output)   *out_3: SI+CT (reusable output)*
out_2: SI+OR (basic output)   *out_4: SI+OR+CT (basic reusable output)*

Moreover, the following four throughput operators are defined as closure operators on the set S. out_i^+: out_i+Id (throughput) We write $out(Q)$ for any of these output operations and $out^+(Q)$ for any of these throughput operations. We also write $l \in out(Q, L)$ iff $L \rightarrow l \in out(Q)$, and $l \in out^+(Q, L)$ iff $L \rightarrow l \in out^+(Q)$.

Example 1. Given $RD(Q) = \{a \rightarrow x, x \rightarrow z\}$ the output of Q contains $x \wedge a \rightarrow z$ using the rule SI. Using also the CT rule, the output contains $a \rightarrow z$. $a \rightarrow a$ follows only if there is the Id rule.

A technical reason to distinguish pointers from rules is to facilitate the description of the priority ordering we introduce in the following definition.

The notorious contrary-to-duty paradoxes such as Chisholm's and Forrester's paradox have led to the use of constraints in input/output logics [22]. The strategy is to adapt a technique that is well known in the logic of belief change - cut back the set of norms to just below the threshold of making the current situation inconsistent.

Definition 2 (Constraints).

Let $\geq$: $2^Q \times 2^Q$ be a transitive and reflexive partial relation on the powerset of the pointers to rules containing at least the subset relation and $RD : Q \rightarrow Rul(X)$ a function from the pointers to the set of rules. Moreover, let out be an input/output logic:

- *$maxfamily(Q, P)$ is the set of $\subseteq$-maximal subsets Q' of Q such that $out(Q', P) \cup P$ is consistent.*
- *$preffamily(Q, P, \geq)$ is the set of $\geq$-maximal elements of $maxfamily(Q, P)$.*
- *$outfamily(Q, P, \geq)$ is the output under the elements of preffamily, i.e., $\{out(Q', P) \mid Q' \in preffamily(Q, P, \geq)\}$.*
- *$P \rightarrow x \in out_\cup(Q, \geq)$ iff $x \in \cup outfamily(Q, P, \geq)$*
  $P \rightarrow x \in out_\cap(Q, \geq)$ iff $x \in \cap outfamily(Q, P, \geq)$

Example 2. Let $RD(\{a, b, c\}) = \{a = (\top \rightarrow m), b = (p \rightarrow n), c = (o \rightarrow \neg m)\}$, $\{b, c\} > \{a, b\} > \{a, c\}$, where by $A > B$ we mean as usual $A \geq B$ and $B \not\geq A$.
$maxfamily(Q, \{o\}) = \{\{a, b\}, \{b, c\}\}$,
$preffamily(Q, \{o\}, \geq) = \{\{b, c\}\}$,
$outfamily(Q, \{o\}, \geq) = \{\{\neg m\}\}$

The $maxfamily$ includes the sets of applicable compatible pointers to rules together with all non applicable ones: e.g., the output of $\{a, c\}$ in the context $\{o\}$ is not consistent. Finally $\{a\}$ is not in $maxfamily$ since it is not maximal, we can add the non applicable rule b. Then *preffamily* is the preferred set $\{b, c\}$ according to the ordering on set of rules above. The set *outfamily* is composed by the consequences of applying the rules $\{b, c\}$ which are applicable in o (c): $\neg m$.

Due to space limitations we have to be brief on details with respect to input/output logics, see [21, 22] for the semantics of input/output logics, further details on its proof theory, its possible translation to modal logic, alternative constraints, and examples.

5.2 Multiagent systems

We assume that the base language contains boolean variables and logical connectives. The variables are either *decision variables* of an agent, which represent the agent's actions and whose truth value is directly determined by it, or *parameters*, which describe both the state of the world and *institutional facts*, and whose truth value can only be determined indirectly. Our terminology is borrowed from Lang *et al.* [24] and is used in discrete event systems, and many formalisms in operations research.

Given the same set of mental attitudes, agents reason and act differently: when facing a conflict among their motivations and beliefs, different agents prefer to fulfill different goals and desires. We express these agent characteristics by a priority relation on the mental attitudes which encode, as detailed in [20], how the agent resolves its conflicts. The priority relation is defined on the powerset of the mental attitudes such that a wide range of characteristics can be described, including social agents that take the desires or goals of other agents into account. The priority relation contains at least the subset-relation which expresses a kind of independence among the motivations.

Background knowledge is formalized by a set of effects E represented by rules.

Definition 3 (Agent set). *An agent set is a tuple $\langle A, X, B, D, G, AD, E, \geq, \geq_E \rangle$, where:*

- *the agents A, propositional variables X, agent beliefs B, desires D, goals G, and effects E are six finite disjoint sets.*
- *B, D, G are sets of mental attitudes. We write $M = D \cup G$ for the motivations defined as the union of the desires and goals.*
- *an agent description $AD : A \to 2^{X \cup B \cup M}$ is a total function that maps each agent to sets of variables (its decision variables), beliefs, desires and goals, but that does not necessarily assign each variable to at least one agent. For each agent $b \in A$, we write X_b for $X \cap AD(b)$, and B_b for $B \cap AD(b)$, D_b for $D \cap AD(b)$, etc. We write parameters $P = X \setminus \cup_{b \in A} X_b$.*
- *the set of effects E represents the background knowledge of all agents.*
- *a priority relation $\geq: A \to 2^B \times 2^B \cup 2^M \times 2^M$ is a function from agents to a transitive and reflexive partial relation on the powerset of the motivations containing at least the subset relation. We write $\geq_b$ for $\geq (b)$.*
- *a priority relation $\geq_E: 2^E \times 2^E$ is a transitive and reflexive partial relation on the powerset of effects containing at least the subset relation.*

Example 3. $A = \{\mathbf{a}\}$, $X_{\mathbf{a}} = \{drive\}$, $P = \{s, catalytic\}$, $D_{\mathbf{a}} = \{d_1, d_2\}$, $\geq_{\mathbf{a}} = \{d_2\} \geq \{d_1\}$. There is a single agent, agent $\mathbf{a}$, who can drive a car. Moreover, it can be sanctioned and the car can be catalytic. It has two desires, one to drive (d_1), another one not to be sanctioned (d_2). The second desire is more important.

In a multiagent system, beliefs, desires, goals and effects are abstract concepts which are described by rules built from literals.

Definition 4 (Multiagent system). *A multiagent system, written as NMAS, is a tuple $\langle A, X, B, D, G, AD, E, RD, \geq, \geq_E \rangle$, where $\langle A, X, B, D, G, AD, E, \geq, \geq_E \rangle$ is an agent set, and the rule description $RD : (B \cup M \cup E) \to Rul(X)$ is a total*

function from the sets of beliefs, desires and goals, and effects to the set of rules built from X. For a set of pointers $S \subseteq B \cup M \cup E$, we write $RD(S) = \{RD(q) \mid q \in S\}$.

Example 4 (Continued). $RD(d_1) = \top \rightarrow drive$, $RD(d_2) = \top \rightarrow \neg s$.

In the description of the normative system, we do not introduce norms explicitly, but we represent several concepts which are illustrated in the following sections. Institutional facts (I) represent legal abstract categories which depend on the beliefs of the normative system and have no direct counterpart in the world. $F = X \setminus I$ are what Searle calls "brute facts": physical facts like the actions of the agents and their effects. $V(x, \mathbf{a})$ represents the decision of agent $\mathbf{n}$ that recognizes x as a violation by agent $\mathbf{a}$. The goal distribution $GD(\mathbf{a}) \subseteq G_{\mathbf{n}}$ represents the goals of agent $\mathbf{n}$ the agent $\mathbf{a}$ is responsible for.

Definition 5 (Normative system). *A normative multiagent system, written as NMAS, is a tuple $\langle A, X, B, D, G, AD, E, RD, \geq, \geq_E, \mathbf{n}, I, V, GD \rangle$ where the tuple $\langle A, X, B, D, G, AD, E, RD, \geq, \geq_E \rangle$ is a multiagent system, and*

- *the normative system $\mathbf{n} \in A$ is an agent.*
- *the institutional facts $I \subseteq P$ are a subset of the parameters.*
- *the norm description $V : Lit(X) \times A \rightarrow X_{\mathbf{n}} \cup P$ is a function from the literals and the agents to the decision variables of the normative system and the parameters.*
- *the goal distribution $GD : A \rightarrow 2^{G_{\mathbf{n}}}$ is a function from the agents to the powerset of the goals of the normative system, such that if $L \rightarrow l \in RD(GD(\mathbf{a}))$, then $l \in Lit(X_{\mathbf{a}} \cup P)$.*

Agent $\mathbf{n}$ is a normative system with the goal that non catalytic cars are not driven.

Example 5 (Continued). There is agent $\mathbf{n}$, representing the normative system.
  $P = \{s, V(drive, \mathbf{a}), catalytic\}$, $D_{\mathbf{n}} = G_{\mathbf{n}} = \{g_1\}$,
  $RD(g_1) = \{\neg catalytic \rightarrow \neg drive\}$, $GD(\mathbf{a}) = \{g_1\}$.
The parameter $V(drive, \mathbf{a})$ represents the fact that the normative system considers a violation agent $\mathbf{a}$'s action of driving. It has the goal that non-ecological vehicles should not be driven by $\mathbf{a}$ and it has distributed this goal to agent $\mathbf{a}$.

In the following, we use an input/output logic *out* to define whether a desire or goal implies another one and to define the application of a set of belief rules to a set of literals; in both cases we use the out_3 operation since it has the desired logical property of not satisfying identity.

We now define obligations and counts-as at the three levels of abstraction.

Regulative norms are conditional obligations with an associated sanction. At the higher level of abstraction, the definition contains three clauses: the first two clauses state that recognitions of violations and sanctions are a consequence of the behavior of agent $\mathbf{a}$, as it is represented by the background knowledge rules E. For an obligation to be effective, the third clause states that the sanction must be disliked by its addressee.

Definition 6 (Obligation (level 1)). *Let NMAS be a normative multiagent system $\langle A, X, B, D, G, AD, E, RD, \geq, \geq_E, \mathbf{n}, I, V, GD \rangle$.*

Agent $\mathbf{a} \in A$ is obliged to see to it that $x \in Lit(X_{\mathbf{a}} \cup P)$ with sanction $s \in Lit(P)$ in context $Y \subseteq Lit(X)$ in NMAS, written as $NMAS \models O^1_{\mathbf{an}}(x, s|Y)$, if and only if:

1. $Y \cup \{\sim x\} \rightarrow V(\sim x, \mathbf{a}) \in out(E, \geq_E)$: if Y and x is false, then it follows that $\sim x$ is a violation by agent $\mathbf{a}$.
2. $Y \cup \{V(\sim x, \mathbf{a})\} \rightarrow s \in out(E, \geq_E)$: if Y and there is a violation by agent $\mathbf{a}$, then it is sanctioned.
3. $Y \rightarrow \sim s \in out(D_\mathbf{a}, \geq_\mathbf{a})$: if Y, then agent $\mathbf{a}$ desires $\sim s$, which expresses that it does not like to be sanctioned.

Example 6. Let: $E = \{e_1, e_2\}$, $D_\mathbf{a} = \{d_2\}$
$RD(e_1) = \{\neg catalytic, drive\} \rightarrow V(drive, \mathbf{a})$
$RD(e_2) = \{\neg catalytic, V(drive, \mathbf{a})\} \rightarrow s$
$RD(d_2) = \neg catalytic \rightarrow \sim s$
$NMAS \models O_\mathbf{an}^1(\neg drive, s \mid \neg catalytic)$, since:

1. $\{\neg catalytic, drive\} \rightarrow V(drive, \mathbf{a}) \in out(E, \geq_E)$
2. $\{\neg catalytic, V(drive, \mathbf{a})\} \rightarrow s \in out(E, \geq_E)$
3. $\neg catalytic \rightarrow \sim s \in out(D_\mathbf{a}, \geq_\mathbf{a})$

Constitutive norms introduce new abstract categories of existing facts and entities, called institutional facts. In [6] we formalize the counts-as conditional as a belief rule of the normative system $\mathbf{n}$. Since the condition x of the belief rule is a variable it can be an action of an agent, a brute fact or an institutional fact. So, the counts-as relation can be iteratively applied. In our model the counts-as relation does not satisfy the identity rule. See [6] for a discussion of the motivations.

Definition 7 (Counts-as relation (level 1)). *Let NMAS be a normative multiagent system $\langle A, X, B, D, G, AD, E, RD, \geq, \geq_E, \mathbf{n}, I, V, GD \rangle$. A literal $x \in Lit(X)$ counts-as $y \in Lit(I)$ in context $C \subseteq Lit(X)$, $NMAS \models$ counts-as$_\mathbf{n}^1(x, y|C)$, iff $C \cup \{x\} \rightarrow y \in out(B_\mathbf{n}, \geq_\mathbf{n})$: if agent $\mathbf{n}$ believes C and x then it believes y.*

Example 7. $P \setminus I = \{catalytic\}$, $I = \{eco\}$, $X_\mathbf{a} = \{drive\}$, $B_\mathbf{n} = \{b_1\}$, $RD(b_1) = catalytic \rightarrow eco$

Consequently, $NMAS \models$ counts-as$_\mathbf{n}^1(catalytic, eco|\top)$. This formalizes that for the normative system a catalytic car counts as an ecological vehicle. The presence of the catalytic converter is a physical "brute" fact, while being an ecological vehicle is an institutional fact. In situation $S = \{catalytic\}$, given $B_\mathbf{n}$ we have that the consequences of the constitutive norms are $out(B_\mathbf{n}, S, \geq_\mathbf{n}) = \{eco\}$ (since out_3 does not include Id).

At the middle level of abstraction, actions of the normative systems are added in the definition of the obligations: the recognition of a violation and sanctions. Since the actions undergo a decision process, desires and goals of the normative system are added. The first and central clause of our definition of obligation defines obligations of agents as goals of the normative system, following the "your wish is my command" metaphor. It says that the obligation is implied by the desires of the normative system $\mathbf{n}$, implied by the goals of agent $\mathbf{n}$, and it has been distributed by agent $\mathbf{n}$ to the agent. The latter two steps are represented by $out(GD(\mathbf{a}), \geq_\mathbf{n})$.

The second and third clause can be read as the normative system has the goal that the absence of p is considered as a violation. The third clause says that the agent desires

that there are no violations, which is stronger than that it does not desire violations, as would be expressed by $\top \rightarrow V(\sim x, a) \notin out(D_{\mathbf{n}}, \geq_{\mathbf{n}})$.

The fourth and fifth clause relate violations to sanctions. The fifth clause says that the normative system is motivated not to sanction as long as their is no violation, because otherwise the norm would have no effect. Finally, for the same reason the last clause says that the agent does not like the sanction. The second and fourth clause can be considered as instrumental norms [9] contributing to the achievement of the main goal of the norm.

Definition 8 (Obligation (level 2)). *Let NMAS be a normative multiagent system* $\langle A, X, B, D, G, AD, E, RD, \geq, \geq_E, \mathbf{n}, I, V, GD \rangle$. *Agent* $\mathbf{a} \in A$ *is obliged to see to it that* $x \in Lit(X_{\mathbf{a}} \cup P)$ *with sanction* $s \in Lit(X_{\mathbf{n}} \cup P)$ *in context* $Y \subseteq Lit(X)$ *in NMAS, written as* $NMAS \models O^2_{\mathbf{an}}(x, s|Y)$, *if and only if:*

1. $Y \rightarrow x \in out(D_{\mathbf{n}}, \geq_{\mathbf{n}}) \cap out(GD(\mathbf{a}), \geq_{\mathbf{n}})$: *if Y holds then agent $\mathbf{n}$ desires and has as a goal that x, and this goal has been distributed to agent $\mathbf{a}$.*
2. $Y \cup \{\sim x\} \rightarrow V(\sim x, \mathbf{a}) \in out(D_{\mathbf{n}}, \geq_{\mathbf{n}}) \cap out(G_{\mathbf{n}}, \geq_{\mathbf{n}})$: *if Y holds and $\sim x$, then agent $\mathbf{n}$ has the goal and the desire $V(\sim x, \mathbf{a})$: to recognize it as a violation by agent $\mathbf{a}$.*
3. $\top \rightarrow \neg V(\sim x, \mathbf{a}) \in out(D_{\mathbf{n}}, \geq_{\mathbf{n}})$: *agent $\mathbf{n}$ desires that there are no violations.*
4. $Y \cup \{V(\sim x, \mathbf{a})\} \rightarrow s \in out(D_{\mathbf{n}}, \geq_{\mathbf{n}}) \cap out(G_{\mathbf{n}}, \geq_{\mathbf{n}})$: *if Y holds and agent $\mathbf{n}$ decides $V(\sim x, \mathbf{a})$, then agent $\mathbf{n}$ desires and has as a goal that it sanctions agent $\mathbf{a}$.*
5. $Y \rightarrow \sim s \in out(D_{\mathbf{n}}, \geq_{\mathbf{n}})$: *if Y holds, then agent $\mathbf{n}$ desires not to sanction. This desire of the normative system expresses that it only sanctions in case of violation.*
6. $Y \rightarrow \sim s \in out(D_{\mathbf{a}}, \geq_{\mathbf{a}})$: *if Y holds, then agent $\mathbf{a}$ desires $\sim s$, which expresses that it does not like to be sanctioned.*

The rules in the definition of obligation are only motivations, and not beliefs, because a normative system may not recognize that a violation counts as such, or that it does not sanction it: it is up to its decision. Both the recognition of the violation and the application of the sanction are the result of autonomous decisions of the normative system that is modelled as an agent.

The beliefs, desires and goals of the normative agent - defining the obligations - are not private mental states of an agent. Rather they are collectively attributed by the agents of the normative system to the normative agent: they have a public character, and, thus, which are the obligations of the normative system is a public information.

Example 8. Let: $\{g_1, g_2, g_4\} = G_{\mathbf{n}}, G_{\mathbf{n}} \cup \{g_3, d_2\} = D_{\mathbf{n}}, \{g_1\} = GD(\mathbf{a}), \{d_2\} = D_{\mathbf{a}}$

$RD(g_2) = \{\neg catalytic, drive\} \rightarrow V(drive, \mathbf{a})$ $RD(g_3) = \top \rightarrow \neg V(drive, \mathbf{a})$
$RD(g_4) = \{\neg catalytic, V(drive, \mathbf{a})\} \rightarrow s$
$NMAS \models O^2_{\mathbf{an}}(\neg drive, s \mid \neg catalytic)$, since:

1. $\neg catalytic \rightarrow \neg drive \in out(D_{\mathbf{n}}, \geq_{\mathbf{n}}) \cap out(GD(\mathbf{a}), \geq_{\mathbf{n}})$
2. $\{\neg catalytic, drive\} \rightarrow V(drive, \mathbf{a}) \in out(D_{\mathbf{n}}, \geq_{\mathbf{n}}) \cap out(G_{\mathbf{n}}, \geq_{\mathbf{n}})$
3. $\top \rightarrow \neg V(drive, \mathbf{a}) \in out(D_{\mathbf{n}}, \geq_{\mathbf{n}})$
4. $\{\neg catalytic, V(drive, \mathbf{a})\} \rightarrow s \in out(D_{\mathbf{n}}, \geq_{\mathbf{n}}) \cap out(G_{\mathbf{n}}, \geq_{\mathbf{n}})$

5. $\neg catalytic \rightarrow \sim s \in out(D_\mathbf{n}, \geq_\mathbf{n})$
6. $\neg catalytic \rightarrow \sim s \in out(D_\mathbf{a}, \geq_\mathbf{a})$

At the middle level of abstraction, the beliefs of the normative system represent only the connections between actions and the consequences of these actions for the normative system. The normative system has the desire and goal that the institutional fact y holds if the fact x holds in context C. The normative system believes that to make y true it has to perform an action z. Thus it is not sufficient the fact x holding in context C for the institutional fact y to be true: it is necessary also a decision to do z by the normative system.

Definition 9 (Counts-as relation (level 2)). *Let NMAS be a normative multiagent system* $\langle A, X, B, D, G, AD, E, RD, \geq, \geq_E, \mathbf{n}, I, V, GD \rangle$. *A literal* $x \in Lit(X)$ *counts-as* $y \in Lit(I)$ *in context* $C \subseteq Lit(X)$, *NMAS* $\models$ *counts-as*$_\mathbf{n}^2(x, y|C)$, *iff:*

1. $C \wedge x \rightarrow y \in out(D_\mathbf{n}, \geq_\mathbf{n}) \cap out(G_\mathbf{n}, \geq_\mathbf{n})$: *it is a desire and goal of the normative system that in context* C *the fact* x *is considered as the institutional fact* y.
2. $\exists z \in X_\mathbf{n}$ *such that* $C \cup \{z\} \rightarrow y \in out(B_\mathbf{n}, \geq_\mathbf{n})$: *there exists an action* z *of the normative system* $\mathbf{n}$ *such that if it decides* z *in context* C *then it believes that the institutional fact* y *follows (i.e.,* counts-as$_\mathbf{n}^1(z, y|C)$ *at the first level of abstraction).*

Example 9. $P \setminus I = \{catalytic\}$, $I = \{eco\}$, $X_\mathbf{a} = \{drive\}$, $X_\mathbf{n} = \{stamp\}$
$D_\mathbf{n} = G_\mathbf{n} = \{d_3\}$, $RD(d_3) = catalytic \rightarrow eco$
$B_\mathbf{n} = \{b_1\}$, $RD(b_1) = stamp \rightarrow eco$

Consequently, $NMAS \models counts\text{-}as_\mathbf{n}^2(catalytic, eco|\top)$. This formalizes that the normative system wants that if a car is catalytic, then it is considered as an ecological vehicle and the normative believes that from system putting a stamp on a catalytic car licence follows the fact that the car is catalytic. In situation $S = \{catalytic\}$, given $B_\mathbf{n}$ we have that the consequences of the constitutive norms are $out(B_\mathbf{n}, S, \geq_\mathbf{n}) = \emptyset$ and thus the goal d_3 remains unsatisfied, while in situation $S' = \{catalytic, stamp\}$ they are $out(B_\mathbf{n}, S', \geq_\mathbf{n}) = \{eco\}$ and the goal d_3 is satisfied.

The institutional facts can appear in the conditions of regulative norms:

Example 10. A regulative norm which forbids driving non catalytic cars can refer to the abstract concept of ecological vehicle rather than to catalytic converters: $O_\mathbf{an}^2(\neg drive, s \mid \neg eco)$.

As the system evolves, new cases can be added to the notion of ecological vehicle by means of new constitutive norms, without changing the regulative norms about it. E.g., if a car has fuel cells, then it is an ecological vehicle: $fuelcell \rightarrow eco \in RD(B_\mathbf{n})$.

This level of abstraction supposes that the normative system is an agent acting in the world. This abstraction can be detailed by introducing agents acting on behalf of the normative system: the normative system wants that an agent a makes the institutional fact y true if x holds in context C and believes that the effect of action z of agent a is the institutional fact y.

Before introducing the more concrete level of abstraction in obligations we discuss the third level of constitutive norms which is based on the notion of delegation of power.

Definition 10 (Counts-as relation (level 3) and delegation of power). *Let NMAS be a normative multiagent system* $\langle A, X, B, D, G, AD, E, RD, \geq, \geq_E, \mathbf{n}, I, V, GD \rangle$.

$a \in A$ is an agent, $z \in X_a$ an action of agent a, $x \in Lit(X)$ is a literal built out of a variable, $y \in Lit(I)$ a literal built out of an institutional fact, $C \subseteq Lit(X)$ the context. Agent a has been delegated the power to consider x in context C as the institutional fact y, $NMAS \models delegated_{\mathbf{n}}(a, z, x, y|C)$, iff:

1. *$C \wedge x \rightarrow y \in out(D_{\mathbf{n}}, \geq_{\mathbf{n}}) \cap out(GD(a), \geq_{\mathbf{n}})$: it is a desire of the normative system and a goal distributed to agent a that in context C the fact x is considered as the institutional fact y.*
2. *$\exists z \in X_a$ such that $C \cup \{z\} \rightarrow y \in out(B_{\mathbf{n}}, \geq_{\mathbf{n}})$: there exists an action z of agent a such that if it decides z then the normative system believes that the institutional fact y follows (i.e., $counts\text{-}as_{\mathbf{n}}^1(z, y|C)$ at the first level of abstraction).*

If $NMAS \models delegated_{\mathbf{n}}(a, z, x, y|C)$, then $NMAS \models counts\text{-}as_{\mathbf{n}}^3(x, y|C)$,

Example 11. $b \in A, P \setminus I = \{catalytic\}, I = \{eco\}, X_{\mathbf{a}} = \{drive\}, X_b = \{stamp\}$
    $D_{\mathbf{n}} = GD(b) = \{d_3\}, RD(d_3) = catalytic \rightarrow eco$
    $B_{\mathbf{n}} = \{b_1\}, RD(b_1) = stamp \rightarrow eco$

Thus, $NMAS \models delegated_{\mathbf{n}}(b, stamp, catalytic, eco|\top)$. Note that with respect to Example 9, the goal d_3 is distributed to agent b and $stamp$ is an action of agent b.

We can now define obligations where agents have been delegated the power of recognizing violations by means of actions which count as such. Differently from the obligation of level 2, clause 2 distributes a goal to agent b who is in charge of recognizing violations and whose action z is believed by the normative system $\mathbf{n}$ to be the recognition of a violation (clause 7).

Definition 11 (Obligation (level 3)). *Let NMAS be a normative multiagent system $\langle A, X, B, D, G, AD, E, RD, \geq, \geq_E, \mathbf{n}, I, V, GD \rangle$. Agent $\mathbf{a} \in A$ is obliged to see to it that $x \in Lit(X_{\mathbf{a}} \cup P)$ with sanction $s \in Lit(X_b \cup P)$ in context $Y \subseteq Lit(X)$ in NMAS, written as $NMAS \models O_{\mathbf{an}}^3(x, s|Y)$, if and only if $\exists b \in A$ and a decision variable $z \in X_b$ such that Definition 8 holds except that:*

2. *$Y \cup \{\sim x\} \rightarrow V(\sim x, \mathbf{a}) \in out(D_{\mathbf{n}}, \geq_{\mathbf{n}}) \cap out(GD(b), \geq_{\mathbf{n}})$: if Y holds and $\sim x$ is true , then agent $\mathbf{n}$ has distributed the goal $V(\sim x, \mathbf{a})$: that it is recognized as a violation in context Y.*
7. *$Y \cup \{z\} \rightarrow V(\sim x, \mathbf{a}) \in out(B_{\mathbf{n}}, \geq_{\mathbf{n}})$: from action z of agent b is believed to follow the recognition of the violation.*

From clause 2 and 7 it follows that agent b has been delegated the power to recognize violations by means of its action z.
    $NMAS \models \exists b \in A, z \in X_b \ delegated_{\mathbf{n}}(b, z, \sim x, V(\sim x, \mathbf{a}) \mid Y)$

Note that clause 2 of the definition above is like the first clause of an obligation $O_{b\mathbf{n}}(V(\sim x, \mathbf{a}), s' \mid Y \cup \{\sim x\})$. The model can thus be extended with obligations towards agents which have to take care of the procedural aspects of law, like prosecuting violations and sanctioning violators. These additional obligations are discussed in [19] and provide a motivation for the prosecuting and sanctioning agents. In the Italian law, for example, it is obligatory for an attorney to start a prosecution process when he comes to know about a crime (art. 326 of *Codice di procedura penale*).

6 Conclusions and related work

In this paper we introduce the notion of delegation of power which elaborates the counts-as relation extending it to cope with some real situations. We show that counts-as relations in some cases depend on the action of agents which are in charge of recognizing facts as institutional facts. Moreover, we show that these agents are motivated to do so by a goal delegated to them by the normative system. If these two conditions are true we say that the agents have been delegated a power. Once we define the delegation of power relation, we can use it to extend our sanction based definition of obligations in order to model agents which prosecute violations.

Our model allows to distinguish three levels of abstractions: at the higher level of abstraction violations, sanctions and institutional facts follow without the intervention of any agent. At the middle level the normative system acts to satisfy its goal to recognize violations, to sanction and to establish institutional facts. At the most concrete level, agents are in charge of achieving the goals of the normative system and are empowered to do so.

The notion of empowerment in normative multiagent systems is widely discussed, but it has not been related yet with the notion of goal delegation.

Pacheco and Santos [25], for example, discuss the delegation of obligations among roles. In particular, they argue that when an obligation is delegated, a corresponding permissions must be delegated too. This rationality constraint inside an institution parallels our notion of delegation of power: when the goal to make true an institutional fact is delegated, the agent must be empowered to do so too. Moreover, in our model we can add to the notion of delegation of power also the permission for the delegated agent to perform the action which counts as the delegated institutional fact. This can be done using the definition of permission given in [8].

Pacheco and Santos consider the delegation process among roles rather than among agents. This feature can be added to our model too, using the framework for roles we discuss in [13]. Note that our model of roles describes roles by means of beliefs and goals; it is, thus, compatible with the distribution of goals to agents described by clause 2 of Definition 11.

Gelati *et al.* [26] combine obligations and power to define the notion of mandate in contracts: "a mandate is a proclamation intended to create the obligation of exercising a declarative power". However, they do not apply their analysis to the definition of constitutive rules but to the normative positions among agents.

Comparison with other models of counts-as is discussed in [6] and [14].

Future work is studying the relation between regulative rules and delegation of power: defining how it is possible to create global policies [8] obliging or permitting other agents to delegate their power. Finally, abstraction in the input/output logic framework has been left for lions or input/output networks. In such networks each black box corresponding to an input/output logic is associated with a component in an architecture. A discussion can be found in [14].

References

1. Searle, J.: The Construction of Social Reality. The Free Press, New York (1995)

2. Grossi, D., Meyer, J.J., Dignum, F.: Counts-as: Classification or constitution? an answer using modal logic. In: Procs. of DEON'06, Berlin, Springer Verlag (2006)
3. Jones, A., Sergot, M.: A formal characterisation of institutionalised power. Journal of IGPL **3** (1996) 427–443
4. Grossi, D., Dignum, F., Meyer, J.: Contextual terminologies. In: Procs. of CLIMA'05, Berlin, Springer Verlag (2005)
5. Boella, G., van der Torre, L.: A game theoretic approach to contracts in multiagent systems. IEEE Transactions on Systems, Man and Cybernetics - Part C (2006)
6. Boella, G., van der Torre, L.: Constitutive norms in the design of normative multiagent systems. In: Procs. of CLIMA'05, Berlin, Springer Verlag (2006)
7. Georgeff, M., Ingrand, F.F.: Decision-making in an embedded reasoning system. In: Procs. of 11th IJCAI. (1989) 972–978
8. Boella, G., van der Torre, L.: Security policies for sharing knowledge in virtual communities. IEEE Transactions on Systems, Man and Cybernetics - Part A (2006)
9. Hart, H.: The Concept of Law. Clarendon Press, Oxford (1961)
10. Castelfranchi, C.: Modeling social action for AI agents. Artificial Intelligence **103(1-2)** (1998) 157–182
11. Bottazzi, E., Ferrario, R.: A path to an ontology of organizations. In: Procs. of EDOC Int. Workshop on Vocabularies, Ontologies and Rules for The Enterprise (VORTE 2005). (2005)
12. Boella, G., van der Torre, L.: From the theory of mind to the construction of social reality. In: Procs. of CogSci'05, Mahwah (NJ), Lawrence Erlbaum (2005) 298–303
13. Boella, G., van der Torre, L.: Organizations as socially constructed agents in the agent oriented paradigm. In: LNAI n. 3451: Procs. of ESAW'04, Berlin, Springer Verlag (2004) 1–13
14. Boella, G., van der Torre, L.: An architecture of a normative system. In: Procs. of DEON'06, Berlin, Springer Verlag (2006)
15. Anderson, A.: A reduction of deontic logic to alethic modal logic. Mind **67** (1958) 100–103
16. Meyer, J.J.C.: A different approach to deontic logic: Deontic logic viewed as a variant of dynamic logic. Notre Dame Journal of Formal Logic **29(1)** (1988) 109–136
17. Artikis, A., Sergot, M., Pitt, J.: An executable specification of an argumentation protocol. In: Procs. of 9th International Conference on Artificial Intelligence and Law, ICAIL 2003, New York (NJ), ACM Press (2003) 1–11
18. Boella, G., van der Torre, L.: Δ: The social delegation cycle. In: LNAI n.3065: Procs. of ΔEON'04, Berlin, Springer Verlag (2004) 29–42
19. Boella, G., van der Torre, L.: Norm governed multiagent systems: The delegation of control to autonomous agents. In: Procs. of IEEE/WIC IAT'03, IEEE Press (2003) 329– 335
20. Broersen, J., Dastani, M., Hulstijn, J., van der Torre, L.: Goal generation in the BOID architecture. Cognitive Science Quarterly **2(3-4)** (2002) 428–447
21. Makinson, D., van der Torre, L.: Input-output logics. Journal of Philosophical Logic **29** (2000) 383–408
22. Makinson, D., van der Torre, L.: Constraints for input-output logics. Journal of Philosophical Logic **30(2)** (2001) 155–185
23. Bochman, A.: Explanatory Nonmonotonic Reasoning. World Scientific Publishing, London (UK) (2005)
24. Lang, J., van der Torre, L., Weydert, E.: Utilitarian desires. Autonomous Agents and Multi-agent Systems **5(3)** (2002) 329–363
25. Pacheco, O., Santos, F.: Delegation in a role-based organization. In: Procs. of ΔEON'04, Berlin, Springer Verlag (2004) 209–227
26. Gelati, J., Rotolo, A., Sartor, G., Governatori, G.: Normative autonomy and normative co-ordination: Declarative power, representation, and mandate. Artificial Intellingence and Law **12(1-2)** (2004) 53–81