

# Norm negotiation in online multi-player games

Guido Boella · Patrice Caire · Leendert van der Torre

Received: 7 November 2007 / Revised: 16 March 2008 / Accepted: 16 May 2008  
© Springer-Verlag London Limited 2008

**Abstract** In this paper, we introduce an agent communication protocol and speech acts for norm negotiation. The protocol creates individual or contractual obligations to fulfill goals of the agents based on the so-called social delegation cycle. First, agents communicate their individual goals and powers. Second, they propose social goals which can be accepted or rejected by other agents. Third, they propose obligations and sanctions to achieve the social goal, which can again be accepted or rejected. Finally, the agents accept the new norm by indicating which of their communicated individual goals the norm achieves. The semantics of the speech acts is based on a commitment to public mental attitudes. The norm negotiation model is illustrated by an example of norm negotiation in multi-player online gaming.

**Keywords** Multi-player online games · Normative multi-agent systems · Dependence networks · Social delegation cycle · Agent communication protocols

## 1 Introduction

Online multi-player games are online spaces in which human and artificial agents interact with each other every day [5]. For example, EverQuest has more than 400,000 subscribers and has developed into a complex system of social dynamics, like the Sims Online or Second Life, something more than just a game. Massively, multi-player online games have now become among the most complex and sophisticated online social spaces in existence [36]. One aspect in which online games move toward social spaces is that players become more independent and autonomous with respect to the game designers, in the sense that they start to play in ways unforeseen by the game designers.

---

G. Boella  
Dipartimento di Informatica, Università di Torino, Turin, Italy

P. Caire · L. van der Torre (✉)  
Department of Computer Science and Communication,  
University of Luxembourg, Luxembourg, Luxembourg  
e-mail: leendert@vandortorre.com

Why, you might ask, would anyone waste four hours of their life doing this? Because a game said it couldn't be done.

This is like the Quake freaks that fire their rocket launchers at their own feet to propel themselves up so they can jump straight to the exit and skip 90% of the level and finish in 2 seconds. Someone probably told them they couldn't finish in less than a minute. Games are about challenges, about hurdles or puzzles or fights overcome. To some players, the biggest hurdle or challenge is how to do what you (the designer) said couldn't happen. If you are making a game, accept this [36].

An important aspect of autonomy, both for individuals and groups, is the creation of one's own norms, as reflected by the literal meaning of autonomy: auto-nomos. In online games, players can create their own norms to coordinate their activities, as we illustrate by a running example in this paper. Moreover, autonomy is a central concept in agent theory, and therefore multi-agent systems technology can be used to develop online games. However, existing social mechanisms for norm creation and change in multi-agent systems are restricted in the sense that either norms are not explicitly represented, or there are no communication primitives, protocols and social mechanisms to create or change norms. Normative multi-agent systems [9, 10, 14, 23] provide agents with abilities to automatically devise organizations and societies coordinating their behavior via obligations, norms and social laws. A distinguishing feature from group planning is that also sanctions and control systems for the individual or contractual obligations can be created. Since agents may have conflicting goals with respect to the norms that emerge, they can negotiate amongst each other which norm will be created. In [11], we assume that the social goals emerge, and in [12] we propose a two-step negotiation process:

- Propose the social goal and argue about that (instead of merging), and
- Negotiate the norm and sanction.

The research question of this paper is how to define a communication protocol for online multi-player games and social spaces based on the social delegation cycle [11, 12] explaining the negotiation of new social norms from cognitive agent goals in three steps. First, individual agents or their representatives negotiate social goals, then a social goal is negotiated in a social norm, and finally the social norm is accepted by an agent [22] when it recognizes it as a norm, the norm contributes to the goals of the agent, and it is obeyed by the other agents. A model of norm negotiation explains also what it means, for example, to recognize or to obey a norm, and how new norms interact with existing ones. This breaks down in the following three sub-questions:

1. How to use agent technology for the social delegation cycle referring to the individual goals of the agents, which are not accessible for other agents and which are not represented in the agent environment? We propose an additional step before the social delegation cycle, in which the agents communicate their individual goals and powers. In our communication model, the social delegation cycle does not refer to the individual private goals of the agents, but to the communicated and therefore public goals.
2. How do agents publicly accept a norm in the social delegation cycle based on public mental attitudes? We assume that each agent has to commit itself to the social norm by indicating which of its publicly communicated goals the norm achieves.
3. How to use the communication protocol in multi-player online games like EverQuest and social spaces like The Sims Online or Second Life, where the kind of norms that must be used cannot be anticipated beforehand, norms are created by human players and users, and we cannot expect that users are able to specify logical formulas as used in

most agent communication languages. We use an abstract power and dependence-based representation of the social delegation cycle.

The semantics of the speech acts is based on public mental attitudes. To keep the formal details of this paper within limits, we do not detail the operational semantics of the speech acts, but sketch informally their effects and illustrate them by a running example. Formalizations of speech acts with public mental attitudes can be found in [27,30,41].

An important subtask of the social delegation cycle is to break down goals in subgoals, using a goal hierarchy. However, to focus on the main social mechanism, in this paper, we do not detail the sub-protocol for subgoaling. Therefore, we assume that the individual goals communicated by the agents are such that they can be seen to by other agents. The extension with a protocol for subgoaling is left for further research.

The layout of this paper is as follows. In Sect. 2, we discuss norm creation in online games and social spaces, and we introduce our running example. In Sect. 3, we define our social-cognitive conceptual model of multi-agent systems and in Sect. 4, we discuss the kind of agent communication language in which we study and formalize the social delegation cycle. In Sect. 5, we formalize individual goal and power communication. In Sect. 6, we define the negotiation protocol, which we use in Sect. 7 to formalize goal negotiation, and in Sect. 8 we formalize norm negotiation. Finally, in Sect. 9, we formalize the acceptance relation. In Sect. 10, we discuss related research of an abstract model of the social delegation cycle in Tennenholtz' game-theoretic artificial social systems, our own more detailed social-cognitive model, and we discuss why these models are less suitable for communication models in online games and social spaces than our power and dependence-based model.

## 2 Multi-player online games

As a running example, we use in this paper, an example discussed by Ludlow [36] from Sony's EverQuest.

### 2.1 EverQuest

EverQuest is a multi-player online game where gamers are supposed to fight each other in a world of snakes, dragons, gods and the Sleeper. Sony intended the Sleeper to be unkillable and gave it an extreme high hit points. However, a combined combat of close to 200 players nearly succeeded to kill the animal. Unfortunately, Sony decided to intervene and rescue the monster. Most of the discussion on this example has highlighted the decrease in trust of the game players in Sony, despite the fact that the next day Sony let the game players beat the Sleeper. However, in this paper, we would like to highlight what this story tells us about the desires of game players, and its consequences for necessary technology in games. The following quote illustrates the excitement in killing the Sleeper.

A supposedly [player-vs.-player] server banded together 200 people. The chat channels across the server were ablaze, as no less than 5,000 of us listened in, with "OMG they're attempting the Sleeper! Good luck d00dz!" Everyone clustered near their screens, sharing the thrill of the fight, the nobility of the attempt and the courage of those brave 200. Play slowed to a crawl on every server as whispers turned to shouts, as naysayers predicted, "It can't be done" or "It will drop a rusty level 1 sword" and most of us just held our breath, silently urging them forward. Rumors abounded: "If they win, the whole EQ world stops and you get the text from the end of Wizardry 1," or "If

they win, the president of Sony will log on and congratulate them.” With thousands watching and waiting, the Sleeper’s health inched ever downward.

...

[EverQuest player] Ghenwivar writes, On Monday, November 17th, in the most amazing and exciting battle ever, [EverQuest guilds] Ascending Dawn, Wudan and Magus Imperialis Magicus defeated Kerafyrm, also known as The Sleeper, for the first time ever on an EverQuest server. The fight lasted approximately three hours and about 170180 players from [EverQuest server] Rallos Zek’s top three guilds were involved. Hats off to everyone who made this possible and put aside their differences in order to accomplish the impossible. Congratulations RZ!!! [36]

The example illustrates that the game had been so well wrought that a real community of players had formed, one that was able to set aside its differences, at least for a night, in pursuit of a common goal. This was not intended or foreseen by Sony, and getting 200 people to focus on accomplishing the same task is a challenge. In this paper, we study multi-agent technology to support these emerging cooperation in online games.

## 2.2 Multi-agent systems

Computer games have traditionally implemented empirical solutions to many AI problems and are now turning to more traditional AI algorithms. Cavazza [19] introduced the role of AI in gameplay, reviewed the main techniques used in current computer games such as finite-state transition networks, rule-based systems and search algorithms, described the implementation of AI in several commercial computer games, as well as academic research in AI targeting computer games applications, and discussed future trends and proposing research directions.

Another important challenge for AI in computer games is the development of on-line and multi-player gaming. In theory, the availability of human opponents around the clock would seal the fate of AI in on-line games. However, this is unlikely to be the case, as for instance autonomous actors would still be needed for the most mundane tasks in large-scale simulations and on-line games, as human players are unwilling to play such parts.

The recent advances in graphic rendering will leave more room for the future development or AI in games. With the level of realism currently achieved, further improvements in realism will come from physics and AI. The development of AI should also foster the development of new game genres based on more sophisticated user interaction with artificial actors [19].

Laird and van Lent [32] proposed that artificial intelligence for interactive computer games is an emerging application area in which this goal of human-level AI can successfully be pursued. Interactive computer games have increasingly complex and realistic worlds and increasingly complex and intelligent computer-controlled characters. They motivated their proposal of using interactive computer games, reviewed previous research on AI and games, presented the different game genres and the roles that human level AI could play within these genres, described the research issues and AI techniques that are relevant to each of these roles. Their conclusion is that interactive computer games provide a rich environment for incremental research on human-level AI.

From a researcher’s perspective, even if you are not interested in human-level AI, computer games offer interesting and challenging environments for many, more isolated,

research problems in AI. We are most interested in human-level AI, and wish to leverage computer games to rally support for research in human-level AI. One attractive aspect of working in computer games is that there is no need to attempt a “Manhattan Project” approach with a monolithic project that attempts to create human-level intelligence all at once. Computer games provide an environment for continual, steady advancement and a series of increasingly difficult challenges. Just as computers have inexorably gotten faster, computer game environments are becoming more and more realistic worlds, requiring more and more complex behavior from their characters. Now is the time for AI researchers to jump in and ride the wave of computer games [32].

### 2.3 Agent communication protocol

We use the EverQuest example in this paper to illustrate our communication model. It consists of the following four steps.

1. Agents communicate their individual goals and powers. They say which monster they would like to kill, how they want it to be killed, and which tasks they are able to see to given their powers.
2. They propose social goals which can be accepted or rejected by other agents, based on the communicated goals of the agents. The negotiation can be done by a subset of all players involved.
3. They propose obligations and sanctions to achieve the social goal, which can again be accepted or rejected. Again, the negotiation can be done by a subset of the agents involved in the coordinated attack.
4. Finally, the agents who like to participate in the coordinated attack accept the new norm by indicating which of their communicated individual goals the norm achieves.

## 3 Power viewpoint on normative multi-agent systems

In this paper, we follow the definition of power as the ability of agents to achieve goals. Thus, an agent is more powerful than another agent if it can achieve more goals.

For example, in the so-called power view on multi-agent systems [6], a multi-agent system consists of a set of agents ( $A$ ), a set of goals ( $G$ ), a function that associates with each agent the goals the agent desires to achieve (*goals*), and a function that associates with each agent the sets of goals it can achieve (*power*). To be precise, since goals can be conflicting in the sense that achieving some goals may make it impossible to achieve other goals, the function *goals* returns a set of set of goals for each set of agents. Such abstract structures have been studied as qualitative games by Wooldridge and Dunne [56], although they do not call the ability of agents to achieve goals their power. To model trade-offs among goals of agents, we introduce a priority relation among goals.

**Definition 1** Let a multi-agent system be a tuple  $\langle A, G, goals, power, \succeq \rangle$  where:

- the set of agents  $A$  and the set of goals  $G$  are two finite disjoint sets;
- $goals : A \rightarrow 2^G$  is a function that associates with each agent the goals the agent desires to achieve;
- $power : 2^A \rightarrow 2^{2^G}$  is a function that associates with each set of agents the sets of goals the set of agents can achieve;
- $\succeq : A \rightarrow \subseteq 2^G \times 2^G$  is a function that associates with each agent a partial pre-ordering on the sets of his goals;

The following example illustrates the notion of a multi-agent system for the running example.

*Example 1* The set of agents are the players that participate in the coordinated attack. We assume that the individual agents have various ways to attack the Sleeper, and they disagree about which is the best way to kill the beast. The goals of the agents are different individual ways to attack the animal. The powers of the agents are the goals each agent can achieve. The priorities of the agents are their preferences for the ways to attack the animal.

To model the role of power in norm negotiation, we extend the basic power view in a couple of ways. To model obligations we introduce a set of norms, we associate with each norm the set of agents that has to fulfill it, and for each norm we represent how to fulfill it, and what happens when it is not fulfilled. In particular, we relate norms to goals in the following two ways.

- First, we associate with each norm  $n$  a set of goals  $O(n) \subseteq G$ . Achieving these normative goals  $O(n)$  means that the norm  $n$  has been fulfilled; not achieving these goals means that the norm is violated. We assume that every normative goal can be achieved by the group, i.e., that the group has the power to achieve it.
- Second, we associate with each norm a set of goals  $V(n)$  which will not be achieved if the norm is violated (i.e., when its goals are not achieved), this is the sanction associated with the norm. We assume that the group of agents does not have the power to achieve these goals.

Since we accept norms without sanctions, we do not assume that the sanction affects at least one goal of each agent of the group the obligation belongs to.

**Definition 2** Let a normative multi-agent system be a tuple  $\langle MAS, N, O, V \rangle$  extending a multi-agent system  $MAS = \langle A, G, goals, power, \geq \rangle$  where:

- the set of norms  $N$  is a finite set disjoint from  $A$  and  $G$ ;
- $O : N \times A \rightarrow 2^G$  is a function that associates with each norm and agent the goals the agent must achieve to fulfill the norm; We assume for all  $n \in N$  and  $a \in A$  that  $O(n, a) \in power(\{a\})$ ;
- $V : N \times A \rightarrow 2^G$  is a function that associates with each norm and agent the goals that will not be achieved if the norm is violated by agent  $a$ ; We assume for each  $B \subseteq A$  and  $H \in power(B)$  that  $(\cup_{a \in A} V(n, a)) \cap H = \emptyset$ .

The normative multi-agent system is illustrated by our running example.

*Example 2* The obligations of each agent are the attacks each individual player must make. Sanctions can be that someone is removed from the coalition, has to pay a penalty (if this is possible in the game), gets a lower reputation, and so on.

An alternative way to represent normative multiagent systems replaces the function *power* by a function representing dependencies between agents. For example, a function of minimal dependence can be defined as follows. Agent  $a$  depends on agent set  $B \subseteq A$  regarding the goal  $g$  if  $g \in goals(a)$ ,  $g \notin power(\{a\})$ ,  $g \in power(B)$ , and there is no  $C \subset B$  such that  $g \in power(C)$ . Note that dependence defined in this way is more abstract than power, in the sense that we have defined dependence in terms of power, but we cannot define power in terms of dependence.

## 4 Agent communication languages

Agent communication languages share the idea that agents communicate by performing speech acts [4, 43], and that these actions can be modeled as some types of planning operators with preconditions and effects [3, 21]. Longer stretches of dialog can be explained by plans which structure attempts of participants to reach their joint goals, e.g., [34]. However, many variations of agent communication languages have been designed on these common speech act foundations, thus making the standardization effort difficult.

There is a distinction in terms of communication primitives between two main traditions in agent communication. The agent's *mental attitudes* like beliefs and intentions are included in preconditions and effects of speech acts used in information seeking dialogs using speech acts such as inform and request. This side has been popularized by the standardization efforts of the Foundation for Intelligent Physical Agents [25], although the concepts of *desire*, *goal* and *intention* are not used uniformly, and *desire* and *goal* are often used interchangeably. Below the agent social semantics tradition, developed as a reaction to FIPA, is depicted [48]. The agent can be either a creditor or a debtor of his social commitments. The semantics of speech acts used in negotiation or persuasion refer to these commitments. Although both kinds of dialog use the terminology of commitment, requesting an action in negotiation means something else than defending a proposition in persuasion after a challenge. We therefore distinguish action commitment from propositional commitment, where the former is used to give a semantics of speech acts like *request* or *propose* in negotiation [26, 48], and the latter is used for speech acts like *assert* or *challenge* in persuasion [29, 55].

In this paper, we use an approach bridging the two traditions based on a reinterpretation of the beliefs and intentions. The idea that meaning cannot be private, but is inter-subjective or based on a common ground, has been accepted for a long time in the philosophy of language. This idea of public meaning has been discussed, for example, by [33] or [49], who stress the importance of the common ground. The crucial point is that a private semantics does not make it possible for a language user to be objectively wrong about an interpretation from a third person point of view. In contrast to FIPA, beliefs and intentions are not interpreted as private mental attitudes, but as some kind of public mental attitudes, for example as grounded beliefs [27] or as ostensible beliefs or public opinions [41].

## 5 Individual goal and power communication

The first phase consists of communicating the individual goals and powers of the agents. Since existing agent communication languages do not cover inform speech acts whose content is a public goal, we introduce the speech act *desire*. The semantics of the speech act

$$desire(a, \phi)$$

is that agent  $a$  has the public goal  $\phi$ . It means that he is committed to this public goal, and, for example, he cannot say later that he desires the opposite.

If hundreds of agents have to communicate their goals, then the communication will become long and cumbersome. So, when an agent informs the others that he desires  $\phi$ , the other agents can respond by a simple 'Me too!'. Moreover, we do not have only speech acts implying that the speaking agent has a goal, but also speech acts implying that all agents have the same goal. Agent  $a$  says that everyone has the desire for  $\phi$ :

$desireall(a, \phi)$

The agents can respond to this speech act by saying ‘Not me!’. If none of them answers in this way, then the *desireall* act is equivalent to agent  $a$  stating his goal for  $\phi$ , and all the other agents saying ‘Me too!’. The principle that by default all agents agree, is sometimes called ‘silence means consent’.

An optional parameter of these two speech acts is the priority of the individual goals. In this running example we assume that there are five priorities: essential, important, intermediate, useful, possible.

*Example 3* Suppose that three agents  $\{a, b, c\}$  are going to negotiate the attack strategy to kill the Sleeper. Moreover, we assume that there are various ways in which the individual players can attack the animal: hitting the feet, throwing stones at the animal, shooting between the eyes, and so on. The three agents are negotiating for three kinds of game players with specific powers. For example, agent  $a$  is negotiating for all agents that have the power to hit the feet, throw stones, and have guns. In this example, each negotiator communicates goals they cannot see to by themselves. For example, agent  $a$  will communicate the desire to scare the Sleeper to death.

The negotiators begin by communicating their individual goals. Each player can type individual or shared goals by entering a text string into his computer. In the text box is also a list of goal priorities, which is default at intermediate. When someone types an individual goal, other agents can press the ‘Me too!’ button, When an agent enters a shared goal, other agents can press the ‘Not me!’ button, or change the suggested priority of the goal. For example, when agent  $a$  enters his goal and agent  $b$  accepts it too, and agent  $c$  communicates a shared goal but agent  $d$  refuses it, then the system translates it into the following speech acts.

$desire(b, \text{‘hitting the feet’})$   
 $desire(b, \text{‘hit him with stones’}, \textit{intermediate})$   
 $desire\ me\ too(c, \text{‘hit him with stones’}, \textit{important})$   
 $desire(c, \text{‘shoot him in the eyes’}, \textit{possible})$   
 $desire(a, \text{‘show him a picture of M.’})$   
 $desireall(a, \text{‘scare him to death’})$   
 $desire\ not\ me(c, \text{‘scare him to death’})$   
 ...

Each agent can indicate which goals it can see to. We use the new speech act  $power(a, \phi)$ , indicating that agent  $a$  communicates that he has the power to see to goal  $\phi$ . As before, the other agents can respond by ‘Me too!’ and there is a speech *powerall* which agents can respond to by saying ‘Not me!’.

*Example 4* Consider the following three speech acts.

$power(a, \text{‘hitting the feet’})$   
 $powerall(a, \text{‘hit him with stones’})$   
 $power\ not\ me(c, \text{‘hit him with stones’})$   
 $powerall(c, \text{‘shoot him in the eyes’})$   
 ...

The communicated goals and powers are represented in a single table for each player, as shown in Table 1. In case the agents are negotiating a norm for a large set of players, then not all individual names of the agents are given, but only aggregate information such as the number of players having the power to see to a goal.



**Table 1** Communicated goals and powers

# Goal	Speaker	Agents and priorities	Goal description	Powers
1	<i>b</i>	<i>b, c</i>	Hitting the feet	<i>a</i>
2	<i>b</i>	<i>b</i> : intermediate, <i>c</i> : important	Hit him with stones	<i>a, b</i>
3	<i>c</i>	<i>c</i> : possible	Shoot him in the eyes	<i>a, b, c</i>
4	<i>a</i>	<i>a</i>	Show him a picture of M	<i>b, c</i>
5	<i>a</i>	<i>a, b</i>	Scare him to death	<i>c</i>
...	...	...	...	...

## 6 Generic negotiation protocol

A negotiation protocol is described by a set of sequences of negotiation actions which lead to either success or failure. In this paper, we only consider protocols in which the agents propose a so-called deal, and when an agent has made such a proposal, then the other agents can either accept or reject it (following an order  $\succ$  of the agents). Moreover, they can also end the negotiation process without any result.

**Definition 3** (*Negotiation Protocol*) A negotiation protocol is a tuple  $\langle Ag, deals, actions, valid, finished, broken, \succ \rangle$ , where

- the agents *Ag*, *deals* and *actions* are three disjoint sets, such that  $actions = \{propose(a, d), accept(a, d), reject(a, d) \mid a \in Ag, d \in deals\} \cup \{breakit(a) \mid a \in Ag\}$ .
- *valid*, *finished*, *broken* are sets of finite sequences of *actions*.

We now instantiate this generic protocol for negotiations in normative multi-agent systems. We assume that a sequence of actions (a history) is valid when each agent does an action respecting the order defined on agents. Then, after each proposal, the other agents have to accept or reject this proposal, again respecting the order, until they all accept it or one of them rejects it. When it is an agent’s turn to make a proposal, it can also end the negotiation by breaking it. The history is *finished* when all agents have accepted the last deal, and *broken* when the last agent has ended the negotiations.

**Definition 4** (*NMAS protocol*) Given a normative multi-agent system  $\langle MAS, N, O, V \rangle$  extending a multi-agent system  $MAS = \langle A, G, goals, power, \geq \rangle$ , a negotiation protocol for *NMAS* is a tuple  $NP = \langle A, deals, actions, valid, finished, broken, \succ \rangle$ , where:

- $\succ \subseteq A \times A$  is a total order on *A*,
- a history *h* is a sequence of actions, and *valid*(*h*) holds if:
  - the *propose*() and *breakit*() actions in the sequence respect  $\succ$ ,
  - each *propose*() is followed by a sequence of *accept*() or *reject*() actions respecting  $\succ$  until either all agents have accepted the deal or one agent has rejected it,
  - there is no double occurrence of a proposal *propose*(*a, d*) in the same deal by any agent *a*  $\in$  *Ag*, and
  - the sequence *h* ends iff either all agents have accepted the last proposal (*finished*(*h*)) or the last agent has broken the negotiation [*broken*(*h*)] instead of making a new proposal.

In theory we can add additional penalties when agents break the negotiation. However, since it is in the interest of all agents to reach an agreement, we do not introduce such sanctions. In this respect norm negotiation differs from negotiation about obligation distribution [8], where it may be the interest of some agents to see to it that no agreement is reached. In such cases, sanctions must be added to the negotiation protocol to motivate the agents to reach an agreement.

*Example 5* Assume three agents and the following history.

$action_1 : propose(a, d_1)$

$action_2 : accept(b, d_1)$

$action_3 : reject(c, d_1)$

$action_4 : propose(b, d_2)$

$action_5 : accept(c, d_2)$

$action_6 : accept(a, d_2)$

We have  $valid(h)$ , because the order of action respects  $\succeq$ , and we have  $accepted(h)$ , because the history ends with acceptance by all agents ( $action_5$  and  $action_6$ ) after a proposal ( $action_4$ ).

The open issue of the generic negotiation protocol is the set of deals which can be proposed. They depend on the kind of negotiation. In social goal negotiation the deals represent a social goal, and in norm negotiation the deals contain the obligations of the agents and the associated control system based on sanctions. This is illustrated in our running example in the following sections.

## 7 Social goal negotiation

We characterize the allowed deals during goal negotiation as a set of goal which contains for each agent a goal it desires. Moreover, we add two restrictions. First, we only allow goals the agents have the power to achieve. Then, we have to consider the existing normative system, which may already contain the norms that look after the goals of the agents. We therefore restrict ourselves to new goals. Additional constraints may be added, for example excluding goals an agent can see to itself. However, since such additional restrictions may be unrealistic in some applications (e.g., one may delegate some tasks to a secretary even when one has the power to see to these tasks oneself), we do not consider such additional constraints (Table 2).

**Definition 5** (*Deals in goal negotiation*) In the goal negotiation protocol, a deal  $d \in deals$  is a set of goals satisfying the following restrictions:

1.  $d \in power(A)$
2. for all  $a \in A$  there exists some  $g \in d$  such that
  - (a)  $g \in goals(a)$
  - (b) there does not exist a norm  $n$  in  $N$  such that  $g \in \cup_{a \in A} O(n, a)$

The following example illustrates social goal negotiation in the running example.

*Example 6* The agents proposing a deal select a couple of individual communicated goals from Table 1 by clicking on them. The system then send them to the other agents, which can either accept them, or reject them. When they reject a communicated goal, they can add a reason to it, which is communicated to the other agents.

Let  $MAS = (\{a, b, c\}, \{g_1, g_2, g_3, g_4, g_5\}, goals, power, \succeq)$  be a multi-agent system representation of Table 1, where

**Table 2** Social goal negotiation

# Goal proposal	Speaker	Goals	Accepted	Rejected	Reason
1	<i>a</i>	2, 5	<i>a</i>	<i>b</i>	...
2	<i>b</i>	2, 4, 5	<i>a, b, c</i>		
...	...	...	...	...	

power:  $power(a) = \{\{g_1\}, \{g_2\}, \{g_3\}\}$ ,  $power(b) = \{\{g_2\}, \{g_3\}, \{g_4\}\}$ ,  $power(c) = \{\{g_3\}, \{g_4\}, \{g_5\}\}$ , if  $G_1 \in power(A)$  and  $G_2 \in power(B)$  then  $G_1 \cup G_2 \in power(A \cup B)$ . Agent *a* has the power to achieve goals  $g_1, g_2, g_3$ , agent *b* has the power to achieve goals  $g_2, g_3, g_4$ , and agent *c* can achieve goals  $g_3, g_4, g_5$ . There are no conflicts among goals.

goals:  $goals(a) = \{g_4, g_5\}$ ,  $goals(b) = \{g_1, g_5\}$ ,  $goals(c) = \{g_1, g_2\}$ . Each agent desires the tasks it cannot perform itself.

Moreover, let  $NMAS = \langle MAS, N, O, V \rangle$  be a normative multi-agent system with  $N = \{n\}$ ,  $O(n, a_1) = \{g_1\}$ . Since there has to be some benefit for agent *b* and *c*, the goals  $g_5$  and  $g_2$  have to be part of the social goal. Therefore, social goals (i.e., possible deals) are  $\{g_2, g_5\}$  and  $\{g_2, g_4, g_5\}$ .

Finally, consider the negotiation. Assuming that agent *a* is first in the order  $\succ$ , he may propose  $\{g_2, g_5\}$ . The other agents may accept this, or reject it and agent *b* will then propose  $\{g_2, g_4, g_5\}$ . The latter would be accepted by all agents, as they know that according to the protocol no other proposals can be made.

The example illustrates that the negotiation does not determine the outcome, in the sense that there are multiple outcomes possible. Additional constraints may be added to the negotiation strategy to further delimit the set of possible outcomes.

### 8 Social norm negotiation

We formalize the allowed deals during norm negotiation as obligations for each agent to see to some goals, such that all goals of the social goal are included. Again, to determine whether the obligations imply the social goal, we have to take the existing normative system into account. We assume that the normative system only creates obligations that can be fulfilled together with the already existing obligations.

**Definition 6 (Fulfillable NMAS)** A normative multi-agent system  $\langle MAS, N, O, V \rangle$  extending a multi-agent system  $MAS = \langle A, G, goals, power, \succeq \rangle$  can be fulfilled if there exists a  $G' \in power(A)$  such that all obligations are fulfilled  $\cup_{n \in N, a \in A} O(n, a) \subseteq G'$ .

Creating a norm entails adding obligations and violations for the norm.

**Definition 7 (Add norm)** Let  $NMAS$  be a normative multi-agent system  $\langle MAS, N, O, V \rangle$  extending a multi-agent system  $MAS = \langle A, G, goals, power, \succeq \rangle$ . Adding a norm  $n \notin N$  with a pair of functions  $\langle o_1, o_2 \rangle$  for obligation  $o_1 : A \rightarrow 2^G$  and for sanction  $o_2 : A \rightarrow 2^G$  leads to the new normative multi-agent system  $\langle MAS, N \cup \{n\}, O \cup o_1(n), V \cup o_2(n) \rangle$ .

Moreover, if every agent fulfills its obligation, then the social goal is achieved.

**Table 3** Negotiated norm

# Goal	Speaker	Agents and priorities	Goal description	Powers	Obligation
2	<i>b</i>	<i>b</i> : intermediate, <i>c</i> : important	Hit him with stones	<i>a, b</i>	<i>a</i>
4	<i>a</i>	<i>a</i>	Show him a picture of M	<i>b, c</i>	<i>b</i>
5	<i>a</i>	<i>a, b</i>	Scare him to death	<i>c</i>	<i>c</i>

**Definition 8** (*Deals in norm negotiation*) In the norm negotiation protocol, a deal  $d \in \text{deals}$  for social goal  $S$  is a pair of functions  $\langle o_1, o_2 \rangle$  for obligation  $o_1 : A \rightarrow 2^G$  and for sanction  $o_2 : A \rightarrow 2^G$  satisfying the following conditions:

1. Adding  $\langle o_1, o_2 \rangle$  to  $NMAS$  for a fresh variable  $n$  (i.e., not occurring in  $N$ ) leads again to a normative multi-agent system  $NMAS'$ ;
2.  $NMAS'$  achieves the social goal,  $\cup_{a \in A} o_1(a) = S$ .
3. If  $NMAS$  is fulfillable, then  $NMAS'$  is too.

The running example illustrates the norm negotiation protocol.

*Example 7* For the social goal  $\{g_2, g_5\}$ , there will always be one of the agents who does not have to see to any goal. Consider norm for the social goal  $\{g_2, g_4, g_5\}$  in Table 3. A possible solution here is that each agent sees to one of the goals.

Sanctions can be added in the obvious way. In the norm negotiation as defined thus far, the need for sanctions has not been formalized yet. For this need, we have to consider the acceptance of norms.

## 9 Norm acceptance

An agent accepts a norm when it believes that the other agents will fulfill their obligations, and the obligation implies the goals the cycle started with. For the former we use the quasi-stability of the norm (e.g., if the norm is a convention, then we require that the norm is a Nash equilibrium). Each agent  $b$  fulfills the norm *given that all other agents fulfill the norm*. Again we have to take the existing normative system into account, so we add the condition that all other norms are fulfilled. In general, it may mean that an agent does something which it does not like to do, but it fears the sanction more than this dislike. We use the trade-off among goals  $\geq$ .

**Definition 9** (*Stability*) A choice  $c$  of agent  $b \in A$  in  $NMAS$  with new norm  $n$  is  $c \in \text{power}(b)$  such that  $\cup_{m \in N \setminus \{n\}} O(m, b) \subseteq d$ . The choices of the other agents are  $oc = \cup_{a \in A \setminus \{b\}, m \in N} O(n, a) \cup V(n, a)$ . The effect of choice  $c$  is  $c \cup oc \cup V(n, a)$  if  $O(n) \subseteq c$ ,  $c \cup oc$  otherwise.  $NMAS$  is stable if  $\forall b \in A$ , there is a choice  $c$  such that  $O(n, b) \subseteq c$ , and there is no choice  $c' \geq (b)c$  with  $O(n, b) \not\subseteq c'$ .

Finally, we have to test whether the new situation is better than the old one for all agents. For example, we may request that the outcome in both the original multi-agent system as in the new multi-agent system is a Nash equilibrium, and we demand that each Nash outcome in the new system is better than each Nash outcome in the original normative multi-agent system. The formalization of these concepts is along the same lines as the definition of acceptance in Definition 9.

**Table 4** Norm Acceptance

# Goal	Speaker	Agents and priorities	Goal description	Powers	Obligation	Norm acceptance
2	<i>b</i>	<i>b</i> : intermediate, <i>c</i> : important	Hit him with stones	<i>a, b</i>	<i>a</i>	<i>c</i>
4	<i>a</i>	<i>a</i>	Show him a picture of M	<i>b, c</i>	<i>b</i>	<i>a</i>
5	<i>a</i>	<i>a, b</i>	Scare him to death	<i>c</i>	<i>c</i>	<i>b</i>

We ask that each agent communicates to the other agents why it accepts the norm. By making a public statement, it thus commits to the norm and gives an explanation of its acceptance. Agent *a* accepts the norm *n* for goal *g*:

$$accept\text{-}norm(a, n, g)$$

Norm acceptance is illustrated in our running example.

*Example 8* Let *n* stand for the norm accepted by the agents. Assume that every agent can select the goal for which he accepts the norm from the list of all goals of the social goal. This can be communicated to the other agents using the following three speech acts.

$$accept\text{-}norm(a, n, 4)$$

$$accept\text{-}norm(b, n, 5)$$

$$accept\text{-}norm(c, n, 2)$$

The resulting table is given in Table 4.

## 10 Related work

In this section, we discuss the existing three formalizations of the social delegation cycle.

### 10.1 Norms are a class of constraints deserving special analysis

Meyer and Wieringa define normative systems as “systems in the behavior of which norms play a role and which need normative concepts in order to be described or specified” [40, preface]. Alchourròn and Bulygin [2] define a normative system inspired by Tarskian deductive systems:

When a deductive correlation is such that the first sentence of the ordered pair is a case and the second is a solution, it will be called normative. If among the deductive correlations of the set  $\alpha$  there is at least one normative correlation, we shall say that the set  $\alpha$  has normative consequences. A system of sentences which has some normative consequences will be called a normative system [2, p. 55].

Jones and Carmo [31] introduced agents in the definition of a normative system by defining it as “sets of agents whose interactions are norm-governed; the norms prescribe how the agents ideally should and should not behave. [...] Importantly, the norms allow for the possibility that actual behavior may at times deviate from the ideal, i.e., that violations of obligations, or of agents’ rights, may occur.” Since the agents’ control over the norms is not explicit here, we use the following definition.

A normative multi-agent system is a multi-agent system together with normative systems in which agents can decide whether to follow the explicitly represented norms, and the normative systems specify how and in which extent the agents can modify the norms [14].

Note that this definition makes no presumptions about the internal architecture of an agent or of the way norms find their expression in agent's behavior.

### 10.1.1 Representation of norms

Since norms are explicitly represented, according to our definition of a normative multi-agent system, the question should be raised how norms are represented. Norms can be interpreted as a special kind of constraint, and represented depending on the domain in which they occur. However, the representation of norms by domain-dependent constraints runs into the question what happens when norms are violated. Not all agents behave according to the norm, and the system has to deal with it. In other words, norms are not hard constraints, but soft constraints. For example, the system may sanction violations or reward good behavior. Thus, the normative system has to monitor the behavior of agents and enforce the sanctions. Also, when norms are represented as domain-dependent constraints, the question will be raised how to represent permissive norms, and how they relate to obligations. Whereas obligations and prohibitions can be represented as constraints, this does not seem to hold for permissions. For example, how to represent the permission to access a resource under an access control system? Finally, when norms are represented as domain-dependent constraints, the question can be raised how norms evolve.

We therefore believe that norms should be represented as a domain independent theory. For example, deontic logic [37–39, 51, 52, 54] studies logical relations among obligations and permissions, and more in particular violations and contrary-to-duty obligations, permissions and their relation to obligations, and the dynamics of obligations over time. Therefore, insights from deontic logic can be used to represent and reason with norms in multi-agent systems. Deontic logic also offers representations of norms as rules or conditionals. However, there are several aspects of norms which are covered neither by constraints nor by deontic logic, such as the relation among the cognitive abilities of agents and the global properties of norms. Meyer and Wieringa explain why normative systems are intimately related with deontic logic.

Until recently in specifications of systems in computational environments the distinction between normative behavior (as it *should be*) and actual behavior (as it *is*) has been disregarded: mostly it is not possible to specify that some system behavior is non-normative (illegal) but nevertheless possible. Often illegal behavior is just ruled out by specification, although it is very important to be able to specify what should happen if such illegal but possible behaviors occurs! Deontic logic provides a means to do just this by using special modal operators that indicate the status of behavior: that is whether it is legal (normative) or not [40, preface].

### 10.1.2 Norms and agents

Conte et al. [22] distinguished two distinct sets of problems in normative multi-agent systems research. On the one hand, they claimed that legal theory and deontic logic supply a theory of norm-governed interaction of autonomous agents while at the same time lacking a model that integrates the different social and normative concepts of this theory. On the other hand, they claim that three other problems are of interest in multi-agent systems research on norms: how agents can acquire norms, how agents can violate norms, and how an agent can be autonomous. Agent decision making in normative systems and the relation between desires and obligations has been studied in agent architecture [16], which thus explain how norms and obligations influence agent behavior.

An important question in normative multi-agent systems is where norms come from. Norms are not necessarily created by legislators, but they can also be negotiated among agents, or they can emerge spontaneously, making the agents norm autonomous [53]. In electronic commerce research, for example, cognitive foundations of social norms and contracts are studied [9]. Protocols and social mechanisms are now being developed to support such creations of norms in multi-agent systems: formal methods designed for protection against security violations [20], communication protocols for transactions [1] and kinship-based cooperation between agents for more effectiveness [42]. Moreover, agents like legislators playing a role in the normative system have to be regulated themselves by procedural norms [13], raising the question how these new kind of norms are related to the other kinds of norms.

When norms are created, the question can be raised how they are enforced. For example, when a contract is violated, the violator may have to pay a penalty. But then there has to be a monitoring and sanctioning system, for example police agents in an electronic institution. Such protocols or roles in a multi-agent system are part of the construction of social reality, and Searle [44] has argued that such social realities are constructed by constitutive norms. This raises the question how to represent such constitutive or counts-as norms, and how they are related to regulative norms like obligations and permissions [10].

### 10.1.3 Norms and other concepts

Not only the relation between norms and agents but also the relation between norms and other social and legal concepts must be studied. How do norms structure organizations? How do norms coordinate groups and societies? How about the contract frames in which contracts live? How about the relation between legal courts? though in some normative multi-agent systems there is only a single normative system, there can also be several of them, raising the question how normative systems interact. For example, in a virtual community of resource providers each provider may have its own normative system, which raises the question how one system can authorize access in another system, or how global policies can be defined to regulate these local policies [10].

## 10.2 Social delegation cycle using artificial social systems

The problem studied in artificial social systems is the design, emergence or more generally the creation of social laws. Shoham and Tennenholtz [45] introduced social laws in a setting without utilities, and they defined *rational* social laws as social laws that improve a social game variable [46]. We follow Tennenholtz' presentation for stable social laws [50]. In Tennenholtz' game-theoretic artificial social systems, the goals or desires of agents are represented by their utilities. A game or multi-agent encounter is a set of agents with for each agent a set of strategies and a utility function defined on each possible combination of strategies. The social goal is represented by a minimal value for the social game variable. Tennenholtz [50] used the maximin value as game variable. This represents safety level decisions, in the sense that the agent optimizes its worst outcome assuming the other agents may follow any of their possible behaviors.

The social norm is represented by a social law, characterized as a restriction to the strategies available to the agents. It is *useful* with respect to an efficiency parameter  $e$  if each agent can choose a strategy that guarantees it a payoff of at least  $e$ . A social law is *quasi-stable* if an agent does not profit from violating the law, as long as the other agent conforms to the social law (i.e., selects strategies allowed by the law). Quasi-stable conventions correspond

to Nash equilibria. The efficiency parameter can be seen as a social kind of *utility aspiration level*, as studied by Simon [47]. Such aspiration levels have been studied to deal with limited or resource-bounded reasoning, and have led to the development of goals and planning in artificial intelligence; we therefore use a goal based ontology in this paper. The three steps of the social delegation cycle in this classical game-theoretic setting can be represented as follows. Goal negotiation implies that the efficiency parameter is higher than the utility the agents expect without the norm, for example represented by the Nash equilibria of the game. Norm negotiation implies that the social law is useful (with respect to the efficiency parameter). The acceptance relation implies that the social law is quasi-stable.

The game-theoretical model has several drawbacks as a basis of our communication model. Due to the uniform description of agents in the game-theoretic model, it is less clear how to distinguish among kinds of agents. For example, the unique utility aspiration level does not distinguish the powers of agents to negotiate a better deal for themselves than for the other agents. Moreover, the formalization of the social delegation cycle does neither give a clue on how the efficiency parameter is negotiated, nor on how the social law is negotiated. For example, the goals or desires of the agents as well as other mental attitudes may play a role in this negotiation. There is no sanction or control system in the model (adding a normative system to encode enforceable social laws to the artificial social system complicates the model [7]). Finally, an additional drawback is that the three ingredients of the model (agent goals, social goals, and social laws) are formalized in three completely different ways.

### 10.3 Power and dependencies

The second formalization of the social delegation cycle is the highly detailed model we have proposed within normative multi-agent systems [11]. The challenge to define social mechanisms, as we see it, is to go beyond the classical game theoretic model by introducing social and cognitive concepts and a negotiation model, but doing so in a minimal way. In the model proposed in this paper, we therefore keep goals and obligations abstract and we do not describe them by first-order (or propositional) logic or their rule structure, we do not introduce decisions, actions, tasks, and so on. Similar concerns are also mentioned by Wooldridge and Dunne in their qualitative game theory [56].

In [12] we introduce a norm negotiation model based on power and dependence structures. Power may affect norm negotiation in various ways, and we therefore propose to analyze the norm negotiation problem in terms of social concepts like power and dependence. Power has been identified as a central concept for modeling social phenomena in multi-agent systems by various authors [15, 17, 24, 35], as Castelfranchi observed both to enrich agent theory and to develop experimental, conceptual and theoretical new instruments for the social sciences [18].

The model in [12] is based on a distinction between social goal negotiation and the negotiation of the obligations with their control system. Roughly, the social goals are the benefits of the new norm for the agents, and the obligations are the costs of the new norm for the agents in the sense that agents risk being sanctioned. Moreover, in particular, when representatives of the agents negotiate the social goals and norms, the agents still have to accept the negotiated norms. The norm is accepted when the norm is quasi-stable in the sense that agents will act according to the norm, and effective in the sense that fulfillment of the norm leads to achievement of the agents' desires—i.e., when the benefits outweigh the costs.

Our new model is based on a minimal extension of Tennenholtz' game theoretic model of the social delegation cycle. We add a negotiation protocol, sanction and control, and besides acceptance also effectiveness. It is minimal in the sense that, compared to our earlier model



[11] in normative multi-agent systems, we do not represent the rule structure of norms, we do not use decision variables, and so on. Also, as discussed in this paper, we do not add goal hierarchy, definitions, etc. The model therefore focuses on various uses of power: the power as ability to achieve goals and in negotiation.

## 11 Summary

Massively, multi-player online games have now become among the most complex and sophisticated online social spaces in existence [5,36]. As the lives people lead in virtual worlds become more and more connected with their offline lives such online spaces are shedding significant light on questions of liberty and responsibility, and on how an online existence reverberates in what we think of as a gamer's real life [28,36]. A restricted version of norm creation is built in some of the games, such as for example *A Tale in the Desert*, but we envision in this paper a gaming environment in which the players can create arbitrary norms to coordinate their behaviors.

In this paper, we introduce an agent communication protocol and speech acts for norm negotiation. It creates individual or contractual obligations fulfilling goals of the agents based on the social delegation cycle. The first problem is that the social delegation cycle refers to the individual goals of the agents, which are not accessible to other agents and which are not represented in the agent environment. This problem is solved by the first step of our communication protocol.

1. First, agents communicate their individual goals and powers.
2. Second, they propose social goals which can be accepted or rejected by other agents.
3. Third, they propose obligations and sanctions to achieve the social goal, which can again be accepted or rejected.
4. Finally, the agents accept the new norm by indicating which of their communicated individual goals the norm achieves.

Consequently, the social delegation cycle does not refer to the individual private goals of the agents, but to the communicated public goals.

The second problem, to define norm acceptance, is solved by the final step of our norm negotiation protocol, in which each agent has to commit itself to the social norm by indicating which of its publicly communicated goals the norm achieves.

Finally, the solution of the third problem, that is, how to use the norm negotiation protocol in online multi-player games and social spaces like *The Sims Online* or *Second Life*, is to use an abstract power and dependence based representation of the social delegation cycle. Clearly the game players cannot detail their individual probabilities and utilities as required by Tennenholtz' game-theoretic model, and they cannot detail the decision variables of the model in [11]. However, they can detail their goals and powers, as they can use natural language to describe their goals, and detail their powers by selecting the relevant goals. This solution is illustrated by our running example of norm negotiation in multi-player online gaming, where the goals are lines of text entered in a table, powers are cross-tables, social goals are sets of individual goals, and norms are distributions of the goals over the agents.

An important subtask of the social delegation cycle is to break down goals in subgoals, using a goal hierarchy. However, to focus on the main social mechanism, in this paper we do not detail the sub-protocol for subgoaling. Therefore, we assume that the individual goals communicated by the agents are such that they can be seen to by other agents. The extension with a protocol for subgoaling is left for further research.

## References

1. Albers M, Jonker CM, Karami M, Treur J (2004) Agent models and different user ontologies for an electronic market place. *Knowl Inf Syst* 6(1):1–41
2. Alchourrón CE, Bulygin E (1971) *Normative systems*. Springer, Wien
3. Allen JF, Perrault CR (1980) Analyzing intention in utterances. *Artif Intell* 15:143–178
4. Austin JL (1962) *How to do things with words*. Harvard University Press, Cambridge
5. Beau F (2007) *Culture d'Univers*. FYP editions, Limoges, France
6. Boella G, Sauro L, van der Torre L (2004) Social viewpoints on multiagent systems. In: *Proceedings of the third international joint conference on autonomous agents and multi agent systems (AAMAS'04)*, pp 1358–1359
7. Boella G, van der Torre L (2005) Enforceable social laws. In: *Proceedings of the 4th international joint conference on autonomous agents and multiagent systems (AAMAS'05)*, pp 682–689
8. Boella G, van der Torre L (2006) Fair distribution of collective obligations. In: *Proceedings of the 17th European conference on artificial intelligence (ECAI'06)*, pp 721–722
9. Boella G, van der Torre L (2006) A game theoretic approach to contracts in multiagent systems. *IEEE Trans Syst Man Cybern C* 36(1):68–79
10. Boella G, van der Torre L (2006) Security policies for sharing knowledge in virtual communities. *IEEE Trans Syst Man Cybern A* 36(3):439–450
11. Boella G, van der Torre L (2007) Norm negotiation in multiagent systems. *Int J coop Inf Syst (IJCIS)* 16(1):97–122
12. Boella G, van der Torre L (2007) Power in norm negotiation. In: *Agent and multi-agent systems: technologies and applications. First KES international symposium, KES-AMSTA 2007, LNCS, vol 4496*. Berlin, Springer, pp 436–446
13. Boella G, van der Torre L (2008) Substantive and procedural norms in normative multiagent systems. *J Appl Logic* 6(2):152–171
14. Boella G, van der Torre L, Verhagen H (2006) Introduction to normative multiagent systems. *Comput Math Organ Theory Spec Issue Normative Multiag Syst* 12(2–3):71–79
15. Brainov S, Sandholm T (1999) Power, dependence and stability in multi-agent plans. In: *Proceedings of the 21st national conference on artificial intelligence (AAAI'99)*, pp 11–16
16. Broersen J, Dastani M, Hulstijn J, van der Torre L (2002) Goal generation in the BOID architecture. *Cogn Sci Q* 2(3–4):428–447
17. Castelfranchi C (1998) Modeling social action for AI agents. *Artif Intell* 103(1, 2):157–182
18. Castelfranchi C (2003) Micro–macro constitution of power. *Protosociology* 18–19:208–268
19. Cavazza M (2000) AI in computer games: survey and perspectives. *Virtual Real* 5:223–235
20. Chen Q, Zhang C, Zhang S (2005) Endl: a logical framework for verifying secure transaction protocols. *Knowl Inf Syst* 7(1):84–109
21. Cohen PR, Perrault CR (1979) Elements of a plan-based theory of speech acts. *Cogn Sci* 3:177–212
22. Conte R, Castelfranchi C, Dignum F (1999) Autonomous norm-acceptance. In: *Intelligent agents V (ATAL'98)*, LNAI, vol 1555. Springer, Berlin, pp 99–112
23. Conte R, Falcone R, Sartor G (1999) Agents and norms: how to fill the gap? *Artif Intell Law* 7(1):1–15
24. Conte G, Sichman J (2002) Multi-agent dependence by dependence graphs. In: *Proceedings of the first international joint conference on autonomous agents and multiagent systems (AAMAS'02)*, pp 483–490
25. FIPA (2002) FIPA ACL communicative act library specification. Technical report SC00037J. Foundation for intelligent physical agents
26. Fornara N, Colombetti M (2004) A commitment-based approach to agent communication. *Appl Artif Intell* 18(9–10):853–866
27. Gaudou B, Herzog A, Longin D (2006) A logical framework for grounding-based dialogue analysis. *Electron Notes Theor Comput Sci* 157(4):117–137
28. Grimmelmann J (2006) Virtual borders: the interdependence of real and virtual worlds. *First Monday*, 11(2) URL: [http://firstmonday.org/issues/issue11\\_2/grimmelmann/index.html](http://firstmonday.org/issues/issue11_2/grimmelmann/index.html)
29. Hamblin CL (1970) *Fallacies*. Methuen & co, London
30. Hulstijn J, Boella G, Damiano R, van der Torre L (2007) A common ontology of agent communication languages: modeling mental attitudes and social commitments using roles. *Appl Ontol* 2(3–4):217–265
31. Jones A, Carmo J (2001) Deontic logic and contrary-to-duties. In: Gabbay D, Guenther F (eds) *Handbook of philosophical logic*, vol 3. Kluwer, Dordrecht, pp 203–279
32. Laird J, van Lent M (2001) Human-level AI's killer application: interactive computer games. *AI Mag* 22(2):15–25
33. Lewis D (1969) *Convention: a philosophical study*. Harvard University Press, Cambridge

34. Lochbaum KE (1998) A collaborative planning model of intentional structure. *Comput Linguist* 24(4):525–572
35. Lopez y Lopez F (2003) Social power and norms: impact on agent behaviour. Ph.D. thesis, University of Southampton
36. Ludlow P, Wallace M (2007) *The Second Life Herald*. MIT Press, Cambridge
37. Makinson D, van der Torre L (2000) Input–output logics. *J Philos Logic* 29:383–408
38. Makinson D, van der Torre L (2001) Constraints for input–output logics. *J Philos Logic* 30:155–185
39. Makinson D, van der Torre L (2003) Permissions from an input–output perspective. *J Philos Logic* 32(4):391–416
40. Meyer J, Wieringa R (1993) *Deontic logic in computer science: normative system specification*. Wiley, Chichester, England
41. Nickles M, Fischer F, Weiss G (2006) Communication attitudes: a formal approach to ostensible intentions, and individual and group opinions. *Electron Notes Theor Comput Sci* 157(4):95–115
42. Oh JC (2005) Emergence of cooperative internet server sharing among internet search agents caught in the  $\langle i \rangle$ - $\langle i \rangle$ -person prisoner's dilemma game. *Knowl Inf Syst* 7(1):23–55
43. Searle JR (1969) *Speech acts: an essay in the philosophy of language*. Cambridge University Press, Cambridge
44. Searle JR (1995) *The construction of social reality*. Free Press, New York
45. Shoham Y, Tennenholtz M (1995) On social laws for artificial agent societies: off-line design. *Artif Intell* 73(1–2):231–252
46. Shoham Y, Tennenholtz M (1997) On the emergence of social conventions: modeling, analysis and simulations. *Artif Intell* 94(1–2):139–166
47. Simon HA (1955) A behavioral model of rational choice. *Q J Econ* 69:99–118
48. Singh MP (2000) A social semantics for agent communication languages. In: *Issues in agent communication (AC'00)*, LNCS 1916. Springer, Berlin, pp 31–45
49. Stalnaker RC (2002) Common ground. *Linguist Philos* 25:701–719
50. Tennenholtz M (1998) On stable social laws and qualitative equilibria. *Artif Intell* 102(1):1–20
51. van der Torre L (2003) Contextual deontic logic: normative agents, violations and independence. *Ann Math Artif Intell* 37(1–2):33–63
52. van der Torre L, Tan Y (1999) Contrary-to-duty reasoning with preference-based dyadic obligations. *Ann Math Artif Intell* 27(1–4):49–78
53. Verhagen H (1999) On the learning of norms. In: *Proceedings of modelling autonomous agents in a multi-agent world (MAAMAW'99)*
54. von Wright GH (1951) Deontic logic. *Mind* 60:1–15
55. Walton DN, Krabbe EC (1995) *Commitment in dialogue: basic concepts of interpersonal reasoning*. State University of New York Press, New York
56. Wooldridge M, Dunne PE (2004) On the computational complexity of qualitative coalitional games. *Artif Intell* 158(1):27–73

## Author biographies



**Guido Boella** is received the Ph.D. degree at the University of Torino in 2000. He is currently professor at the Department of Computer Science of the University of Torino. His research interests include multi-agent systems, in particular, normative systems, institutions and roles using qualitative decision theory. He organized the first workshops on normative multi-agent systems NorMAS, on coordination and organization CoOrg, and the AAAI fall symposium and ECOOP roles.



**Patrice Caire** is currently Ph.D. student at the University of Luxembourg. She received a Master in Computer Science from New York University. Her research interests include digital cities, virtual communities, dependence networks and tools for conviviality.



**Leendert van der Torre** is currently full-time professor in intelligent systems at the University of Luxembourg. He received the Ph.D. degree from Erasmus University Rotterdam in 1997. He has developed the so-called input/output logics and the BOID agent architecture. His current research interests include deontic logic, qualitative game theory and coordination in normative multiagent systems. He is responsible for priority P1 on security and reliability of University of Luxembourg.