Normative Agents in Health Care: Uses and Challenges

Javier Vázquez-Salceda

institute of information and computing sciences, utrecht university

technical report UU-CS-2004-038

www.cs.uu.nl

Normative Agents in Health Care: Uses and Challenges

Javier Vázquez-Salceda [1]

Abstract.
The use of Multi-Agent Systems (MAS) in health-care domains is increasing. Such Agent-mediated Medical Systems are designed to manage complex tasks and have the potential to adapt gracefully to unexpected events. However, in that kind of systems the issues of privacy and security (in the access to patient records), safety and soundness (of the individual agent behaviours and the multiagent system as a whole) and trust (among heterogeneous agents and among users and agents) are particularly sensitive. An additional issue is that health care systems are highly regulated by regional, national and European regulations and policies. Therefore several normative contexts (European, national, regional) should be taken into account while designing agent-mediated health care systems. In this paper the less explored normative problem of the application of agents in Health Care will be presented. The impact on both the individual agents and the agent platforms will be discussed. On the individual agent side, Normative Agents (agents whose behaviour is guided by collections of norms) will be presented. On the agent platforms side, the concept of Electronic Institutions (as normative environments to increase trust and ensure proper agent behaviour) will be presented, along with some methodological considerations, the impact on implementation and some of the technologies needed to be developed in the future.

1 INTRODUCTION

The raise of what is called the *Information Society* (IS) joined with the need of promoting Health services (mainly in the United States of America, where Health is just another market) are the origin of a change in the way the health service providers are managed and presented, from a institutional-centered approach to a patient-centered one, in order to create a personalized environment to attract patients. The connectivity these modern networks provide is also creating new services or modifying existing ones to be able to operate not only at local but also at trans-national level.

1.1 The European e-Health Area

One of the most challenging scenarios is Europe, where health systems and health policies are becoming more and more interconnected and citizen-centered in order to improve efficiency, create added-value services and move towards an European e-Health Area. In the last decade the European Commission aimed and aims to integrate all EU health-related policies[2] and concentrate resources in order to improve patient care and make more efficient and responsive Health Systems, without duplicating the work of the Member States or international organizations. The EU Health Strategy [10] set out in May 2000 defined these goals and thereby impacted in further EU's IS policies and activities related to Health Care.

In 2002, e-Health became a eEurope 2005 policy priority [11], setting targets for both the European Commission and Member States to meet in areas such as:

- creating a European electronic health card, featuring added functionalities, such as medical emergency data and secure access to personal health information;
- developing Health Information Networks to speed the flow of health information through the healthcare system;
- putting health services online such as information on healthy living and illness prevention, electronic health records, teleconsultation and e-reimbursement.

The European Commission adopted on April 2004 an Action Plan [12] which addresses the crucial role of new technologies and new ways of delivering health care in improving access to, quality and effectiveness of care. The Action Plan has two main aims: a) to make the most of new information and communication technologies in the health sector and b) to better integrate a range of e-Health policies and activities.

A key issue in the European e-Health Area is *patient mobility*. EU law gives patients the right to go for treatment in other Member States. But exercising that right can prove difficult. On the patients' side IST might help to empower the patients by giving them better information on how to obtain treatment accross borders, the quality, availability and appropriateness of the treatment and how to het their national Health Care system or Health insurer to pay for the treatment. On the Health Care providers side, IST should give them promptly access to the patients' records, enable effective planning, coordination and interconnection and help control costs. A report by the Health Council on December 2003 makes 19 specific recommendations to promote patient mobility (see [35]).

1.2 Issues when Applying Agents for Health-Care

In this distributed, multi-lingual scenario, agent technologies are considered a feasible technology to improve Health Care management and decision-making. objectives the Health Care sector is interested in. As explained in [39], Agent Technology allows:

- To proactively anticipate the information needs of a patient, and deliver it in a periodical basis
- To support communication and coordination, either synchronous or asynchronous, among all members of a medical team, allowing the share of distributed information and knowledge sources, and providing distributed decision making support.
- To adapt medical services to patients' needs (*personalization*).
- Increase the patient's control over all the data collected in his/her medical records.

[1] Institute of Information and Computing Sciences, Utrecht University, The Netherlands. email: javier@cs.uu.nl

[2] It is important to note here that the integration of EU health-related policies and regulations is only a recommendation that cannot be fully imposed to the Member States, as in Article 152, paragraph 5 of the EU Treaty, it is stated that Community Action in the field of public health must fully respect the competence of the Member States in the field of health care. Therefore, it is unlikely that there will be a single, uniform regulation for the whole EU on Health Care services, but National and Local polices and regulations may co-exist with European ones.

Problem	Standard solutions	Agent-Mediated solutions
Data exchange problem:	standard data interchange formats	Agent Communication Languages, Agent Ontologies
Communication problem:	international notations or translation mechanisms	
Coordination issues:	policies,planners, shared dietaries.	Agent-Mediated Coordination
Variety of regulations:	?	Agent-Mediated Electronic Institutions
Trust:	?	

Figure 1. Problems in a distributed organ and tissue allocation system

However, issues of privacy, security and trust are particularly sensitive in relation to matters such as agents having access to patient records, what is an acceptable behaviour for an agent in a particular role and the development of trust both among (heterogeneous) agents and among agents and users.

An additional issue is that health care systems are highly regulated through different forms of regional, national and European regulations and policies. This means that not only such normative constraints should be taken into account while designing agent-mediated health care systems but that one or several normative contexts (European, national, regional) may apply and therefore taken into account.

We can summarize the issues to solve in a geographically distributed Health Care environment as follows:

- i1 *the data exchange problem:* exchange of information is a major issue, as each of the actors collects different information and stores it in different formats. The obvious, and easily stated, solution is the definition of standard data interchange formats.
- i2 *the communication problem:* countries typically use different languages and terminologies to tag the same items or facts. Either a standard notation or a translation mechanism needs to be created to avoid misunderstandings.
- i3 *the coordination issues:* in order to operate at an international level, there is the need to coordinate geographically distributed teams, and to coordinate the secure delivery of information or physical items (laboratory tests, organs, tissues) at an international level.
- i4 *the variety of regulations:* an additional issue is the necessity to accommodate a complex set of, in some cases conflicting, national and international regulations, legislation and protocols governing the exchange of organs. These regulations also change over time, making it essential that the software is adaptable.
- i5 *trust:* in a geographically distributed environment, Health Care systems should interact with other systems, in a bigger, open, distributed system. While in closed systems all the components of the system can implicitly trust the information and services provided by others, in an open environment *trust* becomes an additional issue. Systems may purposely send incorrect or partial data, or provide unacceptable quality of services. There is the need to create overarching environments to ensure fair interactions and proper behaviour and, therefore, create trust.

Figure 1 relates these issues with possible solutions. It can be seen that issues i1 (data exchange) and i2 (communication) can largely be resolved using standard software solutions. Issue i3 (coordination) may be solved with conventional software. However, Medicine is one of the most difficult fields for coordination, as it is extremely difficult to foresee all the conditions that may occur, leading to unexpected side effects when certain decisions or actions are performed in a unanticipated situation. As Fox & Das argue in [23], these are the kind of ill-defined fields which have historically been the concern of Artificial Intelligence. They also identify software agents as having many strengths (mainly *pro-activity* and *autonomy*) which make Agent Technologies well-suited for medical applications. Last two issues (the variety of regulations changing over time and trust) can hardly be solved with conventional software and therefore underpin our case for the use of so-called *e-institutions*:

- E-Institutions and the norms that govern them are the key to a system that is able to adapt automatically to changes in regulations. The purpose of an *e-institution* is to provide an over-arching framework for agent interaction, where agents may reason about the norms, in the same way as physical institutions do in the real world through social norms. These norms define the *acceptable* actions that each agent may perform depending on the role or roles it is playing, and clearly specifies the data it may access and/or modify in playing those roles.
- Trust in a distributed scenario becomes even more important when agent tecnology is at work. In *open multi-agent systems*, where there are several, heterogeneous agents created by different people, agents may not only have completely different goals but there is also a higher probability that they have conflicting goals, (that is, the individual goals of an agent interfering or competing with other agents' goals.) This raises a problem of *trust* in an agent society.[3] In close-multi-agent systems, where cooperation among agents is included as part of the designing process, there is an implicit trust: an agent a_i requesting information or a certain service from agent a_j can be sure that such agent will answer him if a_j has the capabilities and the resources needed, otherwise a_j will inform a_i that it cannot perform the action requested. However, in an open environment *trust* is not easy to achieve[4], as agents may give incomplete or false information to other agents or betray them if such actions allow them to fulfill their individual goals. E-Institutions should provide a safe environment with mechanisms to defend and recommend right and wrong behaviour (defined by the norms), thereby inspiring trust into the agents that will join such an environment.

1.3 The Organ and Tissue allocation

Through this paper we will use the allocation of Organ and Tissues for transplantation purposes as example. Organ transplantation from human donors is the only option available when there is a major damage or malfunction in an organ. It is also very important in economic terms. For instance let us take the case of Spain, where transplantation of one kidney compared with dialysis would save between 186400 and 240530 Euros.

Over the years, transplant techniques have evolved, knowledge of donor-recipient compatibility has improved and so have immunosuppressant drug regimes, leading to an increase in the number of organs

[3] In [24] Gambetta defines *trust* as a particular level of subjective probability with which an agent a_j will perform a particular action before the action is performed.

[4] There are lots of work to try to define reputation and trust mechanisms in open systems, such as [1, 47, 48, 37].

that can be transplanted, but also in the range of transplants, moving beyond organs (heart, liver, lungs, kidney, pancreas) to tissues (bones, skin, corneas, tendons). However, the allocation process for tissues is quite different from that for organs, because of the time such pieces[5] can be preserved outside the human body. Tissues are clusters of quite homogeneous cells, so the optimal temperature for preservation of all the cells composing the tissue is almost the same. Thus, tissues can be preserved for several days (from six days in the case of corneas to years in the case of bones) in tissue banks. For tissues, the allocation process is triggered when there is a recipient with a need for a certain tissue, at which time some number of tissue banks are searched for a suitable one.

Organs, on the other hand, are very complex structures with several kinds of cell types with different optimal preservation temperatures. That fact leads to quite short preservation times (hours), no need for an organ bank, and an allocation process that is triggered when a donor appears, taking the form of a search for a suitable recipient in some number of hospitals.

Since 1980 the number of requests for the application of transplant techniques has risen so much[6] that the human coordinators—the people at the hospitals who act as the interfaces between the surgeons internally and the organ and tissue banks externally—are facing significant problems in dealing with the volume of work involved in the management of requests and piece assignment and distribution. Transplant-based therapies are the subject of much investigation and increasing application, such that demand for pieces may well rise rapidly in the near future.

1.4 Organ and tissue transplantation as a trans-national problem

Organ and tissue transplantation are widely-used therapies against life-threatening diseases. But there are two issues that make transplantation management a very complex issue: (i) *scarcity* of donors, so it is important to try to maximize the number of successful transplants (ii) *donor/recipient matching*, because of the diversity and multiplicity of genetic factors involved in the response to the transplant.

The relative scarcity of donors has led to the creation of coalitions of transplant organizations. In the case of the United States of America, a new organization called UNOS (*United Network for Organ Sharing* [41]) has appeared in order to join and coordinate the several pre-existing transplant organizations that existed in some states. Also international coalitions of transplant organizations have been created, such as EUROTRANSPLANT [21] (Austria, Belgium, Germany, Luxembourg, the Netherlands and Slovenia) or Scandiatransplant [38] (Denmark, Finland, Iceland, Norway and Sweden). Indeed there is an initiative called *The Donor Action Foundation*[7] [18] which plans to create a world-wide coalition.

Most of the work in the field of transplant allocation (such as EU projects RETRANSPLANT [36][8], TECN [40]) is devoted to the creation of a) standard formats to store and exchange information about pieces, donors and recipients among organizations, b) telematic networks, or c) distributed databases. Project ESCULAPE [19] uses conventional software to help in tissue histocompatibility by developing HLA[9] referencing computer systems and software packages to be used by hospitals and laboratories as a human tissue typing tool. However none of them give support to the coordination and regulatory problems, which should be solved entirely by humans.

The consideration of these factors leads to the question of whether some kind of automation of the assignation process is desirable and if so, whether it is possible. We proposed in [45] an agent-based architecture for the tasks involved in managing the vast amount of data to be processed in carrying out

- recipient selection (e.g., from patient waiting lists and patient records),
- organ/tissue allocation (based on organ and tissue records),
- ensuring adherence to legislation,
- following approved protocols and
- preparing delivery plans (e.g., using train and airline schedules).

There are very few references in the literature about the use of agents in the transplant domain. In [42] Valls *et al.* describe an agent that uses multi-criteria decision techniques in the selection of the best receiver in a transplant, providing the Hospital Transplant Coordinator with a result according to the weights the user assigned to each criteria. Moreno *et al.* present in [31] a hierarchical multi-agent system where the agent on the root node plans transport routes between hospitals using the information obtained from the other agents in the hierarchy, removing routes that will exceed the maximum available time for transportation and avoiding potential fatal delays due to mistakes in coordination of different means of transport. In [30] Moreno *et al.* propose a multi-agent system architecture to coordinate hospital teams for organ transplants. Coordination is achieved through agents that keep track of the personnel schedules and the availability of the facilities (both described as time-tables divided into slots of thirty minutes). Calisti *et al.* present in [7] the *Organ Transplant Management* (OTM) system, an agent-based platform to support medical practitioners in the tasks of data management and decision making. Finally, Aldea *et al.* present in [2] an alternative design for a multi-agent architecture for the Spanish organ allocation process. It identifies the agents needed to solve the problem and organizes them in four levels (*Hospital Level*, *Regional Level*, *Zonal Level* and *National Level*). However, the normative dimension of the problem is explored in none of these works.

2 APPLYING NORMS IN AGENT-MEDIATED HEALTH CARE SYSTEMS

In this section the aim is to discuss the application of norms in agent societies by addressing the following three issues:

- How norms should be expressed (see §2.1).
- How cognitive agents are affected by norms (norms from the agent perspective, see §2.2).
- How to define safe environments for agents to ensure trust and acceptable behaviour (norms from the institutional perspective, see §2.3).

[5] We will use the term *piece* to refer to both *tissues* and *organs*.

[6] The continuous raise in requests is due, among other factors, to the introduction of new immunosuppressors which have significantly decreased rejection in recipients' clinical evolution.

[7] The Donor Action Foundation is an initiative of the Eurotransplant International Foundation, the Spanish National Transplant organization (ONT) and the former Partnership for Organ Donation (U.S.A.).

[8] Unfortunately, there is no information available about the practical results of those projects other than the project URL.

[9] The HLA system is used as one of the tests in matching donor and receptor tissues or organs in the allocation process. HLA stands for *Human Leukocyte Antigen* system, a group of the most important antigens responsible for tissue compatibility, together with the four significant genetic markers (on chromosome 6) that encode them (HLA-A, HLA-B, HLA-C and HLA-D).

In our approach we assume that norms can sometimes be violated by agents in order to keep their autonomy, which can also be functional for the system as a whole as argued in [8].

Then in §3 we will discuss norm implementation on both the agent and the institutional perspective.

2.1 A Language for Norms

In order to express complex norms, a language is needed that can express deontic concepts (OBLIGED, PERMITTED, FORBIDDEN) which can be conditional (IF) and can include temporal operators (BEFORE, AFTER).

In [44] a classification of the different kinds of norms is presented. As a summary, norms can be characterized by wherther a) they refer to a state or an action, b) they are conditional, c) they include a deadline, or d) they are norms concerning other norms:

2.1.1 Norms concerning that agent a sees to it that some condition/predicate P holds.

In this case the norm is timeless, that is, the norm on the value of P is active at all times. There are three possible expressions:

$$\mathsf{OBLIGED}(a,P) \qquad \mathsf{PERMITTED}(a,P) \qquad \mathsf{FORBIDDEN}(a,P)$$

An example of such a timeless norm is the following:

$$\mathsf{FORBIDDEN}(recipient, (in\_waiting\_list(hospital_1) \wedge in\_waiting\_list(hospital_2) \wedge (hospital_1 \neq hospital_2)))$$

2.1.2 Norms concerning agent a performing an action A.

In this case the norm on the execution of A is also timeless, that is, the norm is active at all times.

$$\mathsf{PERMITTED}(a\ \mathsf{DO}\ A) \qquad \mathsf{FORBIDDEN}(a\ \mathsf{DO}\ A)$$

There are no unconditional obligations (OBLIGED), since this would express an obligation to execute an action all the time. An example of an unconditional norm would be the following:

$$\mathsf{FORBIDDEN}(person\ \mathsf{DO}\ sell(organ))$$

Note that action A can be an abstract action, that is, an action that is not present in the repertoire of the agents or defined in the protocol. In such cases A should be translated in more concrete actions to be checked.

2.1.3 Norms concerning a condition P or an action A under some circumstance C.

The norm is conditional under C. A condition C may be a) a predicate about the state of the system, or b) a state of some action (starting, running, done).

$$\begin{array}{ll}\mathsf{OBLIGED}((a,P)\ \mathsf{IF}\ C) & \mathsf{OBLIGED}((a\ \mathsf{DO}\ A)\ \mathsf{IF}\ C)\\ \mathsf{PERMITTED}((a,P)\ \mathsf{IF}\ C) & \mathsf{PERMITTED}((a\ \mathsf{DO}\ A)\ \mathsf{IF}\ C)\\ \mathsf{FORBIDDEN}((a,P)\ \mathsf{IF}\ C) & \mathsf{FORBIDDEN}((a\ \mathsf{DO}\ A)\ \mathsf{IF}\ C)\end{array}$$

An example is the following:

$$\mathsf{FORBIDDEN}((allocator\ \mathsf{DO}\ assign(organ, recipient))\ \mathsf{IF}\ \mathsf{NOT}(hospital\ \mathsf{DONE}\ ensure\_quality(organ)))$$

2.1.4 Conditional norms with deadlines.

This is a special type of conditional norm where the start of the norm is not defined by a condition but by a deadline. Deadlines can be either *absolute* (E.g. 23:59:00 09/05/2004) or *relative* to another event (E.g. $time(done(assign(organ, recipient))) + 5min$).

There are 12 possible expressions with deadlines, by combining the three deontic operators, the temporal operators (BEFORE and AFTER) and applying them to actions or predicates. Examples of such expressions are:

$$\begin{array}{l}\mathsf{OBLIGED}((a,P)\ \mathsf{BEFORE}\ D)\\ \mathsf{PERMITTED}((a\ \mathsf{DO}\ A)\ \mathsf{AFTER}\ D)\\ \mathsf{FORBIDDEN}((a,P)\ \mathsf{BEFORE}\ D)\end{array}$$

An example of a conditional norm with deadline is the following:

$$\mathsf{OBLIGED}((allocator\ \mathsf{DO}\ assign(heart, recipient))\ \mathsf{BEFORE}\ (time(done(extraction(heart, donor))) + 6hours))$$

2.1.5 Obligations of enforcement of norms.

In this case the norms concerning agent b generate obligations on agent a.

$$\begin{array}{l}\mathsf{OBLIGED}(a\ \mathsf{ENFORCE}(\mathsf{OBLIGED}(b...)))\\ \mathsf{OBLIGED}(a\ \mathsf{ENFORCE}(\mathsf{PERMITTED}(b...)))\\ \mathsf{OBLIGED}(a\ \mathsf{ENFORCE}(\mathsf{FORBIDDEN}(b...)))\end{array}$$

An example is the obligation of the ONT[10] to avoid money exchanges for organs:

$$\mathsf{OBLIGED}(ONT\ \mathsf{ENFORCE}(\mathsf{FORBIDDEN}(person\ \mathsf{DO}\ sell(organ))))$$

2.2 Normative Agents

In order to understand how the norms modify the reasoning process of an agent, let us first analyse the relation among the norms and the agent's knowledge about the world. Work on this perspective can be found in [5] [6] [9]. In this section, however, we will analyse such relations in a more intuitive way, by means of Venn diagrams representing the epistemic dimension.

As we saw in §2.1, we have chosen a variation of modal logic (Deontic logic) in order to express norms because of the expressivity such approach gives to the normative framework. So, with such approach, it is natural to express an agent's mental states in terms of *possible worlds* that are the basis of *Kripke semantics*. This will also allow us to describe the agent's mental states by means of BDI logic and then, in §3.1, see how norms impact the BDI internal cycle of the agent.

Our model is composed by the *epistemic dimension*, where allowed behaviour is expressed in terms of possible worlds, and the *normative dimension*, collecting the set of norms. Figure 2a) depicts the elements of our model and the connection between both dimensions:

- W is the set of possible worlds w that define the environment of the agents (the *epistemic dimension*).
- N_W is the set of norms in the normative dimension that apply to the world W (the *normative dimension*).

[10] The Spanish National Transplant Organization

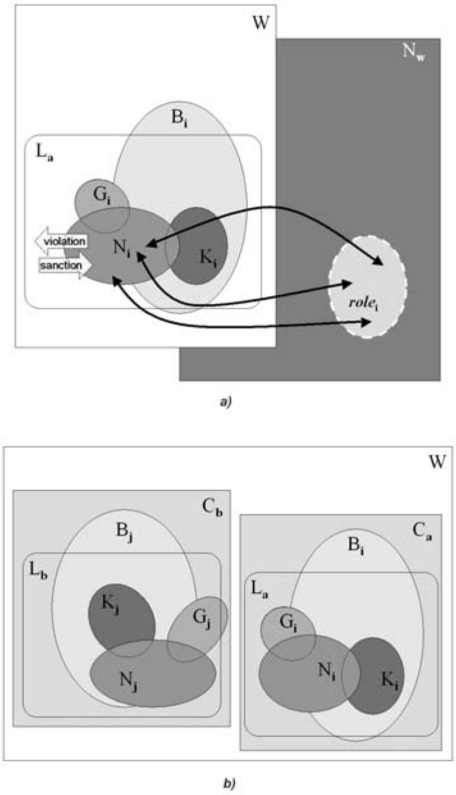

Figure 2. a) The influence of norms in cognitive agents from the Epistemic Perspective. b) Two agents in the same world but in two different contexts.

- The effect of the set N_W in W is expressed by the set L_W. The set $L_W \subseteq W$ is the set of legally accessible worlds for all agents that are situated in W, as defined by the norms in N_W. [11]
- $B_i \subseteq W$ corresponds to the set of worlds that the agent a_i believes with a given certainty level. [12]
- $K_i \subseteq B_i$ is the subset of worlds w that the agent a_i knows with a high certainty level. [13]
- $role_i$ is the subset of norms in N_w that apply to a given role r_i.
- The set $N_i \subseteq L_w$ is the set of possible worlds the agent a_i can legally reach when enacting role r_i, according to the norms defined in the set $role_i \in N_W$.

2.2.1 Norms for Safety and Soundness

The concept of *legally accessible worlds* allows us to describe wanted (legal) and unwanted (illegal) behaviour by defining which states are acceptable (safe) or unacceptable (unsafe).

In this model, the Safety and Soundness of the system is based in two concepts: violation and sanction. A *violation* is a situation where a given agent breaks one or more rules, entering in an illegal (unsafe) state. In order to define the consequence of a violation, sanctions are also defined. A *sanction* can be defined as an action or set of actions whose realization will remove the violation. So sanctions are ways to make agents become *legal* again (enter in a safe state).

From the agent's epistemic perspective, the goal set G_i is limited by the set N_i of *legal accessible worlds*, as the agent will enter in violations when trying to achieve the worlds in G_i that are not in N_i. In the case that the agent enters (willingly or unwillingly) in a violation, the sanction should not only punish the agent but also include those actions needed by the whole system to recover from such violation. Sanctions are therefore the mechanism to ensure soundness of the system in case of violation.

The environment where the agent interaction occurs should therefore detect the violations and apply the proper sanctions. We will discuss this issue in §3.2.1.

2.2.2 Context, Contextual Norms and Multi-Context Ontologies

In most of real domains norms are not universally valid but bounded to a given context. This is the case of norms in Health Care, as they are bounded to trans-national, national and regional regulations, each of them defining a different context.

Any Agent-Mediated Health Care System is, in some degree, a Knowledge-Based system which depends on a model of the medical domain to properly operate. Guha notes in [26] that there are several Knowledge-Based systems which domain models suffer of what he calls *Homogeneity of the KB*:

- the same vocabulary is used uniformly throughout the KB,
- KB contains a single theory of the world that must be kept consistent,
- the single theory approach should be kept generic and independent of particular problems, i.e., the representation should not be tailored specifically towards solving certain problems.

The problem that arises with the *single vocabulary/single theory* approach is that it tries to find a universal vocabulary, theory and representation to model and reason about any situation. The theoretical alternative, first proposed by McCarthy in [29], is to include context as formal objects. Therefore, most theoretical approaches have moved towards having an explicit representation of context. One of the most used approaches is the *box metaphor*, that is, considering context as a box:

> *[...] Each box has its own laws and draws a sort of boundary between what is in and what is out.*[25]

With this idea, we may define context in our model as follows:

- A context C_a is a subset of worlds $w \in W$ where there is a shared vocabulary and a normative framework to be followed by a certain group of agents. A context may contain other contexts inside it.

Applied to the Health Care domain, each context should define a) its vocabulary (by means of an ontology) and b) the (regional/national/translantional) norms that apply in that context.

Figure 2b) shows an example of contexts. For simplicity it only shows two contexts that are totally independent. Agent a_i is situated in a context C_a. The context C_a defines a normative framework, that in terms of possible worlds, it is the set of *physical accessible worlds* for all the agents inside C_a, agent a_i included. The sets B_i and K_i represent the knowledge agent a_i has about the context it is situated in. Based in that knowledge, it has also its set G_i of goals. The same applies for agent a_j, which is situated in the context C_b.

[11] This is based in the concept of *socially acceptable worlds* presented in the Artificial Social Systems' model by Moses and Tennenholtz. In [32] they defined the set W_{soc} as the set of worlds that all agents can legally reach.
[12] The set B_i corresponds in Epistemic Logic to the BEL operator.
[13] The set K_i corresponds in Epistemic Logic to the K operator.

Although it is not depicted in figure 2b), the definition of context also refers to a vocabulary to be shared by agents in a given context. It means that each context is associated with a domain ontology that defines the meaning of the terms that are both present in the norms, the actions the agent may carry and the terms in the communication with others. However, standard ontologies are not enough. As the definition points out, each context (defining their norms and an ontology) may contain other contexts inside (extending and/or modifying the norms and the ontology). This appears in Medicine also, specially in a multi-lingual environment such as Europe. Different contexts may not only use different terms to refer to the same concept (because of linguistic differences, such as *organ/órgano*, *tissue/tejido* or *stem cells/celulas madre*) but also define a given concept differently (e.g., in medical environments *blood* is considered a *tissue*, while Spanish legislation puts out *blood* from the coverage of tissue-related laws and has its own regulations). Therefore, instead of trying to force a single ontological representation for all contexts, it is needed a multi-level, multi-context framework for ontologies that allows the definition and refinement of the concepts in a context connecting them with related concepts in their super-contexts (for instance, connecting spanish definition of blood with the definition given in EU Law). We will return to this issue in §4.

2.3 Electronic Institutions

Given that agents might deviate from expected behavior, open multi-agent systems need mechanisms to systematize, defend and recommend right and wrong behaviour, along with safe environments to support those mechanisms, thereby inspiring trust into the agents that will join such an environment. Norms are commonly used means to describe such expected behavior. Some foundational work in this direction has been done in the ALFEBIITE project [3], in particular in [4].

An Electronic Institution [16] [20] is a safe environment mediating in the interaction of agents. The expected behaviour of agents in such an environment is described by means of an explicit specification[14] of norms, which is a) expressive enough, b) readable by agents, and c) easy to maintain. Such normative model should be supported by organizational models, defining roles, their relationships, their responsibilities and interaction dinamics, and by powerful ontologies to represent concepts and relationships in the domain that are rich enough to cover the necessary contexts of agent interaction while keeping in mind the relevance of those concepts for the global aims of the system.

In [17] we have presented OMNI, which is a suitable framework for complex Electronic Institutions. OMNI is an integrated framework for modelling a whole range of MAS, from closed systems with fixed participants and interaction protocols, to open, flexible systems that allow and adapt to the participation of heterogeneous agents with different agendas. OMNI is composed by three dimensions: **Normative**, **Organizational** and **Ontological** that describe different characterizations of the environment. The model is based on two recent MAS models, OperA [15], and HARMON*IA* [43]. Figure 3 depicts the different modules that compose the proposed framework. The division of the modules into three levels of abstraction aims to ease the transition from the very abstract statutes, norms and regulations to the very concrete protocols and procedures implemented in the system. Different domains have different requirements concerning normative, organizational and communicative characteristics, which means that not always all three modules have the same impact or are even needed: in those domains with none or small normative components, design is mainly guided by the Organizational Dimension, while in highly regulated domains the Normative Dimension is the most prominent.

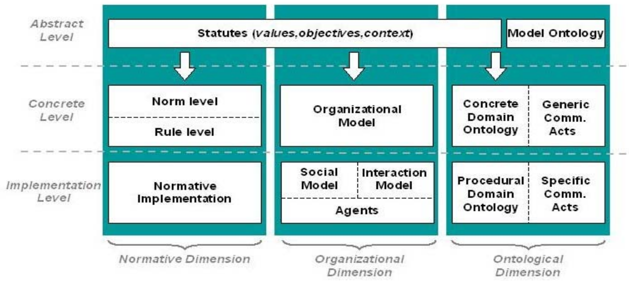

Figure 3. *The* OMNI *framework.*

The **Abstract Level** defines the statutes of the organization to be modelled. *Statutes* indicate, at the most abstract level, the main *objective* of the organization, the *values* that direct the fulfilling of this objective and they also point to the *context* where the organization will have to perform its activities. [15] This level also contains the definition of terms that are generic for any organization in the domain (that is, that are incontextual) ant the ontology of the OMNI model itself. [16]

The **Concrete Level** specifies the analysis and design process, starting from the abstract values defined in the previous level, refining their meaning in terms of norms and rules, roles, landmarks and concrete ontological concepts. In order to check norms and act on possible violations of the norms by the agents within an organization, on the normative dimension, abstract norms have to be translated into actions and concepts that can be handled within such organization. The organizational dimension specifies the means to achieve the objectives identified in the the abstract level as an Organizational Model. The content aspects of communication, or domain knowledge, are specified by Domain Ontologies and Generic Communication Acts define the interactions languages used in the Organizational Model.

The **Implementation Level** describes the implementation of the design in a given multi-agent architecture, including the mechanisms for role enactment and for norm enforcement. The normative dimension provides both the low-level protocols and the related rules that enable agents to comply with organizational norms. OMNI assumes that individual agents are designed independently from the society to model goals and capabilities of a given entity. Agent populations of the organizational model are described in the Social Model in terms of commitments regulating the enactment of roles by individual agents. Depending of the specific agents that will join the organization, several populations are possible for each organizational model. The Interaction Model describes the specific interactions such

[14] The main reason for having explicit representations of norms is that norms may change over time. If norms are embedded in the agents' design and code, all the design steps have to be checked again and all the code verified to ensure compliance with new regulations.

[15] The definition of the statutes is similar to a first step in the requirements analysis.

[16] The *Model Ontology* can be seen as a meta-ontology that defines all the concepts of the framework itself, such as norms, rules, roles, group, violations, sanctions and landmarks.

as agreed upon by the agents. Specific Communication Acts covers the communication languages actually used by the agents as they agree in the interaction contracts. As with the content ontologies, communicative acts defined at a lower level of abstraction implement those defined at a higher level.

All dimensions of OMNI have a formal logical semantics, which ensures consistency and feasibility of system verification. A full description of the OMNI framework is available at [46].

3 IMPLEMENTING NORM ENFORCEMENT

Implementing norms is not implementing a theorem prover that, using the norms semantics, checks whether a given interaction protocol complies with the norms. The implementation of norms should consider a) how the agents' behaviour is affected by norms, and b) how the institution should ensure the compliance with norms. The former is related to the *implementation of norms from the agent perspective*, by analyzing the impact of norms in the agents' reasoning cycle (see §3.1). The latter is related with the *implementation of norms from the institutional perspective*, by implementing a safe environment (including the enforcing mechanisms) to ensure trust among parties. (see §3.2).

As far as we know, the most complete model in literature considering some operational aspects of norms for MAS is the extension of the SMART agent specification framework by López y López, Luck and d'Inverno [27] [28]. The framework aims to represent different kinds of agent societies based on norms. However, no implementation of the architecture applying it to a real problem has been reported in literature, there are no tools to support the development and implementation of a normative multiagent system, and there are no mechanisms defined from the institutional perspective in order to enforce the norms.

In order to implement enforcement mechanisms that are well-found, one has to define some kind of operational semantics first. In general, an operational semantics for norms always comes down to either one of the following:

- Defining constraints on unwanted behaviour.
- Detecting violations and reacting to these violations.

The choice between these two approaches is highly dependent on the amount of control over the addressee of the norms. Prevention of unwanted behaviour can only be achieved if there is full control over the addressee; otherwise, one should define and handle violations (see §2.2.1 and §3.2.2).

Therefore there are two main assumptions in our approach. First of all, as mentioned in §2, we assume that norms can sometimes be violated by agents in order to keep their autonomy. The violation of norms is handled from the organizational point of view by violation and sanction mechanisms. Secondly we assume that from the institutional perspective the internal state of the external agents is neither observable nor controlable (external agents as black boxes). Therefore, we cannot avoid a forbidden action to be in the goals and intentions of an agent, or impose an obligatory action on an agent to be in their intentions.

3.1 Impact on the agent: Normative Agents

In §2.2 we discussed how norms have an effect in the agent from the epistemic point of view. In order to implement *normative agents*, that is, agents whose behaviour is guided by the norms, we should analyse the norms from the operational point of view, seeing the effects of norms in the reasoning cycle of the agent. To do so we will use as basis the abstract interpreter for a BDI agent that Rao and Georgeff presented in [33]. The interpreter describes the control of a BDI agent by means of a processing cycle, from perception to action. There are several versions of the cycle, but it can basically be expressed as follows:

```
B := B_init;
I := I_init;
while (true)
{
    get_perception(perc);
    B := belief_revision(B,perc);
    D := options(B,I);
    I := filter(B,D,I);
    plan = generate_plan(B,I);
    execute(plan);
}
```

The cycle is an infinite loop where the agent starts perceiving its environment (`get_perception` function) and with such perceptions modifies its beliefs `B` about the state of the environment (`belief_revision` function). Then the `options` function generates a set `D` of possible alternatives (based on its beliefs and its current intentions). According to the resulting set of options `D` of the previous function the agent should choose one of them (`filter` function). The result is the current intention `I` the agent is committed to achieve. According to that intention, the agent creates the plan to achieve `I` by means of some kind of *means-end reasoning*. Finally, the plan `plan` is executed (`execute` function).

This version of the interpreter is a useful abstraction of the theoretical model of Rao and Georgeff. However this version is too simple and cannot be applied in real-time systems, as the agent does not perceive the world until it has executed the whole plan. With this first version it may happen that the state of the environment changes in a way that the execution of the plan will not have the result the agent expected, or even it may be impractical. In [34] Rao and Georgeff proposed some changes in the architecture, in representation and added some restrictions (such as having a set of pre-defined plans with invocation conditions instead of a plan generator).

Another extension to be done is to include the norms in the cycle. As we mentioned in §2.2, one of the effects of the norms is that they reduce the set of possible actions the agent can choose. So, in order to link this interpreter with the model we presented in §2.2, we have to identify at which point of the reasoning cycle we must include the norms that apply to the agent (defined in the set $role_i$).

In §2.2 we saw how in our model the goal set G_i is limited by the set N_i of *legal accessible worlds*, as the agent will enter in violations when trying to achieve the worlds in G_i that are not in N_i. Therefore, the restrictions of the normative framework have an effect on the desires and intentions of the agent. In the case of the interpreter's cycle, that means that the rules should be introduced in the intention formation steps, as the agent must, in most cases, check that the intention that it commits to achieve is legal according to the normative framework. There are two ways to do so:

- by modifying the `options` function to allow that only legal options are suggested. With this modification not only *feasible* options (from the point of view of the agent's capabilities) are created but only those that are *legal*.
- to modify the `filter` function in order to ensure that the final intention the agent chooses is a legal one.

In both cases we add an *allowance* criteria to the *feasibility* criteria already present into the interpreter.

However, as F. Dignum *et al* explain in [14], reasoning only in terms of the *allowance* criteria is not enough. Normative frameworks use to have not only restrictive norms that forbid some actions to be taken (such as *"don't kill a human being"*), but also those that impose states or actions to be taken into account into the agent's reasoning cycle (such as *"you are obliged to pay the goods you bid for"* in an Electronic Auction House).

The BDI interpreter presented in [14] adds the norms that impose actions as *deontic events* to be handled by the agent. An example is an agent A having a conditional obligation to B, $O_{AB}(\phi \mid \psi)$: when the precondition ψ becomes true, then a deontic event $O_{AB}(\phi)$ is created.[17] Such deontic events are compared then by the agent with its goals ans desires (seen also as events) and then decides which event will try to handle.

This idea can be applied to the extended BDI interpreter presented in this section by adapting the `option` and `filter` functions:

```
B := B_init;
I := I_init;
while (true)
{
    get_perception(perc);
    B := belief_revision(B,perc);
    D := options(B,I,oblEvents);
    I := filter(B,D,I,oblRestr);
    plan = find_plan(B,I);
    while not( empty(plan) OR succeeded(I,B)
               OR impossible(I,B)
    {
       action = next_action(plan);
       execute(action);
       get_perception(perc);
       B :=belief_revision(B,perc);
       if reconsider(I,B,oblEvents) then
       {
          D := options(B,I,oblEvents);
          I := filter(B,D,I,oblRestr);
       }
       if not(sound(plan,I,B)) then
       {
          plan = find_plan(B,I);
       }
    }
}
```

This version solves most of the real-time problems mentioned above. For instance, in order to avoid *overcommitment* (trying to achieve an intention that is no longer possible or trying to execute a plan that is not valid):

- plans are not executed from beginning to end but one action at a time, checking if the plan is still sound (`sound` function) in the current state of the world.
- intentions are checked and reconsidered at some point in time. As intention reconsideration is a quite complex process where the consistency of current intention is checked, this process is not done for each action: it is the `reconsider` function (usually a random function) the one that tells the agent when to check if the current intention is still valid.

The extended BDI-cycle also includes the effect of norm on the agent:

- the `options(B,I,oblEvents)` function now not only considers beliefs (`B`) and intentions (`I`) but also the obligation events (`oblEvents`) to be handled in the decision making,
- the `filter(B,D,I,oblRestr)` function includes now the restrictive norms (`oblRestr`) as input in order to ensure that the final intention the agent chooses is a legal one.
- the `reconsider(I,B,oblEvents)` function also should include the obligation events (`oblEvents`). With this change the agent, while executing a plan to fulfill a given intention `I`, might reconsider its intention because the appearance of a new obligation event that is more important than the current intention.

[17] In the case of the norm expressions presented in §2.1, *deontic events* appear in the Conditional Norms and in the Obligations of Enforcement of Norms.

3.2 Impact on the Multiagent System: Electronic Institutions

As we saw in §2.2, each agent has its set N_i of norms which defines which worlds are *legal* or *illegal* for the agent, depending on the roles it may enact. We also saw that the set $\{G_i \cap N_i\}$ defines which are the goals that the agent can *legally reach*. However, in our model:

- the internal mental states of the agents or their reasoning process cannot be seen or controled by other agents, only their observable behaviour in terms of (public) messages and (visible) actions can be seen, and
- we allow agents to break the norms in some situations.

So it is not granted that all the agents will follow all the norms anytime. Therefore, some kind of (weak) norm enforcement is needed. In our model norm enforcement is based in the concepts of *violation* and *sanction*, defined in §2.2.1.

As norm enforcement is not granted from the agent perspective, some control mechanisms should be created. The easiest way is to have a centralized controller, a *Coordinator Agent*. In such scenario the goals of all the agents in the system are more or less defined by the tasks the Coordinator Agent assigns to the agents. In this scenario the Coordinator Agent can, directly or indirectly change the other agents' goals in order to ensure that none of them violates the norms. However, the limitations of this approach are too serious to be acceptable (mainly *scalability* and *adaptability* issues.

So it may seem that the enforcement of norms should be completely distributed through all the agents in the system. In this scenario all agents should be aware not only of their rights and obligations, but also about the others' rights and obligations, so when there is an agent that is breaking a norm, the affected agents may detect it and punish the agent's new goal in some way. To do so, the agents should have not only a complete knowledge of the norms that apply to them but also a part of the norms that apply to other agents, and continually reason about the legacy of their own behaviour and the neighbours' one. In our model, this option means that an agent a_i must not only have knowledge about the set of rules associated to $role_n$ (that he must follow according to its role) and the interpretation N_i of such rules, but it has to know about all the sets $role_k$ defined in R_a for the other agents and their corresponding interpretations N_l in C_a. As a result, the agents may expend too many resources (i) checking other agents' behaviour, and (ii) reasoning about the roles such agents should play. Also, it is unclear which agents each agent should check, so it may happen that an agent is checked by more than one agent at the same time while there may be an agent whose behaviour is checked by no one.

F. Dignum proposes in [13] an optimization of this scenario. In his proposal the agents do not have to be aware of all the norms but only:

- the social norms that affect the agent
- the contracts the agent committed with other agent, and the countermeasures that can be taken if such contracts are violated (for instance, to ask for a reward/compensation).

The knowledge about the countermeasures to be taken in case of violation is expressed in terms of authorizations. Dignum's proposal is sound, but it makes the asumption that all agents in the system are able to reason about obligations and authorizations, and that all of them implement the extension of the BDI architecture he proposes.

A solution must be found that makes no assumption of the internal structure of the agent. A good option is the idea of *Guardian Agents* presented by Fox and Das in [22]. In our model this idea is included by defining an institutional role to enforce norms. In this scenario there is a prominent role, the *Police Agent*, that can be enacted by one or more agents. Such agents cannot access the internal code of the agents, but only perceive their actions (they see the other agents as black boxes performing actions in an environment). The Police Agent checks if the behaviour of those agents follows all the norms, (like a policeman checks the behavior of car drivers).

3.2.1 Platform enforcement mechanisms

In order to support the task of these Police Agents, the platform should provide time-efficient services to help those agents to enforce proper behaviour in large agent societies.

Detection of the occurrence of an action. In the case of agent actions, there are three possible points to be detected: a) when the action is going to be performed, b) it is being performed, or c) it is done. In an agent platform with several agents performing different actions at the same time, a question arises on how to implement the detection of the occurrence of actions. Police Agents may become overloaded on trying to check any action on any time. Therefore in [44] we propose to create two plaftorm mechanisms: a) a **black list mechanism** of actions to be checked, and b) an **action alarm mechanism** that triggers an alarm when a given action A in the black list attempts to start, is running or is done. This trigger mechanism has to do no further checks, only to make the Police Agent aware of the occurrence of the action. The action alarm mechanism can only be done with actions defined in the institutions' ontology, which specifies the way each action is to be monitored. For instance, when the performance of the action $assign(organ, recipient)$ should be checked, the action is registered by a Police Agent on the black list. Then as soon as $assign(organ, recipient)$ occurs, the trigger mechanism sends an alarm to the Police Agent, that will check if the action was legal or illegal given the norms for that context.

When actions are performed by users through a user interface, the action alarm mechanism can be placed in the interface itself. In the case of the following norm:

$$\textsf{PERMITTED}((nurse\ \textsf{DO}\ include(donor\_data, register))\\ \textsf{IF}\ (assigned(nurse, transplant\_unit, hospital)))$$

the inclusion of the medical data of the donor is done by all staff through a special form. Therefore the interface knows when the user is filling in donor data, and at the moment of submission of such data to the system it can detect that the nurse is nor part of the transplant unit and send an alarm to the system.

Detection of activation and deactivation of norms. In the case of conditional norms we have to detect the *activation of the norm* (when condition C is true) and *the deactivation of the norm* (when predicate P or action A is fulfilled or C does not hold). An additional issue is to establish the allowed *reaction time* between the activation and deactivation of an obligation, i.e. the time that is allowed for the completion of the obligation when it becomes active (e.g. immediately, in some minutes). [18] The length of the reaction time for each norm is highly dependent on the application domain. A violation does not occur when the norm becomes active but when the reaction time has passed. The way and the moment to checks the norm conditions is highly dependent on the verifiability levels of each check:

- *conditions computationally verifiable*: the verification of all predicates can be done easily, at any moment.
- *conditions not computationally verifiable directly, but by introducing extra resources*: these conditions need some kind of process to be checked that is time consuming. In this case it is advisable to add to the platform **data structures** (such as fast-access indexes) and **processing mechanisms** that efficiently take out the burden on agents to do the check.
- *conditions non-computationally verifiable*: a condition that can be machine-verified but it is computationally hard to verify. In this case verification of the condition is not done all the time, but is delayed, doing a sort of "garbage collection" that detects violations. There are three main families:
  – Verification done when the system is not busy and has enough resources.
  – Verification scheduled periodically (e.g., each night, once a week.).
  – Random Verification (of actions/agents), like random security checkings of passengers in airports.

Deadlines. Deadlines represent an special case in the implementation of conditional norms, as they are not that easy to check. Deadlines require a continuous check (second by second) to detect if a deadline is due. If the institution has lots of deadlines to track, it will become computationally expensive. We propose to include within the agent platform a **clock trigger mechanism** that sends a signal when a deadline has passed. The idea is to implement the clock mechanism as efficiently as possible (some operating systems include a clock signal mechanism) to avoid the burden on the agents.

3.2.2 Connecting the enforcement mechanisms to the norms

As described in §3, we cannot assume to have full control over the agents entering in the institution. Because there may be illegal actions and states which are outside the control of the Police Agents, violations should be included in the normative framework. In order to manage violations, each violation should include a plan of action to be executed in the presence of the violation. Such a plan not only includes sanctions but also countermeasures to return the system to an acceptable state (repairs).

In §2.1 we have introduced a machine-readable format for expressing norm conditions, and then in §3.2.1 we have discussed how to detect the activation and violation of norms. In order to link these detections with the violation management, in [44] is proposed that a norm description includes, at least, the following:

- The norm condition (expressed as seen in §2.1).
- The violation state condition.
- A link to the violation detection mechanism.
- A sanction: the sanction is a plan (a set of actions) to punish the violator.
- Repairs: a plan (set of actions) to recover the system from the violation.

[18] In theoretical approaches, the semantics are defined in a way that when an obligation becomes active, it has to be fulfilled instantly. But this is impractical for implementation, because agents need some time between detection and reaction. This *reaction time* is ignored in norm theories, but has to be addressed when implementing norms.

In this format, the *norm condition*-field is denoting when the norm becomes active and when it is achieved. The *violation* is a formula derived from the norm to express when a violation occurs (e.g. for the norm $\mathsf{OBLIGED}((a, P)\ \ \mathsf{IF}\ C)$ this is exactly the state when C occurs and P does not, that is, the state where the norm is active, but not acted upon). The *detection mechanism* is a set of actions that can be used to detect the violation (this includes any of the proposed detection mechanisms described in §3.2.1). The set of actions contained in the *sanction*-field is actually a plan which should be executed when a violation occurs (which can contain imposing fines, expulsing agents from the system, etc.). Finally, the *repairs* contains a plan of action that should be followed in order to 'undo' the violation. An example (extracted from organ and tissue allocation regulations) is the following:

Norm condition	$\mathsf{FORBIDDEN}(allocator\ \mathsf{DO}\ assign(organ, recipient))$ $\mathsf{IF}\ \mathsf{NOT}(hospital\ \mathsf{DONE}\ ensure\_quality(organ)))$
Violation condition	$\mathsf{NOT}(done(ensure\_quality(organ))\ \mathsf{AND}$ $done(assign(organ, recipient))$
Detection mechanism	$\{detect\_alarm(assign, 'starting');$ $check(done(ensure\_quality(organ)));\}$
Sanction	$inform(board, "\mathsf{NOT}(done(ensure\_quality(organ))$ $\mathsf{AND}\ done(assign(organ, recipient))")$
Repairs	$\{stop\_assignation(organ);$ $record("\mathsf{NOT}(done(ensure\_quality(organ))\ \mathsf{AND}$ $done(assign(organ, recipient))", incident\_log);$ $detect\_alarm(ensure\_quality, 'done');$ $check(done(ensure\_quality(organ)));$ $resume\_assignation(organ);\}$

This example shows how violations and their related plans of action are defined. The violation condition defines when the violation occurs in terms of concrete predicates and actions (in the example, the violation condition uses exactly the predicates and actions in the norm expression as there is no need to refine them). The detection mechanism is defined as a plan (in this case involving an *action alarm* mechanism detecting each time that an assignment is attempted). Sanction plans define *punishment mechanisms*, either direct (fines, expulsion of the system) or indirect (social trust or reputation). In this scenario the punishment mechanism is indirect, by informing the board members of the transplant organization about the incident. Finally, the repairs is a plan to solve the situation (that is, a contingency plan). In action precedence norms (A precedes B), it usually has the same structure: stop action B (*assign*), record the incident in the systems' incident log and then wait (by means of the action alarm mechanism) for action A (*ensure_quality*) to be performed.

4 CONCLUSION

In this paper we have presented the importance of including normative aspects in Agent-Mediated Medical Systems applied to distributed, trans-national scenarios such as the European e-Health Area. We have presented the concepts of *Normative Agents* and *Electronic Institutions* and discussed the impact of norms both in the agents and in the agent platforms.

However, the use of Normative Agents in the Health Care domain still presents some major challenges. The first one is that such tecnologies should become comprehensible, safe and sound in order to increase the trust of patients and practitioners on such technologies. There are also some technological challenges to be met. Building a complex, multi-agent system where the medical knowledge and the medical normatives are included is a major, complex task, as all three

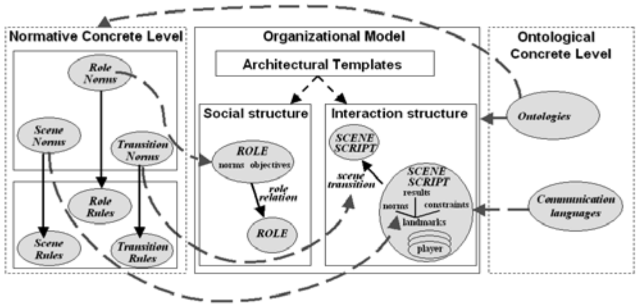

Figure 4. Connection between the multi-level ontology, the norms and the organization structure

dimansions (the Normative, the Organizational and the Ontological) are highly interconnected. In the case of OMNI this interconnection is depicted in figure 4. We can see how the ontology gives meaning to both the role and scene definitions in the Organizational Dimension, and to the terms appearing in the norms and rules in the Normative Dimension. The creation of tools integrating all three dimensions is needed in order to assist the designer in the definition of all the components, to provide ways to inter-connect them easily and to check the consistency of the design as a whole.

It is important to note that when implementing norms in MAS, the Ontology is a key component, not only in Medical applications but in other highly-regulated domains, as ontologies are needed to express the meaning of terms in norms and thereby support the reasoning. Therefore, as discussed in §2.2.2, instead to try forcing a single ontological representation for all contexts, in trans-national scenarios there is a need for multi-level, multi-context frameworks for ontologies that allow the definition and refinement of the concepts in a context connecting them with related concepts in their neightbouring contexts (for instance, connecting spanish definition of blood with the definition given in European Law). This connection may be also valuable to allow the communication of agents coming from different contexts (e.g., Spain, the Netherlands), by using, for instance, the terms that both ontologies inherit from a shared super-context (European Law) in order to build communication.

ACKNOWLEDGEMENTS

This paper compiles results of work done in collaboration with Frank Dignum, Virginia Dignum, Davide Grossi and Huib Aldewereld. It also includes ideas coming from valuable discussions with Ulises Cortés, Julian Padget and Owen Cliffe.

REFERENCES

[1] A. Abdul-Rahman and S. Hailes, 'Supporting trust in virtual communities', in *Proceedings of the Hawaii International Conference on System Sciences*, (2000).

[2] A. Aldea, B. López, A. Moreno, D. Riaño, and A. Valls, 'A multi-agent system for organ transplant co-ordination.', in *Proceedings of the 8th. European Conference on Artificial Intelligence in Medicine, Portugal, 2001.*, eds., Barahona Quaglini and Andreassen, Lecture Notes in Artificial Intelligence 2101: Artificial Intelligence in Medicine, pp. 413–416, (2001).

[3] A Logical Framework for Ethical Behaviour between Infohabitants in the Information Trading Economy of the Universal Information Ecosystem (ALFEBIITE). http://www.iis.ee.ic.ac.uk/˜alfebiite/ab-home.htm.

[4] A. Artikis, *Executable Specification of Open Norm-Governed Computational Systems*, Ph.D. dissertation, Department of Electrical & Electronic Engineering, Imperial College London, November 2003.
[5] G. Boella and L. van der Torre, 'Fulfilling or violating norms in normative multiagent systems', in *Proceedings of IAT 2004*. IEEE, (2004).
[6] G. Boella and L. van der Torre, 'Normative multiagent systems', in *Proceedings of Trust in Agent Societies Workshop at AAMAS'04*, New York, (2004).
[7] M. Calisti, P. Funk, S. Biellman, and T. Bugnon, 'A multi-agent system for organ transplant management', in *Applications of Software Agent Technology in the Health Care Domain*, eds., A. Moreno and J. Nealon, Whitestein Series in Software Agent Technologies. Birkhäuser Verlag, (2003).
[8] C. Castelfranchi, 'Formalizing the informal?: Dynamic social order, bottom-up social control, and spontaneous normative relations', *Journal of Applied Logic*, **1**(1-2), 47–92, (February 2003).
[9] C. Castelfranchi, F. Dignum, C. Jonker, and J. Treur, 'Deliberative Normative Agents: Principles and architecture', in *Proc. of the 6th Int. Workshop on Agent Theories, Architectures, and Languages (ATAL-99)*.
[10] Communication com(2000) 285 final, from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions on the health strategy of the European Community, 16 may 2000.
http://europa.eu.int/comm/health/ph_overview/strategy/health_strategy_en.htm.
[11] Communication COM(2002) 263 final, from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions. eEurope 2005: An information society for all. 28 may 2002.
http://europa.eu.int/information_society/eeurope/2002/news_library/documents/eeurope2005/eeurope2005_en.pdf.
[12] Communication COM(2004) 356 final, from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions. e-Health - making healthcare better for European citizens: An action plan for a European e-Health Area, 30 april 2004.
http://europa.eu.int/information_society/doc/qualif/health/COM_2004_0356_F_EN_ACTE.pdf.
[13] F. Dignum, 'Autonomous Agents with Norms', *AI and Law*, **7**, 69–79, (1999).
[14] F. Dignum, D. Morley, and E.A. Sonenberg, 'Towards socially sophisticated BDI agents', in *DEXA Workshop*, pp. 1134–1140, (2000).
[15] V. Dignum, *A Model for Organizational Interaction: based on Agents, founded in Logic*, SIKS Dissertation Series 2004-1, SIKS, 2004. PhD Thesis.
[16] V. Dignum and F. Dignum, 'Modeling agent societies: Coordination frameworks and institutions', in *Progress in Artificial Intelligence*, eds., P. Brazdil and A. Jorge, LNAI 2258, pp. 191–204. Springer-Verlag, (2001).
[17] V. Dignum, J. Vázquez-Salceda, and F. Dignum, 'OMNI: Introducing Social Structure, Norms and Ontologies in Agent Organizations', in *Second International Workshop on Programming Multiagent Systems Languages and tools (PROMAS 2004), New York, USA*, (2004).
[18] The Donor Action Foundation.
http://www.donoraction.org.
[19] ESCULAPE: Use of computer techniques for tissues matching and analysis as an aid to human transplantation.
http://dbs.cordis.lu/cordis-cgi/srchidadb?CALLER=EN_CORDIS&QZ_WEBSRCH=ESCULAPE.
[20] M. Esteva, J. Padget, and C. Sierra, 'Formalizing a language for institutions and norms', in *Intelligent Agents VIII*, eds., J.-J.CH. Meyer and M. Tambe, volume 2333 of *LNAI*, pp. 348–366. Springer Verlag, (2001).
[21] Eurotransplant International Foundation.
http://www.eurotransplant.nl.
[22] J. Fox and S. Das, *Safe and Sound*, AAAIPress/MIT Press, 1^{st} edn., 1999.
[23] J. Fox and S. Das, 'Guardian agents for safety-critical systems', in *Workshop on Autonomous Agents in Health Care*, ed., V. Shankararaman, pp. 25–34, (2000).
[24] D. Gambetta, *Can We Trust Trust?*, 213–237, Basil Blackwell, Oxford, 1990.
[25] F. Giunchiglia and P. Bouquet, 'Introduction to contextual reasoning', *Perspectives on Cognitive Science*, **3**, (1997).
[26] R.V. Guha, *Contexts: A Formalization and Some Applications*, Standford University, 1991. PhD. Thesis.
[27] F. López y López and M. Luck, 'Towards a Model of the Dynamics of Normative Multi-Agent Systems', in *Proc. of the Int. Workshop on Regulated Agent-Based Social Systems: Theories and Applications (RASTA '02)*, eds., G. Lindemann, D. Moldt, M. Paolucci, and B. Yu, volume 318 of *Mitteilung*, pp. 175–194. Universität Hamburg, (2002).
[28] F. López y Lopez, M. Luck, and M. d'Inverno, 'A framework for norm-based inter-agent dependence', in *Proceedings of The Third Mexican International Conference on Computer Science*, pp. 31–40. SMCC-INEGI, (2001).
[29] J. McCarthy, 'Generality in artificial intelligence', *Communications of the ACM*, **30**(12), 1030–1035, (1987).
[30] A. Moreno, A. Valls, and J. Bocio, 'Management of hospital teams for organ transplants using multi-agent systems.', in *Proceedings of the 8th. European Conference on Artificial Intelligence in Medicine, Portugal, 2001.*, eds., Barahona Quaglini and Andreassen (Eds.), Lecture Notes in Artificial Intelligence 2101: Artificial Intelligence in Medicine, pp. 374–383., (2001).
[31] A. Moreno, A. Valls, and A. Ribes, 'Finding efficient organ transport routes using multi-agent systems.', in *IEEE 3rd International Workshop on Enterprise Networking and Computing in Health Care Industry (Healtcom), L'Aquilla, Italy.*, (2001.).
[32] Y. Moses and M. Tennenholtz, 'Artificial social systems', *Computers and AI*, **14**(6), 533–562, (1995).
[33] A.S. Rao and M.P. Georgeff, 'An abstract architecture for rational agents', in *Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning (KR'92)*, pp. 439–449. Morgan Kaufmann publishers Inc.: San Mateo, CA, USA, (1992).
[34] A.S. Rao and M.P. Georgeff, 'BDI-agents: from theory to practice', in *Proceedings of the First International Conference on Multiagent Systems*, San Francisco, (1995).
[35] Report HLPR/2003/16, from the Health and Consumer Protection Directorate-General. High Level Process of Reflection on Patient Mobility and Healthcare Developments in the European Union, 9 december 2003.
http://europa.eu.int/comm/health/ph_overview/Documents/key01_mobility_en.pdf.
[36] REgional and International Integrated Telemedicine Network for Medical Assistance in End Stage Diseases and Organ TRANSPLANT.
http://retransplant.vitamib.com.
[37] J. Sabater and C. Sierra, 'Reputation and social network analysis in multi-agent systems', *Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems AAMAS-02. Bologna, Italy*, (2002).
[38] Scandiatransplant. http://www.scandiatransplant.org.
[39] V. Shankararaman, V. Ambrosiadou, T. Panchal, and B. Robinson, 'Agents in health care', in *Workshop on Autonomous Agents in Health Care*, ed., V. Shankararaman, pp. 1–11, (2000).
[40] TECN : Transplant Euro Computer Network.
http://www.ejeisa.com/nectar/t-book/html/health.htm#TECN.
[41] United Network for Organ Sharing.
http://www.unos.org.
[42] A. Valls, A. Moreno, and D. Sánchez, 'A multi-criteria decision aid agent applied to the selection of the best receiver in a transplant.', in *4th. International Conference on Enterprise Information Systems, Ciudad Real, Spain.*, (2002).
[43] J. Vázquez-Salceda, *The Role of Norms and Electronic Institutions in Multi-Agent Systems. The HARMONIA framework*, Whitestein Series in Software Agent Technology, Birkhäuser Verlag, 2004.
[44] J. Vázquez-Salceda, H. Aldewereld, and F. Dignum, 'Implementing Norms in Multiagent Systems.', in *Second German Conference on Multiagent System Technologies, Erfurt, Germany.*, (2004).
[45] J. Vázquez-Salceda, U. Cortés, and J. Padget, 'Integrating the organ and tissue allocation processes through an agent-mediated electronic institution', *LNAI-2504*, 309–321, (2002).
[46] J. Vázquez-Salceda, V. Dignum, and F. Dignum, 'Organizing multiagent systems', Technical Report UU-CS-2004-015, Institute of Information and Computing Sciences, Utrecht University, (2004).
[47] B. Yu and M.P. Singh, 'A social mechanism of reputation management in electronic communities', *Proceedings of Fourth International Workshop on Cooperative Information Agents*, 154–165, (2000).
[48] G. Zacharias and P. Maes, 'Trust management through reputation mechanisms', *Applied Artificial Intelligence*, (14), 881–907, (2000).