# Mapping Abstractions of Norms in Electronic Institutions

Ismail Khalil Ibrahim, Gabriele Kotsis and Wieland Schwinger
*Johannes Kepler University Linz, Austria*
*{ismail / gk / wieland}@tk.uni-linz.ac.at*

## Abstract

*Electronic institutions are the agents' counterpart of human organizations, which are specifically designed for providing support, trust, and legitimacy in electronic commerce applications. Two approaches have been advocated for the design and modeling of multi agent systems in an environment that is governed by some kind of (social) norms: in* coordination *strategy, multi agent systems are defined as a set of entities regulated by mechanisms of social order and created by more or less autonomous actors to achieve common goals; in* cooperation *strategy, agents model specific roles in the society and interact with each other as means to accomplish their goals. In this paper, we argue that there is a relative similarity between the two approaches with respect to their use of norms as constraints on the social behavior of multi agents systems.*

## 1. Introduction

The study and modeling of norms has attracted the interest of scientists from different disciplines such as sociology, economics, philosophy, and computer science. [1, 2, 4, 7, 11]

According to sociology, a norm is a rule or standard of behavior shared by members of a social group. Norms are more specific than values or ideals: honesty is a general value, but the rules, defining what is honest behavior in a particular situation, are norms [2, 5, 6, 10].

According to philosophy, a norm is an authoritative rule or standard by which something is judged and on that basis approved or disapproved. Examples of such norms include standards of right and wrong, beauty and ugliness, and truth and falsehood.

According to economics, a norm (from norma, Latin for carpenter's level) is a model of what should exist or be followed, or an average of what currently does exist in some context, such as an average salary among members of a large group.

Several researchers [3, 5, 6, 7] have recognized that the design and modeling of (social) norms in multi agent systems can benefit from abstractions analogously to those employed by sociology, philosophy, and economics.

Thus, norms in the context of multi agent systems can be defined as expectations about what behavior, thoughts, or feelings are appropriate within a given group (of agents) within a given context (environment).

In the same way, agents in multi agents systems can be redefined in the context of social norms as intentional entities who are self-aware (autonomous: free to perceive, free to make decision, free to act), who have goals, who can think about how to match up to their own and other agents standards, and who will act in ways to influence other's view.

Agents in this sense can accept a task, negotiate or reject it. They should be capable to communicate, to coordinate, and to cooperate with other agents in order to solve a problem.

As agents communicate by messages exchange, agent interactions can be either actual: where two or more agents interact directly, or imagined: where different agents have mental representation of each other, or implied: where agents interact indirectly.

In this context, we notice that agents in a multi agent systems 1.) are created for some reason (purpose and orienting towards it), 2.) have values which determine their behavior and based on these values, agents' actions can be decided as good, right, and even moral, 3.) have control over their actions (accept a task, negotiate, or reject) and this control is maximized by reaching challenging goals, and 4.) have positive values (self respect and respect of others)

## 2. Structure of the Paper

The reminder of the paper is organized as follows: in the next section, we review the literature relevant to this line of work. In section 4, we present our approach for modeling norms in electronic institutions that allows us to formulate the cooperation strategy as an institution which will be presented with adequate examples in the next section. Finally, we conclude with some remarks on our plans to design and implement the normative structure of the institution.

Proceedings of the Twelfth IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE'03)
1080-1383/03 $17.00 © 2003 IEEE

## 3. Norms and Institutions

Electronic institutions are the agents' counterpart of human organizations, which are specifically designed for providing support, trust, and legitimacy in electronic commerce applications. They function in the same way our human organizations function to create trust, prevent fraud and reduce deception by verifying rules, regulations which is at the final stage maintaining norms and insuring that all the participants are conforming to these norms.

Two approaches have been advocated for the design and modeling of norms in electronic institutions: in the cooperative normative behavior [2, 4, 6, 7], the norms are defined implicitly through the behavior of the agents and depends on how agents function in the institution; in the coordination strategy [5, 9, 10, 11], where the norms of the electronic institution are defined explicitly and the agents in the institution have to comply to these norms.

Most of the cooperative normative behavior model is based on the assumption that different agents in the institution have some common goal or intention and this common goal enforces some type of cooperative behavior on all agents in the institution. This means that the rules and regulations to which agents adapt their behavior are hard wired in the protocols used by the agents to interact with each other.

In the coordination model, actions of agents in the electronic institution are either rational or norm guided. In this case, norms are nothing but instruments of individual, collective, or generic optimization and to accept norms, as a motivational mechanism is not to violate methodological individualism, or to deny the importance of rational choice. In this case, the outcome of agent actions is a compromise between what the norm prescribes and what rationality dictates.

Both approaches of coordination and cooperative normative behavior for implementing and modeling norms in electronic institutions can be compared with respect to conformity; by conformity we mean yielding to, or going a long with a perceived social norm within an institution. Conformity can be measured based on the social influence of an agent or group of agents on the attitudes, beliefs, or behavior of the others, the conformity of an agent to a norm perceived by other agents, the compliance of an agent to act in accordance with a direct request of another agent, or the obedience of an agent to act in response to direct orders from authoritative institutional agents.

## 4. Modeling a Normative Structure for Electronic Institutions

Society has set up different kinds of mechanisms, methods, and conditions for adopting and shaping norms. In this paper, we distinguish between formal norms, which are defined by rules and regulations and informal norms, e.g. cultural and social behavior.

Both types of norms are socially and culturally perceived as standards for behavior, which guide the action and specify what must be done and what is forbidden in a given situation.

### 4.1. Norms Abstraction

We [8] identify three levels of abstractions for norms in electronic institutions:

- values,
- normative behavioral rules, and
- low level norms.

**Values.** These are the values and ethics which are central to any organization; both human and electronic. Values and ethics are extremely broad terms, but our focus is on the aspects most relevant to the purpose of the creation of the institution and based on these values, decisions and actions are taken. Values are the embodiment of what an organization stands for, and should be the basis for the behavior of its members. Examples of such kind of values are those related to the integrity of the institution, professionalism and excellence, and contractual obligations to mention few.

**Normative behavioral rules**. These are the rules of conduct within the institution. Normative behavioral rules are necessary to judge if the behavior of the members of the institution complies with the values defined on the first level of the abstraction. Examples of these rules are declaration of interest (registration), accuracy of information, and fair competition.

**Low level norms**. This is the lowest level of abstraction. Examples are social norms (organization, roles, authority), communication norms (protocols, dialogues), and behavioral norms (obligations, permissions, interdiction).

There is a broad body of work on norms of the first and the second levels, and a general consensus on what (values) and (normative rules) mean. However, there is less work on the meaning of norms, and certainly no consensus. What is meant by a norm can differ widely. For instance, one may define that norms must be consistent with the institution, or define that they must be provable statements, deducible from the institution. Another view is that norms really represent meta-knowledge – knowledge about the institution itself- and should, perhaps, be written in a non-classical logic.

In the sequel of this paper, we use the standard syntax for first-order logic, with the usual symbols for variables, connectives, quantifiers, punctuation, equality, constants, and predicates.

Now, we redefine the normative structure of an institution to have three components:

- ***values*** which are logical facts about the institution:
  *institution(fair, responsible)*
  *agent(polite, cooperative, helpful)*
  *information(accurate)*
- ***normative rules*** which are clausal formulas:
  *transaction(secure)* ← *institution(fair, -), agent(polite,-,-) ,information(accurate)*
- ***norms*** which are from the perspective of implementation, are restricted to denial constraints; that is, formulas of the form
  ← *L1,..., Ln*
  in which the $L_i$'s are atoms for example
  - *there is no question asked and don't get answer in a multi-agent system*
    ← *agent(polite,-,-), ask($a_i$ , $a_j$, Φ), ans($a_j$ , $a_i$, Φ)*
  - *there is no age discrimination in an electronic institution*
    ← *institution (fair,-), age(young), age(old)*
  - *there is no entrance to the institution without registration*
    ← *institution(-,responsible)*

## 4.2. Semantics of the Normative Structure

It is very useful to find out the relations and interactions between these three levels of abstraction through the use of analogy to the concepts of classes and inheritance in object oriented programming as shown in Figure 1.

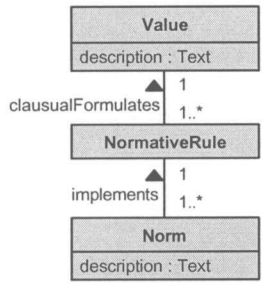


**Figure 1. Hierarchical meta-model of the normative structure**

Thus we formalize the hierarchy of the Normative Structure (represented in a UML class diagram) to have values which are aspects that lead to semantics, and having normative rules which are sort of these values, each norm is a specialization of a normative rule, each normative rule is a generalization of the norms belong to it, all subclasses of normative rules (norms) function the same way, and the normative rules extend the concept of values.

The normative structure of the electronic institution has both structure and ontologies, which describe the semantics of the data. Ontologies allow designers of electronic institution to organize the normative structure into taxonomies of concepts, each with their attributes, and describe relationships between concepts.

When the normative structure is marked using ontologies, agents can better understand the semantics of the normative structure and therefore intelligently decide to comply or violate the norms perceived by this structure. The following two examples illustrate the situation. The previously presented meta-model implemented in terms of UML profile, i.e., through dedicated stereotypes.

**Example 1**

Consider an institution, which set the following vision as a value:

*"Maintaining an impeachable standard of integrity in all business relationships"*

***Integrity***, in terms of data and network security, means the assurance that information can only be accessed or modified by those authorized to do so. This value of "integrity" can be translated into two normative rules; one is related to the ***security*** of the institution; and the second is concerned with the ***consistency*** of the institution. These two normative rules can be further translated into low-level norms. The normative structure of the example is depicted in Figure 2.

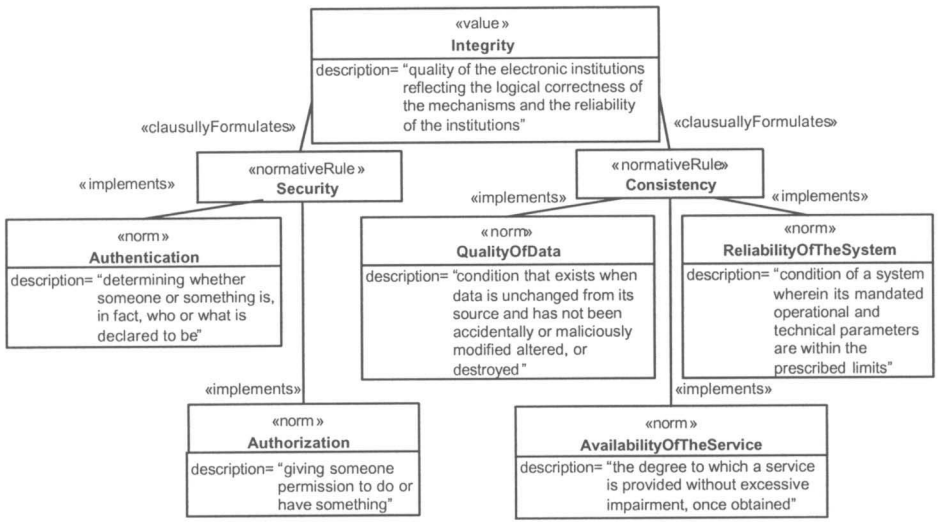


**Figure 2. Class model representing the normative structure of Example 1**

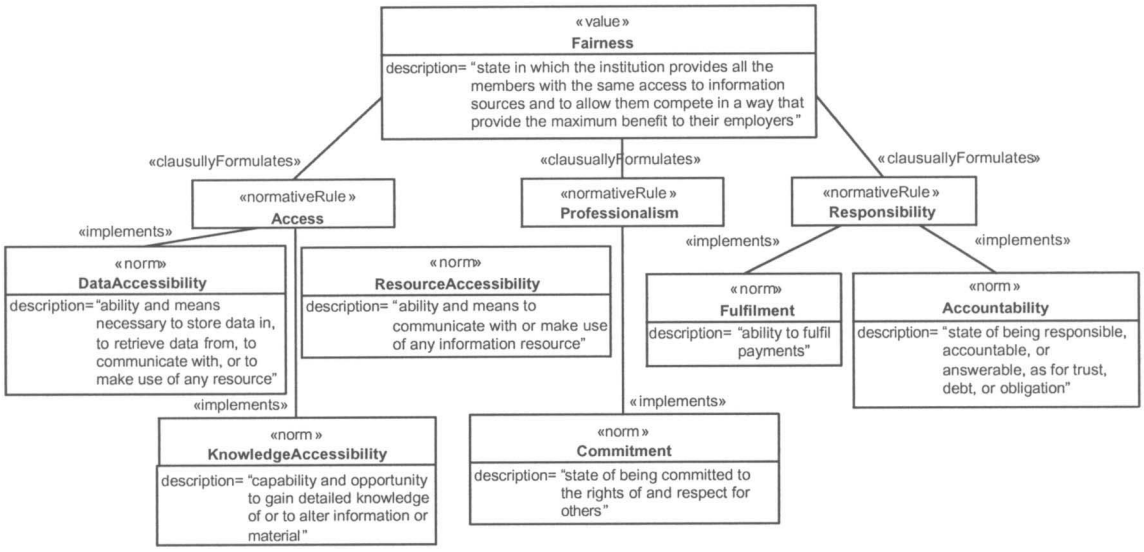


**Figure 3. Class model representing the normative structure of Example 2**

**Example 2**

Now, let us turn to another value in an institution

*"Fostering the highest standards of fair competence amongst those for whom the institution is responsible"*

The notion of ***fairness*** is a very general concept and can be used to coin terms in many different application areas. In electronic commerce the term fair exchange refers to the problem that two parties want to swap goods, services, information, or payments in a way which insures that no party can gain advantage over the other. This means that the value of "fairness" can be translated into three normative rules; one is related to the ***access*** to resources within the institution; and the second is concerned with the ***professionalism*** with which the institution is operating, and the third is ***responsible management***. These three normative rules can be translated further into low-level norms. The normative structure of the example is depicted in Figure 3.

## 5. Norms Abstractions as a Cooperation Strategy

Let us consider a simple scenario to motivate the approach taken in this paper:

*When an agent needs a service from another agent belonging to an institution, it should enter*

*the institution, find about rules governing that institution, and behave accordingly. The agent decides when to enter and when to exist, but once entered, it obeys the institution rules.*

From this simplified scenario, we can abstract what we call scenes such that for each activity that can take place in the institution, there is a corresponding scene, in which interactions between agents are articulated through agent group meetings that follow a well-defined protocol. This set of scenes and the connections between them – what roles agents may play in them, how many of each role, to which scenes they may move – constitute the performative structure for the electronic institution.

The purpose of this performative structure is to show the different scenes, which comprise the institution by means of a transition graph.

## 5.1. Cooperation defined as an institution

For our purpose, an institution is solely defined by the name of the institution and its arity (number of scenes). As we have seen cooperation strategy models the institution as a schema

$$E = (A, I, \Sigma, N)$$

where:

$E$ is the name of the institution

$A$ is the set of agent/role pairs

$I$ is the union of local illocution schemas in the form of $<P, S, R, C>$ where P is an illocutionary particle (request, accept, deny, inform, or pay), $S$ is the sender agent identifier (agent/role pair), $R$ is the receiver agent identifier (agent/role pair), and $C$ is the content of the message.

$\Sigma$ is the set of scene definitions for the component scenes in terms of the institution

and $N$ is a set of norms on the component scenes.

Given a set of tuples for the Agent/Roles combinations, an instance of the institution is defined by the minimal model of $\Sigma \cup A$.

An instance of the institution is consistent if it is a model of $N$, i.e. the institution complies to the set of norms. In our case, we only need to verify that $A$ is also a model of $N$ i.e., , the scenes are only materialized, say $J$ is the actual instance of the *illocution schema*, an instance of the institution is an instance of $A \cup J$ such that it corresponds to the minimal model $\Sigma \cup A$, and $A$ is consistent with $N$. Notice that $A$ may not be unique. For the sake of simplicity, we will assume that the modeler of the solution has been careful and that a minimal instance is guaranteed to exist.

In this model we cannot reason about compliance to norms, because the institution is described at a high level of abstraction. So we introduce a dual view which will allow us to check compliance to norms at run time.

## 5.2. A dual view of the institution

In order to comply with the reality of the situation in the modeling, we would need to construct an institution

$E'' = (A, I, \Sigma'', N'')$ with the same instances as $E$.

Using similar transformation as the one used in [8], we construct the institution as

$E' = (A \cup I, 0, 0, N')$ which has the same instances as $E$.

Given a scene definition for $s$:

- $s(\overline{X}) \leftarrow B(\overline{Y})$

we first make all the implicit quantifications explicit

- $\forall \overline{X}\ s(\overline{X}) \leftarrow \exists (\overline{Y} - \overline{X}), B(\overline{Y})$

the completed axiom is

- $s(\overline{X}) \Leftrightarrow \exists (\overline{Y} - \overline{X}), B(\overline{Y})$

which we can transform into two norms

- (1) $s(\overline{X}) \rightarrow \exists (\overline{Y} - \overline{X}), B(\overline{Y})$
- (2) $B(\overline{Y}) \rightarrow s(\overline{X})$

if we call $\hat{N}$ the set of norms of type 2, the new institution

$$E' = (A \cup I, 0, 0, \Sigma \cup N \cup \hat{N})$$

We claim but do not prove that the original and transformed institutions have the same instances. Let us look at the application of the transformation on an example.

**Example 3**

*Let us consider the following institution*

- *INST = ({w/3}, {v/2}, {v(X,Y) ← p{X,Y,Z)}, 0} where w is the name of the institution with three scenes, two illocution schemas have been uttered which lead that agents X, Y have moved from scene p to scene v and there are no norms specified for scene p. According to the transformation of the institution we have described, we generate the following four norms*
- $n_1 = reg(A_i,-,-) \rightarrow \exists B_j \exists C_j\ auc(A_j,B_j,C_j)$
- $n_2 = s_1(A_i,B_i,C_i) \leftarrow s(Agent, Role)$

- $n_3 = auc(A_i,-,-) \rightarrow \exists B_j \exists C_j\, sett(A_j,B_j,C_j)$
- $n_4 = s_2(A_j,B_j,C_j) \leftarrow s(Agent, Role)$

*we are now considering the institution schema to be:*

$INST' = (\{I/3, s/2\}, 0, 0, \{n_1, n_2, n_3, n_4, n_5\}$

*where* $n_5$

- $s_1(A_i,B_i,C_i), s_2(A_j,B_j,C_j) \rightarrow A_i = A_j$

## 6. Conclusions and Future Work

In order to check norms and act on possible violations of the norms by the agents within an institution, the abstract norms have to be translated into actions and concepts that can be handled within the institution. In this paper, we have categorized the abstract norms in the ways they are trying to regulate the behavior of the agents. Subsequently we have indicated how a translation can be made between the abstract norms and more concrete forms. We have also indicated the relations that exist between the abstract and concrete norms (although we do not get to a complete formalization of this aspect).

The abstract norms try to capture many different situations and therefore are "vague" in several different ways:

1. They are referring to an abstract action that can be implemented in many ways
2. They use terms that are vague and that have to be defined separately
3. They refer to actions that are not (directly) controllable and/or checkable

The purpose of the project from which this research paper stems is to build a complete monitoring system for electronic markets. Our objective is to integrate seamlessly the cooperation and coordination models as a basis for modeling electronic institutions.

### References


[1] Boman, M. (1997). Norms as Constraints on Real_time Autonomous Agent Action. In: M. Boman, W. van de Velde (eds.), Multi-Agent Rationality, Proc. of the 8th Workshop on Modelling Autonomous Agents in a Multi- Agent World, MAAMAW'97, Lecture Notes in AI, vol. 1237, Springer Verlag, Berlin, 1997, pp. 36-44.

[2] Castelfranchi, C., Dignum, F., Jonker, C, and Treur J., (1999). Deliberative Normative Agents: Principles and Architectures, Proc. of the 6th Workshop on Agent Theories, Architectures and Languages, ATAL'99. Orlando, pp. 206-220.

[3] Cohen, P., and Levesque, H. (1991). Teamwork, Nous, vol.35, pp. 487-512.

[4] Conte, R., Castelfranchi, C., and Dignum, F. (1999). Autonomous Norm-Acceptance, In: J.P. Mueller, M.P. Singh, A.S. Rao (eds.), Intelligent Agents V, Proc. of the 5th Workshop on Agent Theories, Architectures and Languages, ATAL'98. Lecture Notes in AI, vol. 1555, Springer Verlag, Berlin.

[5] Conte, R., Falcone, R., Sartor, G. (1999). Agents and Norms: How to fill the gap? In: Conte, R., Falcone, R. and Sartor, G. (eds) Agents and Norms, special issue, AI & Law, Vol.7, No.1, pp. 1-15, 1999.

[6] Dignum, F. (1996). Autonomous Agents and Social Norms. In: R. Falcone and R. Conte (eds.), ICMAS'96 Workshop on Norms, Obligations and Conventions, Kyoto, pp. 56-71. Revised version in In AI & Law, Vol.7, 1999, No.1, pp. 69-79.

[7] Esteva, M., Padget, J., Sierra, C., (2001), Formalizing a language for Institutions and Norms, Proc of the 8th Workshop on Agent Theories, Architectures, and Languages, ATAL-2001, Seattle.

[8] Ibrahim, I. K., Weippl, E., Dignum, F., Schwinger, W., (2002), Modeling Norms in Electronic Institutions, Proc. of the Workshop on Agent-Based Simulation, Passau, Germany.

[9] López y López, F., and Luck, M., and d'Inverno, M. (2002). Constraining autonomy through norms. Proc. of the 1st Joint Conference on Autonomous Agents & Multiagent Systems, AAMAS 2002, ACM, pages 674-681

[10] Omicini, A., and Zambonelli, F., Klusch, M., and Tolksdorf, R. (eds.) (1991). Coordination of Internet Agents. Models, Technologies, and Applications, Springer-Verlag.

[11] Staller, A. and Petta, P. (2000), Introducing Emotions in the Computational Study of Norms, Proc. of the AISB'00 Sympoisum on Starting from Society - the Application of Social Analogies to Computational Systems, Birmingham, pages 101-112.