From Abstract to Concrete Norms in Agent Institutions

Davide Grossi and Frank Dignum

Utrecht University,
The Netherlands
{davide,dignum}@cs.uu.nl

Abstract. Norms specifying constraints over institutions are stated in such a form that allows them to regulate a wide range of situations over time without need for modification. To guarantee this stability, the formulation of norms need to abstract from a variety of concrete aspects, which are instead relevant for the actual operationalization of institutions. If agent institutions are to be built, which comply with a set of abstract requirements, how can those requirements be translated in more concrete constraints the impact of which can be described directly in the institution? In this work we make use of logical methods in order to provide a formal characterization of the *translation rules* that operate the connection between abstract and concrete norms. On the basis of this characterization, a comprehensive formalization of the notion of institution is also provided.

1 Introduction

Electronic institutions, such as auctions and market places are electronic counterparts of institutions that are established in our societies. They are established to regulate interactions between parties that are performing some transaction (see [6] for more details on the roles of institutions). Interactions are regulated by incorporating a number of norms in the institution which indicate the type of behavior each of the parties in the transaction should adhere to within that institution. The main concern of this work is to investigate what formal relation could be specified which accounts for how (abstract) norms can be incorporated in the (concrete) procedures constituting the institution, in such a way that agents operating within the institution either operate in accordance with those norms, or may be punished as they violate them.

That this relation is more complicated than just adding some constraints on the actions in the institution can be seen from the following example. The norm "*it is forbidden to discriminate on the basis of age*" can be formalized in deontic logic as "F(discriminate(x,y,age))" (stating that it is forbidden to discriminate between x and y on the basis of age). The translation of this formula would get down to something like that the action "`discriminate(x,y,age)`" should not occur. However, it is very unlikely that the agents operating within the institution will explicitly have such an action available. The action actually states something far more abstract. We claim that the level on which the norms are specified is more abstract and/or general than the level on which the processes and structure of the institution are specified. From an institutional standpoint norms need, in order to be incorporated in the institution itself, to be therefore "translated" to a level in which their impact on the institution can be described directly. A formal account of these "*translation rules*" constitutes the central aim of this work.

The work is organized in accordance with the following outline. In Section 2 some preliminaries about the notions of norms, normative systems and institutions are set forth; in Section 3 the issue addressed is made concrete by means of two examples, and our line of analysis of the problem is stated; in Section 4 a formal framework is proposed, which allows for formal definitions of the notions of abstract and concrete norms, and of translation rules; in Section 5 these definitions are used in order to provide a formal account of the notion of institution itself able to cope with the issue of abstractness of norms; in Section 6 this formal notion is shown to be embeddable in various formal argumentation systems, thus enabling the possibility of articulate institutional reasoning patterns; finally, in Section 7, some conclusions are drawn.

2 Some Preliminaries

The first concept to introduce is the concept of **norm**. As we will see later in Section 2.2, institutions are defined in terms of norms, which are therefore the basic building block, so to say, of our work. With the term norm we intend whatever in general indicates something ideal and which, consequently, presupposes a distinction between what is ideally the case and what is actually the case. In natural language norms are usually, but not always, expressed by locutions such as: "it is obligatory", "it is forbidden", "it is permitted", etc..

In this paper we will assume norms to be conditional, because that is the form in which they mostly appear in statutes and regulations governing institutions. In conditional norms we recognize the condition of application of the norm, and its normative effect, i.e. the normative consequence the norm subordinates to its condition: "under condition A, it is obligatory (respectively, permitted or prohibited) that B"

Another important concept we will come to take into consideration, though not in detail, is the concept of **procedure**. Here a procedure is seen as an algorithm-like specification describing how a certain activity is carried out. The difference between a norm and a procedure is of extreme relevance for our purposes (see Section 2.2): a norm states that something ought to be the case under certain conditions, while a procedure describes only a way of bringing something about; semantically, norms incorporate a concept of ideality, whereas for procedures it is instead central a notion of transition.

2.1 Normative Systems

In [14] normative systems are defined as follows:

> "a normative system is any set of interacting agents whose behavior can [...] be regarded as norm directed".

According to this view, a normative system is thus a norm directed agency. In this sense, a set of norms meant to direct an agency constitutes a form of (normative) specification of that agency; in other words, a set of norms addressed to a given agency determines that agency as a normative system. As such, normative systems are therefore amenable of formal description in terms of logical theories containing normative expressions[1].

There is wide agreement upon the fact that all normative systems of high complexity, like for example legal systems, cannot be regarded simply as sets of norms ([14, 13]). Besides norms, they consist also of definitional components yielding a kind of contextual definition: "A means (counts as) B in context i". An example: "signing form 32 counts as consenting to an organ donation, in the context of Spanish transplant regulation [26][2]". Normative components of this type are known in legal and social theory as *constitutive norms*, while purely normative components, i.e. what we called norms, are known as *regulative norms* (see for example [12, 19, 25]). Both these components will be logically represented (Section 4) by means of rules: regulative norms via rules having a deontic consequent *normative rules*; constitutive norms via *translation rules*. Concepts introduced are recapitulated in Table 1.

Table 1. Normative systems' components

COMPONENTS	regulative norms	constitutive norms
REPRESENTATION	normative rules	translation rules

2.2 Institutions

The term institution is quite ambiguous. Following [17] we distinguish two senses of the term, which are of significance for our purposes.

– First, an institution can be seen a the set of agents with specific roles, private and common objectives, the activities of which are procedurally determined.

[1] This is precisely how normative systems are conceived in [1], where they are analyzed as sets of sentences deductively connecting normative conditions to normative effects.

[2] These examples have been chosen on the basis of work carried out on the regulations from which they are excerpted.

We speak in this case about institutions seen as **organizations**. As an example, the agents operating Utrecht Hospital, and the set of procedures according to which their activity is planned, constitute an organization.
– Second, an institution can be seen as the set of norms (constitutive and regulative) an organization can instantiate implementing them. We use in this case the term **institutional form**. In this sense the set of regulations holding at Utrecht Hospital defines an institutional form. Also the set of regulations concerning hospitals in The Netherlands defines an institutional form, namely a general institutional form, say, "hospital". The organization of Utrecht Hospital instantiates both these institutional forms.

This distinction between organizations and institutional forms lies in the aforementioned distinction between norms and procedures. While analyzing institutions as organizations emphasizes the procedural aspects involved in operating institutions, an analysis of them in terms of institutional forms stresses instead the normative nature of institutions specifications. This last perspective on institutions is the one underpinning the analysis of abstract and concrete norms that will be carried out in the next sections. Viewing institutions as institutional forms, that is to say, as sets of constitutive and regulative norms, allows for an application of a normative system perspective ([13, 14]) to their analysis and will lead, in Section 5, to a formal definition of institutions as sets of rules[3].

It is instructive to spend still some more words on the distinction proposed. The relation between these two conceptions of institutions constitutes a very interesting issue, which is also of definite relevance in relation with the general problems addressed here. What is at stake is the understanding of how an organization implements an institutional form, or in other words, how can a set of procedures implement a set of norms, what is the formal link between norms and procedures. Answering these questions would lead to a deeper understanding of the variety of aspects characterizing institutionalized agencies. This problem forms nevertheless a separate issue, which will not be explicitly dealt with in the present paper[4].

3 Abstractness of Norms

3.1 Abstract norms and concrete norms

The issuing of norms, as it appears in various statutes or regulations specifying constraints over institutions, has the characteristic of stating norms in such a form that allows them to regulate a wide range of situations and to be stable for a long period of time. The vaguer or abstract norms are, the easier it becomes to keep them stable. The downside of this stability is that normative formulations

[3] The formal analysis of *organizations*, i.e. procedural description of agencies, is therefore left aside in this work. In what follows we will use the terms institution and institutional form interchangeably.

[4] See [7] for some first thoughts on this topic.

seem to be less well defined. In law it is even an explicit task of the judges to interpret the law for specific situations and determine whether someone violated it or not.

It is our thesis that abstract and concrete notions are described within different ontologies. Concrete norms are described in terms of the concepts that are used to specify (possible) procedural descriptions of the concrete institutions. Abstract levels are instead described using a more general ontology.

In order to precisely illustrate the problem we are concerned with, we discuss two examples. The first one is taken from the Dutch regulation about personal data treatment within police registers ([8]). In the mentioned regulation the following norm is stated: "*the inclusion of personal data in a severe criminality register occurs only when it concerns: a) suspect of crimes; b) etc.*" (Article 13a). This norm states that, under certain conditions, personal data may be included in a specific kind of police register. Suppose now that an electronic institution for that register has to be built which fully complies with the norms regulating the use of that register ([5]). The following question comes naturally about: "*what can be concretely included in the register*", that is "*what is classified to be* personal data *in the context of* [8]"? That this is more than just a definitional issue can be seen from the fact that more data may be included as they regard suspects and less as they regard persons which are indirectly connected with a crime: the notion of personal data varies. These "variations" are specified in the model regulations on police registers ([16]).
The second example is instead taken from the Spanish regulation on organ transplantation ([26]): "*a living donor must consent before a transplantation may take place*" (Article 9). An analogous question can be raised: "*what is understood as* consent *in the context of* [26]"? This example shows that abstraction takes place over data (first example) as well as over actions. The *consent* action can be implemented by *signing form 32* within the context of the transplant regulation in Spain. However, this way of implementing consent is only "valid" within that context.

On what basis are we entitled to consider the above translations as complying with the abstract ones? Signing a form seems a reasonable implementation of giving consent, whereas we would probably not accept *wearing a red hat* as a way of implementing consent. What does the connection between abstract and concrete normative formulations consist of, from a formal point of view? This is the central question we are here addressing.

3.2 Connecting abstract and concrete norms

The model regulation on severe criminality registers ([16]) is explicitly conceived to lead to an application of the law in the context of the usage of severe criminality registers. The following norm is stated: "*[In a severe criminality register] the following kinds of data can at most be included: financial and corporate data; data concerning nationality; etc.*" (Article 6). Basically, this article provides the list of data that are allowed to be included in the register, and it therefore consists of a concrete version of Article 13a cited in Section 3.1. Such a "translation", as we called it, is possible because an interpretation of the notion of *personal data* occurring in Article 13a, is somehow presupposed: "*personal data are financial and corporate data; data concerning nationality; etc.*". This rule, defining the notion of personal data within the context of the usage of severe criminality registers, states that if something is a datum concerning the nationality of, for instance, a suspect, then this datum is a personal datum and it can therefore be legally included in the register. We claim these rules to constitute the connection between abstract and concrete norms.

In this example, being a personal datum is an abstract fact exactly because something can be a personal datum in many ways, depending on the context: in the context of the regulation of severe criminality registers, data as specified in Article 6 count as personal data, but within a different context, for example in the regulation about so called provisional police registers, something else can count as a personal datum. Abstract constraints are stable and hold for many situations because they are made concrete in several, possibly different, ways. The contextual nature of these translation rules led us to the logical framework we are going to expose in Section 4.

To understand this contextual nature of institutions it seems useful to see them as regulating facts that hold on specific levels of abstractness: concrete levels are the levels on which facts hold that can be directly handled by the procedures an institution is organized through (something is a datum concerning nationality); abstract levels are the levels on which more abstract facts hold (something is a personal datum), and to which many more concrete levels can be seen to converge via translation rules. We therefore understand institutions as sets of norms and translation rules which regulate facts holding on levels of abstractness[5]. Such a perspective also shows how more particular institutions, such as the ones operating severe criminality registers, are nested in more general ones, such as the one regulating the use of police registers in general. This nesting takes place through the abstractness layering. Picture 1 below provides a graphical account of the intuitions just exposed.

Analogous considerations may be carried out in relation with the second example mentioned in Section 3.1.

4 Formal Framework

4.1 A logic for levels of abstractness

Before presenting a proposal to formally capture the notion of level (context) we have in mind, it is necessary to identify, in further detail, the features of this concept that we would like to be able to express in our formalism.

1. In our view, levels constitute a structure ordered according to the relation "i is strictly less abstract than j". This relation is, reasonably, irreflexive, asymmetric and transitive. Moreover, it seems intuitive to assume it to be

[5] See section 2.

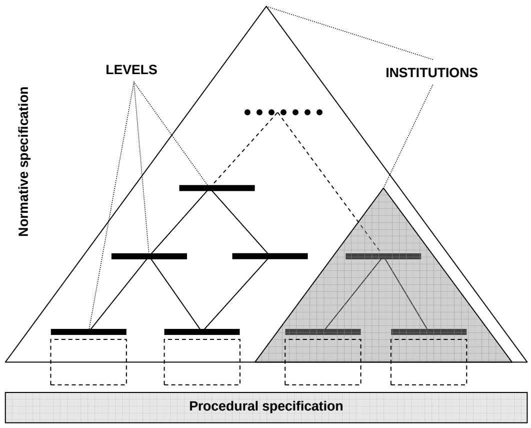

Fig. 1. Institutions and levels

partial. There might be levels i and j both strictly less abstract than a given level k, but such that they remain unrelated with respect to each other[6].

2. Levels are such that what holds in a level holds irrespectively of the level from which that fact is considered: if at level i the donor expresses his/her consent, then at level j it holds that at level i the donor expresses his/her consent and vice versa.
3. No inconsistency holds at any level, levels are coherent.
4. Finally, there exists a trivial "outermost level", representing the absence of context, that is, the level of logical truths.

To capture these features we use a multi modal logic $KD45_n^{i-j}$ ([15]) which corresponds to a propositional logic of n contexts (PLC) with: consistency property (corresponding to feature 3), flatness property (feature 2), outermost context (feature 4) and total truth assignments (see [18, 4, 3])[7].

Language. The alphabet of language $\mathcal{L}^L$ for levels of abstractness expands the language for propositional logic and contains the following sets of symbols:

---

[6] Notice that these are precisely the properties also of the *conventional generation* relation analyzed in [10].

[7] We deemed a multi modal formalism to be better readable than a propositional context logic one. This is the reason why we chose for using a modal logic formulation instead of a contextual logic one. The correspondence result we claimed is guaranteed by results proved in [3]. A word must be spent also about the use of propositional context logic with total truth assignments. In fact, partial truth assignments are one of the most relevant features of context logics as introduced in [4, 3]. However, it has been proved in [18] that every propositional context logic system with partial truth assignments is equivalent to one with total truth assignments. For this reason this aspect has been here disregarded.

the set of logical connectives $\{\neg, \wedge, \vee, \rightarrow\}$; the set of propositional constants $\mathbb{P}$; and the set of modal operators $\{\Box_i\}_{i\in L}$ where L is the set of indexes denoting levels of abstractness, and $||L|| = n$, that is to say, there are as many modal operators as levels of abstractness. The set of well formed formulas $\mathbb{F}$ is then defined as follows:

$$\mathbb{F} := \mathbb{P} \cup (\neg\mathbb{F}) \cup (\mathbb{F} \wedge \mathbb{F}) \cup (\mathbb{F} \vee \mathbb{F}) \cup (\mathbb{F} \rightarrow \mathbb{F}) \cup (\Box_i\mathbb{F}).$$

By means of this language it is possible to express statements about what holds on a level (in a context) via modal formulas.

Semantics. As a semantics for this system we can use very simple models $M = (W, L, <, c, v)$ such that for every level of abstractness (or context) $i \in L$ function c associates a non-empty subset of W $(c : L \longrightarrow Pow^+(W))$, v is the usual valuation function assigning truth values to propositions in worlds. Ordering $< \subseteq L \times L$ is an irreflexive, asymmetric and transitive ordering on L, the intuitive reading of which is: $i < j$ means that i is less abstract than j (feature 1). Using these models we can define the semantics of the levels of abstractness as follows:

$$M, w \vDash \Box_i A \textbf{ iff } \forall w' \in c(i) : M, w' \vDash A$$

We omit here the obvious clauses for satisfaction of propositional formulas. Notice that the truth value of $\Box_i A$ does not depend on the world where it is evaluated. This reflects the intuition that whether A is true at level i does not depend on the place from which you evaluate it. It only depends on the truth of A in that specific level (in this precisely consists the aforementioned flatness property corresponding to feature 3). With respect to the other requirements, we have that: feature 2) is guaranteed by the fact that c delivers non-empty subsets of W, and feature 4) is guaranteed by the fact that there can be worlds not belonging to any $c(i)$[8]. Noticeably, this semantics implements in a straightforward way the thesis developed in context modeling according to which contexts can be soundly represented as sets of possible worlds ([27]).

A final aspect worth stressing is that the ordering of the levels does not play any role in the semantics. One could imagine that the ordering on L imposes an ordering on the sets W_i. E.g. $i < j \Rightarrow W_i \subseteq W_j$. This would imply the following validity: $\Box_j A \rightarrow \Box_i A$ iff $i < j$ i.e. a kind of inheritance from more abstract levels to more concrete levels. We have chosen not to include this property because it would impose many restrictions on the relation between levels, which are not really necessary. We will come back to this point later on in Section 5 where we will indicate some ideas about more subtle relations between levels of abstractness.

[8] It is instructive to notice that this semantics is equivalent with a more standard relational semantics for $KD45_n^{i-j}$ given in terms of Kripke models with a family of accessibility relations $\{R_i\}_{i\in L}$ which are serial, transitive, and i-j euclidean $(wR_iw', wR_jw'' \Rightarrow w'R_iw'')$. The proof can be obtained once the family $\{R_i\}_{i\in L}$ is defined to be such that wR_iw' iff $w' \in c(i)$. The whole proof is worked out in [15].

Axiomatization. $KD45_n^{i-j}$ is obtainable via the following axioms and rules schemas:

$$
\begin{array}{rl}
(P) & \text{all tautologies of propositional calculus} \\
(K) & \Box_i(A \rightarrow B) \rightarrow (\Box_i A \rightarrow \Box_i B) \\
(D) & \neg\Box_i\bot \\
(4^{i-j}) & \Box_i A \rightarrow \Box_j \Box_i A \\
(5^{i-j}) & \neg\Box_i A \rightarrow \Box_j \neg\Box_i A \\
(MP) & A,\ A \rightarrow B\ /\ B \\
(N) & A\ /\ \Box_i A
\end{array}
$$

The system at issue is then a multi modal homogeneous $KD45$ with the two interaction axioms 4^{i-j} and 5^{i-j}[9]. This axiomatization is sound and complete with respect to the semantics presented (see [15]).

4.2 A logic for translation rules

Informally, A *counts as* B iff A at a level i determines the truth of B at a level j, where $i < j$ (see Section 3.2).

Theoretically, our proposal consists in understanding translation rules as *bridge rules* in the sense of theory of contexts (see for example [18]). Translation rules connect truth among different levels of abstractness, and more precisely from more concrete to more abstract levels. In addition, we consider translation rules to be defeasible. The reason for this choice is that different translation rules could have contradictory consequents, and therefore the antecedent of a translation rule cannot be strenghtened: "signing form 32 counts as consenting for organ donation" but "signing form 32 while being under threat does not count as consenting for organ donation".

To model this notion of translation rule we make use of normal prioritized default logic ([2]) defining a normal prioritized default theory $\mathcal{T}_T$ on the system $KD45_n^{i-j}$ for language $\mathcal{L}^L$:

$$\mathcal{T} = (F, D_T, \prec_T)$$

where F is a (possibly empty) set of assumptions, D_T is a set of defaults, $\prec_T$ is a priority ordering on defaults of D_T. By means of this logical machinery the following definition of translation rule can be stated:

Definition 1. (Translation rules)
A translation rule is a default rule of this form:

$$\Box_i A \ \rightsquigarrow\ \Box_j B \quad \textit{with} \quad i < j.$$

[9] Instead of 4^{i-j}, it would be sufficient to assume a simple 4 axiom: $\Box_i A \rightarrow \Box_i \Box_i A$ (see [15]).

Here "$\Box_i A \leadsto \Box_j B$" is a shorthand for $\Box_i A : \Box_j B / \Box_j B$, i.e. a normal default, the meaning of which is that the truth of B can be derived on level j from the truth of A at level i if the truth of B on level j does not result in an inconsistency. This account has several advantages: it has a clear theoretical grounding on context theory; it has a neat semantics; it enables easy non monotonic derivations; it can rely on a broadly investigated logic. Thus, the fact that "signing form 32" is a way of "consenting for organ donation" in a certain hospital can now be formally represented as:

$$\Box_i \; signing\_form\_32 \leadsto \Box_j \; consent$$

where i is a more concrete level of abstraction within the institution of "hospital" than j.

In order to deal successfully with defeasibility we also introduced in definition 1 explicit prioratization ordering $\prec_T$ on the set of defaults:

$$\begin{aligned} &d_1 : \; \Box_i A \leadsto \Box_j B \\ &d_2 : \; \Box_i (A \wedge C) \leadsto \Box_j \neg B \end{aligned}$$

One prioratization criterion is that more specific defaults have the precedence according to a strict partial ordering. So, this means $d_2 \prec_T d_1$.

Note that this prioritization orders only conflicting defaults such that either the prerequisites of the first imply the prerequisites of the second or vice versa. It does not supply a tool for deciding among conflicting defaults the prerequisites of which are logically unrelated. It may be useful, for example, to include a prioritization based on concreteness of the antecedent. This can be used in the following case:

$$\begin{aligned} &d_1 : \; \Box_i A \leadsto \Box_j B \\ &d_2 : \; \Box_k A \leadsto \Box_j \neg B \\ &k < i \end{aligned}$$

obtasining that $d_2 \prec d_1$.

We deem important to stress that specificity and concreteness are only two of the many ways of deciding about conflicting defaults. In normative reasoning especially, conflicts are often decided on the basis of authority hierarchies subsisting on norms, or on the basis of the time of their enactment ([21]). Moreover, conflicts between priority ordering themselves can arise. The specificity and concreteness criteria should therefore only be seen as an exemplification of this range of possible criteria.

4.3 A logic for normative rules

Having defined levels of abstractness and their relations in the previous sections, we now turn to defining the norms themselves that operate on levels. To do this, we have to: first, enable a representation of deontic notions within the framework defined in Section 4.1; then, introduce suitable rules to model the conditional aspect of norms, which has been stressed in Section 2.

Let us focus on the first point. To handle deontic notions (obligation, permission, prohibition), the standard deontic logic system KD (see [28]) suffices our needs here. We can therefore define a fusion[10] $KD \otimes KD45_n^{i-j}$ on a common language $\mathcal{L}^{LO}$ containing the language for expressing the abstractness layering $\mathcal{L}^L$, and the language of standard deontic logic $\mathcal{L}^O$.

Language The language is a propositional logic language the alphabet of which is expanded with an O-operator and a set of indexed $\Box_i$-operators. The set of well found formulas $\mathbb{F}$ is defined as follows:

$$\mathbb{F} := \mathbb{P} \cup (\neg\mathbb{F}) \cup (\mathbb{F} \wedge \mathbb{F}) \cup (\mathbb{F} \vee \mathbb{F}) \cup (\mathbb{F} \to \mathbb{F}) \cup (\Box_i \mathbb{F}) \cup (\Box_i O(\mathbb{F}))$$

Note that we allow deontic modalities to operate only within $\Box_k$-formulas and we do not allow deontic operators to have $\Box_k$ formulas in their scope if they are not under the scope of another $\Box_k$-operator. This expressive limitation is dictated by the fact that we do not want deontic operators to occur if not in the scope of a $\Box_k$-operator. This to capture the idea according to which normative consequences of certain conditions are supposed to be always holding at certain levels of abstractness: normative consequences are always localized.

Semantics Semantics for $\mathcal{L}^{LO}$ is given on structures $M = (W, L, <, c, R, v)$ such that $(W, L, <, c, v)$ is a model for $\mathcal{L}^L$ (see Section 4.1), and (W, R, v) is a model for $\mathcal{L}^O$ with R being a serial accessibility relation on W. We omit here the obvious clauses for satisfaction of propositional formulas. The semantics of $\Box_k$-operators remains the same described in Section 4.1. As to the semantics for the O-operator we use the usual clause obtaining the following expanded clause for formulas in $\Box_i O(\mathbb{F})$:

$$M, w \models \Box_i O(A) \textbf{ iff } \forall w' \in c(i), \forall w'' \in W : R(w', w'') \Rightarrow M, w'' \models A$$

Permission (P-operator) and prohibition (F-operator) can be defined in terms of obligation: $P(A) \equiv \neg O(\neg A)$ and $F(A) \equiv O(\neg A)$.

Axiomatization Logic $KD \otimes KD45_n^{i-j}$ can be easily axiomatized by the union of the set of axioms for $KD45_n^{i-j}$ and the set of axioms for KD. Axiom-

[10] For a detailed exposition of the concept of fusion we refer to [9]. Intuitively, a fusion of two logics is the simple join of them.

atization $KD45_n^{i-j}$ (Section 4.1) should thus be extended as follows:

$$\begin{array}{rl}
(P) & \text{all tautologies of propositional calculus}\\
(K_\Box) & \Box_i(A \to B) \to (\Box_i A \to \Box_i B)\\
(D_\Box) & \neg\Box_i\bot\\
(4_\Box^{i-j}) & \Box_i A \to \Box_j\Box_i A\\
(5_\Box^{i-j}) & \neg\Box_i A \to \Box_j\neg\Box_i A\\
(MP) & A,\ A \to B\ /\ B\\
(N_\Box) & A\ /\ \Box_i A\\
(K_O) & O(A \to B) \to (OA \to OB)\\
(D_O) & \neg O\bot\\
(N_O) & A\ /\ OA
\end{array}$$

Notice that no interaction axioms between $\Box_i$ and O operators are stated. As proved in [9], fusions of systems preserve soundness and completeness, therefore system $KD \otimes KD45_n^{i-j}$ is sound and complete with respect to the semantics presented.

To enable a representation of the aspect of conditionality of norms, and then of *normative rules*, we make again use of normal prioritized default logic defining a normal prioritized default theory $\mathcal{T}_N$ on the system $KD \otimes KD45_n^{i-j}$ for language $\mathcal{L}^{LO}$:

$$\mathcal{T}_N = (F, D_N, \prec_N)$$

where F is a (possibly empty) set of assumptions, D_N is a set of defaults, $\prec_N$ is a priority ordering on defaults of D_N. By means of this logical machinery the following definition of normative rules can be stated:

Definition 2. (Normative rules)
A normative rule is a default rule of the form:

$$\Box_i A \ \rightsquigarrow\ \Box_j OB \quad \textit{with} \quad i < j.$$

Here "$\Box_i A \rightsquigarrow \Box_j OB$" is a shorthand for $\Box_i A : \Box_j OB / \Box_j OB$, i.e. a normal default, the meaning of which is that the truth of OB can be derived on level j from the truth of A at level i if the truth of OB on level j is not leading to an inconsistency.

Conditional permission and prohibition are easily defined by replacing the O-operator by the P and F operators respectively. All remarks underlined in Section 4.2 about prioritizing defaults formalizing translation rules hold also for defaults formalizing normative rules. Given the above definition we can represent the norm that consent is required in order to perform a transplantation, as follows:

$$\Box_i\ \mathit{consent} \rightsquigarrow \Box_i P\ \mathit{transplant}$$

At this point, it is worth remarking that translation rules and normative rules share the same type of defeasibility. This representational choice captures an important analogy which we deem to subsist between the two types of rules composing institutions:

- Translation rules *connect* truth on a level to truth on a more abstract level, and this connection takes place in a defeasible way.
- Normative rules *connect* truth on a level to ideality on another, possibly the same, level, and also this connection takes place defeasibly.

That connection is what they share and what we represented here by means of normal defaults [11].

Within this framework, definitions of abstract and concrete normative rules, representing respectively abstract and concrete norms, can be also stated:

Definition 3. (Concrete Normative Rules)
A concrete normative rule is a default $\Box_i A \rightsquigarrow \Box_j O\ (B)$ *s.t. there is no default* $\Box_h C \rightsquigarrow \Box_k D$ *with* $h < k$ *s.t.* $A \equiv D$ *and* $i = k$ *or* $B \equiv D$ *and* $j = k$*.*

Definition 4. (Abstract Normative Rules)
An abstract normative rule is a normative rule which is not concrete.

In the next section we put this articulate framework at work, providing the reader with an example.

4.4 An example

The example we are going to model is chosen again from [8, 16].

Example 1. (Personal data in severe criminality registers)
Part of the abstract norm "*the inclusion of personal data in a severe criminality register occurs only when it concerns: a) suspect of crimes; b) etc.*" can be modeled as follows:

$$\Box_a(personal(datum) \wedge suspect(datum)) \rightsquigarrow \Box_c P\ include(datum)$$

Part of the concrete norm "*personal data are financial and corporate data; data concerning nationality; etc.*" might be represented as follows:

$$\Box_c(nationality(datum) \wedge suspect(datum)) \rightsquigarrow \Box_c P\ include(datum)$$

The translation rule "*personal data are financial and corporate data; data concerning nationality; etc.*" is representable as follows:

$$\Box_c nationality(datum) \rightsquigarrow \Box_a\ personal(datum)$$

where $c < a$.

[11] In this respect, our approach is close to the proposal in [11], though we carried it out by means of different formal tools.

The first norm is more abstract because it operates between level a and level c. The second one is instead more concrete. The connection among the two of them is expressed by the translation rule connecting c to a with respect to the states of affairs $nationality(datum)$ and $personal(datum)$[12]. It may be worth noticing a reasoning pattern straightforwardly available on the basis of this representation: assuming $\Box_c(nationality(datum) \wedge suspect(datum))$, by means of default $\Box_c nationality(datum) \leadsto \Box_a personal(datum)$ and validities for $\Box$, we can infer $\Box_a(personal(datum) \wedge suspect(datum))$; we can then infer the normative consequence $\Box_c P\ include(datum)$ by means of default $\Box_a(personal(datum) \wedge suspect(datum)) \leadsto \Box_c P\ include(datum)$[13].

5 Institutions Defined Formally

On the basis of the formal analysis just presented we are now in a position to provide a formal definition of the concept of institution in terms of default theories. However, before getting to this, a related issue should be considered, that is: how to rigorously relate institutions and levels of abstractness. In other words, at what level of abstractness does the institution end? If one includes only the levels explicitly specified for the institution, then the norms possibly coming from more abstract levels would not come to belong to the institutional theory. I.e. if $i < j$ and j is a level that does not belong to the institution then the norms operating on level j also are not "inherited" by the institution. On the other hand, incorporating all levels of abstractness connected to the levels explicitly defined within the institution would include the complete layering in which the institution is merged.

We therefore choose to propose two definitions, one corresponding to an "explicit" view on institutional theories and one corresponding to the "implicit" one.

Let us consider the default theory $\mathcal{T} = (F, D_N \cup D_T, \prec_N \cup \prec_T)$, i.e., a default theory for both translation and normative rules, and let L be the set of abstractness levels and $<$ their ordering. Let then L_I be the set of levels of abstractness on which institution I works. Let then $<_{L_I}$ be the sub-ordering of $<$ on L_I. The following definitions can be stated.

Definition 5. (Explicit Institutional Theories)
An explicit institutional theory I^{expl} is defined as a triple $(N_I, T_I, \prec_I)$ where:

$$N_I \equiv \bigcup_{i \in L_I} N_i$$

with $N_i \equiv \{\Box_i A \leadsto \Box_j O\ B \mid \Box_i A \leadsto \Box_j O\ B\ \in D_N\ \ \&\ \ j \in L_I\}$. *And where:*

$$T_I \equiv \bigcup_{i \in L_I} T_i$$

[12] Notice that we presupposed the state of affairs $include(datum)$ to be a concrete one.
[13] Notice that this argument is nothing but a normal defaults proof.

with $T_i \equiv \{\Box_i A \rightsquigarrow \Box_j B \mid \Box_i A \rightsquigarrow \Box_j B \;\in D_T \;\;\&\;\; j \in L_I\}$*. The third element of the triple consists in the prioritization ordering* $\prec_I \subseteq \prec_N \cup \prec_T$ *on defaults in* N_I *and* T_I*.*

Intuitively, an institution is described as the set of all normative and translation rules defined between the levels explicitly belonging to that institution.

Definition 6. (Implicit Institutional Theories)
An implicit institutional theory I^{impl} *is defined as a triple* $(N*_I, C*_I, \prec *_I)$ *where:*

$$N*_I \equiv N_I \cup \bigcup_{k \in L} N_k$$

with $N_k \equiv \{\Box_k A \rightsquigarrow \Box_l O\ B \mid \Box_k A \rightsquigarrow \Box_l O\ B \;\in D_N \;\;\&\;\; \exists j \in L_I, j < k\}$*. And where:*

$$T*_I \equiv T_I \cup \bigcup_{k \in L} T_k$$

with $N_k \equiv \{\Box_k A \rightsquigarrow \Box_l B \mid \Box_k A \rightsquigarrow \Box_l B \;\in T_N \;\;\&\;\; \exists j \in L_I, j < k\}$*. The third element of the triple consists in the prioritization ordering* $\prec *_I \subseteq \prec_N \cup \prec_T$ *on defaults in* $N*_I$ *and* $T*_I$*.*

Intuitively, an implicit theory of an institution I is nothing but a sort of closure of the explicit theory I^{expl} of I along the abstractness ordering $<$, leading the explicit theory to incorporate every normative and translation rules defined between more abstract levels than the levels explicitly belonging to I. From definitions 5 and 6 obviously follows that: $N_I \subseteq N*_I$ and $T_I \subseteq T*_I$. Let us consider now a simple example excerpted again from [26].

Example 2. (Rules inheritance within institutions)
In order to extract an organ from a living donor each hospital in Spain ought to ascertain the legal age of the donor. The state of affairs $legal\_age$ is not a concrete one; let the level of abstractness it holds on to be s_3. The institution "hospital in Spain" I_S inherits a rule from Spanish general law according to which $legal\_age$ supervenes on $being\_eighteen\_years\_old$. Neither this last state of affairs can be properly seen as concrete; let its level be s_2. Then the institution "Valencia hospital" I_V contains another rule according to which $being\_eighteen\_years\_old$ supervenes on $ID\_testifies\_legal\_age$. This can be deemed as concrete; let its level be s_1. We then have three ordered levels and two institutions constituted by rules operating on those levels. One institution is general, namely I_S, and it works between levels s_1, s_2 and s_3, the other one, namely I_V, is more particular and it operates between s_1 and s_2.
Theory I_S^{expl} would be a triple $(N_S, T_S, \prec_S)$ such that:

$$\Box_{s_1} extract \rightsquigarrow \Box_{s_2} O\ (being\_eighteen\_years\_old) \in N_S,$$
$$\Box_{s_2} being\_eighteen\_years\_old \rightsquigarrow \Box_{s_3} legal\_age \in T_S$$

Theory I_V^{expl} would instead be a triple $(N_V, T_V, \prec_V)$ such that, basically:

$$\Box_{s_1}\ (ID\_testifies\_legal\_age) \rightsquigarrow \Box_{s_2}(being\_eighteen\_years\_old) \in T_V.$$

To understand the sense of this rule in the context of I_V it is necessary to consider the explicit account I_V^{impl} of this institution: $(N{*}_V, C{*}_V, \prec {*}_V)$. We then obtain what follows:

$$\Box_{s_1} extract \rightsquigarrow \Box_{s_2} O\ (being\_eighteen\_years\_old) \in N{*}_V,$$
$$\Box_{s_2} being\_eighteen\_years\_old \rightsquigarrow \Box_{s_3} legal\_age \in T{*}_V$$

This means that I_V^{impl} and I_S^{expl} share something: in this case $N *_V \cap N_S \neq \emptyset$ and $T *_V \cap T_S \neq \emptyset$. This exactly shows how I_V inherits rules from I_S, and more noticeably how I_V concretizes norms belonging to I_S by means of translation rules.

6 Reasoning with Institutional Theories

In this section we show how our formal approach to institutions, that led to Definitions 5 and 6, can be straightforwardly merged in formal argumentation frameworks specifically developed to account for legal reasoning, such as [24, 20, 22]. This will display some guidelines on how to enable articulate reasoning patterns within our approach.

Logical systems for argumentation formalize "a particular group of patterns of inferences, namely those where arguments for and against a certain claim are produced and evaluated, to test the tenability of the claim" ([23]). In [24] an argumentation framework is presented, which is based on normal default logic and which accounts for reasoning with both what we called, in Section 2, regulative and constitutive norms of normative systems. Within this setting, the central concept on which the argumentation system is based is the concept of *deontic context*, that is, a set of facts on which the set of default rules can be applied inferring the relevant normative consequences to that set of facts. In that work, anyway, no attention is given to the issue of abstractness and concreteness of norms, and consequently the logic on which default theories are built upon is just a standard deontic logic system KD. Defaults are therefore rules of this type: $A \rightsquigarrow B$ and $A \rightsquigarrow O\ B$. If we assume the multi-modal system exposed in Section 4.3 as the logic on which to apply normal defaults, and recalling Definitions 5 and 6, this useful notion can be adapted to our approach and modified as follows.

Definition 7. (Institutional Contexts)
An explicit institutional context $\mathcal{I}^{expl} = (F, I^{expl})$ consists of a set F of propositional sentences on a language $\mathcal{L}^{LO}$, and an explicit institutional theory I^{expl}.
An implicit institutional context $\mathcal{I}^{impl} = (F, I^{impl})$ consists of a set F of propositional sentences on a language $\mathcal{L}^{LO}$, and an implicit institutional theory I^{impl}.

By means of these notions of institutional contexts, scenarios in which an institution I is made operative on the set of facts F can be formalized: through the rules of which institution I consists normative consequences at different levels of abstractness can be defeasibly established from F. The whole formal argumentation machinery exposed in [24] can then be put at work on *institutional contexts* instead of on deontic contexts, thus providing definitions of the notions of: *argument*, *conflict* and *defeat relations* between arguments, and *justified*, *defensible* and *overruled* arguments[14].

Analogous observations can be carried out in relation with the argumentation framework for legal reasoning presented in [20, 22], which is also based on normal default logic and therefore, in principle, perfectly suitable to handle our notion of institutional theory.

7 Conclusions and Future Work

In this work we discussed the problem of incorporating abstract norms into institutions that regulate the interactions between agents. We have shown by means of several examples that the level of abstraction of the norms is different from that of the procedures operating the institution. For this reason it does not suffice to just formalize the norms and procedures and then validate or verify the procedures against the norms. We therefore proposed to use explicit translation rules (formalized by normal defaults), corresponding to the so-called constitutive rules in legal and social theory, to formally characterize this translation. In order to capture the idea of a translation from the abstract level to the concrete level we chose to represent those levels explicitly, modeling them as contexts. Translation rules played then a kind of bridging role between levels/contexts.

Two research lines are particularly worth investigating in order to further develop the results presented here. First, as underlined in Section 2.2, an adequate understanding of the relation of implementation of a set of norms via a set of procedures deserves an accurate analysis in order to fully understand how norms are translated to an operational dimension, and therefore how institutions are instantiated by specific organizations. Secondly, although the logical formalism proposed gives the tools to describe the relations between norms on different abstraction levels, it does not in itself account for the restrictions which apply to this relation. As already noticed in Section 3.1, "wearing a red hat" is probably not acceptable as an implementation of "consenting for organ donation", or analogously the "daily temperature" can not count as a "personal datum". We intend to use formal ontological descriptions to account for this kind of restrictions constraining translation rules.

References

1. C. E. Alchourrón and E. Bulygin. *Normative Systems.* Springer Verlag, Wien, 1986.
2. G. Antoniou. *Nonmonotonic Reasoning.* MIT Press, Cambridge, 1997.
3. S. Buvač, S. V. Buvač, and I. A. Mason. The semantics of propositional contexts. *Proceedings of the eight ISMIS. LNAI-869*, pages 468–477, 1994.

[14] For an exhaustive account of the role of these concepts in argumentation logics we refer to [23].

4. S. V. Buvač and I. A. Mason. Propositional logic of context. *Proceedings AAAI'93*, pages 412–419, 1993.
5. F. de Jonge, W. Wiegerinck, E. Akay, M. Nijman, J. Neijt, and A. van Beek. Proanita: A multi-agent solution for legitimate information retrieval. In *Proceedings BNAIC'03, Nijmegen*, pages 453–454, 2003.
6. F. Dignum. Agents, markets, institutions, and protocols. In *Agent Mediated Electronic Commerce, The European AgentLink Perspective.*, pages 98–114. Springer-Verlag, 2001.
7. F. Dignum. Abstract norms and electronic institutions. In *Proceedings of the International Workshop on Regulated Agent-Based Social Systems: Theories and Applications (RASTA '02), Bologna*, pages 93–104, 2002.
8. Wet politieregisters (dutch police records act). Staadsblad, 1990.
9. D.M. Gabbay, A. Kurucz, F. Wolter, and M. Zakharyaschev. *Many-dimensional modal logics. Theory and applications*. Elsevier, 2003.
10. A. I. Goldman. *A Theory of Human Action*. Princeton University Press, Princeton, 1976.
11. G. Governatori, J. Gelati, A. Rotolo, and G. Sartor. Actions, institutions, powers. preliminary notes. In *International Workshop on Regulated Agent-Based Social Systems: Theories and Applications(RASTA'02)*, pages 131–147, 2002.
12. H. L. A. Hart. *The Concept of Law*. Clarendon Press, Oxford, 1961.
13. A. J. I. Jones and M. Sergot. Deontic logic in the representation of law: towards a methodology. *Artificial Intelligence and Law 1*, 1992.
14. A. J. I. Jones and M. Sergot. On the characterization of law and computer systems. *Deontic Logic in Computer Science*, pages 275–307, 1993.
15. A. Lomuscio and M. Sergot. Deontic intepreted systems. *Studia Logica 75*, pages 63–92, 2003.
16. Modelreglement register zware criminaliteit. Sttatscourant, 2002.
17. D. C. North. *Institutions, Institutional Change and Economic Performance*. Cambridge University Press, Cambridge, 1990.
18. L. Serafini P. Bouquet. wo formalizations of context: a comparison. *Modeling and Using Contexts. LNAI 2116*, pages 468–477, 2001.
19. A. Peczenik. *On Law and Reason*. Kluwer, Dordrecht, 1989.
20. H. Prakken. A logical framework for modeling legal argument. In *Proceedings of the Fourth International Conference on Artificial Intelligence and Law*. ACM Press, 1993.
21. H. Prakken. Two approaches to the formalization of defeasible deontic reasoning. *Studia Logica 57*, 1996.
22. H. Prakken. *Logical Tools for Modelling Legal Arguments*. Kluwer, 1997.
23. H. Prakken and G. Vreeswijk. Logics for defeasible argumentation. *Handbook of Philosophical Logic vol. IV*, pages 218–319, 2002.
24. L. Royakkers and F. Dignum. Defeasible reasoning with legal rules. *Defeasible Deontic Logic*, pages 263–286, 1997.
25. J. Searle. *The Construction of Social Reality*. Free Press, 1995.
26. Ley 30/1979, de 27 de octubre, sobre extracción y transplante de órganos. Boletín Oficial del Estado 266, 29th april 1986.
27. R. Stalnaker. On the representation of context. In *Journal of Logic, Language, and Information 7*, pages 3–19. Kluwer, 1993.
28. G. H. Von Wright. Deontic logic. *Mind 60*, pages 1–15, 1951.