

MOSTRO deliverable n°3- WP1
**Ontology of Organizations and
Security
(final version)**

“An Ontology of the Teleological Dimension in
Organizations”

ISTC-CNR

Contents

1	Introduction	2
2	Time	8
3	Agentivity, actions and organizations	8
3.1	Basic and complex actions	11
3.2	Collective actions	11
3.3	Interactions	13
3.4	Taxonomies of action types	14
3.5	Capabilities and dependencies	14
4	Plans	16
4.1	Delegation	18
4.2	Functions	21

5	Constraints	23
5.1	Roles	24
5.2	Roles and sub-organizations	26
5.3	Normative constraints	27
6	Design and existence of organizations	33
6.1	Creating organizations	34
6.2	The notion of environment	36
7	Conclusions	37

1 Introduction

Social organizations shape many of the most important aspects of our existence. Starting from our birth, and even sometimes from our intrauterine life, we are inserted in a complex network of rules and institutions, where a large amount of rights and obligations are laid down by organizations as – for instance – modern States. But not only these: General Registry Offices, schools, universities, firms, public offices, free associations, parties, hospitals and cemeteries are just some of the many institutions that we have to deal with during our life [Guarino, 2005].

Roughly speaking, these are the kinds of objects we want to represent in our model. Thus, at an intuitive level, we want to capture individuals like the FIAT company, the Italian State or Al Qaeda but, at least for the time being, we don’t commit to represent “emergent” entities or self-organized groups, like a group of friends meeting every thursday at Mollie’s pub. This latter kind of organizations are in a way more difficult to deal with, because many of the components that we will especially address (like roles, internal structure, rules, objectives) are mostly left implicit. In any case, there is still a lot of work to be done in order to give a good model of non emergent organizations and here we only emphasize that recently a multidisciplinary renaissance of studies that have a flavor of institutionalism have taken place:

Furthermore, the concepts of self-organization or spontaneous order are insufficient for an understanding of all institutions. Menger ([1871] 1981) himself recognized a distinction between organic (self-organizing) and pragmatic (designed) institutions. But many subsequent authors ignore the latter to concentrate

on the former. Indeed, much of the existing literature on institutions exhibits an excessive emphasis on the (albeit essential) idea of self-organization, to the detriment of other vital mechanisms of institutional emergence and sustenance. With institutions that are not self-organizing, there is a stronger dependence on other institutions that are required to enforce the internal rules. [Hodgson, 2006]

Anyway, despite their importance and the large amount of literature devoted to them, ranging from sociology to economics, organizations (as explicitly constituted entities) are very difficult to define, at least from an ontological point of view. Even if giving a syntentic answer to the question “what is an organization?” is nearly impossible, it is necessary to offer at least a general idea of what it is going to be the subject of this work. For example, if we consider a sentence like “The Fiat company has hired 300 persons”, what is the referent of “Fiat”? Are we referring to the sum of individuals that are members of the organization? Are we referring to a sort of fictitious entity that people take as objectively existent in the social realm? Is this fictitious entity something capable to act? How is it possible for a fictitious entity to act?

There are different possible ways to hint at some characterization of these complex entities. First of all, it is necessary to understand the meaning of the question “what is an organization?”. An option can be that of interpreting it as something like: “what are the key elements of an organization?”. In this direction there is a general agreement – in sociology, philosophy and computer science – that the notion of organization is always linked with the idea of a plurality of individuals that play some roles according to some rules (see [Scott, 2001], [Tollefsen, 2002a], [Dignum, 2004]). Other scholars (as [Miller, 2007, Miller, 2002]) give some more emphasis on the fact that there must be some purpose necessarily linked to organizations:

Organisations consist of an (embodied) formal structure of interlocking roles. These roles can be defined in terms of tasks, procedures (in the above sense) and conventions. Moreover, unlike social groups, organisations are individuated by the kind of activity that they undertake, and also by their characteristic ends. So we have governments, universities, business corporations, armies, and so on. Perhaps governments have as an end or goal the or-

dering and leading of societies, universities the end of discovering and disseminating knowledge, and so on. [Miller, 2007]

We hold a similar position: organizations are entities that are *designed* in order to achieve some *goals* or *ends*. According to this teleological view, organizations are the result of a *decision-making* process that is necessarily based on goals and on how these goals can be achieved, even if sometimes they are only very roughly specified, as, for example, at the beginning of the design process or when they are still in the head of some high level managers. In section 2 we will focus on the concepts needed to give this kind of characterization. The latter will be obtained by going through the concepts of action and execution of an action, those of plan, rule and role, in order to work out a formal way to express how these elements are intertwined and tailored to the pursuing of some goals.

Another way to partially answer the question “ what are organizations?” is to try and see how these entities can be classified according to a previously defined ontology or conceptualization. We made an attempt in this direction in [Bottazzi and Ferrario, 2005, Bottazzi and Ferrario, 2007], where we relied on a specific ontology: DOLCE [Masolo et al., 2002]. In the present work we won’t commit to a specific ontology and we will try to show some hypotheses of classification of this entity by using some primitive concepts as social object, artifact and agent.

The fact that organizations are social objects – something that is the result of some agreement or convention among a community of humans – is strongly pointed out in the whole literature that we have taken into consideration¹. This is a very important point in the philosophical analysis, but we will not go into details on this. But there is a side debate that has an import from the foundational point of view, that is to say the fact that social objects are seen by many as artifacts². In this perspective, also organizations can be seen as *artifacts*, whose function is to coordinate some collective behavior. If we take an artifact, for instance a chair, we can see that each part of the chair contributes to the main function of the chair, that is something to sit on. In the same way, we can imagine that every part of an organization has

¹See for instance [Bloor, 1997, Boella and van der Torre, 2006, Hindrinks, 2005, Miller, 2007, Miller, 2002, Tuomela, 1997, Tuomela, 1995, Tuomela and Balzer, 2003].

²For a deep philosophical analysis see [Kroes, 2003] and [Kroes, 2003]. In cognitive science see [Tummolini and Castelfranchi, 2006], especially devoted to institutions and organizations.

a function that contributes to the general goal of the organization itself. For the time being we take this only as a metaphor, leaving aside some more detailed ontological questions on artifacts, namely whether organizations are *literally* artifacts and to what extent they are different from material artifacts as chairs and hammers³.

Organization designers try to develop a plan for solving the need / achieving the goals of the organization. Note that ‘designed’ does not necessarily mean that there exists a single designer that defines an organization on the basis of a top-down process. Designing can be a (public) process in which a community of agents that organize themselves in order to achieve some goals is involved. In this case the ‘designer’ is a group of agents. In this sense the design can be itself a distributed and collective process and the plan for generating the plan and the plan to obtain the goals seem to be quite intertwined.

In the case of the physical artifacts the process of design is well separated from the design itself. In particular, the specific structure of the artifact is the result of a sequence of *refinement steps* made during the design process. This is because normally the artifacts do not have the capability of designing themselves, contrary to organizations. The designers of the (upper-)organizations can delegate part of the design, part of the decisional process, to the designers of the sub-organizations, i.e. the sub-organizations can have autonomy to refine the design of a part of the whole design. In this sense, the organization exists before its design is completed, or, better, the structure of the organization can evolve in time, and the designer can be the organization (or part of it). Therefore, the structure of an organization encodes not only the different roles of the sub-organizations designed for achieving the goals of the (upper-)organization, but also the responsibilities these sub-organizations have in the design of the organization itself. In a sense there can be a sort of delegation of design among organizations. If we describe the snapshot of an organization, the fact that a specific sub-plan is delegated to a sub-organization with a specific autonomy, and that this sub-organization has chosen a particular implementation of this plan, represents both a structural information and a design information.

Now let us turn to another issue. In order to achieve their goals, organizations need to act in the reality, and therefore they cannot be seen just as

³It is also possible to draw a parallel with the algebraic specification and program development, see for example [Sannella and Tarlecki, 1997], [Sannella, 2000]

abstract entities, i.e. they need to have a physical structure and they need to be embedded in the physical world. Similarly to artifacts, organizations have a double nature; they have a structural and a functional dimension. The functional dimension expresses the goals for which the organizations can be employed in an bigger organization. I.e. the functional description identifies how the goals (that can be potentially) achieved by a (sub-)organization can be employed in order to achieve the goals of the (upper-)organization.

In our approach, the plan of the global organization is decomposed in sub-plans that are delegated to different sub-organizations. The way these sub-plans are coordinated/linked together to obtain the global plan is established by the global organization via some constraints/norms that regulate the interactions between the sub-organizations. These constraints and the chosen sub-plans constitute the structure of the global-organization. Note that the interactions imposed by the global organization are the necessary ones to realize the global plan, while the sub-organizations can autonomously decide to set up additional interactions that are needed in order to realize the chosen refinement of the plan.

Another way to classify organizations is by considering them as agents.

The agentic nature of organizations is discussed in length in philosophy, especially in philosophy of sociology and in philosophy of law but also in logic and computer science. They are often seen as artificial persons:

In Law, in addition to natural persons (the human beings), there exist artificial persons. These artificial persons are collective entities that aggregate several natural persons allowing them to collectively pursue some interests. They have a real existence in human society: they have juridical personality, which means that they may be the subject of obligations and rights and they also have legal qualification, which means that they can exercise their rights and be responsible for the unfulfillment of their obligations. [Pacheco and Carmo, 2003]

Since many have pointed out that organizations act in a very peculiar way, namely through the actions of some agents who, in virtue of the roles they play, are delegated to act on their behalf, it is not very obvious that they can really be classified as agentic entities. Not only this: their actions (the actions these agents perform on their behalf) are of a particular form, that we can call “institutional”. The President doesn’t hit a piece of wood with a

stick on behalf of the organization he's president of (unless this is a symbolic gesture with some further meaning), but he can very easily sign a contract on behalf of it. In other terms, every act which is indirectly performed by an organization must be institutional.

Anyway, we don't want to enter in these details in this work and we will assume that organizations are agents in a broad sense. This choice is mainly taken for the sake of simplicity, but it is also due to the fact that in many contexts – as in some legal ones – humans tend to simplify their view on organizations by directly ascribing to them some actions [Miller, 2002, Tollefsen, 2002b].

Summing up, we consider *organizations* as *designed* agents that act in the *physical world* with a specific *mission*. As we previously said, in our view an organization has some *goals*, i.e. some desired states of the world and, in order to achieve its goals, it adopts a (set of) plan(s) corresponding to a specific strategy for achieving these goals. Referring to the classical distinction between desires and intentions, we can say that an organization *intends* to achieve the states of the world that correspond to its goals only when it has a plan for achieving these goals.

A *plan* is a description of a sequence of processes or actions to obtain a final goal (state) starting from given initial conditions and passing through intermediate sub-goals. A plan can be decomposed in sub-plans, the atomic sub-plans are basic transitions from the initial to the final state. Some sub-plans can be 'under-specified' in the sense that the kind of actions needed for a transition can only be very generally specified (a very general kind of actions) or, as limit case, completely unspecified, i.e. the organization needs to achieve the final state starting from the initial one, but it does not have any specific knowledge/commitment on the way of achieving that final state. Therefore the entire structure of the organization is more or less affected by its global plan at any level via the assignment of sub-plans to other (sub)organizations. This assignment seems a sort of *delegation* of a plan, i.e. a delegation not only of a goal but also of the way of achieving that goal, with a degree of control and power on the delegate.

Now let's see how this general ideas can be formally implemented. In the next sections we will formally introduce our model of organizations. More specifically, we will clarify the basic notions the model is based on, namely those of *agent* and *action*, those of *plan*, *role*, *constraint*, *contract* and *norm*.

2 Time

In the following, we will assume a first-order multi-modal logic with identity. Therefore *states of affairs* (also called *states of the world*) are represented by propositions (closed first-order formulas). The ‘material implication’ between two propositions represents the specialization between states of affairs, i.e. if $\phi \rightarrow \psi$ (where ϕ and ψ are propositions) then we will say that ϕ is more specific than ψ .

In order to make things simple, we take a *discrete branching time*, i.e. a partial order between *moments* of time, that we represent by the following modal operators that applies to a generic proposition:

$\Box\phi$: “ ϕ is true at all the moments”;

$\vec{\Box}\phi$: “ ϕ is true at all the moments (including the present) in the future”;

$\overleftarrow{\Box}\phi$: “ ϕ is true at all the moments in the past”;

$\diamond\phi$: “there exists at least a moment where ϕ is true”;

$\vec{\diamond}\phi$: “there exists at least a moment in the future (including the present) where ϕ is true”;

$\overleftarrow{\diamond}\phi$: “there exists at least a moment in the past where ϕ is true”;

$\bigcirc\phi$: “ ϕ is true at the next moment”;

$\overline{\bigcirc}\phi$: “ ϕ is true at the previous moment”.

This implies that we do not refer directly to moments in the theory but we only via these modal operators.

3 Agentivity, actions and organizations

We claimed that, in our approach, organizations are seen as specific kind of agents. Therefore we consider a *finite* set of agents (AG) of which organizations are a subset.

Agents can have *mental attitudes*, in particular here we assume that agents (we will see in the following how additional mental attitudes can be introduced):

- have *goals*, i.e. they pursue some aims, ends, missions; the modal operator $\mathbf{G}_a\phi$ stands for “the agent a aims at achieving the state ϕ (called the goal)”; and they
- have *beliefs*, i.e. they have an internal point of view on the world; the modal operator $\mathbf{B}_a\phi$ stands for “the agent a believes that the state/proposition ϕ obtains/is true”.

Clearly, $\mathbf{G}_a\phi \rightarrow \text{AG}a$ and $\mathbf{B}_a\phi \rightarrow \text{AG}a$. Note that we do *not* assume neither (i) $\mathbf{B}_a\phi \rightarrow \phi$, nor (ii) $\phi \rightarrow \mathbf{B}_a\phi$, i.e agents are not god and the belief operator is clearly different from the knowledge one for which at least (i) is normally accepted. In addition we refuse the *realist* property: $\mathbf{B}_a\phi \rightarrow \mathbf{G}_a\phi$ and we just assume that $\mathbf{G}_a\phi \rightarrow \mathbf{B}_a\overleftrightarrow{\phi}$.

One fundamental aspect of agents concerns the capabilities they have to act in the world to achieve their goals according to their beliefs. As stated before, we want to explore the *agentivity* organizations have and the strategies they intend to follow to achieve their goals. To take into account these aspects, the actions agents can execute to achieve a state of the world need to be considered, therefore

- we introduce a set of *actions* (ACT) and a finite set of *types of actions* (α, β, \dots) where $\alpha x \rightarrow \text{ACT}x$; and
- we extend the *achievement stit* operator (see [Belnap et al., 2001]) explicitly introducing the type of actions performed to achieve the state of affair: $\mathbf{E}_a^\alpha\phi$ stands for “the agent a has completed the execution of one action of kind α and ϕ is now (at the present moment) true”. From a semantic oriented point of view, $\mathbf{E}_a^\alpha\phi$ means: ϕ is true now because at some moment in the past, agent a chosen to execute an action of type α to achieve ϕ .

We write $\mathbf{E}_a^\alpha\top$ simply to mean “agent a executed an action of type α ” without focusing on the specific goal a had. In addition we can abstract from specific agents:

$$\mathbf{E}_A^\alpha\phi \triangleq \exists a(A(a) \wedge \mathbf{E}_a^\alpha\phi),$$

$$\mathbf{E}^\alpha\phi \triangleq \exists a(\mathbf{E}_a^\alpha\phi).$$

To be executed, actions require the truth of some *preconditions*, i.e. the state of the world can prevent or allow for the execution of an action. We characterize these preconditions (see below) by a sort of dual of the \mathbf{E} operator that represents what was true at the beginning of the action: $\bar{\mathbf{E}}_a^\alpha \phi$ stands for “when agent a started the execution of action α , ϕ was true” and its generalizations:

$$\begin{aligned}\bar{\mathbf{E}}_A^\alpha \phi &\triangleq \exists a(A(a) \wedge \bar{\mathbf{E}}_a^\alpha \phi) \wedge \forall a(A(a) \rightarrow \bar{\mathbf{E}}_a^\alpha \phi), \\ \bar{\mathbf{E}}^\alpha \phi &\triangleq \exists a(\bar{\mathbf{E}}_a^\alpha \phi) \wedge \forall a(\bar{\mathbf{E}}_a^\alpha \phi).\end{aligned}$$

Each action type α have a specific precondition π^α that can be characterized by:

$$\Box((\mathbf{E}^\alpha \top \rightarrow \bar{\mathbf{E}}^\alpha \phi) \leftrightarrow (\pi^\alpha \rightarrow \phi))$$

but some very general conditions too can be made explicit:

$$\begin{aligned}\mathbf{E}_a^\alpha \phi &\rightarrow \bar{\mathbf{E}}_a^\alpha \mathbf{G}_a \phi \\ \mathbf{E}_a^\alpha \phi &\rightarrow \bar{\mathbf{E}}_a^\alpha \mathbf{G}_a \mathbf{E}_a^\alpha \phi \\ \mathbf{E}_a^\alpha \phi &\rightarrow \bar{\mathbf{E}}_a^\alpha \mathbf{G}_a \mathbf{E}_a^\alpha \top\end{aligned}$$

Note that, given a action type α , its preconditions π^α are independent from the executor, i.e. they represent the general mechanistic behavior of the world, a sort of deterministic causation. But agents can have subjective point of views on the rules that govern the world:

$$\mathbf{B}_a \Box((\mathbf{E}^\alpha \top \rightarrow \bar{\mathbf{E}}^\alpha \phi) \leftrightarrow (\pi_a^\alpha \rightarrow \phi))$$

In the case π_a^α is equivalent to π^α their beliefs match the real world, otherwise wrong behavior can be experimented. Assuming that $\pi_a^\alpha \rightarrow \mathbf{B}_a \overleftrightarrow{\mathbf{E}}^\alpha \top$ we are sure that these wrong behaviors are ‘sincere’.

3.1 Basic and complex actions

Basic actions are the most specific actions in the model. From a temporal perspective they are simple transactions between two moments that are one the direct successor of the other. The agent(s) that executes one basic action is exactly determined as well as all the other ‘parameters’ that define what an action is (the other participants, the spatial extension, additional temporal properties, etc.). We do not discuss here which parameters determine actions (actually this is a very hard and interesting topic) but we simply introduce a finite number of *disjoint* basic action *types* (which instances are therefore atomic transitions) together with the associated preconditions.

Basic actions can be composed to form *complex* actions, e.g. sequences of basic actions or parallel basic actions. Here we do not introduce specific composition primitives, but we generally indicate that the action *type* α is a complex action coming from a structured composition of the action *types* $\alpha_1, \dots, \alpha_n$ by $\text{cmp}(\alpha|\alpha_1, \dots, \alpha_n)$. Formally this corresponds to an axiom that constraints how the actions of kind $\alpha_1, \dots, \alpha_n$ are structured, and it has implication at the level of execution:

$$\text{cmp}(\alpha|\alpha_1, \dots, \alpha_n) \wedge E^\alpha \top \rightarrow \overleftarrow{\diamond} E^{\alpha_1} \top \wedge \dots \wedge \overleftarrow{\diamond} E^{\alpha_n} \top$$

The specific way of composing the basic actions allow for reconstructing the temporal extension of the actions and their preconditions for which we simply write $\text{cmp}(\pi^\alpha|\pi_1^\alpha, \dots, \pi_n^\alpha)$.

3.2 Collective actions

Right now we discussed (from a quite general point of view) how basic types of actions can be combined independently from executors. The structure of complex actions explains how basic actions can be coordinated to achieve some goals. The coordination basically considers how the execution of some actions depends on the execution of other actions, i.e. it expresses how the results achievable by actions of a specific type can contribute to make true the preconditions necessary for the execution of actions of other types in order to contribute to the goal of the whole complex action.

But organizations coordinate agents that play roles with different competences and tasks, therefore collaborative and coordinated actions are fundamental here. In addition to the structural constrains already considered in the complex types, for this kind of actions, called *collective actions*, it is

essential to constraint who are the executors of the single components, i.e. collective actions are complex actions that are composed by basic actions executed by different agents (that interact in a specific way).

We will write $\text{cmp}(\alpha|\alpha_1:A_1, \dots, \alpha_n:A_n)$ to indicate that the component α_i needs to be executed by an agent of kind A_i . This simply means that α is a structured composition of actions $\alpha_1, \dots, \alpha_n$ with the additional constraint that these actions need to be executed by agents of a specific kind (respectively A_1, \dots, A_n). As in the case of complex actions, we can minimally characterize this constraint using the E operator:

$$\text{cmp}(\alpha|\alpha_1:A_1, \dots, \alpha_n:A_n) \wedge E^\alpha \top \rightarrow \overleftarrow{\diamond} E_{A_1}^{\alpha_1} \top \wedge \dots \wedge \overleftarrow{\diamond} E_{A_n}^{\alpha_n} \top$$

But who executes α ? If A_1, \dots, A_n have a non empty intersection, then one individual agent can perform α (assuming that there are no additional structural constraints that, for example, require that α_i and α_j are executed at the same time in different places, or that explicitly says that different agents have to be involved). But if A_1, \dots, A_n have an empty intersection then α requires the involvement of more than one individual agent. Therefore, what does it mean ‘individual agent’? Is a group an individual agent? Is an organization an individual agent? For the moment we do not want to answer to this very difficult question, we just assume that there exists an operator grouping agents in *collectives*: $a = [a_1, \dots, a_n]$ means that a is a collective whose members are a_1, \dots, a_n (all the members will be considered as different). Note that a is still an agent (AGa) and that a_i can be a collective too, i.e. it is possible to have collectives as members, for example, assuming $a_1 = [b_1, \dots, b_m]$, we have $a = [[b_1, \dots, b_m], a_2, \dots, a_n]$. This mechanism allows for taking into account multi-level agents and multi-layer actions.

Coming back to the E operator, we can introduce the following general axiom schema:

$$\text{cmp}(\alpha|\alpha_1:A_1, \dots, \alpha_n:A_n) \wedge E^\alpha \top \rightarrow \overleftarrow{\diamond} E_{A_1}^{\alpha_1} \top \wedge \dots \wedge \overleftarrow{\diamond} E_{A_n}^{\alpha_n} \top$$

and more specific ones that regards specific executors with specific goals (where $\text{cmp}(\phi|\phi_1, \dots, \phi_n)$ means that ϕ is true because of a composition, the same structure of the action, of the sub-goals achieved by means of the sub-actions):

$$\text{cmp}(\alpha|\alpha_1:A_1, \dots, \alpha_n:A_n) \wedge E_{[a_1, \dots, a_n]}^\alpha \phi \rightarrow \overleftarrow{\diamond} E_{a_1}^{\alpha_1} \phi_1 \wedge A_1(a_1) \wedge \dots \wedge \overleftarrow{\diamond} E_{a_n}^{\alpha_n} \phi_n \wedge A_n(a_n) \wedge \text{cmp}(\phi|\phi_1, \dots, \phi_n)$$

Note that we have not introduced any constraint on the kinds of agents that perform *basic* actions. Because the level of granularity of our analysis can influence the choice of which basic actions types are included in the model, to be general enough, we allow for basic collective actions, i.e. basic actions that need to be performed by a collective. As consequence collective actions cannot be always decomposed in individual actions.

In addition, note that the characterization of collective agents is only based on the their agentivity. We leave open if and how the goals and the beliefs of collective agents can be reduced to the goals and beliefs of their members. This means that the collective agent can be a simple group, which goals/beliefs are the union of the members' goals/beliefs, or something more complex where emergent goals and beliefs need to be taken into account.

3.3 Interactions

In the previous section we claimed that collective actions coordinate what different agents perform in order to achieve a global goal. We said that relevant components of collective actions are *interaction actions*, i.e. actions that regulate and establish how the autonomous members of the collective split the whole task in simpler subtasks that are assigned to the members via an interactive process.

There are at least three interaction types that can be considered as the basic ingredients of the cooperation: *delegation*, *adoption*, and *transfer*. Delegation actions refer to *requests* to an agent of something another agent needs, while adoption actions regard *acceptance of* and *commitment to* satisfy these requests. The contracts can be seen as a sort of public and transparent way of fixing some delegation/adoption interaction. In some cases, delegations and adoptions are conditioned by the availability of resources, or in a environment submitted to regulations, to rights and permissions that limit or empower agents. The transfer actions normally go together with delegation/adoption interactions and allow for establishing the migration from an agent to another one of (the control of) some resources, or the conferment of rights and permissions.

In order to better characterize these interaction types we need more notions that we will consider in the next sections. We will come back later to interaction actions.

3.4 Taxonomies of action types

It is possible to generalize basic and complex action types building taxonomies of *abstract* action types. Intuitively abstract types collect actions that belong to different basic/complex types and have something in common. They are useful for representing more general ways of achieving goals that play an important role in modeling the design process where we can start from a very vague idea of the action needed to achieve the goals (and then from a highly abstract kind), and then, step-by-step, refine it to quite specific actions. Note that at the end every action is reducible to basic actions, but the reduction can be quite complex as we will see.

One important abstraction concerns the collections of all actions that are executable by a specific kind of agents. These actions can belong to very different basic/complex types because the unitary condition is uniquely based on the kind of executors. Therefore, the taxonomy of the kinds of agents induces a taxonomy on the actions that can be executed by these agents.

We will indicate the more general kind of actions (the class that contains all the actions) with Λ and the fact that the action type α is a *specialization* of the action type β with: $\alpha \Rightarrow \beta$. This simply means that for all action types α we have $\alpha x \rightarrow \Lambda x$ and that if $\alpha \Rightarrow \beta$ then $\alpha x \rightarrow \beta x$.

3.5 Capabilities and dependencies

Acting in the world, i.e. executing some actions, agents are able to attain some state of the worlds. *Capabilities* regard which states of the world agents can *potentially* achieve i.e. their definition involves a modal notion. We distinguish what an agent can achieve performing only a specific kind of actions from what it can achieve without any limitation on the executable actions, i.e.

$$C_a^\alpha \phi \triangleq \overrightarrow{\diamond} E_a^\alpha \phi \quad \text{from} \quad C_a \phi \triangleq C_a^\Lambda \phi;$$

and, as usual, we can abstract from specific agents:⁴

$$C_A^\alpha \phi \triangleq \forall a (A(a) \rightarrow C_a^\alpha \phi) \quad \text{and} \quad C_A \phi \triangleq C_A^\Lambda \phi.$$

⁴It is possible to weaken the first definition assuming just the existence of a capable agent of kind A : $C_A^\alpha \phi \triangleq \exists a (A(a) \rightarrow C_a^\alpha \phi)$. Here we consider the stronger version.

The actions that an agent is able to perform, i.e. the executable action library of an agent, can be easily defined by means of a conjunction of statement of the kind $C_a^\alpha \top$.

In some cases it is useful to consider what agents are able to achieve acting under some initial constraints θ that represent a restriction of the preconditions of performable actions (analogously for the other kinds of capabilities above defined):

$$C_a^{\alpha|\theta} \phi \triangleq \overrightarrow{\diamond} (E_a^\alpha \phi \wedge \overline{E}_a^\alpha \theta).$$

The capabilities above introduced consider only the actions directly performed by individual or collective agents but they do not consider what an agent can achieve with the help of other agents. We call these capabilities, *collective capabilities* of an agent and they include the states of affairs that an agent is able to achieve participating to a collective action that involves other agents a_1, \dots, a_n :

$$cC_a^\alpha \phi \triangleq \exists [a, a_1, \dots, a_n] (C_{[a, a_1, \dots, a_n]}^\alpha \phi);$$

$$cC_a \phi \triangleq \exists [a, a_1, \dots, a_n] (C_{[a, a_1, \dots, a_n]} \phi).$$

Note that collective capabilities often highlight dependencies of an agent on other agents to achieve a goal and therefore they are often the good motivations for interacting with other agents via delegation and transfer. If an agent want to achieve a goal ϕ but it is not individually able to achieve ϕ , then it needs to ask the help of other agents to satisfy its goal, in this sense it is dependent (we indicate with AG the set of all agents):

$$D_{\{a|a_1, \dots, a_n\}} \phi \triangleq G_a \phi \wedge \neg C_{AG \setminus \{a_1, \dots, a_n\}} \phi.$$

In some cases the agent is able to achieve its goal performing some actions that are not permitted. In this case it depends on other agent for achieving a goal executing some permitted action α :

$$D_{\{a|a_1, \dots, a_n\}}^\alpha \phi \triangleq G_a \phi \wedge \neg C_{AG \setminus \{a_1, \dots, a_n\}}^\alpha \phi.$$

Similar definitions can be introduced for capabilities of agent types and for capabilities under constraints.

Note that this notion of dependence is quite weak. In particular from $D_{\{a|a_1, \dots, a_n\}} \phi$ it does not follow that $C_{[a, a_1, \dots, a_n]} \phi$. This means that an agent a can be dependent on an agent b because without the help of b it is impossible for a to obtain ϕ , but b it is not enough to achieve ϕ and an additional help is required.

4 Plans

Agents have beliefs about how they can act in order to achieve their goals under given constraints. “[T]he purpose of planning is to find which actions to apply to which states in order to achieve some objective when starting from some given situation. A *plan* is a structure that gives the appropriate actions.” ([Ghallab et al., 2004], p.6)

At the beginning of the analysis, the plan an agent has can be *incomplete*, i.e. the agent can have only a very generic idea of *how* its goal can be obtained. Deeper analyses can step-by-step make precise the plan until a complete description of the needed complex action is obtained. This analysis needs to take into account the restrictive constraints the agent is subject to, therefore its goal includes the satisfaction of these constraints, i.e. it is a sort of *extended* goal composed by both what the agent really wants and the satisfaction of the constraints.

During the planning process, the agent can follow alternative ways for achieving the goal and at each planning step it needs to chose which alternative solution it prefers. Therefore the planning is a *decision-making* process based on constraints but also on a ‘cost-benefit analysis’ (preference analysis) that we do not consider here.

If the agent can reach its goal independently, then it can build a plan that involves only actions performed by itself. But if it is dependent on some others’ action, it needs to delegate some actions and therefore, in its plan, it needs to consider other performers. Note that even in the former case (independently reachable), it can prefer to delegate to other agents.

Let us suppose that a collective or individual agent a has a goal σ and that it starts to develop a plan to achieve σ from a situation where the constraints θ hold.

First of all, we assume that the searching for a plan presupposes the a ’s belief that σ is achievable under the given constraints θ , i.e. $B_a \overrightarrow{\varphi}(\theta \wedge \overrightarrow{\varphi} \sigma)$, or, more specifically, a believes that there exists a way of achieving σ individually or collectively: $B_a C_a^{\Lambda|\theta} \sigma \vee B_a c C_a^{\Lambda|\theta} \sigma$. Or, more generally, a believes that is not impossible to individually or collectively achieve σ : $\neg B_a \neg C_a^{\Lambda|\theta} \sigma \vee \neg B_a \neg c C_a^{\Lambda|\theta} \sigma$.

Secondly, as stated before, a can refine its beliefs about the way of achieving σ , i.e. it can make more precise the sketchy and vague idea it had at the beginning. This means that the planning is a dynamic activity, and that

$\diamond(\mathbb{B}_a\mathbb{C}_a^{\alpha|\theta}\sigma \wedge \bigcirc\mathbb{B}_a\mathbb{C}_a^{\beta|\theta}\sigma \wedge \beta \Rightarrow \alpha)$. Similarly in the case of collective plans. In addition to refinement, other kinds of processes can be considered, as for example revision. Here we consider only refinement that, as we will see, is central in the context of organizations.

Thirdly, a can believe that there are alternative ways of achieving its goal, for example: $\mathbb{B}_a\mathbb{C}_a^{\alpha|\theta}\sigma \wedge \mathbb{B}_a\mathbb{C}_a^{\beta|\theta}\sigma \wedge \mathbb{B}_a(\alpha \not\Rightarrow \beta \wedge \beta \not\Rightarrow \alpha)$ or $\mathbb{B}_a\mathbb{C}_a^{\alpha|\theta}\sigma \wedge \mathbb{B}_a\mathbb{c}\mathbb{C}_a^{\beta|\theta}\sigma$. Therefore, a needs to choose which way it *intends* to follow to achieve σ , i.e. it needs to commit to a specific strategy among the possible ones. For example, if a *intends* to achieve σ by executing an action of type α instead than another action of type β we will write $\mathbb{I}_a^{\alpha|\theta}\sigma$, for which we assume the following basic axiom:

$$\mathbb{I}_a^{\alpha|\theta}\sigma \rightarrow \mathbb{G}_a\sigma \wedge \mathbb{B}_a(\mathbb{C}_a^{\alpha|\theta}\sigma \vee \mathbb{c}\mathbb{C}_a^{\alpha|\theta}\sigma) \wedge \alpha \neq \Lambda;$$

in this case, we will say that α is a 's *plan* to achieve σ given the constraints θ . Therefore, in our framework, a *plan* is an action type that an agent *intends* to execute (more correctly the agent intends to execute an instance of that action type) to achieve a goal.

Note that, for the axiom above, intentions presuppose a strategy ($\alpha \neq \Lambda$) to obtain a goal, i.e. a cannot have a totally incomplete plan, it must have at least a vague idea about the way of achieving its goal. While beliefs can be in some way ‘unjustified’, intentions require some motivated choices. But this constraint is really weak and, for example, α could also be a disjunction of two alternative action kinds that do not imply a real decision. It is possible to strengthen this constraint by requiring a *complete* plan, i.e. requiring that the action α is specified in all the details. This constraint seems too strong because, when an agent commits to a plan it leaves some successive refinements open even if it is independent w.r.t. the goal, i.e. some decisions are taken during the execution of the plan itself. In this case, in order to consider the plan as complete, we need to add this decision action as a basic action. This is an interesting move that we do not consider here.

Another interesting point regards the exact meaning of the expression $\mathbb{B}_a\mathbb{C}_a^{\alpha|\theta}\sigma$ when a is a collective agent. The notion of capability is an objective one: $\mathbb{C}_{[a_1, \dots, a_n]}^{\alpha|\theta}\sigma$ means that in principle $[a_1, \dots, a_n]$ can execute an action α . But here we are not considering which role the members a_i have in the execution of α and if all the a_i are willing to use their abilities (in executing a part of α) for the global goal σ . This discussion leads to understand how the goals and the beliefs of a collective agent relate to the goals and the beliefs

of its members. This is a very difficult topic that we do not address in this deliverable. We just assume that if a collective agent a intends to follow a plan then all the members act ‘in the same direction’, i.e. we assume that all the a_i will execute their respective sub-plans to achieve they sub-goals independently of having or not having σ as personal goal.

4.1 Delegation

As already said, *delegation* and *adoption* are two important types of interactions. Since delegation is a sort of request, it is normally motivated by the fact that the delegator a believes to depend on the delegatee b for a goal or for achieving that goal in a specific way: $\mathbf{B}_a \mathbf{D}_{\{a|b\}} \sigma$ or $\mathbf{B}_a \mathbf{D}_{\{a|b\}}^\alpha \sigma$. In other cases a simple preference is enough, i.e. a prefers to rely on b in order to obtain a specific sub-goal.

We include delegation in the basic action types, and $\mathbf{del}(b, \phi, \beta)$ identifies the action of delegating to agent b the achievement of ϕ by the execution of an action of kind β . In this sense, when a performs a delegation action ($\mathbf{E}_a^{\mathbf{del}(b, \phi, \beta)} \top$), it delegates not only a sub-goal, but also a way of achieving it. We say that a delegates a sub-plan in the sense that a intends that b executes an action of kind β to achieve ϕ . This means that the delegator a believes that it is possible that the delegatee accepts and it is able to execute the delegated sub-plan, i.e. the final goal of the delegation is the execution by agent b of action β to achieve ϕ : $\mathbf{E}_a^{\mathbf{del}(b, \phi, \beta)} \mathbf{E}_b^\beta \phi$. In this sense, we can say that b ’s execution of an action of kind β to achieve ϕ is true because of the delegation action of agent a , i.e., in some sense, a ’s delegation caused b ’s execution of action β .

But intermediate goals can also be considered, for example a goal can be exactly the implicit adoption of the sub-plan by b : $\mathbf{E}_a^{\mathbf{del}(b, \phi, \beta)} \mathbf{I}_b^\beta \phi$. Or, if we consider the additional basic adoption actions, $\mathbf{adp}(a, \phi, \beta)$ identifies the action of adoption by agent a of the execution of an action of type β to achieve ϕ , then the goal of a delegation action can be an explicit intermediate adoption action by b : $\mathbf{E}_a^{\mathbf{del}(b, \phi, \beta)} \mathbf{E}_b^{\mathbf{adp}(a, \phi, \beta)} \mathbf{E}_b^\beta \phi$, i.e. the execution of β to achieve ϕ is caused by the adoption that is in turn caused by the delegation.

Some authors ([Castelfranchi and Falcone, 1998]) consider *tacit or weak delegation*, i.e. delegation that presupposes neither an explicit request nor an adoption. Tacit delegation is based on the belief that the delegatee intends to perform the action needed by the delegator in any case: $\mathbf{B}_a \mathbf{I}_b^\beta \phi$. Note that it is not possible to model this kind of delegation just with $\mathbf{E}_a^{\mathbf{del}(b, \phi, \beta)} \mathbf{E}_b^\beta \phi$, because

$E_b^\beta \phi$ has not been caused by the delegation: it is an action independent from the delegation. While explicit delegation and explicit adoption ‘force’ the delegatee to execute a specific action, tacit delegation is just a decision of a who, for a given goal ϕ , intends to exploit what b intends to execute for its goals. Tacit delegation is then more a synchronization action, i.e., to achieve ϕ , a intends to wait for b ’s independent action.

In some very regulated contexts, some agents have the right of delegate to other agents that cannot refuse to adopt the delegation. In this case, the delegation action does not require an explicit adoption, but is still an action intended to force the execution of a delegatee’s action.

[Pacheco and Santos, 2004] enumerates different forms of delegation:

- delegation by command (where the delegating agent has authority over the delegatee and the delegation is immediately turned into a command)
- delegation by joint action (where each agent delegates “a part” of the joint action to the other participating agents)
- institutional delegation (where there’s no direct agreement between the agents but it is institutionally stated that, for instance, in the absence of an agent playing a certain role, another agent, in his/her role, is delegated to act on his/her behalf)
- informal delegation (not interesting because devoid of normative effect).

Note that it is possible to represent goal delegation and adoption (independently on the way used for the achievement of the goal) simply by $\text{del}(b, \phi, \Lambda)$ and $\text{adp}(a, \phi, \Lambda)$.

We have not considered yet the preconditions of an action type that regards which *resources* are necessary for the execution of an action. Resources can be physical but also ‘legal’ in the sense that in some cases permissions and rights are similar to physical resources in a system governed by laws. Agents (and more generally roles, as we will see) have permissions/rights and they control or own some physical resources. In some cases, when they delegate a sub-plan, the *transfer* of some resources from the delegator to the delegatee is also necessary, i.e. a delegation action is typically coupled with a transfer of means that guarantee the preconditions for the execution of the delegated action to be fulfilled. Differently from Secure Tropos [?], we prefer to use the

term *transfer* for resources and the term *delegation* for sub-plans because in general only the latter need an adoption. In addition, note that, differently from [Pacheco and Santos, 2004], delegation is not reduced to transference of obligations, permissions and the like but these transfers are just the result or are coupled with delegation and adoption actions.

The kind of transfer coupled with a delegation identifies different delegations. Delegation can either keep a shared responsibility (when obligations and permissions are transferred from an agent to another, but at the same time are kept by the former agent) or imply a complete transfer of responsibility (when the agent who is delegating abandons the obligations and permissions he/she is transferring).

What is the reason for delegating a kind of action to a specific agent? Trust. Is it possible to trust all the agents of a specific kind, i.e. is it possible to characterize the agents from a general point of view some classes of agents that I will trust just because they satisfy these constraints? Yes, I think. Therefore we need an additional mental attitude that is trust, and this attitude can be towards a specific agent (*personal trust*) or towards a class of agents A , but it is not the case that we trust A because we trust all the agents of kind A , it is the opposite: we trust an agent of kind A simply because he is of type A . In addition it is maybe also possible do introduce trust between two kinds of agents.

Is trust just a general relation between agents/kind of agents, or we need also to specify with respect what we trust? We trust somebody with respect to a specific task, i.e. with respect to a specific way of achieving a goal. Clearly we can then abstract from the way, and then trust just about a goal.

Delegated agents can decide (if they have the rights) to *refine* the sub-plan they adopt. Let us suppose that $I_a^\alpha \sigma$ and that in α we have a delegation action to agent b : $E_a^{\text{del}(b,\phi,\beta)} E_b^{\text{adp}(a,\phi,\beta)} E_b^\beta \phi$. b 's adoption requires that b satisfies the constraints, i.e. it needs to achieve ϕ with an action of kind β . But if ϕ and β are not completely specified, than b can choose a specific way of satisfying them, for example it can achieve a more specific goal by executing a more specific action: $E_b^\delta \psi \wedge (\delta \Rightarrow \beta) \wedge (\psi \rightarrow \phi)$.

In this case we can say that b 's plan of $(I_b^\delta \psi)$ is a refinement of the delegated sub-plan. This means that b has the right and the competences to understand which is the best way to take in charge the sub-plan it adopted.

This aspect is particularly important in organizations where the delegation is often motivated by a lack of competences of the delegator who prefers

the delegatee to execute a specific sub-plan because of its capabilities. The expertise of the delegatee implies an understanding of the exact actions needed to achieve the delegated goal, i.e. given the constraints it has, the delegatee is able to choose the best way to achieve the delegated goal. In some more extreme cases, the delegator gives the power to the delegatee to *build* the optimal context to execute the delegated sub-plan, i.e. the delegatee has the power of structuring a sub-organization.

4.2 Functions

The notion of *function* has been extensively investigated in the literature of design of artifacts. A complete analysis of functions and functionalities is out of the scope of this deliverable, but here we want to point out some links with this field that, at least from our viewpoint, are quite enlightening for the case of organizations.

According to Lind, “functions represent the roles the designer intended a system to have in the achievement of goals of the system(s) of which it is part.” ([Lind, 1994], p.33).

Simon [Simon, 1969] pays attention to the difference between what artifacts are capable of doing from what artifacts are expected to do within some contexts, and this latter aspect is used to define technology and its function.

Cummins [Cummins, 1975] states that the function of a part of an artifact is its capacity that causally contributes to the capacity of the whole artifact.

Houkes and Vermaas claim that “using an artifact A (...) can be described as executing a *use plan* P for achieving a goal g ; the artifact’s *function* is the role it has in this plan.” ([Houkes and Vermaas, 2006], p.205)

Taking into account these quotations, we consider functionality as a relative notion linked to the structure of an artifact: something has a function only within a context (the structure of the artifact encoded in its plan), and it can have different functions in different contexts even though its capabilities are unchanged, i.e. the same part can be reused with different roles in different contexts. Our idea is that an organizational function is what specifies how (part of) the objectives of one role or organization contributes to the objectives of another organization, i.e. the function of the objectives of a role or organization in the context of the other role organization.

The *function* of an organization o in another organization O is represented by the sub-plan (i.e. the sub-goal and the way of achieving this sub-goal) o is executing for O , i.e. *functions* are sub-plans (ways of obtaining an input-

output transition) assigned to other organizations. In this sense functions are relative to an upper-organization O . Interesting question: can the same organization have different functions in different organizations? I.e. is the function that an organization has in an upper-one part of the identity criteria of that organization or not? In addition, note that, the function defined in this way cannot be simply reduced to a teleological notion in the sense that it is not a goal-directed behavior. A goal-directed behavior is constrained by the plan, our notion of function regards more the role the behavior of a sub-organization has in the context of an upper-organization. In this sense functions (like roles, as we shall see) belongs more to the social realm than to the private/mental one.

The notion of plan considered in our framework allows us to (at least partially) characterize this relative notion of function in terms of delegation and adoption of sub-plans. Following this idea, given the plan of an agent a , we can distinguish the function of a delegatee b , i.e. the sub-plan (of a 's plan) b adopt⁵, from the function of a sub-plan a intends to delegate or to execute, i.e. the role of the sub-plan in the context of a 's plan as a whole. For example, if $I_a^\alpha \sigma$ and $E_b^{\text{adp}(a,\phi,\beta)} \top$, then the function of b in a 's plan α to obtain σ is the achievement of ϕ by means of an action of kind β , or, in other words, it is the execution of an action of kind β to achieve the sub-goal ϕ . Moreover, the function of the delegation action $E_b^{\text{del}(b,\phi,\beta)} \top$ is just the achievement of the sub-goal ϕ by the execution of an action of kind β . Therefore, once the notion of plan and interaction are available, the notion of function becomes quite trivial.

In some cases the second kind of function (the function of a sub-plan in the context of a whole plan) is described by a general expression that establishes what are the links of a sub-plan with the other parts of the whole plan, i.e. it is a general description of how the sub-plan contributes to the whole plan. In addition, in general delegates are considered as autonomous agents that have adopted a sub-plan according to a more or less constrained choice, therefore, the previous definition of function does not rule out the possibility that the same agent can have different functions with respect to the same or to different plans. For these reasons, as we have seen in the quotations, the notion of function is often linked to the one of *role*. Here we

⁵In the case of organizations (and of artifacts) we often talk about the function of an organization o in another organization o' . In this case we just intend the function of o in the plan of o' .

prefer to use the term ‘role’ to identify the basic atoms of organizations (see below).

It is also possible to talk about the functionality of actions (action types) in a plan. In the previous example, we can say that the function of the execution of an action β (in a ’s plan α to obtain σ) is the achievement of the subgoal ϕ . Note that in any case the function is relative to a plan that an agent has in order to achieve a specific goal.

Note also that in the case of delegated sub-plans, the function of the delegatee does not change if the plan is refined, because it is fixed by the delegation/adoption step and not by the specific way the delegatee chooses to ‘implement’ the delegate’s sub-plan.

5 Constraints

In the real world not every action is possible to achieve. The world of organizations is not an exception to this. Actually there are more limitations in the social world of organizations than in the physical world we are immersed into. On top of the limitations imposed by reality there are other limitations imposed by sociality. We deal with this by considering the notion of *constraint* and the notion of *executable action under constraints*. Let us assume a set of constraints represented as the specific state of world θ . These constraints limit the executable actions: an action of kind α is executable only if the precondition π^α are compatible with the constraints θ , i.e. if θ does not imply $\neg\pi^\alpha$.

This is particularly interesting if we assume a subjective view, in this case an action of type α is executable under constraints θ , if an agent believes that the preconditions π_a^α are compatible with θ .

We write $X_{\alpha/\theta}$ for “actions of kind α are executable under constraints θ ”, that can be defined as:

$$X_{\alpha/\theta} \triangleq \vec{\diamond}(\mathbf{E}^\alpha \top \wedge \bar{\mathbf{E}}^\alpha \theta).$$

There are several kinds of constraints that an organization has to deal with: constraints that come out from the physical world, constraints that come from the social world. The latter can be distinguished in internal and external, depending on the fact that they are made by the organization itself to control its elements or they come from another organization that is controlling it. We will see at the end of this section how a set of external constraints

shapes the social space where an organization is forced to move: what we called its social or institutional *environment*. There are two main elements to which organizational constraints are applied: roles and sub-organizations. Therefore, before going a bit deeper into the notion of constraint, we need to spend some words about these important notions.

5.1 Roles

Social roles can be seen as concepts used in order to abstract from specific agent and to reason in general about goals and plans. In general, when an agent starts to specify a plan it is especially concentrated on the steps (sub-plans) necessary to achieve the goal. In some cases it also takes from the beginning a clear commitment on the involvement of specific agents, i.e. it already knows that it wants to rely on some agents, but in general it is just interested in finding some agents able to execute the sub-plans it intends to delegate. In this sense it wants to restrict the delegation only to agents that have the necessary capabilities to execute the task, or it wants to be much more restrictive by delegating to agents who it completely trusts.

In any case, the general plan an agent a designs for its goal σ is similar to a structure of actions that refers more to types of agents, rather than to specific agents:

$$\mathbf{E}_{[a, A_1, \dots, A_n]}^\alpha \sigma \leftrightarrow \mathbf{E}_a^{\alpha_0} \phi \wedge \mathbf{E}_{A_1}^{\alpha_1} \phi_1 \wedge \dots \wedge \mathbf{E}_{A_n}^{\alpha_n} \phi_n \wedge \alpha = \mathbf{str}(\alpha_1, \dots, \alpha_n) \wedge (\phi_1 \wedge \dots \wedge \phi_n) \rightarrow \sigma.$$

The agent a is then defining which kinds of agents it wants to execute the actions it needs. Going back to the notion of function, we can then state that, let us say, the function of A_1 in a 's plan α to obtain σ is the achievement of ϕ_1 by means of actions of kind α_1 . We already said that in this case we can also talk about the *role* of A_1 in a 's plan α to obtain σ . But note that there are a lot of authors who use the word 'role' to indicate a more general notion that abstracts from the specific function A_1 has in a 's plan α to obtain σ and refers only to a characterization of the players of this role in terms of the capabilities and powers they have. In this case the same role, the same characterization of the players can be used to restrict the possible players in different plans, i.e. the notion of role is abstracted from the one of plan.

In [Bottazzi et al., 2006] and especially in [Masolo et al., 2004] they have already been described and analyzed at length, while in [Bottazzi and Ferrario, 2007]

one can find a general characterization of them as linked to an organization. Starting from these works, we can summarize their main general features as follows:

- a role can be played by different entities, at different times or even simultaneously;
- conversely, the same entity can play different roles, even simultaneously, so
- there's no necessary relation between a role and its player(s), an entity can change role and also play the same role more than once;
- roles are intrinsically relational, in the sense that, at a definitional level, they depend on the definition of other roles; a definition of a role cannot be given "in isolation" (let's think about the roles employer/employee, buyer/seller...);
- finally, they are linked to some specific kinds of entities that provide explicit definitions for them; in the case of organizations, we can think about these entities as norms and descriptions.

At this level there are two important aspects to consider: requirements and constraints; the main distinction is given by a temporal parameter: while requirements need to be satisfied by the agent before playing the role and they are *conditions* that must be necessarily present in order to play the role, constraints need to be satisfied during the whole period in which the agent is playing the role.

A similar definition can be found in [Demolombe and Louis, 2005], where a role is then defined in terms of the conditions that an agent must satisfy in order to be able to play it and of the (static and dynamic) norms that apply to it.

Among requirements we can have both properties and abilities. Properties can be of various kinds, like physical properties (for instance, an air company can require that, in order to play the role of hostess, a candidate must be taller than 1,70m), or certified titles (for instance, a university can require that, in order to play the role of assistant professor, a candidate must hold a PhD), etc.

Abilities are skills that an agent must possess, like for instance being able of programming in Java, etc. Often there's no sharp distinction between

properties and abilities, in the sense that in some cases what is really needed is a capability, but the requirement is expressed in terms of a property. As an example, one could presuppose that a graduate student in computer science is able to program in Java and thus the requirement is set on the title, which is something easier to verify than an ability.

In [Cavedon and Sonenberg, 1998] there's a notion of capability quite similar: "Adopting a role involves certain responsibilities. These responsibilities may be associated with a role in a number of ways, including ensuring that the necessary *capabilities* for playing the role in a team plan are provided by the agent.

As for constraints, we can distinguish between physical and behavioral constraints. Physical constraints are physical features that an agent must possess in order to play role and continue to play it. With behavioral constraints, as it is easy to infer, we mean the *behaviors* that an agent is expected to display when playing a certain role.

5.2 Roles and sub-organizations

Roles and sub-organizations share many features and in some sense they seem to be different just in their degree of complexity, being suborganizations a sort of complex roles. Roles (and the other organizations too) are 'ideal and designed' agents, i.e. agents that satisfy all the imposed constraints, in some sense they are 'fictional entities' that act as they are supposed to act and therefore they are needed in order to specify the structure of an organization and to simulate how the organization works. This is quite similar to the design of an artifact where we always consider the parts of the artifact as functioning. Real agents (persons for example) playing a specific role may be malfunctioning, i.e. their behaviors can be different from what expected as specified in the role, and roles have specific conditions in order to limit this malfunctioning, exactly as the specification of the parts of an artifacts needs to satisfy some requirements that are intended to guarantee a certain level of quality, accuracy, performance, etc.

Aside of these similarities there are some differences among roles and sub-organizations. This is a very hard topic that needs further investigation. For the time being we list some working hypothesis:

- The players of a role are individual specific agents (a human or maybe some kind of "intelligent" device) or collective (another organization),

while the players of a sub-organization have the form $[a_1, \dots, a_n]$ where a_i can be a collective or individual agent.

- The role does not delegate to its player, i.e. the *plays* relation is not a delegation, is something different, something that is related from a metaphysical point of view to a classification process and from a social one to the notion of responsibility (there is no responsibility).
- An hypothesis could be that the difference between roles and sub-organization is just a matter of power of the delegator on the delegates: a role (or the player of the role?) who delegates to another role has less power on the delegatee than that of an organization that delegates has to its sub-organization or role.
- During the design process what at the beginning was considered as a role can then be transformed in a sub-organization, i.e. we can add some structure and the players have then the form $[a_1, \dots, a_n]$: that is to say that the structural difference among organizations and roles emerges in the process of planning, by recognizing that the means to reach that end need are more complex than what was thought at the beginning.
- It seems that sub-organizations can play roles in other organizations but they cannot be played by agents; on the other hand, roles can be played by agents, but they cannot play roles. An hypothesis could be that sub-organizations are instances, i.e. they are specific agents, while roles are classes, i.e. they need to be instantiated by agents. If we do not accept this difference and we think that ‘the professor of University of Trento’ is a role and at the same time an agent, then, at this point, the difference with respect to a sub-organization can be established only on the basis of the *plays* relation: both roles and organizations are agents, but only roles can be played by other agents. But again if we think that both organizations and roles can be played by agents, then it is difficult to keep such a distinction.

5.3 Normative constraints

In order to pursue its goals, an organization tries to enforce the behavior of its agents by giving to them some normative constraints. Normative con-

straints are more specific than constraints in general. They are normative in the sense that there is an irreducible deontic dimension linked to them [von Wright, 1963]. The deontic world is the world of ought [Searle, 1995a]. It is the world of permissions and obligations with respect to actions undertaken by agents and there is a large amount of studies devoted to this kind of topic, going back at least to the last half of the past century. Here we want to try to sketch some ideas on norms coming from the study of organizations partially influenced from the literature on this specific issue.

In this normative layer we can distinguish two main types: contracts and proper norms. Contracts and proper norms are as we said ways to enforce behaviors by an “ought”. The main difference between contracts and norms is that contracts once accepted are constraints among specific individuals. When an agent, say Paolo Rossi, is hired by a company, say FIAT, he acquires new rights and obligations that are partially specified in the contract he signs. To be more precise contracts (once accepted) need to have both sides of the deontic relations filled by individuals, for norms (once valid) have just one side of the deontic relation. A contract is among two specific subject that commit to behave in such and such a way:

- persons adopt roles
- organizations adopt structured organizations (i.e. in some sense they adopt a specific structure to achieve a goal)

A norm is in a sense more general: there is a specific organization, suppose the Italian State, that fixes the constraints in the behavior for a vast number of persons. This is done by applying norms to roles, as in the case of Italian citizens: for all Italian citizens this regulation applies.

Agents’ behavior is not only forced by regulative or primary norms. In what we can call ‘realizations’ of organizations (e.g.. organizations where roles are filled by agents), agents are often linked to the organizations via agreements or contracts. A realization is then a particular instance of a designed organization in which all the roles are assigned to specific agents. For instance, all the persons that actually have a (direct or indirect) employment contract with FIAT constitute its actual realization. In this sense, a contract can be conceived of as the bridge between the descriptive level of designed organizations and the concrete level of agents, i.e. a sort of norm that links abstract roles with specific individual agents. When an agent, say Paolo Rossi, is hired by a company, say FIAT, he acquires new rights and

obligations that are partially specified in the contract he signs. The adoption of a role by an agent via a contract specifies or modifies the environment (right, obligations, permissions, resources, etc.) established at the moment of the creation of the role. This phenomenon is due to the fact that in the negotiation phase, the specific agent that will play the role can have the power of personalizing the environment of the role, i.e. the organization that employs the agent can be ready to weaken the constraints on the role, because it thinks that this specific agent is very important and that in any case probably it will be able to perform the delegated plan. To employ a particularly interesting agent, the organization can be also ready to modify its plan and structure in order to adapt to the capabilities of that interesting agent. On the other hand, the agent can decide to play a role and revise its own goal (just because it wants to enter in this organization, for example).

Another difference is that contracts are can be more easily seen as similar to promises⁶ than norms, except if we believe in a social contract view on norms that is not widely accepted [Searle, 2005].

Organizations are entities that strongly depend on norms. As we pointed out in [Bottazzi et al., 2007] in some accounts, as shown in [Miller, 2007], organizations are intended as completely made up of norms. In our account norms are essential in order to execute the plan of an organization. We understand norms as deontic constraints on behavior that apply on the possible actions that can be executed by the agents that are involved in a plan of an organization. By means of norms the objectives of the organizations are linked to their roles. The relations among these objectives and the roles can be permissions or obligations: as an example, a president is allowed to enter in some area but it is mandatory for the president not to interfere with the job of the technicians in that area.

They can be understood as what Searle [Searle, 1995b, Searle, 1969] calls “regulative rules” and Hart calls “primary rules” [Smith and Zaibert, 2006]. This kind of rules regulate antecedently existing forms of behavior. For instance, a rule like “drive on the right hand side of the road” regulates the driving without defining it: the driving exists before the rule that imposes duties to individuals by the way of the role that these individuals play.

Anyway they are classically seen as being:

- *Rights*: for instance, to use some resources, to have 2 weeks of holiday in the year, to have access to some information, etc.

⁶For a philosophical analysis of the ontology of the promises see [Mulligan, 1987]

- *Obligations* which constrain what the player of the role is obliged to do or bring about. Under this interpretation *objectives* can be seen as obligations, as they are what the agent playing the role is expected to bring about. Sometimes the constraints are set only on what must be achieved (the objective itself) and how to achieve it is left to the initiative of the agent playing the role, other times also the procedure to obtain the objective is constrained.

Following Fasli[Fasli, 2005], another interesting notion that can be seen as falling under the domain of obligations is *responsibility*. This is what Fasli says about it:

[...] an obligation to do something with the expectation that some act or output will result. [...] responsibility ultimately derives from the owner(s). The owner hires or appoints a group, often a board of directors, to be responsible for managing the organization, making decisions and reaching the targets set by the owner. A downward chain of responsibility is then established [...] The chain extends throughout the organization because each manager has its own obligation to fulfill which plays a part in meeting the owners expectations. [...] each manager is still held accountable for the outputs of those to whom s/he delegates tasks. [Fasli, 2005]

Nonetheless, this notion seems to be more linked to that of delegation than to that of obligation, as in some particular cases obligation (and responsibility) can be seen as deriving from a delegation.

In general these constraints can involve other roles (with respect to the same agent), for example it is possible to require that in order to play a role you cannot play another role, or, viceversa, you need to play another role.

We add constraints on the way the players of a role/org interact with the players of other roles/orgs. Here we need to consider a lot of different interaction relations. One important relation is *control/power* that can be a sort of marker of functional analysis, i.e. the “components” of organizations need to be considered only among the controlled roles/orgs.

if we look at the literature, this general notion of control/power is specified by refining it in more detailed differentiation, even if sometimes these distinctions are not very clear.

For instance, Fasli distinguishes between the following relations:

- *Power* – “the potential ability of an agent or group to exercise control over another agent or group. Thus, an agent may have the potential to control others, but may choose not to exercise this control” [Fasli, 2005].
- *Authority* – “In organizational settings, authority is power that has been legitimized within a specific social context” in short authority is “legitimate power” [Fasli, 2005]. Authority concerns two levels, firstly it “involves the right to take action, make decisions or use resources to accomplish certain goals” [Fasli, 2005] and secondly it “legitimizes the exercise of power and enables the issuance of instructions to others” [Fasli, 2005]. Authority stays in close relation to obligations, rights and responsibility. An agent responsible for or obliged to accomplishing certain tasks must have the authority to use resources to achieve those tasks. On the other hand the fact that some player of the role has the authority (power) to delegate some task over the players of the other role, assumes the possession of the right of delegating over other agents. After b is authorized by a , it can be said that b is obliged to a to bring about this task (compare Fasli).

Fasli pointed out that whereas “responsibility cannot be delegated down to others, but authority can”. Authority can be used to build a hierarchy of roles.

- *Obligations, Rights* – Besides the obligations and rights concerning one player, so the ones of the form “an agent is obliged (permitted) to do something (or bring about something)”, there are the norms which establish the relations between two agents, as the ones of the form “an agent a is obliged to agent b to do (bring about) something” and “and agent a has the right ψ over another agent b ” (Fasli).

[Grossi et al., 2005] lists, as properties and relations between “agent roles”: power, delegation of tasks, responsibilities, permissions, access to resources and communication, but, at the level of the organizational structure, it singles out three main relations: *power*, related to delegation, *coordination*, related to information and knowledge, and *control*, related to monitoring. This is because roles are conceived in terms of objectives, norms and information.

Objectives are what the agents commit to when they enact a role; norms are the constraints imposed over an agent when playing a certain role and

information refers to what an agent must know in order to be able to enact a role.

Power is seen as a chain of delegations, in the sense that an agent playing a certain role has the possibility of delegating to an agent playing another role.

The delegation of a task or a goal entails the obligation for the agent to whom the task/goal has been delegated to execute the task or achieve the goal; thus, together with the delegation, the relevant information that enables the addressee of the delegation to achieve the goal or execute the task must also be communicated.

Moreover, someone must control whether the goals are achieved and the norms respected.

In order to express these relations, [Grossi et al., 2005] uses *special propositions*:

- $Power(r, s)$ means that the agent playing role r can delegate goals or tasks to the agent playing role s ;
- $Coordination(r, s)$ means that the agent playing role r has access to all information accessible to the agent playing role s ;
- $Control(r, s)$ means that the agent playing role r controls (is responsible for) the agent playing role s .

According to [Demolombe and Louis, 2005], the primitive notions, with respect to normative relations between agents enacting roles are the notion of *obligation* and that of *action*; even the notion of *right* and *institutional powers* should be expressible in terms of obligations and actions.

Firstly, to say that an agent i has a *right* means to say that i has the *power*, when he asks to exercise his right on another agent j , showing that he indeed has this right, to create some obligation for the agent j . Then, [Demolombe and Louis, 2005] postulates a distinction between different kinds of right:

- the right of i to execute an action implies that j is obliged to make this action realizable for i or to prevent this action from being unrealizable for i ;
- the right of i to know an information implies that j is obliged to see to it that i knows this information;

- the right of i to be in a certain state implies that j is obliged to see to it that i is in that state.

Then, a semi-formal definition of *institutional power* is given:

$$power(i, s, cond, proc, n) = IF\ cond\ AND\ makes(i, proc)\ THEN\ n\ inside\ s$$

If some conditions $cond$ are satisfied, then, inside a certain institution s , an agent i has the *power* of creating a certain normative situation, represented by a proposition n , by realizing a certain procedure $proc$. From the constraint and interaction analyses it is possible to understand the possible effects of a group of players playing a role in the pattern considered. In this sense this can be seen as a sort of definition of (a weak notion of) group: a set of interacting agents with no common goals/objectives.

6 Design and existence of organizations

As already said, we consider organizations as agents with a strong teleological connotation, i.e. organizations exist for doing/achieving something. In our model, and in most theories of artifacts, the teleological aspect is intertwined with the *design* process: organizations are agents designed for doing/achieving something. These two aspects, teleology and design, are captured in our model by the plans organizations have. A plan specifies how the constraints coming from the environment (i.e. the external and physical constraints) and from the intentional dimension are taken into account to pursue the mission of the organization. While the design process and the designer(s) of an artifact can be quite clearly individuated and, before existing, an artifact needs to be completely specified, the case of organizations is more complex. Who are the creator and the designer of an organization is not always clear and the initial plan is often very rough and general. After its creation or foundation, an organization can change its plan, and refine its design, i.e. it can create and design new (sub) organizations to which it delegates some sub-plans. From the point of view of the delegator, the delegated sub-plan can be seen as a sort of ‘black-box’: the delegator has a sort of means-end perspective and therefore it is only interested in the achievement of a sub-goal according to some constraints. This is a quite typical feature of the functional analysis of artifacts. But notice that in the case of organizations the delegatee can refine the adopted sub-plan, i.e. it can ‘fulfill’ the black-box in a specific way. This fulfillment is exactly the aim of the design

process of artifacts that must specify how the opaque functional description can be realized by choosing a particular *structure* for the artifact. In this sense the structure of the organization can be specified in a dynamical and distributed way through delegation steps. In some cases the delegator can put very strong constraints that fix how the delegates interact, which are their powers and rights, etc. Thus, the delegator is designing its structure. In other cases the delegator imposes only very weak constraints and the choice of the needed (sub) structure is transferred to the delegatee. Note that in doing that the delegating organization is creating a new organization (the delegated organization) that needs to respect some constraints but normally has a certain degree of autonomy.

From the above considerations it is clear that the creation and the existence of an organization are central notions. We will try to address these complex problems in the following.

6.1 Creating organizations

Let us suppose that in the plan of an organization a there is a delegation to an agent b . At this design step (the one that includes the delegation we are talking about), the organization a can be interested in including in its plan somebody with given characteristics without committing to a specific person or ‘real’ agent. From the point of view of a , b is a sort of ‘ideal’ agent that behaves according to the imposed constraints, i.e. it is a creation of a necessary to the implementation of a ’s plan. b ’s adoption of the delegated sub-plan is then presupposed in the creation of b itself.

This example applies to the case of an institution i that creates a department or an office d with specific competences, rights and obligations in order to implement some specific task the institution needs. At the moment of the legal act that establishes the department d , the department itself is not autonomous, in the sense that d cannot decide anything and its design is given by i . But when d starts functioning, some ‘real’ agents need to cover at least some roles in d , and d itself can start to have some autonomy and decisional/acting power. Starting from a completely virtual dimension, d is becoming a more real and concrete agent. In this sense there are at least two levels of existence for organizations and roles: the *virtual* one and the *concrete* one. The virtual existence concerns the legal or intended establishment of the ideal agent that is supposed to do specific things, while the concrete existence concerns how actual agents that cover specific roles act

according to (or contrary to) the expectations included in the specification of the virtual agent. While i has complete control on the virtual agent, it has only a partial control on the concrete agent, i.e. the virtual agent enacted by some real and autonomous agents. The transformation from virtual to concrete agent of d passes through a negotiation phase that ends up with the signature of a *contract* between the real agent that *represents* i and the real agent (or the representative of the concrete agent) that will play the role in d . Depending on the constraints in the contract, the concrete agent is more or less free in acting. In particular, the contract establishes the environment in which the plan that the concrete agent has adopted (the contract is a sort of public statement of adoption) needs to be executed and the constraints on the possible way of refinement of that plan.

But, when we create a new organization from scratch, who establishes the initial plan of the organization? In this case there is no explicit delegation from an upper-organization. The designers of the new organization pre-exist to the organization and they decide the structure and the objectives of the organization before its establishment. Therefore, when the organization is founded it already has an initial plan to obtain its goals, and in general some roles (like manager) are enacted from the beginning, i.e. there are not completely virtual founded organizations.

Organizations exist although not all the roles involved in their plans have been covered, i.e. it is possible that sub-organizations responsible for some sub-plans are still not committed to specific agents in specific roles, i.e. they are still virtual. If the uncovered roles are necessary to the achievement of the goal of the organization, then, at this stage, the organization is unable to achieve its goal. Is that enough to say that the organization does not exist? Or can we say that the organization exists but “it cannot work”? I.e. is it enough to cover just some roles in order for an organization to exist?

Can we push the reasoning even farther? Is it necessary that the process of design is completed in order for an organization to exist? Or can organizations exist even when they have very rough and general plans? When an organization is established, in general it does not already have either a precise plan or all the employees, therefore it seems quite normal to accept this ‘vagueness’ in organizations. In any case, probably, it is also true that it is not possible to have an organization with all roles empty, i.e. somebody, some specific agent needs to be considered in the organization.

6.2 The notion of environment

Organizations are situated in an environment of other organizations that can be internal or external with respect to it. When two organizations are linked, this often affects the normative dimension. This kind of relations are considered in [Bottazzi and Ferrario, 2007] and in [Boella and van der Torre, 2004]. In [Bottazzi and Ferrario, 2007] an example of a specific interorganizational normative relation is given, namely the relation that holds between the Italian State and the University of Torino [Bottazzi and Ferrario, 2007]:

We could say that the University of Torino is in a way “nested” into the Italian State. The normativity of the relation relies on the fact that the descriptive system of the “contained” organization is, in some sense, more specialized with respect to the descriptive system of the “containing” one: all the norms that are valid in the Italian State must also be valid in the University of Torino.

There are other kinds of norms and normative relations in the institutional setting, for example we can consider again contracts and agreements that make alliances possible, as in the case of military alliances among countries, like in the NATO organization.

Anyway, here again we follow the artifact metaphor that underlies the whole deliverable. According to Peter Kroes: “The behavior of these systems [complex systems] is driven in a significant way by their technical elements, but the functioning of the whole system depends as much on the functioning of these technical components as on the functioning of a social infrastructure (legal systems, billing systems, insurance systems, etc.) and on the behaviour of humans actors. From an engineering point of view this raises the issue of whether the social infrastructure is to be regarded as part of the outer environment and is to be modelled as constraints for the design of technical systems, or is to be taken as a part of the system to be designed (...) [P]roblems arise about where to draw the boundary between the system under consideration and its environment. ”⁷.

Maybe this is less critical in the case of organizations because the normative environment is necessary and it is always present. For example the upper-organization of a sub-organization can also be included in its environment, in the sense that the upper-organization influence and has some

⁷Peter Kroes, *Engineering design*, manuscript

impact in the sub-organization. Note that in any case, the specific agents are not really considered in the environment, only the roles are included.

There is a difference between the design of classical artifacts, and the design of organizations that, at the end, have a specific level of autonomy. In this case the functioning of an organization can be established only by means of rules that control the emergent phenomena that can be born inside these organizations.

7 Conclusions

The main contribution of this research is to give an preliminary model of organizations by taking an ontological perspective on a well known issue in literature: that is their teleological dimension. This was done also by looking at fresh literature on artifacts from the point of view of philosophy, computer programming and technical design.

Taking into account these results allow us to give a compact model that explains in a simpler way how organizations are made: there is not just a hierarchical structure or a net of roles that participate in a huge amount work-flows inside an organization: all of these components exists for achieving some goal. For sure this approach has a long history⁸, but it was too often taken into account without trying a to build a formalism at a level of generalization that allows reusability and semantic transparency.

A designer may decompose the objective of an organization into sub-objectives assigned to sub-organizations or roles that are purportedly created to accomplish those objectives. Then it is possible possible to perform a check on the feasibility of the overall objective by the analysis of the chosen structure (*validation* of the design).

It is a long way in order to have a good and complete ontology of organization, but we believe that we are in a good direction. What we want now is to build an ontology of organization that:

- better suits the process of abstraction from planning to hierarchal structuration;
- that is able to check for coherence among goals and plans;

⁸[Simon, 1964]

- that allows to check compatibility among goals and structure in order to see if organizational means are compatible with its general and specific ends.

References

- [Belnap et al., 2001] Belnap, N., M., P., and M., X. (2001). *Facing the Future: Agents and Choices in our Indeterminist World*. Oxford Univeristy Press.
- [Bloor, 1997] Bloor, D. (1997). *Wittgenstein, Rules and Institutions*. Routledge, London and New York.
- [Boella and van der Torre, 2004] Boella, G. and van der Torre, L. (2004). Organizations as socially constructed agents in the agent oriented paradigm. In *LNAI n. 3451: Procs. of ESAW'04*, pages 1–13, Berlin. Springer Verlag.
- [Boella and van der Torre, 2006] Boella, G. and van der Torre, L. (2006). A foundational ontology of organizations and roles. In *Procs. of DALT'06 workshop at AAMAS'06*.
- [Bottazzi et al., 2006] Bottazzi, E., Catenacci, C., Gangemi, A., and Lehmann, J. (2006). From collective intentionality to intentional collectives: An ontological perspective. *Journal of Cognitive Systems Research*, 7(2-3):192–208.
- [Bottazzi and Ferrario, 2005] Bottazzi, E. and Ferrario, R. (2005). A path to an ontology of organizations. In *EDOC International Workshop on Vocabularies, Ontologies and Rules for The Enterprise (VORTE 2005)*.
- [Bottazzi and Ferrario, 2007] Bottazzi, E. and Ferrario, R. (2007). Preliminaries to a DOLCE ontology of organizations. *International Journal of Business Process Integration and Managemen*, (forthcoming).
- [Bottazzi et al., 2007] Bottazzi, E., Ferrario, R., Masolo, C., and Trypuz, R. (2007). Designing organizations: Towards a model. In Boella, G., van der Torre, L., and Verhagen, H., editors, *Normative Multi-agent Systems*, number 07122 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany. Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI), Schloss.

- [Castelfranchi and Falcone, 1998] Castelfranchi, C. and Falcone, R. (1998). Towards a theory of delegation for agent-based systems. *Robotics and Autonomous Systems, Special issue on Multi-Agent Rationality*, 24(3-4):141–157.
- [Cavedon and Sonenberg, 1998] Cavedon, L. and Sonenberg, E. (1998). On social commitment, roles and preferred goals. In *International Conference on Multi-Agent Systems*, Paris.
- [Cummins, 1975] Cummins, R. (1975). Functional analysis. *Journal of Philosophy*, 72:741–765.
- [Demolombe and Louis, 2005] Demolombe, R. and Louis, V. (2005). Normes, pouvoirs et rôles : vers une formalisation en logique. In *Journées Francophones sur les Systèmes Multi Agents*, Calais.
- [Dignum, 2004] Dignum, V. (2004). *A Model for Organizational Interaction: based on Agents, founded in Logic*. PhD thesis, Universiteit Utrecht.
- [Fasli, 2005] Fasli, M. (2005). On the interplay of roles and power. In *International Conference on Intelligent Agent Technology (IAT'05)*. IEEE.
- [Ghallab et al., 2004] Ghallab, M., Nau, D., and Traverso, P. (2004). *Automated Planning*. Morgan Kaufmann.
- [Grossi et al., 2005] Grossi, D., Dignum, F., Royakkers, L., and Dastani, M. (2005). Foundations of organizational structures in multi-agent systems. In Press, A., editor, *Proceedings of AAMAS'05, Fourth International Conference on Autonomous Agents and Multiagent Systems*, pages 690–697, Utrecht.
- [Guarino, 2005] Guarino, G. (2005). *L'uomo-istituzione*. Laterza, Roma-Bari.
- [Hindriks, 2005] Hindriks, F. (2005). *Rules and Institutions*. PhD thesis, Erasmus Universiteit Rotterdam.
- [Hodgson, 2006] Hodgson, G. M. (2006). What are institutions? *Journal of Economic Issues*, 15(1):1–25.

- [Houkes and Vermaas, 2006] Houkes, W. N. and Vermaas, P. E. (2006). Planning behavior : technical design as design of use plans. In Verbeek, P.-P. and C., B. J., editors, *User behavior and technology development*, chapter 20, pages 203–210. Springer Verlag.
- [Kroes, 2003] Kroes, P. (2003). Screwdriver philosophy. *Techné*, 6(3):22–35.
- [Lind, 1994] Lind, M. (1994). Modeling goals and functions of complex industrial plant. *Applied Artificial Intelligence*, 8(2).
- [Masolo et al., 2002] Masolo, C., Gangemi, A., Guarino, N., Oltramari, A., and Schneider, L. (2002). Wonderweb deliverable d17: The wonderweb library of foundational ontologies. Technical report.
- [Masolo et al., 2004] Masolo, C., Vieu, L., Bottazzi, E., Catenacci, C., Ferrario, R., Gangemi, A., and Guarino, N. (2004). Social roles and their descriptions. In *Ninth International Conference on the Principles of Knowledge Representation and Reasoning*, Whistler Canada.
- [Miller, 2002] Miller, S. (2002). *Social Action: a Teleological Account*. Cambridge University Press, Cambridge.
- [Miller, 2007] Miller, S. (2007). Social institutions. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*.
- [Mulligan, 1987] Mulligan, K. (1987). Promisings and other social acts: Their constituents and structure. In Mulligan, K., editor, *Speech Act and Sachverhalt*, pages 29–90. M. Nijhoff, Dordrecht.
- [Pacheco and Carmo, 2003] Pacheco, O. and Carmo, J. (2003). A role based model for the normative specification of organized collective agency and agents interaction. *Journal of Autonomous Agents and Multi-Agent Systems*, 6(2):145–184.
- [Pacheco and Santos, 2004] Pacheco, O. and Santos, F. (2004). Delegation in a role-based organization. In Lomuscio, A. and Nute, D., editors, *Proceedings of 7th International Workshop on Deontic Logic in Computer Science, DEON’04*, volume LNAI 3065, pages 209–227.

- [Sannella, 2000] Sannella, D. (2000). Algebraic specification and program development by stepwise refinement. In *Proc. 9th Intl. Workshop on Logic-based Program Synthesis and Transformation, LOPSTR'99*, volume 1817 of *Lecture Notes in Computer Science*, pages 1–9. Springer.
- [Sannella and Tarlecki, 1997] Sannella, D. and Tarlecki, A. (1997). Essential concepts of algebraic specification and program development. *Formal Aspects of Computing*, 9:229–269.
- [Scott, 2001] Scott, W. R. (2001). *Institutions and Organizations*. Sage, Thousand Oaks, CA.
- [Searle, 1969] Searle, J. R. (1969). *Speech Acts: An Essay on the Philosophy of Language*. Cambridge University Press, Cambridge.
- [Searle, 1995a] Searle, J. R. (1995a). *The Construction of Social Reality*. The Free Press, New York.
- [Searle, 1995b] Searle, J. R. (1995b). *The Construction of Social Reality*. The Free Press, New York.
- [Searle, 2005] Searle, J. R. (2005). What is an institution? *Journal of Institutional Economics (to appear)*.
- [Simon, 1964] Simon, H. A. (1964). On the concept of organizational goal. *Administrative Science Quarterly*, 9(1).
- [Simon, 1969] Simon, H. A. (1969). *The Sciences of the Artificial*. MIT-Press, Cambridge, Massachusetts, first edition.
- [Smith and Zaibert, 2006] Smith, B. and Zaibert, L. (2006). Hart, rawls, and searle: Social ontology and the problem of normativity. In Tsohatzidis, S. L., editor, *Intentional Acts and Institutional Facts: Essays on John Searle's Social Ontology*. Springer.
- [Tollefsen, 2002a] Tollefsen, D. (2002a). *Interpreting Organizations*. PhD thesis, Ohio State University.
- [Tollefsen, 2002b] Tollefsen, D. (2002b). Organizations as true believers. *Journal of Social Philosophy*, 33(3):395–410.

- [Tummolini and Castelfranchi, 2006] Tummolini, L. and Castelfranchi, C. (2006). The cognitive and behavioral mediation of institutions. *Cognitive System Research*, 7(2-3):307–323.
- [Tuomela, 1995] Tuomela, R. (1995). *The Importance of Us: a Philosophical Study of Basic Social Notions*. Stanford University Press, Stanford.
- [Tuomela, 1997] Tuomela, R. (1997). Searle on social institutions. *Philosophy and Phenomenological Research*, 57(2):435–441.
- [Tuomela and Balzer, 2003] Tuomela, R. and Balzer, W. (2003). Social institutions, norms, and practices. In *Social Order in Multiagent Systems*. Dordrecht:Kluwer.
- [von Wright, 1963] von Wright, G. H. (1963). *Norm and Action : a Logical Enquiry*. Routledge and Kegan Paul, London.