Accountability and Responsibility in Agent Organizations

Matteo Baldoni[1[0000-0002-9294-0408]] (✉), Cristina Baroglio[1[0000-0002-2070-0616]], Olivier Boissier[2[0000-0002-2956-0533]], Katherine M. May[1], Roberto Micalizio[1[0000-0001-9336-0651]], and Stefano Tedeschi[1[0000-0002-9861-390X]]

[1] Università degli Studi di Torino - Dipartimento di Informatica, Torino, Italy
firstname.lastname@unito.it

[2] Laboratoire Hubert Curien UMR CNRS 5516, Institut Henri Fayol, MINES Saint-Etienne, Saint-Etienne, France
Olivier.Boissier@emse.fr

Abstract. We discuss the limits of current agent organizations, and the benefits of introducing an explicit account of responsibility and accountability. We, then, illustrate how through such notions it is possible to design both organization specifications and organization entities, that are guaranteed to properly distribute responsibilities, that is, not only to own but also to connect the needed, distributed control over the goal so as to enable its achievement.

Keywords: Accountability · Responsibility · Agent Organizations.

1 Introduction

Multiagent Systems (MAS) provide a programming paradigm for the development of complex systems, which are characterized by multiple autonomous threads of execution that run in parallel, interact and coordinate with each other. Several design methodologies and programming platforms that have been proposed (e.g., [1]) are grounded on the metaphor of the *organization*. Such agent organizations represent strategies decomposing complex organizational goals into simpler sub-tasks and allocating them to roles. By adopting roles in the organization, agents acquire responsibilities and execute the corresponding tasks in a distributed, coordinated and regulated fashion.

However, even if current models are targeting open systems by allocating and enforcing rights and duties to agents about the tasks to realize, they lack an explicit representation of the relationships between the agents, resulting in the following limits: (*i*) difficulty for the agents to identify who should give restitution to whom for a certain state of the organization, (*ii*) even if agents who enter the organization are under the regulation of norms, that stipulate their rights and duties, the organization has no guarantee that they will provide all the accompanying proofs, that are induced by their responsibilities. We claim that the introduction of accountability relationships could enable a more fruitful participation of agents to the organization both from the agent and the organization perspectives. Accountability, indeed, is a fundamental concept that could help to overcome these limitations. However, this term is little understood and is often used to refer to answerability for one's actions or behavior. As [26] explains referring to Public Administrations, accountability plays a greater role in organizational processes than indicated by the idea of answerability. As underlined by the authors, accountability involves the *means* (i.e., the control on the necessary resources within and outside the organization) by which organizations and their members manage the expectations on fulfillment of their duties.

The objective of this paper is to introduce this broader perspective in agent organizations and demonstrate that it is an important ingredient that agent organizations should encompass. We propose thus to use accountabilities as explicitly taken social relationships, between an *account-giver* and an *account-taker* within an agent organization. Such relationships are mutually agreed by the parties and concern, in our proposal, roles, agents, goals (either complex or atomic). Accountability relationships are important in the design of agent organizations since, as we will explain, when they are properly defined the system properly distribute responsibilities, i.e., not only to own but also to "connect" the needed, distributed control over the goal so as to enable its achievement. In particular, we extend the specification of an organization in a way that enables the verification of the feasibility, for a group of agents, to incarnate the organizational roles properly, i.e., by respecting the accountabilities the agents can cooperate so as to achieve the organizational goal and discharge their responsibilities. The organizational model, thus, should no longer be a structure that distributes goals to its agents, but it should become *a way for coordinating responsibility assumption by the agents*.

The paper is so structured. Sec. 2 reviews the current existing agents organizational models and analyzes their current limitations w.r.t. accountability Sec. 3 proposes a formal definition of accountability and of the accompanying concept of responsibility. This proposal is then applied to agent organizations, defining accountability and responsibility relationships between roles and between agents. From these definitions, Sec. 4 shows how such relationships may help the design and enactment of agent organization definitions within open MAS. We illustrate this in the context of the $\mathcal{M}$OISE organization model which is part of the JaCaMo MAS-oriented programming platform [8]. The choice is representative of those approaches where the organization is described in terms of roles, goals and norms, and the organization issues obligations to push agents to pursue the assigned goals at the right moment.

2 Lack of accountability in Agent Organizations

To face the inherent need of coordination among autonomous agents, the *organization* metaphor has been used for a long time in MAS. When looking back, a set of initial proposals [13, 18] have defined an explicit structure of roles and relations, through which responsibilities of tasks are distributed by adoption of roles, among the agents participating in the organization. Such models are well adapted to "closed agent organization" where benevolent agents, always complying, coordinate with each other to achieve their responsibility assumptions (actions, goals or interactions). A second generation of organization models, following the electronic institution pioneering approaches [21], has introduced norms in the structure of roles and tasks, giving birth to *normative organization* [17, 19, 22, 16, 8]. These social coordination frameworks [1] are targeting "open socio-technical systems". Thanks to norms, the structures of distributed responsibilities among agents have been enriched with structures of *social expectations*: besides being the source of task responsibility assumption, roles have become the anchoring point of social expectations on the behavior of the agents who will play them in the organization. As for normative MAS [7], normative organizations are equipped with a set of mechanisms to publish, enact, adapt, monitor and enforce normative behaviors. Thus, once decided to adopt a role with the accompanying norms and thus participating to the organization, agents assume the responsibility of the targeted tasks. Moreover, they are expected by the organization to accomplish their duties. In case of violation, they are enforced to do so through sanctions.

However, while addressing the requirement of assigning duties and rights to agents, agent organizations are obfuscating accountability as pointed by [10, 11]. Agent autonomy demands a different way of conceptualizing coordination by clearly constraining them in terms of responsibilities that are explicitly taken on by them, and by establishing a directed relationship from one agent to another, that reflects the legitimate expectations the second principal has of the first. Agent organizations have not established yet the foundational facts of accountability, i.e., following [6] who has control over the situation and who is responsible of acting (or not) in accordance with established expectations.

Lack of control for the agents. Current agent organizations are lacking an easy way, for the agents participating to them but also for the organization designers, to check who has control over the situation. That is to say, checking that the means to execute the expected tasks are properly provided to the agents who become in charge of them through their adopted roles, i.e., enabling agents to have control/power on resources, on other tasks, on other agents on which their duties depend. For instance, in an organization coordinating the building of a house, a bricklayer who depends on some other worker in charge of preparing the site does not have the means to ask about occurring delays. Not even the house owner has such means. Indeed, even if each co-worker, by reasoning on the organization specification, may know about the existence of others with whom it should coordinate, the co-worker has no explicit endorsement from the organization to do so. Assignment of roles to tasks in organization specifications and then roles to agents via their enactments are not sufficient to explain the control structure of the expected coordinated tasks in the organization.

Agents' responsibility is not well captured. In most of the current normative organization approaches, when norms are enacted through adoption by agents of the role on which they bear, they are translated into deontic modalities or social commitments. Deontic modalities only constrain the agent who is in charge of fulfilling the norm. Targeting the control of its autonomy, they are lacking all what concerns the act of assuming responsibility in the broader context of the organization such as role adoption, detachment of duties. Social commitments [9, 28] help to capture the deliberate act of the agent that takes on a duty by adopting the role, but still lack to capture the adoption of role itself, and the detachment. Moreover, besides detachment and adoption, there is a lack of an explicit social ground that clearly models in terms that are known and agreed upon by all agents participating in the organization, what duties an agent has accepted to bring about in interaction with others [15]. That is to say, when a failure occurs in the agent organization, it is not possible for an agent to attribute causal responsibilities nor to identify the causes of the failure.

In most of the approaches, then, it is assumed that the sanctions associated with the violation (or fulfillment) of an obligation are a sufficient tool for constraining agents' behaviors. However, in order to apply sanctions, there is a need to conduct an inquiry process, to investigate on the reasons of the failure or success in order to properly apply adequate and justified sanctions. Agents participating to the organization are thus also expected to provide proofs and explanations of their behaviors in the organization w.r.t. their responsibilities.

Organizations are dynamic structures with a life cycle chaining design, role adoption, execution with fulfillment and enforcement of corresponding social expectations. When tasks participating in the definition of social expectations, connected to roles, change, the organization is changed, restarting a new cycle of design, adoption and fulfillment. All current models in MAS are assuming such a life cycle. However, there is no explicit commitment that organizations will not change the tasks attached to social expectations during the time agents play the corresponding role. In case of malevolent organization, for instance, it may be possible that the organization dynamically changes the set of allocated tasks to roles, roles already adopted by agents. Thus when the normative organization issues an obligation towards an agent, that agent may not have the desire or a proper capability for satisfying that obligation (see [5, 4] for instance). This short example demonstrates that currently social expectations are mainly directed from the organization to the agents, stating the expected agents' behaviors when adopting roles within the organization. The inverse relation, where the organization is engaging in expected behaviors with respect to the agents, does not exist.

3 Accountability in Agent Organizations

The accountability model that we will define aims at being used in the context of agent organizations. Before presenting this model, let's first introduce in Sec. 3.1, a definition of its components in the context of collective execution of plans, decomposition of tasks into subtasks with temporal relations as presented in Fig. 1. Based on this model, we will then define in Sec. 3.2, accountability relation in agent organizations.

3.1 Preliminary definitions

As a first step, it is necessary to provide a language for expressing those conditions and behaviors to which accountabilities refer. To this aim, we rely upon *precedence logic* [29]. Precedence logic is an event-based linear temporal logic devised for modeling and reasoning about Web service composition. The interpretation of such a logic deals with occurrences of events along runs (i.e., sequence of instanced events). Event occurrences are assumed to be non-repeating and persistent: once an event has occurred, it has occurred forever. The logic has three primary operators: '$\vee$' (choice), '$\wedge$' (concurrence), and '$\cdot$' (before). The *before* operator allows constraining the order with which two events must occur, e.g., $a \cdot b$ means that a must occur before b, but the two events

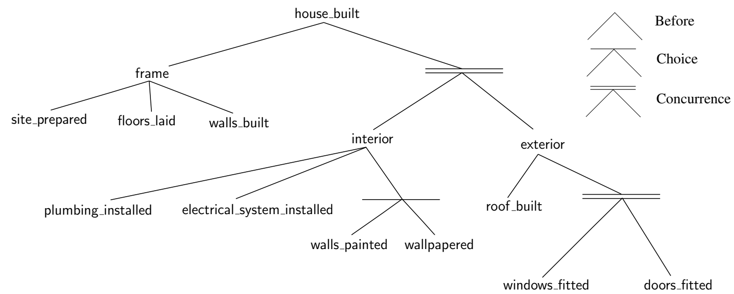


Fig. 1: The building-a-house goal's functional decomposition.

do not need to occur one immediately after the other. Such a language, thus, allows us to model complex expressions, whose execution needs to be coordinated as they are under the responsibility of different agents. Let e be an event. Then $\overline{e}$, the complement of e, is also an event. Initially, neither e nor $\overline{e}$ hold. On any run, either e or $\overline{e}$ may occur, not both. Intuitively, complementary events allow specifying situations in which an expected event e does not occur, either because of the occurrence of an opposite event, or because of the expiration of a time deadline.

Example 1 (Building a house). For the sake of explanation, we rely on the *building-a-house* example introduced in [8] for JaCaMo. We represent by means of precedence logic the functional specification of the organization:

- house_built $\doteq$ frame $\cdot$ (interior $\wedge$ exterior)
- frame $\doteq$ site_prepared $\cdot$ floors_laid $\cdot$ walls_built.
- interior $\doteq$ plumbing_installed $\cdot$ electrical_system_installed $\cdot$ (walls_painted $\vee$ wallpapered).
- exterior $\doteq$ roof_built $\cdot$ (windows_fitted $\wedge$ doors_fitted).

The main goal, house_built, requires the site to be prepared and then both the interior and exterior of the house to be built. The two activities can be performed in any order or even in parallel. All such sub-goals amount to complex processes. Most activities need to be carried out one after the other (e.g. site_prepared $\cdot$ floors_laid $\cdot$ walls_built) but concerning the walls, it will be up to the performer to decide whether to paint them or to lay paper on them. The decomposition of house_built is graphically shown by Figure 1.

We also rely on the notion of *residuation*, inspired by [25, 29]. Residuation allows tracking the progression of temporal logic expressions, hopefully arriving to their satisfaction, i.e., the completion of their execution. The *residual* of a temporal expression q with respect to an event e, denoted as q/e, is the remainder temporal expression that would be left over when e occurs, and whose satisfaction would guarantee the satisfaction of the original temporal expression q. Residual can be calculated by means of a set of rewrite rules. The following equations are due to Singh [29, 25]. Here, r is a sequence expression, and e is an event or $\top$. Below, Γ_u is the set of literals and their complements mentioned in u. Thus, for instance, $\Gamma_e = \{e, \overline{e}\} = \Gamma_{\overline{e}}$ and $\Gamma_{e \cdot f} = \{e, \overline{e}, f, \overline{f}\}$.

$$0/e \doteq 0 \qquad \top/e \doteq \top \qquad (r \wedge u)/e \doteq ((r/e) \wedge (u/e))$$
$$(r \vee u)/e \doteq ((r/e) \vee (u/e)) \qquad (e \cdot r)/e \doteq r, \text{ if } e \notin \Gamma_r \qquad r/e \doteq r, \text{ if } e \notin \Gamma_r$$
$$(e' \cdot r)/e \doteq 0, \text{ if } e \in \Gamma_r \qquad (\overline{e} \cdot r)/e \doteq 0$$

Using the terminology in [3], we say that an event e is *relevant* to a temporal expression p if that event is involved in p, i.e. $p/e \not\equiv p$. Let us denote by $\boldsymbol{e}$ a sequence $e_1, e_2, \ldots, e_n$ of events. We extend the notion of residual of a temporal expression q to a sequence of events $\boldsymbol{e}$ as follows: $q/\boldsymbol{e} = (\ldots((q/e_1)/e_2)/\ldots)/e_n$. If $q/\boldsymbol{e} \equiv \top$ and all events in $\boldsymbol{e}$ are relevant to q, we say that the sequence $\boldsymbol{e}$ is an *actualization* of the temporal expression q (denoted by $\widehat{q}$).

Example 2. Let $(a \cdot b)/a$ be the temporal expression b, while $(a \cdot b)/\overline{a}$ and $(a \cdot b)/b$ cause the temporal expression to become false, in the first case because the opposite event of a occurs, in the second because event b occurs in the wrong order. Referring to Example 1, the residual of (plumbing_installed · electrical_system_installed· (walls_painted $\vee$ wallpapered)) / plumbing_installed is electrical_system_installed · (walls_painted $\vee$ wallpapered). Instead, the residual of the latter temporal expression with respect to walls_painted would be false because the event occurrence disrupted the order, captured by the temporal expression. Finally, the residual of the temporal expression (plumbing_installed · electrical_system_installed· (walls_painted $\vee$ wallpapered)) with respect to the sequence {plumbing_installed, electrical_system_installed} is (walls_painted $\vee$ wallpapered).

3.2 Specifying Accountability in Agent Organizations

No unique and standard organization specification model exists yet to specify agent organization in the MAS domain. However, as pointed in [14], generally, their specifications are based on the concepts of *roles*, which have to be adopted by agents, *tasks* (e.g. actions, goals, interactions) assigned to roles through *norms*, usually expressed with deontic modalities. Agents' organizations' life cycle chains design, enactment (i.e., adoption of roles by agents, building what is usually called *organization entity*), execution (i.e., coordination of agents realizing their duties, monitoring, enforcement). The proposal that we explain fits equally well the specification of accountability in the context of organization specification and in the context of organization entity, i.e., accountability between roles at the specification level, and accountability between agents at the entity level. We let the structure of roles at the simplest expression (i.e., we won't consider groups or relations among roles).

In the following we use the notations $\mathsf{A}(x, y, r, u)$ and $\mathsf{R}(x, q)$ in order to explicitly represent accountabilities and responsibility assumptions respectively. By $\mathsf{A}(x, y, r, u)$ we express that x, the account-giver, is accountable towards y, the account-taker, for the condition u when the condition r (*context*) holds. Both r and u are temporal expressions, given in precedence logic. If we think of a process being collectively executed, we can say that when the r part of the process is done, then x becomes accountable of the u part. When u is true, x is considered to have satisfied the expectation that was put on it by exercising its control, which means that it has built a proof that can be supplied to the account-taker. A proof here is intended as a set of recorded facts, that demonstrate the achievement of the specified condition. Indeed, the account-taker can ask at any time for a proof to the account-giver, provided that r is true (in this case the accountability is detached). Such a proof of the partial execution will amount to the set of facts collected that far. Along with the execution, expectation and control will evolve and will run out with the satisfaction of the accountability, and only the final proof will be left. When, instead, u is false, the expectation was violated, and x's control failed. When r is false, instead, the accountability expires. This means that those conditions, which subtend both the expectation about u and its control, do not hold anymore. Instead, by $\mathsf{R}(x, q)$ we capture the responsibility assumption by x of the temporal expression q. When q is true the responsibility is fulfilled, when it is false, it is neglected.

We denote by **A** a set of accountabilities, calling it an *accountability specification*, and by **R** a *responsibility distribution*, that is a set of responsibility assumptions.

We use residuation to compute the progress of both accountabilities and responsibility assumptions: the idea is that, even though they are not temporal expressions, such relationships progress with the progress of their temporal expressions. $\mathsf{A}(x, y, r/\boldsymbol{e}, u/\boldsymbol{e})$ denotes the residual of $\mathsf{A}(x, y, r, u)$ with respect to the sequence of events $\boldsymbol{e}$. On the other hand, when $r/\boldsymbol{e} \doteq 0$, we say that the accountability expires; when $r/\boldsymbol{e} \doteq \top$ and $u/\boldsymbol{e} \doteq 0$, the accountability is violated; when $u/\boldsymbol{e} \doteq \top$ it is satisfied. Similarly, $R(x, q/\boldsymbol{e})$ denotes the residual of $R(x, q)$ with respect to the sequence of events $\boldsymbol{e}$, while when $q/\boldsymbol{e} \doteq \top$ the responsibility is fulfilled, and when $q/\boldsymbol{e} \doteq 0$, it is neglected.

As explained since the introduction, $\mathsf{A}(x, y, r, u)$ is grounded on *control* and *expectation*. While expectation is naturally conveyed with the accountability itself, the control needs to be recursively verified on the structure of u. In fact, x controls u either directly or indirectly by relying on accountabilities by other parties. In the following discussion we adopt the convention in [25] and limit sequences to just two events each. This is done to simplify the formalization and without loss of generality, because $e_1 \cdots e_n \equiv (e_1 \cdot e_2) \wedge \ldots \wedge (e_{n-1} \cdot e_n)$.

Definition 1 (Control). *Let* **A** *be an accountability specification, we denote by* $\xi(x, r, u)$ *the control in* **A** *of* x *over* u *in the context* r *(*$\xi(x, r, u)$ *in* **A** *holds). For control, the following rules hold:*

- $\xi(x, r, u)$ *in* **A** *if* $u/r = \top$*;*
- $\xi(x, r, u' \wedge u'')$ *in* **A** *if* $\xi(x, r, u')$ *in* **A** *and* $\xi(x, r, u'')$ *in* **A***;*
- $\xi(x, r, u' \vee u'')$ *in* **A** *if* $\xi(x, r, u')$ *in* **A** *or* $\xi(x, r, u'')$ *in* **A***;*
- $\xi(x, r, u)$*, where* $u/r = u' \cdot u''$*, in* **A** *if* $\xi(x, r, r \cdot u')$ *in* **A** *and* $\xi(x, r \cdot u', r \cdot u' \cdot u'')$ *in* **A***;*
- $\xi(x, r, u)$ *in* **A** *if there exists* $\mathsf{A}(y, x, r', u) \in \mathbf{A}$ *such that* $\xi(x, r, r')$ *in* $\mathbf{A} - \{\mathsf{A}(y, x, r', u)\}$.

Notice that having control does not mean having the ability of making a temporal expression become true in any case, but that x has the possibility of realizing it. Moreover, the control relation on atomic temporal expressions cannot be checked from the specification only. The check depends on the responsibility assumption by the agent who has adopted the role.

Definition 2 (Accountability Closure). *Let* **A** *be an accountability specification,* **A** *is* closed under control *if for each* $\mathsf{A}(x, y, r, u) \in \mathbf{A}$*, such that* u/r *is not atomic, we have* $\xi(x, r, u)$ *in* **A**.

Residuation preserves control, indeed the following proposition holds.

Proposition 1. *Let* $\mathbf{A}$ *be an accountability specification that is closed under control, and let* e *be an event, then* $\mathbf{A}/e$=$\{\mathsf{A}(x,y,r/e,u/e)\ |\mathsf{A}(x,y,r,u)\in\mathbf{A}\}$ *is still closed under control.*

Proof. The proof proceeds by induction. To correctly define the base case, we select, at each inductive step, the subset $\mathbf{A}$ of the accountabilities that are relevant to a specific control expression. More precisely, given a control expression $\xi(x,r,u)$ in $\mathbf{A}$, let us denote with $\mathbf{A}|_{r,u}$ a set of accountabilities $\{\mathsf{A}(z,w,p,q)\}\subseteq\mathbf{A}$ such that either $p/r\not\equiv p$, $p/u\not\equiv p$, $q/r\not\equiv q$, or $q/u\not\equiv q$. Now, it is sufficient to assume that the control rules in Definition 1 hold in $\mathbf{A}|_{r,u}$ (if this is true, they trivially hold in $\mathbf{A}$, too). Now let us show that $\mathbf{A}/e$ is closed under control if also $\mathbf{A}$ is closed under control. To demonstrate this, we show that the progression caused by an event e preserves control $\xi(x,r,u)$ in $\mathbf{A}|_{r,u}$. By induction over the size of $\mathbf{A}|_{r,u}$ and the length of u/r.

- Base case: $\xi(x,r,u)$ in $\mathbf{A}|_{r,u}$ holds and $u/r\equiv\top$, then it is obvious that also $\xi(x,r/e,u/e)$ holds in $\mathbf{A}$
- Base case: $\xi(x,r,u)$ in $\mathbf{A}|_{r,u}$=$\{A(z,w,r',u')\}$ and $\xi(x,r,r')$ in $\{\}$ hold. Given the rules of control in Definition 1, it must be the case that $u\equiv u'$, $w\equiv x$. In addition, since $\xi(x,r,r')$ cannot base control upon another accountability relation, it must be the case that $r\equiv r'$ (falling in the previous base case). It follows that if $\xi(x,r,u)$ in $\mathbf{A}|_{r,u}$=$\{A(z,x,r',u)\}$ and $\xi(x,r,r)$ in $\{\}$ hold, also $\xi(x,r/e,u/e)$ in $(\mathbf{A}|_{r/e,u/e})/e$=$\{A(z,x,r'/e,u/e)\}$ and $\xi(x,r/e,r/e)$ in $\{\}/e$ hold.
- Inductive step, cases $\xi(x,r,u'\wedge u'')$ and $\xi(x,r,u'\vee u'')$ follow from definition of residuation, and from the rules of control.
- Inductive step, case $\xi(x,r,u)$, where $u/r=u'\cdot u''$. We have that $\xi(x,r,r\cdot u')$ in $\mathbf{A}|_{r,r\cdot u'}$ and $\xi(x,r\cdot u'/e,r\cdot u'\cdot u''/e)$ in $(\mathbf{A}|_{r\cdot u'/e,r\cdot u'\cdot u''/e})/e$ from the definition of control. Now, by inductive hypothesis, we have that $\xi(x,r/e,r\cdot u'/e)$ in $(\mathbf{A}|_{r/e,r\cdot u'/e})/e$ and $\xi(x,r\cdot u'/e,r\cdot u'\cdot u''/e)$ in $(\mathbf{A}|_{r\cdot u'/e,r\cdot u'\cdot u''/e})/e$. Thus, we conclude that $\xi(x,r/e,u/e)$ in $(\mathsf{A}|_{r/e,u/e})/e$. In fact if e is not relevant to r nor to u, than no change occurs, and control is trivially preserved. If e is relevant to both r and u, then e cannot be but the first element of r since otherwise it would progress the consequent condition to 0.
- We have $\xi(x,r,u)$ in $\mathbf{A}|_{r,u}$, then there is $\mathsf{A}(y,x,r',u)\in\mathbf{A}|_{r,u}$ s.t. $\xi(x,r,r')$ in $(\mathbf{A}|_{r,u}-\{\mathsf{A}(y,x,r',u)\})|_{r,r'}$. By inductive hypothesis, $\xi(x,r/e,r'/e)$ in $(\mathbf{A}|_{r,u}-\{\mathsf{A}(y,x,r',u)\})|_{r,r'}/e$, that is $(\mathbf{A}|_{r/e,u/e}-\{\mathsf{A}(y,x,r',u)\})|_{r,r'}$. Now, $\xi(x,r/e,r'/e)$ in $(\mathbf{A}|_{r/e,u/e}-\{\mathsf{A}(y,x,r',u)\})$ because of the last set includes $\mathbf{A}|_{r/e,u/e}-\{\mathsf{A}(y,x,r',u)\})|_{r,r'}$ and, for the same reason, $\xi(x,r'/e,r/e)$ in $\mathbf{A}|_{r/e,u/e}$ that proves the case. ■

In words, this property means that the possibility of realizing a temporal expression is not disrupted by the occurrence of events, and that possibility remains step after step. Of course, agents, in their autonomy will maintain the decision about what to do (e.g., make an accountability expire) but this will remain in the proof.

We now show how accountabilities are complemented with responsibilities to the aim of developing full organization specifications and organization entities. In the following, we denote by $\mathbb{A}$ a set of accountability specifications $\mathbf{A}_i$, each of which is closed under control. Intuitively, each $\mathbf{A}_i$ in $\mathbb{A}$ represents a proper way to achieve the organizational goal. $\mathbb{A}$ is therefore the set of alternative solutions the organization designer considers as acceptable at runtime. In this paper, we assume that the designer has specified $\mathbb{A}$ in a way that complies with the design aims. In particular, we assume that for each $\mathbf{A}_i$ there is at least a sequence of events $\boldsymbol{e}$ that satisfies all the accountabilities in $\mathbf{A}_i$, allowing the achievement of the organizational goal. We denote by $[\![\mathbf{A}_i]\!]$ the set of event sequences that satisfy all the accountabilities in $\mathbf{A}_i$.

Any actual set of agents enacting roles within the organization, should therefore be such to satisfy at least one of the accountability specifications $\mathbf{A}_i$ in $\mathbb{A}$. To verify whether this occurs, we approach the problem in general terms by means of the responsibility characterization. Intuitively, assuming that agents are willing to take on a set of responsibilities, each declaring what is willing to bring about within the organization, the problem becomes to verify whether such a set of responsibility declarations fits at least one of the accountability specifications in $\mathbb{A}$. Part of such responsibilities will be due to the roles agents will enact, thus they can be considered as part of the specification of the organization (role responsibility), e.g., deduced from the definition of norms that connect roles to goals. Part of them may, instead, have as a source the agents themselves – constraints they pose on the organization for playing roles (agent responsibility). Depending on the source of responsibilities that is considered, thus, it will be possible either to check the consistency of the specification of the organization or to check the feasibility for a group of agents to incarnate the foreseen roles properly, i.e., respecting the accountabilities and preserving closure under control, which means that the agents can cooperate so as to achieve the organizational goal and discharge their responsibilities. The problem is formalized as follows.

Definition 3 (Accountability fitting). *Given a set of accountability specifications* $\mathbb{A}$ *and a responsibility distribution* $\mathbf{R}$*, we say that* $\mathbf{R}$ fits $\mathbb{A}$*, denoted by* $\mathbf{R} \rightsquigarrow \mathbb{A}$*, if there is* $\mathbf{A} \in \mathbb{A}$ *such that for each accountability* $\mathsf{A}(x, y, r, u) \in \mathbf{A}$*, there is a responsibility* $\mathsf{R}(x, q) \in \mathbf{R}$ *such that, for some actualization* $\widehat{q}$*,* $(u/r)/\widehat{q} \equiv \top$.

In particular, the following propositions hold.

Proposition 2. *Given a set of accountability specifications* $\mathbb{A}$*, and a responsibility distribution* $\mathbf{R}$*, such that* $\mathbf{R} \rightsquigarrow \mathbb{A}$*, let* e *be an event, then* $\mathbf{R}/e \rightsquigarrow \mathbb{A}/e$*.*

Proof. By Definition 3 (Accountability Fitting), we know that there exists in $\mathbb{A}$ at least one accountability specification $\mathbf{A}$ such that, for each $\mathsf{A}(x, y, r, u)$ there is one responsibility declaration $\mathsf{R}(x, q) \in \mathbf{R}$ such that $(u/r)/\widehat{q} \equiv \top$ some actualization $\widehat{q}$ of q. To show that $\mathbf{R}/e \rightsquigarrow \mathbb{A}/e$ we have just to show that $\mathbf{R}/e \rightsquigarrow \mathbf{A}/e$, for any possible event e. If e is irrelevant to $\mathbf{R}$, then $\mathbf{R}/e \equiv \mathbf{R}$, and hence the same actualizations that make $\mathbf{R}$ fit $\mathbf{A}$ are still possible in $\mathbf{R}/e$, and hence $\mathbf{R}/e \rightsquigarrow \mathbf{A}/e$. If e is relevant to some responsibility declaration $\mathsf{R}(x, q)$ in $\mathbf{R}$ we consider two cases:

- e is not relevant to $\mathbf{A}$, this happens when the responsibility taken on by x covers a wider set of events than actually required by $\mathbf{A}$. The actualizations of the residual expression q/e, thus, can still bring some accountabilities in $\mathbf{A}$ to satisfaction, namely for some $\mathsf{A}(x, y, r, u) \in \mathbf{A}$ it must happen $(u/r)/\widehat{q/e} \equiv \top$, and hence $\mathbf{R}/e \rightsquigarrow \mathbf{A}/e$.

- e is relevant to **A**, that is, there exists at least one accountability $\mathsf{A}(x, y, r, u)$ such that e is relevant for r, u, or both. Since by hypothesis we know that $(u/r)/\widehat{q} \equiv \top$ holds, it must also hold that $((u/e)/(r/e))/\widehat{q/e} \equiv \top$, in fact the actualizations of the residual expression q/e are just suffixes of the actualizations of expression q.

Thus, since whichever event e occurs $\mathbf{R}/e \rightsquigarrow \mathbf{A}/e$, we conclude that $\mathbf{R}/e \rightsquigarrow \mathbb{A}/e$. ■

Proposition 3. *Given a set of accountability specifications* $\mathbb{A}$*, and a responsibility distribution* $\mathbf{R}$ *such that* $\mathbf{R} \rightsquigarrow \mathbb{A}$*, then, there exists* $\boldsymbol{e}$ *such that: (1)* $\boldsymbol{e} = \widehat{q}$ *where* $q = \bigwedge_{\mathsf{R}(x,q_i)\in\mathbf{R}} q_i$ *(2)* $\boldsymbol{e} \in [\![\mathbf{A}_i]\!]$*, for some* $\mathbf{A}_i$ *in* $\mathbb{A}$*.*

Proof. Let us assume, by absurd, that the sequence $\boldsymbol{e}$ does not exist. This means that for any sequence $\boldsymbol{e}$ we obtain by the actualizations of the responsibility declarations in **R**, and for all accountability specification $\mathbf{A}_i \in \mathbb{A}$, there is at least one accountability $\mathsf{A}(x, y, r, u) \in \mathbf{A}_i$, that does not progress to satisfaction when $\mathsf{A}(x, y, r/\boldsymbol{e}, u/\boldsymbol{e})$. That is to say, there is a gap in the responsibilities due to the fact that $\mathsf{R}(x, p)$, such that $(u/r)/\widehat{p} \equiv \top$ is missing. Of course, this is not possible as we are assuming by hypothesis that $\mathbf{R} \rightsquigarrow \mathbb{A}$ and hence, there must exist at least one accountability specification **A** in $\mathbb{A}$, such that, for accountability $\mathsf{A}(x, y, r, u)$ in **A** there exists $\mathsf{R}(x, p)$, such that $(u/r)/\widehat{p} \equiv \top$. It follows that, when a sequence $\boldsymbol{e}$ is an actualization of all the responsibilities in **R**, it will also be an actualization of each accountability in at least one $\mathbf{A} \in \mathbb{A}$. ■

Schlenker's well-known triangle model of responsibility [27] states that responsibility depends on three linkages called prescription-identity, identity-event, and event-prescription. Only when the three linkages are drawn will an individual feel responsible for something. We resort on this model to summarize what our proposal adds to organization specifications and organization entities.

Responsibility assumptions in **R** describe which duties agents take on when playing some roles inside an organization. From an organization designer's perspective, such duties would be captured in the simplest case through norm specification, or, in a richer form, norms would be complemented with requirements the agents have to comply with for adopting roles concerned by the norms. Still, this is on the organization side. On the agent side, obligations (per se) are received by fiat; following [20], they succeed in directing individual behavior only when they agree with the sensitivity of the individuals. Our proposal fills this design gap through explicitly declared/taken responsibility assumptions and accountability relationships, which give agents the means for reasoning about the implications of role enactment, and give designers the means for specifying organizations that show the good characteristics expressed by Proposition 3.

For a normative organization to function well, its agents should interiorize the norms in their behavior but when can this happen in open organizations? If the agent considers a norm (say, an obligation) as a prescription concerning one of its identities (i.e., one of the roles it plays in the organization having that norm) the norm would start being something more than "given by fiat". In our proposal this can be done because, even though here we do not focus on the process, **R** can be derived from the organization norms; in some case, the norm specification could even reduce to the specification of the responsibility distribution – which duties are up to which roles. That would not, however, be enough if the same agent cannot see the connection between the prescription and some events it concerns (in our setting, the prescription would apply in a context), and also between the event and the identified identity (the context as one in which the role has control over something). It is the co-presence of the three linkages (1) to create in the agent the urge to tackle that context, abiding by the prescription, by virtue of its role, should the prescription apply; (2) that helps the designer to create organizations where role specification and goal distribution combine well.

On the other hand, $\mathbb{A}$ is focused on accountability, basing it on the coordination aspect, and specifies alternative ways to be accountable in the achievement of the organizational goals. This separation of concerns encourages both modularity and reuse. In fact, the accountability specifications can be defined and verified w.r.t. responsibility concerning roles independently of the actual agents that will play roles in the organization itself. The separation of concerns is at two levels. First, the organization specification level: a same organization can be characterized by several accountability specifications and several responsibility distributions, that fit with each other. Second, the organization entity level. Here, the same set of agents can take responsibility and be accountable in different organizations specifications, as well as, different sets of agents could take responsibility and be accountable in the same organization specification. The proposed formalization enables the check that the responsibility, taken by agents according to the responsibility specification in the organization specification, fits the accountability in the organization entity, that is enacted from the accountability specification in the organization specification.

4 Example: Building a House

Let's consider Example 1, originally presented in [8], and relying on the three dimensions of the $\mathcal{M}$OISE organizational model [23]. The structural dimension specifies roles, groups and links between roles in the organization. The functional dimension is composed of one or more schemes that elicit how the global organizational goals are decomposed into subgoals and how these subgoals are grouped in coherent sets, called missions, to be distributed to roles in the normative dimension. This latter binds the two previous dimensions by specifying the roles' permissions and obligations for missions. While the model presented in the previous section is independent of a particular agent organization model, we demonstrate its use and interest, focusing on the roles, functional and normative dimensions of the $\mathcal{M}$OISE model.

As in the original *building-a-house* example, an agent wants to build a house on a plot. To achieve this goal the companies, he has contracted with, must coordinate and execute various tasks, part of which can be executed in parallel, while part depends on other tasks. The temporal order is specified by the functional specification of the process (cf. Fig. 1). Its translation in precedence logic is given in Example 1. The structural specification defines a group which includes the following roles: *House Owner* (ho), *Boss* (bo), *Frame Manager* (fm), *Interior Exterior Manager* (iem), *Site Prep Contractor* (spc), *Bricklayer* (bl), *Roofer* (ro), *Fitter* (ft), *Plumber* (pl), *Electrician* (el), *Painter* (pa). The normative specification defining how each of the goals are allocated to roles will be described later when mapping it in terms of responsibilities.

The top-level organizational goal is house_built, of which bo should be in charge. On this basis, the designer can define the accountability relationship $\mathsf{A}(bo, ho, \top, \mathsf{frame} \cdot (\mathsf{interior} \wedge \mathsf{exterior}))$. Below, we report an example accountability specification $\mathbf{A}_1$ (see Fig. 2a) that includes the accountability of interest:

- $\mathtt{a1}_1$. $\mathsf{A}(bo, ho, \top, \mathsf{frame} \cdot (\mathsf{interior} \wedge \mathsf{exterior}))$
- $\mathtt{a1}_2$. $\mathsf{A}(fm, bo, \top, \mathsf{frame})$
- $\mathtt{a1}_3$. $\mathsf{A}(spc, fm, \top, \mathsf{site\_prepared})$
- $\mathtt{a1}_4$. $\mathsf{A}(bl, fm, \mathsf{site\_prepared}, \mathsf{site\_prepared} \cdot \mathsf{floors\_laid})$
- $\mathtt{a1}_5$. $\mathsf{A}(bl, fm, \mathsf{site\_prepared} \cdot \mathsf{floors\_laid}, \mathsf{site\_prepared} \cdot \mathsf{floors\_laid} \cdot \mathsf{walls\_built})$
- $\mathtt{a1}_6$. $\mathsf{A}(iem, bo, \mathsf{frame}, \mathsf{frame} \cdot \mathsf{interior})$
- $\mathtt{a1}_7$. $\mathsf{A}(iem, bo, \mathsf{frame}, \mathsf{frame} \cdot \mathsf{exterior})$
- $\mathtt{a1}_8$. $\mathsf{A}(pl, iem, \mathsf{frame}, \mathsf{frame} \cdot \mathsf{plumbing\_installed})$
- $\mathtt{a1}_9$. $\mathsf{A}(el, iem, \mathsf{frame} \cdot \mathsf{plumbing\_installed},$
  $\mathsf{frame} \cdot \mathsf{plumbing\_installed} \cdot \mathsf{electrical\_system\_installed})$
- $\mathtt{a1}_{10}$. $\mathsf{A}(pa, iem, \mathsf{frame} \cdot \mathsf{plumbing\_installed} \cdot \mathsf{electrical\_system\_installed},$
  $\mathsf{frame} \cdot \mathsf{plumbing\_installed} \cdot \mathsf{electrical\_system\_installed} \cdot \mathsf{walls\_painted})$
- $\mathtt{a1}_{11}$. $\mathsf{A}(ro, iem, \mathsf{frame}, \mathsf{frame} \cdot \mathsf{roof\_built})$
- $\mathtt{a1}_{12}$. $\mathsf{A}(ft, iem, \mathsf{frame} \cdot \mathsf{roof\_built}, \mathsf{frame} \cdot \mathsf{roof\_built} \cdot \mathsf{windows\_fitted})$
- $\mathtt{a1}_{13}$. $\mathsf{A}(ft, iem, \mathsf{frame} \cdot \mathsf{roof\_built}, \mathsf{frame} \cdot \mathsf{roof\_built} \cdot \mathsf{doors\_fitted})$

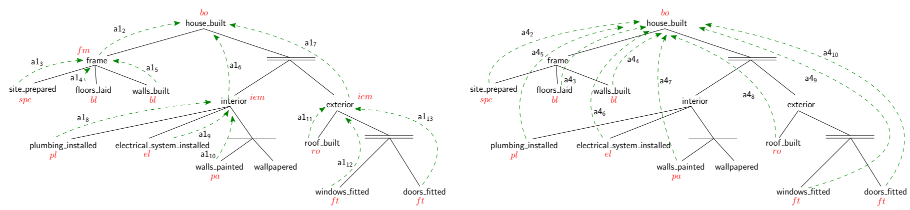
It is easy to see that $\mathbf{A}_1$ is closed under control (see Definition 2). Let us start with $\mathtt{a1}_1$. We must verify if $\xi(bo, \top, \mathsf{frame} \cdot (\mathsf{interior} \wedge \mathsf{exterior}))$ holds. To this aim, by Definition 1, we should have that $\xi(bo, \top, \mathsf{frame})$ and $\xi(bo, \mathsf{frame}, \mathsf{frame} \cdot (\mathsf{interior} \wedge \mathsf{exterior}))$, which is true because of the accountabilities $\mathtt{a1}_2$, $\mathtt{a1}_6$, and $\mathtt{a1}_7$. Similarly for every relationship in $\mathbf{A}_1$. The choice walls_painted $\vee$ wallpapered in Example 1 enables an alternative accountability specification $\mathbf{A}_2$, by substituting $\mathtt{a1}_{10}$ with $\mathsf{A}(pa, iem, \mathsf{frame} \cdot \mathsf{plumbing\_installed} \cdot \mathsf{electrical\_system\_installed}, \mathsf{frame} \cdot \mathsf{plumbing\_installed} \cdot \mathsf{electrical\_system\_installed} \cdot \mathsf{wallpapered})$. We could, then, define $\mathbb{A}$ as the set $\{\mathbf{A}_1, \mathbf{A}_2\}$ if both are considerd adequate by the designer.

$\mathbf{A}_1$ and $\mathbf{A}_2$ rely on two managers, fm and iem, who act as intermediaries between their account-givers and bo: bo controls the overall process through the accountability relationships in which it is account-taker and the managers are account-givers. Accountability specification $\mathbf{A}_3$ shows a more substantial change. Here, fm is removed and spc and bl are directly accountable towards bo:

- $\mathtt{a3}_1$. $\mathsf{A}(spc, bo, \top, \mathsf{site\_prepared})$
- $\mathtt{a3}_2$. $\mathsf{A}(bl, bo, \mathsf{site\_prepared}, \mathsf{site\_prepared} \cdot \mathsf{floors\_laid})$
- $\mathtt{a3}_3$. $\mathsf{A}(bl, bo, \mathsf{site\_prepared} \cdot \mathsf{floors\_laid}, \mathsf{site\_prepared} \cdot \mathsf{floors\_laid} \cdot \mathsf{walls\_built})$

The extreme is when all accountabilities, though having the already seen shape, show bo as account-taker. This leads to $\mathbf{A}_4$ (see Fig. 2b) which includes:

- $\mathtt{a4}_1$. $\mathsf{A}(bo, ho, \top, \mathsf{frame} \cdot (\mathsf{interior} \wedge \mathsf{exterior}))$
- $\mathtt{a4}_2$. $\mathsf{A}(spc, bo, \top, \mathsf{site\_prepared})$
- $\mathtt{a4}_3$. $\mathsf{A}(bl, bo, \mathsf{site\_prepared}, \mathsf{site\_prepared} \cdot \mathsf{floors\_laid})$
- $\mathtt{a4}_4$. $\mathsf{A}(bl, bo, \mathsf{site\_prepared} \cdot \mathsf{floors\_laid}, \mathsf{site\_prepared} \cdot \mathsf{floors\_laid} \cdot \mathsf{walls\_built})$
- $\mathtt{a4}_5$. $\mathsf{A}(pl, bo, \mathsf{frame}, \mathsf{frame} \cdot \mathsf{plumbing\_installed})$
- $\mathtt{a4}_6$. *and so forth ....*



(a) $\mathbf{A}_1$ (b) $\mathbf{A}_4$

Fig. 2: Two accountability specifications for the building-a-house organization. Green arrows depict who is accountable towards whom.

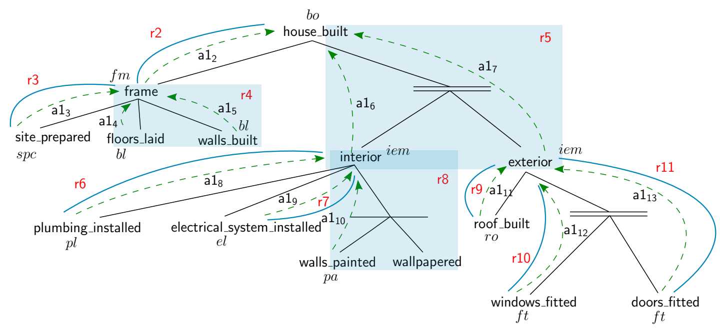


Fig. 3: Accountability specification $\mathbf{A}_1$ fitted by $\mathbf{R}$

Also $\mathbf{A}_3$ and $\mathbf{A}_4$, if deemed adequate, may be included in $\mathbb{A}$.

Finally, let us consider $\mathbf{A}_5$, which is similar to $\mathbf{A}_1$, but for $\mathsf{a1}_3$, substituted by $\mathsf{a5}_3$: $\mathsf{A}(spc, bo, \top, \mathsf{site\_prepared})$, and where $\mathsf{a1}_8$ is not defined. Two main problems can be identified. First, for fm to have control over frame (i.e., site_prepared $\cdot$ floors_laid $\cdot$ walls_built), there should be three accountability relationships, one for each event, with fm as account-taker. In particular, spc should be accountable to fm rather than to bo (as, instead, encoded in $\mathsf{a5}_3$) for site_prepared. Second, there is no accountability concerning plumbing_installed.

Let us now consider the set of responsibility assumptions $\mathbf{R}$, depicted in Fig. 3:

r1. $\mathsf{R}(bo, \mathsf{frame} \cdot (\mathsf{interior} \wedge \mathsf{exterior}))$
r2. $\mathsf{R}(fm, \mathsf{frame})$
r3. $\mathsf{R}(spc, \mathsf{site\_prepared})$
r4. $\mathsf{R}(bl, \mathsf{floors\_laid} \cdot \mathsf{walls\_built})$
r5. $\mathsf{R}(iem, \mathsf{interior} \wedge \mathsf{exterior})$
r6. $\mathsf{R}(pl, \mathsf{plumbing\_installed})$
r7. $\mathsf{R}(el, \mathsf{electrical\_system\_installed})$
r8. $\mathsf{R}(pa, \mathsf{walls\_painted})$
r9. $\mathsf{R}(ro, \mathsf{roof\_built})$
r10. $\mathsf{R}(ft, \mathsf{windows\_fitted})$
r11. $\mathsf{R}(ft, \mathsf{doors\_fitted})$

This set of responsibilities can be deduced from the normative specification of the $\mathcal{M}$OISE organization specification that connects roles to goals of the functional specifications through missions. Thus, for instance, $\mathsf{R}(bo, \mathsf{frame} \cdot (\mathsf{interior} \wedge \mathsf{exterior}))$ is deduced from the norm stating that *bo* has the obligation of achieving interior ∧ exterior after frame. Due to space limitations we cannot provide this normative specification.

It can easily be shown that **R** fits $\mathbb{A}$. Recalling Definition 3, there should be at least one $\mathbf{A}_i \in \mathbb{A}$ such that, for each accountability belonging to $\mathbf{A}_i$ there is a responsibility declaration belonging to **R** with an actualization that allows to satisfy the accountability, thus discharging the responsibility. This holds in particular with respect to $\mathbf{A}_1$. Indeed, let's consider, for instance, $\mathtt{al}_6$, $\mathtt{al}_7$ and r5. An actualization of interior is {plumbing_installed, electrical_system_installed, walls_painted}. Similarly, a possible actualization of exterior is {roof_built, windows_fitted, doors_fitted}. It's important to point out that these are not the only two allowed actualizations. Moreover, any interleaving which preserves the partial ordering of the two sequences is an actualization of interior ∧ exterior. It is easy to show that such an actualization would bring both the consequents of $\mathtt{al}_6$, $\mathtt{al}_7$ (residuated w.r.t. the antecedent) to $\top$. The same holds for every accountability in $\mathbf{A}_1$. As a remark, it is important to highlight that the responsibility distribution **R** would fit $\mathbf{A}_4$, as well. However, the absence of the two managers in the accountability specification would make r2 and r5 unnecessary.

5 Conclusions and Future Work

In this paper we have proposed the foundational facts for specifying accountability within agent organizations. Such a specification complements the responsibility assumption, coming from the normative specification, which expresses social expectations on the rights and duties of the agents, participating to the organization. Based on precedence logic, this model of accountability involves roles in an organization specification, or agents in an organization entity. It is based on control relations and on social expectations about tasks contributing to the achievement of organizational goals. We have shown how it is possible to check that responsibility distribution in roles (or agents while playing roles) fits the accountability relationships coming from the accountability specification of the organization. As such, the proposal has many application fields, like software development, agent reasoning, organization management. Since the proposal relies on the constitutive elements of the approaches to agent organizations, i.e. roles, goals and norms, we could illustrate how this model can be used in the context of the $\mathcal{M}$OISE organization model, component of the JaCaMo platform. While here mainly demonstrated to support an organization designer–to check the coherence between an accountability specification and a responsibility distribution, the model could be used also at the agent level for providing agents with reasoning capabilities on those accountability relationships that come from the organizations to which they (may) participate.

Besides what already discussed, accountability relationships and responsibility, and in particular the presented proposal, will help to enrich the expression of social expectation by enlarging their scope and by introducing expectations from agents towards organization – thus turning organizations into structures of bilateral social expectations. The scope of social expectations are limited to the task to be executed. They are currently missing all what concerns the adoption of roles by agents. It is, usually, assumed that an agent who is going to adopt a role will necessarily have proper capabilities for each task it will ever receive. Organizations are trusting the agents for having the right capabilities for realizing what is expected from them when playing roles. There is no way of expressing social expectations on requirements for playing a role and of checking them. Few models have addressed this question by proposing modeling of contracts [12] or [24, 2] involving agents and organization when adopting roles.

In the future, we intend to extend the proposal by including other forms of accountability relationships, such as the negative accountability, i.e., to capture that someone is expected not to impede social progress and negatively impact others. We will also study ways for leading agents to an agreement on a specific fitting of accountabilities when more than one exists.

References

1. Huib Aldewereld, Olivier Boissier, Virginia Dignum, Pablo Noriega, and Julian Padget, editors. *Social Coordination Frameworks for Social Technical Systems*, volume 30 of *Law, Governance and Technology Series*. Springer International Publishing, August 2016.
2. Huib Aldewereld, Virginia Dignum, Catholijn M. Jonker, and M Birna van Riemsdijk. Agreeing on role adoption in open organisations. *KI-Künstliche Intelligenz*, 26(1):37–45, 2012.
3. Matteo Baldoni, Cristina Baroglio, Federico Capuzzimati, and Roberto Micalizio. Type Checking for Protocol Role Enactments via Commitments. *Journal of Autonomous Agents and Multi-Agent Systems*, 32(3):349–386, May 2018.
4. Matteo Baldoni, Cristina Baroglio, Katherine M. May, , Roberto Micalizio, and Stefano Tedeschi. Computational Accountability in MAS Organizations with ADOPT. *Journal of Applied Sciences, special issue "Multi-Agent Systems"*, 8(4):489, March 2018.
5. Matteo Baldoni, Cristina Baroglio, Katherine M. May, Roberto Micalizio, and Stefano Tedeschi. ADOPT JaCaMo: Accountability-Driven Organization Programming Technique for JaCaMo. In *PRIMA 2017: Principles and Practice of Multi-Agent Systems, 20th Int. Conf.*, number 10621 in LNCS, pages 295–312. Springer, 2017.
6. Matteo Baldoni, Cristina Baroglio, Katherine M. May, Roberto Micalizio, and Stefano Tedeschi. An Information Model for Computing Accountabilities. In Chiara Ghedini, Bernardo Magnini, Andrea Passerini, and Paolo Traverso, editors, *Proc. of 17th International Conference of the Italian Association for Artificial Intelligence (AI*IA 2018)*, LNAI, Trento, Italy, 2018. Springer.
7. Guido Boella, Leendert W. N. van der Torre, and Harko Verhagen. Introduction to the special issue on normative multiagent systems. *J. of AAMAS*, 17(1):1–10, 2008.
8. Olivier Boissier, Rafael H. Bordini, Jomi F. Hübner, Alessandro Ricci, and Andrea Santi. Multi-agent oriented programming with JaCaMo. *Sci. Comput. Program.*, 78(6):747–761, 2013.
9. Cristiano Castelfranchi. Principles of Individual Social Action. In *Contemporary action theory: Social action*, volume 2, pages 163–192, Dordrecht, 1997. Kluwer.
10. Amit K. Chopra, Fabiano Dalpiaz, Fatma Basak Aydemir, Paolo Giorgini, John Mylopoulos, and Munindar P. Singh. Protos: Foundations for engineering innovative sociotechnical systems. In *IEEE 22nd Int. Requirements Engineering Conf.*, pages 53–62, 2014.
11. Amit K. Chopra and Munindar P. Singh. From social machines to social protocols: Software engineering foundations for sociotechnical systems. In *Proc. of the 25th Int. Conf. on WWW*, 2016.
12. Alan Colman and Jun Han. Operational management contracts for adaptive software organisation. In *Proceedings of the 2005 Australian conference on Software Engineering*, pages 170–179. IEEE Computer Society, 2005.

13. Daniel D. Corkill and Victor R. Lesser. The use of meta-level control for coordination in distributed problem solving network. In Alan Bundy, editor, *Proceedings of the 8th International Joint Conference on Artificial Intelligence (IJCAI'83)*, pages 748–756, Los Altos, CA, 1983. William Kaufmann.
14. Luciano R. Coutinho, Jaime S. Sichman, and Olivier Boissier. Modelling dimensions for agent organizations. In *Handbook of research on multi-agent systems: Semantics and dynamics of organizational models*, pages 18–50. IGI Global, 2009.
15. Antônio Carlos da Rocha Costa. Proposal for a notion of modularity in multiagent systems. In Fabiano Dalpiaz, Jürgen Dix, and M. Birna van Riemsdijk, editors, *Pre-proceedings of Engineering Multi-Agent Systems Workshop at AAMAS, Paris*, pages 21–40, 2014.
16. Mehdi Dastani, Nick AM Tinnemeier, and John-Jules Ch Meyer. A programming language for normative multi-agent systems. In *Handbook of Research on Multi-Agent Systems: semantics and dynamics of organizational models*, pages 397–417. IGI Global, 2009.
17. Virginia Dignum. *A model for organizational interaction: based on agents, founded in logic*. PhD thesis, Utrecht University, 2004. Published by SIKS.
18. Virginia Dignum. Handbook of research on multi-agent systems: Semantics and dynamics of organizational models, 2009.
19. Virginia Dignum, Javier Vázquez-Salceda, and Frank Dignum. OMNI: introducing social structure, norms and ontologies into agent organizations. In *Programming Multi-Agent Systems, Second International Workshop ProMAS, Selected Revised and Invited Papers*, volume 3346 of *Lecture Notes in Computer Science*, pages 181–198. Springer, 2004.
20. Emile Durkheim. *De la division du travail social*. PUF, 1893.
21. Marc Esteva, Juan-Antonio Rodríguez-Aguilar, Carles Sierra, Pere Garcia, and Josep L. Arcos. On the formal specification of electronic institutions. In Frank Dignum and Carles Sierra, editors, *Agent Mediated Electronic Commerce: The European AgentLink Perspective*, pages 126–147. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001.
22. Nicoletta Fornara, Francesco Viganò, Mario Verdicchio, and Marco Colombetti. Artificial institutions: a model of institutional reality for open multiagent systems. *Artificial Intelligence and Law*, 16(1):89–105, 2008.
23. Jomi F. Hübner, Olivier Boissier, Rosine Kitio, and Alessandro Ricci. Instrumenting multi-agent organisations with organisational artifacts and agents. *Autonomous Agents and Multi-Agent Systems*, 20(3):369–400, 5 2010.
24. Rosine Kitio Teussop. *Gestion de louverture au sein dorganisations multi-agents: une approche basée sur des artefacts organisationnels*. PhD thesis, Saint-Etienne, EMSE, 2011.
25. Elisa Marengo, Matteo Baldoni, Cristina Baroglio, Amit K. Chopra, Viviana Patti, and Munindar P. Singh. Commitments with regulations: reasoning about safety and control in REGULA. In *Proc. of the 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS)*, volume 2, pages 467–474, 2011.
26. Barbara S. Romzek and Melvin J. Dubnick. Accountability in the Public Sector: Lessons from the Challenger Tragedy. *Public Administration Review*, 47(3), 1987.
27. Barry R. Schlenker, Thomas W. Britt, John Pennington, Murphy Rodolfo, and Kevin Doherty. The triangle model of responsibility. *Psychological Review*, 101(4):632–652, October 1994.
28. Munindar P. Singh. An ontology for commitments in multiagent systems. *Artif. Intell. Law*, 7(1):97–113, 1999.
29. Munindar P. Singh. Distributed Enactment of Multiagent Workflows: Temporal Logic for Web Service Composition. In *The Second International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS 2003, July 14-18, 2003, Melbourne, Victoria, Australia, Proceedings*, pages 907–914. ACM, 2003.