Norms as a Basis for Governing Sociotechnical Systems

MUNINDAR P. SINGH, North Carolina State University

We understand a sociotechnical system as a multistakeholder cyber-physical system. We introduce *governance* as the administration of such a system by the stakeholders themselves. In this regard, governance is a peer-to-peer notion and contrasts with traditional management, which is a top-down hierarchical notion. Traditionally, there is no computational support for governance and it is achieved through out-of-band interactions among system administrators. Not surprisingly, traditional approaches simply do not scale up to large sociotechnical systems.

We develop an approach for governance based on a computational representation of norms in organizations. Our approach is motivated by the Ocean Observatory Initiative, a thirty-year $400 million project, which supports a variety of resources dealing with monitoring and studying the world's oceans. These resources include autonomous underwater vehicles, ocean gliders, buoys, and other instrumentation as well as more traditional computational resources. Our approach has the benefit of directly reflecting stakeholder needs and assuring stakeholders of the correctness of the resulting governance decisions while yielding adaptive resource allocation in the face of changes in both stakeholder needs and physical circumstances.

1. CHALLENGES IN SOCIOTECHNICAL SYSTEMS

A cyber-physical system comprises a number of computational and physical resources, usually in a specific social context. A major value of such systems is in expanding human and social capabilities in dealing with a complex environment. First-generation cyber-physical systems and research have been largely focused on low-level aspects such as sensors and effectors. Existing approaches assume that a single organization (e.g., a hospital or a disaster recovery team) owns or controls all the resources in question.

In contrast, our interest lies in *sociotechnical systems*, which we define as multistakeholder cyber-physical systems. Sociotechnical systems feature autonomous stakeholders whose interests are at best imperfectly aligned. Administering systems, which is difficult and expensive at the best of times, becomes complicated when multiple stakeholders are involved. Further, sociotechnical systems contend with complexity and change in *both* the cyber and the physical worlds. Traditional approaches are not only expensive but also preclude extracting the most value from the systems in question. We address the challenge of enabling stakeholders to administer or *(self-)govern* such systems in a manner that respects their autonomy. A particular benefit is adaptability in accommodating the exceptions and opportunities that arise in a complex environment.

Our participation in the recently launched Ocean Observatories Initiative (OOI) [Arrott et al. 2009], a paradigmatic sociotechnical system, has reinforced our motivation for addressing the above

challenges. The OOI facilitates the efforts of scientists and research institutions in acquiring, storing, analyzing, and sharing information from the world's oceans. Its stakeholders include oceanographers, educators, members of the public as well as research laboratories and universities. The stakeholders own and share resources such as Underwater Autonomous Vehicles (UAVs), buoys, ocean sensors, and research databases.

The OOI is designed to be operated for decades with initial funding for a thirty-year period. Consequently, we expect that nearly all implementation technologies deployed today will be obsolete within the lifetime of the system. Thus, not only must we accommodate changes in stakeholder needs, we cannot rely upon any specific technology to provide a stable notion of correctness. Further, the OOI is a large system: about $40 million or 10% of its budget is for IT. It is worth noting that the OOI itself would not own most resources involved in the collaborations it will help administer. The OOI is conceived of as a system with thousands of stakeholders, tens of thousands of physical resources such as ocean gliders, and potentially millions of cyber resources such as datasets. At those scales, adaptation is essential for administering resources according to the preferences of the stakeholders.

How can we accommodate stakeholder needs that are continually *changing*? How can multiple stakeholders function collaboratively in a sustainable, efficient manner? How can individual ownership and control be respected as autonomous parties interoperate? How can resources be added or dropped dynamically at runtime? How can dynamic coalitions be constructed and enacted to optimally share resources while entertaining challenges such the stakeholders' needs changing unexpectedly, as in an emergency? How may we accomplish all of the above *adaptations* over a wide of range of resource granularities and timescales?

1.1. Self-Governance: Interplay of Norms and Organizations

The above challenges come together in the problem of self-governance or *governance* for short. Briefly, governance is how autonomous entities administer themselves. Governance contrasts with traditional top-down management, which presumes authority (superior to subordinate) relationships. In the systems of interest, the collaborating parties are autonomous peers and none has authority over the others. Today, governance is carried out "by phone call"—by ad hoc negotiations among humans. Such manual techniques can work in small settings where a few resources need to be shared over long timescales. In contrast, the (pervasive) sociotechnical systems of interest involve large numbers of resources and require decision making at fast timescales. Manual negotiations would simply not scale up to such settings. Governance as a research theme is coming of age. For instance, Brazier et al. [2012] report on a working group of a recent Dagstuhl seminar, and emphasize the importance of governance.

We observe that from the perspective of governance, the stakeholders of a sociotechnical system are themselves *participants*. Recognizing the autonomy of the participants of sociotechnical systems, we observe that we cannot prescribe a decision-making strategy for each participant. Instead, each system can prescribe its rules of encounter via a set of norms. Informally, a *norm* characterizes sound or "normal" interactions among the participants of a social group, reflecting their mutual expectations. We emphasize *interactions* for governance because we have no interest in restricting a participant's private behavior that has no effect on other participants. Two examples of norms in a scientific setting are putting an instrument in a power-save mode at the end of an experiment and closing unneeded datastreams from sensors. We are not concerned here with how norms arise, whether through top-down legislation or bottom-up conventions emerging from norms implicit in participant strategies [Savarimuthu et al. 2009]. Excellent work by others on policies and modeling, e.g., [Johnson et al. 2011] on interdependent "coactive" participants, is addressing some of these challenges. We further restrict ourselves to norms that have some contractual force, so that their satisfaction or violation is significant.

Based on the above intuition, we formalize a sociotechnical system as an *organization* that involves two or more roles, each specified in terms of the norms applying to it. To this end, we formalize norms not as amorphous properties of the "system"—whatever that might be—but as directed

normative relationships between participants in the context of an organization. Our formal model reflects this *essential duality* of organizations and norms: an organization is defined via norms and a norm is defined in an organization. Importantly, our approach accommodates *open* settings where a party may live and act outside the scope of a sociotechnical system while remaining subject to the norms defined in the system.

1.2. Principles of Adaptation in Sociotechnical Systems

Our approach seeks to engineer a sociotechnical system in such a manner as to support adaptation, both (1) in its configuration (and implementation) and (2) in its enactment realized through the interactions of its participants. The twin challenges of ensuring adaptation and achieving rigor lead us to adopt the following main principles.

— *Centrality of Norms:* A normative, as opposed to an operational, characterization of acceptable interactions is minimally constraining and thus essential for capturing the "invariants" of a long-lived system (the OOI's lifetime being decades).
— *Autonomy and Policies:* The participants are autonomous, though subject to applicable norms. Each participant applies its internal *policies* to decide how to interact given the norms; its policies reflect its autonomy (the OOI's members being autonomous).

The foregoing emphasis on autonomy and adaptation suggests that our computational system must incorporate *agents*: active computational entities that represent individual participants and organizations. The internal implementations of the agents are not relevant to governance, but their interactions (subject to norms) are. The agents are only partially regimented. Where appropriate, we prefer to develop agents that respect the applicable norms, but recognizing the autonomy of the agents means that any agent may violate a norm. Therefore, norms provide a rigorous basis for *coherence*, which we view as a relaxed notion of correctness that accommodates restoring a "good" state after a violation.

1.3. Contributions and Claims

We develop a novel approach for governance that is computationally realized and deals well with complexity and dynamism. Our contributions are two-fold.

— A *formal model* for governance that incorporates a rich set of normative clauses promoting adaptability and reuse. This model provides a natural mapping to computations and can be realized generatively. It also supports useful kinds of analyses of particular organizations and norms.
— An *architecture* that realizes the above model and helps demonstrate our approach on significant use cases arising in the OOI setting.

We claim that our model and architecture (1) enable the construction of a *flexible* sociotechnical system that can naturally (2) adapt in its *configuration*, thereby accommodating changes in stakeholder needs by reconstituting its rules of encounter and (3) adapt in its *operations*, thereby accommodating the dynamics of a sociotechnical system.

For simplicity and brevity, we limit the scope of this paper to the aspects of the model and architecture that specifically focus on governance. In particular, we elide the substantial efforts within the OOI project on ontologies, resource models, instrumentation, data management, and an advanced cloud-based computing infrastructure.

Section 2 introduces important governance scenarios from OOI with an emphasis on adaptation. Section 3 describes our formal model for a sociotechnical system. Section 4 shows how to enact specifications in our formal model. Section 5 evaluates our approach with respect to the scenarios of Section 2. Section 6 discusses some general themes along with interesting directions for future research.

2. APPLICATION SCENARIOS AND VARIETIES OF ADAPTATION

Let us consider some simple OOI scenarios to convey our conception of a sociotechnical system being put to use, and illustrate the tension between regimentation and adaptability that is an essential characteristic of sociotechnical systems. These scenarios help distinguish our work from traditional approaches.

2.1. Collaboration through Resource Sharing

The stakeholders of OOI include a broad range of users, such as researchers, educators, students, and enthusiasts, with varying interests and expectations. Say, a teacher from a school near Chesapeake Bay would like to have his students conduct a project that exposes them to real-world data from their local environment. The teacher discovers an OOI member willing to share data from her salinity sensors in the Bay. Elsewhere, a researcher plans a comparative study of seasonal variations in salinity in Chesapeake Bay and Monterey Bay and its effect on algae. Although both the teacher and the researcher seek collaborations, the two engagements would differ in duration, exclusivity, and permissions over data use. Configurational adaptation: The researcher observes that scientific and educational engagements only account for 40% of her instrument's capacity. To maximize her instrument's value, she begins to participate in a community of enthusiasts formed of members of the public. Operational adaptation: Because of an oil spill, suddenly numerous requests arrive from researchers and enthusiasts. The instrument owner preemptively prioritizes new requests from researchers over ongoing engagements with enthusiasts. Accordingly, she pulls back her instrument from the enthusiasts' community but lets the enthusiasts continue to access a datastream from the instrument.

2.2. Affiliation

Research institutions and laboratories are central to the scientific effort and are first-class participants in OOI. Recognizing the benefits of sharing ocean instruments and curated datasets on a regular basis, the Chesapeake and Monterey laboratories, become *affiliates* of each other. Hence, the research staff of one laboratory can access resources from the other. But, each laboratory would keep some data and analytical tools private, e.g., because such data and tools are part of an ongoing study whose results the laboratory wishes to be the first to publish. Configurational adaptation: The laboratories expand their affiliation to include their respective zoological databases and students on a reciprocal basis. Operational adaptation: The Monterey laboratory learns that the Chesapeake laboratory has hired a researcher who was involved in some controversy about publishing premature results. At the Monterey laboratory's behest, the two modify their affiliation to forbid unilateral publishing of results arising from collaborative studies.

2.3. Sanction

Individual collaborators or laboratory affiliates agree to specific terms, some of which restrict their actions. For example, a collaborator may be forbidden from changing the firmware on an instrument that is temporarily checked out to him or from externally publishing the results of a joint experiment. The participants in OOI are autonomous, meaning that they have an existence outside of the OOI system. Thus they can potentially violate the terms of an agreement through actions that OOI cannot prevent, e.g., because they have physical control of an instrument or use an external web site to publish some data. However, such breaches may eventually be detected by the concerned parties, who can complain to the OOI, viewed as an authority. In such cases, OOI would subject the responsible party, if identified and found culpable, to specified sanctions, such as having to replace the instrument or issue a public retraction. OOI could cancel the account of a malfeasant participant. Configurational adaptation: The above engagement may be modified to allow revealing the data externally, though only as necessary to fulfill a research sponsor's deliverable requirements. Operational adaptation: When a severe algae bloom occurs hidden beneath the surface of the Bay, a researcher unilaterally reports it to the press. The sanctioning process absolves the researcher be-

cause of extenuating circumstances: in this case, the researcher's violation was necessary to protect the health and safety of the public.

2.4. Requirements Induced from the Foregoing Cases

The above scenarios indicate the need for flexibility in configuring engagements among individuals and institutions, because no static solution would accommodate the dynamic nature of stakeholder needs. For example, a researcher must be able to specify her requirements for sharing her ocean instruments. Even though such requirements would fall into a few typical patterns, the patterns of best practices themselves would change over the course of years, if not of decades. Therefore, instead of legislating fixed policies, we must provide a flexible means to govern collaborations that naturally supports adaptation while ensuring a rigorous, though relaxed, notion of correctness. In essence, we must lift the architecture from considerations of control or data flow among software components to considerations of norms among autonomous participants. In particular, given the autonomy of the participants, we cannot assume that no norm will be violated. This is because it would often be impossible to regiment all interactions of the participants. Thus each participant should potentially have recourse in case one of the other participants violates a norm, even if it does so outside the operational scope of the OOI.

Singh et al. [2009] identify three main elements of a service engagement: *transactional* or what the engagement accomplishes for its participants; *structural* or how the engagement is organized; and, *contextual* or the broader rules of encounter to which the engagement is subject. We adopt the idea of Desai et al. [2009] to classify changes in requirements in terms of the above three elements. Whereas Desai et al. consider cross-organizational business processes, here we consider norms broadly and consider more subtle situations of how the engagements in question are arranged. Viewed in the above light, the adaptations in the resource usage, affiliation, and sanctioning scenarios correspond to the transactional, structural, and contextual elements, respectively.

3. MODELING A SOCIOTECHNICAL SYSTEM

The foregoing use cases suggest two main requirements: the need for adaptivity and the need for rigor. On the one hand, the autonomy of the participants and the fact that they carry out long-lived collaborations across institutional scopes means that we must accommodate change. On the other hand, the same features mean that we must do so in a rigorous manner because otherwise it would be impossible to guarantee appropriate outcomes in such a complex setting. We develop a normative approach to address the above challenges. The norms are founded upon the idea of stakeholders being modeled as autonomous principals, who are represented computationally as agents that carry out loosely coupled interactions.

The plan of our technical development is as follows. We begin from a general organizational model for sociotechnical systems. We further refine this model to introduce a small set of norm types. From the organization model, we develop a conceptual model of a vocabulary in which to express norms and thus to specify an organization for a sociotechnical system. To enact such a system, we introduce an agent architecture based on policies, also expressed using the above vocabulary and additional relevant predicates.

3.1. Conceptual Model of a Sociotechnical System

Figure 1 illustrates the conceptual model that underlies our approach for governance. The notion of an Org is crucial in formulating interactions in terms of norms. Indeed, in our approach, all norms arise with an Org as a backdrop. In simple terms, an Org is recursively constructed: its members are principals that could themselves be Orgs. A principal may be a member of more than one Org: thus Orgs can have overlapping memberships. For simplicity, we assume that the membership relation between Orgs and principals is well-founded as in our setting. For example, two Orgs are not members of each other,.

Principals communicate and collaborate within the scope of an Org of which they are members. The most important purpose of an Org in our architecture is that an Org systematizes the norms
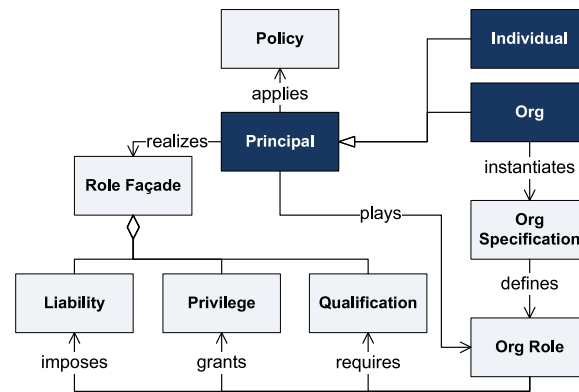
Fig. 1.   Simple conceptual model geared toward governance. (Here, boxes with white text indicate active entities; those with black text indicate representations.) This model is centered on the notion of *principal*, which corresponds to a participant in a system. A principal may be an *individual* (a researcher or even a laboratory if understood as an atom) or an *Org* (a structured entity such as a resource sharing community, an affiliation, or even the OOI itself). In either case, a principal is a locus of autonomy and potentially chooses its own *policies*, which reflect its autonomy. For example, a researcher may decide through her internal policies whether to contribute usage of her ocean glider to a community of enthusiasts. A laboratory may decide whether to admit an educator to access a datastream from a sensor installed on a buoy.

among its members and potentially provides an authority to which the members may complain regarding norm violations by others. An Org may apply any appropriate sanctions on any of its members; such sanctions typically include canceling the membership of, or further escalating a complaint against, a principal it judges malfeasant.

Orgs are finely structured through the notion of a *role*, which codifies a set of related interactions that a member of an Org may enact. To be a member of an Org means to play at least one role in that Org. In principle, a principal may concurrently play more than one role in the same Org or in different Orgs. However, some roles may limit such flexibility, e.g., to ensure a separation of duties. Each Org is specified by defining the rules of encounter for each of its roles. Together these rules of encounter may be understood as a multiparty contract. However, the elements that concern an individual role are most relevant to a principal who plays that role. For each role, we collect these elements into what we term the *façade* of that role. Each façade comprises three major components.

> *Qualification.* A prerequisite or eligibility requirement for a principal to play the specified role. Example: Only a credentialed and currently employed teacher may participate as an *educator* in a continuing education Org for school teachers must be.
> *Privilege.* A liberty, broadly understood, accorded to a principal who plays the specified role. Example: A teacher who is admitted as an *educator* to a summer camp Org is authorized to access all datasets available within the camp and is empowered to further admit a student as a *pupil* to the camp.
> *Liability.* A demand imposed on a principal who plays the specified role. Example: Continuing with the above scenario, a teacher who becomes an *educator* must entertain help requests from a student who is a *pupil*. Likewise, a *pupil* who introduces a virus into the camp's computers would risk sanctions, including possible expulsion.

Importantly, privileges and liabilities map naturally to the normative relationships that the principals enter into, some of which accord liberties and some of which impose demands on the principal who adopts the specified role. Adopting a role in an Org is thus a natural path for a principal to enter into norms with other principals. Moreover, principals may form additional norms through explicit negotiation. However, even such negotiated norms are governed by the norms that arise from the roles in an Org. For example, a teacher as an *educator* gains access to datasets but not to instruments. To be able to use an instrument owned by a scientist, a teacher may agree on additional

terms and conditions, such as that he would not reboot the instrument. Such agreements would arise in the scope of the same Org, and their violation could have consequences such as the imposition of sanctions defined in the *educator* façade.

The model of Figure 1 posits that an Org is a principal and can thus participate in another Org by playing a role therein. We now further posit that an Org *qua* principal may also interact with and enter into norms with its own members. For example, when researcher Ryzard joins OOI, not only is he subject to OOI's norms but he may also expect OOI to keep his private information safe. We capture the above intuition by postulating a distinct *self* role for each Org. In any Org, this role is played by exactly one principal, namely, the Org itself. Further, this role is instantiated simultaneously with the Org coming into being. In conceptual terms, an Org as *self* interacts with all its members, handles their requests to discover other members and resources, entertains their complaints about each other, adjudicates on the norms between them (in its capacity as the context for such norms), and enforces any applicable sanctions upon them.

3.2. Normative Concepts

Based on an analysis of sociotechnical systems, especially the OOI, we postulate the following normative concepts as the key elements of a role façade. Because these concepts are familiar to people, we can use them to model relevant situations in a way that stakeholders can easily comprehend.

When employed as a design construct, a norm codifies desired properties of interactions among principals. In simple terms, a norm captures the sense of how an interaction *ought* to proceed and thus regulates the interactions of the principals involved. By providing a rich set of constructs with which to express the norms, we enable encoding the essential properties of interactions in a manner that is flexible (any enactment that satisfies the norms is acceptable) yet rigorous (there is a precise computational notion of when a norm is violated). The flexibility helps ensure correctness while supporting adaptation in configuration (to accommodate changes in stakeholder requirements) and during enactment by the principals. During enactment, the norms progress because of the principals' interactions: e.g., they may be activated, satisfied, or violated. A snapshot of the norms taken together constitutes the *normative state* of the sociotechnical system. Figure 2 shows how our norm representation generalizes over the representations of Singh [1999; 2008].
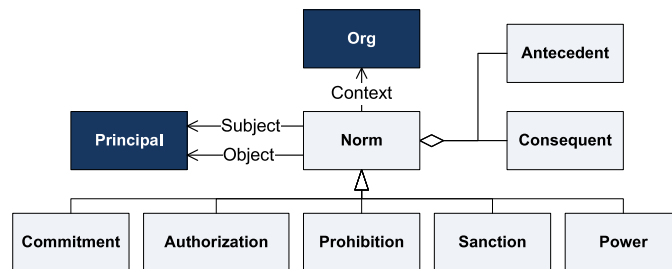

Fig. 2. The unified logical form of a norm. Each norm involves a *subject* (the principal on whom the norm is focused), an *object* (the principal with respect to whom the norm arises), a *context* (the Org within whose scope the norm arises), an *antecedent* (expressing the conditions under which the norm is fully activated and brought into force), and a *consequent* (expressing the conditions under which the norm is fully satisfied and deactivated). In many practical cases, we set the antecedent to true to indicate an unconditional norm. Further, the context could be distinct from or the same as the subject or the object, whereas the subject and object are always distinct.

Each norm is *directed* from a subject to an object thus making clear upon whom it applies, and enhancing modularity by supporting multiple normative patterns such as reciprocal commitments and prohibitions with or without a sanction for violation. We place norms in an organizational *context* and support their *manipulation*. In this way, we combine the benefits of (1) a precise declarative characterization of normative state with (2) a clear statement of institutional actions. Section 4 shows

how to operationalize norms in a way that applies naturally to sociotechnical systems. We consider five types of norms.

Commitment. Within the scope of the organizational context [Singh et al. 2009], the subject (i.e., debtor) commits *to* the object (i.e., creditor) that if the antecedent holds, the debtor will bring about the consequent. When the consequent holds, the commitment is satisfied and deactivated. Example: A researcher who borrows an instrument for a study commits to returning it within one hour of being requested to do so.

Authorization. With respect to the given context, the object authorizes (i.e., permits) the subject to bring about the consequent when the antecedent holds. Bringing about the consequent if the antecedent remains false is a violation. Example: An instrument owner authorizes a colleague to use the instrument between 7:00PM and 9:00PM today.

Prohibition. With respect to the given context, the object prohibits (i.e., forbids) the subject from bringing about the consequent provided the antecedent holds. Bringing about the consequent if the antecedent holds is a violation. Examples: An instrument owner prohibits a borrower from changing the firmware on the instrument. A dataset curator prohibits a reader from publishing any of the data on an external web site.

Sanction. With respect to the given context, the object would sanction (i.e., punish) the subject by bringing about the consequent provided the antecedent holds. Examples: An instrument owner would sanction a borrower who illicitly changes the firmware on a borrowed instrument by giving the borrower a poor rating. A dataset curator would sanction a reader who publishes any of the data externally by complaining to the Org. The resource sharing Org would sanction a reader who publishes any of the data externally by ejecting him from the Org.

Power. With respect to the given context, when the antecedent holds, the object empowers the subject to bring about the consequent at will. Loosely following Hohfeld [1919], we treat a power as the ability to alter the norms between two or more principals, usually those playing specific roles. We follow Jones and Sergot [1996] in treating power as an *institutional* construct, meaning that a power exemplifies the so-called *counts-as* relation between a low-level (physical) ability and a high-level (institutional) action. This is a particular form of making a norm concrete [Aldewereld et al. 2010]. Importantly, a principal may be empowered to do something but not be authorized to do so. Our setting supports the simplification that the physical action is a communication: thus when the antecedent holds, the subject need only "say so" to bring about the consequent. Examples: The Chesapeake Bay Org is empowered to admit or eject its members by declaring so. An instrument owner is empowered to contribute her instrument to a resource sharing Org, also by declaring so. A system administrator is empowered to admit new people into OOI by creating their accounts, but is—crucially—prohibited from creating accounts (and thus effectively admitting members) without approval from the membership department. However, because the administrator has the power, her creation of a new account will succeed, though it might later be deemed illicit and revoked, and the administrator sanctioned for exercising the power illicitly. Here, the power is misused and the prohibition is violated.

Table I shows how the norms map to the components of a role façade. It reflects the intuition that liability and privilege are two faces of the same coin: a liability for a principal in one role is a privilege for a principal in the "counter" role. It is quite intuitive that the subject of a commitment, prohibition, or sanction is a liability since it can only lead to the subject investing effort or having its freedom curtailed or suffering a penalty. In the same spirit, being the subject of an authorization or a power is a privilege since the subject obtains an option to perform additional actions without being required to do so. As remarked above, whenever the subject has a privilege, the object has a liability and vice versa. Note that qualifications do not feature in this table because they are formed of the credentials of the principals, such as their participation in specified Orgs in specified roles (see Figure 3 below).

As Section 2.3 illustrates, a sociotechnical system is inherently open in that its autonomous participants have an external existence. In general, each Org is open and cannot regiment all the actions

Table I. Mapping normative concepts to role façades.

Normative Concept	Subject's Façade	Object's Façade
Commitment	Liability	Privilege
Authorization	Privilege	Liability
Power	Privilege	Liability
Prohibition	Liability	Privilege
Sanction	Liability	Privilege

of its participants. We address the above challenge through a simple approach that consists of two parts: (1) representing the appropriate norms for each Org, as below, and (2) enacting the norms appropriately, as in Section 4.

In general, a principal ought to perform only those actions for which it is authorized and not prohibited. We distinguish interactions that occur within the scope of an Org from those that occur without. We adopt the following "design" pattern that simplifies modeling and enactment. We treat authorizations as applying exclusively to the internal interactions and prohibitions as applying exclusively to the external interactions. The internal interactions are architecturally regimented based on authorizations (analogous to access control) and therefore never occur unless authorized. The external interactions are subject to prohibitions but cannot be architecturally regimented. Therefore, for each prohibition we need to specify a sanction in case it is violated, but not so for any of the authorizations.

Figure 3 summarizes our (extensible) vocabulary for antecedents and consequents of norms. This vocabulary provides the predicates we use to state norms, formulate communications, and state agent policies, the last of which are discussed in Section 4. The foregoing Org and norm models provide a principled way to organize and, if necessary, extend this vocabulary. Specifically, in an Org description, interagent communication, or agent implementation, we can refer to actions such as admitting or ejecting a principal (Org participation), contributing a resource to an Org (Org resource registration), modifying a norm (norm operation), controlling an underwater vehicle (resource capability), and submitting a request (communicating). We can also refer to relevant elements of the state such as whether a principal plays a specified role in a specified Org (participation stative), whether a sensor can provide salinity information (resource stative), and whether a commitment has been satisfied (norm state). Specifications of Orgs and of the policies by which agents participate in Orgs all rely upon this extensible vocabulary.

Notice that the state of a norm can be referenced from another norm. For example, consider a commitment $c_1 = \mathsf{C}(d, c, o, p, q)$ in the form introduced in Figure 2. Then we can express a commitment from Org context o to creditor c that if c_1 is violated, Org o will compensate creditor c by ensuring r as $c_2 = \mathsf{C}(o, c, \mathsf{OOI}, \mathsf{vio}(\mathsf{C}(d, c, o, p, q)), r)$ [Singh et al. 2009]. Our approach does not support self-referential or mutually referring norms.

Governance involves modeling not only the norms but also how the norms are *manipulated*. For example, when a school teacher joins a resource sharing Org as a *user*, he acquires the norms specified in the *user* façade. We generalize Singh's [1999] commitment manipulation operations for all norms. Any norm may be *created* (directly by the liable party or via a causal chain leading back to the creation of another norm by the same party), *discharged* (satisfied by the liable party), *canceled* (terminated by the liable party, though at risk of violation), *released* (terminated by the privileged party, because it does not care), *delegated* (by the liable party to a new liable party), or *assigned* (by the privileged party to a new privileged party). Notice that qualifications are treated merely as credentials even when they happen to refer to privileges in other Orgs.

Norms are expressed as schemas that involve roles and parameterized expressions, the latter featuring as antecedents and consequents. During enactment, the norms that arise would be instantiated with agents playing roles and constants in place of parameters. Such instance norms progress based on the operations on norms as well as events in the Org. Figure 4 shows the life cycle of a norm in terms of its key states and transitions. Taking on a role creates the associated norms; exiting a role terminates as appropriate some (but not necessarily all) of the associated norms that are active

Action → Property ← Stative

Organizational Capability | Resource Capability | Communicative Act | Participation | Norm State

Participation Capability | Resource Registration Capability | Norm Operation | Inform, Request, Query, ... | Resource

Eject, Admit, ... | Contribute, Withdraw, ... | Create, Delegate, Assign, ... | Member, Registrand, Contributed, Owns, ... | Capability, In State, ... | Created, Satisfied, Violated, ...
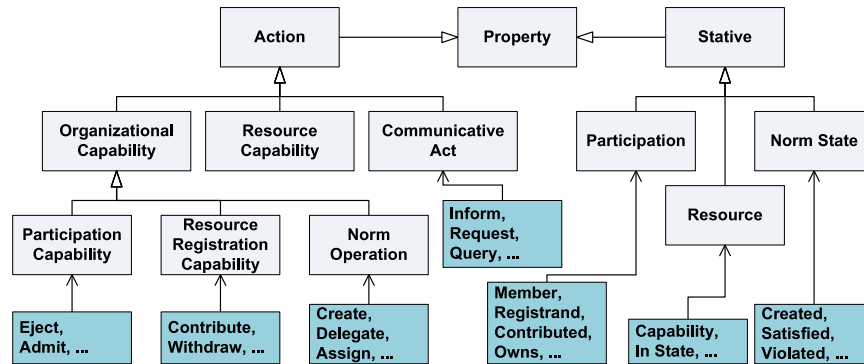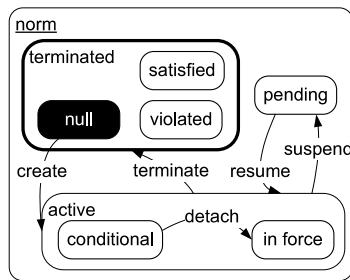
Fig. 3. A fragment of the vocabulary used in expressions (antecedents and consequents). Each property is applied with the requisite number of arguments. The unshaded boxes are general classes corresponding to our model; the shaded boxes are important examples, some specific to OOI. Resource capabilities would be highly domain specific. The actions bring about changes in state and the statives help refer to relevant states from within norms. Nominally, the antecedents of norms would involve statives and their consequents actions.

or pending; a sanction may create an additional commitment to pay a penalty and cancel current authorizations to use any instruments within the Org; and so on. The table in Figure 4 specifies the substate of a terminated norm. For example, a commitment enters vio if it terminates when its antecedent is true and consequent false.

Further considerations of semantics and logical properties are out of our present scope. An instance of such a property is that we would consider only the maximally strongest norms of each type in determining their state. For instance, we would disregard apparent violations by an authorization when another authorization is satisfied. For example, with a fixed subject, object, and context, consider two authorizations: if r then u and if $r \vee s$ then u. The latter is more general and would be satisfied when , $\neg r$, s, and u hold, in which case the first authorization is violated. However, such a violation is irrelevant because the more general authorization is satisfied.

norm

terminated: satisfied, null, violated

pending

create | terminate | resume | suspend

active | detach

conditional | in force

If terminated in		Then				
ant	con	Com	Aut	Pro	San	Pow
false	false	null	null	null	null	null
false	true	sat	vio	null	null	null
true	false	vio	null	sat	null	vio
true	true	sat	sat	vio	sat	sat

Computing the substate of a terminated norm (abbreviated to three letters). In the case of a power, a vio occurs upon the failure of an attempt to bring about the consequent.

Fig. 4. The unified life cycle for a norm as expressed in the UML state diagram notation. (A nested state indicates that the norm could be in any of its substates.) Here null is the initial state and terminated is the final state. A norm is active when created: it is in force when its antecedent holds and conditional otherwise. A norm may be suspended, e.g., when a subject of a commitment delegates it, and resumed, e.g., when the delegate fails and the subject activates the commitment again. A norm is terminated because its subject, object, or context explicitly deactivate it or because of timeouts. The table specifies the appropriate terminating substate of a norm depending whether its antecedent and consequent are true or false therein. In other words, when a norm enters the terminated state, there is a choice of three substates for it to enter. The table provides the basis for making that choice on purely logical grounds. For example, a prohibition is considered violated if terminates with both its antecedent and its consequent true.

3.3. Outline of a Modeling Methodology and Specification Snippets

To understand our modeling methodology, consider the resource sharing scenario of Section 2.1 again. A resource sharing Org admits principals who may play one or both of *user* and *owner* roles. Any principal who owns resources may accept the *owner* façade and thus enroll in the Org. An *owner* may contribute a resource to the Org: the Org would list it in a directory. Similarly, a *user* may search the resource directory maintained by the Org, request access to, and use resources contributed by others. A *user* and an *owner* may negotiate usage terms, possibly creating additional norms. The negotiation may be as simple as an *owner* requiring a *user* to accept a disclaimer about the quality of the resource. An *owner* may withdraw a resource it previously contributed, but only when no *user* is actively using a resource. Further, a *user* may not share a resource obtained from this Org with any entity external to the Org. However, if the *user* wishes to share a resource externally, the Org cannot prevent it. Therefore, we express a prohibition against external sharing along with a sanction of possibly ejecting violators from the Org.

Applying our methodology on this scenario yields a specification of the Org in our formal language. This formal language makes the foregoing model concrete. Specifically, we define each Org in terms of its roles, each role in terms of its façade captured via qualifications and norms. In writing these elements, we use the vocabulary of Figure 3. For brevity, we embed some illustrative snippets of the specification below (a leading question mark indicates a variable).

— *Identify the roles in the scenario: user* and *owner* as well as *self* (needed for each Org).
— *Identify the interactions:* A principal interacts with the OOI Org to discover an Org for accessing data about the water chemistry of Chesapeake Bay. The principal discovers resources contributed by the members of this Org. Alternatively, or in addition, the principal may contribute resources to the Org. The foregoing yields interactions for discovering, negotiating for, using, contributing, and withdrawing resources.
— *Identify resource capabilities:* A glider may be dived, surfaced, recharged, and read.

```
Capability : Resource (? aResource , DepthControl )
```

— *Identify the façade of each role:* The *user* and *owner* façades include the following.
  — Qualifications: A *user* must be a member of OOI.

```
Qualification OOI: Public (? user )
```

  — Privileges: An *owner* is empowered to contribute or withdraw a resource. As Section 3.2 explains, we only need write the consequent of the power (in the Then part); the antecedent implicitly conjoins the proposition that the subject says the Then part. An *owner* may only contribute a resource that she owns and may withdraw a resource that she owns only when it is not currently in use. An *owner* may authorize a *user* to apply a resource capability.

```
Power {
  If SharableResource (? resource )
   AND Owns (? owner , ? resource )
   AND Supports (? resource , ? capability )
  Then Contribute (? owner , self , ? resource , ? capability )}

Power {
  If Contributed ( self , ? resource , ? capability )
   AND Owns (? owner , ? resource )
  Then Withdraw (? owner , self , ? resource , ? capability )}

Authorization {
  If Contributed ( self , ? resource , ? capability )
   AND NOT inState ( inUse , ? resource , ? capability )
  ThenMay Withdraw (? owner , self , ? resource , ? capability )}

Power {
  If Contributed ( self , ? resource , ? capability )
```

```
  AND Owns(?owner, ?resource)
  AND Participant(self, ?user, ?role)
 Then Authorize(?owner, ?user, Apply(?role, ?resource, ?capability))}
```

In the above example, the *owner* is the subject and the resource sharing Org is the object because the *owner* is respectively empowered or authorized by the Org. Notice that the last power expression above includes an authorization. In that authorization, the *user* is the subject and the *owner* is the object. That is, the *owner* has the power (i.e., a privilege) to grant an authorization, i.e., to create a liability upon itself.

Another important case is of the distinguished *self* role that we discussed above. The Org (viewed as its *self* role) is empowered to admit members into the Org. Thus, *self* is the subject and the OOI (viewed as the Org under whose aegis the resource sharing Org exists) is the object.

— Liabilities: A *user* may not externally share a capability on a resource accessed through this Org. A *user* who violates the above prohibition is subject to the sanction of being potentially ejected from the Org. Here, we use an expanded syntax to refer to the prohibition from the sanction; we could equally well use an abbreviated syntax that referred to the prohibition without repeating its entire definition. In these examples, the *user* is the subject and the Org (i.e., its *self* role) is the object.

```
Prohibition {
 If NOT Participant(self, ?outsider, ?norole)
 AND Contributed(self, ?resource, ?capability)
 ThenMayNot Share(?user, ?outsider, ?resource, ?capability)}

Sanction {
 If Violated
  (Prohibition {
   If NOT Participant(self, ?outsider, ?norole)
   AND Contributed(self, ?resource, ?capability)
   ThenMayNot Share(?user, ?outsider, ?resource, ?capability)})
 Then Eject(self, self, OOI:member, ?user)}
```

— *Validate the set of norms:* No principal should be prohibited from satisfying a commitment. A sanction must be applied by a principal who possesses the requisite power and authorizations. For example, an aggrieved principal may sanction a malfeasant principal by escalating its dispute to the Org, which upon verifying the violation would impose its own sanctions on the malfeasant principal.

Formally, given the design pattern introduced above, we need an authorization for every power. Therefore, the following permissive authorization is automatically generated for a power to bring about P, for which no other authorization is specified.

```
Authorization {
 If true
 ThenMay P}
```

Our language supports role inheritance so that one role may extend another role. This enhances reusability. Specifically, *owner* extends *user* since it grants additional privileges and imposes additional liabilities.

4. ENACTING A SOCIOTECHNICAL SYSTEM

Notice that the above model helps us specify an Org. It describes how the roles of the Org would interact. However, the above model must be supplemented by specifications of the executing entities, i.e., agents, in order to enact the system. Because we understand sociotechnical systems as involving autonomous participants, an essential requirement is that we enact such a system in a conceptually decentralized manner.

To explain our contribution, we describe a simple approach that generalizes Desai et al.'s [2005] approach for commitments. An *agent* is a computational surrogate of a principal. An agent is not autonomous with respect to its principal, but is autonomous as viewed from the perspective of other agents. In deciding how to interact with other agents, an agent applies its internal *policies*, presumably based on its principal's preferences. The policies of an agent capture its decision making and thus reflect the autonomy of the agent (and of its principal). The policies are kept internal, i.e., private, to encapsulate its internal reasoning and thus to promote heterogeneity.

Each principal's agent helps with the bookkeeping of the norms in which it features as subject, object, or context. The agent helps determine if the principal itself is complying with the applicable norms and, equally importantly, if the principals with whom it deals are complying as well. The agent maintains its local view of the normative state by continually updating the relevant norms. We can thus address the following challenges: (1) developing an agent so that it respects the façades of the roles its designer would have it play; (2) having an agent judge if an interaction complies with the specified norms; and (3) during enactment, having an agent compute what actions it ought or ought not to perform.

4.1. Computing with Rules

In architectural terms, our approach is neutral as to whether an agent is implemented in a more or a less restrictive manner, ranging from traditional software to a general-purpose planner. As a practical matter, we adopt a rule-based approach because it offers a happy middle between flexibility and ease of implementation. Note that any domain-specific reasoning could be realized through a traditional imperative language even though we account for norms through a rule-based language. To this end, we model each agent as maintaining a *belief store*, which represents the agent's local view. An agent acts according to its beliefs, but norms are inherently interactive and compliance in general is not based on what an agent believes but solely on how it interacts with others. Hence, we obtain a design requirement to ensure both that agents have true beliefs and can reason properly from them.

An agent updates its belief store by asserting or deleting beliefs as it communicates. We capture actions on resources as messages sent and observations from the environment as messages received. The beliefs residing in an agent's belief store represent the current snapshot of the physical state of the system, e.g., that a glider is broken or that a network connection to a buoy has a throughput of 2kbps. We separate out elements of the normative state, such as (1) that the agent has an active commitment to report the failure of the glider to the (agent of the) owner of the glider and (2) that its commitment is pending because it was delegated to another agent. The beliefs occur as propositions within the antecedents and consequents of a norm. Each agent ideally tracks each norm in which its principal features, whether as subject, object, or context. Potentially, any action that an agent chooses to perform or omit may have repercussions on the satisfaction or violation of its norms: in some cases immediately and in other cases a long time into the future. Therefore, an agent may evaluate and filter its options with respect to the norms it tracks.

Figure 5 illustrates our reference agent architecture. The decision maker attempts actions. The normative filter checks all of the agent's attempted actions for proper authorization and forwards along exactly those that it judges to be in compliance with the applicable norms. The communicator receives and sends messages, thereby applying the agent's attempted action if approved by the filter. In either case, it updates the beliefs accordingly.

4.2. Mapping Norms to Generative Rules

We now discuss how we systematically map a role façade to an enactable agent specification. The agents, being autonomous, apply their internal policies. However, each role that an agent plays constrains it based on the role's façade. Since we conceive of sociotechnical systems in which the agents are broadly cooperative though not fully trusted, we propose a straightforward means by which we can ensure that an agent complies with its norms, assuming its principal wishes it to. This involves placing some regimentation into the computational system as a way of ensuring that each
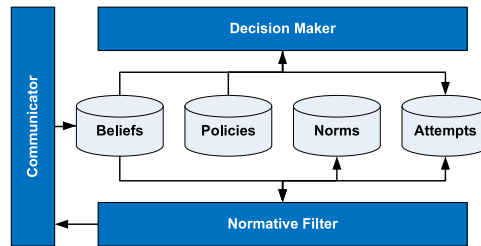
Fig. 5.   A simplified representation of our agent architecture. An agent has four main stores (for beliefs, policies, norms, and attempts) and three active modules. The communicator reports incoming and outgoing messages to the belief store. The decision maker applies its policies given the beliefs to compute possible actions, which it attempts. The normative filter maintains the agent's norm store. In particular, it adjudicates on the suitability of the attempted actions and determines whether other agents are complying with their norms that are of relevance to the agent. The agent acquires or modifies norms, including when it adopts or exits any role.

agent respects its authorizations. However, we leave open the possibility of an agent not complying with an applicable norm, such as some prohibitions, of which it is the subject.

Since the commitments where an agent is the subject (i.e., debtor) specify what conditions it must bring about, we use the commitments to structure the decision maker component of the above architecture. A commitment maps to the following forward-chaining rule template for its subject. Here the variables in the antecedent are bound when the commitment becomes in force, i.e., detached, and additional variables needed in the consequent are bound through the agent's policy. The way to interpret this template is that the antecedent would involve one or more variables, the policy may refer to one or more of the variables in the antecedent and would introduce zero or more additional variables, and all variables that occur in the consequent must occur in the antecedent or policy.

```
If  Antecedent (?x ?y)
 AND Policy (?x ?z)
Then Attempt (Consequent (?x ?y ?z))
```

Each policy is itself captured through one or more backward-chaining rules accounting for how the programmer wishes the agent to reason in this case. As the above snippet shows, the policy would involve overlapping sets of variables with the antecedent and consequent and provide any variable bindings needed in the consequent that are not set in the antecedent. In general, each commitment included in the role façade should have at least one policy for enacting it, else there would be no way to enact it—suggesting that either the Org specification is excessive or the agent implementation is incomplete. When multiple policies apply, we must deal with any potential conflicts between them, such as by prioritizing them. Many instances of the same commitment in the specification are possible, one for each tuple of bindings of the variables in the antecedent. They would all be treated by the same forward-chaining rules and the same policies.

The above formulation helps us ensure that the commitments in the specification guide the implementation of the agent policies. Since the communications have fixed meanings in an Org, it is not possible for a commitment to arise that involves an unexpected antecedent or consequent. Potentially, however, an agent can have multiple policies for dealing with the same commitment and may rely on prioritization among policies to handle such. We defer the study of the associated challenges of correctness to future work.

For a debtor, if the antecedent holds and the policy evaluates to true for some bindings of variables, the decision maker attempts the consequent action. In cases where the consequent of a commitment involves exercising a power, i.e., the consequent of a commitment includes the consequent of a power, we generate an alternative rule template whose Then clause describes the necessary outgoing communication corresponding to the consequent of the commitment and the antecedent of the power.

For a creditor, the rules might help determine if the debtor is complying, e.g., by checking whether the consequent of a commitment has become true provided its antecedent has become true.

An additional rule template corresponds to handling messages received from others. Each such template checks if the sender is suitably empowered and authorized for the given interaction. Also, where access is given to a resource (as by an *owner* to a *user* for an instrument), the authorization is placed on a computational object that acts as a proxy for the resource and verifies that an incoming request is valid before the resource acts on it. In either case, failure flags a violation.

The normative filter verifies if the action being attempted is authorized and passes it along to the communicator if and only if it is. We model two kinds of attempts: *now or never* (discarded on failure) or *good till canceled* (retried repeatedly until it executes once or the decision maker annuls it). In addition, in our default operational model, the normative filter also verifies whether the action being attempted would violate any prohibitions. Doing so improves the quality of a collaboration. In general, an agent cannot assume others will not violate their prohibitions, because the agents are not implemented or controlled uniformly. Specifically, users may cause their agents to violate a norm or, as explained above, act externally to the Org. Thus the "trust but verify" dictum applies in this setting.

Determining whether an attempted action is authorized is nontrivial, because some actions can have additional consequences, and some of those consequences might not be authorized. In particular, when the agent is empowered to create a new norm, it may not exercise such a power unless the norm to be created is authorized. For example, the Monterey Org should not commit Ryzard to reboot Alice's instrument without her authorization. To this end, the normative filter computes the *power closure* of an action and verifies that all actions in the closure are authorized.

Additionally, the normative filter tracks the states of all applicable norms in which the given agent features. Specifically, it updates the normative state based on any powers (of this or other agents) that might be exercised when an outgoing or incoming message occurs. If it detects a violation of a norm by another agent, it applies the specified sanction. In particular, a common sanctioning pattern is for the agent to generate an escalation, i.e., a complaint, to the Org that is the context of the violated norm, and for the Org to exercise stricter sanctions of its own. In addition, depending on the applicable role façade, the given agent may also carry out a sanction such as giving the violating agent a poor rating.

5. EVALUATION

We now address the claims asserted in Section 1.3 by returning to the OOI application scenarios introduced in Section 2.

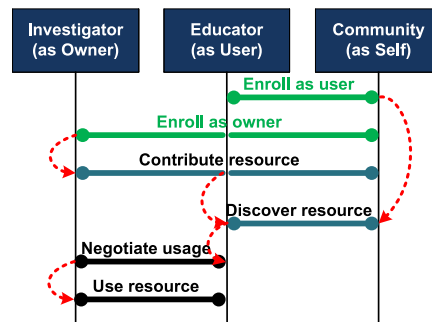### 5.1. Realizing the Motivating Scenarios of Section 2

Fig. 6.   Illustration of governance in a resource sharing Org. The notation is loosely based on message sequence charts. The horizontal lines show governance interactions that create or modify norms among the parties whose life-lines they connect. (In general, a governance interaction may involve more than two parties.) We can think of each horizontal line as describing a modular protocol in meaning-based terms by describing what norms among the associated principals that it creates or manipulates. Any causality requirements across such interactions are captured via the dashed arrows that connect some pairs of the horizontal lines [Singh 2011].

5.1.1. Collaboration. Figure 6 shows how governance may be flexibly captured. It shows various interactions, modularized in terms of their effect of the normative states of the principals involved. For example, enrolling as *user* creates norms on both the Educator as the new user and the Community as the Org.

Capturing governance requirements using norms yields modularity and clarity in understanding and validating the modeled governance as well as flexibility in operational terms. Specifically, even as simple an interaction as enrollment can potentially be operationalized in multiple ways, including by having either (1) the prospective enrollee or (2) the prospective enroller take initiative by, respectively, requesting membership or inviting the enrollee. In effect, each governance interaction may be operationalized as multiple message sequence charts. Such flexibility is one of the reasons it is appropriate to use the undirected notation, which indicates an association among the specified principals but avoids (1) having to say who initiates the transaction and (2) over-specifying the operational constraints on the messages sent and received.

5.1.2. Affiliation. Figure 7 illustrates the structural and contextual scenarios. The Chesapeake and Monterey Orgs *qua* principals play the *affiliate* role in the Salinity and Algae Org, whose norms support the formation and maintenance of norms between Ryzard, a Monterey *user*, and Bejan, a Chesapeake *owner*.
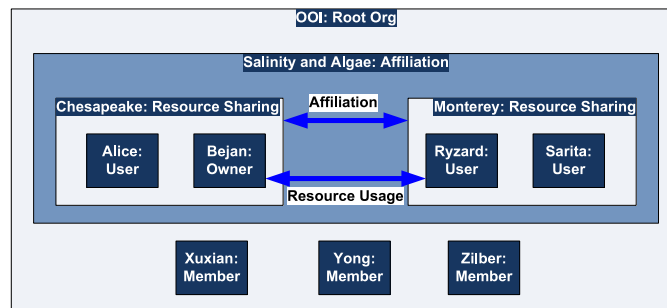

Fig. 7. A schematic of two resource sharing Orgs affiliating with each other through the aegis of another Org. Of the root Org's members, two have enrolled in the Chesapeake Org and two in the Monterey Org. The Chesapeake and Monterey Orgs form their affiliate, the Salinity and Algae Org, and function as principals within it.

5.1.3. Sanction. Figure 7 treats OOI itself as a principal that acts as an overarching authority for all interactions within its scope. As the *root* Org, OOI defines the identities for the principals involved and provides the basic rules of encounter for all constituent Orgs. In this setting, if Ryzard misuses Bejan's instrument, Bejan can complain to the Chesapeake Org; his complaint is forwarded via the Salinity and Algae Org to the Monterey Org, which may sanction Ryzard or risk itself being ejected from the Salinity and Algae Org.

5.2. Flexibility in Configuration

5.2.1. Transactional Adaptation. The researcher merely enrolls in a community for enthusiasts to which she contributes instruments that have spare capacity. She limits the contributed capabilities so an inexperienced user cannot inadvertently harm her instrument.

5.2.2. Structural Adaptation. The Salinity and Algae affiliation Org is expanded so that each laboratory (1) exposes its zoological databases, enabling their discovery by members of the other laboratory, and (2) entertains discovery and usage requests from principals playing the *student* role in the other laboratory.

5.2.3. Contextual Adaptation. The collaborators decide that a prohibition against sharing data externally would prove onerous. They decide to remove that prohibition with respect to deliverables of datasets to satisfy their respective research sponsors. This modifies the prohibition and in essence reconfigures the subsequent engagement.

5.3. Flexibility during Operation

5.3.1. Transactional Adaptation. The researcher simply applies a policy that leads her to exercise her power as *owner* to withdraw the instrument that is attracting high demand from the community of enthusiasts. According to the rules of encounter, she can deny additional usage requests for the instrument immediately, but must wait to withdraw the instrument until its current usage sessions have ended.

5.3.2. Structural Adaptation. In the middle of the ongoing affiliation, one of the parties proposes a modification of one or more role façades. Each participating Org, viewed as a principal, acts autonomously with respect to the other and the modification takes place only if both agree. If they do not agree, the proposing Org may terminate the affiliation according to the existing norms. Notice that the Orgs are autonomous with respect to each other, but need not be autonomous with respect to their members. For instance, based on each Org's membership norms, a designated role could have the power to decide on its behalf or perhaps the Org could conduct a referendum of its members. Moreover, the two Orgs may use completely different mechanisms to arrive at their respective decisions.

5.3.3. Contextual Adaptation. We support this scenario by dynamically modifying the sanctioning norm of the Chesapeake Org. The Org's configuration would not change but the policy under which the Org exercises applicable sanctions is altered, through a decision mechanism similar to the one alluded to above under Structural Adaptation.

6. DISCUSSION: LITERATURE AND FUTURE WORK

This paper has made the case that self-governance is a natural metaphor for the administration of multistakeholder sociotechnical systems, treating the stakeholders as active participants. First, our approach respects the autonomy of the participants while supporting adaptations in their mutual interactions. Thereby, it enables the automation of what would otherwise be manual out-of-band administrative processes. Second, being founded in norms, our approach naturally provides an elegant way to realize governance by providing a measure of correctness that emphasizes interactions and is independent of implementation details. Third, a benefit of our approach is that through the composition of Orgs, it supports modularizing the norms and the agents' policies with respect to norms. As a result, it simplifies reusing Org specifications as well as the policies through which agents enact their roles. We have applied our approach on real-life scenarios from the specification and operation of a large sociotechnical system. As Table II shows, it is conceptually quite straightforward to accommodate a rich variety of adaptations in our approach.

6.1. Relevant Literature

Because of space limitations, we can review only representative publications from the two major classes of literature that are relevant to our approach.

6.1.1. Autonomics and Policy. Our principles and approach for adaptation respect but enhance autonomic computing [Kephart and Chess 2003]. In particular, our configurational adaptations capture their notion of self-configuration. Brazier et al. [2009] identify synergies between autonomic computing and multiagent systems, which this paper partly illustrates. A point of distinction from the above works is that we emphasize multistakeholder systems, where self-governance is a better metaphor than self-management.

Shankar et al. [2006] generalize ECA policies by explicitly modeling the pre- and postconditions of actions. Doing so facilitates computing a correct order in which to apply the policies. Our

Table II. Summary of the main types of adaptations accommodated by our approach.

	Element	**Example** (Section 2)	**Approach** (Sections 5.2 and 5.3)
Configuration	*Transactional*	Support enthusiasts with spare capacity instruments	Join an additional Org corresponding to a community of enthusiasts
	Structural	Expand an affiliation to introduce new resources and participants	Agree to make additional resources available to the affiliate and to entertain requests from an expanded pool of roles
	Contextual	Allow deliverables to be sent to a research sponsor	Weaken the prohibition against sharing and no longer prohibit sharing data with a sponsor
Operation	*Transactional*	Reallocate resources during oil spill	Withdraw a resource with high demand from the (low priority) community of enthusiasts
	Structural	Modify the affiliation to forbid unilateral publishing of results	Introduce a new norm into the façades of the appropriate roles
	Contextual	Disregard the prohibition on sharing during an algae bloom emergency	Modify the sanctioning policy to account for situations threatening public safety

approach goes beyond previous work in modeling a system with multiple autonomous parties, capturing norms explicitly, and in having each participant base its actions on the applicable norms, and on its *own* policies. In particular, by representing norms explicitly, we can (1) decouple creating a norm from acting on the norm and (2) support manipulating norms.

6.1.2. Norms and Organizations. Dignum's [2004] OperA model captures similar intuitions to ours; it is applied in supporting organizational adaptivity in terms of norms [Álvarez-Napagao et al. 2009]. The key points of difference with the OperA organization model are our directed representation of (conditional) norms within an organizational context, our explicit treatment of operations on norms, and our emphasis on the duality of organizations and norms. Our representations support multiple perspectives because they enable principals to interact as peers. In this manner, they contrast with workflows (as of "scenes") expressed from a central perspective. A benefit of our approach is that it recognizes the inherent autonomy of principals and avoids the situation where one principal controls the interactions or actions of another. Further, our approach supports organizations themselves being principals that can nest other organizations. Thus we can replace OperA's organizational, social, and interaction models by just one set of abstractions.

The OperA models capture an organization and a role's goals and objectives, which we do not model since the internal details of an autonomous party are not relevant to modeling how it interacts with others. We capture the interaction requirements purely in normative terms—in this task we are assisted by having norms modeled as directed and conditional, and having placed them in an Org context. Doing so makes our norms more expressive than traditional representations because we can naturally capture who is responsible to whom, what nuanced conditions apply on a norm, and how to handle various exceptions through the Org context. OperA's notion of the right of a role as a capability that agents playing that role acquire are quite similar to our notion of privilege. However, we treat a liability in the same framework as simply a counter perspective to a privilege. Further, we capture both privileges and liabilities via our small set of norms, which yields a more refined treatment than OperA's. In passing, we note that, in our approach, we reserve the notion of "capabilities" for physical abilities in contrast with powers that can arise as privileges in a role façade. Capabilities are not the focus of this paper.

Fornara et al. [2008] propose the OCeAN metamodel motivated by similar intuitions about norms as us, especially that agents should not be regimented and should be able to violate their norms, albeit at the risk of facing sanctions. Fornara et al. give a special status to commitments, like [Singh 1999], but unlike this paper—we conjecture that a first-class representation of a variety of norm concepts simplifies the elicitation and representation of stakeholder requirements. However, they also consider traditional deontic relations such as obligations, which are unnecessary given commitments, and treat powers and authorizations as synonymous (p. 91). Fornara et al. consider in-

stitutional actions, which we map to operations on norms to be performed by suitably empowered parties. They do not consider the expanded operations on norms, such as delegate, as we do. Our architecture offers the benefit of combining regimentation, which is efficient and reliable but incomplete, with sanctioning, which is essential in open settings. Fornara et al.'s more extensive temporal representation can possibly be combined with our approach.

Vasconcelos et al. [2007] apply norms for modeling e-Science organizations. They treat norms as being applied by the organization and do not emphasize the autonomy of the principals; also, their norms are not directed from one to another thereby losing some expressive power. Vasconcelos et al. do not provide an account of external violations and sanctioning as a way to resist such violations. However, Vasconcelos et al. address the important problem of resolving conflicts among norms, which can arise when a principal plays two or more roles. It would be useful to combine their approach with ours so that an agent can analyze its norms before taking on any roles.

6.1.3. Norms and Adaptation. Campos et al. [2009] propose an adaptation mechanism for electronic institutions that employs "staff" agents to monitor members' behavior and if necessary autonomically reconfigure the system. Overbeek et al. [2010] propose an approach that supports both direct control by a regulator and self-regulation as ways to ensure norm fulfillment. In contrast, our approach emphasizes the participants' autonomy, so no staff or regulators can control an agent or reconfigure an Org. We formulate the *self* role, which captures the Org as a participant, and projects the Org's identity uniformly inside and outside of its scope. We address adaptations through decentralized creation and manipulation of norms. As a result, we can accommodate configurational and operational adaptations in a simplified, uniform framework. However, Overbeek et al.'s value-based methodology is compelling, especially for a multistakeholder system. It would be useful to layer it over our governance model.

Tinnemeier et al. [2010] study schema and instance changes in norms, which correlate with contextual adaptation in terms of configuration and operation. They assume that a norm change is somehow specified, but do not consider the governance processes by which principals would agree to a specific norm change. Derakhshan et al.'s [2012] approach deals with changing norms but it considers norms that are not directed and appear to be centrally assigned: the "system" relies on a central knowledge base and assigns new roles to agents. Our approach is inherently more flexible because of the directionality of norms and the local representation of normative state by each party. In principle, an agent may adopt or exit any role as long as doing so does not violate any norms. Developing an engineering methodology to accommodate such flexibility is an important future task.

The dynamic aspects of norms may additionally be approached from a bottom-up perspective, where norms may be implicit in the interactions of agents and may emerge through mutual reinforcement [Hollander and Wu 2011; Savarimuthu et al. 2009]. The norms handled in this body of work are generally of a simplistic and a universal flavor, such as driving on the left or the right side of the road. There is no reason why the logically structured norms of the present paper could not also arise bottom up: as indeed they did and still do in human society: the technical challenge involved in accommodating the emergence of such complex norms appears nontrivial and interesting.

Further, the dynamic approaches associate naturally with quantitative aspects of norms, since they often consider the probabilistic or utility-based analyses of norms. Such quantitative aspects have been studied in connection with commitments [Desai et al. 2008; Yolum and Singh 2007]; additional theoretical bases for norms could potentially be developed exploiting their formal semantics based on computation paths, e.g., as motivated for commitments by Singh [2008].

6.2. Directions for Future Work

We expect that an agent would satisfy all the norms that apply on its interactions. Therefore, detecting inconsistencies among norms and computing acceptable actions for consistent sets of norms is a crucial challenge. We can expect the designers of an Org to create consistent norms. However, a principal may play roles in multiple Orgs. Further, privileges can sometimes function as liabilities.

For example, assume principal Yong, but not Zhang, is empowered to publish a report merely by sending an email to a web site. Then Yong's power could prove undesirable for him because he might violate a prohibition, whereas Zhang would be protected from such a violation. Therefore, support for norm consistency (as inspired by [Vasconcelos et al. 2007]) and other validity checks are a key challenge.

Techniques for authoring agent policies and verifying them with respect to the norms that govern a given agent are crucial. A related challenge is verifying whether the specified norms are supported by a given operational description such as might be described via sequence diagrams. Telang and Singh [2012] address this problem for commitments; we leave it as future work to extend it to the full range of norms introduced in this paper.

We outlined a simple methodology for the design of sociotechnical systems. However, a more extensive approach is needed that would accommodate considerations such as stakeholders' goals, which underlie governance (and other) requirements, as in Tropos [Bresciani et al. 2004]. Our approach agrees in spirit with Tropos but begins from a first-class status for norms among autonomous principals. Penserini et al. [2007] address the challenge of high-variability design from the standpoint of Tropos. Therefore, it is only natural that in future work we attempt to develop a methodology based on Tropos but extended to deal with the special challenges of norms and organizations.

The centrality of norms also brings new problems to light. The procedure for arbitration of conflicts arising between peers is more loosely defined than those for enacting or forming norms. We envision a model to capture the utility derived from a contract and using this to reason how a conflict may be settled by a neutral party. Such adjudication is clearly not tractable and requires formalization. Desai et al. [2008] propose a model to assess the safety and benefits of participating in a contract.

REFERENCES

ALDEWERELD, H., ÁLVAREZ-NAPAGAO, S., DIGNUM, F., AND VÁZQUEZ-SALCEDA, J. 2010. Making norms concrete. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. IFAAMAS, Toronto, 807–814.

ÁLVAREZ-NAPAGAO, S., CLIFFE, O., PADGET, J. A., AND VÁZQUEZ-SALCEDA, J. 2009. Norms, organisations and semantic web services: The ALIVE approach. In *Proceedings of the 2nd Multi-Agent Logics, Languages, and Organisations (MALLOW)*. CEUR Series, vol. 494. CEUR-WS.org, Torino.

ARROTT, M., CHAVE, A. D., FARCAS, C., FARCAS, E., KLEINERT, J. E., KRUEGER, I., MEISINGER, M., ORCUTT, J. A., PEACH, C., SCHOFIELD, O., SINGH, M. P., AND VERNON, F. L. 2009. Integrating marine observatories into a system-of-systems: Messaging in the US Ocean Observatories Initiative. In *Proceedings of Oceans, the MTS-IEEE Conference on Marine Technology for our Future: Global and Local Challenges*. IEEE Computer Society, Biloxi, Mississippi, 1–9.

BRAZIER, F., DIGNUM, V., HUHNS, M. N., DERKSEN, C., DIGNUM, F., LESSNER, T., PADGET, J., QUILLINAN, T., AND SINGH, M. P. 2012. Agent-based organisational governance of services. *Multiagent and Grid Systems: An International Journal (MAGS) 8,* 1, 3–18.

BRAZIER, F. M. T., KEPHART, J. O., PARUNAK, H. V. D., AND HUHNS, M. N. 2009. Agents and service-oriented computing for autonomic computing: A research agenda. *IEEE Internet Computing 13,* 3, 82–87.

BRESCIANI, P., PERINI, A., GIORGINI, P., GIUNCHIGLIA, F., AND MYLOPOULOS, J. 2004. Tropos: An agent-oriented software development methodology. *Journal of Autonomous Agents and Multi-Agent Systems 8,* 3, 203–236.

CAMPOS, J., LÓPEZ-SÁNCHEZ, M., RODRÍGUEZ-AGUILAR, J. A., AND ESTEVA, M. 2009. Formalising situatedness and adaptation in electronic institutions. In *Proceedings of the International Workshop on Coordination, Organizations, Institutions and Norms in Agent Systems (COIN 2008)*. LNCS Series, vol. 5428. Springer-Verlag, Berlin, 126–139.

DERAKHSHAN, F., BENCH-CAPON, T., AND MCBURNEY, P. 2012. Dynamic assignment of roles, rights and responsibilities in normative multi-agent systems. *Journal of Logic and Computation 22*, 1–18. Published online. doi: 10.1093/logcom/exr027.

DESAI, N., CHOPRA, A. K., AND SINGH, M. P. 2009. Amoeba: A methodology for modeling and evolution of cross-organizational business processes. *ACM Transactions on Software Engineering and Methodology (TOSEM) 19*, 2, 6:1–6:45.

DESAI, N., MALLYA, A. U., CHOPRA, A. K., AND SINGH, M. P. 2005. Interaction protocols as design abstractions for business processes. *IEEE Transactions on Software Engineering 31*, 12, 1015–1027.

DESAI, N., NARENDRA, N. C., AND SINGH, M. P. 2008. Checking correctness of business contracts via commitments. In *Proceedings of the 7th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. IFAAMAS, Estoril, Portugal, 787–794.

DIGNUM, V. 2004. A model for organizational interaction: Based on agents, founded in logic. Ph.D. thesis, Universiteit Utrecht.

FORNARA, N., VIGANÒ, F., VERDICCHIO, M., AND COLOMBETTI, M. 2008. Artificial institutions: A model of institutional reality for open multiagent systems. *Artificial Intelligence and Law 16*, 1, 89–105.

HOHFELD, W. N. 1919. *Fundamental Legal Conceptions as Applied in Judicial Reasoning and other Legal Essays*. Yale University Press, New Haven, CT. A 1919 printing of articles from 1913.

HOLLANDER, C. D. AND WU, A. S. 2011. Using the process of norm emergence to model consensus formation. In *Proceedings of the 5th IEEE International Conference on Self-Adaptive and Self-Organizing Systems (SASO)*. IEEE, Ann Arbor, Michigan, 148–157.

JOHNSON, M., BRADSHAW, J. M., FELTOVICH, P. J., JONKER, C. M., VAN RIEMSDIJK, M. B., AND SIERHUIS, M. 2011. The fundamental principle of coactive design: Interdependence must shape autonomy. In *Proceedings of the 6th AAMAS Workshop on Coordination, Organization, Institutions and Norms (COIN 2010)*. LNCS Series, vol. 6541. Springer, Toronto, 172–191.

JONES, A. J. I. AND SERGOT, M. J. 1996. A formal characterisation of institutionalised power. *Logic Journal of the IGPL 4*, 3, 427–443.

KEPHART, J. O. AND CHESS, D. M. 2003. The vision of autonomic computing. *IEEE Computer 36*, 1, 41–50.

OVERBEEK, S., DIGNUM, V., AND TAN, Y.-H. 2010. Towards a value-sensitive system to support agents in norm fulfillment and enforcement. TR, Delft University. June.

PENSERINI, L., PERINI, A., SUSI, A., AND MYLOPOULOS, J. 2007. High variability design for software agents: Extending Tropos. *ACM Transactions on Autonomous and Adaptive Systems (TAAS) 2*, 4, 16:1–16:27.

SAVARIMUTHU, B. T. R., CRANEFIELD, S., PURVIS, M. K., AND PURVIS, M. A. 2009. Norm emergence in agent societies formed by dynamically changing networks. *Web Intelligence and Agent Systems 7*, 3, 223–232.

SHANKAR, C., TALWAR, V., IYER, S., CHEN, Y., MILOJICIĆ, D., AND CAMPBELL, R. 2006. Specification-enhanced policies for automated management of changes in IT systems. In *Proceedings of the 20th Large Installation System Administration Conference (LISA)*. Usenix, Washington, 103–118.

SINGH, M. P. 1999. An ontology for commitments in multiagent systems: Toward a unification of normative concepts. *Artificial Intelligence and Law 7*, 1, 97–113.

SINGH, M. P. 2008. Semantical considerations on dialectical and practical commitments. In *Proceedings of the 23rd Conference on Artificial Intelligence (AAAI)*. AAAI Press, Chicago, 176–181.

SINGH, M. P. 2011. Information-driven interaction-oriented programming: BSPL, the blindingly simple protocol language. In *Proceedings of the 10th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. IFAAMAS, Taipei, 491–498.

SINGH, M. P., CHOPRA, A. K., AND DESAI, N. 2009. Commitment-based service-oriented architecture. *IEEE Computer 42*, 11, 72–79.

TELANG, P. R. AND SINGH, M. P. 2012. Specifying and verifying cross-organizational business models: An agent-oriented approach. *IEEE Transactions on Services Computing 5*, 1–14. In press; available at http://www.csc.ncsu.edu/faculty/mpsingh/papers.

TINNEMEIER, N. A. M., DASTANI, M., AND MEYER, J.-J. C. 2010. Programming norm change. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. IFAAMAS, Toronto, 957–964.

VASCONCELOS, W., KOLLINGBAUM, M. J., AND NORMAN, T. J. 2007. Resolving conflict and inconsistency in norm-regulated virtual organizations. In *Proceedings of the 6th International Joint Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. IFAAMAS, Honolulu, 632–639.

YOLUM, P. AND SINGH, M. P. 2007. Enacting protocols by commitment concession. In *Proceedings of the 6th International Joint Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*. IFAAMAS, Honolulu, 116–123.