Towards a Model of Social Coherence in Multi-Agent Organizations

Erick Martínez
SIAT*
Simon Fraser University
Vancouver, Canada
emartinez@sfu.ca

Ivan Kwiatkowski
ISIMA†
Clermont-Ferrand, France
kwiatkow@poste.isima.fr

Philippe Pasquier
SIAT
Simon Fraser University
Vancouver, Canada
pasquier@sfu.ca

ABSTRACT

We propose a *social coherence*-based model and *simulation framework* to study the dynamics of multi-agent organizations. This model rests on the notion of *social commitment* to represent all the agents' explicit inter-dependencies including *roles* and *organizational structures*. A local coherentist approach is used that, along with a *sanction policy*, ensures *social control* in the system and the *emergence* of *social coherence*. We illustrate the model and the simulator with a simple experiment comparing two *sanction policies*.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*intelligent agents, multiagent systems*

General Terms

Algorithms, Experimentation, Theory

Keywords

Social and organizational structure, social commitments, agent reasoning, social control

1. INTRODUCTION AND MOTIVATIONS

Research in the area of Computational Organization Theory [3] and multi-agent systems (MAS) has resulted in a large number of models capturing different aspects of organizational behaviour [18, 19, 1, 6]. This paper presents a model and *simulation framework* to study the social dynamics of multi-agent organizations. The model uses the notion of *social commitment* (defined in Section 2) as the main building block to represent all the inter-dependencies between *social entities*. *Sanction policies* provide *social control* mechanisms (defined in Section 3) to regulate the enforcement of *social commitments*. Our model extends previous work on *cognitive coherence* [15, 16] by showing how the *coherence principle* can drive the emergence of *social behaviour*. In particular, by organizing agent behaviour in a way that makes global *social coherence* (formalized in Section 4) emerge from the local *cognitive coherence* of interacting agents.

We strive to build a simple minimalist model, where *social behaviour* emerges from local behaviour, enabling us to study the social dynamics of multi-agent organizations. This paper advances the state of the art by proposing a unified yet computational and operational view of the social aspects of multi-agent systems. We also present a sample pizza delivery domain (Section 5), and illustrate the use of the model and simulator with a simple experiment (Section 6) to investigate *social control* mechanisms while comparing two *sanction policies*. Then, we discuss our work while relating it to other research (Section 7). Finally, we conclude and discusses future work (Section 8).

*School of Interactive Arts and Technology.

†Institut Supérieur d'Informatique, de Modélisation et de leurs Applications.

2. SOCIAL MODELLING

2.1 Handling Actions

We represent atomic actions as (possibly) parametrized predicate formulas with unique names. We use a discreet instant-based sequential model of time where actions are assumed to be instantaneous. However, each action requires a *preparation time* expressed in time steps.

Definition 1. ***(Primitive or Atomic Action)*** *Given the non-empty set $\mathcal{X}$ of all atomic actions in the system, a* ***primitive action*** $\alpha \in \mathcal{X}$ *is represented as a tuple* $\alpha = \langle \alpha(\vec{x}), \Delta_\alpha \rangle$, *where:*

- $\alpha(\vec{x})$ *is a predicate formula s.t.* $\alpha(\vec{x}) \neq \beta(\vec{x})$, *and* $\alpha(\vec{x}) = \alpha(\vec{y}) \Rightarrow \vec{x} = \vec{y}$; *and*
- $\Delta_\alpha > 0$ *specifies the* preparation time *of action* $\alpha(\vec{x})$ *measured in time steps.*

In our model, *exogenous* events are treated as actions not necessarily performed by agents in the system. Therefore, in the rest of the paper events and actions are used interchangeably. We model an *exogenous* event as an action recurring within certain period of time.

Definition 2. ***(Exogenous Action)*** *Given the set $\hat{\mathcal{X}}$ of all exogenous actions in the system, an* ***exogenous action*** $\hat{\alpha} \in \hat{\mathcal{X}}$ *is represented as a tuple* $\hat{\alpha} = \langle \alpha^{exog}(\vec{x}), \varepsilon \rangle$, *where:*

- $\alpha^{exog}(\vec{x})$ *is a predicate formula s.t.* $\alpha^{exog}(\vec{x}) \neq \beta^{exog}(\vec{x})$, *and* $\alpha^{exog}(\vec{x}) = \alpha^{exog}(\vec{y}) \Rightarrow \vec{x} = \vec{y}$; *and*
- $\varepsilon \geq 0$ *specifies the maximum period within which the event* $\alpha^{exog}(\vec{x})$ *will occur once.*

2.2 Social Commitment

This section briefly presents a formal model of social commitment (henceforth abbreviated s-commitment). Concretely, commitments have proven useful to represent all the agent inter-dependencies: social norms, roles, authority relations

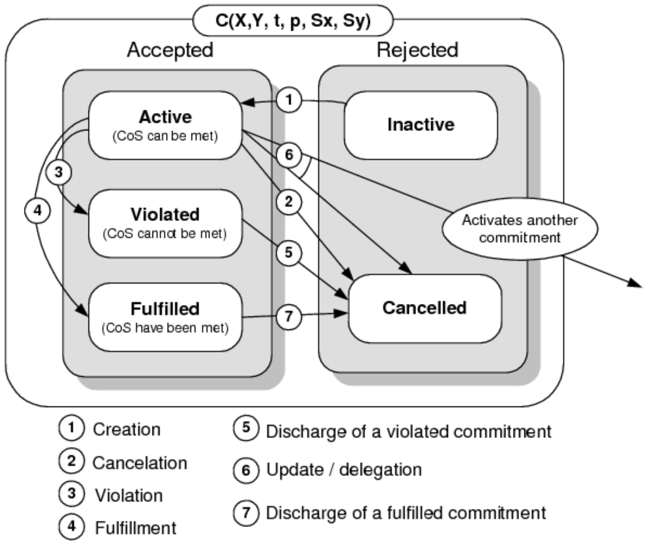

Figure 1: Social commitment finite state transition machine.

and the semantics of agent communication [4, 20]. Conceptually, commitments are oriented responsibilities contracted by a *debtor* towards a *creditor*.[1] One can distinguish *action commitments* from *propositional commitments* [24]. Propositional commitments entail complications and for that reason, following a number of other researchers [4, 8, 7], we will only consider action commitments in the rest of this paper. That is, commitments where a *debtor* is committed towards a *creditor* to bring about the effects of some *atomic action*. We adopt the model of Pasquier et al. [16] in which the dynamics of social commitments is formalized as a finite state machine (FSM). Figure 1 illustrates the different ways s-commitments can be manipulated. Note that update and delegation will not be considered in the rest of this paper.

Definition 3. ***(Action Commitment Schema)*** *Given the non-empty set $\mathcal{SC}$ of action commitments in the system, a particular* **action commitment schema** $\bar{c} \in \mathcal{SC}$ *is represented as a* unique *rule of the form:*

$$\bar{c} = \phi \Rightarrow C(x, y, \alpha, t_d, \mathcal{S}_x, \mathcal{S}_y) \quad (1)$$

where:

- *The antecedent ϕ is a formula representing any general trigger condition, i.e. a primitive action, an exogenous action, or any other complex condition[2];*
- *The consequent $C(x,\ y,\ \alpha,\ t_d,\ \mathcal{S}_x,\ \mathcal{S}_y)$ is a predicate formula with an arity of 6; representing the fact that the* debtor *enacting role x is committed towards the* creditor *enacting role y to achieve the effects of action α within $t_d > 0$ time steps of the creation time, under the sanctions sets $\mathcal{S}_x$ and $\mathcal{S}_y$, which specify the different sanctions that will be applied to x and y according to the states and transitions applicable to this commitment; and*
- $\alpha = \langle \alpha(\vec{x}), \Delta_\alpha \rangle$*, with* $t_d \geq \Delta_\alpha$.

[1]Social commitments share a great deal with the notion of directed obligation as defined in deontic logic and as also used by some researchers in the context of agent communication.

[2]In this paper we restrict the antecedent formula ϕ to be either a primitive or exogenous event.

Note that, *action commitment schema* (1) can only be valid, if its total duration time t_d is at least as long as the *preparation time* (i.e. Δ_α) of the *atomic action* α. In this paper, we only consider action commitments involving atomic actions. We can look at *action commitment schemes* as *abstract* place holders describing generic oriented responsibilities contracted by a *debtor* towards a *creditor*. Social commitment *schemes* are ultimately instantiated by agents.

Definition 4. ***(Instantiated Action Commitment)*** *Given an* action commitment schema

$$\bar{c} = \phi \Rightarrow C(x, y, \alpha, t_d, \mathcal{S}_x, \mathcal{S}_y) \quad (2)$$

where the trigger condition ϕ is satisfied, we define an ***instantiation*** *of $\bar{c}$ as a* unique grounded *predicate formula with an arity of 7:*

$$c = C(x, y, \alpha, t_s, t_f, \mathcal{S}_x, \mathcal{S}_y) \quad (3)$$

where x and y are debtor and creditor agents, respectively. Formula (3) *results from:*

- *Removing the (satisfied) antecedent ϕ from formula* (2)*;*
- *Removing parameter t_d from the consequent of* (2)*, then adding parameters t_s, t_f where: t_s represents the creation time when ϕ occurs and schema $\bar{c}$ gets* instantiated*; and $t_f = t_s + t_d$;*
- *Instantiating every free variable from formula* (3)*.*

Note that both *schemes* and *instantiated action commitments* must be distinctly identified in the system. *Accepted* action commitments take the form of a grounded predicate formula: $C(x, y, \alpha, t_s, t_f, \mathcal{S}_x, \mathcal{S}_y)$. *Rejected* commitments, meaning that *debtor* x is not committed towards *creditor* y to achieve action α, take the form $\neg C(x,\ y,\ \alpha,\ t_s,\ t_f,\ \mathcal{S}_x,\ \mathcal{S}_y)$.

Our model also accounts for *ordering constraints* between *instantiated* social commitments.

Definition 5. ***(Time Constraint)*** *Given two* distinct instances *of social commitments*

$$\begin{aligned} c_i &= C(x_i, y_i, \alpha, t_{s_i}, t_{f_i}, \mathcal{S}_{x_i}, \mathcal{S}_{y_i}) \\ c_j &= C(x_j, y_j, \alpha, t_{s_j}, t_{f_j}, \mathcal{S}_{x_j}, \mathcal{S}_{y_j}) \end{aligned} \quad (4)$$

where: $t_{f_i} < t_{s_j}$ (resp. $t_{f_j} < t_{s_i}$); we use the ***time (ordering) constraint*** *notation $c_i \prec c_j$ (resp. $c_j \prec c_i$) to represent* disjoint *time intervals between $\{c_i, c_j\}$, where c_i (resp. c_j) temporally precedes c_j (resp. c_i). Otherwise, we use the notation $c_i \preccurlyeq c_j$ (resp. $c_j \preccurlyeq c_i$) to represent a time interval* overlapping constraint *between $\{c_i, c_j\}$.*

Active social commitments raise action expectations, and the enforcement of social commitments can take place through various *social control* mechanisms instead of through assumptions of sincerity and cooperativeness [16]. Social commitments, when modelled with their enforcement mechanism [16], are not necessarily sincere and do not require the agents to be cooperative. From this perspective, social commitments serve to coordinate the agents whether or not they are cooperative and whether or not they are sincere.

2.3 Social entities

In this paper, we only consider three types of social entities: agents, social roles, and social organizations. While numerous refinements are possible, we take a minimalist approach to define these entities. Formally:

Definition 6. ***(Social Entity)*** *Given the non-empty set $\mathbb{D} = Ag \cup Role \cup Org$, a particular* **social entity** *d is represented as $d \in \mathbb{D}$ where:*

- *Ag, $Role$, Org are sets that stand for all the agents, social roles, and organizations respectively;*
- *And $Ag \cap Role = \emptyset$, $Ag \cap Org = \emptyset$, $Role \cap Org = \emptyset$.*

Definition 7. ***(Organization)*** *Given the set Org of all organizations in the system, a particular* ***organization*** *$o \in Org$ is represented as a tuple $o = \langle A_o, R_o, \rho_o \rangle$ where:*

- *A_o is the set of agents that belongs to the organization, with $A_o \neq \emptyset$, $A_o \subseteq Ag$;*
- *R_o is the set of roles relevant to the organization, with $R_o \neq \emptyset$, $R_o \subseteq Role$; and*
- *ρ_o is a binary relation that assigns to each agent that belongs to the organization, one or several roles from R_o, noted $\rho_o : A_o \longrightarrow R_o^n$ $(1 \leq n \leq |R_o|)$, s.t. $\forall ag_i \in A_o$ $\rho_o(ag_i) \neq \emptyset$.*

Definition 8. ***(Social Role)*** *Given the set $Role$ of all social roles and the set $\mathcal{X}$ of all primitive actions in the system, a particular* ***social role*** *$r \in Role$ is represented as a tuple $r = \langle \mathcal{X}_r, \mathcal{SC}_r \rangle$ where:*

- *$\mathcal{X}_r$ is the set of* primitive actions *that define the capabilities of this role, with $\mathcal{X}_r \neq \emptyset$, $\mathcal{X}_r \subseteq \mathcal{X}$; and*
- *$\mathcal{SC}_r$ is the set of* s-commitment schemes *specifying the inter-dependencies between this role and every other* debtor *or* creditor, *with $\mathcal{SC}_r \neq \emptyset$.*

Definition 9. ***(Agent)*** *Given the set Ag of all agents in the system, a particular* ***agent*** *$ag \in Ag$ is represented as a tuple $ag = \langle R_{ag}, \varkappa_{ag} \rangle$ where:*

- *R_{ag} is the set of roles the agent is assigned to, with $R_{ag} \neq \emptyset$, $R_{ag} \subseteq Role$; and*
- *$\varkappa_{ag}$ is a binary relation that assigns a probabilistic* reliability *value to each primitive action $\alpha_i \in \mathcal{X}_{ag}$ within the capabilities of agent ag, capturing the probability of agent ag succeeding at performing primitive action α_i, noted $\varkappa_{ag} : \mathcal{X}_{ag} \longrightarrow [0, 1]$, with $\mathcal{X}_{ag} = \bigcup \{\mathcal{X}_{r_j} \mid \langle \mathcal{X}_{r_j}, \mathcal{SC}_{r_j} \rangle \in R_{ag}\}$.*

Organizations and roles are abstract constructs enacted by actual agents. When representing *instantiated commitments* we use a notation inspired by Carabelea and Boissier [2] to capture the role being enacted by the creditor and debtor agents respectively. So, we can now rewrite Formula (3) as follows:

$$c = C(ag_i : r_x, ag_j : r_y, \alpha, t_s, t_f, \mathcal{S}_{ag_i}, \mathcal{S}_{ag_j}) \quad (5)$$

meaning that agent ag_i enacts role r_x and agent ag_j enacts role r_y.

The *capabilities* of an agent are determined by all the primitive actions which define the capabilities of each *role* the agent is assigned to. For example, besides being a *cook* within organization Ω, agent ag_1 could also play the role of a *volunteer firefighter* within a different organization. In such a case, the individual capabilities of the agent ag_1 will clearly span beyond those determined by the scope of his/her role within organization Ω.

There might be instances where the same agent plays several roles within an organization. There might be other instances where several agents play the same role within an organization. In the latter case, we follow a *fair allocation principle* so that (on average) all agents have a similar chance to enact the same role they were assigned to. In our implementation of the model, the Agent Allocation Manager (AAM) module handles the system-wide allocation of agents. It is actually implemented as a wrapper to the *Mersenne Twister* (MT19937 implementation) pseudo-random number generator, which provides fast generation of high-quality pseudo-random numbers. For each role r_i the AAM keeps track of which agents are available (resp. unavailable). When instantiating a s-commitment, the AAM will randomly pick an agent from the *pool* of available agents enacting role r_i until all agents have been allocated a s-commitment and the *pool* is empty. Then, the AAM 'replenishes' the *pool* by flagging all agents enacting role r_i as available and repeats the same process again.

3. SOCIAL CONTROL MECHANISMS

Theories of *social control* [13, 9] focus on the strategies and techniques that help to regulate agent behaviour, and lead to conformity and compliance with the rules of society (at both the macro and the micro level). In the remainder of this section, we detail the main elements used in the enforcement of social commitments: *sanctions*, which are considered in their general sense of positive or negative incentives.

Most s-commitment-based approaches assume that the agents will respect their social commitments (thus applying regimentation). This assumption is unrealistic since unintended commitment violation is likely to occur and unilateral commitment cancellation as well as commitment modification are desirable. Intuitively, sanctions should meet the following base criteria. Violation and *cancellation* are either associated with (possibly) *negative* sanctions, *fulfilment* is associated with a (possibly) *positive* sanction and *violation* carries either a harsher or similar sanction than *cancellation*.

In previous work [16], we have proposed an ontology of sanction types and punishment policies. Here we will only present the basic mechanism by which the enforcement of s-commitment is ensured in our model of *social coherence*. A *sanction policy* determines the type of sanctions (and their magnitude) that are assigned to social commitments at creation time. For simplicity, we assume that sanctions are not delayed through time and are applied at the time of occurrence as specified in the sanction policy.

Definition 10. ***(Sanction Policy)*** *Given an organization $o = \langle A_o, R_o, \rho_o \rangle$; the set $\mathcal{SC}_o$ of all social commitment* schemes *related to the organization; and the set $\mathbb{T}$ of all the transitions applicable to s-commitments. For every* schema *$\bar{c} \in \mathcal{SC}_o$ of the form*

$$\bar{c} = C(r_x, r_y, \alpha, t_s, t_f, \mathcal{S}_x, \mathcal{S}_y)$$

we specify the sanction sets $\mathcal{S}_x = \{s_x^f, s_x^c, s_x^v\}$, and $\mathcal{S}_y = \{s_y^c\}$ using the following function (where z is the transition consumed in the FSM from Figure 1):

$$\sigma_{sc}(z) = \begin{cases} s_x^f & \textit{if } z = 7, \textit{ // discharge of fulfilment} \\ s_x^v & \textit{if } z = 5, \textit{ // discharge of violation,} \\ s_x^c & \textit{if } z = 2, \textit{ // cancellation by debtor,} \\ s_y^c & \textit{if } z = 2, \textit{ // cancellation by creditor} \\ nil & \textit{if } z \notin \{2, 5, 7\} \end{cases} \quad (6)$$

where:

- *$\sigma_{sc} : \mathbb{T} \longrightarrow [-1, 1]$;*
- *s_x^f represents the sanction value applied to* debtor x *when* fulfilling *commitment* c*;*
- *s_x^v represents the sanction value applied to* debtor x *when* violating *commitment* c*;*
- *s_x^c represents the sanction value applied to* debtor x *when* cancelling *commitment* c*; and*
- *s_y^c represents the sanction value applied to* creditor y *when* cancelling *commitment* c.

4. SOCIAL COHERENCE

In cognitive sciences and social psychology most cognitive theories appeal to the *coherence principle* which puts coherence as the main organizing mechanism: *the individual is more satisfied with coherence than with incoherence.* In this section, we build on and extend previous work on *cognitive coherence* [15, 16] by showing how to use the *coherence principle* as the driving force that makes *social behaviour* emerge from the local *cognitive coherence* of interacting agents.

4.1 Formal characterization of social coherence

We present a constraint satisfaction based model of social coherence resulting in a symbolic-connexionist hybrid formalism. In our approach, the cognitions of a social entity are represented through the notion of elements (i.e. instantiated s-commitments). We denote $\mathbb{E}$ the set of all elements. *Elements* are divided in two sets: the set $\mathcal{A}$ of *accepted elements* and the set $\mathcal{R}$ of *rejected elements.* We adopt a closed-world assumption which states that *every non-explicitly accepted element is rejected.* Since not all s-commitments are equally modifiable, a *resistance to change* is associated to each element. Formally:

Definition 11. ***(Resistance to Change)*** *We specify the* ***resistance to change*** *of an element (i.e. instantiated s-commitment) through the function:*

$$Res : \mathbb{E} \times \mathbb{T} \longrightarrow \mathbb{R} \equiv -\sigma_{sc}(z) \tag{7}$$

where $\mathbb{E}$ is the set of all elements, $\mathbb{T}$ is the set of all the transitions applicable to social commitments, and $\sigma_{sc}(z)$ ($z \in \mathbb{T}$) is the sanction policy.

Note that, we equate the *resistance to change* with the sanctions corresponding to the transitions (i.e. *fulfilment*, *cancellation*, *violation*) of the s-commitment as specified in the *sanction policy* (Formula (6)). The higher the punishment (resp. reward) for cancelling/violating (resp. fulfilling) a s-commitment, the higher (resp. lower) the *resistance to change* will be.

S-commitments can be related or unrelated. When they are related, positive compatibility relations like facilitation and entailment are represented as *positive constraints.* Negative incompatibility relations like mutual exclusion (e.g. critical time overlap), hindering, and disabling are represented as *negative constraints.* We use $\mathcal{C}^+$ (resp. $\mathcal{C}^-$) to denote the set of positive (resp. negative) constraints and $\mathbb{C} = \mathcal{C}^+ \cup \mathcal{C}^-$ to refer to the set of all constraints. For each of these constraints, a weight reflecting the importance degree for the underlying relation is attributed (our constraint generation mechanism is described in Section 4.2). Those weights can be accessed through the function $Weight : \mathbb{C} \longrightarrow \mathbb{R}$. Constraints can be satisfied or not.

Definition 12. ***(Constraint Satisfaction)*** *A positive constraint is satisfied if and only if the two elements that it binds are both accepted or both rejected, noted $Sat^+(x,y) \equiv (x,y) \in \mathcal{C}^+ \wedge [(x \in \mathcal{A} \wedge y \in \mathcal{A}) \vee (x \in \mathcal{R} \wedge y \in \mathcal{R})]$. On the contrary, a negative constraint is satisfied if and only if one of the two elements that it binds is accepted and the other one rejected, noted $Sat^-(x,y) \equiv (x,y) \in \mathcal{C}^- \wedge [(x \in \mathcal{A} \wedge y \in \mathcal{R}) \vee (x \in \mathcal{R} \wedge y \in \mathcal{A})]$. Satisfied constraints within a set of elements $\mathcal{E}$ are accessed through the function:*

$$Sat : \mathcal{E} \subseteq \mathbb{E} \longrightarrow \left\{ \begin{array}{l} (x,y) \mid x,y \in \mathcal{E} \wedge \\ (Sat^+(x,y) \vee Sat^-(x,y)) \end{array} \right\} \tag{8}$$

In that context, two elements are said to be *coherent* (resp. *incoherent*) if and only if they are connected by a relation to which a satisfied (resp. non-satisfied) constraint corresponds. The main interest of this type of modelling is to allow defining a metric of cognitive coherence that permits the reification of the coherence principle in a computational calculus.

Given a partition of elements among $\mathcal{A}$ and $\mathcal{R}$, one can measure the *coherence degree* of a non-empty set of elements $\mathcal{E}$. We use $Con()$ to denote the function that gives the constraints associated with a set of elements $\mathcal{E}$. $Con : \mathcal{E} \subseteq \mathbb{E} \longrightarrow \{(x,y) \mid x,y \in \mathcal{E}, (x,y) \in \mathbb{C}\}$.

Definition 13. ***(Coherence Degree)*** *The* ***coherence degree*** *$\mathcal{C}(\mathcal{E})$, of a non-empty set of elements, $\mathcal{E}$ is obtained by adding the weights of constraints linking elements of $\mathcal{E}$ which are satisfied divided by the total weight of concerned constraints. Formally:*

$$\mathcal{C}(\mathcal{E}) = \frac{\sum_{(x,y) \in Sat(\mathcal{E})} Weight(x,y)}{\sum_{(x,y) \in Con(\mathcal{E})} Weight(x,y)} \tag{9}$$

Note that $\mathcal{C}(\mathcal{E}) \in [0,1]$ since $Sat(\mathcal{E}) \subseteq Con(\mathcal{E})$. The general social coherence problem is then:

Definition 14. ***(Coherence Problem)*** *The general* ***coherence problem*** *is to find a partition of the set of elements $\mathcal{E} \subseteq \mathbb{E}$ (i.e. instantiated s-commitments) into the set of accepted elements $\mathcal{A}$ and the set of rejected elements $\mathcal{R}$, such that, it maximizes the coherence degree $\mathcal{C}(\mathcal{E})$ of the set of elements $\mathcal{E}$.*

The coherence problem is a constraint optimization problem shown to be NP-complete [22]. The state of a social entity can be defined as follows:

Definition 15. ***(Social Entity's State)*** *A* ***social entity's state*** *is characterized by a tuple $W = \langle \mathcal{SC}, \mathcal{C}^+, \mathcal{C}^-, \mathcal{A}, \mathcal{R} \rangle$, where:*

- *$\mathcal{SC}$ is a set of elements that stand for the social entity's agenda, that stores all the social commitments from which the social entity is either the debtor or the creditor;*
- *$\mathcal{C}^+$ (resp. $\mathcal{C}^-$) is a set of non-ordered positive (resp. negative) binary constraints over $\mathcal{SC}$ such that $\forall (x,y) \in \mathcal{C}^+ \cup \mathcal{C}^-$, $x \neq y$;*
- *$\mathcal{A}$ is the set of accepted elements and $\mathcal{R}$ the set of rejected elements and $\mathcal{A} \cap \mathcal{R} = \emptyset$ and $\mathcal{A} \cup \mathcal{R} = \mathcal{SC}$.*

Finally, the overall degree of *social coherence* of an organization can be formally defined as follows:

Definition 16. ***(Organization's Social Coherence)*** *The degree of* ***social coherence*** *of an organization o is calculated over the set of elements $\mathcal{E}_{int} \cup \mathcal{E}_{ext} \subseteq \mathbb{E}$, where:*

- *$\mathcal{E}_{int}$ is the set of instantiated s-commitments where both the* debtor *and the* creditor *are members of organization o; and*
- *$\mathcal{E}_{ext}$ is the set of instantiated s-commitments where either the* debtor *or the* creditor *(XOR) is member of organization o.*

4.2 Constraints generation

Our social coherence model does provide a systematic mechanism for generating the constraints between social commitments. Our approach draws from TÆMS [11], a domain-independent framework for environment centred analysis and design of coordination mechanisms. This very well studied

Table 1: Weights and precedence order between *hard* and *soft* constraints.

Hard Constraints		Soft Constraints	
Disabling	$w = 3$	Hindering	$w = 1$
Overlapping	$w = 2.5$	Facilitating	$w = 1$
Enabling	$w = 2$		

framework, provides a comprehensive taxonomy of elements (i.e. tasks, methods, resources) and their interrelationships for modelling open MAS. We adapted their taxonomy of constraints between tasks and constraint precedence to generate constraints between action commitments, as follows:

1. *Disabling.* Given two distinct instances of social commitments c_i, c_j, involving primitive actions α_i, α_j respectively, such that (i) there is a strict *ordering constraint* $c_i \prec c_j$ (resp. $c_j \prec c_i$); and (ii) the execution of α_i (resp. α_j) *disables* α_j (resp. α_i); we say there is a negative constraint $c^-_{ij} \in \mathcal{C}^-$ between c_i and c_j.
2. *Overlapping (duration).* Given two distinct instances of s-commitments c_i, c_j involving the same *debtor*, such that $c_i \not\prec c_j$; we say there is a negative constraint $c^-_{ij} \in \mathcal{C}^-$ between c_i and c_j.[3]
3. *Enabling.* Given two distinct instances of social commitments c_i, c_j, involving primitive actions α_i, α_j respectively, such that (i) there is a strict *ordering constraint* $c_i \prec c_j$ (resp. $c_j \prec c_i$); and (ii) the execution of α_i (resp. α_j) *enables* α_j (resp. α_i); we say there is a positive constraint $c^+_{ij} \in \mathcal{C}^+$ between c_i and c_j.
4. *Hindering.* Given two distinct instances of social commitments c_i, c_j, involving primitive actions α_i, α_j respectively, such that (i) there is a strict *ordering constraint* $c_i \prec c_j$ (resp. $c_j \prec c_i$); and (ii) the execution of α_i (resp. α_j) somewhat *diminishes* the way (e.g. cost, duration) α_j (resp. α_i) can get executed; we say there is a negative constraint $c^-_{ij} \in \mathcal{C}^-$ between c_i and c_j.
5. *Facilitating.* Given two distinct instances of social commitments c_i, c_j, involving primitive actions α_i, α_j respectively, such that (i) there is a strict *ordering constraint* $c_i \prec c_j$ (resp. $c_j \prec c_i$); and (ii) the execution of α_i (resp. α_j) somewhat *improves* the way (e.g. cost, duration) α_j (resp. α_i) can get executed; we say there is a positive constraint $c^+_{ij} \in \mathcal{C}^+$ between c_i and c_j.

We assign weights to *hard* (i.e. *disabling*, *overlapping*, *enabling*) and *soft* (i.e. *facilitating*, *hindering*) constraints to capture the degree of importance of underlying relations between social commitments (Table 1). The constraints are generated automatically at instantiation time based on the constraints between actions (See Example 1, Formula 13). As can be expected *hard* constraints always have a higher precedence than *soft* ones. Note that, *hard* constraints have a strict ordering while *soft* constraints have the same precedence.

4.3 Local search algorithm

Decision theories as well as micro-economical theories define utility as a property of some valuation functions over some states of interest (e.g. consumption bundles, outcome of actions, state of the world). A function is a *utility function* if and only if it reflects the agent's preferences over these states. In our model, according to the afore-mentioned *coherence principle*, *social coherence* is preferred to *incoherence* which allows us to define the following expected utility function:

Algorithm 1 Recursive Local Search Algorithm

Function LocalSearch(W)

Require: $W = \langle \mathcal{SC}, \mathcal{C}^+, \mathcal{C}^-, \mathcal{A}, \mathcal{R} \rangle$; // current agent state
Ensure: List: *Change*; // ordered list of elements to change
  Local:
  Float: $G, Gval, C, Cval$; // expected utility value of best move
  Set: $\mathcal{A}', \mathcal{R}'$;
  Element: y, x;
  State: J; // agent state buffer

for all $x \in \mathcal{SC}$ **do**
  if $x \in \mathcal{A}$ **then**
    $\mathcal{A}' \leftarrow \mathcal{A} - \{x\}$; $\mathcal{R}' \leftarrow \mathcal{R} \cup \{x\}$;
  end if
  $W' \leftarrow \langle \mathcal{SC}, \mathcal{C}^+, \mathcal{C}^-, \mathcal{A}', \mathcal{R}' \rangle$;
  // expected utility of flipping x with transition z
  $G \leftarrow C(W') - C(W) - Res(x, z)$;
  $C \leftarrow C(W') - C(W)$; // pure coherence gain
  if $G > Gval$ **then**
    $J \leftarrow W'$; $y \leftarrow x$; $Gval \leftarrow G$; $Cval \leftarrow C$;
  end if
end for
if ($Cval < 0$ and $Gval < 0$) **then**
  return *Change*; // stop when coherence is not raising anymore and the expected utility is not positive
else
  $Dialogue(y)$;
  Update ($Res(y)$); Add (J, *Change*);
  LocalSearch(J); // recursive call
end if

DEFINITION 17. ***(Expected Utility Function)*** *The* ***expected utility*** *for an agent to attempt to reach the state* W'[4] *from the state* W *(which only differs by the change of state of* ***one*** *s-commitment* X *through the consumption of transition* Z*) is expressed as the difference between the incoherence before and after this change minus the cost of the change (expressed in term of the resistance to change of the modified s-commitment for the given transition, that is in term of sanctions). Formally:*

$$G(W') = C(W') - C(W) - \sum_{X \in \mathbb{E}, Z \in \mathbb{T}} Res(X, Z) \qquad (10)$$

Note that, our expected utility function does not include any probabilities. This reflects the case of equi-probability in which the agent has no information about the probabilities that an actual change of the social commitment will occur. For now, agents do not take into account any uncertainty measures into their coherence calculus. For example, they do not have knowledge of their own *reliability*, nor about others'. Since *sanction policies* provide the *social control* mechanisms to regulate the enforcement of social commitments; Formula (10) explicitly integrates *social control* into the coherence calculus.

At each step of his reasoning, an agent will search for a cognition acceptance state change which maximizes this expected utility. That is, the agent will attempt to change an instantiated social commitment that maximizes the utility value through dialogue. A recursive version of the local search algorithm the agents use to maximize their *social coherence* is presented in Algorithm 1. While this is an approximation algorithm for solving the *coherence problem* (Def. 14), it behaved optimally on tested examples. Since it does not make any backtracking, the worst-case complexity of this algorithm is polynomial: $\mathcal{O}(mn^2)$, where n is the

[3] Note that, we make the assumption that for any agent, two instantiated s-commitments whose time intervals overlap are negatively constrained (i.e., agents do not multi-task).

[4] See Definition 15.

number of elements considered and m the number of constraints that bind them.[5]

Note that, we have no need to encode agents' behaviour as it automatically emerges from the coherence calculus. Although the model provides a computational metric for measuring organizational coherence (Def. 16), the overall behaviour of the system is solely driven by the local behaviour of agents. That is, macro-level social order is a coherence-driven emergent phenomena resulting from the local *cognitive coherence* of interacting agents.

5. EXAMPLE: PIZZA DELIVERY DOMAIN

EXAMPLE 1. *Lets consider a domain involving a pizza delivery organization* Ω*; social roles* $\{r_k = \text{cook}\}$, $\{r_{dp} = \text{delivery-person}\}$, $\{r_{mt} = \text{maintenance- technician}\}$, *and* $\{r_c = \text{customer}\}$*; and agents* $\{ag_1, ag_2, ag_3, ag_4, ag_5\}$ *as follows:*

- *Primitive actions* (Def. 1)*:*

$$\mathcal{X} = \left\{ \begin{array}{l} \alpha_1 = \langle \textbf{order-pizza}(ag_i : r_c, pid), 1\rangle, \\ \alpha_2 = \langle \textbf{cook-pizza}(ag_i : r_k, pid), 7\rangle, \\ \alpha_3 = \langle \textbf{clean-oven}(ag_i : r_k, oid), 5\rangle, \\ \alpha_4 = \langle \textbf{pack-pizza}(ag_i : r_{dp}, pid), 2\rangle, \\ \alpha_5 = \langle \textbf{deliver-pizza}(ag_i : r_{dp}, c, pid), 20\rangle, \\ \alpha_6 = \langle \textbf{pay-order}(ag_i : r_c, ag_j : r_{dp}, price, pid), 1\rangle, \\ \alpha_7 = \langle \textbf{repair-oven}(ag_i : r_{mt}, oid), 30\rangle, \end{array} \right\} \tag{11}$$

- *Exogenous events* (Def. 2)*:*

$$\hat{\mathcal{X}} = \left\{ \begin{array}{l} \hat{\alpha}_8 \; = \langle \textbf{break-oven}^{exog}(oid), 200\rangle, \\ \hat{\alpha}_9 \; = \langle \textbf{make-oven-dirty}^{exog}(oid), 100\rangle, \\ \hat{\alpha}_{10} = \langle \textbf{become-hungry}^{exog}, 20\rangle \end{array} \right\} \tag{12}$$

- *Constraints between actions* (Section 4.2)*:*

$$\mathcal{X}_{cons} = \left\{ \begin{array}{l} \textbf{order-pizza} \; \textit{enables} \; \textbf{cook-pizza}, \\ \textbf{break-oven}^{exog} \; \textit{disables} \; \textbf{cook-pizza}, \\ \textbf{make-oven-dirty}^{exog} \; \textit{hinders} \; \textbf{cook-pizza}, \\ \textbf{clean-oven} \; \textit{disables} \; \textbf{cook-pizza}, \\ \textbf{repair-oven} \; \textit{disables} \; \textbf{cook-pizza}, \\ \textbf{cook-pizza} \; \textit{enables} \; \textbf{delivery-pizza}, \\ \textbf{delivery-pizza} \; \textit{enables} \; \textbf{pay-order}, \end{array} \right\} \tag{13}$$

- *Organization* (Def. 7)*:*

$$\Omega = \left\langle \begin{array}{l} \{ag_1, ag_2, ag_3, ag_4\}, \\ \{r_k, r_{dp}, r_{mt}\}, \\ \{(ag_1, r_k), (ag_2, r_{dp}), \\ \;(ag_3, r_{dp}), (ag_4, r_{mt})\} \end{array} \right\rangle \tag{14}$$

- *Social roles* (Def. 8)*:*[6]

$$Roles = \left\{ \begin{array}{l} r_k = \langle \{\alpha_2, \alpha_3\}, \{\bar{c}_1, \bar{c}_2, \bar{c}_5, \bar{c}_6\}\rangle, \\ r_{dp} = \langle \{\alpha_4, \alpha_5\}, \{\bar{c}_1, \bar{c}_2, \bar{c}_3, \bar{c}_4, \bar{c}_6\}\rangle, \\ r_{mt} = \langle \{\alpha_7\}, \{\bar{c}_4\}\rangle, \\ r_c = \langle \{\alpha_1, \alpha_6\}, \{\bar{c}_3, \bar{c}_4\}\rangle \end{array} \right\} \tag{15}$$

- *Agents* (Def. 9)*:*

$$Ag = \left\{ \begin{array}{l} ag_1 = \langle \{r_k\}, \{(\alpha_2, 1), (\alpha_3, 1)\}\rangle, \\ ag_2 = \langle \{r_{dp}\}, \{(\alpha_4, 1), (\alpha_5, 1)\}\rangle, \\ ag_3 = \langle \{r_{dp}\}, \{(\alpha_4, 1), (\alpha_5, 1)\}\rangle, \\ ag_4 = \langle \{r_{mt}\}, \{(\alpha_7, 1)\}\rangle, \\ ag_5 = \langle \{r_c\}, \{(\alpha_1, 1), (\alpha_6, 1)\}\rangle \end{array} \right\} \tag{16}$$

- *Social commitment schemes* (Def. 3)*:*

$$\mathcal{SC} = \left\{ \begin{array}{l} \bar{c}_1 = \alpha_1 \Rightarrow C(r_k, r_{dp}, \alpha_2, 8, [0,0,0], [0]) \\ \bar{c}_2 = \alpha_2 \Rightarrow C(r_{dp}, r_k, \alpha_4, 3, [0,0,0], [0]) \\ \bar{c}_3 = \alpha_4 \Rightarrow C(r_{dp}, r_c, \alpha_5, 21, [0,0,0], [0]) \\ \bar{c}_4 = \alpha_5 \Rightarrow C(r_c, r_{dp}, \alpha_6, 2, [0,0,0], [0]) \\ \bar{c}_5 = \hat{\alpha_8} \Rightarrow C(r_{mt}, r_k, \alpha_7, 31, [0,0,0], [0]) \\ \bar{c}_6 = \hat{\alpha_9} \Rightarrow C(r_k, r_{mt}, \alpha_3, 6, [0,0,0], [0]) \end{array} \right\} \tag{17}$$

[5] n coherence calculus (sum over m constraints) for each level and a maximum of n levels to be searched.

[6] Roles $\{r_k = \textit{cook}\}$, $\{r_{dp} = \textit{delivery-person}\}$, and $\{r_{mt} = \textit{maintenance-technician}\}$ are part of Ω, but role $\{r_c = \textit{customer}\}$ is external to the organization.

This example comprises 1 *cook* agent (ag_1), 2 *delivery-person* agents (ag_2, ag_3), 1 *maintenance-technician* agent (ag_4), and 1 *customer* agent (ag_5). Note that, the social commitment schemes in Formula (17) implicitely define the following pizza delivery workflow: **order-pizza** → **cook-pizza** → **pack-pizza** → **deliver-pizza** → **pay-order**; which is initiated when exogenous event **become-hungry**^{exog} occurs, making the *customer* agent perform the action **order-pizza**.

6. INITIAL VALIDATION

A *SC-sim* simulator has been implemented as a Java applet, which provides some flexibility in terms of deployment and facilitates sharing results with the research community. To illustrate the use of the model and the simulator, we introduce a simple experiment involving two *sanctions policies*:

- ***SPol*** $\emptyset$. $\mathcal{S}_d = \{0, 0, 0\}$, and $\mathcal{S}_c = \{0\}$. Debtors receive no rewards. Both debtors and creditors have no penalties. This policy entails no social control; and

- ***SPol 1***. $\mathcal{S}_d = \{0, -1, -1\}$, and $\mathcal{S}_c = \{-1\}$. Debtors receive no rewards and high violation penalties. Both debtors and creditors have high cancellation penalties.

Experiment. We ran the experiment on the pizza delivery domain presented in Example 1. We varied the *periodicity* (Def. 2) of the exogenous event **become-hungry**^{exog} (starting from 80 time steps, down to 40, 20, 10, 5, 2, and 1 time steps). As a result, the *customer* agent starts placing orders more frequently. Note that we assume neither agents, nor actions can fail. We measured the overall *efficiency* (i.e. percentage of s-commitments fulfilled) of the system. For each parametrization, we ran 15 simulations of 750 time steps each and computed the standard sample mean. Figure 2 presents the results.

Observation 1. As expected, the *efficiency* of the organization degraded from nearly optimal as the *frequency* of orders and the corresponding *level of activity* (i.e. number of s-commitments per agent per time step, not shown here) was increased.

Observation 2. We can observe drastic differences between the evaluated policies. These two *sanction policies* had a distinct effect on the performance of the system. Under policy ***SPol 1*** the organization was more *efficient* than without having any *social control* (i.e., ***SPol*** $\emptyset$).

Observation 3. Desirable (sometimes nearly optimal) agent behaviour results from local *coherence maximization*, without explicitly encoding agents behaviour. More importantly, macro-level *social coherence* does emerge from local *coherence maximization*.

Although this paper focuses on presenting the model, we think these experimental results are encouraging as they provide some preliminary validation. Of course, there is still much work to be done in terms of running more experiments, analyzing results and evaluating the scalability of the model. Our work takes on the problem of modelling *desirable and (relatively) predictable emergent social behaviour* from the local actions of the agents [5]. Observation 3, provides some preliminary evidence to support the suitability of our model for running social simulations, where complex *emergent social patterns* can be obtained and reproduced from the dynamics of local interactions among agents. There is

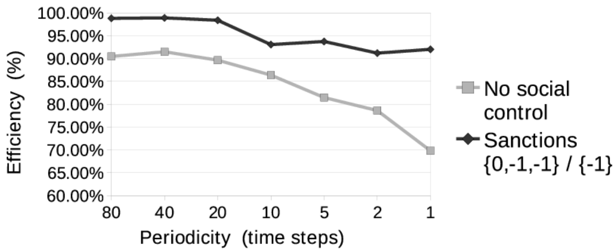

Figure 2: Experimental Results (Efficiency %).

complexity happening that cannot be fully explained analytically, thus justifying an empirical simulation-based approach. Similarly, Observation 2 provides some evidence to support the effectiveness of integrating *social control* mechanisms (for the enforcement of *s-commitments*), into the coherence calculus. Note that, when agents have neither positive, nor negative incentives their local *coherence-driven deliberation* might eventually lead them to unilaterally cancelling, or even violating *social commitments* as there are no consequences. Some authors have suggested [5] that *social cooperation* does not necessarily require an agent's understanding, agreement, nor even awareness. Our proposal aligns with this view, and Observation 2 shows that we are able to re-produce *desirable cooperation-like behaviour*, through the implementation of an appropriate *sanction policy* (e.g., ***SPol 1***).

7. DISCUSSION AND RELATED WORK

There have been several approaches [4, 20, 16, 2] to formalizing *social commitments*. The proposal of Carabelea and Boissier [2] relies on *social commitments* for coordinating agents within the context of organizational interactions. Like us, they do define *social entities* and *organizational structures* entirely based on *social commitments*. However, in our proposal all the dynamics of *social commitments* are captured by a generic state-transition model which is associated with *social control* mechanisms for the enforcement of *social commitments*. In addition, we choose not to explicitly specify *authority relations* between roles. Instead, we capture them as implicitly resulting from *social commitments schema* associated with roles. Thus, we can get a more compact representation without compromising expressiveness.

One other coherence-based framework inspired by early work on cognitive coherence in MAS [15] has been proposed by Joseph et al. [18, 19]. Their framework builds on the BDI model of agency and the coherence theory [21]. Their approach, is also based on a *coherence maximization* model of agent rationality implemented as a constraint satisfaction problem. However, their proposal substantially differs from ours as (i) their main motivation seems to be the study of the interactions between the agent's internal cognitions (BDI) and some social aspects of MAS such as: norm evaluation [19], and the behaviour of institutional agents [18]; and (ii) their approach uses coherence graphs to represent each BDI modality resulting in a more complex model that has not been validated nor implemented and thus does not allow to derive new knowledge. In contrast, we are interested in modelling and evaluating the general dynamics of social systems. We claim that our model not only is more compact and decreases the computational overhead incurred when calculating coherence, but also is expressive enough to represent complex social systems. Although, in this paper, we do not consider *social norms*, we can certainly model them by representing s-commitments from a role towards an organization.[7]

Other organizational approaches to social modelling have been reported in the literature [17, 23]. The former, is a knowledge-based approach to automated organizational design, which enables efficient role selection to match organizational goals, as well as agent-to-role allocation. Like us, they define organizational structures in terms of *agents* enacting *roles* in *organizations*. However, their focus is on designing effective organizations which can change forms depending of varying performance requirements. Instead, our *simulation framework* focuses on evaluating the emergent social dynamics and performance of multi-agent organizations from the local *coherence*-driven interactions among agents. The latter proposal [23], presents an agent-oriented language (endowed with an operational semantics) for developing multi-agent organizations. *Organizations* are defined in terms of *roles*, *norms*, and *sanctions*. Although structurally close from an abstract organizational standpoint; our models also differ as theirs specify *roles* in terms of the same mental attitudes attributed to BDI agents. Instead, we define *roles* in terms of *capabilities* which can be enacted by *agents*. Moreover, the *state* of a *role* makes no reference to mental attitudes.

8. CONCLUSIONS AND FUTURE WORK

In this paper, we presented a simple operational model capturing the dynamics of social systems. Our work advances the state of the art by proposing a unified yet computational view of the social aspects of multi-agent organizations. In previous work [15], we proposed a constraint satisfaction-based model of *cognitive coherence* within the context of agent communication pragmatics. Here, we built on this work and extended it to consider *social coherence*. We introduced the notion of *social coherence* as the main social organizing principle in MAS. Moreover, our model relies on the notion of *social commitment* to represent all the inter-dependencies between *social entities*. Together with the notion of *sanction policy*, *social coherence* reify the notion of *social control*. In our model, *social control* is actually integrated into the coherence calculus (Def. 11). Local *coherence* is the driving force that organizes agents' behaviour and from which *social coherence* emerges. Finally, we illustrated our model and simulator by running a simple experiment to investigate the effects of two *social control* mechanisms (reified by *sanction policies*) on a sample domain.

As future work, we will refine our model using an action language such as *event calculus* [12]. We want to evaluate the benefits of introducing a more comprehensive treatment of *time*, as well as reasoning about *actions*. We also plan to address the issue of handling *complex actions*. Another immediate extension to our model will be the introduction of *uncertainty reasoning* into the *coherence calculus*. For now, agents do not take into account any uncertainty measures. Since both *actions* and *agents* can fail (as reflected by the

[7]An *institution* can be seen as a particular type of organization.

reliability probability value in Def. 9), agents should be able to incorporate these information into their expected utility calculus. Agents with different levels of knowledge should also be modelled, such as: agents with no knowledge, with partial knowledge, or with complete/shared knowledge. Furthermore, various machine learning mechanisms would allow agents to progressively learn these probabilities.

Finally, we want to run more experiments and evaluate the scalability of our model. For instance, we should model social domains with multiple organizations and greater number of agents, where agents can play several roles possibly in different organizations. We also want to investigate how our coherentist approach might be used to evaluate the functionality and behaviour of typical *organizational structures* reported in the literature (e.g., hierarchies, holarchies, societies, federations [10]). Furthermore, since no single organizational design is suitable for all domain applications we want to cross-validate our model by running simulations involving different organizational structures. Last but not least, we want to continue studying the effects of *social control* mechanisms.

9. ACKNOWLEDGMENTS

The authors would like to thank Marek Hatala and the anonymous reviewers for their useful comments.

10. REFERENCES

[1] S. Bandini, S. Manzoni, and G. Vizzari. Agent based modeling and simulation. In Meyers [14], pages 184–197.

[2] C. Carabelea and O. Boissier. Coordinanting Agents in Organizations Using Social Commitments. In *Proceedings of the 1st International Workshop on Coordination and Organisation*, Namur, April 2005.

[3] K. M. Carley and L. Gasser. *Multiagent Systems: a modern approach to distributed Artificial Intelligence*, chapter Computational Organization Theory, pages 299–330. MIT Press, 2001.

[4] C. Castelfranchi. Commitments: from Individual Intentions to Groups and Organizations. In *Proceedings of ICMAS95*, pages 41–48, June 1995.

[5] C. Castelfranchi. Engineering Social Order. In *Proceedings of the First International Workshop on Engineering Societies in the Agents World*, volume 1972, pages 1–18, 2000.

[6] P. Davidsson and H. Verhagen. Social phenomena simulation. In Meyers [14], pages 8375–8379.

[7] R. A. Flores and R. C. Kremer. To Commit or Not to Commit: Modeling Agent Conversations for Action. *Computational Intelligence*, 18(2):120–173, 2002.

[8] N. Fornara and C. Colombetti. Operational Specification of a Commitment-Based Agent Communication Language. In C. Castelfranchi and W. L. Johnson, editors, *Proceeding of the First Autonomous Agents and Multi-Agents Systems Joint Conference (AAMAS'02)*, volume 2, pages 535–543. ACM Press, 2002.

[9] M. Hechter and K. D. Opp. Introduction. *Social Norms*, pages xi–xx, 2001.

[10] B. Horling and V. Lesser. A Survey of Multi-Agent Organizational Paradigms. *The Knowledge Engineering Review*, 19(4):281–316, 2005.

[11] B. Horling, V. Lesser, R. Vincent, T. Wagner, A. Raja, S. Zhang, K. Decker, and A. Garvey. The TAEMS White Paper, 1999.

[12] R. A. Kowalski and M. Sergot. A Logic-Based Calculus of Events. *New Generation Computing*, 4:67–95, 1986.

[13] D. Martindale. *Social Control for the 1980s: A Handbook for Order in a Democratic Society*, chapter The Theory of Social Control, pages 46–58. Greenwood Press, 1978.

[14] R. A. Meyers, editor. *Encyclopedia of Complexity and Systems Science*. Springer, 2009.

[15] P. Pasquier and B. Chaib-draa. The Cognitive Coherence Approach for Agent Communication Pragmatics. In *Proceedings of The Second International Joint Conference on Autonomous Agent and Multi-Agents Sytems (AAMAS'03)*, pages 544–552, Melbourne, 2003.

[16] P. Pasquier, R. A. Flores, and B. Chaib-draa. Modelling Flexible Social Commitments and their Enforcement. In *Proceedings of the Fifth International Workshop Engineering Societies in the Agents World (ESAW'04)*, volume 3451 of *Lecture Notes in Artificial Intelligence*, pages 153–165. Springer-Verlag, 2004.

[17] M. Sims, D. Corkill, and V. Lesser. Automated Organization Design for Multi-agent Systems. *Autonomous Agents and Multi-Agent Systems*, 16(2):151–185, 2008.

[18] J. Sindhu, C. Sierra, and M. Schorlemmer. A Coherence Based Framework for Institutional Agents. In *Proceedings of the Fifth European Workshop on Multi-Agent Systems (EUMAS'07)*, December 2007.

[19] J. Sindhu, C. Sierra, M. Schorlemmer, and P. Delunde. Formalizing Deductive Coherence: An Application to Norm Evaluation. *Logic Journal of the Interest Group in Pure and Applied Logic (IGPL)*, 2008.

[20] M. P. Singh. An Ontology for Commitments in Multiagent Systems: Toward a Unification of Normative Concepts. *Artificial Intelligence and Law*, 7:97–113, 1999.

[21] P. Thagard. *Coherence in Thought and Action*. MIT Press, 2000.

[22] P. Thagard and K. Verbeurgt. Coherence as Constraint Satisfaction. *Cognitive Science*, 22:1–24, 1998.

[23] N. Tinnemeier, M. Dastani, and J.-J. Meyer. Roles and Norms for Programming Agent Organizations. In *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems*, pages 121–128, Richland, SC, 2009.

[24] D. N. Walton and E. Krabbe. *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning.* Suny Press, 1995.