Maintenance of Social Commitments in Multiagent Systems

Pankaj Telang[1] **Munindar P. Singh**[2] **Neil Yorke-Smith**[3]
[1] SAS Institute, Cary, NC 27513, USA
[2] North Carolina State University, Raleigh, NC 27695, USA
[3] Delft University of Technology, Delft, 2600GA, The Netherlands
ptelang@gmail.com, singh@ncsu.edu, n.yorke-smith@tudelft.nl

Abstract

We introduce and formalize a concept of a maintenance commitment, a kind of social commitment characterized by states whose truthhood an agent commits to maintain. This concept of maintenance commitments enables us to capture a richer variety of real-world scenarios than possible using achievement commitments with a temporal condition. By developing a rule-based operational semantics, we study the relationship between agents' achievement and maintenance goals, achievement commitments, and maintenance commitments. We motivate a notion of coherence which captures alignment between an agents' achievement and maintenance cognitive and social constructs, and prove that, under specified conditions, the goals and commitments of both rational agents individually and of a multiagent system are coherent.

1 Introduction

Social commitments enable flexible coordination between agents. Research has primarily focused on achievement commitments (Castelfranchi 1995; Singh 1991, 2012).

Consider an agent, such as an aircraft operator, who wishes to maintain a condition, such as an aircraft being fit to fly. Fig. 1 shows a high-level process flow of aerospace aftermarket services (van Aart et al. 2007). Its participants are an airline operator (OPER), an aircraft engine manufacturer (MFR), and a parts manufacturer (PMFR). Such situations highlight the need for understanding *maintenance*. We motivate a new family of social commitments wherein a *debtor* agent commits to a *creditor* agent that if some *antecedent* condition holds it would *maintain* a *consequent* condition until some *discharge* condition holds. Maintenance here means ensuring that the condition does not become false or, if it does become false, then to re-establish its truthhood. Specifically, we address how maintenance arises in connection with goals and commitments, as needed for multiagent systems. In this manner, our work contrasts with previous work on maintenance, which emphasizes single-agent settings and primarily addresses maintenance goals.

Commitments are a natural basis for modelling interaction between agents by representing the meanings of communication (Chocron and Schorlemmer 2018; Mallya and Singh 2007; Singh 2000) and for reasoning about safety and control (Marengo et al. 2011). Understanding maintenance commitments opens up the realm of social interaction. For example, maintaining a green lawn or paying down a mortgage are naturally modelled as maintenance commitments.

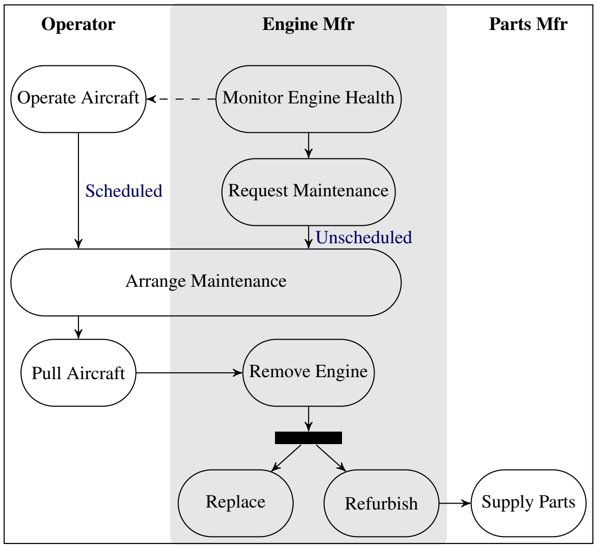

Figure 1: Aerospace aftermarket (van Aart et al. 2007).

A major theme of this paper is capturing the dynamic relationships between an agent's beliefs and goals (i.e., cognitive state) and its commitments (i.e., social state). Maintenance commitments enable richer relationships than otherwise possible, thereby supporting expanded forms of collaboration. Specifically, a commitment can relate to both end and means goals. For example, (1) an end goal of paying for a house may lead you to a maintenance commitment of mortgage: paying down a loan incrementally until it's paid up. (2) The mortgage commitment may lead you to a maintenance (means) goal of making loan payments, which (3) may lead you to commit to doing a job to get paid every month.

Prior research has not adequately addressed maintenance commitments, treating them instead as achievement commitments for temporal formulae of the form "always in the future p" (Mallya, Yolum, and Singh 2003; Chesani et al. 2013). Such a formulation cannot capture the lawn example above. In contrast, we treat maintenance commitments as a

---

first-class construct, accommodating both reactive and proactive interpretations, and incorporating enactments wherein the condition may be negated and resurrected.

Recent research characterizes *coherence* between a rational agent's (achievement) commitments and its (achievement) goals (Telang, Singh, and Yorke-Smith 2019). Building on this approach, we motivate an enhanced notion of coherence and with it study the synergy between an agent's maintenance commitments and its achievement and maintenance goals. We also show how coherence applies to a multiagent system as a whole with respect to specific maintenance commitments and achievement and maintenance goals.

Our formal operational semantics develops two sets of conditional rules. First, *life cycle* rules specify the mandatory progression of goal and commitment states as the agents update their beliefs or perform 'social actions' on the goals and commitments. Second, *practical* rules represent patterns of reasoning that specify potential social actions for an agent based on its goals and commitments.

This paper advances the state of the art as follows. First, we introduce a new powerful type of social commitment, along with its life cycle. Second, we specify a formal semantics of multiagent system configuration, encompassing both achievement and maintenance commitments and goals. Third, we provide a methodology and an exemplar set of practical rules. Fourth, we define coherence and prove conditions under which it is maintained in the multiagent system.

2 Preliminaries and Necessary Background

We suppose a finite set of agents, $x_1, x_2, \ldots \in \mathscr{A}$, and a finite set of propositional atoms, $a_1, a_2, \ldots \in \Omega$. We write Ψ for the set of all propositional formulae over Ω. The symbol $\top$ abbreviates $a \vee \neg a$ for any atom a, and the symbol $\bot$ abbreviates $\neg\top$. We assume classical propositional logic. Specifically, given a set of propositions $\Phi \subseteq \Psi$ and a proposition $\psi \in \Psi$, $\Phi \models \psi$ denotes that Φ entails ψ. We say that a set of propositions Φ is *consistent* iff $\Phi \not\models \bot$.

2.1 Beliefs

Definition 1. *A* belief state function $\mathcal{B} : \mathscr{A} \times \Psi \rightarrow \{\top, \bot\}$ *applies to agent-atom pairs and returns* $\top$ *exactly if the agent believes the atom. We lift* $\mathcal{B}$ *to all propositions and close it under entailment: if* $\mathcal{B}(x, p)$ *and* $p \models q$ *then* $\mathcal{B}(x, q)$*. An agent's beliefs are consistent:* $\neg\mathcal{B}(x, \bot)$*. We write* $\mathcal{B}_x = \{p \in \Psi : \mathcal{B}(x, p)\}$ *for the set of all current beliefs of* x*.*

An agent x may have no belief about p or $\neg p$, meaning that $B(x, p) = \bot$ and $B(x, \neg p) = \bot$ can coexist.

Definition 2. *The* belief addition function $+ : \mathcal{B} \times \Psi \rightarrow \mathcal{B}$ *adds a belief to the belief set.*

E.g., $\mathcal{B}' = +(\mathcal{B}, p)$ means p is added to $\mathcal{B}$; hence $\mathcal{B}'(x, p) = \top$.

2.2 Achievement Commitments and Goals

We adopt achievement commitment and achievement goal as defined by Telang, Singh, and Yorke-Smith (2019), hereinafter TSY, denoting achievement commitments by C and achievement goals by G. We adopt and enhance TSY's definitions of state functions, maximal sets, and consistency.

2.3 Maintenance Goals

We define a maintenance goal and its life cycle based on Duff, Thangarajah, and Harland (2014). Let $M = \mathsf{M}(x, m, s, f)$ be agent x's *maintenance goal* for condition m. M's success condition is s and its failure condition f. $\mathcal{B}_x \models \neg m$ means that x believes that m is false: thus, x adopts a recovery achievement goal. Let π_x capture x's lookahead mechanism (provided by the agent designer), independent of M (Duff, Harland, and Thangarajah 2006). Then, $\mathcal{B}_x \models m \wedge \pi_x(\neg m)$ means that x believes that m will become false unless it acts appropriately: thus, x adopts a preventive achievement goal.

Figure 2 depicts a state-based life cycle for a maintenance goal. We do not include the state `Suspended` from Duff et al., for reasons explained in the next section. Note the labels denote events (e:fail, e:succeed) and actions (all others). The actions are performed by the agent whereas the events are observed in the environment.

Definition 3. *A* maintenance goal *is a tuple* $\langle x, m, s, f\rangle$*, where* $x \in \mathscr{A}$ *is an agent, and* $m, s, f \in \Psi$ *are the goal's* maintenance*,* success*, and* failure *conditions, respectively, where* $s \wedge f \models \bot$ *and* $m \wedge f \models \bot$*.*

We write a maintenance goal as $M = \mathsf{M}(x, m, s, f)$.

3 Maintenance Commitments

Informally, in a maintenance commitment, debtor x commits to creditor y that if the antecedent holds true, then until the discharge condition d becomes true, agent x will sustain the maintenance condition m. The symbol S stands for *sustains*.

Definition 4. *A* maintenance commitment *is a tuple* $\langle x, y, l, m, d\rangle$*, where* $x, y \in \mathscr{A}$ *are agents,* $x \neq y$*, and* $l, m, d \in \Psi$ *are formulas. We call* x *and* y debtor *and* creditor *of this commitment and call* l*,* m*, and* d *its* antecedent*,* maintenance condition*, and* discharge condition*, respectively. We write a maintenance commitment as* $S = \mathsf{S}(x, y, l, m, d)$*. We write* $\mathcal{S}_x$ *for the set of all non-*`Null` *maintenance commitments in which agent* x *is either debtor or creditor.*

Figure 3 shows the life cycle of a maintenance commitment. Both debtor x and creditor y represent the changing states of a commitment according to this life cycle. We elide the concerns of communication and alignment for simplicity.

We define *commitment strength*, enhancing prior definitions (Singh 2008; Chopra and Singh 2009), to enable defining the important commitment closure properties below.

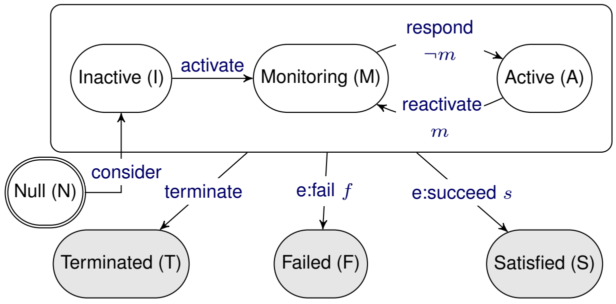

Figure 2: Life cycle of a maintenance goal.

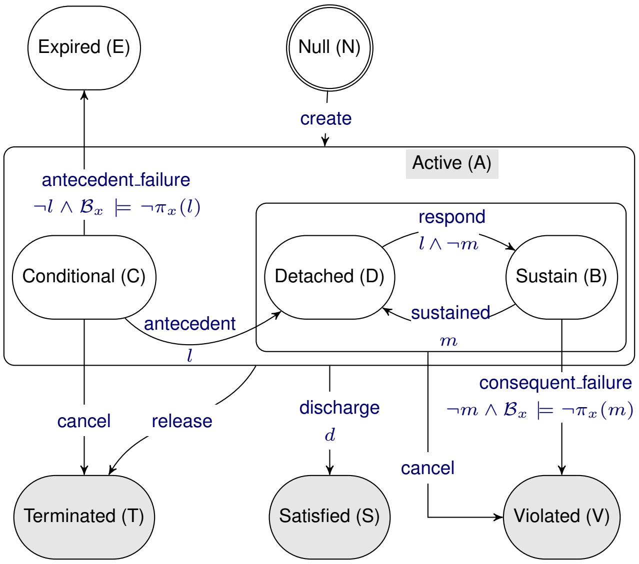

Figure 3: Life cycle of a maintenance commitment.

Definition 5. *A maintenance commitment* $S_1 = \mathsf{S}(x, y, l_1, m_1, d_1)$ *is stronger than* $S_2 = \mathsf{S}(x, y, l_2, m_2, d_2)$*, written* $S_1 \succeq S_2$ *or* $S_2 \preceq S_1$*, iff* $l_2 \models l_1$*,* $l_2 \wedge m_1 \models l_2 \wedge m_2$*, and* $l_2 \wedge d_1 \models l_2 \wedge d_2$*.*

For example, let $S_1 = \mathsf{S}(x, y, \mathsf{pay}, \mathsf{green\_lawn} \wedge \mathsf{bug\_free\_lawn}, \mathsf{year\_end})$, meaning that x commits to maintaining the yard green. Let $S_2 = \mathsf{S}(x, y, \mathsf{pay} \wedge \mathsf{provide\_fertilizer}, \mathsf{green\_lawn}, \mathsf{year\_end})$. Then S_1 is stronger than S_2. Note commitment strength is a partial order.

Definition 6. *Let* $\chi_S = \{\mathtt{N}, \mathtt{C}, \mathtt{E}, \mathtt{D}, \mathtt{B}, \mathtt{T}, \mathtt{V}, \mathtt{S}\}$ *be the set of states in Figure 3. The* maintenance commitment state function $\mathcal{S}$ *maps each maintenance commitment to a state in* χ_S*. For simplicity, we write* $\mathcal{S}(\mathsf{S}(x, y, l, m, d))$ *as* $\mathcal{S}(x, y, l, m, d)$*. This function satisfies the following closure properties:*

- *If* $\mathcal{S}(S_1) \in \{\mathtt{C}, \mathtt{S}, \mathtt{E}\}$*,* $S_1 \succeq S_2$*, then* $\mathcal{S}(S_2) = \mathcal{S}(S_1)$*.*
- *If* $\mathcal{S}(S_1) \in \{\mathtt{D}, \mathtt{B}, \mathtt{V}, \mathtt{T}\}$*,* $S_2 \succeq S_1$*, then* $\mathcal{S}(S_2) = \mathcal{S}(S_1)$*.*

The closure properties ensure that an agent configuration is semantically viable. The states assigned to maintenance commitments in any configuration respect the following property: if a commitment is in one of the states `Conditional`, `Satisfied`, and `Expired`, then so is any weaker commitment; whereas if a commitment is in one of the states `Detached`, `Sustained`, `Violated`, `Terminated`, then so is any stronger commitment. The properties are valuable for states that are not based on the content of a commitment. For example, when a commitment S is cancelled, all stronger commitments must be cancelled, else the S would be immediately resurrected from a stronger commitment. Using S_1 and S_2 as above, we see that if S_2 is in `Detached`, then so must S_1 be.

Intuitively, a *maximal maintenance commitment* w.r.t. a state σ is a commitment in state σ such that no strictly stronger maintenance commitment is in the same state σ. We express practical rules over such commitments. Below, we identify sets of maximal commitments w.r.t. sets of states.

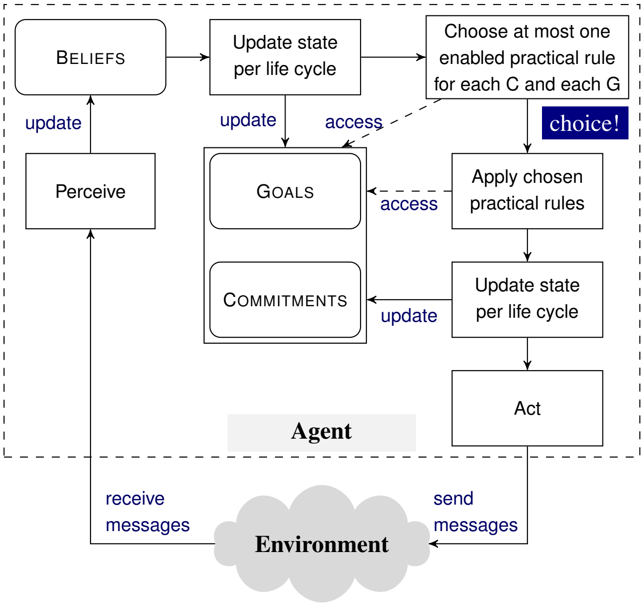

Figure 4: Simple agent architecture and operations.

Definition 7 (Maximal m-commitment set, $maxc(\cdot)$)**.** *Let* $\Sigma \subseteq \chi_S$ *be a set of maintenance commitment states.* $maxc(\Sigma) = \{S_1 \in \Sigma \mid (\forall S_2 \in \Sigma\colon S_2 \succeq S_1 \Rightarrow S_2 = S_1)\}$*.*

Some of our theorems require an agent's configuration to have *consistent maintenance commitments*. Informally, an agent will not try to maintain a condition and its complement.

4 Configurations and Life Cycle Rules

Figure 4, adapted from TSY, describes how an agent operates with respect to its beliefs, goals, and commitments. The simple agent architecture provides an illustrative context for our semantics; it is not intended to be an alternative to the fully-fledged architectures in the literature.

Although the agent may consider multiple actions, in each deliberation cycle the agent can *choose* at most one action (based on an enabled practical rule) for each commitment and goal. Suspension and reactivation of goals and commitments occurs through the operational model of Figure 4. A goal or commitment is deemed *suspended* if no practical rule pertaining to it is chosen. When the agent subsequently chooses a practical rule pertaining to that goal or commitment, that means it is reactivated. We treat the agent's operations on its cognitive and social state through our practical rules. We do not treat the agent's plans or domain actions (box 'Act'). Since we do not model an agent's domain actions, we do not reason about the agent's success or failure with its goals and commitments, just about the coherence of the goals and commitments of a single agent or of a multiagent system.

We now define the configuration of a multiagent system and begin to study its consistency according to the life cycle rules of commitments and goals.

4.1 Agent Configuration

An agent's configuration comprises elements both of its cognitive state (i.e., beliefs and goals) and its social state (i.e., commitments of which the agent is creditor or debtor).

Definition 8. *The* configuration *of an agent x is the tuple $S(x) = \langle \mathcal{B}_x, \mathcal{G}_x, \mathcal{M}_x, \mathcal{C}_x, \mathcal{S}_x \rangle$ where $\mathcal{B}_x$ is state function for x's beliefs, $\mathcal{G}_x$ and $\mathcal{M}_x$ are state functions of achievement and maintenance goals respectively, and $\mathcal{C}_x$ and $\mathcal{S}_x$ are state functions for achievement and maintenance commitments respectively, in which agent x is either debtor or creditor.*

To reduce clutter, we write the configuration of agent x as $\langle \mathcal{B}, \mathcal{G}, \mathcal{M}, \mathcal{C}, \mathcal{S} \rangle_x$ instead of $\langle \mathcal{B}_x, \mathcal{G}_x, \mathcal{M}_x, \mathcal{C}_x, \mathcal{S}_x \rangle$.

An agent's goals and commitments must be consistent with its beliefs. For example, if agent x believes in the success condition of a goal, then the goal's state must be `Null` (i.e., whereupon it is not in $\mathcal{G}_x$) or `Satisfied`. We also assume the goals are mutually consistent (Winikoff et al. 2002).

How goals and commitments cohere, within and across agents, is a main theme of this paper—see Section 5.

4.2 System Configuration and Traces

In our model, computation in a multiagent system is realized entirely in its member agents. A goal is private to an agent. Each commitment is represented by both its creditor and its debtor. For simplicity, we assume for each commitment that its creditor and debtor agree on its state.

Definition 9. *The* system configuration *of a multiagent system of agents $\mathscr{A} = x_1, \ldots, x_n$ is given by n-tuple $\langle S(1), \ldots, S(n) \rangle$, where $S(i)$ is the configuration of x_i.*

When required, we write the multiagent system configuration with each agent's configuration expanded to its beliefs, goals, and commitments: $\langle \langle \mathcal{B}, \mathcal{G}, \mathcal{M}, \mathcal{C}, \mathcal{S} \rangle_1$, $\langle \mathcal{B}, \mathcal{G}, \mathcal{M}, \mathcal{C}, \mathcal{S} \rangle_2, \ldots \langle \mathcal{B}, \mathcal{G}, \mathcal{M}, \mathcal{C}, \mathcal{S} \rangle_n \rangle$. A *trace* is a (possibly infinite) sequence of system configurations.

The rules we introduce in the coming sections apply to each agent's internal representation separately. These rules constitute a labelled transition system, with the actions being the labels and the multiagent system configuration being the state, i.e., $S \xrightarrow{\alpha} S'$ where α is an action on a cognitive or social construct. As explained above, we do not model an agent's domain actions or plans. Thus, a single transition could potentially correspond to zero or more domain actions.

4.3 Life Cycle Rules

We now define formally the life cycle of goals and commitments. For this, we need action sets for beliefs and achievement and maintenance goals and commitments For all agents combined, we define $\mathbb{B}$, $\mathbb{G}$, $\mathbb{M}$, $\mathbb{C}$, and $\mathbb{S}$ as the sets of beliefs, achievement goals, maintenance goals, achievement commitments, and maintenance commitments.

Each agent can act on its own elements of the system configuration. The *belief actions* set is BACTS $= \{+\}$. The *achievement goal actions* set is GACTS $= \{$consider-G, activate-G, terminate-G$\}$. The *maintenance goal actions* set is MACTS $= \{$consider-M, activate-M, terminate-M$\}$. The *achievement commitment actions* set is CACTS $= \{$create-C, cancel-C, release-C$\}$. The *maintenance commitment actions* set is SACTS $= \{$create-S, cancel-S, release-S$\}$.

	TO			
↓ FROM	a-goal	m-goal	a-comm	m-comm
a-goal	TSY	N/A	TSY	practical
m-goal	practical	closure	N/A	N/A
a-comm	TSY	N/A	TSY	N/A
m-comm	practical	practical	N/A	closure

Table 1: Possible interactions between components. TSY = achievement-only case: not the topic of this paper. Closure = follows from closure properties such as Def. 7. Practical = treated in this paper in Sect. 5.2. N/A = not applicable.

An *action instance* pairs an action and a corresponding belief, goal, or commitment. For example, the action instance $\langle$activate$, G_1 \rangle$ corresponds to the action of activating goal G_1. Valid action instances are consistent across the components. Where an action concerns a goal or commitment condition such as a consequent, it must be consistent with changes to the agent's beliefs, and actions corresponding to that belief change must occur on all goals and commitments. When a goal or commitment is affected, so are weaker or stronger goals and commitments to preserve consistency and closure properties, e.g., as in Definition 6.

An *action set* is a set of concurrent action instances of the same agent. We define a *life cycle* rule as a mapping from a system configuration and an action set into a resulting system configuration. We adopt TSY's life cycle for achievement goals and commitments, with `Pending` and `Suspended` removed, respectively. The life cycles of maintenance goals and commitments capture Figures 2 and 3 in logical terms.

An illustrative life cycle rules concerns agent x's maintenance goals and the belief action $\forall \langle +, b \rangle \in \mathbb{A}$, $b = \mathcal{B}(x, p)$ corresponding to x believing p. Then: if $\mathcal{M}(x, m, s, f) \in \{\mathtt{I}, \mathtt{A}, \mathtt{M}\}, p \models s$, then $\mathcal{M}'(x, m, s, f) = \mathtt{S}$. The intuition is that if $p \models s$, each maintenance goal $\mathsf{M}(x, m, s, f)$ that is `Inactive`, `Active`, `Monitoring` succeeds.

5 Relating Commitments and Goals

An agent's practical rules reflect its decision-making. To organize the practical rules, we note that beliefs do not directly give rise to actions. Therefore, we consider direct interactions between achievement and maintenance goals and commitments, giving rise to the $4^2 = 16$ possibilities in Table 1.

Practical rules capture an agent's rational behaviour, for example: an agent would adopt commitments to help achieve or maintain its end goals and given its commitments, would create means goals to satisfy or sustain them. Figure 5 captures the relationships of a maintenance commitment or goal as pairs of functions. Let S, G, M respectively be a maintenance commitment, achievement goal, and maintenance goal. Then, for instance, GAS(S) identifies achievement goals such that S's antecedent models the success condition of the goals. The goals created by the creditor to detach S are in GAS(S).

For each of these functions, we define an 'inverse' as a function in the reverse direction.

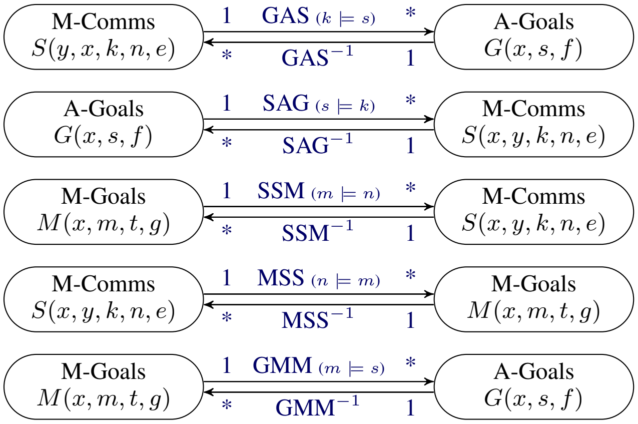

Figure 5: Functions relating (s-)comms and (m-)goals.

5.1 Coherence and Convergence

Goals and commitments, respectively, reflect the cognitive and social states of agents. How well these constructs cohere indicates how well a multiagent system is being enacted. Ideally, an agent should enter into commitments in accordance with its goals and take on goals that would lead to its commitments being satisfied or sustained. But an agent being autonomous may drop its goals and commitments arbitrarily.

We say an agent configuration is *coherent* if it satisfies the stated coherence properties over beliefs, goals, and commitments of an agent. Informally, the goals and commitments in a coherent configuration reflect the agent's rationality in that their existence may be justified based on another element. For example, when an end goal is satisfied, an agent may drop its corresponding commitment and if a means goal fails, it may adopt another goal or decide to give up on the commitment.

If a trace $S_1, S_2, \ldots$ converges infinitely often to a configuration S_k, and S_k is coherent, then the trace *sustains a coherent configuration*. Note that this repeated converge contrasts with the 'one time' convergence that is adequate for achievement commitments in TSY.

A judicious set of practical rules would ensure that goals and commitments in a multiagent system remain coherent even though the agents act autonomously. We demonstrate such a set below. As discussed at the end of the paper, our methodology is generic in that the same approach can be used for alternative sets of practical rules.

5.2 Practical Rules

We use TSY's syntax of a practical rule template of the form $E \xrightarrow{\text{RULENAME}} \alpha$. The expression E is a conjunction of the form of: this goal is (or is not) in some state and that commitment is (or is not) in some state, and about commitment and goal sets computed by the functions of Fig 5 and their states. The expression α is a commitment (or goal) action to be performed on one or more commitments (or goals). We write $ant(\cdot)$ for the antecedent of a (m-) commitment, $maint(\cdot)$ for the maintenance condition of m-comm or m-goal, and $succ(\cdot)$ for the success condition of a (m-) goal.

A-goal to m-comm We first describe the rules in which one or more m-commitments support an a-goal.

- S-CREATE: Suppose agent x has an active achievement goal $G = \mathsf{G}(x, ., .)$. Then create one or more maintenance commitments that can satisfy the goal G. Let $\omega = \bigwedge_i ant(S_i)$, where $S_i = \mathsf{S}(x, y, ., ., .)$ and $S_i \in \text{SAG}(G)$, and Φ be a set of commitments such that $\bigwedge_j ant(S_j) \wedge \omega \models succ(G)$ and $S_j \in \Phi$.

$$\mathcal{G}(G) = \mathtt{A} \wedge \omega \not\models succ(G) \overset{\text{S-CREATE}}{\rightsquigarrow} \mathsf{create}(\Phi)$$

- S-TERMINATE: Suppose a goal $G = \mathsf{G}(x, ., .)$ fails or is terminated. Then cancel each maintenance commitment supporting the goal that is not supporting some other goal.

$$\mathcal{G}(G) \in \{\mathtt{F}, \mathtt{T}\} \wedge S \in \text{SAG}(G) \wedge \text{SAG}^{-1}(S) \setminus G = \emptyset \overset{\text{S-TERMINATE}}{\rightsquigarrow} \mathsf{terminate}(S)$$

M-goal to a-goal We describe the rules in which one or more achievement goals support a maintenance goal.

- A-CONSIDER: Suppose a m-goal M is in the active state, that is $maint(M)$ is false. Then consider one or more a-goals to restore $maint(M)$ to true. Let $\omega = \bigwedge_i succ(G_i)$, where $G_i \in \text{GMM}(M)$, and $\Phi = \{G_j\}$ is a set of new goals such that $\bigwedge_j succ(G_j) \wedge \omega \models maint(M)$.

$$\mathcal{M}(M) = \mathtt{A} \wedge \omega \not\models maint(M) \overset{\text{A-CONSIDER}}{\rightsquigarrow} \mathsf{consider}(\Phi)$$

- A-TERMINATE: Suppose a goal m-goal M fails or is terminated. Then terminate each achievement goal G supporting M that is not supporting some other m-goal.

$$\mathcal{M}(M) \in \{\mathtt{F}, \mathtt{T}\} \wedge G \in \text{GMM}(M) \wedge \text{GMM}^{-1}(G) \setminus M = \emptyset \overset{\text{A-TERMINATE}}{\rightsquigarrow} \mathsf{terminate}(G)$$

M-comm to m-goal We describe the rules in which one or more maintenance goals support a maintenance commitment.

- M-CONSIDER: Suppose a m-comm S is in the detached state. Then consider one or more m-goals to maintain the condition $maint(\mathsf{S})$. Let $\omega = \bigwedge_i maint(M_i)$, where $M_i \in \text{MSS}(S)$, and $\Phi = \{M_j\}$ is a set of new goals such that $\bigwedge_j maint(M_j) \wedge \omega \models maint(S)$.

$$\mathcal{S}(S) = \mathtt{D} \wedge \omega \not\models maint(S) \overset{\text{M-CONSIDER}}{\rightsquigarrow} \mathsf{consider}(\Phi)$$

- M-TERMINATE: Suppose a m-comm S is expired or terminated. Then terminate each m-goal supporting S that is not supporting some other m-comm S'.

$$\mathcal{S}(S) \in \{\mathtt{E}, \mathtt{T}\} \wedge M \in \text{MSS}(S) \wedge \text{MSS}^{-1}(M) \setminus G = \emptyset \overset{\text{M-TERMINATE}}{\rightsquigarrow} \mathsf{terminate}(M)$$

M-goal to M-comm We describe the rules in which one or more maintenance commitments support a maintenance goal.

- C-CREATE: Suppose a m-goal $M = \mathsf{M}(x, m, ., .)$ is in the monitoring state. Then create one or more commitments $C_j = \mathsf{C}(x, y_j, \mathcal{S}(y_j, x, \top, m_j, d_j) = \mathtt{D}, q_j)$ to persuade agent y_j to maintain the condition m_j. Note that the antecedent of C_j is a condition that the m-comm $S(y_j, x, \top, m_j, d_j)$ is detached. Thus this enhancement conforms to the structure of an a-comm (TSY) . Let $\omega = \bigwedge_i maint(S_i)$, where $S_i \in \mathrm{SSM}(M)$, and $\Phi = \{C_j\}$ is a set of new commitments such that $\bigwedge_j m_j \wedge \omega \models m$.

$$\mathcal{M}(M) = \mathtt{M} \wedge \omega \not\models maint(M) \xrightarrow{\text{C-CREATE}} \mathsf{create}(\Phi)$$

- MS-TERMINATE: Suppose a m-goal M fails or is terminated. Then cancel each m-comm supporting the goal M that is not supporting some other m-goal.

$$\mathcal{M}(M) \in \{\mathtt{F}, \mathtt{T}\} \wedge S \in \mathrm{SSM}(M) \wedge$$
$$\mathrm{SSM}^{-1}(S) \setminus M = \emptyset \xrightarrow{\text{MS-TERMINATE}} \mathsf{terminate}(S)$$

M-comm to a-goal Last, we describe the rules in which one or more a-goals support a maintenance commitment.

- AD-CONSIDER: Suppose a m-comm S is in the conditional state. Then the debtor of S considers one or more a-goals to detach S. Let $\omega = \bigwedge_i succ(G_i)$, where $G_i \in \mathrm{GAS}(S)$, and $\Phi = \{G_j\}$ is a set of new goals such that $\bigwedge_j succ(G_j) \wedge \omega \models ant(S)$.

$$\mathcal{S}(S) = \mathtt{C} \wedge \omega \not\models ant(S) \xrightarrow{\text{AD-CONSIDER}} \mathsf{consider}(\Phi)$$

- AD-TERMINATE: Suppose a m-comm S is expired or terminated. Then terminate each a-goal supporting S that is not supporting some other m-comm S'.

$$\mathcal{S}(S) \in \{\mathtt{E}, \mathtt{T}\} \wedge G \in \mathrm{GAS}(S) \wedge \mathrm{GAS}^{-1}(G) \setminus S = \emptyset$$
$$\xrightarrow{\text{AD-TERMINATE}} \mathsf{terminate}(G)$$

6 Results on Coherence and Convergence

We prove repeated convergence of traces under two assumptions. First, *action fairness* means that all agents act towards achieving their commitments and goals. Hence, all a-goals eventually reach a terminal state, either positive (e.g., `Satisfied`) or negative (e.g., `Failed`); and no m-goals remain indefinitely in a 'response' state, i.e., `Active`.

Second, for convergence we cannot have forever-cycling commitments or goals. TSY define cycling in the achievement case; the next two definitions provide a definition in the maintenance case.

Definition 10. *Let $S = \mathsf{S}(x, y, p, q)$ be a m-comm and τ be a trace of states $\langle S_0, S_1, \ldots \rangle$. Suppose $\mathcal{S}(S) = \sigma$ in some state S_i and in some subsequent state S_j, where $j > i$. If τ contains infinite pairs of $\langle S_i, S_j \rangle$, $S_j \neq \mathtt{S}$, then we say that S is* cycling *on τ.*

Definition 11. *Let $M = \mathsf{M}(x, m, s, f)$ be a m-goal and τ be a trace of states $\langle S_0, S_1, \ldots \rangle$. Suppose $\mathcal{G}(G) = \sigma$ in some state S_i and in some subsequent state S_j, where $j > i$. If τ contains infinite pairs of $\langle S_i, S_j \rangle$, $S_j \neq \mathtt{A}$, then we say that G is* cycling *on τ.*

Results focusing on one agent These theorems relate one agent's (s-) commitments and (a- and m-) goals. The first theorem says that a m-comm does not remain in `Sustain` infinitely long on a trace. The second says that the trace sustains a coherent configuration.

Theorem 1. *Let $S = \mathsf{S}(x, y, l, m, d)$ be an m-comm and τ be a trace of states $\langle S_0, S_1, \ldots \rangle$. Suppose in state S_c that $\mathsf{S}(S) = \mathtt{B}$. Then $\exists$ a state $S_h, h > c$ such that $\mathsf{S}(S) \neq \mathtt{B}$.*

Theorem 2. *Let $\tau = \langle S_0, S_1, \ldots \rangle$ be a trace. Then for any state S_i in τ, if S_i is not coherent, there is a subsequent state S_j, $j > i$, in τ such that S_j is coherent.*

Results focusing on many agents These theorems concern the m-comms in a multiagent system. They state that the agents together maintain their m-goals and commitments in a rational way.

Theorem 3. *Suppose agent x in a multiagent system $\mathcal{M}$ has a m-goal $M = \mathsf{M}(x, m, s, f)$. Then the agents in $\mathcal{M}$ create minimal sets of m-comms, m-goals, and a-goals necessary to maintain the condition m.*

Theorem 4. *Suppose agent x in a multiagent system $\mathcal{M}$ has a m-comm $S = \mathsf{S}(x, y, l, m, d)$. Then the agents in $\mathcal{M}$ create minimal sets of m-goals and a-comms and a-goals necessary to maintain the condition m.*

7 Applying the Theory

We illustrate the value of integrated reasoning over maintenance commitments and goals with the aerospace aftermarket scenario of Figure 1. Due to space, we apply our approach to a portion of the scenario, and compact a few steps by combining goal consideration (`Null` to `Inactive`) and activation (`Inactive` to `Active`). Table 2 shows a possible progression of the operator and manufacturer configurations.

In Step 1, the operator employs C-CREATE rule to create an a-comm to the manufacturer to paying if the latter maintains the engine: C = C(OPER, MFR, S, $\mathsf{maint\_paid}$), where S = S(MFR, OPER, $\top$, $\mathsf{engine\_running}$, expiry). In Step 2, MFR employs DETACH rule (TSY) to create m-comm S, which detaches a-comm C. In Step 3, MFR employs M-CONSIDER rule to consider and activate the m-goal M = M(MFR, $\mathsf{engine\_running}$, expiry, $\mathsf{engine\_dead}$). In Step 4, OPER pays ($\mathsf{main\_paid}$) OPER, which satisfies commitment C (TSY). In Step 5, suppose the engine fails and stops running. This causes the m-goal M to transition to `Active` and m-comm S to transition to `Sustain`. In Step 6, MFR employs A-CONSIDER to consider and activate a goal R=G(MFR, $\mathsf{engine\_running}$, $\mathsf{engine\_dead}$) to restore the engine. In Step 7, MFR fixes the engine, which satisfies R and causes M to transition to `Monitoring`, and S to `Sustain`.

8 Related Work

This paper draws on Telang, Singh, and Yorke-Smith's (2019) study of a-comms and a-goals. That work does not consider maintenance of either construct. M-goals are handled by, for instance Duff, Thangarajah, and Harland (2014), who do not consider commitments. The developments we provide to handle maintenance of commitments are non-trivial, including

Step	Rule	OPER Action	OPER State	MFR Action	MFR State
1	C-CREATE–Sec 5.2	create(C)	$\langle C^A \rangle$		$\langle C^A \rangle$
2	DETACH–TSY		$\langle C^D, S^D \rangle$	create(S)	$\langle C^D, S^D \rangle$
3	M-CONSIDER—Sec 5.2		$\langle C^D, S^D \rangle$	activate(M)	$\langle M^M, C^D, S^D \rangle$
4	Life cycle—TSY	maint_paid	$\langle C^S, S^D \rangle$		$\langle M^M, C^S, S^D \rangle$
5	Life cycle—Figures 2, 3		$\langle \neg\text{engine\_running}, S^B \rangle$		$\langle \neg\text{engine\_running}, M^A, S^B \rangle$
6	A-CONSIDER–Sec 5.2		$\langle \neg\text{engine\_running}, S^B \rangle$	activate(R)	$\langle \neg\text{engine\_running}, R^A, M^A, S^B \rangle$
7	Life cycle—Figures 2, 3		$\langle \text{engine\_running}, S^D \rangle$	engine_running	$\langle \text{engine\_running}, R^S, M^M, S^D \rangle$

Table 2: Progression of configurations in an aerospace scenario. C = C(OPER, MFR, S, maint_paid), S = S(MFR, OPER, $\top$, engine_running, expiry), M = M(MFR, engine_running, expiry, engine_dead) R = G(MFR, engine_running, engine_dead)

new definitions of support functions, closure and coherence, and new life cycle rules and theorems. Other differences from TSY are that we handle suspension operationally rather than with life cycle states and practical rules, which reduces the complexity of our model. The theorems, naturally, are unique to the presence of maintenance constructs.

The few works that address m-comms reduce them to a-comms albeit with more complex formulae. However, such a representation is inadequate because it does not capture potentially repeated interventions by the debtor to re-establish m should it fail. Further, these works do not study the life cycle of m-comms, nor the connection to goals. Mallya, Yolum, and Singh (2003) write maintenance commitments as achievement commitments for temporal formulae of the form "always m". Chesani et al. (2013) define commitments with universally quantified properties during a time interval. They employ a Reactive Event Calculus framework, which supports greater temporal expressiveness than ours. Their formulation does not have an explicit notion of m-commitment, having only a limited maintenance life cycle. We anticipate that an approach such as ours could be developed for their technical framework. Chopra and Singh (2015) define a commitment-specification language, Cupid, that is first-order and maps to event expressions using relational algebra. Cupid can capture commitments of the form of "for each insurance claim, I will provide a payment" but does not capture maintenance in that it handles only the achievement of the consequent: it does not handle a consequent being forever or repeatedly made true.

Günay, Winikoff, and Yolum (2015) propose a framework to enable agents to create a commitment protocol dynamically. Such an approach to agent interaction provides for runtime coordination between agents' goals. The framework admits achievement commitments. Determining whether an agent keeps to its (achievement) commitments is known as monitoring (Dastani, van der Torre, and Yorke-Smith 2017). The act of monitoring has a maintenance-like ongoing nature. Extending the approach of, e.g., Kafalı and Torroni (2018) to understand monitoring and responding to failures is future work, as is handling disputes between agents as to the facts (Telang et al. 2015) and their effects on coherence.

Criado, Black, and Luck (2016) discuss a notion of coherence where they seek to identify consistent sets of norms. They formulate coherence as constraint satisfaction where an agent can compute preferences across its norms with respect to its cognitive state. In a somewhat similar approach, Desai, Narendra, and Singh (2008) evaluate sets of commitments, viewed as contracts, from the perspective of the preferences of the participants. In contrast to these approaches, we seek not to evaluate coherence but to show how each agent is individually coherent and how, linked through their commitments, the agents are collectively coherent over time.

Al-Saqqar, Bentahar, and Sultan (2016) develop a logic that considers both agents' knowledge and commitments. In contrast, we develop an operational semantics relating achievement and maintenance goals and commitments.

9 Conclusion

This paper studies the dynamic relationships between a rational agent's cognitive state (i.e., beliefs and goals) and its social state (i.e., commitments). First-class maintenance commitments are a powerful new type of social commitments that enable richer relationships than otherwise possible, thereby supporting expanded forms of collaboration.

By formalizing the concept of a maintenance commitment, we defined an operational semantics based on life cycle rules and practical rules. Further, by motivating an extended notion of coherence, we proved that a system of rational agents following the practical rules will have coherence between the achievement and maintenance goals and commitments.

We proposed a set of practical rules that captures certain intuitions and leads to coherence results. One can conceive of alternative sets of such rules: our methodology is generic.

Future work is, first, to allow explicit representation of time in commitments, as indicated by Chesani et al. (2013); Fornara and Colombetti (2009). Particularly interesting is how temporally qualified propositions interact with the nature of maintenance commitments. Second, our approach assumes propositional goals and commitments: a future direction is to consider enhanced representations that involve decidable fragments of first-order logic.

Another interesting future direction is to consider maintenance as an essential component of norms broadly, not only commitments. Such an approach would be an enabler of a new theory of resilient norms that maintain their coherence with each other and the cognitive constructs, supporting both probabilistic reasoning about norms (Cranefield et al. 2016) and new verification techniques (Drawel et al. 2020).

Acknowledgments

We thank the anonymous reviewers for their suggestions. MPS thanks the US National Science Foundation (grant IIS-1908374) for support. NYS was partially supported by TAILOR, a project funded by EU Horizon 2020 research and innovation programme under GA No 952215.

References

Al-Saqqar, F.; Bentahar, J.; and Sultan, K. 2016. On the soundness, completeness and applicability of the logic of knowledge and communicative commitments in multi-agent systems. *Expert Systems with Applications* 43: 223–236. doi:https://doi.org/10.1016/j.eswa.2015.08.019.

Castelfranchi, C. 1995. Commitments: From Individual Intentions to Groups and Organizations. In *Proceedings of International Conference on Multiagent Systems (ICMAS)*, 41–48.

Chesani, F.; Mello, P.; Montali, M.; and Torroni, P. 2013. Representing and monitoring social commitments using the event calculus. *Autonomous Agents and Multi-Agent Systems* 27: 85–130. doi:10.1007/s10458-012-9202-0.

Chocron, P.; and Schorlemmer, M. 2018. Inferring Commitment Semantics in Multi-Agent Interactions. In *Proceedings of the International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, 1150–1158. doi: 10.5555/3237383.3237867.

Chopra, A. K.; and Singh, M. P. 2009. Multiagent Commitment Alignment. In *Proceedings of the International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, 937–944. Budapest. doi:10.5555/1558109.1558143.

Chopra, A. K.; and Singh, M. P. 2015. Cupid: Commitments in Relational Algebra. In *Proceedings of the Conference on Artificial Intelligence (AAAI)*, 2052–2059. Austin, Texas.

Cranefield, S.; Meneguzzi, F.; Oren, N.; and Savarimuthu, B. T. R. 2016. A Bayesian Approach to Norm Identification. In *Proceedings of the 22nd European Conference on Artificial Intelligence (ECAI)*, 622–629. Amsterdam: IOS Press.

Criado, N.; Black, E.; and Luck, M. 2016. A Coherence Maximisation Process for Solving Normative Inconsistencies. *Autonomous Agents and Multi-Agent Systems* 30(4): 640–680. doi:10.1007/s10458-015-9300-x.

Dastani, M.; van der Torre, L. W. N.; and Yorke-Smith, N. 2017. Commitments and interaction norms in organisations. *Autonomous Agents and Multi Agent Systems* 31(2): 207–249. doi:10.1007/s10458-015-9321-5.

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)*, 787–794. Estoril, Portugal. doi:10.5555/1402298.1402333.

Drawel, N.; Qu, H.; Bentahar, J.; and Shakshuki, E. M. 2020. Specification and Automatic Verification of Trust-Based Multi-Agent Systems. *Future Generation Computer Systems* 107: 1047–1060. doi:10.1016/j.future.2018.01.040.

Duff, S.; Harland, J.; and Thangarajah, J. 2006. On Proactivity and Maintenance Goals. In *Proceedings of the International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, 1033–1040. Hakodate, Japan. doi: 10.1145/1160633.1160817.

Duff, S.; Thangarajah, J.; and Harland, J. 2014. Maintenance Goals in Intelligent Agents. *Computational Intelligence* 30(1): 71–114. doi:10.1111/coin.12000.

Fornara, N.; and Colombetti, M. 2009. Specifying and Enforcing Norms in Artificial Institutions. In *Declarative Agent Languages and Technologies VI, Revised Selected and Invited Papers*, volume 5397 of *LNCS*, 1–17. Berlin: Springer.

Günay, A.; Winikoff, M.; and Yolum, P. 2015. Dynamically Generated Commitment Protocols in Open Systems. *Autonomous Agents and Multi-Agent Systems* 29(2): 192–229. doi:10.1007/s10458-014-9251-7.

Kafalı, Ö.; and Torroni, P. 2018. COMODO: Collaborative Monitoring of Commitment Delegations. *Expert Systems with Applications* 105: 144–158.

Mallya, A. U.; and Singh, M. P. 2007. An Algebra for Commitment Protocols. *Autonomous Agents and Multi-Agent Systems* 14(2): 143–163. doi:10.1007/s10458-006-7232-1.

Mallya, A. U.; Yolum, P.; and Singh, M. P. 2003. Resolving Commitments Among Autonomous Agents. In *Advances in Agent Communication: Proceedings of the International Workshop on Agent Communication Languages (ACL'03)*, 166–182. doi:10.1007/978-3-540-24608-4_10.

Marengo, E.; Baldoni, M.; Chopra, A. K.; Baroglio, C.; Patti, V.; and Singh, M. P. 2011. Commitments with Regulations: Reasoning about Safety and Control in REGULA. In *Proceedings of the International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, 467–474. Taipei, Taiwan. doi:10.5555/2031678.2031684.

Singh, M. P. 1991. Social and Psychological Commitments in Multiagent Systems. In *Proceedings of AAAI Fall Symposium on Knowledge and Action at Social and Organizational Levels*, 104–106.

Singh, M. P. 2000. A Social Semantics for Agent Communication Languages. In *Proceedings of the 1999 IJCAI Workshop on Agent Communication Languages*, number 1916 in Lecture Notes in Artificial Intelligence, 31–45. Berlin: Springer. doi:10.1007/10722777_3.

Singh, M. P. 2008. Semantical Considerations on Dialectical and Practical Commitments. In *Proceedings of the Conference on Artificial Intelligence (AAAI)*, 176–181. Chicago.

Singh, M. P. 2012. Commitments in Multiagent Systems: Some History, Some Confusions, Some Controversies, Some Prospects. In Paglieri, F.; Tummolini, L.; Falcone, R.; and Miceli, M., eds., *The Goals of Cognition: Essays in Honor of Cristiano Castelfranchi*, chapter 31, 591–616. London: College Publications.

Telang, P. R.; Kalia, A. K.; Madden, J. F.; and Singh, M. P. 2015. Combining Practical and Dialectical Commitments for Service Engagements. In *Proceedings of the 13th International Conference on Service-Oriented Computing (ICSOC)*, number 9435 in Lecture Notes in Computer Science, 3–18. Goa, India: Springer. doi:10.1007/978-3-662-48616-0_1.

Telang, P. R.; Singh, M. P.; and Yorke-Smith, N. 2019. A Coupled Operational Semantics for Goals and Commitments. *Journal of Artificial Intelligence Research* 65: 31–85. doi: 10.1613/jair.1.11494.

van Aart, C. J.; Chábera, J.; Dehn, M.; Jakob, M.; Nast-Kolb, K.; Smulders, J. L. C. F.; Storms, P. P. A.; Holt, C.; and Smith, M. 2007. Use Case Outline and Requirements. Deliverable D6.1, IST CONTRACT Project. Http://tinyurl.com/6adejz.

Winikoff, M.; Padgham, L.; Harland, J.; and Thangarajah, J. 2002. Declarative and Procedural Goals in Intelligent Agent Systems. In *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning (KR)*, 470–481. Toulouse, France.