From Signed Information to Belief in Multi-Agent Systems

Laurent Perrussel
IRIT – Université de Toulouse
laurent.perrussel@irit.fr

Emiliano Lorini
IRIT – Université de Toulouse
emiliano.lorini@irit.fr

Jean-Marc Thévenin
IRIT – Université de Toulouse
jean-marc.thevenin@irit.fr

***Abstract*—The aim of this paper is to propose a logical framework for reasoning about signed information. That is, as long as agents receive information in a multi-agent system, they keep track of the information source. The main advantage is that by considering a reliability relation over the sources of information, agents can justify their own current belief state. Agents believe at first information issued from the most reliable sources. Keeping track of belief's origin also enables agents to improve communication by asking and gaining details about exchanged information. This is a key issue in trust handling and improvement: an agent believes some statement because it may justify the statement's origin and its reliability.**

I. Introduction

An agent embedded in a multi-agent system gets information from multiple origins; it captures information from its own sensors or, through some communication channels it may receive messages issued by other agents. Based on this set of basic information the agent then defines its beliefs and performs actions [1]. As long as it gets information, the agent has to decide what it should believe and also which beliefs are dropped [2], [3]. In order to decide which beliefs should hold, the agent needs some criteria. A common criterion consists of handling a reliability relation on its beliefs w.r.t. their origins [4], [5]. According to its opinion about the reliability of the information source, the agent decides to adopt or not the received piece of information. By keeping track of information and its origin, agents can justify their beliefs: agent a believes φ because agent b has provided φ and b is reliable [6]. This explicit representation helps agents to enrich their dialogs: they cannot only provide information but they may also mention the third party at the origin of information. Let us consider again agent a and information provided by b: a may then ask b the underlying source of φ and a may then ask to this source. Hence, this issue is a key one for trust characterization: keeping track of agents involved in information broadcasting enables agents to evaluate, from their own point of view, whether they are all reliable, i.e. believable [7].

The aim of this paper is to propose a modal framework for representing agent's belief state and its dynamics by considering signed information, that is information associated to its source. If many work has been made in order to show how an agent can merge information issued from multiple origins [8], [9], very few work has focused on the explicit representation of the origins of information [10], [6] in the context of BDI-based systems with communication actions. But we advocate that this explicit representation is necessary since it represents the underlying rationale of agents' beliefs.

The dynamics is usually described in terms of performative actions based on KQML performatives [11] or speech acts [12], [13]. Hereafter, we propose to consider *tell* actions as private announcements from an agent (the sender of the message) to another agent (the receiver of the message). Private announcements enable to stress up how agents "restrict" their belief state as they receive information. More precisely, they shrink the space of information with their origins and then according to that space, they build up their beliefs.

The paper is structured as follows: In section II, we present the intuitive meaning of signed information and belief state. Next in section III, we present the technical details of the modal logic framework. In section IV, we then represent an intuitive and common policy for relating signed information and belief which consists in the adoption as belief of all consistent information. Next, in section V, we extend the logical system with actions of the form "agent a tells to agent b that a certain fact p is true". We conclude the paper in section VI by summing up the contribution and considering some open issues.

II. Setting the Framework

Handling the source of information leads to the notion of *signed statement*, that is some statement is true according to some source. From a semantics perspective, we want to be able to represent, w.r.t. some initial state of affairs, for each agent, what are the possible states that can be signed by each source. Agents build their own belief state using information signed by each source and the reliability of the source.

Example 1 *Suppose a car accident involving three cars which are blue (bc), red (rc) and yellow (yc). Now suppose a police detective who is interviewing the witnesses of the accident. Let po be the police detective. The first witness w_1 tells to the police detective that the blue car is responsible of the accident while the second one (w_2) states that the red car has caused the collision. Both of them tell to the police detective that yc is not responsible of the accident. In that context of information gathering, the police detective does not need to assume that the witnesses tell the truth or believe in information they provide. The police detective just needs to assume that w_1 provides or signs information $bc \wedge \neg rc \wedge \neg yc$ and w_2 provides or* *signs information $\neg bc \wedge rc \wedge \neg yc$. Next, based on these pieces of information, the police detective will build his opinion, i.e. his belief about the accident. The police detective faces contradicting information about the blue and red cars, but because the witnesses both agree about the yellow one, the police detective should believe that the yellow car is not the responsible of the collision. That is, the detective is willing to root his belief upon the set of signed statements he handles.*

A. Representing signed statements

Signed statements can be represented through Kripke models using one accessibility relation per source of information. Let $\mathsf{Sign}(b, p)$ be a modal operator stating that statement p is true according to source b. $\mathsf{Sign}(b, p)$ is true in state w if p holds in all states reachable from w through a relation denoted S_b describing the possible information states issued from b.

Example 2 *Let us consider the initial example. Information which might be signed by the two witnesses are bc and rc which leads to the signed statements $\mathsf{Sign}(w_1, bc)$, $\mathsf{Sign}(w_1, rc)$, $\mathsf{Sign}(w_1, bc \wedge rc)$,... With respect to our example, hereafter we will focus on the two signed statements $\mathsf{Sign}(w_1, bc \wedge \neg rc \wedge \neg yc)$ and $\mathsf{Sign}(w_2, \neg bc \wedge rc \wedge \neg yc)$.*

B. Interpreting signed statements

The aim is to represent formulas such as $\mathsf{Bel}(a, \mathsf{Sign}(b, \varphi_0))$ or, in a more general way $\mathsf{Bel}(a, \varphi_0)$, which respectively stands for agent a believes that agent b signs φ_0 and agent a believes φ_0. As for signatures, we use an accessibility relation denoted B_a to represent the possible belief states of agent a.

We assume that signed statements represent the rationales for beliefs. That is, if agent a believes φ_0 it is because some signed statement $\mathsf{Sign}(b, \varphi_0)$ holds in every possible belief state of agent a and agent a is willing to commit to this signed statement. Let a, b and c be three agents and p be a propositional symbol; Figure 1 illustrates the possible belief states of agent a w.r.t. some initial state w_0 using accessibility relation B_a (if p holds in a state, p is mentioned between brackets). Agent a considers two possible belief states, w_1 and w_2. In state w_1, the two possible states given by S_b contain p which entails that p is signed by b. On the other hand, the two possible states given by S_c contain p and $\neg p$: no information can be signed by agent c. In all states related to w_2 with S_b and S_c, p is true. From this figure we can conclude that in state w_0

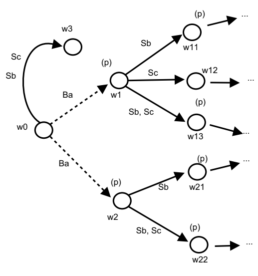

Fig. 1: Relating belief state and signatures

agent a believes that p is signed by b that is $\mathsf{Bel}(a, \mathsf{Sign}(b, p))$ while it does not believe that p or $\neg p$ is signed by c. Since p is signed by b and agent c says nothing about p, agent a should believe p: $\mathsf{Bel}(a, p)$. Hence, it follows that in order to prevent adoption of inconsistent statements, hereafter we will assume that signed statements are always consistent (and thus relation S_a is serial).

Notice that the way we consider the link between beliefs and signed statements differs from the way this link is defined in [6]. That is, signed states are considered from each belief state while C. Liau [6] considers informational states and belief states in an independent way. This is due to the fact that informational states in [6] reflect communication actions while our notion of signed statement is more considered as an epistemic notion.

Example 3 *Let us pursue our motivating example. As mentioned, we assume that the detective is willing to adopt as belief statements signed by the witnesses: $\mathsf{Bel}(po, \mathsf{Sign}(w_1, bc \wedge \neg rc \wedge \neg yc))$ and $\mathsf{Bel}(po, \mathsf{Sign}(w_2, \neg bc \wedge rc \wedge \neg yc))$. Since both witnesses agree on $\neg yc$, agent po also adopts as belief $\neg yc$. Meanwhile, he cannot set his belief about the two other cars since po faces contradicting signed statements.*

C. Preferences over information sources

In order to know how to handle mutually inconsistent signed statements, agents consider extra information stating which signed statement they prefer. Agents may determine themselves their preferences by considering the sources of information [4], [9], temporal aspects or the topics of the statements [14].

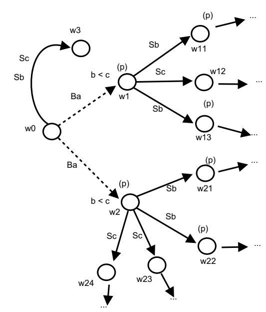

Fig. 2: Contradicting signed statements

In this paper, for the sake of conciseness and following numerous contributions such as [5], we propose to consider extra information about the reliability of sources of information as illustrated by Figure 2. That is, we assume that the agents consider information about only one topic. Consequently, handling competencies or different kinds of reliability (such as suggested in [15]) is out of the scope of the paper.

That is, if agent a believes that b is more reliable than c, then agent a adopts statement p as a belief even if agent c has signed $\neg p$. Suppose that reliability is represented with the help of a pre-order relation $\leqslant$ (or $<$): $a \leqslant b$ stands for a is at least as reliable as b. In semi formal terms, we get that:

$$\mathsf{Bel}(a, (\mathsf{Sign}(b,p) \wedge \mathsf{Sign}(c,\neg p) \wedge b < c)) \Rightarrow \mathsf{Bel}(a,p)$$

It follows that in each state, we do not only consider the value of propositional symbols but also a pre-order relation which characterizes a reliability order over information sources. Using extra-information on reliability and by considering signed statements rather than statements, the problem of belief change [2] is almost rephrased in terms close to the ones used in belief merging [16], [17]. Reliability order over sources of information enables us to stratify signed information and then by merging this stratified information in a consistent way the agents get "justified" beliefs [18].

Example 4 *Let us go on with our motivating example. Suppose agent po considers that the first witness is at least as reliable as the second one and he is himself willing to adopt as belief the signed statements issued by the two witnesses, i.e. we have the following belief:*

$$\mathsf{Bel}(po, w_1 \leqslant w_2 \leqslant po)$$

Hence, according to the previous semi formal axiom schema previously given, the police detective should believe that the blue car (bc) has caused the accident. Notice that the willingness attitude is translated in terms of preferences: po has no opinion and considers as more important information provided by w_1 and w_2.

D. Representing tell statements

Dynamics is viewed as restriction on agents' belief states. We interpret the *tell* performative as a private announcement [19] rather than with help of actions and transitions between states. A private announcement consists of an information flow from one agent to a second one with a propositional statement as content. Figure 3 illustrates how agent a's belief state changes after agent c tells p. According to this example, after the performative $Tell(c,a,p)$, agent a has restricted its possible belief states to the states in which c signs p. In the initial situation (the left part of the figure), at w_0, agent a believes $\mathsf{Sign}(b,p)$, does not believe $\mathsf{Sign}(c,p)$ and does not believe p (since p does not hold in w_2). After receiving agent c's message (right part of the figure), states where p is not signed by c are no longer possible states for agent a and thus, at w_0, agent a believes $\mathsf{Sign}(b,p)$, $\mathsf{Sign}(c,p)$ and finally also believes p. That is, the performative $Tell(c,a,p)$ (agent c tells to agent a that p is true) is responsible for updating a's beliefs in such a way that a believes that c signs p. In other words, private announcements stress up the information gathering aspect: possible worlds accessible through relation B_a represent the ignorance of agent B_a and by shrinking this set of possible believable worlds, we represent how agent a gains information. Let us stress that this way of handling the dynamics entails as a drawback that agent's belief cannot always be consistent: updating a model might lead to a model where seriality cannot be guaranteed.

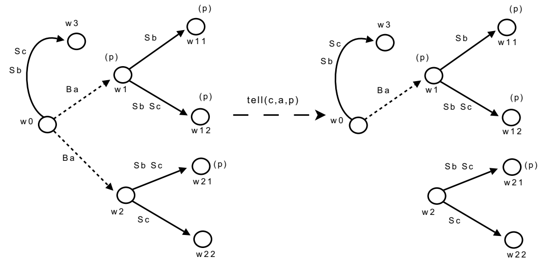

Fig. 3: Agent c tells p to agent a

Example 5 *In the context of our motivating example, the dynamics is represented by the sequence of interviews. For instance, agent po interviews at first w_1, action represented by $Tell(w_1, po, bc \wedge \neg rc \wedge \neg yc)$ and then interviews the second witness ($Tell(w_2, po, \neg bc \wedge rc \wedge \neg yc)$). After these two announcements, the detective believes:* $\mathsf{Bel}(po, \mathsf{Sign}(w_1, bc \wedge \neg rc \wedge \neg yc))$ *and* $\mathsf{Bel}(po, \mathsf{Sign}(\neg bc \wedge rc \wedge \neg yc))$.

III. Formal Framework

The proposed language for reasoning about signatures, beliefs and preferences is a restricted first order language which enables quantification over agent ids. In this section, we focus on these three notions, *tell* actions will be introduced later. Quantification allows agents to reason about anonymous signatures. For the sake of conciseness, we restrict signed statements to propositional statements. Let $\mathcal{L}_0$ be the propositional language built over a set of propositional symbols $\mathcal{P}$ and $\mathcal{L}$ be the logical language. Language $\mathcal{L}$ is based on doxastic logic. Modal operator Bel represents beliefs: $\mathsf{Bel}(a,\varphi)$ means agent a believes $\mathcal{L}$-formula φ. Modal operator Sign represents signed statements: $\mathsf{Sign}(t,\varphi_0)$ means t (an agent id or a variable of the agent sort) signs propositional statement φ_0. In order to represent agent's opinion about reliability, we introduce the notation $a \leqslant b$ which stands for: agent a is said to be at least as reliable as b.

Definition 1 (Syntax of $\mathcal{L}$) *Let $\mathcal{P}$ be a finite set of propositional symbols. Let $\mathcal{A}$ be a finite set of agent ids. Let $\mathcal{V}$ be a set of variables s.t. $\mathcal{A} \cap \mathcal{V} = \emptyset$. Let $\mathcal{T} = A \cup \mathcal{V}$ be the set of agent terms. The set of formulas of the language $\mathcal{L}$ is defined by the following BNF:*

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \mathsf{Sign}(t,\varphi_0) \mid \mathsf{Bel}(a,\varphi) \mid \forall x \varphi \mid t \leqslant t'$$

where $p \in \mathcal{P}$, $t \in \mathcal{T}$, $\varphi_0 \in \mathcal{L}_0$, $a \in \mathcal{A}$ and $x \in \mathcal{V}$.

Writing $a < b$ stands for a is strictly more reliable than b: $a \leqslant b \wedge \neg(b \leqslant a)$. Writing $a \sim b$ means that a and b are equally reliable. Operators $\rightarrow$ and $\exists$ are used according to their usual meaning.

A. Semantics

The semantics of $\mathcal{L}$-formulas is defined in terms of possible states and relations between states [20]. Those relations respectively represent the notion of signatures and beliefs. In each state, propositional symbols are interpreted and total pre-orders representing agents' reliability are set.

Definition 2 (Model) *Let M be a model defined as a tuple:*

$$\langle W, \bigcup_{i\in\mathcal{A}} S_i, \bigcup_{i\in\mathcal{A}} B_i, I, \preceq\rangle$$

where W is a set of possible states. $S_i \in W \times W$ is an accessibility relation representing signatures, $B_i \in W \times W$ is an accessibility relation representing beliefs. I is an interpretation function of the propositional symbols w.r.t. each possible state, $I : W \times \mathcal{P} \mapsto \{0,1\}$. $\preceq$ is a function which represents total pre-orders; these pre-orders are specific to each state, that is $\preceq : W \mapsto 2^{A\times A}$.

A variable assignment is a function v which maps every variable x to an agent id. A t-alternative v' of v is a variable assignment similar to v for every variable except t. For $t \in \mathcal{T}$, $[\![t]\!]_v$ belongs to A and refers to the assignment of agent terms w.r.t. variable assignment v, such that:

if $t \in A$ then $[\![t]\!]_v = t$ if $t \in \mathcal{V}$ then $[\![t]\!]_v = v(t)$

We define the satisfaction relation $\models$ with respect to some model M, state w and variable assignment v as follows.

Definition 3 ($\models$) *Let M be a model and v be a variable assignment: $v : \mathcal{V} \to A$. M satisfies an $\mathcal{L}$-formula φ w.r.t. a variable assignment v and a state w, according to the following rules:*

- *$M, v, w \models t \leqslant t'$ iff $([\![t]\!]_v, [\![t']\!]_v) \in \preceq(w)$.*
- *$M, v, w \models p$ iff $p \in \mathcal{P}$ and $I(w, p) = 1$.*
- *$M, v, w \models \mathsf{Sign}(t, \varphi_0)$ iff $M, v, w' \models \varphi_0$ for all w' s.t. $(w, w') \in S_{[\![t]\!]_v}$*
- *$M, v, w \models \mathsf{Bel}(a, \varphi)$ iff $M, v, w' \models \varphi$ for all w' s.t. $(w, w') \in B_a$*
- *$M, v, w \models \forall t\varphi$ iff for every t-alternative v', $M, v', w \models \varphi$.*

We write $\models \varphi$ iff for all M, w and v, we have $M, v, w \models \varphi$. The semantics for operators $\neg$, $\to$, $\vee$, $\wedge$ and $\exists$ is defined in the standard way. Let us now detail the constraints that should operate on the model. We only require that signature has to be consistent which entails that all relations S_i have to be serial. Belief operator is a *K45* operator and thus all B_i are transitive and euclidian. Interwoven relations between signatures and beliefs are detailed in the next section.

1) Constraining the Reliability Relations: We assume that every agent holds belief about reliability without any uncertainty. That is, agent's beliefs about reliability can be represented as a total pre-order. However, it does not mean that we consider a fixed notion of reliability: we propose to handle multiple pre-orders by indexing reliability with worlds. That is, in each possible world or believable world, an agent considers how it ranks the agents. In that context, each rank is considered as a possible rank and thus it is natural that each of them should be total. However, we enforce a stronger notion of totality which states that the aggregation of all believable ranks over agents (which are total) leads to a total preorder. This will then help the agent to integrate all signed statements. In other words, we require that the integration (or merging) of signed statements should be based on an underlying total preorder over statements (as it is commonly assumed in the belief revision and merging areas—see [2], [21], [17]). In terms of constraints on states and relations between them, it means that:

1) for all states w, $t \preceq(w)\, t'$ or $t' \preceq(w)\, t$ and,
2) suppose wB_iw' and $t \preceq(w')\, t'$, then for all states w'' s.t. wB_iw'', $t \preceq(w'')\, t'$.

The first constraint enforces that pre-orders are total in all states; the second constraint expresses that totality should hold in all belief states. Moreover, preorder definition entails that reflexivity and transitivity hold.

(K_S)	$\mathsf{Sign}(a, \varphi_0 \to \psi_0) \to (\mathsf{Sign}(a, \varphi_0) \to \mathsf{Sign}(a, \psi_0))$
(D_S)	$\mathsf{Sign}(a, \varphi_0) \to \neg\mathsf{Sign}(a, \neg\varphi_0)$
(K_B)	$\mathsf{Bel}(a, \varphi \to \psi) \to (\mathsf{Bel}(a, \varphi) \to \mathsf{Bel}(a, \psi))$
(4_B)	$\mathsf{Bel}(a, \varphi) \to \mathsf{Bel}(a, \mathsf{Bel}(a, \varphi))$
(5_B)	$\neg\mathsf{Bel}(a, \varphi) \to \mathsf{Bel}(a, \neg\mathsf{Bel}(a, \varphi))$
$(R_\leqslant)$	$t \leqslant t$
$(Tr_\leqslant)$	$t \leqslant t' \wedge t' \leqslant t'' \to t \leqslant t''$
$(T_\leqslant)$	$t \leqslant t' \vee t' \leqslant t$
$(To_\leqslant)$	$\mathsf{Bel}(a, t \leqslant t') \vee \mathsf{Bel}(a, t' \leqslant t)$
(MP)	From φ and $\varphi \to \psi$ infer ψ
(G)	From φ infer $\forall t\varphi$
(N_S)	From φ_0 infer $\mathsf{Sign}(t, \varphi_0)$
(N_B)	From φ infer $\mathsf{Bel}(a, \varphi)$

TABLE I: Logic $\mathcal{L}$ axioms and inference rules

B. Axiomatics

Let us now translate these constraints in terms of proof theory. Axiomatization of logic $\mathcal{L}$ includes all tautologies of propositional calculus. Table I details the axioms and inference rules describing the behavior of belief, signed statement and reliability. Notice axiom schema $(To_\leqslant)$ which reflects that reliability relations have to be believed as total. Let $\vdash$ denotes the proof relation. We conclude by giving results about soundness and completeness.

Theorem 1 *Logical system $\mathcal{L}$ is sound and complete*[1].

IV. Linking Signatures and Beliefs

There are multiple ways to switch from information to beliefs. These different ways may follow principles issued from the belief merging principle [16], [17], [5] or epistemic attitudes such as trust [7], [6]. As previously mentioned, we do not require that an agent has to believe that others

[1]In this paper all proofs have been skipped; however a longer version of the paper with all proofs is downloadable at the URL http://www.irit.fr/~Laurent.Perrussel/lads2010-long.pdf.

believe in information they provide. This is a key issue when information is propagated from one agent to another. At some stage, an agent may just broadcast some information without committing to that information in terms of belief.

A common and rational way to proceed is to consider as belief all non mutually inconsistent signed statements. All signed statements are considered in an incremental way, that is " from the most reliable to the less reliable statements". To describe the signed statements adoption stage, we first characterize agents which are equally reliable. Agents can be ranked since we always consider a total preorder; agents which are equally can be gathered in a same group. Each group can then be ranked. Let us at first characterize the most reliable set of agents; this set is denoted as C_1:

$$a \in \mathsf{C}_1 =_{def} \forall t(a \leqslant t)$$

The formula characterizing members of C_1 can then be used for characterizing membership to a set C_i such that $i > 1$.

$$a \in \mathsf{C}_i =_{def} (\neg(a \in \mathsf{C}_{i-1}) \wedge \forall t \neg(t \in \mathsf{C}_{i-1})) \rightarrow (a \leqslant t)$$

Hence, all agents belonging to a set C_i are equally reliable and for all $a \in \mathsf{C}_i, b \in \mathsf{C}_j$ if $i <_{\mathbb{N}} j$ then $a \leqslant b$. Next, the following definition stands for each agent t^k belonging to some specific set C_i believes statement φ_0^k:

$$\bigwedge_{t^k \in \mathsf{C}_i} \mathsf{Sign}(t^k, \phi_0^k) =_{def} \bigwedge_{t^k \in A} (t^k \in \mathsf{C}_i) \rightarrow \mathsf{Sign}(t^k, \phi_0^k)$$

Using these shortcuts, we can now describe the merging process. The following axiom states that if a propositional statement φ_0 is believed by agent a if the conjunction of the statements signed by the agents belonging to the same set C_i entails φ_0 is believed by agent a (line 1), if statement $\neg\varphi_0$ is not already believed by a(line 2) and $\neg\varphi_0$ cannot be entailed with the help of statements signed by agents which are at least as reliable as agents belonging to C_i (line 3).

$$\begin{aligned}(\mathsf{Bel}(a, \bigwedge_{t^k \in \mathsf{C}_i} \mathsf{Sign}(t^k, \varphi_0^k)) \wedge \mathsf{Bel}(a, \bigwedge \varphi_0^k \rightarrow \varphi_0) \wedge \\ \neg\mathsf{Bel}(a, \neg\varphi_0) \wedge \\ (\bigwedge_{0<j<i} \neg\mathsf{Bel}(a, \bigwedge_{t^l \in \mathsf{C}_j} \mathsf{Sign}(t^l, \varphi_0^l) \wedge \bigwedge \varphi_0^l \rightarrow \neg\varphi_0))) \\ \rightarrow \mathsf{Bel}(a, \varphi_0) \quad (\mathbf{IB})\end{aligned}$$

In terms of semantics, it means that, w.r.t. some initial state w_0, all belief states are related to some signed states. Hence, it requires to consider the state's interpretation, that is to express the relation by using worlds, i.e. a state and its associated interpretation. We represent a world as a set of propositional symbols, symbols that hold in the associated state. Let w be a state and $[w]$ the associated world:

$$[w] = \{p | I(w, p) = 1\}$$

In a more general way, if W is a set of states, then $[W]$ denotes the set of associated worlds. At first, from the belief states, we rank agent ids based on reliability relations believed by the agent. Suppose an agent a and a world w_0; using relation B_a, we extract the total preorder representing reliability relation believed by agent a at w_0. Notice that the constraints shown section III-A1 ensures that this preorder is total and thus agents ids could be ranked for building a partition of set of agents. Let $\mathcal{C}$ be a partition of $\mathcal{A}$ such that in every set $\mathcal{C}_i$ of $\mathcal{C}$, all agents are equally reliable and for all $a \in \mathcal{C}_i, b \in \mathcal{C}_j$ if $i <_{\mathbb{N}} j$ then $a \prec b$. Second, from each set $\mathcal{C}_i$, we consider common information, that is statements that are signed by every agents belonging to $\mathcal{C}_i$. Let $[\mathcal{C}_i]^{w_0}$ be the set of worlds commonly signed by all agents belonging to $\mathcal{C}_i$ and related to w_0:

$$[\mathcal{C}_i]^{w_0} = \bigcap_{a \in \mathcal{C}_i} \{[w] \mid (w_0, w) \in S_a\}$$

Next all sets of worlds $[\mathcal{C}_i]^{w_0}$ are merged in a consistent way, the resulting set of worlds is denoted as *reliable worlds*. By consistent way, we mean an incremental process which considers as reliable worlds at first the whole set of possible worlds $[W]$. Next for each part $\mathcal{C}_i$, the set of reliable worlds is intersected with $[\mathcal{C}_i]^{w_0}$ only if it does not lead to an empty set, i.e. an inconsistent result.

Definition 4 (Reliable worlds) *Let M be a model and w_0 a state such that $w_0 \in W$. The set of reliable worlds Ω^{w_0} is defined in an incremental way such that:*

- $\Omega^0 = [W]$
- $\Omega^i = \Omega^{i-1} \cap [\mathcal{C}_i]^{w_0}$ *if* $\Omega^{i-1} \cap [\mathcal{C}_i]^{w_0} \neq \emptyset$ *and* $i > 0$
- $\Omega^i = \Omega^{i-1}$ *if* $\Omega^{i-1} \cap [\mathcal{C}_i]^{w_0} = \emptyset$ *and* $i > 0$

The resulting set Ω^{w_0} is equal to Ω^k such that $k = |\mathcal{C}|$.

Since the sets of worlds and agents are finite, we do not have to consider the infinite case. Reliable worlds represent information that should be actually believed. Let us consider agent a and an initial world w_0; from w_0, we extract the belief states, and from these belief states, the set of reliable worlds. Beliefs of agent a are *rational* if all its believed worlds are included in its set of reliable worlds:

$$\bigcup_{(w_0,w) \in B_a} [w] \subseteq \bigcap_{(w_0,w) \in B_a} \Omega^w \qquad (\mathsf{IB})$$

The following theorem relates formula (**IB**) and constraint (IB). Let $\vdash_{\mathbf{IB}}$ denotes proof relation of the $\mathcal{L}$-sytem augmented with axiom schema (**IB**) and $\models_{\mathsf{IB}}$ be the satisfaction relation where for all models, constraint (IB) holds.

Theorem 2 $\vdash_{\mathbf{IB}} \varphi$ *iff* $\models_{\mathsf{IB}} \varphi$

V. Acquiring Information

In the previous section, we have detailed a policy for building belief based on signed information. This policy considers belief and signed information from a static point of view. Let us now consider a more dynamic view by introducing actions of the form "agent a tells to agent b that a certain fact φ_0 is true" (*alias* tell actions). This kind of action ensures that agent b will believes that agent a signs p, that is, a *tell* action is responsible for updating an agent's beliefs about other agents' signatures and, consequently, for the agent's acquisition of new information and for updating the agent's beliefs about objective facts. We note tell actions by $\texttt{Tell}(a, b, \varphi_0)$. Let $\mathcal{L}_T$ be the extended language which embeds *tell* statements.

Definition 5 (Syntax of $\mathcal{L}_T$) *The set of formulas of the language $\mathcal{L}_T$ is defined by the following BNF:*

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi' \mid \mathsf{Sign}(t, \varphi_0) \mid \mathsf{Bel}(a, \varphi) \mid \forall x\varphi \mid t \leqslant t' \mid [\texttt{Tell}(a, b, \varphi_0)]\varphi$$

where $p \in \mathcal{P}$, $t \in \mathcal{T}$, $\varphi_0 \in \mathcal{L}_0$, $a \in \mathcal{A}$ and $x \in \mathcal{V}$.

In other terms, $\mathcal{L}_T$ just extends $\mathcal{L}$ with dynamic operators $[\texttt{Tell}(a, b, \varphi_0)]$. The intuitive meaning of statement $[\texttt{Tell}(a, b, \varphi_0)]\varphi$ is after a tells φ_0 to b, φ holds.

The truth conditions are those given above for the formulas p, $\neg\varphi$, $\varphi \wedge \varphi'$, $\mathsf{Sign}(t, \varphi_0)$, $\mathsf{Bel}(a, \varphi)$, $\forall x\varphi$, $t \leqslant t'$ and $[\texttt{Tell}(a, b, \varphi_0)]\varphi$. The truth condition for $[\texttt{Tell}(a, b, \varphi_0)]\varphi$ is defined in a way which is closed to the semantics of dynamic epistemic logic [19]. More precisely, after agent a tells to agent b information φ_0, agent b removes from its belief state all states in which agent a does not sign φ_0. Therefore, after agent a tells to agent b information φ_0, agent b believes that agent a signs φ_0. In our framework, a tell action of agent a (the sender) towards agent b (the receiver) that φ_0 is true is considered as a *private announcement* in the sense of [22], that is, after agent a tells to agent b information φ_0, only agent b's belief state should change whereas the belief states of the other agents are not changed. In other words, a *tell* action $\texttt{Tell}(a, b, \varphi_0)$ characterizes a *private communication* from a sender to a specific receiver of the sender's message, where the content of the speaker's message is nothing else than the content of the speaker's signature (i.e. $\mathsf{Sign}(t, \varphi_0)$).

Definition 6 (Announcement Semantics) *Let $M = \langle W, \bigcup_{i\in\mathcal{A}} S_i, \bigcup_{i\in\mathcal{A}} B_i, I, \preceq\rangle$ be a model and let w be a state in W. We have:*

- $M, v, w \models [\texttt{Tell}(a, b, \varphi_0)]\varphi$ *iff* $M|_{\langle a,b,\varphi_0\rangle}, v, w_1 \models \varphi$.

$M|_{\langle a,b,\varphi_0\rangle} = \langle W^*, \bigcup_{i\in\mathcal{A}} S_i^*, \bigcup_{i\in\mathcal{A}} B_i^*, I^*, \preceq^*\rangle$ *is defined as follows:*

- $W^* = \{w_1 | w \in W\} \cup \{w_2 | w \in W\}$;
- $B_b^* = \{(w_1, w_1') | (w, w') \in B_b$ *and* $M, v, w' \models \mathsf{Sign}(a, \varphi_0)\} \cup \{(w_2, w_2') | (w, w') \in B_b\}$;
- $B_i^* = \{(w_1, w_2') | (w, w') \in B_i\} \cup \{(w_2, w_2') | (w, w') \in B_i\}$ *for all* $i \in \mathcal{A}$ *such that* $i \neq b$;
- $S_i^* = \{(w_1, w_2') | (w, w') \in S_i\} \cup \{(w_2, w_2') | (w, w') \in S_i\}$ *for all* $i \in \mathcal{A}$;
- $\preceq^*(w_1) = \preceq^*(w_2) = \preceq(w)$ *for all* $w \in W$;
- $I^*(w_1, p) = I^*(w_2, p) = I(w, p)$ *for all* $w \in W$.

Basically, the effect of a's action of telling to b that φ_0 is to shrink the set of belief accessible states for b to the states in which a signs φ_0, while keeping constant the set of belief accessible states for all other agents. Note that a's action of telling to b that φ_0 also keeps constant agents' signatures and the reliability order over agents.

(T_{AP})	$[\texttt{Tell}(a, b, \varphi_0)]p \leftrightarrow p$
(T_N)	$[\texttt{Tell}(a, b, \varphi_0)]\neg\varphi \leftrightarrow \neg[\texttt{Tell}(a, b, \varphi_0)]\varphi$
(T_C)	$[\texttt{Tell}(a, b, \varphi_0)](\varphi \wedge \varphi') \leftrightarrow ([\texttt{Tell}(a, b, \varphi_0)]\varphi \wedge [\texttt{Tell}(a, b, \varphi_0)]\varphi')$
(T_B)	$[\texttt{Tell}(a, b, \varphi_0)]\mathsf{Bel}(b, \varphi) \leftrightarrow \mathsf{Bel}(b, (\mathsf{Sign}(a, \varphi_0) \to [\texttt{Tell}(a, b, \varphi_0)]\varphi))$
$(T_{B\neq})$	$[\texttt{Tell}(a, b, \varphi_0)]\mathsf{Bel}(i, \varphi) \leftrightarrow \mathsf{Bel}(i, \varphi)$ if $i \neq b$
(T_S)	$[\texttt{Tell}(a, b, \varphi_0)]\mathsf{Sign}(t, \varphi_0') \leftrightarrow \mathsf{Sign}(t, \varphi_0')$
$(T_{\leq})$	$[\texttt{Tell}(a, b, \varphi_0)](t \leq t') \leftrightarrow (t \leq t')$
$(T_\forall)$	$[\texttt{Tell}(a, b, \varphi_0)]\forall x\varphi \leftrightarrow \forall x[\texttt{Tell}(a, b, \varphi_0)]\varphi$

TABLE II: Logic $\mathcal{L}_T$ axioms and inference rules

Theorem 3 *If M is a $\mathcal{L}$-model then $M|_{\langle a,b,\varphi_0\rangle}$ is also a $\mathcal{L}$-model.*

Theorem 4 *If M is a $\mathcal{L}$-model in which constraint IB holds then $M|_{\langle a,b,\varphi_0\rangle}$ is also a $\mathcal{L}$-model in which constraint IB holds.*

Let us now focus on the axiomatics of the logic $\mathcal{L}_T$. Table II details the reduction axioms describing the behavior of the operator $[\texttt{Tell}(a, b, \varphi_0)]$. (T_{AP}) denotes the atomic permanence, (T_N) denotes negation handling, and (T_C) denotes conjunction handling. (T_B) describes the interplay between a *tell* action and the beliefs of the message receiver. $(T_{B\neq})$ describes the interplay between a *tell* action and the beliefs of all agents different from the message receiver. In particular, $(T_{B\neq})$ highlights the permanence of the beliefs of all agents different from the message receiver. (T_S) describes signature permanence, $(T_{\leq})$ describes preferences permanence, and $(T_\forall)$ describes the interplay between *tell* action and quantification over variable assignments.

Theorem 5 *The schemata in table II are valid.*

We then state the theorem about completeness of the logic $\mathcal{L}_T$.

Theorem 6 *The logic $\mathcal{L}_T$ is completely axiomatized by principles of the logic $\mathcal{L}$ together with the schemata in Table II and the rule of replacement of proved equivalence.*

We write $\vdash_T$ to denote the proof relation for the logic $\mathcal{L}_T$ determined by the principles of the logic $\mathcal{L}$, the schemata in table II and the rule of replacement of proved equivalence.

For instance, the following theorem of the logic $\mathcal{L}_T$ captures the essential aspect of the *tell* action. It says that, after agent a tells to agent b information φ_0, agent b believes that agent a signs φ_0:

$$\vdash_T [\texttt{Tell}(a, b, \varphi_0)]\mathsf{Bel}(b, \mathsf{Sign}(a, \varphi_0))$$

Once agent b starts to believe that agent a signs φ_0 (as an effect of a's act of telling to b that φ_0), agent b might also start to believe that φ_0. As we have shown above, this depends on the reliability of agent a according to agent b and on principles linking signatures with beliefs such as principle (**IB**).

Example 6 *Let us go back to our initial example and let us represent in the system $\vdash_T$, how agent po concludes that the blue car has caused the collision. At first, assume that* (**IB**)*. Second, assume the following preferences:* $\mathsf{Bel}(po, w_1 \leqslant w_2)$*. Then it follows that after the two announcements (we focus on the blue car), preferences are unchanged:*

$$\vdash_T [\mathtt{Tell}(w_1, po, bc)][\mathtt{Tell}(w_2, po, \neg bc)]\mathsf{Bel}(po, w_1 \leqslant w_2)$$

And the detective believes the received information

$$\vdash_T [\mathtt{Tell}(w_1, po, bc)][\mathtt{Tell}(w_2, po, bc)] \\ \mathsf{Bel}(po, \mathsf{Sign}(w_1, bc) \wedge \mathsf{Sign}(w_2, \neg bc))$$

Finally, axiom (**IB**) *entails that*

$$\vdash_T [\mathtt{Tell}(w_1, po, \neg bc \wedge rc)][\mathtt{Tell}(w_1, po, bc \wedge \neg rc)]\mathsf{Bel}(po, bc)$$

VI. Conclusion

In this paper we have shown how information and its source can be processed by an agent so that at first, it just acquires information from sensors or other agents and second, it builds its belief state by considering signed information. By splitting information and belief, an agent is able to handle clear rationales to construct its belief state both from a static and dynamic perspectives. From a static perspective we have applied our formal framework to characterize a possible attitude for agents in the process of building their belief state from the basic signed information they hold. From this perspective this work is close to what has been done in belief merging [16], [17], [5]. The key difference with existing work in the belief merging area is the introduction of merging in a modal based framework at first (this is also a common characteristic with [5]); second a clear distinction between belief and signed statement and third a dynamic view on belief construction. These last two characteristics differ in two ways from existing work [16], [17], [5]: (i) it is usually assumed that belief and information are almost similar; we have shown that we do not have to assume this hypothesis; (ii) beliefs are almost not viewed as a primitive concepts but rather as the result of some information processing which gives a flexible framework (e.g. axiom **IB**). Our work is also related to the work of [23] in which agents' mental attitudes and agent's ostensible (expressed) attitudes are distinguished and a formalism capturing this distinction is proposed. In particular, our notion of signed information is close to the notion of ostensible belief of Nickles et al. However, Nickles et al. do not consider reliability of information sources. Moreover, their approach does not deal with dynamics of information by means of communicative actions. The latter is a central aspect of our proposal (see Section V).

Concerning the dynamic perspective we have shown how the basic signed information held by an agent may change as it receives tell statements from another agent processed in a similar way to private announcements in the sense of dynamic epistemic logic (DEL) [19], [22].

Our short term goal is to consider more sophisticated ways to set the reliability relations. That is, our aim is to consider agent skills [15] so that agent can consider multiple reliability relations at the same time. At this time, even if agent can consider multiple alternative reliability relations, they cannot mixed them. Our goal is to avoid this limit.

References

[1] A. Rao and M. Georgeff, "Modeling rational agents within a bdi-architecture," in *Proc. of KR'91*, 1991, pp. 473–484.

[2] P. Gärdenfors, *Knowledge in flux: Modeling the Dynamics of Epistemic States*. MIT Press, 1988.

[3] A. Herzig and D. Longin, "Belief dynamics in cooperative dialogues," *J. of Semantics*, vol. 17, no. 2, 2000, vol. published in 2001.

[4] A. Dragoni and P. Giorgini, "Revising beliefs received from multiple sources," in *Frontiers of Belief Revision, Applied Logic*. Kluwer, 1999.

[5] L. Cholvy, "A modal logic for reasoning with contradictory beliefs which takes into account the number and the reliability of the sources," in *Proc of ECSQARU'05*, ser. LNCS, vol. 3571. Springer, 2005, pp. 390–401.

[6] C. Liau, "Belief, information acquisition, and trust in multi-agent systems–a modal logic formulation," *Artificial Intelligence*, vol. 149, no. 1, pp. 31–60, 2003.

[7] E. Lorini and R. Demolombe, "From Binary Trust to Graded Trust in Information Sources: A Logical Perspective," in *Trust in Agent Societies*, ser. LNAI. Springer-Verlag, 2008, vol. 5396, pp. 205–225.

[8] W. van der Hoek and M. Wooldridge, "Towards a logic of rational agency," *Journal of the IGPL*, vol. 11, no. 2, pp. 133–157, 2003.

[9] L. Perrussel and J. Thévenin, "(Dis)Belief Change based on Messages Processing," in *Proc. of CLIMA'IV*, 2004.

[10] ——, "A logical approach for describing (dis)belief change and message processing," in *Proc. of AAMAS'04*. IEEE C.S., 2004, pp. 614–621.

[11] T. Finin, Y. Labrou, and J. Mayfield, "KQML as an agent communication language," in *Software Agents*, J. Bradshaw, Ed. MIT Press, 1997.

[12] P. Cohen and H. Levesque, "Rational Interaction as the Basis for Communication," in *Intentions in Communication*, P. Cohen, J. Morgan, and M. Pollack, Eds. MIT Press, 1990, pp. 221–256.

[13] ——, "Communicative actions for artificial agents," in *Proceedings of the First International Conference on Multi-Agent Systems (ICMAS'95)*, V. Lesser and L. Gasser, Eds. San Francisco, CA, USA: The MIT Press, 1995, pp. 65–72.

[14] L. F. del Cerro, A. Herzig, D. Longin, and O. Rifi, "Belief reconstruction in cooperative dialogues," in *Proc. of AIMSA'98*, ser. LNCS, F. Giunchiglia, Ed., vol. 1480. Springer, 1998, pp. 254–266.

[15] L. Cholvy, "Automated reasoning with merged contradictory information whose reliability depends on topics," in *Proc. of ECSQARU'95*, ser. LNCS, C. Froidevaux and J. Kohlas, Eds., vol. 946, 1995, pp. 125–132.

[16] P. Liberatore and M. Schaerf, "Arbitration (or how to merge knowledge bases)," *IEEE Transactions on Knowledge and Data Engineering*, vol. 10, no. 1, pp. 76–90, 1998.

[17] S. Konieczny and R. Pérez, "Propositional belief base merging or how to merge beliefs/goals coming from several sources and some links with social choice theory," *European Journal of Operational Research*, vol. 160, no. 3, pp. 785–802, 2005.

[18] S. Benferhat and L. Garcia, "Handling locally stratified inconsistent knowledge bases," *Studia Logica*, vol. 70, pp. 77–104, 2002.

[19] H. van Ditmarsch, W. van der Hoek, and B. Kooi, *Dynamic Epistemic Logic*, ser. Synthese Library. Springer, 2007, vol. 337.

[20] R. Fagin and Y. M. J. Halpern, *Reasoning About Knowledge*. MIT Press, 1995.

[21] H. Katsuno and A. Mendelzon, "Propositional knowledge base revision and minimal change," *Artificial Intelligence*, vol. 52, no. 3, pp. 263–294, 1991.

[22] J. Gerbrandy and W. Groeneveld, "Reasoning about information change," *J. of Logic, Language and Information*, vol. 6, no. 2, 1997.

[23] M. Nickles, F. Fischer, and G. Weiss, "Communication attitudes: A formal approach to ostensible intentions, and individual and group opinions," in *Proc. of LCMAS'05*, ser. Electronic Notes in Computer Science, vol. 157(4). Elsevier, 2005, pp. 95–115.