Public Group Announcements and Trust in Doxastic Logic

Elise Perrotin[1], Rustam Galimullin[2], Quentin Canu[3], and Natasha Alechina[2]

[1] IRIT, Toulouse, France
`elise.perrotin@irit.fr`
[2] University of Nottingham, UK
{`rustam.galimullin, natasha.alechina`}`@nottingham.ac.uk`
[3] ENS Paris-Saclay, France
`quentin.canu@ens-paris-saclay.fr`

Abstract. We present a doxastic logic for multi-agent systems with public group announcements. Beliefs are represented using belief bases and a dynamic of trust is introduced in order to handle belief change under contradictory announcements. We provide a complete axiomatization for this logic and illustrate its expressive power with a simple example.

Keywords: Syntactic Beliefs · Group Announcements · Belief Change · Trust.

1 Introduction

Receiving contradictory pieces of information from different sources is a common occurrence of daily life. However, deciding what to believe and who to trust as a result of those announcements is a task that is usually not that straightforward. Neither is representing these interactions between announcements, trust and beliefs.

We here propose a simple logic which can express these interactions following group announcements. All of the individual elements we have mentioned have been studied rather extensively, and often with various different approaches. We will now give an overview of this literature and what we are taking from it. As our focus is to allow for interaction of these elements in a single, manageable logic, we will often be choosing the less expressive, but easier to work with, options.

The first thing to consider is representation of beliefs, which may be false, and of their evolution. Two standard approaches exist for this: Dynamic Epistemic Logic (DEL, [10]), in which beliefs are represented using Kripke models with possible worlds, and the AGM approach ([3]), which uses sets of formulas, or *belief bases*, for that purpose, and from which follows Dynamic Doxastic Logic (DDL), as introduced in [21]. Public announcements and the resulting reorganization of beliefs, with or without trust, are built into DEL. However, these announcements, while they may be of false beliefs or even lies ([9]), are given one formula at a time and therefore do not allow for contradiction within that formula. While group announcement logics based on DEL exist ([1]), they handle announcements of what agents know rather than believe, and therefore these announcements are also necessarily consistent.When dealing with beliefs, plausibility models ([4]) can be used and the worlds reorganized after announcements, but those reorganizations depend on the order in which announcements are received, and it is not clear how they should function in the case of several simultaneous announcements, especially contradictory.

We will here work with belief bases, as we wish to focus first on what agents should believe after an announcement, rather than delving into all of the expressive ramifications allowed by DEL. The AGM approach is ideal for this, though, as pointed out in [5], it usually does not give much thought to where new information comes from, focusing instead on how to integrate it into existing beliefs. The dynamics expressed by DDL could, however, be interpreted as the results of announcements. AGM and DDL usually deal with only one agent, but multi-agent systems using belief bases, such as in [18], have also been proposed.

Handling contradicting statements in AGM is one of the topics of paraconsistent logic ([20]), which deals with identifying and isolating contradictions and extracting useful information from a belief base or an announcement. Closely related to this is belief merging ([15]), the aim of which is to merge several belief bases into one while preserving consistency. While weighted belief merging, as presented in [7], offers a mechanism allowing for different levels of reliability of sources, which can be interpreted as trust, it does not handle the associated evolution of this trust that we wish to represent. We will here avoid many difficulties brought up in paraconsistent logic and belief merging by using simplified belief bases, in which precise sources of contradiction are clearly identifiable.

Representing trust is the subject of yet another rather extensive body of work. The word 'trust' can have many meanings, and many corresponding models (see, e.g., [13, 16, 12, 8, 17, 14]). We here focus on trust as belief that what the other says is true, that is, trust in the reliability of another agent. For the sake of simplicity, we restrict ourselves to this definition of trust only, and assume in particular that agents only announce what they actually believe. The closest account of trust is the one given in [17], in which the focus is on whether information given by a source should be believed or not by an agent following the trust of the agent in that source, and in which contradicting announcements lead to loss of trust. Another interesting study of the evolution of trust is given in [14], in which trust is seen to be gained and lost as a result of so-called *trust-positive* and *trust-negative* experiences.

We present here a very basic notion of trust, which is binary: an agent either fully trusts or fully distrusts another agent, with no variation depending on topic (as opposed to [17]) and no gradual trust (as opposed to [14]). We also work with a memory-less trust, once again to present the most basic version of our framework. As such, trust-negative experiences are quite easy to identify (viz. contradictions in announcements), but trust-positive experiences are more tricky to define. For this reason we will here present a framework in which trust can only be lost. While quite basic, we argue that the notion of trust that we use can still be relevant in practice.

In the rest of the paper, we first present our logic, which we call *Syntactic Dynamic Doxastic Logic with Trust* (SDDLT) and the notion of trust that we use, and show that we can express coalition announcements. We then give an axiomatization and show that it is complete. We finish by giving a simple application example in order to justify our choice of trust dynamics and illustrate what can be expressed within our framework, and further discussing some of the choices we have made.

2 Syntax

Let Agt be a finite set of agents, and $\mathcal{P}$ a countable set of propositional variables. Consider $p \in \mathcal{P}$, $a, b \in Agt$ and $G \subseteq Agt$. The language $\mathcal{L}_{\mathsf{SDDLT}}$ of SDDLT is described by the following grammar:

$$\mathcal{L}_B \ni \varepsilon ::= \top \mid p \mid \neg p \mid B_a \varepsilon$$
$$\mathcal{L}_{\mathsf{SDDLT}} \ni \varphi ::= T_{a,b} \mid \varepsilon \mid \neg\varphi \mid \varphi \wedge \varphi \mid [A]\varphi$$

$B_a\varepsilon$ reads "agent a believes ε", $T_{a,b}$ reads "agent a trusts agent b", and $[A]\varphi$ reads "after announcement A (defined below), φ holds".

We also introduce some useful notations for the rest of the paper. The letter l will be used to denote literals (*i.e.* a variable or its negation). Given $a_1, \ldots, a_n \in Agt$ and $p \in \mathcal{P}$, if $\varepsilon_1 = B_{a_1} \ldots B_{a_n} p$ and $\varepsilon_2 = B_{a_1} \ldots B_{a_n} \neg p$, then we denote $\varepsilon_1 = \overline{\varepsilon_2}$ and $\varepsilon_2 = \overline{\varepsilon_1}$. Moreover, we denote the subformula relation by $\preceq$.

Announcements in our setting are *group announcements*, that is, they are public announcements consisting of statements given simultaneously by a group of agents. Given a group of agents $G \subseteq Agt$, an announcement by group G is a collection of pairs (a, A_a) where a is in G and A_a is a subset of $\mathcal{L}_B$ consisting of the formulas announced by agent a. Only one such set of formulas is allowed for each agent of G, and conversely, there is a set of announced formulas for each agent of G, though this set may be empty. We identify an announcement A with the corresponding function, that is, if A is an announcement by a group G and a is in G, we call $A(a)$ the set of formulas such that $(a, A(a))$ is in A.

We add a requirement on announcements: if A is an announcement by a group G, and if a is an agent of G, we require that for any formula ε_a of $A(a)$, no agent b of G simultaneously announces $B_a\overline{\varepsilon_a}$. That is, for all b in G, $B_a\overline{\varepsilon_a} \notin A(b)$. These kinds of announcements could be dealt with so that trust in the announcer is lost, but would lead to a much more complex axiomatization. For the sake of clarity down the line it is simpler to consider that these situations do not happen. This and other choices concerning announcements and trust dynamics will be further discussed in section 7.1.

For any group G, we call Ann_G the set of all possible announcements by agents of G, that is, the subset of $(2^{\mathcal{L}_B})^G$ following the above requirement. We call Ann the set of all possible announcements.

3 Semantics

3.1 Belief states and trust models

A belief state is a tuple $s = (\{BB_a^s\}_{a \in Agt}, \{T_a^s\}_{a \in Agt})$, where BB_a^s and T_a^s respectively denote the belief base and trust set of agent a at s. Belief bases are subsets of $\mathcal{L}_B$ and trust sets are subsets of Agt.

We require for all agents to trust at least themselves, and for all belief bases to contain at least $\top$. Belief bases must also be *consistent*, that is, they should not contain both ε and $\overline{\varepsilon}$ for any formula ε of $\mathcal{L}_B$. Furthermore, let $\dashrightarrow_a^s$ be defined in the following manner:

$$\varepsilon_1 \dashrightarrow_a^s \varepsilon_2 \text{ iff } \exists b \in T_a^s, \varepsilon_1 = B_b \varepsilon_2.$$

We denote by $\rightarrow_a^s$ the reflexive and transitive closure of $\dashrightarrow_a^s$. That is, if ε_1 and ε_2 are in $\mathcal{L}_B$, then $\varepsilon_1 \rightarrow_a^s \varepsilon_2$ if and only if there exist agents $a_1, \dots, a_k$ in T_a^s (for some $k \geq 0$) such that $\varepsilon_1 = B_{a_1} \dots B_{a_k} \varepsilon_2$. We require BB_a^s to be closed under $\rightarrow_a^s$ for any agent a at any belief state s: if $\varepsilon_1 = B_{a_1} \dots B_{a_k} \varepsilon_2$, a believes ε_1, and a trusts a_1, then by that trust a should believe $B_{a_2} \dots B_{a_k} \varepsilon_2$, but then if a trusts a_2, a should also believe $B_{a_3} \dots B_{a_k} \varepsilon_2$, and so on. More generally, we denote by $Cl_a^s(B)$ the closure of a set B under $\rightarrow_a^s$.

A *trust model* is a pair (S, V) where S is the set of all belief states and $V \subseteq \mathcal{P}$ is a valuation representing the actual state of the world.

We will work with pointed models (M, s) where $s \in S$.

3.2 Contradictions

When announcements are made, integrating the announced formulas to the belief bases of agents as is may render those belief bases inconsistent. Our restricted language, however, allows us to identify contradictions within announcements as well as between announcements and agents' beliefs rather easily.

If a is an agent, we say that two formulas ε_1 and ε_2 are *contradictory according to a at s* if $Cl_a^s(\{\varepsilon_1, \varepsilon_2\})$ is inconsistent. We say that a formula ε *contradicts a's beliefs at s* if there is a formula ε' in the belief base of a at s such that ε and ε' are contradictory according to a at s. Moreover, ε is *supported by a's beliefs at s* if $\overline{\varepsilon}$ contradicts those beliefs, and ε is *neutral w.r.t. a's beliefs at s* if it neither contradicts nor is supported by them.

We introduce the following notation: given $\varepsilon \in \mathcal{L}_B$, $\min_a^s(\varepsilon)$ is the shortest suffix ε' of ε such that $\varepsilon \rightarrow_a^s \varepsilon'$, that is, the shortest formula that a can deduce from ε at s. For example, if a trusts b and not c at s, we have that $\min_a^s(B_b B_c p) = B_c p$.

Lemma 1. *Let ε and ε' be two formulas of $\mathcal{L}_B$, a an agent, and s a belief state.*

1. *ε and ε' are contradictory according to a at s iff $\min_a^s(\varepsilon) = \overline{\min_a^s(\varepsilon')}$.*
2. *ε contradicts a's beliefs at s iff $\overline{\min_a^s(\varepsilon)} \in BB_a^s$, and ε is supported by a's beliefs at s iff $\min_a^s(\varepsilon) \in BB_a^s$.*

Proof. For the first statement, the interesting proof is that of the left-to-right direction. Suppose that $Cl_a^s(\{\varepsilon, \varepsilon'\})$ is inconsistent. This means that there exists a formula ε_0 such that $\varepsilon_0 \in Cl_a^s(\{\varepsilon, \varepsilon'\})$ and $\overline{\varepsilon_0} \in Cl_a^s(\{\varepsilon, \varepsilon'\})$. By definition of Cl_a^s, and because ε_0 and $\overline{\varepsilon_0}$ have different literals, we have either $\varepsilon \rightarrow_a^s \varepsilon_0$ and $\varepsilon' \rightarrow_a^s \overline{\varepsilon_0}$, or $\varepsilon' \rightarrow_a^s \varepsilon_0$ and $\varepsilon \rightarrow_a^s \overline{\varepsilon_0}$. The second case is reducible to the first by replacing ε_0 by $\overline{\varepsilon_0}$, so consider that $\varepsilon \rightarrow_a^s \varepsilon_0$ and $\varepsilon' \rightarrow_a^s \overline{\varepsilon_0}$. By definition of $\min_a^s(\varepsilon')$, the latter implies that $\overline{\varepsilon_0} \rightarrow_a^s \min_a^s(\varepsilon')$, that is, $\varepsilon_0 \rightarrow_a^s \overline{\min_a^s(\varepsilon')}$. By transitivity of $\rightarrow_a^s$, we get that $\varepsilon \rightarrow_a^s \overline{\min_a^s(\varepsilon')}$, and once again by minimality of $\min_a^s(\varepsilon')$, we have that $\min_a^s(\varepsilon) = \min_a^s(\varepsilon')$.

For the second statement, suppose that ε contradicts a's beliefs at s. Then there is a formula ε' in BB_a^s such that ε and ε' are contradictory according to a at s. By the first statement, this means that $\min_a^s(\varepsilon)' = \overline{\min_a^s(\varepsilon)}$, and by closure of BB_a^s under $\rightarrow_a^s$, $\min_a^s(\varepsilon')$ is in BB_a^s.

3.3 Update of trust

Given an announcement A made by a group G at a state s, we wish to define the updated state $s \cdot [A]$. For this we must define $BB_a^{s\cdot[A]}$ and $T_a^{s\cdot[A]}$ for any a. We begin with updates of trust.

When an announcement is made, agents first update their trust in other agents. In our framework, trust can only be lost. An agent a will stop trusting other agents when contradictory information is given in the announcement.

Let s be a state, a an agent, A an announcement by a group G, b and c two agents of G, and let ε_b and ε_c be two formulas such that $\varepsilon_b \in A(b)$ and $\varepsilon_c \in A(c)$. What a is learning is that $B_b\varepsilon_b$ and $B_c\varepsilon_c$, and a problem occurs when these two formulas are contradictory according to a at this state.

In order to choose which of the contradicting agents is no longer to be trusted, a will look at the statements as well as their own beliefs. If $B_b\varepsilon_b$ contradicts a's beliefs at s (which is equivalent to $B_c\varepsilon_c$ being supported by a's beliefs at s), then b is no longer trusted. Otherwise, if none of the statements are supported by a's beliefs, then a has no means of discrimination between the statements and both b and c are no longer trusted.

Finally, the new trust set is defined by:

$$
\begin{aligned}
T_a^{s\cdot[A]} = T_a^s \setminus \{b \in G \mid & \exists c \in G, \varepsilon_b \in A(b), \varepsilon_c \in A(c), \\
& \min_a^s(B_b\varepsilon_b) = \overline{\min_a^s(B_c\varepsilon_c)} \text{ and } \min_a^s(B_b\varepsilon_b) \notin BB_a^s\}
\end{aligned}
$$

3.4 Update of belief bases

Once the trust sets are updated, we can update the belief bases. As agents believe in the sincerity of all announcements, all formulas $B_b\varepsilon_b$ where $\varepsilon_b \in A(b)$ will be added to all belief bases. The update of trust ensures that there are no longer any conflicts between these formulas.

In case of conflict not with other statements in the announcement, but with previous beliefs of an agent, priority is given to the new information. We therefore remove formulas in the belief base which still contradict the announcement after updating trust. More formally, we obtain:

$$\begin{aligned} BB_a^{s\cdot[A]} = Cl_a^{s\cdot[A]}((BB_a^s \setminus \{\varepsilon \in BB_a^s \mid \exists b \in G, \exists \varepsilon_b \in A(b), \\ \min_a^{s\cdot[A]}(\varepsilon) = \min_a^{s\cdot[A]}(\overline{B_b\varepsilon_b})\}) \\ \cup \{B_b\varepsilon_b \mid b \in G \text{ and } \varepsilon_b \in A(b)\}) \end{aligned}$$

By removing all conflicts, we have ensured that this new belief base is indeed consistent.

3.5 Examples

To illustrate these dynamics, we study the effects of announcements on an agent a such that $T_a^s = \{a, b, c, d\}$ and $BB_a^s = \{p\}$, where $Agt = \{a, b, c, d, e\}$.

- If b announces $B_d p$ and c announces $\neg p$ ($A_1 = ((b, B_d p), (c, \neg p)))$, then b's statement is supported by a's beliefs, and therefore a loses trust in c: $T_a^{s\cdot A_1} = \{a, b, d\}$ and $BB_a^{s\cdot A_1} = \{p, B_b B_d p, B_d p, B_c \neg p\}$.
- If b announces q and c announces $\neg q$ ($A_2 = ((b, q), (c, \neg q)))$, a has no way of discriminating between the two announcements and therefore a loses trust in both b and c: $T_a^{s\cdot A_2} = \{a, d\}$ and $BB_a^{s\cdot A_2} = \{p, B_b q, B_c \neg q\}$.
- If c announces $\neg p$ and e announces p ($A_3 = ((c, \neg p), (e, p)))$, there is no conflict according to a, because a does not trust e and therefore cannot deduce p from $B_e p$. In this case, because c is trusted by a, c's announcement takes precedence over a's previous beliefs: $T_a^{s\cdot A_3} = \{a, b, c, d\}$ and $BB_a^{s\cdot A_3} = \{\neg p, B_c \neg p, B_e p\}$.

3.6 Semantics

Finally, we can define the semantics of SDDLT. Let (M, s) be a pointed trust model, $p \in \mathcal{P}$ a variable, $a \in Agt$ an agent, and $\varphi, \psi \in \mathcal{L}_{\mathsf{SDDLT}}$. Let A be an announcement by a group $G \subseteq Agt$. We introduce the shorthand $B_G A := \bigwedge_{g\in G} \bigwedge_{\varepsilon_g \in A(g)} B_g \varepsilon_g$. Then,

$$\begin{array}{ll}
(M, s) \models p & \text{iff } p \in V \\
(M, s) \models \neg\varphi & \text{iff } (M, s) \not\models \varphi \\
(M, s) \models \varphi \wedge \psi & \text{iff } (M, s) \models \varphi \text{ and } (M, s) \models \psi \\
(M, s) \models B_a\varepsilon & \text{iff } \varepsilon \in BB_a^s \\
(M, s) \models T_{a,b} & \text{iff } b \in T_a^s \\
(M, s) \models [A]\varphi & \text{iff } ((M, s) \models B_G A \Rightarrow (M, s \cdot [A]) \models \varphi)
\end{array}$$

4 Announcements by Groups and Coalitions

Announcements considered in the paper are made by groups of agents. Quantification over such announcements in a setting of epistemic logic has been studied in [1, 2]. The resulting formalisms – group announcement logic and coalition announcement logic – expand Public Announcement Logic [19] with operators $\langle G\rangle\varphi$ and $\langle\!\langle G\rangle\!\rangle\varphi$ correspondingly. The former is read as 'there is a joint public announcement by agents from group G such that φ holds in the resulting model,' and the latter means that 'there is a joint public announcement by agents from coalition G such that whatever agents from $A \setminus G$ announce at the same time, φ holds in the resulting model.'

Group and coalition announcements have been so far studied only from the epistemic perspective, i.e. agents in groups and coalitions announce what they *know*. Treatment of these operators in the doxastic setting is an open research problem.

Due to agents' limited reasoning in our framework, it is possible to define group and coalition announcements in SDDLT *provided we restrict the maximal depth of nestings of belief operators in any announcement.* This means that if the depth of nesting is restricted to some number, say 3, then the agents can make announcements of the form $B_aB_bB_ap$, but they are not allowed make announcements of the form $B_aB_bB_aB_bp$. Such a restriction is commonly made when formalising resource-bounded reasoning (see, e.g., [11]). We refer to the restricted logic where the depth of announced formulas cannot be higher than m as SDDLT^m.

We denote by $\mathrm{Lit}(\phi)$ the literals (positive and negative variables) appearing in a formula ϕ, and by $\mathrm{lit}(\varepsilon)$ the single literal appearing in a formula ε. The restriction on the maximal depth of formulas means we can now consider finite numbers of possible announcements. We denote the set of possible announcements by a group G relevant to φ by $PA(G, \varphi) = \{A \in Ann_G \mid \forall b \in G, \forall\varepsilon \in A(b), \mathrm{depth}(\varepsilon) \leqslant m \text{ and } \mathrm{lit}(\varepsilon) \in \mathrm{Lit}(\varphi)\}$, where for any formula ε, $\mathrm{depth}(\varepsilon)$ is the depth of ε defined as the length of the sequence of belief operators in ε.

We now define the group and coalition announcement operators for SDDLT^m:

$$\langle G\rangle\varphi \leftrightarrow \bigvee_{A\in PA(G,\varphi)} \Big(\bigwedge_{\substack{g\in G\\ \varepsilon_g\in A(g)}} B_g\varepsilon_g \wedge [A]\varphi\Big)$$

$$\langle\!\langle G\rangle\!\rangle\varphi \leftrightarrow \bigvee_{A_G\in PA(G,\varphi)} \ \bigwedge_{A_{\overline{G}}\in PA(\overline{G},\varphi)} \Big(\bigwedge_{\substack{g\in G\\ \varepsilon_g\in A_G(g)}} B_g\varepsilon_g \wedge \Big(\bigwedge_{\substack{g'\in \overline{G}\\ \varepsilon_{g'}\in A_{\overline{G}}(g')}} B_{g'}\varepsilon_{g'} \to [A_{Agt}]\varphi\Big)\Big),$$

where $\overline{G} = Agt \setminus G$ and $A_{Agt} = A_{\overline{G}} \cup A_G = \{(a, A_{\overline{G}}(a) \cup A_G(a)) \mid a \in Agt\}$.

5 Axiomatization

Now that SDDLT is completely defined, the next step is to give a sound and complete axiomatization for it. As group and coalition announcements are definable from the other operators, we do not consider these types of announcements in the axiomatization. Completeness will be proved for the fragment of the logic with no announcements, and we will give axioms reducing SDDLT to that fragment.

5.1 The first set of axioms

We give in table 1 the first nine axioms and the two inference rules of our system, where A represents an announcement by a group G. We will state two more axioms later as they require additional definitions.

$$\text{Propositional tautologies} \tag{A0}$$
$$B_a\top \tag{A1}$$
$$\varepsilon \to \neg\overline{\varepsilon} \tag{A2}$$
$$T_{a,b} \to (B_aB_b\varepsilon \to B_a\varepsilon) \tag{A3}$$
$$T_{a,a} \tag{A4}$$
$$[A]\neg\varphi \leftrightarrow (B_GA \to \neg[A]\varphi) \tag{A5}$$
$$[A](\varphi \wedge \psi) \leftrightarrow [A]\varphi \wedge [A]\psi \tag{A6}$$
$$[A]p \leftrightarrow (B_GA_G \to p) \tag{A7}$$
$$[A]T_{a,b} \leftrightarrow T_{a,b} \text{ for } b \notin G \tag{A8}$$
$$\vdash \varphi, \varphi \to \psi \Rightarrow \vdash \psi \tag{I0}$$
$$\vdash \varphi \Rightarrow \vdash [A]\varphi \tag{I1}$$

Table 1. The first set of axioms of SDDLT

The proofs of soundness of these axioms are rather straightforward. In order to have completeness, we need two more reduction axioms. The cases left to deal with are those of $[A]T_{a,b}$ when $b \in G$ and $[A]B_a\varepsilon$.

5.2 Trust and announcement

We first introduce a few notations. Given $\varepsilon, \varepsilon' \in \mathcal{L}_B$, we call $\max(\varepsilon, \varepsilon')$ the longest common suffix of ε and ε', that is, the longest formula μ such that $\mu \preceq \varepsilon$ and $\mu \preceq \varepsilon'$. This may be the empty formula. If $\varepsilon = B_{a_1} \dots B_{a_n} l$, then for all $1 \leq k \leq n$, we denote $Agt(\varepsilon \setminus B_{a_k} \dots B_{a_n} l) = \{a_1, \dots, a_{k-1}\}$. If $k = 1$, then $Agt(\varepsilon \setminus \varepsilon) = \emptyset$. Moreover, $Agt(\varepsilon \setminus l) = \{a_1, \dots, a_n\}$.

Recall the definition of the updated trust set for agent a after an announcement A by a group G at s:

$$T_a^{s\cdot[A]} = T_a^s \setminus \{b \in G \mid \exists c \in G, \varepsilon_b \in A(b), \varepsilon_c \in A(c), \\ \min{}_a^s(B_b\varepsilon_b) = \overline{\min{}_a^s(B_c\varepsilon_c)} \text{ and } \min{}_a^s(B_b\varepsilon_b) \notin BB_a^s\}$$

We need to express $\min_a^s(\varepsilon_b) = \overline{\min_a^s(\varepsilon_c)}$ (“ε_b and ε_c are contradictory from the point of view of a”) and $\min_a^s(\varepsilon_b) \notin BB_a^w$ (“a has no previous beliefs backing up b’s claim”).

We have the following result:

Lemma 2. *Let* ε_1 *and* ε_2 *be formulas of* $\mathcal{L}_B$*, a be an agent, and s a state. Define* ε_0 *as* $\varepsilon_0 = \max(\varepsilon_1, \overline{\varepsilon_2})$*. We have that* $\min_a^s(\varepsilon_1) = \overline{\min_a^s(\varepsilon_2)}$ *iff* $\mathrm{lit}(\varepsilon_1) = \mathrm{lit}(\overline{\varepsilon_2})$, $\varepsilon_1 \rightarrow_a^s \varepsilon_0$ *and* $\varepsilon_2 \rightarrow_a^s \overline{\varepsilon_0}$.

Using lemma 2, we get the following formula expressing contradiction of ε_1 and ε_2 from the point of view of a:

$$CO(a, \varepsilon_1, \varepsilon_2) = \bot$$

if $lit(\varepsilon_1) \neq \overline{lit(\varepsilon_2)}$, and

$$CO(a, \varepsilon_1, \varepsilon_2) = \Big(\bigwedge_{\alpha \in Agt(\varepsilon_1 \setminus \max(\varepsilon_1, \overline{\varepsilon_2}))} T_{a,\alpha}\Big) \wedge \Big(\bigwedge_{\beta \in Agt(\varepsilon_2 \setminus \max(\varepsilon_2, \overline{\varepsilon_1}))} T_{a,\beta}\Big)$$

otherwise.

Now to express that $\min_a^s(\varepsilon_b) \notin BB_a^s$, we use the following result:

Lemma 3. *If* ε *is a formula of* $\mathcal{L}_B$*, a is an agent, and s is a state, we have that:*

$$\min_a^s(\varepsilon) \notin BB_a^s \Leftrightarrow (\forall \mu \preceq \varepsilon, \varepsilon \rightarrow_a^s \mu \Rightarrow \mu \notin BB_a^s)$$

Using this we express the condition about previous beliefs:

$$PB(a, \varepsilon) = \bigwedge_{\mu \preceq \varepsilon} \Big(\Big(\bigwedge_{\alpha \in Agt(\varepsilon \setminus \mu)} T_{a,\alpha}\Big) \rightarrow \neg B_a \mu\Big)$$

Finally, we give the reduction axiom $(A9)$:

$$[A]T_{a,b} \leftrightarrow T_{a,b} \wedge \neg \bigvee_{\varepsilon_b \in A(b)} \bigvee_{\substack{c \in G \\ \varepsilon_c \in A(c)}} (CO(a, B_b\varepsilon_b, B_c\varepsilon_c) \wedge PB(a, \varepsilon_b)) \tag{A9}$$

if $b \in G$.

5.3 Belief and announcement

The last reduction axiom we need is for $[A]B_a\varepsilon$ where A is an announcement given by a group G, a is an agent, and ε is a formula of $\mathcal{L}_B$. Looking at the semantics, we have:

$$\begin{aligned} BB_a^{s\cdot[A]} = Cl_a^{s\cdot[A]}((BB_a^s \setminus \{\varepsilon \in BB_a^s \mid \exists b \in G, \exists \varepsilon_b \in A(b), \\ \min_a^{s\cdot[A]}(\varepsilon) = \min_a^{s\cdot[A]}(\overline{B_b\varepsilon_b})\}) \\ \cup \{B_b\varepsilon_b \mid b \in G \text{ and } \varepsilon_b \in A(b)\}) \end{aligned}$$

Hence ε is in $BB_a^{s\cdot[A]}$ if:

- There exists a formula ε_b announced by an agent b such that $B_b\varepsilon_b \rightarrow_a^{s\cdot[A]} \varepsilon$ (we say that ε is *successfully announced* to a)

– or ε was believed by a and not contradicted by the announcement from a's point of view after updating trust. Formally, this means that ε is in BB_a^s and for any formula ε_b announced by an agent b, we have that $\min_a^{s\cdot[A]}(\varepsilon) \neq \min_a^{s\cdot[A]}(\overline{B_b\varepsilon_b})$, or equivalently, $\min_a^{s\cdot[A]}(\overline{\varepsilon}) \neq \min_a^{s\cdot[A]}(B_b\varepsilon_b)$. That is, $\min_a^{s\cdot[A]}(\overline{\varepsilon})$ is not successfully announced to a.

We express the fact that ε is successfully announced to a through announcement A by the following formula:

$$Ann(A,\varepsilon,a) = \bigvee_{\substack{b\in G\\ \varepsilon_b\in A(b):\varepsilon\preceq B_b\varepsilon_b}} \bigwedge_{\alpha\in Agt(B_b\varepsilon_b\setminus\varepsilon)} [A]T_{a,\alpha}$$

Using this, we can also express the fact that ε was not contradicted in A from a's point of view. This is equivalent to no formula $\mu \preceq \overline{\varepsilon}$ such that $\overline{\varepsilon} \to_a^{s\cdot[A]} \mu$ being successfully announced:

$$NC(A,\varepsilon,a) = \bigwedge_{\mu\preceq\overline{\varepsilon}}((\bigwedge_{\alpha\in Agt(\overline{\varepsilon}\setminus\mu)} [A]T_{a,\alpha}) \to \neg Ann(A,\mu,a))$$

Finally, the axiom $(A10)$ is defined as:

$$[A]B_a\varepsilon \leftrightarrow Ann(A,\varepsilon,a) \vee (B_a\varepsilon \wedge NC(A,\varepsilon,a)) \tag{A10}$$

5.4 Completeness

The reduction axioms follow the semantics quite closely, and therefore we will not dwell on the proofs of soundness of each axiom and instead move on to completeness of our axiom system. As we have reduction axioms allowing us to translate formulas of the full language to that of the static language $\mathcal{L}^*_{\mathsf{SDDLT}}$, it suffices to show completeness of the axiom system constituted by the axioms $(A0)-(A4)$ and the inference rule $(I0)$ for the corresponding static logic SDDLT^*, that is, the logic without announcements.

The proof of completeness of SDDLT^* is a standard canonical model proof using maximal consistent sets, as given in [6]. Because of space constraints, we do not detail this proof but only give a few indications of the details specific to this logic. First, the canonical model used is the following:

Definition 1. *Let Γ be a consistent set of formulas of $\mathcal{L}^*_{SDDLT}$ that is maximal for inclusion. The canonical model for Γ is defined as $M^\Gamma = (S, V^\Gamma)$ where $V^\Gamma = \mathcal{P} \cap \Gamma$.*

We consider the state s_Γ such that $BB_a^{s_\Gamma} = \{\varphi \mid B_a\varphi \in \Gamma\}$ and $T_a^{s_\Gamma} = \{b \in Agt \mid T_{a,b} \in \Gamma\}$ for all a in Agt, verifying that it is indeed a belief state. A truth lemma stating that the formulas true at (M^Γ, s_Γ) are exactly the formulas of Γ is then shown by induction. This, conjointly with the Lindenbaum lemma, gives us the completeness of SDDLT^*.

Theorem 1. *For every* $\varphi \in \mathcal{L}^*_{\mathsf{SDDLT}}$*, if* $\models \varphi$*, then* $\vdash_{\mathsf{SDDLT}^*} \varphi$*.*

Corollary 1 (Completeness of SDDLT). *For every* $\varphi \in \mathcal{L}_{\mathsf{SDDLT}}$*, if* $\models \varphi$*, then* $\vdash_{\mathsf{SDDLT}} \varphi$*.*

Completeness of SDDLT^m (the logic with the belief depth of announcements bound to a fixed m) is a straightforward corollary, which gives us an axiomatisation of a logic with coalition announcements.

Corollary 2 (Completeness of SDDLT^m). *Given* $\varphi \in \mathcal{L}^m_{\mathsf{SDDLT}}$*, we have*

$$\models \varphi \Rightarrow \vdash_{\mathsf{SDDLT}} \varphi$$

6 A simple example: of bad influences and the importance of speaking out

We here give a concrete example to argue that our choice of trust dynamics is relevant, even though the agents' reasoning is quite basic. When considering real-life situations, group announcements do not consist of simultaneous announcements, but can be seen as statements proclaimed over a short period of time.

We see many news and articles about the importance of speaking up against bullying or harassment, or about how media is a bad influence to children. Everyone has their own story of something nobody talked to them about when they were a child, which led them to believe ridiculous –in hindsight– ideas they got from the television or magazines. The importance of speaking out against these wrong opinions to prevent the spread of their influence can be illustrated using our framework.

Say we have a group of agents Agt. In this group, there is a bad influence b, a group of gullible agents Gul, and a group of watchers Wat. The gullible agents trust everybody, the watchers only trust other watchers, and the bad influence trusts only themselves. For example, the bad influence could be the television, the gullible demographic the younger audience, and the watchers the parents of this audience. The bad influence could also be someone being uncivil, the gullible group could be foreigners still learning about the local culture, while the watchers would be local bystanders. Another example would be a bully, or a group of bullies acting as one, other students, and teachers at a school.

While all other agents in the group think $\neg p$ (for example, the incivility going on is not normal), the bad influence believes that p. We equate b carrying out the uncivil act to their announcing that p (there is nothing wrong with that action). While there is no risk of the bystanders starting to believe p, the foreigners could be led to believe that this is how things are done in this country. If nobody speaks out against b's actions, this is what the situation will lead to. However, as soon as one person speaks up, they will confirm others' belief that there is a problem, make the gullible agents lose trust in the troublemaker, and ensure that the incivility does not spread.

Formally this situation is described in the following model. Let $Agt = Gul \cup Wat \cup b$ and $M = (S, V)$ for some V (the actual state of the world is not important here). Consider s in S such that:

- $\forall a \in Gul \cup Wat, BB_a^s = \{\neg p\}$
- $BB_b^s = \{p\}$
- $\forall a \in Wat, T_a^s = Wat$
- $\forall a \in Gul, T_a^s = Agt$
- $T_b^s = \{b\}$

Then we have that

$$(M, s) \models [\{(b, \{p\})\}] \bigwedge_{g \in Gul} B_g p$$

and

$$(M, s) \models \bigvee_{a \in Wat} [\{(b, \{p\}), (a, \{\neg p\})\}] \bigwedge_{g \in Gul} (B_g \neg p \land \neg T_{g,b})$$

That is, if b announces p and nothing else is said, then gullible agents will start believing that p. However, if any one watcher a states that $\neg p$ as b claims the opposite, then the gullible agents will both retain the belief that $\neg p$ and learn that b is not to be trusted.

7 Discussion

7.1 On the choice of dynamics

Throughout this paper we have made choices in order to attempt to find a balance between non-trivial trust dynamics and clarity of the formalization. More complex dynamics could, of course, be envisioned: for example, in our framework, if an agent a receives a false announcement about their own beliefs from a trusted agent b (say, b announces $B_a p$ but a actually believes $\neg p$), they will blindly trust b and change these beliefs, so that a will start to believe p. This kind of blind trust may easily be considered too strong, and we may wish in this situation for a to stop trusting b instead, thus eliminating instances of contradictions with b's announcements . This can be expressed in the update of the trust set, which would become, after announcement A by a group G: if we call $C(a, s, A)$ the set $\{b \in G \mid \exists \varepsilon \notin BB_a^s, B_a \varepsilon \in A(b)\}$, we have

$$T_a^{s \cdot [A]} = T_a^s \setminus (C(a, s, A) \cup \{b \in G \mid \exists c \in G \setminus C(a, s, A), \varepsilon_b \in A(b), \varepsilon_c \in A(c), \\ \min_a^s(B_b \varepsilon_b) = \overline{\min_a^s(B_c \varepsilon_c)} \text{ and } \min_a^s(B_b \varepsilon_b) \notin BB_a^s\})$$

We have shown in section 5 how to express as formulas all of the properties necessary to amend the reduction axioms to follow these new dynamics. The constraint on announcements given in section 2 could also be lifted in a similar manner. However, this would lead to longer and less legible reduction axioms. For this reason we have presented dynamics which are more naive, but which suffice to make our point and are adaptable enough to work with more complex dynamics.

7.2 Sequential and simultaneous announcements

Say an agent a announces p, and afterwards an agent b announces $\neg p$, and say agent c trust both a and b. This situation can be modeled either as a group announcement $((a, \{p\}), (b, \{\neg p\}))$, which will lead c to lose trust in at least one of the two other agents, or as two announcements $(a, \{p\})$ and $(b, \{\neg p\})$ happening one after the other, leading c to believe first p, then $\neg p$, without losing trust in either a or b. In general, while simultaneous announcements are understandable as is in the context of, say, a search query, it is less clear what they represent when dealing with social interactions. Our understanding of group announcements in this context is that of announcements given over a short amount of time. This notion of short may depend on the agents and the situation, and while there is no clear-cut time stamp we can put on this, we can imagine that if enough time has gone by after a's announcement, c will no longer associate p with a, or they will accept that the situation may have changed since that first announcement, and accept b's announcement without feeling that there is too much of a conflict.

8 Conclusions and Future Work

We have defined a framework for reasoning about the evolution of trust and beliefs as triggered by group announcements, in which contradictions within the group can lead agents to lose trust in the speakers involved, and in which agents' beliefs can help them pick a side in case of conflict. We have given an axiomatization for our logic SDDLT, and shown that it is sound and complete. We have also shown that two operators for coalition announcements are definable in SDDLT.

Our work could be furthered in several directions. First, it would be interesting to expand on announcements. Though we have only considered public announcements here, we believe it would be quite simple to generalize these semantics to private announcements. It would also be interesting to no longer assume sincerity of the agents.

The notion of trust could also be expanded upon, following the existing literature, in particular the trust functions in [14]. For example, we could have several 'degrees' of trust, rather than simply binary trust. It would also be interesting to add a mechanism for gaining trust, the inner workings of which are less clear than those for loss of trust. The latter could also be refined: for instance, in case of conflict where no previous beliefs help the agent choose whom to trust, trust in both agents could be lost "until confirmation" of one of the two theses.

Finally, agents in our setting have very limited reasoning capabilities. Allowing more complex reasoning, and more complex formulas in belief bases, would make identifying contradictions stemming from announcements less straightforward. However, belief merging techniques could perhaps be applied to our framework to allow for these expansions.

References

1. Ågotnes, T., Balbiani, P., van Ditmarsch, H., Seban, P.: Group announcement logic. Journal of Applied Logic **8**(1), 62–81 (2010)
2. Ågotnes, T., van Ditmarsch, H.: Coalitions and announcements. In: Padgham, L., Parkes, D.C., Müller, J.P., Parsons, S. (eds.) Proceedings of AAMAS 2008, Volume 2. pp. 673–680. IFAAMAS (2008)
3. Alchourrón, C.E., Gärdenfors, P., Makinson, D.: On the logic of theory change: Partial meet contraction and revision functions. The journal of symbolic logic **50**(2), 510–530 (1985)
4. Baltag, A., Smets, S.: Dynamic belief revision over multi-agent plausibility models. In: Proceedings of LOFT. vol. 6, pp. 11–24. University of Liverpool (2006)
5. van Benthem, J.: Dynamic logic for belief revision. Journal of applied non-classical logics **17**(2), 129–155 (2007)
6. Blackburn, P., de Rijke, M., Venema, Y.: Modal Logic. Cambridge University Press (2001)
7. Delgrande, J.P., Dubois, D., Lang, J.: Iterated revision as prioritized merging. KR **6**, 210–220 (2006)
8. Demolombe, R., Liau, C.J.: A logic of graded trust and belief fusion. In: Proc. of the 4th Workshop on Deception, Fraud and Trust in Agent Societies. pp. 13–25 (2001)
9. van Ditmarsch, H.: Dynamics of lying. Synthese **191**(5), 745–777 (Mar 2014)
10. van Ditmarsch, H., van der Hoek, W., Kooi, B.: Dynamic Epistemic Logic, Synthese Library, vol. 337. Springer (2008)
11. Fisher, M., Ghidini, C.: Exploring the future with resource-bounded agents. Journal of Logic, Language and Information **18**(1), 3–21 (2009)
12. Ghosh, S., Velázquez-Quesada, F.R.: Agreeing to agree: Reaching unanimity via preference dynamics based on reliable agents. In: Proceedings of AAMAS 2015. pp. 1491–1499. IFAAMAS (2015)
13. Herzig, A., Lorini, E., Hübner, J.F., Vercouter, L.: A logic of trust and reputation. Logic Journal of the IGPL **18**(1), 214–244 (2010)
14. Jonker, C.M., Treur, J.: Formal analysis of models for the dynamics of trust based on experiences. In: European Workshop on Modelling Autonomous Agents in a Multi-Agent World. pp. 221–231. Springer (1999)
15. Konieczny, S., Pérez, R.P.: Logic based merging. Journal of Philosophical Logic **40**(2), 239–270 (2011)
16. Leturc, C., Bonnet, G.: A normal modal logic for trust in the sincerity. In: André, E., Koenig, S., Dastani, M., Sukthankar, G. (eds.) Proceedings of AAMAS 2018. pp. 175–183. IFAAMAS (2018)
17. Liau, C.J.: Belief, information acquisition, and trust in multi-agent systemsa modal logic formulation. Artificial Intelligence **149**, 31–60 (2003)
18. Lorini, E.: In praise of belief bases: Doing epistemic logic without possible worlds. In: Thirty-Second AAAI Conference on Artificial Intelligence (2018)
19. Plaza, J.: Logics of public communications (reprint of 1989's paper). Synthese **158**(2), 165–179 (2007)
20. Priest, G., Tanaka, K., Weber, Z.: Paraconsistent logic. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, summer 2018 edn. (2018)
21. Segerberg, K.: Belief revision from the point of view of doxastic logic. Logic Journal of the IGPL **3**(4), 535–553 (1995)