A Logic of Negative Trust

G. Primiero
Department of Philosophy, University of Milan, Italy.
giuseppe.primiero@unimi.it

Abstract

We present a logic to model the behaviour of an agent trusting or not trusting messages sent by another agent. The logic formalizes trust as a consistency checking function with respect to currently available information. Negative trust is modelled in two forms: distrust as the rejection of incoming inconsistent information; mistrust, as revision of previously held information becoming undesirable in view of new incoming inconsistent information, which the agent wishes to accept. We provide a natural deduction calculus, a relational semantics and prove soundness and completeness results. We overview a number of applications which have been investigated for the proof-theoretical formulation of the logic.

1 Introduction

The notion of trust has been central in both epistemology and computing, where its conceptual definition as well as applications have been explored. For the former, debates on third-person perspective knowledge, testimony and belief about other agents' epistemic states exemplify situations in which trust plays a role, either explicitly or implicitly. In various areas of the computational sciences, characterizations of trust are used to identify relevant, secure or preferred sources, channels and contents. A long list of applications can be mentioned where trust is involved in some form: software management systems and web certificates, cryptography and authentication protocols, design and analysis of social networks, data analytics, reputation systems to name some. It seems therefore obvious to expect logical analyses of trust that can answer the need for a formal and theoretical foundation in these fields.

An intuitive way of formalizing trust is as a first order relation between agents in the form $trust(A, B)$. But such a basic translation is ripe with problematic aspects. First, as debated at large in the computational trust community and reported in Section 5, trust interpreted as a first order relation between agents notoriously induces propagation based on transitivity, an undesirable property in many epistemic and security contexts:

Example 1.1 (Trust Transitivity)**.** *If Alice trusts Bob and Bob trusts Carol, then Alice trusts Carol.*

Consider for example security contexts, where authorization relations cannot be considered to hold transitively across agents. Consider also software management, where agents are servers holding software repositories and clients: when desired packages have dependencies across several non-directly authorised locations, it is usually safe to constrain trusting operations. A variant of the transitivity property in the definition of trust as first order relation is that it applies to all contents generated by the trusted sender in a generalised way (Alice always trusts Bob), and a number of additional strategies need to be deployed to mitigate such situation, like contextual or situational descriptions. Note that here a relation of transmission of information is emerging, which can be preceded by one of access authorization across agents.

Another important problem is the definition of the semantics of negative trust. A notion of negative trust requires reference to intentions when distinguishing between misplacement of trust and betrayal. An intentional characterization of trust distinguishing mistrust from distrust is presented in Primiero and Kosolosky (2013). Let us consider two examples:

Example 1.2 (Intentionally untrustworthy transmission)**.** *Alice receives data d from Bob; she considers the transmission untrustworthy because she believes that Bob has sent false data intentionally.*

In this form of untrustworthiness attribution, the obvious meaning should operationally be reflected by a rejection of the received data.

Example 1.3 (Unintentionally untrustworthy transmission)**.** *Alice receives data d from Bob; she considers the transmission trustworthy because she believes Bob to be sincere (about d); she now considers her data ¬d false, albeit she held such data unintentionally as true.*

In this form of untrustworthiness attribution, the obvious meaning should operationally be reflected by a rejection of the previously held data.

The combination of negation and transitivity generates the problem of trust muliplication. Propagation for negative (first-order) trust is formulated as follows (Jøsang and Pope (2005)):

Example 1.4 (Negative trust Multiplication)**.** *If Alice does not trust Bob and Bob does not trust Carol, then Alice trusts Carol.*

A way to avoid the aforementioned problems is to revise the view of trust as a first order relation between agents, and define it as a second-order property of first-order relations (e.g. of communication) between agents (Primiero and Taddeo (2012)):

Definition 1.1 (Trust as a Property)**.** *Assume a a first-order relation of information transmission between Alice and Bob. If Alice chooses to rely on Bob about the content of the transmission and acts on that basis, then we say that trust qualifies the first-order relation of information transmission between Alice and Bob.*

In order to provide a formally viable translation, we formulate trust as a function that is exercised on a message sent by Bob under the knowledge state that qualifies Alice (which we call her profile):

$$(the\ profile\ of\ A)\ trusts\ (message - from - B) \tag{1}$$

In such a system, we refer to trust as a bridging function between the content that an agent can read from other agents and the content which the agent is allowed to write (i.e. to make available for reading to other agents). Under this interpretation, trusting a message means to check its consistency with respect to the current profile. As a result, we are considering a trust operation that constrains message passing operations, see Primiero and Raimondi (2014). The model is currently limited to first-order messages: agents cannot send messages about messages they trust or distrust.

Definition 1.2 (Trust as Consistency Check Function)**.** *If Alice reads message ϕ from Bob and ϕ is consistent with her profile, Alice trusts ϕ, i.e. she can add it to her profile and can write it as a message to other agents.*

Consider the following modified example:

Example 1.5 (Intentional Untrust Multiplication)**.** *Alice does not trust ϕ from Bob: she believes he sends her intentionally false information. Bob does not trust $\neg\phi$ from Carol: he believes she sends him intentionally false information. Assume Alice is aware of that, should she trust $\neg\phi$ from Carol?*

The question whether Alice is safe in assuming the original $\neg\phi$ to be valid is now specified more precisely, and the related epistemic action of *distrust*, reformulating the previous notion from Definition 1.2, has the following intuitive semantics:

Definition 1.3 (Distrust)**.** *If Alice reads ϕ from Bob, and ϕ is inconsistent with Alice's profile, Alice distrusts ϕ and writes $\neg\phi$.*

A distinct case for trust misplacement can be formulated as follows:

Example 1.6 (Unintentional Untrust Multiplication)**.** *Alice reads ϕ from Bob, false in view of her current information: she believes she has unintentionally held false information $\neg\phi$. Bob has received ϕ from Carol, who can confirm it to Alice. Should Alice trust ϕ from Carol?*

This operation has an intuitive semantic meaning which underpins an act of trust (of the new information) but leads to an act of negative trust (of the old information held):

Definition 1.4 (Mistrust)**.** *If Alice reads ϕ from Bob, ϕ is inconsistent with Alice's profile and Alice wants to maintain consistency, then she mistrusts $\neg\phi$.*

To accept or reject such contradicting information (i.e, to distrust or mistrust the received information) might depend on the number and role of other agents available for confirmation, or from selected parameters.

The aim of the present contribution is to provide a natural deduction calculus and a semantics for these notions of trust and their negative counterpart, including meta-theoretical soundness and completeness results. In particular, this article proceeds as follows: in Section 2 we present a generalised version of the natural deduction calculus (un)SecureND and analyse its structural properties including normalization under trust preservation; in Section 3 we present a relational semantics based on filter Kripke models for the calculus; and in Section 4 we offer soundness and completeness results. We conclude in Section 5 with an overview of the related literature, compared to the applications studied so far for (un)SecureND. The Section 6 sets the next steps of this research.

2 The Proof Theory of (un)SecureND

We introduce in this Section the proof theory of (un)SecureND. This calculus was initially introduced as SecureND in Primiero and Raimondi (2014), with a basic application to the resolution of problems generated by transitive trust operations, where one wishes to block trust applications among agents where consistency is lost, according to a trust function reflecting the intuitive meaning presented in Definition 1.2. SecureND resolves unintended transitive trust by requiring explicit localisation of trusted messages in the agents' profiles, similar to what suggested in Clarke, Christianson, and Xiao (2009). (un)SecureND extends the former logic by allowing negation in the language: it was first introduced in Primiero (2016) and it was aimed at resolving the problem of untrust multiplication by the definition of functions of mistrust and distrust according to the intuitive meaning presented in Definitions 1.3 and 1.4.

We start with introducing the language of our logic:

Definition 2.1. *(**Syntax of** (un)SecureND)*

$$\begin{aligned}
&\mathcal{S} := \{A \leq B \leq \cdots \leq \Omega\}\\
&\phi^S := a^S \mid \neg\phi_i^S \mid \phi_i^S \rightarrow \phi_j^S \mid \phi_i^S \wedge \phi_j^S \mid \phi_i^S \vee \phi_j^S \mid \bot \mid\\
&\qquad Read(\phi^S) \mid Write(\phi^S) \mid Trust(\phi^S)\\
&\Gamma^S := \{\phi_i^S, \ldots, \phi_n^S\}
\end{aligned}$$

$\mathcal{S}$ is a finite set of agents, ordered by $\leq$: we use S in the definition above as a metavariable for any $S \in \mathcal{S}$. The expression $A \leq B$ over $\mathcal{S} \times \mathcal{S}$ intuitively expresses that agent A has access on agent B's content. This relation models therefore only the information access across agents and it says nothing about the trust that A might express or not over a message she reads from B. The accessibility relation $A \leq B$ refers to the authorization that A has to read the content that agent B issues or produces. Typical examples of this kind of conditions are expressed by access control systems or organizations where e.g. a top-down authorization relation on reading messages could be defined. Hence, this order can be defined according to the requirements of a specific application and reflects an access policy. Here we simply take a total order over all agents, while a partial one could be defined, or multiple access order over subsets of

the whole hierarchy. From a logical viewpoint, this expresses the validity of formulas expressing the readability of belief bases of agents according the order relation $\leq$, but not otherwise. We denote with $A \sim B$ the case in which A, B have each access to the other agent's state.

ϕ^S is a metavariable for formulae, defined from a denumerable set of atoms a^A (or possibly a finite one, if allowed by the application context) with their logical composition inductively defined by connectives, including functions to read, trust and write, with S here a metavariable for elements in $\mathcal{S}$. Atomic formulas, a_i^A says that formula a_i is signed by agent $A \in \mathcal{S}$, either because asserted or because it is derived in her profile; for complex formulas, we use the convention that ϕ_i^A if and only if a_i^A, for all $a_i \in \phi_i$, and otherwise they can be composed by formulas issued by distinct agents.

The partial order defined over elements $S \in \mathcal{S}$ allows for branching in the hierarchy, so that e.g. $S < S' < S''$ and $S < S' < S'''$, i.e. S'', S''' both are accessible from S' and transitively from S, but S'', S''' could be unaccessible to each other. The language includes $\bot$ to express conflicts: we formulate $\neg\phi_i^A$ as an abbreviation for $\phi_i^A \to \bot$, meaning that formula ϕ_i^A induces a contradiction.

A user profile Γ^S is the list of all formulas issued by the same agent S. A judgement $\Gamma^S \vdash \phi^{S'}$ states that a formula ϕ signed by agent S' holds in the context Γ of formulas signed by agent S. A profile is consistent if it prevents contradictions, i.e. it does not include formulas $\phi^S, \neg\phi^S$, or formulas ϕ^S, ψ^S such that ψ^S implies $\neg\phi^S$. A formula which does not depend on any other formula, is derivable under any context, hence a judgement $\vdash \phi_i^S$ says that a formula ϕ_i signed by agent S holds in *any* context.

Along with standard complex formulas built by negation, implication, conjunction, disjunction, access formulas are built by three additional operators: $Read(\phi^S)$ means "the message ϕ issued by S is read"; $Trust(\phi^S)$ means "the message ϕ issued by S is trusted"; $Write(\phi^S)$ means "the message ϕ issued by S is written"; for all these formulas, the reading is in the passive form. A judgdment of the form $\Gamma^S \vdash Read(\phi^{S'})$ (respectively: $\Gamma^S \vdash Trust(\phi^{S'})$ and $\Gamma^S \vdash Write(\phi^{S'})$) means "the message ϕ issued by S' is read under the profile of agent S" (respectively: "the message ϕ issued by S' is trusted under the profile of agent S"; and "the message ϕ issued by S' is written under the profile of agent S").

For all the rules presented below, we assume $A \leq B \leq C$, and that the instances of such rules are trivially valid for formulas with the same agent's index on the left and right side of the turnstile. Profiles are constructed inductively from the empty profile, see Figure 1. Profiles can be extended $\Gamma^S; \Gamma^{S'} = \{\phi_i^S, \ldots, \phi_n^S; \phi_{n+1}^{S'}\}$, by the rule `Profile Extension`, where the semi-colon is commutative; this formulates a profile composed by formulas $\phi_i, \ldots, \phi_n$ issued by agent S and formula ϕ_{n+1} issued by agent S'. When such extension comes from the same agent, we use a comma: Γ^S, ϕ_i^S. Profile extension by deductive closure is guaranteed by the rules for $\to$ under the same agent's index.

The operational rules in Figure 2 formulate closure under compositionality by logical connectives. As it is standard in proof-theoretic semantics, the meaning of our binary (connectives) and unary functions is given by a pair of

$$\frac{}{\{\} : profile}\ \texttt{Empty Profile}$$

$$\frac{\Gamma^A : profile \qquad \vdash \psi_j^B}{\Gamma^A; \psi_j^B : profile}\ \texttt{Profile Extension}$$

Figure 1: The System `(un)SecureND`: Profile Construction Rules

introduction and elimination rules, the former intended to express how a formula with that connective is obtained, the latter how it can be dispensed with. The rule `Atom` establishes derivability for any formulas available within the same profile, or across any profile accessible according to $\leq$, as reflected by the side condition expressing validity of the rule for any profile $B \geq A$. $\perp$ is used to express inconsistency of a profile, inducing admissibility of any formula, reflecting a form of *ex falso sequitur quodlibet*. We use implication to $\perp$ as a way to introduce $\neg$. The introduction rule for conjunction $\wedge$-I allows composition of formulas from distinct profiles; by the corresponding elimination rule $\wedge$-E, each composing formula is derivable under the combined profiles. The introduction rule for disjunction $\vee$-I says that a combined profile can derive any formula from each of the composing profiles; by the corresponding elimination $\vee$-E, each formula derivable from each individual profile can also be derived from the extended profile. $\rightarrow$-Introduction expresses inference of a formula from a combined profile as inference between formulas (Deduction Theorem), which reads "If the profile of agent A is extended with ϕ_i issued by agent B, then ϕ_j is issued by agent C", and it generalizes the format with a single agent; its elimination $\rightarrow$-E allows to recover such inference as profile extension (Modus Ponens).

We equip the system with a set of rules describing derivability for the access functions `read, trust, write` and their negations. This fragment of rules is presented in Figure 3. `read` says that from any consistent context Γ^A a formula ϕ_i^B can be read, provided the access policy is valid; this is expressed by the side condition which reflects the order relation on users' profiles. This side condition can be reformulated as a proper premise in the rule if so required. `trust` works as an elimination rule for `read`: it says that if a formula ϕ_i^B can be read under context Γ^A and its inclusion preserves profile consistency, then it can be trusted. `write` works as an elimination rule for *trust*: it says that a readable and trustable formula can be written (made available to other users). `derive` allows to reduce the access process to derivability: any formula that is safely written in a consistent profile can be derived in it.

The following set of rules extend the $Trust$ function by negation. The rules for Distrust are intended to preserve the current profile in view of conflicting external information. The Introduction rule for distrust $\texttt{DTrust} - \texttt{I}$ expresses the following principle: a formula ϕ_i^B whose reading is inconsistent with the current context Γ^A is untrustworthy, i.e. the rule for `trust` is not applied. The corresponding elimination rule $\texttt{DTrust} - \texttt{E}$ uses $\rightarrow$-introduction: it derives any formula ψ_j^C which is consistent with the profile in which access to the

$$\frac{\Gamma^A;\Gamma^B : profile}{\Gamma^A;\Gamma^B \vdash \psi_i^B}\ \texttt{Atom}, \text{ for any } \psi_i^B \in \Gamma^B, s.t. B \geq A$$

$$\frac{\Gamma^A \vdash \phi^B \to \bot}{\Gamma^A \vdash \neg\phi^B}\ \bot, \text{ for any } B \in \mathcal{S}$$

$$\frac{\Gamma^A \vdash \phi_i^A \qquad \Gamma^B \vdash \phi_j^B}{\Gamma^A;\Gamma^B \vdash \phi_i^A \wedge \phi_j^B}\ \wedge\text{-I}$$

$$\frac{\Gamma^A;\Gamma^B \vdash \phi_i^A \wedge \phi_j^B}{\Gamma^A;\Gamma^B \vdash \phi_i^A}\ \text{r-}\wedge\text{-E} \qquad \frac{\Gamma^A;\Gamma^B \vdash \phi_i^A \wedge \phi_j^B}{\Gamma^A;\Gamma^B \vdash \phi_j^B}\ \wedge\text{-E}$$

$$\frac{\Gamma^A;\Gamma^B \vdash \phi_i^A}{\Gamma^A;\Gamma^B \vdash \phi_i^A \vee \phi_j^B}\ \text{r-}\vee\text{-I} \qquad \frac{\Gamma^A;\Gamma^B \vdash \phi_j^B}{\Gamma^A;\Gamma^B \vdash \phi_i^A \vee \phi_j^B}\ \text{l-}\vee\text{-I}$$

$$\frac{\Gamma^A;\Gamma^B \vdash \phi_i^A \vee \phi_j^B \qquad \phi_{i/j}^I \vdash \psi_k^C}{\Gamma^A;\Gamma^B \vdash \psi_k^C}\ \vee\text{-E}$$

$$\frac{\Gamma^A;\phi_i^B \vdash \phi_j^C}{\Gamma^A \vdash \phi_i^B \to \phi_j^C}\ \to\text{-I} \qquad \frac{\Gamma^A \vdash \phi_i^B \to \phi_j^C \qquad \Gamma^A \vdash \phi_i^B}{\Gamma^A;\phi_i^B \vdash \phi_j^C}\ \to\text{-E}$$

Figure 2: The System `(un)SecureND`: Operational Rules

conflicting formula ϕ_i^B is blocked. The rules for Mistrust are intended to modify the current user profile to accommodate conflicting external information. To do so, the removal of one or possibly more currently derivable formulas is necessary. The Introduction rule for mistrust $\texttt{MTrust} - \texttt{I}$ expresses the following principle: given a formula ψ_i^B whose reading is inconsistent with the current context Γ^A, identify the subset $\Delta^A \subseteq \Gamma^A$ by removal of formula ϕ_j^A in Γ^A so that the profile is still a valid one when ψ_i^B is added to it. In the new profile Δ^A, the formula ϕ_j^A generating inconsistency with ψ_i^B is not trusted. To reach such a $\Delta^A \subseteq \Gamma^A$ which allows to trust ψ_i^B, it is possible that several iteration of the $\texttt{MTrust} - \texttt{I}$ have to be applied, and more than one formula be removed. The corresponding elimination rule $\texttt{MTrust} - \texttt{E}$ expresses the following procedure: given a consistent profile resulting from removal of a formula and addition of a previously inconsistent one ψ_i^B, identify the set of formulas from profiles higher than the one of B (if any) which are consistent with ψ_i^B, and possibly required to confirm trust in the latter.

By the latter set of rules, `Distrust` is a flag for preventing admissibility of conflicting external information, while `Mistrust` is a flag for facilitating removal

$$\frac{}{\Gamma^A \vdash Read(\phi_i^B)}\ \texttt{read},\ \text{for any } \phi_i^B \in \Gamma^B, s.t. B \geq A$$

$$\frac{\Gamma^A \vdash Read(\phi_i^B) \qquad \Gamma^A; \phi_i^B : profile}{\Gamma^A \vdash Trust(\phi_i^B)}\ \texttt{trust}$$

$$\frac{\Gamma^A \vdash Read(\phi_i^B) \qquad \Gamma^A \vdash Trust(\phi_i^B)}{\Gamma^A \vdash Write(\phi_i^B)}\ \texttt{write}$$

$$\frac{\Gamma^A \vdash Write(\phi_i^B)}{\Gamma^A \vdash \phi_i^B}\ \texttt{derive}$$

$$\frac{\Gamma^A \vdash Read(\phi_i^B) \to \bot}{\Gamma^A \vdash \neg Trust(\phi_i^B)}\ \texttt{DTrust} - \texttt{I}$$

$$\frac{\Gamma^A \vdash \neg Trust(\phi_i^B) \qquad \Gamma^A \vdash \neg Trust(\phi_i^B) \to \psi_j^C}{\Gamma^A \vdash Write(\psi_j^C)}\ \texttt{DTrust} - \texttt{E}$$

$$\frac{\Gamma^A \vdash Read(\psi_i^B) \to \bot \qquad \Delta^A; \psi_i^B : profile}{\Delta^A; \psi_i^B \vdash \neg Trust(\phi_j^A)}\ \texttt{MTrust} - \texttt{I},\ \text{with } \Delta^A \subseteq \Gamma^A \ni \phi_j^A \to \neg\psi_i^B$$

$$\frac{\Delta^A; \psi_i^B \vdash \neg Trust(\phi_j^A) \qquad \Gamma^C; \psi_i^B : profile}{\Delta^A; \Gamma^C \vdash Trust(\psi_i^B)}\ \texttt{MTrust} - \texttt{E},\ \forall C \leq B$$

Figure 3: The System `(un)SecureND`: Access Rules

of conflicting formulas present in the current installation profile. For negated trust, double negation elimination does not hold: from $\neg\neg Trust(\phi)$ one cannot infer $Trust(\phi)$, as it depends on the choice of negated trust rules applied, $DTrust - E$ or $MTrust - E$. In particular, note that this semantics is nondeterministic with respect to which rule to apply in the presence of a derived formula $\Gamma^A \vdash Read(\psi_i^B) \to \bot$. The choice might be determined by the ranking of the agents: in the case of an agent highest in the ranking, distrust will always be the standard protocol; for an agent reading from higher sources, mistrust will always be the standard protocol. Otherwise, the choice of which negative trust protocol to select can be contextually defined. The order of negated trust operations is also relevant and the choice is not always open. Consider two scenarios.

Example 2.1. *At some step of a derivation tree the following rule occurs:*

$$\dfrac{\dfrac{\Delta^A; \psi_i^B \vdash \neg Trust(\phi_j^A) \qquad \Gamma^C; \psi_i^B : profile}{\Delta^A; \Gamma^C \vdash Trust(\psi_i^B)} \texttt{MTrust} - \texttt{E}, \forall C \leq B \qquad \Delta^A; \Gamma^C \vdash Read(\phi_i^D)}{\Delta^A; \Gamma^C \vdash \neg Trust(\phi_i^D)}$$

where $\psi_i^B \rightarrow \neg\phi_i^B$. *In this case, a second mistrust operation would be impossible as there are no agents higher than* D *who have a base consistent with* ϕ. *Consider now a second different scenario as follows:*

$$\dfrac{\Gamma^B \vdash \neg Trust(\phi_j^C) \qquad \Gamma^B \vdash \neg Trust(\phi_j^C) \rightarrow \psi_j^C}{\Gamma^B \vdash Write(\psi_i^C)} \texttt{DTrust} - \texttt{E}$$

followed by:

$$\dfrac{\dfrac{\Gamma^B \vdash Read(\neg\psi^A) \rightarrow \bot \qquad \Delta^B; \neg\psi^A : profile}{\Delta^B; \neg\psi^A \vdash \neg Trust(\phi^B)} \texttt{MTrust} - \texttt{I}, \forall A \leq B \qquad \Gamma^A; \neg\psi^A : profile}{\Delta^B; \Gamma^A \vdash Trust(\neg\psi^A)}$$

where $\psi_i^B \rightarrow \neg\phi_i^C$ *and* $\Delta^B \subseteq \Gamma^B$. *Here the choice between the contradictory formulas allows an alternation between distrust and mistrust, as the sources from which information is received is consistent with such procedure.*

Example 2.2. *Another scenario which shows the dynamics of mistrust and distrust is illustrated as follows: consider two mistrustful agents exchanging messages with each other, e.g. two partners* $A \sim B$, *each with low self-esteem and high reliance on the other. In this scenario the first agent to receive a message from the other will accept it*

$$\dfrac{\Gamma^A \vdash Read(\phi^B) \rightarrow \bot \qquad \Delta^A; \phi^B : profile}{\Delta^A; \phi^B \vdash \neg Trust(\neg\phi^A)} \texttt{MTrust} - \texttt{I}$$

It is then entirely possible that a message from another source $C < A \sim B$ *will contradict the previous message:*

$$\dfrac{\dfrac{\Delta^A \vdash Read(\neg\phi^C) \rightarrow \bot \qquad \Delta'^A; \neg\phi^C : profile}{\Delta'^A; \neg\phi^C \vdash \neg Trust(\phi^A)} \qquad \Delta'^A; \neg\phi^C \vdash Trust(\neg\phi^A)}{\Delta'^A \vdash Write(\neg\phi^C)}$$

As a consequence everyone's content could be restored:

$$\dfrac{\Gamma^B \vdash Read(\neg\phi^C) \rightarrow \bot \qquad \Delta^B; \neg\phi^C : profile}{\Delta^B; \neg\phi^C \vdash \neg Trust(\phi^B)}$$

$$\frac{\Gamma^A \vdash Write(\phi_i^A) \qquad \Gamma^A \vdash Trust(\phi_j^B)}{\Gamma^A; \phi_j^B \vdash Write(\phi_i^A)}\ \texttt{Weakening}$$

$$\frac{\Gamma^A, \phi_i^A; \phi_i^B \vdash Write(\psi_k^A)}{\Gamma^A, \phi_i^A \vdash Write(\psi_k^A)}\ \texttt{Contraction}$$

$$\frac{\Gamma^A, \phi_i^A, \phi_j^A \vdash Write(\phi_k^A)}{\Gamma^A, \phi_j^A, \phi_i^A \vdash Write(\phi_k^A)}\ \texttt{Exchange}$$

$$\frac{\Gamma^A \vdash \phi_i^B \qquad \Gamma^B, \phi_i^B \vdash \phi_j^B}{\Gamma^A; \Gamma^B \vdash \phi_j^B}\ \texttt{Cut}$$

Figure 4: The System `(un)SecureND`: Structural Rules

2.1 Structural Rules

Structural rules for (`un`)`SecureND` hold with restrictions, as illustrated in Figure 4. They all hold under the assumption that $A \leq B$ and in particular for the Weakening rule, an instance of the *trust* rule is explicitly required. As a result, the system qualifies as substructural, see e.g. Restall (2000).

`Weakening` usually is formulated as to guarantee that consistency of a derived formula is not affected by adding assumptions; in order for this to work, we need to guard profiles against conflicting extensions. Hence, the rule is constrained by an instance of *trust*: it says that a valid derivation of ϕ_i^A in the profile Γ^A is preserved under a profile extension in view of formula ϕ_j^B if and only if such formula is trustworthy, i.e. one extending consistently the current profile. `Contraction` normally expresses the principle that copies can be safely ignored, as the information they provide is already available. Under our interpretation, profiles are ordered by the corresponding order on agents. Hence, `Contraction` requires a constrained to preserve such dependencies: it says that a valid formula ϕ_k^A is preserved when removing one instance of two identical formulas $\phi_i^A; \phi_i^B$, provided one preserves the fromula from the agent higher in the order (if one exists), so as to guarantee that any further dependency below is preserved. `Exchange` holds usually for set of formulas, where no order is present and it expresses the principle that formula derivability is preserved across sets of assumptions in which two formulas are swapped. In the present language, the rule is constrained by the ordered structure of agents' profiles: it says that a valid derivation of ϕ_k^A is preserved under reorder of formulas ϕ_i^A, ϕ_j^A coming from the same agent A. Finally, the `Cut` rule expresses valid derivation under profile extension, and it holds under the order relation of access $A \leq B$ being preserved: if a formula ϕ_i^B is valid under profile Γ^A and a profile Γ^B including ϕ_i^B allows deriving a formula ϕ_j^B, then assuming $A \leq B$ the extended profile

$\Gamma^A; \Gamma^B$ allows derivation of ϕ_j^B. We show here the admissibility of this rule for the most relevant cases:

- for $\phi_i^B \equiv \neg Trust(\phi_j^B)$: consider the first premise of the rule to be the conclusion of a $\texttt{DTrust} - \texttt{I}$ rule, then the whole rule collapses in a form of $\texttt{DTrust} - \texttt{E}$ rule, where ϕ_j^B is any formula consistent with the removal of ϕ_i^B;
- for $\phi_i^B \equiv \neg Trust(\phi_i^A)$: consider the first premise of the rule to be the conclusion of a $\texttt{MTrust} - \texttt{I}$ rule, then $\Gamma^A \equiv \Gamma^A \backslash \{\phi_i^A\}; \psi_j^B$ for some formula ψ_j^B for which access is desired; then the second premise of the Cut rule is an instance of $\texttt{DTrust} - \texttt{I}$. The conclusion illustrates then the situation presented in Figure 5, valid for all $C < B$.

$$\frac{\Gamma^A \setminus \{\phi_i^A\}; \psi_j^B \vdash \neg Trust(\phi_i^A) \qquad \Gamma^C; \neg Trust(\phi_i^A) \vdash \xi_k^C}{\Gamma^A \setminus \{\phi_i^A\}; \psi_j^B; \Gamma^C \vdash \xi_k^C}\ \texttt{Cut}$$

Figure 5: An instance of the Cut Rule

This instance of the rule shows that any formula ξ_k^C which is valid under a profile $\Gamma^C; \neg Trust(\phi_i^A)$ which mistrusts a given formula ϕ_i^A, will be preserved by any consistent extension of the profile $\Gamma^A \setminus \{\phi_i^A\}; \psi_j^B; \Gamma^C$, and this should be preserved for any further profile containing formulas accessible from the highest profile A in the order.

Theorem 2.2 (Cut-Elimination Theorem)**.** *Any* (un)SecureND *derivation with an occurrence of the* Cut *rule can be transformed into another derivation with the same end sequent without* Cut *using only* Weakening *and* Trust*.*

Proof. By induction on the derivation D which is the redex of the cut-elimination. The standard proof holds for the logical rules. We consider here the specific cases for the last step being obtained by an access rule or a $\perp$ rule (for the specific instance of the (un)trust rules).

1. The left premise of the cut rule is the conclusion of $\perp$, then cut is of the form:

$$\frac{\dfrac{\Gamma^A \vdash \phi^B \to \perp}{\Gamma^A \vdash \neg\phi^B} \qquad \Delta^B, \neg\phi^B \vdash \psi^B}{\Gamma^A; \Delta^B \vdash \psi^B}\ \texttt{Cut}$$

   then also $\Gamma^A \vdash Trust(\neg\phi^B)$ holds and the conclusion can be obtained from the second premise by Weakening.

2. The right premise of Cut is the conclusion of $\perp$, then Cut is of the form:

$$\dfrac{\Gamma^A \vdash \neg\phi^B \qquad \dfrac{\Delta^B \vdash \phi^B \to \bot}{\Delta^B, \neg\phi^B \vdash \psi^B}}{\Gamma^A; \Delta^B \vdash \psi^B}\ \texttt{Cut}$$

then if $\psi^B \in \Delta^B$, $\Gamma^A; \Delta^B \vdash \psi^B$ is from *Atom*; if $\psi^B = \neg\phi^B$ then $\Gamma^A \vdash \psi^B$ and $\Gamma^A; \Delta^B \vdash \psi^B$ is obtained by `Weakening`.

3. The left premise of `Cut` includes a `read` operation, then cut requires a detour through `trust` of the form:

$$\dfrac{\dfrac{\overline{\Gamma^A \vdash Read(\phi^B)} \qquad \Gamma^A, \phi^B : profile}{\Gamma^A \vdash Trust(\phi^B)} \qquad \Delta^B, \phi^B \vdash \psi^B}{\Gamma^A; \Delta^B \vdash \psi^B}\ \texttt{Cut}$$

then the conclusion is obtained from the second premise by `Weakening`.

4. The right premise of `Cut` includes a `read` operation, then cut requires a detour through `trust` of the form:

$$\dfrac{\Gamma^A \vdash \phi^B \qquad \dfrac{\Delta^B, \phi^B \vdash Read(\xi^C) \qquad \Delta^B, \xi^C : profile}{\Delta^B \vdash Trust(\xi^C)}}{\Gamma^A; \Delta^B \vdash Trust(\xi^C)}\ \texttt{Cut}$$

so the conclusion is obtained again from the second premise by `Weakening`.

5. The left premise of `Cut` is the conclusion of `write`, then cut is of the form:

$$\dfrac{\dfrac{\dfrac{\Gamma^A \vdash Read(\phi^B) \qquad \Gamma^A \vdash Trust(\phi^B)}{\Gamma^A \vdash Write(\phi^B)}}{\Gamma^A \vdash \phi^B} \qquad \Delta^B, \phi^B \vdash \psi^B}{\Gamma^A; \Delta^B \vdash \psi^B}\ \texttt{Cut}$$

then again the conclusion is obtained by `Weakening` from the second premise.

6. The right premise of `Cut` is the conclusion of `write`, then cut is of the form:

$$\dfrac{\Gamma^A \vdash \phi^B \qquad \dfrac{\Delta^B, \phi^B \vdash Read(\xi^C) \qquad \Delta^B, \phi^B \vdash Trust(\xi^C)}{\Delta^B; \phi^B \vdash Write(\xi^C)}}{\Gamma^A; \Delta^B \vdash Write(\xi^C)}\ \texttt{Cut}$$

then the conclusion is obtained again by `Weakening` from the second premise.

□

3 Relational Semantics for (un)SecureND

A relational semantics appropriate for the calculus (un)SecureND can be formulated in terms of an interpretation of the accessibility relations on worlds. In this context, the meaning of the access rules for $read, trust, write$ express conditions on those accessibility relations. To distinguish between the information held by agents, and the access operations across agent's states, we will denote the former as local states, and to information validated across distinct agents as global states. These states are defined below respectively by associated satisfaction relations in Definitions 3.2 and 3.5. The conditions of access operations have the following informal meanings:

- for *read*: a local state β_i for an agent B is accessible from a local state α_j for an agent A if and only if A is authorized to access information from B and the information available at β_i is issued by B at a time earleier or at most as late as the time of state α_j for agent A.
- for *trust*: the same condition as for *read*, with the addition that the information available at β_i is consistent with the information available at α_j.
- for *write*: the same condition as for *trust*, with the addition that the information available at α_j becomes visible at successive states of any authorised agent.
- for *dtrust*: the failure of the additional consistency requirement on the *trust* condition and hence the impossibility to add the visibility condition given by *write*.
- for *mtrust*: the failure of the additional consistency requirement on the *trust* condition, followed by the removal on the visibility condition given by *write* on the locally available information.

A major difference of the semantics with respect to the proof theory is that the former makes it explicit the temporal order among agents' states, which only remains implicit in the latter, in the form of occurrence of judgements at given steps of a derivation tree. In this section, we spell out the formal details of this semantics.

Throughout we assume a denumerable set of atomic proposition $AP = \{a, b, c, \dots\}$ (or finite, if allowed and required by the context of application). We define a model as follows:

Definition 3.1 (Relational Model)**.**

$$\mathcal{M} = \langle \mathcal{A}, \leq, \Lambda_{A \in \mathcal{A}}, \preceq, \alpha_n, \omega_1, U^{\Lambda_{I,\dots,J}}, v \rangle$$

such that:

1. $\mathcal{A} := \{A, B, \Gamma, \dots, \Omega\}$ *is a finite set of agents.*

2. $\leq \subseteq \mathcal{A} \times \mathcal{A}$ *is a partial relation over* $\mathcal{A}$*. When* $A \leq B$*, we say that* A *has authorized access to* B*'s information, hence expressing a ranking. This relation is reflexive, but we do not assume in general that* $\leq$ *is symmetric, while transitivity is conditional on trust (see Definition 3.5).*

3. $\Lambda_{I \in \mathcal{A}} := \{\lambda_1, \ldots, \lambda_n\}$ *is a finite set of states for each agent* $I \in \mathcal{A}$*, and* $i, \ldots, n \in \mathtt{N}$*. We use the convention that* α_i *is used to denote the* i*th local state of agent* $A \in \mathcal{A}$*.*

4. $\preceq \subseteq \Lambda_A \times \Lambda_B$ *(with possibly* $A = B$*) is the total temporal relation over local states of agents* A, B*. When* $\alpha_i \preceq \beta_j$*, we say that the information holding at state* α_i *is issued at a time earlier or equivalent to the time at* β_j*. This relation is assumed to be reflexive, transitive and serial.*

5. *a designated state* α_n *representing the latest state of the highest ranked agent (i.e. the agent with the most authorized access over other agents).*

6. *a designated state* ω_1 *representing the earliest state of the lowest ranked agent (i.e. the agent with the least authorized access over other agents).*

7. $U^{\Lambda_{A,\ldots,\Omega}} := \Lambda_A \times \cdots \times \Lambda_\Omega$ *is the Cartesian Product of the sets* $\Lambda_{I \in \mathcal{A}}$*. We call such set a universe of states, and its elements global states. For brevity of notation, in the following* U^{Λ_I} *is denoted by* U^I *and* $U^{\Lambda_{A,\ldots,\Omega}}$ *is denoted by* $U^{\mathcal{A}}$*. Note that: the maximal set* $U^{\Lambda_{A,\ldots,\Omega}}$ *includes all formulas up to the latest temporal state of the most authorized agent that are satisfied in the Filter Model by Global Satisfaction, and its cardinality is denoted by* 1*; the minimal set* U^{Ω} *includes only formulas up to the earliest temporal state of the least authorized agent that are satisfied in the Model by Local Satisfaction, and its cardinality is denoted by* min*. Note that at each state we account only for the formulas valid in that state by the agent and, as mentioned above, this can be constrained to a finite set of propositional variables of interest, further constrained by the signature of the agent.*

8. $v : AP \to U^{\mathcal{A}}$ *is the labelling function. Intuitively* $v(a^A)$ *identifies the set of states in* $U^{\mathcal{A}}$ *where* a *indexed by agent* A *holds. Despite the fact that a selection of states in* $U^{A,\ldots,\Omega}$ *might not present an hereditary function* v *with respect to* $\leq$*, that is,* $\alpha_i \in v(a^A)$ *and* $\beta_i \leq \alpha_i$ *is not enough to establish* $\beta_i \in v(a^A)$*, we show that a final selection on the model always satisfies the hereditary condition.*

In the semantics we define the local satisfaction relation as the evaluation of formulas within an agent's A set of states

Definition 3.2 (Local Satisfaction)**.** *Given a* `(un)SecureND` *formula* ϕ *and a model as above, we define the satisfaction of* ϕ *at a local state* α_i *for an agent* A *by induction as follows:*

- $\alpha_i \vDash a^A$ *iff* $\alpha_i \in v(a^A)$

- $\alpha_i \vDash \top$ *for every* α_i
- $\alpha_i \vDash \bot$ *never*
- $\alpha_i \vDash \phi^A \vee \psi^A$ *iff* $\exists \alpha_h \preceq \alpha_i$ *such that* $\alpha_h \vDash \phi^A$ *or* $\alpha_h \vDash \psi^A$
- $\alpha_i \vDash \phi^A \wedge \psi^A$ *iff* $\alpha_i \vDash \phi^A$ *and* $\alpha_i \vDash \psi^A$
- $\alpha_i \vDash \phi^A \rightarrow \psi^A$ *iff* $\exists \alpha_j \succeq \alpha_i \in \Lambda_A$, *such that* $\alpha_j = \{Cn(\alpha_i \cup \phi^A)\}$ *and* $\alpha_j \vDash \psi^A$

According to this definition, an atom is satisfied at a local state if it is in the set of valuations at that state; every local state is consistent and never inconsistent. A disjunction is satisfied at a local state if there is an earlier local state for the same agent such that either one or the other disjunct is satisfied at that state; a conjunction is satisfied at a local state if both conjuncts are satisfied at that state. Note both conjunction and disjunction rely on the temporal relation, as follows: the latter admits any previously held content to satisfy a disjunction, including content at the current state; the former looks at the current state only, where hereditary validity of formulas is preserved. An implication is satisfied at a local state if at some next local state the set of valid formulas is in the consequence set of the union set of the current state and the antecedent of the implication, then and at that next state the consequent of the implication must be satisfied. The negation connective is defined in terms of implication and $\bot$. Note that $\alpha_i \vDash 1$ if and only if $\alpha_i \equiv \alpha_n$, i.e. the set of satisfied formulas has the largest cardinality if the local state of evaluation is the designated state of the most authorised agent at its latest temporal state; and $\alpha_i \vDash min$ if and only if $\alpha_i \equiv \omega_1$, i.e. the set of satisfied formulas has the least cardinality if the designated state of evaluation is that of the least authorised agent at its earliest temporal state.

The notion of satisfiability corresponds to validity in the local states of any given agent:

Definition 3.3 (Satisfiability)**.** *A formula* ϕ_i^A *is true in a model* $\mathcal{M}$*, denoted* $\mathcal{M} \vDash \phi_i^A$ *if and only if* $\alpha_i \in U^A \vDash \phi_i^A$ *for every* $\alpha_i \succeq \alpha_1 \in U^A$*.*

The relation of local satisfaction is monotonic, i.e. if $\alpha_i \in v(\phi^A)$, for all $\alpha_j \succeq \alpha_i$ it holds $\alpha_j \in v(\phi^A)$. Given α_n denotes the the maximally consistent set of formulas for agent A, it follows that if $\vdash \phi_i^A$ then $\phi_i^A \in \alpha_n$.

When extending a local state by a local state by a distinct agent, it is conceivable that they might include contradictory formulas. In the following, when validating an instance of a distrust or mistrust operation, the monotonicity of the model requires that some local states be dismissed in view of incoming contradictory information, in order to preserve global monotonicity. The dismissed states from the model can be states of the sender (distrust) as well as of the receiver (mistrust). In either case, an operation of filtering out these states from the model is required. The notion of filter model satisfies this requirement:

Definition 3.4 (Filter Model)**.** *A filter model $\mathcal{M}'$ of $\mathcal{M}$ is a structure constructed according to Definition 3.1 such that $U^{\mathcal{A}} \in \mathcal{M}'$ is obtained by $U^{\mathcal{A}} \in \mathcal{M}$ by a new selection in $\Lambda_A \times \cdots \times \Lambda_\Omega$. Such selection of states and the addition of possibly new local states in $U^{\mathcal{A}}$ results from the Global Satisfaction Relation in Definition 3.5. Filter models of a given class are defined as those which select the same subset from $U^{\mathcal{A}} \in \mathcal{M}$.*

A global satisfaction is defined across distinct agents in $\mathcal{A}$ for the access rules:

Definition 3.5 (Global Satisfaction)**.** *Given an* `(un)SecureND` *formula ϕ, a filter model as by Definition 3.4 above and the notion of local satisfaction it inherits, we define global satisfaction of ϕ at a state α_i for an agent A in a universe $U^{\mathcal{A}}$ by induction as follows:*

- $\alpha_i \in U^{\mathcal{A}} \vDash Read(\phi^B)$ *iff*
  1. $A \leq B$ *and*
  2. $\exists \beta_i \in U^{\mathcal{A}}$ *s.t.* $\beta_i \preceq \alpha_i$ *and*
  3. $\beta_i \vDash \phi^B$

- $\alpha_i \in U^{\mathcal{A}} \vDash Trust(\phi^B)$ *iff*
  1. $A \leq B$ *and*
  2. $\exists \beta_i \in U^{\mathcal{A}}$ *s.t.* $\beta_i \preceq \alpha_i$ *and*
  3. $\beta_i \vDash \phi^B$ *and*
  4. $\exists \alpha_j \in U^{\mathcal{A}}$ *s.t.* $\alpha_i \preceq \alpha_j$ *and*
  5. $\alpha_j = \{Cn(\alpha \cup \{\phi^B\})\}$

- $\alpha_i \in U^{\mathcal{A}} \vDash Write(\phi^B)$ *iff*
  1. $A \leq B$ *and*
  2. $\exists \beta_i \in U^{\mathcal{A}}$ *s.t.* $\beta_i \preceq \alpha_i$ *and*
  3. $\beta_i \vDash \phi^B$ *and*
  4. $\exists \alpha_j \in U^{\mathcal{A}}$ *s.t.* $\alpha_i \preceq \alpha_j$ *and*
  5. $\alpha_j = \{Cn(\alpha_i \cup \{\phi^B\})\}$ *and*
  6. $\exists \alpha_k \in U^{\mathcal{A}}$ *s.t.* $\alpha_j \preceq \alpha_k$ *and*
  7. $\alpha_k \vDash \phi^A$

- $\alpha_i \in U^{\mathcal{A}} \vDash DTrust(\phi^B)$ *iff*

1. $A \leq B$ *and*
2. $\exists \beta_i \in U^{\mathcal{A}}$ *s.t.* $\beta_i \preceq \alpha_i$ *and*
3. $\beta_i \vDash \phi^B$ *and*
4. $\exists \alpha_j \in U^{\mathcal{A}}$ *s.t.* $\alpha_i \preceq \alpha_j$ *and*
5. $\alpha_i = \{Cn(\alpha_j \cup \{\neg\phi^B\})\}$

– $\beta_i \in U^{\mathcal{A}} \vDash MTrust(\phi^B)$ *iff*

1. $\exists \beta_h \preceq \beta_i$ *s.t.* $\beta_h \vDash \phi^B$ *and*
2. $B \leq A$ *and*
3. $\exists \alpha_i \in U^{\mathcal{A}}$ *s.t.* $\beta_i \succeq \alpha_i$ *and*
4. $\alpha_i \vDash \neg\phi^B$
5. $\exists \beta_j \in U^{\mathcal{A}}$ *s.t.* $\beta_i \preceq \beta_j$ *and*
6. $\beta_j = \{Cn(\beta_i \setminus \{\phi^B\})\}$

According to these definitions, a local state α_i for an agent A in a given universe $U^{\mathcal{A}}$:

– can read ϕ from B if and only if A is higher in the hierarchy of accesses than B, and there is a state for B in the same universe at which ϕ holds and is temporally anterior to the state for A; item can trust ϕ from B if and only if it can read it and it is consistent in at least one posterior state;

– can write ϕ from B if and only if it can read it, trust it and relabel it so that it is satisfied in at least one posterior state;

– distrusts ϕ from B if and only if it can read it, but there is a posterior state in the same universe in which ϕ is not consistent with the consequences of the information held;

– mistrusts ϕ on its own state if and only if can read $\neg\phi$ from a successive state and so extends its own state to remove ϕ from its consequence set.

The notion of satisfiability is generalised in a universe of local state as truth in a given class of filter models:

Definition 3.6 (Validity)**.** *A formula ϕ_i^A is valid in a class of filter models, denoted $\mathcal{M}' \vDash \phi_i^A$, if and only if $\alpha_i \in U^{\mathcal{A}} \vDash \phi_i^A$ for every $\alpha_i \succeq \omega_1$ and every $U^{\mathcal{A}}$ in that class.*

4 Meta-Theory

The main results of this Section are proofs of soundness and completeness. In order to obtain them, we need to list conditions on models of our semantics. These conditions express a basic global monotonicity property for canonical models; the last two conditions express a global non-monotonicity property, which is accounted for in terms of canonical filter models below.

Definition 4.1 (Conditions on Models)**.** *In the following, we assume a universe U^i of local states appearing in the relevant condition:*

c0. *If $\alpha_i \in v(\phi^A)$ and $\alpha_i \preceq \alpha_j$, then $\alpha_j \in v(\phi^A)$. [Local Monotonicity]*

c1. *If $\beta_i \preceq \alpha_n$ and $A \leq B \leq \Gamma$ then $\exists \gamma_i \preceq \beta_i \preceq \alpha_n$. [Temporality Order over Access]*

c2. *If $\Omega \leq \Xi$, then $\omega_1 \prec \xi_i \preceq \alpha_n$ and $\xi_i \prec \omega_1 \preceq \alpha_n$. [Minimality]*

c3. *If $A \leq B \leq \Gamma$ and $\gamma_i \preceq \delta_j \preceq \epsilon_k$, then $\exists \xi_i, \xi_j$ s.t. $\alpha_i \preceq \delta_j \preceq \xi_i$ and $\beta_i \preceq \delta_j \preceq \xi_j$ and $\Xi \leq E$. [Implicative Closure]*

c4. *If $A < B$ and $A = B$ then $B < A$ [Reflexive Access].*

c5. *If $A \leq B$, $\beta_i \preceq \beta_j$ and $\beta_j \preceq \alpha_i$, then $\beta_i \preceq \alpha_i$ [Transitivity over Local Accesses].*

c6. *If $\beta_i \preceq \beta_j$ and $\omega_1 \preceq \beta_j \preceq \alpha_n$, $\exists \gamma_j$ s.t. $\gamma_j \preceq \beta_j \preceq \alpha_j$ and $\omega_1 \preceq \beta_i \preceq \gamma_j$ [Continuous Alternative Temporal Paths].*

c7. *$\forall \alpha_i . \exists \alpha_j$ s.t. $\alpha_i \preceq \alpha_j$ [Temporal Seriality]*

c8. *$A \leq B \leq \cdots \leq \Omega$ implies $\exists \alpha_i, \beta_i$ such that $\beta_i \preceq \alpha_i$ [Instantiation].*

c9. *$\exists \Omega$ such that $A \leq \Omega < M$ implies $A \leq M < \Omega$. [Finality]*

c10. *$\exists A$ such that $A \leq N$, for all N. [Primality]*

c11. *If $A < B$ and $\beta_i \preceq \alpha_i \prec \alpha_j$ and $\alpha_j \prec \alpha_i$, then $\beta_i \preceq \alpha_j \preceq \alpha_i$. [Temporal Equivalence of Local States]*

c12. *$A \leq B \leq \Gamma$ and $\gamma_i \preceq \beta_i$ imply $\gamma_i \preceq \alpha_i$ [Extension of Temporal States by Access].*

c13. *$A \leq B \leq \cdots \leq M$, $A \leq \Gamma \leq \Xi$, $A \leq \Delta \leq \Xi$, and $\mu_j \preceq \delta_j \preceq \gamma_j$ imply $\exists \xi_i, \xi_j$ such that $\mu_j \preceq \xi_i \preceq \xi_j \preceq \beta_j \preceq \alpha_j$. [Alternative Temporal Accesses]*

c14. *$A \leq B \leq \cdots \leq M$ and $M \leq N \leq \cdots \leq \Omega$ imply $\exists T$ s.t. $B \leq N \leq \cdots \leq T$ and $A \leq T \leq \cdots \leq \Omega$. [Compactness on Order Decomposition]*

c15. *If $A \leq B \leq \ldots \Omega$ and $A \leq B \leq \ldots M$, then $\exists T, \Upsilon$ s.t. $A \leq B \leq \ldots M \leq \ldots T$ and $T \leq \Upsilon \leq \ldots \Omega$. [Compactness on Order Composition]*

c16. $A \leq B$ *and* $\alpha_i \preceq \beta_i \preceq \alpha_j$ *imply* $\alpha_i \preceq \alpha_j$. *[Temporal Contraction]*

c17. $A \leq B$, $\beta_i \preceq \alpha_j$, $\beta_i \in v(\phi^B)$ *and* $\alpha_i \notin v(\neg\phi^A)$ $\forall \alpha_i \preceq \alpha_j$ *imply* $\alpha_j \in v(\phi^A)$. *[Global Monotonicity]*

c18. $A \leq B$, $\beta_i \preceq \alpha_j$, $\beta_i \in v(\phi^B)$, *and* $\alpha_j \notin v(\phi^A)$ *imply* $\exists \alpha_j \succeq \alpha_i$ *s.t.* $\alpha_i \in v(\neg\phi^A)$, *imply* . *[Local Monotonicity by Distrust]*

c19. $B \leq A$, $\beta_i \succeq \alpha_i$, $\alpha_i \notin v(\phi^A)$, $\beta_i \in v(\phi^B)$ *imply* $\exists \alpha_j \succeq \alpha_i$ *s.t.* $\alpha_j \in v(\phi^A)$. *[Local Non-Monotonicity by Mistrust]*

Local Monotonicity is guaranteed for any local state with no access to other agents: in such condition, an agent A who holds ϕ_i true at some state, will hold it at any other state.

Temporality Order over Access establishes that there is always a temporal order matching the authorization access, with a latest local state for the agent with most authorization: this condition is made valid simply by the two designated states in the model.

Minimality guarantees that if the least authorised agent has access to some other agent's states, then there are temporal states of each that make their states accessible from one another and both are accessible from the latest state of the most authorized agent (i.e. the designated state α_n).

Implicative Closure guarantees closure of the temporal order over accessible states of agents.

Reflexive Access is valid for symmetrically accessible identical states.

Transitivity over Local Accesses grants transitivity of the temporal order over authorised accesses across agents.

Continuous Alternative Temporal Paths guarantees that given a local temporal path between the two designated states (earliest with minimal access, latest with most access), there is always a local state that connects them.

Temporal Seriality is the equivalent of the seriality axioms for normal modal logic: for every temporal local state there is a later one within that local state.

Instantiation guarantees that for each agents there are temporal states satisfying the order imposed by the authorizations.

Finality establishes the existence in the model of one agent which admits overall access to its states from other agents. This will be the agent whose earlier state is the designated ω_1.

Primality establishes the existence in the model of one agent with overall access to all other local states. This will be the agent whose latest state is the designated α_n.

Temporal Equivalence of Local States establishes that global access is preserved by by exchanging equivalent local local states.

Extension of Temporal States by Access says that access across agents' states implies a temporal order over such states.

Alternative Temporal Accesses states the existence of a temporally compact access across two alternative series of states that give access to the same local state.

Compactness on Order Decomposition states that given two access orders with some common term, there must be one at least local states that makes their chaining compact.

Compactness on Order Composition states that given an access order and a subchain of it, is it possible to identify two sets of local states in it that decompose the access in two compact chains of access with a shared set of local states.

Temporal Contraction states that given two temporal states of the same agent ordered through an external state, there is always a way to establish the temporal order among them.

Global Monotonicity is guaranteed between local states when the earliest one satisfies the content of the agent with the most access.

Local Monotonicity by Distrust expresses the dynamic of distrust as an access condition on temporal local states.

Local Non-Monotonicity by Mistrust expresses the dynamic of mistrust as an access condition on temporal local states.

We first show that a derivation constructs formulas that are true in some class of models, according to a filter obtained by some non-monotonic condition which needs resolution:

Theorem 4.2 (Soundness)**.** *If a judgement* $\Gamma^A \vdash \psi_j^B$ *is provable in* `(un)SecureND`*, then it is true in all filter models of a given class.*

Proof. First, we demonstrate when the following condition holds for all formulas ψ_j^N:

* If $U^{\mathcal{A}} \supseteq \{\gamma_i \preceq \beta_i \preceq \alpha_i\}$, $\gamma_i \vdash \psi_j^N$ and $\beta_i \vdash \psi_j^N$ and $A \leq B \leq \Gamma$, then $\alpha_i \vdash \psi_j^N$.

The proof proceeds by induction, with the cases for connectives requiring the conditions above to be restricted over one agent's domain according to the Definition 3.2 of Local Satisfaction, reflected in the conditions for models if condition [$c17$] is preserved:

l1 Use condition $c1$ iff $\psi_j^N \equiv 1$.

l2 Use condition $c2$ iff $\psi_j^N \equiv min$.

l3 Use condition $c3$ iff $\psi_j^N \equiv \phi_i^N \rightarrow \xi_j^N$.

l4 Use conditions $c4 - c6$ iff $\psi_j^N \equiv \phi_i^N \vee \xi_j^N$.

l5 The formula $\psi_j^N \equiv \phi_i^N \wedge \xi_j^N$ reduces to atomic instances of the two conjuncts.

For access rules generalised to a universe according to Definition 3.5 of Global Satisfaction, condition [$c17$] could be upheld: then if [$*$] fails, the formula at hand is an instance of distrust or mistrust; in either case, the following respectively must hold:

** If $U^{\mathcal{A}} \supseteq \{\gamma_i \preceq \beta_i \preceq \alpha_i\}$, $\gamma_i \vdash \psi_j^N$ and $\beta_i \vdash \psi_j^N$ and $A \leq B \leq \Gamma$, and $\alpha_i \nvdash \psi_j^N$, then $\exists \psi_j'^A$ s.t. $\alpha_i \vdash \psi_j'^A$ and $\gamma_i \preceq \beta_i \vdash \psi_j'^A$.

*** If $U^{\mathcal{A}} \supseteq \{\gamma_i \preceq \beta_i \preceq \alpha_i\}$, $\gamma_i \vdash \psi_j^N$ and $\beta_i \vdash \psi_j^N$ and $A \leq B \leq \Gamma$, and $\alpha_i \nvdash \psi_j^N$, then $\exists \psi_j'^A$ s.t. $\alpha_j \nvdash \psi_j'^A$ and $\gamma_i \preceq \beta_i \preceq \alpha_i \preceq \alpha_j$.

Then corresponding conditions on models are as follows:

g1 Use any of the required previous conditions, with the appropriate inverse order on $\preceq$ for $read(\psi_j^N)$.

g2 Use any of the required previous conditions, the appropriate inverse order on $\preceq$, and satisfaction of condition $c0$ for $trust(\psi_j^N)$.

g3 Use any of the required previous conditions, the appropriate inverse order on $\preceq$, condition $c0$ and condition $c8$ for $write(\psi_j^N)$.

g4 On the basis of failing condition $c17$ which expresses $dtrust(\psi_j^N)$-intro, use conditions $c18$ and $c8$ for $dtrust(\psi_j^N)$-elimination.

g5 On the basis of failing condition $c17$ which expresses $dtrust(\psi_j^N)$-intro, use conditions $c19$ and $c8$ for $mtrust(\psi_j^N)$-elimination.

Now we can prove the validity of initial formulas of `(un)SecureND` and the preservation by its rules. Suppose $\Gamma^A = \{\phi_1^A, \ldots, \phi_n^A\}$, then if the judgement $\Gamma^A \vdash \psi_j^B$ is valid in all models, it must be the case that for every $\alpha_i \in U^A$ it holds $\alpha_i \in U^A \vdash \{\phi_1^A, \ldots, \phi_n^A\} \rightarrow \psi_j^B$. Note that for $B \neq A$, $(*)$ and conditions for the access rules are always required, while for $B = A$, only conditions on the operational rules are needed.

- If $\Gamma = \emptyset$, Profile Extension corresponds to the identity judgement $\psi_j^B \rightarrow \psi_j^B$, which follows from $c4$ and $c8$.
- For the Atom Rule, $\psi_j^B \equiv \phi_i^A$, for some $\phi_i^A \in \Gamma^A$, which then follows from $c4$.
- For the $\bot$ rule, condition $c0$ fails for some $\phi_i^A \in \Gamma^A$; use $c9$ to imply any order;
- $\wedge$-I requires no specific conditions besides $c0$; $\wedge$-E preserves $c0$;
- $\vee$-I requires conditions $c8$–$c10$; $\vee$-E requires condition $c13$;
- for $\rightarrow$-I, use condition $c14$; for $\rightarrow$-E, use conditions $c10$ and $c15$;
- $\bot$ follows from $c0$;
- *read* follows from $*$ and $c8$;
- *trust* follows from $*$, $c0$;

- *write* follows from $*$, $c0$ and $c8$;
- *derive* follows from *write*;
- If a formula introduced by $DTrust$-I is valid by $c18$ and $**$, the negated atom is valid for $DTrust$-E by $c21$ and the appropriate conditions for *write*;
- If a formula introduced by $MTrust$-I is valid by $c19$ and $***$, the negated atom is valid for $DTrust$-E by $c23$ and the appropriate conditions for *write*;
- Weakening follows from $c11 - c12$, assuming $*$ and $c0$;
- Contraction from $c8$, $c10$ and $c16$;
- Exchange from $c8$ and $c11$;
- Cut from $c4, c5, c16$.

□

Definition 4.3 (Canonical Model)**.** *Let* $\Gamma^A \vdash \phi_j^B$ *be a judgement of* `(un)SecureND`*. A canonical model of* `(un)SecureND` *for* $\Gamma^A \vdash \phi_j^B$ *is the structure*

$$\mathcal{M} = \langle \mathcal{A}, \leq, \Lambda_{A \in \mathcal{A}}, \preceq, \alpha_n, \omega_1, U^{\Lambda_{A,\ldots,\Omega}}, v \rangle$$

as in Definition 3.1 where

1. $\mathcal{A}$ *is a non-empty set of agents corresponding to* $\mathcal{S}$ *in* `(un)SecureND`*;*
2. Λ_A *for each* $A \in \mathcal{A}$ *is non-empty;*
3. $\forall A, B \in \mathcal{A}$*,* $A \leq B$ *holds iff so is in* $\mathcal{S}$ *and* $\exists \alpha_i \in \Lambda_A, \beta_j \in \Lambda_B$ *such that* $\beta_j \preceq \alpha_i$*;*
4. $\alpha_n \in \Lambda_A$ *and* $\omega_1 \in \Lambda_\Omega$*;*
5. $U^{\mathcal{A}} \subset \Lambda_A \times \cdots \times \Lambda_\Omega$*;*
6. *for all* ϕ_i^A *s.t.* $\alpha_i \in v(\phi_i^A)$*, then for all* $\beta_j \preceq \alpha_i$ *and* $A \leq B$*, it holds* $\beta_j \in v(\phi_j^B)$*;*
7. *for all* ϕ_i^B *s.t.* $\beta_j \in v(\phi_i^B)$*, then for all* $\alpha_i \succeq \beta_j$*,* $A \leq B$ *and* $\alpha_i \in v(trust(\phi_i^B))$ *it holds* $\alpha_i \in v(\phi_i^A)$*;*

A canonical model is one that satisfies exactly all conditions $c0 - c17$ from Definition 4.1: this means that a Standard Model is one that does not require any instance of Mistrust or Distrust operation, as all local satisfaction operations are monotonic with respect to the global satisfaction relation. In the Definition: the first clause establishes the set of agents as they occur in the proof-theory; the second clause defines their local states; the third clause establishes that the

order across agents in the model is the same as those in the derivations, and that for every local state of each agent the accessibility relation matches the temporal relation, i.e. an agent with higher position in the accessibility relation will have information issued at states that are later or equivalent to the states of agent with lower position in the order; in the fourth clause, the latest state of the agent with highest access is the designated state with most access, and the earliest state of the agent with lowest access is the designated state with least access; by the fifth clause, the universe is the Cartesian product of all the local states of all agents; according to the sixth clause, monotonicity of the satisfaction relation in the model holds across local states of agents accessible from higher to lower position in the order relation; according to the seventh clause, a trust relation holds across local states of agents accessible from lower to higher position in the order relation. Hence, a canonical model is the maximally consistent set of formulas generated by local models accessing each other, when no filtering is required.

When the global monotonicity condition fails, a canonical model can be obtained as a filter model by satisfying also conditions $[c18 - c19]$: this means that some states in the model are removed (by either a mistrust or distrust operation) to restore monotonicity according to conditions $**$ or $***$.

Definition 4.4 (Canonical Filter Model)**.** *A canonical filter model $\mathcal{M}'$ of* `(un)SecureND` *for $\Gamma^A \vdash \phi_j^B$ is the structure obtained by a canonical model $\mathcal{M}$ where for all ϕ_i^B s.t. $\beta_i \in v(\phi_i^B)$, then for all $\alpha_i \succeq \beta_i$, $A \leqslant B$ and $\alpha_i \in v(trust(\neg\phi_i^A))$, it holds*

1. *either $\beta_i \notin U^{\mathcal{A}}$ and $\exists \beta_j \succeq \beta_i$ s.t. $\beta_j \in v(trust(\neg\phi_i^A))$*
2. *or $\alpha_i \notin U^{\mathcal{A}}$ and $\exists \alpha_j \succeq \beta_i$ s.t. $\alpha_j \in v(\phi_j^B)$.*

The canonical filter model restores monotonicity by satisfying one of two conditions:

1. for distrust by agent A: removing a local state β_i from the lower ranked, receiving agent B, validating a formula ϕ which is not trusted by another higher agent A; adding a state β_j validating such formula;
2. for mistrust by agent A: by removing a local state from a higher agent A, adding a later one which validates the formula received from a lower ranked agent B.

Lemma 4.5 (Truth Lemma)**.** *Let $\mathcal{M}'$ be a canonical filter model for* `(un)SecureND` *such that $U^A \in \mathcal{M}'$. Then for $\alpha_n \in U^A$ and every formula ϕ_i^A,*

$$\phi_i^A \in \alpha_n \leftrightarrow \mathcal{M}' \models \phi_i^A$$

Proof. By induction on the structure of ϕ_i^A. Let us start considering the simple cases of connectives (corresponding to the Operational Rules in `(un)SecureND`) within a given local state:

– It holds trivially if ϕ_i^A is an atomic proposition or $\bot$: this follows by the first clause in the Definition 3.2 of Local Satisfaction, Definition 3.3 and condition $c0$ of Local Monotonicity in Definition 4.1, so that $\phi_i^A \in \alpha_i$ for each $\alpha_i \in U^A$.

– If ϕ_i^A is a disjunction, it holds by definition of the connective, i.e. by atomic reduction to a previous state $\alpha_h \preceq \alpha_n \in U^A$ and condition $c0$ of Local Monotonicity.

– If ϕ_i^A is a conjunction, it holds by definition of the connective, i.e. by atomic reduction to the current state α_n.

– If ϕ_i^A is of the form $\phi_j^A \rightarrow \phi_k^A$: assume $\phi_j^A \rightarrow \phi_k^A \in \alpha_n$, by definition of the designated state $\alpha_n \succeq \alpha_j \in U^A$ and $\alpha_j \in U^A \models \phi_j^A$; then $\alpha_n \in U^A \models \phi_k^A$ and hence $\mathcal{M}' \models \phi_k^A$; as α_j is an arbitrary anterior state of α_n, $\mathcal{M}' \models \phi_j^A \rightarrow \phi_k^A$. Assume now that $\alpha_n \in U^A \not\models \phi_j^A \rightarrow \phi_k^A$: by deductive closure it must be the case that $Cn(\alpha_j \cup \phi_j^A) \not\supset \phi_k^A$, for some $\alpha_j \preceq \alpha_n$. Then there is α_k s.t. $Cn(\alpha_j \cup \phi_i^A) \preceq \alpha_k$ and $\alpha_k \not\models \phi_k^A$. Hence, by induction hypothesis $\mathcal{M}' \models \phi_j^A$ and $\mathcal{M}' \not\models \phi_k^A$. Because $\alpha_k \preceq \alpha_n \in U^A$, then the hypothesis $\alpha_n \in U^A \not\models \phi_j^A \rightarrow \phi_k^A$ is contradicted.

Let us now consider the access operators, which will allow to generalize to models with different local states:

– If ϕ_i^A is of the form $read(\psi_j^B)$: by induction hypothesis $read(\psi_j^B) \in \alpha_n$ means $\alpha_i \in U^{\mathcal{A}} \vDash read(\psi_j^B)$ for all α_i, and $\beta_j \in U^{\mathcal{A}} \vDash \psi_j^B$, by model condition $c4$; by the canonical model construction $A \leq B$, and condition $c10$ makes $read(\psi_j^B)$ holds for all $N \geq A$, hence $\mathcal{M}' \vDash read(\psi_j^B)$. For the opposite direction: assume $\mathcal{M}' \vDash read(\psi_j^B)$: then it must be the case that for every $A \in \mathcal{A}$, $\exists \alpha_i \succeq \beta_j$ s.t $\alpha_i \in U^{\mathcal{A}} \vdash read(\psi_j^B)$, hence $B \geq A$, and as this holds for any arbitrary α_i up to α_n, the left-to-right implication holds.

– If ϕ_i^A is of the form $trust(\psi_j^B)$: by induction hypothesis, $\alpha_n \in U^{\mathcal{A}} \vDash trust(\psi_j^B)$ and $\beta_j \in U^{\mathcal{A}} \vDash \psi_j^B$; by the canonical model construction $A \leq B$, and by the definition of the operator there is $\alpha_j \preceq \alpha_n$ such that $\alpha_n = Cn(\alpha_i \cup \psi_j^B)$. As this holds for any arbitrary α_j satisfying the previous assumptions, then $\mathcal{M}' \vDash trust(\psi_j^B)$. For the opposite direction: assume $\mathcal{M}' \vDash trust(\psi_j^B)$: then $\alpha_i \in U^{\mathcal{A}} \vDash trust(\psi_j^B)$ for any arbitrary α_i for which $read(\psi_j^B)$, hence for α_n.

– If ϕ_i^A is of the form $write(\psi_j^B)$: this follows directly from the previous case.

– If ϕ_i^A is of the form $dtrust(\psi_j^B)$: for the left-to-right direction, this follows from $\alpha_i \in U^{\mathcal{A}} \vDash trust(\neg\psi_j^B)$, for some $\alpha_i \preceq \alpha_n$ and some $\beta_i \preceq \alpha_i$ s.t $\beta_i \in U^{\mathcal{A}} \vDash trust(\psi_j^B)$, but such that $\exists \beta_j \succeq \beta_i \in U^{\mathcal{A}} \vDash dtrust(\psi_j^B)$: in

this case, remodulate $U^{\mathcal{A}}$ so as to obtain the filter model which includes b_j and purges b_i; for the right-to-left direction, pick the filter model of $\mathcal{M}$ which satisfies $trust(\psi_j^B)$ and such that for some $\beta_i \in U^{\mathcal{A}} \vDash trust(\psi_j^B)$ is purged.

– If ϕ_i^A is of the form $mtrust(\psi_j^B)$: for the left-to-right direction, this follows from $\alpha_i \in U^{\mathcal{A}} \vDash trust(\neg\psi_j^B)$, for some $\alpha_i \preceq \alpha_n$ and some $\beta_i \succeq \alpha_i$ s.t $\beta_i \in U^{\mathcal{A}} \vDash trust(\psi_j^B)$, but such that $\exists\beta_j \succeq \beta_i$ and $\beta_j \in U^{\mathcal{A}} \vDash trust(\neg\psi_j^B)$: in this case remodulate U^i so as to obtain the filter model which includes b_j and purge b_i; for the right-to-left direction, pick the filter model of $\mathcal{M}$ which satisfies $trust(\psi_j^B)$ for some β_i such that $\beta_i \in U^{\mathcal{A}} \vDash trust(\psi_j^B)$ is purged.

□

Theorem 4.6 (Completeness)**.** *If $\mathcal{M}' \vDash \phi_i^A$, then there is a branch of a derivation tree in* `(un)SecureND` *terminating in $\Gamma^A \vdash \phi_i^B$, for some $A \leq B$.*

Proof. By contradiction, assume $\alpha_j \in U^{\mathcal{A}} \vDash \phi_i^A$ for every $\alpha_j \succeq \omega_1$ and every $\Lambda_I \in U^{\mathcal{A}}$ in the class of filter models $\mathcal{M}'$ of $\mathcal{M}$ of interest, but $\Gamma^A \nvdash \phi_i^A$ in `(un)SecureND`. Then there must be a state $\omega_1 \preceq \alpha_i \preceq \alpha_j$ and $\alpha_i \in U^{\mathcal{A}}$ of $\mathcal{M}$ such that $\alpha_i \models \phi_i^A$. Now construct a relevant canonical filter model $\mathcal{M}'$ from Definition 4.4: ϕ_i^A will be distrusted or mistrusted and therefore the relevant state α_i will have been purged from the given filter model $\mathcal{M}'$ of $\mathcal{M}$. Hence, it will not hold at α_n and hence is not true in $\mathcal{M}'$ and cannot be valid. □

5 Discussion

In logic, trust has been object of many analyses, see e.g. Castelfranchi and Falcone (2010), most of them semantic ones, much less proof-theoretically. Also, most of them agree on the fact that trust is a form of the truster's belief in some trustee's property. Studies range from computational settings inspired by applications like security and access control, to more philosophically inspired problems. In the Bell-LaPadula Model (BLPM, Bell and LaPadula (1973)) a subject can read resources if the content's access group is dominated by the subject's access group ("no read up"), and can write only resources whose access group dominates the subject's access group ("no write down"); a trusted subject is allowed to violate the writing constraint above, if it is not against security by design. A domination relation among subjects (or groups of subjects) defines therefore the possibility for one to access by writing or reading resources from another subject. Trust is simply intended as a property of agents, while security is a property of the system defined independently of the former. In the Role Based Access Control security model (RBAC, Ferraiolo, Sandhu, Gavrila, Kuhn, and Chandramouli (2001)), subjects on lower integrity levels are not permitted to write resources on higher integrity levels ("no write up"); and subjects on higher integrity levels cannot be corrupted by accessing resources on lower

integrity levels ("no read down"). Trustworthiness of resources corresponds to prevention of unauthorized change. Trust-aware RBAC models have been recently explored, in which trustworthiness is either defined by temporal-spatial constraints (e.g. Bertino, Bonatti, and Ferrari (2001); Suroop Mohan Chandran (2005)), level-constraints (S. Chakraborty and Ray (2006)) or by explicitly requiring role related assessments based on the behavior history of the user, see e.g. Oleshchuk (2012) for a subjective logic model. A constrained transitive trust model for the latter logic is also offered in Jøsang and Pope (2005). In authentication logics, starting with Abadi, Burrows, Lampson, and Plotkin (1993), beliefs are explicitly used: its `says` modality for principals can be interpreted to include a trust relation, as fully explored in Genovese (2012), with application to a distributed setting, see Barker and Genovese (2011).Trust, and its transitivity, is treated by means of a modal logic approach in Demolombe (2004, 2011, 2017); Lorini and Demolombe (2009), by specifying several kinds of trustee's properties which are relevant in the context of communication. Another modal logic approach to trust is presented in Liau (2003), where a combination of belief, information acquisition and trust are modelled semantically and axiomatically: in this logic, $B_i(w)$ is the set of worlds that agent i considers possible under world w according to her belief; $I_{ij}(w)$ is what the agent i considers possible according to the information received from j; and for any $S \subseteq W$, $S \in T_{ij}(w)$ indicates that agent i trusts j's judgement on the truth of the proposition corresponding to S. The three operators are axiomatically related: information received from a source reputed trustworthy about that content is believed; and a source reputed trustworthy about that content is believed to be such; moreover, if an agent trusts another agent's judgement on some content, then her trust is independent of the syntactic form of that content. Trust is characterised by some useful properties: if an agent acquires contradictory information from two sources, then sche cannot be trusting both; if one source is at least as trustworthy as the other, then the latter is not trusted; trust is not closed under consequence, nor distributivity of conjunction, but trust of finer content can be extracted from negative trust on more expressive contents; here there is no distinction between different forms of negative trust: in fact, a particular aspect of this logic (differing from ours) is that agents can be trust both on positive formulas and their negative, irrespective of the current state of the trusting agent or any order on local states. Finally, here the problem of transitivity is treated in the form of transferability: if i believes that j trusts k, then i will also trust k due to the endorsement of j, which is considered different from strict trust transitivity, because it is expressed relatively to an actual belief and information transmission. This notion of trust shares with ours the assumption that agents are honest with respect to the information they share; but it differs from ours in that we do not force trusted contents to hold in the universe (i.e. across different local states: obviously, in our system if there are two local states with contradictory contents holding, in the filter model of a universe only one such content will be holding; note that a different filter model might be designed where the opposite content holds). The problem of trust transitivity has been largely discussed in the literature, see e.g. P. S. Chakraborty and Karform (2012);

Christianson and Harbison (1996); Jamali and Ester (2010); Jøsang, Marsh, and Pope (2006). Solutions include decentralised trust (Abdul-Rahman and Hailes (1997)), bounded-transitivity in authorization contexts (Chapin, Skalka, and Wang (2008)), and a constraint by guarantors in Clarke et al. (2009). Transitivity of trust is also analysed in the context of cryptographic applications, see e.g. Maurer and Schmid (1996). Trust is defined as occurrent and dispositional in a logic of time, action, beliefs and choices in Herzig, Lorini, Hübner, and Vercouter (2010).

Recently, research has started considering the different meanings of negative trust (Guha, Kumar, Raghavan, and Tomkins (2004); Marsh and Dibben (2005); McKnight and Chervany (2000); McKnight, Kacmar, and Choudhury (2003); Ziegler and Lausen (2005)). In the social sciences, distrust is response to lack of information and mistrust is former trust destroyed or healed, see Cvetkovich (1999); Cvetkovich and Lofstedt (1999); Sztompka (1999). The contextual account of Marsh and Dibben (2005) presents mistrust as misplaced trust, untrust as little trust and distrust as no trust. This approach designs a continuum between the positive and negative evaluations (with some blurry limit at trust value zero) but it abstracts from the reasons behind the attribution of these evaluations, in favour of a purely quantitative approach. On the other hand, it clearly refers to intentions when it distinguish between misplacement of trust and betrayal. Another account of misplaced trust is provided in Singh (2011). This work formulates trust as a generic modality not directly related to information or belief, although constrained by commitments and beliefs. Trust is defined by a possible worlds semantics, while commitments are expressed as abstractions; a notion of reality in the world is used to express a specific path of execution in the multi-agent setting of the semantics. In this way, well-placed trust means trust in a proposition occurring in a real path across worlds; negative trust is reduced to misplaced trust, which on the contrary means trust in a proposition which does not occur in a real path. Note that in this case the trustee is not responsible for failing to deliver the content of the trusted proposition, as commitment wa snot necessarily formulated. This is a substantial difference with our system, where the trustor can apply different negative trust strategies, for example on the assumption that the trustee as the source of information is intentionally betraying the trustor's trust.

In fields like Human-Computer Interaction and Cyber-Physical Systems, the intentional requirement on trust emerges more clearly. In Hoffman et al. (2009), a scale is considered:

> *unjustified trust* (antitrust) $\rightarrow$ *justified trust* (skeptical) $\rightarrow$ *conditional trust* (contingent) $\rightarrow$ *unconditional trust* (faith)

Among the few semantic formal models that present a similar scalar understanding of trust is A. Baltag and Smets. (2012), where the informational stance of an agent is dependent from her attitude towards the source of information, encoded as strategies for belief change. Propositional attitudes are distinguished among (Irrevocable) knowledge, (Simple) belief, Strong belief, Triviality and

Inconsistency, to each of which corresponds an upgrade of the knowledge state translated into a dynamic attitude:

> *isolation* $\rightarrow$ *neutrality* $\rightarrow$ *semi-positive minimal trust* $\rightarrow$ *minimal trust* $\rightarrow$ *strong trust* $\rightarrow$ *infallible trust*

The logic of negative trust `(un)SecureND` introduced in this work is formulated in a general proof-theoretic format, comprehensive of structural rules, a semantics, and meta-theoretical results. The proof-theoretic fragment of the logic has also been presented in a number of more constrained and application-oriented formats. A first application of `(un)SecureND` has been to the context of software management. In Boender, Primiero, and Raimondi (2015), we considered security threats in software installation processes, posed by transitively trusted dependencies between packages from distinct repositories. To analyse them, a Coq implementation has been defined, using an explicit trust function to bridge repository access and software package installation rights and for verification of its formal correctness. Thereby, we resolve a version of the minimum install problem under trust conditions on repositories. In Primiero and Boender (2017, 2018) we extended the analysis to negative trust as considered in the present paper to analyse the complementary issue of packages' removal, both in case of conflicts and of security issues. We identify packages that are undesirable in view of the current installation profile; and currently installed packages that become inconsistent with a new intended installation. That formulation differs from the present more general format in that the access rules are completed by a $\neg-$distribution rule, essential to preserve completeness across the positive and negative fragments of the language: it ensures that if an operation is not possible on a formula ϕ_i^B under context Γ^A, then the same operation must be possible for a contradictory formula in that same context. This formula implements a closed-world assumption: it assumes that the set of atomic formulas on which operations can be performed is finite and defined for every agent, and that each agent can explicitly formulate the operations allowed on each formula. While this rule indicates a logic stronger than intuitionistic, the system in the current formulation remains weaker than classical: the operational rules from Figure 2 are inspired by intuitionistic connective, and for the operational rules of Figure 3 double negation elimination cannot hold, as $\neg\neg Trust(\phi)$ can be inferred from the two distinct protocols of distrust and mistrust. A second variant of our calculus of negative trust dubbed $\texttt{SecureND}^{sim}$ has been applied to contradictory information transmissions in networks, presented in Primiero, Raimondi, Bottone, and Tagliabue (2017). The networks of interest are built by ranked agents with different epistemic attitudes. In this context, positive trust is a property of the communication between agents required when message passing is executed bottom-up in the hierarchy, or as a result of a sceptic agent checking information. These two situations are associated with a confirmation procedure that has an epistemic cost. Negative trust results from refusing verification, either of contradictory information or because of a lazy attitude. The procedural semantics of the logic is implemented in a NetLogo simulation to test experimentally

its formal properties. Results suggest that a sceptic approach is favourable when maximisation of consensus is the goal; a lazy approach should be pursued when minimisation of costs is the goal. Ranking of initial nodes is only of little relevance to consensus reaching, while a rigidly structured network (linear) is the most expensive in this respect. Finally, the experimental analysis shows that trust is a better mean to information propagation than distrust. The presence of contradictory information is by itself the cause of distrust generation, independently from the initial attitude of the agents. The format of $\texttt{SecureND}^{sim}$ can be further explored in view of different types of agents (e.g. paranoid agents, distrusting all information coming from an agent higher in the hierarchy). The general format of the logic abstracts away from both epistemic attitudes and networks of information transmission, relying only on the access relation defined across agents. Finally, a last field of application for our logic of negative trust has been communication protocols in vehicular ad-hoc networks (VANET). In the context of such networks, security requirements need to rely on a combination of reputation of communicating agents and trust relations over the messaging framework. This is crucial in order to maintain dynamic and safe behaviour under all circumstances. In Primiero, Raimondi, Chen, and Nagarajan (2017) we have adapted the logic to a reputation and trust model for VANETs, exploiting its formal verification through the existing translation into the Coq proof assistant, so as to guarantee consistency of messaging protocols and security of transitive transmissions. In Primiero, Martorana, and Tagliabue (2018) we focused on the networks' vulnerability to attacks by malicious users. Despite their characterization as dynamically reconfigurable networks, for security reasons it is nonetheless essential to identify topology and population properties that can optimise mitigation protocols' deployment. We have provided an algorithmic definition and simulation of a trust and mitigation based protocol inspired by `(un)SecureND` to contain a Black Hole style attack, i.e. one where malicious agents divert all the traffic toward themselves without forwarding the (non-malicious) data packets to the neighbouring nodes. We experimentally showed its optimal working conditions using repeated broadcasting, opportunistic message forwarding and testing on real data. This application, in particular, relies on the configuration of several agents acting as purposefully distrustful ones, to implement a common strategy. The variety of applications and different contexts in which we have modelled `(un)SecureND` shows the general nature of the trust protocol it expresses, as formulated in the present paper.

We have stressed that the semantics of negative trust implemented by `(un)SecureND` can be seen as non-deterministic. Its resolution corresponds to the problem of defining strategies to establish which negative trust operation should an agent perform when faced with a message that fails the consistency check in the `trust` rule. This corresponds to establish under which conditions an agent should perform revision on her own belief base, or when she should reject incoming information. The extension of `(un)SecureND` presented in Ceolin and Primiero (2019) offers a number of such strategies based on calculating a value of trustworthiness for each source, and partitioning such set for each receiver into higher and lower ranked sources. In this case, a measure of trustworthiness for

each profile is computed as a function of three parameters:

- *Knowledgeability*: the number of profiles accessible from an agent's profile A.
- *Popularity*: the number of sources having access to A.
- *Reputation*: the proportion between trust and negative trust operations on formulas indexed by A.

On the basis of an initial computational evaluation of such notion of trustworthiness, different strategies are proposed to establish further negative trust strategies by any agent; for all of them, consider the set of agents profiles with trustworthiness higher than the profile currently involved in an inconsistent reading operation:

- distrust: an agent distrusts received content if sent from a source with lower trustworthiness;
- weak mistrust: the agent accepts incoming information and removes from its own profile any conflicting information by the simple presence of the sender in the set of sources with higher reputation;
- majority mistrust: requires computing the partitions of the set of sources with higher trustworthiness than A and comparing their cardinality: any content held by the larger partition will be kept by the receiver (including the case in which this reduces to an application of a distrust rule);
- weighted majority: the essential condition is expressed by the higher average reputation of the partition.
- complete mistrust: the agent requires that every element in the set of sources with higher reputation agrees on the received message to accept it and remove conflicting information on her own profile.

This assessment is used to model a real-case scenario of source assessment to decide the agent position in an online debate, and the selection strategies are applied to the resulting source hierarchy. We show that a linear combination of these parameters presents a decent correlation with user-provided assessments.

6 Conclusions

We have presented the logic `(un)SecureND`, its proof theory, with a number of structural properties, its relational semantics and soundness and completeness results.

The main limitation of the logic `(un)SecureND` as developed so far is its binary treatment of trust. A more interestingly and dynamic extension of the

logic can be formulated through a probabilistic assessment of source trustworthiness. Starting from the existing model of trustworthiness assessment developed in Ceolin and Primiero (2019), we aim at providing first a probabilistic method to reason systematically on the trustworthiness of information sources. While semantic approaches and constraints on sharing abilities of online platforms are difficult and inconvenient, such a method would focus on determining which information sources are most trustworthy, by reasoning on their knowledgeability, popularity and reputation. This model can be further implemented in a distributed ledger structure for trustworthiness assessment in public debates, developing the DAG structure considered in Bottone, Raimondi, and Primiero (2018). The blockchain technology seems not to have yet been tested in this context. The ledger can provide opinion debates with a register which accounts for previously held positions by participants, maintaining a traceable history of the evolution of the debate; the model guarantees always to each participant the right to access the debate with a probability which is never zero, at the same time establishing priority on the basis of an acquired level of trustworthiness; the model grants a dynamic debate, with a revision option for previously held positions; the structure models a dynamic notion of source trustworthiness, depending on other agents' assessments and from the source ability to remain updated on the opinions available in the debate and it finally requires participants to the debate to be informed and knowledgeable about other participants' opinions, thus increasing the overall reliability of the debate.

From the point of view of the semantic analysis of negative trust by `(un)SecureND`, one main task is the simplification and generalization of the relational filter models presented in this paper. One way this could be achieved is by modelling accessibility and temporality among agents' states by a unique relation. While this would undoubtedly simplify the formal treatment of our relations of interest, it would make accessibility dependent from the temporal order. One sort of applications which could be usefully interpreted by such simplified model are operations across servers where data accessibility could be temporally restricted, e.g. in conditions of upgrade or for security reasons.

The development of such models and applications for the probabilistic extension of `(un)SecureND` are the next steps of this research.

References

Abadi, M., Burrows, M., Lampson, B., & Plotkin, G. (1993, September). A calculus for access control in distributed systems. *ACM Trans. Program. Lang. Syst.*, *15*(4), 706–734. Retrieved from `http://doi.acm.org/10.1145/155183.155225`

A. Baltag, B. R., & Smets., S. (2012, August). Doxastic attitudes as belief revision policies. In *Online pre-proceedings of esslli 2012 special workshop on "strategies for learning, belief revision and preference change", opole, poland.* Retrieved from `http://www.ninagierasimczuk.com/LBP2012/accepted.html`

Abdul-Rahman, A., & Hailes, S. (1997). A distributed trust model. In T. Haigh, B. Blakley, M. E. Zurko, & C. Meodaws (Eds.), *Proceedings of the 1997 Workshop on New Security Paradigms, Langdale, Cumbria, United Kingdom, September 23-26, 1997* (pp. 48–60). ACM. Retrieved from http://doi.acm.org/10.1145/283699.283739

Barker, S., & Genovese, V. (2011). Socially constructed trust for distributed authorization. In V. Atluri & C. Díaz (Eds.), *Esorics* (Vol. 6879, p. 262-277). Springer.

Bell, D. E., & LaPadula, L. J. (1973). *Secure computer systems: Mathematical foundations* (Tech. Rep. Nos. MTR-2547, Vol. 1). Bedford, MA: MITRE Corp.

Bertino, E., Bonatti, P. A., & Ferrari, E. (2001, August). Trbac: A temporal role-based access control model. *ACM Trans. Inf. Syst. Secur.*, *4*(3), 191–233. Retrieved from http://doi.acm.org/10.1145/501978.501979

Boender, J., Primiero, G., & Raimondi, F. (2015). Minimizing transitive trust threats in software management systems. In A. A. Ghorbani et al. (Eds.), *13th Annual Conference on Privacy, Security and Trust, PST 2015, Izmir, Turkey, July 21-23, 2015* (pp. 191–198). IEEE. Retrieved from http://dx.doi.org/10.1109/PST.2015.7232973

Bottone, M., Raimondi, F., & Primiero, G. (2018). Multi-agent based simulations of block-free distributed ledgers. In L. Barolli, M. Takizawa, T. Enokido, M. R. Ogiela, L. Ogiela, & N. Javaid (Eds.), *32nd international conference on advanced information networking and applications workshops, AINA 2018 workshops, krakow, poland, may 16-18, 2018* (pp. 585–590). IEEE Computer Society. Retrieved from https://doi.org/10.1109/WAINA.2018.00149

Castelfranchi, C., & Falcone, R. (2010). *Trust theory: A Socio-cognitive and computational model*. Wiley.

Ceolin, D., & Primiero, G. (2019). A granular approach to source trustworthiness for negative trust assessment. In W. Meng, P. Cofta, C. D. Jensen, & T. Grandison (Eds.), *Trust management XIII - 13th IFIP WG 11.11 international conference, IFIPTM 2019, copenhagen, denmark, july 17-19, 2019, proceedings* (Vol. 563, pp. 108–121). Springer. Retrieved from https://doi.org/10.1007/978-3-030-33716-2_9

Chakraborty, P. S., & Karform, S. (2012). Designing Trust Propagation Algorithms based on Simple Multiplicative Strategy for Social Networks. *Procedia Technology*, *6*(0), 534–539. Retrieved from http://www.sciencedirect.com/science/article/pii/S2212017312006093 (2nd International Conference on Communication, Computing & Security [ICCCS-2012])

Chakraborty, S., & Ray, I. (2006). Trustbac: Integrating trust relationships into the rbac model for access control in open systems. In *Proceedings of the eleventh acm symposium on access control models and technologies* (pp. 49–58). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/1133058.1133067

Chapin, P. C., Skalka, C., & Wang, X. S. (2008). Authorization in trust

management: Features and foundations. *ACM Comput. Surv.*, *40*(3). Retrieved from http://doi.acm.org/10.1145/1380584.1380587

Christianson, B., & Harbison, W. S. (1996). Why Isn't Trust Transitive? In T. M. A. Lomas (Ed.), *Security Protocols, International Workshop, Cambridge, United Kingdom, April 10-12, 1996, Proceedings* (Vol. 1189, pp. 171–176). Springer. Retrieved from http://dx.doi.org/10.1007/3-540-62494-5_16

Clarke, S., Christianson, B., & Xiao, H. (2009). Trust*: Using local guarantees to extend the reach of trust. In B. Christianson, J. A. Malcolm, V. Matyas, & M. Roe (Eds.), *Security protocols workshop* (Vol. 7028, p. 171-178). Springer.

Cvetkovich, G. (1999). The attribution of social trust. In G. Cvetkovih & R. Lofstedt (Eds.), *Social Trust and the Management of Risk* (pp. 53–61). Earthscan.

Cvetkovich, G., & Lofstedt, R. E. (1999). Social trust and culture in risk management. In G. Cvetkovih & R. Lofstedt (Eds.), *Social Trust and the Management of Risk* (pp. 9–21). Earthscan.

Demolombe, R. (2004). Reasoning about trust: A formal logical framework. In C. D. Jensen, S. Poslad, & T. Dimitrakos (Eds.), *Trust management, second international conference, itrust 2004, oxford, uk, march 29 - april 1, 2004, proceedings* (Vol. 2995, pp. 291–303). Springer. Retrieved from https://doi.org/10.1007/978-3-540-24747-0_22

Demolombe, R. (2011). Transitivity and propagation of trust in information sources: An analysis in modal logic. In J. Leite, P. Torroni, T. Ågotnes, G. Boella, & L. van der Torre (Eds.), *Computational logic in multi-agent systems - 12th international workshop, CLIMA xii, barcelona, spain, july 17-18, 2011. proceedings* (Vol. 6814, pp. 13–28). Springer. Retrieved from https://doi.org/10.1007/978-3-642-22359-4_2

Demolombe, R. (2017). Reasoning about trust and aboutness in the context of communication. *Journal of Applied Non-Classical Logics*, *27*(3-4), 292–303. Retrieved from https://doi.org/10.1080/11663081.2017.1420316

Ferraiolo, D. F., Sandhu, R., Gavrila, S., Kuhn, D. R., & Chandramouli, R. (2001, August). Proposed nist standard for role-based access control. *ACM Trans. Inf. Syst. Secur.*, *4*(3), 224–274. Retrieved from http://doi.acm.org/10.1145/501978.501980

Genovese, V. (2012). *Modalities in access control: Logics, proof-theory and applications* (Unpublished doctoral dissertation). University of Luxembourg and University of Torino.

Guha, R. V., Kumar, R., Raghavan, P., & Tomkins, A. (2004). Propagation of trust and distrust. In S. I. Feldman, M. Uretsky, M. Najork, & C. E. Wills (Eds.), *Proceedings of the 13th international conference on World Wide Web, WWW 2004, New York, NY, USA, May 17-20, 2004* (pp. 403–412). ACM. Retrieved from http://doi.acm.org/10.1145/988672.988727

Herzig, A., Lorini, E., Hübner, J. F., & Vercouter, L. (2010). A logic of trust and reputation. *Logic Journal of the IGPL*, *18*(1), 214–244. Retrieved

from https://doi.org/10.1093/jigpal/jzp077

Hoffman, R. R., Lee, J. D., Woods, D. D., Shadbolt, N., Miller, J., & Bradshaw, J. M. (2009, November/December). The dynamics of trust in cyberdomains. *IEEE Intelligent Systems*, 5–11.

Jamali, M., & Ester, M. (2010). A Matrix Factorization Technique with Trust Propagation for Recommendation in Social Networks. In *Proceedings of the Fourth ACM Conference on Recommender Systems* (pp. 135–142). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/1864708.1864736

Jøsang, A., Marsh, S., & Pope, S. (2006). Exploring Different Types of Trust Propagation. In K. Stølen, W. Winsborough, F. Martinelli, & F. Massacci (Eds.), *Trust Management* (Vol. 3986, pp. 179–192). Springer Berlin Heidelberg. Retrieved from http://dx.doi.org/10.1007/11755593_14

Jøsang, A., & Pope, S. (2005). Semantic constraints for trust transitivity. In *Proceedings of the 2nd asia-pacific conference on conceptual modelling - volume 43* (pp. 59–68). Darlinghurst, Australia, Australia: Australian Computer Society, Inc. Retrieved from http://dl.acm.org/citation.cfm?id=1082276.1082284

Liau, C. (2003). Belief, information acquisition, and trust in multi-agent systems–a modal logic formulation. *Artif. Intell.*, *149*(1), 31–60. Retrieved from https://doi.org/10.1016/S0004-3702(03)00063-8

Lorini, E., & Demolombe, R. (2009). From trust in information sources to trust in communication systems:an analysis in modal logic. In J.-J. C. Meyer & J. Broersen (Eds.), *Knowledge representation for agents and multi-agent systems* (pp. 81–98). Berlin, Heidelberg: Springer Berlin Heidelberg.

Marsh, S., & Dibben, M. (2005). Trust, Untrust, Distrust and Mistrust – An Exploration of the Dark(er) Side. In P. Herrmann, V. Issarny, & S. Shiu (Eds.), *Trust Management* (Vol. 3477, pp. 17–33). Springer Berlin Heidelberg. Retrieved from http://dx.doi.org/10.1007/11429760_2

Maurer, U. M., & Schmid, P. E. (1996). A calculus for security bootstrapping in distributed systems. *Journal of Computer Security*, *4*(1), 55-80.

McKnight, D. H., & Chervany, N. L. (2000). Trust and Distrust Definitions: One Bite at a Time. In R. Falcone, M. P. Singh, & Y. Tan (Eds.), *Trust in Cyber-societies, Integrating the Human and Artificial Perspectives* (Vol. 2246, pp. 27–54). Springer. Retrieved from http://dx.doi.org/10.1007/3-540-45547-7_3

McKnight, D. H., Kacmar, C., & Choudhury, V. (2003). Whoops...Did I Use the Wrong Concept to Predict E-Commerce Trust? Modeling the Risk-Related Effects of Trust versus Distrust Concepts. In *36th Hawaii International Conference on System Sciences (HICSS-36 2003), CD-ROM / Abstracts Proceedings, January 6-9, 2003, Big Island, HI, USA* (p. 182). IEEE Computer Society. Retrieved from http://dx.doi.org/10.1109/HICSS.2003.1174393

Oleshchuk, V. A. (2012). Trust-aware rbac. In I. V. Kotenko & V. A. Skormin (Eds.), *Mmm-acns* (Vol. 7531, p. 97-107). Springer. Retrieved from http://dblp.uni-trier.de/db/conf/

mmmacns/mmmacns2012.html#Oleshchuk12

Primiero, G. (2016). A Calculus for Distrust and Mistrust. In S. M. Habib, J. Vassileva, S. Mauw, & M. Mühlhäuser (Eds.), *Trust Management X - 10th IFIP WG 11.11 International Conference, IFIPTM 2016, Darmstadt, Germany, July 18-22, 2016, Proceedings* (Vol. 473, pp. 183–190). Springer. Retrieved from http://dx.doi.org/10.1007/978-3-319-41354-9_15

Primiero, G., & Boender, J. (2017). Managing software uninstall with negative trust. In J. Steghöfer & B. Esfandiari (Eds.), *Trust management XI - 11th IFIP WG 11.11 international conference, IFIPTM 2017, gothenburg, sweden, june 12-16, 2017, proceedings* (Vol. 505, pp. 79–93). Springer. Retrieved from https://doi.org/10.1007/978-3-319-59171-1_7

Primiero, G., & Boender, J. (2018). Negative trust for conflict resolution in software management. *Web Intelligence*, *16*(4), 251–271. Retrieved from https://doi.org/10.3233/WEB-180393

Primiero, G., & Kosolosky, L. (2013). The Semantics of Untrustworthiness. *Topoi*, *35*(1), 253–266. Retrieved from http://dx.doi.org/10.1007/s11245-013-9227-2

Primiero, G., Martorana, A., & Tagliabue, J. (2018). Simulation of a trust and reputation based mitigation protocol for a black hole style attack on vanets. In *2018 IEEE european symposium on security and privacy workshops, euros&p workshops 2018, london, united kingdom, april 23-27, 2018* (pp. 127–135). IEEE. Retrieved from https://doi.org/10.1109/EuroSPW.2018.00025

Primiero, G., & Raimondi, F. (2014). A typed natural deduction calculus to reason about secure trust. In A. Miri, U. Hengartner, N. Huang, A. Jøsang, & J. García-Alfaro (Eds.), *2014 Twelfth Annual International Conference on Privacy, Security and Trust, Toronto, ON, Canada, July 23-24, 2014* (pp. 379–382). IEEE. Retrieved from http://dx.doi.org/10.1109/PST.2014.6890963

Primiero, G., Raimondi, F., Bottone, M., & Tagliabue, J. (2017). Trust and distrust in contradictory information transmission. *Applied Network Science*, *2*, 12. Retrieved from https://doi.org/10.1007/s41109-017-0029-0

Primiero, G., Raimondi, F., Chen, T., & Nagarajan, R. (2017). A proof-theoretic trust and reputation model for VANET. In *2017 IEEE european symposium on security and privacy workshops, euros&p workshops 2017, paris, france, april 26-28, 2017* (pp. 146–152). IEEE. Retrieved from https://doi.org/10.1109/EuroSPW.2017.64

Primiero, G., & Taddeo, M. (2012). A modal type theory for formalizing trusted communications. *J. Applied Logic*, 92–114.

Restall, G. (2000). *An Introduction to Substructural Logics*. Routledge.

Singh, M. P. (2011). Trust as dependence: a logical approach. In L. Sonenberg, P. Stone, K. Tumer, & P. Yolum (Eds.), *10th international conference on autonomous agents and multiagent systems (AAMAS 2011), taipei, taiwan, may 2-6, 2011, volume 1-3* (pp. 863–870). IFAAMAS. Retrieved from http://portal.acm.org/citation.cfm?id=

2031741&CFID=54178199&CFTOKEN=61392764

Suroop Mohan Chandran, J. (2005). Lot-rbac: A location and time-based rbac model. In A. H. H. Ngu, M. Kitsuregawa, E. J. Neuhold, J.-Y. Chung, & Q. Z. Sheng (Eds.), *Wise* (Vol. 3806, p. 361-375). Springer. Retrieved from http://dblp.uni-trier.de/db/conf/wise/wise2005.html#ChandranJ05

Sztompka, P. (1999). *Trust: a sociological theory.* Cambridge University press.

Ziegler, C.-N., & Lausen, G. (2005, December). Propagation Models for Trust and Distrust in Social Networks. *Information Systems Frontiers*, *7*(4-5), 337–358. Retrieved from http://dx.doi.org/10.1007/s10796-005-4807-3