Second hand information and imperfect information sources

Laurence Cholvy

ONERA Centre de Toulouse
2 avenue Edouard Belin
31055 Toulouse, France,
cholvy@onera.fr

Abstract. In a situation of communication, a rational agent who receives a piece of information of interest has to decide if it can adopt it as a new belief. This paper shows that, for doing so, the agent may use its own beliefs about information sources. Symmetrically, the agent may want to know what it can assume about information sources if it wants to believe information they provide. This is the second point addressed in this paper. This work addresses the case when information of interest is provided directly by an agent. But it also addresses a more general case by considering second-hand information i.e when the source of information does not provide directly information of interest but when it cites another source. Furthermore, this study assumes that information sources may be imperfect and deliver false information.

1 Introduction

In a situation of communication, a rational agent who receives a piece of information has to decide if it can trust it before revising its own belief base [1, 26, 14, 16]. Obviously, the agent may believe this piece of information if it trusts the agent who reported it for reporting true information i.e if it trusts the information source (or source for short) for reporting true information. We also claim that the agent may believe that this piece of information is false if it trusts the source for always reporting false information.

For instance, assume that a friend who lives in the mountains calls me and tells that the outside temperature has risen recently. If I trust this friend for always telling me true information, then I can believe what he is reporting i.e I can believe that the temperature near the glacier has risen recently. Now, assume a different situation: I watch TV forecast and they say that the temperature in the mountains has risen recently. If I think that the weather forecasts are always false on this channel, then I can infer that what they report is false i.e I can conclude that the temperature has not risen in the mountains.

In [9], Demolombe studied several properties of communicating agents among which *validity* and *completeness* are noteworthy. Validity and completeness of a communicating agent depends on the piece of information it provides and on the agent it informs. More precisely, an agent i is valid for a piece of information

p and towards another agent j if i informs j of p only if p is true; an agent i is complete for a piece of information p and towards another agent j if, if p is true than i informs j of p. Notice how validity and completeness are important in the question we address. If the agent who receives a new piece of information trusts the information source for being valid, then it can believe this piece of information. At the opposite, an agent who does not receive a given piece of information from an information source assumed to be complete can conclude that this piece of information is false.

Demolombe's work assumes that the information of interest is directly provided by the information sources. But it may happen that an information source does not directly provide the piece of information of interest but cites another source which provides it. This is the case when, for instance, I am informed by my neighbour that one of his friends who lives near the glacier told him that the temperature has risen. *In this case, the information my neighbour reports is that one of his friends has reported that the temperature has risen. My neighbour does not report himself that the temperature has risen.* This is a more delicate issue. Indeed, trusting my neighbour for giving true forecast is not useful here. However, trusting him not to lie and trusting his friend for giving true forecast can allow me to conclude that, indeed, the temperature has risen. We insist on the fact that the very information which interests me (has the temperature risen ?) is reported via two agents: my neighbour's friend and my neighbour, the second citing the first.

In this work, we first present a model that shows how an agent can use its beliefs about sources in order to decide if it believes information they deliver. We address this question in the general case of second-hand information, i.e information reported via several agents who cite each others. We also assume that information sources may be imperfect and deliver false information.

To our knowledge, the only work which has addressed the case of second-hand information is [13]. There, the author deals with analysis of press dispatches like for instance: "according to minister X, the crash has been caused by bad weather conditions". Here the press dispatch does not inform about the cause of the crash but it cites X's report giving the cause of the crash. According to [13], the main factors used to evaluate such a report are: the quality of the sources, the opinion of the sources towards what they report (I strongly believe, it could be that...) and the relationships existing between actors (hostility, neutral, alliance). In this work, each factor is defined as a criteria and finally, a multi-criteria aggregation method is used to estimate the credibility of the report.

In this present paper, we adopt a different point of view and propose a logical model ([3]) which is based on Demolombe's model. Not only this model emphasizes the properties of validity and completeness of information sources but it also considers their negative counterparts. The negative counterpart of validity characterizes agents we call *misinformers* and the negative counter part of completeness characterize agents we call *falsifiers*. Roughly speaking, a misinformer for a proposition reports it only if it is false and a falsifier for a given proposition reports it if it is false.

We also show that this logical model can be a basis to abductively infer sources properties that allow the agent to believe information they provide. This part of the work has not been addressed before.

The paper is organised as follows. In section 2, we introduce some properties of communicating agents. A modal logic is used for giving them formal definitions. Section 3 shows how this logical model can be used by an agent to decide if it believes a piece of information it receives. Section 4 shows how this logical model can be a basis to abductively generate the sources properties which ensure that an agent can believe information they provide. Finally section 5 concludes the paper.

2 Properties of agents

2.1 Starting point

The work which influenced us is Demolombe's work [9] which, in particular, formalizes in modal logic the relations which exist between a piece of information, its truth and the mental attitudes of the agent which produces this piece of information. The operators of the modal logic used in this paper are: B_i ($B_i p$ means "agent i believes that p"), K_i ($K_i p$ means "agent i strongly believes that p); I_i^j ($I_i^j p$ means "agent i informs agent j that p"). Operator B_i obeys KD system which is quite usual for beliefs and operator I_i^j only obeys rule of equivalence substitutivity ([6]) i.e., if $p \leftrightarrow q$ is a theorem then $I_i p \leftrightarrow I_i^j q$ is a theorem too. K_i obeys KD system plus the axiom $K_i(K_i p \to p)$. Furthermore, the following axioms are considered: $K_i p \to B_i p$, $I_i^j p \to K_j I_i^j p$ and $\neg I_i^j p \to K_j \neg I_i^j p$.

According to Demolombe, agents can have the following properties:

- *Sincerity*: Agent i is sincere with regard to agent j for information p iff, if i informs j that p, then i believes p. I.e. a sincere agent believes what it says. Thus $sincere(i, j, p) \equiv I_i^j p \to B_i p$.
- *Competence*[1]: Agent i is competent about p iff, if i believes p then p is true. I.e., the beliefs of a competent agent are true. Thus $competent(i, p) \equiv B_i p \to p$.
- *Validity*: Agent i is valid with regard to j for p iff, if i informs j about p, then p is true. I.e., a valid agent tells the truth. Thus $valid(i, j, p) \equiv I_i^j p \to p$.
- *Cooperativity* : Agent i is cooperative with regard to agent j for p iff if a believes p then it informs b about p. I.e., i does not hide p to j. Thus: $cooperative(i, j, p) \equiv B_i p \to I_i^j p$.
- *Vigilance* Agent i is vigilant about p iff if p is the case then a belives it. I.e., $vigilant(i, p) \equiv p \to B_i p$.
- *Completeness* Agent i is complete with regard to j for p iff if p is the case then i informs j about p. Thus: $complete(i, j, p) \equiv p \to I_i^j p$.

[1] Demolombe also calls it credibility

Notice that for any i,j and p,

$$sincere(i,j,p) \wedge competent(i,p) \rightarrow valid(i,j,p)$$
$$vigilant(i,j,p) \wedge cooperative(i,p) \rightarrow complete(i,j,p)$$

All these notions are then used to derive the beliefs of an agent who receives a piece of information.

For instance, $I^i_j p \wedge K_i valid(j,i,p) \rightarrow K_i p$ is a theorem. It shows that an agent's strong belief about the validity of the agent who emits the new information influences its own strong belief about this information. An instance of this theorem is: $I^i_f risen \wedge K_i valid(f,i,risen) \rightarrow K_i risen$ which means that if my friend tells me that the temperature has risen in the mountains and if I strongly believe that my friend is valid with regard to the weather report in the moutains, then I strongly believe that the temperature has rised there.

2.2 Our proposal

The six previous properties characterize "perfect" information sources since we expect that agents which provide information fulfill most of these properties. However, there are situations in which information sources are not as perfect as this. For instance, in intelligence domain [18], information is often obtained by unreliable, non-cooperative or lying sources. Consequently, negative counterparts of these properties must also be considered. They have first been introduced in [3] and they are:

- *Hypocrite agent:* Hypocrisy is the counterpart of sincerity. Roughly speaking, an agent is a hypocrite when he believes the opposite of what he says, i.e if the agent says something, then, he believes that it is false.
- *Incompetent agent:* Incompetence is the counterpart of competence. Roughly speaking, an agent is incompetent when what he believes is false in the world.
- *Misinformer*[2] *agent:* Misinformerness aims to be the counterpart of validity. Roughly speaking, an misinformer agent only reports false information.
- *Mythomaniac agent:* Mythomania is the counterpart of cooperativity. Roughly speaking, an agent is mythomaniac when he says the opposite of what he believes, which means that if he believes something then he will say the contrary.
- *Agent subject to hallucination:* Hallucination is the counterpart of vigilance. Roughly speaking, an agent is subject of hallucination when he believes the opposite of what is true in the world.
- *Falsifier agent:* This is the couterpart of completeness. Roughly speaking, an agent is a falsifier when he reports the opposite of what is true.

In this present paper, we restrict our study to valid agents, complete agents, misinformers and falsifiers.

[2] In [3] the term "invalid" was used.

For giving a formal definition to these properties, we consider a propositional modal logic with two families of modal operators: B_i and R_i (where i are agents). $B_i p$ means that the agent i believes the proposition p to be true. $R_i p$ means that the agent i reports that the proposition p is true. Formulas are defined as usual. B_i obeys KD axiomatics as before and R_i obeys the rule of substitutivity like I_i did before. Thus the axioms of the logic are:

- Axioms schemata of classical propositional logic
- (D_B) $\quad B_i \neg p \to \neg B_i p$
- (K_B) $\quad B_i p \wedge B_i(p \to q) \to B_i q$

And inference rules are:

$$(MP)\ \frac{p \quad p \to q}{q} \qquad (Nec_{B_i})\ \frac{p}{B_i p} \qquad (RE_{R_i})\ \frac{p \leftrightarrow q}{R_i p \leftrightarrow R_i q}$$

Notice that operator R_i is unary and does not allow one to take the addresee into account like I_i^j previously mentionned could. This is adequate to model public utterances where the speaker (writer) utters a piece of information without specifying any addressee. Notice also that there is no relation between belief and reporting. Obviously, we do not require that an agent believes what it reports. Finally, since the operators we consider obey KD axiomatics, this logic can be provided with a Kripke semantics in which there are as many serial accessibility relations as operators so that the axiomatics is sound and complete towards this semantics.

Like in [4], this logic can be given a semantics according to which models are made of: a set of worlds, a valuation function, as many serial accessibility relations as modalities B_i and as many neigborhood functions as R_i. Like in [4] the logic can be proved to be sound and complete by using a method based on canonical minimal models [6].

The four properties considered in this paper are given by the following definition in which φ is a formula of the language.

Definition 1.

$$\begin{gathered} valid(i,\varphi) \equiv R_i\varphi \to \varphi \\ misinformer(i,\varphi) \equiv R_i\varphi \to \neg\varphi \\ complete(i,\varphi) \equiv \varphi \to R_i\varphi \\ falsifier(i,\varphi) \equiv \varphi \to R_i\neg\varphi \end{gathered}$$

Thus, an agent is *valid* for φ iff, if it reports φ then necessarily φ is true; it is a *misinformer* for φ iff, if it reports φ then necessarily φ is false; it is *complete* for φ iff, if φ is true then it reports it or equivalently, if it does not report φ then necessarily φ is false; Finally, an agent is a *falsifier* for φ iff, if φ is true then it reports its contrary.

Notice that these definitions do not use the concept of intention [20]. In particular they do not require that the information source intends to provide false information nor intend to let the receiver believe false information. Thus, strict comparisions with concepts introduced in dishonest communications ([23, 24, 21, 22]) is not possible.

Theorem 1. *Let i be an agent and φ a formula of the language. Any property of i defined by $X \to Y$ or by $Y \to X$ so that X belongs to $\{R_i\varphi, R_i\neg\varphi\}$ and Y belongs to $\{\varphi, \neg\varphi\}$ can be expressed by one of these four properties.*

This theorem, proved by listing exhaustively all the different possible cases of X and of Y, shows the completeness of the proposed model: there is no other property of the required form that the information sources could have and which cannot be expressed by these four ones.

Besides, it must be noticed that some combinations of properties are impossible, by definitions. For instance,

- $\neg(R_i\varphi \wedge valid(i,\varphi) \wedge misinformer(i,\varphi))$ is a theorem. Consequently an agent who reports a piece of information cannot be valid and misinformer for it.
- $\varphi \to \neg(complete(i,\varphi) \wedge falsifier(i,\varphi))$ is a theorem. Consequently an agent cannot be complete and falsifier for a piece of information which is true.

3 First application of the model: can a reported information be believed?

We consider an agent i, different information sources j, k, and an item of information(i.e., a formula) φ. In this section, we show that i can use its beliefs about information sources properties in order to accept φ as a new belief. For doing this, we first study the case when φ is directly reported by an information source. Then, we study the case of second-hand information.

In this section, the question which is addressed is:

$$(Q_1)\text{: can } i \text{ believe } \varphi \text{ ?}$$

We answer this question by giving some theorems, the proofs of which do not pose any difficulty.

3.1 First hand information

Here we consider that agent i is in direct contact with the information source named j which is supposed to provide φ. There are two cases.

- First case. j indeed reports φ and i is aware of it i.e we have $B_iR_j\varphi$. The following proposition answers the question (Q_1).

  Theorem 2. *The following formulas are theorems and their premisses are exclusive.*

$$B_iR_j\varphi \wedge B_ivalid(j,\varphi) \to B_i\varphi$$
$$B_iR_j\varphi \wedge B_imisinformer(j,\varphi) \to B_i\neg\varphi$$

Thus if agent i believes that j reported φ and if it believes that j is valid (resp, misinformer) for information φ then it can conclude that φ is true (resp, false). Furthermore, i cannot believe j both valid and misinformer. Thus it cannot infer both φ and $\neg\varphi$.

– Second case. j did not report φ and i is aware of it, i.e we have $B_i \neg R_j \varphi$. The following proposition answers the question (Q_1).

Theorem 3. *The following formulas are theorems and their premisses are exclusive.*

$$B_i \neg R_j \varphi \wedge B_i complete(j, \varphi) \rightarrow B_i \neg\varphi$$
$$B_i \neg R_j \varphi \wedge B_i falsifier(j, \neg\varphi) \rightarrow B_i \varphi$$

Thus, if i believes that j has not reported φ while it believes that j is complete for φ then it can conclude that φ is false. As an example, assume that I know that my friend who lives in the mountains is complete and always informs me when the temperature changes. If he does not report any temperature rise then I can conclude that the temperature did not rise.
Furthemore, if agent i believes that agent j has not reported φ while it believes that j is a falsifier for $\neg\varphi$ then it can conclude that φ is true. As an example, assume now that I am informed by a person known to falsify information related to temperature. If this person does not tell me that the temperature did not rise, this is because the temperature has not risen. Indeed, if the temperature had risen, this falsifier would have told me it had not. Thus I can conclude that indeed, the temperature did not rise.

3.2 Second-hand information

Here we consider that agent i is not in direct contact with the agent which is supposed to provide information, named k, but there is a go-between agent, named j. There are four cases.

– First Case. j reports that k reports φ and i knows it, i.e $B_i R_j R_k \varphi$. The following proposition answers the question (Q_1).

Theorem 4. *The following formulas are theorems and the premisses are exclusive.*

$$B_i R_j R_k \varphi \wedge B_i valid(j, R_k \varphi) \wedge B_i valid(k, \varphi) \rightarrow B_i \varphi$$
$$B_i R_j R_k \varphi \wedge B_i valid(j, R_k \varphi) \wedge B_i misinformer(k, \varphi) \rightarrow B_i \neg\varphi$$
$$B_i R_j R_k \varphi \wedge B_i misinformer(j, R_k \varphi) \wedge B_i complete(k, \varphi) \rightarrow B_i \neg\varphi$$
$$B_i R_j R_k \varphi \wedge B_i misinformer(j, R_k \varphi) \wedge B_i falsifier(k, \neg\varphi) \rightarrow B_i \varphi$$

For instance, my neighbour told me that one of his friends who lives in the moutains told him that the temperature has risen. Suppose I know that when my neighbour says such a sentence, it is false i.e, I can infer that his friend did not tell him that the temperature has risen. But suppose that I know that his friend always inform him when the temperature rises. I can conclude that the temperature had not risen.

– Second Case. j does not report that k reported φ and i knows it, i.e $B_i \neg R_j R_k \varphi$. The following proposition answers the question *(Q_1)*.

Theorem 5. *The following formulas are theorems and their premisses are exclusive.*

$$B_i \neg R_j R_k \varphi \wedge B_i complete(j, R_k \varphi) \wedge B_i complete(k, \varphi) \rightarrow B_i \neg \varphi$$
$$B_i \neg R_j R_k \varphi \wedge B_i complete(j, R_k \varphi) \wedge B_i falsifier(k, \neg \varphi) \rightarrow B_i \varphi$$
$$B_i \neg R_j R_k \varphi \wedge B_i falsifier(j, \neg R_k \varphi) \wedge B_i valid(k, \varphi) \rightarrow B_i \varphi$$
$$B_i \neg R_j R_k \varphi \wedge B_i falsifier(j, \neg R_k \varphi) \wedge B_i misinformer(k, \varphi) \rightarrow B_i \neg \varphi$$

– Third Case. j reports that k did not report φ and i knows it, i.e $B_i R_j \neg R_k \varphi$. The following proposition answers the question (Q_1).

Theorem 6. *The following formulas are theorems and their premisses are exclusive.*

$$B_i R_j \neg R_k \varphi \wedge B_i valid(j, \neg R_k \varphi) \wedge B_i complete(k, \varphi) \rightarrow B_i \neg \varphi$$
$$B_i R_j \neg R_k \varphi \wedge B_i misinformer(j, \neg R_k \varphi) \wedge B_i valid(k, \varphi) \rightarrow B_i \varphi$$
$$B_i R_j \neg R_k \varphi \wedge B_i misinformer(j, \neg R_k \varphi) \wedge B_i misinformer(k, \varphi) \rightarrow B_i \neg \varphi$$
$$B_i R_j \neg R_k \varphi \wedge B_i valid(j, \neg R_k \varphi) \wedge B_i falsifier(k, \neg \varphi) \rightarrow B_i \varphi$$

– Fourth Case. j did not report that k did not report φ and i knows it, i.e $B_i \neg R_j \neg R_k \varphi$. The following proposition answers the question (Q_1).

Theorem 7. *The following formulas are theorems and their premisses are exclusive.*

$$B_i \neg R_j \neg R_k \varphi \wedge B_i complete(j, \neg R_k \varphi) \wedge B_i valid(k, \varphi) \rightarrow B_i \varphi$$
$$B_i \neg R_j \neg R_k \varphi \wedge B_i complete(j, \neg R_k \varphi) \wedge B_i misinformer(k, \varphi) \rightarrow B_i \neg \varphi$$
$$B_i \neg R_j \neg R_k \varphi \wedge B_i falsifier(j, R_k \varphi) \wedge B_i complete(k, \varphi) \rightarrow B_i \neg \varphi$$
$$B_i \neg R_j \neg R_k \varphi \wedge B_i falsifier(j, R_k \varphi) \wedge B_i falsifier(k, \neg \varphi) \rightarrow B_i \varphi$$

4 Second application of the model: generation of hypothesis about the source properties

4.1 Motivations and objectives

In this section, we apply the model proposed in section 3 to generate the hypothesis an agent can make about the sources properties so that this agent can believe information they provide. A motivating example is the following.

Example 1. Consider for instance that agent a is aware that agent b reports that agent c has reported that a given atomic proposition A, i.e., $B_a R_b R_c A$. Consider also that agent a is aware that agent d reported $\neg A$, i.e., $B_a R_d \neg A$. Finally, suppose that a considers b valid for what it reports, i.e, $B_a valid(b, R_c A)$. a's beliefs about the sources are not sufficient to believe A. Consequently, an interesting question is : *What can agent a assume about the different information sources so that it believes A ?*

In this section, we propose a method for answering this question. It is illustrated on example 1.

Example 1. (continued) Let us denote BB_a the explicit beliefs of a about the information sources i.e., here $BB_a = \{B_a R_b R_c A,\ B_a R_d \neg A,\ B_a valid(b, R_c A)\ \}$. Let us denote $\models$ the logical consequence relation of the previous logic.

Finding the hypothesis agent a can make about the different information sources so that it believes A comes to answering the following question:

$$(Q_2) : \text{what are the formulas } H \text{ such that: } BB_a \cup \{H\} \models B_a A \ ?$$

Obviously, several constraints apply on formulas H.

- First of all, H must be composed of atomic formulas of the form $B_a valid(.,.)$, $B_a misinformer(.,.)$, $B_a complete(.,.)$ or $B_a falsifier(.,.)$. Indeed, we want to find what hypothesis a can make on the sources properties.
- Secondly, $BB_a \cup \{H\}$ must be consistent. Indeed, any H which is inconsistent with BB_a satisfies the previous equation but does not characterize interesting answers.
- Finally, H must be minimal. I.e., if there is another H' which statisfies the two previous constraints so that $\models H \rightarrow H'$ then $\models H' \rightarrow H$.

The problem is thus to find some missing premisses for deducing a given conclusion. This is abductive reasoning.

4.2 Abductive reasoning and SOL-resolution: summary

Abductive reasoning has initially been introduced by Peirce [2] for finding plausible explanations to observations. The study of this problem in Artificial Intelligence has been introduced by Morgan [17] and Pople [19] and a huge number of works have been focused on the complexity of this reasoning or its automation. More formally, abductive reasoning is a particular type of reasoning which allows one to deduce plausible information, H, from some given information E so that $E \cup H$ allows to deduce some given facts, F. I.e, H satisfies: $E \cup H \models F$

Obviously, question (Q_2) is an instance of this problem in which, E is BB_a, F is $B_a A$. Thus answering question (Q_2) can be done by using an abductive reasoning engine.

According to us, one of the most interesting inference rule for abduction is the SOL Resolution defined by K. Inoue in [12]. It is defined in the context of first-order logic. It is used for generating the clauses[3] which are logical consequences of a given set of clauses. Indeed, the problem $E \cup H \models F$ is equivalent to $E \cup \neg F \models \neg H$, thus the abductive inference of H can be realized by the deductive generation of $\neg H$ from $E \cup \neg F$. More specifically, the SOL-resolution is a complete inference rule for generating first-order clauses minimal for subsumption and which belong to a given language and which satisfy a given condition.

[3] A clause is a disjonction of literals e.g., $P(x) \vee \neg Q(x)$. The empty clause is denoted $\square$

For this, Inoue introduces the notion of production field. Inoue's main definitions are recalled in the following.

Definition 2.
A production field P is defined by a language $\mathrm{L_P}$ *and a condition* $\mathrm{Cond_P}$*;*
A clause belongs to P iff it belongs to language $\mathrm{L_P}$ *and it satisfies* $\mathrm{Cond_P}$*;*
A production field P is stable if any clause which subsumes[4] *a clause belonging to P belongs to P too;*
A structured clause is a pair <P,Q> where P and Q are clauses.

The main result proved in [12] is the following:

Theorem 8. (Soundness and completeness of SOL). *Let Σ be a set of clauses and C a clause. Let $P = \{\mathrm{L_P}, \mathrm{Cond_P}\}$ be a stable production field.*

- *(1) Let $< \mathrm{S}, \Box >$ be a structured clause generated by a SOL deduction from Σ, with $< \Box, \mathrm{C} >$ as top clause, and with production field P. Then, S is a consequence of $\Sigma \cup \{C\}$ and it belongs to P*
- *(2) Let T be a clause which is a consequence of the set $\Sigma \cup \{C\}$, but not a consequence of Σ only, and which belongs to P. Then, there is a clause S which subsumes T such that $< \mathrm{S}, \Box >$ is generated by a SOL deduction, from Σ with $< \Box, \mathrm{C} >$ as top clause, with production field P.*

Our aim is thus to use SOL-resolution for solving question (Q_2) defined before. For doing so, we need to translate the problem we face, which is expressed in modal logic, into a first order logic problem. This is shown in the following.

4.3 Reformulation of the problem in first order logic

We consider a first-order language with: some constant symbols (called agent-constants) in bijection with the agents. Their set is denoted Ag; some constant symbols (called proposition-constants) in bijection with the propositional letters. Their set is denoted $Prop$; a binary predicate symbol: B; five binary function symbols R, $valid$, $misinformer$, $complete$, $falsifier$; a unary function symbol not.

In the following, we define a function t which translates some particular formulas of the modal logic defined in section 2 into formulas of this first-order language.

Definition 3.
$t(B_i F) = B(i, t'(F))$
$t'(P) = P$ *if P is a propositional letter*
$t'(\neg F) = not(t'(F))$
$t'(R_j F) = R(j, t'(F))$
$t'(valid(i, F) = valid(i, t'(F))$

[4] A clause C subsumes a clause D iff there exists a substitution of variables σ such that $C\sigma \subseteq D$

$t'(misinformer(i, F) = misinformer(i, t'(F))$
$t'(complete(i, F) = complete(i, t'(F))$
$t'(falsifier(i, F) = falsifier(i, t'(F))$

Definition 4. *If* $S = \{F_1, ..., F_n\}$ *a finite set.* $t(S)$ *denotes* $\{t(F_1), ..., t(F_n)\}$.

Example 1. (Continued) According to this definition, $t(BB_a)$ is:
$t(BB_a) = \{B(a, R(b, R(c, A))), B(a, R(d, not(A))), B(a, valid(b, R(c, A)))\}$

Let us consider now the following set of first order formulas:

Definition 5.

$$\begin{aligned} Def_a = \{&\forall x \forall y \quad B(a, valid(x, y)) \rightarrow (B(a, R(x, y)) \rightarrow B(a, y)), \\ &\forall x \forall y \quad B(a, misinformer(x, y) \rightarrow (B(a, R(x, y)) \rightarrow B(a, not(y))), \\ &\forall x \forall y \quad B(a, complete(x, y) \rightarrow (B(a, not(R(x, y))) \rightarrow B(a, not(y))), \\ &\forall x \forall y \quad B(a, falsifier(x, y) \rightarrow (B(a, not(R(x, not(y)))) \rightarrow B(a, not(y)))\} \end{aligned}$$

Let us now consider the following question [5]:

(q_2): what are the formulas h such that: $Def_a \cup t(BB_a) \cup h \models B(a, A)$ and

- h is composed of atomic formulas of the form $B(a, valid(., .))$, $B(a, misinformer(., .))$, $B(a, complete(., .))$ or $B(a, falsifier(., .))$.
- $Def_a \cup t(BB_a) \cup h$ is consistent
- h is minimal. I.e., if there is another h' which statisfies the two previous constraints so that $\models h \rightarrow h'$ then $\models h' \rightarrow h$.

Even if we have the intuition that answering query (Q_2) is equivalent to answering query (q_2), we did not formally prove it yet. This is expressed by the following conjecture we consider as true from now.

Conjecture 1. H answers query (Q_2) if and only if $t(H)$ answers query (q_2)

If S is a set of first-order formulas, let us denote $Cl(S)$ the set of clauses which correspond to the formulas of S.

Definition 6. *Consider the following production field:*

$$P = \{L_P \ , \ \{c : \neg c \wedge Cl(Def_a \wedge t(BB_a)) \text{ is consistent}\}\}$$

Where L_P *is the set of clauses only written with literals* $B(a, valid(., .))$, $\neg B(a, valid(., .))$, $B(a, misinformer(., .))$, $\neg B(a, misinformer(., .))$, $B(a, complete(., .))$, $\neg B(a, complete(., .))$, $B(a, falsifier(., .))$ *or* $\neg B(a, falsifier(., .))$.

Theorem 9. P *is stable.*

Consequently, we can apply the SOL-resolution to answer question (q_2) by considering $\Sigma = Cl(Def_a \cup t(BB_a))$ and by considering the top clause $< \Box, \neg B(a, A) >$.

[5] Notice that here, formulas h are first order and $\models$ is the first order logical consequence.

4.4 Implementation and restrictions

The previous method has been implemented and restrictions have been made. More specifically, it must be noticed that $Cl(Def_a)$ is recursive and there is an infinite numbers of answers which are potentially generated by the SOL-resolution.

Thus, we have restricted the problem to a finite and manageable case. Instead of considering $Cl(Def_a)$, we have considered $ground(Cl(Def_a), i)$, whose clauses are ground and obtained by replacing variables in clauses of $Cl(Def_a)$ by constants. More precisely, the x variable in Def_a is replaced by constants in Ag. The y-variable is replaced by constants in $Prop_i$, defined as follows.

- $Prop_0 = Prop$
- $Prop_i = Prop_{i-1} \cup \{not(u)$ where $u \in Prop_{i-1}\} \cup \{f(t, u)$ where $t \in Ag$ and $u \in Prop_{i-1}$ and f being R, $valid$, $misinformer$, $complete$, or $falsifier\}$

In this case, the number of answers which are generated is finite.

The method implements a width-first development of the tree which generates all the SOL-resolutions from the top clause $< \Box, \neg B(a, A) >$ and with $\Sigma = ground(Cl(Def_a), i) \cup Cl(t(BB_a)))$, i being arbitrarly fixed.

Example 1. (Continued) With $ground(Cl(Def_a), 2)$, this method generates the two following formulas: $B(a, valid(c, A))$ and $B(a, misinformer(d, not(A)))$. This means that if a assumes that c is valid for A, or if it assumes that d is misinformer for $\neg A$, then it can believe A. These two answers can help a to investigate further about the sources c or d.

5 Conclusion

This paper showed that an agent's beliefs about the piece of information it receives are influenced by its beliefs about the information sources properties.

First we assumed that information sources could deliver false information. Secondly, not only we dealt with information provided directly by an information source but we also dealt with second-hand information i.e, information reported by a source which cites another one. Even if we did not insist on it, the model defined here does not restrict to the case of two levels of imbrication of sources but also applies to "n-handed" information i.e information reported by a source which cites another one which itself cites another source etc.

Notice that the case of second-hand information has received little attention in the litterature. As far as we know, the first work which addressed this question is [13]. The only other work we know which mentions this kind of information is [11] and concerns trust management. The reason why second-hand information has received little attention in the litterature is maybe because this question can be confused with the question of estimating how an agent can believe a piece of information when it is passed from an agent to another one. In this case, each agent can pass it correctly or can deform it. The example given in the introduction showed that these two questions must not be confused. Indeed, when my neighbour tells me that according to his friend, the temperature has risen, the information my neighbour reports *is not* that the temperature has risen.

This present work can be extended in many different ways.

First, the formal proof establishing the validity of the reformulation of the problem in first order logic must be done. More precisely, we must find the conditions (if any) under which the two questions (Q_2) and (q_2) are equivalent. For doing this, we think to prove that a set S of modal formulas is inconsistent in our modal model, if and only if its reformulation $t(S)$ is inconsistent in first-order logic. For doing so, we have the intuition that a result proved in [5] will be useful. This result shows that proving the inconsistency of a set of first order modal formulas can be reduced to proving the inconsistency of a set of classical propositional formulas.

Secondly, as for the model, it would be more general to define information sources properties according to sets of propositions and not to a given proposition. For instance, we could express that an agent is valid for any proposition related to the topic "weather forecast" and complete for any proposition related to the topic "mountains". That would imply that this agent is valid and complete for any information related to weather forecat in the mountains. However, such a reasoning has to be formally defined and describing these topics by means of an ontology is a possible solution.

More, in the present model, the reporting action operator R is unary and the receiver agent is omitted. Considering a binary operator would allow us to propose a finer model in which, for instance, we could express that an agent is a misinformer for a proposition (or some propositions) when it reports it to some particular agent.

Richer would be a model which could help the user to decide if a source satisfies such or such property. For instance, *self-interested* information sources [15] are misinformers. The illustrative example borrowed from [15] is a self-interested vehicule-agent who wants to make the other vehicule-agents to believe that its own path is heavy so that they let it free. Determining if an agent is self-interested needs to model the relation between the agent, the proposition it utters and the benefit it gets when uttering it.

Finally, allowing to express some kind of uncertainty and reasoning with it would be a generalisation of this work. First, the agent receiving pieces of information may have graded beliefs and not strict beliefs in the properties of the sources. For instance, I can believe that my neighbour is highly valid as regard to what it tells me and poorly complete. Or the agent receiving information may only have preferences [16] on the sources and expresses relative validity, relative completeness... It is the case for instance when I believe that my neighbour is more valid than its friend. Secondly, an agent who provides information may include in its utterance its own uncertainty. For instance, my neighbour can tell me that he is almost certain that his friend has mentioned a temperature rise. Expressing these kinds of uncertainty and taking them into account is an interesting extension of the present work. For this, two directions are foreseen. The first one is to check if the graded modal logic introduced in [10] can be applied. The other approach is to use Theory of Evidence [25], an approach we initiated in [7]. Notice however that adapting the abductive reasoning in this case remains an open question.

Acknowledgements. I am deeply grateful to the anonymous referees whose pertinent questions led me to improve this paper.

References

1. C. E. Alchourron, P. Gardenfors, and D. Makinson. *On the logic of theory change: Partial meet contraction and revision functions.* Journal of Symbolic Logic, vol 50, Issue 2, 510-530, 1985.
2. J. Buchler, *Philosophical writings of Peirce.* Dover, New York 1955.
3. J. Chauvet-Paz and L. Cholvy. *Reasoning with Information reported by Imperfect Agents*. Proceedings of First International Workshop on Uncertainty Reasoning and Multi-Agent Systems for Sensor Networks (URMASSN'11), Belfast, Northern Ireland, 2011.
4. L. Cholvy and R. Demolombe and A. Jones *Reasoning about the safety of information: from logical formalization to operational definition* Proceedings of 8th International Symposium on Methodologies for Intelligent Systems (ISMIS '94), Charlotte, USA, 1994.
5. M. Cialdea. *Resolution for some first-order modal systems.* Theoretical Computer Science vol 85, 213-229, 1991.
6. B. F Chellas. *Modal logic: An introduction.* Cambridge University Press, 1980.
7. L. Cholvy. *Collecting information reported by imperfect information sources.* in: Proceedings of the 14^{th} International Conference on Information Processing and Uncertainty Management in Knowledge-Based Systems (IPMU'12), 501-510, 2012.
8. C. Chang an R. Lee *Symbolic Logic and Mechanical Theorem Proving* Academic Press, inc, 1973.
9. R. Demolombe, *Reasonig about trust: a formal logical work.* in: Proceedings of the 2^{nd} second International Conference iTrust, 291-303, 2004.
10. R. Demolombe *Graded trust.* Workshop on Trust in Agent Societies, 2009.
11. R. Demolombe, *Transitivity and propagation of Trust in Information Sources. An analysis in Modal Logic.* in: Proceedings of the 12^{th} International Workshop on Computational Logic in Multi-Agent Systems (CLIMA'11), 13-28, 2011.
12. K. Inoue, *Linear Resolution for Consequence Finding.* Artificial Intelligence 56(2-3), 301-353, 1992.
13. J. Jacquelinet, *Pertinence et cotation du renseignement.* Master thesis, ENSIIE, Paris, 2007.
14. H. Katsuno and A. Mendelzon. *On the Difference between Updating a Knowledge Base and Revising it.* Proceedings of Principles of Knowledge Representation and Reasoning(KR'91), 1991.
15. R. Lin, S. Kraus, Y. Shavitt, *On the Benefits of Cheating by Self-Interested Agents in Vehicular Networks.* in: Proceedings of International Conference on Autonomous Agents and Multiagent Systems, 183-190, 2007

16. E. Lorini and L. Perrussel and J.M. Thévenin. *A Modal Framework for Relating Belief and Signed Information.* in: Proceedings of 13^{th} International Workshop on Computational Logic In Multi-Agent Systems (CLIMA'11), LNAI 6814,58-73, Springer-Verlag, 2011.
17. C. Morgan, *Hypothesis generation by machine.* Artificial Intelligence 2, 179-187, 1971.
18. North Atlantic Treaty Organizations. *Standardization Agreement, Intelligence Report STANAG n° 2511*, 2003.
19. H. Pople, *On the mechanization of abductive logic.* in: Proceedings of International Joint Conference on Artificial Intelligence, 147-152, 1973.
20. A. Rao and M. Georgeff. *Modeling rational agents within a bdi-architecture.* in: Proceedings of Principles of Knowledge Representation and Reasoning(KR'91), 473-484, 1991.
21. Chiaki Sakama *Logical Definitions of Lying.* Workshop on Trust in Agent Societies, 2011.
22. C. Sakama, *A formal model of dishonest communication.* Workshop on Formal Models of Communication, 2012.
23. C. Sakama, M. Caminada, A. Herzig, A logical account of lying, in: Proceedings of 12^{th} European Conference on Logics in Artificial Intelligence (JELIA'10), 286-299, 2010.
24. C. Sakama, M. Caminada, *The many faces of deception.* NonMon@30: Thirty Years of Nonmonotonic Reasoning, 2010.
25. G. Shafer, *A mathematical Theory of Evidence.* Princeton University Press, 1976.
26. M. Winslett. *Updating Logical Databases.* Cambridge University Press, 1990.