Formalizing the reliability of agent's information

Robert Demolombe
CERT/ONERA
France
Robert.Demolombe@cert.fr

We are never guaranteed that information stored in data or knowledge bases is a correct representation of the world. Nevertheless, there are many situations where we have strong supports for the validity and/or completeness of parts of the information.

For instance, we may know that salaries of people in a company are valid if they are inserted by people from accounting department, and that data about absenteeism are complete if they are inserted by people from the department of human ressources. In that example validity and completeness depend on reliability of information sources, in other examples they may only depend on the type of data. For instance, we may know that data about health of people are valid.

In these contexts, if one asks a complex query it is not easy to infer from the meta information about validity and completeness which parts of the answer to the query are valid, and which parts are complete. This paper presents some results of an ongoing work [1] which is intended to formalise this kind of meta information and to formalise associated queries about validity and completeness of data. We start with a formalisation in modal logic of situations where a distinction is made between, on one hand, data that have been explicitly inserted by information sources who are supposed to be, or not to be, reliable and, on the other hand, data infered from the previous ones. Then, formalisation is adapted to situations where reliability only depends on some types of information. Finally, the formalism is extended to situations where we can define several levels of reliability.

[1]This work was initiated in collaboration with Andrew J. Jones (see [5]). Most of the ideas presented in section 1 are the result of this work, even if they are presented in a slightly different form.

1 Reliable information sources

In this section we consider situations where some information sources are supposed to be reliable in regard to validity or to completeness of some data.

A distinction is made between data base content, which is viewed as a set of beliefs and a correct description of the world. Sentences of the kind $B_{db}p$ can be read "the data base believes p", and sentence p can be interpreted as "p is true of the world", where B_{db} is a normal modal operator that obeys axiom schemas (K) and (D), and inference rule (Nec) [2]:

$$\text{(K)} \quad B_{db}(p \to q) \to (B_{db}p \to B_{db}q)$$

$$\text{(D)} \quad B_{db}p \to \neg B_{db}\neg p$$

$$\text{(Nec)} \quad \frac{\vdash p}{\vdash B_{db}p}$$

An information source "i" is reliable in regard to validity of sentence p iff the fact that i has inserted p in the data base implies that p is true of the world. The fact that i has inserted p is expressed in the form: "agent i has brought about that the data base believes p", which is formalised, using the action operator E_i, by the sentence: $E_i(B_{db}p)$. The operator E_i is a classical (not normal) modal operator that obeys axiom schemas ($\neg$N) and inference rule (RE):

$$(\neg\text{N}) \quad \neg E_i(\text{true})$$

$$\text{(RE)} \quad \frac{\vdash p \leftrightarrow q}{\vdash E_i p \leftrightarrow E_i q}$$

Since no information about the world can be guaranteed to be true, in an absolute sense, the fact that an information source is reliable, itself, cannot be guaranteed to be true. However, it is assumed that this fact of being a reliable information source has a different status than other database beliefs, in the sense that the data base believes that it is a true belief. For this reason another normal modal operator K_{db} is introduced, and sentences of the kind $K_{db}p$ can be read "the data base "knows" p". For operator K we accept inference rule (Nec), and axiom schemas (K), (D) and the additional following schemas (T') and (KB):

$$(T') \quad K_{db}(K_{db}p \rightarrow p)$$

$$(KB) \quad K_{db}p \rightarrow B_{db}p$$

Notice that $K_{db}p$ does not imply p, that is, we do not have the axiom schema (T): $K_{db}p \rightarrow p$. Now, reliability of information source i in regard to sentence p is defined in that way:

$$RV_i(p) \stackrel{def}{=} K_{db}(E_i(B_{db}p) \rightarrow p)$$

This definition is extended to reliability for all the sentences of the form p(x):

$$RV_i(p(x)) \stackrel{def}{=} K_{db}(\forall x(E_i(B_{db}p(x)) \rightarrow p(x)))$$

In a similar way an information source "i"is defined as a reliable information source in regard to completeness of sentence p iff the fact that p is true of the world implies that "i" has brought about that the data base believes p. The formal definition is:

$$RC_i(p(x)) \stackrel{def}{=} K_{db}(\forall x(p(x) \rightarrow E_i(B_{db}p(x)))$$

It is assumed that the data base "knows" whether an information source has performed, or not, the insertion of some sentence p. This assumption is formalised by the axiom schemas (OBS1) and (OBS2):

$$(OBS1) \quad E_i(B_{db}p) \rightarrow K_{db}(E_i(B_{db}p))$$

$$(OBS2) \quad \neg E_i(B_{db}p) \rightarrow K_{db}(\neg E_i(B_{db}p))$$

Moreover, it is assumed that if the data base knows that some information source has brought about that the data base believes p, then the data base does accept belief p. This is formalised by the axiom schema (B) [2] :

$$\text{(B)} \quad K_{db}(E_i(B_{db}p)) \rightarrow B_{db}p$$

In this approach data base content db is represented by a set [3] of sentences of the form: $E_i(B_{db}p)$ and $\neg E_i(B_{db}p)$. As a consequence of that representation it is possible to distinguish a situation where agent i has inserted p and he has inserted r, formally represented by: $db_1 = \{E_i(B_{db}p), E_i(B_{db}r)\}$, and another situation where agent i has inserted inserted $p \wedge r$, which is represented by $db_2 = \{E_i(B_{db}(p \wedge r))\}$. The reason for such a distinction is that each insertion is supposed to have its own independent justification. From an intuitive point of view, explicit content of the data base is considered as the result of performed insertions. Of course, if in the history of the data base, p has been inserted, which is formally represented by $E_i(B_{db}p)$, and then it has been deleted, which is formally represented by $E_i(\neg B_{db}p)$, then neither $E_i(B_{db}p)$ nor $E_i(\neg B_{db}p)$ should be in db.

Meta information mdb about reliability of information sources is represented by a set of sentences of the form: $RV_i(p)$ and $RC_i(p)$. For practical reasons it is impossible to represent in db all the insertions that have *not* been performed. A solution to this problem is to accept the following inference rule (COMP) whose meaning is: if from db it is not possible to infer that i has inserted p, then it is the case that i has not inserted p. That means that db is supposed to be a complete description of all the performed insertions.

$$\text{(COMP)} \quad \frac{\nvdash db \rightarrow E_i(B_{db}p)}{\vdash db \rightarrow \neg E_i(B_{db}p)}$$

For a given query q(x) the formal definition of the **standard answer** is the set:

$$\{ a \ : \ \vdash db \rightarrow B_{db}q(a)\}$$

[2]Notice that from (OBS1) and (B) we have: $E_i(B_{db}p) \rightarrow B_{db}p$, which is a restricted form of (T).

[3]In some contexts, db is supposed to represent the conjunction of all the formulas in this set.

The formal definition of a **valid anwer** is the set:

$$\{ a \;:\; \vdash db \wedge mdb \rightarrow K_{db}q(a)\}$$

Information about completeness of an information source in regard to p is used to infer information about the validity of $\neg p$. For instance if we have: $db = \{E_i(B_{db}p)\}$, $mdb = \{RV_i(p), RC_j(r)\}$, and the query is $q = p \wedge \neg r$, from (COMP) we have: $\vdash db \rightarrow E_i(B_{db}p) \wedge \neg E_j(B_{db}r)$, and from the definitions of RV and RC we have: $\vdash mdb \rightarrow K_{db}(E_i(B_{db}p) \rightarrow p) \wedge K_{db}(\neg E_j(B_{db}r) \rightarrow \neg r)$, then, from (OBS1) and (OBS2), we have: $\vdash db \wedge mdb \rightarrow K_{db}(p \wedge \neg r)$.

This can be illustrated by a situation where p and r respectively mean "John is enrolled in course C1" and "John has successfully passed examination of course C1", and agent i is reliable for validity of p while agent j is reliable for completeness of r. Then, the database "knows" that John has not successfully passed his examination. It is worth noting that the reason why the database can draw this conclusion is that j is reliable for the completeness of r, though j is not necessarily reliable for the validity of r. Intuitively, if John would have succeeded, agent i would not have "forgotten" that he succeeded.

2 Reliable parts of the information

In this section we suppose that reliability is defined for some parts of the information, and we do not distinguish inserted data and derived data. For instance, if it assumed that data base content is reliable in regard to validity of $p \wedge r$, then, if the data base believes $p \wedge r$, we can infer that data base "knows" $p \wedge r$, whatever this data base belief is a consequence of an insertion of $p \wedge r$, or the consequence of an insertion of p and of an insertion of r, and whatever are the information sources who have performed the insertions.

The technical consequence of this new approach [4] is that reliability of parts of the data base is defined in that way:

$$RV'(p(x)) \stackrel{\text{def}}{=} K_{db}(\forall x(B_{db}p(x) \rightarrow p(x)))$$

[4]This approach corresponds to the approach adopted by Ami Motro in [7, 8].

$$RC'(p(x)) \stackrel{def}{=} K_{db}(\forall x(p(x) \rightarrow B_{db}p(x)))$$

and axiom schemas (OBS1) and (OBS2) are replaced by:

$$(OBS1') \quad B_{db}p \rightarrow K_{db}(B_{db}p)$$

$$(OBS2') \quad \neg B_{db}p \rightarrow K_{db}(\neg B_{db}p)$$

Comment: for (OBS1') we have adopted a stronger axiom schema than: $B_{db}p \rightarrow B_{db}(B_{db}p)$, because we assume that the data base believes that he is correctly informed about what he believes; the same comments applies to (OBS2').

The data base content is represented here by a set of sentences db' of the form $B_{db}p$, and the meta data base is represented by a set of sentences mdb' of the form $RV'_i(p)$ and $RC'_i(p)$. For the same reason as the one mentioned in the previous section, we accept the inference rule:

$$(COMP') \quad \frac{\not\vdash db' \rightarrow B_{db}p}{\vdash db' \rightarrow \neg B_{db}p}$$

The definition of a standard answer to a query is defined like in the previous section. A **valid answer** to a query q(x) is a sentence $p_1(x)$ such that:

$$\vdash \forall x(p_1(x) \rightarrow q(x)) \quad \text{and} \quad \vdash mdb' \rightarrow RV'(p_1(x))$$

and such that $p_1(x)$ is maximal for consequence relation, in the sense, that if there is another sentence $p'_1(x)$ that satisfies the same properties, and which is a consequence of $p_1(x)$, then $p'_1(x)$ is logically equivalent to $p_1(x)$.

If $p_1(x)$ is a valid answer to q(x), then we have: $\vdash mdb' \rightarrow \forall x(B_{db}p_1(x) \rightarrow K_{db}q(x))$. This means that, if the data base believes $p_1(a)$ then the data base "knows" q(a), or, in other terms, the standard answer to the query $p_1(x)$ gives a subset of the standard answer to q(x) which is "guaranteed" to be true of the world (this is represented in figure 1 by the fact that the extension of $B_{db}p_1(x)$ is included in the extension of $K_{db}q(x)$).

A **complete answer** to a query q(x) is a sentence $p_2(x)$ such that:

$$\vdash \forall x(q(x) \rightarrow p_2(x)) \quad \text{and} \quad \vdash mdb' \rightarrow RC'(p_2(x))$$

and such that $p_2(x)$ is minimal for consequence relation, in the sense, that if there is another sentence $p_2'(x)$ that satisfies the same properties, and which implies $p_2(x)$, then $p_2'(x)$ is logically equivalent to $p_2(x)$.

If $p_2(x)$ is a complete answer to q(x), then we have: $\vdash mdb' \rightarrow \forall x(\neg B_{db}p_2(x) \rightarrow K\neg q(x))$. This means that, if the data base does not believe $p_2(a)$ then the data base "knows" that q(a) is false, or, in other terms, the standard answer to the query $p_2(x)$ gives a superset of the standard answer to q(x) such that all the element that are not in this set are "guaranted" not to satisfy q(x) (this is represented by the fact that the extension of $K_{db}q(x)$ is included in the extension of $B_{db}p_2(x)$).

An intuitive interpretation of valid answer $p_1(x)$ and complete answer $p_2(x)$, is that the answer to q(x), if it would be evaluated on a valid and complete description of the world, would be respectively bounded down and up by the answers to $p_1(x)$ and to $p_2(x)$ (see figure 1).

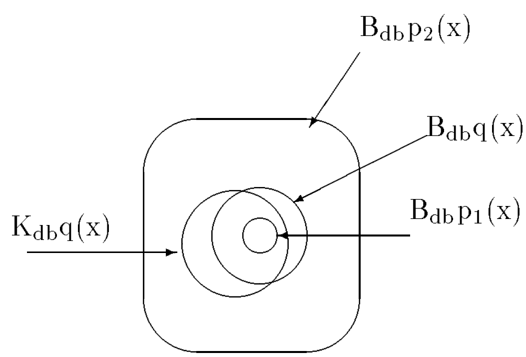

Figure 1: The standard answer q(x) is bounded up and down by $p_1(x)$ and $p_2(x)$.

3 Multiple reliability levels

In previous sections we have considered that there is only two agent reliability levels: the first level corresponds to agents who are supposed to be reliable, and the second level corresponds to the other ones. Here we extend this logical framework to several levels which can be structured by a partial order. To define the notion of relative reliability in this framework we gradually transform the former definitions.

For the first transformation, it is considered that the database does not play a specific role, and that there are several agents who may send or receive data. That leads to definitions of the form $RV_i^k(p) \stackrel{def}{=} K_k(E_i(B_k p) \rightarrow p)$ and $RC_i^k(p) \stackrel{def}{=} K_k(p \rightarrow E_i(B_k p))$.

Second transformation results from adoption of a simplified kind of communicative actions. Instead of actions of the form $E_i(B_k p)$, which means "agent i has brought about that agent k believes p", we consider actions of the kind "agent i has asserted p", formally represented by $A_i p$. The fact that agent k has been informed about p by agent i is formally represented by $K_k(A_i p)$, that is "agent k knows that agent i has asserted p". That gives two definitions of the form: $RV_i^k(p) \stackrel{def}{=} K_k(A_i p \rightarrow p)$ and $RC_i^k(p) \stackrel{def}{=} K_k(p \rightarrow A_i p)$.

Third transformation, the most important one, is based on the idea that an assumption of the form $RV_i^k(p)$ about agent i's reliability, can be viewed as an assumption that informs us not about the state of the world, represented by p, but about what an agent g (say "god"), who is supposed to be the most reliable information source, would have asserted about the world. This "virtual assertion" is formally represented by $A'_g p$. That means that sentence $A_i p \rightarrow p$, whose intuitive meaning is "if agent i asserts p then p is true" is replaced by $A_i p \rightarrow A'_g p$, whose intuitive meaning is "if agent i asserts p, then, in the same circumstances, agent g would have asserted p". That leads to the definitions $RV_i^k(p) \stackrel{def}{=} K_k(A_i p \rightarrow A'_g p)$ and $RC_i^k(p) \stackrel{def}{=} K_k(A'_g p \rightarrow A_i p)$. If it is assumed that $A'_g p \leftrightarrow p$, these definitions are logically equivalent to the previous ones.

Fourth transformation allows us to go from the notion of "absolute reliability" to "relative reliability". For this purpose we generalise the idea that agent i is as reliable as agent g, formally represented by $A_i p \rightarrow A'_g p$ and $A'_g p \rightarrow A_i p$, to the idea of reliability in reference to some other agent j, who is not supposed to be perfectly reliable. That leads to the definitions $RV_{i,j}^k(p) \stackrel{def}{=} K_k(A_i p \rightarrow A'_j p)$ and $RC_{i,j}^k(p) \stackrel{def}{=} K_k(A'_j p \rightarrow A_i p)$.

The last transformation comes from observation that reliability definitions are in conditional form, like $A_i p \rightarrow A'_j p$, and that agent reliability is not based on the assertions that agents have actually preformed, but on assertions they might perform, that is virtual assertions. Then, in the definitions, actions of the from $A_i p$ are replaced by actions of the form $A'_i p$. That leads to the final definitions:

$$RV^k_{i,j}(p) \stackrel{\text{def}}{=} K_k(A'_i p \rightarrow A'_j p)$$

$$RC^k_{i,j}(p) \stackrel{\text{def}}{=} K_k(A'_j p \rightarrow A'_i p)$$

For the action operators A_i and A'_i we adopt the same axiomatics as for action operator E_i, that is $(\neg N)$ and (RE). In addition to this axiomatics we have the following axiom schemas which represents the fact that all the consequences of virtual assertions should also be consequences of actual assertions:

$$(V) \quad A_i p \rightarrow A'_i p$$

We also have axiom schemas to represent the fact that agents know what has been asserted, or has not been asserted, by other agents:

$$(OBS1'') \quad A_i p \rightarrow K_k(A_i p)$$

$$(OBS2'') \quad \neg A_i p \rightarrow K_k(\neg A_i p)$$

A given situation is represented by a set s of sentences of the form $A_i p$, which represents what has been actually asserted by all the agents, and for practical reasons, like in previous sections, we accept the inference rule:

$$(COMP'') \quad \frac{\not\vdash s \rightarrow A_i p}{\vdash s \rightarrow \neg A_i p}$$

The meta information about agent reliability is represented by a set ms of sentences of the form $RV^k_{i,j}(p)$ or $RC^k_{i,j}(p)$. At the present time we have not adpated the definitions of queries and answers to this extended framework. This deserves further investigations.

Notice that the notion of relative reliability is transitive, since we have $RV^k_{i,j}(p)$ and $RV^k_{j,l}(p)$ implies $RV^k_{i,l}(p)$, and $RC^k_{i,j}(p)$ and $RC^k_{j,l}(p)$ implies $RC^k_{i,l}(p)$. Also, if for two agents i and j we accept the (strong) assumptions $\neg A'_i p \leftrightarrow A'_i \neg p$ and $\neg A'_j p \leftrightarrow A'_j \neg p$, then we have $RV^k_{i,j}(p)$ iff e have $RC^k_{i,j}(\neg p)$.

To give an indication of how to use these notions of relative reliability we consider an example where several information sources may give information about meteorological expectations. These information sources are radio or television broadcasting channels like, for instance: RAI1, RAI2, BBC, RFI, CANAL+, ARTE, FR2, FR3, CNN, RTL,...etc...Information about relative reliability may be, for example: CANAL+ is as reliable as FR2 for expectations about France, BBC is as reliable as ARTE for expectations about Europe, RFI is as reliable as CNN for expectations about the world...Now, we can select some of these informations sources as reference sources, and we can define a partial order among these reference source, which represents the fact that some reference sources are supposed to be more reliable than other ones. For instance, if we accept the idea that local reference sources are more reliable than general reference sources for local expectations, we shall say that, for expectation about South of France, FR3 is more reliable than FR2, which is a french national television, FR2 is more reliable than ARTE, which is a european television, and ARTE is more reliable than CNN. Then, it will be possible to ask questions about expectations about South of France which have the same level of reliability as if they would have been broadcasted by FR3, or by ARTE, or by CNN... Of course, different information sources may broadcast inconsistent data. For instance, ARTE may asserts that tomorrow it will be sunny in France, while CNN asserts that it will be cloudy.

This example shows that we should not confuse two different, even if not independent, problems. The first one is to determine the reliability level of a given sentence p, which corresponds to the reference source r such that we can infer, for some k, $K_k(A'_r p)$; meta information about relative reliability of information sources is used for this purpose. The second one is to merge data which have different, but comparable, reliability levels, using the partial order on reference sources. In particular, in the case where these sets of data are inconsistent, we are faced to the well known problem of belief revision [1].

4 Conclusion

We have presented formal definitions of valid information sources and complete information sources in modal logic. Two different definitions have been presented which are based either on properties of the agents who insert data in a database, or on the type of data stored in a database. Then, we have considered a more general problem, where several agents can exchange data, and where they are supposed to have different reliability levels. This has led to the notion of reference sources who play a role similar to what is called uncertainty factors. For each different framework we have shown how the notion of validity and completeness are related. In [4, 3] we have also shown how the approach presented in section 2 can be adapted to the context of relational databases, where sentences are expressed in relational algebra.

One of the possible extensions of this work is to have a more synthetic characterization of reliability of information sources in terms of topics instead of sentences. The logical framework defined in [6] for reasoning about the links between topics and sentences might be used for this extension.

References

[1] J. Cantwell and H. Svensson. Resolving conflicting information . Technical report, Uppsala Prints and Preprints in Philosophy, 1995.

[2] B. F. Chellas. *Modal Logic: An introduction.* Cambridge University Press, 1988.

[3] R. Demolombe. Answering queries about validity and completness of data: from Modal Logic to Relational Algebra. In *Proc. of the Workshop on Flexible Query-Answering Systems.* Roskilde University, 1996.

[4] R. Demolombe. Validity Queries and Completeness Queries. In *Proc. of 9th International Symposium on Methodologies for Intelligent Systems*, 1996.

[5] R. Demolombe and A.J.I. Jones. Deriving answers to safety queries. In R. Demolombe and T. Imielinski, editors, *Nonstandard queries and nonstandard answers.* Oxford University Press, 1994.

[6] R. Demolombe and A.J.I. Jones. On sentences of the kind "sentence "p" is about topic "t": some steps toward a formal-logical analysis. In H-J. Ohlbach and U. Reyle, editor, *Logic, Language and Reasoning. Essays in Honor of Dov Gabbay.* Kluwer Academic Press, 1996.

[7] A. Motro. Completeness information and its application to query processing. In *Proc. of 12th International Conference on Very Large Data Bases*, 1986.

[8] A. Motro. Integrity = validity + completeness. *ACM TODS*, 14(4), 1989.

References

[1] J. Cantwell and H. Svensson. Resolving conflicting information . Technical report, Uppsala Prints and Preprints in Philosophy, 1995.

[2] B. F. Chellas. *Modal Logic: An introduction.* Cambridge University Press, 1988.

[3] R. Demolombe. Answering queries about validity and completness of data: from Modal Logic to Relational Algebra. In *Proc. of the Workshop on Flexible Query-Answering Systems.* Roskilde University, 1996.

[4] R. Demolombe. Validity Queries and Completeness Queries. In *Proc. of 9th International Symposium on Methodologies for Intelligent Systems*, 1996.

[5] R. Demolombe and A.J.I. Jones. Deriving answers to safety queries. In R. Demolombe and T. Imielinski, editors, *Nonstandard queries and nonstandard answers.* Oxford University Press, 1994.

[6] R. Demolombe and A.J.I. Jones. On sentences of the kind "sentence "p" is about topic "t": some steps toward a formal-logical analysis. In H-J. Ohlbach and U. Reyle, editor, *Logic, Language and Reasoning. Essays in Honor of Dov Gabbay.* Kluwer Academic Press, 1996.

[7] A. Motro. Completeness information and its application to query processing. In *Proc. of 12th International Conference on Very Large Data Bases*, 1986.

[8] A. Motro. Integrity = validity + completeness. *ACM TODS*, 14(4), 1989.