Distributed Belief Revision

Aldo Franco Dragoni
University of Ancona

Paolo Giorgini
University of Trento

Abstract. In this paper, a distributed approach to belief revision is presented. It is conceived as a collective activity of a group of interacting agents, in which each component contributes with its own local beliefs. The integration of the different opinions is performed not by an external supervisor, but by the entire group through an election mechanism. Each agent exchanges information with the other components and uses a local belief revision mechanism to maintain its cognitive state consistent. We propose a model for local belief revision/integration based on what we called: "Principle of Recoverability." Computationally, our way to belief revision consists of three steps acting on the symbolic part of the information, so as to deal with consistency and derivation, and two other steps working with the numerical weight of the information, so as to deal with uncertainty. In order to evaluate and compare the characteristics and performance of the centralized and of the distributed approaches, we made five different experiments simulating a simple society in which each agent is characterized by a degree of competence, communicates with some others, and revise its cognitive state. The results of these experiments are presented in the paper.

Keywords: Belief Revision, Information Fusion, Multi-agent systems

1. Introduction

Some time ago, an eminent biologist, asked whether there will ever be, on earth, organisms more complex than human brain, answered that such an astonishing super-brain already exists and it is the society of the minds interacting on the planet. Indeed, the claim that "thinking" should be regarded as a social phenomenon is a well known theory of cognitive psychology:

> "…the assumption that what we know is [only] a direct reflection of what we can perceive in the physical world has largely disappeared. …People also build their knowledge structures on the basis of what they are told by others, orally, in writing, in pictures, and in gestures. The social context in which cognitive activity takes place is an integral part of that activity, not just the surrounding context for it. …The social invisibly pervades even situations that appear to consist of individuals engaged in private cognitive activity. …The history of a culture -an inherently social history- is carried into each individual act of cognition. …individuals' beliefs constrain what they will observe empirically."[1]

The social view of cognition could help to understand how new scientific theories emerge and old philosophical currents disappear; but it could also explain simple phenomena as: how was it possible that an entire people believed that thunderbolts were thrown by a god, *even if none of its members ever saw him* (presumably)? And: how is it possible that the authors of this paper do no longer believe in Vulcan (as their ancestors probably did) *even if they never entered in a volcano's crater?* No doubt that most of our opinions are not elaborated from the others but, simply, adopted. But here we anticipate two of the main questions of this research:

Q1 how do we choose (possibly unconsciously) the sources of information to believe in, among the many possible conflicting ones?

Q2 how do individuals' criteria for evaluating the others' reliability affect the global social cognitive behavior?

Question Q1 is common to all of us. Q2 is more academic and, probably, could interest few people but sociologists and ... computer scientists. In fact, both questions could have a *descriptive* answer (how people decide) and a *normative* answer (how people should decide in order to improve performances, satisfy requirements or reach goals). The latter kind of answer could interest software engineers building distributed problem-solving systems in which each node is affected by some degree of incompetence. So, we are focusing on a very specific problem of collective intelligence: the *distributed elicitation of knowledge*, i.e., how is it possible that from a variety of inferential schemas and judging capabilities, from different opinions and dogmas, from distinct perspectives and opposite point of views, after a continual interaction, a more uniform (if not unique) vision of the world emerges? Under which extraordinary circumstances the emergent representation of the world results a correct one? Our hope is that of capturing some successful mechanisms in order to replicate them in a world of intelligent highly-engineered (programmed or trained) interacting cognitive agents. As anticipated, a central question is: how should each agent ascribe a relative degree of reliability to any member of its group? Each agent should even be able to evaluate its own reliability, since:

- the sensing equipment from which it experiences the external world might be faulty
- its ability to interpret sensorial data might be insufficient
- its inferential ability might be dubious.

In conclusion, even when agents are supposed not to lie, it is important to define methods for assessing each member's relative degree of reliability. We limited our attention to groups in which this ascription is performed under "liberal policies", i.e., each one is permitted to stand on its own opinions, evaluating him/herself and the others on the basis of the reciprocal experience and acquaintance. From a global normative perspective we distinguish two desiderata:

convergence: are there "*local* cognitive strategies" which favor the convergence of the opinions (independently from their correctness w.r.t the real world)?

correctness: which of these local cognitive strategies favor also the correctnes of the opinions?

where by "local cognitive strategies" we mean:

LCS1 policies of communication

- *when* to communicate
- *what* to communicate
- communicate to *whom*.

LCS2 methods that form a personal opinion about the credibility of the information and the reliability of the sources

Convergence is not trivial; think, for instance, a criterion for LCS2 such as: "always believe the last information to arrive among conflicting information"; it does not guarantee convergence under any policy of communication. These criteria are almost conscious to humans. On the contrary, the way we form our opinions, from directly perceived material and from information received from the others, seems to be partially unconscious. Perhaps, one makes at least five kinds of check after the incoming of a new information from an external source:

- T0 "although I was not aware of it, is this new information *a logical consequence* of my beliefs about the world?"
- T1 "although I was not aware of it, is this information *in accordance with* my beliefs about the world?"
- T1a "although I was not aware of it, is this information in accordance with *my direct experience* of the world?"
- T2 "how many people believe it?"
- T3 "how much reliable are those people?"
- T4 "what's the source's goal when saying that to me?"
- T5 "how much relevant is that to me?"

T0 represents a mere confirmation, while T1 and T1a yield a consistent expansion of the agent's knowledge. People who stress the importance of T1a are confident in themselves. Those who prefer T2 are rather conformist, while check T3 is preferred by suspicious people; noticeable, T3 links the problem of establishing the credibility of the various pieces of information to the problem of evaluating the reliability of their respective sources. Perhaps T4 should be the most important to humans; we are continually addressed by commercial advertisments whose obvious goals should be taken into account when evaluating their truthfulness. However, this paper avoids the problem of goals recognition and goals treatment in order to evaluate the credibility of an information because:

- the artificial agents we have in mind have no hidden agendas and most of them only have inherent aims (as the sensors of the application discussed in [2])
- we experienced the need to simplify the scenario to reach some conclusions, even if they will be partial

We also completely escaped from dealing with relevance (T5); we simply assumed that all the pieces of information running through the network were relevant for each one of its nodes.
Question Q2 before could be rephrased as follows: if all the individuals adopt the same local criteria to evaluate an incoming information, how is the global process of knowledge elicitation affected? For instance, if all the individuals only adopt test T1, then we would expect little gains from the interaction, each staying in his native degree of correctness. On the other hand, if all the individuals exclusively adopt T2, with no regard toward his own and the others' competence, then we would expect a global flattening to the medium degree of correctness of the agency. Perhaps we should adopt a reasonable mix of these criteria. Unfortunately, T1-T3 are too vague to be studied on a statistical simulation basis. We need more precise rules. Summarizing, opinion formation deals with the way we integrate pieces of evidence

coming from different sources (expecially when they are conflicting), and the way we estimate the reliability of the various sources (ourselves included). In the next section we introduce the techniques we adopted to make our agents able to perform these cognitive operations. Question Q2 is discussed in section 3, specifically: what happens to the emergent group's opinions when its members perform as described in section 2 and all obey a common communication policy? The group's performance could be evaluated under different perspectives:

local perspective: by measuring each individual's derived benefit from having been part of the group

global perspective: by comparing the group's global opinions under different strategies of belief revision and different policies of communication

In particular, we are interested in comparing the global group performances under a really *distributed architecture* (where agents are permitted to talk to each other) vs. what we consider a quasi-distributed architecture, or, better, a *centralized architecture* (where a single agent receives information from the others and has the task to reconcile and recapitulate the various opinions for all the group). We tried to make these evaluations by means of simulation. Results are reported in section 4, where we also try some tentative interpretations. Future work and conclusions are sketched in the sections 5 and 6.

2. Revising Beliefs in a Multi-Source Environment

Derived from researches in Multi-Agent [3] and investigative domains [5], the model for beliefs revision/integration adopted in this research is a novel assembly of known techniques to the treatment of consistency and uncertainty. Let us recapitulate here the main ideas.
Defined as a symbolic model-theoretical problem [6] [7] [8] [9], belief revision has also been approached both as a qualitative syntactic process [10] [11] and as a numerical mathematical issue [12]. Both the cognitive state and the incoming information can be represented either as sets of weighted sentences or as sets of weighted possible worlds (the models of the sets of sentences). Weights can be either reals (normally between 0 and 1), representing explicitly the credibility of the sentences/models, or ordinals, representing implicitly the believability of the sentences/models w.r.t. the other ones. Essentially, belief revision consists in the redefinition of these weights in the light of the incoming information.
We think that to revise beliefs in a Multi-Agent scenario, where many sources refer about a same *static* situation (i.e., we are not dealing with "updating"), the framework should posses some requisites.

- Ability to reject an incoming information.
  A belief revision system for a multi-source environment should drop the rationality principle of "priority to the incoming information" which is no more acceptable since there is no strict correlation between the chronology of the informative acts and the credibility of their contents [14]; it seems more reasonable treating all the available pieces of information as they had been collected at the same time. This principle implies that even an improbable piece of information could be part of the current cognitive state just because, *therein*, it is not contradicted by other beliefs. Roughly speaking, we tend to distinguish "credibility", which is a property of beliefs, from "consistency with ...", which is a relation between beliefs and cognitive states (i.e., between beliefs and beliefs). Let CCS denote the *current*

cognitive state, and let p denote a given belief. Qualitatively speaking, there could be four possibilities:

- p has a *high* degree of credibility and belongs to CCS:
  this is a quiet case, the agent does not experience any "cognitive friction"
- p has a *low* degree of credibility and does not belong to CCS:
  this is another quiet case
- p has a *low* degree of credibility and belongs to CCS:
  this case has been discussed above
- p has a *high* degree of credibility and does not belong to CCS:
  this is the most problematic case; the fact is that p contradicts part of the beliefs inside CCS and the conjunction of these beliefs is more credible than p.

- Ability to recover previously discarded beliefs.
  Cognitive agents should be able to recover previously discarded pieces of knowledge after that new evidence redeems them. This should be done not only when the new information r directly "supports" a previously rejected belief p, but also when r *indirectly* supports it by disclaiming the beliefs q that caused its ostracism. More formally, let us denote with K_q^* the cognitive state K revised in the light of q. The "principle of recoverability" [14] says that for every K and every sentences p and q such that $K \vdash p$ and $K_q^* \nvdash p$, there can always be another piece of information r such that $(K_q^*)_r^* \vdash p$, even if $r \nvdash p$. An obvious case should be $r = \neg q$. The rationale for this principle is that, if someone gave us a piece of information (sometime in the past) and currently there is no reason to unbelieve it, then we should still accept (or *re*accept) it! Of course, this principle does not hold for updating, where changes may be irrevocable. This feature could also be subtitled: "revocable treatment of consistency".

- Ability to combine contradictory and concomitant evidences.
  The notion of beliefs integration should blend that of revision [15]. Every incoming information changes the cognitive state. Rejecting the incoming information does not necessarily mean leaving the cognitive state unchanged since, in general, the new information alters the distribution of the credibility weights. Suppose that A is believing p while B comes saying $\neg p$. If B is not convincing, then A will stay in its opinion regarding p, but, probably, she will be less sure about it! Changes will probably propagate to other logically-related beliefs (like in Williams's transmutation [16]) till an equilibrium point is reached. Things become more complex if we accept the principle that *changes in the credibility of an information affect the reliability of its source and vice-versa*. In this case, changes in the credibility of p provided by A yields corresponding changes in the reliability of A, which, in turn, provokes changes in the credibility of other pieces of information provided by the same source A, even if they are not logically related with p.
  As a consequence of this alteration, a completely different selection of preferred sentences might result. This new selection might reconsider some previously discarded beliefs, whether the income information would be accepted or not. Probably, the last come information decreases the credibility of the beliefs with whom it got in contradiction, even in the case that it has been rejected. The same when receiving a piece of information of which we were already aware of; it is not the case that nothing happened (as the fourth postulate in [6] states) since we are now, in general, more sure about that belief. Furthermore, there is no reason to limit the changes introduced by the new information

into an insertion in a pre–established relative order with consequent rearrangement of the rankings to accomplish the logical relations between beliefs (as Williams' transmutation does [16]). If it is true that newcoming information affects the old one, it is likewise true that the latter affects the former. In fact, an autonomous agent (where "autonomous" means that his cognitive state is not determined by other agents) judges the credibility of a new information on the basis of its previous cognitive state.

- Ability to deal with couples <source, information> rather than with <information> alone.
  The way the credibility ordering is generated and revised must reflect the fact that beliefs come from different sources of information, since the reliability and the number of independent informants affect the credibility of the information and vice-versa [17].

- Ability to maintain and compare multiple candidate cognitive states.
  This ability is part of humans intelligence which does not limit its action to comparing single pieces of information but goes on trying to reconstruct alternative cognitive scenarios as far as it is possible.

- Sensibility to the syntax.
  Despite Dalal's "principle of irrelevance of the syntax" [18], syntax plays an important role in everyday life. The way we pack (and unpack) pieces of information reflects the way we organize thinking and judge credibility, importance, relevance and even truthfulness. A testimony of the form $\alpha \wedge \beta \wedge \cdots \wedge \neg\alpha$ from a defendant A in a trial has the same semantic truth value than the testimony $\beta \wedge \neg\beta$ from a defendant B, but we remember of many cases in which B has been condemned while A has been absolved, being regarded his/her testimony "partially true", contrasting with the B's one regarded as "absolutely contradictory". A set of sentences seems not to be logically equivalent to their conjunction and we could change a cognitive state by simply clustering the same beliefs in a different way.

According to us, "Revising beliefs" should mean "dealing with a new broader set of pieces of information". Our sentence-based approach to the revision of a cognitive state envisages two knowledge repositories:

knowledge background KB , which is the set of all the propositions available to the reasoning agent; it may contain all the information received by the agent during its cognitive life, but it may be as well one of the partitions obtainable from the multiple independent knowledge scenarios in which the agent was involved before; KB may be inconsistent;

knowledge base $B \subseteq KB$, which is the maximally consistent, currently preferred piece of knowledge that should be used for reasoning and decision supporting.

Given incoming information p, computationally, our way to belief revision consists of five steps:

S1 detection of the minimally inconsistent subsets of $KB \cup \{p\}$ (*nogoods*)

S2 generation of the maximally consistent subsets of $KB \cup \{p\}$ (*goods*)

S3 revision of the credibility weights of the sentences in $KB \cup \{p\}$

S4 choice of a preferred maximally consistent subset of $KB \cup \{p\}$ as the new revised base B'

S5 selection of the derived sentences which are derivable from B'

The incoming information p, with its weight of evidence, is confronted not just with the current base B, but with the overall knowledge background KB, so that the degrees of credibility of the sentences in $KB \cup \{p\}$ are reviewed on a broader and less prejudicial basis (S3). The main advantage is that we can rescue sentences from KB by virtue of the maximal consistency of B' (S4). If we'd revise only B by p we could not recover information from KB. S4 might select a new base B' syntactically equal to the previous B (meaning that p has been rejected) but, in general, B' will have a different credibility distribution than B. p might be rejected even if S4 chooses a base B' different from B, but that still contains sentences incompatible with p.

When p is consistent with B, not necessarily $B' = B \cup \{p\}$, since S3 may yield a totally different choice at S4. Previously rejected pieces of knowledge $R \subset KB$ can be rescued simply by determining some upsetting between the credibility of a set $S \subset KB$ and the credibility of R, this may happen if p supports R against S.

S1, S2 and S5 deal with consistency and derivation, and act on the symbolic part of the information. Operations are in ATMS style (Assumption-Based Truth Maintenance System, see [19]); to find out nogoods and goods, we could adopt a set-covering algorithm such the one presented in [20]. Notwithstanding this, even in the propositional case, determining all the minimal inconsistencies can be very hard. However, such a condition can be relaxed (the consequence is that some of the goods are not really consistent). In practical applications dealing with commonsense knowledge, such minimal inconsistencies could be provided interactively by the external user.

S3 and S4 deal with uncertainty and work with the numerical weight of the information. Both contribute to the choice of the revised knowledge space so their reasonableness should be evaluated as a couple. Numerical formalisms are able to perform both of them since the credibility of a single sentence p is determined in the same way as the credibility of a set of sentences B by the weights attached to the models of p and B, respectively.

Flexibility is an advantage in separating the two steps; for instance, depending on the characteristics of the knowledge domain under consideration and the kind of task and/or decision that should be taken on the basis of the revision outcome, the selection function could consider also one (or a combination) of the methods described in the previous section. Probabilistic methods with uncertain inputs seem inadequate for the strong dependence that they impose on the credibility of a sentence and that of its negation. We see that the belief-function formalism, in the special guise in which Shafer and Srivastava apply it to auditing [21], could work well because it treats all the pieces of information as they had been provided at the same time. The method has the following I/O (see [13]):

INPUT	list of pairs <source, information>
	list of pairs <source, "a priori" reliability>
OUTPUT	list of pairs <information, credibility>
	list of pairs <source, "a posteriori" reliability>

Let $S = \{s_1, \cdots, s_n\}$ be the set of the sources, and let kb_i be the subset of KB received from s_i. Each source s_i is associated with a *reliability* $R(s_i)$, that is regarded as the probability that s_i is faithful in everything s/he says. The main idea with this multi-source version of the belief

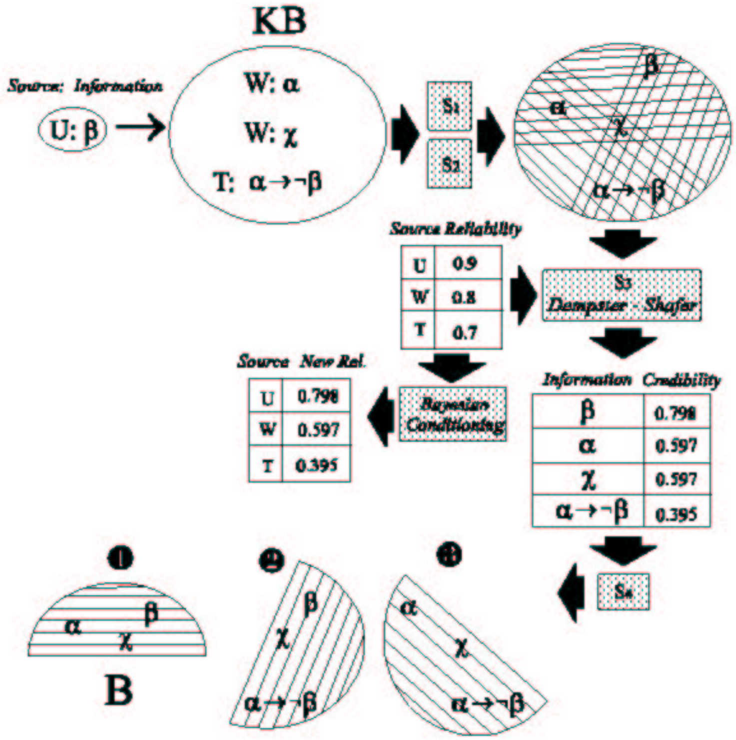


Figure 1. Dempster's and Bayes rules at work.

function framework is that *a reliable source cannot give false information, while an unreliable source may give correct information*; the hypothesis that s_i is reliable is compatible only with the models of kb_i, denoted by $[kb_i]$, while the hypothesis that s_i is unreliable is compatible with all the models of the language adopted to represent the beliefs in KB; let Ω be that set of models. Each source s_i is an evidence for KB and generates the following basic probability assignment $P_i(\cdot)$ on 2^Ω:

$$P_i(X) = \begin{cases} R(s_i) & if X = [kb_i] \\ 1 - R(s_i) & if X = \Omega \\ 0 & otherwise. \end{cases}$$

All these pieces of evidence will be then combined through the Dempster Rule of Combination:

$$P(X) = \frac{\sum\limits_{X_i \cap X_j = X} P_1(X_i) \cdot P_2(X_j)}{\sum\limits_{X_i \cap X_j \neq \emptyset} P_1(X_i) \cdot P_2(X_j)}$$

This rule, extensible to combine n probability assignments, reinforces concordant evidences and weakens conflicting ones. It can be applied only if evidences are independent and referred to the same Ω. Is is important to note here that, because of the commutativity of the product, the rule is independent from the sequence $P_1 \cdots P_n$, so it violates the principle of priority to the incoming information.
Bayesian Conditioning seems, instead, a natural way to estimate the real reliability of the

sources. If the sources are independent, the reliability of $S' \subseteq S$ (i.e., the "a priori" probability that only the sources in S' are reliable) is:

$$R(S') = \prod_{s \in S'} R(s) \cdot \prod_{s \notin S'} (1 - R(s))$$

It holds that:

$$\sum_{S' \in 2^S} R(S') = 1$$

Let $R^*(S')$ denote the "a posteriori" reliability of a set of sources S' (i.e., estimated in the light of all the information contained in KB) . $R^*(S') = 0$ for each S' containing sources in conflict. The others subsets of S are subjected to a bayesian conditioning so that their new reliability sum up again to 1. The "a posteriori" reliability of a source s is defined as:

$$R^*(s) = R^*(\{s\}) = \sum_{S' \supseteq \{s\}} R(S')$$

If a source s has been involved in some contradictions, then $R^*(s) < R(s)$, otherwise $R^*(s) = R(s)$.
S4 translates such ordering on the *sentences* in $KB \cup \{p\}$ into an ordering on the *goods* of $KB \cup \{p\}$. The best classified good is selected as the preferred revised knowledge base. A question is: should S4 consider only the *relative* ordering of the sentences in $KB \cup \{p\}$ or could it take advantage of the *numerical* weights. The first approach seems closer to the human cognitive behavior. The second one seems more informative (it takes into account not only relative positions but also the gaps between the items). Among the qualitative methods described in [11], the "inclusion-based method" seems the most reasonable since it eliminates always the least credible one among conflicting pieces of knowledge. A simple example of a numerical way to perform S4 is ordering the goods according to their average credibility (with this method the preferred good may not contain the most credible sentence). By regarding a set of formulae as a single formula, the belief-function formalism is capable of attaching directly a degree of credibility to each good (thus bypassing S4). However, in [13] we show that if a good contains only part of the information supplied by a source, then its credibility is null. This is unreasonable and, unfortunately, the event is frequent. S5 is not particularly significant since, theoretically, it simply consists in applying classical entailment on the preferred good to deduce plausible conclusion from it. We adopted an ATMS and we stored each sentence derived by the Theorem Prover with an *origin set*, i.e., a set of basic assumptions which are all necessary to derive it. Practically, the last step consists in selecting from the derived sentences, all those whose origin set is subset of the preferred good. We could relax the definition of origin set to that of a set of basic assumptions used to derive the sentence. This is easier to compute and does not have pernicious consequences; the worst that can happen is that, being this relaxed origin set a superset of the real one, it is not sure that it will be a subset of the preferred good as the real one is, and so some derived logical consequences of the preferred good may be not recognized (at first). Besides recoverability, this computational model for belief revision overcomes various limitations of other classic approaches, in particular:

- revision can be iterated
- inconsistent incoming information does not yield inconsistent revised knowledge spaces
- numerical revision is performed on a broader base (the overall KB)

- revision is more flexible; for instance, the incoming information could be rejected even if it is consistent with the current knowledge base
- complete numerical ordering renders the revision as little drastic as possible

Furthermore, the splitting between the symbolic treatment of the inconsistencies and the numerical revision of the credibility weights, provides a clear understanding of what is going on and lucid explanations for the choices.

3. Example

It is an extreme simplification of a case which run of the Inquiry Support System presented in [5].

```
INITIAL SITUATION
 B asserts: A is a partner of R
 M asserts: A was driving the car of S
 A asserts: it is false that A knows S
```

The "a priori" reliability of A, B and M is 0.5. There are no contradictions; there is only one good containing the three propositions and the degrees of reliability don't change.

```
NEW INFORMATION
 U asserts: A was driving the car of S implies A knows S
```

The "a priori" reliability of U is 0.3. The hypothetical rule introduced by U causes a contradiction with the testimonies of A and M. Steps 1 and 2 yield three goods (see Table I); step 3 recalculates the credibility of the beliefs and step 4 sorts the good (in this case the best-out and the numerical average methods agree); the preferred good does not change and the hypothetical rule given by U has been discarded (for the moment). As a byproduct, the degrees of reliability of the sources are recalculated through Bayesian Conditioning. The sources which have been involved in the conflict lose part of their reputation.

```
NEW RELIABILITY OF THE SOURCES
 B : 0.500
 M : 0.459
 A : 0.459
 U : 0.243
```

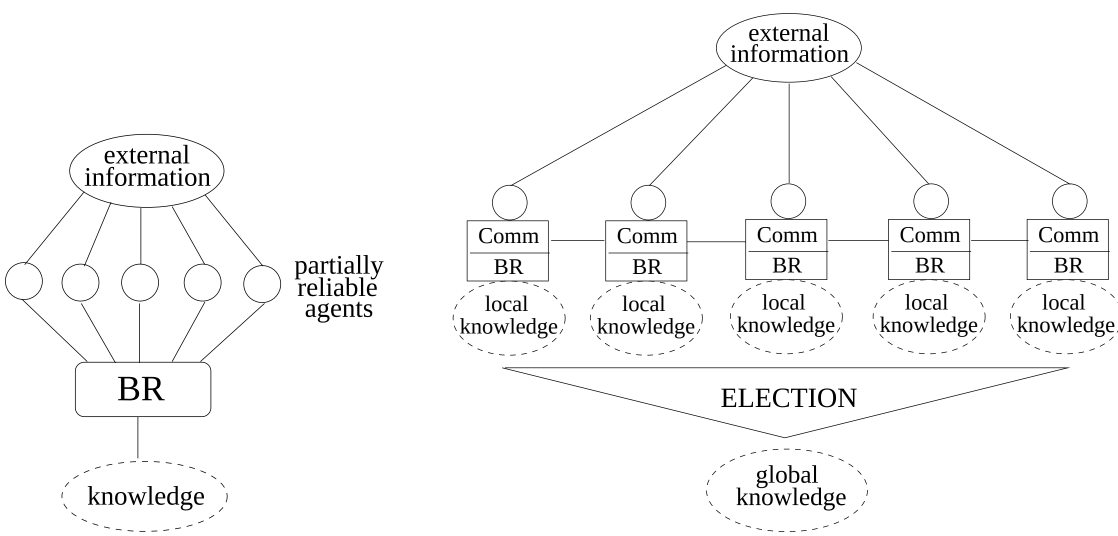
4. The Global perspective

4.1. Centralized vs. Distributed Architectures

Till now we considered the integration/revision of the information under a centralized perspective (Fig 2.a).

Table I. Example. Goods as candidate knowledge bases.

GOOD 1	Average 0.473
A is a partner of R	0.500
A was driving the car of S	0.459
it is false that A knows S	0.459
GOOD 2	Average 0.401
A is a partner of R	0.500
A was driving the car of S	0.459
A knows S or it is false that A was driving the car of S	0.243
GOOD 3	Average 0.401
A is a partner of R	0.500
A knows S or it is false that A was driving the car of S	0.459
it is false that A was driving the car of S	0.243



a. Centralized (BR). b. Distributed (DBR).

Figure 2. The two architectures

Fig 2.b shows what we mean by Distributed Belief Revision. Here nodes exchange information with each others. They need a communication module dealing with the three fundamental policies LCS1 presented in the introduction. Furthermore, communication can either be spontaneous (nodes offer information to each others) or on-demand (nodes ask each others for information). Among the various thinkable criteria, six of them are worth noticing.

Offering information spontaneously to

1. the node considered the *most* reliable
2. the node considered the *least* reliable
3. a *randomly* selected node

Asking information from

1. the node considered the *most* reliable
2. the node considered the *least* reliable
3. a *randomly* selected node

Being guided by an "esprit de corps", one might think that Offering.2 is the best strategy, since it increases the quality of the cognitive state of the less competent members. However, the same collaborative spirit might lead to Offering.1: one should send its best information to the most competent node, so that, sooner or later, she will be recognized also by the others as the most reliable one and the good knowledge will be spread more easily over the group. The latter criterion seems to imply that unreliable nodes will be gradually isolated from the rest of the group. Unfortunately, things are much more complicated since each agent might be wrong in estimating his own and the others' reliability. We introduced two assumptions:

1. nodes do not communicate to the others the sources from where they received the data, but they present themselves as completely responsible for the knowledge they are passing onto the others; a receiver considers the sender as the original source of the information it is sending

2. nodes do not exchange opinions regarding the reliability of the other nodes with whom they got in touch.

With 1 we extend the scope of responsibility: a node is responsible not just for the information that it provides to the network as its original source, but also for the information that it receives from other nodes and, *retaining it credible*, passes it on to the others. With 2 we limit the range of useful information: an agent's opinion regarding the others' (and its own) reliability is drawn out from pure data regarding the knowledge domain under consideration, not from indirect opinions. This also prevents us from falling into the bimillenarian paradox of the liar, i.e., how should we take into account a communication from agent X saying: "agent X is *not* reliable"?
By comparing its opinion with the others' ones, each node produces its own local belief. The effects of the others' opinions depend on the rules adopted by the BR module. Although not necessary, one might want to extract an "emergent" global opinion regarding the information treated by the group. To preserve the decentralized nature of DBR, this opinion should be synthesized not by an external supervisor, but by the entire group through some form of election: the group elects what it believes the global output to be returned to the external world should be. However, nothing prevents getting information directly from a single node's output, since the election does not change the node's personal opinions. The election of the group's emergent output could be done in several ways. We have no room here to explore sufficiently this matter, however at the extreme positions we see two distinct kinds of election:

"data driven" election: the candidates are the pieces of information; the elected will be part of the global output (direct synthesis of the global output, which has, in any case, to be checked for consistency)

"node driven" election: the candidates are the members of the group; the elected will be charged to represent the global opinion (synthesis "by proxy" of the global output)

Many strategies can be conceived by mixing these two kinds of election. We do not expect that the global output emergent from DBR will be better than the output of BR, neither for the quality nor for the quantity of the information provided. What we expect is that the distributed architecture will be:

more efficient: since communications from each agent are not broadcasted to all the others, each agent receives less communications than the superagent in the centralized architecture. Hence each local BR module should manage less information than the global BR module in the centralized architecture, and this should be very important as BR shows exponential complexity

more fault-tolerant: it should be able to offer an acceptable output even in cases that BR fails due to nodes seriously compromised.

less bandwidth consuming: the Comm modules exchange partially elaborated high-level (symbolic) information, which is more synthetic than rough data.

On the other hand, nowadays DBR is a viable alternative to BR since the prices of hardware and communication have been dramatically cut down. We believed that the comparison of the characteristics and performances of the two paradigms illustrated in Fig. 2 could be done only on a simulation basis.

5. The Simulation Experiment

We simulated a simple society [4], in which each agent is characterized by a degree of competence, communicates with some others (through Comm), and revises its cognitive state (through BR). Following Ng and Abramson: "panels with more than five members will generally cost more than they are worth" ([22], p. 927), we limited at six the number of agents. The experiments were of five different types:

Centralized. A single agent adopts BR to revise/integrate information received from the others.

Centralized with communication. Each single component of the society adopts BR and communicates with each other.

Centralized with communication and teacher. The same as before, but with an agent that is regarded as a reliable source by the others.

Centralized with communication and dynamic competence. The same as before, but with agents that change their competence during the experiment.

Distributed. A society of agents adopting locally BR, Comm and global strategies of elicitation of knowledge.

The knowledge domain is simply constituted of two different files, one containing 100 *correct* atomic propositions $\{a, b, \ldots\}$, the other containing their negations $\{\neg a, \neg b, \ldots\}$, which are to be considered *false*. by considering only atomic propositions we reduce the complexity of the BR module without loosing generality w.r.t. the goals of the simulation. Agents, cyclically, access one of these two files, communicate with another agent, and, if necessary, perform a

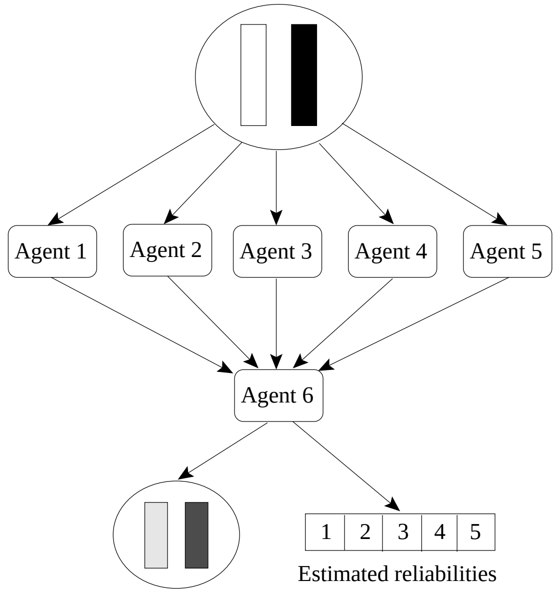


Figure 3. The structure of the experiment 1.

session of belief revision on its cognitive state. Each agent is characterized by a degree of competence (between 0 and 1) that is adopted as the relative frequency with which it accesses (unconsciously) the correct file. The kind of communication depends on the type of experiment. Finally, we adopt the BR method described in section 2 for all the agents.

5.1. Experiment 1: the centralized case

The aim of the experiment was that of evaluate the performances of the BR mechanism. Figure 3 shows the structure of the experiment. Cyclically, $Agent_1 \div Agent_5$ access one of the two files according to their respective degree of competence; they read, randomly, one of the 100 propositions and send it to the "decision maker" $Agent_6$; practically, the five agents work like sensors, monitoring a world simulated by the two files. Within each cycle, $Agent_6$ receives the five propositions and runs BR. The outputs of the experiment are:

- the value of reliability of $Agent_1 \div Agent_5$ estimated by $Agent_6$
- the two files as reconstructed by $Agent_6$.

The experiment's length is 15 cycles and each experiment is repeated 20 times. In order to reduce the effects of casualness, we calculated the average reliability, for each cycle, over the twenty simulations. By differentiating the agents' competence, we made a series of experiments imposing that the initial reliability assigned by $Agent_6$ to the others were all equal to "0.9". Results concerning the integration and the choice of the preferred good show that the BR mechanism is able to guarantee a high percentage of correct choices for the Current Cognitive State. However, we have not sufficiently room to present completely these results, so we refer to [2] for details. We prefer to present here some concerns regarding the sources' reliability. Table II reports the final results of some among the most meaningful experiments.

We could summarize these results as follows.

Table II. Competence and estimated reliability of the five agents in some meaningful experiments.

	Competence					Estimated Reliability				
Exp	1	2	3	4	5	1	2	3	4	5
1	0.8	0.8	0.8	0.8	0.8	0.4	0.4	0.4	0.3	0.4
2	0.6	0.6	0.6	0.6	0.6	0.4	0.3	0.4	0.4	0.4
3	0.4	0.4	0.4	0.4	0.4	0.4	0.3	0.4	0.3	0.4
4	0.2	0.2	0.2	0.2	0.2	0.4	0.4	0.4	0.4	0.4
5	1.0	1.0	1.0	1.0	0.8	0.9	0.9	0.8	0.8	0.3
6	1.0	1.0	1.0	0.8	0.6	0.8	0.9	0.8	0.6	0.1
7	1.0	1.0	0.8	0.6	0.4	0.8	0.8	0.3	0.2	0.2
8	1.0	0.8	0.6	0.4	0.2	0.5	0.5	0.4	0.4	0.3
9	0.8	0.6	0.4	0.2	0.0	0.3	0.3	0.3	0.5	0.5
10	0.6	0.4	0.2	0.0	0.0	0.3	0.4	0.4	0.7	0.8
11	0.4	0.2	0.0	0.0	0.0	0.0	0.5	0.9	0.9	0.8
12	0.2	0.0	0.0	0.0	0.0	0.3	0.9	0.9	0.9	0.9
13	1	0.9	0.8	0.7	0.6	0.6	0.5	0.5	0.5	0.4

- When agents have the same degree of competence, their estimated reliability are very low and similar (see experiments $1 \div 4$). BR is not able to estimate their degrees of competence (all are judged very unreliable) but only that all are, more or less, at the same level of competence.
- When few agents have a degree of competence clearly lower than all the others, BR is able to detect them (see experiments $5 \div 8$)
- When the average competence of the five agents is higher than 0.5, then there is an acceptable correspondence between estimated relative degrees of reliability and real competence. Oppositely, if the average competence is lower than "0.5", then there is an inversion of trend, that is, the more competent agents are estimated less reliable and vice-versa (compare experiment 7 with 11, and the experiment 8 with 10). We called this effect "*majority effect*".

When you do not have certainties regarding the reliability of information sources, Bayesian Conditioning is an acceptable way to estimate their *relative* degrees of reliability. These results are also confirmed by experiments we made in the police investigative domain [5].

5.2. Experiment 2: Centralized with communication

The aim of this second series of experiments was that of studying how communication influences the cognitive state of each component of the society. The structure of the experiment is showed in figure 4.
Here agents were permitted to talk each other on the basis of the assumptions discussed in section 4. Furthermore, communication is peer-to-peer, the receiver is selected randomly, and the piece of information to communicate is also selected randomly in the Current Cognitive State. For each agent we evaluated three parameters: its average *reliability*, the *quality* and the *quantity* of the beliefs in its preferred good G.

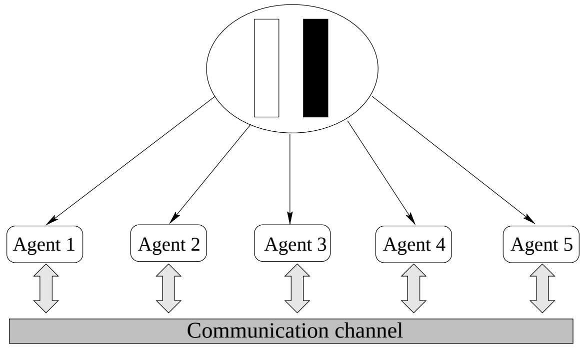


Figure 4. The structure of the experiment 2.

Reliability of the i^{th} agent: $R_i = \frac{\sum_{k=1}^{|agents|} r_{ki}}{|agents|}$ where r_{ki} is the reliability of i estimated by k.

Quality $= Ql_{with-commun.} - Ql_{without-commun.}$ where $Ql = (\frac{|true\ propositions\ in\ G| + |false\ propositions\ in\ G|}{|propositions\ in\ G|})$

Quantity $= Qt_{with-commun.} - Qt_{without-commun.}$ where $Qt = |true\ propositions\ in\ G|$

Quality and Quantity are *local* parameters because they are calculated by using the elements of the knowledge of the single agent. On the contrary, Reliability is a *global* parameter because it is calculated by using the judgments of each member of the group. Each experiment has been repeated 20 times, and we calculated the average of each parameter, for each cycle, over the twenty repetitions. After few tentatives with 40 cycles, we realized that 25 cycles were sufficient to reach stable values. Finally, we imposed that the initial reliabilities assigned by each agent to the other ones were all the same, but in this case we assigned them the value of "0". By differentiating the agents' degrees of competence we made three different types of experiments (A, B and C of table III). In A, agents are equally competent; in B the agents' degrees of competence decrease gradually from $Agent_1$ to $Agent_5$ and, in C the agents' competence is differentiated as reported in table III.

Results can be summarized as follows.

Quality Interaction increases the quality of incompetent agents' cognitive state and decreases the quality of the competent ones. The average quality (Av) stands at zero. This means that if there are few incompetent agents, then they gain much in correctness while the others lose very little. On the contrary, if there are few competent agents, then they lose much in correctness while the others gain very little.

Quantity Interaction always increases quantity. However, incompetent agents gain more than the competent ones. Quantity depends on both the average competence of the group and the particular distribution of competence. In case of low values of average competence it is possible to have also negative value of quantity for the most capable agents.

Reliability All the agents lose reliability, but it is evident what we called "majority effect". Roughly, if the average competence of the group is higher than "0.5", then the competent agents lose less than the incompetent ones. On the other hand, if the average competence is lower than "0.5", then the competent agents lose more than the incompetent ones. A detailed analysis allowed to distinguish three types of results in correspondence of three different sets of average competence:

Table III. Agents' competence in three series of experiments.

	A					B					C				
Exp	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	0	0	0	0	0	0.2	0	0	0	0	1	0	0	0	0
2	0.2	0.2	0.2	0.2	0.2	0.4	0.2	0	0	0	1	1	0	0	0
3	0.4	0.4	0.4	0.4	0.4	0.6	0.4	0.2	0	0	1	1	1	0	0
4	0.5	0.5	0.5	0.5	0.5	0.8	0.6	0.4	0.2	0	1	1	1	1	0
5	0.6	0.6	0.6	0.6	0.6	1	0.8	0.6	0.4	0.2	0.8	0	0	0	0
6	0.8	0.8	0.8	0.8	0.8	1	1	0.8	0.6	0.4	0.8	0.8	0	0	0
7	1	1	1	1	1	1	1	1	0.8	0.6	0.8	0.8	0.8	0	0
8						1	1	1	1	0.8	0.8	0.8	0.8	0.8	0
9											1	0.2	0.2	0.2	0.2
10											1	1	0.2	0.2	0.2
11											1	1	1	0.2	0.2
12											1	1	1	1	0.2
13											0.8	0.2	0.2	0.2	0.2
14											0.8	0.8	0.2	0.2	0.2
15											0.8	0.8	0.8	0.2	0.2
16											0.8	0.8	0.8	0.8	0.2

Average competence $\in$ **[0, 0.4]** there is an exact inversion of the estimated ranking of reliability with respect to the real ranking of competence.

Average competence $\in$ **[0.4, 0.6]** the situation is confused; for some distribution of competence there is correspondence with the estimated degrees of reliability, but for others there is not such a correspondence.

Average competence $\in$ **[0.6, 1]** there is an exact correspondence between (ranking of) real competence and estimated reliability.

Figure 5 shows the trend of the estimated average reliability with respect to the average competence of the group. A high value of average reliability means that the group is confident on its member' competence; on the contrary, a low value means that the group is uncertain on their capacities. This confirms the results: high values of average reliability correspond to low and high degrees of average competence, while low degrees of average reliability correspond to middle values of average competence.

"Majority effect" was an expected result, in fact, evidence in support or against any piece of information is given exclusively by the components of the group, whose degrees of reliability are estimated on a distributed reciprocal basis; there is no firm evidence from the outside of the group, that could be considered as a reference for estimating the reliability of the group's members. The group looks like a class without teacher, or a scientific community without experimental evidence. We will see in the next experiment what happens when a "teacher" agent is present in the society of agents.

All the previous experiments have been repeated with files of 200 propositions and for 100 cycles. The general trends follow those presented before. However, we observed a stabilization of the "quality", a saturation for the "quantity" and higher precision in the estimates of the degrees of "reliability" of the members of the group.

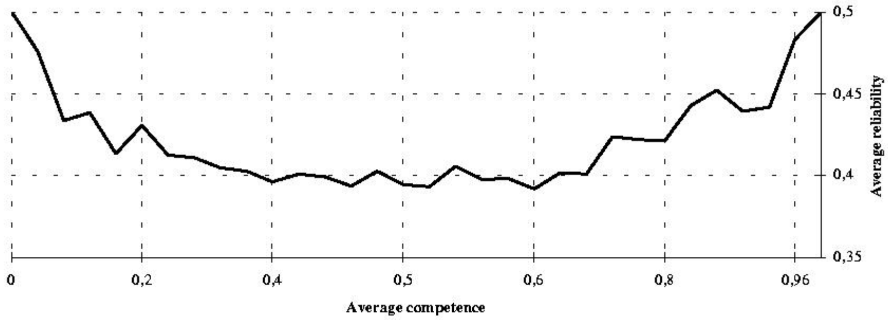


Figure 5. Trend of the esimated average reliability w.r.t. the average competence of the group in the experiment 2.

5.3. Experiment 3: Centralized with communication and teacher

The structure of this trivial experiment is the same of experiment 2 plus an "oracle" agent with the following features:

- $Competence = 1$ (it accesses only the correct file)
- All the agents know that it is absolutely reliable
- It transmits but does not receive information (it's cognitive state will not be contaminated by the others)

Figure 6 shows the structure of the experiment.

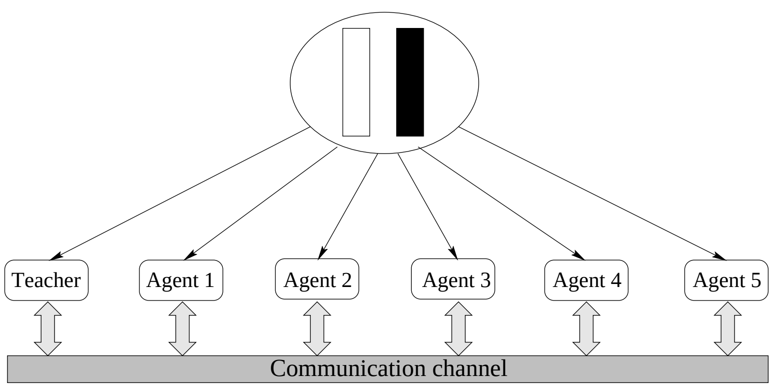


Figure 6. The structure of the experiment 3.

We repeated the experiments of table II with the teacher/oracle. We obtained results that we summarize here:

Quality Now the average quality of the cognitive states is positive and slightly increasing. This is due to the presence of the oracle, which, occasionally (communications took place randomly), gave the possibility to the others to make the right choice.

Quantity Even in this case, interaction increases the amount of correct data in the cognitive state of each agent, and the gain is inversely proportional to the competence of the agent. The effect of the oracle is that the gain is higher than before for any agent

Reliability The correspondence between estimated reliability and real competence holds when the average competence is higher than "0.6". Under that value the situation becomes confused. What we called majority effect in experiment 2 is no longer appreciable when there is a teacher in the group. Figure 7 confirms the absence of majority effect and shows the confused situation for average competence lower than "0.6".

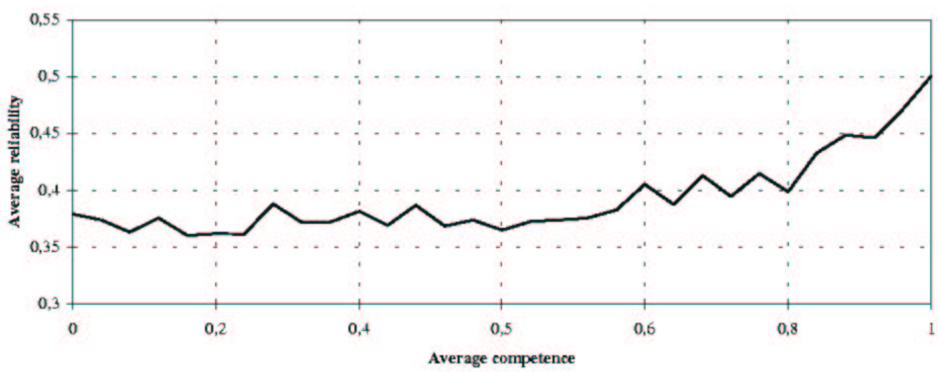


Figure 7. Trend of the esimated average reliability w.r.t. the average competence of the group in the experiment 3.

5.4. Experiment 4: Centralized with communication and dynamic competence

Till now, the characteristics of the agents were static; their competence did not change during the experiments. The aim of this experiment was that of checking if the group would have been able to realize that some of its member changed their degree of competence, and how long it would have taken to the group to be aware of that change. The structure of the experiment is the same of experiment 2. After several simulations with an agent decreasing or increasing its competence, we realized that only high quality groups were able to perceive the change. For groups with an average competence less "0.", the situation at the time of the change was sufficiently chaotic to hide it. This implies that only a decreasing of competence will be perceived by the group. We also tried to study the effects of the depth of the reciprocal acquaintance: the longer the period spent together, the more reluctant should be the agents in changing their opinions regarding their companion who decreased its competence. So we decreased an agent's competence (from "0.98" to "0.8") at different cycles: 0, 5, 10, 15, 20, and 25. As a matter of fact, figure 8 shows its reliability trends, estimated by its companions. We can see that the later the change, the more moderate the slope; the group which appreciated the agent's reliability for a longer period has more inertia to change its opinion regarding that agent.

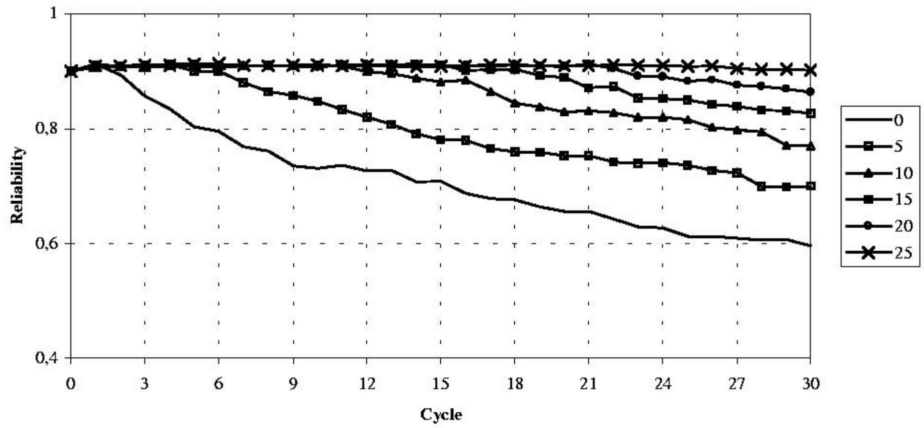


Figure 8. Trend of the estimated reliabilities in the experiment 4.

5.5. Experiment 5: Distributed Revision

The previous experiments were focused on how communication affects the cognitive state of each single element of the group. Now, we move towards a distributed perspective: the focus will be no longer on the single but on the global cognitive behavior. In particular, we will concentrate on comparing the results obtained with the distributed architecture with those obtained in the centralized case (experiment 1). The structure of the experiment is showed in figure 9.

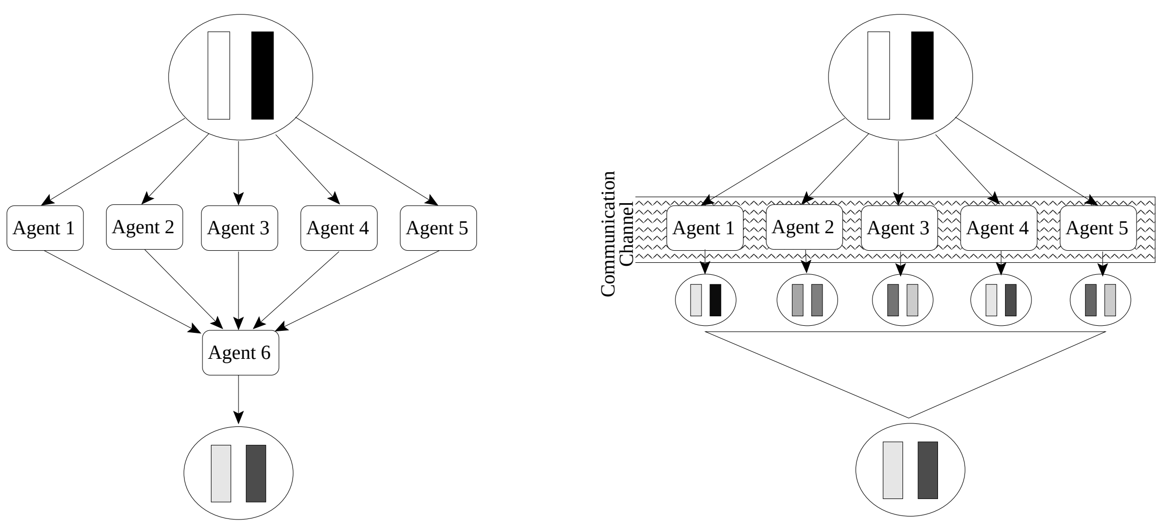


Figure 9. Structure of experiment 5: comparison between BR and DBR

We already discussed the advantage of the distributed architecture. Now, we want to see if these advantages have a price to pay, for instance, a reduction in the correctness of the global representation. Specifically, the experiment aims to compare the ability of the centralized and of the distributed architectures in reconstructing the two files (i.e., comparing the two outputs). Agents adopt a peer to peer communication with one of the rules of the table IV. Each of these fifteen rules are a combination of spontaneous and on-demand communication. A *spontaneous* (Sp) communication consists in offering the most credible piece of information to a randomly chosen agent (R), or to the agent which is retained the most reliable one (+), or to the least reliable one (-). Analogously, an *on-demand* (Od) communication consists in sending the query: "I believe α: do you agree with me or not?", to a randomly chosen agent (R), or to the agent which is retained the most reliable one (+), or to the least reliable one (-). Agents will reply "yes" if they believe α, "no" if they do not, and they do not reply if they have no information about α. The "Nil" in table IV means that the respective kind of communication is not adopted in the experiment.

Table IV. Communication rules P1-P15

		Od			
		+	-	R	Nil
	+	P3	P12	P2	P1
Sp	-	P10	P6	P5	P4
	R	P9	P13	P8	P7
	Nil	P11	P14	P15	

The global opinion is attained by merging the different local beliefs through one of the following four kinds of voting mechanism.

yes/no vote: each agent j votes 1 for a believed proposition i (i.e., the proposition is in its current cognitive state) and 0 for unbelieved pieces of information ($v_{ji} = 0/1$): $V_i = \sum_{j=1}^{|agents|} v_{ji}$

numerical vote: each agent j votes c_{ji} for a believed proposition i (i.e., its own opinion regarding the credibility of i): $V_i = \sum_{j=1}^{|agents|} c_{ji}$

weighted yes/no vote: each agent j's yes/no (0/1) vote is weighted with its average reliability $\overline{R_j} = \frac{\sum_{i \neq j}^{|agents|} r_{ij}}{|agents|-1}$ as estimated by the other agents: i.e.,$V_i = \sum_{j=1}^{|agents|} \overline{R_j} \cdot v_{ji}$

weighted numerical vote: each agent j's numerical vote is weighted with its average reliability $\overline{R_j}$ as estimated by the other agents: i.e.,$V_i = \sum_{j=1}^{|agents|} \overline{R_j} \cdot c_{ji}$

The first two are examples of *data driven* election, whereas the latters are combinations of data driven and *node driven* election. To obtain the global cognitive state by selecting a preferred good among the elected beliefs, we adopt the "best-out" algorithm [11] which is equivalent to the "inclusion-based method" when dealing with simple atomic propositions and their negations. We compared the outputs of the two architectures through the percentage Q of propositions in the correct place (believed if true and unbelieved if false) on the total number of propositions treated by the entire group. Let KBg and Gg be, respectively, the complete knowledge base handled by the group and its final global cognitive state. Then:

$$Q = \left(\frac{|true\ propositions\ in\ Gg| + |false\ propositions\ outside\ Gg|}{|propositions\ in\ KBg|} \right)$$

The comparison is made by the percentage of the difference between Q centralized and Q distributed: $\Delta Q = (Q_{DBR} - Q_{BR}) \cdot 100$

We made many experiments with 24 different distributions of competence among the agents (see [25] and [26] for an extensive description of the experiments). To limit the size of the tables we report here the results for only the following eight distributions of competence.

Table V. Eight significant distributions of competence for the fifth experiment

	1	2	3	4	5
1	0.9	0.9	0.9	0.9	0.9
2	0.8	0.8	0.8	0.8	0.8
3	0.6	0.6	0.6	0.6	0.6
4	0.9	0.8	0.7	0.6	0.5
5	0.9	0.9	0.9	0.9	0.2
6	0.8	0.8	0.8	0.8	0.2
7	0.9	0.9	0.9	0.2	0.2
8	0.8	0.8	0.8	0.2	0.2

We repeated each experiment twenty times to reduce the effects of casualness. Table VII in the Appendix compares the two architectures for each policy of communication under the yes/no election mechanism. Table VIII does the same in the case of numerical vote. Table IX considers the weighted yes/no vote and Table X ends with the weighted numerical vote.

Some considerations:

- The performances of the centralized and of the distributed architectures are similar (this was confirmed by limited series of experiments with many more cycles). Generally the former performs better than the latter.

- There is no voting policy clearly better than the others. However, it seems that weighted vote performs better in groups where the average competence is high but there are few agents with low degrees of competence. Probably, in this case those incompetent agents are properly figured out and the weighting mechanism correctly reduces the importance of their opinion.

- Typically, the distributed architecture performs better under the communication policies P13, P10 and P8.

Let us now focus on two particularly significant distribution of competences: the first (1), where agents are all pretty competent, and the fifth (5), where there is a sharp minority of agents (only one) with a very low degree of competence. Since the voting mechanisms exhibit little variance, we can here mediate on them and leave they out of our consideration. The results (approximated at the second decimal point) are reported in the following Table VI

Some conclusions:

Table VI. Performances of the various policies of communication when all the agents are competent (1) and when there is a limited number of seriously corrupted agents (5).

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	AvΔQ
1	-1,16	-1,62	**0,41**	-1,14	**1,3**	-3,23	-0,76	**0,35**	**1,97**	**1,88**	-1,17	-0,82	**0,22**	-0,19	-2,34	-0,42
5	-2,59	-1,58	**0,2**	-2,33	-0,6	-1,07	-1,59	-1,12	-0,54	-1,5	-2,74	-3,02	-0,19	-3,73	**1,23**	-1,45

- When agents are all competent (row 1), the centralized and the distributed architectures exhibit similar performances. The worst communication policy is P6, i.e., giving and asking information to the least reliable agent. The best communication policy are P9 and P10, which both require asking information to the most reliable agent.
- When few agents are netly incompetent (row 5) the centralized architecture performs better; an easy justification is that the centralized external superagent has immediately a good perception of the faulty agent since it has an overall perspective and a strong majority of reliable agent to rely on. However, as a strange (unjustifiable?) exception, the distributed architecture performs better under the communication policy P15, which requires not giving spontaneously information and asking randomly.
- In general, renouncing to ask for information, i.e., forcing agents to adopt one of the communication policies P1, P4 and P7, is not a good choice. However, in this case the best policy seems to be P7 (giving spontaneously information to agents choosen randomly).
- Even renouncing to spontaneous communication, i.e., forcing agents to adopt one of the communication policies P11, P14 and P15, seems to be not a good choice (with the already remarked exception of P15 but only when there are few incompetent agents).
- If it is not possible to exclude that there are incompetent agents, then adopting the policy of communication P3, which means both giving and asking information to the most reliable agent, assures anyway good group performances.

6. Related Work

Mason's Seismic Event Analyzer (SEA) [27] was initially regarded as an application of distributed revision techniques. Conceived against the proliferation of nuclear weapons, the system main task was that of integrating information coming from a geographically distributed network of seismographs in order to discriminate natural seismic events from underground nuclear explosions. Distributed Truth Maintenance System (DTMS) [28] is one of the theoretical backgrounds of the latest versions of the system. A limit of DTMS is that it presupposes the trustworthiness of the network's nodes. For cases in which the nodes may be mutually inconsistent, Mason proposed a Distributed Assumption-Based Truth Maintenance System (DATMS). By supporting multiple contexts, DATMS was a step toward what they called *liberal belief revision policy*: "it is better let agents stand by their beliefs based on their own view of the evidence". However, Mason's SEA is not a distributed architectures since the data integration is attained in a centralized way, as in every apparatus for sensors' data fusion.
Ng and Abramson [22] studied the derivation of consensus among a panel of qualified experts,

humans or artificial (multi-knowledge base systems). We recognize the need they pointed out of assigning weights to source, information pairs; we also learned that simple aggregation techniques work better, and that the marginal benefit of additional experts, after the fifth, declines rapidly. However there are some major differences:

- our agents do *communicate* each others, so they could reach a real consensus, if possible
- we *estimate* the different degrees of expertise of the sources starting from "a priori" weights
- we do not require our agents to be *sincere*
- the notion of "maximal consistency" lets us accept even *limited portions* of the contributions of the various sources.

As long as "expert opinion aggregation" can be related to "sensor data fusion", one can find in [23] an interesting classification of various techniques, included Bayesian and Dempster-Shafer approaches; in [24] the former technique was shown to be "better performing" than the latter for combining statistical information from multiple sensors. However, all these works refer to a centralized architecture, where a decision maker integrates different opinions from multiple sources; the main contribution of this paper is in the different perspective of the distributed architecture, in which agents communicate each others and decide by themselves standing on their own opinions formed from their partial view of the situation.

7. Conclusions

We were able to address a very limited portion of the items discussed in the introduction. Furthermore, the very specific nature of our simulations prevents us from drawing out general conclusions. Our hope is that of having contributed to stimulate further researches in the field of distributed cognition under realistic assumptions of partial incompetence and/or untruthfulness of some individuals. However, we try here to give some partial answers to the "desiderata" discussed in the introduction.

convergence: our approach to revision/integration of information coming from different sources guarantees the convergence of the opinions regardless of the communication policy adopted by the members of the group

correctness: as expected, the soundness of the opinions heavily depends on the medium degree of competence of the individuals.

From a local perspective, without the help of an external oracle it is not possible to estimate the exact degrees of competence of the information sources. However, the BR method showed a nice ability at estimating their *relative* ranks of reliability, also in case of dynamic competence. That is, BR allows us to know if an agent is more or less competent than others, but not how much it effectively is. Of course, this ranking is as much accurate as more clear is the distinction between competent and incompetent sources. However, this ability of the BR method heavily depends on the average competence of the group. If this is much higher than 50% (this should be the normal case), then the estimated ranking corresponds to the real one. On the opposite, if the average competence is much lower than 50%, then there is an inversion of trend (what we called *majority effect*), that is, the most competent agents receive the worst reputation and vice-versa. The experiments also showed that communication always produces a flattening of the quality of knowledge. Also for quality there exists a majority effect: if the average competence of the group is much higher than 50%, then the quality of the more competent agents decreases, and that of the less competent ones increases, but if the average competence is much lower than 50%, then there is the inversion of trend. The majority effect disappears when one of the agents is considered an oracle from the others (of course, the effect produced by the oracle is positive if, and only if, it is really competent).

When comparing the centralized and the distributed architectures from the perspective of the quality of the knowledge extracted from the group, they seem substantially equivalent (at most a difference of four points per cent), and the voting mechanisms seem to be not so crucial as expected. Of course experiments confirmed that the main advantage of DBR is *efficiency*. This is due to the fact that DBR is a parallel asynchronous architecture and its local BR modules handle at most the 30% of the information managed by the BR module of the external superagent in the centralized case (in our experiments with five agents; this percentage would be much smaller with more agents). Having that module an exponential complexity (both the Assumption-Based Reasoning and the Dempster Rule of Combination contribute to this complexity) it is evident the reduction in memory consumption and CPU time.

If we restrict our attention on two cases, when agents are all competent and when there are very few currupted agents, then we can reach further conclusions. In the first case (where agents are all competent but not completely), the worst we can do is imposing on our agents to give and ask information to the companion they retain the least reliable (of course). On the contrary, asking information to the most reliable agent, and giving spontaneously information to any other assures good results. In the latter case (when there is an unknown corrupted agent), not giving spontaneously information and asking randomly seems to be the best rule of interaction Of course, the normal situation is that we don't know if there are (few) corrupted agents or not! In this case, both giving and asking information to the most reliable companion, assures anyway good group performances, while renouncing to ask for information seems to be not a good strategy of communication.

8. Appendix

References

1. Resnick, L.B. Shared Cognition: Thinking as Social Practice. In L.B. Resnick, J.M. Levine and S.D. Teasly, editors, *Perspectives on Socially Shared Cognition*, Washington DC. American Psychological Association.
2. A. F. Dragoni and P. Giorgini Sensor Data Validation for Nuclear Power Plants through Bayesian Conditioning and Dempster's Rule of Combination. *Computers and Artificial Intelligence*, 17:2–3, pp. 151-168, 1998.
3. A. F. Dragoni and P. Giorgini Belief Revision through the Belief Function Formalism in a Multi-Agent Environment. *Intelligent Agents III*, LNAI n 1193, Springer-Verlag, 1997.
4. A.F. Dragoni, P. Giorgini and M. Baffetti, Distributed Belief Revision vs. Belief Revision in a Multi-Agent Environment. in M. Boman and W. Van de Welde editors,*Multi-Agent Rationality*, LNAI n 1237, Springer-Verlag, 1997.
5. A. F. Dragoni and S. Animali Maximal Consistency and Theory of Evidence in the Police Inquiry Domain. to appear on *Artificial Intelligence and Law Journal*, Kluwer Academic Publisher, 2001.
6. Alchourrn C.E., Gärdenfors P., and Makinson D. On the Logic of Theory Change: Partial meet Contraction and Revision Functions. *The Journal of Simbolic Logic*, 50, pp. 510-530, 1985.

Table VII. ΔQ in the yes/no vote for each policy of communication and competence distribution.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	
1	-1.3	-1.5	0.5	-1.1	1.2	-3.3	-0.7	0.5	1.9	1.7	-1.1	-0.9	0.1	-0.3	-2.2	-0.4
2	0.1	0.3	0.5	-0.3	1.3	0.1	1.1	0.4	0.4	-0.0	-0.9	-0.5	1.4	0.6	-1.0	0.2
3	-0.7	-0.6	-3.2	-0.6	0.2	-0.4	-1.0	-0.1	0.5	0.3	-0.9	-2.5	0.2	-0.7	-2.4	-0.8
4	-1.2	-1.7	-1.0	-1.4	-3.1	-0.5	-1.7	-1.4	-0.6	-2.2	-0.3	-0.6	-0.8	-1.7	-2.3	-1.4
5	-3.2	-2.0	-0.3	-2.5	-0.6	-1.7	-1.6	-1.8	-0.7	-1.8	-3.5	-3.5	-0.3	-4.3	0.7	-1.8
6	-2.8	-1.0	-5.4	-1.8	-4.1	-1.7	-1.6	1.7	-1.8	-0.2	-3.3	-1.5	-0.1	-1.5	1.2	-1.6
7	1.3	1.1	2.1	0.5	-1.4	1.3	-0.3	0.7	0.1	0.4	-1.2	2.1	2.1	-0.1	0.6	0.6
8	2.3	-0.4	-0.8	-1.3	1.6	-2.6	-2.4	0.4	-1.2	2.0	0.4	-3.0	1.2	-2.4	0.4	-0.4
	-0.7	-0.7	-1.0	-1.0	-0.6	-1.1	-1.0	0.3	-0.2	0.0	-1.3	-1.3	0.5	-1.3	-0.6	Av

Table VIII. ΔQ in the numerical vote for each policy of communication and competence distribution.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	
1	-1,1	-1,9	0,5	-0,9	1,2	-3,3	-0,6	0,1	2,0	1,9	-1,3	-0,9	0,4	-0,4	-2,5	-0,5
2	0,4	0,4	0,7	-0,2	1,1	0,5	0,7	0,6	0,7	0,7	-1,1	-0,6	1,5	0,9	-0,9	0,4
3	-0,8	-0,1	-3,0	-0,7	0,3	0,4	-1,5	-0,2	0,6	0,5	-1,4	-2,7	0,2	-0,2	-2,9	-0,8
4	-1,6	-1,6	-0,9	-1,9	-2,9	-0,3	-1,1	-1,4	-0,8	-0,4	0,6	-0,1	0,1	0,8	0,1	-0,8
5	-2,2	-1,9	0,3	-1,8	-0,7	-0,5	-1,6	-1,3	-0,1	-1,3	-2,9	-2,9	-0,4	-4,0	1,5	-1,3
6	-2,8	-1,2	-4,3	-1,9	-4,0	-1,5	-2,1	2,3	-1,3	-0,4	-2,5	-1,5	-0,4	-1,5	1,0	-1,5
7	1,7	1,7	1,9	0,7	-1,1	1,4	0,7	0,8	0,4	0,5	-1,2	2,4	2,7	0,8	0,2	0,9
8	2,7	-0,5	-1,1	-1,0	1,9	-1,3	-2,3	0,6	-1,6	2,3	0,6	-3,6	1,5	-2,6	0,3	-0,3
	-0,5	-0,6	-0,8	-1,0	-0,5	-0,6	-1,0	0,2	0,0	0,5	-1,2	-1,2	0,7	-0,8	-0,4	Av

7. Gärdenfors P. Knowledge in Flux: Modeling the Dynamics of Epistemic States. Cambridge, Mass., MIT Press, 1988.
8. Gärdenfors P. Belief Revision. Cambridge University Press, 1992.
9. M.A. Williams and H. Roth Frontiers in Belief Revision. Kluwer Academic Publisher, 2000.
10. B. Nebel, Base Revision Operations and Schemes: Semantics, Representation, and Complexity. *Proceedings of the "11th European Conference on Artificial Intelligence" (ECAI 94)*, Amsterdam August 8-12. John Wiley & Sons. 1994.
11. Benferhat S., Cayrol C., Dubois D., Lang J. and Prade H., Inconsistency Management and Prioritized Syntax-Based Entailmen. *Proceedings of the "13th Inter. Joint Conf. on Artificial Intelligence" (IJCAI 93)*, pp. 640-645, 1993.
12. Dubois D. and Prade H. A Survey of Belief Revision and Update Rules in Various Uncertainty Models. *International Journal of Intelligent Systems*, 9, pp. 61-100, 1994.
13. A. F. Dragoni Belief Revision: from theory to practice. *The Knowledge Engineering Review*, 12:2, pp. 147-179, Cambridge University Press, 1997.
14. A. F. Dragoni. F. Mascaretti and P. Puliti A Generalized Approach to Consistency-Based Belief Revision. in Gori, M. and Soda, G. (Eds.), *Topics in Artificial Intelligence*, LNAI n 992, Springer-Verlag, 1995.
15. A. F. Dragoni and P. Giorgini Distributed Knowledge Revision-Integration. Proceedings of the *Sixth ACM International Conference on Information Technology and Management*, Las Vegas, ACM Press, november 1997.
16. M.a. Williams Iterated Theory Base Change: A Computational Model. *Proceedings of the "14th Inter. Joint Conf. on Artificial Intelligence" (IJCAI 95)*, pp. 1541-1547, 1995.

Table IX. ΔQ in the weighted yes/no vote for each policy of communication and competence distribution.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	
1	-1,1	-1,4	0,3	-1,3	1,4	-3,1	-0,9	0,6	2,0	1,8	-1,2	-0,8	0,1	0,2	-2,3	-0,4
2	0,6	0,3	0,5	-0,7	1,2	0,0	0,8	0,1	0,5	0,0	-0,7	-0,7	1,3	0,7	-0,5	0,2
3	-1,0	-0,6	-3,2	-0,3	0,0	-0,1	-0,9	-0,4	0,1	0,6	-0,7	-2,9	0,2	-0,7	-2,0	-0,8
4	-0,8	-1,2	-1,2	-1,3	-2,5	-0,6	-1,6	-1,7	-1,1	0,7	-0,4	0,3	-0,1	-0,4	0,2	-0,8
5	-3,8	-1,2	0,4	-2,7	-0,5	-1,4	-1,6	-0,5	-1,2	-1,5	-3,4	-2,8	0,2	-2,9	1,4	-1,4
6	-3,0	-0,8	-5,4	-2,2	-4,2	-1,5	-1,0	2,2	-1,8	-0,4	-2,2	-1,3	0,9	-0,9	1,1	-1,4
7	1,4	0,4	2,5	0,7	-0,7	1,6	-1,3	1,7	-0,6	0,3	-0,5	2,9	2,8	1,3	0,8	0,9
8	2,7	0,0	-0,6	-1,8	2,4	-2,6	-2,2	0,6	-1,2	1,1	0,5	-3,6	1,4	-2,8	1,1	-0,4
	-0,6	-0,6	-0,9	-1,2	-0,4	-1,0	-1,1	0,3	-0,4	0,3	-1,1	-1,1	0,8	-0,7	0,0	Av

Table X. ΔQ in the weighted numerical vote for each policy of communication and competence distribution.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	
1	-1,2	-1,7	0,5	-1,2	1,4	-3,3	-0,8	0,2	2,0	2,1	-1,2	-0,7	0,4	-0,3	-2,5	-0,4
2	0,5	0,6	0,6	-0,3	1,2	0,4	0,5	0,7	0,5	0,4	-1,2	-0,5	1,7	1,0	-0,9	0,3
3	-0,7	-0,5	-3,0	-0,8	0,1	0,3	-1,4	-0,3	0,4	0,5	-1,4	-2,7	0,3	-0,1	-2,7	-0,8
4	-1,7	-1,6	-1,1	-1,9	-2,9	-0,3	-1,0	-1,3	-1,0	-0,7	0,5	-0,3	0,6	0,9	0,1	-0,8
5	-2,2	-1,3	0,4	-2,3	-0,6	-0,7	-1,6	-0,9	-0,1	-1,4	-2,5	-2,8	-0,2	-3,8	1,3	-1,3
6	-2,7	-1,1	-4,9	-2,1	-4,3	-1,4	-1,6	2,3	-1,5	-0,4	-2,5	-1,5	-0,3	-2,0	0,8	-1,6
7	2,2	1,7	1,8	0,8	-1,4	1,4	0,7	1,5	-0,2	1,2	-0,1	2,1	2,7	1,1	0,5	1,1
8	2,7	-0,3	-1,1	-1,0	2,2	-1,7	-3,1	0,4	-1,6	2,1	0,4	-3,6	1,4	-2,7	0,4	-0,4
	-0,4	-0,5	-0,9	-1,1	-0,5	-0,7	-1,0	0,3	-0,2	0,5	-1,0	-1,3	0,8	-0,7	-0,4	Av

17. A. F. Dragoni A Model for Belief Revision in a Multi-Agent Environment. in Werner E. and Demazeau Y. editors, *Decentralized A. I. 3*, North Holland Elsevier Science Publisher, 1992.
18. M. Dalal, Investigations Into a Theory of Knowledge Base Revision. *Proceedings of the "7th National Conf on Artificial Intelligence"*, pp. 475-479, 1988.
19. J. de Kleer An Assumption Based Truth Maintenance System. *Artificial Intelligence*, Vol 28, pp. 127-162, 1986.
20. R. Reiter A Theory of Diagnosis from First Principles. *Artificial Intelligence*, Vol 32, n. 1, pp 57-95, 1987.
21. G. Shafer and R. Srivastava The Bayesian and Belief-Function Formalisms a General Perpsective for Auditing. in G. Shafer and J. Pearl editors, *Readings in Uncertain Reasoning*, Morgan Kaufmann Publishers, 1990.
22. Keung-Chi Ng and B. Abramson, Consensus Diagnosis: A Simulation Study. *IEEE Transactions On Systems, Man, And Cybernetics*, 22(5), pp. 916-928, September/October 1992.
23. I. Bloch, Information Combination Operators for Data Fusion: A Comparative Review with Classification. *IEEE Transactions On Systems, Man, And Cybernetics*, 26(1), pp. 52-67, January 1996.
24. D.M. Buede, P. Girardi, A Target Identification Comparison of Bayesian and Dempster-Shafer Multisensor Fusion. *IEEE Transactions On Systems, Man, And Cybernetics*, 27(5), pp. 569-577, September 1997.
25. M. Baffetti, Un'architettura Distribuita per la Revisione /Integrazione di Conoscenze Incerte. master Thesis (in italian), Universit di Ancona, 1997.
26. P. Giorgini, Belief revision in Multi-Agent Systems. PhD Thesis, Universit di Ancona, 1998.
27. C. Mason, An Intelligent Assistant for Nuclear Test Ban Treaty Verification. *IEEE Expert*, vol 10, no 6, 1995.

28. M. N. Huhns, D. M. Bridgeland, Distributed Truth Maintenance. In Dean, S. M., editor, *Cooperating Knowledge Based Systems*, Springer-Verlag, pp.133-147, 1990.