DETECTING AND CORRECTING MISTAKES IN INFORMATION FUSION

Vivek Bharathan* and John R. Josephson
Laboratory for Artificial Intelligence Research,
The Ohio State University,
Columbus, OH, 43210

ABSTRACT

This article scrutinizes one inherent pitfall in automated information fusion - its fallibility as an artifact of dealing with the real world. Since it is highly desirable to avoid contamination of further processing by past mistakes, we investigate the nature of the recovery process in a prototype agent that performs the Level One Information Fusion task of entity tracking and re-identification, Smart-ASAS. Smart-ASAS attempts to solve this fusion task by treating it as one of Abductive Inference or Inference to the Best Explanation. We discover that the problem space bounding recovery from errors is exponential in nature, but investigate the possibility of handling this computationally complex problem with some proposed heuristics that would result in some satisficing solution.

1. INTRODUCTION

Information dominance is central to conceptions of the Future Force, and information fusion is critically important for information dominance. Unless fusion becomes highly automated, all those sensors and databases, all that bandwidth and computing power, all will be wasted. Decision-making will actually be degraded by information overload.

Now, theoretically, information fusion is "ampliative" inferencing. That is, sometimes fusion is required to produce interpretations (i.e., form conclusions) that go beyond the available evidence. Unlike purely deductive inferencing, fusion sometimes has to guess. In that regard, fusion is like perception, and forming generalizations. If fusion has to guess, it will make mistakes. Some fusion mistakes will be harmless, some will be immediately corrected, and some will never make any difference. However, some mistakes will poison subsequent processing, unless they can be found, and corrected. This paper is concerned with how certain mistakes in information fusion can be found and corrected by detecting and exploiting anomalies that are caused by them.

In this article, we analyze the computational complexity of finding and correcting a mistaken previous conclusion by considering which among previous conclusions is best at making the problem disappear if it is revised, without the revision making worse problems appear. We conclude that the problem is exponential, and thus computationally infeasible in general. However, certain reasonable heuristics may help. This paper outlines two such heuristics that attempt to handle the complexity. The problem is further concretized in this paper to entity tracking and re-identification in All-Source Analysis Systems (ASAS).

2. INFORMATION FUSION AND ABDUCTIVE INFERENCE

An abductive argument has a pattern that can be described as follows:

D is a collection of data (facts, observations, givens).
Hypothesis H explains D (would, if true, explain D).
No other hypothesis explains D as well as H does.

Therefore, H is probably correct.

Note that the conclusion is justified, not simply as a possible explanation, but as the best explanation in contrast with alternatives. The strength of the conclusion H, the force of the probably in the conclusion statement, reasonably depends on the following considerations:

- How decisively the leading hypothesis surpasses the alternatives,
- How good this hypothesis is by itself, independently of considering the alternatives,
- How thorough was the search for alternative explanations.

Besides the judgment of likelihood, willingness to accept the conclusion also reasonably depends on pragmatic considerations, including:

Report Documentation Page						*Form Approved OMB No. 0704-0188*

1. REPORT DATE **01 NOV 2006**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE **Detecting And Correcting Mistakes In Information Fusion**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Laboratory for Artificial Intelligence Research, The Ohio State University, Columbus, OH, 43210**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM002075.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**7**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

- How strong the need is to come to a conclusion at all, especially considering the possibility of gathering further evidence before deciding,
- The costs of being wrong and the rewards of being right.

Humans intuitively recognize these five considerations. Moreover, humans sometimes actually come up to the standards set by these considerations in their actual reasoning. That is, whether or not the arguments are made explicit, people commonly come to conclusions that can be defended using strong arguments addressed to these considerations, with arguments for the distinct superiority of the leading hypothesis, for the deficiencies of alternative explanations, and so on. Thus, this serves both as a descriptive account of human reasoning and argumentation, and also as a prescriptive account, describing characteristics of logically strong abductions, and suggesting that logically strong abductions are indeed good reasoning, and tend toward truth.

As argued by (Josephson et al, 2003, and Llinas et al, 2004), many aspects of high-level information fusion, levels 2 and 3 as described by the Joint Directors of Laboratories (JDL) Data Fusion Model, may be construed as replicating this pattern of best explanation reasoning. On this basis, we attempt to exploit certain advantages from the explanatory relations available to this approach; namely, an expectation-based evaluation of the plausibilities of fusing information, and the ability to competently revise past decisions based on new information. (Bharathan and Josephson, 2006) argues for, and describes some of the advantages of, construing Level One Information Fusion, as a task of Abductive Inference or Inference to the Best Explanation. A prototype entity tracking and re-identification system, Smart-ASAS, is described that employs these ideas, and enjoys these benefits. Such an approach also introduces several relevant dimensions to reasoning based on the explanatory relations between hypotheses and data that are closed to traditional probabilistic approaches.

3. INFORMATION FUSION BY DYNAMIC ABDUCERS

Most fusion agents, by virtue of the task they handle, are typically dynamic abducers, i.e. an entity that by virtue of interaction with the world, handles a steady input information stream, and tries to make sense of this information by producing its best explanation for it at any point in time. The dynamic abducer processes a stream of inputs, $O_1, O_2,\ldots,O_n$, each reporting an "observation". It generates and maintains a changing estimate, $S_1, S_2,\ldots,S_n$, of the current state of the situation (that is, of the reality underlying the observations, at a certain level or levels of description). It does so by explaining, if possible, each input, by making an abductive inference, by which we mean Inference to the Best Explanation, which results in accepting (believing) a best-explanation hypothesis for that input, in the light of explanations for previous inputs. Such a definition closely mirrors the description of the information fusion process. The estimate of the current situation is made available to support other processes of decision-making.

A dynamic abducer may also maintain beliefs about the past, especially the recent past, to constrain possibilities and thus help to interpret new observations. The current and previous estimates of the situation are, in effect, the accumulated beliefs of the agent and these, however, may be subject to hindsight revision in the light of new information. The abducer (Smart-ASAS) we describe in this article proceeds monotonically, that is, it accumulates beliefs, until anomalous data raises significant doubt (abductive dilemma) regarding the validity of either its beliefs, or the incoming data, and causes it to revise its beliefs in an attempt to improve its estimate of the situation. This belief-revision process makes use of abductive reasoning, where the anomalous data are the to-be-explained, and where mistakes in previous reasoning steps are considered among the alternative hypotheses for doing the explaining.

4. PROBLEM SPACE COMPLEXITY FOR BELIEF REVISION IN INFORMATION FUSION

Our model of abduction for information fusion is that of an explanatory process that generates hypotheses that could potentially explain a set of data, comparatively evaluates them and chooses and accepts (believes) a subset of those hypotheses as the best explanation. We shall call the accepted hypotheses Abductive Beliefs (shortened to Beliefs) which are, in effect, the fused conclusions the agent has made in the past, in trying to explain some sensory data. It is this set of Beliefs that is the relevant belief system of the agent. The Beliefs in this set are ordered on the basis of time according to when each was adopted. It is important to note that apart from the intuition that Beliefs formed ahead of others could possibly influence the latter ones, there are no concrete dependencies that may be exploited to identify a subset of tainted decisions that may be causally related to any impugnable Belief. Such a dependency-free belief system is characteristic of most ampliative reasoning systems, except for rare cases (subsets) of reasoning involving logical implications or possibly, statistical or probabilistic inferences. Consequently, the search involved in identifying culprits (past Beliefs whose alteration or suspension could resolve the abductive dilemma) has to deal with a search space without any structure, which as we shall go on to show is also exponential in nature.

The past situation influences and constrains the present, however, and an agent that has a correct estimate of the past has an advantage in reasoning about the present. Thus, a dynamic abducer that is able to correct previous estimates of the past has a reasoning advantage in estimating the current situation. We describe a reasoning strategy whereby a dynamic abducer is able to appropriately revise previous estimates, under some circumstances, and use the revised estimates to improve its estimates of the current situation. In (Bharathan & Josephson, 2005), we describe how mistakes in abductive reasoning may be repaired using abductive meta-reasoning, that is, by abductively diagnosing apparent faults in the abductive agent's system of beliefs, and repairing apparent faults by adopting alternative explanations. It is "meta-" both in reasoning about reasoning states and steps, instead of reasoning only about observations, and also in relying on meta-data in its reasoning, namely, the confidences associated with the reasoning steps (which it uses to guide the generation of hypotheses). Such an approach is supported by the conviction that the best explanation for the discrepancies in the belief system would correspond to the actual faults in the system.

The following table parameterizes the dimensions relevant to meta-abductive belief revision in a dynamic abducer.

Table 1:
Parameters describing Abductive Belief Revision

Symbol	Semantics	Worst Case
k	Number* of atomic hypotheses considered as alternatives in one abductive decision	Same as average case
c	Cost of evaluating and assessing confidence of one atomic hypothesis	Same as average case
n	Number* of decisions made i.e. number of reports received	Same as average case
k'	Number* of alternate next best hypotheses that may be substituted for the revision candidate	k' = k
n'	Number* of revision candidates, the alteration of one of which would resolve the dilemma	n' = n

* Estimate of average (or maximum per) corresponding to a typical number for that situation

The computational effort involved in making an abductive decision would therefore be O(kc + PEIRCE[1]), since it would involve the analysis of each atomic hypothesis in determining the confidence with which each could explain the new data (or a subset of it) in the worst case.

The processing of the i^{th} Belief (B_i), i.e. the determination of whether a revision of B_i would resolve the dilemma, would involve the analysis of conjunctions of the k' possible revisions of B_i (the k' revisions of B_i are the still-plausible alternate explanations that were comparatively evaluated with B_i, in the process of accepting B_i), along with the revisions of one or more of its successor Beliefs and the corresponding abductive processing. The revision of each successor is also constrained by k' such choices. Therefore the computational cost of

1 O(PEIRCE) is the complexity of the hypothesis assembly task (See the PEIRCE algorithm in Abductive Inference – Chapter 9)

comprehensive belief revision, i.e. the evaluation of all revision candidates is given by $O(k'^{(n'+1)})$, ignoring the lower order terms. In the worst case it is $O(k^{(n+1)})$.

A strong case has previously been made for the exigency of belief revision in numerous situations by intelligent agents and more particularly in the fusion process (Bharathan & Josephson, 2005). Therefore it would be useful to be able to perform this revision processing in a computationally efficient manner. As just described, the exhaustive search strategy of brute-force enumeration of all possible revision candidates and their comparative, abductive evaluation in determining the best revision is clearly exponential and impractical. One way of addressing this issue would be to devise satisficing approaches that enable an agent to handle the complex process of revising its beliefs.

5. A SPECIFIC EXAMPLE OF FUSION: ENTITY TRACKING AND RE-IDENTIFICATION

A legacy ASAS system performs entity re-identification and tracking in an area under surveillance. This works by monitoring an Area Of Interest (AOI), taking input from all available ISR resources, and processing reports of sightings of entities of various types at various locations and times. This information is fused to form a consistent picture of the AOI, which is updated as subsequent reports are digested. The process occurs under human control, and usually requires a fair amount of human input to resolve ambiguities and inconsistencies. Smart-ASAS is a prototype next-generation ASAS that explicitly uses abductive inference (best-explanation reasoning) to associate reported entities with known entities as much as possible. Diagrammatic reasoning is used to generate and evaluate hypotheses regarding previously known entities that might help explain an unassociated sighting (Chandrasekaran et al, 2005).

Abductive processing includes the generation and evaluation of hypotheses, including their consistency with available information, and their evaluation as rival explanations for observations. It also includes the decision to accept hypotheses when their values as explainers are sufficiently high, and are sufficiently better than alternative explanations. (Bharathan and Josephson, 2006) provides a more detailed description of the workings of Smart-ASAS.

The challenge then, in the case of such a reasoner attempting to tackle a problem of information fusion, is to confidently identify the incorrect, but abductively supported, conclusions, to retract them, and those conclusions that follow from them, and provide convincing alternative explanations for the previous data and the new data that brought about this self-analysis, in such a way as to maintain as high a degree of consistency, coherence, plausibility and explanatory coverage in the agent's view of the universe, as possible. In Smart-ASAS, this belief revision is performed by treating the discrepancies as the to-be-explained (explanandum), by searching for candidate hypotheses (explanans) among the previous abductive conclusions, and comparatively evaluating them over the degree of confidence of each decision.

6. TOWARD EFFICIENT BELIEF REVISION

On reflection, it seems obvious that the effort involved in belief revision may be measured as the optimization of two factors (a) the confidence of the best revision candidate (explanation) and (b) the computational cost. The search for a potential, highly confident, revision candidate may sometimes be sacrificed for a less promising one, if the computational cost is deemed to be prohibitive in some sense (i.e. time, resources, etc.). Since the strategy under discussion is abduction at a meta-level, it is not surprising that these factors are similar to traditional hypotheses assembly where intuitively, the goal is to maximize explanatory coverage, with as high a level of confidence as possible.

The maximal effort, in this form of belief revision, is expended on the process of identifying plausible, suitable hypotheses whose revision (and subsequent reprocessing of future reports) resolves the dilemma. This is by way of saying that it is highly desirable to form some kind of ranking among the confidences of the past decisions chosen as possible revision candidates, which coupled with some inexpensive test for determining relevancy to the current dilemma would cut through a lot of the exponential processing involved in exhaustive abductive belief revision. The ordering of revision candidates by confidence also allows the technique of pruning away from the search list those Beliefs that are either irrelevant or immutable (i.e. above a certain confidence threshold) with respect to the current dilemma. Our efforts along these directions aim to reduce the burden in the computational complexity of abductive belief revision $O(k'^{(n'+1)})$ by curtailing by a few orders of magnitude, both k' and n' in the O-notation.

6.1 Heuristic 1: Relevance Measure – Ghosts

Inspired by the general principle of memory recall, in which a 'small' subset of relevant information is made available for situation specific problem solving, an attempt was made to capture some similar measure of relevance through which the agent would be able to inexpensively construct an index to access past Beliefs relevant to a particular situation that raises significant doubt regarding the validity of an agent's past beliefs (an abductive dilemma).

The Entity Re-identification algorithm in Smart-ASAS has been augmented with a small routine that records a trace of the old location of an entity when it has been correlated with (and subsequently moved to the location of) a new report. These traces or ghosts, of where entities are supposed to have been at certain times, are then used in determining the relevance of a past decision to an abductive dilemma. This in turn is then used to identify a subset of plausible revision candidates from the entire set of past decisions. In effect, a space-time index is constructed through this heuristic. The idea is effectively communicated through the following illustration. We describe a simple case in which one past decision is fingered as a candidate while the other is ignored as irrelevant, based on our relevance heuristic - ghosts.

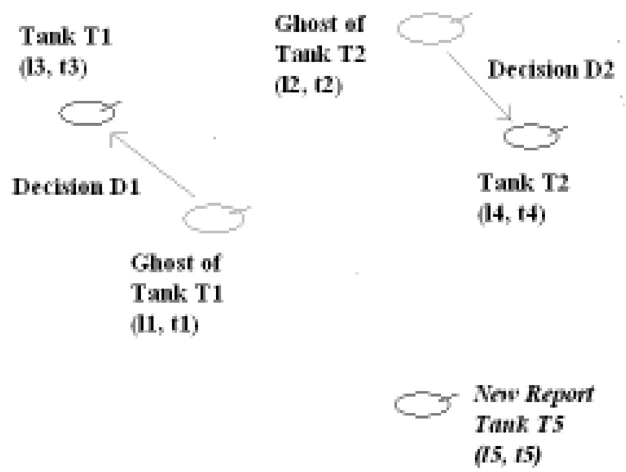

Fig. 1
Illustration of Relevance Measures using Ghosts

Assume that in the agent's past, two correlation decisions have been made:

- Decision D1, associating tank T1, that had been at location l1, at time t1, with a report at location l3, at time t3 &
- Decision D2, associating tank T2, that had been at location l2, at time t2, with a report at location l4, at time t4.

Assume further that when a new report (Tank T5) is received there is no plausible candidate in the region that could satisfactorily explain it. Say it is established through independent reasons that T5 could not be a previously unseen entity at that location. Smart-ASAS thus runs into an abductive dilemma on receiving the new report labeled tank T5, since there are no good hypotheses to explain the available data. On commencement of retraction processing, ASAS-Smart polls its past history for Beliefs relevant to the current location. This is where the ghosts are invoked. On revisiting decision D2, Smart-ASAS, from the ghost of T1, extracts the information that if it were retracted, viable hypothesis would emerge that could explain the report labeled T5. Decision D2 however, is not recognized as a potential revision candidate since neither the current location of its entity, nor the location of its ghost is relevant to the dilemma. Hence, decision D1 is a plausible revision candidate and the previous report at location l3, at time t3, could be reconciled to accepting one of the contending alternates of D1 depending on plausibility. This substitution of the alternate hypothesis is, of course, subject to further fusion processing in the light of all available information.

6.2 Heuristic 2: Entrenchment

The concept of "entrenchment", which refers loosely to some ordering, in terms of attachment and confidence, of the concepts in the human mind, is a notion that may be borrowed from Epistemology (Nayak, 1994) in tackling the complexity of belief revision. A measure of entrenchment might be construed as paralleling the notion of coherence, which in this instance, may be realized as a combination of the elapsed time since the abductive decision was made and how consistent and confident the subsequent decisions that depended on it, turned out to be. This produces an ordering (at least a partial one) of the revision candidates which leads to the less confident decisions being considered before others as revision candidates, and in the case of restrictions of time, it is reasonable to ignore candidates that do not cross a confidence threshold, just as described with ghosts.

It might seem self-defeating, at first glance, to hope to exploit dependencies between past Beliefs in realizing measures of entrenchment while it has already been argued strongly that no such dependencies exist in a purely abductive reasoner. The proposal actually involves re-estimation of the intrinsic confidence of the Belief in the light of subsequent evidence. Typically the abducer stores a Belief and the alternate plausible hypotheses with which it was comparatively evaluated in the abduction process and we can think of the confidence of the abduction decision as the initial measure of entrenchment for that Belief. It is quite possible that sensor information received later enables the abducer to re-estimate the confidence of the decision, by virtue of eliminating one or more of its alternate explainers (of course, the elimination itself has, associated with it, a measure of confidence). The elimination of alternate explainers is made possible by Beliefs adopted subsequently by the abducer (possibly independent of the older Beliefs). The confidence scores of earlier Beliefs could be bolstered by the decision-making process by utilizing them in the adoption of later Beliefs (i.e. the later Beliefs are contingent on some of the earlier ones being true) or by making implausible (possibly by means of a mutual exclusion relation, or a softer version of it, between them), some of the contending alternate explainers to earlier Beliefs. These iterative re-estimations provide the required measure of entrenchment that intuitively, also includes a measure of coherence by acknowledging the other Beliefs that were influenced by the older Belief and its alternate explainers.

7. SUMMARY & FUTURE WORK

In this article we argue that efficient, extensive, information fusion, and thereby automated fusion, is critical to the notion of information dominance. However by the very nature of the task, such processing is fallible and therefore it is necessary for most such agents to incorporate some mechanism for revising past incorrect estimates or decisions. A brute-force approach to such a problem is infeasible due to the exponential nature of the search space and thus some intelligent heuristics are needed to arrive at some satisficing solution set. With the aid of a prototype system, Smart-ASAS, that performs Level One Information Fusion (Entity Tracking), we outline two such ideas for (partially) tackling the computational complexity of the task of a fusion agent recovering from past errors. Quantitative estimates of the step up in efficiency are still pending investigation.

Acknowledgements

Support for this work was provided through participation in the Advanced Decision Architectures Collaborative Technology Alliance sponsored by the U.S. Army Research Laboratory under Cooperative Agreement DAAD19-01-2-0009. The views and conclusions presented in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U. S. Army, or of the U. S. Government.

REFERENCES

B. Chandrasekaran, Unmesh Kurup and Bonny Banerjee. *A Diagrammatic Reasoning Architecture: Design, Implementation and Experiments*. In Proc. AAAI Spring Symposium 2005, Reasoning with Mental and External Diagrams: Computational Modeling and Spatial Assistance, 2005.

James Llinas, Christopher Bowman, Galina Rogova, Alan Steinberg, Ed Waltz and Frank White. *Revisiting the JDL Data Fusion Model II*. In Proceedings of the 7th International Conference on Information Fusion, 2004, Stockholm, Sweden.

John R. Josephson, B. Chndrasekaran, and Mark T. B. Carroll. *Toward a Generic Architecture for Information Fusion*. Proceedings of the Collaborative Technology Alliances Conference 2003, on Advanced Decision Architectures pp. 111-116.

Josephson, J. R. & Josephson, S. G. (Eds.). *Abductive Inference: Computation, Philosophy, Technology*. New York: Cambridge University Press, (1994, 1996). The first chapter is available online at: http://www.cis.ohiostate.edu/jj/pubs/AbdInfCh1.pdf.

Nayak, A. C., 1994, *Iterated belief change based on epistemic entrenchment*, Erkenntnis 41:353-390.

Vivek Bharathan & John. R. Josephson, 2005, *Belief Revision controlled by Meta-Abduction*. Special Issue of Logic Journal of IGPL (Oxford University Press), ABDUCTION, PRACTICAL REASONING, AND CREATIVE INFERENCES IN SCIENCE, edited by L. Magnani, issue 1 or issue 2 in 2006 (Volume 14).

Vivek Bharathan & John R. Josephson, 2006, *An Abductive Framework for Level One Information Fusion*. In Proceedings of FUSION 2006, July 10-13, 2006. Florence, Italy.