A Formal Analysis of Solution Quality in FA/C Distributed Sensor Interpretation Systems*

Norman Carver
Computer Science Department
Southern Illinois University
Carbondale, IL 62901
(carver@cs.siu.edu)

Victor Lesser
Computer Science Department
University of Massachusetts
Amherst, MA 01003
(lesser@cs.umass.edu)

Abstract

The functionally-accurate, cooperative (FA/C) distributed problem-solving paradigm is one approach for organizing distributed problem solving among homogeneous, cooperating agents. The idea behind the FA/C model is that agents should produce tentative, partial results based on only local information and then exchange these results, exploiting the constraints that exist among their local subproblems to resolve the uncertainties and global inconsistencies that result from the use of incomplete information. While several FA/C systems have been implemented, there has been little formal analysis of the quality of the solutions that can be produced using the approach or of the conditions that are necessary for the approach to be effective. This paper reports on work we have done to formally analyze the FA/C model in the context of distributed sensor interpretation (SI). Several results are presented that compare the quality of solutions produced by a distributed FA/C system to those produced by an equivalent centralized system, based on particular agent problem-solving and coordination strategies. We first establish that while it is possible for an FA/C system to produce the same solution as a centralized system, this requires the use of interpretation and coordination strategies that are impractical for most SI applications. Because of this we then consider the effect of "approximate" interpretation and coordination strategies, given some assumptions about the characteristics of the domain.

* This work was supported in part by the Department of the Navy, Office of the Chief of Naval Research, under contract N00014-92-J-1450. The content of the information does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

1 Introduction

In the *functionally accurate, cooperative* (FA/C) paradigm for distributed problem solving [8, 10], agents need not have all the information necessary to completely and accurately solve their subproblems. Instead, agents are designed to produce tentative, partial results based on only local information and to then exchange these results with the other agents to resolve local uncertainties and global inconsistencies. The basic intuition behind this approach is that for many applications there exist (inter-agent) constraints among the subproblems, and these constraints can be exploited to resolve the inconsistencies and uncertainties that occur in local problem solving due to the lack of accurate, complete, and up-to-date information.

While several systems that use the FA/C approach have been built (e.g., [3, 9]), there has never been any formal analysis of the quality of the solutions that can be produced by the approach or of the effect of domain conditions and coordination strategies on solutions. This paper reports on our efforts in formally characterizing the FA/C paradigm. We will present some results about the quality of solutions that can be produced by an FA/C-based system relative to particular agent problem solving and coordination strategies. Since most FA/C applications have been in distributed sensor interpretation (SI) (particularly distributed vehicle monitoring) our analysis focuses on this domain.[1] Though we concentrate on this domain, we believe that our basic results will be applicable to other domains in which local problem solving involves approximate, satisficing search.

First, it is important to be clear about what we will and will not consider. [10] referred to the issues of *solution uncertainty* and *control uncertainty* in FA/C problem solving, and these concepts are useful in delineating the focus of our efforts. In this paper we concentrate on the issue of solution uncertainty: uncertainty about the correctness of partial solutions due to agents having incomplete local information. Control uncertainty is largely concerned with what actions agents should take and when they should take them to achieve efficient, coherent global problem solving. *Coordination* of agent activities involves both solution and control uncertainty, but we will ignore the issue of when information should be communicated and the possible communication of meta-level information. Thus, while we examine the conditions that are *necessary* for effective FA/C problem solving, these conditions are not sufficient to guarantee it. Coordination strategies for *efficient* FA/C problem solving remain an important area of research.[2] Another important issue in the design of FA/C systems is the role that agent architectures play in supporting effective FA/C problem solving. For example, some architectures allow a wide range of inconsistencies to be resolved using very directed, limited communication of data and hypotheses among the agents [3, 4]. Finally, there are other

[1] Distributed sensor interpretation plays an important role in many military situation assessment (decision support) systems.

[2] We are pursuing both empirical and analytic approaches to address this issue [6, 7].

factors that might prevent effective FA/C problem solving in certain domains, but they are beyond the scope of this paper.[3]

The main aspect of FA/C performance that we address is the quality of solutions produced by an FA/C system. Several theorems are presented that compare the solutions of an FA/C system to the solutions that would be produced by an "equivalent" centralized system, given particular agent problem-solving and coordination strategies. A key issue in this evaluation is the fact that approximate, satsificing techniques are often required to identify SI solutions (see Section 3). Because of this, there are several standards against which we might judge solution quality. We have chosen to define the "globally best solution" as the solution that would be produced by an *equivalent* centralized system: a centralized system using the same interpretation strategy as the individual distributed agents use for processing their own local data. The advantage of this definition is that it partially separates the effects of approximate local strategies from the effect of distributed processing.

We also consider the effect of different coordination strategies, in terms of the completeness with which subproblem (local interpretation) interactions are pursued. The analysis assumes the capabilities of an abstract model of the DRESUN system for distributed SI [3, 4]. In particular, we assume that all possible subproblem interactions can be easily identified and that they can be selectively (and incrementally) pursued (see Section 4). Again, there are several alternative strategies that might be used. These range from "complete strategies" that effectively eliminate the effects of the distribution of data, to "incomplete strategies" that do not.

Our results begin by showing that it is possible for a DRESUN-based distributed SI system to produce exactly the same global interpretation as would be produced by a centralized system and that this can indeed be the optimal solution. However, this is true only when the centralized system and the agents are doing complete, exact interpretation and subproblem interactions are completely resolved. As we point out, these are unrealistic assumptions for SI in general and for distributed SI. A key assumption of previous FA/C research has been that solutions can be produced without the need for "excessive" communication among the agents. This is the main reason for using incomplete coordination strategies: to make FA/C problem solving efficient enough to be effective. The rest of our results focus on the effects of incomplete, approximate coordination strategies.

In exploring the use of incomplete coordination strategies, the most important issue we will explore is whether FA/C systems can focus on the subproblem interactions involving only the local solutions produced by the agents. This can greatly increase the efficiency of FA/C problem solving. Unfortunately, this strategy does not in general guarantee that the global solution would be be

[3] For example, in some domains it could be substantially more difficult to represent and develop *partial* solutions than it would be to deal with only complete solutions.

the optimal solution or even the same as what would be produced by a centralized system. In conjunction with this result, we discuss certain domain characteristics that would allow an FA/C distributed SI system to produce solutions that are "satisfactory." The key concept that we consider is what we term *approximate monotonicity.*

The final set of results focus on the use of approximate interpretaton and coordination strategies where solutions are selected using a satisficing approach in which hypotheses are judged part of the solution if their belief surpasses an "acceptance threshold." This is a common strategy for many SI systems. Here the coordination strategy is incremental in that it uses the same criteria to determine the degree to do pursue additional inter-agent suproblem interactions or not. We show that there are conditions under which it is possible to guarantee that the distributed system produces a solution that is comparable to the centralized solution, and other conditions under which there is merely some probability of obtaining such a solution.

In the next two sections we describe our model of (distributed) sensor interpretation, discuss why SI is different from much studied "diagnosis problems," and then examine the use of approximate strategies for determining SI solutions. Section 4 introduces the DRESUN distributed SI model that we assume for our analysis, followed in Section 5 by a description of the coordination strategies that will be considered. Section 6 then defines the notation that will be used in the remainder of the paper. Section 7 contains the theorems and proofs related to FA/C solution quality for distributed SI. The paper concludes with a summary of the major results in the paper and our future research plans.

2 Distributed Sensor Interpretation

By sensor interpretation, we mean the determination of high-level, conceptual explanations of sensor data. Our model of the interpretation process will be essentially that used in [2]: interpretation hypotheses are incrementally constructed via *abductive inferences*, based on a causal domain model that defines the relationships among the data types and abstraction types. For each type T, the causal model defines the type's support, $\mathcal{S}_T$, and its possible explanations, $\mathcal{E}_T$. $\mathcal{S}_T = \{\mathsf{S}_i : \mathsf{S}_i$ is a *type instance specification*, which defines a legal support for a hypothesis of type $T\}$ (a type instance specification has an associated interpretation type and a set of constraints that define the legal attribute values for hypotheses of that type to support a hypothesis of type T). $\mathcal{E}_T = \{\mathsf{E}_i : \mathsf{E}_i$ is an interpretation type that might explain some type T hypothesis$\}$. If T is a *data* type then $\mathcal{S}_T$ is empty. If T is a *top-level* type then $\mathcal{E}_T$ is empty.

An interpretation system makes abductive inferences that identify possible explanations for a piece of data or a hypothesis. Thus, every hypothesis h with $type(h) = T$ is the result of a set of abductive inferences each of the form $h_i \Rightarrow h$, where $type(h_i) \in \mathcal{S}_T$. Conversely, each hypothesis

h with $type(h) = T$ may be part of a set of abductive inferences of the form $h \Rightarrow h_j$, where $type(h_j) \in \mathcal{E}_T$. We will use the following notation: *support-of*$(h) = \{h_i\}$, the (immediate) support for h; $supports(h, h_i)$; and $explains(h, h_j)$. In addition to its immediate support, each hypothesis must ultimately be supported by sensor data via a chain of abductive inferences $d \Rightarrow h_1 \Rightarrow \ldots \Rightarrow h_m$, where d is data and $supports(h_1, d)$, $supports(h_l, h_{l-1})$. To deal with these inference chains, we will use the notation *support-of**(h) to refer to that data that ultimately supports h, as well as the analogous *supports**(h_m, d), and *explains**(d, h_m).

Abductive inferences are uncertain rather than logically correct inferences that provide *evidence* for the hypotheses rather than conclusively proving them. The key source of uncertainty for any hypothesis is the possibility of alternative explanations for the data that supports the hypothesis. Even if *complete* support can be found for a hypothesis (i.e., there is a hypothesis or datum in *support-of*(h) that corresponds to each $\mathsf{S}_i \in \mathcal{S}_T$), the hypothesis may still be uncertain as a result of competing, alternative explanations for *support-of**(h). Furthermore, because hypotheses are incrementally constructed, they may not have complete support at a given point in the processing even if complete support could be found in the available data (complete support might not be able to be found even if the hypothesis is correct due to masking, environmental disturbances, or sensor errors).

An *interpretation* of a data set is an explanation what caused all of the data. In general, an interpretation will be a *composite* of a set of hypotheses whose types are from a specified subset of the interpretation types (the *explanation corpus* [12]), each of which explains some subset of the data, and which together explain all of the data set. Formally, given data set D and interpretation I, if $h_i \in I$ then *support-of**$(h_i) \in D$ and $D =$ the union of *support-of**(h_i) forall $h_i in I$. A *solution* to an interpretation problem is an interpretation of the data that is judged best according to some criteria (we will return to this issue below).

In a centralized SI system, all of the sensor data is available to the single agent. In a distributed system, each agent has (direct) access to data from only a subset of the sensors, and each sensor is associated with a single agent. As a result, each agent monitors only a portion of the overall area of interest, and agents' local solutions must be combined in order to construct a global solution. Construction of a global solution may not be straightforward, however, because the local solutions are often not independent and may in fact be inconsistent because they are based on different incomplete subsets of the data. Agent solutions are interdependent whenever data (evidence) for hypotheses are spread among multiple agents or when agent areas of interest overlap as a result of overlapping sensor coverage.

In this paper we are interested in analyzing the quality of the global solutions that can be produced by FA/C distributed interpretation systems. To do this we need some standard for comparison—i.e., the "globally best solution." Usually one would like this to be the *most probable* *explanation* (MPE) [12] given all of the globally available data.[4] The problem with defining the globally best solution to be the global MPE is that for most real-world SI problems it is impractical to compute the (true) MPE. For this reason and because our main interest is the effect of FA/C distributed problem solving, we have chosen to compare distributed SI solutions to those of an *equivalent* centralized system. By this we mean the solution that would be produced by a centralized (single-agent) system that has access to all of the data of the distributed system and that uses the same (possibly approximate) interpretation strategies as are used by the distributed agents.

The complexity of determining the true MPE can be understood by considering the differences between SI problems and the kinds of problems that are typically studied in research on abductive inference and probabilistic network inference (e.g., [12, 13]). For simplicity, we will refer to these problems as *diagnosis problems*. One key difference is that diagnosis problems are propositional while SI is not, in general. In other words, diagnosis problems have a fixed set of causes and possible findings, with fixed relations among them. The problem of identifying the MPE in such systems has been studied extensively (e.g., [12, 13]). By contrast, while SI problems have a fixed set of "top-level" cause *types* and a fixed set of data *types*, they can have an indeterminate number of instances of any of the types. For example, in a vehicle monitoring system, an unknown number of vehicles will have been responsible for the overall data set and each vehicle will produce a "track" of data points that can be of varying, unknown lengths.[5]

Because there can be an indeterminate number of instances of any top-level cause, a key problem for SI is what is known in the target tracking literature as the *data association problem* [1]: which target should data be associated with? The data association problem and the possibility of an indeterminate number of top-level causes lead to the problem of *correlation ambiguity* (it is ambiguous/uncertain which potential explanation hypothesis a data evidence should be associated with) and to a combinatorial explosion of possible explanations for a data set (potentially, every single piece of data could have a unique source instance as its cause).[6]

Other problems faced by SI systems, but not by most diagnosis problems relate to the nature of sensor data: evidence from sensor data is not conditionally independent even given a source/explanation hypothesis and there can be massive amounts of data resulting from multiple

[4]The MPE is not necessarily the *correct* interpretation—i.e., what actually produced the data. It is simply the most likely interpretation given the actual data and the model of how data can be produced by different events. The MPE can be incorrect when the characteristics of the data are unusual or when the system has a poor domain model. Also, the MPE may not be the "best" solution once one considers what will be done with the interpretation. For instance, failing to warn of a possible attacking aircraft has much more severe negative consequences (utility) than failing to identify a possible friendly one.

[5]Obviously the number of vehicles is usually not completely unknown because it can be bounded based on physical constraints or assumptions about the number vehicles that might appear. However, these constraints are not always very useful: large-scale military surveillance systems may have to deal with many hundreds of simultaneous vehicles.

[6]In general it is often impossible to *conclusively* rule out most of these interpretations, especially in applications where targets may be poorly sensed (i.e., often missed by the sensors).

passive sensors continuously operating in a noisy environment.[7] One way that we might sum up the differences is to say that in diagnosis problems the key source of complexity is the need to consider all combinations of causes for the evidence, while in SI problems it is the explosion of possible interpretations as large amounts of data are considered. Multiple cause combinations are often less of a problem for SI because data can often be assumed to have single causes, so most source hypotheses are alternatives.

Despite the differences between SI and diagnosis problems, it is useful to think of the networks of data and hypotheses that SI systems construct as being very similar to probabilistic belief networks. Key differences to keep in mind are: the SI hypothesis network will be incrementally constructed as data is interpreted, the complete network of possible interpretations may not be constructed, belief computations are often approximate, a noisy-OR model is usually not appropriate for modeling source/explanation interactions, and while networks may be drawn in such a way that data appears to be conditionally independent this is often not the case. Also, instead of random variables, in SI we talk about data and hypotheses, which can be complex, multi-attribute entities.

3 Approximate Interpretation Strategies

To deal with the characteristics discussed in the previous section, SI systems must be *constructive* [2, 5] and they must use approximate, satisficing strategies to determine solutions. It is not the point of this paper to investigate approximate strategies for SI, however to understand the significance of some of our results one must understand that that most real-world SI systems are approximate—even without the issues of distributed SI. We address this in our results by considering the effect of approximate interpretation strategies as well as the effect of different coordination strategies. Interpretation strategies are specified in general terms, since the particular strategies employed will be largely domain-specific. We describe alternative strategies via assumptions about the characteristics of the interpretations that result from the strategies that agents use for their local data. In the remainder of this section, we will give some sense of the kinds of approximation techniques that SI systems can use and their general effects. The distributed SI model we describe in the next section is based on the RESUN/DRESUN architecture, which provides great flexibility for implementing approximate interpretation strategies [2, 3].

There are five basic approximation techniques that can be used by SI systems: process only part of the available data, construct only some of the possible interpretation hypotheses for the processed data, compute approximate belief ratings (conditional probabilities) for the hypotheses,[8] consider only some of the possible composite interpretations (hypothesis combinations), and use

[7]The probabilistic belief network community is just starting to become aware of the data association problem and the issues it raises that have not been addressed in the work on diagnosis problems—e.g., [11].

[8]We will use the term "belief" to mean the degree-of-belief accorded to a hypothesis or interpretation. This will be either the conditional probability or approximate conditional probability of the object.

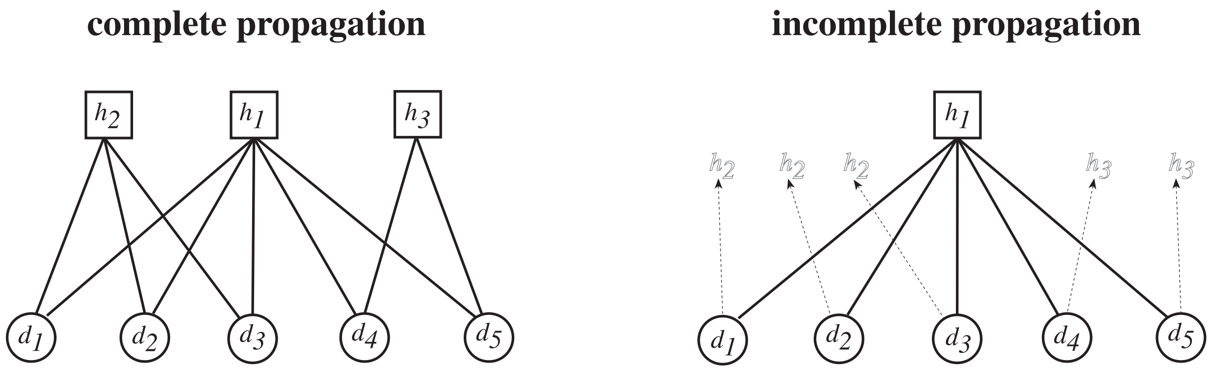

Figure 1: An example of approximate interpretation due to incomplete evidence propagation. In the complete propagation case, the system not only has created the most probable explanation, h_1, it also has created the alternative explanations, h_2 and h_3 (using the most complete support possible). This allows the system to determine the conditional probability of h_1 given the available data $\{d_1, \ldots, d_5\}$. In the incomplete propagation case, the alternative explanations for h_1 have not been created. This means that the belief computed for h_1 is only an approximation of the true conditional probability since the likelihood of the alternative explanations has been only approximately considered. h_1 is still uncertain, though, because the possibility of alternative type 2 and 3 explanations for each piece of supporting data is known, as are the a priori likelihoods of these explanations. (Note that for simplicitly, the figure does not show the numerous incompletely supported versions of the hypotheses that also might have been created in the case of complete propagation.)

some method other than maximum joint probability to select the solution interpretation. For example, the RESUN SI framework allows the use of heuristic knowledge to control the data that is processed and the hypotheses that are constructed, and it allows solutions to be assembled from hypotheses whose belief ratings simply surpass some *acceptance threshold.*

Obviously, these techniques are not independent. If a system does not process all of the data then it cannot in general create all possible interpretations of the complete data set. If a system does not create every possible interpretation of its data (i.e., every possible hypothesis h such that *explains**(D, h)), this not only limits the interpretations that can be considered, it also results in hypothesis belief ratings being only approximations of the true conditional probabilities of the hypotheses. This is because incomplete hypothesis construction results in incomplete propagation of the effects of evidence (we are using "evidence propagation" in the same basic sense that [12] refers to "belief propagation"). Figure 1 provides an example of this situation.

The bottom line is that these approaches will result in SI solutions that are only approximations of the MPE: they may be incomplete or they may not be the most probable composite interpretation. In the following sections we will consider both exact and approximate interpretation strategies being applied by agents to process their locally available data. We will use the name *exact local interpretation strategy* to refer to a strategy in which an agent's local solution (ignoring global interactions) would be the true MPE of the complete local data set. The term *approximate local interpretation strategy* will be used to refer to any strategy that is not exact. We will be more precise about the type and effect of approximate local strategies only as necessary.

One issue that we have not yet discussed, but which is of great importance with many approximate interpretation and coordination strategies is the *termination problem*: has the system done enough work to get a satisfactory solution? For SI this means has the system processed enough data, performed sufficient evidence propagation, and evaluated enough interpretations to have a reasonably good solution? We will return to the issue of termination criteria in our discussion of coordination strategies and in the results.

4 The DRESUN Distributed SI Model

In an FA/C system, there must be some mechanism to drive interactions among the agents so that incorrect and inconsistent local solutions can be detected and dealt with. Ideally, this would be accomplished with a mechanism that allowed agents to understand where there are constraints among their subproblems, so that information interchange could be highly directed. DRESUN [2, 3, 4] provides this capability, and it will form the basis for our model of the capabilities of an FA/C agent.

DRESUN agents create symbolic *source of uncertainty* statements (SOUs) to represent all of the reasons why their hypotheses are uncertain based on their local evidence. Whenever it is determined that a hypothesis can obtain evidence from another agent—i.e., whenever a subproblem interaction (constraint) is detected—DRESUN agents create *global consistency SOUs* (GSOUs) to represent the potential interactions. Each GSOU denotes that another agent might be able to provide additional evidence that would corroborate or contradict the associated local agent hypothesis.[9] GSOUs are viewed as sources of uncertainty about the correctness of an agent's local solution because until a GSOU is "resolved," it is uncertain whether the associated hypothesis is globally consistent. In this way, the basic DRESUN problem-solving goal of resolving interpretation uncertainty drives agents to communicate and produce solutions that are *globally consistent*.

To see that it is possible to identify all possible subproblem interactions, consider that in SI there are just three classes of evidential interactions between a local interpretation hypothesis and an external agent: the external agent might have duplicate support evidence due to overlapping sensor coverage, it might have support evidence for the hypothesis, or it might have explanation evidence. All instances of these situations can be easily detected given a local hypothesis, the SI domain model, and knowledge of the organization of agent interest areas. DRESUN uses three GSOUs to denote instances of these global interactions: *consistent-overlapping-model*, *consistent-global-extension*, and

[9] The GSOUs sometimes denote that an agent (or group of agents) is a *definite* source of evidence and sometimes only that an agent (or agents) is a *possible* source of evidence. Uncertainty arises from things like only partial coverage by an agent of the region from which evidence for a hypothesis might originate—e.g., the possible locations for the source vehicle for a ghost track might include area not completely monitored any agent.

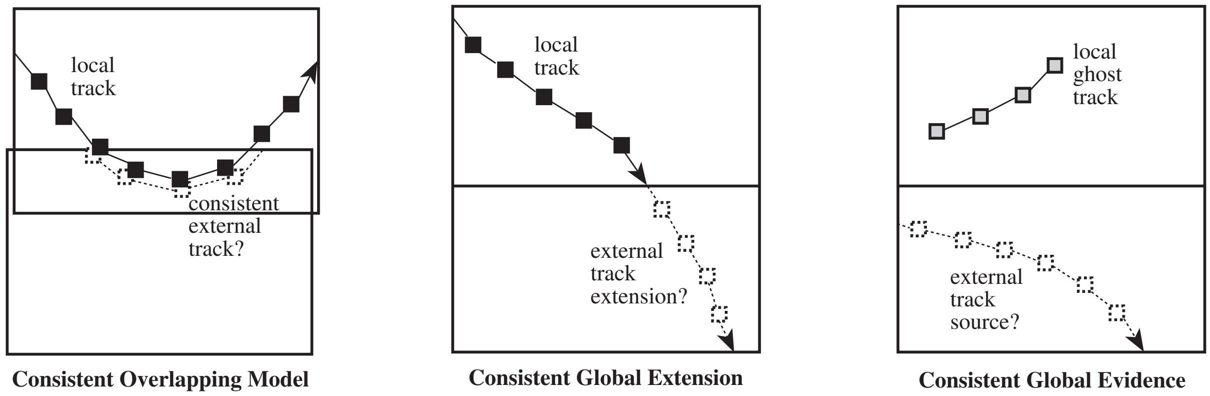

Figure 2: Examples of global consistency SOUs for vehicle monitoring.

consistent-global-evidence. Note that because of the special role of "continuous" support evidence in SI (a key type of support evidence that is not conditionally independent), a special GSOU is used to represent these support interactions while "individual" evidence interactions are represented with another GSOU (that is used for both support and explanation evidence). Figure 2 shows examples of vehicle monitoring situations involving each of these GSOUs.

To understand the effect of approximate local problem solving (interpretation) strategies on global solution quality, it is important to be clear that in DRESUN GSOUs are created based only on the interpretation hypotheses that are constructed. In other words, they do not result from direct analysis of the data.[10] Thus, approximate local interpretation strategies that do not construct all possible interpretations may result in a DRESUN system failing to detect the existence of potential global evidential interactions (which might affect the solutions that are chosen).

5 GSOU Resolution Strategies

While DRESUN agents have a representation of all inter-agent interactions (for the set of locally created hypotheses), we must now consider what should be done with this information in order to develop global solutions. In other words, what does *resolving* a GSOU mean, what are appropriate strategies for resolving the GSOUs, and how do these strategies affect the quality of the global solutions? Basically, resolution of a GSOU involves exchanging information between the associated agents so as to propagate the effects of each agent's local evidence between the agents. An example of the resolution of a GSOU is shown in Figure 3 (the figure will be discussed further below). Resolution of a GSOU is analogous to local (intra-agent) evidence propagation discussed in Section 3 in

[10]While interactions could be postulated directly from the data, this would result in large numbers of highly uncertain possible interactions. For example, in a large-scale military surveillance system it is possible for many targets over the entire area being monitored to be acting in concert, as in some attack scenario. As a result, a priori, there is some (non-zero) probability of *any* two pieces of data being interrelated in such a system.

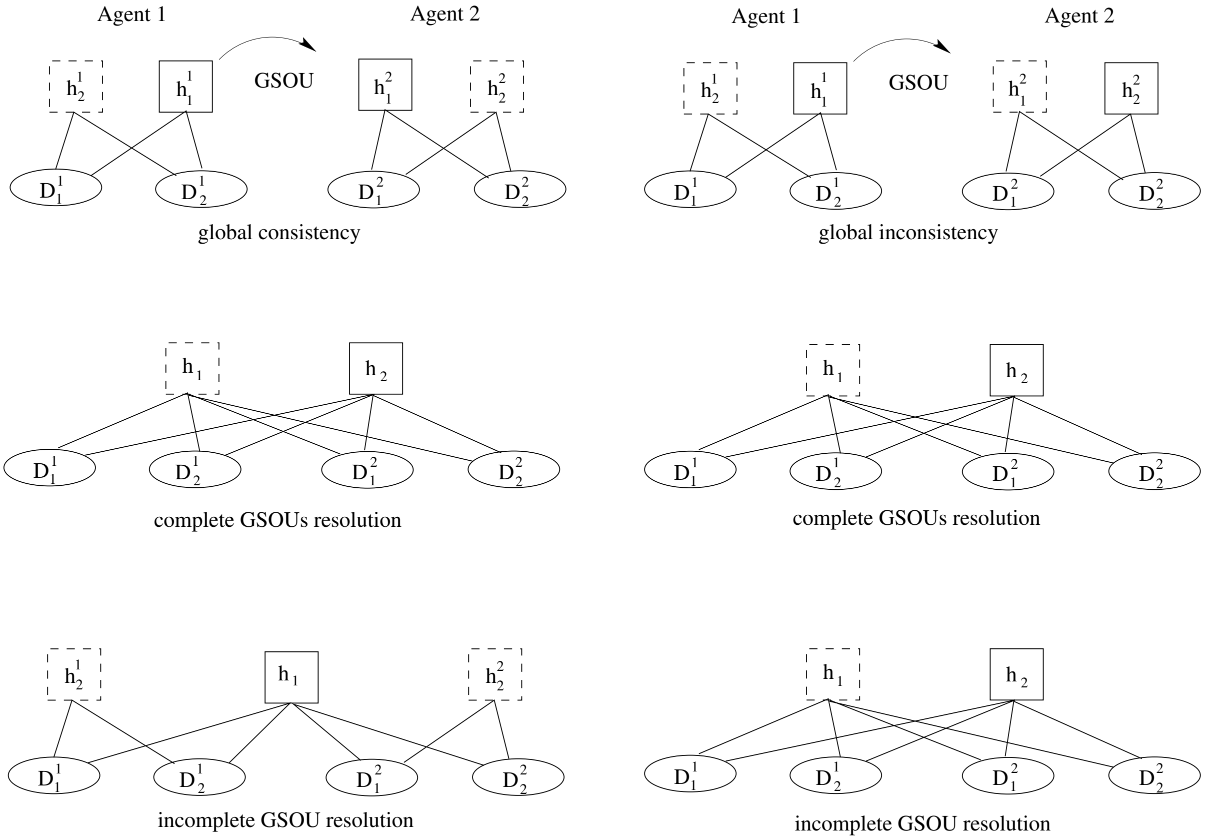

Figure 3: Examples of resolving a global consistency SOU (GSOU).

connection with the local interpretation strategies. As in local propagation, resolution of a GSOU may not be a one shot process, but may require periodic propagation as part of an incremental interpretation process. Thus, conceptually we can think of the GSOU resolution process as establishing evidential links among the associated agents' hypothesis networks (analogous to the links in a belief network).[11]

As with the local interpretation process, there are a range of GSOU resolution strategies that may be used. These strategies differ in how completely and accurately global interactions are considered, based on different choices about which global SOUs to pursue and how completely to assess the evidential effects of external agent data. In considering what are appropriate GSOU resolution strategies, we return to the issue of termination criteria. A key criterion that we wish to impose on FA/C agents is that their local solutions are "consistent" at termination. This allows the local solutions to be merged into a global solution without the need for additional problem solving. Our definition of global consistency of local solutions is an evidential one: solutions are

[11]Since inter-agent evidence propagation requires inter-agent communication, these links are only conceptual. In DRESUN, when evidence has been communicated to an external agent it is tagged. These tags alert the agent to communicate later modifications.

consistent if their hypotheses are pairwise identical, independent, or corroborative. Local solutions are inconsistent when any of their hypotheses are contradictory—i.e., have a negative evidential relationship.

The most comprehensive resolution strategy is for *all* global SOUs to be *completely* resolved—i.e., for all GSOUs in all agents to be resolved using the entire set of possible interpretation hypotheses constructed by the relevant external agents and creating all possible interpretation hypotheses from the joint interpretations (and their supporting data/evidence). We will refer to this as the *complete GSOUs resolution strategy.* An approximate coordination strategy that meets the global consistency criterion is for each agent to resolve only those GSOUs that are associated with its local solution ("best" interpretation) and to pursue evidence propagation using only the local solution hypotheses of the relevant external agents. We will refer to this as the *local solutions GSOUs resolution strategy.*

As we noted earlier, the overall effect of these strategies is dependent on the local interpretation strategy being used by the agents, since resolution is based on the interpretation hypotheses actually constructed by the agents.[12] In fact, there are other reasons why we cannot consider the effects of the global strategies in isolation from the local interpretation strategies. First, we must consider how the GSOUs resolution strategy interacts with the local interpretation strategy when the propagation of evidence among the agents results in an agent changing its belief ratings and/or creating new hypotheses. Our assumption will be that such changes trigger the agent to revise its solution, using the same criteria that it had been using for its local interpretation strategy. Thus, an agent using the *exact local interpretation strategy* will revise its solution to be the true MPE given its previous evidence and its newly acquired evidence. This is a reasonable assumption since SI systems often must be able to incrementally develop their solutions when they are operating continuously or providing users with current interpretations in real-time.

The second interaction issue, which is raised by the need to update the local solutions, is what information does the local interpretation strategy require be propagated between the agents. The local interpretation strategy can have a major impact on the amount of information that must be transferred among the agents to "resolve GSOUs" in the sense needed to meet the global consistency criteria. For example, approximate local strategies may simply require that global resolution provide sufficient information to compute the revised hypothesis beliefs that result from any interactions. By contrast, the *exact local interpretation strategy* has to consider the joint probabilities of the possible interpretations since it is to compute the MPE. In this case, if we are

[12] Thus, a more comprehensive strategy for resolving GSOUs would be one that considers not just the possible interpretations created by the external agents but also any of their data that could serve as evidence for the local hypotheses. We will not consider this alternative—it completely confuses the effects of the local and global strategies because the global strategy can effectively override the local strategy.

to truly have *complete* resolution of the GSOUs then enough information must be transferred to allow each agent to determine "their part" of the MPE given the joint evidence of the agents. When the MPE can be computed using a local computation, message passing scheme (see the discussion of belief revision in [12]) this is not a major issue. However, if the agents require access to all of the possible component hypotheses of the possible interpretations then substantial additional evidence may have to be communicated (beyond that necessary to compute revised hypothesis beliefs).

Figure 3 shows examples of resolving a GSOU for both the *exact local interpretation strategy* and the *local solutions GSOUs resolution strategy*. One thing to note from this figure is the effect that global solution consistency/inconsistency has on the amount of inter-agent propagation that occurs using the two strategies. As the figure shows, the *exact local interpretation strategy* performs the same level of inter-agent evidence propagation regardless of whether the local solutions are consistent or not. By constrast, the *local solutions GSOUs resolution strategy* is an approximate strategy because it does not completely propagate the effects of the agents evidence when the local solutions are consistent. However, it performs exactly the same level of propagation as the *exact local interpretation strategy* when the local solutions are inconsistent.

6 Notation

This section defines the notation that will used in presenting the results and their proofs that appear in the following section. For the centralized (single-agent) case, we will use the following notation:

- $\mathcal{D}$ is the complete data set available to the system.
- D_i denotes some subset of $\mathcal{D}$ ($D_i \subseteq \mathcal{D}$).
- C_i denotes an interpretation *context*—i.e., a particular stage of processing. Each context refers to a specific data subset D_i that has been processed and a particular set of hypotheses and evidential links that have been created.
- $BEL(h, C_i)$ denotes the true degree-of-belief in hypothesis h in context C_i—i.e., $P(h \mid D_i)$, where D_i is the data subset associated with context C_i. This may not be the belief rating actually associated with h—see $\widehat{BEL}(h, C_i)$ below. Also, in general $BEL(h, D_i) \neq BEL(h, \mathcal{D})$ when $D_i \subset \mathcal{D}$, and because interpretation domains are not typically monotonic $BEL(h, D_i)$ may be less than or greater than $BEL(h, \mathcal{D})$. When the interpretation context is clear, we will simply use $BEL(h)$.
- $\widehat{BEL}(h, C_i)$ is the belief rating that is actually associated with h given the interpretation strategy being used. In general $\widehat{BEL}(h, C_i) \neq BEL(h, C_i)$ unless the *exact local interpretation strategy* is being used. As above, $\widehat{BEL}(h, C_i)$ will be written as $\widehat{BEL}(h)$ when the context is clear.
- $\mathcal{I}(\mathcal{D})$ is the true MPE of the complete data set $\mathcal{D}$.
- $\mathcal{I}(C_i)$ is the true MPE of the data subset associated with context C_i. This may not be the solution that is selected by the interpretation strategy—see $\hat{\mathcal{I}}(C_i)$ below.

- $\hat{\mathcal{I}}(C_i)$ is the "best interpretation" solution of the data subset associated with context C_i given the local interpretation strategy being used. We do not specify how the solution interpretation is selected, but it need not be the MPE given the approximate processing represented by C_i (the solution may for instance have been selected using acceptance thresholds).

For the distributed, multi-agent case, modifications must be made to account for the distribution of data and hypotheses among multiple agents. In particular, the concept of an interpretation context must be extended for the distributed case to deal with there being multiple contexts related to each agent and with evidence communication/propagation among the agents as part of GSOUs resolution. The following notation will be used for the distributed case:

- $\mathcal{A}$ is the set of agents $\{A_1, A_2, \ldots\}$, with their interest area specifications.
- $\mathcal{D}^i$ is the complete data set available (directly) to only agent A_i—i.e., the complete data set that is available from agent A_i's own sensors.
- $\mathcal{D}^G$ refers to the complete, globally available data set—i.e., the combined data from all of the agents, $\bigcup_{A_i \in \mathcal{A}} \mathcal{D}^i$. $\mathcal{D}^G = \mathcal{D}$ for an equivalent centralized system.
- D_j^i denotes some subset of agent A_i's local data ($D_j^i \subseteq \mathcal{D}^i$).
- $D_k^{j \to i}$ refers to external data from agent A_j, D_k^j , that is available to agent A_i because it has been communicated from A_j as part of GSOUs resolution. Note that the evidential effects of this data may only partly be known to A_i depending on the local interpretation strategy being used by A_j.
- $D_k^{G \to i}$ refers to external data from all the other agents, $\bigcup_{A_l \neq A_i} D_k^l$, that is available to agent A_i because it has been communicated by these agents as part of GSOUs resolution.
- C_j^i denotes an interpretation context for agent A_i. As with C_i above, each context refers to a specific subset of the local data D_j^i that has been processed and a particular set of hypotheses and evidential links that have been created. In addition, it also refers to any available external evidence that has been communicated from other agents in the process of resolving GSOUs, $D_j^{G \to i}$.
- C_j^G will be used to refer to a global context—i.e., $C_j^G \equiv \bigcup_{A_i \in \mathcal{A}} C_j^i$.
- $BEL^i(h, C_j)$ denotes the true degree-of-belief in hypothesis h in agent A_i's context C_j^i. It is just $P(h \mid D_j^i, D_j^{G \to i})$. This may not be the belief rating that agent A_i actually associates with h—see $\widehat{BEL}^i(h, C_j)$ below. When the interpretation context is clear, we will simply use $BEL^i(h)$.
- $\widehat{BEL}^i(h, C_j)$ is the belief rating that agent A_i actually associates with h in context C_j^i given the interpretation strategy being used. $\widehat{BEL}^i(h, C_j) \neq BEL^i(h, C_j)$ unless the agents are using the *exact local interpretation strategy*. As above, $\widehat{BEL}^i(h, C_j)$ will be written as $\widehat{BEL}^i(h)$ when the context is clear.
- $\mathcal{I}^i(C_j)$ is the MPE for the data subset associated with context C_j^i.
- $\hat{\mathcal{I}}^i(C_j)$ is agent A_i's "best interpretation" solution of the data subset associated with context C_j^i—given the local interpretation and global resolution strategies being used. As in the

centralized case, we do not specify how the solution interpretation is selected but it need not be the MPE given the approximate processing represented by C_j^i.

- $\hat{\mathcal{I}}^G(C_j)$ denotes the combined, global "best interpretation" solution of the data subset associated with the global context C_j^G. That is, $\hat{\mathcal{I}}^G(C_j) = \bigcup_{A_i \in \mathcal{A}} \hat{\mathcal{I}}^i(C_j)$. $\hat{\mathcal{I}}^G(C_j)$ is defined only when all the $\hat{\mathcal{I}}^i(C_j)$ are consistent, as in a final context that satisfies the consistency termination criteria.

7 Solution Quality Results

In this section we will examine the quality of solutions that can be produced using an FA/C approach, relative to those that could be produced using an equivalent centralized approach. To rephrase the main question we are interested in addressing: what is the relationship between $\hat{\mathcal{I}}^G(C_{f_d}^G)$ and $\hat{\mathcal{I}}(C_{f_c})$, where $C_{f_d}^G$ and C_{f_c} represent possible final contexts (contexts that satisfy the termination criteria) for the distributed and centralized cases, respectively?

7.1 Theorem 1

In this section we will examine the quality of FA/C global solutions assuming that the centralized system and the distributed agents use the *exact local interpretation strategy* to process their local data and that the distributed agents use the *complete GSOUs resolution strategy* (the strategies are described in Sections 3 and 5).

Theorem 1: *Given a centralized system that uses the* exact local interpretation strategy*, a set of agents $\mathcal{A}$ each of which uses the* exact local interpretation strategy *and the* complete GSOUs resolution strategy*, $\forall \mathcal{D}$ and $\forall \{\mathcal{D}^i\}$ with $\mathcal{D}^G = \mathcal{D}$ (all possible divisions of the data among the agents), $\hat{\mathcal{I}}^G(C_{f_d}^G) = \hat{\mathcal{I}}(C_{f_c}) = \mathcal{I}(\mathcal{D})$.*[13] In other words, when agents use the *exact local interpretation strategy* and the *complete GOUs resolution strategy*, the FA/C distributed system produces the same exact interpretation as the centralized system, which is the true MPE of all the globally available data.

Proof: Clearly $\hat{\mathcal{I}}(C_{f_c}) = \mathcal{I}(\mathcal{D})$ by the definition of the *exact local interpretation strategy* and the fact that the centralized system (agent) has access to the complete data set $\mathcal{D}$. To show that $\hat{\mathcal{I}}^G(C_{f_d}^G) = \mathcal{I}(\mathcal{D})$ we must show that if $h \in \hat{\mathcal{I}}^G(C_{f_d}^G)$ then $h \in \mathcal{I}(\mathcal{D})$ and if $h \in \mathcal{I}(\mathcal{D})$ then $h \in \hat{\mathcal{I}}^G(C_{f_d}^G)$. Suppose that $h \in \hat{\mathcal{I}}^G(C_{f_d}^G)$. This means that there is at least one $A_i \in \mathcal{A}$ such that $h \in \hat{\mathcal{I}}^i(C_{f_d}^i)$. We specified that when using these two strategies in conjunction, enough evidential

[13] Note that we can refer to the global interpretation $\hat{\mathcal{I}}^G(C_{f_d}^G)$ rather than just individual agent interpretations $\hat{\mathcal{I}}^i(C_{f_d}^G)$ since in the final context $C_{f_d}^G$ individual agent solutions must be consistent (mergeable).

information must be transferred among the agents so that the local agent solutions are not only consistent, but together represent the MPE given the evidence based on their local hypotheses. If we assume that every agent creates every possible interpretation of its data as part of the *exact local interpretation strategy*, then this will mean that every possible inter-agent evidential relation will be represented by some GSOU and so will be resolved. Thus, in a final context where all GSOUs have been completely resolved and exact local interpretations have been computed, h must be in $\mathcal{I}(\mathcal{D})$, the MPE given the globally available data. Now, suppose that $h \in \mathcal{I}(\mathcal{D})$. If we again assume that agents are creating every possible interpretation hypothesis as part of the *exact local interpretation strategy*, h will have been created by at least one agent A_i. There are two cases to consider: 1.) $BEL^i(h, \mathcal{D}^i) = BEL(h, \mathcal{D})$ and 2.) $BEL^i(h, \mathcal{D}^i) \neq BEL(h, \mathcal{D})$. In the first case, the belief in h is independent of the data in all external agents. This means that A_i can locally assess the belief in h and whether it is part of the MPE without any resolution of GSOUs. Thus, if $h \in \mathcal{I}(\mathcal{D})$ then $h \in \hat{\mathcal{I}}^i(C^i_{f_d})$ and so $h \in \hat{\mathcal{I}}^G(C^G_{f_d})$ (by the global consistency criteria). In the second case, A_i cannot locally assess the belief in h. However, if we assume that all possible interpretation hypotheses are created and all GSOUs are completely resolved, then agent A_i must have the information necessary to assess both the belief in h and whether it is in the MPE. Thus again, if $h \in \mathcal{I}(\mathcal{D})$ then $h \in \hat{\mathcal{I}}^i(C^i_{f_d})$ and so $h \in \hat{\mathcal{I}}^G(C^G_{f_d})$ (by the global consistency criteria).

This result is what we would expect given the definitions of the *exact local interpretation strategy* and the *complete GOUs resolution strategy*, and given the assumptions inherent in the DRESUN model (that the GSOUs represent all possible interactions for the set of created hypotheses). However, it serves to formally establish the theoretical capabilities of a DRESUN-based FA/C distributed SI system. Other FA/C distributed SI architectures have not had this property because they have lacked a representation of all the inter-agent interactions and a mechanism for controlling their resolution. This result shows that an incremental, distributed approach to SI need not produce poorer quality solutions than a centralized system.[14] Furthermore, the result is not trivially obvious. We had to make the assumption that agents create all possible interpretation hypotheses and we also had to be very careful in how we defined the interactions between the local interpretation and GSOUs resolution strategies for the result to hold.

Unfortunately, the practical implications of this result are very limited since few SI systems (centralized or distributed) can afford to do exact local interpretation and complete GSOUs resolution. The only time such a situation would be practical would be if enough agents are used so that each agent's local processing requirements are adequate to completely and exactly process its

[14]We assumed that both systems had access to exactly the same data. This is reasonable if only passive sensors are being used or if the SI is not being done in real-time, but is much less reasonable once active sensors can be used in real-time SI.

local data and if the global interactions are limited in both number and scope. In other words, if few GSOUs are produced and those that are can be resolved with limited processing. This latter point is important, since even if there are few GSOUs, it may be very costly to completely resolve every one if they can each potentially interact with large numbers of interpretation hypotheses in the associated external agent(s).

It is easy to show that the use of approximate local interpretation strategies or an incomplete GSOUs resolution strategy results in the systems failing to be guaranteed of producing the MPE. However, if we accept that even centralized SI systems will have to settle for sub-optimal solutions, then what we are really interested in determining is whether distributed systems can produce solutions that are equivalent with (or at least sufficiently close to) those produced by approximate centralized systems, and whether they can do it efficiently.[15] It certainly is possible to come up with approximate local interpretation strategies for an approximate FA/C-based SI system to produce solutions that are equivalent to those of an approximate centralized system. However, this depends on the details of the approximate interpretation strategy since there are only certain approximation strategies whose effects can be duplicated by a GSOUs resolution strategy (the agents in the distributed system will still have limited views of the globally avaiable raw data, for instance). In the remaining results subsections, we will begin to consider approximate interpretation and coordination strategies.

7.2 Theorem 2

We are not yet at the point where we can formally define and analyze the effects of arbitrary interpretation and coordination strategies. However, in this section we will undertake some evaluation of the effects of the *local solutions GSOUs resolution strategy*. This is an important strategy because it can be very efficient when local agent solutions are consistent. In addition, the developers of the FA/C paradigm clearly saw this as a useful coordination strategy. For example, [8] refers to *consistency checking* of the tentative local solutions with results received from other nodes as "an important part of the FA/C approach."

Theorem 2: *Given a centralized system that uses an arbitrary (possibly approximate) local interpretation strategy and a set of agents $\mathcal{A}$ each of which uses this same local interpretation strategy and the* local solutions GSOUs resolution strategy*, $\exists \mathcal{D}$ and $\exists \{\mathcal{D}^i\}$ with $\mathcal{D}^G = \mathcal{D}$ (some*

[15] To clarify this discussion, we should point out that a distributed SI system does not necessarily do more "work" than the centralized system does to get the MPE solution. Completely resolving the GSOUs results in the distributed system performing precisely the same evidential propagation as would have to be performed by a centralized system. Of course, doing this propagation via inter-agent communication will be more costly both in terms of processing and elapsed time than it would be in a centralized system—that is the issue. The basic complexity of SI is determined by the interrelatedness/independence of the possible data interpretation hypotheses, though.

division of the data among the agents), such that $\hat{\mathcal{I}}^G(C^G_{f_d}) \neq \hat{\mathcal{I}}(C_{f_c})$ *and* $\hat{\mathcal{I}}^G(C^G_{f_d}) \neq \mathcal{I}(\mathcal{D})$. In other words, the *local solutions GSOUs resolution strategy* does not guarantee that the FA/C distributed system produces the same exact interpretation as the centralized system or the true MPE of all the globally available data.

Proof: Regardless of the local interpretation strategy being used, there must be some data and some distribution of that data so that there is some hypothesis h such that $h \in \hat{\mathcal{I}}^i(\mathcal{D}^i)$ and $h \in \hat{\mathcal{I}}^j(\mathcal{D}^j)$, but not $h \in \hat{\mathcal{I}}(C_{f_c})$. The only case in which this would not be true would be a domain in which there was complete solution monotonicity—i.e., $\forall D_i \subset \mathcal{D}$, if $h \in \mathcal{I}(D_i)$ then $h \in \mathcal{I}(\mathcal{D})$.

For example, it is entirely possible to have $P(H_a \mid D_1) > P(H_b \mid D_1)$ and $P(H_a \mid D_2) > P(H_b \mid D_2)$, but $P(H_a \mid D_1, D_2) < P(H_b \mid D_1, D_2)$ (where H_a and H_b are competing hypotheses and D_1 and D_2 are data sets in different agents). In other words, the solution that is locally most likely in both agents (H_a) may not be the globally most likely solution—even though the local solutions are consistent. This is simply a consequence of the nonmonotonicity inherent in reasoning based on incomplete and uncertain information.

7.3 Approximate Monotonicity

Because "consistency" of local solutions does not in general guarantee that the merged, global solution is the best solution, the *local solutions GSOUs resolution strategy* may not produce solutions that are equivalent to an approximate centralized system, let alone the MPE. This is unfortunate since this coordination strategy can be very efficient. It also conflicts with the intuitions of ourselves and other other FA/C researchers that this is a reasonable approximate FA/C coordination strategy.

We believe that the reason for this apparent contradiction is that many real world domains have characteristics that allow the this strategy to work quite well. For example, in vehicle monitoring, a considerable amount of evidence must typically be acquired to achieve a vehicle track hypothesis with high belief. While additional evidence can negatively affect the belief in such a track, it will take substantial additional evidence to have a major effect on the belief. In other words, while belief is nonmonotonic with increasing evidence, in some situations the effects of additional evidence are not totally unpredictable.

What this observation means for the "local solutions strategy" is that if a track hypothesis is part of a local solution, has "fairly high" belief, and is globally consistent, then it is pretty likely that the track is correct. We will refer to this characteristic as *approximate monotonicity*. We have formalized this notion in terms of probability distributions for the correct global probability of a hypothesis in terms of its belief based on local agent data only. A key focus of our future research will be to use this concept to design a coordination strategy whose effects are predictable in terms of the probability of its global solution being optimal.

7.4 Theorem 3

In this section and the next, we will consider the use of another approximate GSOUs resolution strategy, one that is *incremental* in the sense that the completeness with which the GSOUs are resolved is not prespecified but depends on the situation. The termination criteria will now be phrased in terms of whether a solution is *acceptable* or not. We define the predicate *acceptable*$(\mathcal{I}(C))$ to mean that $\forall h \in \mathcal{I}(C), \widehat{BEL}(h) \geq AT$, where AT is the *acceptance threshold* $(0 < AT \leq 1)$. Based on this definition we define the *incremental GSOUs resolution strategy* as: each agent will resolve the GSOUs as in the *local solutions GSOUs resolution strategy*, but if its resulting local solution is not *acceptable* it must continue to resolve global SOUs associated with the hypotheses in its solution until the solution is judged acceptable or all the relevant global SOUs have been resolved (i.e., as when using the *complete GSOUs resolution strategy*).

In this section we will assume that the agents use the *exact local interpretation strategy*. For the centralized case, this means that the final context C_{f_c} refers to the complete data set $\mathcal{D}$, and that $\hat{\mathcal{I}}(C_{f_c}) = \mathcal{I}(\mathcal{D})$. Because we are now considering a satisficing approach to solution selection, instead of comparing the centralized and distributed solutions based on whether they are idential, we will judge them based on whether they are *acceptable*.

Theorem 3: *Given the* exact local interpretation strategy *and the* incremental GSOUs resolution strategy, *$\forall \mathcal{D}$ if acceptable$(\hat{\mathcal{I}}(C_{f_c}))$ then acceptable$(\hat{\mathcal{I}}^G(C^G_{f_d}))$, where C_{f_c} and $C^G_{f_d}$ are final contexts that meet the specified centralized and distributed termination criteria for the $\mathcal{D}$, respectively.* In other words, if a centralized SI system is able to produce an acceptable solution for a data set then so should a distributed DRESUN system using the specified strategies.

Proof: Assume that the theorem is false. This means that even though *acceptable*$(\hat{\mathcal{I}}(C_{f_c}))$, $\exists h \in \hat{\mathcal{I}}^G(C^G_{f_d}))$ such that $\widehat{BEL}^G(h, C^G_{f_d}) < AT$. Now, because *acceptable*$(\hat{\mathcal{I}}(C_{f_c}))$, $\exists \{h_j\} \subset \hat{\mathcal{I}}(C_{f_c})$ such that *explains*(support-of**$(h), \{h_j\})$ and $\forall h_j \in \{h_j\}, BEL(h_j) \geq AT$. In other words, the centralized system has one or more hypotheses that are part of its solution, that have acceptable belief ratings, and that explain the data underlying the unacceptable hypothesis h in the global interpretation. Since these hypotheses are alternative explanations for *support-of**(h), each of these hypotheses would also have been created by at least one of the agents that produced h.[16] So, for

[16]Actually, the exact same hypothesis may not have been produced by a single agent since each of the agents may have only a portion of the data necessary to produce the version of the hypothesis with the most *complete* support possible. This issue arises because interpretation problems are not propositional, so as additional evidence is generated for a hypothesis the evidence both modifies the hypothesis' belief and refines the values of its attributes—creating

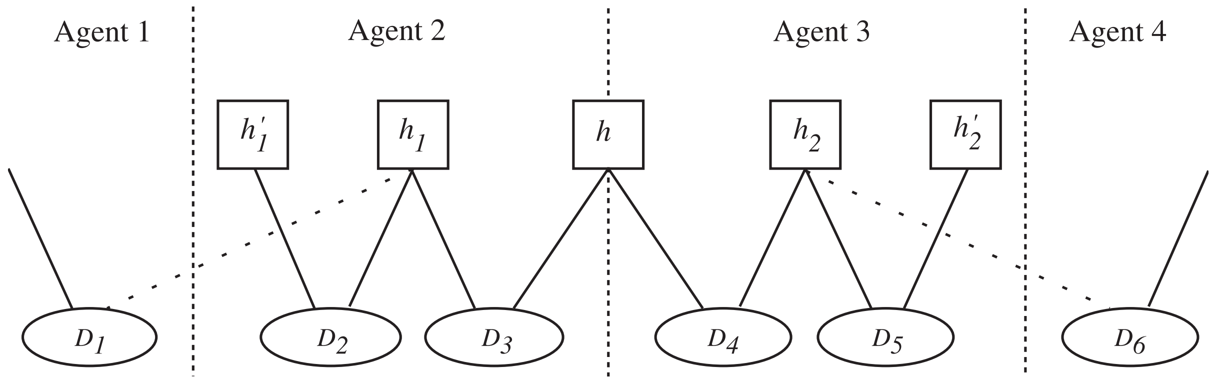

Figure 4: An example of the effect of partial resolution of the global SOUs.
Here, resolution of a GSOU btween agents 2 and 3 produced the hypothesis h, which resulted in $\{h, h'_1, h'_2\}$ being judged as an acceptable explanation for agents 2 and 3. Because of this, the GSOUs associated with h_1 and h_2 were not resolved. However, resolution of these GSOUs could result in corroborating evidence being obtained from agents 1 and 4, which might make $\{h_1, h_2\}$ the best explanation.

each h_j there is a corresponding hypothesis h_j^k that would have been created by A_k. Now, why is $\widehat{BEL}^k(h_j^k) < AT$ when $BEL(h_j) \geq AT$? There are three possibilities: (1) there exists data in one or more agents other than A_k that would corroborate h_j^k, but this data was not propagated to h_j^k;[17] (2) there exists data in one or more other agents that would contradict one or more alternatives to h_j^k; (3) both of the above may be true. Consider case 1: If there exists data in other agents that could support h_j^k, then A_k would have had GSOUs associated with h_j^k that identified the possible existence of this data (since GSOUs identify all interrelationships among the agents given the constructed hypotheses), but these SOUs were not resolved. This leads to a *contradiction* with our specified coordination strategy, however, since these SOUs should have been pursued when no acceptable explanation was found for the data common to h and h_j^k (remember that $h \in \hat{\mathcal{I}}^G(C_{fd}^G)$ but $\widehat{BEL}^G(h, C_{fd}^G) < AT$). Cases 2 and 3 lead to similar contradictions.

Thus, given the specified strategies if a centralized interpretation system can find an interpretation above a particular threshold then so can a DRESUN-based distributed interpretation system. This is a somewhat weak result, but it serves to establish a basic capability of the DRESUN architecture. Other FA/C distributed problem-solving architectures have not had this property because they have lacked a representation of the inter-agent constraints.

It is important to note that while we have shown that a distributed DRESUN system would find an "acceptable" solution, this interpretation may not be identical to that produced by the

new versions of the hypothesis (what [2] terms *extensions*). Thus, what we really mean is that a precursor, partial version of each alternative hypothesis would have been created by one or more of the agents. This is true because interpretation hypotheses can be created based on incomplete support, and our specified strategy was for each agent to create all the possible interpretations that are (even partially) supported by its own local data.

[17]The data cannot be in A_k since we have assumed complete intra-agent propagation.

centralized system and may not be the MPE of the global data set. Because an acceptable solution might be found without all the GSOUs having been resolved, there could have been incomplete evidence propagation among the agents. As a result, hypothesis belief ratings would be only approximations of the true conditional probabilities and this could lead to alternative explanations being (incorrectly) rated more highly—see Figure 4. This is the reason that the proof used $BEL(h)$ for the centralized system, but $\widehat{BEL}^G(h)$ for the distributed system. As Theorems 1 and 2 show, the only way to guarantee that a distributed system produces a solution that is *identical* to that of the centralized system (and is the MPE of the globally available data) is to require that all of the GSOUs associated with every hypothesis in each agent be completely resolved.

7.5 Theorem 4

In this section we will again examine the nature of global solutions when using the *incremental GSOUs resolution strategy.* However, unlike in Theorem 3, here we will assume that the agents use an *approximate local interpretation strategy.* This means that agents do not necessarily construct all possible interpretation hypotheses and so may compute only approximate belief ratings (conditional probabilities) for the hypotheses. Analyzing the solution quality given incomplete evidence propagation is complicated by the fact that a system's solutions can depend on the details of the strategies being used. We are not yet at the point where we can formally define and analyze the effects of different strategies. Thus, we will not specify a particular strategy for controlling evidence propagation here. Instead, we will simply assume that the same strategy is used in the single agent of the centralized case and in all of the agents of the distributed case. We will continue to assume that all of the data directly available to each agent is (at least partially) processed by that agent.

Theorem 4: *Given an* approximate local interpretation strategy *and the* incremental GSOUs resolution strategy*, it is not the case that,* $\forall \mathcal{D}$ *if* $acceptable(\hat{\mathcal{I}}(C_{f_c}))$ *then* $acceptable(\hat{\mathcal{I}}^G(C^G_{f_d}))$*, where* C_{f_c} *and* $C^G_{f_d}$ *are final contexts that meet the specified centralized and distributed termination criteria for the data set* $\mathcal{D}$*, respectively.* In other words, even if a centralized system is able to produce an acceptable solution for a data set, a distributed DRESUN system may not (even when the same local interpretation strategy is being used).

Proof: Assume that the theorem is false: $\forall \mathcal{D}$ $acceptable(\hat{\mathcal{I}}(C_{f_c}))$ implies $acceptable(\hat{\mathcal{I}}^G(C^G_{f_d}))$. It is easy to show the general existence of counter examples (the counter example is illustrated in Figure 5). Suppose that, the centralized system, which has a view of the complete data set $D_1 \cup D_2 \cup D_3 \cup D_4$, initially chooses to construct hypotheses h_1 and h_2. If this is an acceptable explanation ($\widehat{BEL}(h_1)$ and $\widehat{BEL}(h_2) \geq AT$) no further propagation may be done. In the distributed case, the agents might initially produce the hypotheses h_3 and h_4 (based on their own local data).

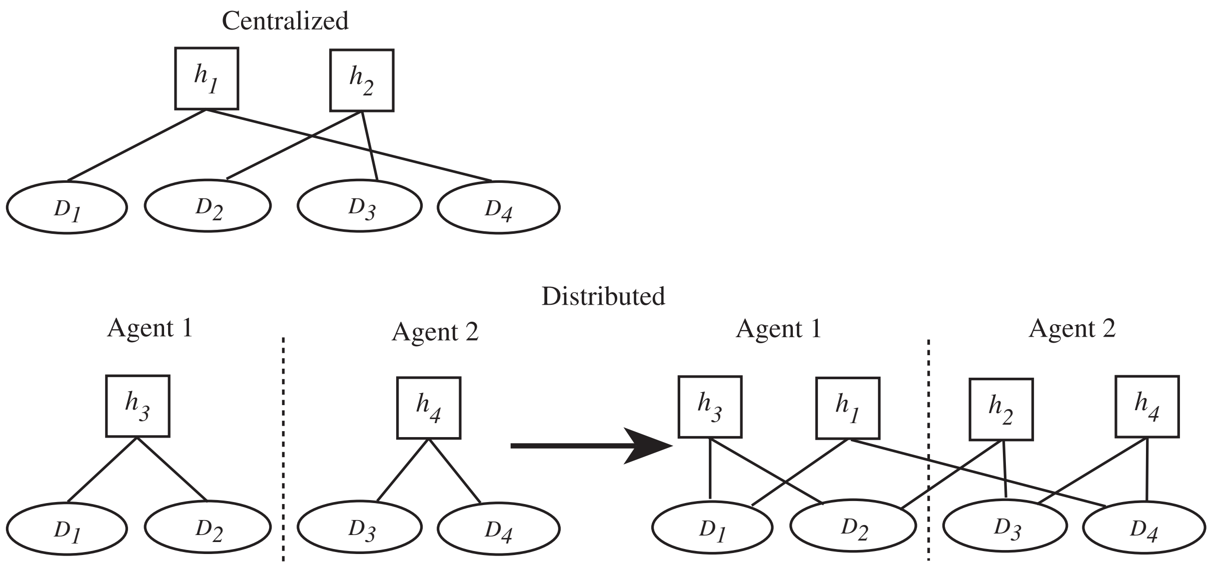

Figure 5: A counter example for the proof of theorem 4.

In general, there will certainly exist data scenarios for which this interpretation would not be acceptable. However, even if the agents communicate and construct h_1 and h_2, $\widehat{BEL}(h_1)$ and $\widehat{BEL}(h_2)$ may be different than the approximate values computed in the centralized case, and could in fact lead to this alternative interpretation also being judged unacceptable.

The reason that an acceptable centralized solution does not guarantee an acceptable distributed solution is that the division of data among the distributed agents may lead to the construction of a different set of hypotheses from those constructed by the centralized system. The construction of different hypotheses produces different belief approximations. This could cause the distributed system to produce a different acceptable solution or it could cause it to fail to find any acceptable solution. Even if the distributed system eventually constructs the same interpretation hypotheses as the centralized system, it may have done additional evidence propagation that causes the interpretation to be judged unacceptable. Furthermore, just because the distributed system may be driven to do more evidence propagation than the centralized system, this does not mean that its best solution is more likely to be the (true) global MPE (as long as its propagation is still incomplete).

7.6 Theorem 4 Bounds

Another way of stating theorem 4 is that when approximate GSOUs resolution strategies are used, there is merely some probability of a distributed system producing an acceptable solution when an acceptable solution is produced by an equivalent centralized system. This probability is a function of the interpretation and coordination strategies being used, the acceptance threshold, the characteristics of the domain, and the organization of agent interest areas. Understanding the effects that these factors have on the probability of achieveing comparable global solutions is a major focus of our future research. As a first step, we have explored how to decompose this probability so that we could determine bounds on its components based on reasonable probabilistic characterizations of a control strategy and domain.

So what we want to determine is: $P(acceptable(\hat{\mathcal{I}}^G(C^G_{f_d})) \mid acceptable(\hat{\mathcal{I}}(C_{f_c})))$, where C_{f_c} and $C^G_{f_d}$ are final contexts that meet the specified centralized and distributed termination criteria for the $\mathcal{D}$, respectively. (For simplicity, we will use $\hat{\mathcal{I}}^G$ for $\hat{\mathcal{I}}^G(C^G_{f_d})$ and $\hat{\mathcal{I}}^C$ for $\hat{\mathcal{I}}(C_{f_c})$ from this point on.) One way to decompose the probability is:

$$\begin{aligned}
&P(acceptable(\hat{\mathcal{I}}^G) \mid acceptable(\hat{\mathcal{I}}^C)) = \\
&\quad P(acceptable(\hat{\mathcal{I}}^G) \mid acceptable(\hat{\mathcal{I}}^C) \wedge \hat{\mathcal{I}}^G = \hat{\mathcal{I}}^C) * P(\hat{\mathcal{I}}^G = \hat{\mathcal{I}}^C \mid acceptable(\hat{\mathcal{I}}^C)) + \\
&\quad P(acceptable(\hat{\mathcal{I}}^G) \mid acceptable(\hat{\mathcal{I}}^C) \wedge \hat{\mathcal{I}}^G \neq \hat{\mathcal{I}}^C) * P(\hat{\mathcal{I}}^G \neq \hat{\mathcal{I}}^C \mid acceptable(\hat{\mathcal{I}}^C))
\end{aligned}$$

Here, we have decomposed the situation based on whether the distributed system finds that the most likely interpretation is the same interpretation as was found by the centralized system or not (remember that though the distributed system may have produced the same interpretation as the centralized system it may have done different propagation, so its belief ratings for the hypotheses of the interpretation may be different from those of the centralized system). Based on this decomposition, it is possible to identify some bounds on the component probabilities.

Consider the first conditional probability term in the decomposition:

$$\begin{aligned}
&P(acceptable(\hat{\mathcal{I}}^G) \mid acceptable(\hat{\mathcal{I}}^C) \wedge \hat{\mathcal{I}}^G = \hat{\mathcal{I}}^C) \geq \\
&\quad P(\forall h_i \in \hat{\mathcal{I}}^C, BEL(h_i) \geq AT \mid acceptable(\hat{\mathcal{I}}^C))
\end{aligned}$$

Proof: To prove this, we need to show that $acceptable(\hat{\mathcal{I}}^C) \wedge (\forall h_i \in \hat{\mathcal{I}}^C, BEL(h_i) \geq AT)$ implies $acceptable(\hat{\mathcal{I}}^G)$ (when $\hat{\mathcal{I}}^G = \hat{\mathcal{I}}^C$). Assume that this is false, so that $acceptable(\hat{\mathcal{I}}^C) \wedge (\forall h_i \in \hat{\mathcal{I}}^C, BEL(h_i) \geq AT) \wedge \neg acceptable(\hat{\mathcal{I}}^G)$. Based on the specified propagation strategy, $\neg acceptable(\hat{\mathcal{I}}^G)$ would cause the distributed system to do further evidence propagation until either $acceptable(\hat{\mathcal{I}}^G)$—which would contradict the assumption—or complete propagation has been done for all hypotheses in $\hat{\mathcal{I}}^G$ that are not acceptable—which also leads to a contradiction, as we will show. The complete propagation situation means that $\forall h_i \in \hat{\mathcal{I}}^G$ such that $\widehat{BEL}^G(h_i) < AT, \widehat{BEL}^G(h_i) = BEL^G(h_i)$. However, this contradicts the premise assumptions that $\forall h_i \in \hat{\mathcal{I}}^C, BEL(h_i) \geq AT$ since $\hat{\mathcal{I}}^G = \hat{\mathcal{I}}^C$ here.

$P(\forall h_i \in \hat{\mathcal{I}}^C, BEL(h_i) \geq AT \mid acceptable(\hat{\mathcal{I}}^C))$ is the probability that the hypotheses in the centralized interpretation are "acceptable" given these hypotheses true belief ratings whenever the partial propagation strategy finds that the hypotheses are acceptable given the approximate beliefs. This probability could be determined experimentally for domains in which complete propagation is at least possible (even if it is not possible for real-time problem solving), and could be experimentally estimated for domains in which complete propagation is absolutely intractable. It may also be possible to directly derive this probability from appropriate domain and strategy models. This would be easier than trying to derive $P(acceptable(\hat{\mathcal{I}}^G) \mid acceptable(\hat{\mathcal{I}}^C) \wedge \hat{\mathcal{I}}^G = \hat{\mathcal{I}}^C)$ since it would not require modeling the relationship between the organization of agent interest areas and the uncertainty/incompleteness of the locally available data, and how such conditions interact with a particular control strategy.

For the corresponding conditional probability term in which $\hat{\mathcal{I}}^G \neq \hat{\mathcal{I}}^C$, there is the following bound:

$$P(acceptable(\hat{\mathcal{I}}^G) \mid acceptable(\hat{\mathcal{I}}^C) \wedge \hat{\mathcal{I}}^G \neq \hat{\mathcal{I}}^C) \geq$$
$$P(\forall I_j (I_j \neq \hat{\mathcal{I}}^C) \Rightarrow (\forall h_i \in I_j, BEL(h_i) \geq AT) \mid acceptable(\hat{\mathcal{I}}^C) \wedge I_j \neq \hat{\mathcal{I}}^C)$$

Proof: The proof is analogous to the last proof.

$P(\forall I_j (I_j \neq \hat{\mathcal{I}}^C) \Rightarrow (\forall h_i \in I_j, BEL(h_i) \geq AT) \mid acceptable(\hat{\mathcal{I}}^C) \wedge I_j \neq \hat{\mathcal{I}}^C)$ is the probability that all interpretations of the data other than the most likely interpretation have hypotheses whose true belief ratings would make them "acceptable" interpretations, given that the most likely interpretation has been found to be acceptable using a partial propagation strategy. While this does provide a lower bound on the component conditional probability, it is a weak bound because it is unlikely to be very large for reasonable values of the acceptance threshold AT. In the future we will be working to improve this bound by exploring how different strategies affect the likelihood of locating acceptable interpretations when such interpretations exist. This will also assist in characterizing the relative likelihood that the distributed system will find the same or different interpretations than the centralized system—a key issue in assessing the target conditional probability.

8 Conclusions

This paper began with a substantial discussion of SI and the kinds of interpretation and coordination strategies that can be used for FA/C distributed SI. Several basic results about the quality of solutions that can be produced by an FA/C-based distributed SI system were then presented. First, while it is possible for a distributed SI system to produce the same solution as a centralized system (and the true MPE of the available data) this requires strategies that are typically impossible for real-world SI. Second, while coordination that focuses only on local solutions has been popular in FA/C systems and can be efficient, it is not guaranteed to produce centralized-equivalent or optimal solutions. Finally, if a satisficing approach to constructing solutions is adopted then an FA/C-based system can produce results that are "comparable" to those of a centralized system, but again only under conditions that are typically impractical.

Our future research will concentrate on pursuing the issues that appear in Sections 7.3 and 7.6. This is will involve characterizing the effects of practical approximate FA/C strategies for SI, based on probabilistic models of the domains. We believe that results based on the notion of approximate monotonicity will support some of the intuitions that people have had about approximate FA/C-based SI: certain types of hypotheses are better to construct and communicate, and it is useful to achieve a certain local level of belief in hypotheses before communicating them. In each of these cases we see that the likelihood of global solutions being correct when approximate strategies are being used will be improved.

As we pointed out in the introduction, different coordination strategies for how and when to communicate the information necessary to resolve a GSOU can greatly affect the efficiency of the process. This issue is related to the question of local and global strategy interactions. For example, suppose that an agent wants to resolve a certain GSOU, but the appropriate external agent has not yet processed (interpreted) the data and created the hypotheses necessary to do this. Should the initiating agent wait for the external agent to get around to doing this? Should it instead be sent the raw data and do with it what it wants, even if that might produce a different result than would have been produced in the first scenario? Since our focus in this paper is on solution quality that can attained by FA/C distributed SI systems, we tried to separate the effects of the local and global strategies as much as possible. In practice, however, their interactions can have a profound effect on efficiency and solutions. These are other issues that we will be pursuing as part of our efforts to analyze the FA/C model.

References

[1] Yaakov Bar-Shalom and Thomas Fortmann, *Tracking and Data Association*, Academic Press, 1988.

[2] Norman Carver and Victor Lesser, "A New Framework for Sensor Interpretation: Planning to Resolve Sources of Uncertainty," *Proceedings of AAAI-91*, 724–731, 1991.

[3] Norman Carver, Zarko Cvetanovic, and Victor Lesser, "Sophisticated Cooperation in FA/C Distributed Problem Solving Systems," *Proceedings of AAAI-91*, 191–198, 1991.

[4] Norman Carver, Victor Lesser, "The DRESUN Testbed for Research in FA/C Distributed Situation Assessment: Extensions to the Model of External Evidence," *Proceedings of the International Conference on Multiagent Systems*, June, 1995.

[5] William Clancey, "Heuristic Classification," *Artificial Intelligence*, vol. 27, 289–350, 1985.

[6] Keith Decker, and Victor Lesser, "An Approach to Analyzing the Need for Meta-Level Communication," *Proceedings of IJCAI-93*, 360–366, 1993.

[7] Keith Decker, and Victor Lesser, "Designing a Family of Coordination Algorithms," *Proceedings of the International Conference on Multiagent Systems*, June, 1995.

[8] Victor Lesser and Daniel Corkill, "Functionally Accurate, Cooperative Distributed Systems," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 11, no. 1, 81–96, 1981.

[9] Victor Lesser and Daniel Corkill, "The Distributed Vehicle Monitoring Testbed: A Tool for Investigating Distributed Problem Solving Networks," *AI Magazine*, vol. 4, no. 3, 15–33, 1983 (also in *Blackboard Systems*, Robert Engelmore and Tony Morgan, editors, Addison-Wesley, 1988).

[10] Victor Lesser, "A Retrospective View of FA/C Distributed Problem Solving," *IEEE Transactions on Systems, Man, and Cybernetics*, special issue on Distributed Artificial Intelligence, vol. 21, no. 6, 1347–1362, 1991.

[11] Ann Nicholson and Michael Brady, "Dynamic Belief Networks for Discrete Monitoring," *IEEE Transactions on Systems, Man, and Cybernetics*, special issue on Knowledge-Based Construction of Probabilistic and Decision Models, vol. 24, no. 11, 1593–1610, 1994.

[12] Judea Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, 1988.

[13] Yun Peng and James Reggia, *Abductive Inference Models for Diagnostic Problem-Solving*, Springer-Verlag, 1990.