Collaborative Decision Making and Intelligent Reasoning in Judge Advisor Systems

David C. Wilkins[1], Ole Mengshoel[1], Oleksandr Chernyshenko[1],
Patricia M. Jones[1], Caroline C. Hayes[2], Robin Bargar[1],

[1]Beckman Institute, University of Illinois, 405 N. Mathews Avenue, Urbana, IL 61801
[2]Dept. of Mechanical Engineering, University of Minnesota
dcw@uiuc.edu, p-jones5@uiuc.edu, hayes@lilac.me.umn.edu, r-bargar@uiuc.edu
chernysh@uiuc.edu, mengshoe@uiuc.edu

Abstract

This paper presents a judgement and decision making analysis of collaborative problem solving. This analysis is done with respect to the Raven and CoRaven decision-making tools for filtering, interpreting, and visualizing large amounts of uncertain data in the domain of intelligence analysis. Raven and CoRaven are multimodal advisory decision aids and their inferential reasoning is based on Bayesian networks. Human decision makers and information sources interact with these decision making systems in many ways, during their design, construction, refinement, and usage. This paper analyzes the collaborative aspects of the use of Raven and CoRaven using the Judge Advisor System model.

1. Introduction

Real-time decision-making that involves very large amounts of uncertain data is challenging for human decision makers. Filtering, interpreting, and visualizing massive large amounts of data is a central challenge in battlefield reasoning, especially in intelligence analysis. The Raven and CoRaven systems are currently under development to allow abstract representation of large amounts of uncertain data, and to facilitate collaborative problem solving.

This paper is organized as follows. In Section 2 a description is given of the Raven/CoRaven systems, which are multimodal advisory decision aids for making decisions under uncertainty from sensor data. In Section 3, the decision making cycle is described using the terminology of Judge Advisor Systems. In Section 4, it is shown how the Judge Advisor System (JAS) is relevant to an improved analysis, design, and understanding of the Raven and CoRaven systems.

There are two collaborative efforts that relate to CoRaven. The approach of [Jones, et al, 1998] begins with user studies, and based on an analysis of user behavior creates interactive tools that assist with collaboration, such as whiteboards, the ability to annotate visual displays with notes, and a range of editing tools. By contrast, the effort described in this paper is based on a judgement and decision making approach from psychology, in particular the Judge Advisor System model.

2.0 The Raven and CoRaven Systems

The purpose of Raven is to filter and interpret large amounts of uncertain data relating to battlefield reasoning using Bayesian networks and present the information to a decision maker via a multimodal interface. CoRaven ("Collaborative" Raven), extends Raven to include tools to assist with collaboration. As it is the more general system, we will use the term CoRaven throughout this paper, although the analysis and discussion applies to both Raven and CoRaven.

The rest of this section is organized as follows. First, Bayesian networks are briefly presented. Second, a scenario for military intelligence analysis is given. Third, we discuss how to formally represent this scenario as a Bayesian network which can be used for filtering, fusion, and intelligent selection of information. Finally, we conclude and outline directions for future research.

2.1 Bayesian Networks

Bayesian networks are used for reasoning and learning under uncertainty [Kim & Pearl, 1983] [Pearl, 1988] [Lauritzen & Spiegelhalter, 1988]. Bayesian networks are also known as causal networks, belief networks, or influence diagrams. Probability theory and graph theory form their basis: random variables are nodes and conditional dependencies are edges in a directed acyclic graph. Edges typically point from cause to effect. Temporal Bayesian networks can be used in dynamic environments [Mengshoel & Wilkins, 1997a].

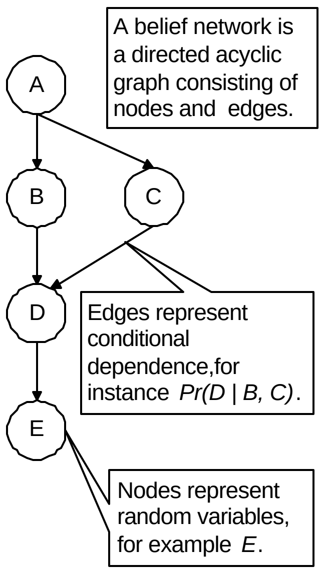

Figure 1: Example Bayesian network.

A simple Bayesian network consisting of the five nodes *A*, *B*, *C*, *D*, and *E* is shown in Figure 1. In addition to the graph, there are conditional probability tables associated with each node *V* and its parents *pa(V)*, expressing the conditional probability $Pr(\mid pa(V))$. For Figure 1, assuming discrete binary nodes with values {0,1}, $Pr(D=0 \mid B=1, C=0))$ is one of the entries in *D*'s conditional probability table.

The inference task of belief updating amounts to the following: Given evidence at node *E* and query node *Q*, infer posterior probability $Pr(Q \mid E=e_i)$. Any nodes in the network can be evidence or query nodes. For the example BN, this leads to different types of inference: (i) diagnostic – as in $Pr(A \mid E=e_i)$; (ii) causal - as in $Pr(E \mid A=a_j)$, and (iii) mixed – as in $Pr(D \mid E=e_i, A=a_j)$. A variety of approaches to Bayesian network inference have been investigated [Kim & Pearl, 1983] [Pearl, 1988]. Much of the Raven project focues on improved methods of inference [refs], but this subject is outside the scope of the current paper.

2.2 Domain Scenario: Intelligence Analysis

The purpose of intelligence analysis is to answer priority intelligence requests (PIRs). There is a doctrinal one to one mapping between PIRs and decisions that are executed during battle, so PIRs play a central role in battlefield reasoning. The central component of the intelligence analysis task is diagnosis (or abduction), where the PIRs are diagnoses and intelligence reports are symptoms. Intelligence analysis takes as input reports concerning enemy activity, and an intelligence analysis BN can be used to perform this filtering task automatically. In particular, consider the situation shown in Figure 2. The highest abstraction level is shown to the left, the lower abstraction to the right.

The battlefield reasoning challenge is twofold: First, sensor data is low-level, high-volume, and uncertain. Second, model knowledge is uncertain and incomplete. The Bayesian network (BN) approach, as illustrated in Figure 3, is to perform data fusion and filtering using belief network and output and visualize the results.

We will distinguish between two classes of scenarios. In enemy defensive scenarios, the enemy is defending, friendly is attacking. Here, a standard (atemporal) belief network is used to model the enemy and to filter uncertain information. In enemy offensive scenarios, the enemy is attacking, friendly is defending. Temporal belief networks are used to model the enemy and to filter uncertain information for these scenarios. In the following we will focus on enemy defensive scenarios, however the approach can easily be extended to enemy offensive scenarios.

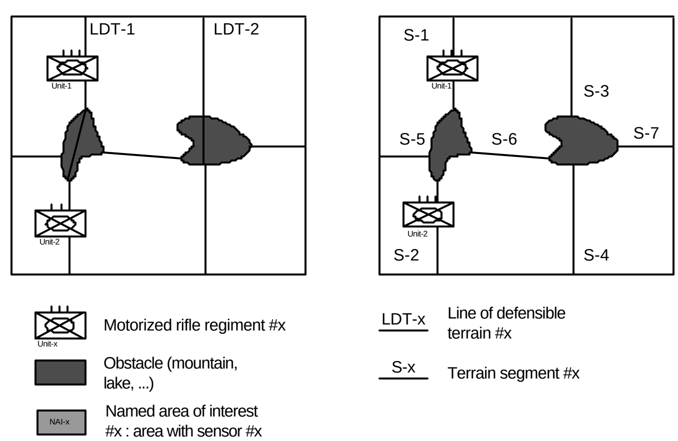

Figure 2: Mpas at different levels of abstraction for intelligence scenario analysis.

A partial Bayesian network model for battlefield reasoning is shown in Figure 4. These networks consists of three main layers. First, the diagnosis layer corresponds to priority intelligence requests (PIRs). Second, the filtering layer corresponds to specific intelligence requests (SIRs). Third, the message layer corresponds to observable SIRs.

2.3 Representation Using Bayesian Networks

The battlefield reasoning challenge is twofold: First, sensor data is low-level, high-volume, and uncertain. Second, model knowledge is uncertain and incomplete. The Bayesian network approach, as illustrated in Figure 3, is to perform data fusion and filtering using a Bayesian network and output and visualize the results.

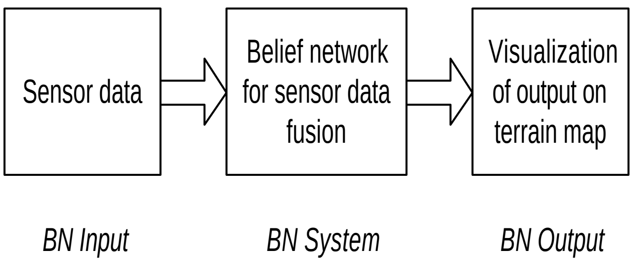

Figure 3. Bayesian network (BN) system for filtering battlefield data.

We will distinguish between two classes of scenarios. In defensive enemy scenarios, the enemy is defending, friendly is attacking. The scenario presented above is of this type. Here, a non-temporal belief network is used to model the enemy and to filter uncertain information. In offensive enemy scenarios, the enemy is attacking, friendly is defending. Temporal belief networks are used to model the enemy and to filter uncertain information in these situations [Mengshoel & Wilkins, 1997a]. In the following we will focus on the defensive enemy scenario from the previous section, however the approach generalizes to other scenarios.

Figure 1 Overview of the Decision Making Situation

A partial Bayesian network model for battlefield reasoning, the RAVEN BN, is shown in Figure 4. This network consists of three main layers, which can be mapped to the military intelligence community's terminology [Schlabach, 1997]. Starting from the top of Figure 4, the layers are as follows. First, the diagnosis layer corresponds to priority intelligence requests (PIRs). Second, the filtering layer corresponds to specific intelligence requests (SIRs). Third, the message layer corresponds to observable SIRs. Also notice the correspondence to the three levels of abstraction visualized in Figure 2.

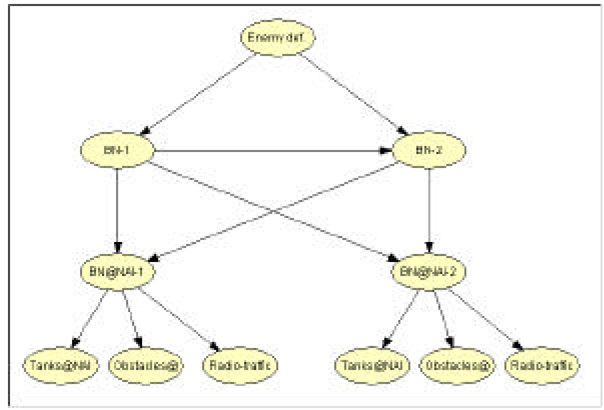

Figure 4: Bayesian network for filtering battlefield data.

Due to space limitations, the conditional probability tables of the RAVEN BN are not shown here. The states of the nodes are presented in Figure 5 – the numbers in this figure may be ignored for now. The PIR in this BN is the node Enemy-def., which ranges over states {LDT-1, LDT-2}. This node describes the enemy's two possible LDTs – see Figure 2. The nodes BN-1 and BN-2 in the filtering layer model where the two battalions are, so they both range over the line segments {S-1, S-2, S-3, S-4}. The nodes BN@NAI-1 and BN@NAI-2 describe the allocation of enemy BNs to map locations, and in particular to NAIs. Finally, the six nodes in the message layer represent whether certain signatures are or are not observed in a particular NAI. For example, Radio-traffic@NAI-1 has states {Yes, No} and represents whether or not there is radio traffic in NAI-1.

2.4 Inference Using Bayesian Networks

The BN model described above can be used to draw probabilistic inferences. First, consider the probabilities (or confidences) over the node Enemy-def. based on prior knowledge - before any messages arrive. LDT-1 is the most probable line of defense, as can be seen to the left in Figure 5. Note that probabilities are given in percentages rather than on the usual [0,1] scale. So based on prior knowledge, there is an 80% chance that Enemy def. is LDT-1, while there is a 20% chance that it is LDT-2.

Now suppose three messages arrive: "No obstacles at NAI-1", "Radio traffic at NAI-1", and "No tanks at NAI-1". To reflect these messages, the corresponding three nodes are clamped to the appropriate values as shown to the right in Figure 5. Because of these updates of the BN, the inference procedure updates the probabilities over the other nodes in the network as shown to the right in Figure 5. In particular, the Enemy-def. state LDT-2 is now the most probable, with 60.42% chance. The chance of LDT-1 is 39.58%. This is consistent with the intuition that not seeing tanks or obstacles in any NAI on LDT-1 makes LDT-2 more probable - see Figure 2.

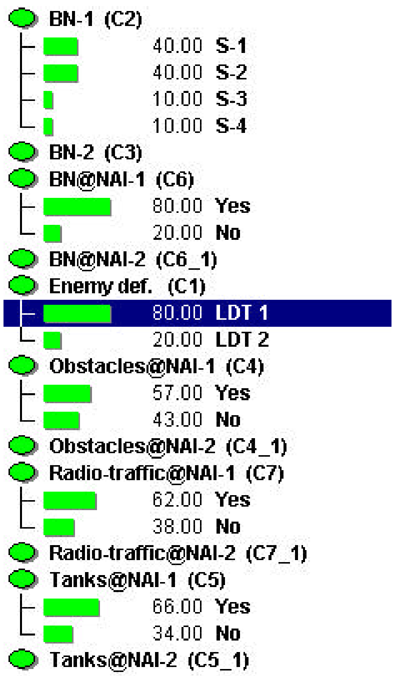

Figure 5. Probabilities over Raven Bayesian Network after processing three messagse.

The inference algorithm of the BN tool used above, HUGIN, is an exact algorithm known as clustering [Lauritzen & Spiegelhalter, 1988]. For sparse BNs this algorithm works well, however for large and strongly connected BNs it doesn't. For this reason, a heuristic approach to belief revision in BNs is investigated [Mengshoel, 1997] [Mengshoel & Wilkins, 1998]. More specifically, we assume that the GA fitness function is a joint probability density function represented as a BN. This is a restriction on the fitness function, but probability theory in general and BNs in particular have proven sufficiently rich to make this an interesting restriction. Using a GA has three advantages: simplicity, robustness, and efficiency. First, a GA is a relatively simple algorithm. Second, the search provided by a GA should give robust inference across a spectrum of BNs. Third, there might be BNs where the GA is more efficient than other inference algorithms, in particular compared to stochastic simulation algorithms when there is low-probability evidence [Welch, 1996].

3.0 The Battlefield Domain

In the battlefield, the commanding officer is responsible for choosing a course of action. Traditionally the commanding officer has made decisions in the battlefield by convening one or more advisors on location. Thus the natural structure of the decision making unit involves multiple persons in different roles.

The environment of the commanding officer is a dynamic, turbulent one with great uncertainties about the enemy's troops, location, activities, and plan. To plan and execute a course of action, the commanding officer needs to reduce these uncertainties. It is the function of intelligence analysts to learn the information requirements of the commanding officer, develop a plan to obtain that information, and monitor information is acquired. This function is ongoing over the course of battlefield operations. Because data collection resources are limited, it is necessary to schedule collection efforts over requirements of varying priority. This first stage of intelligence gathering leads to the accumulation of a large set of data that is difficult to use.

One officer has responsibility for organizing the information and making inferences based on it for the commander's operational decision making. Given that the data form multiple layers of probabilistic linkages between initial observations and conclusions, the reasoning task is onerous. Human judgment under uncertainty in general, and probabilistic reasoning in particular are susceptible to numerous limitations and biases [Tversky & Kahneman, 1973]. The challenge in managing the information is made more difficult by the fact that information may be acquired, adjusted, or processed by other advisors in parallel.

3.1 The Roles of Judges and Advisors in CoRaven

The Judge Advisor System can be used in research on decision making with a user and an interactive intelligent system. Unlike other generations of decision aids, an intelligent interactive system allows for bi-directional, dynamic influence between the two entities comprising the decision making body. Others have shown the advantages of considering advanced components such as intelligent agents as a members of a human team [Bui & Sankaran, 1997; Sycara, Lewis, Lenox, & Roberts, 1998].

In this section, we describe the complete decision-making cycle, and the various roles played by judges and advisors in this decision making cycle. All judges and advisors are human; the Raven and CoRaven systems provide a basis for them to share information and communicate information related to the decision-making task.

A decision maker interacts with CoRaven through its four multimodal interfaces. These are: (1) the Raven

Netviewer, which provides a graphical user interface for a Bayesian Belief Network; (2) a VSS Sonification View of the world, wherein the patterns in the network are communicated to the user using sound; and GIS MapView, that allows the user to visualize the areas of interest on a terrain map, and lastly a collaborative viewer under development, which has the goal of allowing users to share their views of the evolving state of the world, and to communicate these views.

Step 1. **Judge:** decide which decisions need to be made (e.g., defend a particular hill on the East side or West side)

Step 2. **Judge:** specify a set of factual data requests relevant to the decisions enumerated in Step 1 (e.g., estimated number of trucks within 1 mile of the particular hill)

Step 3a. **Sensor Advisors**: collect and synthesize data from a variety of sources. Supply probabilities on likelihood that data is correct (e.g., estimated number of trucks within 1 mile of hill based on satellite data). The probabilities don't have to be numerical; a set of linguistic constructs can be used (e.g, confidence in answer is definite, high, medium, or low).

Step 3b. **Structure Advisors**: for each decision that needs to be made, one or more structure advisors construct a Bayesian belief network that relates the data relevant to the decision. This may involve retrieving a previously constructed BBN from a library of BBNs, modifying an existing BBN that is similar in terms of inputs and outputs to that specified by the judge, or constructing a BBN from scratch. Using collaborative editing tools, different advisors may construct different parts of a single BBN, or there may be a one-to-one correspondence between structure advisors and the BBNs that need to be constructed.

Step 3c. **Likelihood Advisors**: for each arc in each Bayesian Belief Network, specify a probabilistic strength. For each node that has more than one arc, specify the conditional probabilities for the set of arcs.

Step 4c. **Interpretation Advisors:** there may be one or more advisers who interpret the network with the judge. This may involve assisting with changing sliders that increase or decrease the strength of the probabilistic links between arcs in the BBN. They changes may, for example, be based on a judge's past experience with the reliability of different sensor advisors or sensor types. Another role of the interpretation advisers are to decide whether the incomplete data received is strong enough to reach a overall decision.

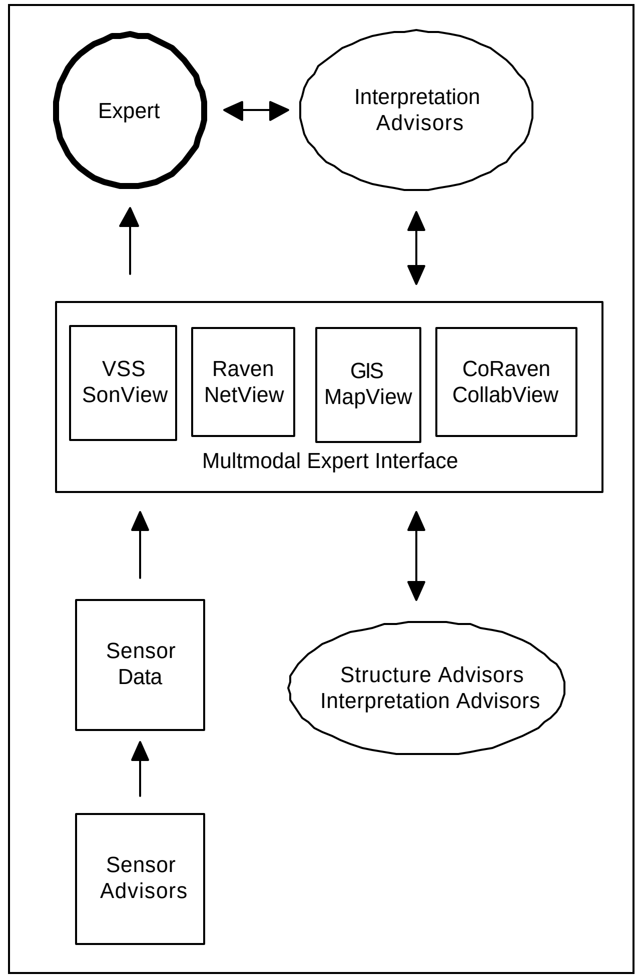

Figure 6. Overview of the Decision Making Situation

Step 4. **Judge**: for each BBN, monitor the BBN in real-time to see the evolving answers to the questions. The BBN is monitored through a visual display of the BBN (called an evidence tree), through a terrain map that represents the answers as a probabilistic footprint, and through sonification that changes a Judge's focus of attention, with respect to which BBN is being monitored, or which parts of a BBN are being monitored. Summarize the answers for the ultimate judge, namely the field commander.

Due to the way that BBNs work, the BBN will have an answer to the decision before any sensor data is input, based on the *a priori* probabilities associated with the BBN. The BBN dynamically updates itself in real-time as each piece of sensor data is received. Thus the judge always has the BBNs tentative answers to the decisions that need to be made based on the data received to date.

8. Evaluation

As is necessary for any system designed for a well-defined class of users, there is ongoing evaluation of CoRaven system components by experts in military intelligence. Design modifications routinely follow. As system development progresses, it will be possible to obtain data and feedback from members of the user population as they use it interactively in real time. Scientific assessment of user reactions is desirable (e.g., Hikz & Johnson, 1989; Mahmood & Sniezek, 1989), but not sufficient. To satisfactorily demonstrate the effectiveness of CoRaven for the purpose for which it is intended, it is necessary to statistically analyze samples of judgments and decisions made with and without the individual and combined features of intelligent collaborative support technology.

The first formal evaluation of CoRAVEN has been the validation of the actual BNs. The initial probabilities were obtained from a domain expert via unstructered knowledge elicitation techniques. It is well known that subjective probabilities generated by experts in a wide variety of domains have some features that are undesirable for purposes of computation in a structure such as the Bayesian belief net. For example, the probabilities assigned to represent one's belief in the accuracy of one's choices or judgments are too high given accuracy levels determined by objective criteria (Lichtenstein, Camerer, 1981). Before proceeding to use probabilities elicited from experts it is wise to demonstrate acceptable levels of reliability, validity, and calibration bias. We are obtaining probabilities for the Bayesian network structures used in CoRAVEN from multiple domain experts, and assessing their inter-rater agreement. In addition, probabilities for choice accuracy are being elicited for a series of choice items from manuals that document current doctrine in the experts' domain of military intelligence. Initial results indicate problems concerning reliability, validity, and calibration that are typical of experts in most domains. Statistical adjustment procedures are being implemented to attenuate these problems.

9.0 Research Questions within the Judge Advisor System Paradigm

There are four major categories of research questions. Each is described below, with examples of questions that can be investigated with this or similar collaborative intelligent systems.

9.1 Judge Advisor Compatibility

Questions of this type center around how the factors concerning the Judge or Advisor and their context affect system effectiveness.

(a) Do the Judge and designer of the Raven Advisor have compatible goals? Identical goals? Designer goal is to choose from small set of discrete alternatives and determine probabilities. Judge goal is to implement decision or act on one of the alternatives. Goal compatibility cannot be taken for granted in JASs (Jungerman, 1999; Sniezek, Heath, & Van Swol, 1998).

(b) Do the Judge and Advisors have the same mental model? How closely does the BBN authored in response to the Judge's information needs match the question in the mind of the Judge? The theory of Shared Mental Models (Cannon-Bowers, Salas, Converse, 1993) assumes that team decision making will be enhanced if members have the same mental representation of the problem and the roles held by each member.

9.2 Selection of Advisors and Advice

(a) Given a decision problem, i.e., question, how does the Judge determine what information sources (sensory Advisors) are needed? Is actual selection affected by perceived validity more than need? Judges have been found to discriminate varying degree of quality in Advisors (Hedlund, Ilgen, & Hollenbeck, 1998). What is the influence of Advisor consistency or calibration on selection? High confidence can make one an attractive Advisor to an uncertain Judge, even if one has knowledge inferior to tht of the Judge (Pritchard & Sniezek, 1992). In addition to raising one's chances of being consulted, high confidence can increase one's influence (Zarnoth & Sniezek, 1996).

(b) Are the Judge's questions or framings of the decision problems constrained by available Advisors, or the characteristics of their advice? That is, does the Judge ask the questions the Judge knows can be addressed with available Sensor, Structure, and Likelihood Advisors, instead of asking a more relevant question that requires a search for an additional Advisor?

9.3 Influences on Advice Quality

(a) How can Advisors be encouraged to contribute their unique advice in lieu of repeating advice given by all Advisors? The failure for uniquely held information to be used in group decision making has been well documented. (Stasser & Stewart, 1985). Although the relative contribution of unqiue to common information has been found to be better in JASs than in groups (Savadori, Van Swol, Sniezek, 1997), much unique information is withheld, or mentioned briefly without influencing the final decision.

(b) Do the Likelihood Advisors (who supply conditional probability estimates etc. once the BBN is authored) agree with each other? Are they consistent? Well calibrated? Experts in many domains are known to have inconsistency in their judgments from probabilistic data (Camerer, 1981), and probabilities more extreme that warranted by objective criteria (Lichtenstein, Fischhoff, & Phillips,1982).

(c) Do Advisors modify their information, recommendations, or qualifications to reflect their own personal risk preferences?

(d) Should Likelihood Advisors communicate before setting the likelihoods? Or only after, to discuss and resolve differences of opinion? If so, is it better if they communicate electronically than face-to-face? A similar process to cueing of Judges occurs for all members of a group if independent individual assessments are not made prior to discussion (Sniezek & Henry, 1990). Estimates that are generated independently will generally have greater variance. And greater variance has been associated with greater improvements in judgment quality over the level of the average group member (Sniezek & Henry, 1989; Valacich, Mennecke, Wachter, & Wheeler, 1993).

9.4 Judge's Use of Advice

(a) How does the Judge interpret advice qualification? For a given probability from the BBN; is Judge's subjective probability lower, higher, or the same? Human decision makers often adjust results from models and human advisors prior to using them in the decision process.

(b) What leads the Judge to mistrust, trust, or overtrust individual Advisors, or the BBN Advisor? Consistency? Variability? Accuracy? Confidence? Explanation has been found to increase trust in expert systems (Ye & Johnson, 1995).

(c) How does the Judge use the sliders to adjust input from Advisors? What factors lead to increases or decreases in probabilities? How sensitive is the BBN to the adjustments? I.e., are adjustments too small to change conclusions or sufficiently large to allow a single subjective assessment by the Judge to overrule the BBN?

(d) Does communication of uncertainty in the form of probabilities interfere with or facilitate choice? Implementation of choice?

(e) What is the optimal timing of advice? Should it precede or follow tentative appraisals by the Judge? "Cueing" of a judge by directing attention to an Advisor's recommendation before independent appraisal by the Judge may depress Judge accuracy while elevating Judge confidence (Sniezek, Paese, & Switzer, 1990; Sniezek & Buckley, 1995). If advice is available to the Judge at intermediate stages, does it alter the Judge's acceptance of the final advice? That is, are there primacy effects of viewing early "conclusions" by the BBN?

The quality of advisor input will depend on many factors, such as the nature of the task [cf., Benbasat & Lim, 1993; McLeod, 1992]. Compared to computer-mediated communications, face-to-face interactions have been shown to lead to higher quality advice by subordinates in teams [Hedlund, Ilgen, & Hollenbeck, 1998]. Empirical studies making direct comparisons in Judge Advisor Systems are few; thus we must draw on the group literature. Computer-mediated communication has been found to be beneficial in tasks requiring greater diversity of input, such as brainstorming [Nunamaker,]. Presumably this occurs because members initially approach the task independently, avoiding being cued or anchored by the judgments of others (Sniezek & Buckley, 1995; Sniezek, 1992). But contribution rates are far less, which can adversely affect the overall quality of the advice available to the Judge. The quality of advice obtained from an expert system will generally be superior, due in part to the opportunity to evaluate and select input from experts based on quality.

Computer-mediated communications have clear advantages over face-to-face interactions in terms of the cost and speed of acquiring advice. (Although computer-mediated groups may spend more time on the task, cf., Fjermestad & Hiltz, 1997, it is far easier to convene them over large geographic distances.) Except for the initial high cost of developing an expert system, it is even easier to use for advice acquisition. These considerations are particularly important for decisions made under time pressure.

One of the challenges for face-to-face and distributed JASs is the development of effective work relationships. For JASs that face problems with high stakes and severe time pressure, typical developmental patterns may not apply [Gersick, 1988]. If JASs form and perform in a brief time period, the opportunity to learn about Advisors over time is likely to be inhibited [Gabarro, 1990]. Learning to work in distributed teams via collaborative technology is not the same as learning to work in face-to-face groups [Knoll & Jarvenpaa, 1995]. Familiarity with electronic technology may not be as important as experience with work relationships [Iacono & Weisband, 1997]. One of the requirements for JAS interactions that enhance decision making is the development of sufficient interpersonal trust. That is, for Judges to take actions based in part on their Advisors' input, they must trust these Advisors. Trust is likely to be facilitated to the extent that the Judge is dependent on an Advisor [Sniezek & Van Swol, in review]. It follows that differential expertise or access to information sources will be associated with greater trust. Although many of the problems in human relationships are avoided with the use of an intelligent system as an Advisor, there remains the serious problem of promoting appropriate trust in the technology.

The activities involving both Judge and one or more Advisors fall into five categories: (a) solicitation of advice (e.g., revision of a likelihood based on background research) by the Judge, (b) offering of input (intelligence information from the Advisor's source) by the Advisor, (c) delegation of a task (e.g., determination of vehicle capabilities given novel obstacles) to an Advisor by the Judge, (d) sharing of information (e.g., terrain maps) between Judge and Advisor, and (e) sharing of a task (e.g., setting alarms) by both Judge and Advisor.

It is true that the sheer number of advisors can be increased with the opportunity for collaboration with remote as well as near advisors. But more is not necessarily better; too many advisors can contribute to the problem of information overload [Hiftz & Turoff, 1985.]. One important implication follows from the fact that the criteria for advisor selection can now shift. Several factors that can be used in determining whom a Judge can consult. In the absence of technology for remote collaboration, the selection set is restricted, and cost criteria such as ease of access dominate. With collaboration, benefit criteria become more important. The Judge can use various aspects of the Advisors (e.g., their reputations, relative success rates), their relationships with the Advisors (e.g., trust, shared values), or the advice itself (e.g., detail of explanation, qualification).

There are potential consequences of collaborative technology for the environment as well. Not only does the Judge have greater access to a wider variety of Advisors, but the actual set of Advisors may change. One possibility is increased specialization of expertise among individual advisors as a result of timely communication and coordination among the multiple Advisors and the Judge. Given that heterogeneity in groups is often associated with greater performance gains over individuals [Sniezek & Henry, 1989; Valacich, Mennecke, Wachter, & Wheeler, 1993], increased specialization of Advisors may be beneficial to the Judge's decision making.

8. Conclusions

This objective of this project is to improve decision making in this setting by means of intelligent reasoning and collaborative technology to support the commanding officer. The goal is to create a system that reasons over large amounts of data, and allows for participation by numerous users.

Judge Advisor Systems for battlefield decision making are distributed, and faced with crucial, high risk decisions under time pressure in an extremely uncertain environment. With the implementation of the technology developed in this project, their activities will include synchronous and asynchronous communication, task delegation and task sharing, sharing of visual information, and sharing of the products of an intelligent reasoning system as it is updated. Unique characteristics of this project are a collaborative graphical user interface to support the activities of the Judge and Advisors, and an intelligent system that uses Bayesian belief networks to structure inferences from observable data as they become available.

The support technology of interest is intended to support the reasoning process and multiparticipant synchronous and asynchronous interactions. Consequently, battlefield collaborative technology can allow for fewer personnel in harm's way in the battlefield itself. Further, the greater portability afforded by replacement of decision making materials such as maps by electronic technology provides those who are in the field with greater mobility and protection from enemy detection. In addition to reducing these risks, inefficiencies, and difficulties of battlefield decision making, intelligent, collaborative support technology is appealing because it offers the potential to improve decision quality. No longer is solicitation of advice restricted to those personnel who are physically present and able to function. It now becomes possible to include advisors any distance away in any time zone in addition to those who are available for face-to-face, radio, or phone consultation. And the capacity to organize and structure information helps with problems of information overload [Hiftz & Turoff, 1985]. Finally, the fact that the system can perform intelligent reasoning operations on the information overcomes many limits to human information processing capabilities. This is especially important in aiding human judgments with probabilistic data [Sniezek & Reeves, 1984, 1986].

The CoRAVEN project seeks to develop a proof-of-concept tool to support battlefield decision making by

Judge Advisor Systems. Judges have access to multiple visual displays, showing spatial data (maps), temporal data (the *synchronization matrix* showing scheduling of collection assets) and graph-based models for fusing evidence (Bayesian networks). These multiple views can be coordinated by a single user, and shared with users. Individual situation awareness and team mental models are facilitated by auditory displays resulting from sonification techniques.

Previous collaborative technology projects have achieved dual goals. For example, Rae, Suresh, and Turoff (1998) report a system to enhancing both group communication and individual reasoning in medical decision making.

The CoRaven system is still under development; thus, no satisfactory effectiveness data are available. However, it is expected that the advising and collaboration capacities of the system will made battlefield decision making less risky, more efficient, and better informed.

Acknowledgements

The research was supported in part by ARL Grant DAAL01-96-2-0003, ONR Grant N00014-95-1-0749, and NRL Contract N00014-97-C-2061. We are grateful for the domain assistance provided by Maj. J. Schlabach and Michael Barnes. Our thanks to Jenet Sniezek for detailed comments on earlier drafts of this paper. Other contributors to the Raven and CoRaven projects include Peter Assaro, Insok Choi, David Fried, Song Han, Debra Kessler, Tim Harahan, Aaron Flish, and Ryan Fox.

References

Bargar, R. and Choi, I. (1998). Sonification of probabilistic belief networks for a distributed environment. *Proceedings of the IEEE 1998 International Conference on Systems, Man, and Cybernetics*, to appear.

Bargar, R., I. Choi, S. Das, and C. Goudeseune. (1994). Model-based interactive sound for an immersive virtual environment. *Proceedings of the 1994 International Computer Music Conference*, 471-474. Aarhus, Denmark: International Computer Music Assn.

Benbasat, I. and Lim, L. H., "The Effects of Group, Task, Context, and Technology Variables on the Usefulness of Group Support Systems: A Mets-Aualysis of Experimental Studies, Small Group Research, 24,4, 1993,430-462.

Bui, T. & Sankaran, S. "Group Decision and Negotiation Support in Telemedicine: An Application of Intelligent Mobile Agents as Non-Human Teleworkers." Proceedings of the 30th HICSS Conference, 1997.

DeSanctis, G. and Gallupe, R. B., "A Foundation for the Study of Group Decision Support Systems," Management Science, 33,5,1987,589-609.

Fjermestad, J., & Hiltz, S. R. (1997). Experimental Studies of Group Decision Support Systems: An Assessment of Variables Studied and Methodology. Proceedings of the 30th HICSS Conference. 1997.

Gabarro, J.J. The Development of Working Relationships. In J. Galegher, R.E. Kraut and C. Egido (eds.), Intellectual Teamwork (pp.79-l 10). Hillsdale, NJ: Erlbaum, 1990.

Gersick, C. Time and Transition in Work Teams: Toward a New Model of Group Development. Academy of Management Journal, 1988, 3 l(l), pp. 9-41.

Hiltz, S. R., Johnson. K., "Measuring Acceptance of Computer-Mediated Communication Systems", Journal of the American Society for Information Science, Vol. 40, No. 6, pp. 386-397, 1989.

Hiltz, S. R., Turoff, M. "Structuring Computer-Mediated Communication Systems to Avoid Information Overload", Communications of the ACM, Vol. 28, No.7, pp. 682-689, July 1985.

Iacono, S. & Weisband, S. "Developing Trust in Virtual Teams"Proceedings of the 30th Annual Hawaii International Conference on Systems Sciences, 1997.

João, H., Legatheaux, J., & Paulo, J. A Flexible Object-Group-Oriented Framework to support Large Scale Collaborative Applications. 1998.

-Jones, P.M., Hayes, C., Wilkins, D. C. , Bargar, R. , Sniezek, J.A., Asaro, P. , Mengshoel, O., Kessler, D. , Lucenti, M., Choi, I., Tu, N. & Schlabach Maj. J. "CoRAVEN: Modeling and Design of a Multimedia Intelligent Infrastructure for Collaborative Intelligence Analysis."

Knoll, K. and Jarvenpaa, S.L. Learning to Work in Distributed Global Teams. Proceedings of the 28th HICSS Conference, 1995, pp.92- 101.

Mahmood, M.A., & Sniezek, J.A. (1989). Defining decision support systems: An empirical assessment of end-user satisfaction. INFOR: Canadian Journal of Operations Research and Information Processing, 27 (3), 253-271.

McGrath, J.E. and Hollingshead, A. B., Groups Interacting with Technology, Thousand Oaks, CA: Sage Publications, 1994.

McLeod, P. , "An Assessment of the Experimental Literature on the Electronic Support of Group Work: Results of a Meta-Analysis, " Human Computer Interaction, 7,3, 1992,257-280.

Mengshoel, O. J. and Wilkins, D. C. (1998a). Abstraction for Belief Revision: Using a Genetic Algorithm to Compute the Most Probable Explanation. *Proc. AAAI Spring Symposium Series on Satisficing Models*, March 1998, Stanford, CA.

Mengshoel, O. J. & Wilkins, D. C. (1998b). Genetic Algorithms for Belief network inference: The Role of Scaling and Niching. *Proc. Seventh Annual Conference on Evolutionary Programming*, March 1998, San Diego, CA.

Rae, G.R., Suresh, B. A., and Turoff, M. "A Group Decision Support System Framework for Medical Decision Making Incorporating Cognitive-Aid Structures and Cognitive Appropriation." 1998.

Sniezek, J. A. (1992). Groups under uncertainty: An examination of confidence in group decision making. In James H. Davis (Ed.) special issue of Group Decision Making for *Organizational Behavior and Human Decision Processes*, 52, 124-155.

Sniezek, J. A. and Buckley, T. (1995). Choice accuracy and confidence in Judge Advisor Decision Making Systems. Organizational Behavior and Human Decision Processes, 62(2), 159-174.

Sniezek, J.A. , & Henry, R. (1989). Accuracy and confidence in group judgment. *Organizational Behavior and Human Decision Processes*, 43(1),1-28.

Sniezek, J.A., & Reeves, A. (1984). Aiding decision making with aggregate information in a laboratory study. Proceedings of the Hawaii International Conference on Systems Sciences, January, 603-608.

Sniezek, J.A. & Reeves, A.P. (1986). Feature cues in probability learning: Database information and judgment. *Organizational Behavior and Human Decision Processes*, 37, 297-315.

Sniezek, J. A. and Van Swol, L. (In review). Trust and the resource exchange in Judge Advisor Systems. *Organizational Behavior and Human Decision Processes*.

Sycara, K.P., Lewis, M., Lenox, T., & Roberts, L. Calibrating trust to integrate intelligent agents into human teams. Proceedings of the 31st HICCS Conference, 1998.

Tversky, A. and Kahnernan, D. “Judgment under Uncertainly: Heuristics and Biases”, Science, Vol. 185, pp. 1124-1131, September 1974.

Valacich, J.S., B.E. Mennecke, R. Wachter, and B.C. Wheeler, “Computer-Mediated Idea Generation: The Effects of Group Size and Group Heterogeneity,” Proceedings of the 26th Hawaii International Conference on Systems Science, IV, (1993), 152-160.

DeSanctis, G. and Gallupe, R. B., “A Foundation for the Study of Group Decision Support Systems,” Management Science, 33, 5, 1987, 589-609.

McGrath, J.E. and Hollingshead, A. B., Groups Interacting with Technology, Thousand Oaks, CA: Sage Publications, 1994.

-Nunamaker, Jr., J. F., Dennis, A. R., Valacich, J. S., Vogel, D.R. and George, J.F. “Electronic Meeting Systems to Support Group Work”, Communications of the ACM, Vol. 34, No. 7, pp. 40-61, Jrdy 1991.

Hiltz, S.R., Dufner, D,, Holmes, M. and Poole, S. “Distributed Group Support Systems: Social Dynamics and Design Dilemmas”, Journal of Organizational Computing, Vol. 2, No 1, pp. 135-159, 1991.

McGrath, J.E. “Groups: Interaction and Performance”, Prentice Hall, Englewood Cliffs, NJ. 1984.

Turoff, M. and Hikz, S.R. “Computer Support for Group Versus Individual Decisions”, IEEE Transactions On Communications, VoLCom-30, No. 1, January 1982.

-Paese, P. W. and Sniezek, J. A. (1991). Influences on the appropriateness of confidence in judgments of performance: practice, information, and decision making. Organizational Behavior and Human Decision Processes, 48(1), 100-130.

Trafimow, D., and Sniezek, J. A. (1994). Perceived expertise and its effect on confidence. *Organizational Behavior and Human Decision Processes*. 57 (2), 290-302.

Kuhn, K. M., & Sniezek, J. A. (1996). Differential effects of scenario generation on confidence and uncertainty. Journal of Behavioral Decision Making. Vol. 9, 231-247.

Cooper, R. S. and Sniezek, J. A. Consensus and Accountability in the Judge Advisor *System*. Annual meeting of the Judgment and Decision Making Society, Los Angeles, November, 1995.

Kessler, D. & Sniezek, J. A. Expert Probability Judgments and CoRaven: Application of Psychological Theory to Battlefield Decision Aid. Fourth Conference on Naturalistic Decision Making. The Airlie Center, Warrenton, VA. May, 1998.

Pritchard, E. T. and Sniezek, J. A. Effects of consulting advisors on choice and confidence. Annual meeting of the Midwestern Psychological Society, Chicago, May,1992.

Lichtenstien, S., Fischhoff, B. & Phillips, L.D. (1982). Calibration of probabilitiles: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases. Cambridge: Cambridge University Press.

Savadori, L., Van Swol, L. M. , & Sniezek, J. A. (1997, November). Insights from process data. In J. A. Sniezek (Chair), Information flow in groups. Symposium conducted at the meeting of the Judgment and Decision-making Society, Philadelphia, PA.

Sniezek, J.A. , & Henry, R. (1990). Revision, weighting, and commitment in consensus group judgment. Organizational Behavior and Human Decision Processes, 45(1)

Sniezek, J. A., Paese, P. W., & Switzer, F. S. III. (1990). The effect of choosing on confidence in choice. Organizational Behavior and Human Decision Processes, 46, 264-282.

Ye, R. L. & Johnson, P. E. (1995). The impact of explanation facilities on user acceptance of expert systems advice. *MIS Quarterly*, 157-172.