PERGAMON

Safety Science 35 (2000) 193–208

SAFETY SCIENCE

www.elsevier.com/locate/ssci

Timely knowledge elicitation from geographically separate, mobile experts during emergency response

D. Mendonça [a,*], R. Rush [b], W.A. Wallace [a]

[a] *Department of Decision Sciences and Engineering Systems, Rensselaer Polytechnic Institute, Troy, New York, USA*

[b] *Queues Enforth Development, Cambridge, Massachusetts, USA*

Abstract

Two crucial factors for effective emergency response are the speed with which the response strategy is implemented and the quality of the expert knowledge on which the response is based. We propose a method for using communication and computing technologies for eliciting and aggregating the knowledge of multiple, geographically separate experts that improves our ability to address these two issues. In our methodology, we send a description of the problem to the experts who in turn submit graphical representations of the responses they propose for addressing it. The representations are then aggregated into a hypothesized central representation. A cycle of voting then begins from which emerges either a consensus problem representation or a diagnosis of impasse. The result is a clear definition of the response problem for resolution. Results from an implementation of the algorithm over the Internet are presented. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Knowledge elicitation; Influence diagram; Emergency response; Network inference

1. Introduction and background

The creation of an effective response strategy is key to controlling harmful effects of unforeseen disasters or unexpected public policy crises. Two factors are particularly important in determining effectiveness: (1) the speed with which the response strategy is implemented and (2) the quality of the expert knowledge on which the response strategy is based. Unfortunately, these two factors often exist at cross-purposes, insofar as addressing the second factor may require amalgamating the input of geographically separate experts. Consider, for example, the *Exxon Valdez*

* Corresponding author.

PII: S0925-7535(00)00031-X

oil spill (Harrald et al., 1990). Clearly, immediate action was warranted. In the face of this urgency, the time available for assembling proven experts to determine the appropriate combination of mechanical and chemical damage control methods was limited.

In this paper we propose a means of resolving this dilemma. Specifically, we describe a methodology for incorporating the knowledge of multiple experts into a single coherent representation of the disaster without physically bringing the experts together. Decreased reaction time is the obvious benefit. The success of the methodology requires that two criteria be satisfied. First, we suggest that the relevant experts must have access to communication technology that allows them to transmit and receive text, image, and voice data (Katz, 1995). Second, we must be able to depict the knowledge of a single expert as a directed graph in which the nodes represent problem components and the arcs influencing relationships, i.e. an influence diagram. These being satisfied, the methodology provides a structured vehicle for electronic voting and collaboration that allows the experts to build consensus on the required response without having to convene physically.

The elicitation of knowledge from a group of physically separate experts requires an algorithm that can gather and summarize this knowledge in a systematic way to provide the basis for further action. In the remainder of this section, we provide the background necessary to construct such an algorithm and a review of the relevant literature on aggregating knowledge in the form of directed graphs and influence diagrams. In Section 2 we present the details of our methodology. As stated above, the algorithm allows experts to contribute knowledge, evaluate other experts' ideas, and finally decide on a consensus representation of the problem situation. In Section 3, we describe an implementation of the algorithm designed to test the methodology. We employed the Oil Wildcatter problem (Clemen, 1996) as the basis of this test. Section 4 discusses the performance of the methodology in this experiment. Section 5 offers concluding remarks and possible directions for further research.

1.1. Eliciting expert knowledge

In the scenario under consideration, geographically separate experts must be consulted as part of the decision-making process. These experts are a potentially rich source of insight into problems (Massey and Wallace, 1991). When these problems are ill-structured, the need to access and organize this knowledge is especially acute (Massey and Wallace, 1996). Because expert knowledge is voluminous and sometimes difficult to cohere, structured methods for eliciting it have been developed (see Boose, 1986, for a review). Any such method of knowledge elicitation must: (1) allow each expert to submit ideas on the problem's structure and (2) yield a logical and defensible encapsulation of these submissions (Rush and Wallace, 1997). Graphical representations, particularly directed graphs (Robinson and Foulds, 1980), not only satisfy these two criteria but also are apt vehicles for communicating ideas to other experts (Howard, 1989). They have been used to elicit knowledge in a variety of domains, including emergency management (Harrald et al., 1990) and policy development (Massey and Wallace, 1996), and are widely accepted in the decision analysis literature (e.g. Shachter, 1988; Howard, 1989; Browne et al., 1997). We next describe in greater detail the structure and use of influence diagrams.

1.2. Influence diagrams

An influence diagram captures relationships among problem components. A problem component is defined as a decision, chance event or value of an outcome. In an influence diagram, a problem component is depicted as a labeled node with a shape specific to its type (Clemen, 1996). The relationship between any two components is represented by a directed arc: the direction of the arc indicates the direction of influence. A directed cycle is "a sequence of nodes that we could visit by moving along the arcs in the proper direction" but which "starts and ends with the same node" (Shachter, 1988). Shachter's definition of an influence diagram, which we adopt, is as follows: "An influence diagram is a network consisting of a directed graph with no directed cycles and detailed data stored within the nodes of the graph" (Shachter, 1988). This restriction on directed cycles precludes the occurrence of instances of indeterminate causality, such as occur when, for three concepts A, B and C, the influence diagram indicates that A causes B, B causes C and C causes A (Black, 1963).

The development of an influence diagram typically consists of three stages: first, a graphical representation of the problem is constructed in which nodes represent concepts of the problem and arcs represent relationships among nodes; second, node type (i.e. decision, chance, deterministic or value) is specified; third, measures of uncertainty and values of outcomes are incorporated (McGovern et al., 1994). The algorithm proposed in this paper addresses the first stage of this development.

An equivalent representation of an influence diagram is given by its adjacency matrix. For each element a_{ij} in the adjacency matrix A, the element a_{ij} is 1 if i influences j and is 0 otherwise. Clearly, the influence diagram and its adjacency matrix are interchangeable: our algorithm employs both.

To summarize, the influence diagram represents one individual's representation of the problem under consideration. A typical first step in the construction of the diagram is the elicitation of key concepts from an individual, after which the influence diagram is expressed as a directed graph.

1.3. Aggregating expert knowledge

We address the aggregation of knowledge from geographically separate experts and describe a method for deciding whether any one diagram may serve as the consensus problem representation.

An expert's influence diagram represents a preferred arrangement of problem components (Goddard, 1983). The problem of aggregating influence diagrams has been addressed by Rush and Wallace (1997), who extended work by Banks and Carley (1994). The method is useful because it allows us to draw statistically defensible inferences from the aggregate diagram — a necessary tool in diagnosing consensus. We will now review these statistical properties (see Rush and Wallace, 1997, and Banks and Carley, 1994, for a complete discussion).

Let M be the set of components in the problem description, with $|M|$ the cardinality of M. The difference between any two graphs (or influence diagrams) g_1 and g_2 is given by:

$$\mathrm{d}(g_1, g_2) = \mathrm{tr}\lfloor (A_1 - A_2)^T (A_1 - A_2) \rfloor, \tag{1}$$

where A_1 and A_2 are the graphs' corresponding adjacency matrices. The result is the number of discrepant edges between the two graphs. Based on this difference equation, the following probability measure is defined:

$$\mathrm{P}_{G,s}(g) = c(s) \times \exp\{-s \times d(g, G)\}, \tag{2}$$

where g is an arbitrary member of the space of all networks, G is the central network of the distribution (described below), s is the dispersion parameter, and $c(s)$ is a normalizing constant. G is analogous to the mean of a distribution; s is a measure of spread.

Construction of the maximum likelihood estimate G^* for the central network G is governed by the so-called majority rule: For N experts, let $g_1, g_2, \ldots, g_N$ be the experts' respective influence diagrams. If an edge (i.e. a connected pair of nodes) is present in more than 50% of the influence diagrams, include it in G^*. The resultant graph is denoted the Multiple Expert Influence Diagram (MEID). The maximum likelihood estimate s^* for the dispersion parameter s is calculated as:

$$s^* = -1\mathrm{n}\left\{\frac{(r \times N)^{-1} Z}{1 - (r \times |M|)^{-1} Z}\right\}, \tag{3}$$

where

$$Z = \sum_{j=1}^{N} d(g_j, G^*) \quad \text{and} \quad r = |M| \times (|M| - 1).$$

We employ a non-parametric bootstrap to construct a confidence interval for G. The bootstrap sample is taken from the set of diagrams under consideration; the sample distance is the average distance from the elements of the bootstrap sample to their MEID. Any influence diagram falling within this confidence interval is a legitimate expression of the problem situation. We therefore adopt the following as our criteria for consensus: If all influence diagrams in the current set of influence diagrams are within the confidence interval for G, consensus has been achieved. Otherwise, consensus has not been achieved.

1.4. Achieving consensus or diagnosing impasse

The method for knowledge elicitation reviewed above offers a means of constructing a statistically defensible central representation but does not address what action to take when a sufficient degree of consensus is absent. A mechanism is required by which experts can express their preferences across graphs. With such a mechanism, experts could more clearly indicate which diagram should be included in the estimation of G and whether the process might reach consensus after some time. Such a mechanism, based on work by Borda (1781) and Kendall (1962), follows.

Assuming that consensus is not reached, the following elimination procedure is implemented. At the beginning of the procedure, influence diagrams outside the confidence interval for the current MEID are eliminated and each expert is asked to rank the remaining diagrams excluding his or her own and including the MEID. A given candidate's total Kendall score, a modification of Borda's marks (Borda, 1781), is the sum of these points across all voters. Candidates are then clustered by Kendall score, where the distance between any two candidates is defined as the absolute value of the difference in their respective Kendall scores.

The set of diagrams in the cluster of most-preferred diagrams then forms the basis for the construction of the next MEID. If the diagrams fall within the confidence interval for the new MEID, the process halts with a consensus representation; otherwise, it continues. Because we require either rapid achievement of consensus or quick diagnosis of impasse, we limit the number of iterations for the process.

The following section describes the complete algorithm for the scenario in which multiple, geographically separate experts must formulate a problem representation in a timely fashion.

2. Description of algorithm

Assume that the knowledge of some number N of experts in the field must be elicited in order to construct a consensus problem representation. During execution of the algorithm, communication with the experts is managed by coordinators at a single location. Coordinators receive information from the experts and intervene where specified. The algorithm has three stages, corresponding to the tasks of concept identification, construction of individual influence diagrams and either achievement of consensus or diagnosis of impasse. These are described next.

2.1. Stage 1: identify problem components and initialize

The N experts are first provided with a problem description. The coordinator sets the maximum number of iterations for the elimination process to R and decides on a confidence level α for the MEID confidence interval described earlier.

An expert i extracts a set of concepts M_i from M. The coordinator then generates the set $P = M_1 \cup M_2 \cup \ldots \cup M_N$. (Note that P is a subset of M.) The set P is then sent to the experts, along with the problem statement and instructions for constructing an influence diagram.

2.2. Stage 2: construct the individual influence diagrams

Let G_P denote the set of all possible influence diagrams constructed from the elements of P, subject to the restriction that there be no directed cycle in any influence diagram. The ith expert then employs some choice function on G_P to produce an influence diagram g_i. Let the set of all distinct influence diagrams in use during the rth stage of the process be W_r (note that the elements of this set may change during the algorithm). Again, the only restriction on any g_i is that it does not include a directed cycle.

2.3. Stage 3: check for consensus or declare impasse

The MEID G_r^* at round r ($r = 1, 2, \ldots, R$) is constructed by applying the majority rule on W_r. The G_r^* diagram will then have some number of arcs $m^* \leqslant m$. The coordinator constructs the MEID's $100(1-\alpha)\%$ confidence interval using the bootstrap procedure. Denote the confidence interval of the MEID at round r by CONF_r. If the distances between each graph and the MEID are within CONF_r, then the procedure stops with consensus. Otherwise, the procedure moves to round $r+1$, as the coordinator initiates the following voting procedure to define W_{r+1}.

Let C represent the set comprised of the MEID and all graphs in the MEID's confidence interval. Each expert receives all elements in the set excluding his/her own, then ranks these graphs according to preference (Black, 1963). The result is a set of rankings that is then sent to the coordinator. The Kendall score τ_j for diagram j is the sum of the ranks across all experts for that diagram (Borda, 1781):

$$\tau_j = \sum_{i \in C}^{N} \pi_{ij}.$$

The distance between the scores for any two diagrams is defined as $d_{ij} = |\tau_i - \tau_j|$. The scores are clustered to maximize inter-cluster homogeneity and intra-cluster homogeneity, via the single linkage method (Johnson and Wichern, 1992). W_{r+1} is set equal to the set of diagrams in the top cluster, and the process repeats.

Finally, we note that impasse is declared if consensus has not been reached after the Rth round of voting.

3. Implementation of the algorithm

We implemented the algorithm over the Internet to (1) assess its feasibility (e.g. could it be implemented with existing technology), (2) identify opportunities for improving its efficiency (e.g. through software development or algorithm re-design), and (3) determine whether it provides participants with a satisfactory means of considering other participants' perceptions and communicating their own understanding of the problem.

Some general restrictions were applied at each stage. The only form of communication available to participants and coordinators was electronic mail ("email"), and all email was time-stamped on transmission and receipt.

3.1. Preliminary stage

After a trial run with a single participant, in which opportunities for making the process more efficient were identified, we solicited participation from nine individuals, five of whom agreed to participate. All participants were pursuing or had recently completed Master's level study in industrial engineering, statistics or a related field. During execution of the algorithm, the participants were situated in different locations and had no direct communication with each other.

As a preliminary task, participants were required to use a World Wide Web browser to access the graph-drawing program Visualizing Graphs with Java (VGJ) (Barowski et al., 1997). With VGJ, they were to draw a particular configuration of three nodes and two arcs, then mail the text representation of that configuration to the coordinators. Finally, participants were sent an email message containing an image of the configuration and asked to reply to the message if they were not able to read it. No difficulties were reported in using browsers, VGJ or email clients.

3.2. Implementation

On the day of implementation, participants were reminded of the time their task would begin, its approximate duration (2 h) and that commencement of the task would be signaled via email.

3.2.1. Stage 1: identify problem components and initialize

To begin, participants received email directing them to a web site that contained a statement of the Oil Wildcatter problem (adopted from Langel and Kann, 1992):

> Your organization, an oil exploration company, is facing a problem that must be solved in the next two hours. The organization is considering drilling a well. Because of budget constraints, only one well can be drilled. The organization is not entirely sure how much it will cost to drill the well, and, of course, is not sure whether it will find oil. In addition, the organization has the option of conducting one of two tests to determine the geologic structure of the drilling site. One of the tests is more expensive, but provides more useful and reliable information. The organization's only costs for this project are the costs to test and drill, and the only revenues for the project are the revenues received from selling the oil the organization finds. What should the organization do?

They also received a request that they submit a list of "elements of the problem" which they thought "should be considered in solving it" and a brief description of each concept. They were given 20 min for this task. Mean time to respond was approximately 18 min and average number of problem components submitted was 4.8.

In the implementation of the algorithm, participants were asked "which elements of the problem do you think should be considered in solving it?" The concepts as elicited from the participants and summarized by the coordinators are shown in Table 1. This task took approximately 25 min. The summarized concepts were

Table 1
Concepts and corresponding node labels sent to participants

Node label	Concept
Node 0	Geologic structure
Node 1	Amount of oil
Node 2	Revenue
Node 3	Should we do test 1?
Node 4	Should we do test 2?
Node 5	Should we drill?
Node 6	Cost of test 1
Node 7	Cost of test 2
Node 8	Cost of drilling
Node 9	Profit
Node 10	Company
Node 11	Results of test 1
Node 12	Results of test 2

entered into the VGJ interface and the program recompiled. Participants were then told to point their browser at a certain web page where they were asked to construct an influence diagram using any of the 13 concepts.

Given the time taken to complete the trial run of the algorithm, the maximum number of iterations was set to four and a significance level α equal to 0.05 was chosen.

3.2.2. Stage 2: construct the individual influence diagrams

Participants constructed their individual influence diagrams using the concepts from Table 1 they deemed "important." (As an example, Participant 5's graph is shown in Fig. 1; the remaining graphs appear in the Appendix as Figs. A1–A4.) Two participants experienced some difficulty at this stage. One could not see the complete list of concepts. Upon her request, the complete list was emailed to her. The other emailed an incomplete diagram. Once notified of his error, he emailed the complete diagram. No submitted graphs contained cycles, so all were admissible. The graphs themselves were considerably more complex than the submitted concepts might have indicated. Indeed, on average a participant used twice as many concepts in his or her diagrams as had been submitted in Stage 1. Graphs on average had 14.8 arcs. The average time required to complete an influence diagram was approximately 25 min.

3.2.3. Stage 3: diagnose consensus or declare impasse

The current MEID G_1^* and its associated confidence interval were computed, as were the distances of individual graphs to the MEID. The graph G_1^* at this stage was, compared to the rich graphs provided by the experts, a skeletal structure that identified some core concepts and relations (Fig. 2). The graphs g_1 and g_3 fell well outside the confidence interval, mainly because they included far more edges than did the MEID. The voting cycle therefore began with the set W consisting of g_2, g_4, g_5 and G_1^*. The correct subsets of W_r were then sent to participants for ranking. These tasks required a significant portion of the coordinators' time (approximately 1 h).

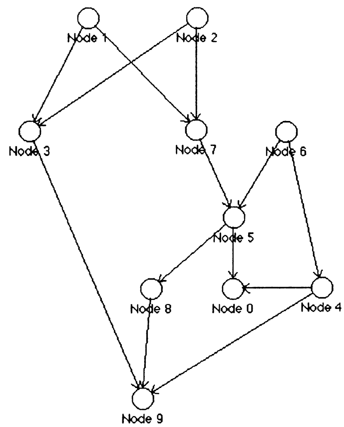

Fig. 1. Oil Wildcatter influence diagram submitted by Participant 5.

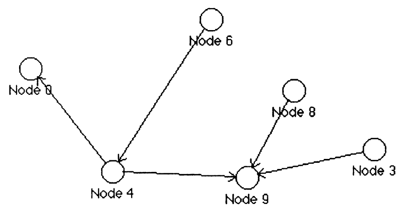

Fig. 2. G_1^*, MEID from Round 1.

Four sets of rankings were submitted to the coordinators (one participant failed to respond). The rankings revealed two clusters, one containing g_1, g_3 and g_4, the other containing G_1^*. The respective Kendall scores were 8, 7, 6 and 16, indicating that G_1^* was the least preferred alternative. This counter-intuitive result might be due, in part, to the rule governing the MEID's construction. As discussed in Section 4, exploration of alternative rules for the MEID's construction is an open area of research.

At this point, the experiment had nearly gone over time and the coordinators asked the participants to stay online an additional hour. Four of the five agreed to do so. The next MEID (i.e. G_2^*) and its associated confidence interval were computed from the top-ranked cluster. Obviously, because fewer graphs were involved in the computation of the MEID, the MEID itself was more complex. Only one graph (g_2) was contained in the confidence interval, though the others (g_2, g_4, g_5) were fairly close (distances from MEID of 2, 9 and 9, respectively). Coordinators' time on task was approximately 40 min.

Participants were asked to rank only the graphs g_2 and G_2^*. One of the four remaining participants was unable to complete the experiment due to schedule conflict. The other participants required an average of 10 min to complete the ranking. Based on these results, G_2^*, shown in Fig. 3, was ranked as the top graph and the process was halted with consensus. The total time for algorithm execution was approximately 3 h.

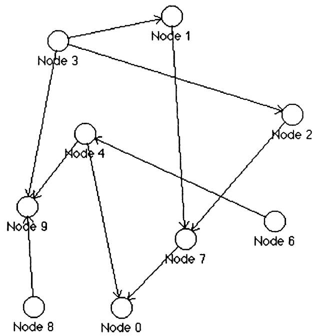

Fig. 3. G_2^*, MEID from Round 2.

Finally, participants were sent a brief questionnaire, the results of which are summarized, along with coordinators' observations on the implementation, in the following section.

4. Assessment of implementation

The questionnaire submitted to the participants was intended to help determine whether the algorithm had fulfilled the purpose given in Section 3. Participants were asked to assess the method, technology, and result of the implementation, per the following questions:

1. Please comment on the method employed for constructing the problem representation:
   Stage One: Developing the List of Concepts
   Stage Two: Developing Problem Representation
   Stage Three: Ranking the Graphical Representations
2. Did you experience any difficulties in using the software (browser, graph drawing or email software)?
3. Graphical Representation of the Problem
   The consensus graphical representation is given below (see Figure 3). Are there any concepts not present in the final representation that you feel should be present? If so, what are these concepts and how should they be related to each other and/or other concepts in the final representation? Also, are there any relationships that should be added to the final representation?

4.1. Assessment of the method

Participants were first asked for each stage to comment on "the method employed for constructing the problem representation". Their comments indicate that some additional consideration should be given to modifying the software interface and to improving communication between coordinators and participants.

Stage 1 was designed to elicit important concepts from participants and to summarize these concepts for use in Stage 2. Some participants seemed to have different perceptions of the summarized concepts. For example, one participant interpreted the concept "geologic structure" as the actual physical characteristics of the site; another interpreted it as the company's prediction of the physical characteristics based on the result of either test. Providing participants with a brief description of each concept may help reduce these discrepancies.

At Stage 2, participants often omitted certain concepts in their influence diagrams. Without additional information, it must be assumed that they thought these concepts irrelevant to the problem. Such behavior could, in certain cases, lead to undesirable outcomes. For example, if a simple majority agrees that concept A influences concept B, yet the minority is completely silent on the issue, that edge would be included. Though some techniques for analyzing such incompletely specified preferences have been developed (e.g. Critchlow, 1985), this work has yet to be considered in the present context. It should also be noted that cases of ties would, of course, not be included in the MEID — even if every expert registered an opinion. A fuller investigation of these challenges with respect to influence diagram aggregation is an open area of research.

Some participants were unsure about the exact nature of their task. One participant thought that she erred by representing "the entire domain relationship" rather than "the solution to our problem", when in actuality she performed the task correctly. Another said "I did not know if I was supposed to include nodes for all of the concepts listed or just the ones (he) thought were important". However, he also noted "This is probably my fault". A clearer definition of the requisite tasks may be warranted.

Participants' opinions on the difficulty of tasks for Stage 3 were varied. One participant thought that ranking the candidate graphs was "easy". Two participants had difficulty completing the stage because nodes were labeled with numbers rather than key words. Also, the problem description, at least in later stages, could only be seen by reading the web page from the experiment's first stage. Some simple modifications at the software interface should help alleviate these difficulties.

A clearer announcement of stage commencement and termination might have reduced confusion among participants. Based on their comments, some were unsure whether the algorithm had terminated or whether additional tasks remained. Similarly, others felt they had spent too much time waiting for the coordinators. It seems likely that, in future implementations of the algorithm, experts will in fact be occupied with tasks to accomplish during waiting periods.

4.2. Assessment of the technology

Participants reported some difficulties in using VGJ to construct the representations, but these were resolved by coordinators via email. No difficulties were reported in using browsers or email clients.

Coordinators required a considerable amount of time to accomplish their tasks during Stage 3, mainly because many different types of software had to be used to process incoming and outgoing correspondence. For example, upon coordinators' receipt of the textual representations of the graphs, it was necessary to translate the text into an adjacency matrix for subsequent processing by a Visual Basic program then by VGJ. To reduce coordinators' time on this task, VGJ might be programmed to accept adjacency matrices as input and give them as output, thus greatly reducing coordinators' effort (and possible error) in transcribing these matrices. Some additional efficiencies might be achieved by integrating the various functions performed by coordinators into a single software package.

In sum, the algorithm as witnessed by participants progressed slowly but straightforwardly, though communication might be clarified. To coordinators, some conceptually simple tasks required considerable time. Some possible improvements in efficiency, such as software redevelopment, are readily identifiable.

4.3. Graphical representation of the problem

In the final set of questions, participants were asked to comment on the final (i.e. consensus) graphical representation of the problem. If they thought some important concepts were missing, they were asked to describe them and specify how they should be included in the final representation. Similarly, they were asked to specify any relationships they thought were missing.

Participants made some salient observations which, in part, we share. Some diagrams clearly differ, in significant ways, from our conception of the problem. As one participant noted, "geologic structure" was "at the end of an edge, indicating that the item leading to it influenced it. I don't see how the cost of a test could influence the geologic structure, but I do see how the geologic structure could influence the cost of a test". A possible reason why someone might draw such a diagram was suggested by another participant who said, "there is a huge hole in the final graph (i.e. G_2^*). It does not represent the effect of the tests (or test results) on the amount of oil found". In both cases, the source of disagreement seems to be grounded in participants' understanding of the concepts: for some, "geologic structure" meant "our knowledge of the geologic structure"; for others, it meant "the geologic structure as it exists". Some further clarification of the concepts may have alleviated this difficulty, but semantic incongruities of this type are pervasive in problem structuring for decision analysis.

5. Conclusions and recommendations

We believe the algorithm presented in this paper is a robust method for aggregating the knowledge of geographically separate experts in a timely manner, and as such possesses significant potential for improving emergency response capability. Technical improvements ought to improve timeliness; examples from actual emergency response cases ought to contribute to the success of training exercises in which the algorithm is employed.

Thus, room for progress exists. The process by which the MEID is generated deserves particular attention. Other rules for creating the MEID may result in a graph that preserves more of the information present in an individual expert's influence diagram. Additionally, we note that some well-known results from theories of social choice, such as Arrow's impossibility theorem (Arrow, 1966), present certain challenges. For example, the process of aggregating valid influence diagrams could generate an MEID that contains cycles (Black, 1963). Various attempts to minimize the pervasiveness of these cycles have been made (e.g. Black, 1963; Fishburn, 1971) but none fully address the problem in reference to influence diagram aggregation. Finally, by applying appropriate techniques from experimental design, it might be possible to accommodate larger groups of decision makers without overly taxing the cognitive processes of participants.

Appendix

We provide the influence diagrams generated by Participants 1–4:

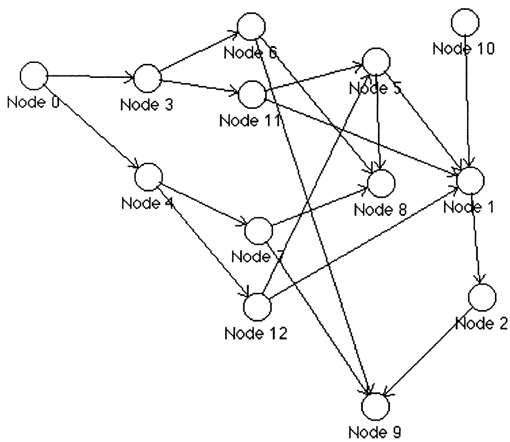

Fig. A1. Participant 1's graph.

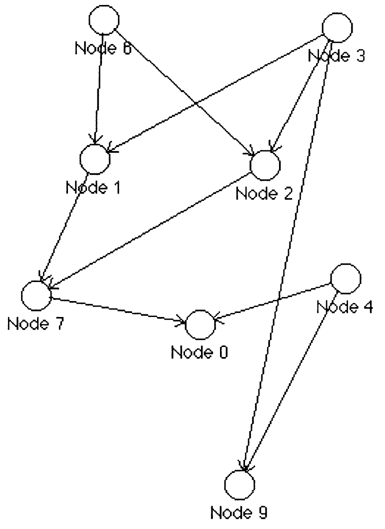

Fig. A2. Participant 2's graph.

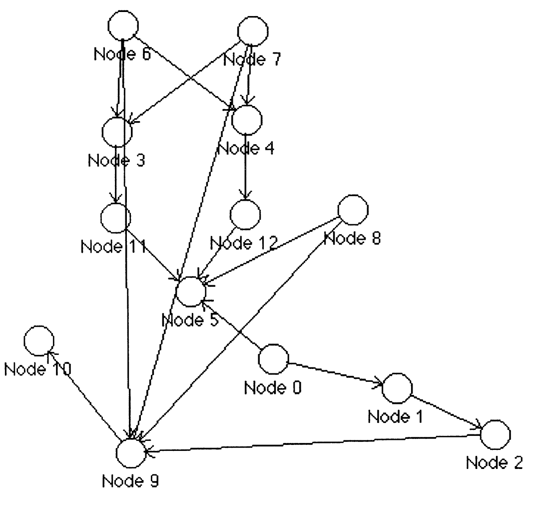

Fig. A3. Participant 3's graph.

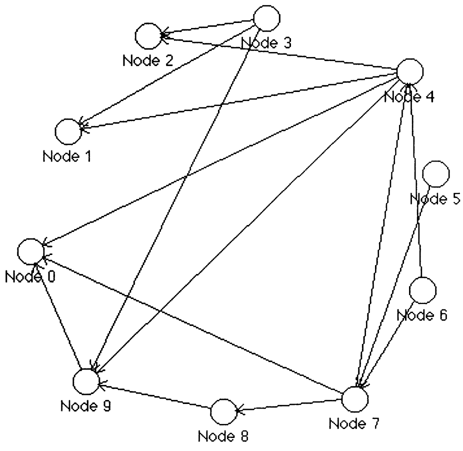

Fig. A4. Participant 4's graph.

References

Arrow, K.J., 1966. Social Choice and Individual Values. Cowles Foundation for Research in Economics at Yale University, New Haven, CT.

Banks, D., Carley, K., 1994. Metric inference for social networks. Journal of Classification 11, 121–149.

Barowski, L., Carter, G., Langford, M., Stutzman, S., 1997. VGJ: Visualizing Graphs with Java. Auburn University, Auburn.

Black, D., 1963. The Theory of Committees and Elections. Cambridge University Press, Cambridge.

Boose, J.H., 1986. Expertise Transfer for Expert Systems Design. Elsevier, New York.

Borda, J.C., 1781. Mémoire sur les Elections au Scrutin. Historie de l'Académie Royale de Science.

Browne, G.J., Curley, S.P., Benson, P.G., 1997. Evoking information in probability assessment: knowledge maps and reasoning-based directed questions. Management Science 43 (1), 1–14.

Clemen, R.T., 1996. Making Hard Decisions. Duxbury Press, Albany.

Critchlow, D.E., 1985. Metric Methods for Analyzing Partially Ranked Data. Springer-Verlag, New York.

Fishburn, P.C., 1971. A comparative analysis of group decision methods. Behavioral Science 16, 538–544.

Goddard, S.T., 1983. Ranking in tournaments and group decision making. Management Science 29 (12), 1384–1392.

Harrald, J., Marcu, H., Wallace, W., 1990. The Exxon Valdez: an assessment of crisis prevention and management systems. Interfaces 25 (5), 14–30.

Howard, R.A., 1989. Knowledge maps. Management Science 35 (8), 903–922.

Johnson, R.A., Wichern, D.W., 1992. Applied Multivariate Statistical Analysis. Prentice Hall, Upper Saddle River, NY.

Katz, R.H., 1995. Adaptation and Mobility in Wireless Information Systems (unpublished: available at URL http://daedalus.cs.berkeley.edu).

Kendall, M., 1962. Rank Correlation Methods. Hafner, New York.

Langel, J., Kann, A., 1992. DPL Advanced Version User Guide. ADA Decision Systems, Menlo Park.

Massey, A.P., Wallace, W.A., 1991. Focus groups as a knowledge acquisition technique: an exploratory study. IEEE Transactions on Knowledge and Data Engineering 3 (2), 193–207.

Massey, A.P., Wallace, W.A., 1996. Understanding and facilitating group problem structuring and formulation: mental representations, interaction and representation aids. Decision Support Systems 17, 253–274.

McGovern, J., Samson, D., Wirth, A., 1994. Using case-based reasoning for basis development in intelligent decision systems. European Journal of Operational Research 77, 40–59.

Robinson, D.F., Foulds, L.R., 1980. Digraphs: Theory and Techniques. Gordon and Breach, Science Publishers, New York.

Rush, R., Wallace, W.A., 1997. Elicitation of knowledge from multiple experts using network inference. IEEE Transactions on Knowledge and Data Engineering 9, 688–696.

Shachter, R.D., 1988. Probabilistic inference and influence diagrams. Operations Research 36 (4), 589–604.

Vitae

David Mendonça is a PhD student in the Department of Decision Sciences and Engineering Systems at Rensselaer Polytechnic Institute. He holds an MS from Carnegie Mellon University and a BA from the University of Massachusetts at Amherst.

Robert Rush earned a BS in Mathematics from Fordham University in May 1989. He then worked for the Central Intelligence Agency for 4 years, focusing on submarine detectability analysis. Returning to graduate school, he earned an MS in Operations Research and Statistics (1994) and a PhD in Decision Sciences and Engineering Systems (1998) from Rensselaer Polytechnic Institute. His research interests include risk management, optimization under uncertainty, and decision support system development.

William A. Wallace is a Professor of Decision Sciences and Engineering Systems at Rensselaer Polytechnic Institute in Troy, New York. He has more than 20 years experience in research and development in management science and decision support systems, with particular emphasis on crisis management. He is presently engaged in research on intelligent decision aids and the process of modeling.