Adapting to User Preferences in Crisis Response

Wayne Iba and **Melinda Gervasio**
Institute for the Study of Learning and Expertise
2164 Staunton Court Palo Alto, CA 94306
(650) 723-1684
{iba,gervasio}@isle.org

ABSTRACT

The domain of crisis planning and scheduling taxes human response managers due to high levels of urgency and uncertainty. Such applications require assistant technologies (in contrast to automation technologies) and provide special challenges for interface design. We present INCA, the INteractive Crisis Assistant, that helps users develop effective crisis response plans and schedules in a timely manner. INCA also adapts to the individual users by anticipating their preferred responses to a given crisis and their intended repairs to a candidate response. We evaluate our system in HAZMAT, a synthetic hazardous materials incident domain. The results show that INCA tailors itself to individual users and provides effective support for the timely generation of effective responses.

Keywords

adaptive interfaces, collaborative scheduling, user modeling

INTRODUCTION

Intelligent assistants hold great promise for agencies such as public health and safety organizations or the military establishment which are chartered with responding to crisis situations an effective and timely manner. The three primary elements of crisis – threat, uncertainty, and urgency [*cogsci97*?] – suggest that assistance mechanisms, rather than automation, will be most beneficial to those humans who must plan and schedule responses to crises. The user must participate in the response generation process in order to quickly and accurately evalutate an assistant's suggestions. Ideally, for the assistant should relieve the user of planning and scheduling details but also support the user by double-checking responses for weaknesses that that could prevent satisfactory resolution of the crisis.

Designing and implementing assistant mechanisms introduce unique problems generally ignored in artificial intelligence work where automating a problem solving activity is typically the goal. In the assistant context, a solution that is effective with respect to an objective criteria in necessary but not sufficient. Responses to a problem must also be *acceptable* to a user's subjective criteria which may include simple preferences that are actually irrelevant in terms of the objective criteria. However, these criteria may also constrain the space of solutions in ways that take advantage of unique characteristics of either the domain, the problem at hand, or the resources available. This expanded specification of desired solutions in the assistant context suggest the employment of user modeling mechanisms that can represent and reason about the problem-solving process as the user views it. That is, the assistant should infer a direction that the user is going with the solution and provide advice and guidance in accordance with this model of the user. However, it is impractical to hand-code such models; therefore we employ learning techniques to acquire models of user behavior through observation of previous solutions.

An intelligent assistant should *adapt* it's problem solving search according to the user's previous and present behavior. One benefit of an adaptive approach is that advice facilitating an effective response can be anticipated and presented in a timely manner, thereby addressing the urgency inherent in a crisis situation. We also want an adaptive assistant to adapt quickly to users' styles or preferences. That is, they should form an accurate user model as soon as possible so as to begin helping the user's response process.

In the following pages, we describe an intelligent user interface for crisis response that adapts its advice to the needs and preferences of individual users. In particular, we examine the task of responding to chemical spills and fires, which we describe in the following section. We describe INCA, our intelligent interface to this crisis domain, which addresses the problems inherent in assistant contexts by forming a user model capturing user preferences. We evaluate our approach to crisis response through experimental studies with INCA and human subjects working together to construct a response to a crisis.

HAZMAT CRISIS DOMAIN

We have developed HAZMAT, a synthetic world involving hazardous material incidents based on the 1996 North American Emergency Response Guidebook (NAERG) [*naerg*], for the purpose of evaluating computational assistants for crisis response. A HAZMAT incident is a spill, and possibly a fire, involving a material with properties such as toxicity and flammability. There are 4000 different classes of HAZMAT incidents, varying in the material involved and the magnitude of the spill and fire.

For every incident, NAERG defines an applicable (*legal*) subset of 49 primitive actions, each of which minimally requires a specified complement of resources from among the 25 types of resources represented in HAZMAT. Responding to a HAZMAT problem involves choosing a subset of the legal primitive actions and scheduling them on the available resources without violating any quantity or capacity constraints.

The effectiveness of a HAZMAT response is determined by simulating the effects on the incident of the scheduled actions and comparing it to what would have occurred without any intervention. Specifically, the HAZMAT simulator tracks the amounts and growth rates of the material spilled/burned and the fire and health hazard levels. The improvement measure reported in the experiments measures how much less material was spilled and burned and how much lower the hazards were because of the user's response.

INCA: INTERACTING WITH HAZMAT

We have developed INCA, an intelligent and adaptive assistant for the rapid construction of high-quality responses in the context of crisis incidents in HAZMAT. INCA serves as the interface between the user and our synthetic hazardous materials domain and contains several assistant mechanisms that enable it to advise the user on actions to select based on previous experience with either that user or other users. Figure X sketches the response task for HAZMAT incidents and shows the respective relationships and responsibilities for INCA and the user. We briefly describe three facets of INCA: the graphical user interface, the assistance mechanisms, and the adaptive mechanisms.

Graphical User Interface

A hazmat user learns of and responds to an incident through the graphic front end of INCA. There are three conceptual components to the graphical interface: world state information pertaining to the incident, an interactive work area for constructing and adapting plans, and a corresponding work area for schedules. The incident description is presented in the upper left hand corner of the screen and the available resources and their amounts are listed down the right side of the screen. The rest of the screen is used for constructing and adapting plans during planning mode and the same area is used for constructing and adapting schedules during scheduling mode. These modes are selected via a top-level menu control feature and the user may move back and forth between modes.

In planning mode, plans are modified by manipulating the hierarchical task network. The plan is displayed as an indented text tree. Tasks in the hierarchy are selected with the mouse

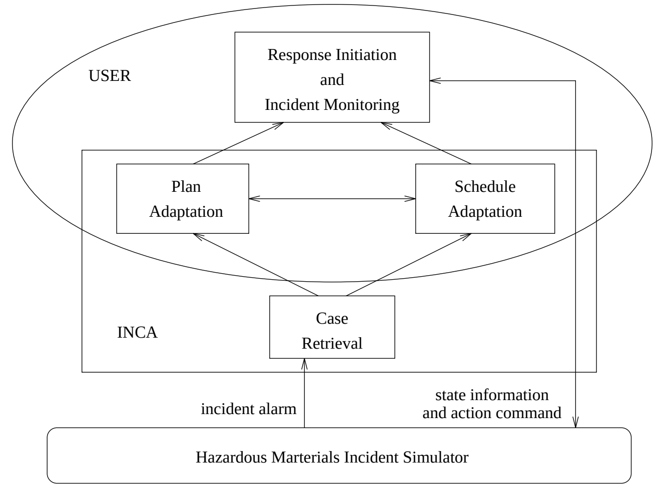

Figure 1: Mixed-initiative response to HAZMAT crises involving INCA (the INteractive Crisis Assistant) and a human user.

and either expanded and deleted. When the user is satisfied with the current plan as given by the expanded primitive actions (non-primitive actions cannot be directly executed), she exits planning mode and enters scheduling mode.

Those jobs comprising the schedule are displayed as color coded blocks on lines corresponding to the specific resources assigned and positioned left to right according to the start time and duration. Schedules are modified by either adding an unscheduled job to the schedule or repairing a particular job that is already scheduled. When jobs are added, the user is led through a dialog that fills in the necessary information to schedule a job: number of resources to allocate, which specific resources to assign, start time, and duration. Jobs can be repaired by selecting a colored rectangle in the schedule and either switching specific resources assigned to the job, changing the starting time of the job, changing the duration of the job, or deleting the job altogether.

Assistant Mechanisms

As an intelligent user interface, INCA provides a number of mechanisms that assist the user in response generation. These mechanisms help in one of three categories: administrative, technician, and collaborator assistance. Adminitrative assistance includes basic record keeping, checking the legality (with respect to the Emergency Response Handbook) of planned actions, and checking the feasibility (i.e., that resources are not over-subscribed and not over-allocated) of final schedules.

Technical assistance is given by initializing, or seeding, the plan and schedule. This seeding is based by default on cases stored from previous solutions to similar crisis incidents (however we have also implemented heurisitc planning rules and an automated scheduling system that can be used to initialize solutions). The seed case is retrieved by comparing the incident description, the available resources, and the legal actions (given the incident) with those of previously seen incidents stored in the case library. The plan and schedule from this retrieved case is minimally adapted according to the specifics of the current situation. The case-based aspects of INCA, as well as other details of the system are described in greater detail in [CogSci98 and AAAI98].

More collaborative or peer assistance is provided in the form of advice during the repair process. Whenever advice is offered by INCA, it is presented by highlighting the anticipated or recommended menu selection. If the user selects another action, the system updates it's model and make a new recommendation based on the action the user has selected. This level of advice is based on both heuristic rules and acquired models of user behavior given similar circumstances.

Adaptive Mechanisms

There are two adaptive mechanisms that enable INCA to acquire a model of the user's preferences and behavior. The first is through the collection of a case library. Cases are collected whenever an incident response has been completed. Since similar cases are retrieved for a new incident, the seed for the initial response will be strongly influenced by the user's previous behavior and solution style.

In addition to adding cases, INCA forms user models based on individual repair actions selected by the user. Every repair is treated as a training instance for a machine learning algorithm (we have experimented with numerous algorithms with largely similar results). The action selected by the user is the class name (i.e., the feature to be predicted in the future) and the incident description, available resources, and schedule (represented as scheduled and unscheduled jobs) are used as the features. During the repair of the seeded solution, the learned classifier is used to predict, given the current situation, resources and state of the schedule, which of the available actions the user is most likely to select. When INCA's predictions are correct, the user's task is simplified from *generating* her next action to that of *recognizing* the next appropriate action. Of course, we need to know how accurate and helpful are all of the assistant mechanisms and we turn now to our experimental evaluation of INCA.

EXPERIMENTAL RESULTS

In a previous study [CogSci], we evaluated the usefulness of INCA's case-based seeding mechanism. We tested subjects on multiple hazmat incidents and evaluated their response effectiveness with and without initial responses derived either from case libraries or from an automated planning and scheduling tool. In another study [AAAI], we showed that the learning mechanism captured regularities in user response behavior. In this paper, we report significant new findings based on those studies.

In our previous study, we explored several different means of generating seeds for users to repair. The dependent variables measured the effectiveness of the response and the time required by the user to adapt or repair the initial seed. The results showed that case-based seeding methods enabled responses to be completed in significantly less time without sacrificing quality although there were no significant differences in quality (z-test, $p < 0.05$).

Table 1: Average situation improvement due to case seed and additional improvement due to user's modifications.

	situation improvement		
Seed	base	A	B
case seed A	22.3	33.0	34.8
case seed B	22.5	36.9	34.5

However, it is not sufficient to demonstrate that INCA helps users complete crisis repsonses more quickly [CogSci]; we also want to determine the respective contributions from the assistant mechanism and the user's repairs. Perhaps, the user was responsible for all the improvement making the initial seed irrelevant or perhaps detrimental. On the other hand, the user might be wasting time trying to repair a seed without measurable improvement. To address this, we compared the effectiveness of the seeded solutions alone to that of the repaired solutions generated by the users. In all conditions, the difference between the seed and the final repaired response was significant (paired t-test, $p < 0.05$). Table X shows the average percentage *improvement* in crisis problem outcomes for two different case libraries, A and B. The improvement in effectiveness is given for the case-seeding alone and the repaired responses for two subjects. In light of our previous studies, these new results show that users are contributing significant improvements to INCA's seeds but that these seeds are also contributing resulting in a net reduction in response time.

Adaptation to Individual Users

In another earlier study [*aaai98*], we showed that machine learning can be used to acquire user models that let INCA *adapt* its assistant behavior to individual users. Specifically, we demonstrated that after learning, INCA's accuracy in predicting a user's next schedule modification operation increased significantly over random guess and guessing the most frequent class. We also showed that learning on data from one user and testing on data from another was inferior to learning and testing on data from the same user, demonstrating the usefulness of adapting to individual users.

In further analyses of the learning results, we identified a possible confounding factor: our data sets had been extracted from traces of users interacting with INCA over the variety of schedule seeding conditions investigated in [*cogsci98*]. Because different initial schedules may lead to very different schedule modification patterns, we hypothesized that dividing the data sets by condition and then learning within individual conditions would result in better performance.

For each resulting data set, we ran 20 trials on 2 different random training and test splits, with the number of training instances varying from 100 to 400, depending on the size of the data set. Figure 2 compares the learning curves averaged

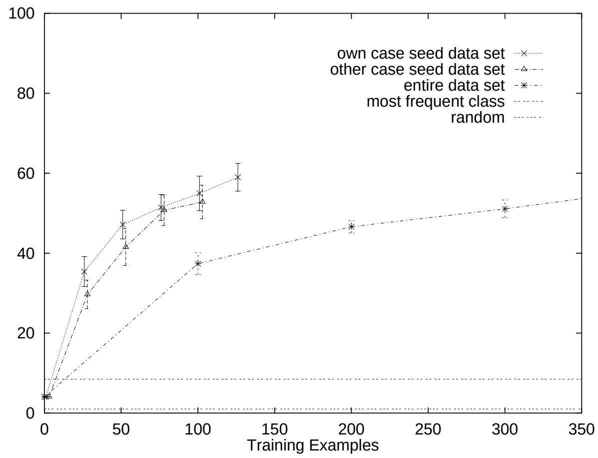

Figure 2: Learning curves, showing improved accuracy when learning on specialized data subsets over learning on entire data sets.

over the 40 total trials of the two users on the data sets extracted under the case-seeded conditions to original learning curves on the entire data sets. The results support our hypothesis: learning was significantly faster on the case-seeded data sets than on entire data sets. Learning on the entire data set required 300 examples before it was comparable to the level of accuracy achieved after just 100 examples in the specialized data sets. Further tests on additional data is needed to verify our hypothesis that the asymptote with the specialized data sets is no worse than with entire data sets, however, even if it were, the higher learning rate may be sufficient for adaptive interfaces such as INCA, where rapid adaptation is desirable.

DISCUSSION

There are numerous dimensions along which we could consider related research efforts. We will only mention the crisis domain, and adaptive interfaces. The underlying goal of our research – adaptive interfaces that act as both apprentice and advisor – is most similar to the body of work on intelligent tutoring systems summarized in [*sleeman*]. There are numerous recent efforts addressing the task of crisis response, perhaps the most similar of which [*desJardins*]. They are addressing interactive crisis response planning using case-based techniques and are also addressing the distributed, collaborative nature of the response task for large agencies.

We are currently pursuing several extensions to this work and have others planned for the near future. We are in the process of significantly revising the graphical user interface in order to remove order dependencies in the dialog when adding and repairing actions in scheduling mode. We are also continuing our data analysis of experimental results collected to date toward answering open questions regarding the effect of incident difficulty and the interaction of urgency with response quality. In the coming months, we intend to run revised experiments using the updated graphical user interface with the intention of gathering larger amounts of higher fidelity data. We also intend to explore an alternative approach to user modeling and solution repair advice that involves learning the user's evaluation function on complete schedules (in contrast to our current approach of learning what amounts to situated control rules).

In closing, it is the claim of our research program that intelligent interfaces must focus on *appropriate* levels of assistance, and this requires that these interfaces adapt to their users. Considerable work remains before we can substantiate this claim in its strong prescriptive sense. However, our work to date is consistent with our claim and shows significant promise for general methods for learning user models for adaptive interfaces.

ACKNOWLEDGEMENTS

This work was supported by the Office of Naval Research under Grant N000014-96-1-1221. We thank Pat Langley, Greg Trafton, Michael Fehling, Gregg Courand, for their contributions to this research with special thanks to Stephanie Sage for her assistance with the graphical interface.

REFERENCES

1. desJardins, M., Francis, A., & Wolverton, M. (1998). Hybrid planning: An approach to integrating generative and case-based planning. *Working Notes of the AAAI-98 Workshop on Case-Based Reasoning Integrations*.

2. Gervasio, M. T., Iba, W., & Langley, P. (1998). Learning to predict user operations for adaptive scheduling. *Proceedings of the Fifteenth National Conference on Artificial Intelligence*.

3. Iba, W., Gervasio, M., Langley, P. & Sage, P. (1998). Evaluating computational assistance for crisis response. *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*.

4. Sleeman,, D., & Brown, J. S. (Eds.) (1982). *Intelligent tutoring systems*. New York: Academic Press.

5. Miyashita, K., & Sycara, K. (1995). CABINS: A framework of knowledge acquisition and iterative revision for schedule improvement and reactive repair. *Artificial Intelligence*, *76*, 337–426.

6. Transport Canada, the U.S. Department of Transportation, & the Secretariat of Communications and Transportation of Mexico. *1996 North American emergency response guidebook*.