The cognitive systems engineering of automated medical evacuationscheduling and its implications

Conference Paper · September 1996

DOI: 10.1109/HUICS.1996.549516 · Source: IEEE Xplore

The Cognitive Systems Engineering of Automated Medical Evacuation Scheduling and its Implications

Richard Cook†, David Woods‡, Marie Walters‡, Klaus Christoffersen‡
†Cognitive Technologies Laboratory, Department of Anesthesia and Critical Care, University of Chicago, Chicago, IL 60637 and ‡Cognitive Systems Engineering Laboratory, The Ohio State University, Columbus, OH 43210

Abstract

Analyzing the cognitive tasks of operators of future novel systems in order to provide design guidance represents a major challenge to cognitive systems engineering. Cognitive task analysis based on extant systems may not adequately reflect operator tasks after radical organizational and technical change. We describe the results of an effort to analyze a proposed, highly automated system for patient evacuation movement scheduling. The new automation will relieve operators of much of the work of schedule generation, but giving up this task to the machine makes it necessary for users to independently generate estimates of the location of the current problem in the problem space. These estimates were previously obtained as a byproduct of manual scheduling. In this context, decision support largely concerns construction of this estimate which then determines operator strategy. This need for problem space estimates is a comprehensible central theme that can serve as an organizing metaphor for system designers.

1: Introduction

The process of technology change in any given work domain produces a gap between the *as is* world and the *as envisioned* world. Developers of new technology make claims about how their new systems will have beneficial impacts on human and organizational performance in the domain of application--in other words, they envision or predict what the target world will be like after the process of technology change takes place. However, studies of technology change have revealed that such predictions are often tenuous; the actual impact of new technology is often quite different than that envisioned by technology advocates. The introduction of new technology for use by practitioners in a given work domain, regardless of the scale of change, almost always produces some surprises (e.g., [5]).

The gap between the *as is* world and the *as envisioned* world presents a recurrent challenge to Human Factors. How can one use results from studying the current world to better predict the impact of proposed technological and organizational changes on the cognitive system of agents, tools, and problem demands? How does one use information gathered about the current system to inform a process that is changing the very system under study?

An example of the introduction of new technology representing a fundamental change in operators' work environment is the TRAC2ES system for U.S. military medical evacuation planning. TRAC2ES includes a state of the art, computationally intensive algorithm designed to supplant many of the current human operator functions in the scheduling of patient evacuation. This new information technology is only one part of a larger change to the process of military medical evacuation that includes basic changes to the organization itself. The hardware and software designs are based, in part, on anticipation of this new system's organizational characteristics. All parties involved recognize that the technological and organizational change will produce a new world of scheduling and planning aeromedical evacuation that is far different from the *as is world*.

As Cognitive Engineers, we faced the challenge of examining the developers' *as envisioned world*. In particular, we examined how the introduction of the new automated schedule generator (ASG) would change the roles and activities of the people involved in scheduling patients needing aeromedical evacuation. Our goal was to employ Cognitive Systems Engineering (CSE) techniques to develop a model of the cognitive life of practitioners after the introduction of an ASG. The results of these investigations, which we present here, are a prediction about the *as envisioned world*. These predictions have been used as tools to attempt to help designers modify the ASG and how it works with people to create an effective decision support system.

0-8186-7493-8/96 $5.00 © 1996 IEEE

Proceedings of the 3rd Symposium on Human Interaction with Complex Systems (HICS '96)
0-8186-7493-8/96 $10.00 © 1996 IEEE

2: CSE and the *as envisioned world* of Aeromedical Evacuation Scheduling

In semantically complex, dynamic, high-risk processes such as commercial aviation and nuclear power, the introduction of new technologies typically takes the form of conservative, gradual refinements and extensions of current systems. Development and implementation of the new technology does not radically change the underlying technical aspects of the process, operator expertise is already highly refined, and system performance is at a high level. In such cases, CSE can proceed largely through empirical study of operators' current environment. Cognitive Task Analysis (CTA) of the current system will reveal important features of operators' current cognitive tasks that will remain largely unchanged after the new technology is in place.

When new system designs envision changes in technology and organization that will radically transform the operators' environment, the envisioned world is different from the *as is world* in ways which significantly limit the ability of empirical study of current practice to inform evaluations of hypothetical designs. The extensive extrapolation required in attempting to bridge the conceptual gap between the current and envisioned systems can be highly brittle, and designers' predictions about the impacts of such technological interventions will be accordingly more tenuous. The challenge to Cognitive Systems Engineers is to discover and describe significant, likely characteristics of the future environment and present them in ways useful to the designers.

In contrast to system designers, who are often necessarily focused on how to make the new technology operate successfully in canonical cases, CSE attempts to examine hypothetical systems at the margins of performance--trying to discover how systems might break down. Learning takes place by examining how future practitioners will handle anomalies and exceptions.

Through analytical techniques (e.g., [3], [7]), in combination with empirical observation, a picture of the underlying constraints or problem demands in the operators' world can be built. This facilitates the proposal of concrete "future" scenarios specifically designed to reveal the stress points in the new system. Working through these scenarios with system designers forces them to explicitly project how the underlying constraints will map to the envisioned technological and organizational environment. This process can provide a great deal of valuable information about how the problem demands faced by operators will manifest themselves and shape the impacts and usage of the new technology.

In designing scenarios and projecting practitioner behavior, CSE can draw on substantial empirical experience from studies of the impacts of technology change on practitioner cognition over a wide variety of domains. This experience, combined with detailed information about the proposed system for specific scenarios, can inform predictions about operators' ability to control the underlying process, detect and diagnose faults, recover from failure, and predict the future course of system operations.

As part of a project to provide design guidance to TRAC2ES system developers, we attempted to develop an operators' perspective of an *as envisioned world* for aeromedical evacuation scheduling. The effort included study of the proposed system as it would be seen from the operators' perspective using descriptions of the system obtained from designers and engineers. The design and implementation teams knew they wanted to create a decision support system, but their background and project organization were focused on the details of the ASG algorithm and its functional capabilities. For the most part, designers assumed that the decision support goal would be met simply by asking people to evaluate the schedule produced by the ASG and by allowing people to run the ASG multiple times. The effect of CSE was largely to redirect attention of designers towards the deeper implications of what is needed to achieve cooperative, aided aeromedical scheduling decision making rather than simply supplanting it with automation [1].

The self-assured optimism of designers that they understand the impact of the new technology on cognitive life of practitioners makes CSE both necessary and difficult. Studies of the impact of technology change on practitioner cognition and performance have revealed that designers tend to over-simplify the practitioner's new roles, often overlooking the need for human operators to direct automation, make decisions, exercise judgment, and communicate the results generated by the automation to others.

In our studies of the *as envisioned world* of aeromedical evacuation, we concentrated on defining what operations would remain for human operators to perform, the sorts of decisions that operators would be required to make, how such decisions could be supported, and how anomalies and failures would be investigated and repaired. In particular, experience with automation in other contexts (e.g., [4]; [6]) suggests that operators will be confronted with the need to understand, qualify, validate, and modify the automated system's plans. Therefore, we attempted to describe ways in which those functions could be supported within the system. This investigation led us to the discovery of several general characteristics of the problem of matching heterogeneous demands with potentially limited resources under time pressure.

We used the notion of a *problem space* as the means to communicate our findings back to the designers. In particular, the analysis focused attention on those situations in which operators will be called upon to justify or repair schedules that leave urgent requests for patient movement unsatisfied. In these situations, determining the location of the current problem in the problem space is a prerequisite to deciding how to proceed.

Proceedings of the 3rd Symposium on Human Interaction with Complex Systems (HICS '96)
0-8186-7493-8/96 $10.00 © 1996 IEEE

During the course of the study, the system designers began to move to take explicit account of the ways that operators would interact with schedules generated by the automation. Proceeding from this, they began to include interface display and navigation elements to support post-schedule-generation investigation of the resource bottlenecks that shaped the automated schedule. The insights provided by CSE study of the *as envisioned world* thus appear to have had a tangible impact on the design of this complex system.

3: Methods

Over a nine-month period we interviewed designers, engineers and management of the TRAC2ES software development, a project that incorporates a distributed, object oriented database and a highly refined scheduling algorithm that is intended to supplant most of a complex manual scheduling system. The goal of the system development effort is to produce efficient management of transportation and medical resources in peacetime and especially in conflict situations. Design of the objects and relations was complete at the time of our first contact and interface issues were under intense scrutiny. In a series of meetings and demonstrations, supplemented by ongoing discussion and, in some instances, review of source code, we developed a model of what the automated scheduling algorithm at the core of TRAC2ES would offer to users and what it would demand of them. One of us (MW) observed an exercise intended to simulate casualty evacuation and also recorded details of an operating evacuation management center doing peacetime patient movement (see Walters et al., this volume).

Because the new system was under development and unavailable for testing with future users, we paid particular attention to the ways in which designers and engineers described the expected user interaction with the system, in effect attempting to elicit their model of the user, and contrasted that with the types of interaction we anticipated on the basis of observations of users and experience with other automation [2], [4], [8]. Midway through the project we provided the engineers and designers with an interim summary of our findings that described ways in which the ultimate users of the system would be expected to respond to its algorithmic features, i.e. the strategies that we anticipated users would employ when using the new system under demanding real world conditions. The report concentrated on three primary features: (1) the nature of the problem space that encompassed all possible scheduling problems, (2) ways in which manual scheduling activity helps locate the current problem in that space and (3) ways in which users were likely to seek to use the new automated system to establish the location of the current problem when confronted with incomplete or undesirable schedules.

4: The *problem space*

Medical evacuation planners will be confronted with a variety of scheduling situations (problems) with widely varying characteristics. Among these characteristics are the scale of demand (number of patients to be moved), complexity of demand (the different locations of patients, medical characteristics, etc.) and number and disposition of resources (available slots in transport aircraft, bed availability in hospitals, etc.). An individual scheduling problem may be regarded as representing one of an infinite number of possible problems, that is, as defining a single point within the problem space containing all possible problems.

The problem space for real world activities like medical evacuation is multidimensional but for conceptual purposes may be projected onto a two dimensional diagram showing the relationship between supply (or capability, capacity, availability, etc.) and demand (production targets, patient movement requests, etc.); as illustrated in figure 1. The location of a problem within the problem space determines whether the problem has any "good" solutions. In the present and foreseeable future U.S. military medical evacuation system, a "good" schedule is one that provides a timely movement path for all "urgent" patient movement requests (for a discussion of the implications of "urgent" classification see Walters, et al., this volume).

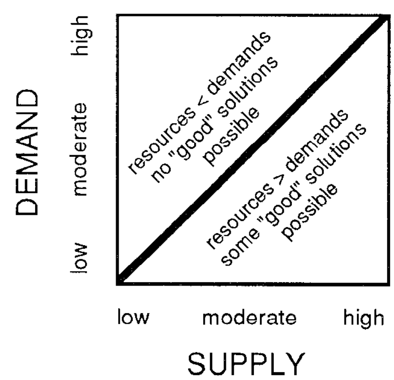

Figure 1

It is quite possible for the human scheduler to be uncertain about where in the problem space the current problem lies or to be certain but mistaken about its location. Especially when the exact location is thought to be within a region that includes both potentially soluble and absolutely insoluble problems, determining the precise location of the problem in the problem space may be a crucial part of solution generation. When operators realize that there is no "good" solution they will shift from trying to create a "good" solution to trying to

Proceedings of the 3rd Symposium on Human Interaction with Complex Systems (HICS '96)
0-8186-7493-8/96 $10.00

discover what additional resources might be added to the current collection of demands to make a "good" solution possible. Clearly, locating the current problem in the problem space is critical to appropriate shift of attention during schedule evaluation.

Significantly, what makes a problem difficult to solve depends on how closely the resources and demands are matched, not simply on the problem size (number of requests for patient movement). Although large problems are generally more difficult than similarly configured but smaller ones, other aspects of a problem, especially the relative excess of supply factors over demand elements, determine the effort required to obtain a good solution (figure 2) and suggest the presence of what might be described as *iso-effort lines* (dotted lines) along which the work required to generate a good solution is constant. With a large excess of resources, a variety of solutions are easily generated and the effort required to produce a "good" solution is low (#1). As resources and requests are more closely matched, the ease with which a "good" solution can be generated falls (#2). Along the line of identity, where resources precisely equal demand, only a single solution exists and the effort required to generate a solution is maximal (#3). Above the line of identity, no amount of effort will produce a "good" solution; only "bad" solutions are possible in this region (#4). However, a problem located near the line of identity (#5) is more amenable to repair by the provision of extra resources than is one far away (#4).

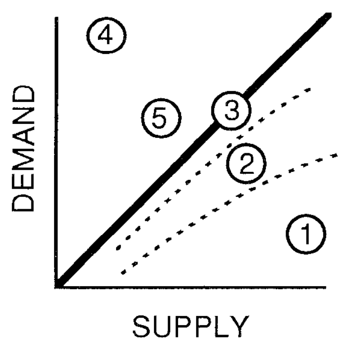

Figure 2

Interestingly, human manual schedule generation produces, as a byproduct of the scheduling effort, some appreciation of how the resource supply and movement request demands combine to form resource bottlenecks. The sequential satisfaction of requests and consumption of resources that characterizes manual planning helps the human planner to "see" the specific features of the planning problem and the potential solutions. This allows the planner to produce both a schedule and specific, focused notions of what resources are in short supply and what additional resources might be used to overcome these limitations. An optimal schedule is by no means assured with manual scheduling; especially in the case of large problems it may be difficult to keep track of all the bottlenecks and subtle interactions between resources and demands that prevent formation of a "good" solution. Nevertheless, the principle of becoming informed of the approximate location in the problem space by direct contact with the characteristics of the problem is fundamental to understanding how human operators understand the problem and the resulting schedule.

Indeed, *planning* (as distinguished from *scheduling*) is a process that produces both a schedule and the informed appreciation of its features that makes it possible for the planner to comprehend, rationalize, and justify the schedule in terms that are suited to the domain. Thus a significant challenge to those who would automate 'he scheduling portion of the human planner's task is to provide, along with the schedule, information that locates the current problem in the problem space.

5: The *problem space* and bad solutions

When a problem lies in the excess resource portion of the problem space (figure 2, #2), the various possible "good" schedules may be contrasted in terms of efficiency (e.g. cost) or resilience (e.g. vulnerability to loss of a specific resource) and a choice of schedule made on those grounds. In this region, planners compare varieties of success.

In contrast, a "bad" schedule represents failure because it leaves unplanned the movement of urgent (presumably critically ill) patients. It is reasonable to predict that "bad" schedules will receive more scrutiny than "good" ones; this is characteristic of process control environments - operators focus their attention on anomalies. Such schedules cannot be adopted as plans without justification. Part of the planning process is the production of such justification.

It is entirely possible that a "bad" schedule reflects either the best possible use of resources or better use of resources than any other scheduler might achieve in a limited time. But human subordinates preparing "bad" schedules will perforce have to present these to the command authority with the assurance that every effort to accommodate the more urgent requests for patient evacuation has been made. In using automated tools to generate schedules, an important role of human operators is to describe to the command authority what resources, if added to the available resources, would convert the "bad" schedule to a "good" one. The same problem characteristics that define the location in the problem space are those that are the focus of attention of human operators confronted with a "bad" schedule. The process of reviewing a schedule, recovering the problem characteristics, and rationalizing the proposed schedule in light of the characteristics, necessarily precedes embracing

Proceedings of the 3rd Symposium on Human Interaction with Complex Systems (HICS '96)
0-8186-7493-8/96 $10.00 © 1996 IEEE

and adopting the schedule, i.e. to turning the *schedule* into a *plan*.

The need to justify "bad" plans to the command authority and to advise the command on precisely what resources should be added to overcome bottlenecks is likely to lead to use of automation *as an analysis tool* to explore the region of the problem space surrounding the current problem. By altering the collection of resources offered to the automated scheduler, operators may be able to discover a resource mix that can produce a "good" plan. Once discovered, the difference between the available resources and those that produce a good solution is exactly the new resources needed to convert the schedule from "bad" to "good". It is then possible for the human operators of the system to present the "bad" schedule to the command authority along with a precise statement of what is needed to convert that schedule to a "good" one. The limitations on this sort of exploration are, in part, narrowly technical: to explore the problem space in this way, operators must be able to create hypothetical resource sets, to run the automation on several different sets, and to track and compare the resulting schedules. Failure of this sort of exploration to produce results, time pressure, or difficulty handling the bookkeeping associated with multiple automated scheduler executions will halt the exploration.

It is also possible to convert "bad" schedules to "good" ones by reducing demand for resources until the extant resource pool can satisfy the demand. In the *as envisioned world* of TRAC2ES, for example, physicians may constrain patients to specific destinations rather than allowing patients to move to any available hospital. However, these requirements may substantially hobble the ability of the algorithm to produce a "good" schedule. This may motivate planners to try to engage in negotiations with physicians to try to re-classify patients for scheduling purposes. For example, given a patient who has been constrained by the attending physician to move only to hospital X, a planner may say: "I can't get your patient to hospital X for two days, but I can get him to hospital Y by tomorrow morning. Will you drop the destination requirement?" Eliminating these sorts of constraints relaxes demand and may allow the automation to produce a "good" schedule. Exploratory running of the automation with different forms of relaxation of demand may thus form the basis for repair of "bad" schedules.

Supporting decisions to add resources or relax the characteristics of demand is largely a matter of being able to demonstrate the precise consequences of undertaking these actions. This is more likely in the setting where a "bad" schedule has been proposed and repair or justification is required. This suggests that an array of features to support multiple executions of an automated scheduler would support the operators' cognitive tasks. Such features include the ability to run the scheduler with hypothetical resource or demand mixes, to track and, especially, to contrast the results of executions.

Interestingly, situations where demand outstrips supply by a large margin are likely to be handled in other ways. These situations reflect overwhelming demand (e.g. catastrophic numbers of casualties) and the disparity will be so large that little formal justification will be required. This class of problems (figure 2 #4) is not solved by scheduling or incremental adjustments to resources or demands but by wholesale provision of large amounts of resource.

6: From schedules to plans

The decision support aspects of an automated schedule generation tool for military medical evacuation planning are related to the users ability to comprehend the location of the present problem in the problem space. The schedule itself is an artifact that assigns resources to demands but, by itself, says nothing about the way in which the current demands and available resources shaped the final solution. The process of *manual* scheduling *by its nature* generates a great deal of specific, detailed information which constitutes an important form of decision support. The fact that this problem space information is a *byproduct* of the manual scheduling process makes it no less valuable. Grappling firsthand with the details of resources and demands directs attention to those components that are worth close examination, worth exploring for defects, and worth the concentrated effort necessary to improve the situation by recruiting more resources.

ASGs relieve operators of the task of producing the schedule artifact and may even produce superior schedules to those that would have been created by manual scheduling (at least for large scale problems). But the automation, if it is uncommunicative about how it encounters, assesses, and resolves the conflicts inherent in scheduling, provide no support to operators in understanding the location of the problem in problem space. The operator must work to recover the problem characteristics in some other way. In well designed systems, recovery may be included as part of the display that accompanies the schedule artifact, i.e. the operator can "see" the factors that shape the schedule generated by the automated agent. Achieving this sort of design is synonymous with producing decision support. Creating such displays is difficult and remains the subject of debate.

In the absence of insight into the function of the automation, operators are reduced to recovering problem characteristics in other ways. We have suggested exploratory execution of hypothetical problems as one means for exploring the problem space. How such explorations will be guided is uncertain. Confronted with a list of unscheduled movement requests, operators will need to look for similarities between them (and also the differences between them and successfully scheduled requests) in order to fashion the hypothetical resource mix they then feed to the automation. Based on other

Proceedings of the 3rd Symposium on Human Interaction with Complex Systems (HICS '96)

automation studies (e.g., [2]), it is likely that experienced operators may "game" the system by deliberately understating available resources, holding these in reserve and performing manual scheduling off-line for residual unscheduled movement requests.

Alternatively, manual planning of "important" patient movement requests (however these are defined) may be performed and the resources "locked in" before running the automation. Although designers of TRAC2ES anticipated a purely automated approach to schedule generation, it is more likely that a hybrid approach involving both manual and automated schedule generation will evolve. In some cases operators will use the automation primarily and manually schedule what the automation leaves unscheduled. Alternatively and, in our view, more probably, operators will begin with manual scheduling and use automation to assign whatever resources remain after the most important requests are satisfied.

However produced, the adoption of a schedule as a plan for action involves commitment by the organization to a course of action. Such commitments are not taken lightly. Converting a schedule to a plan is the decision that requires support and this process itself depends heavily on the appreciation of where in the problem space the current situation lies.

7: The cognitive engineering of *as envisioned worlds*

Coping with the complexity of advanced automation is a problem first for designers, then for the engineers who implement the system, and finally for the users who operate it. It was the consideration of an automated schedule generating system that led to the realization that the manual scheduling process provides much more than just the artifact of a completed schedule. Thus the process of design and development itself is one means for exploring the cognitive world of the operator. In a sense, the design and implementation details constitute a hypothesis about the expected roles of human and machine in the world as envisioned. CSE, in this setting, involves going beyond the details of the system as envisioned to explore the ways that operators may be expected to cope with the complexity of the world and the automation.

The present study involves a system undergoing fundamental change, one where the value and even the possibility of Cognitive Engineering is contentious. In such a setting, the role of Cognitive Systems Engineers is not well defined and the available tools crude. Whether studies such as the present one provide useful guidance is sure to be the subject of much debate. Against the noise of a large, complicated project where deadlines, milestones, and the very real problems of software engineering tend to dominate development, it is difficult to make the cognitive voice heard. Attention to the boundaries of the system and the concept of recovering the location in the problem space is one way of giving that voice a distinctive quality.

Acknowledgments

The authors acknowledge the assistance of Scott Potter, Viktor Saks, Jeff Stonebrook, Ron Ball, Gretchen Braidic, Bill Elm, and others at Carnegie Group Inc. who patiently explained TRAC2ES to us. A number of military personnel inside and outside USTRANSCOM allowed us to observe their work and to ask questions. Their commitment, seriousness of purpose, diligence, and adaptability were remarkable and we thank them for allowing us to see a small part of their world. Preparation of this paper was made possible by support via Carnegie Group Inc. with funds for TRAC2ES development. The opinions expressed in this paper reflect those of the authors and not necessarily those of the sponsor or any agency or department of the government of the United States.

References

[1] Billings, C, Aviation Automation: The Search for a Human-Centered Approach. NJ: Lawrence Erlbaum, in press.

[2] R. I. Cook and D. D. Woods, "Adapting to New Technology in the Operating Room," Human Factors, in press.

[3] E. Hollnagel and D. D. Woods, "Cognitive Systems Engineering: New Wine in New Bottles," Int. J. Man. Mach. Studies, vol. 18, pp. 583-600, 1983.

[4] N. Sarter and D. D. Woods, "How in the world did we ever get into that mode?" Human Factors, vol. 37, pp. 5-19, 1995.

[5] N. Sarter, D. D. Woods, and C. Billings, "Automation Surprises," in G. Salvendy, ed., Handbook of Human Factors/Ergonomics, (2ed.). NY: Wiley, in press.

[6] D. D. Woods, R. I. Cook, and C. M. Billings, "The Impact of Technology on Physician Cognition and Performance," J. Clin. Monitoring, vol. 11, pp. 92-95, 1995.

[7] D. D. Woods and E. M. Roth, "Cognitive Systems Engineering," in M. Helander, ed., Handbook of human-computer interaction. NY: North-Holland, 1988.

[8] E. M. Roth, K. Bennett, and D. D. Woods, "Human Interaction with an 'Intelligent' Machine," Int. J. Man-Mach. Studies, vol. 27, pp. 479-525, 1987.

Proceedings of the 3rd Symposium on Human Interaction with Complex Systems (HICS '96)