PUTTING TIME (BACK) INTO DYNAMIC FUNCTION ALLOCATION[1]

Michael Hildebrandt and Michael Harrison
Department of Computer Science
University of York, York YO10 5DD, UK
{hilde, mdh}@cs.york.ac.uk

This position paper discusses extensions to concepts of Dynamic Function Allocation (DFA) that could improve our understanding of the trade-offs involved in designing and operating human-automation systems. We suggest that current DFA paradigms, focusing predominantly on allocation along the human-automation resource dimension, may provide an insufficient basis for design decisions as they fail to take account of alternative function management strategies. Of these strategies, Dynamic Function Scheduling (DFS), the allocation of functions along the temporal dimension, is of particular interest, not least because scheduling is both a mature engineering discipline and a ubiquitous aspect of human behavior. Understanding these scheduling decisions requires consideration of the temporal properties of functions (e.g. continuous, periodic, sporadic, pre-emptable, interleavable), temporal requirements (e.g. deadlines), and the temporal properties of the agents, human or automatic (e.g. service rates, interruption handling, task switch costs, temporal reasoning abilities, control modes). The paper reviews engineering and human factors approaches that could support the representation, analysis and design of DFS.

INTRODUCTION

Function Allocation, the process of dividing or sharing responsibilities among humans and automation at the design stage, is driven by several, partly conflicting, motivations. Automation promises to extend or support human performance, to compensate for human performance deficits, to relieve the human of routine tasks, or to replace the human altogether. At the same time, it transforms the role of the human to strategic decision maker and supervisory controller of an ever-increasing multitude of functions. The operator may have to switch between functions, react to unforeseen events or demands, and compensate for automation failures. To fulfill this role, the human needs to be kept informed of, and involved in, the operation of the system. Consistent and transparent automation behavior can help to maintain the compatibility of the operation strategies of the human and the automation, and mismatches between these strategies can lead to failure. The challenge for automation designers lies in providing automation that can act rapidly and dependably under hard deadlines, while allowing the human to control, configure or intervene in the operation of the system.

Dynamic Function Allocation (DFA) addresses the problems of static allocation by providing multiple levels of automation, and decision rules (potentially also automated) to switch between them at runtime. This creates a workload balancing mechanism that makes the system more adaptive to a wide range of operational parameters, as the agents – human(s) and automation – can supply mutual back-up in case of performance degradation or changing demand characteristics. Changes in automation levels may occur as part of a planned operation strategy (e.g. engage autopilot during cruise phase, disengage during take-off / landing), but most empirical studies in this area have been concerned with allocation switches as a reaction to unforeseen events or workload changes (e.g. Endsley and Kaber, 1999). It is important to notice that these allocation

[1] To appear in *Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society. Denver, Colorado, October 13-17, 2003*.

decisions are dynamic in time, but in a form that is best described as "snapshot allocation". At any instant, the allocation algorithm (relying, for instance, on triggers from critical-event monitoring, performance or physiological measurement) assesses the need for re-distribution of functions and suggests or implements the required changes by moving up or down the automation scale. In other words, the human-automation resource scale moves orthogonally along the timeline, but has itself little or no temporal extension. It is largely oblivious to the temporal characteristics of the current system state, the requirements and options for temporal function allocation, and the temporal effects of an allocation decision. Relying on snapshot information helps to keep the allocation decision computationally simple, but it also disconnects automation design considerations from another important class of workload management strategies: scheduling.

FROM DYNAMIC FUNCTION ALLOCATION TO DYNAMIC FUNCTION SCHEDULING

The focus of DFA research on automation is, at first glance, not surprising, given its origins in Function Allocation approaches. As a concept for guiding empirical research, DFA has provided important insights into the benefits and problems of adaptive automation, but unlike Function Allocation it has not yet been translated into a mature design approach (see, however, Hancock & Scallen, 1998, or Scerbo, 1996, for suggestions). Function Allocation design methods (e.g. Dearden, Harrison and Wright, 2000; Parasuraman, Sheridan and Wickens, 2000) may form the basis of such an approach, but to understand fully the design space for dynamic multitask environments, adaptive automation will have to be seen in the context of other operation strategies such as satisficing and scheduling.

A closer look at the literature reveals that this is just the view adopted by earlier multitask studies (e.g. Chu and Rouse, 1979; Tulga and Sheridan, 1980). With Sheridan and Verplank's (1978) much-cited paper on levels of automation appearing roughly at the same time, it is reasonable to argue for a historical link between automation and scheduling concepts. The variables considered in these early studies are just as relevant to current dynamic allocation research as they where then. They include task arrival rates, their predictability and distribution, the task's deadline, its duration, and the service rates of the available agents. Of particular importance is the notion of "value" as a measure of a task's relative importance or benefit, which allows the assessment of priority structures among concurrent tasks. Based on these parameters, scheduling and queuing theory can be used to optimise task servicing according to criteria such as, for example, maximising the number of tasks per time unit, minimising the number of tasks failing their deadline, minimising slack time, or maximising total value. As Tulga and Sheridan (1980), among others, have demonstrated, these normative engineering models also present a background against which human performance may be interpreted. Their particular appeal lies in making explicit the parameters of the decision procedure and the criteria for success. Since these early papers, our understanding of human performance on scheduling tasks has improved significantly (see Sanderson, 1989, for a review). However, this promising strand of Human Factors research has failed to enter into the DFA research and design agenda, so the challenge for contemporary workload management research is to integrate scheduling with other allocation options into a comprehensive design method.

DYNAMIC FUNCTION SCHEDULING

We use the term Dynamic Function Scheduling to refer to a perspective on work analysis that emphasises the temporal characteristics of multi-task environments and the temporal options for workload distribution (c.f. Hildebrandt and Harrison, 2002a). Scheduling involves value-based choices among concurrent functions, as well as value-based trade-offs in choosing among strategies for implementing a given function within its deadline (these strategies could, for instance, involve automation). To illustrate these concepts, consider the following four examples (for detailed analysis, see Hildebrandt and Harrison, 2002b):

(a) In a hospital, an expert system may be available for providing online diagnoses. Assume that the system's diagnosis may not always be as accurate as that of a senior doctor, but might significantly support an unaided junior staff member. If the condition of a patient is not time-critical, a 'manual' diagnosis by a senior doctor would be the strategy of choice, but in a time-critical situation or in the absence of such a doctor, staff may have to rely on the automation's advice.

Note that the decision to rely on the automation may be influenced by more subtle temporal factors. For instance, the junior doctor may arrive at a highly accurate diagnosis, but it may take a considerable amount of time. This time/quality trade-off may be compared with the expected performance of the automation to choose the strategy that best suits the treatment objectives and constraints. This would suggest that time might moderate the effects of trust (in the reliability of automation) on automation use: Under high time pressure, users might be more likely to accept the advice of an automatic system that is not perfectly reliable than when there is sufficient time to check the system's output.

(b) In a chemical factory, leaks may appear in the pipes connecting different reactors. To repair the leak, the operator may have to close the pipe, which would interrupt the production process. Depending on the size of the leak, the operator may decide to postpone the repair and tolerate the fluid loss until the current production task is finished and the leak can be repaired without affecting production. Other measures such as pumping off the spilled chemicals may be taken to contain the situation.

(c) Figure 1 illustrates a fault-servicing scenario for an aviation hydraulics system, where leaks may occur in any of the reservoirs or servos. Three strategies for dealing with the problem can be distinguished. The first strategy uniquely identifies the faulty element by going through a checklist procedure of selectively setting the valves connecting the reservoirs with the servos. The alternative strategy does not go through an initial diagnostic process, but switches to the redundant reservoir immediately. This fixes the problem quickly, but if the problem was located in one of the servos, then an intact reservoir has been disconnected and is unavailable. One or the other strategy may be used depending on the operational parameters (are there other high-value functions waiting to be serviced? May the reservoir be needed during the rest of the mission?). The pilot may even decide to take no action at all (i.e. drop the function) if the leak is unlikely to become critical before touchdown, and there are other urgent functions waiting.

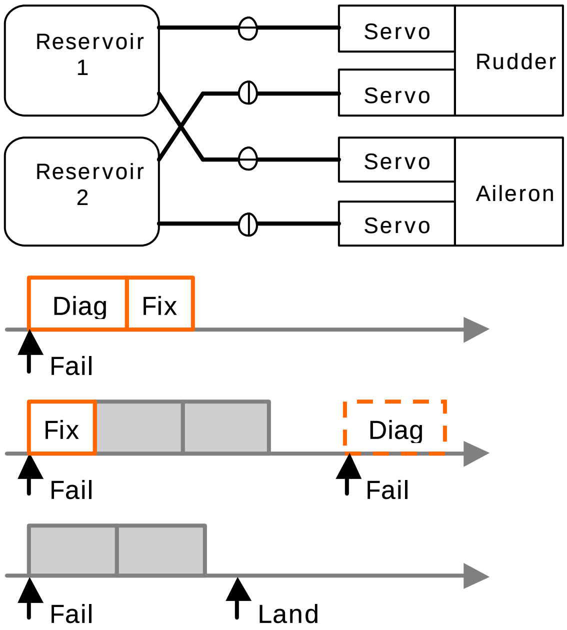

Fig. 1. Hydraulics system (top). Leaks can occur in the servos or reservoirs. Lower diagram shows scheduling options.

(d) Finally, consider the example of a supermarket checkout, where the prices are scanned by the checkout operator and the goods are packed by the customer (Fig. 2). If the customer cannot pack quickly enough, products build up waiting to be packed, which reduces the overall service rate of the checkout system. From the cashier's perspective, the value of the scanning function relative to the packing function is reduced. Therefore, the cashier may decide to postpone scanning to assist the customer with the packing, and resume scanning only after the backlog has been cleared.

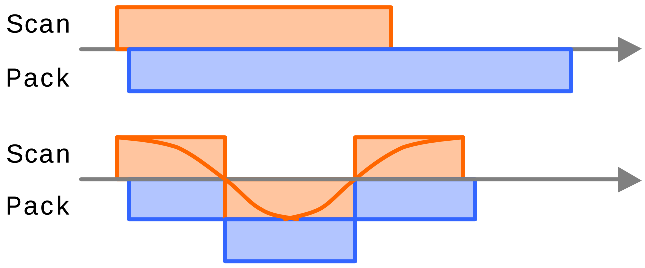

Fig. 2. Fixed (top) and flexible (bottom) scheduling, involving function postponement. Curved line represents value of scanning function

REPRESENTATION, ANALYIS AND DESIGN OF DYNAMIC FUNCTION SCHEDULING

The aim of Dynamic Function Scheduling is to make scheduling decisions and their associated trade-offs transparent to the designer, and to enable him/her to support the operator's online control decisions. This support may involve the visualisation of timing constraints, or the design of sequentially flexible functions, e.g. functions that allow pre-emption or interleaving (an important contribution in this area is McFarlane's "taxonomy of human interruption" and the associated empirical HCI research on different interruption methods; immediate, negotiated, scheduled and mediated interruption). Automation has a place in these considerations both as an operation strategy with particular temporal properties and as a support system for making online scheduling decisions. Formal analysis using scheduling or queuing models is a first step towards understanding these design choices, as it allows the representation of workload levels, their distribution, and the required and available processing time of different function servicing strategies. On a lower level of analysis, task notations with a set of temporal operators are required to express the compatibility of different strategies (one candidate may be the User Action Notation, c.f. Hartson and Gray, 1992).

However, in addition to formal analysis and representation methods, design of DFS needs to be informed by an understanding of the characteristics of human temporal cognition, and in particular temporal errors. As Hollnagel (1991) notes in his discussion of the phenotypes of human error, "[s]urprisingly, few of the existing action and error taxonomies include the aspect of time, but rather describe and classify human error on an atemporal (static) basis […] In many domains it is, however, necessary to include time in a much more conspicuous way, as, perhaps, one of the principal 'mechanisms' or 'error areas' of human action […] This is particularly true with respect to planning and scheduling." Human performance in temporal control tasks can deviate strongly from the normative model. These performance deficits may be related to low-level biases in the psychophysics of time perception, such as under or over estimation of durations. They may also be caused by decreased accuracy of decision-making due to time stress (see Edland and Svenson, 1993, for a review). Advances in the cognitive psychology of time (e.g. Block, 1990) have fostered an interest in higher-level aspects of temporal cognition among the human factors community (c.f. De Keyser, 1995). For instance, Moray et al. (1991), using Tulga and Sheridan's (1980) experimental paradigm, found evidence that operators faced with various tasks of different duration prefer to start the longest task first, even though the normative model would require the shortest task to be started first. In a multi-task micro world experiment (combined manual and supervisory control), Kirlik (1993) found evidence that operators are sensitive to the temporal costs of automation and refuse to use an automatic aid if its engagement is associated with a considerable delay. Both these studies discuss human performance data against the background of a normative model of the task and argue for the importance of combining empirical data and formal analysis. This approach is in the spirit of the early multitask studies mentioned above, and could serve as a blueprint for future Dynamic Function Scheduling research.

Finally, a growing literature on temporal factors in judgement and decision making suggests that conventional utility models, where costs and benefits are usually described in terms of money or similar commodities, may not be valid models for describing the perception of temporal costs. For instance, Soman (2001) found that the "sunk cost" effect (where a prior investment has an irrational influence on a current decision) disappeared when the investment was described in terms of time instead of money. Varey and Kahneman (1992) studied subjects' retrospective evaluation of extended periods of aversive experiences and found evidence that these judgements are influenced by the peak discomfort, the discomfort towards the end of the episode, and the trend over the episode, but not by the duration of the episode. If such effects translate into the Human Factors domain, they may bias the operators' assessment of the value of concurrent functions or the available function servicing strategies (e.g. various automation options).

CONCLUSION

This position paper argues for the introduction of a temporal perspective on Dynamic Function Allocation. A comprehensive understanding of workload management choices requires consideration of allocation along the human – automation resource dimension as well as the temporal dimension. Value-based scheduling and strategy selection enables both the designer and the operator to consider priority structures among functions and the trade-offs involved in different function servicing strategies. Automation has an important role to play in making these trade-offs explicit and in supporting scheduling decisions. A design method for DFS may combine representations of the temporal structure of functions with formal workload management approaches such as queuing models. These normative models have also provided a useful basis for interpreting empirical data on human scheduling performance. Future Human Factors research into Dynamic Function Scheduling can draw on a growing body of literature on the cognitive psychology of time to explore temporal control strategies.

REFERENCES

Block, R.A. (Ed.). (1990). *Cognitive Models of Psychological Time*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Chu, Y.-Y. & Rouse, W.B. (1979). Adaptive allocation of decisionmaking responsibility between human and computer in multitask situations. *IEEE Transactions on Systems, Man, and Cybernetics, 12*, 769-778.

Dearden, A., Harrison, M.D. & Wright, P.C. (2000) Allocation of function: scenarios, context and the economics of effort. *International Journal of Human-Computer Studies, 52*, 289-318.

De Keyser, V. (1995). Time in ergonomics research. *Ergonomics, 38*, 1639-1660.

Edland, A. & Svenson, O. (1993). Judgement and decisions making under time pressure. In O. Svenson & A.J. Maule (Eds.). *Time Pressure and Stress in Human Judgment and Decision Making* (pp. 27-40). New York: Plenum Press.

Endsley, M.R. & Kaber, D.B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, *42*, 462-492.

Hancock, P.A. & Scallen, S.F. (1998). Allocating functions in human-machine systems. In R.R. Hoffman, M.F. Sherrick & J.S. Warm (Eds.), *Viewing psychology as a whole: The integrative science of William M. Dember* (pp. 509-539). Washington, DC: APA Press

Hartson, H.R. & Gray, P.D. (1992). Temporal aspects of tasks in the User Action Notation. *Human-Computer Interaction*, *7*, 1-45.

Hildebrandt, M. & Harrison, M.D. (2002a). The temporal dimension of Dynamic Function Allocation. In S. Bagnara, S. Pozzi, A. Rizzo & P.Wright (Eds.), *Proc. of the 11th European Conference on Cognitive Ergonomics* (pp. 283-292). Rome: Consiglio Nazionale delle Ricerche.

Hildebrandt, M. & Harrison, M.D. (2002b). Time-related trade-offs in Dynamic Function Scheduling. In C. Johnson (Ed.). *Proc. of the 21st European Annual Conference on Human Decision Making and Control*. (pp. 89-95).

Hollnagel, E. (1991). The phenotype of erroneous actions: Implications for HCI design. In G.R.S. Weir & J.L. Alty (Eds.), *Human-Computer Interaction and Complex Systems* (pp 73-121). London: Academic Press.

Kirlik, A. (1993). Modeling strategic behaviour in human-automation interaction: Why an "aid" can (and should) go unused. *Ergonomics*, *35*, 221-242.

McFarlane, D. (1989). *Interruption of People in Human-Computer Interaction*. PhD thesis, George Washington University.

Moray, N., Dessouky, M.I., Kijowksi, B.A. & Adapathya, R. (1991). Strategic Behavior, workload, and performance in task scheduling. *Human Factors, 33*, 607-629

Parasuraman, R., Sheridan, T.B. & Wickens, C.D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics, 30*, 286-297.

Sanderson, P (1989). The human planning and scheduling role in advanced manufacturing systems: An emerging human factors domain. *Human Factors, 31*, 635-666.

Scerbo, M.W. (1996). Theoretical perspectives on adaptive automation. In R. Parasuraman & M. Mouloua (Eds.), *Automation and Human Performance: Theory and Applications* (pp. 38-63). Hillsdale, NJ: Lawrence Erlbaum.

Sheridan, T.B. & Verplank, W.L. (1978). *Human and computer control of undersea teleoperators* (Technical Report). Man-Machine Systems Laboratory, Department of Mechanical Engineering, MIT, Cambridge, MA.

Soman, D. (2001). The mental accounting of sunk time cost: Why time is not like money. *Journal of Behavioral Decision Making, 14*, 169-185.

Tulga, M.K. & Sheridan, T.B. (1980). Dynamic decisions and work load in multitask supervisory control. *IEEE Transactions on Systems, Man, and Cybernetics*, *10*, 217-232.

Varey, C. & Kahneman, D. (1992). Experiences extended across time: Evaluation of moments and episodes. *Journal of Behavioral Decision Making*, *5*, 169-185.