HUMAN COGNITIVE SIMULATION FOR EVALUATION OF HUMAN-ROBOT INTERFACE
A trade-off between flexibility in robot control and mental workload

Hiroshi Furukawa
Department of Risk Engineering, University of Tsukuba, 1-1-1 Tenoudai, Tsukuba, Japan

Keywords: mental workload, cognitive modeling, flexibility, adaptable automation, delegation, human-robot interaction, manual control, Playbook, multiple robots.

Abstract: Adaptable automation is a scheme that human operators can modify function allocations among human and machines (or robots) dynamically depending on situations. The concept is that operators should be able to delegate tasks to autonomous agents at times of their own choosing. Playbook is an example of a delegation architecture based on a team's book of approved plays that provides a "common language" for efficient and effective communication between human operators and agents. The author attended an empirical study examined the efficacy of Playbook interface using the Roboflag simulation platform. The results confirmed the benefits, compared to less flexible interfaces which are susceptible to negative effects due to suboptimal automation or unexpected events. This benefit was somewhat reduced, however, when the number of robots was increased. At this higher load, the benefit may have been reduced due to the greater workload demand imposed by full flexibility. This paper described a probabilistic simulation method to estimate behaviors of human operators as a tool for evaluating human-robot interfaces for operation of multiple robots. Through its application to the multiple robots simulation, advantages and costs of different design alternatives has been investigated in terms of cognitive workload indexes of the human operators. The results may suggest the validity of the hypothesis that there is a trade-off between flexibility in operational alternatives and operator's mental workload.

1 INTRODUCTION

Proper function allocation between humans and machines is one of the design requirements to optimize the advantages of automation, such as extraordinary precision, higher processing capabilities, and extension of the operator's perceptual and cognitive capabilities (Sheridan, 2000). The function allocation may foster human-machine interactions where "human performance and machine performance are not a zero-sum game, implying that the combination can be much better that either by itself" (p. 204, Sheridan, 2000).

To apply the function allocation strategy to actual system designs, it may be necessary to conduct quantitative evaluations of design alternatives that consider specific factors in task situations, such as time constraints, timing of actions, error rates, and limitation of resources. Particularly in complex systems, e.g. operations of multiple robots, most situations are not static but dynamic, and proper allocation may differ by situation. For example, cognitive experiment with such design alternatives is one of typical methods for the quantitative evaluation.

Adaptable automation is one of promising schemes for human interaction with automated agents, which allows human operators to modify function allocations among human and machines (or robots) dynamically depending on situations. The concept claims that human operators should be able to delegate (or not) tasks to autonomous agents at times of their own choosing, and receive feedback on their performance. There is growing number of evidence that shows efficacy of the scheme (Crandall & Goodrich, 2002; Parasuraman, Galster, & Miller, 2003).

Playbook is an example of the delegation architecture based on a team's book of approved plays that provides a "common language" for efficient and effective communication between human operators and agents (Miller, Pelican, & Goldman, 2000; Miller & Parasuraman, 2002). This

Furukawa H. (2005).
HUMAN COGNITIVE SIMULATION FOR EVALUATION OF HUMAN-ROBOT INTERFACE - A trade-off between flexibility in robot control and mental workload.
In *Proceedings of the Second International Conference on Informatics in Control, Automation and Robotics - Robotics and Automation*, pages 277-282
DOI: 10.5220/0001190702770282

"common language" affords human and agents a shared understanding of a goal to be achieved, and actions or methods for achieving the goal with varying levels of granularity.

Squire and his colleagues, including the author, conducted the cognitive experiment using the Roboflag simulation platform to examine the efficacy of Playbook interface using eight different control architectures (Squire, et al., 2004). The results confirmed the benefits provided by a flexible Playbook interface in which operators are empowered to delegate (or not) tasks to automated agents, compared to less flexible interfaces which are susceptible to negative effects due to suboptimal automation or unexpected events. This benefit was somewhat reduced, however, when the number of robots was increased. At this higher load, the benefit may have been reduced due to the greater workload demand imposed by full flexibility.

This paper shows the results of human cognitive simulation of Roboflag operators to validate the hypothesis that the performance benefit of the flexible interface may have been countered due to the greater management workload demand imposed by the flexibility in the condition with 8 robots, which the study suggested.

To achieve the goal, a method using human cognitive simulation was used as a framework for evaluating alternative designs of human-robot interface. Because of their ability to capture the dynamic interrelationships among system entities, events, and processes, computer simulation models have been recognized as an important class of techniques for systems design and function allocation (Sharit, 1997, Furukawa, et al., 2001, Furukawa, et al., 2004, Inagaki & Furukawa, 2004).

2 EXPERIMENTAL EVALUATION OF "PLAYBOOK" INTERFACE

This section describes the previous study which examined the efficacy of Playbook interface using the Roboflag simulation platform and the results about the benefits and costs compared to less flexible interfaces (Squire, et al., 2004).

2.1 Playbook interface

The empirical study investigated the two dimensions of adaptation flexibility offered by the Playbook: abstraction and aggregation. The abstraction can be thought of as in the multiple levels of a task hierarchy, where primitive robotic behaviors, such as waypoint-to-waypoint movement, are at the lower level of abstraction. More complex behaviors such as patrol border or defence (continuous planned movement action and reaction to events such as opponent attack) are at a higher level of abstraction. Aggregation can be defined as the number of robotic agents controlled as a group. Low aggregation refers to commands given to individual agents, whereas high aggregation refers to commanding all available agents with the same command. An intermediate level of aggregation is also possible where commands can be given to groups of robots smaller or equal to that of the whole team.

2.2 Comparative evaluation of alternative interface designs

The Roboflag simulation was used as the platform to examine the human performance effects using eight different types of human-robot interface, each corresponding to a combination of the abstraction and aggregation dimensions. The task of participants was to catch the opponent's flag using own robots against opponent's robots. The number of robots was equal on both side, i.e., either four or eight in this study.

Figure 1 depicts all possible interface combinations, and the eight experimental interfaces (#1 - #8) used in the study. Interfaces #1 - #4 represented highly restricted Playbook interfaces, and Interfaces #5 - #8 represented flexible Playbook interfaces. The three levels of abstraction (y-axis in Figure 1) included: waypoint control (user selects endpoint destinations, point to point guidance), plays (pre-programmed robotic behaviors, options being

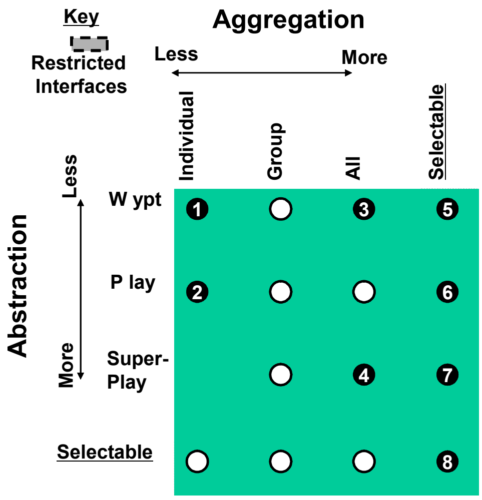

Figure 1: All possible interface combinations of the dimensions of abstraction and aggregation. (Squire, et al., 2004)

circle defense, circle offense, patrol border), and super-play (higher level group plays comprised of more than a single play, requiring selection of more than one robot, options being offense, defense, mixed). The three levels of aggregation (x-axis in Figure 1) were the selection of individual (one robot at a time), a sub-group (number selectable from zero to all), or all robots. Conditions within the highlighted portion in Figure 1 represent flexible control for either level of abstraction or level of aggregation. According to this taxonomy, Interface #8 had flexible control for both factors, whereas the other seven were progressively less flexible.

2.3 Results and hypothesis

Performance was superior with flexible interfaces for four robots, but this benefit was eliminated when eight robots had to be supervised.

Analysis of the subjective mental workload ratings revealed that the main effect of "Number of Robots" and "Interface Type" were significant, as was the interaction. Planned comparisons revealed that participants reported significantly higher mental workload when supervising eight robots compared to four robots (see Figure 2). At this higher load, the benefit of the flexible interfaces may have been reduced due to the greater workload demand imposed by full flexibility.

Also, participants reported higher workload for conditions where only waypoint control was available (individual waypoint, all waypoint, flexible waypoint) than when automated plays were available (individual play, all superplay, flexible play, flexible superplay).

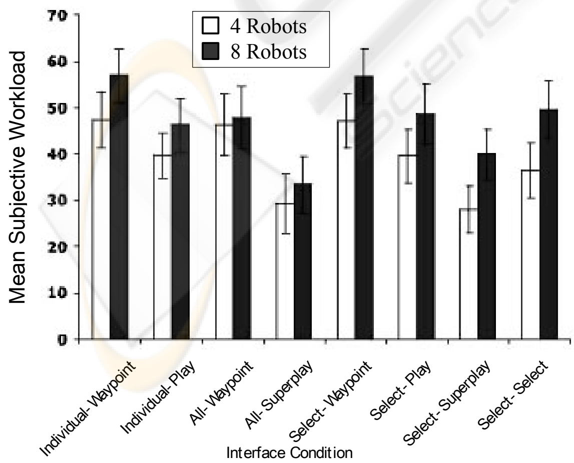

Figure 2: Mean subjective workload for the eight interfaces representing combinations of the abstraction and aggregation dimensions

3 COGNITIVE SIMULATION OF HUMAN-ROBOT INTERARCION

This section shows human cognitive simulation of human-robot interaction and the results of the simulation of Roboflag operators to validate the hypothesis about a trade-off between flexibility in operational alternatives and operator's mental workload.

3.1 Human cognitive simulation

A simulation code for estimating behaviors of human operators is developed to evaluate the eight types of interface designs. The reason for using simulation is that it is possible to have knowledge with variety of situations with performing parametric simulation. The parameters can be selected for describing states of robots, environment, and human operators. This approach may discover effective parameters that change system behaviors, which can not be revealed through empirical studies, and/or which are not predicted in advance.

Through cognitive task analysis, processes in the work are resolved into sets of cognitive tasks and paths, which are called *cognitive task network models* (Laughery, 1999). The granularity of one task should be set so that it is reasonable to be assumed that cognitive resources used by operators are unchanged during the task. Figure 3 shows the network model of robot control in Select-Select condition, developed for this research.

The cognitive simulation is based on task network models that consist of cognitive tasks by human operators, tasks assigned to machines, and sequential paths among them (Laughery, 1999).

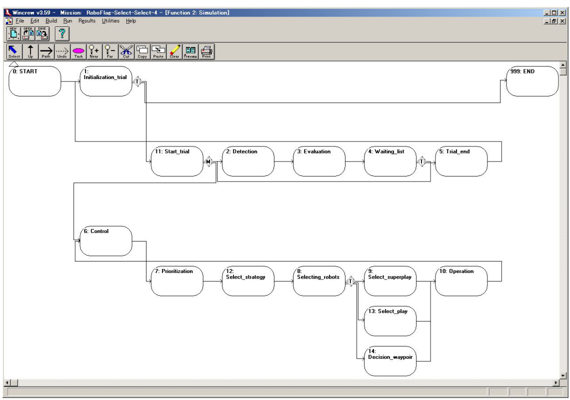

Figure 3: A cognitive task network of robot control in Select-Select condition

Task completion time for each task is estimated by a probabilistic model using a normal distribution. To avoid cases where negative values are provided, the time domain was limited to truncate the distribution. Momentary states of cognitive resources are simulated with the Multiple Resource Model proposed by Wickens and Yeh (Wickens & Yeh, 1986), in which five types of resources are defined: visual, auditory, cognitive, motor, and speech. Indexes describing cognitive workload, which are necessary to use the resources in achieving tasks, are assigned in a model using data from a reported database (Micro Analysis and Design, 1997). Monte Carlo simulations provide quantitative time data and total momentary workload indices based on the estimated cognitive resources.

3.2 Task network models for human-robot interaction

The task network model of the Roboflag operation was constructed considering three fundamental steps of the operators' cognitive work: state recognition, decision making, and operation. Fundamental information about necessary tasks in the work was defined through cognitive task analysis on the experimental data.

We considered two major factors in the model that would affect cognitive work across the different Roboflag interface types. One was the probability of the need for the user to intervene manually in controlling robots in order to ensure mission success. For example, an operator assigned whole robots an option of super-play, and then reassigns some of the robots another play after he/she recognized the former strategy was not appropriate or turned inappropriate because the situation had changed. We assumed that this probability would, by definition, be relatively high for the waypoint-only interface, lower for play operations, and much lower for super-play operations. Also, we assumed a higher probability when eight rather than four robots had to be supervised, because the time constraint must be highly related with the number. The actual values assigned for the probabilities are shown in Table 1.

Table 1: Probability of need for user's manual intervention in controlling robots (assumption in this study).

	The level of aggregation			
	All		Select	
# of robots	4	8	4	8
Waypoint	0.15	0.3	0.2	0.4
Play	0.075	0.15	0.1	0.2
Superplay	0.0375	0.75	0.05	0.1

The second factor was the operator's time for decision-making. We assumed that this would be shorter when the number of operational alternatives was small, and longer when the number was large. For example, in the Select-Select condition, an operator should choose from among three options, waypoint, play, or super-play, then select the number of robots to which the option should be applied, and finally execute the plan.

The cognitive model estimates relative indexes for mental workload of a human operator playing Roboflag simulation with one of the alternative interfaces.

3.3 Monte Carlo simulation

This simulation is implemented on a PC using WinCrew (Micro Analysis and Design, Inc.), a discrete event simulation-modeling tool (Laughery, 1999).

One thousand trials of Monte Carlo simulations were performed for each of the eight human-robot interfaces examined in the cognitive experiment with either four or eight robots. To compare the results under the different conditions, the simulated operational time was set equally at 60 seconds.

4 RESULTS AND DISCUSSION

Figure 4 shows expected values and the standard deviations of total time-integrated workload for each of the eight interfaces.

As expected, workload was higher when eight rather than four robots were supervised ($F(1, 15996)=60.297$, $p<.001$). This finding accords with the result of the previous experiment.

Also, workload was higher when only waypoint control was available (individual waypoint, all waypoint, flexible waypoint) compared to when automated plays could be used, with the exception that relatively low values were found for the "All-Waypoint" interface. Finally, workload was high in the "Select-Select" interface, particularly when eight robots were supervised. These simulation findings closely parallel those reported in the empirical data from the cognitive experiment (see Figure 2).

One exception is that the estimated workloads were much lower than the subjective data in the conditions that the level of aggregation is all (all waypoint, all superplay). A possible reason of this validation is that the assumption on the probability of the need for the user's manual intervention is not appropriate, where the probability would be relatively lower when the level of aggregation is "All" and higher when the level is "Select." To evaluate the possibility, the simulation was conducted with the condition where the probability in “Select” was used for that of “All” (see Table 1). Figure 5 shows that the validations were improved for both of “All-Waypoint” and “All-Superplay”. This result may suggest that the probability of the need for the user’s manual intervention does not depend on the level of aggregation.

The cognitive model was constructed based on empirical and objective data from experiments. The results of the simulations agree with the qualitative estimations about the interrelationships between factors in human interface design and mental workload that is necessary for human operators to achieve their tasks. This may suggest the validity of the hypothesis that the performance benefit of the flexible interface may have been countered due to the greater management workload demand imposed by the flexibility in the condition with eight robots.

It is clear that the strengths and weaknesses of each design alternative must be discussed quantitatively considering appropriate contributing factors. This case is also an example that the computer modeling technique can be used as a framework providing effective knowledge for improvement of designs of human-robot systems.

Though fidelity of models developed in this simulation is not high, it can be said that the models include appropriate factors to achieve the purpose of this simulation. Thus, this method can provide valuable information for preliminary designing of human-robot interface, such as necessary cognitive workload for each of operators when a type of interface is introduced in a simulation of multiple robots.

5 CONCLUSIONS

The probabilistic simulation method to estimate behaviors of human operators has been proposed as a tool for evaluating human-robot interfaces for operation of multiple robots.

Through its application to multiple robots game, Roboflag, with one of the eight human-robot interface based on Playbook design concept, advantages and costs of different design alternatives has been investigated in terms of cognitive workload indexes of the human operators. The condition parameters applied in this research are the number of robots and the types of human-robot interface. The results show that evaluation of interface is not straightforward, and that situation dependency and multiple indexes must be considered with great care (Olson & Goodrich, 2003, Goodrich & Olsen, 2003, Yanco, Drury & Sholtz, 2004).

ACKNOWLEDGEMENT

The author would like to express his thanks to Prof. Raja Parasuraman and Mr. Peter Squire, George Mason University, for informative discussions. This work has been partially supported by Grants-Aid for Science Research 16500076 of the Japanese Ministry of Education, Science, Sports and Culture.

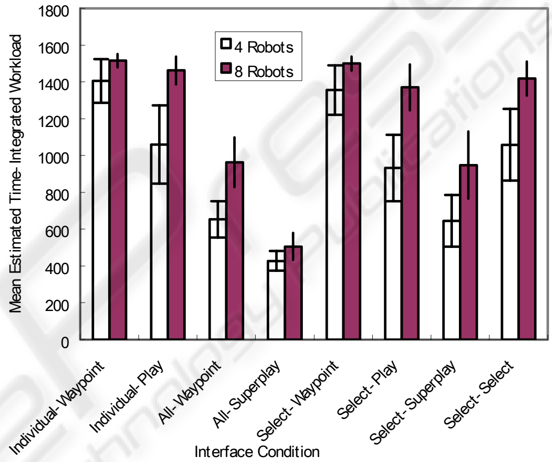

Figure 4: Mean expected values of time-integrated workload from the simulation analysis for the eight interfaces

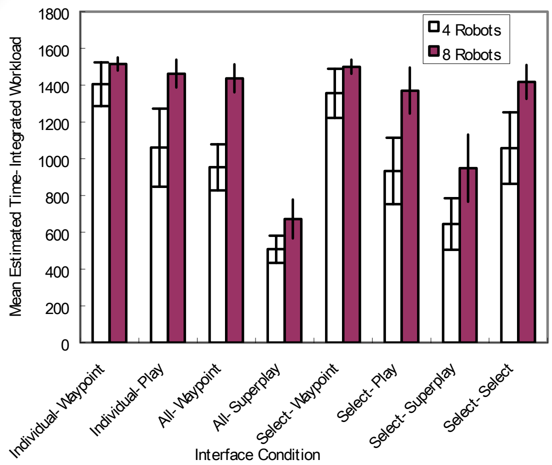

Figure 5: Mean expected values of time-integrated workload from the simulation analysis for the eight interfaces (results of re-simulations in All-Waypoint and All-Superplay conditions with new parameter settings).

REFERENCES

Sheridan, T., (2000). Function allocation: algorithm, alchemy or apostasy? *Int. J. Human-Computer Studies*, 52, 203-216.

Fitts P., (1951). Human engineering for an effective air navigation and traffic control system. *Ohio State University Foundation Report*, Columbus.

Crandall, J., & Goodrich, M., (2002). Principles of adjustable interactions. *Human-Robot Interaction, 2002 AAAI Fall Symposium* (A.C. Schultz, Chair). Menlo Park, CA: AAAI Press.

Parasuraman, R. Galster, S., & Miller, C. (2003). Human control of multiple robots in the RoboFlag simulation Environment. *Proceedings of the 2003 IEEE International Conference on Man, Systems and Cybernetics*. Washington, D.C.

Miller, C., Pelican, M., & Goldman, R. (2000). "Tasking" interfaces for flexible interaction with automation: Keeping the operator in control. *Proceedings of the Conference on Human Interaction with Complex Systems*. Urbana-Champaign, IL

Miller, C. & Parasuraman, R. (2002). *Designing for flexible human-automation interaction: Playbooks for supervisory control*, Technical Report, Minneapolis, MN: Smart Information Flow Technologies.

Squire, P., Furukawa, H., Galster, S., Miller, C., & Parasuraman, R., (2004). Adaptability and flexibility are key! Benefits of the "playbook" interface for human supervision of multiple unmanned vehicles. *Proc. Human Factors and Ergonomics Society 48th Annual Meeting*, New Orleans, USA, September, 61-65.

Sharit, J., (1997). Allocation of functions. In: Salvendy G. (ed) *Handbook of Human Factors and Ergonomics.* John Willy & Sons, New York, 301-339

Furukawa, H., Sato, M., & Inagaki, T., (2001). A probabilistic simulation method for evaluation of task allocation schemes in ship management. *Proc. Twentieth IASTED International Conference: Modelling, Identification and Control 2001*, Innsbruck, Austria, February, 533-538.

Furukawa H, Niwa, Y, & Inagaki, T. Optimization of task allocation through human cognitive simulation: levels of automation and human behaviors. *Proceedings of 7th International Conference on Probabilistic Safety Assessment and Management & European Safety and Reliability International Conference 2004*, Berlin, 260-265.

Inagaki, T., & Furukawa, H., Computer simulation for the design of authority in the adaptive cruise control systems under possibility of driver's over-trust in automation. *Proc. 2004 IEEE International Conference on Systems, Man & Cybernetics*, The Hague, Netherlands, October.

Laughery K., (1999). Modeling human performance during system design. Salas E (ed) *Human/technology interaction in complex systems.* JAI Press, Stamford, 147-174.

Wickens C,, & Yeh Y., (1986). A multiple resource model of workload prediction and assessment. *Proc. the IEEE Conference on SMC*, 1044-1048.

Micro Analysis and Design, (1997). *User's manual of WinCrew: Windows-based workload and task analysis tool*, Micro Analysis and Design.

Olson, D., & Goodrich, M., (2003). Metrics for evaluating human-robot interactions performance metrics for intelligent systems. *Proceedings of Performance Metrics for Intelligent Systems PerMIS '03*

Goodrich, M., &. Olsen, D, (2003). Seven principles of efficient human robot interaction. *Proceedings of the 2003 IEEE International Conference on Man, Systems and Cybernetics*.

Yanco, H., & Drury, J., & Sholtz, J., (2004). Beyond usability evaluation: Analysis of human-robot interaction at a major robotics competition. *Human-Computer Interaction*, vol. 19, 117-149.