Measuring Coordination Demand in Multirobot Teams

Michael Lewis
School of Information sciences
University of Pittsburgh
Pittsburgh, PA 15260

Jijun Wang
Quantum Leap Innovations, Inc.
3 Innovation Way, Suite 100
Newark, DE 19711

Conventional models of multirobot control assume independent robots and tasks. This allows an additive model in which the operator controls robots sequentially neglecting each until its performance deteriorates sufficiently to require new operator input. This paper presents a measure of coordination demand, CD, and experiments intended to extend the neglect tolerance model to situations in which robots must cooperate to perform dependent tasks. In the first experiment operators controlled 2 robot teams to perform a box pushing task under high coordination demand, teleoperation, moderate demand (waypoint control/ heterogeneous robots), and low demand (waypoint control/homogeneous robots) conditions. In the second experiment participants performed a search and rescue task requiring cooperation between robots creating maps and others carrying cameras. Measured demand and performance were largely consistent with the CD model's predictions.

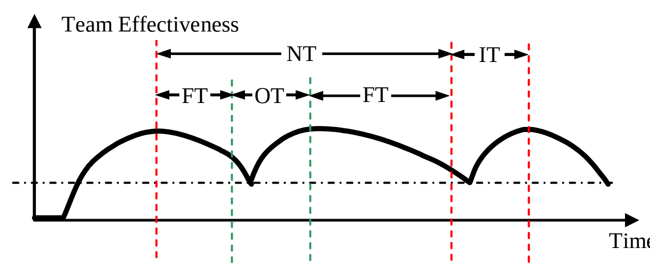
INTRODUCTION

The performance of human-robot teams is complex and multifaceted reflecting the capabilities of the robots, the operator(s), and the quality of their interactions. Recent efforts to define common metrics for human-robot interaction (Steinfeld et al., 2006) have favored sets of metric classes to measure the effectiveness of the system's constituents (human and robot) and their interactions as well as the system's overall performance. In this paper we present measures of the demand coordination places on operators of multirobot systems. Two experiments test the usefulness and validity of these measures.

In the simplest case of multirobot control an operator controls multiple independent robots interacting with each as needed. A search task in which each robot searches its own region would be of this category. Control performance at such tasks can be characterized by the average demand of each robot on human attention (Crandall et al., 2005).

For more strongly cooperative tasks and larger teams the round-robin control strategy used for controlling individual robots no longer works because the operator must predict and synchronize actions between multiple robots. Estimating the costs of this coordination, however, is difficult. Gerkey and Mataric (2004) have shown the computational complexity of choosing *n* out of *m* robots to perform a task, the iterated role assignment problem, to be of order, O(nm). Yet even such combinatorial arguments fail to capture the actual difficulty of controlling coordinating robots that arises because of the demands those robots impose on one another. In the box-pushing task used in our first experiment, for example, each movement by one robot demands a compensating movement by the other to maintain the box on course. Because these reciprocal demands cycle continuously an operator's attention may be fully occupied in controlling as few as two robots.

Established methods of estimating multirobot system (MRS) control difficulty, neglect tolerance and Fan-out (Crandall et al., 2005) are predicated on the independence of robots and tasks. In neglect tolerance the period following the end of human intervention but preceding a decline in performance below a threshold is considered time during which the operator is free to perform other tasks. If the operator services other robots over this period this model provides an estimate of Fan-out, the number of robots that might be controlled. This approach presumes that operating an additional robot imposes an additive demand on cognitive resources and that the performance threshold has a fixed value. These measures are particularly attractive because they are based on readily observable aspects of behavior: the time an operator is engaged controlling the robot, interaction time (IT), and the time an operator is not engaged in controlling the robot, neglect time (NT).



NT: Neglect Time; IT: Interaction Time;
FT: Free Time, time off task; OT: Occupied Time
IT+OT: time on task

Figure 1. Extended neglect tolerance model for cooperative robot control

The neglect tolerance model describes an operator's interaction with multiple robots as a sequence of control episodes in which an operator interacts with a robot for period IT raising its performance above some upper threshold after which the robot is neglected for the period NT until its performance deteriorates below a lower threshold when the operator must again interact with it. To accommodate dependent tasks we introduce occupied time, OT, to describe the time spent controlling other robots in order to synchronize their actions with those of the target robot. The episode depicted in Figure 1 starts just after the first robot is serviced. The ensuing FT preceding the interaction with a second dependent robot, the OT for robot-1 (that would contribute to IT for robot-2), and the FT following interaction with robot-2 but preceding the next interaction with robot-1 together constitute the neglect time for robot-1. Coordination demand, CD, is then defined as:

$$CD = 1 - \frac{\sum FT}{NT} = \frac{\sum OT}{NT} \qquad (1)$$

Where, CD for a robot is the ratio between the time required to control cooperating robots and the time still available after controlling the target robot, i.e.; the portion of a robot's free time that must be devoted to controlling cooperating robots. Note that OT_n associated with $robot_n$ is less than or equal to IT_n because OT_n covers only that portion of NT_n needed for synchronization. A related measure, team task demand (TAD), adds IT's to both numerator and denominator to provide a measure of the proportion of time devoted to the cooperative task; either performing the task or coordinating robots. While straight forward, this measurement is difficult to determine in practice since it requires discriminating between interactions with a second robot needed to coordinate it with the target robot (OT) and interactions with the second robot taken for its own benefit.

Most MRS research has investigated homogeneous robot teams where additional robots provide redundant (independent) capabilities. Differences in capabilities such as mobility or payload, however, may lead to more advantageous opportunities for cooperation among heterogeneous robots. These differences among robots in roles and other characteristics affecting IT, NT, and OT introduce additional complexity to assessing CD.

Where tight cooperation is required, as in box-pushing, task requirements dictate choice of robots, the interdependence of their actions, and the distinction between IT and OT since interactions are almost exclusively of type OT. In the more general case requirements for cooperation can be relaxed allowing the operator to choose the subteams of robots to be operated in a cooperative manner as well as the next robot to be operated. This general case of heterogeneous robots cooperating as needed characterizes the types of field applications our research is intended to support. To accommodate this more general case the Neglect Tolerance model must be further extended to measure coordination between different robot types. This leads to a more complicated expression of CD_i (Wang et al., 2008), coordination demand for robots of type i that can be interpreted as the CD a robot of type i exerts on other robot types. Again, if all actions are presumed to be in response to those of a cooperating robot CD_i can be determined from the interaction patterns.

In the first experiment robots perform a box pushing task in which CD is varied by control mode and robot heterogeneity. The second experiment attempts to manipulate coordination demand by varying the proximity needed to perform a joint task in two conditions and by automating coordination within subteams in the third.

CD EXPERIMENTS

USARSim and MrCS

Both experiments were conducted in the high fidelity USARSim (Lewis, Wang, & Hughes, 2007) simulator developed as a simulation of urban search and rescue (USAR) robots and intended as a research tool for the study of human-robot interaction (HRI) and multi-robot coordination. The MrCS (Multi-robot Control System, Figure 2), a multirobot communications and control infrastructure with accompanying user interface developed for experiments in multirobot control and RoboCup competition (Balakirsky et al., 2007), was used with appropriate modifications in both experiments. MrCS provides facilities for starting and controlling robots in the simulation, displaying camera and laser output, and supporting inter-robot communication through Machinetta (Scerri et al., 2004) a distributed mutiagent system. The distributed control enables us to scale robot teams from small to large.

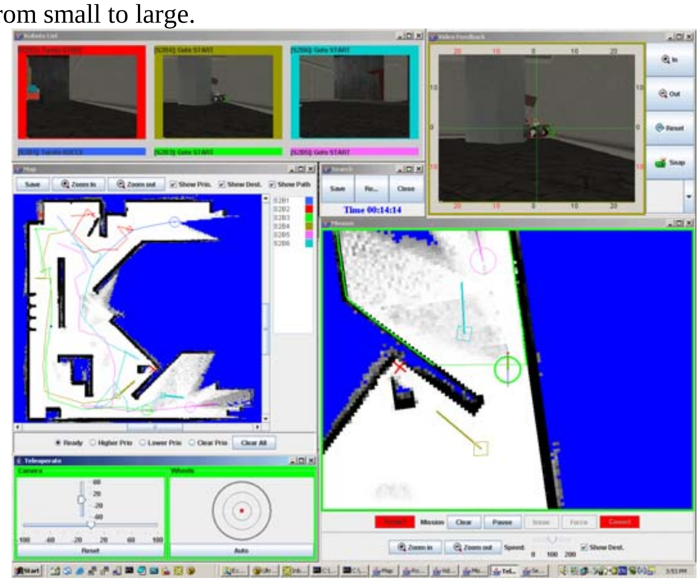


Figure 2 Standard MrCS GUI

Experiment 1

In this experiment, we investigated CD by comparing performance across three conditions selected to differ substantially in their expected coordination demands. We selected box pushing, a typical cooperative task that requires the robots to coordinate. When an operator teleoperates the robots one by one to push the box forward, he must continuously interact with one of the robots because neglecting both would immediately stop the box. Because the task allows no free time (FT) we expect CD to be 1. However, when the user is able to issue waypoints to both robots, the operator may have FT before she must coordinate these robots again because the robots can be instructed to move simultaneously. In this case CD should be less than 1. Intermediate levels of CD should be found in comparing control of homogeneous robots with heterogeneous robots. Higher CD should be found in the heterogeneous group since the unbalanced pushes from the robots would require more frequent coordination. In the present experiment, we measured CDs under these three conditions.

Figure 3. Box pushing task

Figure 3 shows our experiment setting simulated in USARSim. The controlled robots were either two Pioneer P2AT robots or one Pioneer P2AT and one less capable three wheeled Pioneer P2DX robot. When a robot pushes the box, both the box and robot's orientation and speed will change. Furthermore, because of irregularities in initial conditions and accuracy of the physical simulation the robot and box are unlikely to move precisely as the operator expected. In addition, delays in receiving sensor data and executing commands were modeled presenting participants with a problem very similar to coordinating physical robots.

We introduced a simple number matching task as a secondary task to allow us to estimate the FT available to the operator. Participants were asked to perform this secondary task as possible when they were not occupied controlling a robot. Every operator action and periodic timestamped samples the box's moving speed were recorded for computing CD .

A within subject design was used to control for individual differences in operators' control skills and ability to use the interface. To avoid having abnormal control behavior, such as a robot bypassing the box bias the CD comparison, we added safeguards to the control system to stop the robot when it tilted the box.

The operator controlled the robots using a modified multi-robot control system (MrCS) with separate teleoperation widgets for the left and right robots. The bottom center is a map based control panel that allows the user to monitor the robots and issue waypoint commands on the map. On the bottom right corner is the secondary task window where the participants were asked to perform the matching task when possible.

Participants and Procedure

14 paid participants, 18-57 years old were recruited from the University of Pittsburgh community. None had prior experience with robot control although most were frequent computer users.

The experiment started with collection of the participant's demographic data and computer experience. The participant then read standard instructions on how to control robots using the MrCS. In the following 8 minute training session, the participant practiced each control operation and tried to push the box forward under the guidance of the experimenter. Participants then performed three testing sessions in counterbalanced order. In two of the sessions, the participants controlled two P2AT robots using teleoperation alone or a mixture of teleoperation and waypoint control. In the third session, the participants were asked to control heterogeneous robots (one P2AT and one P2DX) using a mixture of teleoperation and waypoint control. The participants were allowed eight minutes to push the box to the destination in each session. At the conclusion of the experiment participants completed a questionnaire about their experience.

Results

Figure 4 shows a time distribution of robot control commands recorded in the experiment. As we expected no free time was recorded for robots in the teleoperation condition and the longest free times were found in controlling homogeneous robots with waypoints. The box speed shown on Figure 4 is the moving speed along the hallway that reflects the interaction effectiveness (IE) of the control mode. The IE curves in this picture show the delay effect and the frequent bumping that occurred in controlling heterogeneous robots revealing the poorest coordination performance.

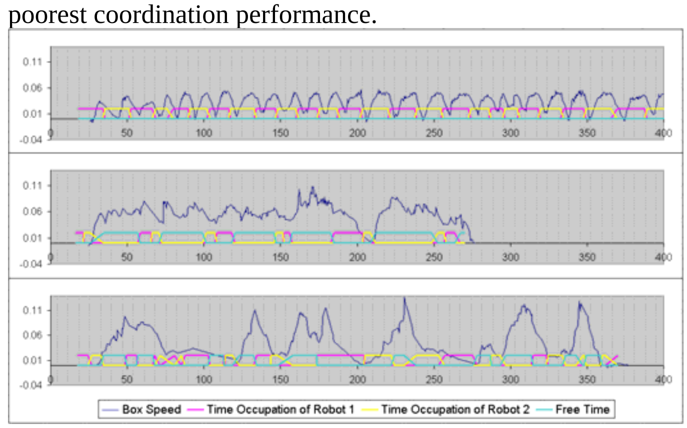


Figure 4 Time distribution curves for teleoperation (upper) and waypoint control (middle) for homogeneous robots, and waypoint (bottom) for heterogeneous robots

None of the 14 participants were able to perform the secondary task while teleoperating the robots. Hence, we uniformly find TAD=1 and CD=1 for both robots under this condition. Within participants comparison found that under waypoint control the team attention demand in heterogeneous robots was significantly higher than the demand in controlling homogeneous robots, $t(13)=2.213$, $p=0.045$. A two-tailed t-test shows that when a participant controlled a P2AT robot, lower CD was required in homogeneous condition than in the heterogeneous condition, $t(13)=-2.365$. $p=0.034$.

Experiment 2

To test the usefulness of the CD measurement for a weakly cooperative MRS, we conducted an experiment assessing coordination demand using an Urban Search And Rescue (USAR) task requiring high human involvement (Murphy & Burke, 2005) and of a complexity suitable to exercise heterogeneous robot control. In the experiment participants were asked to control *explorer* robots equipped with a laser range finder but no camera and *inspector* robots with only cameras. Finding and marking a victim required using the *inspector's* camera to find a victim to be marked on the map generated by the explorer. The capability of the robots and the cooperation autonomy level were used to adjust the coordination demand of the task.

Experimental design

Three simulated Pioneer P2AT robots and 3 Zergs (Balakirsky et al., 2007), a small experimental robot were used. Each P2AT was equipped with a front laser scanner with 180 degree FOV and resolution of 1 degree. The Zerg was mounted with a pan-tilt camera with 45 degree FOV. The robots were capable of localization and able to communicate with other robots and control station. The P2AT served as an *explorer* to build the map while the Zerg could be used as an *inspector* to find victims using its camera. To accomplish the task the participant must coordinate these two types of robot to ensure that when an *inspector* finds a victim, it is within a region mapped by an *explorer* so the position can be marked.

Three conditions were designed to vary the coordination demand on the operator. Under condition 1, the *explorer* had 20 meters detection range allowing *inspector* robots considerable latitude in their search. Under condition 2, scanner range was reduced to 5 meters requiring closer proximity to keep the *inspector* within mapped areas. Under condition 3, *explorer* and *inspector* robots were paired as subteams in which the *explorer* robot with a sensor range of 5 meters followed its *inspector* robot to map areas being searched. We hypothesized that CDs for *explorer* and *inspector* robots would be more even distributed under condition-2 (short range sensor) because *explorer*s would need to move more frequently in response to *inspectors'* searches than in condition-1 in which CD should be more asymmetric with *explorer*s exerting greater demand on *inspectors*. We also hypothesized that lower CD would lead to higher team performance. Three equivalent damaged buildings were constructed from the same elements using different layouts. Each environment was a maze like building with obstacles, such as chairs, desks, cabinets, and bricks with 10 evenly distributed victims. A fourth environment was constructed for training.

A within subjects design with counterbalanced presentation was used to compare the cooperative performance across the three conditions. The standard control interface shown in Figure 2 allowing participants to control robots through waypoints or teleoperation was used in all conditions.

Participants

19 paid participants, 19-33, years old were recruited from the University of Pittsburgh community. None had prior experience with robot control although most were frequent computer users. 6 of the participants (31.5%) reported playing computer games for more than one hour per week.

Procedure

After collecting demographic data the participant read standard instructions on how to control robots via MrCS. In the following 15~20 minute training session, the participant practiced each control operation and tried to find at least one victim in the training arena under the guidance of the experimenter. Participants then began three testing sessions in counterbalanced order with each session lasting 15 minutes. At the conclusion of the experiment participants completed a questionnaire.

Results

Overall performance was measured by the number of victims found, the explored areas, and the participants' self-assessments. To examine cooperative behavior in finer detail, CDs were computed from logged data for each type robot under the three conditions. We compared the measured CDs between condition 1 (20 meters sensing range) and condition 2 (5 meters sensing range), as well as condition 2 and condition 3 (subteam). To further analyze the cooperation behaviors, we evaluated the total attention demand in robot control and control action pattern as well. Finally, we introduce control episodes showing how CDs can be used to identify and diagnose abnormal control behaviors.

Overall performance

A paired t-test shows that in condition-1 (20 meters range scanner) participants explored wider areas, $t(16) = 3.097$, $p = 0.007$, as well as found more victims, $t(16) = 3.364$, $p = 0.004$, than under condition-2 (short range scanner). In condition-3 (automated subteam) participants found marginally more victims, $t(16) = 1.944$, $p = 0.07$, than in condition-2 (controlled cooperation) but no difference was found for the extent of regions explored.

Coordination effort

During the experiment we logged all the control operations with timestamps. From the log file CDs were computed for each type robot. The CD for *explorers* was found to be roughly twice the CD for *inspectors*. After the participant controlled an *explorer*, he needed to control an *inspector* multiple times or multiple *inspectors* since the *explorer* has a long detection range and large FOV. In contrast, after controlling an *inspector*, the participant needed less effort to coordinate *explorers*.

We predicted that when the *explorer* has a longer detection range, operators would need to control the *inspectors* more frequently to cover the mapped area. Therefore a longer detection range should lead to higher CD for *explorers*. This was confirmed by a two tailed t-test that found higher coordination demand, $t(18) = 2.476$, $p = 0.023$, when participants controlled *explorers* with large (20 meters) sensing range.

We did not, however, find a corresponding difference, $t(18)=.149$, $p=0.884$, between long and short detection range conditions for the CD for *inspectors*. This may have occurred because under these two conditions the *inspectors* have exactly the same capabilities and the difference in *explorer* detection range was not large enough to impact *inspectors'* CD for *explorers*. Under the subteam condition, the automatic cooperation within a subteam decreased or eliminated the coordination requirement when a participant controlled an *inspector*. Within participant comparisons shows that the measured CD of *inspectors* under this condition is significantly lower than the CD under condition 2 (independent control with 5 meters detection range), $t(18) = 6.957$, $p < 0.001$. Because the *explorer* always tries to automatically follow an *inspector*, we do not report CD of *explorers* in this condition. The action pattern, the ratio of control times between *inspector* and *explorer*, was found to be significantly larger under long sensing condition than under short range scanner condition, $t(18) = 2.193$, $p = 0.042$.

DISCUSSION

The Neglect Tolerance model for describing human-machine interaction as a series of periodic interventions needed to restore the system to proper functioning is conceptually very general and appears to apply across the board to a wide variety of human interactions with automated systems. While the model was presented in terms of regular cyclic interaction and neglect intervals, in practice these durations can vary greatly making predictions such as Fan-out (Crandall et al., 2005) an upper bound rather than realistic estimate. The model, however, can be extended to meet these objections using scheduling and queuing theory (Cummings et al., 2006) to accommodate the distributions of arrival and service times. Another difficulty involves the definition of a "task". Missions of the sort envisioned for robotic teams, for example, flying to a wilderness area and then searching for a lost hiker often consist of discernable phases with substantially different requirements. An accurate model would need to take these phases into account.

While NT based prediction of Fan-out provides an elegant way for describing multirobot control, it fails to account for the various ways in which robots might be controlled dependently or in the aggregate. In a wide variety of situations such as designating areas to be searched, paths to be followed, or targets to be sought, aggregated commands are both more direct and simpler than single robot alternatives. In addition most current work in coordinating robot teams whether it be biologically inspired control laws for swarms, optimizing path planners (Gerkey & Mataric, 2004), or BDI agents (Scerri et al., 2004) involve commands to an aggregate.

Our work to develop measures of coordination demand are part of a wider effort to extend the Neglect Tolerance model beyond independently controlled robots to aggregate commands and varying degrees of dependent multirobot control. We envision control over robotic teams that involves commands of all these different types with some issued to the entire team, such as a search area, others involving a single robot, such as verifying a reported victim, and still others to subteams such as our explorer/inspector dyads. We have started by extending the Neglect Tolerance model to allow us to evaluate coordination demand in applications where an operator must coordinate multiple robots to perform dependent tasks. Results from the first experiment that required tight coordination conformed closely to our hypotheses with the teleoperation condition producing CD=1 as predicted and heterogeneous teams exerting greater demand than homogenous ones. The CD measure proved useful in identifying abnormal control behavior as well (Wang et al., 2008). As most target applications such as construction or search and rescue require weaker cooperation among heterogeneous platforms the second experiment extended NT methodology to such conditions. Results in this more complex domain were mixed, although the predicted effects on CD for sensor range and autonomous coordination were found.

These experiments have demonstrated the utility of measuring the process of human-robot interaction as well as outcomes to diagnosing operator performance and identifying aspects of the task, particularly for multiple robots, that might benefit from automation. While our approach to measuring CD was supported in both experiments the second experiment suggests the need for more sophisticated measures that can take into account strategies and patterns of actions as well as their durations.

ACKNOWLEDGMENTS

This research has been sponsored in part by AFOSR FA9550- 07-1-0039 and AFOSR FA9550-08-1-0356.

REFERENCES

Balakirsky, S., Carpin, S., Kleiner, A., Lewis, M., Visser, A., Wang, J., & Zipara, V. (2007). Toward hetereogeneous robot teams for disaster mitigation: Results and performance metrics from RoboCup Rescue, *Journal of Field Robotics*, 2007.

Crandall, J., Goodrich, M., Olsen, D., & Nielsen, C. (2005). Validating human-robot interaction schemes in multitasking environments. IEEE Transactions on Systems, Man, and Cybernetics, Part A, 35(4):438–449.

Cummings, M., Nehme, C., & Crandall, J., (2006). Predicting Operator Capacity for Supervisory Control of Multiple UAVs, tech report, MIT.

Gerkey, B. & Mataric, M. (2004). A formal framework for the study of task allocation in multi-robot systems. International Journal of Robotics Research, 23(9):939–954.

Lewis, M., Wang, J. & Hughes, S. (2007). USARSim : Simulation for the Study of Human-Robot Interaction. Journal of Cognitive Engineering and Decision Making, 2007. 1(1): p. 98-120.

Murphy, R., & Burke, J. (2005). Up from the Rubble: Lessons Learned about HRI from Search and Rescue, in Proceedings of the 49th Annual Meetings of the Human Factors and Ergonomics Society.

Olsen, D. & Wood, S. (2004). Fan-out: measuring human control of multiple robots, in Proceedings of the SIGCHI conference on Human factors in computing systems. 2004, ACM Press: Vienna, Austria. p. 231-238.

Scerri, P., Xu, Y., Liao, E., Lai, G., Lewis, M., & Sycara, K. (2004). "Coordinating large groups of wide area search munitions," In D. Grundel, R. Murphey, and P. Pandalos (Ed.), *Recent Developments in Cooperative Control and Optimization*, Singapore: World Scientific, 451-480.

Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Schultz, A. & Goodrich, M. (2006). Common Metrics for Human-Robot Interaction, *2006 Human-Robot Interaction Conference*, ACM, March, 2006.

Wang, J. & Lewis, M. (2007) Human control of cooperating robot teams, 2007 *Human-Robot Interaction Conference (HRI'07)*, ACM.

Wang, J., Wang, H., & Lewis, M. (2008). Assessing Cooperation in Human Control of Heterogeneous Robots,Proceedings of the Third ACM/IEEE International Conference on Human-Robot Interaction (HRI'08), ACM.