A Systematic Approach to Predict Performance of Human–Automation Systems

Ayanna M. Howard, *Senior Member, IEEE*

***Abstract*—This paper discusses an approach for predicting system performance resulting from humans and robots performing repetitive tasks in a collaborative manner. The methodology uses a systematic approach that incorporates the various effects of workload on human performance, and estimates resulting performance attributes derived between teleoperated and autonomous control of robotic systems. Performance is determined by incorporating capabilities of the human and robotic agent based on accomplishment of functional operations and effect of cognitive stress due to continuous operation by the human agent. This paper provides an overview of the prediction system and discusses its implementation on a simulated rendezvous/docking task.**

***Index Terms*—Human–robot interaction, performance prediction, task allocation.**

I. Introduction

ONE OF THE key issues in human–robot interaction scenarios is determining which tasks are best done with humans, or robotic systems, or a combination of each. As human–robotic systems are increasingly deployed in various applications such as telesurgery, military applications, and personal care robots, there is a corresponding need to develop methods that optimally partition the task space to ensure mission success. Typically, the process of selecting an appropriate technique for evaluation of human and automated systems requires knowledge of the objectives of a task and a realistic environment in which to assess performance. In addition, assessment of systems having both human and robotic agents must focus on the capability of both agents. If the human operator is overloaded, but the human agent is still required to perform during a crisis, the system should be capable of estimating performance, accordingly and allow the redistribution of tasks such that the human can deal with the high-threat task, while the automated system tries to manage the more repetitive workload. Therefore, the first step in estimating system performance consists of two primary factors: 1) approximating the relative decline in performance associated with the constant mental/resource load required to complete a task and 2) quantifying how well an agent (whether human or automated) achieves a task.

Although research in human–robot performance assessment is expanding, an approach that integrates the contributions of both human and robot agents to estimate future performance has been addressed only to a limited extent. As such, a method is presented to enable systematic estimation of system performance for human–robot scenarios.

Manuscript received January 18, 2005; revised May 6, 2005. This work was carried out under a contract with the National Aeronautics and Space Administration and funded through the Draper Laboratory Research and Development Program. This paper was recommended by Associate Editor Z. Zdrahal.

The author is with the Human–Automation Systems Lab, School of Electrical and Computer Engineering Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: ayanna.howard@ece.gatech.edu).

Digital Object Identifier 10.1109/TSMCC.2007.897505

II. Background

Typically, research that focuses on performance assessment of systems having both human and robotic agents tends to disregard the capability of one of the agents. In [1], a human-centered approach is used to understand the role of human–robotic teamwork in future human space exploration missions. In this work, a framework is developed in which robots become functional tools that assist the human rather than replace the human operator. In this regard, the autonomy levels of the agents are adjusted to maintain system performance, which is associated with the requirement that agents always operate within established constraints, and are always responsive to human control. In [2] and [3], the focus is to optimize the overall performance by designing systems that use adjustable autonomy to dynamically change the autonomy of an intelligent agent. Different criteria are used to determine how the autonomy level, and thus, the performance of the system is adjusted, and ranges from using analysis based on human physiological responses [2] to determining autonomy level based on reasoning about the costs of decisions [3].

Currently, there is limited research that focuses on predicting system performance by considering both the human and robot as integral contributors to performance. For example, automated systems tend to not perform well in unexpected, high-crisis situations. In such cases, if the human operator is overloaded, the human agent is still required to perform in order to ensure mission success. Ideally, in such cases, there should be a redistribution of tasks such that the human can deal with the high-threat task, while the robotic system tries to manage the more repetitive workload. Performance should be measured as an aspect of both human and robotic system performance, i.e., the capability of the robotic system to implement tasks should be understood, as well as the human's ability to perform. Recent work [4] has focused on evaluating human and robot teams through an analytical framework that decomposes tasks into independent functional primitives. Currently, the performance analysis proposed is in a generalized form that presents a concept of how to perform performance evaluation, but does not provide validated experimental results nor does it discuss the type of metrics needed for evaluation. In [5] and [6], complementary research is presented that introduces taxonomies and metrics useful for human–robot performance evaluation. Fong *et al.* [6] attempt to address the wide dispersion found in this area and develop common metrics for task-oriented human–robot interaction in terms of five task categories dependent on the level of human interaction.

Although research in human–robot performance assessment is expanding, an approach that integrates the contributions of both human and robot agents to estimate performance has been only limitedly addressed. This work attempts to address these limitations by developing a systematic approach that incorporates the various effects of workload on human performance, and predicts system performance derived from allocation of tasks between human-controlled and autonomous robotic systems.

III. Methodology for Human–Automation System Performance (HumAnS)

This section presents a methodology for human–automation system performance (HumAnS) that evaluates the various effects of workload on human performance, and estimates performance derived from task allocation between human-controlled and automated systems [4], [7]. HumAnS consists of four primary steps: 1) decompose scenario into set of major functional task primitives and define performance metrics for each primitive; 2) estimate the performance of all agents (human, robot) in performing each task primitive; 3) calculate a performance score based on satisfaction of task primitives and effect on agents (i.e., is human agent stressed?); and 4) compute a composite task score to enable tradeoff studies to be made for allocation of tasks between humans and robots.

A. Scenario Decomposition

Perhaps the greatest contributor to human error in many systems is the extensive workload placed upon the human operator [8]. Workload studies are used to characterize human performance in terms of total mental demand placed on a person implementing a task. Developing a methodology to assess workload using actual human subjects is a time consuming process, which must adequately deal with the inherent discrepancies found in the different subjects. To address this limitation, research efforts have focused on developing workload assessment models without the use of human subjects [9]. These efforts focus on decomposing tasks into a series of subtasks [10], and assigning workload values by pairwise comparing the level of effort required to implement each subtask. Following this approach, human–robot scenarios are first decomposed into a set of functional task primitives, i.e., activities that need to be implemented by the human or the robotic system for goal achievement. Different criteria can be used to determine how a human–robot scenario is decomposed, including analyzing task parallelism, temporal sequences, or spatial resolution [11]. For this work, the focus is on decomposing scenarios into nonconcurrent tasks that can be executed sequentially.

The method used for scenario decomposition is based on the task diagram interview sequence developed as part of the applied cognitive task analysis (ACTA) technique [12]. ACTA provides practical methods to identify mental demands placed on a human operator while performing tasks within a given scenario. This work, which is based on the task diagram interview technique in which a broad overview of a scenario is constructed by interviewing subject matter experts, follows the same concept to decompose human–robot scenarios using the following processes.

1) Break the scenarios into three to six functional primitives based on commonalities found in robotic systems operating in the real world.[1]
2) Determine cognitive skills/mental demand derived from functional primitive.
3) Create task diagram to give broad overview of mental demand derived from scenario.

TABLE I
Elementary Functional Primitives and Activity Type

	Primary Activity Type		
Primitives	**Motor**	**Cognitive**	**Sensory**
Grasp/Release	X		
Identify			X
Lift/Unload	X		
Locate/Localize		X	
Mate/Unmate	X		
Model/Represent		X	
Plan		X	
Track			X
Traverse	X		

In other works [4], [7], [13], an inclusive set of functional primitives in various robotic scenarios was constructed for assessing system performance. Using this as a basis, an elementary set of functional primitives is constructed, and the cognitive skills associated with each are identified. To identify cognitive skills, the cognitive architecture construct [14] is used to break human information processing into three macrolevel mechanisms: perception, cognition, and motor activities. Primitives are then selected to be as independent from each other as possible, and to emphasize different aspects of cognitive, motor, and sensory (i.e., perception) skills associated with mental demand (Table I).

Table I depicts the elementary functional primitives that classify various robotic scenarios. To shift this representation into the space of human–robot scenarios, the primitive set is expanded into more specific human–robot applicable operations, and those with similar characteristics are grouped into one primitive representation (Fig. 1). To determine similarity in characteristics, general low-level control steps are instantiated based on work in [15] and [16]. As an example, “mate with stationary object” (a primitive expansion of the mate/unmate functional primitive) involves the general control steps: 1) align with the object and 2) push toward the object. On the other hand, “grasp stationary object” involves the general control steps: 1) align with the object and 2) grip the object. Although these steps implement different control algorithms, the similarities in operations (that of two relevant motor activities) place these two expanded primitives into the same class. This expansion and regrouping exercise is implemented for the class of elementary functional primitives depicted in Table I. Fig. 1 shows a subset of the expansion/regrouping stage for extracting human–robot related functional primitives. The resulting functional primitives

[1] Scenario decomposition is kept to between three and six primitives to match with the task diagram technique [12] in which tasks are limited to between three and six steps in order to ensure that time is not wasted digging into small details.

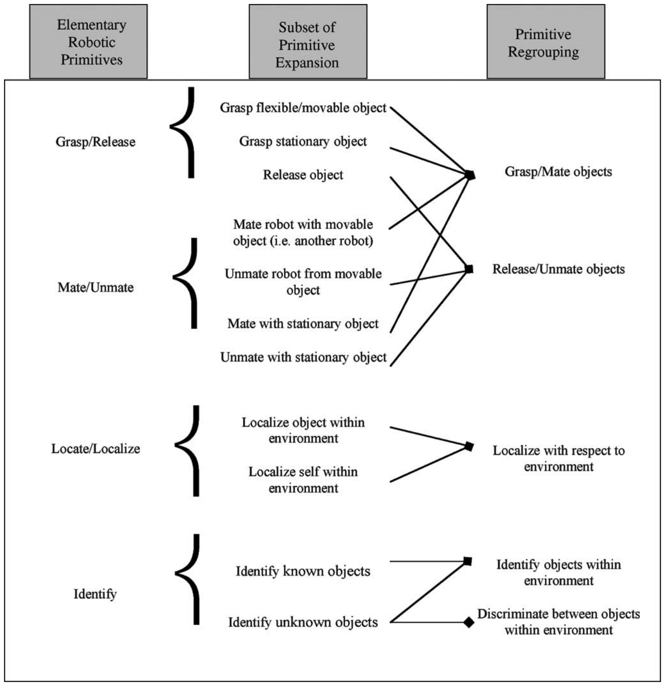


Fig. 1. Shifting robotic functional primitives into the space of human–robot scenarios.

TABLE II
MOTOR ACTIVITY

Functional Primitive
No Motor Activity
Traverse within environment
Grasp/Mate objects
Release/Unmate objects

TABLE III
COGNITIVE ACTIVITY

Functional Primitive
No Cognitive Activity
Localize with respect to environment
Represent/Model objects within environment
Represent/Model environment
Plan path within environment

TABLE IV
SENSORY ACTIVITY

Functional Primitive
No Visual Activity
Identify objects within environment
Discriminate between objects within environment
Visually track/follow object

based on activity type are then classified into three tables of relevance to human–robot scenarios (Tables II–IV).

By utilizing this redefined set of functional primitives, task scenarios can be defined by linking the primitives into a primitive hierarchy (or tree), which provides the necessary task diagram information for human–robot scenarios. This process

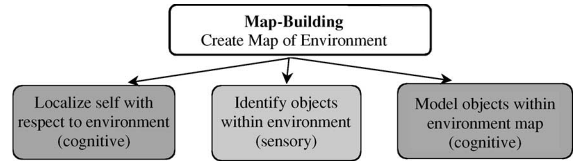


Fig. 2. Primitive hierarchy for map-building scenario.

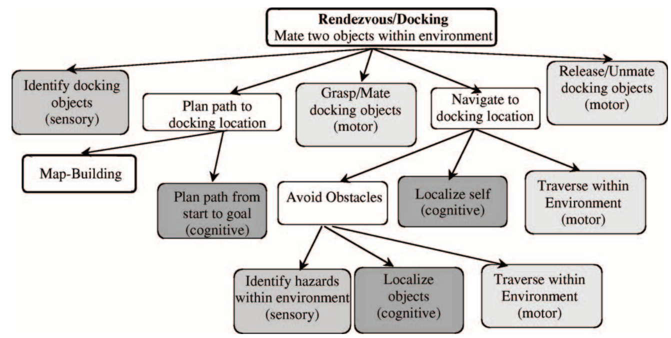


Fig. 3. Primitive hierarchy for rendezvous/docking scenario.

provides a broad understanding of the cognitive skills/mental demands required for each scenario. As an example, two primary scenarios of relevance to automation are shown in Figs. 2 and 3. Each scenario is decomposed into its lowest level, such that the last node (or leaves) of the primitive hierarchy consists solely of the functional primitives identified in the activity type tables listed above.

B. Calculating Performance Metrics for Human–Robot Scenarios

To determine the overall mental effort required to complete a task scenario, workload values must be calculated for each functional primitive that exists in the scenario decomposition. For this work, there is also an interest in the performance of both human and robot agents in implementing each identified task primitive. Performance metrics are therefore defined as consisting of both workload values and execution time components. Workload values are used to determine the relative decline in performance associated with the constant mental/resource load required to complete the task, while execution times quantify how well the agent achieves the primitive operation. Performance metrics are calculated for each elementary functional primitive, and use relative ranking measures for human and robotic agents based on a pairwise comparison method. To calculate human performance metrics, values are extracted by pairwise comparing the level of effort required to implement each subtask based on the calculated performance values. Execution ranking times for robotic systems are determined from evaluation of current robotic systems implemented in real time.

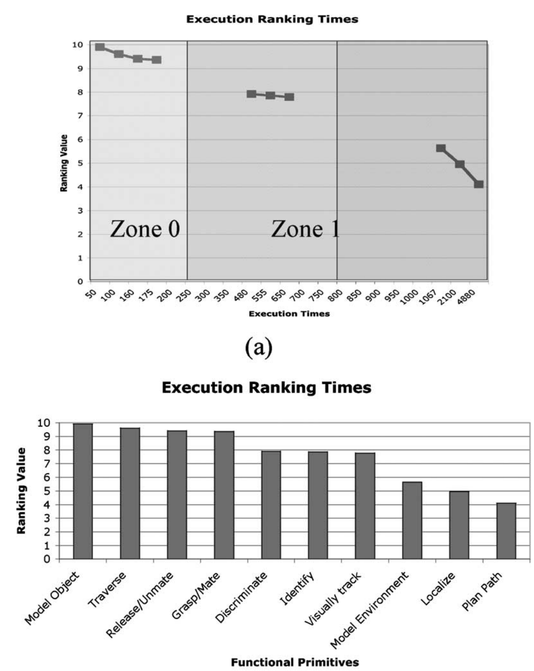
ACT-R [17] is a framework that models how human cognition works in various task scenarios based on assumptions derived from psychology experiments. Various researchers have utilized the ACT-R infrastructure to predict performance of human participants performing particular tasks in complex environments. The ACT-R model has been successfully used and

TABLE V
HUMAN PERFORMANCE DATA ASSOCIATED WITH FUNCTIONAL PRIMITIVES

ACT-R (Human) Primitives	Functional Primitives	Activity Type	Environmental Complexity	
			Low	High
Move to new location and punch key or mouse [18]	Manipulate new object (grasp/mate)	Motor	175 ms	f(distance)
Punch key or mouse click [18]	Manipulate object (release/unmate)	Motor	160 ms	
Move mouse to new location [18]	Traverse within environment	Motor	100 ms	f(distance, direction)
Encode Visual Position of Single Object [19]	Model object within environment	Cognitive	50 ms	
Visually code map/maze [20]	Model environment	Cognitive	1067 ms	
Plan path through encoded map [20]	Plan path within environment	Cognitive	4.88 sec	14.80 sec
Establish correspondence between two views of a space [21]	Localize with respect to environment	Cognitive	2.1 sec	6 sec
Visually tracking multiple objects – switching time [22, 23]	Visually track moving object	Sensory	650 msec	1100 msec
Recognize object in a scene [24]	Identify object within environment	Sensory	550 msec	
Search among objects for specific object [24, 25]	Discriminate between objects	Sensory	480 msec	3500 msec

TABLE VI
ASSOCIATING PERFORMANCE METRICS WITH FUNCTIONAL PRIMITIVES

Workload Ranking Value (0-10)	**Execution Ranking Time (0-10)**		**Functional Primitives**
Human		Robot	
			Motor Activity
0.0	10.0	10.0	No Motor Activity
4.0	9.6	5.0	Traverse within environment
2.0	9.4	9.0	Release/Unmate objects
3.0	9.4	8.0	Grasp/Mate objects
			Cognitive Activity
0.0	10.0	10.0	No Cognitive Activity
1.0	9.9	8.5	Represent/Model objects within environment
7.0	5.6	3.0	Represent/Model environment
9.0	5.0	2.0	Localize with respect to environment
10.0	4.1	3.5	Plan path within environment
			Sensory Activity
0.0	10.0	10.0	No Visual Activity
4.0	8.0	6.0	Discriminate between objects within environment
3.0	7.9	7.0	Identify objects within environment
3.0	7.8	7.0	Visually track/follow object



(a)

(b)

Fig. 4. (a) Zone of operations for performance data. (b) Graph depicting execution ranking times computed from performance data.

compared with traditional measures of cognitive psychology, such as the time required to perform a task, accuracy in a task, and neurological performance data. Based on an analysis of ACT-R performance data derived from various task scenarios, execution times are extracted for the functional primitives defined for each activity type previously discussed. Table V depicts ACT-R primitives associated with human activities, and links them with functional primitives defined for the scenario decompositions. Associated with each primitive are the approximate time values of human performance based on published research results. In some cases, execution times depend on the environmental complexity in which the task is performed. The table reflects this dependence in columns 4 and 5.

Given these approximate values of human performance, a comparison rating is constructed to derive performance metrics using relative values in the [0.0, 10.0] range, where 0.0 represents the minimum ranking value and 10.0 is the maximum ranking value (Table VI). To compute the ranking for execution time, the performance data (Table V) is first segmented into three zones of operation [40 ms, 250 ms], [251 ms, 800 ms], and [801 ms, 5500 ms]. This segmentation naturally corresponds to the execution time granularity inherent in the performance data set (Fig. 4). As execution times are used to establish a metric to compare both human and robot agents, it is assumed that 1) the robot agent performs no better than the human agent (given the human's greater capability to deal with anomalies within the environment); 2) ranking times for the robotic system can be derived from comparing current robotic systems implemented in real time [7], [13] to human performance; 3) a ranking time of 4.0 is the minimum ranking value (i.e., slowest execution time) associated with a human agent; and 4) "no [motor/cognitive/visual] activity" is associated with the maximum ranking time of 10.0, as it requires the least time for execution (i.e., idle activity time is estimated at 40 ms). Based on these assumptions, the boundary conditions for the execution times are then established such that 40 ms corresponds to a ranking time of 10.0, and 5.5 s corresponds to a ranking time of 4.0. For each zone of operation, a logarithmic function representative of the performance data granularity is also computed, such that

$$Z_0 : \mathit{ExecutionRank} = R_m - \log(T) + \varepsilon_0 \quad (1)$$

$$Z_1 : \mathit{ExecutionRank} = R_m - \log(T) + \varepsilon_1 \quad (2)$$

$$Z_2 : \mathit{ExecutionRank} = R_m - \ln(T) + \varepsilon_2 \quad (3)$$

where Z_n represents one of the three zones of operation, *ExecutionRank* is the execution ranking time, R_m represents the maximum ranking time of 10.0, T represents the minimum time values associated with each functional primitive (from Table V), and ε_n is determined based on the boundary conditions, such that

$$R_m - \ln(40) + \varepsilon_0 = 10.0 \Rightarrow \varepsilon_0 = 1.6 \quad (4)$$

$$R_m - \ln(5500) + \varepsilon_2 = 4.0 \Rightarrow \varepsilon_2 = 2.6 \quad (5)$$

$$\frac{\|Z_0\|}{\|Z_1\|}\varepsilon_0 = \varepsilon_1 \Rightarrow \varepsilon_1 = 0.6 \quad (6)$$

where Z_0 and Z_1 represent the difference between the starting and ending time values associated with the respective zone of operation. Fig. 4 depicts the ranking values calculated from the performance data that is represented in Table VI.

Workload represents the cognitive load placed on the human operator while performing a task. It is designed to increase with regard to the amount of time a human requires to implement a task. The workload ranking value is thus determined based on the execution time and the decrease in performance associated with environmental complexity, such that

$$\begin{aligned} &Z_0, Z_1: \\ &\quad \textit{If}\,(E_h - E_l = 0), \textit{Workload} = \log(T - T_0) \\ &\quad \textit{Else Workload} = \log((T - T_0) + (E_h - E_l)) \quad (7) \\ &Z_2: \\ &\quad \textit{If}\,(E_h - E_l = 0), \textit{Workload} = \ln(T - T_0) \\ &\quad \textit{Else Workload} = \ln((T - T_0) + (E_h - E_l)) \quad (8) \end{aligned}$$

where Z_n represents one of the three zones of operation, *Workload* is the workload ranking value, T represents the minimum time values associated with each functional primitive, T_0 is the idle activity time, E_l and E_h are the execution times associated with *low* and *high* environmental complexity, respectively. It is assumed that for execution times dependent on distance and direction, the operator will be required to move to a position at least ten times the minimum. Table VI depicts the workload ranking values calculated from this assessment.

C. Performance Evaluation for Human–Robot Scenarios

In multi-agent coordination, dynamic task allocation involves determining which agent should execute which task at which time in order to achieve a global goal. The optimal assignment problem [26] addresses this problem by estimating how well each robot agent can be expected to perform each task given a system of n robots and m prioritized single-robot tasks. Markov decision processes (MDP) allow each agent to choose individual actions based on maximizing an optimization function for the entire system [27]. This concept of an optimization function is utilized to calculate a composite task score using the detailed functional decomposition of a task scenario. The optimization function incorporates both attributes of workload ranking values and execution ranking times. To determine the effect of workload on performance, the work of Dinges and Mallis [28] was examined. In [28], studies were performed to examine the capability of various techniques for determining the relationship between various validation criteria (such as percentage of eye closure, brain wave activity, head positioning, etc.), and the performance of a human operator executing complex tasks

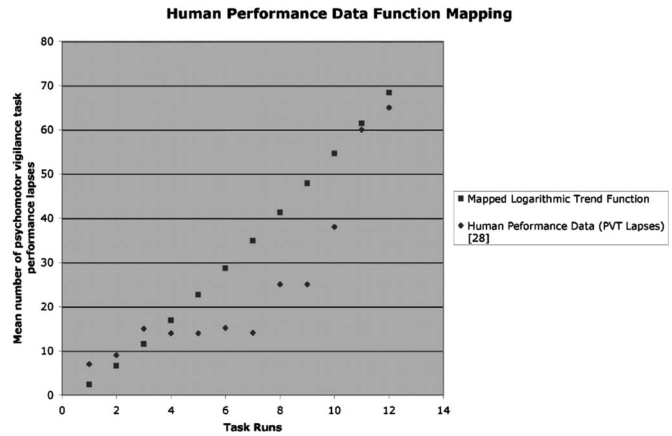


Fig. 5. Mapping human workload effect using a logarithmic function.

over time. To evaluate each algorithm, experimental data were extracted to monitor the performance of human operators performing tasks every 2 h over a 42-h time cycle. In order to determine the relative effect of workload on performance, a logarithmic function was roughly mapped to obtain a time-dependent performance trend associated with human task implementation (Fig. 5). This trend reflects the effect of workload in the optimization function. For each task scenario, a composite task score can thus be constructed to estimate overall system performance while incorporating the decreases in performance associated with consistent work operation, such that $\forall$ primitive $= 1{:}\,n$

$$r = \text{Execution Rank(agent)agent} = \{\text{human, robot}\}$$

$$w = \{\begin{cases} \dfrac{s\ln(s \times \text{workload})}{\text{workload}} & \text{agent} = \text{human} \\ 0, & \text{agent} = \text{robot} \end{cases}$$

$$\text{Composite Task Score(s)} = \sum_{i=1}^{n} (\rho_i - \omega_i)$$

where s designates the repetitive number of scenario runs that have occurred, *Workload* is the workload ranking value associated with primitive i, and *ExecutionRank* is the execution ranking time associated with primitive i. The composite task score is summed over all functional primitives for the task scenario, and can be calculated for each repeated scenario run. A final composite task score provides an overall evaluation of the relative performance for the scenario.

IV. Results for Evaluation of Human–Robot Scenarios

A. Test Environment for Human–Robot Scenarios

The first rendition of the test environment (Fig. 6) consists of a graphical user control panel that enables the human operator to control a robot operating in the real world. For this work, the Sony ERS-210 robot is utilized for task implementation. The control panel allows the human operator to view the world through the robot's eyes, as well as command the robot to move forward, backward, and turn either left or right. The human operator can also toggle between teleoperated control and autonomous behavior of the robot.

Fig. 6. Virtual environment consisting of human operator unit and 3-D World Environment.

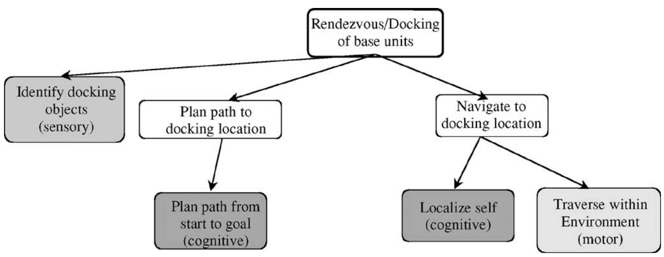


Fig. 7. Primitive hierarchy for simplified rendezvous/docking scenario.

TABLE VII
PERFORMANCE METRICS ASSOCIATED WITH SIMPLIFIED TASK DECOMPOSITION

Workload Value (0-10)	Performance Score (0-10)		Functional Primitives
	Human	Robot	
			Motor Activity
4.0	9.6	5.0	Traverse within environment
			Cognitive Activity
9.0	5.0	2.0	Localize with respect to environment
10.0	4.1	3.5	Plan path within environment
			Sensory Activity
3.0	7.9	7.0	Identify objects within environment

B. Performance Evaluation of Human–Robot Scenarios

To validate the HumAnS prediction system, the methodology is applied to estimate the performance of human and robotic agents performing a simulated rendezvous/docking task. The first step in the HumAnS process is to decompose the task into functional operations, and associate relative performance scores and workload values to each operation. As the focus is on documenting the applicability of HumAnS to a representative task scenario, branches of the rendezvous/docking scenario are pruned in Fig. 3 to involve two primary operations: locating a target base unit in an obstacle-free environment, and navigating to a position for subsequent transportation of the base units into a desired configuration. Fig. 7 displays the primitive hierarchy associated with this simplified task, and the corresponding performance metrics for the functional primitives in this task are depicted in Table VII.

In the current analysis, the performance of human agents is directly compared to the performance of robot agents in the rendezvous/docking scenario (Fig. 8). The two setups constructed for assessment are 1) direct teleoperated control of the robot by a human operator (via the graphical user interface) and 2) fully autonomous control of the robot, without direct human intervention. The autonomous behavior programmed onto the robot allows the robot to search for and locate the target base unit within the environment, and navigate toward the corresponding goal position for subsequent transportation of the base unit. This involves implementing a vision-based algorithm to locate the base unit via color information and extracting object size and location from the image data [29]. This information is then fed into a stored table that associates the two extracted image parameters with 3-D world position to which the robot is directed.

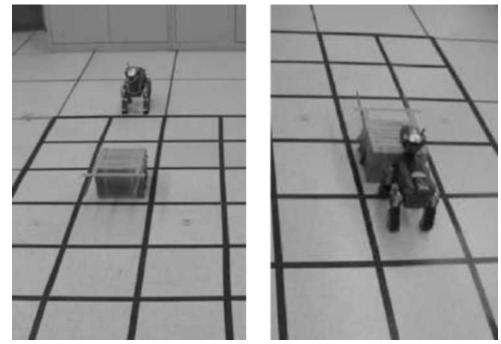


Fig. 8. Initial and final robot configuration for rendezvous/docking scenario.

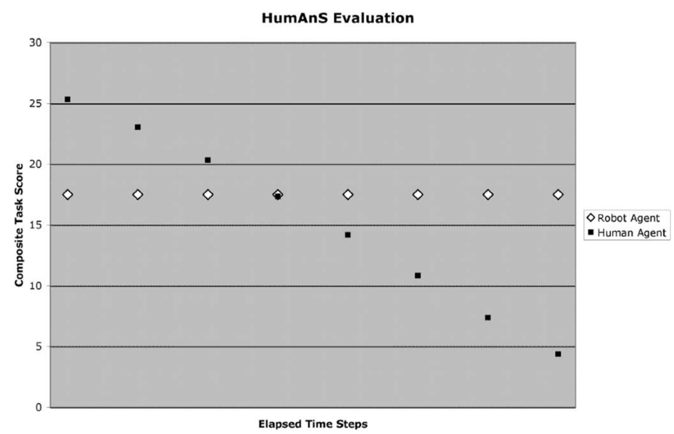


Fig. 9. Composite task scores calculated by HumAnS.

Fig. 9 documents the composite task score calculated by HumAnS for each setup based on the performance metrics and workload values extracted from Table VII.

To compare the prediction results with real-world implementation, the testing process consists of running through each scenario ten times (with the target base unit located at different sites) for six to ten continuous runs (i.e., time steps) for each scenario, and documenting the execution time. Fig. 10 shows the resulting outcome from two of the sample runs.

To map execution time to composite task score, the elapsed time steps are correlated and scaled to execution ranking times that match with the composite task score calculated for the robot agent. The time steps are selected to begin after the learning cycle for each scenario run (typically, the first two to four time steps). This process is acceptable because the prediction system is interested in understanding the *relative* performance of humans versus robots, and in capturing the corresponding decline in human performance associated with workload during real-time implementation. Implementing this normalization process results in the outcome depicted in Fig. 11.

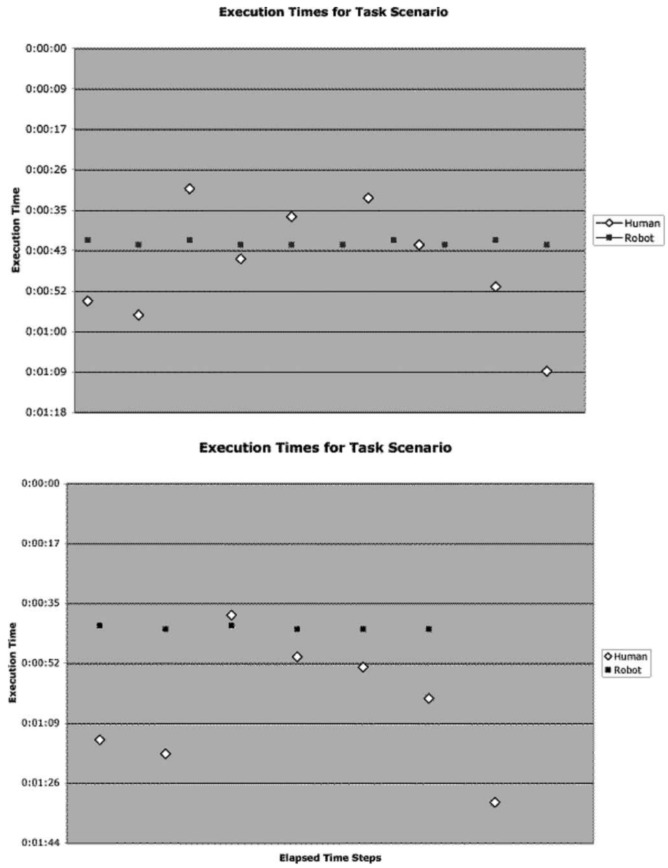


Fig. 10. Sample runs depicting execution times for direct teleoperated control and fully autonomous control.

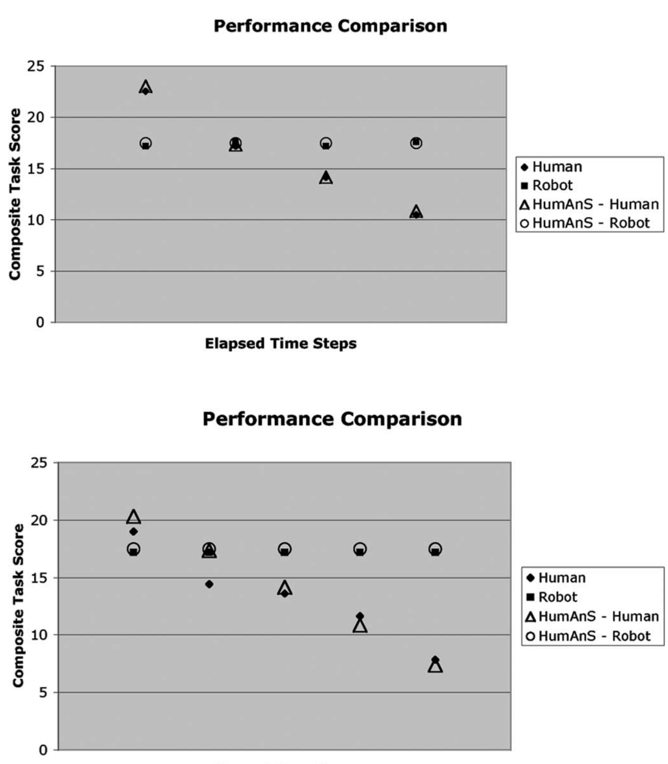


Fig. 11. Comparison of HumAnS prediction system versus real-world implementation data.

As shown in Fig. 11, the relative trend displayed by the HumAnS prediction system compares favorably to the actual performance data collected during real-time implementation. As time elapses, the time for task completion by the human agent increases in the real-world implementation, and the task score decreases in the HumAnS prediction system. In essence, HumAnS is capable of incorporating aspects of both human and robotic system performance and comparing the capability of both agents in a realistic scenario.

V. Conclusion and Future Work

In this paper, a prediction methodology HumAnS that evaluates the various effects of workload on human performance and estimates performance derived from task allocation between human-controlled and autonomous robotic systems is presented. HumAnS utilizes a two-tier process involving performance metrics and performance evaluation that can be applied to a wide range of human–robotic activities performed in complex environments. The HumAnS prediction system has been discussed in detail and its implementation compared on a representative scenario. The experimental setup was designed in order to provide measures for validating the theory for performance prediction. The implementation of the method is shown to provide a correlated comparison that reflects the actual performance of huma–robotic systems operating in the real world.

The ultimate objective of HumAnS is to predict the performance and various effects of workload on human and machine performance. The current version of the performance prediction system uses the pairwise comparison method to rank execution times and workload values. This assumes ideal operating conditions, and limits the ability of the system to handle unplanned discrepancies, such as extreme environmental complexity in the task space or untrained human operators. Future work for the prediction system will thus involve learning from the actual implementation data, and allowing refinement of the execution ranking times and workload ranking values in real time. In addition, performance is calculated based on execution times. To allow full evaluation, other attributes, such as accuracy and repeatability will be incorporated into the composite task score calculation. Lastly, the system assumes that the human agent is an expert in implementation of the task operations, and does not incorporate the learning cycle required for a human operator to first become efficient in a new task. Future work will thus involve incorporating a parameter to acknowledge the learning lag necessary to correlate with real-world performance.

Acknowledgment

The author would like to thank Aryen Moore-Alston for managing the data collection process.

References

[1] M. Sierhuis, J. M. Bradshaw, A. Acquisti, R. V. Hoof, R. Jeffers, and A. Uszok, "Human-Agent Teamwork and Adjustable Autonomy in Practice," presented at the 7th Int. Symp. Artif. Intell., Robot. Autom. Space (i-SAIRAS). Nara, Japan, 2003.

[2] P. Rani, N. Sarkar, and C. A. Smith, "Affect-sensitive human–robot cooperation-theory and experiments," in *Proc. IEEE Int. Conf. Robot. Autom.* Sep. 14–19, 2003, vol. 2, pp. 2382–2387.

[3] P. Scerri, D. Pynadath, and M. Tambe, “Towards adjustable autonomy for the real-world,” *J. AI Res.*, vol. 17, pp. 171–228, 2002.
[4] G. Rodriguez and C. R. Weisbin, “A new method to evaluate human–robot system performance,” *Auton. Robots*, vol. 14, no. 2–3, pp. 165–178, 2003.
[5] H. Yanco and J. Drury, “A taxonomoy for human–robot interaction,” in *Proc. AAAI Fall Symp. Hum.-Robot Interact.*, 2002, pp. 111–119.
[6] T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, and A. Steinfeld, “Common metrics for human–robot interaction,” [Online]. Available: http://vrai-group.epfl.ch/page9329.html, 2004.
[7] A. Howard *et al.*, “A methodology to determine impact of robotic technologies on space exploration missions,” presented at the 10th Int. Symp. Robot. Appl.. Seville, Spain, Jun. 2004.
[8] B. Hahler, S. Dahl, and R. Laughery, “Crewcut — A tool for modeling the effects of high workload on human performance,” presented at the Hum. Factors Soc. 35th Annu. Meeting. Santa Monica, CA, Sep. 1991.
[9] C. D. Wickens, “Resource management and time sharing,” in *Human Performance Models for Computer-Aided Engineering*. NRC Washington, D.C.: National Academy Press, 1989.
[10] J. Keller, “Human Performance Modeling for Discrete-Event Simulation: Workload,” presented at the 2002 Winter Simul. Conf., San Diego, CA, Dec. 2002.
[11] S. Szabo, H. A. Scott, K. N. Murphy, and S. A. Legowik, “Control system architecture for a remotely operated unmanned land vehicle,” presented at the 5th IEEE Int. Symp. Intell. Control. Philadelphia, PA, Sep. 1990.
[12] L. G. Militello and R. J. B. Hutton, “Applied cognitive task analysis (ACTA): A practitioner’s toolkit for understanding cognitive task demands,” *Ergonomics*, vol. 41, no. 11, pp. 1618–1641, 1998.
[13] A. Howard and G. Rodriguez, “Validating mission relevance of autonomy technologies through increased science return,” presented at the 20th Int. Conf. Mach. Learn. Washington, D.C., Aug. 2003.
[14] W. Zachary, J. Ryder, and J. Hicinbothom, “Building cognitive task analyses and models of a decision-making team in a complex real-time environment,” *Cogn. Task Anal.,* New Jersey, 2000.
[15] I. A. Nesnas, *et al.*, “CLARAty: An Architecture for Reusable Robotic Software,” presented at the SPIE Aerosense Conf., Orlando, FL, Apr. 2003.
[16] T. Huntsberger *et al.*, “CAMPOUT: A control architecture for tightly coupled coordination of multi-robot systems for planetary surface exploration,” *IEEE Trans. Syst., Man, Cybern., A: Syst. Hum.*, vol. 33, no. 5, pp. 550–559, Sep. 2003.
[17] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin, “An integrated theory of the mind,” *Psychol. Rev.*, vol. 111, no. 4, pp. 1036–1060, 2004.
[18] M. D. Byrne, “ACT-R/PM and menu selection: Applying a cognitive architecture to HCI,” *Int. J. Hum.-Comput. Stud.*, vol. 55, pp. 41–84, 2001.
[19] N. A. Taatgen, C. Lebiere, and J. R. Anderson, “Modeling paradigms in ACT-R,” *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation*, E. Sun, Ed., 2004, Submitted for publication.
[20] D. Fum and F. Del Missier, “Climbing the Mazes: A cognitive model of spatial planning,” in *Proc. 3rd Int. Conf. Cogn. Model.*, Veenandal, The Netherlands, 2000, pp. 126–133.
[21] G. Gunzelmann and J. R. Anderson, “Spatial orientation using map displays: A model of the influence of target location,” presented at the 24th Ann. Conf. Cogn. Sci. Soc.. Mahwah, NJ, Aug. 2004.
[22] D. E. Kieras, D. E. Meyer, J. A. Ballas, and E. J. Lauber, “Modern computational perspectives on executive mental processes and cognitive control. Where to from here?,” in *Control of Cognitive Processes: Attention and Performance XVIII* Cambridge, MA: MIT Press.
[23] A. K. Chavez and D. D. Salvucci, “An ACT-R Model of the Wickens Tracking Task,” presented at the 25th Annu. Meeting Cogn. Sci. Soc., Boston, MA, Jul. 2003.
[24] J. R. Anderson, M. Matessa, and C. Lebiere, “ACT-R: A theory of higher level cognition and its relation to visual attention,” *Hum. Comput. Interact.*, vol. 12, no. 4, 1997.
[25] M. D. Fleetwood and D. D. Byrne, “Modeling icon search in ACT-R/PM,” *Cogn. Syst. Res.*, vol. 3, pp. 25–33, 2002.
[26] B. P. Gerkey and M. J. Mataric, “A formal analysis and taxonomy of task allocation in multi-robot systems,” *Int. J. Robot. Res.*, vol. 23, no. 9, pp. 939–954, 2004.
[27] R. Becker, S. Zilberstein, V. Lesser, and C. V. Goldman, “Transition-independent decentralized markov decision processes,” in *Proc. 2nd Int. Joint Conf. Auton. Agents Multi Agent Syst.*, Jul. 2003, pp. 41–48.
[28] D. F. Dinges, M. M. Mallis, G. Maislin, and J. W. Powell, “Evaluation of techniques for ocular measurement as an index of fatigue and as the basis for alertness management,” U.S. Dept. Transportation. Rep. DOT HS 808 762.
[29] A. Howard and H. Seraji, “Vision-based terrain characterization and traversability assessment,” *J. Robotic Syst.*, vol. 18, no. 10, pp. 577–587, 2001.

Ayanna M. Howard (S’92–M’99–SM’04) is an Associate Professor at the Georgia Institute of Technology (Georgia Tech), Atlanta. From 1993 to 2005, she was at NASA’s Jet Propulsion Laboratory, where she led research efforts on various robotic projects. Following this, she joined the Systems and Controls Group at Georgia Tech and founded the Human–Automation Systems (HumAnS) Laboratory. Her area of research is centered around the concept of humanized intelligence, the process of embedding human cognitive capability into the control path of autonomous systems. This work, which addresses issues of autonomous control as well as aspects of interaction with humans and the surrounding environment, has been applied to a number of projects—from autonomous rover navigation for planetary surface exploration to intelligent terrain assessment algorithms for landing on Mars. To date, her unique accomplishments have been documented in over 12 featured articles—including being named as one of the world’s top young innovators of 2003 by the prestigious MIT’s journal *Technology Review* and in *TIME* magazine’s “Rise of the Machines” article in 2004.