

# Physically-Proximal Human-Robot Collaboration: Enhancing Safety and Efficiency Through Intent Prediction

Catharine L. R. McGhan\* and Ella M. Atkins†  
University of Michigan, Ann Arbor, Michigan, 48109

Future space missions may be accomplished with greater efficiency through human-robot collaboration, with robots expected to autonomously schedule and execute repetitive or easily-automated tasks. This paper explores issues associated with a human and robotic manipulator operating in a shared physical workspace. A multi-layer autonomous control system is presented that enables a robot to determine near-term intent of a human companion and translate that intent into manipulator motions. These motions must be planned and adapted to avoid collisions with and minimize distraction to the human as both manipulator and human accomplish their mission tasks. This paper presents results from initial experiments designed to assess the impact of a physically-proximal manipulator on human task performance and workload. Human tasks are recognized upon initiation and translated to constraints on manipulator motion to avoid blocking the humans's expected path to a target or visual gaze to a computer screen used to solve a sequence of simple cognitive (math) tasks. Preliminary results indicate the robot is able to accomplish tasks with minimal impact on a human companion's performance or workload. Future work is planned to mature the robotic system's perceptual and prediction capabilities and to complete a more comprehensive set of human-robot collaboration experiments.

## I. Introduction

Humans and robots will ultimately work together as teams in potentially hazardous environments. Efficiency will be maximized through specialization and cooperation with safety as a persistent requirement regardless of environment or mission. Space is one such harsh environment that poses extreme hazards to humans. These hazards may be reduced and overall productivity increased by the inclusion of robotic "partners" that can work alongside human astronauts during extravehicular activities (EVA) in which human perception and mobility are inherently constrained by a spacesuit. Task completion efficiency can be increased by leveraging the human's and robot's differing capabilities: robots are good at repetitive tasks that are well-modeled, while humans are good at devising and executing innovative solutions in a less-predictable, open-ended environment. On Earth, humans typically would also be better at tasks that require complex sensing or a high level of physical dexterity; however, on EVA, the spacesuit "levels the playing field" to an extent. Full human-robot collaboration may require humans and robots to operate in close proximity, requiring coordination of *motions* as well as *tasks*. Most existing vehicle and manipulator platforms, however, have been designed to operate alone or exclusively with other robotic platforms, particularly when those robotic platforms can apply sufficient forces and torques to injure a human companion. This paper explores the problem of close-proximity human-robot collaboration from the viewpoint of the robotic system. Rather than being merely reflexive or reactive to observed positions or motions of its human companion once observed, the robot is goal-seeking: it uses its knowledge of the human's goals (intent) to anticipate human motion and perceptual needs so as to plan its own paths in a manner that it neither collides with nor annoys its human companion. This work presumes the robot and human have simple, decoupled, distinct tasks that can be planned and executed easily to enable research focus on the use of intent information for short-term prediction (single-task recognition in our initial work) of the human's motion to determine constraints on the short-term goals of the robot and its motion. Our case study is a 4-DOF manipulator working in a common environment with a seated human who is executing tasks that require only (a) visual perception of the workspace and (b) motion of the right arm and hand.

---

\* PhD Candidate, Aerospace Engineering Dept., University of Michigan, Ann Arbor, MI 48109, Student Member.

† Associate Professor, Aerospace Engineering Dept., University of Michigan, Ann Arbor, MI 48109, Associate Fellow.

Researchers have previously studied human-robot collaboration primarily at the task level. Automated planning/scheduling tools have been applied long-term to strictly robotic systems. More recently, comparable tools have been developed to assist human space exploration teams in the development of pre-scripted procedures and daily schedules. These tools are generally viewed as valuable for organizing routine tasks and critical activities that can be rehearsed prior to launch,<sup>1</sup> but astronauts prefer the option to exercise ingenuity to manage anomalies and capitalize on scientific discovery opportunities as they arise.<sup>2</sup> Task assignment and division of labor between collaborators are nontrivial problems when activities are not pre-scripted. Safety issues are also nontrivial with highly-capable robots that could injure their companions, mandating hard constraints on applied forces and torques and also their traversal speeds. The inclusion of sense-and-react technologies can help mitigate these safety issues: vehicles could then sense companions to match their motions.<sup>3</sup> Reactive sense and avoid technologies can enable a moving vehicle to remain clear of companions<sup>4</sup>; a reflexive “smart skin” can enable a manipulator to slow down or stop when contacting another object.<sup>5</sup> A robot can further enhance mission safety and efficiency by adapting its own behaviors (motions) based on predictions of its human companion’s intent.

Our research objective is to integrate models of human cognition and motion into a multi-layer control system able to plan the robot’s tasks and motions in a manner that accounts for sensed and anticipated human motion and perceptual focus (e.g., visual gaze). Summarized in Figure 1, this architecture adopts a somewhat traditional three-tier structure. In this figure,  $x_h$  represents physical state of the human, specifically the right arm position and 3-D visual gaze vectors in our experiments,  $x$  represents the physical state of robot, the vector of manipulator joint angles in our experiments, and  $X_G$  represents the matrix of 3D waypoint goal locations, end effector task execution stations in our experiments.  $x_d$  is the 3-D trajectory the manipulator will follow, designated as a sequence of joint-space or end-effectors waypoints,  $e = x_d - x$  is the manipulator feedback control error vector, and  $u$  is the vector of control inputs, joint speed commands sent to the embedded manipulator servos in our system. The top architectural layer is deliberative, using an algorithm to translate expected or known goals to human intent, specified in our current system as the next physical waypoint. We then compare the current sensed human state – physical right arm motion or gaze direction – to the constraints and determine conflicting goal states. From this, the robot’s current goal is decided. This information is sent to the middle layer, which performs *path planning* and any necessary guidance. (Because we do not include reactive behavior to human state data in this early work,  $x_h$  is not yet used in the planner.) Planning algorithms could range from an optimal control-based method for a vehicle platform to a manipulator tooltip motion planner that computes Cartesian space trajectories around obstacle models before conversion to feasible joint-space motions via inverse kinematics. Our experiments employ a series of pre-chosen waypoints and poses to reach them, including stowed and unstowed positions, and a “direct path” joint-space procedure to transition between them, to maximize predictability. A traditional feedback controller is found at the lowest level. As will be described below, our manipulator system relies on built-in servos that accept joint angle (servo) positions and speeds as control inputs.

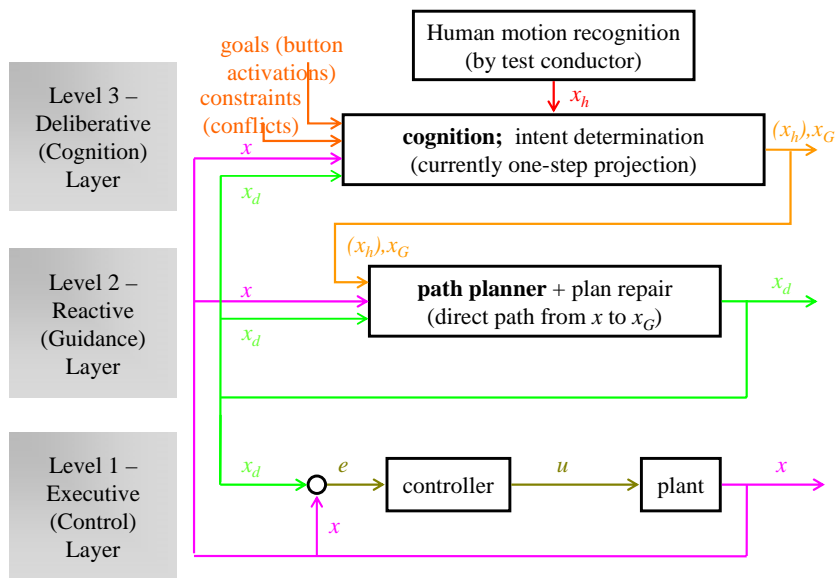


Figure 1: Multi-layer Control Architecture for Proximal Human-Robot Operation

The Figure 1 architecture enables a robot to predict near-term human motion, then build and execute a conflict-free motion sequence without explicit communication. We presume that any necessity for the human to explicitly convey their intent through words, gestures, or keystrokes would be a distraction from nominal task execution. Conversely, while the robot could inform the human of its next goals with minimal overhead (e.g., a GUI indicator of the robot's next waypoint target), such a display would need to be an unobtrusive source of information rather than requiring routine monitoring for safety or goal achievement. For this paper, we assume the human and robot do not communicate except through physical movements executed to achieve task-level goals or meet avoidance constraints. Further, we presume the human and robot can persistently observe the other's motion as needed, and that these observations convey sufficient information for safe interaction. We also assume for simplicity that neither human nor robot exhibits significant learning of the other's behaviors, requiring that our experiments be sufficiently intuitive to be understood and efficiently completed after few trials.

Our hypothesis is that if the robot achieves its goals unobtrusively, the human can go about his/her work without a need or desire to divert attention to the robot. Below we describe a series of experiments to assess the validity of our hypothesis. Furthermore, the experiments were designed to offload tasks also executable by the robot from the human, to enable performance comparison between a human conducting all tasks alone versus in "parallel" with a robotic companion. Our experimental setup, models, and test plan are described below, followed by a presentation of the test results and conclusions. Throughout the presentation, assumptions required for the simplification of models and experimental procedures are stated.

## II. Test Description, Assumptions, and Models

Our experiments involve a single human and a fixed-base manipulator with dimensions comparable to those of a human arm. The human test subject is seated and we assume his/her torso location is static. This allows the manipulator to approximate human motion only through estimation of his/her arm motion and visual gaze from "fixed" coordinates relative to the seat. The seated human subject is tasked with viewing math problems on a computer screen and typing solutions into a keyboard comfortably situated on his/her lap. Simultaneously, he/she is asked to sporadically drink soda from a can adjacent to their right arm or eat chips from a bowl on the opposite side of the workspace. The robot and human share the majority of their workspace volume. The robot has a model of the static objects in its environment, specifically the seated human's torso, desk to which the manipulator is clamped, task targets, food, and drink locations. The robot must observe and anticipate human right arm motion to ensure no collisions occur when the human takes action to drink some soda or eat a chip. Although not sensed directly in this effort, the robot must also anticipate visual gaze to avoid disrupting the human's perceptual focus-of-attention; such disruption would result in reduction of efficiency and/or annoyance for the human. Those predictions of human motion and gaze are then incorporated into the robot's task and motion planning algorithms in the top two layers to devise a plan that achieves its goals while meeting collision avoidance and human disruption constraints. The sole robotic manipulation task-level goal is to reposition and hold its end effector at different target waypoints ("buttons") when they are activated, since with independent tasks only the physical manipulator links in motion can impact the human in a common physical workspace.

Experiments were conducted using a 4-DOF manipulator, as shown in Figures 2 and 4, developed from low-cost components by University of Michigan students. This manipulator has size and speed similar to a human arm and its D-H parameters approximately correspond to both a human arm and NASA's Robonaut arm.<sup>6</sup> The arm is encased in sufficient padding to prevent injury should a human-robot collision actually occur during testing. The human test subject is asked to use only his/her right arm for all tasks, moving between a "rest position" on the keyboard to/from the bowl or soda can and to their mouth. Given no distractions, the human is expected to focus his/her visual attention between the computer screen and the food/drink locations. The robot

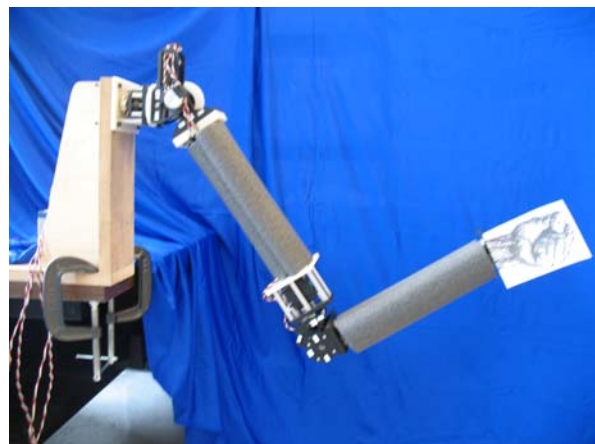
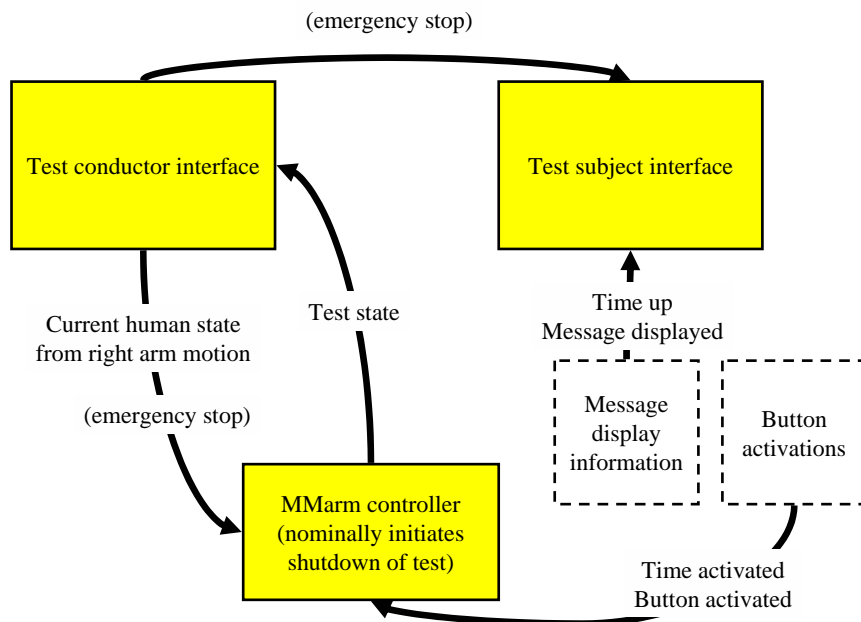


Figure 2: MichiganMan(ipulator) Arm

is constrained to move between its stowed position and three indicated waypoint targets where simulated “button depression” tasks are executed.<sup>‡</sup> Results are derived from a series of quantitative and qualitative datasets. Baseline statistics are collected in terms of task completion efficiency (timeliness) for human and robot conducting their tasks alone. Proximity operations results are then compared to the baseline data to assess efficiency losses or gains throughout the test series. Performance is assessed over a matrix of cases ranging from a baseline where the robot and human have no conflicting tasks, to worst-case scenarios when all tasks the robot is tasked to complete conflict with the human’s designated tasks or when there are simply too many tasks to physically complete in the allotted time. Safety is assessed qualitatively by visual identification of collisions and near-collisions; the robot is not to impact the human in this testing; it is programmed to defer completion of its tasks to completion of the human’s tasks at all times. Efficiency is assessed in terms of task completion times and a human distraction metric based on quantitative data (solution frequency and accuracy of basic math problems, in our case a series of three and two digit integer addition problems). Qualitative data is also collected after each test using the NASA TLX (Task Load Index) questionnaire<sup>7</sup> and via a test subject exit survey completed at the end of each session. These are used to assess the subject’s perceived workload, comfort level, distraction, and annoyance for later system improvements.

Our current software implementation of the Figure 1 architecture uses an “external” human observer rather than automated sensor system to recognize human motion and intent at the cognitive layer. The human observer, not otherwise involved in the testing, conveys changes in the test subject’s intent to the robot, after which the robot’s decisions, primarily motions, are autonomously computed and executed. A C++ interface reads in this information from the observer and communicates the human state data through TCP sockets to an embedded C++ implementation of the path planner, which in turn commands the manipulator navigation and control laws running on a PC/104+ system. The PC/104+ system directly interfaces with all manipulator actuator hardware. Figure 3 shows the three program software infrastructure.



**Figure 3: Software Infrastructure**

**A. Manipulator Description and Kinematics**

The MichiganMan(ipulator) arm (MM-arm) illustrated in Figures 2 and 4 is a low-power, lightweight, low-cost robotic manipulator developed at the University of Michigan. The joints of the arm are driven by high-torque Hitec HS-5955TG and HSR-5995TG digital servos controlled by a PONTECH SV203 servo motor controller board and powered by two 6 V lead-acid batteries in parallel. Command signals are sent to the arm via RS/232 serial port connecting the motor controller board and an Advantec PCM-3370F PC/104+ CPU board running Linux Fedora Core 3 and all embedded manipulator control software.

<sup>‡</sup> The manipulator has no end effector capable of completing dexterous tasks, so for this work we presume the manipulator linkages are the primary source of distraction and conflict.

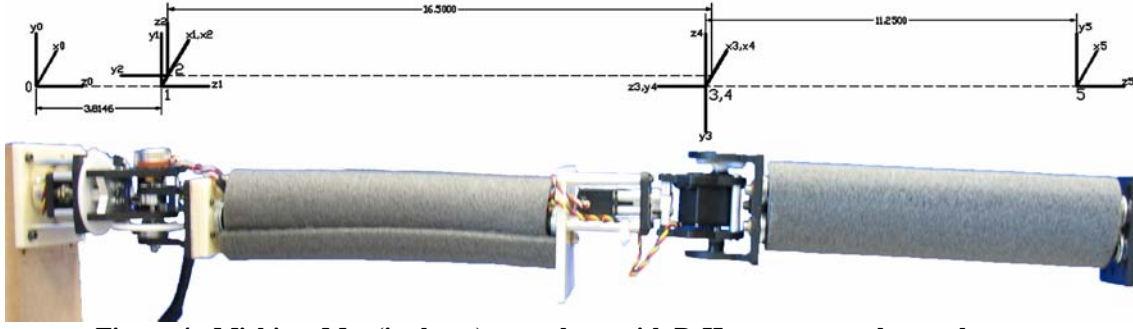


Figure 4: MichiganMan(ipulator) arm photo with D-H parameters shown above arm

The fixed-based MM-arm manipulator kinematics model adopts J. Craig’s Denavit-Hartenberg (D-H) parameter convention<sup>8</sup> shown in Table 1 below. Note that for our implementation, we specify a transformation to a fixed tooltip frame at  $d_5 = -11.25$  inches representing the center of soft padding affixed to the end of the wrist.

Table 1: MM-arm D-H parameters

$i$	$\alpha_{i-1}$ (degrees)	$a_{i-1}$ (inches)	$d_i$ (inches)	$\theta_i$ (degrees)
1	0	0	3.814	$\theta_1$
2	-90	0.345	0	$\theta_2$
3	-90	-0.345	-16.5	$\theta_3$
4	90	0	0	$\theta_4$

The forward kinematics of the manipulator, from base frame 0 to tooltip frame 5 can then be represented by the following transformation matrix:

$${}^0_5T = \begin{bmatrix} (c_1c_2c_3 + s_1s_3)c_4 - c_1s_2s_4 & -(c_1c_2c_3 + s_1s_3)s_4 - c_1s_2c_4 & c_1c_2s_3 - s_1c_3 & 11.25((c_1c_2c_3 + s_1s_3)s_4 + c_1s_2c_4) \\ (s_1c_2c_3 - c_1s_3)c_4 - s_1s_2s_4 & -(s_1c_2c_3 - c_1s_3)s_4 - s_1s_2c_4 & s_1c_2s_3 + c_1c_3 & -0.345(c_1c_2 - c_1) + 16.5c_1s_2 \\ s_2c_3c_4 - c_2s_4 & s_2c_3s_4 - c_2c_4 & -s_2s_3 & 11.25((s_1c_2c_3 - c_1s_3)s_4 + s_1s_2c_4) \\ 0 & 0 & 0 & -0.345(s_1c_2 - s_1) + 16.5s_1s_2 \\ & & & 3.814 - 11.25(s_2c_3s_4 - c_2c_4) \\ & & & + 0.345s_2 + 16.5c_2 \\ & & & 1 \end{bmatrix} \quad (1)$$

Note that trigonometric functions are abbreviated  $\sin(\theta_i) = s_i$ ,  $\cos(\theta_i) = c_i$ .

For our experiments, we imposed software constraints on the MM-arm joint angles to emulate “human arm” joint ranges. For example, the elbow joint, link 4, is not allowed to bend back past zero degrees. Arm linkages were covered in an extra layer of dark soft fur over the nominal foam covering to provide more cushioning and help the manipulator appear more arm-like.<sup>8</sup> The manipulator wrist was extended with a ball of white cotton batting material to soften the end as well as make it the visual focus of attention for the human. The arm control software was configured to move the manipulator joints at speeds well below their maximum rates, to ensure safety and minimize distraction for the human test subject. The arm was mounted to the corner of a table, about a foot higher than the human test subject’s shoulder when he/she is seated in the chair. The workspace of the arm overlaps the human’s workspace over a significant volume in front of the human, but the MM-arm is positioned so that the seated test subject’s head, torso, and legs are outside its reachable workspace.

<sup>8</sup> The fur was originally part of a gorilla costume sold at Halloween. We presume an arm that looks animal-like will be perceived as safer and “friendlier,” and more compatible with operation in a common workspace.

## B. Workspace Setup

The workspace area is shown in Figures 5 and 6. “Buttons” b1, b2, and b3 are task target locations marked for human visual identification by blue reflectors. The work area was configured to favor right-handed test subjects, since for modeling simplicity the test subject was expected to type answers and perform manipulation tasks only with their right hand. This is acceptable, as all our test subjects were right-handed. Math problems requiring the test subject’s cognitive and visual attention were displayed in a large graphics font on the monitor and reset to a new problem each time a test subject entered an answer. A wireless keyboard held on the test subject’s lap enabled the user to enter solutions to math problems and also when they began and completed a manipulation task. Logging software captured all keystrokes and timings, as well as correct and incorrect math answers, for later results evaluation. This information was not made available to the MMarm planner during testing.

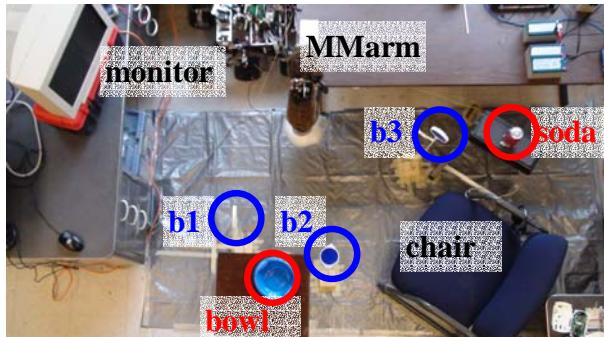


Figure 5: Layout of Experimental Test Area, Overhead View

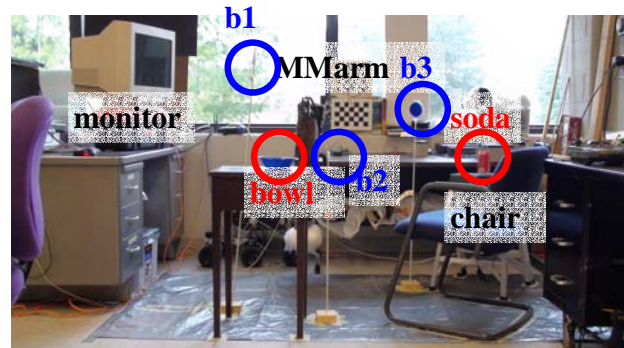


Figure 6: Layout of Experimental Test Area, Side View

Since we wanted tests to remain as simple as possible to facilitate dataset comparison, task sets were decoupled: the cognitive process involved with the math was not related to the manipulation tasks. To sidestep known causes of cognitive delay, our cognitive task was straightforward: addition of a two-digit number (10-99) to a three-digit number (100-999). With this task the human need only interpret numbers, not switch between mathematical operations (e.g., addition vs. subtraction). The interruptive tasks for the human were pressing buttons (1, 2, or 3), eating chips, or drinking soda. Test subjects were told that solving the addition problems was their main task and that they were to solve them as quickly as possible, but not at the cost of giving incorrect answers, which would not raise their score. (Correctness was valued most, with speed as a modifier weight on efficiency in the scoring algorithm.) Test subjects were instructed to use the number line at the top of the keyboard rather than the numberpad to input their math solutions. We hypothesized that this would better balance test subjects since there is large variation in numberpad typing speeds based on previous experience. Users entered solutions right to left (ones, tens, hundreds, thousands place) to facilitate step-by-step entry of addition solutions given frequent problems requiring carry operations.

In this testing, the human was constrained to solve math problems by default or else complete the higher-priority eat/drink tasks before returning to the math problems. Messages that requested the user to complete a non-cognitive task (e.g., eat-chip) were shown on the monitor below the addition problem. Once a request was issued the number keys were locked, and the program would not allow a user to resume work on the math problem until the interruptive task was completed. Task initiation and completion was logged by the user pressing the spacebar twice – once when the task was first noticed, and once at task completion. We used “eat-chip” and “drink-soda” requests in this test series to equalize activities across test subjects in order to control task switch frequency and the potential for conflict between human test subject and manipulator actions. Test subjects were also instructed to complete the physical movements of interruptive tasks at their preferred pace, not as fast as possible. This allowed us to capture nominal human movement timings and creates a more realistic and relaxed environment.

Test subjects were told how the arm sensed them – through input specified by test conductor observing the experiment – but were instructed to focus on their own tasks rather than the manipulator’s motions or task progress. Subjects were given information on which of their targets conflicted with the robot’s movements, and told that the robot would defer to their movements during testing. They were then told that they could pause or take longer to complete a nonconflicting task to allow the arm more time to complete its own tasks, although were also told not to worry whether the arm could complete its tasks or not. They were also told not to complete a robot’s tasks for it. Although clear, these instructions had an interesting effect to be discussed in our results section.

The test laboratory had background noise, and the manipulator servos emitted a distinct acoustic signature that appeared to be a function of both applied servo torques and manipulator pose in the gravity field. To remove sound

as a distraction, or possible source of manipulator motion information not available to a suited astronaut, test subjects were instructed to wear noise-cancelling headphones during all testing. Instrumental music with low dynamic range was played through the headphones, since it has been noted that such music can block external noise without introducing new distraction.

### C. Experimental Plan

To best compare the impact of human and robot on the other’s performance, we conducted five sets of experiments with each test subject, allowing the test subject to observe the robot operating alone for orientation purposes:

1. Nominal robot action without consideration for human test subject actions
2. Nominal human math problem-solving with the manipulator at a static workspace position
3. Human math problem-solving with interruption to eat chips, drink soda, or press buttons
4. Collaborative operation baseline: Test 3 and 1 with manipulator sporadically avoiding the human
5. Collaborative operation “stress tests”: Test 4 configured to induce frequent task conflicts

Test set 1 is the best-case scenario for the robot, one in which it need not take human motion into account. Test set 2 is necessary to establish a best-case cognitive task performance index for each test subject. Test set 3 is the worst-case scenario for the human test subject acting alone, representing the case in which the human is frequently interrupted from their main (cognitive) task(s) to conduct a physical task, in our experiments eating a chip or drinking soda or pushing buttons, in some cases all three over the duration of an experiment. Test set 4 is the best-case scenario for human-robot collaboration, collaboration without conflicting tasks. Test set 5 represents a worst-case scenario in our setup for human-robot collaboration. In this scenario, conflicting tasks are frequent and efficiency for both human and robot were hypothesized to be at minimum values.

To regulate the degree to which conflicts arise, the manipulator pose necessary to press button b1 was configured to directly intersect the gaze of the human to the monitor. The manipulator trajectory to and configuration at button b2 intersected with the human test subject’s path to the bowl containing chips. The motion of the manipulator to press button b3 was free from conflict, and the test subject could always reach their keyboard and soda without conflict. Figures 7, 8, and 9 illustrate the poses associated with the various conflict cases.

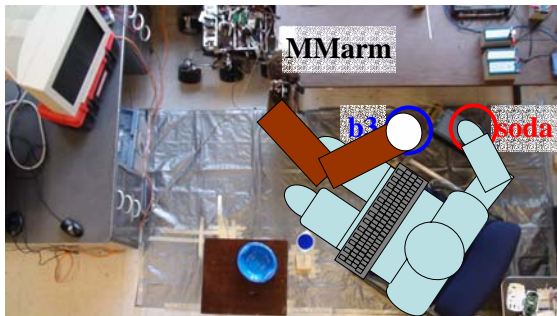


Figure 7: No-conflict, Robot-b3 and Human-soda Poses

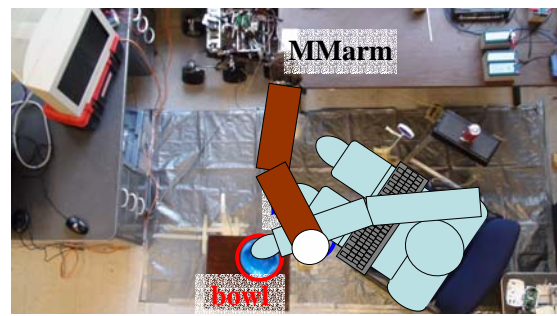


Figure 8: Conflict, Robot-b2 and Human-chip Poses

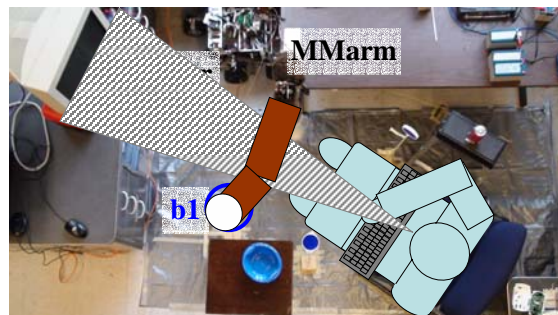


Figure 9: Conflict, Robot-b1 and Human-gaze Poses

During a typical test, the test conductor sits to the left of the test subject with manipulation workspace and test subject both in view. The test conductor currently shortcuts the process required to sense and determine the next action initiated by the human, a process to be eliminated as our automated cognitive intent determination algorithms

are better developed. The path planner then uses this information to decide whether a conflict could occur between the manipulator and human, currently making the binary decision to continue the ongoing task or else discontinue its current conflicting task and go to the next task on its queue or else a nominal “stow” task if no conflict-free task is queued. The logic for the robot’s task scheduler is a simple FIFO (first-in first-out) queue built from button press activities stored in a file for each test. Button press actions are specified by a button designation and activation time. If a new button activation occurs while an earlier activation for that button is still on the queue, the new button activation is listed as failed and immediately discarded. Although this is a crude planner it responds to avoid conflict with human actions as it attempts to complete its own tasks, our baseline requirement. Nominally, once a goal is selected by the planner, a trajectory arc in joint space is identified from our pre-compiled path database, and the controller begins to send new commands to the manipulator. The planner does not complete the entire motion before selecting a new goal. Instead, it frequently updates human arm and visual attention state to verify that the current trajectory is conflict-free. If the current goal/trajectory can induce conflict, the FIFO planner abandons the current activity in favor of either the next conflict-free task on the queue or a default conflict-free pose executed if either no activities are queued or all queued activities may cause a conflict.

Meanwhile, the seated test subject attempts to solve as many addition problems correctly as possible. Messages (button or eat/drink tasks and time of activation) are read from a script file, with single messages displayed to the monitor at the designated activation time. When a message is displayed, the user reads the on-screen task request and completes this task, with their motions noted by the test conductor. If a new message becomes valid while an earlier message remains on the queue, the earlier message is discarded (task is listed as incomplete) and the new message displayed instead. Given our desire to allow the operator to accomplish tasks at a nominal rather than rushed pace, the timings for our button/eat/drink tasks were spaced sufficiently that tasks were always completed before again activated.

Each of four test subjects completed the ten tests shown below, with the ten tests repeated in two sessions separated by at least several hours. Tasks and timings were the same between sessions, for all subjects. However, the addition problems were always randomized.

1. Nominal robot action showing all arm movements (b1→b2, b2→b3, b3→b1, b1→b3, b3→b2, b2→b1)
2. Nominal human cognition only (to familiarize the test subject with the interface)
3. Human problem-solving with interruption (no manipulator motion)
  - a. human cognition + food/drink (timing eating/drinking tasks only)
  - b. human cognition + buttons (timing button pushing tasks only)
  - c. human cognition + food + buttons (all tasks occurring frequently)
4. Collaborative operation baseline: human + robot performing all tasks with no conflicts
5. Collaborative operation “stress tests”:
  - a. human + robot with some conflicts (button b2 vs. chip eating)
  - b. human + robot with some conflicts (button b1 vs. gaze, allowed only during eat/drink activities)
  - c. human + robot with all conflicts (robot always given button conflicting with current human task)
  - d. human + overtasked robot (new button tasks appear on queue every 5 seconds)

Test 1 was shown to the test subjects to establish familiarity with the manipulator movements, to verify the button 1 pose would occlude the subject’s view of the monitor, and to ensure the subject was comfortable with the [adjustable] speed of the arm. Note that no test subjects asked for the arm to be slowed further, although as discussed below there was a recommendation for the manipulator to move away from a conflict region more quickly. Test 2 gave the subject familiarity with the cognitive task, and provided baseline data for the nominal, no-interruption scenario. Tests 3a, 3b, and 3c were progressively complex scenarios involving only human activity. Test 4 was the nominal collaborative case with task sets engineering to avoid conflicts. Tests 5a, 5b, 5c, and 5d were set up as progressively more challenging collaboration scenarios, the last of which was so challenging the manipulator could not even accomplish all tasks if working alone. It should be noted that if a human tried to complete all Test 5 tasks alone, in most cases they would also not have been able to do so.

Each test subject was tested in two approximately two-hour sessions, each comprised of the above ten tests. Each test was three to four minutes long. At the start of a session, a set of directions was given to the test subject regarding the test setup and what tasks they would be completing. At the end of every session, the test subject was given a NASA TLX survey to complete as well as a short exit survey.

### III. Results

Task completion times for the human subject represent the primary quantitative measure by which we can assess the impact of the manipulator on human performance. Table 2 summarizes the task completion time statistics



averaged over all tests, while Table 3 summarizes cognitive (math) task completion times distinguished by whether or not the robot was also in motion in the workspace. Our hypothesis, that if the robot achieves its goals unobtrusively the human can go about their work without the need or desire to divert their attention to the robot, is supported by the Table 3 data. When menial ‘interruptive’ tasks were reassigned to the robot, the human task completion times increased for three of the four test subjects, but in all cases the change was within the measurement standard deviation.

**Table 2: Averages and Standard Deviations of Task Completion Across All Tests (in seconds)**

	Subject 1		Subject 2		Subject 3		Subject 4	
	average	stdev	average	stdev	average	stdev	average	stdev
Math	3.71	3.39	5.61	4.67	4.32	2.96	6.83	4.34
Eating/drinking	9.08	2.31	7.79	1.14	6.57	1.31	9.18	1.29
Button press	3.34	0.27	3.79	0.57	3.63	0.58	5.44	1.01

**Table 3: Averages and Standard Deviations of Cognitive Task Completion Across Similar Tests (in seconds)**

	Subject 1		Subject 2		Subject 3		Subject 4	
	average	stdev	average	stdev	average	stdev	average	stdev
Math, solo	3.09	2.25	5.13	3.62	4.24	2.73	7.54	4.47
Math, robot	4.09	3.89	6.49	6.03	4.46	3.30	5.71	3.91

Results for the four test subjects are listed separately to facilitate comparison. Standard deviations were uniformly larger for the math task than for the other tasks. We believe this was a product of the variance in math problem complexity, with problems ranging from simple solutions (e.g.,  $121 + 22$ ) to more complex problems with multiple carries and potentially four-digit solutions (e.g.,  $987 + 75$ ). Math problem completion time variation also is a function of interruption for tests in which the subject was instructed to eat a chip, for example, in the midst of a math calculation. The interruptive task completion times generally did not vary substantially for each test subject, but do vary appreciably between test subjects. As shown in Table 3, averages and standard deviations of math problem completion distinguished by the presence of a moving manipulator (or not) indicated no statistically-significant distinction.

As an alternative measure of test subject performance in terms of cognitive task completion, Table 4 shows cognitive task completion rate in math problems solved per second, with Subject 2 and Subject 3 results separated into two test sessions. Higher rates mean better performance by the test subject. Table 4 shows that there are not strong learning effects, and that in most cases completion times did not vary by test type.

**Table 4: Rates of Cognitive Task Completion (in problems solved / second)**

	Subject 1	Subject 2		Subject 3		Subject 4
Test 2	--	--	0.19	0.15	0.16	0.09
Test 3a	--	0.12	0.15	0.14	0.16	0.09
Test 3b	--	0.14	0.16	0.17	0.16	0.09
Test 3c	0.21	0.15	0.17	0.17	0.16	0.11
Test 4	0.21	0.15	0.18	0.19	0.17	0.14
Test 5a**	0.18	0.13	0.14	0.20	0.17	0.11
Test 5b**	0.20	0.14	0.13	0.16	0.17	0.12
Test 5c**	0.18	0.15	0.15	0.17	0.17	0.12
Test 5d**	0.35	0.12	0.15	0.16	0.17	0.11

\*\* There were conflicts with the robotic manipulator during these tests.

It should be noted that the Table 4 rates are of math problem completion, not correctness. More testing and more in-depth data parsing may uncover significant differences in the rate of correct solution response, which should help indicate during which cases the subject was more or less distracted.

Some interesting qualitative results also were compiled during testing. First, all test subjects found it necessary to lean in towards the chip bowl when eating a chip. This suggests that in future testing, we may need to take torso location into account rather than presuming the human’s right arm is attached to a fixed base. In the exit surveys, three of the four responders indicated they wanted the robot to move faster, or that it was too slow. The current speed of the manipulator is far slower than the robot’s capabilities for safety reasons and to intentionally create

limits on the manipulator's ability to move away from the human test subject once a conflict was identified. Speeding up the arm would reduce the time required for the arm to back up from a conflict, but we hypothesize such a speed-up might increase the level of distraction or discomfort associated with the human test subject operating in the shared workspace. Future tests are planned over a variety of manipulator speeds, or speeds that vary within a test as a function of action type (e.g., task completion vs. conflict avoidance).

There were also interesting patterns observed as test subjects attempted their tasks. When the robot and human had conflicting physical tasks, such as the manipulator pressing button b2 as the human is asked to eat a chip, despite the fact that this was cited as a case where the arm was moving too slowly, the subjects did not appreciably increase task completion time to account for this conflict. However, when the robot's physical motions conflicted with the human's cognitive tasks – intersecting the gaze of the human – the human was forced to wait until the robot moved fully away from the monitor. Since the switching time from physical interruptive task to cognitive task was near-instantaneous once the physical task was complete, the robot had far less time to respond to the change in state than when the reverse was true (moving out of the way when the human transitioned from cognitive task to the physical task requiring the arm to move from the keyboard to the chip bowl). This observation motivates the utility of an intent-discerning mechanism with predictive capabilities. The current FIFO planner the robot uses has no lookahead – it cannot predict what the human will do beyond the currently-initiated action. In the cognitive-to-physical task transition, the robot will be able to react with appropriate action and execution speed. However, in the physical-to-cognitive task transition, the robot must anticipate the “gaze vector” in advance of the transition since gaze can be redirected more quickly than the manipulator can move. With intent prediction, the robot could predict human task selection and timings to the extent necessary for the robot to remain clear of expected areas of conflict when the human switches to their next or next series of tasks.

Given the manipulation delay following a physical-to-cognitive task transition, a behavior dichotomy was observed. During some tests, two of the subjects sat higher in their chair to see over the manipulator in an attempt to continue their work. The other two test subjects would either wait patiently for the arm to move out of their gaze, or else attempt to slow down their movements during other non-conflicting tasks perhaps expecting the manipulator would finish its button b1 task before they switched their attention back to the monitor. The delay strategy was effective during the tests in which the manipulator was lightly-tasked because the manipulator would in fact complete its button b1 task and move on. With high utilization levels, the manipulator more frequently occluded the monitor, in which case patience was of minimal use. This dichotomy suggests that in future tests the subjects either be instructed in how to trade off acceptance of delays with identification of ways to work around whatever causes the delay-inducing conflict. In this case, the tradeoff was between the relative importances of completing the designated cognitive tasks in the timeliest manner versus conserving energy by patiently waiting.

One final case is worthy to note, in which a substantial delay was incurred in communicating human state information to the robot. This delay occurred while the robot was reaching for button b2 and the subject had just received an “eat chip” message. In this case, the manipulator continued to block b2, introducing a conflict with the eat chip action. The test subject quickly noticed the robot remained at button b2, paused to wait for the robot, then completed their own chip-eating motion and verbally commented upon the lack of response. This situation suggests that the human is easily able to recognize differences in the robot's behavior between cases in which it reacts to the human versus cases when it does not. This provides anecdotal evidence that the robot's intent can be communicated to the human solely through its motions. Our next step is to more fully translate this recognition process to the complementary process in which the robot characterizes intent of its human companion.

#### **IV. Conclusions and Future Work**

In this paper we presented a multi-layer architecture to enable a robot to effectively collaborate with a human companion through intent characterization, translation to robot motion constraints, and subsequent task selection and control. We devised a series of experiments in which a human test subject and fixed-base manipulator operate in a shared workspace to characterize human performance given a series of cognitive and physical tasks conducted with and without the manipulator executing its own tasks. The test set progressed from basic experiments with only robot or human acting in the workspace, to more complex experiments in which both robot and human are over-tasked with a variety of conflicting actions. The goal of these experiments was to determine the extent to which the robot impacts human task performance, as a beginning to a continuing test sequence aimed at both characterizing this impact and understanding how to mitigate or minimize this impact through improved algorithms for the robot to recognize and act appropriately based on its companion's intent. Based on our tests, the robot's presence and motions do not have a significant effect on human performance. However, as the standard deviations are rather high, more testing and data processing are needed to establish statistically-significant conclusions. Qualitative

results suggest the human test subjects felt safe with their robotic companion, with subjects in fact recommending higher manipulator speeds to minimize conflicts/delays in future tests. Human test subjects also demonstrated varied responses to conflicts when they occurred and provided evidence that they could easily interpret whether the robot was continuing to pursue its goals versus executing a conflict avoidance action.

Future work will involve more extensive testing to provide statistically-significant results, including more subjects, more varied tests, and additional data collection/processing. We also intend to comparatively characterize human performance, expected to exhibit clear degradation, when the robot strictly executing its task queue without regard to conflict with the human. Development of algorithms to predict future intent of a human companion presents a major challenge, but will be necessary particularly in cases in which the human test subject focuses his/her attention back to a cognitive task more quickly than the manipulator can move to un-occlude the test subject's view of the computer monitor. Ultimately, the test conductor, who currently acts in place of a human pose and motion sensor system, must be replaced with an autonomous system to complete the same function, most likely employing multiple sensors that together characterize human arm pose and human gaze direction.

### Acknowledgments

The authors thank several graduate students from the Department of Aerospace Engineering Dynamics and Control group for volunteering to be test subjects for the MMarm experiments.

### References

- <sup>1</sup>Kortenkamp, D., "A day in an astronaut's life: reflections on advanced planning and scheduling technology," *IEEE Intelligent Systems*, Volume 18, Issue 2, March-April 2003, pp. 8-11.
- <sup>2</sup>Jaap, J., and Phillips, S., "On using an incremental scheduler for human exploration task scheduling," *2005 IEEE Aerospace Conference*, IEEE, Big Sky, MT, 5-12 March 2005, pp. 4062-4069.
- <sup>3</sup>Gavrila, D. M., "The Visual Analysis of Human Movement: A Survey," *Computer Vision and Image Understanding*, Volume 73, No. 1, January 1999, pp. 82-98.
- <sup>4</sup>Thrun, S., Bennewitz, M., Burgard, W., Cremers, A. B., Dellaert, F., Fox, D., Hahnel, D., Rosenberg, C., Roy, N., Schulte, J., and Schulz, D., "MINERVA: a second-generation museum tour-guide robot," *Proceedings of the 1999 IEEE International Conference on Robotics and Automation*, Vol.3, IEEE, Detroit, MI, 10-15 May 1999, pp. 1999-2005.
- <sup>5</sup>Rekimoto, J., "SmartSkin: an infrastructure for freehand manipulation on interactive surfaces," *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, Minneapolis, MN, 20-25 April 2002, pp. 113-120.
- <sup>6</sup>Nickels, K., "Hand-Eye Calibration of Robonaut," *San Antonio Chapter IEEE Computer Society Past Meeting Presentations September 16, 2004* [online archive], URL: [http://www.ieee-cs-cts.org/past\\_meetings.htm](http://www.ieee-cs-cts.org/past_meetings.htm) [cited 10 September 2008].
- <sup>7</sup>"NASA TLX Paper and Pencil Version Instruction Manual," *NASA TLX Homepage* [online archive], URL: <http://humansystems.arc.nasa.gov/groups/TLX/> [cited 10 March 2009].
- <sup>8</sup>Craig, John J., *Introduction to Robotics: Mechanics and Control*, 3<sup>rd</sup> ed., Pearson Education, Inc., Upper Saddle River, NJ, 2005, pp. 28, 35-36, 68-69.