Cost-Based Anticipatory Action Selection for Human–Robot Fluency

Guy Hoffman and Cynthia Breazeal

***Abstract*—A crucial skill for fluent action meshing in human team activity is a learned and calculated selection of anticipatory actions. We believe that the same holds for robotic teammates, if they are to perform in a similarly fluent manner with their human counterparts. In this work, we describe a model for human–robot joint action, and propose an adaptive action selection mechanism for a robotic teammate, which makes anticipatory decisions based on the confidence of their validity and their relative risk. We conduct an analysis of our method, predicting an improvement in task efficiency compared to a purely reactive process. We then present results from a study involving untrained human subjects working with a simulated version of a robot using our system. We show a significant improvement in best-case task efficiency when compared to a group of users working with a reactive agent, as well as a significant difference in the perceived commitment of the robot to the team and its contribution to the team's fluency and success. By way of explanation, we raise a number of fluency metric hypotheses, and evaluate their significance between the two study conditions.**

***Index Terms*—Algorithms, anticipatory action selection, fluency, human factors, human–robot interaction, teamwork.**

I. INTRODUCTION

TWO PEOPLE repeatedly performing an activity together naturally reach a high level of coordination, resulting in a fluent meshing of their actions. In contrast, human–robot interaction is often structured in a stop-and-go fashion, inducing delays and following a rigid turn-taking pattern. Aiming to design robots that are capable peers in human environments, we try to attain a more fluent meshing of human and machine activity.

In recent years, the cognitive mechanisms of joint action have received increasing attention [1]. Among other factors, successful coordinated action has been linked to the formation of expectations of each partner's actions by the other and the subsequent acting on these expectations [2], [3]. We argue that the same holds for collaborative robots: if they are to go beyond stop-and-go interaction, agents must take into account not only past events and current perceived state but also expectations of their human collaborators.

In this paper, we present an adaptive anticipatory action selection mechanism for a robotic teammate. We analyze our model of anticipatory action in a cost-based framework of coordinated shared-location action, and compare it to a purely reactive agent acting within a traditional perception–action loop, demonstrating a theoretical improvement in joint task efficiency.

We then present results from a study involving untrained human subjects working with a simulated version of a robot using our anticipatory system. We show a significant improvement in best-case task efficiency when compared to users working with a purely reactive agent. However, we were not able to show this difference being significant when measuring the mean score over repetitions. We attribute this, in part, to the small number of repetitions used in our study.

That said, we are not interested solely in efficiency, but also in the qualitative notion of fluency in coordinated action meshing, ultimately leading to more appropriate collaborative behavior. In a poststudy survey we found a significant difference in the perceived contribution of the robot to the team's fluency and success, as well as its commitment to the team. Given that there are no generally accepted measures of teamwork fluency, we raise three fluency metric hypotheses, and evaluate these between the two conditions. We find the groups to differ significantly in two (time between human and robot action, and time spent in concurrent motion), but not in a third (human idle time).

The remainder of the paper is structured as follows: In Section II, we briefly describe the cost-based Markov process in which our agent is set, and, in Section III, we outline a reactive action-selection mechanism for an agent in this world. In Section IV, we introduce our adaptive cost-optimizing anticipatory agent and analyze its behavior vis-a-vis a simulated human teammate. Section V presents and discusses results from the human subject study; Section VI discusses related work, and we conclude in Section VII with future research directions.

II. WORLD DESCRIPTION

We model the team fluency problem as a discrete time-based deterministic decision process including two agents, a *robot* and a *human*, working together on a shared task.

Both robot and human share a common workspace, which at any time point is in one of a finite number of states $\Sigma_W = \{s^w_0, \ldots, s^w_n\}$, and is initially in state s^w_0. The agents also have a number of states $\Sigma_H = \{s^h_0, \ldots, s^h_n\}$ (the human's states) and $\Sigma_R = \{s^r_0, \ldots, s^r_n\}$ (the robot's states). In our model, the robot can only perceive the state of the workspace if it is in state s^r_0 (the *perceptive* state). We denote a full state of the system $s_n \equiv < s^w_i, s^h_j, s^r_k >$, and similarly, $\Sigma = \Sigma_W \times \Sigma_H \times \Sigma_R$.

Human and robot have distinct abilities, described as two sets of actions, $A_H = \{a^h_1, \ldots, a^h_k\}$ for the human, and $A_R = \{a^r_1, \ldots, a^r_l\}$ for the robot.

$T : ((A_H \cup A_R) \times \Sigma) \mapsto \Sigma$ is a transition function that maps certain state-action pairs to new states. We denote a particular

Manuscript received October 16, 2006; revised May 23, 2007. This paper was recommended for publication by Associate Editor Y. Nakauchi and Editor H. Arai upon evaluation of the reviewers' comments. This work was supported in part by the Office of Naval Research under Grant YIP N000140510623.

The authors are with the MIT Media Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: guy@media.mit.edu; cynthiab@media.mit.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TRO.2007.907483

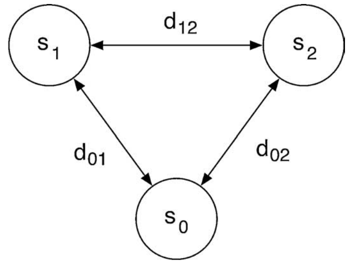


Fig. 1. Transition costs between two states as a directed graph.

state-action pair transition in T

$$\tau_i^x(s_k) = s_l \equiv < a_i^x, s_k > \mapsto s_l$$

meaning that if agent $x \in \{h, r\}$ (human or robot) performed action i while the world is in state s_k, the world would transition into state s_l after applying the action.

A central motivation of our model is to investigate aspects of *time* associated with actions of two collaborating agents. Therefore, state transitions are not atomic, and the decision to take a particular action does not result in an immediate state transition. Instead, moving between states takes time, and is associated with a known discrete cost, which is a function of the states before and after the action. This cost can be thought of as the "distance" between states, or, more generally—the duration it takes to transition between states. We denote the cost of transitioning between states s_k and s_l with $d(s_k, s_l)$.

$$D : \left(\bigcup_{i \in W,H,R} < \Sigma_i \times \Sigma_i > \right) \mapsto \mathbb{N}$$

Thus when at time t, agent $x \in \{h, r\}$ decides to take action a_i^x on state s_k and $\tau_i^x(s_k) = s_l$, the world will be in state s_l only at time $t + d(s_k, s_l)$. While the other agent may take more actions during this time, the next time-step agent x will be able to take another action is $t + d(s_k, s_l)$. It can be useful to depict the state transitions as a directed graph, with the nodes representing the states and the edges the transitions between the states, weighted by the duration/cost function D (see Fig. 1).

For the sake of simplicity, we will, sometimes, denote $d(s_k, s_l)$ as d_{kl}, as indicated in the figure.

Agents cannot change the other agent's states with their actions, but they operate on a common workspace. Therefore, our model is clearly ill-defined with regards to race conditions on the Σ_W state space. There are several possible solutions (such as the use of semaphores and other synchronization mechanisms). In this work, for the sake of simplicity, we will assume that actions that change the workspace are locking with regard to actions that operate on the common workspace *for both agents*. In the implementation of our model described earlier, we solve this race condition by making all state transitions affecting the workspace atomic.

A. The Factory World

In our experiments, we use a simulated factory setting (Fig. 2). The goal of the team is to assemble a cart made of a *Body*, a *Floor*, two *Axle*s, and four *Wheel*s. The various parts have particular ways to be attached to each other—the *Body* is welded to the *Floor*, *Axle*s are riveted to the *Floor*, and *Wheel*s are attached to *Axle*s using a wrench of matching color. A *component* is a partially assembled cart segment that includes one or more individual parts attached to each other, for example, $Axle1 + Body + Floor$.

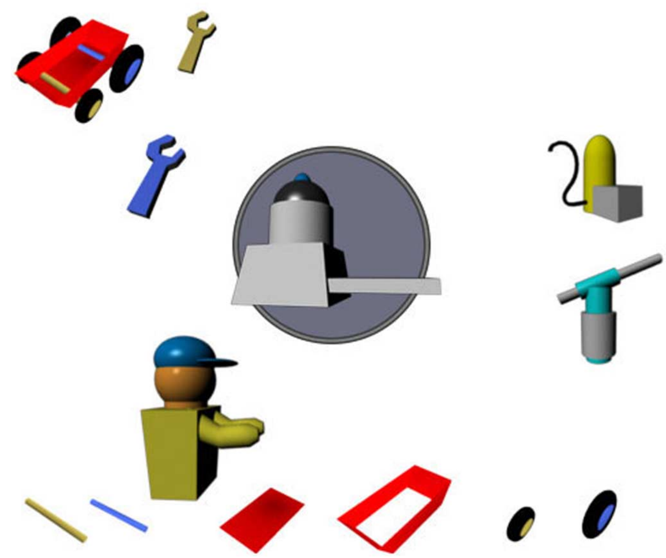


Fig. 2. Simulated factory setting with a human and a robot building carts, while sharing a workbench (gray circle), but dividing their tasks. The robot has access to the tools (right and top-left of workbench), whereas the human is responsible to bring the parts (below the workbench). Top left shows a completed cart.

The labor is divided between the human and the robot: the human has access to the individual parts, and is capable of carrying them and positioning them on the workbench. The robot is responsible for fetching the correct tool and applying it to the currently pertinent component configuration in the workbench. Each part has a stock location (with an infinite supply of parts), and each tool has a storage location, to which it has to be returned for the robot to be able to find it again. The workbench can, at any one time, contain at most two components.

In the earlier-described framework, the workbench state space $\Sigma_W = Comps^2$, where $Comps$ is the space of all possible components.[1] In our case, $|Comps| = 42$, and thus $|\Sigma_W| = 42 \times 42 = 1764$. The robot's state space includes its position at one of the four tools' storage areas or at the workbench, and whether the robot is holding one of the tools or not. Therefore, $|\Sigma_R| = 25$. Similarly, with six kinds of parts, $|\Sigma_H| = 49$. Thus, the size of the state-space in this simulation is 2,160,900.

The action-space of the robot is

$$A_R = \{Workbench, Welder, Rivet, Wrench1, Wrench2, PickUp, PutDown, Use\}.$$

The first five actions are mobility actions, moving to one of the five locations in the factory. $PickUp$ and $PutDown$ are

[1] Note that this is not $2^{|\text{Parts}|}$, since not all parts can be attached to each other, and some parts can appear multiple times in a component.

operational only in the tool locations, with the latter only available at the correct storage location of the currently held tool. A_H is a similar space with two more navigation actions, and no Use action. To illustrate the state transitions, here are two examples of transitions brought about by actions in A_R (Here, W is $Wrench1$, R is $RivetGun$)

$$\tau^r_{Pickup}(\langle s^w_i, s^h_j, \langle W, \emptyset\rangle\rangle)$$
$$= \langle s^w_i, s^h_j, \langle W, W\rangle\rangle$$
$$\tau^r_{Use}(\langle\langle Floor, Body\rangle, s^h_j, \langle Wrkspc, Welder\rangle\rangle)$$
$$= \langle\langle Floor + Body, \emptyset\rangle, s^h_j, \langle Wrkspc, Welder\rangle\rangle$$

and

$$\tau^r_{RivetGun}(\langle s^w_i, s^h_j, \langle W, \emptyset\rangle\rangle) = \langle s^w_i, s^h_j, \langle R, \emptyset\rangle\rangle.$$

The duration cost of a state transition that involves navigation is the distance between the previous and the new location. The duration cost for state transitions involving the inventory of an agent, or changes to the workbench, is 1 in this implementation, but could, theoretically, be different for each tool.[2] The robot can perceive the state of the workbench only when it is located in it. Workbench state changes that happen while the robot is in any other state are not applied to its internal representation.

Moreover, we assume that the robot has a function Φ that maps the workbench state to the appropriate tool required to bond the two components on the workbench. For example: $\phi(\langle Floor + Axle1, Wheel1\rangle) = Wrench1$. This can be a lookup table, or a decision process. In our implementation, the agents models the components as having open and closed male and female attachments to deduct Φ. Also note that some workspace states do not warrant any tool, because they either have an empty component, or two components that cannot be attached. We mark these function values as $\phi(s^w_i) = \emptyset$.

III. Reactive Agents

A baseline agent that is purely responsive to its environment and internal state, can be defined by an action policy that waits in the workbench when $\Phi(WorkBench) = \emptyset$, and fetches tool x, uses it, returns it, and returns to the workbench when $\Phi(WorkBench) = x$.

The obvious fallacy of this policy occurs when the same tool is needed twice in a row (which can happen with the wheels and axles, in the factory domain), resulting in a superfluous sequence of returning and then fetching the same tool. If the distance between the workspace and the tool is d, and under the assumption that the time it takes for the human to bring the next part h is smaller than $4d + 3$, the total cost of this sequence is $6d + 5$.

The naïve policy can, therefore, be improved by delaying the decision to return a tool until the state of the workbench changes. This prevents the agent from returning a tool before it is certain that it is not needed again in the next step. We call this policy *conservative tool return*. Given the time delay between the two $Uses$

$$\delta = \begin{cases} h - (2d+2), & \text{if } h > 2d+2 \\ 0, & \text{otherwise.} \end{cases}$$

The total cost of the sequence is $2d + 3 + \delta$. The gain in performance is $4d + 2 - \delta$.

However, it is straightforward to demonstrate that there is a negative impact of the "conservative tool return" strategy in the case where the next tool needed is different from the current tool. Note that the cost effect of conservative tool return is dependent not only on the known world configuration, but also on the turnaround time of the human action h, a quantity that cannot be known but only estimated by the robotic agent. Additionally, the overall expected cost effect is dependent on the probability distribution on the workbench configuration over time. It, therefore, makes sense to discuss an action selection policy based on these factors, which is the topic of the following section. We will then frame the two reactive policies discussed here as a subset of the proposed anticipatory policy.

IV. Anticipatory Action Selection

As discussed in the introduction, humans are remarkably adaptive and increasingly effective when performing repetitive trials of an identical task collaborating with a consistent teammate. The use of educated anticipatory action based on expectations of each other's behavior may be a key ingredient in the achievement of this action fluency. In this section, we will attempt to adopt this insight in the human–robot interaction domain within the discussed framework.

A necessary assumption for anticipatory action selection in our agent is that the human collaborator will follow a roughly consistent action pattern, i.e., will make similar decisions under similar circumstances.

The agent, thus, models the workbench as a first-order Markov Process.[3] The probability of the workbench state at time t, σ^w_t, is thus conditional on σ^w_{t-l} and denoted as

$$p^w_{i|j} \equiv Pr(\sigma^w_t = s_i | \sigma^w_{t-l} = s_j)$$

The agent can learn the parameters of this Markov process using a naïve Bayesian estimate. To do this, the agent keeps a one-step history of the state transitions of the workbench. A change from state s_j to state s_i increases the counter $n_{i|j}$. Consequently, $p^w_{i|j}$ is computed as

$$p^w_{i|j} = \frac{n_{i|j}}{\sum_{k=1}^{|\Sigma_W|} n_{k|j}}$$

However, in order to estimate the cost of preemptive action, as described in the following section (which is ill-defined for nonconstructive workbench states), and also to reduce the decision

[2]For example, welding can take longer than riveting, and picking up the wrench could be faster than picking up the welder.

[3]A presumably more realistic model would be to view the collaboration as a hidden Markov model, with the human state transitions being hidden, and the workbench transitions being the evidence layer of the model. However, since many of the human's state transitions do not affect the workbench state, and the probability of workbench transitions conditional on the human state transitions $Pr(\sigma^w_t = s_i | \sigma^h_t = s_j)$ are not independent of σ^w_{t-l}, it is unclear whether such a model would indeed be of value in our domain, and is, therefore, left to future investigation.

state space, the robot in our factory domain can alternatively model the probability of the tool needed based on the previous state: if $Q(x) = \{s_i : \phi(s_i) = x\}$ is the set of workbench states that warrant tool x, the new probability model learned is now

$$p_{x|j} \equiv Pr(\sigma_t^w \in Q(x) | \sigma_{t-l}^w = s_j).$$

We estimate this model as follows: a change from state s_j to state $s_i \in Q(x)$ increases the counter $n_{x|j}$. Using a Laplace correction of 1, $p_{x|j}$ is, then, estimated by

$$p_{x|j} = \frac{n_{x|j} + 1}{\sum_{k=1}^{|Tools|} n_{k|j} + 1}.$$

A. Action Selection

As the agent only perceives the workbench state (and, therefore, information about the transition distribution) when it is in the workbench state, it makes sense to make decisions in terms of *action sequences*. The acquisition of these sequences is beyond the scope of this paper, but suffice to say that, in our scenario, the agent needs only to consider action sequences that begin and terminate while it is in the workbench state.

In the discussed factory domain, we can identify four protosequences (state transitions for the full sequences are presented in brackets)

1) Pick up a tool and use it
$$[\langle s_0^r, \emptyset \rangle \rightarrow \langle s_0^r, x \rangle].$$
2) Return a tool and return to workbench
$$[\langle s_0^r, x \rangle \rightarrow \langle s_0^r, \emptyset \rangle].$$
3) Return a tool, bring a new tool, and use it
$$[\langle s_0^r, x \rangle \rightarrow \langle s_0^r, y \rangle].$$
4) Do nothing and wait
$$[s_i^r \rightarrow s_i^r].$$

The action selection process operates as follows: at any time the robot is in the workbench state, it evaluates the cost of each of the protosequences. Protosequence 1 needs to be grounded for each tool and protosequence 3 needs to be grounded for each of the currently not held tools. Given the probability distribution, the robot can compute the expected cost for choosing each of the strategies, and selects a grounded sequence optimizing for cost. In calculating the expected cost for protosequences 1–3, we need to assume that $\forall i, [h < 2d_{0i}]$. Also note that the cost in our calculations includes performing the correct action afterward. Denoting the current state of the workbench s_j, and the workbench position 0, the expected duration costs of protosequence 1–3 are

$$Cost_1(x) = p_{x|j}(2d_{0x} + 2) + \sum_{y \neq x}[p_{y|j}(3d_{0x} + d_{xy} + d_{0y} + 4)]$$

$$Cost_2(x) = \sum_{y=1}^{|Tools|}[p_{k|j}(2d_{0x} + 2d_{0y} + 3)]$$

$$Cost_3(x, y) = p_{y|j}(d_{0x} + d_{xy} + d_{y0} + 3) + \sum_{z \neq y}[p_{z|j}(d_{0x} + d_{xy} + 2d_{y0} + d_{yz} + d_{z0} + 5)].$$

Action sequence 4 is unique insofar as it is dependent not only on the state transitions in the workbench but also on the behavior of the human teammate. If the human's next workbench-changing action is at time $t + h$, the cost of waiting is the cost of performing the correct action with complete confidence, plus h. For the case that the robot is holding a tool z

$$Cost_4 = p_{z|j} + \sum_{y \neq z}[p_{y|j}(d_{0z} + d_{zy} + d_{y0} + 3)] + h.$$

For the case that the robot is not holding a tool

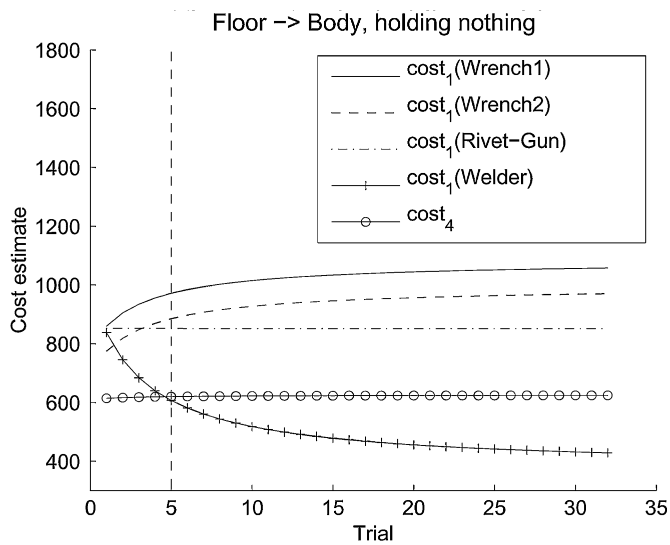
$$Cost_4 = \sum_{y=1}^{|Tools|}[p_{y|j}(2d_{0y} + 2)] + h.$$

However, since h is not directly accessible to the robotic agent, its estimate can be used as a confidence parameter, adjusting between an aggressively anticipatory behavior and a more cautious approach (see also the following text).

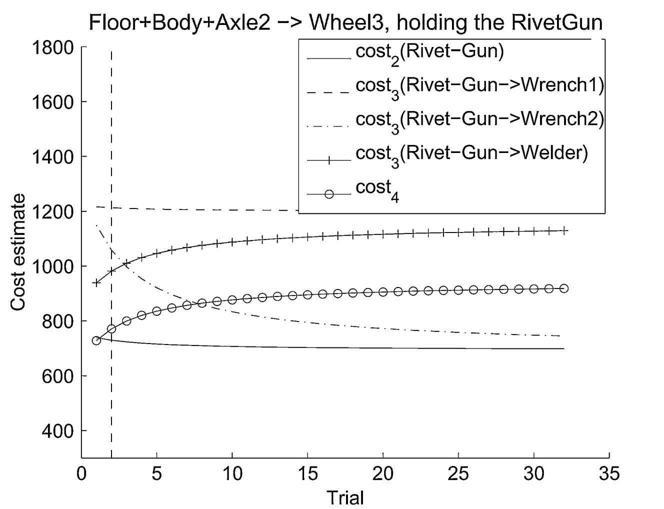
Using this notation, we can now rephrase the previously discussed reactive agent behaviors. The naïve agent's policy can be viewed as selecting protosequence 2 whenever it is holding a tool in the workbench, and selecting protosequence 1 whenever a tool is warranted. The agent employing conservative tool return can be rephrased as selecting protosequence 4 whenever no tool is warranted, and selecting protosequence 1 or 3 if a workbench state warrants a tool. This rephrasing enables comparison between the different policies, as described in the following section.

B. Analysis

Fig. 3(a) and (b) demonstrates the adaptation of cost-optimizing anticipatory action vis-a-vis a theoretical human teammate. In these figures, the factory layout is as depicted in Fig. 2, the human's action is simulated to be constant given a specific configuration, and the agent is using $h = 250$. Fig. 3(a) depicts the expected cost for the five available action sequences when the robot perceives the $Floor$ in the workbench, holding nothing, over 31 trials in which the human is simulated to consistently bring the $Body$ in this situation. We can see that Sequence 4 (waiting) is the cost-optimizing action for trials 1–4, and that getting the $Welder$ becomes the optimal anticipatory action from trial 5 onward. In contrast, Fig. 3(b) shows that when holding the $RivetGun$ and perceiving $Floor + Body + Axle2$ (with a human consistently bringing $Wheel3$ to the workbench), Sequence 2—returning the $RivetGun$—becomes optimal starting from the second trial. This difference becomes apparent considering the location of the $Wrench$ on the opposite side of the workbench, making it considerably more expensive to wait, the more confident the robot gets that the $Wrench$ is needed next. It it interesting to note that due to the particular tool arrangement, returning the $RivetGun$ and prefetching the $Wrench$ does not become cost-optimizing even after 31 trials.



(a)



(b)

Fig. 3. Cost evaluation with a simulated consistent human teammate. In (a), the robot is holding nothing and perceiving the $Floor$, with a human simulated to always bring the $Body$ next; in (b), the robot is holding the $RivetGun$ and perceiving $Floor + Body + Axle2$, with a human simulated to always bring $Wheel3$.

While it does become more optimal than waiting after eight consistent trials, the cost of an erroneous prediction, even as it becomes extremely unlikely, is still too high, resulting in a preference for Sequence 2 over Sequence 3. Note that this does not hold for other decision junctions. For example, while holding the $Welder$ with a consistent need for the $RivetGun$, prefetching it on the way back from the $Welder$ location becomes cost-optimizing on the sixth trial (not shown in the figure).

Using the analysis in the previous section, we can now compare the reactive agents to our proposed method. In the case described in Fig. 3(a), our algorithm is equal to the reactive agents (equivalent to Sequence 4) in trials 1–4, and outperforms them increasingly as the amount of evidence increases. In the case described in Fig. 3(b), the naïve reactive agent is equivalent to Sequence 2, slightly outperforming our method in the first trial, and then matching it, while the conservative

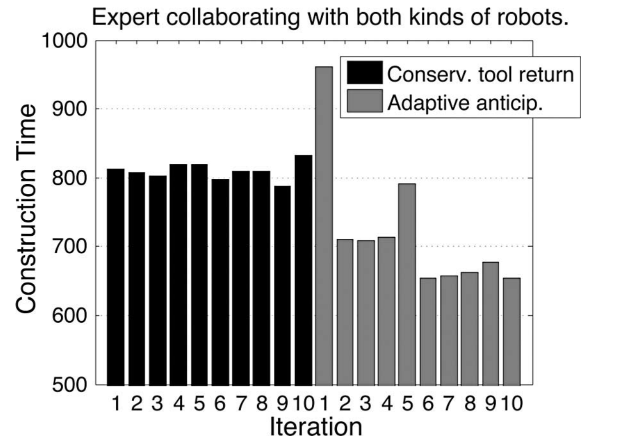


Fig. 4. Change in per-cart construction time with an expert consistent human in a pilot experiment vis-a-vis (left) the reactive agent and (right) the adaptive anticipatory agent.

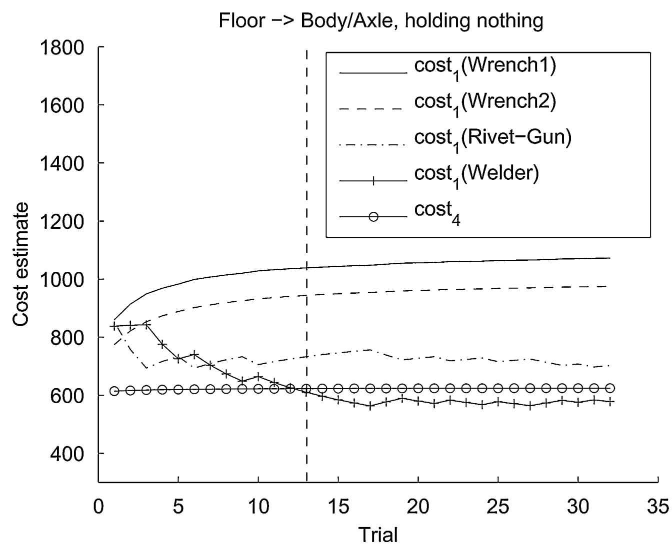


Fig. 5. Effect of an inconsistent human teammate: Graph shows perceiving the $Floor$ with a simulated human teammate producing the $Body$ with a probability of 70% and an $Axle$ with a probability of 30%.

tool return agent (equivalent to Sequence 4) chooses a more costly approach than our method from trial 2 onward. Generally speaking, using $h = 250$ in the factory scenario, we usually see the agent outperforming the reactive agents within two trials, and converging into full anticipatory behavior within 10 trials.

In an actual pilot run vis-a-vis a real-life, experienced, and consistent human teammate, we can see evidence to that effect. Whereas the reactive agent with conservative tool return remains constant at a construction cost[4] of *circa* 800, the anticipatory adaptive agent shows a significant improvement after the first trial and again at the sixth trial, finally settling at a lower per-cart construction cost of *circa* 650 (see Fig. 4).

Fig. 5 shows the adaptation of the cost-function vis-a-vis a theoretical inconsistent human teammate. In this case, given the $Floor$ in the workbench, the simulated human action is

[4]The cost units, when measured with a human teammate, are in simulation frames, running at 30 frames/s.

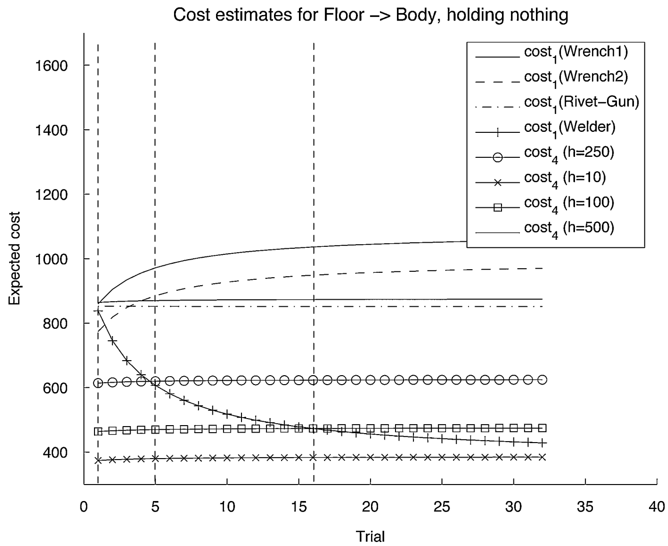


Fig. 6. Cost analysis perceiving the $Floor$ with a simulated consistent human teammate producing the $Body$, but varying the estimated human turnaround cost h.

a random variable with a fixed probability distribution, bringing the $Body$ with a probability of 70%, and an $Axle$ with a probability of 30%. The result is that waiting for the human's next move remains cost-optimizing for 12 iterations, delaying the anticipatory behavior of the agent and resulting in slower convergence into a fluent and efficient activity pattern.

A final note regarding the risk-taking parameter h, which we defined as the estimated time the human's turn takes: varying h affects the relative optimality of Sequence 4 (waiting for the human). Lowering h significantly corresponds to an expectation that the human returns very quickly with the next part, resulting in a risk-averse policy (Fig. 6). At the junction depicted in Fig. 3(a), for example, lowering h to 10 will render the decision function equivalent to the "conservative tool return" reactive agent discussed earlier. Setting $h = 500$ results in an agent performing anticipatory actions as soon as trial 1. Fixing h at 100 results in taking the correct anticipatory action at trial 16, instead of trial 5.

Ideally, h should be specific per state, as well as learned over time as the agent collects more data regarding the turnaround time of the human teammate.

V. Human Subject Study

To further investigate the effect of adaptive anticipatory action selection, we conducted a human subject study. We expected to see an increase in efficiency as predicted by the theoretical analysis, as well as an increase in the perceived contribution of the robot to the team's fluency and success.

A. Experimental Design

We recruited 32 participants (15 female) from the Massachusetts Institute of Technology (MIT) community through email solicitation and posters. Participants arrived at our laboratory and were arbitrarily assigned to one of two experiment

TABLE I
Total Cart Completion Metrics for Untrained Human Subjects in the Reactive (Group A) and Adaptive Anticipatory Condition (Group B).

Score metric	Group A mean	Group A std.dev.	Group B mean	Group B std.dev	T(25)
Best	1091.6	200.5	930.1	105.6	2.59; $p < 0.02$
Mean	1423.5	328.6	1233.3	227.5	1.73; not signif.
Final	1182.4	274.3	1030.7	154.8	1.75; not signif.

We compare each subject's best score in ten trials, mean score over ten trials, and tenth trial.

conditions. Subjects in *Group A* interacted with a reactive agent using the "conservative tool return" policy; those in *Group B* interacted with an anticipatory agent.

All participants (from both groups) received identical instructions, which described the factory setup as a video game, and were told that the human–robot team's goal is to build 10 carts, with "each team member [having] their own role in this joint effort." Also, subjects were instructed to "build carts in the least amount of time." The instructions were phrased so as to imply the importance of the team as a joint performing entity. To control for instruction bias, neither group was told whether the robot will adapt to their behavior. All participants were allowed to practice with the system before beginning the experiment.

The experimental protocol was reviewed and approved by the Institutional Review Board of the MIT.

B. Results

Of the participants, five had to be eliminated from the study. Two violated the experimental protocol, one experienced a software crash, one was significantly inattentive, resulting in scattered behavior, and, for one subject, the logging functionality was not working, resulting in a loss of data. This left us with 27 subjects, 14 in Group A and 13 in Group B. All 32 completed a poststudy survey regarding their experience.

Table I shows total cart construction measures for the population. Cost units are in simulation frames at 30 frames/s.

Each subject's best performance is significantly better at a confidence level of 98% in the adaptive anticipatory case compared to the reactive case. Measuring the mean construction time over ten trials, as well as the time for construction of the tenth cart, we find the subjects in the anticipatory case to be better (at $p < 0.1$), but not significantly at a 95% confidence level. We believe that this is, in part, due to the fact that several subjects in Group B took a number of inconsistent trials to identify that the robot was adaptive, leading to a convergence to a stable construction pattern only in the last few carts (see also Section V-D.2). According to this hypothesis, both the mean and the final cart construction cost would be lower in the anticipatory case if there were more trial runs per subject. This claim needs to be investigated in subsequent research.

1) Survey: In the postexperimental survey, we found significant differences between participants in the two groups . On a seven-point Likert scale, subjects in the anticipatory action agent "Group B" selected a significantly higher mark than those in the reactive agent "Group A" when asked whether:

1) "The robot's performance was an important contribution to the success of the team.":
   Group A : *4.88*; Group B : *6.38*; $T(30) = 2.87$;
   $p < 0.01$.
2) "The robot contributed to the fluency of the interaction.":
   Group A : *4.125*; Group B : *5.6875*; $T(30) = 2.99$;
   $p < 0.01$.
3) "It felt like the robot was committed to the success of the team.":
   Group A : *2.8*; Group B : *5.0*; $T(30) = 3.21$;
   $p < 0.005$.

The two groups did not differ significantly when subjects were asked whether they themselves were "committed to the success of the team," or whether they "trusted the robot to do the right thing at the right time." Both groups averaged between 6 and 7 on these two questions.

C. Measures of Fluency

In sum, we found significant differences between the two conditions in the subjects' *perception* of fluency as well as in their perception of the robot's commitment and contribution to the team's success. This conclusion is further embellished by the qualitative findings described in Section V-D later. At the same time, the mean (and convergent) task efficiency of the team was not significantly different between the conditions. This contradictory phenomenon could suggest that the notion of fluency, commitment, and appropriate teamwork are separate from those of simple task-time efficiency. If this is the case, we would like to discern possible quantitatively measurable causes for the above-mentioned perceptual differences.

However, while there is a large body of work measuring verbal fluency, there are no generally accepted measures of fluency in shared-location joint action, even for human teams. We, therefore, propose three fluency metric hypotheses, and compare the mean performance of the two groups along these measures in a posthoc analysis of our study.

1) *Hypothesis I: Concurrent motion*—In postexperiment interviews, some of our participants noted a sense that the team was well synchronized when "both team members were constantly in motion." We tested the hypothesis that the amount of human–robot concurrent motion was different between the anticipatory and the reactive condition. To do so, we measured the percentage of frames within each trial in which both human and robot were in motion (i.e., in transition between two location-based internal states), and, indeed, found those to be significantly different between the two groups (A : *0.227*; B : *0.322*; $T(25) = 3.11$; $p < 0.005$). Fig. 7(a) shows the mean percentage of concurrent motion for each of the 10 trials, averaged over

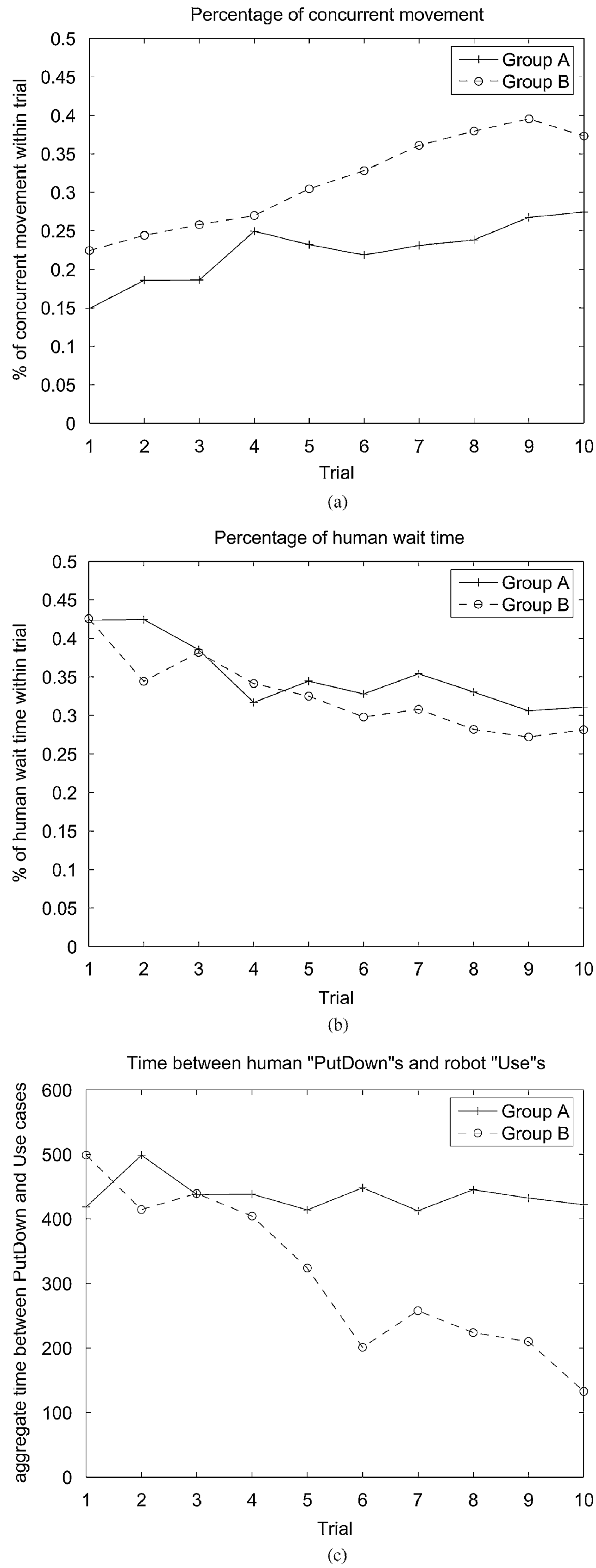


Fig. 7. Three measures of fluency per cart over ten trials, averaged over study groups A (reactive) and B (anticipatory). (a) percentage of concurrent motion within trial; (b) percentage of human idle time; (c) aggregate time between human $PutDown$ to robot Use delay.

subjects in each group. The graph shows that while the percentage of concurrent motion is improving for both groups, it does so at a higher rate in the anticipatory action condition.

2) *Hypothesis II: Human idle time*—Our second hypothesis for a measure of fluency is the amount of time the human spent waiting for the robot. We postulated that if the human was to spend much time waiting, it would feel like the team was not working fluently. However, measuring the percentage of frames in which the human waiting (i.e., not doing anything, and not in transition between two location-based states), we found no significant difference between the two groups [Fig. 7(b)]. This was true for both the mean and the convergent human idle time. Both groups decreased the human waiting time at an approximately equal rate, and with similar results.

3) *Hypothesis III: Functional delay*—We denote our third hypothesis "functional delay," i.e., we postulate that the amount of time passing between the human's action and the robot's consequent action was different between the two conditions. To test this, we measured the time between the human's $PutDown$ action and the robot's subsequent Use action. We found this measure to be significantly lower for Group B(A : *436.78*; B : *310.64*; $T(25) = 5.04$; $p < 0.001$), and more decidedly so for the second half of each subject's trial sequence, after the robot has adapted to the human's construction pattern (A : *432.07*; B : *205.08*; $T(25) = 6.28$; $p < 0.001$) [(Fig. 7(C))]. In the reactive case, there is virtually no change across trials.

While not ruling out additional factors, this evidence points in a promising direction with regards to possible quantitative measures affecting fluency in human–robot joint action. However, these findings are only initial and lay the groundwork for future research, in which each of these hypotheses needs to be separately controlled for, and evaluated for its effect on the human team member's perception of the robot's fluency, commitment, and task contribution.

D. Discussion

The open-ended segment of the postexperiment questionnaire reveals a qualitative difference between the two conditions. Several subjects in Group B noticed the anticipatory behavior and remarked on it positively, e.g., "it was nice when [the robot] anticipated my next move," or "[the] robot's anticipation of my actions was impressive and exciting." Negative remarks in Group B usually referred to a desire for even more anticipatory behavior, such as "[the robot] could do better by getting the first tool before/while I take the first part, because it was a consistent process and could be predicted," or "the robot should watch what I'm grabbing in advance."

Somewhat surprisingly, many subjects in Group A—without having been informed that the study was related to anticipatory action or that the robot was meant to be adaptive—noted with frustration that the robot did not predict their actions. We view this tendency as indicative of the fact that adaptiveness and anticipatory action are natural expectations of a robotic teammate in a repetitive task. Quotes from Group A included: "I was hoping that the robot would learn to anticipate more," "I expected more predictive behavior from the robot," "[the] robot was not able to anticipate [the] human's actions," and "it might have been more efficient if after a few carts the robot could pick up on the order in which i was bringing in the parts and be prepared with the equipment to join it."

Group A's positive comments regarding the robot's performance were limited to remarks shaped by a low level of expectation from the agent: "The robot seemed to do what was expected," "the robot did not mess up," and "the robot was highly responsive and never let the human down with its predictability," were representative responses in this condition.

1) Notions of Teamwork: It is interesting to note that several subjects in Group A noted that the team felt "lopsided," that "the human was the one who strategized, the robot just sat there," that the human "was more important than the robot," and that "the team's performance was highly dependent on human innovation." Subjects in this group concluded that "the robot seemed more like an assembly tool than a team member," that they "didn't see the robot as a team player," that the robot was used "as a tool," and one subject said that they "didn't get a sense that the robot really cared about the success of the team." In contrast, in Group B, only one subject noted that they "felt that the success or failure of the task was [their] responsibility." Conversely, one other stated that they "trusted [the robot] more over time, as it seemed to anticipate what [they were] going to do." The rest of the subjects in Group B did not address the balance of the team, the issue of trust, or that of commitment, in any way.

2) Effect of Repetition Size: As noted in Section V-B, we believe that the relatively minor improvement in mean task efficiency through anticipatory action is related to the small number of repeating trials in the experiment. Appraisal of server logs, as well as user testimony, reveals that, in many cases, subjects experimented with various construction strategies in the first few runs, which caused the Bayesian model to converge more slowly. This seemed to be particularly true when subjects noticed that the robot changed its behavior, causing them to experiment with different construction sequences in an attempt to reveal the robot's *modus operandi*. One reason for this behavior was the experiment's insistence on identical instructions for both groups, not revealing that the robot would adapt to the human's consistent behavior. Several subjects explicitly noted that the team would have performed better had they known in advance that the robot learned to anticipate their actions. Another possible way to counter this effect would be to discount the learning over time (see also Section VII).

3) Effect of "Best Score" Indicator: We also believe that the display of the game's all-time "Best Score" in the user interface was detrimental to the experiment as it might have caused subjects to experiment with different strategies instead of forming a consistent behavior pattern. Originally intended to motivate subjects to faster performance, the exceedingly good record time (only possible with a well-adapted agent) provoked subjects to question their strategy attaining a significantly worse score, and, subsequently, to change it several times over the course of the experiment.

VI. Related Work

Most work related to joint action—whether in philosophy, psychology, or artificial intelligence—has been concerned with a goal-oriented view of the problem, paying little attention to the *quality* of action meshing and fluency of teamwork, both as it is perceived by the team members, and as it affects the quantitative measures of the task.

In this body of work, joint action is usually described as solving a problem where the participants share the same goal and a common plan of execution. Grosz pointed out, in this context, that collaborative plans do not reduce to the sum of the individual plans, but consist of an interplay of actions that can only be understood as part of the joint activity [4].

In Bratman's detailed analysis of shared cooperative activity, he defines certain prerequisites for an activity to be considered shared and cooperative [5]. He stresses the importance of mutual responsiveness, commitment to the joint activity, and commitment to mutual support. Supporting Bratman's guidelines, Cohen and Levesque propose a formal approach to building artificial collaborative agents [6]. Their notion of joint intention is viewed not only as a persistent commitment of the team to a shared goal, but also implies a commitment on part of all its members to a mutual belief about the state of the goal. These principles have been used in a number of human–robot teamwork architectures [7], [8].

Much work has been done in the field of discourse theory, investigating discourse as a collaborative activity. Grosz and Sidner have analyzed the structure of discourse and, subsequently, modeled shared plans as a separate *extension*, rather than a *composition* of simple, single-agent plans [9]. Later work has further elaborated the workings of collaborative discourse, in terms of plans, beliefs, goals, and actions (e.g., [10], [11]). Collaborative discourse systems have been developed and implemented on screen-based and robotic dialog systems, taking into account both the verbal and the nonverbal aspects of discourse (e.g., [12], [13]). Still, the question of fluency in action meshing has not been part of this corpus. Moreover, as these works focused mainly on linguistic dialog, they have not addressed the case of nonverbal shared-location teamwork, or the improvement thereof through repetitive joint execution of a task.

Human–robot teamwork has also remained mostly in the turn-taking domain. Some have studied a robotic arm assisting a human in an assembly task [14]. Their work addressed issues of vision and task representation, but does not investigate joint adaptation, and does not address the timing issue. Other works study human–robot collaboration with an emphasis on dialog and control, aimed primarily at teleoperation [15], [16]. Some frame human–robot collaboration in the context of mixed-initiative control and shared autonomy, arbitrating between the robot's autonomy and direct human control, but also fail to address the question of shared-location fluency [17], [18].

Some work in shared-location human–robot collaboration has been concerned with the mechanical coordination of robots in shared tasks with humans (e.g., [19]). This work is predominantly concerned with single-action control and safety issues.

We have previously presented work in shared-location human–robot teamwork, investigating the role of nonverbal behavior on teamwork [7], [20]. While this task-level work included turn-taking and joint plans, anticipatory action and fluency have not been addressed.

Timing and synchronization have been reviewed on the motor level in the context of a human–robot synchronized tapping problem [21]. Anticipatory action, without relation to a human collaborator, has been investigated in robot navigation work, e.g., [22].

VII. Conclusion

We have presented work investigating the effect of adaptive anticipatory action on the efficiency and fluency of action in human–robot teamwork. Through this, we hope to initiate an interest in the question of shared-location action timing and fluency.

In the work contained herein, we introduced a framework for evaluating shared human–robot fluency, and have presented a cost-based anticipatory action selection mechanism. We showed initial results on both the theoretical analysis of this method and its effect on untrained humans, showing significant differences in the subject's perception of the robot's fluency, commitment, and contribution, while showing only a small difference in mean and convergent task efficiency. In order to explain this discrepancy, as well as quantitatively evaluate the notion of fluency, we proposed three fluency metric hypotheses and compared these between conditions, finding significant differences along two of these metrics.

Several improvements to our method present themselves: in the discussed framework, the robot has no knowledge of the structure of the task. Domain-specific knowledge can decrease the action space at each decision point and fortify the accuracy of the probabilities of subsequent states.

We believe that our system can also be made more robust by introducing a discount factor in the learned state transition distribution, making more recent moves by the human teammate more salient to the robot.

Furthermore, the estimate of the human's turnaround time h should be state-specific and could be learned by the robot during the collaboration.

In future work, we would like to evaluate the relative effect achieved by the state transition distribution learning, as opposed to the cost analysis during action selection. Also, the scalability of our method should be evaluated by increasing task complexity.

Additionally, the effects of anticipatory action vis-a-vis an expert—instead of a naïve—human teammate, is of interest, as is a controlled evaluation of the effects of the proposed fluency metrics on the efficiency of the task and the perceived fluency and commitment of the robot.

Finally, this anticipatory framework is now being implemented on a physical robot and we are currently conducting studies evaluating the effects of our method on human–robot fluency in a task involving a human–robot hybrid team. Through this ongoing research, we hope to evaluate the applicability of our model to real-life human–robot teamwork applications.

REFERENCES

[1] N. Sebanz, H. Bekkering, and G. Knoblich, "Joint action: Bodies and minds moving together," *Trends Cogn. Sci.*, vol. 10, no. 2, pp. 70–76, 2006.

[2] G. Knoblich and J. S. Jordan. (2003, Sep.). Action coordination in groups and individuals: Learning anticipatory control. *J. Exp. Psychol.: Learn., Mem., Cogn.* [Online] 29*(5)*, pp. 1006–1016. Available: http://dx.doi.org/10.1037/0278-7393.29.5.1006

[3] M. Wilson and G. Knoblich, "The case for motor involvement in perceiving conspeci.cs," *Psychol. Bull.*, vol. 131, pp. 460–473, 2005.

[4] B. J. Grosz, "Collaborative systems," *AI Mag.*, vol. 17, no. 2, pp. 67–85, 1996.

[5] M. Bratman, "Shared cooperative activity," *Philos. Rev.*, vol. 101, no. 2, pp. 327–341, 1992.

[6] H. J. Levesque, P. R. Cohen, and J. H. T. Nunes, "On acting together," in *Proc. AAAI 1990*, Boston, MA, pp. 94–99.

[7] G. Hoffman and C. Breazeal, "Collaboration in human–robot teams," in *Proc. AIAA 1st Intell. Syst. Tech. Conf.*, Chicago, IL, Sep. 2004.

[8] R. Alami, A. Clodic, V. Montreuil, E. A. Sisbot, and R. Chatila, "Task planning for human–robot interaction," in *Proc. 2005 Joint Conf. Smart Objects Ambient Intell.* New York, NY, pp. 81–85.

[9] B. J. Grosz and C. L. Sidner, "Plans for discourse," in *Intentions Communication*, P. R. Cohen, J. Morgan, and M. E. Pollack, Eds. Cambridge, MA: MIT Press, 1990, ch. 20, pp. 417–444.

[10] L. Lambert and S. Carberry, "A tripartite plan-based model of dialogue," in *Proc. 29th Conf. Assoc. Comput. Linguistics*, 1991, pp. 47–54.

[11] J. F. Allen. (1995). The TRAINS project: A case study in building a conversational planning agent. *J. Exp. Theoretical AI* vol. 7, pp. 7–48. Available: citeseer.ist.psu.edu/allen94train.html

[12] C. Rich, C. L. Sidner, and N. Lesh, "Collagen: Applying collaborative discourse theory to human–computer collaboration," *AI Mag.*, vol. 22, no. 4, pp. 15–25, 2001.

[13] C. Sidner, C. Lee, and N. Lesh, "Engagement rules for human–robot collaborative interaction," in *Proc. IEEE Int. Conf. Syst., Man Cybern.*, 2003, vol. 4, pp. 3957–3962.

[14] H. Kimura, T. Horiuchi, and K. Ikeuchi, "Taskmodel based human robot cooperation using vision," in *Proc. IEEE Int. Conf. Intell. Robots Syst.*, 1999, pp. 701–706.

[15] T. W. Fong, C. Thorpe, and C. Baur, "Multirobot remote driving with collaborative control," *IEEE Trans. Ind. Electron.*, vol. 50, no. 4, pp. 699–704, Aug. 2003.

[16] H. Jones and S. Rock, "Dialogue-based human–robot interaction for space construction teams" in *Proc. IEEE Aerosp. Conf.*, 2002, vol. 7, pp. 3645–3653.

[17] D. J. Bruemmer, D. D. Dudenhoeffer, and J. Marble, "Dynamic autonomy for urban search and rescue," presented at the 2002 AAAI Mobile Robot Workshop, Edmonton, AB, Canada, Aug.

[18] M. Goodrich, D. Olsen, J. Crandall, and T. Palmer, "Experiments in adjustable autonomy," in *Proc. IJCAI Workshop Auton., Delegation Control: Interact. Intell. Agents*, 2001.

[19] H. Woern and T. Laengle, "Cooperation between human beings and robot systems in an industrial environment," in *Proc. Mechatron. Robot.*, 2000, vol. 1, pp. 156–165.

[20] C. Breazeal, C. D. Kidd, A. L. Thomaz, G. Hoffman, and M. Berlin, "Effects of nonverbal communication on efficiency and robustness in human–robot teamwork," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Edmonton, AB, Canada, Aug. 2005, pp. 708–713.

[21] T. Komatsu and Y. Miyake, "Temporal development of dual timing mechanism in synchronization tapping task," in *Proc. 13th IEEE Int. Workshop Robot Hum. Interact. Commun.*, Sep. 2004, pp. 181–186.

[22] Y. Endo, "Anticipatory and improvisational robot via recollection and exploitation of episodic memories," in *Proc. AAAI Fall Symp.*, 2005.

Guy Hoffman received the M.Sc. degree in computer science from Tel Aviv University, Tel Aviv, Israel, and the Ph.D. degree from the MIT Media Laboratory, Massachusetts Institute of Technology, Cambridge.

His current research interests include human–robot teamwork, with a focus on fluency mechanisms in joint action.

Cynthia Breazeal received the B.S. degree in electrical and computer engineering from the University of California, Santa Barbara, and the M.S. and Sc.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, in 1993 and 2000, respectively.

She is currently an Associate Professor of media arts and sciences with the MIT Media Laboratory, MIT. Her interests include human-like robots that can interact, cooperate, and learn in natural social ways with humans.