Who Takes the Lead? Automated Scheduling for Human-Robot Teams

Brenda Castro,[1*] **Montana Roberts,**[1*] **Karla Mena,**[2] **Jim Boerkoel**[1]
Human Experience & Agent Teamwork Laboratory (https://cs.hmc.edu/HEAT/)
[1]Harvey Mudd College, Claremont, CA; [2] Mountain View High School, El Monte, CA
[1]{bcastro, mwroberts, boerkoel}@hmc.edu, [2] karlamena946@gmail.com

Abstract

Scheduling interactions between humans and robots presents unique challenges—while robots do not have humans' natural ability to improvise and adapt to new setbacks, humans are not able to work with the same precision as robots. Additionally, hesitation, interruptions, and anticipatory action all influence a human's perception and efficiency in social tasks, but are not inherent features of current algorithms. This paper explores both the challenges and opportunities of automated scheduling as a useful tool for human-robot interactions. We contribute an initial exploratory pilot study that suggests that when a robot takes the lead in dictating a schedule, there are gains in team efficiency without loss of humans' perceived comfort.

Introduction

As collaborative robots become commonplace, learning effective social scheduling strategies will be crucial (Shah et al. 2011; Lemaignan et al. 2017). First, humans' execution of tasks tends to be highly sensitive to timing (e.g., a driver hesitating to take their turn at an intersection). Second, planning a course of action that achieves a team's goal can be computationally expensive; effectively scheduling that plan can improve its robustness and usefulness in real-world human-robot settings. Finally, precise and intuitive scheduling of human-robot teams will promote better collaboration and resource utilization.

Humans successfully and intuitively manage difficult scheduling problems on a daily basis (e.g., navigating a busy hallway or intersection, executing a team activity at work, etc.) without having to explicitly represent or reason about the various scheduling constraints in play. This motivates the main thrust of our research—*how do we use the AI tool of automated scheduling to facilitate human-robot interactions that are as intuitive and fluid as human-human interactions?*

This short paper explores solving human-robot teamwork problems using automated scheduling tools, highlighting technical and practical advantages over other approaches. Our exploratory pilot study considers the high-level algorithmic challenge of who should dictate the schedule for human-robot codependent tasks. Our motivating hypotheses are (1) that overall team efficiency can be improved when the robot takes the lead, but (2) humans' sense of team, safety, and comfort will be higher when the human takes the lead. We design two human-robot tasks involving Rethink Robotics' Sawyer platform that allow us to explore these hypotheses and we also include a preliminary analysis that will guide future investigations.

Automated Scheduling

As highlighted by recent workshops on the topic of timing in human-robot interactions (Hoffman, Cakmak, and Chao 2014), precise timing is critical for the efficacy, efficiency, and fluidity of human-robot teams. Finding temporally feasible plans optimal according to some objective function is a well-studied, albeit extremely difficult, problem. However, the best plan may be of limited or no use once something unexpected occurs. This work focuses on scheduling the execution of events of an existing plan to best hedge against the uncertainty introduced by a human. This introduces two critical challenges: (1) representational—humans introduce new sources and types of temporal uncertainty, and (2) algorithmic/strategic—humans may have different expectations on interpreting and responding to robot timing.

Temporal Constraint Networks

Temporal constraint networks assist in the scheduling and execution of events by maintaining a network of constraints that restrict the duration or passage of time between events, such as the start or end times of various tasks or activities. A *Simple Temporal Network (STN)* is an example of a temporal constraint network where the ordering and timing between events are constrained by single intervals of time (Dechter, Meiri, and Pearl 1991). STNs are popular choices for the monitoring and execution of schedules because they are flexible to disruptions in the schedule and poly-time efficient to maintain. Disjunctive versions of temporal networks (e.g., allowing activities to be reordered) also exist, but are much more computationally expensive. While disjunction is clearly important for human-robot tasks (e.g., the desire to decide the order in which to complete tasks), for clarity we will discuss the various opportunities and challenges posed by human teammates in terms of STNs.

[*]Primary authors, listed alphabetically, contributed equally.

Human-based Temporal Uncertainty

How to best represent the scheduling uncertainty introduced by interactions with a human teammate is an open question. One approach would be to treat humans as any another agent. However, we contend that humans' interactions with robots should function like their day to day interactions with other humans, and thus not require a human to explicitly model or reason over their schedules. Thus, we take a robot-centric view of a human as primarily a source of scheduling *uncertainty*. An *STN with Uncertainty (STNU)* permits this by noting that some durations are unknowable before execution and thus will be set by nature (e.g., a human teammate) at execution time (Vidal and Ghallab 1996). *Probabilistic STNs* (Tsamardinos 2002; Brooks et al. 2015) instead assume that the temporal uncertainty can be modeled as a probability density function.

Open questions that we hope to address over the course of this project include: Which, if either, of these models is better suited for modeling the uncertainty of a human? Are there models of temporal uncertainty that generalize across a population—or even across repeated interactions with a single human teammate? Finally, how are such models elicited, learned, or adapted in practice?

Scheduling Strategies for Social Tasks

Automated scheduling technologies are particularly useful at monitoring the execution of a plan and flexibly dispatching scheduling advice in response to how the plan is unfolding (Shah and Williams 2007). Indeed, execution algorithms that mimic human-human interactions have been shown to be effective in improving the efficacy and fluency of human-robot tasks (Shah et al. 2011). We believe such systems can be improved by explicitly modeling and reasoning over the uncertainty introduced by human teammates. However, this presents the additional challenge of proactively hedging against the uncertainty of events that it does not control. One strategy for dealing with this in STNUs is to first make the temporal network *controllable*—that is, finding timings of events that an agent controls so that the schedule is guaranteed to work regardless of how the uncertain events are chosen (e.g., Vidal and Ghallab 1996). In PSTNs, there is similar work that tries to control for as much of the uncertainty as possible to maximize the robustness of the schedule (Fang, Yu, and Williams 2014; Santana et al. 2016; Lund et al. 2017).

We suspect that uncertainty-aware scheduling methods will be effective for human-robot teams. However, there are some unique challenges and opportunities when interacting with humans. On one hand, humans tend to lack both awareness of the detailed scheduling constraints and ability to time their execution with precision. On the other hand, humans are notably adaptable, often using innate heuristics to account for and adjust to deviations and disruptions in their tasks. This motivates the primary question of our exploratory pilot study: who should take the lead in human-robot team tasks? Or alternatively, should the robot passively try to adapt to the uncertainty introduced by a human teammate or proactively try to nudge their human teammate to influence or mitigate the uncertainty? Our goal in this pilot study is to provide a proof of concept that simply changing the scheduling paradigm can influence both team performance and humans' perceptions of team performance. The pilot study will provide general guidance as to which kinds of general automated scheduling representations and strategies we should pursue in future studies.

Who Takes the Lead?

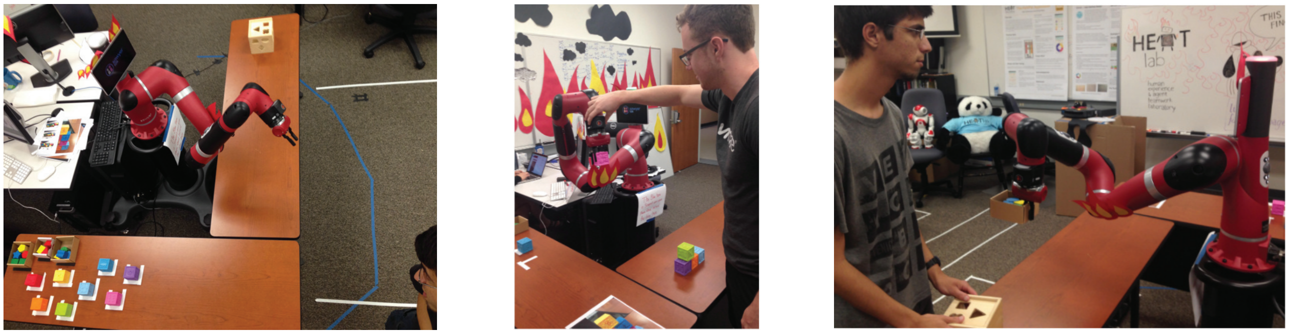
Reconciling the very different natures of humans and robots is challenging. Robots are far more precise than humans, and thus task completion times are predictable and reliable, but potentially brittle to disturbances. Humans, by contrast, bring more temporal uncertainty and adaptability, often making both unanticipated mistakes and clever fixes. Our overarching goal is to optimize team performance by utilizing the complementary strengths of humans and robots. The motivating hypotheses of our exploratory study are (1) that overall team efficiency can be improved when the robot takes the lead by dictating the schedule, but (2) humans' comfort, safety, and sense of team will be higher when the human takes the lead.

Experimental Setup

For the study, we developed two tasks that require one robot and one human agent to work together. In both, Sawyer and the participant share a resource and assembly station (see Figure 1a), which are set up in the same manner for all participants. For the **stacking task** (Figure 1b), participants are given a photograph of a block-based structure to build, with the restriction that only one agent can use the assembly station at a time. The participant is instructed to alternate placing blocks with Sawyer in the designated order. The task is considered successful if the structure is complete at the end and at no point are both agents in the assembly station at once. In the **shapes task** (Figure 1c), the robot brings the human three boxes of wooden shapes that the human must place in the designated holes in a child's sorting box before Sawyer brings the next set of shapes, at which point the participant must leave the unfinished blocks and work only on the next set of shapes. This task is considered successful if the human agent finishes each box before Sawyer sets the next box down on the table. Every participant completed both tasks three times in a row.

The 32 participants, all adults (primarily students) affiliated with Harvey Mudd College, were divided into two groups of 16. The first group participated in the **human-led** setting (Figure 1b), in which the human takes the lead and indicates to Sawyer (by pressing a button) when to complete its next predetermined subtask–which might be stacking the next block or fetching the next set of shapes–and then simply waits for the human to indicate for it to move again. This group acted as a control–they were not given any time constraints, but were instructed to be as efficient as possible and shown a large timer that kept track of task time.

The second group participated in the **robot-led** version (Figure 1c), in which Sawyer determines the pace at which the team proceeds and communicates with participants by



(a) For both tasks, the resource station is on the left and assembly station is on the right. (b) Human-led: Sawyer is signaled to proceed using the cuff button. (c) Robot-led: Sawyer flashes a red light when working and green light when idling.

Figure 1: Photos of our experimental setup (a): the stacking task is illustrated in (b) and shapes task in (c).

flashing a red light when completing its subtask and a green light when waiting for the human to do theirs. In the robot-led setting, Sawyer's pace was determined by adjusting the time it waited between its subtasks for human participants to complete their respective subtask. We set these wait times to be the median of the human-dictated wait times from successful human-led runs. Thus, Sawyer encourages the human to follow the a strategy that is at least as efficient as the median participant from the human-led setting. However, this increases the risk that a participant may fail to complete their portion of the task in time (i.e., before Sawyer encroaches on the assembly station).

To the extent possible, we attempt to control for learning effects. For the stacking task, all participants repeat the same task three times regardless of condition, thus mitigating differences in learning effects. For the shapes task, we used three unique sets of blocks for the three iterations (consistent across all participants) to mitigate the learning effects.

Measuring Team Success Our design allows us to explore how different levels of autonomy impact both team performance and human perceptions. The primary metric of team performance that we use in our initial pilots is overall *team efficiency*, which we measure with overall task completion time and success rate. However, we recognize that there are many additional metrics that will also be critical for the success of automated scheduling in human-robot teams. In particular, Hoffman and Breazeal (2007) suggest using concurrent motion, human idle time, and time between human and robot actions as objective metrics that contribute to human teammates' perception of *team fluency*. We also measure the qualitative experiences of human teammates by adopting survey items that measure aspects of job strain (Ostry et al. 2001), sense of team (Hoffman and Breazeal 2007), as well as team satisfaction and perceived comfort and safety (Lasota and Shah 2015).

Trends and Insights

Next, we discuss insights that will direct our future studies.

Human sources of temporal uncertainty Human participants tended to exhibit a large variability with many outliers in the amount of time they would make the robot wait in the human-led version of the study. The uncertainty tended to exhibit a strong positive skew with high variance; however, both the skew and the variance dampened over repeated iterations. We suspect that temporal uncertainty introduced by humans will change as a function of task difficulty and human autonomy.

This pilot demonstrates that while it may be difficult to accurately model the temporal uncertainty of first encounters with participants (suggesting the use of agnostic models such as STNUs), humans quickly settle into trends that improve and become easier to capture over time (e.g., as PSTNs). Thus, an effective automated scheduler should learn to quickly tailor its model of temporal uncertainty to its teammate to promote fluid, intuitive interactions. We also suspect that a robot could subtly influence the nature of the scheduling uncertainty using the same techniques that humans do (e.g., anticipatory action, hesitation, etc.).

Team efficiency and Fluency Since we used median wait times from the successful human-led tasks to decide Sawyer's schedule for the robot-led task, we would expect roughly half of the participants to successfully complete tasks in the robot-led version if there was no difference in human efficiency. However, the overall success rate was 85.4% across all iterations of all tasks and tended to improve with each iteration for both tasks (starting at 81.2% for the first iteration and improving to 90.6% by the last iteration). As shown in Figure 2, which plots the median and lower-/upper-quartile participant completion times, the robot-led versions of both tasks tended to be faster across a range of participants. We successfully reject the null hypothesis—that the robot-led vs. human-led setting leads to no difference in completion time—using a two-way ANOVA, where the two factors are the robot- vs. human-led setting and the task iteration ($p < 0.005$). Thus, despite encouraging human-led participants to complete the team tasks as fast as possible, our hypothesis that having a robot dictate pace would improve overall efficiency, measured as average task completion time, is validated.

As a measure of fluency in human-robot teams, we also measured concurrent time for each task—that is, how much

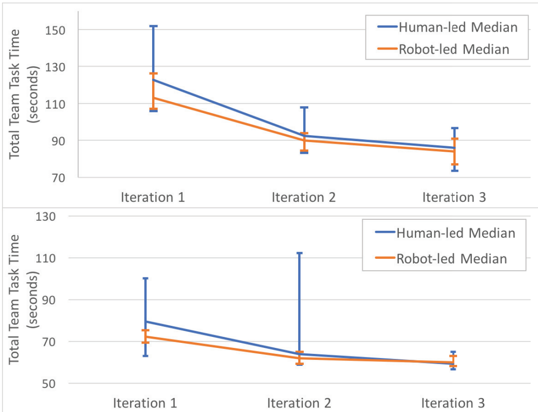


Figure 2: The median task completion time for both our stacking (top) and shape (bottom) tasks. Error bars represent the upper and lower quartiles.

time during tasks both the human and robot were actively working. In the stacking task, the proportion of time spent in concurrent motion was nearly identical (12.3 in the human-led vs. 12.1 in the robot-led). Interestingly, for the shapes task, the robot-led setting led to nearly *twice* the proportion of time spent in concurrent motion (33.4% vs. 17.9%). We suspect that this is due, in part, to the differences in level of difficulty between the two tasks. In the relatively simple stacking task, initially we see the robot-led condition lead to twice as much concurrent motion, but this flips by round 2 as we see participants quickly adapt and better tailor the robot timing to their work flow. In the more cognitively and physically demanding shapes task, the robot-led version resulted in consistent gains in concurrent motion across all iterations. This points to the usefulness of anticipatory scheduling, particularly for more demanding tasks, and also the need to quickly adapt to the pace of human teammates.

Comfort, Safety, and Sense of Team Another important aspect affecting the long-term productivity of a team is how the team members feel working together. After each iteration of each task, all participants filled out a short job strain questionnaire comprised of five questions on 5-point Likert scales ranging from 1, meaning "Strongly Disagree," to 5, meaning "Strongly Agree." The questionnaire inquired about the worker's freedom to complete the task, time allotted, workload, and effort involved.

We expected that participants in the robot-led group would have higher levels of job strain than those in the human-led group and that job strain would increase with each iteration, as Sawyer moved more quickly. However, we found very little difference between the two groups. Interestingly, the only appreciable difference was on the statement "On this job, the worker has a lot of freedom to decide how to do the work." Participants in the human-led group responded with an averages of 3.64 over all tasks and iterations. Participants in the robot-led group had an average of 3.16. Using a two-tailed Mann-Whitney U-test, this statement yielded a statistically significant difference between the two groups ($p < 0.01$). No other statement yielded a difference of $p < 0.05$. We are unable to reject the null hypothesis–that participants in both groups felt similar levels of job strain–except for in the case of perceived freedom. This could reflect the sample population's comfort with robots, or it could mean that participants did not feel threatened or stressed taking orders from a robot teammate, highlighting greater opportunity for robot-led scheduling interventions. It also highlights that a human-responsive robot-led team might provide a nice balance of curated, comfortable interaction, while balancing and encouraging overall task efficiency. Our findings are consistent with previous related studies (Gombolay et al. 2015).

Future Directions

In this paper, we explore the usefulness of automated scheduling for human-robot teamwork. We explore existing representations and algorithms for dealing with the temporal uncertainty humans introduce. Our exploratory pilot study suggests that humans respond favorably to having a robot dictate the schedule by anticipating its teammate's likely completion time both to promote increased concurrency and to subtly nudge human participants towards more optimal team performance. Surprisingly, we saw these gains without appreciable losses in human teammates' comfort level or sense of team. Our investigation suggests that temporal uncertainty is generally high during first encounters with a new human teammate, but quickly settles into predictable trends. We hypothesize that the best scheduling approach is one where a robot adapts its prior model of scheduling uncertainty to its teammate, learning to dynamically adjust to, and opportunistically nudge, its human teammate in a way that balances team efficiency and fluency.

In the future, we hope to conduct more carefully controlled versions of this study to test whether robots can meaningfully capture useful, accurate models of the temporal uncertainty of their teammates and whether we can design scheduling algorithms that promote better team efficiency and fluency. We are also interested in comparing our uncertainty-aware algorithms to previous robot-directed HRI executives (Shah et al. 2011; Gombolay et al. 2015). More broadly, while much work in the field of HRI has focused on how robots should adapt to humans, this paper suggests that we may also want to exploit the strength of humans to adapt to robots. We believe that by carefully designing the algorithms we use to schedule interactions between humans and robots, we can augment current HRI planning tools to be more effective, leading to robots that better understand their roles in team tasks.

Acknowledgments

Funding for this work was graciously provided by the National Science Foundation under grants IIS-1651822 and CNS-1659805, Harvey Mudd College, the Mellon Mays Undergraduate Fellowship Program, and U.S. Department of Education Upward Bound grant program. Thanks to the anonymous reviewers, HMC faculty, staff and HEATlab members for their support and constructive feedback.

References

Brooks, J.; Reed, E.; Gruver, A.; and Boerkoel, J. C. 2015. Robustness in probabilistic temporal planning. In *Proc. of AAAI-15*, 3239–3246.

Dechter, R.; Meiri, I.; and Pearl, J. 1991. Temporal constraint networks. In *Knowledge Representation*, volume 49, 61–95.

Fang, C.; Yu, P.; and Williams, B. C. 2014. Chance-constrained probabilistic simple temporal problems. In *Proc. of AAAI-14*, 2264–2270.

Gombolay, M. C.; Gutierrez, R. A.; Clarke, S. G.; Sturla, G. F.; and Shah, J. A. 2015. Decision-making authority, team efficiency and human worker satisfaction in mixed human–robot teams. *Autonomous Robots* 39(3):293–312.

Hoffman, G., and Breazeal, C. 2007. Effects of anticipatory action on human-robot teamwork efficiency, fluency, and perception of team. In *Proc. of HRI*, 1–8.

Hoffman, G.; Cakmak, M.; and Chao, C. 2014. Timing in human-robot interaction. In *Proc. of HRI*, 509–510.

Lasota, P. A., and Shah, J. A. 2015. Analyzing the effects of human-aware motion planning on close-proximity human–robot collaboration. *Human factors* 57(1):21–33.

Lemaignan, S.; Warnier, M.; Sisbot, E. A.; Clodic, A.; and Alami, R. 2017. Artificial cognition for social human-robot interaction: An implementation. *Artificial Intelligence* 247:45–69.

Lund, K.; Dietrich, S.; Chow, S.; and Boerkoel, J. 2017. Robust execution of probabilistic temporal plans. In *AAAI*, 3597–3604.

Ostry, A.; Marion, S.; Demers, P.; Hershler, R.; Kelly, S.; Teschke, K.; and Hertzman, C. 2001. Measuring psychosocial job strain with the job content questionnaire using experienced job evaluators. *American journal of industrial medicine* 39(4):397–401.

Santana, P.; Vaquero, T.; Toledo, C.; Wang, A.; Fang, C.; and Williams, B. 2016. Paris: a polynomial-time, risk-sensitive scheduling algorithm for probabilistic simple temporal networks with uncertainty. In *Proc. of ICAPS-16*, 267–275.

Shah, J., and Williams, B. 2007. A fast incremental algorithm for maintaining dispatchability of partially controllable plans. In *Proceedings of the Seventeenth International Conference on Automated Planning and Scheduling (ICAPS 2007)*, 296–303.

Shah, J.; Wiken, J.; Williams, B.; and Breazeal, C. 2011. Improved human-robot team performance using chaski, a human-inspired plan execution system. In *Proceedings of the 6th international conference on Human-robot interaction*, 29–36. ACM.

Tsamardinos, I. 2002. A probabilistic approach to robust execution of temporal plans with uncertainty. In *Methods and Applications of Artificial Intelligence*. Springer. 97–108.

Vidal, T., and Ghallab, M. 1996. Dealing with uncertain durations in temporal constraint networks dedicated to planning. In *Proc. ECAI-96*, 48–54.