

Generating Plans that Predict Themselves

Jaime F. Fisac^{1*}, Chang Liu^{2*}, Jessica B. Hamrick^{3*},
Shankar Sastry¹, J. Karl Hedrick², Thomas L. Griffiths³, Anca D. Dragan¹

¹ Department of Electrical Engineering and Computer Sciences

² Department of Mechanical Engineering

³ Department of Psychology

University of California, Berkeley, CA 94720 U.S.A.

{jffisac, changliu, jhamrick, tom_griffiths, anca}@berkeley.edu
sastry@eecs.berkeley.edu, khedrick@me.berkeley.edu

Abstract. Collaboration requires coordination, and we coordinate by anticipating our teammates’ future actions and adapting to their plan. In some cases, our teammates’ actions early on can give us a clear idea of what the remainder of their plan is, i.e. what action sequence we should expect. In others, they might leave us less confident, or even lead us to the wrong conclusion. Our goal is for robot actions to fall in the first category: we want to enable robots to select their actions in such a way that human collaborators can easily use them to correctly anticipate what will follow. While previous work has focused on finding initial plans that convey a set goal, here we focus on finding two portions of a plan such that the initial portion conveys the final one. We introduce t -predictability: a measure that quantifies the accuracy and confidence with which human observers can predict the remaining robot plan from the overall task goal and the observed initial t actions in the plan. We contribute a method for generating t -predictable plans: we search for a full plan that accomplishes the task, but in which the first t actions make it as easy as possible to infer the remaining ones. The result is often different from the most efficient plan, in which the initial actions might leave a lot of ambiguity as to how the task will be completed. Through an on-line experiment and an in-person user study with physical robots, we find that our approach outperforms a traditional efficiency-based planner in objective and subjective collaboration metrics.

1 Introduction

With robots stepping out of industrial cages and into mixed workspaces shared with human beings, human-robot collaboration is becoming more and more important [4, 8, 18]. There is unfortunately a history of serious and sometimes tragic failures in human-automation systems due to inadequate interaction between machines and their operators [12, 16, 17]. The most common reasons for this are mode confusion and “automation surprises”, i.e. *misalignments* between what the automated agent is planning to do and what the human believes it is planning to do.

* These authors contributed equally.

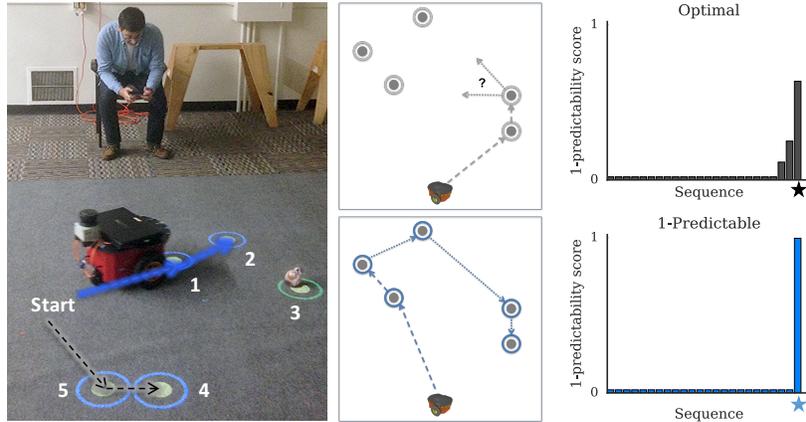


Fig. 1. Left: A participant controls the toy robot to clear the target that he believes will be third in the autonomous robot’s sequence. Center: Schematic of the trajectory followed by the optimal (top) and 1-predictable (bottom) planner. Right: Predictability scores for each sequence under the optimal and 1-predictable planners after the robot has visited a single target. Here, optimizing for 1-predictability leaves a single high-score sequence, making it easier to identify.

Our goal is to eliminate such misalignments: we want humans to be able to infer what a robot is planning to do during a collaborative task. This is important even beyond safety reasons [1], because it enables the human to adapt their own actions to the robot and more effectively achieve common goals [14, 21, 22].

Traditionally, human-robot collaboration work has focused on inferring human plans and adapting the robot’s plan in response [11, 13]. In contrast, here we are interested in the *opposite*—making sure that humans can make these inferences about the robot. We envision that ultimately these two approaches will need to work in conjunction to achieve fluent collaboration. Though it is possible for robots to explicitly communicate their plans via language or text, we focus here on what the beginning of the plan itself implies because 1) people make inferences based on actions [2], and 2) certain scenarios such as the outdoors or busy factories might render explicit channels undesirable.

We introduce a property of a robot plan that we call *t-predictability*: a plan is *t-predictable* if a human can infer the robot’s remaining actions in a task from having observed only the first *t* actions and from knowing the robot’s overall goal in the task. We make the following contributions based on this property:

An algorithm for generating *t-predictable* plans. To generate robot plans that are *t-predictable*, we introduce a model for how humans might infer future plans, building on models for action interpretation and plan recognition [2, 3]. We then propose a planning algorithm that optimizes for *t-predictability*, i.e. it optimizes plans for how easy it will be for a human to infer the last actions from the initial ones.

Prior work has focused on legibility: generating trajectories that communicate a *set goal* via the initial robot *motion* [5, 9, 19, 20]. This is less useful in task planning situ-

ations, where the human already knows the task goal. Instead, t -predictability is about communicating the *sequence of future actions* that the robot will take given a *known goal*. The important difference is that these actions are not set *a priori*: optimizing for t -predictability means *changing not just the initial, but also the final part* of a plan. It is about finding a final part that is easy to communicate, along with an initial part that communicates it.

Our insight is that initial actions can be used to clarify what future actions will be. We find that in many situations, the robot can select initial actions that might seem somewhat surprising at first, but that make the remaining sequence of actions trivial to anticipate (or “auto-complete”). Fig. 1 shows an example. If the robot acts optimally, it will go to target 5 first, making target 4 an obvious candidate as a next step, but leaving the future ordering of targets 1, 2, 3 somewhat ambiguous (high probability on multiple sequences). On the other hand, if the robot instead chooses target 1, users can with high probability predict that the remaining sequence will be 2-3-4-5.

An online user study testing that we can generate t -predictable plans. We conduct an online user study in which participants observe a robot’s partial plan, and anticipate the order of the remaining actions. We find that participants are significantly better at anticipating the correct order when the robot is planning for t -predictability. We also find a $r = 0.87$ correlation between our model’s prediction of the probability of success, and the participants’ actual success rate (see Sec. 3.6 for details).

An in-person user study testing the implications on human-robot collaboration. Armed with evidence that we can make plans t -predictable, we move on to an in-person study that puts participants in a collaborative task with the robot, and study the advantages of t -predictability on objective and subjective collaboration metrics. We find that participants were more effective at the task, and prefer to work with a t -predictable robot than with an optimal robot.

2 Defining and Optimizing for t -Predictability

t -Predictability. We consider a task planning problem from a starting state $S \in \mathcal{S}$ with an overall goal $G \in \mathcal{G}$ that can be achieved through a series of actions, called a *plan*, within a finite horizon T . Let \mathcal{A} denote the space of all feasible plans of length (up to) T that achieve the goal.

Definition 1. The t -predictability \mathcal{P}_t of a feasible plan $\mathbf{a} = [a_1, a_2, \dots, a_T]$ that achieves an overall goal G is the probability with which an observer can correctly infer $[a_{t+1}, \dots, a_T]$ after observing $[a_1, \dots, a_t]$, and knowing the overall goal G . The t -predictability score, defined as $\mathcal{P}_t(\mathbf{a}) = P(a_{t+1}, \dots, a_T | S, G, a_1, \dots, a_t)$, quantifies the level of t -predictability.

Definition 2. A t -predictable planner generates the plan that maximizes t -predictability out of all those that achieve the overall goal G . That is, a t -predictable planner generates the action series \mathbf{a}^* such that $\mathbf{a}^* = \arg \max_{\mathbf{a} \in \mathcal{A}} \mathcal{P}_t(\mathbf{a})$.

This is equivalent, by the general product rule, to:

$$\mathbf{a}^* = \arg \max_{\mathbf{a} \in \mathcal{A}} \frac{P(a_1, \dots, a_T | S, G)}{\sum_{[\tilde{a}_{t+1}, \dots, \tilde{a}_T]} P(a_1, \dots, a_t, \tilde{a}_{t+1}, \dots, \tilde{a}_T | S, G)} \quad (1)$$

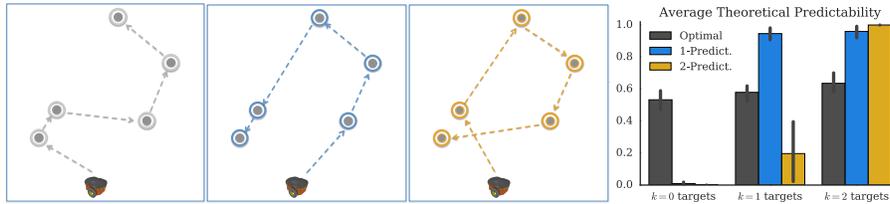


Fig. 2. Average theoretical predictability. Left: a typical task layout where the t -predictable planners generate three different sequences. Right: average theoretical predictability of task layouts used in the online experiment under different numbers of observed targets (k) of sequences generated by different t -predictable planners. In all cases, the highest value corresponds to $k = t$.

Illustrative Example. Fig. 2 shows the outcome of optimizing for t -predictability for $t = 0, 1$, and 2, and plots the t -predictability score for each plan. The 0-predictable plan (gray, left) is best when the observer sees no actions, because it is the optimal plan. However, this is no longer the best plan when the observer gets to see the first action: whereas there are multiple low-cost remaining sequences after the first action in this plan, there is only one low-cost remaining sequence after the first action in the 1-predictable plan (blue, center). The first action in the 2-predictable (orange, right) seems irrational, but this plan is optimized for when the observer gets to see the first two actions: indeed, after the first two actions, the remaining plan is maximally clear.

Relation to Predictability. t -predictability generalizes predictability [7]. For $t = 0$, the t -predictability of a plan simply becomes its predictability, that is, the ease with which the entire sequence of actions can be inferred with knowledge of the overall goal G , i.e. $\mathcal{P}_0 = P(a_1, \dots, a_T | S, G)$.

Relation to Legibility. Legibility [5] as applied to task planning would maximize the probability of the goal given the beginning of the plan, i.e. $P(G | S, a_1, \dots, a_t)$. In contrast, with t -predictability the robot is given a high-level goal describing some state that the world needs to be brought into (for example, clearing all objects from a table), and the observer is *already* aware of this goal. Instead of communicating the goal, the robot conveys the remainder of the plan using the first few elements, maximizing $P(a_{t+1}, \dots, a_T | S, G, a_1, \dots, a_t)$.

One important implication is that for t -predictability, unlike for legibility, *there is no a-priori set entity to be conveyed*. The algorithm searches for *both* a beginning and a remainder of a plan such that, by observing the former, the observer can correctly guess the latter.

Furthermore, legibility and t -predictability entail a different kind of information encoding: in legibility, the robot uses a partial trajectory or action sequence to indicate a single goal state, whereas in t -predictability the robot uses a partial action sequence to indicate *another* action sequence. Therefore, one entails a mapping from a large space to a small set of possibilities (the finite candidate goal states), whereas the other entails a mapping between spaces of equivalent size.

The distinction between task-level legibility and t -predictability is crucially important, particularly in collaborative settings. If you are cooking dinner with your household robot, it is important for the robot to act legibly so you can infer *what* goal it has when it grabs a knife (e.g., to slice vegetables). But, it is equally important for the robot to act in a t -predictable manner so that you can predict *how* it will accomplish that goal (e.g., the order in which it will cut the vegetables).

Relation to Explicability. Explicability [23] has been recently introduced to measure whether the observer could assign labels to a plan. In this context, explicability would measure the existence of any remainder of a plan that achieves the goal, as opposed to optimizing the probability that the observer will infer the robot’s plan.

Boltzmann Noisy Rationality. Computing t -predictability entails computing the conditional probability from (1). We model the human as expecting the robot to be noisily optimal, taking approximately the optimal sequence of actions to achieve G . Boltzmann probabilistic models of such noisy optimality (also known as the Luce-Shepard choice rule in cognitive science) have been used in the context of *goal* inference through inverse action planning [2]. We adopt an analogous model for modeling the inference of *task plans*.

We define optimality via some cost function $c : \mathcal{A} \times \mathcal{S} \times \mathcal{G} \rightarrow \mathbb{R}^+$, mapping each feasible plan, from a starting state and for a particular goal, to a scalar cost. In our experiment, for instance, we use path length (travel distance) for c . Applying a Boltzmann policy [2] based on c , we get:

$$P(\mathbf{a}|S, G) = \frac{e^{-\beta c(\mathbf{a}, S, G)}}{\sum_{\tilde{\mathbf{a}} \in \mathcal{A}} e^{-\beta c(\tilde{\mathbf{a}}, S, G)}}. \quad (2)$$

Here $\beta > 0$ is termed the *rationality coefficient*. As $\beta \rightarrow \infty$ the probability distribution converges to one for the optimal sequence and zero elsewhere; that is, the human models the agent as rational. As $\beta \rightarrow 0$, the probability distribution becomes uniform over all possible sequences \mathbf{a} and the human models the agent as indifferent.

t -Predictability Optimization. We make the assumption that cost is linearly separable, i.e. $c(\mathbf{a}, S, G) = \sum c(a_t, S, G)$. Incorporating (2), (1) becomes:

$$\mathbf{a}^* = \arg \max_{\mathbf{a} \in \mathcal{A}} \frac{\exp(-\beta c(\mathbf{a}_{t+1:T}, S_{\mathbf{a}}^t, G))}{\sum_{\tilde{\mathbf{a}}_{t+1:T} \in \mathcal{A}_{\mathbf{a}}^t} \exp(-\beta c(\tilde{\mathbf{a}}_{t+1:T}, S_{\mathbf{a}}^t, G))}, \quad (3)$$

with $S_{\mathbf{a}}^t$ denoting the state reached by executing the first t steps of plan \mathbf{a} , and $\mathcal{A}_{\mathbf{a}}^t$ denoting the set of all feasible plans that achieve G from state $S_{\mathbf{a}}^t$ in $T - t$ steps or less.

Approximate Algorithm for Large-Scale Optimization. The challenge with the optimization in (3) is the denominator: it requires summing over all possible plan remainders. Motivated by the fact that plans with higher costs contribute exponentially less to the sum, we propose to approximate the denominator by only summing over the lowest-cost l plan remainders.

Many tasks have the structure of Traveling Salesman Problems (TSP), where there are a number of subgoals whose order is not important but influences total cost. Van der Poort et al. [15] showed how to efficiently compute the l best solutions to the standard

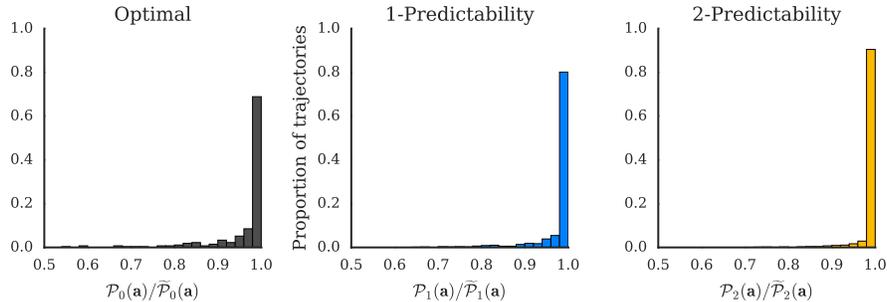


Fig. 3. Approximate t -predictability. Each subplot shows a histogram of ratios between exact t -predictability (\mathcal{P}_t) and approximate t -predictability ($\tilde{\mathcal{P}}_t$) of all possible sequences across 270 unique layouts. The subplots show histograms of ratios for 0-, 1-, and 2-predictability, respectively. In the majority of cases, the exact and approximate t -predictabilities are nearly identical.

(cyclic tour) TSP using a branch-and-bound algorithm. The key mechanism is successively dividing the set of feasible plans into smaller subsets for which a lower bound on the cost can be computed by some heuristic. When a subset of solutions has a lower bound higher than the smallest l costs evaluated so far, it is discarded entirely, while the remaining subsets continue to be broken up. The process continues until only l feasible plans remain. This method is guaranteed to produce the l solutions with the least cost and can significantly reduce time complexity over exhaustive enumeration. In particular, it has been shown that for the standard TSP, computation is in $O(l(T-t)^3 2^{T-t})$ [15], while exhaustive enumeration requires computation in $O((T-t)!)$. While heuristics are domain-specific, we expect that this method can be widely applicable to robot task planning problems. Further, we expect $T-t$ to be a small number in realistic applications, limited by people’s ability to reason about long sequences of actions.

To empirically evaluate the consequences of this approximation of t -predictability, we computed the exact and approximate (using $l = 2$) t -predictability for all possible plans in 270 randomly generated unique scenes (Fig. 3). If we choose the maximally t -predictable sequences for each scene using both the exact and approximate calculations of t -predictability, we find that these sequences agree in 242 (out of 270) scenes for 1-predictability and in 263 for 2-predictability⁴. For the sequences that disagree, the exact t -predictability of the sequence chosen using the approximate method is 89.5% of the optimal t -predictability in the worst case, and 99% of the optimal t -predictability on average. This shows that the proposed approximation is highly effective at producing t -predictable plans.

⁴ For 0-predictability, the denominator is the same in all plans, so all 270 scenes agree trivially.

3 Online Experiment

We set up an experiment to test that our t -predictable planner is in fact t -predictable. We designed a web-based virtual human-robot collaboration experiment in a TSP setting, where the human had to predict the behavior of three robot avatars using different planners. Participants watched the robots move to a number of targets (either zero, one, or two) and had to predict the sequence of remaining targets the robot would complete.

3.1 Independent Variables

We manipulated two variables: the t -predictable planner (for $t \in \{0, 1, 2\}$) and the number of observed targets k (for $k \in \{0, 1, 2\}$).

Planner. We used three different planners which differed in their optimization criterion: the number t of targets assumed known to the observer. Each participant interacted with three robot avatars, each using one of the following three planners:

Optimal (0-predictable): This robot chooses the shortest path from the initial location visiting all target locations once; that is, the “traditional” solution to the open TSP. This robot solves (3) for $t = 0$.

1-predictable: This robot solves (3) for $t = 1$; the sequence might make an inefficient choice for the first target in order to make the sequence of remaining targets clear.

2-predictable: This robot solves (3) for $t = 2$; the sequence might make an inefficient choice for the first *two* targets in order to make the sequence of remaining targets clear.

Number of observed targets. Each subject was shown the first $k \in \{0, 1, 2\}$ targets of the robot’s chosen sequence in each trial and was asked to predict the remainder of the sequence. This variable was manipulated between participants; thus, a given participant always saw the same number of targets (k) on all trials.

3.2 Procedure

The experiment was divided into two phases: a training phase to familiarize participants with TSPs and how to solve them, and an experimental phase. We additionally asked participants to fill out a survey at the end of the experiment.

In the training phase, subjects controlled a human avatar. They were instructed to click on targets in the order that they believed would result in the quickest path for the human avatar to visit all of them. The human avatar moved in a straight line after each click and “captured” the selected target, which was then removed from the display.

For the second phase of the experiment, participants saw a robot avatar move to either $k = 0$, $k = 1$, or $k = 2$ targets. After moving to these targets, the robot paused so that participants could predict the remaining sequence of targets by clicking on the targets in the order in which they believed the robot would complete them. Afterwards, participants were presented with an animation showing the robot moving to the rest of the targets in the sequence determined by the corresponding planner.

Stimuli. Each target layout displayed a square domain with five or six targets. There were a total of 60 trials, consisting of four repetitions of 15 unique target layouts in random order: 1 for the training phase, and the 3 experimental conditions. The trials

were grouped so that each participant observed the same robot for three trials in a row before switching to a different robot. In the training trials, the avatar was a gender-neutral cartoon of a person on a scooter, and the robot avatars were images of the same robot in different poses and colors (either red, blue, or yellow).

Layout Generation. The layouts for the 15 trials were based from an initial database of 270 randomly generated layouts. This number was reduced down to 176 in which the optimal sequence was different between all three planners so that the stimuli were distinguishable. We also discarded some scenarios in which the robot’s trajectory approached a target without capturing it, to avoid confounds. Out of these valid layouts, we chose the ones with the highest theoretical gain in 1-predictability to 2-predictability, to avoid scenarios where the information gain was marginal.

Attention Checks. After reading the instructions, participants were given an attention check in the form of two questions asking them the color of the targets and the color of the robot that they *would not* be evaluating.

Controlling for Confounds. We controlled for confounds by counterbalancing the colors of the robots for each planner; by using a human avatar in the practice trials; by randomizing the trial order; and by including attention checks.

3.3 Dependent Measures

Objective measures. We recorded the proportion of correct predictions of the robot’s sequence of targets out of all 15 trials for each planner, resulting in a measure of *error rate*. We additionally computed the *Levenshtein distance* between predicted and actual sequences of targets. This is a more precise measure of how similar participants’ predictions were to the actual sequences produced by the planner.

Subjective measures. After every ninth trial of the experiment, we asked participants to indicate which robot they preferred working with. At the end the experiment, each participant was also asked to complete a questionnaire to evaluate their perceived performance of three robots. An informal analysis of this questionnaire suggested similar results as those obtained from our other measures (see Section 3.6). Thus, because of space constraints, we have omitted specifics of the survey in this paper.

3.4 Hypotheses

H1 - Comparison with Optimal. *When showing 1 target, the 1-predictable robot will result in lower error than the optimal baseline. When showing 2 targets, the 2-predictable robot will result in lower error than the optimal baseline.*

H2 - Generalization. *The error rate will be lowest when $t = k$: the number of targets shown, k , equals the number of targets assumed by the t -predictable planner, t .*

H3 - Preference. *The perceived performance of the robots will be highest when $t = k$.*

3.5 Participants

We recruited a total of 242 participants from Amazon’s Mechanical Turk using the psiTurk experimental framework [10]. We excluded 42 participants from analysis for

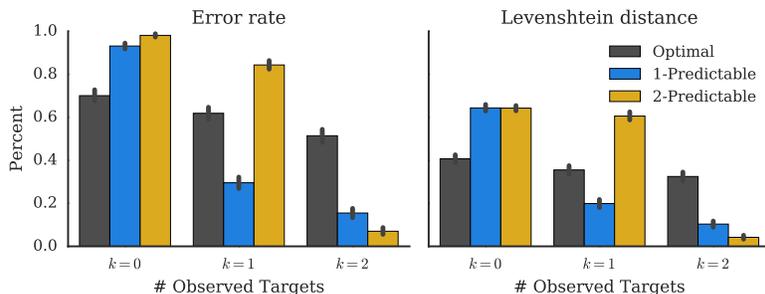


Fig. 4. Error rate and edit distance. The left subplot shows the average proportion of incorrect predictions for different numbers of observed targets (k) of sequences generated by different t -predictable planners. The right subplot shows the average Levenshtein distance between predicted sequences and actual sequences. In both plots, the lowest error rates occur when $k = t$.

failing the attention checks, leaving a total of $N = 200$ participants whose data we used. All participants were treated in accordance with local IRB standards and were paid \$1.80 USD for an average of 22 minutes of work, plus an average bonus of \$0.47.⁵

3.6 Results

Model validity. We first looked at the validity of our model of t -predictability with respect to people’s performance in the experiment. We computed the k -predictability scores (probability of correctly predicting the robot’s sequence from the k targets the user saw) for each task layout under each planner and number of targets the users observed. We also computed people’s actual prediction accuracy on each of these layouts under each condition, averaged across participants.

We computed the Pearson correlation between k -predictability scores and participant accuracy, finding a correlation of $r = 0.87$, 95% CI [0.81, 0.91]; the confidence interval around the median was computed using 10,000 bootstrap samples (with replacement). ***This high correlation suggests that our model of how people predict action sequences of other agents is a good predictor of their actual behavior.***

Accuracy. To determine how similar people’s predictions of the robots’ sequences were to the actual sequences, we used two objective measures of accuracy: first, overall error rate (whether they predicted the correct sequence or not), as well as the Levenshtein distance between the predicted and correct sequences (Fig. 4).

⁵ Bonuses could range from \$0.00 to \$1.35 depending on performance. In the practice trials, participants could get a \$0.03 bonus on each trial if they chose the shortest possible sequence; a \$0.02 bonus if they chose a sequence within the top 5% shortest sequences; or a \$0.01 bonus if they chose a sequence within the top 10% shortest sequences (relative to all possible sequences). They received no bonus if they chose a sequence that was longer than the top 10% shortest sequences. In the experimental trials, participants could get a \$0.02 bonus on each trial if they correctly predicted the robot’s sequence of targets on that trial.

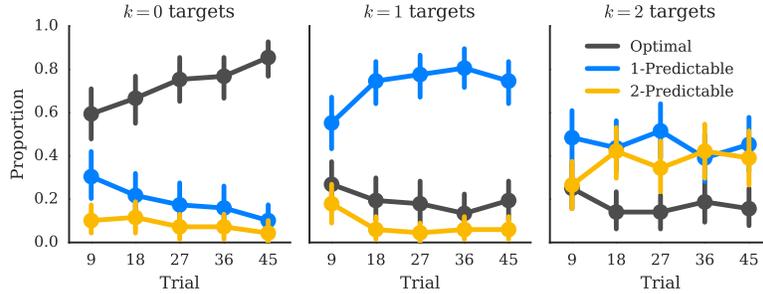


Fig. 5. Preferences over time. Participants prefer the 0-predictable (optimal) robot for $k = 0$ and the 1-predictable robot for $k = 1$, as well as $k = 2$: despite performing better with it, they are frustrated by how inefficient its first 2 actions are.

As the two measures have qualitatively similar patterns of results, and the Levenshtein distance is a more fine-grained measure of accuracy, we performed quantitative analysis only on the Levenshtein distance. We constructed a linear mixed-effects model with the number of observed targets k (k from 0 to 2) and the planner for t -predictability (t from 0 to 2) as fixed effects, and trial layout as random effects.

This model revealed significant main effects of the number of observed targets ($F(2, 10299) = 1894.75, p < 0.001$) and planner ($F(2, 42) = 6.59, p < 0.01$) as well as an interaction between the two ($F(4, 10299) = 554.00, p < 0.001$). We ran post-hoc comparisons using the multivariate t adjustment. Comparing the planners across the same number of targets, we found that in the 0-targets condition the optimal (or 0-predictable) robot was better than the other two robots; in the 1-target condition, the 1-predictable robot was better than the other two; in the 2-target prediction, the 2-predictable robot was better than the optimal and 1-predictable robots. All differences were significant with $p < 0.001$ except the difference between the 2-predictable robot and the 1-predictable robot in the 2-target condition ($t(50) = 1.85, p = 0.56$). Comparing the performance of a planner across number of targets, we found significant differences in all contrasts, with one exception: the accuracy when using the optimal planner was not significantly different when seeing 1 target vs 2 targets ($t(10299) = 2.65, p = 0.11$). Overall, these results support our hypotheses **H1** and **H2**, that *accuracy is highest when t used in the planner equals k , the number of observed targets*.

Preferences over time. Fig. 5 shows the proportion of participants choosing each robot planner at each trial. We constructed a logistic mixed-effects model for binary preferences (where 1 meant the robot was chosen) with planner, number of observed targets, and trial as fixed effects and participants as random effects. The planner and number of observed targets were categorical variables, while trial was a numeric variable.

Using Wald’s tests, we found a significant main effect of the planner ($\chi^2(2) = 13.66, p < 0.01$) and trial ($\chi^2(1) = 9.30, p < 0.01$). We detected only a marginal effect of number of targets ($\chi^2(2) = 4.67, p = 0.10$). However, there was a significant

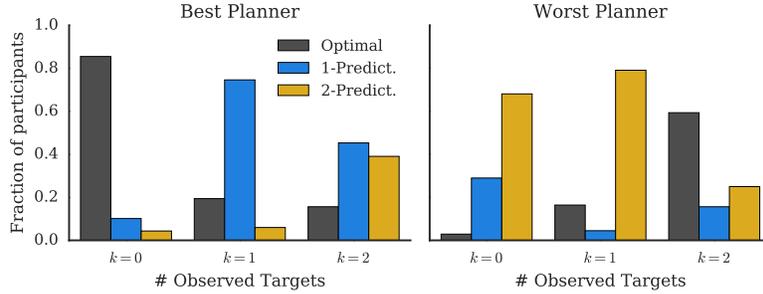


Fig. 6. Final rankings. Participants ranked the planners differently depending on how many targets were observed. For $k = 0$, people preferred the optimal planner; for $k = 1$ and $k = 2$, they preferred the 1-predictable planner.

interaction between planner and number of targets ($\chi^2(4) = 20.26, p < 0.001$). We also found interactions between planner and trial ($\chi^2(2) = 24.68, p < 0.001$) and between number of targets and trial ($\chi^2(2) = 16.07, p < 0.001$), as well as a three-way interaction ($\chi^2(4) = 39.43, p < 0.001$). Post-hoc comparisons with the multivariate t adjustment for p -values indicated that for the 0-targets condition, the optimal robot was preferred over the 1-predictable robot ($z = -13.22, p < 0.001$) and the 2-predictable robot ($z = -14.56, p < 0.001$). For the 1-target condition, the 1-predictable robot was preferred over the optimal robot ($z = 12.97, p < 0.001$) and the 2-predictable robot ($z = 14.00, p < 0.001$). In the two-task condition, we did not detect a difference between the two 1-predictable and 2-predictable robots ($z = 2.26, p = 0.29$), though both were preferred over the optimal robot ($z = 7.44, p < 0.001$ for the 1-predictable robot and $z = 5.40, p < 0.001$ for the 2-predictable robot).

Overall, these results are in line with our hypothesis **H3** that *the perceived performance is highest when t used in the planner equals k , the number of observed targets*. This is the case for $k = 0$ and $k = 1$, but not $k = 2$: even though users tended to perform better with the 2-predictable robot, its suboptimal actions in the beginning frustrated the users, e.g. “*This robot mostly starts out in the worst way and then goes in weird directions but eventually starts to make sense.*”

Final rankings. The final rankings of “best robot” and “worst robot” are shown in Fig. 6. For each participant, we assigned each robot a score based on their final rankings. The best robot received a score of 1; the worst robot received a score of 2; and the remaining robot received a score of 1.5. We constructed a logistic mixed-effects model for these scores, with planner and number of observed targets as fixed effects, and participants as random effects; we then used Walds tests to check for effects.

We found significant main effects of planner ($\chi^2(2) = 41.38, p < 0.001$) and number of targets ($\chi^2(2) = 12.97, p < 0.01$), as well as an interaction between them ($\chi^2(4) = 88.52, p < 0.001$). We again performed post-hoc comparisons using the multivariate t adjustment. These comparisons indicated that in the 1-target condition,

people preferred the optimal robot over the 1-predictable robot ($z = 3.46$, $p < 0.01$) and the 2-predictable robot ($z = 5.60$, $p < 0.001$). In the 1-target condition, there was a preference for the 1-predictable robot over the optimal robot, however this difference was not significant ($z = -2.18$, $p = 0.27$). The 1-predictable robot was preferred to the 2-predictable robot ($z = -6.54$, $p < 0.001$). In the 2-target condition, both the 1-predictable and 2-predictable robots were preferred over the optimal robot ($z = -4.85$, $p < 0.001$ for the 1-predictable robot, and $z = -3.85$, $p < 0.01$ for the 2-predictable robot), though we did not detect a difference between the 1-predictable and 2-predictable robots themselves ($z = -1.33$, $p = 0.84$). Overall, these rankings are in line with the preferences over time.

Summary. Our t -predictability planner worked as expected, with the t -predictable robots leading to the highest user prediction accuracy given the first t targets. However, focusing on just 2-predictability at the expense of 0-predictability frustrated our users. Overall, we believe t -predictability will be important in a task for all t s, and hypothesize that optimizing for a weighted combination will perform best in practice.

We note that β is problem-specific and can be expected to decay as the difficulty of the task increases; in each setting, it can be estimated from participant data. Although β was chosen ahead of time in our experiment to be $\beta = 1$, our results are validated by the $r = 0.87$ correlation between expected and observed human error rates.

The optimal choice of t is also a subject for further investigation and is likely context-specific. Depending on the particular task, there should be a different trade-off between predictability of later actions and that of earlier actions.

4 User Study

Having tested our ability to produce t -predictable sequences, we next ran an in person study to test their implications. Participants used a smartphone to operate a remote-controlled Sphero BB-8 robot, and had to predict and adapt to the actions of an autonomous Pioneer 3-DX robot in a collaboration scenario (Fig. 1).

4.1 Independent Variables

We manipulated one single variable, *predictability*, as a within-subjects factor. Having confirmed the expected effects of the different planners in the previous experiment, and given the good overall performance of the 1-predictable planner across different conditions, we decided to omit the 2-predictable agent and focus on testing the implications of 1-predictable with respect to optimal in a more immersive collaborative context.

4.2 Procedure

At the beginning of the experiment, participants were told that they and their two robot friends were on a secret mission to deactivate an artifact. In each of 4 trials, the autonomous 3-DX navigated to the 5 power sources and deactivated them in sequence; however, security sensors activated at each power source after 3 or more had been powered down. The subject’s mission was to use BB-8 to jam the sensors at the third, fourth

and fifth power sources before the P3DX arrived at them, by steering BB-8 into the corresponding sensor for a short period of time.

After an initial practice phase in which participants had a chance to familiarize themselves with the objective and rules of the task, as well as with the BB-8 teleoperation interface, there were two blocks of 4 trials whose order was counterbalanced across participants. In each block, the subject collaborated with the P3DX under a different task planner which we referred to as different robot “personalities”.

Stimuli. Each of the 5 power sources (targets) in each trial was projected onto the floor as a yellow circle, using an overhead projector (Fig 1). Each circle was initially surrounded by a projected blue ring representing a dormant sensor. When the P3DX reached a target, both the circle and the ring were eliminated, except When the P3DX reached the third target, in which case the blue circles turned red symbolizing their switch into active state. Whenever BB-8 entered a ring, the ring turned green for 2 seconds and then disappeared, indicating successful jamming. If the P3DX moved over a red ring, a large red rectangle was projected, symbolizing capture and the trial ended in failure. Conversely, if the P3DX completed all 5 targets without entering a red ring, a green rectangle indicated successful completion of the trial.

Layout Generation. The 4 layouts used were taken from the larger pool of 15 layouts in the online experiment. There was a balance between layouts where online participants had been more accurate with the optimal planner, more accurate with the 1-predictable planner, or similarly accurate.

Controlling for Confounds. We controlled for confounds by counterbalancing the order of the planners; by using a practice layout; and by randomizing the trial order.

4.3 Dependent Measures

Objective measures. We recorded the number of successful trials for each subject and robot planner, as well as the number of trials where participants jammed targets in the correct sequence.

Subjective measures. After every block of the experiment, each participant was also asked to complete a questionnaire (adapted from [6]) to evaluate their perceived performance of the P3DX robot. At the end of the experiment, we asked participants to indicate which robot (planner) they preferred working with.

4.4 Hypotheses

H4 - Comparison with Optimal. *The 1-predictable robot will result in more successful trials than the optimal baseline.*

H5 - Preference. *Users will prefer working with the 1-predictable robot.*

4.5 Participants

We recruited a total of 14 participants from a university community. Participants were treated in accordance with local IRB standards and were paid \$10 USD. The study took about 30 minutes on average.

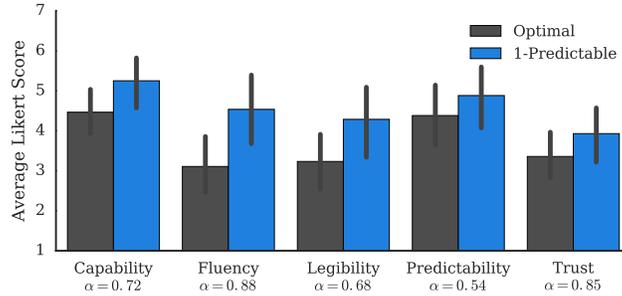


Fig. 7. Perceptions of the collaboration. Over all measures, participants ranked the 1-predictable planner as being preferable to the optimal planner.

4.6 Results

Successful completions. We first looked at how often participants were able to complete the task with each robot. We constructed a logistic mixed-effects model for completion success with predictability as a fixed effect and participant and task layout as random effects. We found a significant effect of predictability ($\chi^2(1) = 11.17, p < 0.001$), with the 1-predictable robot yielding more successful completions than the optimal robot ($z = 3.34, p < 0.001$). This supports **H4**.

Prediction accuracy. We also looked at how accurate participants were at predicting the robots' sequence of tasks, based on the order in which participants jammed tasks. We constructed a logistic mixed-effects model for prediction accuracy with predictability as a fixed effect and participant and task layout as random effects. We found a significant effect of predictability ($\chi^2(1) = 9.49, p < 0.01$), with the 1-predictable robot being more predictable than the optimal robot ($z = 3.08, p < 0.01$).

Robot preferences. We asked participants to pick the robot they preferred to collaborate with. We found that 86% ($N = 12$) of participants preferred the predictable robot, while the rest ($N = 2$) preferred the optimal robot. This result is significantly different from chance ($\chi^2(1) = 7.14, p < 0.01$). This supports **H5**.

Perceptions of the collaboration. We analyzed participants' perceptions of the robots' behavior (Fig. 7) by averaging each participant's responses to the individual questions for each robot and measure, resulting in a single score per participant, per measure, per robot. We constructed a linear mixed-effects model for the survey responses with predictability and measure type as fixed effects, and with participants as random effects. We found a main effect of predictability ($F(1, 117) = 16.42, p < 0.001$) and measure ($F(4, 117) = 5.45, p < 0.001$). Post-hoc comparisons using the multivariate t method for p -value adjustment indicated that participants preferred the predictable robot over the optimal robot ($t(117) = 4.05, p < 0.001$) by an average of 0.87 ± 0.21 SE points on the Likert scale.

5 Discussion

In what remains, we summarize our work and discuss future directions, including the application of t -predictability beyond task planning.

Summary. We enable robots to generate t -predictable plans, for which a human observing the first t actions can confidently infer the rest. We tested the ability to make plans t -predictable in a large-scale online experiment, in which subjects’ predictions of the robot’s action sequence significantly improved. In an in-person study, we found that t -predictability can lead to significant objective and perceived improvements in human-robot collaboration compared to traditional optimal planning.

t -predictability for Motion. Even though t -predictability is motivated by a task planning need, it does have an equivalent in motion planning: find an initial trajectory $\xi_{0:t}$, such that the remainder $\xi_{t:T}$ can be inferred by a human observer with knowledge of both the start state $\xi_0 = S$ and the goal state $\xi_T = G$. Modeling this conditional probability with a Boltzmann model yields

$$\xi^* = \arg \max_{\xi \in \Xi} P(\xi_{t:T}|G, \xi_{0:t}) = \arg \max_{\xi \in \Xi} \frac{e^{-\beta c(\xi_{t:T})}}{\int e^{-\beta c(\hat{\xi}_{t:T})} d\hat{\xi}_{t:T}},$$

where Ξ is the set of feasible trajectories from S to G . Using a second order expansion of $c(\hat{\xi}_{t:T})$ about the optimal remaining trajectory we get:

$$\begin{aligned} \max_{\xi} P(\xi_{t:T}|G, \xi_{0:t}) &\approx \max_{\xi_{0:t}} \frac{e^{-\beta c(\xi_{t:T}^*)}}{e^{-\beta c(\xi_{t:T}^*)} \int e^{-\frac{1}{2}(\hat{\xi}_{t:T} - \xi_{t:T}^*)^T H(\xi_{t:T}^*)(\hat{\xi}_{t:T} - \xi_{t:T}^*)} d\hat{\xi}_{t:T}} \\ &\equiv \max_{\xi_t} \det(H(\xi_{t:T}^*)), \end{aligned} \quad (4)$$

where H is the Hessian. This implies that generating a t -predictable trajectory means finding a configuration for time t such that the optimal trajectory from that configuration to the goal is in a *steep*, high-curvature minimum: other trajectories from ξ_t to the goal would be significantly higher cost. For instance, if the robot had the option between two passages, it would choose the more *narrow* passage because that enables the observer to more accurately predict the remainder of its trajectory.

Limitations and Future Work. Our work is limited by the focus in our experiments on TSP scenarios (though we emphasize that t -predictability as it is formulated in Section 2 is not inherently limited to TSPs). This work is also limited by the choice of a user study that involved a tele-operated avatar to mediate the human’s physical collaboration with the robot. Applications to other scenarios that involve direct physical collaboration and preconditions would be interesting topics to investigate. Additionally, while our work showcases the utility of t -predictability, a main remaining challenge is determining what t or combination of ts to use for arbitrary tasks. This decision requires more sophisticated human behavior models, which are the topic of our ongoing work.

References

- [1] R Alami et al. “Safe and Dependable Physical Human-Robot Interaction in Anthropic Domains : State of the Art and Challenges”. *Society* 6.1 (2006).

- [2] C. L. Baker, R. Saxe, and J. B. Tenenbaum. “Action understanding as inverse planning”. *Cognition* 113.3 (2009).
- [3] E. Charniak and R. P. Goldman. “A Bayesian Model of Plan Recognition”. *Artif. Intell.* 64.1 (1993).
- [4] M. B. Dias et al. “Sliding autonomy for peer-to-peer human-robot teams”. *10th Intelligent Conference on Intelligent Autonomous Systems (IAS 2008), Germany*. 2008.
- [5] A. Dragan and S. Srinivasa. “Integrating human observer inferences into robot motion planning”. *Autonomous Robots* (2014).
- [6] A. Dragan et al. “Effects of Robot Motion on Human-Robot Collaboration”. *ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (2015).
- [7] A. D. Dragan, K. C. T. Lee, and S. S. Srinivasa. “Legibility and predictability of robot motion”. *ACM/IEEE International Conference on Human-Robot Interaction* (2013).
- [8] T. Fong et al. “The peer-to-peer human-robot interaction project”. *AIAA Space 2005* (2005).
- [9] M. J. Gielniak and A. L. Thomaz. “Generating anticipation in robot motion”. *2011 RO-MAN*. IEEE. 2011.
- [10] T. M. Gureckis et al. “psiTurk: An open-source framework for conducting replicable behavioral experiments online”. *Behavioral Research Methods* (2015).
- [11] C. Liu et al. “Goal Inference Improves Objective and Perceived Performance in Human-Robot Collaboration”. *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*. International Foundation for Autonomous Agents and Multiagent Systems. 2016, pp. 940–948.
- [12] D. T. McRuer. “Pilot-Induced Oscillations and Human Dynamic Behavior” (1995).
- [13] S. Nikolaidis and J. Shah. “Human-robot cross-training: Computational formulation, modeling and evaluation of a human team training strategy”. *International Conference on Human-Robot Interaction (HRI)* (2013).
- [14] G. Pezzulo, F. Donnarumma, and H. Dindo. “Human sensorimotor communication: A theory of signaling in online social interactions”. *PLoS One* 8.11 (2013).
- [15] E. S. Van der Poort et al. “Solving the k-best traveling salesman problem”. *Computers & operations research* 26.4 (1999).
- [16] M. Saffarian, J. C. F. de Winter, and R. Happee. “Automated Driving: Human-Factors Issues and Design Solutions”. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 56.1 (2012).
- [17] N. B. Sarter and D. D. Woods. “Team play with a powerful and independent agent: operational experiences and automation surprises on the Airbus A-320.” *Human factors* 39.4 (1997).
- [18] J. Shah and C. Breazeal. “An empirical analysis of team coordination behaviors and action planning with application to human-robot teaming.” *Human factors* 52.2 (2010).
- [19] D. Szafir, B. Mutlu, and T. Fong. “Communication of Intent in Assistive Free Flyers”. *ACM/IEEE International Conference on Human-robot Interaction* (2014).
- [20] L. Takayama, D. Dooley, and W. Ju. “Expressing thought: improving robot readability with animation principles”. *ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (2011).
- [21] M. Tomasello et al. “Understanding and sharing intentions: the origins of cultural cognition”. *Behavioral and brain sciences* 28.05 (2005), pp. 675–691.
- [22] C. Vesper et al. “A minimal architecture for joint action”. *Neural Networks* 23.8 (2010).
- [23] Y. Zhang et al. “Plan Explicability for Robot Task Planning”. *RSS Workshop on Planning for Human-Robot Interaction: Shared Autonomy and Collaborative Robotics* (2016).