Modeling Human Helpfulness with Individual and Contextual Factors for Robot Planning

Amal Nanavati*, Christoforos Mavrogiannis*, Kevin Weatherwax†, Leila Takayama†, Maya Cakmak* and Siddhartha S. Srinivasa*

*University of Washington, Seattle, Washington 98195
Emails: {amaln, cmavro, mccakmak, siddh}@cs.washington.edu
†University of California, Santa Cruz, California 95064
Emails: {wax, takayama}@ucsc.edu

Abstract—Robots deployed in human-populated spaces often need human help to effectively complete their tasks. Yet, a robot that asks for help too frequently or at the wrong times may cause annoyance, and a robot that asks too infrequently may be unable to complete its tasks. In this paper, we present a model of humans’ helpfulness towards a robot in an office environment, learnt from online user study data. Our key insight is that effectively planning for a task that involves bystander help requires disaggregating individual and contextual factors and explicitly reasoning about uncertainty over individual factors. Our model incorporates the individual factor of latent helpfulness and the contextual factors of human busyness and robot frequency of asking. We integrate the model into a Bayes-Adaptive Markov Decision Process (BAMDP) framework and run a user study that compares it to baseline models that do not incorporate individual or contextual factors. The results show that our model significantly outperforms baseline models by a factor of 1.5X, and it does so by asking for help more effectively: asking 1.2X times less while still receiving more human help on average.

I. INTRODUCTION

As robots are increasingly deployed in dynamic and uncertain human environments, situations will arise that they are not fully equipped to handle. Consider the scenario of a robot performing tasks in an office building, like delivering mail or escorting someone to an appointment. Due to limitations in hardware and computation, environmental uncertainties, or a lack of domain knowledge, the robot might get lost, get knocked over, or be asked to go to a location it is unfamiliar with. One way to address these challenges is for the robot to ask for human help [38, 20, 35, 42]. However, the decision of who and when to ask for help is nuanced. A robot that asks the same user for help too frequently may annoy them to the point that they stop helping [20]. A robot that repeatedly interrupts the user at inappropriate times—such as when they are engaged in another primary task—may annoy them to the point that they want to physically harm the robot [34]. On the other hand, a robot that does not ask for help often enough may take too long to complete its tasks, or fail.

In this paper, we address the following research question: How can a robot model human helpfulness, in order to balance the dual objectives of efficiently completing its tasks and minimizing the number of times it asks for help? We consider two types of factors: individual factors, which are unchangeable (and often unobservable) characteristics of a human, and contextual factors, which can change over time and embed the human within their broader physical and temporal context. This work builds on two threads of past work. The state-of-the-art approach to modeling human helpfulness [38] attributes all variance in human helpfulness to the individual factor of “availability,” and does not account for contextual factors. Further, their associated planning framework does not reason about the robot’s uncertainty over the individual factor, thereby preventing the robot from engaging in information-gathering behaviors to intentionally learn the user’s “availability” during real-time execution. On the other hand, the state-of-the-art approach to modeling human interruptability [3] accounts for a rich array of contextual factors—the human’s body pose at that time, audio data, and objects in the scene—but does not account for individual factors that may be learnt through repeated interactions with the same person. Our key insight is that effectively planning for a task that involves bystander help requires disaggregating individual and contextual factors and explicitly reasoning about uncertainty over individual factors.

In this work, we develop a predictive model of human helping behaviors (Sec. III), learned from data collected in a
1. The robot needs help finding a room.

2. The human help model, which is updated based on past interactions with each human, allows it to predict the probability a person will help.

3. For each human the robot encounters, it uses the human help model to decide whether to ask them for help.

Fig. 2: Our human help model, used by the robot when it needs help finding a room, predicts the probability of a human (with user ID UID) helping the robot. The model takes in the human’s latent helpfulness \( h_{UID} \) (learned from past interactions with that human), the human’s current busyness \( b \), and the robot’s past frequency of asking the human for help \( \rho_{ask} \). The robot integrates the human help model into its planning framework to decide whether to ask the humans it encounters for help.

virtual office environment (Fig. [1]) and integrate the model into a Bayes-Adaptive Markov Decision Process (BAMDP) planning framework (Sec. [V]). Our model and planning framework incorporates an individual factor (the human’s latent helpfulness) and contextual factors (the human’s busyness and the robot’s past frequency of asking for help). To understand the importance of both types of factors, consider the example in Fig. [2]. When the robot encounters Human 1, it has learned from past interactions that the human has relatively high latent helpfulness \( h_{UID} \), and therefore asks for help. When the robot encounters Human 2, it has learnt in the past that they have relatively low latent helpfulness, so does not ask. With Human 3, the robot has high uncertainty in its belief about their latent helpfulness. When it sees Human 1 again, the robot’s belief over their latent helpfulness has shifted positively because they helped last time; despite that, because the human is busy, the robot decides not to ask. These behaviors are the direct result of disaggregating individual and contextual factors and explicitly reasoning about the robot’s uncertainty over individual factors.

We evaluate our model with two user studies, comparing the emergent behavior of a robot that uses our model to the emergent behavior of one that uses baseline models that do not have individual or contextual factors (Sec. [V]). The results show that our model significantly outperforms both baselines by a factor of 1.5X, and that it does so by asking for help more effectively: asking 1.2X times less while still receiving more help on average. This paper makes two key contributions:

1. A model of human helpfulness that integrates individual and contextual factors, as well as the associated methodology that can be extended to other models of helpfulness with both types of factors.
2. A BAMDP planning framework that integrates the human help model, as well as an in-depth investigation into the robot behaviors that emerge from different human help models integrated into that framework.

II. RELATED WORK

A. Autonomously Generating Help-Seeking Behaviors

Several past works have focused on robots autonomously generating help-seeking behavior, some of which are in this survey on failures in human-robot interaction [25]. Many of these works assume the human is always available to provide help: navigating with an oracular teacher [35], learning from demonstrations [12], and active learning [10]. Works that do not make that assumption seek to model the uncertainty in human behavior. One work models how humans interpret natural language help requests, and then generates the requests likely to result in desired human actions [28]. Another work models the accuracy of humans given a visual input and uses that model to determine when and how to ask for help [2].

The work of Rosenthal [38] is most similar to ours; it also focuses on asking for help from bystanders in an office.
environment using POMDPs. Yet, their model attributes all variability in whether humans help to the \textit{individual} factor of “availability,” whereas our model decomposes “availability” into its constituent individual \textit{and} contextual factors. Their planning framework assumes the individual factor is fully observable. In contrast, our framework assumes it is unobservable, enabling the robot to engage in information-gathering actions to improve its belief over the individual factor. Notably, however, their user studies were conducted on a physical robot, whereas ours are conducted in a virtual office environment.

Human help is also related to interruptability, covered in this survey \cite{44}. A key interruptibility work in robotics uses features such as body pose, audio data, and nearby objects to predict how interruptable a human is \cite{2}. These are all \textit{contextual} features, because they change over time and are not particular to individuals. This focus on contextual factors extends to other interruptibility work. Of the 41 works whose model input features are surveyed in \cite{44}, only one incorporates fixed, individual factors, and that work’s focus is not the model \cite{44}. In contrast, our model incorporates both individual and contextual factors, enabling the robot to personalize its help-seeking behavior to the user and their context.

\subsection*{B. Human Responses to Help-Seeking Behavior}

Other works that investigate human helpfulness study human responses to help-seeking behavior. These works have revealed that humans are more likely to help robots that ask politely \cite{42}, provide justifications \cite{6}, indicate urgency \cite{9}, require less help \cite{38}, display emotion \cite{13}, offer people desired items \cite{3}, maintain appropriate interpersonal distance \cite{46}, establish smooth communication with the person \cite{47}, and are actively doing tasks for the person \cite{1}. Others showed that humans are less likely to help robots while they are doing a primary task \cite{26, 18} or if they have been exposed to robot failures \cite{33}. Yet other works found that microlinguistic factors (e.g., social atmosphere) influences whether humans help a robot \cite{17} and that help-seeking robots had a higher perceived usability \cite{30}. Finally, multiple works found that robots that ask for help too frequently, especially while the human is doing a task, can annoy their helpers \cite{20, 84, 29}. We use insights from these works to design our robot’s help queries—for example, by using negative politeness \cite{42} in the query—and to inform the factors we include in our model.

\subsection*{C. POMDPs in Human-Robot Interaction}

Several works in human-robot interaction have used Partially Observable Markov Decision Processes (POMDPs) and their derivatives (e.g., MOMDPs, BAMDPs) to plan robot actions. Some of these works used pre-specified models of human behavior within their POMDP \cite{5, 47}. Others learnt their models from data \cite{11, 45, 36}. One such work gathered labels for the unobservable factor—human trust—from a user study and used \textit{supervised} techniques to learn the model \cite{11}. Another work pre-specified the values of the unobservable factor—human subgoals—and used \textit{unsupervised} techniques to infer that factor’s role in the model \cite{43}. Yet another work used unsupervised clustering to \textit{jointly} learn the values of the unobservable factor—human type—and that factor’s role in the model \cite{55}. Like the latter approach, we use an unsupervised technique to \textit{jointly} infer the values of the unobservable factor and its role in the model. However, we use an alternate technique, generalized linear mixed models, that enables our model to: 1) incorporate contextual factors; and 2) learn a continuous, not discrete, values for the unobservable factor.

\section*{III. Modeling Human Helpfulness}

We develop a model of humans’ helpfulness using data from an IRB-approved online user study. Our model predicts the human’s probability of helping the robot based on the human’s busyness, the robot’s past frequency of asking for help, and the human’s latent helpfulness. We selected human busyness and robot past frequency of asking because past works have indicated their importance \cite{38, 20, 34}, but have not modeled their impact on whether humans help. We decided on an online user study to adapt to the COVID-19 pandemic and to provide us additional control over potentially confounding variables.

\subsection*{A. Scenario}

Our scenario takes place in a virtual online office environment (Fig. \ref{fig:scenario}), where a human user and a robot are doing unrelated tasks that both benefit the office generally. Created in Phaser\textsuperscript{2} this 2D top-down environment contains 31 rooms—offices, conference rooms, lounges, and restrooms—based on an academic building. The user controls an IT administrator who goes between rooms performing routine computer maintenance tasks. Every task has an associated time limit to start the task, after which the user starts losing points. Periodically, the user is given a break, where they have no time limit and are asked to go to a lounge or restroom. The office building also contains a mail delivery robot that navigates between the rooms; users are told that the robot is new and still learning about the building. The robot periodically asks the user to lead it to its destination. When asked, the user can either ignore the robot, click “Can’t Help,” or click “Yes” and lead the robot until they click “Stop Following.” The robot moves to targets, including the human, using the A* algorithm, and follows the human 1 cell away. Users use the arrow keys to move, use the mouse to click buttons, and hold the spacebar for 10 seconds to perform computer maintenance tasks (not included in the time limit). Note that the time limit includes every action the user takes between the end of the previous task and the beginning of the next task, including time they contemplate helping the robot. Further note that the time limit is the same regardless of whether users completed their previous task early or late.

\subsection*{B. Experimental Design}

Our user study studies two independent variables, the human’s busyness \(b \in [0, 1]\) and the robot’s frequency of asking

\\[ \text{https://phaser.io/} \]
\( \rho_{ask} \in [0, 1] \), and one dependent variable, human\(_{helps} \):

\[
\rho_{ask} = \frac{b}{k} \quad \text{(1)}
\]

\[
b = \frac{\text{min\_time\_for\_task}}{\text{available\_time\_for\_task}} \sum_{t=T-k}^{T} \mathbb{1} [\text{did\_robot\_ask}_t] \quad \text{(2)}
\]

\[
human\_helps = \begin{cases} 
1 & \text{if the human helped accurately} \\
0 & \text{else} 
\end{cases} \quad \text{(3)}
\]

\( b \) is the ratio of the minimum time it would take the human to get to their goal and the time they have available. \( b = 0 \) corresponds to no time limit (i.e., free time) and \( b = 1 \) corresponds to no time to spare. \( \rho_{ask} \) is the frequency of the robot asking for help, over the last \( k \) times it saw the person (\( t \) in Eq. 2 indexes the discrete timesteps when the robot saw the human). We set \( k := 5 \), as an experimental design consideration to balance between covering \( \rho_{ask} \)’s range while having a manageable number of conditions in the study.

We use a \( 3 \times 5 \) study design, with a within-subjects factor of busyness \( b \) (3 levels: \( \frac{1}{3} \) or “high,” \( \frac{1}{6} \) or “medium,” and 0 or “free time”) and a between-subjects factor of frequency of asking \( \rho_{ask} \) (5 levels: 0.2, 0.4, 0.6, 0.8, 1.0). “High” busyness is barely enough time to complete the task, “medium” is barely enough time to help the robot complete the task, and “free time” has no time limit. The user completes 28 tasks, where the busyness follows a fixed, repeating order of “medium,” “free time,” “high,” etc. The task time limit is calculated as

\[
\text{player\_speed} \cdot \text{shortest\_path\_to\_goal} \quad \text{where the player speed is fixed.}
\]

The robot appears during a predefined 20 of those 28 tasks and either walks past the human or asks the human for help, depending on the frequency condition. To account for people’s varied navigation techniques, the robot does not appear at a specific time or location, but rather whenever the human is a fixed proportion of the distance to their goal. All factors other than busyness and frequency are controlled: the human and robot sequence of goals are fixed, the distance between consecutive human goals is within a small range, and the deviation required of humans to help the robot is also within a small range.

After reading the study description, the user does a tutorial that introduces the mechanics of the office environment: moving the character, finding rooms, interacting with the robot, etc. They then complete the study, after which they are taken to a survey that includes the NASA-TLX [22], RoSAS [8], open-ended questions, and demographic questions. Participants were recruited on Amazon Mechanical Turk, had an approval rate \( \geq 95\% \), were from predominantly English-speaking countries, and were compensated $5 for this 35 minute study.

C. Hypotheses

- **Hypothesis A.1**: The higher the busyness \( b \), the lower the probability of helping. This is based on past work that found that a robot repeatedly interacting with a human who is doing another task annoys the human [24].

  - **Hypothesis A.2**: The higher the frequency \( \rho_{ask} \), the lower the probability of helping. This is based on a survey that found that office workers were less willing to help a robot that asked for help more frequently [38].

D. Dataset

Our final dataset had \( n = 140 \) participants (mean age 37.6, 47 F, 91 M, 2 other), evenly balanced across the 5 frequency levels. This corresponds to 1260 requests for help, evenly balanced across the 3 busyness levels. This excludes participants who did not follow tutorial instructions. Participants helped the robot accurately 29.3% of the time, by leading the robot to its requested room. Participants clicked “can’t help” 21.3% of the time, and ignored the robot 45.3% of the time. This is on par with past works, which found that 44% of surveyed people [47] and 37% of office inhabitants [38] noticed but did not interact with a help-seeking robot. The other 41% of the time the user either inaccurately or mistakenly helped the robot (e.g., clicked “yes” and immediately “stop following”). Because inaccurate help occurred infrequently, we treat human helping as a binary—either they helped accurately or not.

E. The Human Help Model

We want a model-learning approach that does the following:

1. Takes in (user ID (UID), \( b \), \( \rho_{ask} \), human\(_{helps} \)) datapoints;
2. Infers humans’ unobservable latent helpfulness, \( h_{UID} \), a factor that explains individual variation in behavior;
3. Predicts users’ probability of helping given their latent helpfulness, busyness, and the robot frequency of asking.

To meet these requirements, we use generalized linear mixed models (GLMM) with a logistic link function. We build up the model using stepwise regression, a procedure that iteratively adds the most significant factor—main factors first and then interaction factors—until no factor significantly improves the model. The resulting model has user ID (UID) as a random effect, busyness as a main effect, and busyness and frequency as an interaction effect. Mathematically, the model is

\[
\mathbb{P}(human\_helps \mid h_{UID}, b, \rho_{ask}) = \frac{1}{1 + \exp (-f(h_{UID}, b, \rho_{ask}))}, \quad (4)
\]

\[
f(h_{UID}, b, \rho_{ask}) = h_{UID} + c_1 + c_2 \cdot b + c_3 \cdot b \cdot \rho_{ask}, \quad (5)
\]

\[
h_{UID} \sim \mathcal{N}(0, \sigma^2), \quad (6)
\]

where \( c_1, c_2, c_3, \) and \( \sigma^2 \) are learnt parameters.

Fig. 3 shows the human’s likelihood of helping the robot by the frequency of asking and busyness, both for the dataset and the fitted model (for the average human). Wald’s Test reveals that busyness significantly influenced whether the human helped (\( \chi^2(1) = 161.25, p < 0.001 \)), with users being less likely to help the robot when they are busier. This validates **Hypothesis A.1**. Frequency of asking does not have a significant main effect, but has a significant interaction with busyness (\( \chi^2(1) = 8.07, p = 0.005 \)). In particular, the robot’s frequency of asking does not significantly influence the probability of
helping during free time, but as the human gets busier, they are less likely to help a robot that asks more frequently. This partially validates **Hypothesis A.2**. Note that although the fitted model does not account for it, the data visually appears to have a non-monotone relation with frequency, particularly at $\rho_{ask}=0.4$ and $0.8$. A more granular analysis of the relation between frequency and human help is required to understand the shape of this curve, and is left for future work.

1) **Understanding Individual Variation**: Fig. 4 shows the predicted probability of helping per user, using the latent helpfulness value, $h_{UID}$, that most predicted that user’s behavior. These graphs reveal that the learnt individual differences are thresholding differences. Each person had different threshold busyness and frequency levels at which they helped. For example, some only helped during free time, while others risked being late in order to help. This behavior could not be modeled without both individual and contextual factors; only-contextual (solid black line) would miss opportunities to ask helpful people, and only-individual (which would be horizontal lines) would ask even people who are very busy.

The open-ended responses reveal factors that may have influenced users’ latent helpfulness. Some users felt they should not have to help robots: “it would be acceptable for the robot...to have a slight change in my activity.” These responses point to avenues to extend our model in future work.

2) **Additional Factors**: The open-ended responses also reveal unmodeled contextual factors that influence whether humans help. Multiple users wrote about the deviation required to help the robot: “I helped...if the place it wanted to go was near where I was going.” Although we can measure that deviation in the online environment, we did not include it in our model.

**F: Model Evaluation**

We use 5-fold cross-validation to compare our model to baselines. Our proposed model, Hybrid, is both individualized and contextual. We compare it with Contextual, which is non-individualized but contextual (excluding $h_{UID}$ in Eq. 3), Individual, which is individualized but non-contextual (excluding $c_2$ and $c_3$ in Eq. 5), and Only Intercept, with is neither (only $c_1$ in Eq. 5). We also compare it with a Random Forest model with the contextual factors as input, a state-of-the-art approach for predicting human interruptability [2]. Individual is similar to a state-of-the-art model of human helpfulness [38].

Unlike common applications of cross-validation, where the input data is fully known, the individualized models take in the unobservable latent helpfulness. We use sequential reasoning to overcome this challenge. When evaluating user $i$ in the test set, the first time the robot asks for help we assume they have the mean $h_{UID}$ of 0 (the prior learnt by GLMM). Every subsequent time, we minimize the cross-entropy loss function to predict their most likely latent helpfulness, given the robot’s past interactions with them.

Table I shows the results of this 5-fold cross-validation, where proportions are inversely weighted by number of interactions to prevent users in higher $\rho_{ask}$ conditions from disproportionately influencing the results. To analyze the results, we ran a repeated measures ANOVA on both the accuracy and F1 scores of these models, with model as a within-subjects factor and fold as a participant identifier. The accuracy data aligned with both the equality of variance (Mauchly’s Test of Sphericity) and the normality (Shapiro-Wilk test) assumptions of the ANOVA. The F1 score data aligned with the equality...
of variance assumption but violated the normality assumption for Random Forest and Contextual; however, we still used ANOVA because past works have shown it to be robust to violations of normality [40, 27]. This analysis revealed that the model had a significant effect on accuracy ($F(4,16)=13.14, p<0.001$) and the F1 score ($F(4,16)=10.18, p<0.001$). A post-hoc paired t-test comparing every model to Hybrid revealed that Hybrid significantly outperformed every baseline on accuracy ($p<0.01$ for all), and significantly outperformed every baseline but Contextual and Random Forest on F1 score ($p\leq 0.001$ for significant differences). The high standard deviation for Contextual and Random Forest’s F1 scores is due to folds where they predicted the human would never help.

### G. Human Perceptions

For the RoSAS, a one-way ANOVA with frequency as a between-subjects factor found no significant differences. This might be because each human experienced one robot, so may not have baseline expectations of robot behavior to compare to. For the NASA-TLX, Mental Demand had a significant difference ($F(4,135)=4.00, p=0.004$), where $\rho_{ask}=0.4$ had significantly higher mental demand than 0.2 ($p=0.002$). This could be because at 0.4 the robot asks frequently enough to not be novel but not frequently enough for the human to be familiar with helping it; investigating this is left to future work.

### IV. PLANNING WITH THE MODEL

Since Hybrid incorporates the unobservable factor of latent helpfulness, it requires a planning framework that can reason about uncertainty. We formalize the problem of a robot completing its task with bystander help as a Bayes-Adaptive Markov Decision Process (BAMDP) $\langle S, \Phi, A, T, R, b_0, \gamma \rangle$ [31], which we selected because the robot’s uncertainty is over a fixed parameter. We present a high-level planning framework, which is invoked when the robot sees a bystander, and outputs whether the robot should ask them for help. Because our subsequent evaluation (Sec. V) analyzes the robot behavior that emerges from different human help models, this planning framework focuses on the case where the robot is heavily dependant on human help—and therefore the model—to achieve its goal.

#### A. BAMDP Formalization

A BAMDP is a 7-tuple $(S, \Phi, A, T, R, b_0, \gamma)$ [31]. The state $(b_t, \rho_{ask,t}, \text{human\_helps}_{t-1}, \text{correct\_room}_{t-1}) \in S$ consists of the human busyness at time $t$, the robot’s past frequency of asking as of time $t$, whether the human helped accurately at time $t-1$ and whether the robot reached the correct room at time $t-1$. Time $t$ is incremented every time the robot sees the human. The latent variable $h_{UID} \in \Phi$ is the human’s latent helpfulness, and the robot’s actions $A = \{a_{ask}, a_{walk\_past}\}$ are either to ask the human for help or to walk past them and attempt to reach its goal on its own.

Within the transition function $T$, the busyness $b$ transitions uniformly at random because the robot assumes no regularity in human busyness (although this could be improved with domain-specific knowledge). The latent helpfulness $h_{UID}$ is fixed, and the frequency of asking $\rho_{ask}$ transitions as expected. If the robot asked for help, the transition function uses the model to predict the human’s probability of helping $P(\text{human\_helps} | h_{UID}, b, \rho_{ask})$. If the robot did not ask or the human does not help, the robot picks a room uniformly at random ($\frac{1}{n}$ chance of success), and sets $\text{correct\_room}$ accordingly. Note that this transition function is modular, so any probabilistic model of human help can integrate into it.

The reward function $R$ gives a penalty of $-r_{ask} \in [0, 1]$ for asking for help and a reward of 1 for reaching a correct room. The initial belief distribution, $b_0(h_{UID}) := \mathcal{N}(0, \sigma^2)$, is the prior learnt by GLMM. The discount factor $\gamma := 0.99$.

1) The Belief Update: At every time $t$, the robot updates its belief over the human’s latent helpfulness $b_t(h_{UID})$. Standard belief updates use Bayes rule with the precise transition probabilities from $T$ as the likelihood function. However, such belief updates assume the model explains all the variability of human behavior, which is untrue. Hybrid mispredicted 22% of human actions, and open-ended responses revealed unmodeled factors and stochasticity that influenced users’ help-giving behaviors (Sec. III-E2). Past works have suggested ways to augment Bayes Rule to account for such model misspecifications [32, 21]. However, in those works the degree of model misspecification is unknown, whereas in our work the training data reveals the amount of variance that is unexplained by the model. To account for this expected unmodeled variability of human behavior, we augment the belief update with the distribution of expected noise from human behavior, $\mathcal{E}$.

$$b_t(h_{UID} | \text{human\_helps}_t, b_t, \rho_{ask,t}) = \frac{1}{\eta} \left( \int_{\mathcal{E}} \left( P(\text{human\_helps}_t | h_{UID}, b_t, \rho_{ask,t}) + \mathbb{E}[\epsilon] \right) \cdot b_{t-1}(h_{UID}) \right) \cdot b_{t-1}(h_{UID}) \cdot d h_{UID},$$

(7)

$$\eta = \int_{h_{UID}} \left( P(\text{human\_helps}_t | h_{UID}, b_t, \rho_{ask,t}) + \mathbb{E}[\epsilon] \right) \cdot b_{t-1}(h_{UID}) \cdot d h_{UID},$$

(8)

where $\mathcal{E}$ is the mean-0 normal distribution of model residuals, truncated so probabilities stay in $[0, 1]$. In practice, augmenting the belief update with expected noise serves to slow the belief update, giving the robot more interactions before inferring that the human’s behavior is due to latent helpfulness.

### TABLE I: 5-fold cross-validation on Hybrid and baseline models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
<th>F1 Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hybrid</td>
<td>0.78 (0.03)</td>
<td>0.63 (0.03)</td>
</tr>
<tr>
<td>Contextual</td>
<td>0.70 (0.03)**</td>
<td>0.42 (0.24)</td>
</tr>
<tr>
<td>Individual†</td>
<td>0.71 (0.05)**</td>
<td>0.41 (0.07)**</td>
</tr>
<tr>
<td>Only Intercept</td>
<td>0.69 (0.06)**</td>
<td>0.00 (0.00)*****</td>
</tr>
<tr>
<td>Random Forest†</td>
<td>0.68 (0.05)**</td>
<td>0.41 (0.23)</td>
</tr>
</tbody>
</table>
tasks, and the robot interacted with the human at the same
time. Both conditions had the same sequence of
task-related pre-defined tasks. To indicate the robots were different, we colored the robot in condition#1 orange and in condition#2 purple. The main differences from the data collection study included the NASA-TLX \[22\] and RoSAS \[8\]. Survey#2 also included the HOP and HOP-POMDP \[37\], with the addition of Bayesian policies that use Hybrid, Contextual, and Individual, trained on the non-parametric Related Samples Wilcoxon Signed Rank Test to analyze the relation between policy (2 levels) and the dependant variables. We preceded this analysis with an independent samples t-test to investigate ordering effects, and found no significant impact between order and any of the dependant variables \((p=0.1–0.9)\).

Table II shows the policy’s mean and standard deviation performance, compared pairwise with Hybrid. Hybrid received a significantly higher (1.5X higher) cumulative reward than Contextual \((Z=-2.405, p=0.016)\) and Individual \((Z=-2.599, p=0.009)\), also shown in Fig.5 (left). This validates Hypothesis B.1 and demonstrates that Hybrid significantly outperformed baselines. Further analysis reveals that Hybrid asked for help significantly less (1.2X less) than both Contextual \((Z=-4.004, p<0.001)\) and Individual \((Z=-3.048, p=0.002)\). Despite asking for help less, Hybrid received more help than Contextual and significantly more than Individual \((Z=-5.250, p=0.012)\). Hybrid also got significantly fewer help requests rejected than Contextual \((Z=-4.848, p<0.001)\) and Individual \((Z=-5.325, p<0.001)\). This validates Hypothesis B.2.

Fig.5 (right) provides additional insight into the policies’ behavior. This graph clusters users by observed helpfulness, measured on Hybrid because every user experienced it. This graph reveals that Hybrid consistently got less rejected help (negative y-axis) than baselines, regardless of how helpful the human was. For less helpful people, Hybrid recognized that sooner and asked less, whereas for more helpful people Hybrid received more help (positive y-axis). Contextual asked (bar height) a similar amount regardless of helpfulness, because it did not individualize. Both Hybrid and Individual asked

We set \(r_{ask}:=-0.2\), based on a parameter sweep using simulated users. At \(r_{ask}=0\) or 1, policies always or never asked, regardless of the human help model. At \(r_{ask}=-0.2\), the policies relied on the model and asked an average of 25–50% of the time, which seemed sufficient for users to gain an idea of what each policy does. We recruited participants on Amazon Mechanical Turk with the same criteria as above, compensating them $7.5 for this 50 minute study.

### B. Hypotheses

- **Hypothesis B.1**: Hybrid will have a higher cumulative reward than Contextual (a) and Individual (b).
- **Hypothesis B.2**: Hybrid’s help requests will be rejected less than Contextual’s (a) and Individual’s (b). The intuition is that Individual will not identify when people are too busy to help and Contextual will not identify which people are less helpful, so both will ask inappropriately.

### C. Results

We collected data from \(n=100\) people (mean age 36.2, 35 F, 65 M), with 50 people per experiment and 25 per order (“Hybrid 1st” or “Hybrid 2nd”). This excludes participants who did not follow tutorial instructions or incorrectly answered survey attention check questions. Because multiple of the dependant variables violated ANOVA’s equality of variance (Box’s Test, Levene’s Test) and normality (Shapiro-Wilk Test) assumptions, we used the non-parametric Related Samples Wilcoxon Signed Rank Test to analyze the relation between policy (2 levels) and the dependant variables. We preceded this analysis with an independent samples t-test to investigate ordering effects, and found no significant impact between order and any of the dependant variables \((p=0.1–0.9)\).

<table>
<thead>
<tr>
<th>Metric</th>
<th>Hybrid ((n=50))</th>
<th>Contextual ((n=25))</th>
<th>Individual ((n=25))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cumulative Reward</td>
<td>2.74 (3.18)</td>
<td>1.87 (2.95)*</td>
<td>1.80 (3.04)**</td>
</tr>
<tr>
<td>Num Correct Rooms</td>
<td>4.29 (3.88)</td>
<td>3.80 (3.70)</td>
<td>3.88 (3.03)</td>
</tr>
<tr>
<td>Num Asks</td>
<td>7.76 (3.76)</td>
<td>9.64 (4.12)**</td>
<td>10.40 (1.16)**</td>
</tr>
<tr>
<td>Num Help Received</td>
<td>3.83 (3.91)</td>
<td>3.26 (3.72)</td>
<td>3.48 (3.15)*</td>
</tr>
<tr>
<td>Num Help Rejected</td>
<td>3.93 (1.31)</td>
<td>6.38 (1.24)**</td>
<td>6.92 (3.36)**</td>
</tr>
</tbody>
</table>

TABLE II: The mean (standard deviation) of cumulative reward, number of correct rooms reached, number of times asked, number of times helped, and number of times not helped. 1-state-of-the-art, \(*p<0.05, **p<0.01, ***p<0.001\)

### V. Evaluation

We conducted a user study to evaluate the tradeoffs between policies that use Hybrid, Contextual, and Individual, trained on the full dataset. Hybrid has parameters \(c_1=-0.05, c_2=-5.90, c_3=-8.60\), \(\sigma^2=5.89\). Note that Individual plugged into our planning framework is similar to the state-of-the-art LM-HOP and HOP-POMDP \[32\], with the addition of Bayesian reasoning over the individual factor.

#### A. Experimental Design

We ran two within-subjects experiments: Hybrid v. Contextual and Hybrid v. Individual. Each study contained the study description, tutorial, condition#1, survey#1, condition#2, and survey#2. Both conditions had the same sequence of 28 human tasks, and the robot interacted with the human at the same 20 pre-defined tasks. To indicate the robots were different, we colored the robot in condition#1 orange and in condition#2 purple. The main differences from the data collection study were: 1) whether the robot asked for help was determined in real-time using the planning framework; and 2) the human busyness for non-free-time tasks was randomized. Users experienced policies in a random order. Both survey#1 and #2 included the NASA-TLX \[22\] and RoSAS \[8\]. Survey#2 also had forced-choice (e.g. “which robot was more annoying?”), open-ended (e.g. “what differences were there between robot #1 and robot #2?”), and demographic questions. For each condition, our dependant variables were the cumulative reward (Eq. 9), number of correct rooms the robot reached, number of times it asked for help, number of times it was helped, and number of times its help request was rejected.

\[
\text{cumulative\_reward} = \text{num\_correct\_rooms} - r_{ask} \cdot \text{num\_asks} \quad (9)
\]

3https://github.com/JuliaPOMDP/BasicPOMCP.jl
more helpful people more, although Individual received more rejected help because it did not account for contextual factors.

D. Human Perceptions

We analyzed human responses to the RoSAS with the same t-test and Wilcoxon test as above. This revealed that Hybrid scored significantly lower on the perceived “discomfort” measure than Contextual ($Z = -2.085, p = 0.037$), which was primarily due to users associating Contextual with “awkward” significantly more than Hybrid ($Z = -3.136, p = 0.002$). This could be because Contextual asked much more than Hybrid, even for humans who are not latently helpful, which can come across as “awkward”. There was a significant ordering effect with user’s perception of Individual’s “warmth” ($t(48) = 2.143$, $p = 0.037$), where users who experienced Hybrid first rated Individual as significantly less warm ($Z = -2.454, p = 0.014$), while others did not have a significant difference. The significant difference was primarily due to those users associating Hybrid with “feeling” significantly more ($Z = -2.300, p = 0.021$). This might be because Hybrid accounted for the human’s busyness, which can come across as “feeling.”

There were no significant effects on the NASA-TLX, and the few significant effects on forced-choice questions had too many ordering effects. However, the open-ended responses revealed that users perceived the differences between policies. When asked what differences there were between robots, users in Hybrid v. Individual experiment wrote that Hybrid asked for help at more appropriate times: “Robot #2 [Hybrid] also stopped me at more appropriate times like when I was on a break or had a lot of time.” Similarly, users in Hybrid v. Contextual identified individualization as a difference between the policies: “Orange [Hybrid] stayed away after telling it no.” Finally, across both experiments, users wrote about Hybrid’s likeability: “Purple [Hybrid] was politer [than Individual]” or “Robot 1 [Contextual] was more annoying and asking (sic) for help than robot 2 [Hybrid].”

VI. DISCUSSION

In this work, we developed a model of human help that disaggregated individual and contextual factors, integrated the model into a BAMDP planning framework, and demonstrated that the resulting policy significantly outperformed baselines.

A. Generalizability of the Model

Our disaggregation of factors into individual factors, which are unchangeable (and often unobservable) characteristics of a human, and contextual factors, which can change and embed the human within their broader physical and temporal context, is a high-level separation that we believe will generalize to other human help scenarios. In a scenario where both factors are important, our method enables the robot to better achieve its goals. In a scenario where one factor dominates, our approach reduces to the individual [37] or contextual [2] types of models in prior work. The specific contextual factors we use—busyness and robot frequency of asking—are also general factors that we believe will impact any instance of human help. Note that although the planning framework assumes individual factors are fixed, because the belief update is Bayesian it should adapt to gradual changes in latent helpfulness. Some non-gradual changes could be explained by contextual factors (e.g., students not being helpful during exam time), although extending our approach to dynamic latent helpfulness is left to future work. Finally, note that the BAMDP planning framework can generalize to multiple humans—by adding more $h_{UID}$, $b$, and $\rho_{ask}$ values—although that scenario may require extending the model for the bystander effect [19].

B. Deploying onto a Physical Robot

 Deploying this approach on a physical robot requires a module to perceive busyness. This can use proxies such as walking speed, known calendar events, or factors linked to interruptability such as eye gaze [2] or time of day [43]. Although such a module would result in partially-observable busyness, initial simulation results revealed that even with partially-observable busyness the comparative performance of models was the same, but more noisy. Deploying this approach on a physical robot also requires a module for perceiving unique individuals. Finally, it requires reasoning about low-level movement actions in addition to the high-level asking actions we consider. One approach is integrating movement
actions into the BAMDP, so the robot can jointly reason about movement and help (e.g., moving in a direction likely to have human helpers). Another approach is maintaining the separation between a high-level help policy and a low-level movement policy, switching when the robot sees a human.

C. Broader Impacts

We discuss broader impacts of this work, following Hecht et al. [23]'s approach of considering “reasonable broader impacts, both positive and negative.” One potential positive impact is increasing robots’ robustness to unexpected scenarios, by enabling them to ask for help in situations they are unequipped for. This may also reduce the cost of robots, by enabling them to do more with limited sensors and computation. One potential negative impact is perpetuating gender bias. Research has found that women tend to be, or are expected to be, more altruistic, prosocial, or unselfish than men [4, 11, 24, 15], a trend that has been mirrored with humans helping robots [45]. Therefore, our robot’s behavior—asking for help from humans who are believed to be helpful—could result in disproportionately burdening women with providing help. A technical approach to mitigate this risk is penalizing the robot for repeatedly asking the same person. However, fully understanding the social impacts of such a robot would involve studying the socio-technical system—including employment conditions and culture—that the robot embeds into (e.g., using the mutual shaping framework [39]).

D. Limitations and Future Work

One limitation of this work is that it was conducted online. Past works indicate that the factors we model are still relevant in real-world help [35, 20, 34], but the parameters may differ. Another limitation is that our study had one person in the office at a time. We believe multiple humans will increase Hybrid’s advantage, since the robot can ask the most latentely helpful human, although that is left to future work. Another limitation is that the model-learning methodology can only account for factors with a monotone relationship to the humans’ likelihood of helping, because of the GLMM’s logistic link function. This can generalize to some but not all of the additional factors users identified (Sec. III-E2). A final limitation is that the robot did not improve its autonomous performance based on the help. The robot did improve its ability to ask humans for help—by learning their latent helpfulness—but in the same scenario in the future it would still be unable to autonomously complete the task. This was necessary to create a scenario where the robot was dependant on human help. However, some users’ open-ended responses showed that their help was conditioned on robot improvement, so enabling the robot to improve its autonomous performance based on human help is a promising direction for future work.

An exciting direction for future work is incorporating additional aspects of human help: human expertise, inaccurate help, and multiple helpers. Another promising direction is extending the approach to asynchronous remote help.

ACKNOWLEDGMENTS

This work was (partially) funded by the Faithful Steward Endowed Fellowship, National Science Foundation IIS (#2007011), National Science Foundation DMS (#1839371), the Office of Naval Research, US Army Research Laboratory CCDC, Amazon, and Honda Research Institute USA. This work used statistical consulting provided by Youngwon Kim at the Center for Social Science Computation and Research. The authors thank Mridula Venkatesan for her help with data processing and developing the office environment tutorial.

REFERENCES


