Assistive value alignment using in-situ naturalistic human behaviors

Benjamin A Newman

CMU-RI-TR-24-66

September 30, 2024

The Robotics Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA, 15213

Thesis Committee:
Henny Admoni *Chair*
Kris Kitani *Co-Chair*
Andrea Bajcsy
Dylan Losey *(Virginia Polytechnic Institute and State University)*
Christopher Paxton *(Hello Robot)*

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Robotics.

Abstract

As collaborative robots are increasingly deployed in personal environments, such as the home, it is critical they take actions to complete tasks consistent with personal preferences. However, determining personal preferences for completing household chores is challenging. Many household chores, such as setting a table or loading a dishwasher, are sequential and open-vocabulary, creating a landscape of almost endless a priori preferences. Taking assistive actions in this domain means that a robot must first determine someone's personal preference from within this expansive space. To do this, robots rely on people to communicate information about their preferences.

Communication about preferences is often collected *ex situ*: A person is presented with an abstract situation with several alternative solutions and gives feedback on which solution they think they would prefer if they were acting *in situ*. This feedback on the preferred solution, combined with similar responses from multiple people in multiple situations, is then used to train a preference model. These data can be burdensome to collect, are based on *ex situ* data collection which does not guarantee alignment with *in situ* preferences, and fails to capture information about changing to preferences that may arise due to the execution of the collaboration.

In this thesis, we argue that robots can provide personalized *in situ* assistance using observations of naturalistic human behaviors. In other words, robotic assistance can be viewed as a process of value alignment and can be achieved during task execution using observations of naturally occurring goal-directed behaviors. To support this argument, we make five main contributions.

First, we define assistive robotics as a value alignment problem and identify the main components in defining such a problem: the people involved, the space (or environment) in which the interaction takes place, and the relative timing of the robot and collaborative partners' actions. Second, we introduce a dataset of naturalistic human-robot collaboration behavior collected in a simple collaborative object rearrangement task. Third, we use this data set to highlight the importance of continued personalization in assistive scenarios. Fourth, we present a method for extending these ideas to complex surface rearrangement tasks with naturalistic data using large internet-scale pretrained multi-modal foundation models. Finally, we present a method for continually finetuning these large foundation

models using naturalistic *in situ* behaviors, demonstrating how we can provide seamless robotic assistance from varying sources of *in situ* human behavior data.

Acknowledgments

Writing this thesis and completing a doctoral degree has been a journey unlike any other. There are many people who have helped me achieve this milestone, and I take this opportunity to thank them.

My advisors Henny Admoni and Kris Kitani have helped immensely as I completed this degree. I thank them for their time and generosity. I have appreciated their perspective and their challenges, as I have learned to formulate my own opinions on robotics and research. Most of all, I am grateful for the opportunity they have given me.

Adam, Brian, Leo, Ankit, and all the remaining Burger Bros whose names I have misplaced, I appreciate the time at and away from work that was more pleasurable in your presence. Whether it was spending time at Stack'd until the wee hours of the morning or finalizing problem sets for RoboMath, it was always more joyful in your company.

HARPLab has been filled with wonderful colleagues and I am grateful for all the time we have spent together in CHOIR presentations, lab socials, and reading groups.

I am also grateful for the opportunities I have had outside of CMU during the course of my Ph.D. I was lucky enough to work at Meta twice as I completed my degree. To Kevin and Ruta, I am thankful for the wonderful 6 month turned 8 month internship turned pandemic I was afforded. This first foray into industrial research opened my eyes to the possibilities of applied research and I am now seeing the fruits of many of the efforts I witnessed during my time there. To Chris and the folks at EAI Pittsburgh, it was a great two years and I am thankful for all of our GrubHub lunches, no matter how many times we had to explain to the delivery driver not to turn onto Neville Rd.

Sam, Brett, Nate, Devon, Dan, how long has it been? It is amazing to have you all in my life still, and to meet your partners and your children, and to watch you all continue to grow into such wonderful people since we met all those years ago in Somerville. You have been there for me in more ways than you can imagine, and I look forward to all the wonderful times that we have ahead.

To my mother and father, without whom I could not be where I am today. They have been a constant source of support through many years of education and I hope they continue to help me navigate the various ups and downs of life. And to my sister Emily, who is there to commiserate when I refuse to listen to my parents and has been a steady support in her own right.

Finally, to my wife, Marie, who is simultaneously the most important

and constant admirer in my life, and the force that keeps me most humble. She makes me laugh and smile when I think it is impossible. She cares for me when I scarcely have the time to take care of myself. And she has given me the three greatest joys of my life: a marriage to a wonderful, supportive, and true partner in life, my silly dog Picasso who keeps me grounded, and my daughter, Anaïs, whom I love and admire beyond limits. I love all three of you in ways that you cannot imagine and thank you for the sacrifices, the comedic relief, and the immense amount of support you have provided me as I pursued this endeavor.

Funding

This work has been supported by grants from the National Science Foundation (IIS-1755823), ONR (award N00014-18-1-2503), the Paralyzed Veterans of America, the Tang Foundation, and the National Science Foundation Graduate Research Fellowship (Grant No. DGE1745016 and DGE1252522). This work was also supported through funding from Meta Platforms, Inc. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

Contents

1 Introduction **1**

1.1 Thesis outline . . . 3

1.1.1 Background and related work . . . 4

1.1.2 Components of robotic assistance: people, space, and time . . . 4

1.1.3 Human behavior during simple surface rearrangement tasks . . . 4

1.1.4 Evidence that value alignment requires continual personalization 5

1.1.5 Personalized feature spaces in complex surface rearrangement tasks . . . 5

1.1.6 Continual personalization using naturalistic corrections during surface rearrangement tasks . . . 6

1.1.7 Conclusion . . . 7

2 Background and related work **9**

2.1 Human behaviors during human robot collaborations . . . 9

2.1.1 Naturalistic human behaviors for robotic control . . . 10

2.1.2 Multimodal human behavior data sets . . . 10

2.1.3 Explicit featurizations of human behaviors for assisted teleoperation . . . 12

2.2 Household tasks as object rearrangement problems . . . 13

2.2.1 Object rearrangement in robotics and embodied AI . . . 14

2.2.2 Formalizing collaborative rearrangement . . . 14

2.3 Cooperative inverse reinforcement learning for collaborative rearrangement . . . 15

2.3.1 Naturalistic behaviors for adaptive collaborations . . . 16

2.3.2 Solving collaborative object rearrangement with cooperative inverse reinforcement learning . . . 17

2.4 Implicit featurizations for collaborative rearrangement . . . 18

2.4.1 Foundation models for robotics . . . 19

2.4.2 Visual prompting in vision language models . . . 19

2.4.3 State and action-conditioned collaboration . . . 20

3 Components of robotic assistance: people, space, and time **23**

3.1 The assistance perspective . . . 25

3.1.1 General human-robot interaction . . . 25

3.1.2 Assistance as a perspective on human-robot interaction 26
3.1.3 Prior categorizations of assistive robotics 28
3.1.4 Overview of this paper 29
3.2 Methods 29
3.3 People 31
3.3.1 Terminology 31
3.3.2 Additional targets of assistance 32
3.3.3 Additional interactants 34
3.3.4 Combinations of roles 35
3.3.5 Implications 35
3.4 Space 36
3.4.1 Environment 37
3.4.2 Human body 39
3.4.3 Human brain 40
3.4.4 Implications 44
3.5 Time 44
3.5.1 Reactive 45
3.5.2 Proactive 46
3.5.3 Simultaneous 48
3.5.4 Implications 49
3.6 Conclusion 49

4 Human behavior during simple surface rearrangement tasks 53
4.1 Data collection procedure 56
4.1.1 Participants 56
4.1.2 Protocol 56
4.1.3 Assistance conditions 57
4.1.4 Sensors 58
4.2 Descriptive statistics 59
4.3 Data streams 59
4.3.1 Timing and synchronization 61
4.3.2 Eye gaze 61
4.3.3 Third person video 62
4.3.4 Additional sensor data 62
4.4 Known issues 63
4.4.1 Missing data 63
4.5 Accessing the data 64
4.6 Conclusion 64

5 Evidence that value alignment requires continual personalization 67
5.1 Problem domain 69

5.1.1 Defining action primitives 70
5.1.2 Naturalistic human behavior data set 72
5.1.3 Synthetic data set 72
5.2 Method 73
5.2.1 Micro and macro action labeling 73
5.2.2 Models and input representations 74
5.2.3 Problem setup 75
5.3 Experimental results 75
5.3.1 Micro action results 75
5.3.2 Macro action results 76
5.3.3 Participant level cross validation 76
5.4 Discussion 76
5.5 Conclusion 78

6 Personalized feature spaces in complex surface rearrangement tasks 83
6.1 Problem specification 85
6.2 Estimating simulated table arrangements using VLMs 86
6.2.1 Producing task plans through visual prompting 87
6.2.2 Simulating preferences for table arrangement 88
6.2.3 Validation metrics 89
6.2.4 Results 89
6.3 Naturalistic tabletop arrangement dataset 90
6.3.1 Data collection 90
6.3.2 Additional baselines and methods for prompting FMs 94
6.3.3 Results 96
6.4 Using corrective actions to refine pretrained feature spaces 97
6.5 Limitations and Generalizability 99
6.6 Conclusion 100

7 Continual personalization using naturalistic corrections during surface rearrangement tasks 103
7.1 Approach 106
7.1.1 Problem setup 107
7.1.2 Datasets 108
7.1.3 Learning preferences in a diverse user population 110
7.1.4 Bootstrapping shallow linear models for fast, online adaptation 111
7.2 Experimental design 112
7.2.1 Zero-shot coordination 112
7.2.2 Test-time adaptation 114
7.3 Results 115
7.4 Discussion, limitations, and future work 117

7.5 Conclusion 118

8 Conclusion **121**
8.1 Generalizability 122
8.2 Future research directions 123
8.2.1 Real robot implementation 123
8.2.2 Modulating between conflicting goals 124
8.2.3 Incorporating multiple modalities of naturalistic human behaviors 125
8.3 Closing thoughts 126

Bibliography **127**

When this dissertation is viewed as a PDF, the page header is a link to this Table of Contents.

List of Figures

3.1 An assistive system can treat people beyond a single user as additional targets of assistance or as interactants, and either choice introduces particular complications into the assistive dynamic. 31

3.2 A robot can provide assistance by acting in several different action spaces. It can assist by giving information to the user, adjusting the user's body, or changing the environment to help complete the task. . 36

3.3 A key axis in assistive robotic systems concerns what type of cue leads to the robot taking actions. Robots can be reactive and respond to explicit input only, be proactive and interpret the general task state to choose to act on their own, or collaborate closely with the user by acting simultaneously with them. 45

4.1 The HARMONIC data set provides multimodal human, robot, and environmental data collected during an assistive human-robot collaboration. 54

4.2 A visualization of several streams from the HARMONIC data set. The top row displays the ZED video with OpenPose skeletons overlaid, then the egocentric video captured from the Pupil camera, left eye video, one second of the calculated gaze dot, the trajectory of the joystick, and finally the Myo activations. For the gaze dot and the joystick, lighter colors represent more recent points in time. Each of these plots represents one second of data, sampled at 30 FPS. 60

5.1 The HARMONIC data set contains (a) third person video, (b) egocentric video, (c) eye gaze fixations, and (d) joystick data from a human-robot co-manipulation task. 69

5.2 Five macro action labels capture combined eye gaze and end effector dynamics. a) Exploration denotes periods of high eye gaze movement and low joystick (robot) movement. b) Pursuit denotes periods of highly correlated eye gaze and joystick movements, where the eye gaze follows the path of the robot. c) Correction denotes successive glances between different parts of the scene while the robot is moving. d) Mode switch denotes when the user is using modal control to cycle through sets of degrees of freedom. e) Toggle denotes periods in which the joystick is being moved in rapid, short, consistent activations. . . 70

5.3 A graphical description of the differences between macro action categories from HARMONIC. Eye gaze sequences are red, while joystick sequences are blue. The y-axis shows the normalized cumulative distance for each sequence. The x-axis shows normalized sequence lengths. Every trial is plotted, with a representative sequence highlighted in bold. 71

5.4 Our synthetic dataset was modeled as a simplified version of the eye gaze and joystick signals from the HARMONIC task. Here, an example of the exploration action, with simulated eye gaze (square), robot position (triangle), and goal (circle). 72

6.1 DegustaBot takes in a single person's preferred table arrangements (shown here as the visual context k_i and order of object placement $\mathcal{T}_i$), the objects O from which the algorithm can select, the table to set s_0 and the task prompt ℓ. The robot then produces a task plan in the form of an object arrangement (visualized as an image s_T) and the order of object placement $\mathcal{T}$. This predicted arrangement, Pred, should match a held-out preference created by the user, GT. 84

6.2 Details of a table arrangement. An arrangement is described by the objects within the arrangement, the order in which they are placed, their features (such as color, shape and material, and their location and orientation. 85

6.3 Object and arrangement lifting functions, from left to right: OaL, a representation of objects as language, for this we use the json representation; GoMO, grid of marked objects, which represents objects visually with referential marks overlaid on each object; UmA, Unmarked Arrangement lifts the arrangement into the image domain; and finally GMA, Grid-Marked Arrangement, which overlays a spatial reference grid on the continuous table top space. 87

6.4 Quantitative results for evaluating DegustaBot on simulated preferences. Left shows the performance of each model's ability to capture the geometry of the table arrangement, as measured by RMSD. Right shows each model and method's accuracy when choosing items to place in the arrangement. GPT-4o and MOGMA perform the best on both metrics. 90
6.5 Qualitative results. On the top line we see ground truth images from our naturalistic preference data and on the bottom we see DegustaBot-LOGMA's predictions, lifted into the image domain. On the left side of the image we see three examples where DegustaBot-LOGMA predicts similar arrangements to the ground truth image. On the right we see some failure cases. 93
6.6 Additional Qualitative results. We show the qualitative performance of all methods on five random experiments from five random participants. From left to right we show the Target image (GT), Copy and Paste (CP, Human Performance (HP), KNN (KNN), DegustaBot-MOGMA (GMA), DegustaBot-LOUMA (UMA), DegustaBot-LLM (LLM), DegustaBot-LvM (LVM), DegustaBot-VLM (VLM). 94
6.7 Highly correlated objectives across qualities (top) and those same objectives blended with noise ($\beta = 0.5$). Rows of an objective represent surface locations and columns represent objects. Elements represent the value associated with placing that object in that location, ranging from $[-1, 1]$. 98
6.8 We report mean average regret (mAR). From left to right we show per ID representations and per Quality representations. From top to bottom, we show training and zero-shot testing on out of domain objects after N iterations. Colors indicate β thresholds as follows: 0 is blue, 0.01 is orange, 0.1 is green, 0.5 is red, and 1 is purple. Both methods perform similarly in training for well-correlated objectives, while per ID representations outperform when preferences are uncorrelated. For per Quality representations, improvements when training on correlated preferences correspond to improvements in testing set, which does not hold true for per ID representations. 102

7.1 One step of an example surface rearrangement task: cupboard organization. From left to right: a person (H) picks an object to place in the dishwasher; the agent (A) initially places this incorrectly; the person corrects the placement. From this, the agent learns that the user likes to place blue objects on the bottom shelf and can place the next, similar object correctly. 104

7.2 From left to right, we first embed the input state and actions using ψ. These are then concatenated and fed into the preference estimator $\mathcal{M}$. This learns to output reward parameters, g_θ which are used to initialize an online learning policy using the policy π, which determines the robot's action a_r. 108
7.3 Learning curves for each test environment for each algorithm. We report the average accuracy over each episode. BLR-HAC is able to achieve the low zero-shot performance of the transformer method, and the fast adaptation of the linear method. Additionally, we can see that as the episode length increases, these differences in performance are more notable, with the linear method failing to catch up to the other two methods over the course of 20 episodes. 115
7.4 Learning curves over each test environment for each algorithm. We report average accuracy over episodes. BLR-HAC is able to perform on par with the transformer method in the small and medium environments and part of the large environment. BLR-HAC outperforms all methods in all environments immediately after the preference switch. In the large environment, though, the transformer recovers more quickly as it has access to more data. 116

List of Tables

3.1 Assistive robots can be explored along three key axes: how the assistive system thinks about additional **people**, what part of the mutual state aligns with its action **space**, and at what **time** it executes its actions during a task. 51

4.1 Descriptive statistics of each data stream in the data set. *Total duration* and *Total frames* refer to the collective amount of data of that signal over all trials and participants. *Total duration* is extracted by dividing the total frames by the *nominal frequency*. *Frames dropped* are based on interpolating from the nominal frame rate and detecting missing data. *Coverage* is computed by dividing the number of data frames by the expected number of data frames from the nominal frequency over the whole data set, *Present* indicates the fraction of trials that have at least one datum of that type, and *Coverage if present* is the total number of data frames divided by the expected number evaluated only if at least one datum is present in the trial. 65

5.1 Distribution of class labels in HARMONIC and synthetic datasets for both the micro and macro classification tasks. 73

5.2 Results for the synthetic and HARMONIC datasets on the micro classification task. We report accuracy (acc) and mean average precision (mAP). We test a variety of input streams: eye gaze only (E), joystick only (J), an early fusion of eye gaze and joystick (E+J-e), and a late fusion of eye gaze and joystick (E+J-l). Additionally, we test different embeddings (embed) of the inputs: binary, real (raw), and difference (del). Finally, we test two different hidden unit sizes ($\|h\|$): 16 and 256. 79

5.3 Results for the synthetic and HARMONIC datasets on the macro classification task. We report accuracy (acc) and mean average precision (mAP). We test a variety of input streams: eye gaze only (E), joystick only (J), an early fusion of eye gaze and joystick (E+J-e), and a late fusion of eye gaze and joystick (E+J-l). Additionally, we test different embeddings (embed) of the inputs: binary, real (raw), and difference (del). Finally, we test two different hidden unit sizes ($\|h\|$): 16 and 256. 80

5.4 Results for the HARMONIC dataset on the macro task using the best representations from Table 5.2. 80
5.5 Micro action cross validation by participant ID (eye only, dx,dy, 256 model). † Data were excluded due to significant noise. ‡ Data were not labeled for macro actions. 81

6.1 We report the percentage of people who find scenes below a specific threshold hold acceptable under both the random and structured noise condition. We find that people are extremely sensitive to random noise, and not very sensitive to structured noise. We see that when the RMSD is below 0.01, between 71.4% and 99.2% of people find the scene acceptable. 91
6.2 We compare our methods' performances against several baselines at several RMSD thresholds. Post Correction Jitter, which we consider our upper bound, produces model acceptable model responses 12.3% of the time at the 0.01 threshold, 78.7% of the time at the 0.05 threshold, and 92% of the time at the 0.1 threshold. We see that the Copy and Paste baseline, which uses privileged sensing information, has the second best response rate. Our Vision Model method performs the best out of methods that do not use privileged information and, interestingly, is closely aligned with the Human Performance baseline. 92
6.3 We compare DegustaBot-LLM to an autoregressive variant. We report thee results on a random subselection of 445 preferences, due to the significant cost increase of running an autoregressive variant. We see siginificant improvements over the LLM model at lower RMSD thresholds, suggesting that autoregressive models perform better than their non-autoregressive counterparts. 92
6.4 We report object prediction accuracy and RMSD of DegustaBot-MOGMA on naturalistic preferences. We see that with increased context length our method achieves higher performance, though perhaps with diminishing returns. 97

7.1 We compare zero-shot performance on the test set of each environment. We have two axes of comparison: model complexity in the rows, and inductive prior in the columns. Results are reported in terms of accuracy. We can see that the highest capacity, attention based model trained with an inductive prior outperforms all other models in every environment. 111

1 Introduction

While realizing the deployment of general purpose household robots will require significant advancements in many sub fields of robotics, a particular challenge arises with respect to human-robot interaction: how should robots come to understand what it is people want them to do?

People can be particular about how household chores are performed. For example, finding the proper way to load a dishwasher has sparked myriad online debates, and a quick trip to online forums will provide you with almost as many uses for a dishwasher as there are uses for a multitool. This is true even though many manufacturers provide optimal dish loading advice in their owner's manuals [338]. This indicates that household robots need to be flexible in the way they complete chores, allowing individuals to express unique preferences and subsequently performing the chore in a manner faithful to this preference. In other words: care should be taken to ensure that the robot's objectives align with the human's objectives. This process, called *value alignment* [113], does not regulate whether it is the robot's preference that is being executed or the person's. As such, we explore algorithms that operate under *assistive value alignment* which introduces an additional restriction that the robot must identify and execute the *human's* preference.

Determining people's individual preferences for completing a task is challenge. Population-level preferences can be captured by collecting large datasets of many different people's preferences. For example, by collecting *ex situ* preferences by choosing a between two trajectories taken by a robot in a lunar lander game [68] or autonomous driving scenario [267], a reward model can be developed that can be used to train policies that solve the underlying task (e.g. lunar lander or autonomous driving). This approach is perhaps most notable for its success in fine-tuning large language models with reinforcement learning through human feedback [89, 90, 91].

Due to their sample inefficiency, employing these approaches in situations where preference data is difficult to gather, for example, determining individual preferences for how to complete household chores, is challenging. Instead of relying on collecting *ex situ* data, we propose using naturalistic *in situ* data. In prior work researchers have used naturalistic behaviors such as joystick commands [138, 184] to specify objectives in simple, non-sequential tasks, or corrective pushes to specify the trajectory a person prefers a robot to take in motion planning tasks with specified start and end states [185]. We propose using *in situ*, naturalistic behaviors due to its availability (e.g. by recording people completing household tasks) and goal directed nature to train algorithms that execute household chores according to personal preference.

This problem is further complicated by the fact that personal preferences in household chores can be very complex. Household chores can be modeled as rearrangement problems [28]. This class of problems assumes access to a goal specification function $g = \phi(s_0, S^*)$ that maps an initial state s_0 and a solution set S^* to a goal g. Then, a goal-conditioned agent $\pi(a_t|s_t, g)$ is tasked with transforming the state from the initial state s_0 to a goal state $s* \in S^*$.

While much work in this area focuses on training the policy π, we are additionally interested in understanding how to personalize the goal specification function ϕ. Determining this specification can be difficult for two reasons: the open vocabulary and sequentially dependent nature of household rearrangement tasks and the communication mechanism robots use to interpret value-aligned goals.

Household object rearrangement tasks operate over unique sets of objects that differ from household to household and change over time. Additionally, people have complex relationships with how certain items, for example, expensive crystal glasses received as a wedding present, should be treated, making it difficult for a robot to predetermine the objects it may encounter or what it should do with those objects in any given task. This requires treating complex household rearrangement tasks as open-vocabulary tasks [351], in which there is no predetermined set of objects from which to choose. Solving such tasks typically requires developing [147] or using pre-trained generalized feature spaces [343] that characterize objects by their attributes as opposed to the class in which they belong.

Adding further complication is that solutions to household tasks can have ambiguous sequential ordering that may induce personal preference: should plates be made and then served during a dinner party, or should they be set first, with dinner served table side? To account for these difficulties, we situate our work within the domain of assistive surface rearrangement: a sequential decision making task that expands household object rearrangement to include under specified goal states and mechanism for people to communicate their desired goal state.

In this thesis, we propose and investigate two main ideas: 1) that general purpose household robots should be viewed as assistive robots that undergo a process of *assistive value alignment*, and 2) the source of information that robots use to determine people's preferences should be task-aligned, easy to express, and readily available.

1.1 Thesis outline

To investigate these ideas we divide our work into seven chapters. First, we give an overview of related and background work in Chapter 2. Then, we introduce the concept of assistive robotics as an assistive value alignment problem in Chapter 3, then explore the feasibility of collecting naturalistic data in assistive robotics tasks in Chapter 4, which we use to identify the need for personalized solution sets in rearrangement problems in Chapter 5. Following this we scale our problem space up to complex rearrangement tasks and determine the efficacy of using large foundation models to serve as personalized solution sets and goal specification functions in Chapter 6, after which we present an algorithm to continually fine-tune abstract preference representations (such as those given by large foundation models) using naturalistic behaviors in Chapter 7. Finally, we conclude with a discussion on the generalizability of the ideas presented in this research and open question and future directions for developing household robots that perform assistive value alignment from naturalistic behaviors in Chapter 8. These chapters are summarized below.

1.1.1 Background and related work

In Chapter 2 we outline the background and related work necessary to contextualize our research. We first begin with a discussion on human behaviors in human robot collaborations. We then formalize household tasks as object rearrangement problems, and discuss the various problem formulations we will use for the remainder of the thesis. Here, we present the various assumptions for each of our major contributions. Then, we formalize household rearrangement tasks as cooperative inverse reinforcement learning tasks. Finally we discuss various ways to represent states and actions as features for use in collaborative rearrangement problems that utilize large foundation models.

1.1.2 Components of robotic assistance: people, space, and time

Chapter 3 introduces our first contribution, which is to define assistive robotics as an assistive value alignment problem. We support this through a comprehensive review of the literature on assistive robotics research. Additionally, we identify three crucial components that determine an assistive system: the people involved in the assistive interaction, the space or environment in which the interaction takes place, and the time at which the human and robot partner's actions affect the environment. By identifying these components, researchers should be able to identify and resolve ambiguities that can arise during assistive value alignment. This work was drawn from our publication in Frontiers in Robotics and AI [221].

1.1.3 Human behavior during simple surface rearrangement tasks

Next, to determine the feasibility of using and collecting *in situ*, naturalistic behaviors for assistive surface rearrangement, we collect a suite of such behaviors in Chapter 4. Our dataset, called HARMONIC, is situated in a simple surface rearrangement task that represents an assisted eating task. In this work, we assume a goal specification function and define the state s_t, as being comprised of the robot's position x_t and the estimate of the user's goal $b(g)$ (a choice from among three possible goals). The robot's task is to observe the user's joystick actions a_t^h and use this information to update its estimate of the user's goal. After it is updated, the robot can use its estimate of the user's goal to determine an action a_t^r that moves the robot's position closer to the estimated goal. The task is complete when a single goal has been reached. During this time, the user exhibits myriad other behaviors that could indicate their goal choice, such as eye gaze, body position, and electromyographic signals. In HARMONIC, we collect these behaviors and introduce a large dataset of synchronized in situ, naturalistic behaviors. These behaviors can aid in the development of generalizable feature spaces that aids in mapping naturalistic human behaviors to unspecified goals in simple surface rearrangement tasks. This work was drawn from our publication in The International Journal of Robotics Research [222].

1.1.4 Evidence that value alignment requires continual personalization

In Chapter 5, we use the behaviors collected in Chapter 4 to attempt to train a generalizable feature space that could be used to perform downstream surface rearrangement tasks. We identify several actions people consistently perform when controlling a robot in the HARMONIC dataset and train a model that maps naturalistic behaviors to these action categories. We find that including multiple behavior modalities improves performance on the activity recognition task. Moreover, however, we find that our feature space, which was trained at a population level, has widely varying performance when generalizing across participants, as identified through leave-one-out validation. This indicates that while these feature spaces can be trained from naturalistic behavior at a population level, individual differences likely necessitate further personalization of the feature space in order to be useful for downstream tasks.

1.1.5 Personalized feature spaces in complex surface rearrangement tasks

In Chapter 6 we take the evidence that we need personalized feature spaces from Chapter 5 and aim to personalize generalizable feature spaces (parameterized by large foundation models) in complex surface rearrangement tasks. To this point, our work has been situated in simple surface rearrangement tasks characterized by a closed vocabulary and a lack of sequential dependence. Typical household tasks, however, are open vocabulary and have ambiguous sequential dependence. To scale our work to these types of tasks, it is imperative that we adopt a more general feature space that can encapsulate these characteristics of household tasks. Recent work in multi-modal foundation models trained over internet scale data has indicated that these models may serve this purpose well [343, 347]. Combining this insight with the insights from our previous work, namely that generalized feature spaces such as those found in pre-trained foundation models need personalized refinement for downstream tasks, we introduce DegustaBot.

In this work we first introduce a complex, open-vocabulary, sequentially dependent surface rearrangement task. Then, we introduce a method to use large-scale, pre-trained, multi-modal foundation models to solve these complex surface rearrangement tasks in a personalized manner. We validate this method on a naturalistic dataset of table setting examples and find that these feature spaces can be successfully personalized along features such as relative object distance and arrangement centroid position. Interestingly, however, we also find preliminary evidence that these seemingly intuitive metrics may not be those that people optimize for when solving a table top-rearrangement task. This work is currently under submission and published on arxiv [225].

1.1.6 Continual personalization using naturalistic corrections during surface rearrangement tasks

While Chapter 6 introduced a method for personalizing large foundation models, we found that these methods still did not match people's preferences exactly. To account for this, we introduce a method for continued personalization from naturalistic behaviors in Chapter 7. To address these shortcomings, we use insights from our earlier work: that goal-directed in situ naturalistic behaviors can be used to specify goals in surface rearrangement tasks. While previously our aim was to use eye gaze and joystick controls to solve a simple surface rearrangement task, our aim now is to use corrective actions to refine a generalized feature space. We develop a simulated surface rearrangement task in which the robot's goal is to place objects into a dishwasher according to an unknown user preference. The person can communicate this preference during task execution by replacing incorrectly placed objects into their correct location, i.e. take a corrective action. To make use of both the population level model and these corrective actions, the robot policy is first trained at a population level with an architectural bottleneck that forces the final layer of the model to be an online linear regression. This way, the model starts with information gained from the pre-trained feature space to initialize a linear regression model which can then be refined online using the information gained from the corrective actions to perform a belief update over the linear regression model. We find this to be an effective method for fine-tuning a pre-trained feature space in simulation. This work was drawn from our publication in The International Conference on Autonomous Agents and Multiagent Systems [226].

1.1.7 Conclusion

Finally, we conclude our work in Chapter 8. With these five main contributions, we show how developing assistive robots can be viewed as a value alignment problem that can be solved by using naturally occurring goal-directed behaviors. We show this in the context of both simple and complex surface rearrangement tasks, and introduce several naturalistic datasets for further study into developing general purpose robots that can assist in completing household tasks by aligning to individual preferences and adapting to changing preferences.

2 Background and related work

Before presenting our investigations into whether we can use naturalistic behaviors to provide value aligned robotic assistance, we some relevant work and background material. We begin this discussion with an overview of various approaches for using naturalistic human behaviors during human robot collaborations. This is followed by an overview of rearrangement tasks in robotics and embodied AI, as well as a formalization of household tasks as rearrangement tasks. We then cover topics related to cooperative inverse reinforcement learning in collaborative rearrangement, and formalize the rearrangement problem as a cooperative IRL problem. Finally, we cover various approaches for featurizing naturalistic behaviors for use in robotics and collaboration.

2.1 Human behaviors during human robot collaborations

Our main thesis statement revolves around whether or not we can use naturalistic human behaviors to provide value-aligned robotic assistance. Toward this end, we begin our coverage of related work with a discussion over how naturalistic behaviors have been used for robotic control. In this section, we aim to give the reader an overview of previous approaches for incorporating naturalistic behavior into robotics and what types of behavior are typically used. Additionally, we discuss how researchers have approached collecting large data sets of naturalistic behaviors in various subfields of robotics. Finally, naturalistic behaviors can be noisy signals. Due to this, it is often necessary for researchers to develop explicit featurizations of these behaviors in order to make them useful for robotics algorithms. In the final subsection, we provide an overview of some of these techniques.

2.1.1 Naturalistic human behaviors for robotic control

Eye gaze, EMG, and body pose have all been useful signals for robotic control. Since eye gaze is a rich signifier of intention during manipulation, both by hand [116, 141, 165] and by robot [19], its use has been explored through numerous robotic collaboration settings, including anticipating which object a user will request [127], and triggering assistive aid during autonomous driving [41]. Electromyography signals have been used for robot control [20] and task monitoring [79].

There has also been work on learning and leveraging human policies (using keyboard input) [256, 257] and attention models (using keyboard input and eye gaze) [356] for assisted and shared robot control in Atari games in an arcade learning environment [31].

Additional behaviors used to specify user intent during robot control tasks include verbal interaction [44, 316], joystick input [137, 138], or even calculated hand and arm movements that show a robot how to complete a manipulation task [155]. While these signals vary drastically in their modalities, all of the control signals require calculated and deliberate actions from the participant.

2.1.2 Multimodal human behavior data sets

Data sets from human robot interaction

Multimodal data sets have attracted interest in many different communities, such as psychology [346], computer vision [75, 95, 241, 290, 292, 293], human-robot interaction [21, 33, 139, 286, 303], and natural language processing [26]. These data sets, though, can be difficult to collect at a large scale. This can be due to the increasing engineering demand required with each additionally desired modality, physically collocating robots and humans, and the need to respect humans' privacy rights. This leads to many multimodal data sets that include either few participants or few data modalities. In addition, these data sets are rarely designed to study direct, physical human-robot collaborations in which the human and robot act in similar spaces. HARMONIC gives researchers the opportunity to study direct human robot collaboration in the form of a large scale data set in both the number of available modalities as well as the number of participants. Here, we compare how HARMONIC relates to other multimodal human-robot interaction in order to illustrate these distinctions and the potential use of HARMONIC.

Robots in Conversational Settings. The majority of publicly released HRI data sets study the inclusion of robots as conversational partners. To successfully incorporate robots as part of a social conversation, it is necessary to perceive human behavior, understand how this relates to the conversation, and be able to synthesize similar behavior to keep the conversation flowing smoothly. Much of this work surrounds determining the human's visual focus of attention (VFOA) [139, 286]. In these works, VFOA is a discrete representation of eye gaze estimated from the user's head position. Other data sets are designed to capture unscripted conversations with a robot [33] by capturing conversations through a robot's third person video recorder. In all of these works, no signal-specific sensors (e.g. an eye gaze camera) were used in order to capture specific human behaviors (e.g. eye gaze).

Other conversational data sets have a linguistic focus [26]. This work designs an interaction in which a human commands a robot to perform a specific task, and contains many different views of the language spoken. Due to the focus on verbal communication, this data set does not give researchers the ability to understand how nonverbal behaviors may be utilized in order to understand the intent behind the human's command.

Finally, perhaps the most similar data set to HARMONIC (in terms of data streams collected) again focuses on predicting VFOA during a conversation [303]. Unlike other works that focus on VFOA, this data set explicitly captures eye gaze using a Tobii eye tracker [320]. Thus, this work studies how gaze changes between structured and unstructured conversation with and without the presence of a robot. This data set is not designed, however, to study these nonverbal behaviors in physical collaborations.

In all of these situations, the behaviors collected for analysis are centered around noncollaborative tasks. HARMONIC provides an opportunity to study how these behaviors can be interpreted to provide better assistance during collaborative tasks.

Robot as Student. The Multimodal Human-Robot Interaction Dataset [21] is designed for interactive object learning through human guidance. This data set presents a situation in which a human uses a small number of task specific behaviors in order to teach the robot about object models. This data set is intended to instruct the robot by leveraging a human's innate teaching ability, as opposed to studying physical human-robot collaboration.

Humans teaching robots has also been studied in psychology [346]. Here, researchers studied how humans' eye gaze patterns changed as a robot displayed gaze patterns that were designed to emulate the gaze patterns displayed by people who employ different styles of learning. Again, this task is different from a direct collaboration, such as our shared autonomy task. Furthermore, this data set does not appear to be publicly available.

Data sets from machine learning and computer vision

Surprisingly, our data set is similar to those from the machine learning and computer vision communities. The tasks studied in these data sets often include non-scripted, egocentric videos of daily activities [75], gaze prediction for egocentric videos [95], action recognition in third person video [241, 293], relating first and third person videos as a proxy for theory of mind [292], and learning about human social affordance from a third person view [290]. These data sets include large amounts of potentially relevant data for human-robot collaboration, but most importantly, they do not contain interactions with a robot. While these data sets may be useful for an initial understanding of human behavior, they do not provide insights into how these behaviors manifest in human-robot collaborations.

2.1.3 Explicit featurizations of human behaviors for assisted teleoperation

Previous work has also modeled unimodal and multimodal primitives in order to aid in human robot co-manipulation. Uni-modal approaches have focused on recording human behavior, generating primitives from these recordings, implementing them as robot actions, and then testing by having a human complete the task with the robot [32, 191]. Multimodal approaches have focused on generating primitives from recordings of various views of the same or similar direct control input (*e.g.*, EMG and human arm manipulability) in periodic tasks (*e.g.*, sawing a board with a robot) [237, 238]. Regardless of the modality, these approaches focus solely on the direct control inputs.

Although these approaches have shown success, they neglect other signals that people naturally display while engaging with the world, with no added burden to the user. For example, previous work has shown that humans naturally elicit a wide range of non-verbal behaviors [40, 141] when performing collaborative tasks. In particular, eye gaze is closely related to hand movements, especially during manipulation tasks. When reaching for an object, the gaze to that object typically precedes the hand motion by about 600ms [165]. Gaze typically moves to the next object before the hand reaches its target [141], and gaze rarely rests on objects that are not involved in the current task [116].

Nonverbal behaviors have previously been used for direct control in remote robot navigation [166], drone teleoperation [352], and human-robot comanipulation in a table carrying task [8]. However, these approaches do not consider using naturalistic gaze as an indirect and supplementary control method, as we do in this work.

In human robot co-manipulation, prior work has begun to characterize hand-eye coordination when operating a robot under shared autonomy [19]. Although this work characterized many important interactions, it did not provide formal primitives or exhaustively analyze the relationships between joystick and eye gaze signals in a data-driven manner.

2.2 Household tasks as object rearrangement problems

We are particularly interested in providing assistance during household tasks, such as loading a dishwasher, setting a table, packing for a trip, or cleaning a room. Although these tasks are diverse in the types of objects involved and how each object is used, they all have a similar structure: each is essentially a generalized pick-and-place task. This means that these tasks can easily be formulated as *rearrangement* problems [27]. We first provide an overview of various approaches to rearrangement problems in robotics and embodied AI, followed by a formalization of collaborative rearrangement, which incorporates unspecified goals and continual specification through human input.

2.2.1 Object rearrangement in robotics and embodied AI

Recent advances in embodied AI have led to a flurry of benchmarks where an embodied agent is tasked with rearranging objects in a real or simulated home environment [27, 87, 103, 219, 222, 246, 311, 336]. However, these benchmarks have specified goal states, as compared to our work, where the preferred object arrangement has to be inferred from context examples.

Another line of work, focused on object rearrangement, in which target objects are specified by pointing gestures [255], eye gaze [220], or target layouts [348] during task execution. However, these works require manually specified target location for every object and are hence human effort intensive. A line of follow-up work addresses this issue by modeling preferences using learned priors about where objects are typically placed [146, 271, 314]. However, these preferences are generic and not personalized.

Previous work on personalized object rearrangement relies on simulated or large crowd-sourced data sets of human preferences to learn fixed latent preference vectors [147] or latent preferences that can be adapted online [226], model spatial relationships [145] or perform collaborative filtering [2]. In contrast, our approach leverages in-context learning with large-scale pre-trained VLMs to perform personalized object rearrangement. While ours is not the first work which leverages foundational models to perform few shot personalized rearrangement, we are first ones to perform fine-grained preference alignment, i.e, spatial preferences. As compared to prior works [343], which operate over a discrete preference space, that is, identifying the correct receptacle for each object.

2.2.2 Formalizing collaborative rearrangement

To study assistive collaborations, we introduce *assistive surface rearrangement*. As discussed in Chapter 1, a typical rearrangement task assumes a solution set S^*, an initial state s^0, a goal specification function $g = \phi(s^0, S^*)$, and a single goal-conditioned agent $\pi(a^t|s^t, g)$ who moves takes action to advance the initial state s^0 to a goal state $s^* \in S^*$ by taking actions a^t.

In assistive surface rearrangement, we assume there is an additional, human, agent π_h that can optionally take actions in the environment to advance the initial state to the goal state. Furthermore, the solution set and goal specification function are determined by the human agent and can only be observed by π_r through actions taken by π_h. This means that to advance towards s^*, π_r must infer the goal, the goal specification function, the solution set or all three from the actions taken by π_h.

In Chapters 4 and 5, we consider a surface rearrangement problem where the solution set known a priori, and consists of all 6 DOF robot positions x from which a goal $g \in G$ is reachable. The goal specification function is given by the human, who chooses a particular g in advance of the task. The robot must estimate g using observations of the person's behavior through actions taken on a joystick.

In Chapter 6 and 7 we consider a rearrangement problem where the solution set, goal, and goal specification function are all known only to the person in advance of the task and must be inferred (implicitly or explicitly) by π_r. The state $s = [O, L, o, l]$ consists of a set of objects that can be placed O, the objects that have been placed $o_i \in O$, the locations in which an object can be placed L, and the locations at which an object has been placed $l_i \in L$, where $l \in \mathcal{R}^2$.

In Chapter 6, the state is represented as images and language. In Chapter 7, the state is represented symbolically. In both chapters, a robot policy, parameterized by a neural network, must estimate s^* by observing prior demonstrations of the task as completed by π_h. In Chapter 7, we also consider observations of actions that π_h takes during the execution of the task.

2.3 Cooperative inverse reinforcement learning for collaborative rearrangement

After formalizing household tasks as collaborative rearrangement tasks, we still need a method for using naturalistic behaviors to specify people's goals. For this, we use the idea that people's *in-situ* behaviors are goal-oriented [22]. This means that we can use the behaviors people exhibit while they are performing a task to infer their goal. We provide an overview of how previous work has approached this and then formalize solutions to collaborative rearrangement as a solution to a cooperative inverse reinforcement learning problem [113].

2.3.1 Naturalistic behaviors for adaptive collaborations

Using inverse reinforcement learning (IRL) for robot control can be difficult, in part, due to the ambiguity that arises from traditional IRL [1]. Maximum entropy IRL facilitates this by using the principle of maximum entropy to order solutions according to how well they match observed user behavior [361]. This solution has also been used in behavioral science to model people's ability to infer others' goals from their behavior, as exhibited during goal-directed plans [22].

These insights have been applied to robot trajectory optimization for shared control. In the difficult task of teleoperating a high-degree of freedom robot arm with a low-degree of freedom input device, such as a joystick, a robot can observe user input commands and infer the user's most likely goal from a set of predetermined goals. The robot then assists the user by moving along a path toward the predicted goal [138]. MaxEntIRL can also be used to interpret less direct forms of user behavior, such as physically pushing a robot out of the way to determine which path the user prefers the robot to take, for example to carry a coffee mug around a laptop computer instead of over it [185], using naturalistic eye gaze in combination with joystick signals to control a robot arm [18, 19, 222], or using corrective actions to learn about features of the environment that relate to a person's preference to increase generalizability and sample efficiency [223]. We are interested in adapting online MaxEntIRL for determining high-level task plans consistent with user preferences in household collaborations from in-task corrective behavior.

IRL has also been applied to learn robot policies in other types of human-robot interaction. For example, to learn people's preferences from observations of independent task demonstrations [342], or by learning assistive social actions for therapy by combining therapists' expertise with expert demonstrations [15], or for social health, such as a robot receptionist learning to give hygiene advice in a shopping mall [65]. Our formulation learns preferences from *in situ*, collaborative behavior for collaborative rearrangement tasks.

Finally, another important aspect of maintaining human-agent assistive collaborations is maintaining collaborative fluency [122]. Maintaining principles of collaborative fluency, such as minimizing agent and human idle time, allows human-agent collaborations to function similarly to human-human collaborations, thereby reducing friction on people to interact with autonomous agents. Furthermore, robots that help people complete collaborative tasks have been shown to affect a person's ultimate decision [220], making it important to continuously monitor and assess people's goals during collaboration. In this work, we will use these ideas as justification for our desire to develop an algorithm that adapts to user preferences in real-time.

2.3.2 Solving collaborative object rearrangement with cooperative inverse reinforcement learning

We formalize the task of surface rearrangement, a specific instance of rearrangement problems [27, 311], as a coorperative inverse reinforcement learning problem [113], which consists, in our case, of a decentralized partially observable Markov decision problem (DEC-POMDP) which is a tuple of $(S, \Pi, A, T, Z, O, r, \gamma)$. Our objective is to find a policy π_r that solves this DEC-POMDP:

- S is the set of all possible states. As in prior work [185], we assume that a particular state $s \in S$ is a tuple of observable and unobservable features: $s = (x, g)$.
- Π is the set of agents. In our initial version of this problem, we assume two agents: a human agent and an assistive agent.
- A_i is the set of actions for a particular agent a_i. We assume that the person both selects objects and corrects object placements, while the robot can only make object placements.
- Z_i is the set of observations used to infer g. The assistive agent's observation space is the person's action space. In this work we assume that the human does not infer the robot's preference.
- $T(s^{t-1}, \mathbf{a}^{t-1}, s^t)$ denotes the transition dynamics that model the probability of entering a particular state given the current state and both agents' actions. As in prior work [185], changes in T are dictated by g. We assume this to be constant and deterministic within a single episode.
- $O_i(s^{t+1}, u_i^t, z^{t+1})$, the observation distribution for agent π_i.

- $r_i(s^t, \{a_i\}^t)$ is the reward function for the each agent. We assume an assistive setting where the agent is trying to estimate and maximize the person's reward function. We therefore assume all agents have the same reward function.
- γ, a discounting factor.

In Chapter 4, the robot estimates the user's goal (as represented by one of three marshmallows on a plate) using observations of a user's joystick input. In Chapter 7, the robot estimates the user's goal (as represented by a completed symbolic arrangement) using observations of corrective actions the user takes during the completion of the task.

Given that we assume two agents and that we are only optimizing one (because the other is assumed to be a person over whose policy we have no control), this problem reduces to a single-agent problem, allowing it to be decomposed to a POMDP. Since POMDPs are computationally intractable to solve exactly, we use the QMDP approximation [178]. Prior work in online human robot collaboration [185] has shown how a QMDP can be solved online using online gradient descent, adapted for our purpose in Alg. 3, where ψ is a state-feauturizing function that maps states to an arbitrary representation.

2.4 Implicit featurizations for collaborative rearrangement

With a formalization of collaborative rearrangement tasks, as well as a formalized solution to these tasks, we still need a method to featurize naturalistic behaviors in order to make them consumable by robotics algorithms. Previously, we discussed explicit representations of these behaviors, but these representations typically require a lot of domain knowledge about each behavior signal and can be difficult to generalize across different behaviors and tasks. Recently, generalized pretraining of large-scale models has provided an opportunity to experiment with generalized representations of images and language in robotics. Our aim is to explore these models to have generalized representations of naturalistic human behaviors in collaborative rearrangement tasks. We first begin by discussing how such models, often called foundation models, have been used in robotics. We would specifically like to encode behaviors as images, so we follow this section with a discussion on how images are specifically used to prompt multimodal foundation models. Finally, we discuss how these models and similar models have been used in collaborations between humans and robots or embodied agents.

2.4.1 Foundation models for robotics

VLMs pre-trained on large-scale datasets have shown commonsense reasoning abilities. Researchers have used these abilities to perform robotics planning and control [97, 125]. Many prior works [10, 128, 129, 176, 179, 182, 224, 254, 294, 300, 333, 355] have used pre-trained LLMs to generate actionable natural language plans for robots. VLMs have also been used to generate sub-goals for navigation [59, 83, 101, 126, 279, 280] and manipulation [74, 288] tasks. Additionally, previous work has also used LLMs to directly generate low-level executable policy code for robots [174, 296]. Another line of work has also used LLMs to generate rewards, which can be for RL [130, 189, 353]. In our work, we use a VLM to generate the policy code to accomplish a continuous preference-aligned novel goal state.

2.4.2 Visual prompting in vision language models

The development of in-context learning for few shot adaptation of LLMs [47], was followed by a flurry of prompt optimization approaches. While one line of work focuses on prompt-tuning [170] or prefix-tuning [171] through backpropagation using numerical gradients, others rely on the generated answer scores to perform optimization with textual gradients [245], prompt search using genetic algorithms [345], or iterative prompt refinement [349]. More recently, approaches have tried to use visual prompting to improve VLM performance by leveraging a VLM's ability to solve multiple choice problems. [180, 217]. Our approach adapts these methods to a complex visual preference learning task using these insights to prompt VLMs to combine information from multiple images to reason about a person's preference.

2.4.3 State and action-conditioned collaboration

Prior approaches to solving long-horizon tasks with complex temporal dependencies and under specified solutions, such as those present in our surface rearrangement domain, can rely on resolving ambiguities through a combination of teleoperation and preprogrammed routines [70], or by suggesting optimal, predetermined solutions [219]. Solutions following the former method can place undue burden on a person to explicitly express their preferences, and render robot action redundant when a demonstration completes. They also require people to continually demonstrate their desired solution as the constraints of the task, such as a person's preference or the environment, vary. Methods following the latter example do not allow for full freedom of expression from the user and assume that all users have the same "optimal" solution.

Zero-shot coordination is a recent field of research that aims to develop models that can successfully and immediately interact with novel partners. This can be done by pretraining models in simulation against agents designed to mimic human behavior [56] or over a diverse population of simulated agents [308]. However, using these methods can lead to overly specific solutions. Others have used large language models trained with web-scale data to propose task plans that are then executed by robots [10]. These task plans are not adapted to an individual user, whose reward function may or may not fit well within the distribution seen during training. These methods place the burden on the person to either accept a less preferred robot behavior or continue to provide actions that increase the likelihood of the robot behavior exhibiting behavior in line with the person's preferences. In this work, we focus on combining these good initializations with online adaptation.

Often, approaches that are based solely on deep neural networks that have been trained specifically require people to generate explicit descriptions of their preferences that can be decoded by the model into robot action [10]. Actions produced from this process are not guaranteed to align with a person's task objective. Although deep networks can potentially be adapted to individual preferences through fine-tuning [63, 118], doing so with large models can lead to challenging and unstable learning that results in variable performance [208]. In this work, we focus on developing an algorithm that can quickly adapt to people's naturally expressed, task-oriented behavior.

3 Components of robotic assistance: people, space, and time

Before developing methods to provide value aligned robotic assistance using naturalistic human behaviors, it is important to understand robotic assistance. To understand this, it is critical to view assistive robotics as a system with multiple components that is designed to support a person in achieving a goal they want to achieve. In this section, we lay out the components that are involved in an assistive robotics system and explain how assistive robotics can be viewed as a value alignment problem. We support these definitions with a review of recent research in assistive robotics.

Smart wheelchairs that navigate easily through crowded rooms, robots that train older adults through stroke rehabilitation exercises, and robotic arms that aid motor-impaired individuals in eating a meal in a restaurant are examples of research in diverse fields: intelligent motion planning, rehabilitative medicine, and robotic manipulation that have been identified as contributing to the development of robots capable of performing helpful tasks for people. This research has been fruitful, but has remained siloed as researchers from these various fields focus on the specific assistive tasks relevant to their own disciplines.

The lack of a common structure in the field of assistive robotics makes it difficult for researchers to incorporate findings from other domains into their own work. For example, how does the relationship between a grocery stocking robot and the surrounding customers relate to the relationship between an airport guide robot and the surrounding crowd? Does a robot designed to autonomously declutter a room convey a similar sense of agency as a virtual robot suggesting an optimal ordering in which you should clean your room? Answers to these and similar questions would form a basis that would provide clarity for research in assistive robotics, but are currently difficult to determine due to the disparate nature of assistive robotics.

In this work, we identify a subset of common challenges and develop themes that begin a conversation about how assistance can be abstracted out of specific problem domains and can be used to answer questions about assistance in general, thus benefiting the entire field of *assistive robotics*. This would enable researchers to explore the underlying principles of assistive robotics and communicate them across domains. To start, we suggest that assistance is not a characteristic of a robotic system as it has been historically treated. Instead, assistance is a task-independent perspective on human robot interaction. Treating assistance as a task-independent perspective on HRI, we can group existing assistive research by its effect on three key axes: people (e.g. who is involved in the system and the roles they play), space (e.g. how the robot's action affects the task), and time (e.g. when the robot performs its actions during the task).

This perspective considers an assistive system as an interaction in which a user and a robot forge a complex, asymmetric relationship guided by the user's goals. This perspective is somewhat different from general HRI because the user is responsible for determining the interaction's end goal while the robot acts in service of this goal. Similar to other collaborative settings, the human-robot pair is then tasked with performing subsequent actions to achieve the human's goal, but unlike in some collaborations, maintaining human autonomy is paramount. In this relationship, the robot has more agency and independence of action choice than a simple tool (i.e. the robot's choice of action is not determined solely by the user), but it must defer to the user's goal and independent actions.

We introduce three design dimensions with which roboticists can begin to reason about the assistive interactions of robots and humans. First, we discuss how the assistive robot's role can be described with respect to the relationship it has with its user, for example, how it weighs priorities when there are multiple potential people it could assist. Second, we propose that an assistive robot's role can be described in terms of how it operates in the execution space, that is, the space in which the robot has its primary effect. Finally, we propose that the same robot's actions can be described in terms of the temporal space, that is, the duration and sequence of the actions. We support these dimensions by reviewing and grouping over 200 recent assistive robotics research papers.

By using assistance as a lens through which to analyze patterns that arise in assistive robotics, we hope to help designers of assistive robots more easily explore the design space and identify similar examples of past solutions, even across application domains. Additionally, we hope this work will motivate researchers to continue to refine this notion of assistance and its effects on human-robot interaction paradigms.

3.1 The assistance perspective

In the field of robotics, defining assistance can be tricky: in a broad sense, every robot is built to assist some person. Therefore, we do not attempt to separate assistive systems from non-assistive systems. Instead, we propose assistance as a particular *perspective* through which many robotic systems can be viewed. This perspective considers robotic agents that are *autonomous* in action but *subordinate* in goal to a human partner. Almost any robot system can in theory be viewed as assistive to someone, so we do not limit this scope. Rather, we explore what this analytic framework provides. This perspective clarifies particular design tradeoffs and trends general to assistive systems whatever their task domain. In this work, we describe several key design axes that arise when considering a robotic system as assistive and discuss implications these axes have on the interaction.

Before discussing these key design axes, we first formalize what we mean by a human-robot interaction, then provide a more detailed description of what it means to view assistance as a perspective. Next, we give a brief synopsis of previous attempts to characterize assistance and assistive robotics, and finally we give an overview of the remainder of this paper.

3.1.1 General human-robot interaction

Before discussing assistance, we first sketch a general framework for human-robot interaction, which we draw broadly from multi-agent systems research. Formalizations of this problem can be found in previous literature [136]; here we only establish enough language to discuss assistance rather than requiring assistive systems to use this exact model.

First, we define a user $u \in U$ as any person involved closely in the interaction. Typically, the user is in close physical proximity to the robot and provides explicit or implicit control signals to the robot. For example, a person teleoperating a robotic arm, getting directions from a social robot, or building a table with a robot helper would be considered a user.

Next, the system has at least one robot $r \in R$. Canonically, a robot is defined as an embodied system that can sense its environment, plan in response to those sensory inputs, and act on its environment. An assistive robot may have a wide array of sensory, planning, and acting capabilities in order to be successful in its task. Some of these capabilities will be critical for the robot's functioning (e.g., LIDAR to avoid hitting obstacles), while others will be critical for providing assistance to the user (e.g., a body pose recognition algorithm to identify the user's location and gestures).

Finally these agents exist in a shared environment, each with their own internal state. These are described in totality by the mutual state $s_m = (s_r, s_u, s_e)$ that defines the individual states of the robot, user, and environment. The robot and user both have goals $g_r, g_u \in G$ and can take actions $a_r \in A_r$ and $a_u \in A_u$ that affect their mutual state. By acting to update their mutual state, each agent has the potential to affect the other agent's behavior resulting in an interaction between the two agents. Depending on the exact scenario, a task will be considered complete when one or more agents has achieved their goal.

3.1.2 Assistance as a perspective on human-robot interaction

Using this formulation, we can more carefully define assistance. Assistive systems interpret the robot as autonomous in its actions but subordinate in its goal. By giving the user the sole responsibility for setting both agents' goals, the two agents now attempt to satisfy some shared goal g by reaching a mutual state where g is true: s_m^g. This framing distinguishes assistive robotics from both traditional assistive technologies like a white cane, which has no control over its actions or goals, and traditional robotics, which develops agents with full control over their actions and goals. This framing gives rise to three key design axes: how assistive robots affect *people* through *space* and *time*. The discussion of these implications is the subject of the rest of this paper.

In HRI, as in assistive robotics, there is no requirement for there to be a single user. In fact, many assistive robotics scenarios involve more than one user. This becomes challenging, as it is the responsibility of one of these users to set the goal for the robot, but selecting which user has this responsibility may change the type of assistance the robot is able to provide. This is especially true when one user's goals may conflict with another user's goals. This highlights the importance of determining the roles of people when considering assistive robotics problems (Section 3.3).

Furthermore, since the user and robot are working towards the same goal, the robot has more freedom over its action space. Instead, the robot can assume that the user would perform the task independently as a baseline, and only then choose how best to provide assistance. In addition to manipulating the environment directly, the robot can assist by altering the user's state space, which helps the user to make more effective task progress. A head-mounted augmented reality device that displays the optimal path for cleaning a room can assist in the world without needing to physically interact with objects. The assistive scenario allows more choice over the robot's action space than would a general robot (Section 3.4).

Finally, in order to advance to the mutual goal state and complete the task, the user and robot complete each sequence of actions ($a_u^1, \ldots, a_u^t$, $a_r^1, \ldots, a_r^t$, respectively) that transition the system to the desired goal state ($s_m = s_m^g$). Given that these actions occur in the mutual state, it is important that the user and the robot time their actions appropriately, so that they do not attempt to solve the same part of the task simultaneously, or worse, provide conflicting actions that result in undoing each other's work. How to time actions is crucial to studying assistive robotics (Section 3.5).

Each of these axes presents researchers with decisions that result in critical trade-offs when designing an assistive robot. Throughout the remainder of this work, we will describe how assistive robots from different application domains fall along these axes.

By taking assistance as a perspective, it is our goal to provide an abstraction that allows for comparing systems from different domains to discover universal challenges that arise from robot assistance. We do not suggest that these axes describe a full assistive system or are a complete set of critical design axes. Rather, viewing assistance along these particular axes of people, space, and time enables some cross-domain comparisons and insights on its own, and it also demonstrates how assistance overall can benefit from a general examination.

3.1.3 Prior categorizations of assistive robotics

By grouping assistive robots along the aforementioned design axes, we view assistance as an abstract concept that illuminates parallel research problems across different application domains. We build on previous literature which categorizes assistive robotics within particular application domains, for example socially assistive robots [99, 199], joint action [132], and physically assistive robots [46].

Some work does try to describe assistance as a whole. Jarrassé et al. [136] categorizes joint action between dyads by positing a cost function for each agent defined on each agent's task error and required energy. Among categories in which both agents are working together towards the same goal, the paper specifies *collaboration* between two equal peers, *assistance* when one agent is subordinate to another, and *education* in which the educator assists the partner but moderates its own effort to encourage increasing effort from its partner. We take this core idea of assistance as subordination and build on it in our definition of the assistance perspective.

Most similar to the current work, perhaps, is the accounting given in Wandke [331]. This overview of assistance in human-computer interaction notes that defining assistance as any system that provides some benefit to the user would include nearly all technical artifacts. Therefore, the paper restricts its attention to systems that bridge the gap between a user and the technical capabilities of the system due to the user's unfamiliarity with the system or excessive burden of use. In contrast to this approach, our work presents assistance as a perspective rather than a definition; it could in principle be applied to any technical artifact but may only be useful for some. Additionally, this definition of assistance focuses on how assistive systems correct a deficiency in a user's understanding of the system or capability to use it. In contrast, our definition of assistance as a perspective admits beneficial actions from the robot of all sorts, not just those repairing the user's ability to use a system.

3.1.4 Overview of this paper

By defining assistance as a perspective, we provide language to discuss ideas about assistance from different domains. This will allow researchers from various areas of assistive robotics to come together to illuminate and discuss common research challenges. Additionally, researchers can make design decisions about how the assistive robot affects people in space and time by using this framework to consider similar approaches to problems from disparate task domains. In the remainder of this paper, we discuss these design axes and explore their implications through a review of existing assistive robotics literature. Section 3 describes our method for collecting these papersSection 4 describes the *people* design axis, Section 5 describes the *space* design axis, and Section 6 describes the *time* design axis. These axes are summarized in Table 3.1. We then conclude the paper with a discussion over the implications of this work.

3.2 Methods

To develop this taxonomy, we conducted a literature review of recent papers on assistive robotics.

Initial search. First, we hand-selected 74 papers from the last five years of the annual Human Robot Interaction conference (HRI 2016–2020). From these papers we generated an initial set of search terms by aggregating titles, abstracts, and author generated keywords using the `R` [248] package `litsearchr` [110]. Using these aggregated keywords, we formed an initial search query.

Refined search. We ran the initial search query on the Web of Science. This search yielded approximately 1500 papers. We repeated the keyword aggregation on this set of keywords, and then hand-selected new keywords from among them based on their prevalence and relevance to assistive robotics. We repeated the Web of Science query with this refined set of keywords, which yielded, again, approximately 1500 papers. The refined search was run on 29th January 2021. We included a paper based on whether this statement evaluated true based on a search of the entire text of the paper.

((assist* NEAR *robot*)

OR (collab* NEAR *robot*))

AND (*human* OR *people* OR *person* OR *subject* OR *user* OR ”elderly people” OR ”older adults” OR ”natural human” OR ”stroke patients” OR ”healthy subjects”)

AND (“human-robot interaction” OR “human-robot collaboration” OR “robot interaction” OR “robot collaboration” OR collaboration OR hri OR “human robot collaboration” OR “physical human-robot interaction” OR “human robot interaction” OR “machine interaction” OR “human-machine interaction” OR “human interaction”)

AND (“collaborat* task*” OR “assembly task*” OR “social interaction*” OR “assembly process*” OR “shared workspace*” OR “manipulation task*” OR “human safety” OR “daily living” OR “service *robot*” OR “production system*” OR “safety standard*” OR “mobile robot*” OR “assisted therap*” OR “collision avoidance” OR “object manipulation” OR “collaborative assembly” OR “socially assistive” OR “assistive *robot*” OR “social *robot*” OR “teleoperat*”))

Paper selection. Starting from the refined Web of Science results, we filtered out all papers from venues with fewer than two related documents and papers that were older than five years, with a small exception. In an attempt to keep papers with significant contributions to the field, papers older than five years were kept if they had more than 10 citations. This process left approximately 465 papers. Each paper in this set was then manually checked for relevance by reading the title and abstract. To be included, we required the paper to include both (1) an assistive interaction with the user and (2) a system capable of taking actions. This step mainly removed papers focused on robotic system development or perception improvements rather than assistance itself. This yielded 313 papers, each of which was again reviewed against the aforementioned exclusion criteria, removing papers that focused on systems development and the performance of perceptual systems over interaction. The entire search process yielded over 200 papers that we classified into our taxonomy.

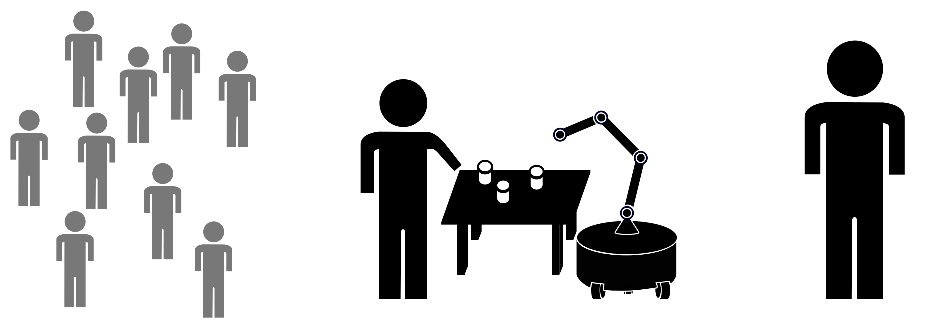


Figure 3.1: An assistive system can treat people beyond a single user as additional targets of assistance or as interactants, and either choice introduces particular complications into the assistive dynamic.

3.3 People

In Section 2, we described assistance with single users. This description works well for situations that have only one user, which is common in laboratory settings. In realistic settings, however, a robot will typically encounter more than one person in the course of their task. These other people can act in a variety of different roles within the interaction. In this section, we explore themes in how assistive interactions incorporate more people into the general human-robot dyad (Fig. 3.1).

3.3.1 Terminology

The simplest approach a system can take towards other people is simply to ignore them completely. While this case tends not to be analyzed explicitly, it is implicit in many systems. It can be appropriate to ignore other people, especially in situations for which additional people will be rare. Even when working with other people, the robot can expect its primary user to control the robot appropriately rather than explicitly considering other users. A robot might severely downplay its relationship with others when doing so conflicts with its primary user's goals, such as an emergency response robot that ignores standard social navigation behaviors to reach its patient as fast as possible.

When the system does choose to reason about other people, its treatment of them can be determined by dividing them into two different roles: the *target of assistance*, whose goals are of equivalent importance as other targets; and *interactants*, who require the attention owed to any other person as explored throughout human-robot interaction research but don't have their goals privileged by the robot.

A *target of assistance* derives directly from the definition of assistance: an assistive scenario must have a person whose goals to primarily support. Consider a scenario in which a person who has a spinal cord injury uses a robotic arm to aid them in eating a meal with friends out at a restaurant. In this scenario, the arm's user sets the goal for the robot: to bring food from their plate to their mouth so they can consume it. Assistive systems may have more than one target of assistance, which can lead to conflicts over what goals to achieve.

The second role a person can play in an interaction is that of *interactant*. An interactant is any other person involved in the scenario who is not a target. Continuing the previous example, the people who are out to dinner with their robot-operating friend are interactants. They have no direct bearing on the robot's goal, but they are potentially affected by the robot's actions and may require some design effort for the system. For example, the robot may have to avoid collisions with them during its operation. While the robot's relationship to interactants is not assistive, the presence of a specific target of assistance can affect how the robot interacts with others.

When considering assistive systems that involve more people than just a single target, the system must determine in which of these roles to consider the additional people. These two roles are not mutually exclusive and there can be more than one of each in a given scenario. Additionally, both targets of assistance and interactants can give explicit control input to the robot. Designating people as additional targets or as interactants bring about different challenges for the assistive system.

3.3.2 Additional targets of assistance

One condition that arises is that a single robot has multiple different targets, often in asymmetrical ways. In the eating scenario, the robot might instead be assisting everyone present, perhaps by both feeding its user and serving food to other people at the table. Here, the robot is presented with a conflict: how should it choose to prioritize the goals given by its targets and reconcile differences between them?

This can be especially challenging in contexts such as education. An educational robot might consider the teacher as its target and work to enrich a student according to a mandated curriculum. It can also consider the student as its target, and try to engage the student with concepts that are interesting to them regardless of the curriculum. Much research in this area aims to make the content proposed by the teacher more enjoyable by developing robotic behaviors that are meant to keep the student engaged. Leite et al. [169] designed a robot puppet show to engage young learners in an educational story, Martelaro et al. [197] designed a robot that encourages students to encourage trust and companionship with their tutor, and Christodoulou et al. [69] designed a robot to give nonverbal feedback to students in response to quiz answers to keep them engaged with the testing material. In contrast, Davison et al. [76] took a different approach and developed the KASPAR robot to look like another student and deployed it in unsupervised, interactions that were totally motivated by the student. In this way, they allowed the student to approach the learning material voluntarily, giving the student more agency to learn what they desired and at their own pace.

This dilemma can again be seen in therapeutic contexts, where a robot must reconcile the goals of the doctor and the patient. Robots can increase a patient's motivation to do mundane, repetitive or uncomfortable exercises through the use of a robot that does the exercise alongside the patient [276, 315]. Alternatively, a robot could be used to give the patient more agency and independence over their own treatment through a robot that helps someone independently practice meditation [12], do independent cognitive behavioral therapy [81], or home therapy for autism [285].

A full analysis of these interactions treats both the teacher and the student, or both the therapist and the patient, as targets of assistance with goals that often align but are not identical. This alignment mismatch can often lead to ethical challenges, which are even more fraught when the capabilities, agency, and relative power of the possible targets vary. While there is no general technical solution, this language encourages designers to explicitly enumerate the multiple targets of the assistance and to reason directly about conflicts in their goals.

3.3.3 Additional interactants

On the other end of the spectrum are robots that treat additional people in the system as interactants. Robots designed with this relationship in mind prioritize the goals of its target of assistance. In our assisted eating scenario, the robot may need to follow basic social norms around the other diners by avoiding collisions with them, but it does not privilege their goals in the same way.

This relationship is typically used in scenarios where some figure of authority (e.g. a teacher or a therapist) needs to relieve themselves of some amount of work. For example, a teacher could employ a robot to teach half of their class in order to reduce the student-to-teacher ratio for a particular lesson [263], or even have the robot teach the class alone if they need to finish other work [242]. In this way, the teacher is the target of assistance, while the students are treated only as interactants. The robot should be able to teach competently enough to achieve the teacher's goals, but the students' preferences about using the robot are not of direct concern.

Similarly in emotional or physical therapy a robot can be employed to lead group sessions in lieu of a doctor, who may have more classes than they can handle [92, 134]. Alternatively, the robot may be better at collecting certain information than the robot. For example a patient who has suffered a stroke may be unable to emit certain social signals expected during social interaction. This could negatively affect a doctor's opinion of this patient, a problem that could be circumvented by having a robot collect this information [43, 326]. The patient here, however, is not asked whether they may prefer the social interaction regardless of the implicit bias the doctor may possess.

These systems don't generally follow an assistance dynamic with interactants. Rather, general human-robot interaction research applies. However, the fact that the system has a target, even if the target is not present, can change the robot's behavior: a robot acting as a proxy for a specific teacher may have different behavior than one employed as a general-purpose robot, which might have bearing on how the general human-robot interaction problem is resolved.

3.3.4 Combinations of roles

If an assistive robot has multiple additional people present in the interaction, it can choose to consider some of them as targets and others as interactants. In this relationship, our assisted eating robot might treat both the user and the companion seated next to them as targets of assistance, while those eating companions seated further away from the user are treated as interactants. In this way the robot can carefully maintain the goals of multiple people in proximity to the robot. This framework can allow for more complex robot behavior near to the user without the additional complication of handling everyone else at the table.

Another example would be a robot that participates in a collaborative scenario with multiple human actors, some of whom serve as both targets of assistance and interactants, while others are only interactants. For example, consider a local repair-person who needs help from a remote repair-person. To give instructions, the remote repair-person can use a robot to highlight the parts of the environment they are discussing [190]. In this way, both actors are interactants in the scenario, but only the local repair-person is the target of assistance.

3.3.5 Implications

These various relationships clarify the design choices involved in developing an assistive system. A particular task, such as assistive eating, does not require a particular relationship between the robot and the people it encounters. Rather, how a robot relates to these people is a design decision that will have implications on the task.

The choice of roles affects how assistive systems with multiple people are evaluated. When treating the user and their eating companions all as targets of assistance, the robot would need to verify that it is helping them all in achieving their independent goals. This type of evaluation may be difficult to actually measure and nearly impossible to succeed on, as the companions have conflicting interests from the user. Identifying what type of relationship the robot should have with its users can help researchers disambiguate otherwise similar systems to determine which evaluations are important.

The choice of which roles to use may also have implications on how much autonomy to imbue in the robot. A robot that balances the goals of many people may require complex sensing, modeling, and planning to carefully moderate between them. A simpler robot might delegate this goal moderation problem to its user and treat additional people as interactants or ignore them entirely. This system gives the target more control over the goals, but requires additional input from the user. If the robot maintains full autonomy in this scenario, but it does not plan for other people's goals, it may in fact endanger them by running into them where another system would have chosen to avoid them. These ideas show how the choice of relationship between the robot and the people it encounters throughout a task can impact the design of the final system.

3.4 Space

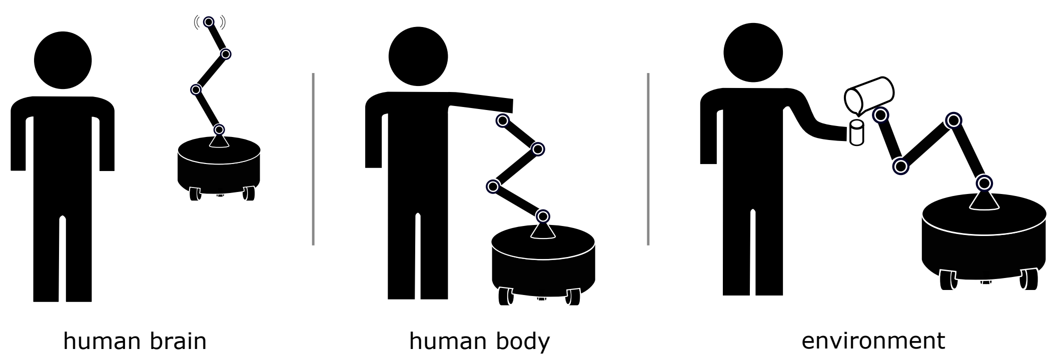


Figure 3.2: A robot can provide assistance by acting in several different action spaces. It can assist by giving information to the user, adjusting the user's body, or changing the environment to help complete the task.

Assistive robotic systems can perform similar tasks by acting in different action spaces. We show in Sec. 3.1 how to represent the mutual state during the interaction as the state of the user s_u, the state of the robot s_r, and the state of the environment s_e. In general, assistive robots help their users perform an action in the environment. Since the robot is fundamentally assisting a user, the same overall task can be assisted by a robot acting in a variety of different ways. In this section, we categorize the action space of assistive robots into the user's mind, the user's body, and the environment (Fig. 3.2).

Consider an assistive eating robot. The robot and its user sit at a table across from one another, with a plate of food between them. The user's goal is to eat the food. The robot can provide assistance by performing a variety of different actions: it can act on the user's mental state by projecting a light onto a morsel of food that would be easy to grab next, it can change the physical state of the user by guiding their hand into an appropriate position, or it can change the environment by picking up the morsel and feeding it to the user. All of these action spaces apply to the same task and the same goal; what differs is in what way the user would most benefit from assistance.

To illustrate this point more broadly, we provide a review of recent assistive robotics literature, grouped by whether the robot is acting on the user's mind, user's body, or environment.

3.4.1 Environment

One straightforward assistive robot is one that simply completes a task for the user. For example, research has focused on autonomous butler robots [301, 302] that perform tasks such as cooking and cleaning. Such a robot assists a user by navigating around the apartment picking up misplaced items such as dirty laundry and dishes and placing them in appropriate locations such as a laundry hamper or dishwasher. The robot provides assistance by directly changing the environment. To meet the minimal requirement of providing assistance (i.e. delivering some benefit to the target of assistance), the robot must shift the environment from an undesirable state configuration to a more desirable one.

Much research surveyed here assists users in exactly this way: by providing autonomous assistance through environmental state manipulations. Researchers have explored how a user can command a robot to organize a messy room [71, 140, 158, 200, 243], fetch misplaced or distant items [127, 131, 323, 339], or even perform more specialized tasks autonomously (under the direction of the user) such as assisted eating [52] and other tasks of daily living [227], search and rescue [84], welding [13], or other industrial tasks [210]. Assistive tasks performed autonomously at the request of a user through environmental manipulation can provide several benefits. This method of task execution requires little user input, which makes it efficient for users who prefer not to spend time on chores and beneficial for users who may not be able to accomplish the task at all.

Environmental assistance is not solely the domain of autonomous robots, however. Collaborative robots, specifically in tasks where the user and the robot take independent actions that jointly manipulate the environment towards a mutual goal state, also perform environmental assistance. Examples of such systems include collaborative cleaning [80] and assembly [272, 358]. A robot working collaboratively with a user can improve its efficiency by modeling the user's behavior, for example by determining specific poses to hold an object in to facilitate fluid collaboration during assembly [11] or by anticipating and delivering the next required item in assembly [114, 115, 192] or cooking [157, 202], or by providing help under different initiative paradigms during assembly [24]. Collaborative environmental assistance can also be used to perform joint actions with a user, such as in handovers [45, 51, 53, 73, 107, 111, 162, 164, 218, 220, 249], where the goal is to transfer an object from the robot's end effector to the user's hand; or co-manipulation [78, 88, 106, 159, 230, 250, 252, 274, 275, 332], where the aim is for the user and the robot to jointly move an object to a specified location or provide redundancy in holding an object in a joint assembly task [234] or safety critical situation such as surgery [309].

So far, all examples of environmental assistance have been provided by standalone robots, commonly taking on a humanoid or robotic arm morphology. These robots affect the environment by changing their own configurations first (e.g., using a robot arm to pick up an object). As such, they are considered decoupled from the environment. Robots can also be designed to be coupled with the environment; in these examples, it is hard to distinguish between the robot's state and the environment state. These robots often take on more conspicuous yet specialized morphologies, such as a mechanical ottoman [298, 357]. For example, a robotic suitcase can assist an airline passenger by following them through an airport [96] and manipulating the user's sense of trust by moving across various proxemic boundaries. A set of robotic drawers containing tools can assist a user in completing an assembly by proactively opening the drawer containing the next required tool [205], and it can also manipulate a user's enjoyment in completing the task by employing emotional drawer opening strategies. Environmentally coupled robots can be designed to be "invisible," [298] or to be modifications to an existing environment or object. Moving away from more traditional robot appearances may mitigate any negative effects from interacting with a robot.

Other approaches include shared control which separates the responsibilities of the user and the robot during the task. For example a teleoperated surgery robot can hold a patient's skin taut so that the surgeon can focus on performing incisions [281]. A telepresence robot [160] can automatically avoid obstacles during navigation [3, 306] or automatically rotate its camera to keep a desired object within view [204]. Finally, a remote, teleoperated space robot can perform as much of a task as is possible before it pings the space station for human intervention [94]. By having the robot configure itself according to some of the task requirements, the robot allows the user to focus on other parts of the task.

3.4.2 Human body

While assistance applied directly to the environment can solve a wide variety of tasks, some tasks require alternate strategies. One such scenario is when some change to the user's physical state is required to perform the task. For example, consider a robot designed to assist a user who has difficulty bathing themselves. While it is technically possible for that robot to transform the environment by bringing a bathtub to the user, this is obviously impractical. The robot can instead transform the user's state by bringing them closer to the bathtub [82, 233]. This strategy of moving a user to assist them is similar to autonomous environmental manipulation, but now the user is being manipulated instead of the environment. This strategy results in limited agency to the user, and is typically only employed when the user has minimal ability to complete the task themselves.

In cases where users can perform some aspects of the task, a robot can also assist by supplementing a user's existing abilities. For example, if a user can walk but has difficulty balancing or navigating, a smart walker can be utilized to help the user navigate between locations [233, 291]. Similarly, if a user has some control over their limbs, an exoskeleton robot can be used to provide extra support for day-to-day usage [23, 67, 175, 214] or in therapeutic scenarios in order to help a user strengthen weakened muscles [55, 362].

In addition to aiding in task execution, physical user state manipulation can also be used to assist in planning, such as when a user's sensing capabilities are diminished. For example, a visually impaired user may wish to solve a Tangram puzzle but must pick up and feel each piece individually. To provide assistance to the user, a robot could sense the puzzle pieces and determine which pieces are viable for the next step of assembly. The robot can then physically guide the user's hand to this piece allowing the user to solve the puzzle [39]. This is an example of human body state manipulation. Instead of manipulating the environment to solve the task, the robot instead changes the user's physical state configuration in order to better position them to solve the task.

Robot assistance that acts on a user's body can also be done by using the resistance of the robot's own joints. A user kinesthetically manipulating a robot arm, for example, may not know the exact path the arm should travel in order to complete a co-manipulation task. The robot can change its admittance or transparency such that it becomes easier [135, 168, 172, 188, 196, 213] or more difficult [37, 49, 50, 163, 344] to move as the robot's end effector deviates from a known, low-cost path. This idea can also be applied to full-scale robots, allowing a user to navigate a robot from one point to another by guiding it as if it were another human [64] or to use the stiffness of the robot's arm as a support while standing up [133]. Admittance control as a body state manipulation allows the user to have a high degree of control when operating the robot, but allows the robot to provide information about which parts of the environment are better to traverse by altering the stiffness of its joints. This strategy can also be used in therapeutic settings, where a patient recovering from a stroke can be given an automatic, smooth schedule of rehabilitation exercises as the robot changes its admittance depending on the force feedback it receives from the user [134].

3.4.3 Human brain

The final location of assistance we identify is the user's mental state. These robots assist by transforming the user's understanding of the world in a helpful way. One common method is for the robot to communicate unknown environmental information to the user. For example, a robot can play particular sounds as it completes its tasks so that a user can track it more easily [57]. A robot can also describe the local environment for a visually impaired user in a navigation task, enabling them to create a semantic map of the environment [61]. Similarly, a robot can provide a visual signal to designate objects it intends to interact with so the user can avoid them [14, 190, 289], areas where the robot expects to move so the user can stay away [121] or areas or paths that the robot thinks the user should take to complete a task in an optimal fashion [219]. In an emergency scenario, a robot can visually indicate the direction of a safe exit [262]. Finally, a robot can provide haptic feedback to indicate when to turn in a navigation task [173, 207]. Robots that provide alerts like these assist by communicating information about the task or the environment directly to the user so that the user can effectively perform the task.

Robots can also assist in the mental state domain by adopting social roles. Generally, these robots are designed to perform socially beneficial functions similar to those that a human would provide, such as a robot that takes the role of a customer service agent [329] or a bingo game leader [186]. In educational settings such as one-on-one tutoring [100, 144, 150, 325] and classroom teaching [150, 232, 242, 253, 263, 337], a robot can deliver lectures in a similar manner to a human teacher. In therapeutic and medical settings, a robot can administer routine medical surveys [326] independent of the doctor's social biases [43], provide therapy sessions for routine cognitive behavioral therapy [81] or physical therapy [201], and perform other general therapeutic tasks [9, 12, 92, 268]. Finally, a robot's assistance can vary based on its social role, such as a concierge robot performing different social behaviors when responding to children or adults [212], an advice-giving robot providing explanations when a user's behaviors become non-optimal [104] or a robot that gives cooking advice varying its strategies so that the advice is more readily received [321].

Instead of performing a procedure itself, a robot can assist a professional when affecting a user's mental state. When a therapist is unable to be physically present with a child, for example, a parrot robot can be employed in the home to entice a child with autism to practice skills learned during a therapy session [35, 285]. During therapy with agitated patients, introducing a pet-like PARO robot can induce mental states more conducive to effective therapy [58, 266, 282, 287]. A child-like robot can allow a young patient to practice social skills with a partner more akin to a peer than the therapist is [4, 108, 152, 229, 313]. Similarly, a child-like robot can assist a teacher by reinforcing a students desire to self-engage in educational material, something students may be more likely to learn with a peer than a teacher [76, 341], or increase a user's ability to recall a story by acting out portions of it [169].

Since robot actions are sometimes interpreted socially and as being intentional, robots can select their actions to influence the user's mental state. For example, predictable and legible motion strategies that indirectly communicate a robot's goals are readily interpreted by people [86]. These same strategies can be used in collaborative tasks to indirectly show the robot's goal to the user [38, 93, 312, 360]. Robots can also mimic human nonverbal behaviors like deictic eye gaze and pointing gestures to indicate task relevant objects during collaborative tasks [42, 98] or to assist in completing mentally taxing tasks [7, 119].

Similarly, robots can use their behavior to suggest their internal emotional state. This strategy can increase rapport, fluidity and reception of a robot's assistance through emotive motions [205, 317] or giving the user feedback regarding a task's success through facial expressions [69, 251, 258]. Using socially meaningful actions enables assistive robots to communicate with the user efficiently and fluidly.

Robots can also use social behaviors to induce specific, beneficial emotional responses from a user. By mimicking human nonverbal behaviors, robots can use their eye gaze to induce social pressure on a user to work more efficiently [261] or to soften its own dominance to allow for better teamwork [239]. Assistive robotic gestures can also increase feelings of openness in people who are discussing negative experiences [123] and motivation in users during medical testing [322], in users during physical exercise [194, 195, 276], and in stroke patients performing rehabilitative exercises [315]. Since people generally view robotic gestures as intentional, robots can use these gestures to induce mental states that assist the user in performing a task.

In addition to nonverbal communication strategies, robots that are capable of speech can converse with users to induce beneficial mental states [154]. Robots can use speech to change the content of the conversation [102] or to answer a question about the surrounding environment [48]. Robots can use dialogue to gather information during collaborative teleoperation [99], to engender trust in an escape room [105], or to facilitate collaboration between two targets of assistance [307]. Robots can also talk about themselves to influence a user's view of themselves. For example, tutoring robots for children can make vulnerable statements about themselves to increase trust with the student and student engagement [197]. Similarly, a robot in a group setting can facilitate group trust by leading with vulnerable statements about itself, so that its teammates feel more comfortable sharing their own vulnerabilities. This effect can cascade as more group members explain their own failures, console each other, and laugh together [277]. Failing to deliver assistance in contexts where the robot is expected to provide assistance can have deleterious effects on a user's mental state, causing users to mistrust the robot negatively their relationship and rapport [156, 265].

Beyond focusing on specific content of speech, conversational robots can further affect the user's mental state in the way they speak. Robots can perform back-channelling to give the appearance of active listening [36, 278], or give informative feedback to improve task performance [112, 167, 284], a user's self-efficacy [354], or their motivation [209, 283]. Robots can choose to only interrupt a distracted user at appropriate times [297, 324]. A robot can also change its tone to project an emotion such as happiness to improve the user's mood and task performance [187, 198, 259, 340]. Finally, a robot can combine these qualities with the content of the conversation to change the user's perception of the robot's social role [25, 34, 206]. Specifically, a robot can act as a student during a tutoring session to induce different learning techniques in a human student [269].

Shared control, especially when an input controller (e.g. a joystick) limits the number of input degrees of freedom [19], can also be made easier for user's by providing assistance that alters the user's mental state. A robot arm can assist its user by maintaining more easily controllable state configurations [19, 137, 222, 319, 330] or by optimizing which degrees of freedom the user can control at any given time [120]. This idea can be extended to supernumerary arms that provide users with an additional appendage but are difficult to control [216, 328], teleoperating robotic arms through electromyography [231, 240] or similar sensing devices [211], or humanoid robots [177, 359]. Additionally, a robot might be able to enter environments that are unavailable to a user allowing the user to teleoperate the robot in these environments effectively extending their reachable environment [124]. These strategies all effectively alter the user's mental state by decreasing the burden of user communication.

Finally, another strategy for robots to assist a user by transforming the robot's own physical configuration into one that is more amenable to task completion. This approach is useful in collaborative scenarios where the robot and user may collide. To avoid this problem, robots can decrease their operating velocity when working in close proximity to users [16, 264, 310] or take paths or actions specifically designed to reduce the likelihood of a collision [77, 117, 181, 228]. Similar to shared control, these strategies to assist the user decrease the user's cognitive burden of planning in the task. By taking responsibility for collisions, a robot can effectively alter its own actions so that the user can be less concerned with monitoring and modelling a robot's behavior and concentrating on completing their portion of the task.

3.4.4 Implications

Choosing which action space the robot should act in is a crucial decision for robot designers. To aid users in room cleaning, for example, researchers have developed robots that alter the environment by directly picking up misplaced objects, while others have developed augmented reality solutions that provide assistance in the user's mental space by showing them routes that, if followed, would lead to the shortest time spent cleaning. Realizing that a given task can be solved by acting in any part of the state allows researchers to develop novel solutions to problems that have historically been restricted to robots that act in a single state.

This realization, however, means that determining the robot's action space is not simply determined by the task that the robot is being built to solve. Instead, a roboticist must carefully consider the capabilities of the users for whom they are designing the robot. The choice of how the robot acts must be tuned to the needs of the user, and it has broader implications on the user's sense of agency and trust in the system. This separation of robot action spaces enables designers to compare robots from different domains that have similar action spaces and develop better assistive solutions.

3.5 Time

The third key design axis we present concerns how assistive robots coordinate the timing of actions with the targets of their assistance. Consider an assisted eating scenario. A robot might only offer food when given an explicit trigger by the user,

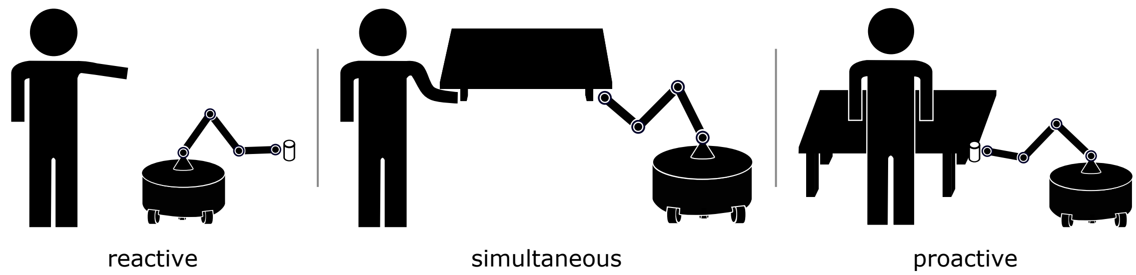


Figure 3.3: A key axis in assistive robotic systems concerns what type of cue leads to the robot taking actions. Robots can be reactive and respond to explicit input only, be proactive and interpret the general task state to choose to act on their own, or collaborate closely with the user by acting simultaneously with them.

or it can monitor the user's behavior to decide when to initiate the action itself. We categorize the timing of assistive actions as *reactive*, *proactive*, or *simultaneous*. Reactive robots act only when given explicit commands. Proactive robots use predictive models or other approaches to understand the world to initiate their actions without an explicit command. Robots acting simultaneously occur in collaborative settings, during which the robot continuously monitors the user for both explicit and implicit information to direct its actions. Choosing how to time the robot's behavior can change the difficulty of the task and how users react to the robot's assistance (Fig. 3.3).

3.5.1 Reactive

Reactive assistance occurs when the assistive action is triggered by an explicit command. Consider a teleoperated robotic arm developed for assistive eating [19, 137, 222]. In these studies, a user uses a two-degree of freedom joystick to control a seven-degree of freedom robot arm and pick up a morsel of food from a plate. Direct control of this robot entails only moving the robot's end-effector while the user is engaging the joystick. The user might also give commands at a higher level of abstraction, perhaps by pressing one button to request food and another for water.

Reactive robots can also respond to more task-specific, contextual triggers. In [53], an assistive robot helps a user to put on their shoes. This interaction is modeled as a complicated handover problem, where the user must have their foot properly positioned and apply enough resistance that the shoe remains on the foot. In this work, the robot responds to a gesture performed by the user through their foot. When they move their foot in the specified way, the robot knows that it is an acceptable time to place the shoe on their foot.

In general, reactive systems give the user more control over the robot and agency in the overall interaction. Additionally, the robot does not generally need sophisticated models of the task, since it can rely on explicit input control. This simplicity means that the robot tends to be less sensitive to the particular task or domain, as it relies on the user to adapt the task to the robot's capabilities. However, this additional control requires the robot's user to spend more time and effort on controlling the robot, which can distract from other tasks. Controlling a robot at this level may also require significant training, as the robot's capabilities may not clearly match the requirements of the task. The control burden grows as the user must explicitly command the robot to begin an interaction [24], and requiring additional control complexity, such as adding modal control to teleoperation, can be cognitively taxing and slow [120]. Furthermore, requiring the user to explicitly cue the robot to act reduces collaborative fluency, which has been shown to be desirable both in terms of the perceived quality of the interaction [123] and by decreasing the time spent during interactions [127].

3.5.2 Proactive

Proactive assistance occurs when the robot predicts that an action would fulfill the user's goals and takes that action without explicit instructions. For example, in assisted eating, the robot may anticipate a user's thirst after eating and choose to reach for the glass of water before receiving explicit input. The robot relies on a model of the task and user behavior to estimate what the user would want next. Proactive assistance generally improves the smoothness of interactions, as the assistance target does not need to spend the training time or cognitive load to provide explicit instructions to the robot. However, this type of assistance is dependent on the model used to cue its actions, so the added complexity may make the system less reliable.

Consider again the task of operating a high degree of freedom robot using a low degree of freedom input device. Instead of using explicit signals from the user, Herlant et al. [120] designed a robot that can proactively switch modes. In a simulated navigation task, a user drives a robot through a two-dimensional maze while restricted to moving the robot either horizontally or vertically at one time. The robot uses a model of the environment to determine whether horizontal or vertical motion is more optimal in the robot's current position. The robot can then switch the mode proactively, allowing the user to simply direct the robot to move, speeding up the overall interaction time and removing the cognitive burden seen in reactive mode-switching.

A robot can build a model of the user to infer the task goal before it has been expressed. For example, a robot can predict the next fruit that a customer wants to add to their smoothie [127]. Before the user explicitly requests this ingredient, the robot can prepare to grab that ingredient, increasing the fluidity of the interaction.

One challenge of proactive assistance is that users can be uncomfortable or even endangered if the robot makes unexpected motion. To mitigate this concern, the robot can communicate its intentions to the user explicitly. This could be done by having the robot show the user its plan directly on the physical environment, for example highlighting the part of a car door it plans to work on [14], or by showing its intended travel path in a virtual reality headset [289].

Proactive assistance enables more robust and general applications than reactive assistance does. However, the added sophistication in assistance requires additional complexity in the robot's models and behavior, which is compounded by the need to act in varied environments to unexpected stimuli. In addition, a purely proactive system can be uncomfortable or dangerous if the user is not prepared for the robot's actions. To mitigate some of these concerns, assistance systems can design some parts of the interaction as reactive and others as proactive. For example, the serving robot in Huang and Mutlu [127] proactively moves closer to its estimate of the user's most likely request, but it does not initiate the actual grasping process until it receives an explicit command.

3.5.3 Simultaneous

Simultaneous assistance exists between these two and includes shared control and collaborative robots. These systems generally function similarly to proactive assistance, but act at the same time as the user. These systems include both shared autonomy systems [137, 138, 183], which fuse the user's direct command with an autonomously generated command and arbitrate between the two according to some schema. It also includes tasks like carrying a table together [78, 230], in which both the user and the robot must act independently for progress to be made.

Simultaneous assistance occurs often in collaborative assembly tasks. The goal and structure of a joint assembly task is often pre-specified, making it easy to determine a user's goal, and a robot can directly assist by, for example, lifting and holding heavy objects steady so that they can be worked [88, 98]. A robot can also assist by orienting a part to optimize construction, for example by following the images found in an assembly manual [11, 332].

Simultaneous assistance often benefits from sophisticated communication strategies. For example, DelPreto and Rus [78] designed a robot to sense electromyographic signals from a user to jointly manipulate a heavy object. To further aid in the task, the robot could communicate back with the user, for example by changing its stiffness during a co-manipulation task in order to alert the user they should not move an object into a specific location [37]. Similarly, the robot could provide the user with cues as to the next step during a complicated assembly task such as by pointing at the next item of interest [7], providing a negative emotive feedback when a user completes an incorrect assembly step [251, 258] or otherwise imbue the robot with emotive capabilities to signal task progress [205, 317].

Simultaneous assistive systems generally require tight collaboration between the user and the robot. The closeness of the collaboration requires the system to have a more complicated strategy for understanding user commands, since it is unlikely that the user will give precise commands while also accomplishing their task. However, these models can be more flexible than pure proactive systems: the robot can gain immediate feedback from the user about whether or not its action is correct, so it can recover from some model failures more quickly.

3.5.4 Implications

Determining when a robot should act has implications on the quality of a robot interaction. Reactive systems use more explicit control which enables more user agency, but it also increases the burden to complete a task. Proactive systems require more sophisticated models and sensing onboard the robot, but they can improve collaborative fluency while decreasing user burden. Systems that act in anticipation of explicit user commands may even be able to influence future user behavior in unforeseen ways, leading to questions about who is in control of setting the task goal [220]. Proactive robots also generally lead to more robot agency, which introduces complex challenges such as safety and trust.

Preferences among how a robot chooses to take action may differ among users even within the same task domain. While one user may prefer a robot that requires less training and complication to operate, another might prefer to have more direct control over the robot to determine its behavior more precisely. If the user is paired with the system they least prefer, the interaction may cease to be assistive. In addition, an assistive system need not be completely proactive, reactive or simultaneous: the system can choose different timing and cueing strategies based on the particular part of the task under consideration. Choosing exactly when a robot executes its actions requires careful thought about the nature of the task, the capability of the robot, and the desires of the user.

3.6 Conclusion

In this work, we describe an overall perspective on robotic systems that emphasizes their assistive intentions. With this perspective, we present three key design axes that compare assistive robotics research across domains: the relationships they develop with people, their action space, and their action timing. We explore these axes through a review of recent assistive robotics research, showing how assistive robots from across domains face similar challenges and make comparable decisions along these axes.

Much of the research discussed in this work is specific to its task domain, due to how the field has been organized and the difficulty of building abstractions. In this work, we propose some abstractions, and we hope that they will enable designers of assistive robots to find systems in other domains that share their problems and to draw deeper connections with them.

For each axis, we discuss design tradeoffs resulting from particular approaches. From among these axes, several themes emerge. Choices in the robot's action space and timing can both affect a user's sense of agency. Similarly, the robot's action space and relationship to the user both impact the structure of the communication between the robot and the user, which alters the quality of the assistance. It is our hope that researchers will explore more themes that span the design axes and provide more structure to the development of assistive robots.

This work is intended to start a conversation about how to understand the specific challenges of assistive robotics within the general area of human-robot interaction. With this framework, we hope to encourage researchers to further explore the nature of assistance as a general concept and its inherent challenges. We do not claim that these axes are complete; rather, we present them as the beginning of a larger effort to develop general principles of assistive robotics.

By understanding assistive robotics as a value alignment problem, as well as the components of an assistive system, we can begin to build a system that provides assistance to people in everyday household tasks. To do this, we can understand that we want to capture behaviors from the person that we wish to assist, that we want to assist by manipulating the environment, and that we will need to incorporate either simultaneous or non-simultaneous actions depending on how the use is collaborating with the robot. Our first step in building this system is to understand people's naturalistic behaviors during rearrangement tasks. Toward this end, we collect a large scale dataset of naturalistic behaviors in a simple, simultaneous action, rearrangement task, introduced in the following chapter.

Key Axis	Description
People (§3.3)	How the robot considers additional people outside the baseline dyad.
Targets of assistance	Additional people whose goals are of comparable importance to the user.
Interactants	Additional people whose goals are not privileged and use general human-robot interaction approaches.
Space (§3.4)	The portion of the mutual state the robot's actions affect.
Environment	The robot affects the environment directly by, e.g., manipulating task objects.
Human body	The robot affects the user's body by physically moving some portion of their body.
Human brain	The robot affects the user's mental state by providing information about the task or reducing the cognitive burden.
Time (§3.5)	The relative timing between a robot's actions and the user's explicit commands during the task.
Proactive	The robot acts before an explicit command.
Reactive	The robot acts in response to an explicit command.
Simultaneous	The robot acts simultaneously with user action.

Table 3.1: Assistive robots can be explored along three key axes: how the assistive system thinks about additional **people**, what part of the mutual state aligns with its action **space**, and at what **time** it executes its actions during a task.

4 Human behavior during simple surface rearrangement tasks

In human-robot collaborations, robots need to perceive, understand, and predict the effects of their own actions, as well as the actions of their human partners. This is especially important for assistive robots, which perform actions toward a (sometimes implicit) human goal. To successfully produce these assistive actions, the robot system must perceive, understand, and predict human mental states (the human's goals, intentions, and future actions, often unknown to external observers) that determine what assistance the robot should provide.

Concretely, when people complete physical tasks, their external behaviors—such as their eye gaze—can reveal insights about their internal mental states. An assistance system that can understand how these behaviors relate to the task can predict which objects and locations of a visual scene the human deems to be task relevant. The system can also use these behaviors to determine whether or not interactions with these objects or locations will take place, and qualities that describe these interactions. This information is not known to the system prior to completing a task and is not relayed to the system by the human via traditional means (e.g. verbal or written communication). Thus, understanding these mental states in order to assist the human requires perceiving and interpreting the human's behavior during human-robot collaborations.

An example of a behavior that has been well studied in physical tasks is eye gaze. People almost exclusively fixate their eye gaze on objects or locations involved in their current task [116], thereby ignoring task irrelevant parts of a scene. Should these objects or locations require a direct interaction, people fixate their gaze on these

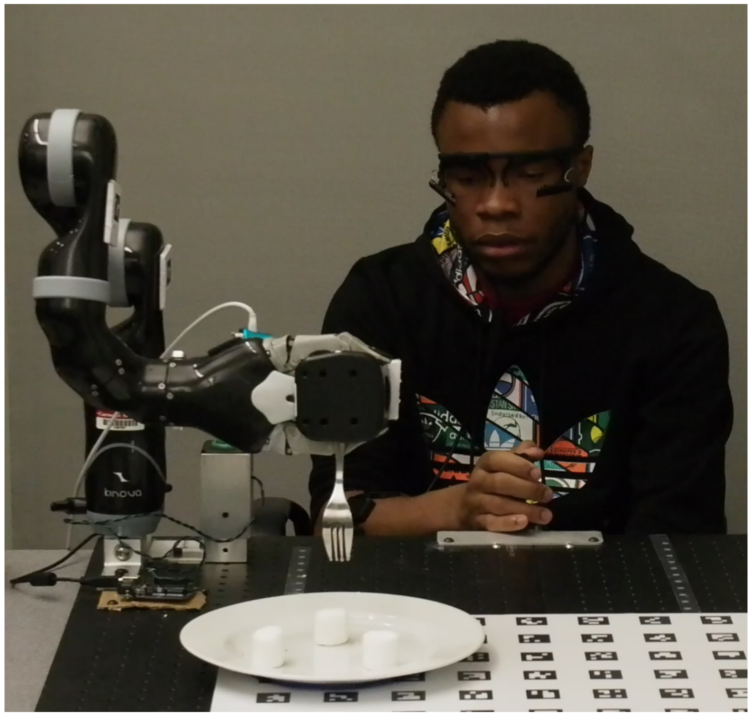


Figure 4.1: The HARMONIC data set provides multimodal human, robot, and environmental data collected during an assistive human-robot collaboration.

objects and locations before moving their hands to complete the interaction [165], thus revealing the intended interaction object in advance of any physical contact. The gaze also lingers on key points in the task, such as obstacles, revealing certain landmarks of manipulation [141]. In addition, people gaze at objects before uttering verbal references, which others can use to disambiguate and predict speech [6, 40].

Other human behaviors can also reveal current mental states. Electromyography (EMG) signals, which record electrical stimulation of muscle fibers, can indicate what action people are attempting to complete with their hands.

Furthermore, pupil size has been correlated with cognitive load [29, 30, 161], and understanding current human body posture can reveal desired tasks and help prevent potentially dangerous collisions [193].

In this paper, we present the Human And Robot Multimodal Observations of Natural Interactive Collaboration (HARMONIC) data set. The HARMONIC data set contains human, robot, and environment data collected during the human-robot collaborative task (Figure 4.1). In this task, people control an assistive robot arm to pick up bites of food in a simple eating scenario. The 6 degree of freedom Kinova Mico robot arm is controlled in three dimensions via a 2 axis joystick and manual mode-switching. In some cases, the robot provides additional assistance through shared autonomy [138].

Although the data were collected during an assistive eating task, their usefulness extends beyond the specific domain of eating. The manual condition can be used to study human teleoperation in the general case, for example, with tasks using simplified grippers such as vacuum tooling. When combined with the shared autonomy conditions, these data can be used to study co-manipulation across individuals and varying levels of robot agency. Included in the data are a wide array of non-verbal behaviors situated in a real-world task defined with a clear goal and thus relevant for a variety of human-robot collaborations.

Human behavioral data include egocentric RGB videos, eye gaze positions relative to these videos, infrared (IR) videos of both eyes, stereo, third person video of the participant, and EMG recordings on the joystick-controlling arm. Robot related data include joystick control inputs from the user, the control input and belief distribution calculated by the assistance algorithm, and the robot position. Environmental data include the 3D locations of food morsels as well as the locations of fiducial markers. Further information and an explanation on how to access these data is provided in the following sections.

Our data set will help researchers study the complex human-robot dynamics of assistive teleoperation, which can vary between individuals and between different levels of robot autonomy. For example, researchers could use this data set to learn correlations between eye gaze and joystick control, in order to improve the goal-inference predictions made by shared autonomy algorithms. Others might be interested in modeling and forecasting the dynamics of joystick inputs under differing amounts of robot assistance. Previous research using similar data has proposed identifying unexpected events (*e.g.*, human errors or task failure) by learning a normative gaze behavior model and identifying anomalies [17]; the higher quality data provided in this data set could continue this line of research as well as extend it to situations where the robot provides variable levels of assistance within a unified framework.

4.1 Data collection procedure

This section presents a brief overview of the user study and the robot system to explain the conditions under which the data streams were recorded.

4.1.1 Participants

Twenty-four participants (13 female) were recruited from the Pittsburgh area. Seventeen were between the ages of 18–24, four between 25–30, one between 31–35, and two between 41–45. The participant pool was screened for prior experience using this robot arm in similar studies and, thus, were novices at the task. The experiment was carried out in the Human And Robot Partners (HARP) Lab on the Carnegie Mellon University campus. The participants were compensated $15 for one and a half hours of their time.

4.1.2 Protocol

The participants controlled a robot arm, attempting to position a fork above one of three marshmallows placed on a plate (see Fig. 4.1). They controlled a robot with a two-axis joystick using modal control: the joystick's two axes moved the end effector of the robot in x and y, z and yaw, or pitch and roll. A joystick button allowed participants to cycle between control configurations when pressed for less than 500 milliseconds. When the task was completed (that is, once a participant was satisfied with the fork's position or had given up on the task), the participant held down the same joystick button for more than 500 milliseconds. This action triggered an autonomously executed plan in which the robot moved down to the height of the plate and speared the marshmallow (conditional on the proper positioning of the fork). Finally, the robot arm moved into a serving position near the participant's mouth. This concluded the trial, and the robot automatically reset to the starting configuration.

Participants were given a brief introduction to the purpose of the study and then began a five-minute familiarization period, in which they controlled the robot in teleoperation mode and the data were not recorded. The participants were then fitted with eye gaze and EMG sensors (described below). They performed the task five times in sequence for each of four assistance modes (described in the next section). The order of the assistance mode was fully counterbalanced among the participants. After each block of five trials, participants were given a brief survey to record their subjective perceptions about the algorithm. Once the final survey was completed, participants were presented with a survey that compared all conditions through ranked preference as well as free response.

4.1.3 Assistance conditions

Participants operated the robot under each of four different assistance conditions: fully teleoperated, two different levels of assistance according to the shared autonomy framework [138], and a fully autonomous robot.

The following is a brief description of how the assistance is calculated; a full description is available in a previous publication [138]. The combined human-robot system is modeled as a Partially Observable Markov Decision Process (POMDP) [143, 299], where the participant's goal is represented as one unknown member of a small set of possible goals. Participant inputs via joystick are treated as observations. The algorithm assumes that the user is noisily optimizing a cost function parameterized by their unknown goal. Therefore, the Maximum Entropy Inverse Optimal Control (MaxEntIOC) [361] framework can be used to evaluate a belief distribution in the known goal set. From this belief state, the overall POMDP is solved by applying the QMDP [178] approximation, which has proved reliable for similar shared control scenarios. Our implementation slightly changed the original formulation to remove the inherent living reward, which can cause the robot to converge on a goal even in the absence of any positive joystick actuation. The resulting robot action consists of a computed assistive action based on the inferred user goal distribution combined with the original applied user action.

To provide different levels of assistance, the shared autonomy transition function was modified slightly from prior work. In Javdani et al. [138], the given transition function applies both the user and the robot control as determined by $a_{applied} = u + a$.

In order to adapt the amount of user control, the applied action was parameterized by a value γ: $a_{applied} = (1 - \gamma)u + \gamma a$, which trades off between the relative strengths of the user command and the robot assistance. Note that the original shared autonomy procedure would correspond to the case $\gamma = 0.5$ and normalizing the vector $a_{applied}$.

The four conditions corresponded to four different levels of γ:

Direct teleoperation, $\gamma = 0$. The assistance signal a was computed but completely discarded, so the user had full manual control over the robot.

Low assistance, $\gamma = 0.33$. The assistance signal was combined with the direct user control, with the user signal weighted double.

High assistance, $\gamma = 0.67$. The assistance signal was combined with direct user control, but the assistance signal was more highly weighted.

Autonomous robot control, $\gamma = 1$. The user control signal was not passed through to the robot control. It was used for goal inference only and the robot was autonomously controlled based on its goal inference results.

4.1.4 Sensors

Eye gaze. Participant eye gaze direction was captured by a Pupil Labs Pupil [149, 247] sensor. This sensor consists of a glasses-like frame with two infrared cameras with infrared illumination mounted below each eye for dark pupil tracking, plus a third RGB camera oriented outward to capture egocentric video. The eye cameras capture video at 120 Hz, and pupil labs software detects the pupil pixel center. Before data were captured, the pupil locations and world camera videos were calibrated by asking the participant to look at the center of the marker held in front of them by the researcher ("manual marker calibration"). This calibration routine was recorded for most participants and is made available in the `calib` folder. The calibration is verified between each condition by asking participants to look at particular places in the scene. These checks are recorded and made available in the `check` folders.

EMG. Participant muscle activation while controlling the joystick was captured using a Myo sensor [318]. Due to initialization failures, these data are only available for about 20% of the runs (see Table 4.1 for full details). It consists of the EMG message, denoting the activation of eight individual EMG sensors, the ORI message, denoting the orientation of the arm in roll/pitch/yaw, and the IMU message, denoting the readings of the IMU attached to the armband.

External video. Participant behavior was captured using a Stereolabs [305] ZED camera. The left and right videos are stored as separate `MP4` files. The ZED camera was placed on a tripod at approximately the same (marked) location for each trial to capture a full-on view of the participant and occasional views of the scene. ZED videos are available for the 10 participants who consented to their images being released. In all cases, offline skeleton and face tracking information is available.

4.2 Descriptive statistics

This data set consists of 480 trials, comprising 20 trials for 24 participants. The data represent about five hours of continuous instrumented robot control. A summary of the available data appears in Table 4.1.

4.3 Data streams

The data are organized first by participant (`p100-p123` reflecting the twenty-four participants). Each participant folder contains folders for three types of recordings: `calib` contains calibration passes, `check` contains intermediate gaze accuracy checks, and `run` contains data collection runs. These folders contain numbered subfolders that indicate the run sequence. A visual representation of selected data streams can be seen in Fig. 4.2.

A single trial capture (a numbered folder) has the following subfolders:

- `text_data` contains exported `CSV` files containing the raw data. The particular raw data streams available are detailed in the following subsections. In addition to raw data, this directory contains the body skeleton, facial, and hand keypoints generated by running OpenPose [54, 295, 335] on the left and right streams of the third-person ZED videos. The outputs from OpenPose are compiled into face, right, and left-hand, and pose files for each stream of the depth camera. For full descriptions, please refer to the OpenPose documentation.
- `stats` contains a number of `YAML` files detailing statistical information about the

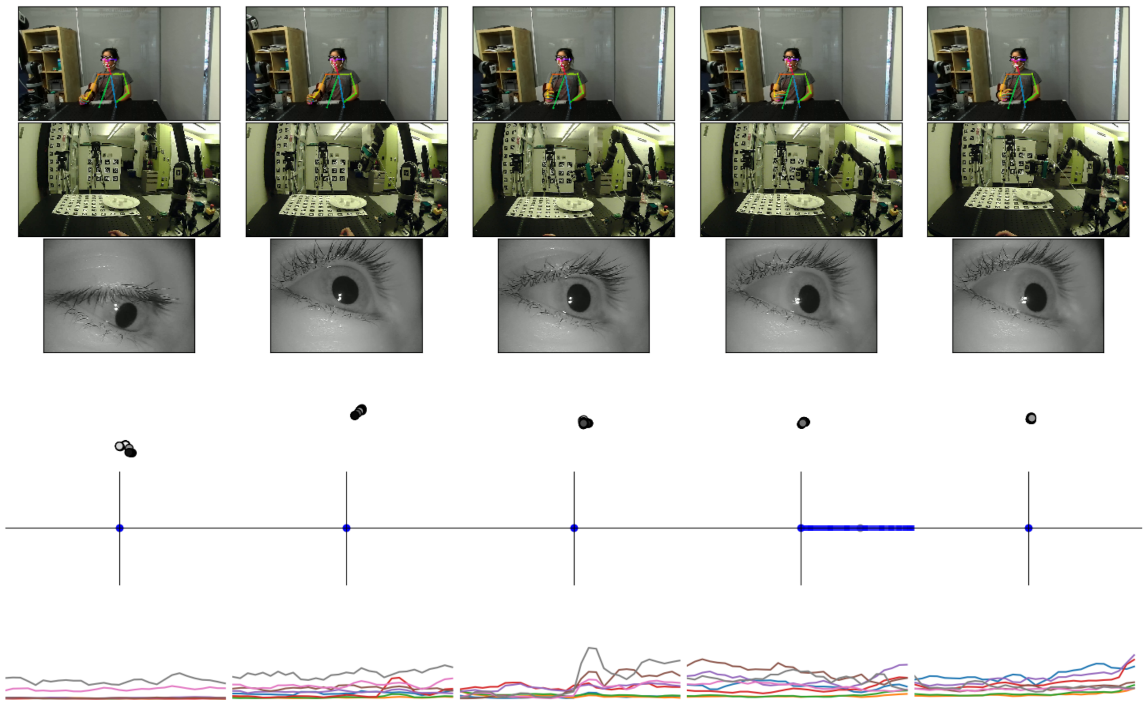


Figure 4.2: A visualization of several streams from the HARMONIC data set. The top row displays the ZED video with OpenPose skeletons overlaid, then the egocentric video captured from the Pupil camera, left eye video, one second of the calculated gaze dot, the trajectory of the joystick, and finally the Myo activations. For the gaze dot and the joystick, lighter colors represent more recent points in time. Each of these plots represents one second of data, sampled at 30 FPS.

trial and the general data stream, including the number of records, approximate time distances between individual records and estimates of times when data points may have been dropped based on the nominal data collection frame rate.

- `videos` contains the Pupil video files (`eye0.mp4`, `eye1.mp4` and `world.mp4`) exported as MP4 files using the H.264 video codec [260]. In addition, the timestamps of each frame as either numpy (`*.npy`) files, raw text (`*.txt`), or `CSV` (`*.csv`).
- `processed` contains a number of new formats of data extrapolated from the underlying data (e.g. a video of the egocentric recording with a dot overlaid at the gaze point).

4.3.1 Timing and synchronization

All data points were timestamped on collection and stored as 32- or 64-bit floating point values in number of nanoseconds from the Unix epoch. The `CSV` files in `text_data` provide these data in several columns.

For ease of use, two common indices are provided for all data streams. The `world_index` field gives the egocentric video frame number corresponding to each data point. A second common index, `world_index_corrected`, provides a second index into the egocentric video, with a correction for dropped video frames. The `world_index_corrected` value approximates a common 30Hz clock running throughout the trial. For more sophisticated data alignment, please use the provided timestamps.

4.3.2 Eye gaze

Eye gaze videos were recorded at 120 Hz and located in the `videos` folder as `eye0.mp4` and `eye1.mp4`, encoded using the H.264 video codec [260]. Frame-level timestamps are available in corresponding NumPy binary files, `eye0_timestamps.npy` and `eye1_timestamps.npy`. The automated pupil detection results for each eye are in the `text_data` folder, under `pupil_eye0.csv` and `pupil_eye1.csv`. The field names correspond to the output of the 3D pupil detection process in Pupil Labs, as described in their documentation.

Egocentric video is available in the `videos` folder as `world.mp4` (encoded using the H.264 codec [260]), with frame-level timestamps located in `world_timestamps.npy`. The calculated gaze position within the corresponding video frame is given in `text_data/gaze_positions.csv`. See the pupil labs documentation for a full description of fields. The fields `norm_pos_x` and `norm_pos_y` correspond to the (x, y) pixel in coordinates normalized to the size of the egocentric video frame, with the origin point on the top left.

Data used to calibrate between pupil data and gaze point are stored in the text files `pupil_cal_eye0.csv`, `pupil_cal_eye1.csv`, and `world_cal_positions.csv`. These data are the same between runs of the same participant and are provided as a convenience to recalculate a calibration if desired. Details of the current calibration method can be found in the Pupil Labs software documentation.

4.3.3 Third person video

ZED videos were recorded using Stereolabs ZED software, version 1.1.0. Data were initially stored as a Stereolabs SVO file, including separate left and right videos and a common timestamp. The videos were extracted from the `videos` directory as `zed_left.mp4` and `zed_right.mp4` encoded using H.264 [260]. The timestamps were rescaled to the Unix epoch and stored as an integer number of nanoseconds from the epoch in `zed_ts.txt`, as well as floating-point NumPy format in `zed_timestamps.npy`. The `zed_corrs.csv` stores the correlations to a common index, as previously explained.

4.3.4 Additional sensor data

The following data streams are available in the `text_data` directory, having been extracted or calculated from the original binary.

- `control_mode.txt` contains one character referring to that trial's assistance condition. Zero represents direct teleoperation and 3 represents robot control.
- `morsel.yaml` is a `YAML` file with the transforms for each detected morsel positions in the robot base frame.
- `ada_joy.csv` stores raw joystick input provided by the user. Joystick input is

only provided when changed from the previous message leading to inconsistent timing in the raw data. To rectify this, joystick data have been resampled to a common 120 Hz frequency and missing data filled by the previous value. Duplicate data are noted by unchanged headers.

- `input_info.csv` contains the user input to the robot. The `robot_mode` field denotes which control mode the robot is in (x/y, z/yaw, or pitch/roll), and the rest of the fields denote the applied twist corresponding to the user's joystick input.
- `assistance_info.csv` contains the outcome of the shared autonomy algorithm. It stores the current probability inferred for each goal and the resultant twist applied to the robot at that timestep.
- `joint_states.csv` contains the information for each joint of the robot.
- `robot_position.csv` contains the cartesian position of each of the robot links, as calculated from the forward kinematics using the data from `joint_states.csv`.
- `myo_emg.csv` contains EMG output of the Myo.
- `myo_imu.csv` contains IMU output of the Myo.
- `myo_ori.csv` contains orientation data received from the Myo sensor.

4.4 Known issues

4.4.1 Missing data

Due to computational load, certain data streams may have periodic dropouts. The `stats` directory contains some information on when and how often these occur, and general statistics are given in Table 4.1. The missing data are particularly exacerbated for the Myo signal because the data recording software failing to start. Finally, due to permissions restrictions, unedited ZED video capture is available for 10 participants, deidentified video (video with faces blurred) is available for 13 participants, and video for 1 participant is unavailable for release. Within the released participants, some initialization failure means that videos of certain trials are occasionally missing.

4.5 Accessing the data

The data will be hosted on the HARP Lab website: http://harp.ri.cmu.edu/harmonic. Several files are provided for download: `harmonic_data.tar.gz`, a compilation of all of the data, (∼ 68 Gb), `harmonic_minimal.tar.gz`, consisting of the `text_data`, `videos`, and `stats` directories, (∼ 15 Gb), `harmonic_text.tar.gz`, consisting of the `text_data` directory, (∼ 4 Gb), and finally `harmonic_sample.tar.gz`, consisting of all of the data for a single participant, (∼ 303 Mb). The data sets will be versioned using semantic versioning and that page will maintain a log of all changes that may be made to the data set after release. Furthermore, our GitHub contains a repository for basic processing tools located here: https://github.com/HARPLab/harmonic_cpp. Finally, for the original robot control code, follow this link to a fork of the publicly available implementation of the shared autonomy code we used: https://github.com/HARPLab/ada_meal_scenario [137]. Our robot control code is on the branch: "adjustable".

4.6 Conclusion

We presented a data set of humans who performed a food acquisition task by controlling a robot manipulator. During this task, a variety of types of participant data were collected, including eye gaze information, electromyography of the controlling arm, stereo video, and robot controller information. This data set enables research into human-robot collaboration and multimodal human behavior analysis.

Using this dataset of naturalistic human behaviors we can begin to understand how to featurize human behaviors in order for a robot to better provide assistance during a simple surface rearrangement task. Using the insight that eye gaze and joystick information may reveal different information about a user's goal during different parts of the task, we used the behaviors collected in HARMONIC to train a coordinated feature representation of these two behaviors for potential use in a downstream assistance task. These experiments are outlined in the following chapter.

	Left Eye	Right Eye	Egocentric Video	ZED Camera
Total duration (h:m:s)	5:19:26	5:10:45	5:33:44	4:44:45
Total frames	2299877	2237380	600728	512569
Nominal frequency (Hz)	120	120	30	30
Frames dropped	133301	195860	7459	94431
Coverage (%)	94.52	91.95	98.77	84.44
Present (%)	100.00	100.00	100.00	87.25
Coverage if present (%)	94.52	91.95	98.77	94.83

	Joystick	Rob. pos.	EMG	IMU	ORI
Total duration (h:m:s)	4:56:00	5:48:05	1:10:49	1:10:53	1:10:53
Total frames	2131160	1670798	212465	212664	212659
Nominal frequency (Hz)	120	80	50	50	50
Frames dropped	114250	1680	802368	802204	802206
Coverage (%)	94.91	99.90	20.94	20.95	20.95
Present (%)	100.00	100.00	21.48	21.48	21.48
Coverage if present (%)	94.91	99.90	99.75	99.83	99.83

Table 4.1: Descriptive statistics of each data stream in the data set. *Total duration* and *Total frames* refer to the collective amount of data of that signal over all trials and participants. *Total duration* is extracted by dividing the total frames by the *nominal frequency*. *Frames dropped* are based on interpolating from the nominal frame rate and detecting missing data. *Coverage* is computed by dividing the number of data frames by the expected number of data frames from the nominal frequency over the whole data set, *Present* indicates the fraction of trials that have at least one datum of that type, and *Coverage if present* is the total number of data frames divided by the expected number evaluated only if at least one datum is present in the trial.

5 Evidence that value alignment requires continual personalization

Our goal of understanding how to use naturalistic behaviors to provide value-assigned assistance requires that we understand how best to incorporate these signals into robotic algorithms. In this work, we explore how we can develop implicit, differentiable representations of these behaviors so that they can be used for downstream assistance tasks.

Without any algorithmic support, it can be extremely taxing for an individual to directly control, *i.e.*, teleoperate such a robot. This is because the number of DOFs of the robot being controlled is generally much larger than that of the input device. Furthermore, it is not sufficient to allow the robot to complete these tasks autonomously, as previous work has shown that retaining explicit user control is especially important in assistive domains, where users strongly prefer systems that allow them to stay in control of assistive robots, even if they are less efficient at completing the task [109, 151].

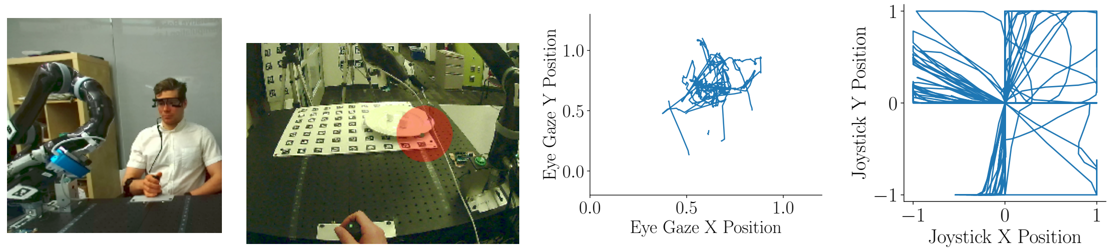
To address this, researchers have developed shared autonomy algorithms that combine user control with autonomous robot behavior, resulting in comanipulation of the robot [85, 109, 137]. For shared autonomy algorithms to be successful, they must have the ability to accurately characterize a user's state to support complex and high-dimensional comanipulation tasks, such as assisting a user with motor impairment to eat with a 6 degree of freedom (DOF) arm. These algorithms use direct control signals, such as the operator's joystick inputs, to predict the operator's goals so that the robot can take a cooperative action to assist the user.

However, basing such goal predictions only on direct control signals, such as explicit joystick behavior, misses the opportunity to use rich human behavior signals that can further reveal user state, specifically user intent. For example, eye gaze is closely related to the target and timing of manipulation actions in people [116, 165]. In human-robot comanipulation tasks, here cooperative eating with a 6 DOF robot arm, eyes can be used to predict user actions or identify errors in teleoperation [5, 17, 19]. Eye gaze is therefore a natural mechanism to supplement the human goal prediction that takes place during shared autonomy.

To successfully incorporate this signal into the shared autonomy paradigm, it is necessary first to understand hand-eye coordination in comanipulation. Specifically, we need to determine the basic building blocks of comanipulation that will allow us to coordinate between a user's eye gaze and how they control the robot's end-effector (here by using their hand to manipulate a joystick). Coordination allows us to relate the varying task-relevant information contained in the different data streams to each other. For example, in assisted eating, the joystick can reveal the immediate vector in which a person wants the robot to travel, but eye gaze can reveal the ultimate bite of food the user wishes to spear; here, it is important to perform an action that does not move too far away from the immediate action while still optimizing for the overall goal. By relating the joystick and eye gaze streams in this example, we can get a more complete vision of the user's state: not just where they want to go, but how they would like to get there.

However, before coordinating the data, it must be processed. From an algorithmic perspective, processing human behavior signals like eye gaze, head pose, or joystick inputs is non-trivial. These signals are noisy and different data streams provide different task-relevant features. In addition to the algorithmic complexities, simply obtaining these signals is challenging, as collecting data from multiple sensors to train data-driven models can require burdensome engineering efforts to set up, calibrate, and synchronize. Fortunately, there have been a few large-scale data set collection efforts for teleoperation and human-in-the-loop comanipulation tasks [222, 334].

In this paper, we use a large-scale multimodal data set called HARMONIC to identify the basic building blocks of hand-eye coordination exhibited by people during comanipulation tasks (Fig. 5.1). We identify semantic, multimodal action primitives that establish the basis for a user's state. Then, we apply modern data-driven techniques to classify multimodal, multiview real world data into these action



(a) Third person view (b) Egocentric view (c) Eye gaze data in one trial (d) Joystick data in one trial

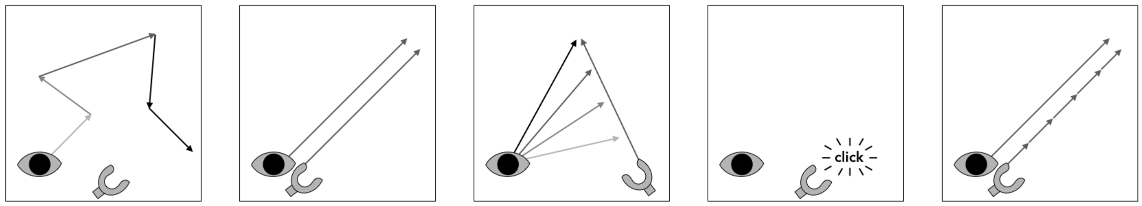
Figure 5.1: The HARMONIC data set contains (a) third person video, (b) egocentric video, (c) eye gaze fixations, and (d) joystick data from a human-robot co-manipulation task.

primitives, to verify our choice in primitives as well as show that coordination between eye gaze and joystick is possible. Finally, we justify the multimodal problem by showing that unimodal analyses are not sufficient to explain joint behaviors.

We provide several novel contributions toward an understanding of hand-eye coordination in comanipulation. We first define macro action primitives, which segment user actions into meaningful sequences of individual user states, and discuss how they differ from physiological gaze primitives (Section 5.1.1). To evaluate our data-driven recognition models for macro actions, we create a synthetic data set that contains these action primitives so that we have full control over the generative process and have perfect access to ground truth annotations (Section 5.1.3). We show how these semantic macro action primitives can be modeled using both the synthetic and real raw data (Section 5.2), and provide a thorough experimental analysis of our models (Section 5.3).

5.1 Problem domain

We build models of hand-eye coordination in human-robot comanipulation to better understand the user state during high-dimensional comanipulation tasks such as assisted eating. To describe hand-eye coordination, we define action primitives that provide a semantic understanding of user state. We draw our action primitive definitions by semantically analyzing the HARMONIC data set, which is described



(a) Exploration (b) Pursuit (c) Correction (d) Mode switch (e) Toggle

Figure 5.2: Five macro action labels capture combined eye gaze and end effector dynamics. a) Exploration denotes periods of high eye gaze movement and low joystick (robot) movement. b) Pursuit denotes periods of highly correlated eye gaze and joystick movements, where the eye gaze follows the path of the robot. c) Correction denotes successive glances between different parts of the scene while the robot is moving. d) Mode switch denotes when the user is using modal control to cycle through sets of degrees of freedom. e) Toggle denotes periods in which the joystick is being moved in rapid, short, consistent activations.

briefly here. Additionally, to test these primitives in a systematic fashion, we construct a synthetic data set, as described below.

5.1.1 Defining action primitives

In our current work, we use the term *micro actions* to refer to three low-level gaze action primitives that can be used to understand attention or user state. *Fixations* are eye movements that focus the eye gaze on a single point in space and are used to gather visual details. *Saccades* are fast, point-to-point movements of the eyes that bring a new area into the center of vision. Finally, *smooth pursuits* are when the eyes track a moving object to keep it in the center of vision.

However, micro actions only partially express hand-eye coordination during co-manipulation, because they do not capture the robot's movement. For this, we manually analyzed the HARMONIC data and identified five common *macro actions* (Fig. 5.2) composed of gaze gaze and joystick movements. Our five macro actions are: exploration, correction, pursuit, mode switch, and toggle.

Exploration is defined by minimal joystick activity and high eye gaze activity. Semantically, this class represents a person exploring the space with their eye gaze, preparing to make an action with the joystick. This sequence starts when the joystick moves into a period of rest and ends once the user activates the joystick. This can be

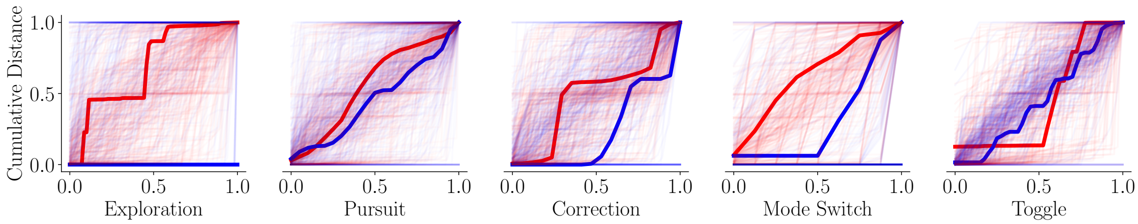


Figure 5.3: A graphical description of the differences between macro action categories from HARMONIC. Eye gaze sequences are red, while joystick sequences are blue. The y-axis shows the normalized cumulative distance for each sequence. The x-axis shows normalized sequence lengths. Every trial is plotted, with a representative sequence highlighted in bold.

seen in Fig. 5.3 where the joystick sits generally at the bottom of the plot, indicating no movement throughout the sequence, and the eye gaze is dispersed throughout the plot.

Pursuit is defined by correlated eye gaze and joystick action. In this class (which is not to be confused with the micro action primitive *smooth pursuit*), the participant moves the joystick and follows the resulting robot action with their gaze. This may result in large eye gaze movements when the robot is moving across the visual scene or little eye gaze movement when the robot's end effector is rotating. This action begins when the eye gaze begins to follow the robot's action (as resulting from the joystick activation) and ends once the eye gaze moves away from the previously fixated position. This relationship can be seen in Fig. 5.3 where the gaze and joystick signals are tightly coupled.

Correction can be categorized by high joystick activity and consistent eye gaze glances between a "home" point and another task relevant scene point. Prior work has called these "monitoring" glances [19]. This action can reveal an operator's goal or the target of their current control input. This sequence begins as eye gaze moves away from a previously fixated position during joystick activation. It ends once the eye gaze has travelled back to the original position (after one or more fixations elsewhere in the scene), the joystick comes to a period of long rest, or the participant enters into one of the other semantic categories. This can be seen in Fig. 5.3 where the eye gaze initially takes a stair step approach indicating fast movement initially, a pause and then fast movement again. This pattern is then followed in the joystick channel.

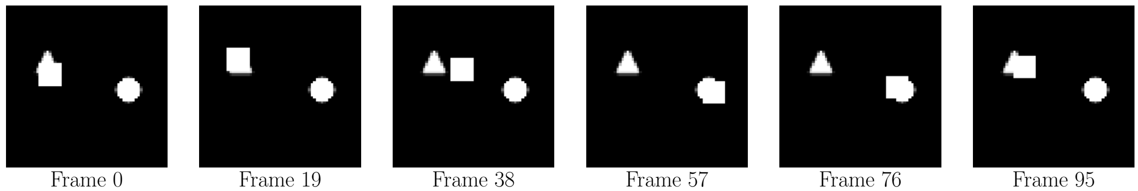


Figure 5.4: Our synthetic dataset was modeled as a simplified version of the eye gaze and joystick signals from the HARMONIC task. Here, an example of the exploration action, with simulated eye gaze (square), robot position (triangle), and goal (circle).

Mode switch represents when the participant switches control modes. This class is programatically generated by taking the five frames before and after the button press that causes a robot control mode switch.

Toggle is defined by quick, successive joystick taps with the eye gaze path closely following the end effector. This begins when the participant makes short bursts with the joystick, and ends either when the joystick comes to a period of inactivity or consistent activity. This can be seen in Fig. 5.3, where the cumulative distance of the joystick takes a stair step pattern, while the gaze initially lags behind and then catches up at the end of the sequence.

5.1.2 Naturalistic human behavior data set

As the source of real-world human-robot comanipulation data, we used the previously released HARMONIC data set [222], described in Chapter 4. Though the full data set includes approximately five hours of data, for the current analysis, we are investigating teleoperation only. Therefore, we only included the five trials per participant where people were fully teleoperating the robot (*i.e.*, the robot assistance signal was set to zero). These data are pictured in 5.1.

5.1.3 Synthetic data set

Real-world data are noisy, so we developed a synthetic data set that allows us to test our models with hypothetically perfect inputs. Additionally, it provides an opportunity to control and experiment under a variety of noise parameters, prototype experiments at scale, have full control over the data generator, and have access to actual ground truth labels.

This synthetic data set was designed to mimic the task in the HARMONIC data set. The robot end effector navigates to a virtual goal, while a virtual eye gaze stream is simultaneously overlaid on the scene. As seen in Fig. 5.4, the robot is represented as a triangle, the eye gaze as a square, and the goal as a circle. The robot aims to navigate to within a threshold of the goal. Trials with fewer than 200 or more than 1000 frames were discarded.

Table 5.1 shows the distribution of micro and macro actions in both HARMONIC and synthetic data sets. We can see that the number of sequences is relatively balanced in the HARMONIC data set, with fewer toggle sequences overall.

Table 5.1: Distribution of class labels in HARMONIC and synthetic datasets for both the micro and macro classification tasks.

		HARMONIC	Synthetic
Micro	Saccade	0.1796	0.4321
	Smooth Pursuit	0.5541	0.1743
	Fixation	0.2663	0.3936
Macro	Exploration	0.2319	0.3377
	Correction	0.2424	0.0993
	Pursuit	0.1765	0.1140
	Mode Switch	0.2181	0.3377
	Toggle	0.1311	0.1113

5.2 Method

5.2.1 Micro and macro action labeling

Micro action labels were automatically classified using Bayesian Decision Theory Identification (I-BDT) [270] which classifies gaze-actions in online settings. For an explanation of this algorithm, we refer the reader to the original paper.

Macro actions were hand-labeled for ten participants in the HARMONIC data set. Sequences with significant amounts of low-confidence eye gaze calculations (given by the eye tracker). Two of the ten labeled participants (p102 and p103) were completely discarded due to significant missing data. For the remaining labeled participants, we dropped 62 of 2931 total sequences (2.1% of sequences) or 2931 of 48135 frames (3.4% of frames) because of missing gaze data. Supplementary information contains more details about our exclusion criteria, as well as our subsampling method.

5.2.2 Models and input representations

We built models for micro and macro actions using a two-layer Gated Recurrent Unit Recurrent Neural Network (GRU-RNN) [66] to encode a given input sequence using the PyTorch neural network library [235]. We tested four input families: eye gaze alone, joystick alone, an early fusion of eye gaze and joystick, and a late fusion of eye gaze and joystick (Table 5.2. We also considered different hidden sizes for these models, which are shown in the *hsize* column of Table 5.2.

To decode these sequences into classification vectors, we collect the context vectors for each step in the sequence, and then feed this into a three-layer Multi-Layer Perceptron. In late fusion models, the eye gaze and joystick signals are each encoded by two separate encoders, and then the context vectors are concatenated prior to being decoded. This is in contrast with the early fusion models, in which the eye gaze and joystick sequences are concatenated along the feature axis and then jointly encoded.

The first layer of the decoder is the product of the maximum sequence length and the hidden layer size. The second layer is half that, and the final layer is the number of classes. This decoder model is fully connected, and ReLU [215] is used for nonlinearity after the first and second layers. All models were trained using the Adam optimizer [153] using a learning rate of 1e-3 and a cross-entropy loss weighted by class distribution (given in Table 5.1).

The *x,y* embedding indicates that the inputs are given to the model as is, without modifications. For synthetic data, this is the x,y position of both the eye gaze and the joystick. For HARMONIC data, eye gaze includes the confidence score from the eye tracker, while the joystick additionally includes a hot vector that indicates the current control mode. The *dx,dy* embedding indicates that the difference (or discrete derivative) of the signal is taken before the input is passed to the model. Finally, the *binary* representation divides the input space into a 10x10 grid and generates a hot vector indicating the pixel closest to the real-valued number. This vector is then passed on to the model as input.

5.2.3 Problem setup

For both tasks, we consider a supervised classification problem. Our goal here is to show a correlation between the segmented raw data and our provided macro labels. Outperforming chance and the zero rule (guessing the majority class) shows that the chosen macro labels are good segmentations of the raw data. In future work, the representations learned by these models could be used as context vectors that can be incorporated into the shared autonomy paradigm. We give the results on the classification problem in Section 5.3.

5.3 Experimental results

The results of the experiment for both real and synthetic data sets for both micro and macro tasks are shown in Table 5.2. Following the initial experiments, an analysis of individual differences in the real world was performed for both the micro and macro task. For this, we used k-fold cross-validation, where each fold was a single participant, as seen in Table 5.5. The accuracy and mean average precision (mAP) are reported for all experiments. For both micro and macro actions, the chance values were calculated by taking the inverse of the number of classes (three for micro action classification and five for macro action classification). The majority class values are given in Table 5.1 for both data sets and tasks. These are calculated by dividing the total number of sequences of a particular class by the total number of sequences in the entire data set.

5.3.1 Micro action results

The best performance for synthetic data in both metrics came from the joint late fusion model with a hidden size of 16 and inputs represented as the difference of the raw signal. The 0.9710 accuracy score outperforms both guessing at chance (0.3333) and consistently guessing the majority class (0.4321). Given that these results are an idealized version of the real-world data, these numbers represent a theoretical upper bound on performance.

The real data also outperformed chance and guessing the majority class (0.3333 and 0.5541). The best results for these data were obtained with the eye-only model, with the best representation being the difference of the input signal. Accuracy was best under the 256 hidden size model, while mAP performed the best under the 16 hidden size model, but both models performed similarly in both metrics.

5.3.2 Macro action results

Synthetic data performed well in all categories for macro action classification (Table 5.2). In all cases, it outperformed chance (0.2) and guessing the majority class (0.3377). The best performance was achieved by the late fusion model with a hidden size of 16 and input streams represented as the difference of the raw signal. Both fusion models significantly outperformed the single-stream models.

The real data also outperformed chance (0.2) and guessing the majority class (0.2424) in all experiments, with the best performance resulting from the early fusion model with hidden size of 256 and the original input stream as the input to the model. Although this model performed the best, performance on the late fusion model and the joystick only model were similar.

5.3.3 Participant level cross validation

Table 5.5 shows the accuracy and mAP scores when each participant is considered as their own test set for the micro and macro tasks. The micro task should be compared to the eye only model, dx, dy, 256 models, and the macro task should be compared to the eye+joy (e), x,y, 256 model. Evidence was found to suggest individual differences in the macro action classification, but not in the micro action classification.

5.4 Discussion

Improved performance in the synthetic data for both classification tasks was realized by jointly modeling eye gaze and joystick data. For real data the best results for micro action classification came by modeling eye gaze alone, with the joystick model only slightly outperforming consistently guessing the majority class. This indicates that in real data, the patterns of joystick behavior do not differ significantly across micro action sequences, further justifying the need to create and analyze macro actions to encompass user control input signals.

In macro action classification, the joystick and both fusion models performed similarly, with slight improvements coming from the fusion models. Thus, outperforming the joystick signal is possible in this task, but the fusion between the eye gaze and joystick signals is non-trivial and should be explored further. All models significantly outperformed guessing the majority class, indicating that eye gaze and joystick display distinct patterns within the proposed primitives.

Another finding was that the model's preferred input representations were consistent across task and data set. The best performance from both synthetic and real data on the micro and macro classification tasks came from representing the eye gaze as the original, real-valued input. The joystick stream was best represented as the difference in the synthetic data and as a real value in the real data, but, since the joystick in the synthetic data actually represents the end effector of the robot, taking the diff of this signal actually results in a signal similar to the joystick data in the real data set.

Additionally, we found no consistent improvement when considering wider hidden representations, indicating that smaller models can be used to achieve this task. This finding is important because assistance algorithms must be processed online, and the use of smaller and lighter models makes this approach more viable. Furthermore, there was no appreciable difference between the early and late fusion models.

Finally, Table 5.5 shows that micro actions are consistent between participants, as the weighted average of accuracy over the 19-fold cross validation compared similarly to the accuracy of the best eye only model (though mAP underperformed). This is in contrast to the macro action categories, Table 5.5, in which the weighted average of both accuracy and mAP significantly under performed the best early fusion model. This indicates that there are individual differences, and the building of effective models to account for these differences should be further studied.

5.5 Conclusion

In this work we are motivated by the need to understand hand-eye coordination for human-robot comanipulation. This problem is especially important for assistive robotics tasks, in which operators could greatly benefit from the introduction of indirect control signals, such as eye gaze, into assistance algorithms. We introduced a novel concept of macro actions, which are semantic action primitives that represent high-level task activities. These macro actions are complementary to micro actions, which represent low-level behavior. We defined five macro actions that combine eye gaze and joystick in an assistive eating task drawn from the HARMONIC data set. We then developed multimodal models of micro and macro actions, and extensively analyzed the models' performance under different parameters. Our analysis further justifies the need for semantic macro primitives, and highlights the benefit of jointly modeling eye gaze and joystick signals within a single task. Finally, we discussed how participants show individual differences. This work will enable future research into the combination of indirect and direct control signals to fully perceive human goals in comanipulation settings.

In this work we additionally found that there is a large variance in the performance of our model as it generalizes across individuals, suggesting that there is a significant need for representations of people's behaviors to be personalized to individuals, i.e. generalized feature spaces are unlikely to be able to provide value-aligned assistance. Toward this end, we explore how we can use large-scale pretrained feature spaces, which have been shown to be easily conditioned using in-context learning, for personalization in object rearrangement tasks. Additionally, through this work, we realized that while the data provided in HARMONIC provide rich naturalistic behavior, the simple surface rearrangement task does not accurately reflect the open-vocabulary and sequentially dependent nature of the household tasks we are aiming to emulate. As such we introduce a new domain in the following chapter and discuss how we can use foundation models to personalize a pretrained feature space.

Table 5.2: Results for the synthetic and HARMONIC datasets on the micro classification task. We report accuracy (acc) and mean average precision (mAP). We test a variety of input streams: eye gaze only (E), joystick only (J), an early fusion of eye gaze and joystick (E+J-e), and a late fusion of eye gaze and joystick (E+J-l). Additionally, we test different embeddings (embed) of the inputs: binary, real (raw), and difference (del). Finally, we test two different hidden unit sizes ($\|h\|$): 16 and 256.

			Micro			
input	embed	hsize	acc (synth)	acc (real)	mAP (synth)	mAP (real)
E	raw	256	0.9186	0.6964	0.9670	0.6700
E	0/1	256	0.9620	0.6648	0.9949	0.6666
E	raw	16	0.9316	0.6774	0.9738	0.6716
E	del	256	0.8656	**0.7076**	0.8703	0.7027
E	del	16	**0.9695**	0.7006	**0.9963**	**0.7028**
J	0/1	256	0.8671	0.5182	0.9068	0.5037
J	raw	256	0.9141	0.5561	0.9464	**0.5248**
J	raw	16	0.9016	0.5610	0.9328	0.5148
J	del	256	0.9456	**0.5638**	0.9895	0.5112
J	del	16	**0.9486**	0.5372	**0.9902**	0.5079
E+J-e	0/1	256	0.9416	**0.6830**	0.9804	0.6668
E+J-e	raw	256	0.8971	0.6669	0.9242	0.6822
E+J-e	raw	16	0.9226	0.6767	0.9555	0.6895
E+J-e	del	256	0.9575	0.6522	0.9955	0.6608
E+J-e	del	16	**0.9650**	0.6669	**0.9956**	**0.6915**
E+J-l	0/1	256	0.9271	0.6767	0.9776	0.6864
E+J-l	raw	256	0.9386	0.6669	0.9861	0.6782
E+J-l	raw	16	0.9051	0.6697	0.9434	0.6869
E+J-l	del	256	0.9640	**0.6957**	0.9954	**0.6930**
E+J-l	del	16	**0.9710**	0.6613	**0.9977**	0.6709

Table 5.3: Results for the synthetic and HARMONIC datasets on the macro classification task. We report accuracy (acc) and mean average precision (mAP). We test a variety of input streams: eye gaze only (E), joystick only (J), an early fusion of eye gaze and joystick (E+J-e), and a late fusion of eye gaze and joystick (E+J-l). Additionally, we test different embeddings (embed) of the inputs: binary, real (raw), and difference (del). Finally, we test two different hidden unit sizes ($\|h\|$): 16 and 256.

			Macro			
input	emb	$\|h\|$	acc (synth)	acc (real)	mAP (synth)	mAP (real)
E	raw	256	0.7763	0.4884	0.5981	0.4522
E	0/1	256	0.8310	0.4884	0.7266	0.4431
E	raw	16	0.8539	0.4653	0.7713	0.4459
E	del	256	0.9185	0.5347	**0.9237**	0.4805
E	del	16	**0.9195**	**0.5644**	0.9050	**0.5054**
J	0/1	256	0.7783	0.6436	0.6274	0.5886
J	raw	256	0.8062	**0.6931**	0.6710	**0.6797**
J	raw	16	0.7962	0.6502	0.6599	0.6642
J	del	256	0.8917	0.5479	**0.8230**	0.5241
J	del	16	**0.9006**	0.5809	0.8203	0.5666
E+J-e	0/1	256	0.7952	0.6205	0.6595	0.6244
E+J-e	raw	256	0.8082	**0.6997**	0.6847	**0.7053**
E+J-e	raw	16	0.8956	0.6832	0.8456	0.6682
E+J-e	del	256	**0.9652**	0.5578	0.9664	0.5452
E+J-e	del	16	0.9543	0.5248	**0.9706**	0.4748
E+J-l	0/1	256	0.8678	0.6271	0.7979	0.6357
E+J-l	raw	256	0.8111	0.6238	0.6702	0.6613
E+J-l	raw	16	0.7873	**0.6964**	0.6680	**0.6997**
E+J-l	del	256	0.9612	0.6139	0.9707	0.5889
E+J-l	del	16	**0.9662**	0.5314	**0.9792**	0.5257

Table 5.4: Results for the HARMONIC dataset on the macro task using the best representations from Table 5.2.

Input	gaze embed	gaze hsize	joy embed	joy hsize	acc	mAP
eye+joy(ef)	dx,dy	16	x,y	16	0.6799	0.6805
eye+joy(ef)	dx,dy	256	x,y	256	0.6403	0.6197
eye+joy(lf)	dx,dy	16	x,y	256	**0.6997**	**0.7147**

Table 5.5: Micro action cross validation by participant ID (eye only, dx,dy, 256 model). † Data were excluded due to significant noise. ‡ Data were not labeled for macro actions.

	Micro		Macro	
ID	Accuracy	mAP	Accuracy	mAP
p101	0.7021	0.7571	0.5587	0.6551
p102	0.6651	0.6396	†	†
p105	0.7317	0.6690	†	†
p106	0.7161	0.7089	0.6818	0.7742
p107	0.6140	0.6658	0.6023	0.5704
p108	0.7901	0.7346	0.7005	0.7010
p109	0.6984	0.6458	0.6447	0.6140
p110	0.6889	0.7492	0.5209	0.5921
p111	0.5305	0.5527	0.6204	0.6111
p112	0.7105	0.7313	0.4638	0.5165
p113	0.6549	0.6851	‡	‡
p114	0.7111	0.7250	‡	‡
p115	0.7333	0.7019	‡	‡
p117	0.8023	0.6709	‡	‡
p118	0.8175	0.6650	‡	‡
p119	0.5987	0.5872	‡	‡
p121	0.7128	0.7097	‡	‡
p122	0.6942	0.6745	‡	‡
p123	0.6915	0.7144	‡	‡
Avg	0.6923	0.6429	0.5951	0.6233

6 Personalized feature spaces in complex surface rearrangement tasks

Assistive home robots should complete tasks in ways that align with user preferences [221]. Consider the task of setting plates and utensils on a table for dinner. The *correct* solution for this task is subject to your preferences for the type of utensils you want to use and how you prefer to arrange them. However, modeling individual preferences, especially those that consider fine-grained features such as the real-valued location of objects and their relative placements, is challenging, as such preferences are subjective, difficult to specify explicitly, and vary from person to person. In this work, we explore this problem through the example of personalized dinner table arrangement and determine a personalized task plan for setting a dinner table given examples of a person's preference.

Prior research on multi-object rearrangement collects problem-specific datasets of simulated or human demonstrations that represent personal preferences for completing a household task. They hypothesize that collecting large data sets of task-specific preferences is sufficient to train a feature space that can generalize to novel participants who were excluded from training data [2, 145, 147, 223]. However, these methods often yield poor generalization to these held-out participants due to the large effects from individual differences. These challenges indicate that it is difficult to collect a large dedicated dataset to train a feature space that covers the unbounded space of personal preferences for even a single task and suggests that generalized pre-training is a potentially fruitful path of research. These ideas are further compounded when considering under specified preferences and preferences that are influenced by cultural norms or decorum.

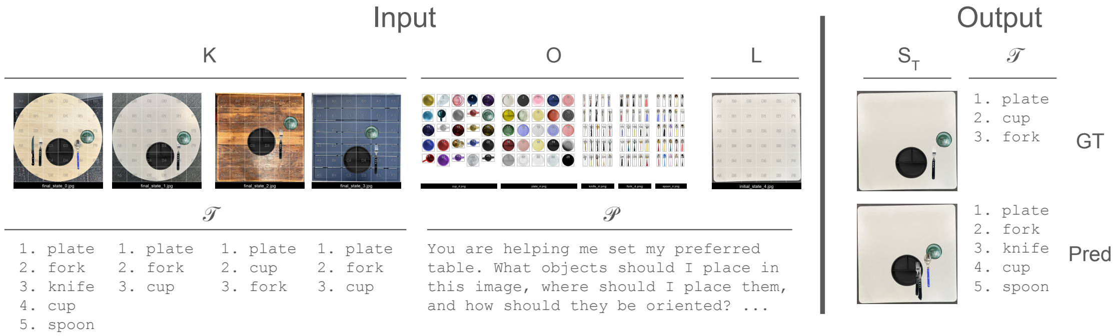


Figure 6.1: DegustaBot takes in a single person's preferred table arrangements (shown here as the visual context k_i and order of object placement $\mathcal{T}_i$), the objects O from which the algorithm can select, the table to set s_0 and the task prompt ℓ. The robot then produces a task plan in the form of an object arrangement (visualized as an image s_T) and the order of object placement $\mathcal{T}$. This predicted arrangement, Pred, should match a held-out preference created by the user, GT.

Large-language models (LLMs) may provide a solution. They are pretrained on internet scale data, and have shown the ability to solve tasks upon which they weren't explicitly trained. They can leverage knowledge about the abstract external concepts that often guide personal preferences. In fact, it has been shown they can infer generalizable user preferences from few examples with in-context learning [343]. Representing multi-object preference learning solely in language, however, removes important visual information about a task. At the same time, these problems cannot be solved *purely* visually, as they do not have a grounded vocabulary of actions over which to reason. Therefore, we developed a new prompting technique that provides such a vocabulary. We built a virtual tabletop setting task, and collected a large dataset of human preferences for different arrangements in order to test this technique. Concretely, in this paper, we make the following contributions:

- DegustaBot: a novel method for implicit visual preference learning to find personalized solutions to fine-grained multi-object rearrangement tasks.
- A novel human evaluation used to evaluate personalized multi-object rearrangement agents.
- A dataset of naturalistic preference data in a simulated multi-object rearrangement task.

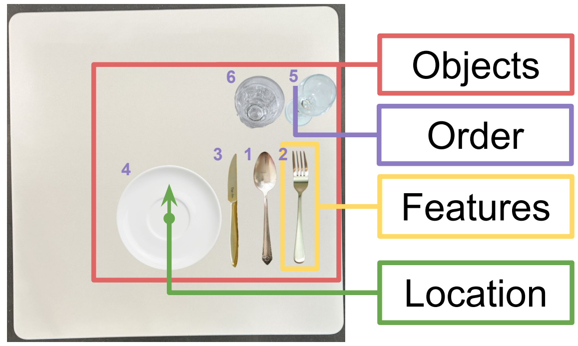


Figure 6.2: Details of a table arrangement. An arrangement is described by the objects within the arrangement, the order in which they are placed, their features (such as color, shape and material, and their location and orientation.

6.1 Problem specification

People have widely varying preferences, making it difficult to develop one-size-fits-all solutions, even for problems that at first glance seem to be well-specified. One person may consider a table to be arranged with a simple fork and a bowl, while another will not sit down to eat until their flower arrangement is perfect and a candle is lit. Furthermore, these preferences cannot easily be specified in natural language: language is often either ambiguous or requires cumbersome description.

Instead, we argue that preferences can be specified implicitly by visual representations of the final state of a multi-object rearrangement task.

To investigate this, we choose table setting as an example application. Concretely, the goal of a preference model $\mathcal{M}$ is to maximize the likelihood of seeing a state $s^* \in S^*$ given an initial state s_0, a preference K, and a language prompt ℓ, $\max_{s \in S^*} \mathcal{M}(s \mid K, O, s_0, \ell)$. A state s is a table arrangement composed of available objects O and locations L, the currently placed objects $o \in O$ and their locations $l \in L$. A state can be represented as an image, for example, a picture of a dining table, or in language, such as the description of a table setting. A preference indicates a solution set $S^* \in S$, which indicates the set of all final states that are acceptable under that preference. We represent a preference as a history of demonstrations $K = [k_0, k_1, \ldots, k_C]$ where each element is the concatenation of the state and object placement order $k_c = \left[s_c^T, \mathcal{T}_c\right]$.

An arrangement s of length T is a sequence of tuples that contain information about positioning $\mathbf{l}$ and object $\mathbf{o}$ information, that is, $s_{1:T} = [(\mathbf{l}_1, \mathbf{o}_1), \cdots, (\mathbf{l}_T, \mathbf{o}_T)]$. $\mathbf{l} \in \mathbb{R}^3$ represents the continuous elements of an arrangement (x- and y-positions of an object on the table surface and r, the rotation of the object), and $\mathbf{o}$ represents the features of an object. Objects can be represented visually, as images, or in text, by a short description of their visual appearance.

As in previous work, [217], we consider a family of functions Ω that lift non-visual information into the image domain. An arrangement can be lifted into the visual domain with $\Omega_s : s \to I_s \times \mathcal{T}_s$, where I_s is a visual representation of the arrangement, and $\mathcal{T}$ is a language description of the order in which objects were placed. We also consider $\Omega_O : O \to I_O$, which jointly maps all objects to the visual domain.

Objective. In this work, we want to investigate whether foundation models trained on internet scale data and conditioned on implicitly defined preferences can produce personalized task plans. In other words, we consider the following problem:

$$\mathcal{M}_g(K, O, L, s_0, \ell) \stackrel{?}{=} \mathcal{M}_H(O, L, s_0, \ell) \text{ where } K \ \sim \mathcal{M}_H,$$

where $\mathcal{M}_\}$ is a model parameterized by parameters g, and $\mathcal{M}_{H_k}$ is a human with preference k. We parameterize M_g as a vision and language model. While there have been a surfeit of large foundation models recently becoming available, we choose to investigate three high-performing, recent models: OpenAI's GPT-4o [91], Anthropic's Claude-3 Haiku [89], and Google's Gemini 1.5 Pro [90].

6.2 Estimating simulated table arrangements using VLMs

Solving our implicit, visually represented tabletop arrangement task requires incorporating visual information across several images (e.g., k, o, s_0). We present and evaluate several different zero-shot methods for prompting VLMs to produce task plans that solve table-top arrangement problems. We do this through *lifting functions* which add extra information to an image before sending it to a VLM like GPT4v along with a query, such as adding a grid or highlighting a point in space.

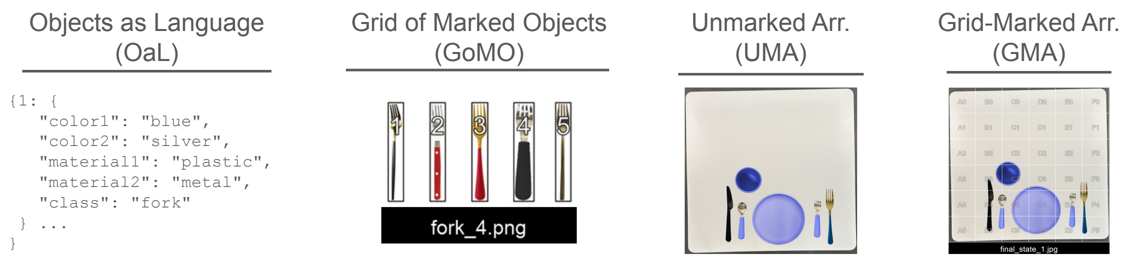


Figure 6.3: Object and arrangement lifting functions, from left to right: OaL, a representation of objects as language, for this we use the json representation; GoMO, grid of marked objects, which represents objects visually with referential marks overlaid on each object; UmA, Unmarked Arrangement lifts the arrangement into the image domain; and finally GMA, Grid-Marked Arrangement, which overlays a spatial reference grid on the continuous table top space.

6.2.1 Producing task plans through visual prompting

We evaluated different lifting functions Ω_s and Ω_O to solve a personalized tabletop arrangement with zero shots. In addition to lifted images, we include a text prompt that describes the problem and the desired format of the output.

Selecting Objects

We first describe the object lifting functions, Ω_O:

Objects as Language We take each object in our dataset and represent it as a dictionary of features, which are concatenated to create a dictionary of all objects and their representations (see Fig. 6.3, left).

Grid-of-Marked Objects. We represent objects visually by creating grids of objects within a category. Each object is overlaid with a spatial reference marker to aid in object detection [350] and the image is labeled with object type and preference number (see Fig. 6.3, second column).

Selecting Locations and Rotations

We represent the arrangements visually through Ω_s in two ways:

Unmarked Arrangements. We assume that positions $\mathbf{x}$ represent an object's centroid. Then, representative images are resized to fit within an initial arrangement (e.g. an empty table) and are pasted at **x**. They are then rotated according to the rotation **r**.

Grid-Marked Arrangements. We turn predicting a continuous location into a multiple choice question over grid cells that serve as spatial reference markers. First, we overlay a spatial grid on each preference image and then ask the model to produce all the grid cells that an object will intersect with on a table after it is placed. We average the centroids of the responses over multiple responses to get a continuous position prediction.

Combining Lifting Functions into Visual Prompting Methods

We combine these lifting functions into four methods that test each models ability to predict task plans in table-top arrangement tasks as follows:

- **DegustaBot-LOUMA.** uses Language Objects with UnMarked Arrangements
- **DegustaBot-LOGMA.** uses Language Objects with Grid-Marked Arrangements
- **DegustaBot-MOUMA.** uses Marked Objects with UnMarked Arrangements
- **DegustaBot-MOGMA.** uses Object Images with Grid-Marked Arrangements.

6.2.2 Simulating preferences for table arrangement

We develop a small data set of simulated preferences for controlled evaluation. These preferences are drawn from three positional preferences, three object preferences, and a single order preference. We design these preferences to approximate a standard, a semi-standard, and a non-standard table setting. To test our model over a distribution of positions, we add positional and rotational noise to these table settings. The object preferences are also drawn from three color preferences: a preference for all red, all blue, or all yellow objects.

We display arrangements on two visually distinct tables. In total, we simulated 18 preferences on two table tops for a total of 36 preferences. We tested these in both a reconstruction task (predict on the same table) and a generalization task (predict on a different table) for 72 total experiments.

6.2.3 Validation metrics

We test two main hypotheses: (1) that large vision and language models develop a feature space that is useful for estimating preferences for real-valued multi-object rearrangement tasks; and (2) that visually grounding preference information aids in estimating complex preferences.

Object selection: this measures a method's ability to select objects consistent with the preference. We compute this as the accuracy of object selection: did the model select exactly the right object?

Position Selection: We take a geometric approach to comparing arrangements. While a point-to-point comparison seems immediately intuitive, the geometric relationship between points in a particular table setting are important, as well. For example, a table settings performing poorly when measured with point-to-point correspondences may still be similar geometrically to the preferred arrangement, only displaced slightly. To account for these issues, we introduce another evaluation metric: root mean squared deviation (RMSD).

We use the Kabsch algorithm [142] to register matched points between predicted and ground truth arrangements, and find the translation vector g and rotation matrix r that minimize the root mean squared deviation between the two paired frames:

$$\min_{r\in R, g\in G} \sqrt{\frac{1}{N}\sum_{n}^{N}(\bar{A}_n^{gt} - R(\bar{A}_n^{pred} - g))^2}.$$

We report this minimized RMSD as our geometric evaluation metric.

6.2.4 Results

We test each model with each combination of visual lifting functions and report results on how well the predicted image matches the ground truth scene's geometry, as reported through RMSD, and chosen utensils, as reported through accuracy. These results are shown in Fig. 6.4.

First, we see that GPT-4o outperforms other choices of VLM in both RMSD and object prediction, under all methods. We can also see that grid-marked methods tend to improve a model's ability to match scene geometry. Finally, we see that visual representations of aid in correct prediction. The results for Gemini-1.5-Pro, LOUMA

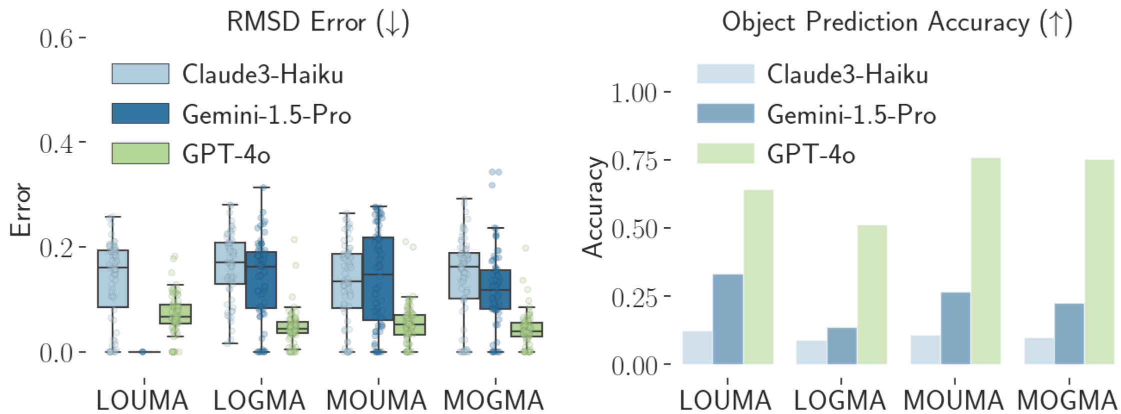


Figure 6.4: Quantitative results for evaluating DegustaBot on simulated preferences. Left shows the performance of each model's ability to capture the geometry of the table arrangement, as measured by RMSD. Right shows each model and method's accuracy when choosing items to place in the arrangement. GPT-4o and MOGMA perform the best on both metrics.

are missing due to a 97% failure rate on this method. Taking these results together, we see that DegustaBot-MOGMA using GPT-4o outperforms all other methods on this task. We will use this method to analyze naturalistic preferences in the next section.

6.3 Naturalistic tabletop arrangement dataset

To understand whether this approach works with naturalistic data, we collected a dataset of preferences for table arrangements through an online study. We then assessed the performance of the model, as in Sect. 6.2.4.

6.3.1 Data collection

Our user study consisted of a task with two stages: *Preference Elicitation* and *Self-Evaluation*. During Preference Elicitation, we prompt participants to create their preferred table-top arrangements. In Self-Evaluation, participants rate these arrangements, as well as several transformations of their arrangements. We used these ratings to develop a subjective analysis of the performance of the model. The

	RMSD Threshold		
	0.01	0.05	0.10
$Acceptance_{random}$	0.714	0.300	0.169
$Acceptance_{structured}$	0.992	0.973	0.965

Table 6.1: We report the percentage of people who find scenes below a specific threshold hold acceptable under both the random and structured noise condition. We find that people are extremely sensitive to random noise, and not very sensitive to structured noise. We see that when the RMSD is below 0.01, between 71.4% and 99.2% of people find the scene acceptable.

participants saw a total of six rounds of this task: a practice round, followed by five experimental trials.

We collected a data set of 995 table arrangements drawn from 199 participants. Participants were paid USD$4.00 for completing the 20 minute task. All participants were located in the United States at the time they completed the task. We did not collect additional demographic information about our participants. This study was approved by the IRB of record.

Preference Elicitation

Participants were first presented with an image of one of five different empty tables, and asked to to set the table by selecting objects from five object categories: plates, spoons, forks, knives, and cups. After selecting an object category, participants were shown all instances in that category simultaneously. Each category contained 25 objects of varying shapes, colors, and materials, yielding an object dataset of 125 Creative Commons licensed and public domain images of table setting utensils.

Participants were then presented with a simple drag-and-drop interface through which they could place the object anywhere on the table's surface and rotate it object to their liking. The participants continued this procedure until they were satisfied with the arrangement. The only restriction placed on table arrangements was that participants were required to set a minimum of three objects. After completing a table arrangement, participants moved on to the self-evaluation portion of the task.

Method	RMSD Threshold 0.01	0.05	0.10
Post Jitter Correction	0.123	0.787	0.920
Copy and Paste	0.059	0.586	0.745
Human Performance	0.010	0.432	0.669
KNN	0.002	0.052	0.182
DegustaBot-MOGMA	0.000	0.219	0.672
DegustaBot-LOUMA	0.009	0.299	0.659
DegustaBot-LLM	0.034	0.446	0.623
DegustaBot-LVM	0.041	0.483	0.666
DegustaBot-VLM	0.036	0.405	0.597

Table 6.2: We compare our methods' performances against several baselines at several RMSD thresholds. Post Correction Jitter, which we consider our upper bound, produces model acceptable model responses 12.3% of the time at the 0.01 threshold, 78.7% of the time at the 0.05 threshold, and 92% of the time at the 0.1 threshold. We see that the Copy and Paste baseline, which uses privileged sensing information, has the second best response rate. Our Vision Model method performs the best out of methods that do not use privileged information and, interestingly, is closely aligned with the Human Performance baseline.

Method	RMSD Threshold 0.01	0.05	0.10
DegustaBot-LLM	0.034	0.443	0.623
DegustaBot-LLM-AR	0.071	0.501	0.630

Table 6.3: We compare DegustaBot-LLM to an autoregressive variant. We report thee results on a random subselection of 445 preferences, due to the significant cost increase of running an autoregressive variant. We see siginificant improvements over the LLM model at lower RMSD thresholds, suggesting that autoregressive models perform better than their non-autoregressive counterparts.

Figure 6.5: Qualitative results. On the top line we see ground truth images from our naturalistic preference data and on the bottom we see DegustaBot-LOGMA's predictions, lifted into the image domain. On the left side of the image we see three examples where DegustaBot-LOGMA predicts similar arrangements to the ground truth image. On the right we see some failure cases.

Subjective Acceptability

We then asked the participants to evaluate several arrangements.

Rating the initial scene. After creating an arrangement, s_{intial} participants were prompted to answer the question, *Please rank your agreement to the following statement: This is my preferred table arrangement*, sliding scale from 0-100 to offer their baseline acceptability score b_0.

Jittering the initial scene. To understand what errors affect people's sense of an acceptable arrangement we randomly transformed s_{intial}. A transformation magnitude between 0 and 1 was randomly selected as t_j. The scene s_{intial} was then translated and rotated by this amount to create s_{jitter}. Participants then gave a second acceptability rating b_{jitter}.

Threshold of Acceptability. We take the difference in ratings $b_{initial} - b_{jitter}$, and correlate this with the distance between $s_{initial}$ and s_{jitter}, giving us a **subjective acceptability** measure. We choose a rating difference threshold of 0.2, which is interpreted as the threshold at which people find s_{jitter} to be at least as acceptable as $s_{intitial}$. We then report the percentage of people who meet this acceptability threshold at various levels of RMSD. This tells us the percentage of people who are

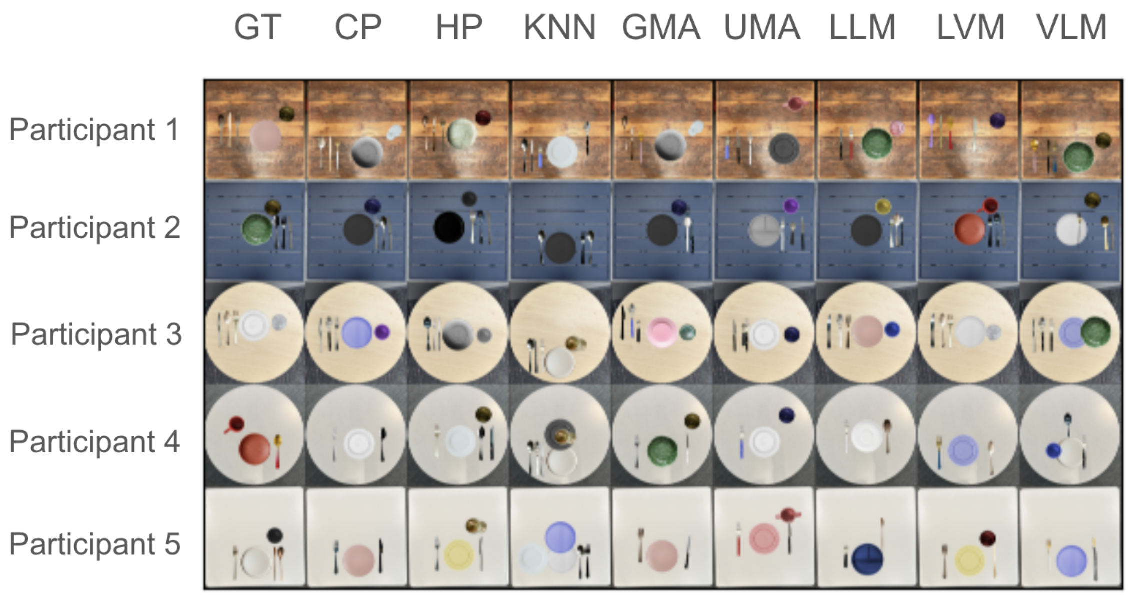


Figure 6.6: Additional Qualitative results. We show the qualitative performance of all methods on five random experiments from five random participants. From left to right we show the Target image (GT), Copy and Paste (CP, Human Performance (HP), KNN (KNN), DegustaBot-MOGMA (GMA), DegustaBot-LOUMA (UMA), DegustaBot-LLM (LLM), DegustaBot-LvM (LVM), DegustaBot-VLM (VLM).

likely to find a scene acceptable below a given RMSD threshold.

6.3.2 Additional baselines and methods for prompting FMs

We introduce several additional baselines and methods for prompting foundation models with naturalistic arrangement data.

Participant Correction We report the RMSD between $s_{correct}$ and $s_{initial}$. This method reflects how an individual participant recreates the scenes after it has been jittered. This method provides a good upper bound for a method's performance, as it reflects what the original participant cared about when correcting the arrangement.

Copy and Paste. We consider a baseline that uses the configuration of a randomly selected input arrangement as the predicted arrangement. This method has privileged information in that we assume that it has perfect perception. This method aims to understand the performance of a method that

Human Performance. We perform a second user study that prompts people to do the same task as the model: given four arrangements, create a new arrangement that matches the preference represented. We collect these annotations for 884 of our original scenes. Participants were paid $4.00 for the completion of the 20 minute task. Each participant annotated 5 unique preferences from the original dataset.

KNN. We divide our preference data set into training and testing sets (80/20 split) and train a nearest neighbor model over the average CLIP [83] features of the input images. We find the closest preference in the training set to the average CLIP features of the prompt from the testing set, and use the held-out arrangement from this set as the prediction for this method. We perform 5 fold cross-validation to cover the whole dataset.

DegustaBot-LLM. DegustaBot Large Language Model represents the entire problem as text. Inspired by TidyBot [343], we represent the inpute preferences as short snippets of code with the objects, a summary of the initial state, several pick and place actions that take an object ID, and a real valued x and y location. We end each snippet with a short summary of the final scene. Objects are represented by a dictionary mapping object ids to a short description of the object, which was annotated by GPT-4o. Summaries of initial and final scenes were also annotated by GPT-4o.

DegustaBot-LLM-AR. We consider an autoregressive version of DegustaBot-LLM. We only consider an autoregressive version of the language variant of our model due to the significant cost of running an autoregressive version of the vision model.

DegustaBot-LVM. DegustaBot Large Vision Model represents the objects, initial scene, and final scene as images. Code snippets to represent the task plan are kept as text. For this method, we annotate object images with their object number. Additionally, we annotate the final scenes with bounding boxes, as provided by OWL-ViT [203], and their object ID.

DegustaBot-VLM. Finally DegustaBot Vision Language Model uses both representation from the previous two methods to prompt the model. Objects are represented both a images and text, as well as initial and final scenes.

6.3.3 Results

We evaluated our method on naturalistic preference data. In Tab. 6.4 we report object prediction accuracy and RMSD as the number of preferences images increases from 0, 2, to 4. We see that increased context length improves performance metrics, but potentially with diminishing returns. This indicates that 1) human preferences are diverse and not easily predicted in a context-free setting, and 2) our method makes efficient use of few context examples to predict a person's preference. We also see that the performance here is well below the performance on the synthetic data, indicating that visual preference learning from naturalistic preference data is a difficult task.

In Tab. 6.1 we report the rate at which people find scenes under certain RMSD thresholds acceptable. In Tab. 6.2 we report the rate at which our various methods and baselines produce arrangements below these thresholds, as measured by the RMSD distance between the scene our model produces and a held out arrangement. From these data, we can see that 1) people are highly sensitive to the jittered scenes we produced, making this a strict metric and 2) our best performing model (DegustaBot-LVM) produces reasonable scenes at both low and high thresholds, when compared to the Copy and Paste and Human Performance baselines. We see that the vision model is the best performing of the language and vision ablations, suggesting that visual information is important for completing this task. We also see that our VLM models are in line with Human Performance, suggesting both that this is a hard task and that our models are doing reasonably well at it. Finally, we see that the Copy and Paste baseline performs the best, suggesting that if you have privileged object detection information and privileged preference information, you can improve performance at this task.

In Tab. 6.3 we compare DegustaBot-LLM with an autoregressive variant, DegustaBot-LLM-AR. In DegustaBot-LLM-AR, we update the state with the output of a previous model iteration until the model produces no more actions to add to the task plan. We see that this method outperforms DegustaBot-LLM on all RMSD thresholds, and suggest that autoregressive methods would improve performance on this task and make them more generalizable to real robotics tasks.

Finally, in Figs. 6.5 and 6.6 we show several qualitative results. In Fig. 6.5, each column depicts a different participant within our data set. The top row displays

Length of k	0	2	4
Object Prediction Accuracy (↑)	0.037	0.090	0.089
RMSD (↓)	0.173	0.109	0.100

Table 6.4: We report object prediction accuracy and RMSD of DegustaBot-MOGMA on naturalistic preferences. We see that with increased context length our method achieves higher performance, though perhaps with diminishing returns.

ground truth images, and the bottom rows depict our method's prediction. When geometric patterns are regular, DegustaBot-LOGMA does well to predict them.

6.4 Using corrective actions to refine pretrained feature spaces

Online IRL relies on people exhibiting behaviors that are highly correlated with their goals, often leading to action representations that do not benefit from shared structure, leading to poor sample efficiency and out-of-domain generalization. To study this, we consider two different choices of action representation: per-ID, ψ^{ID}, which represents actions as one-hot vectors, and per-quality, ψ^{Quality}, which represents actions by shared properties between different actions, such as the material, shape or size of an object in the dishwasher loading domain. We aim to show how ψ^{Quality} can improve sample efficiency and generalizability when people's task objectives are correlated along their actions, for example, people place glass bowls into the dishwasher in a similar manner to glass cups.

We first develop a set of simulated user objectives, g. Eight highly correlated objectives, shown in the top row of Fig. 6.7 are blended with randomly sampled objectives according to $g = g_{\text{correlated}} * (1-\beta) + g_{\text{rand}} * \beta$ along five correlation thresholds $\beta \in \{0, 0.01, 0.1, 0.5, 1\}$, shown in the bottom row of Fig. 6.7, to develop objectives with varying degrees of correlation along actions. We then train a randomly initialized robot objective,s $\hat{g}$, to approximate these user objectives, g, using Algorithm 1.

Furthermore, we create a set of O objects that share features with another object along at least one quality (e.g., a glass mug, a glass bowl, a plastic mug, and a plastic bowl) and represent our surface by a grid of L locations. Objects are divided into

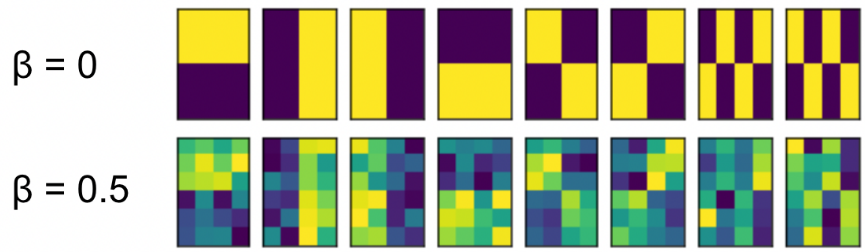


Figure 6.7: Highly correlated objectives across qualities (top) and those same objectives blended with noise ($\beta = 0.5$). Rows of an objective represent surface locations and columns represent objects. Elements represent the value associated with placing that object in that location, ranging from $[-1, 1]$.

training and testing sets.

Per-ID representations return one-hot vectors with the element representing the object or location ID activated. For objects, this results in an action space U_{obj}, that is a square identity matrix of size $|O| \times |O|$. For locations, U_{loc} is a square identity matrix of size $|L| \times |L|$.

Per-quality representations share features across different object categories by concatenating one-hot vectors over each quality (e.g., material or shape), $q \in Q$. Per-quality representations of objects return an action space U_{obj} that is of size $|O| \times \sum_{q \in Q_{\text{obj}}} |q|$, or similarly for locations, U_{loc} of size $|L| \times \sum_{q \in Q_{\text{loc}}} |q|$. Per-ID representations imply a g of size $|O| \times |L|$, while per-quality representations imply a g of size $\sum_{q \in Q_{\text{obj}}} |q| \times \sum_{q \in Q_{\text{loc}}} |q|$. When the total size of the qualities is less than the number of objects they describe, per-quality representations yield a more efficient representation of the space.

Each line in Fig. 6.8 shows a regret curve for a separate objective (ranging from highly correlated across qualities to randomly correlated across features). Shaded regions showing the standard error on the collection of runs. The mean average regret (mAR) for a policy taking random actions is shown in black, and the reference zero line is shown in dotted brown. Per-ID results are shown on the left, and per-quality representations are shown on the right. Training mAR curves are shown on top and zero-shot performance on all objects on the bottom. Each increment along the x-axis is an episode ranging between one and six object placements.

We report regret over other potential metrics, such as accuracy or corrections, because it reflects the underlying reward our algorithm receives. Accuracy and corrections are both overly critical metrics: accuracy penalizes an algorithm for choosing an "incorrect" placement even if it returns the same reward as the "correct" label, while corrections penalize all incorrect placements equally, even when differences in reward received may be negligible.

These results show both representations capture objectives with strong correlations across object qualities and episodes of in-domain objects. For highly correlated objectives, per-quality representations converge slightly faster than per-ID representations, though they underperform as objectives become less correlated.

These results also show that per-quality representations can generalize to out of domain objects, when objectives are highly correlated. This is because these representations can express preferences such as "place all glass objects on the top", something per-ID representations are unable to do. Taking into account the results of in-domain training, we show how choosing an appropriate representation space can improve upon the drawbacks of online IRL, namely that it can be sample inefficient and overfit to in-domain data.

6.5 Limitations and Generalizability

In this work, we developed a method that can serve as a preference perception pipeline for a robot that performs household multi-object rearrangement tasks, DegustaBot. While we only test this method in a table setting domain, it should be generalizable to any surface arrangement task where the preferences can be represented as a planar view of an image. Although this is a strong assumption, many similar works in robotics make a similar assumption, either in simulation [347] or on real robots equipped with suction grippers [148]. Representing visual preferences from non-planar views for foundation models is an open question and important future research. We think this work could also be easily extended to work tasks where the object placement locations would not all fit in frame, for example putting clothes away throughout a house, by including a model of the house and representations of all of its various placement locations.

Although we do not perform this experiment on a real robot, this setting emulates suction cup pick-and-place routines that have been shown to work reliably for similar surface rearrangement tasks [148]. We introduce DegustaBot-LLM-AR, which is an autoregressive version of our model, which takes in the current state information to make a prediction. This method shows success, suggesting that our method would be generalizable to real robotics situations which rely on current state information due to unforeseen errors, such as a cup falling over after being placed.

Finally, our main focus is to develop a method that generalizes to various types of preferences. Demonstrating our methods' performance over our naturalistic data set does demonstrate this ability. Additionally, we demonstrate that our model can generalize across locations, by changing the table representation in the model prompt. However, it would be good to test our methods' ability to generalize a preference across different objects and tasks, as well. This would require collecting more data using different sets of objects, and this is left as future work.

Additionally, we assumed that a robot has access to a person's prior table setting arrangements. Part of our contribution is to release a large dataset of realistic tabletop arrangements for others to train from. In reality, the amount of data that we require from an individual to achieve reasonable performance is quite small compared to prior preference learning approaches [147].

A final limitation is that VLMs are not trained for fine-grained rearrangement tasks like ours. We presented several prompting methods that achieve reasonable performance even under this limitation, but our method would improve substantially if the underlying models were fine-tuned for fine-grained image reasoning tasks or even trained from the ground up for this kind of problem [60].

6.6 Conclusion

We introduced DegustaBot, a novel method for prompting VLMs to perform the challenging task of visual preference learning in multiobject rearrangement tasks. We formalize this task, and then analyze several methods' performance over simulated preferences. We then collect a large dataset of naturalistic preferences and evaluate additional methods on these naturalistic data, finding that, although this is a challenging task, our method produces acceptable table arrangements.

With DegustaBot, we see that it is possible to personalize generalized feature spaces using in-context learning. In this chapter, we additionally present a brief method for using corrective actions to further personalize feature spaces online. This work serves as the basis for the following chapter, which formalizes this idea and rigorously tests it to show how we can combine generalized feature spaces with online personalization from naturalistic human behavior.

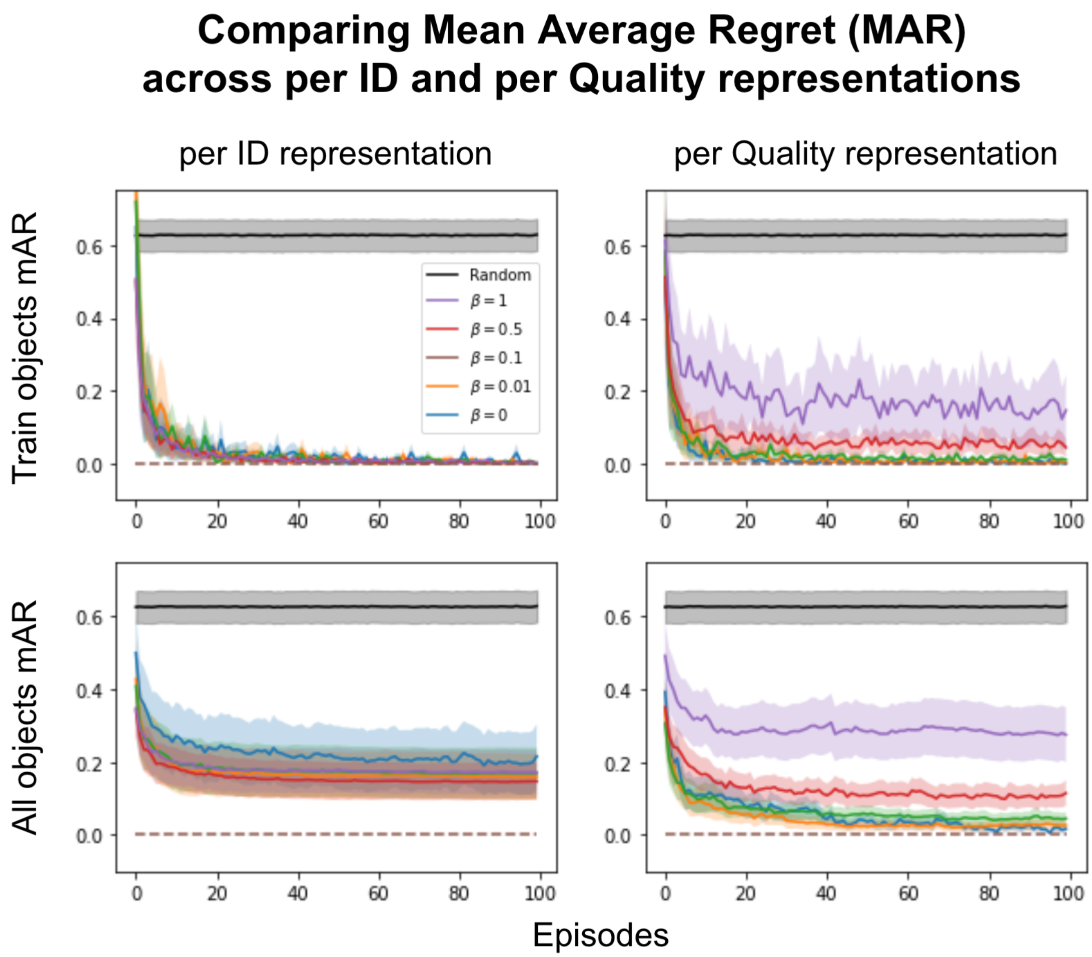


Figure 6.8: We report mean average regret (mAR). From left to right we show per ID representations and per Quality representations. From top to bottom, we show training and zero-shot testing on out of domain objects after N iterations. Colors indicate β thresholds as follows: 0 is blue, 0.01 is orange, 0.1 is green, 0.5 is red, and 1 is purple. Both methods perform similarly in training for well-correlated objectives, while per ID representations outperform when preferences are uncorrelated. For per Quality representations, improvements when training on correlated preferences correspond to improvements in testing set, which does not hold true for per ID representations.

7 Continual personalization using naturalistic corrections during surface rearrangement tasks

Agents that collaborate with people to complete a person's preferred goal cannot always know this preference in advance of an interaction. Although people may initially state these preferences, they may drift, sometimes completely, over the course of multiple interaction episodes. Although there may be no continued explicit communication between collaborative partners, people's *in-situ* behaviors are goal-driven and thus can reveal the up-to-date preference. This means that updating agent policies based on *in situ* behaviors is critical to assisting people during collaborations, that is, to ensure that robot actions are deferential to user goals [221].

Much current research on human-agent collaboration aims to learn zero-shot collaboration policies from offline datasets that are collected from human-human demonstrations [56] or generated synthetically [308]. Instead of using an individual's *in-situ* behavior to update a model online to improve performance with respect to that individual's preference, these approaches train agents offline in collaboration with the population of partner agents represented by the training dataset. They then target good performance in aggregate on task metrics. At test time, these approaches assume that the preferences and behavior of a new human collaborator will fall within the distribution of the collaborators represented by the training data. While these approaches have been shown to be effective on task metrics in general collaboration settings, they do not necessarily transfer to the stricter criteria of assistive collaborations where success in a task is dictated by a personal preference, and people's goals and behaviors can drift away from the training distribution.

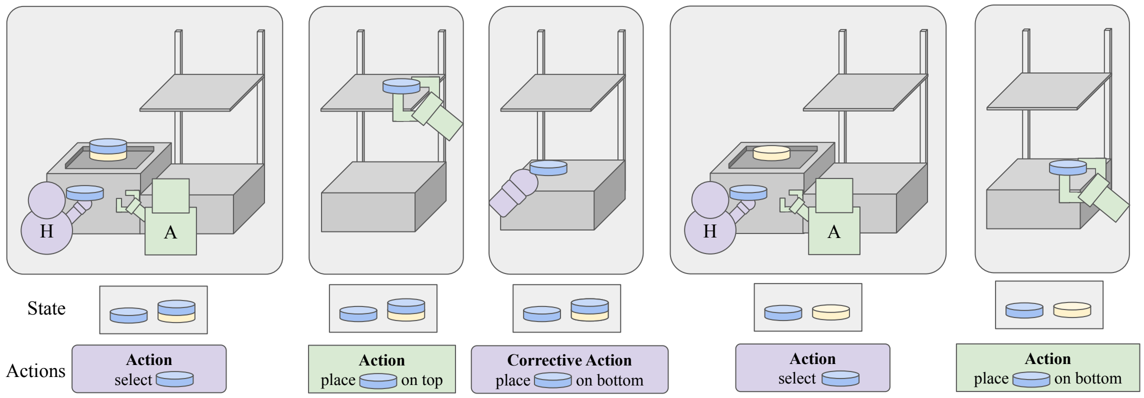


Figure 7.1: One step of an example surface rearrangement task: cupboard organization. From left to right: a person (H) picks an object to place in the dishwasher; the agent (A) initially places this incorrectly; the person corrects the placement. From this, the agent learns that the user likes to place blue objects on the bottom shelf and can place the next, similar object correctly.

Furthermore, the population of personal preferences is substantial and diverse, making it difficult to ensure adequate coverage during training time. Collecting large datasets of human-human data is time-consuming and expensive, while collaboration among populations of procedurally generated agents can yield data that do not tightly match the distribution of the human population. Furthermore, as people repeatedly execute a collaborative task, they may develop new preferences that are unlikely to be captured by the distribution of collaboration data represented in offline datasets.

We propose a method that takes advantage of these advances in zero-shot coordination and applies them to algorithms for fast, online adaptation from *in-situ* behavior. In this way, we hope to achieve both good initial performance when assisting a new partner, but also to continue to adapt to their preference over continued exposure.

However, deciphering people's exact preferences can be difficult, as these preferences are often not explicitly stated and can change over the course of an interaction. Fortunately, *in situ* behaviors are aimed at and can implicitly reveal information about a person's current preference or goal, even when it is not expressly communicated. We suggest that agents engaging in assistive collaborations utilize these goal-directed behaviors to infer and act toward a person's current goal, thereby enabling personalized assistance.

To do develop a model that can utilize these goal-directed behaviors for collaborative assistance, we introduce BLR-HAC: Bootstrapped Logistic Regression for Human Agent Collaboration. This model is trained using a two-stage approach: first, we pretrain a transformer [327] to learn to produce the parameters of a shallow, parameterized policy that second, is updated throughout a human-agent collaboration using online logistic regression. To test BLR-HAC, we first introduce a formalization of a specific instance of a rearrangement task, which we call an assistive surface rearrangement. We then compare BLR-HAC's performance in a simulated version of this task against two baselines: 1) a traditional transformer trained with behavior cloning and 2) a traditional shallow policy trained with online logistic regression.

Our chosen domain of surface rearrangement models household tasks, like dishwasher loading, which have complex, long-term dependencies determined by a combination of a person's environment and their strongly held preferences. For example, a person may prefer to place large dishes before small ones to maximize capacity. Such high-dimensional state and preference spaces lead to an almost infinite number of diverse and equally valid solutions for completing any given household chores. For example, choosing to load a dishwasher based on dish material is just as valid as loading based on dish size; it is a matter of personal preference. Given this diversity, household tasks make especially good testbeds for studying algorithms that require aligning robot policies with people's reward functions, thus mimicking many use cases for assistive robotics.

While some prior approaches to developing autonomous assistants for household tasks rely on people providing full task demonstrations in advance of a collaboration, BLR-HAC aims to operate in real time, utilizing information from each action as it is taken by a person. Furthermore, approaches based on full task demonstrations can introduce additional burden on a person and be redundant with the goal-directed behavior that people exhibit when completing tasks [22]. In contrast, training shallow low-capacity models with logistic regression through MaxEntIRL to utilize *in-situ* behavior has been shown to effectively and quickly adapt to people's objectives in areas such as robot teleoperation [138] and motion planning [185].

We tested BLR-HAC in a simulated version of our surface rearrangement task. We find that BLR-HAC outperforms baseline low-capacity models and large, nonlinear models trained with behavior cloning in zero-shot coordination. We also find that BLR-HAC achieves similar performance but requires a fraction of the compute of a transformer that is fine-tuned online. This finding holds true when considering both preferences that remain the same over time, i.e. are stationary, and those that drift, i.e. are nonstationary. Together, these results show how BLR-HAC is able to take advantage of the strengths of both zero-shot and fast online adaptation methods. It does this by pretraining a large non-linear model to learn the parameters of a shallow policy that can be updated with online logistic regression. This results in a collaborative agent that is both well-initialized and highly adaptable.

In this paper we make the following contributions:

- a formalization of common household tasks as collaborative IRL tasks, which we call surface rearrangement,
- a novel model, BLR-HAC, that combines the strengths of pretrained large, nonlinear models with low-capacity models trained online via logistic regression for efficient learning in human-robot collaborations, and
- evidence from experiments in simulation that BLR-HAC outperforms its component models.

7.1 Approach

Our ultimate goal is to learn an assistive policy that collaborates with a person during a surface rearrangement task. Given that we want our policy to be assistive, it should take actions that align with the person's underlying preference to complete the task. We interpret this as a regret minimization problem, where the policy aims to minimize the regret of its actions with respect to the actions that would be exhibited under the person's true preference for completing the task. Importantly, we assume that the policy does not have prior knowledge of this preference and that the person does not immediately or explicitly reveal it. Additionally, we assume that the space of possible preferences that the person could hold to be extremely large, making disambiguation from limited interaction with the person difficult.

Under these conditions, we have two main ways to perform regret minimization. First, we can ensure that our policy takes good initial actions that are likely to align with the person's preference, often called zero-shot performance. Second, we can adapt the policy online as a history of behavior is accumulated.

Algorithm 1 Surface Rearrangement

Require: π_r, π_h, env, O, L
$s^0 \leftarrow$ env.reset ()
$\xi \leftarrow [\ s^0\ ]$
Initialize g_θ
while ξ.length $< L$ **do**
    $a_h^t \leftarrow \pi_h\left(\cdot | s^{t-1}, g_\theta\right)$
    $a_r^t \leftarrow \pi_r\left(\cdot | a_h^t, s^{t-1}, g_{\hat{\theta}}\right)$
    $s^t, a_c^t \leftarrow$env.step(a_h^t, a_r^t, s^{t-1})
    $\hat{g}_\theta \leftarrow \hat{g}_\theta - \alpha\left(\psi_h(a_h) \cdot \psi_r(a_r) - \psi_h(a_h) \cdot \psi_r(a_c)\right)$
    ξ.append($[a_h^t, a_r^t, a_c^t, s^t]$)
end while

Action inference and policy adaptation do not operate in a vacuum, but rather within the course of the interaction. A common metric in human-robot interaction: collaborative fluency [122], for example, is critical for people considering an interaction with a robot to be "good." An important facet of this metric is related to the amount of time the robot sits idle during task execution. This makes frequently updating large models during an interaction challenging, as both action inference and policy updating require large amounts of computation, leading to high robot idle times. We aim to develop a method that can take advantage of the good performance of large nonlinear models while being able to quickly adapt to user preferences, as expressed through their *in-situ* behavior, without causing the robot to idle.

To learn an assistive policy that solves the DEC-POMDP discussed in Sec. 2.3.2, we first generate a simulated data set of diverse high-level preferences (Sect. 7.1.2). Using these preferences, we collect a dataset of collaborative demonstrations in a simulated surface rearrangement task on a range of difficulties (Sect. 7.1.2). We then train our two-stage algorithm by first learning to mimic the collected expert demonstrations (Sec. 7.1.3) and second, using the preference representations learned in Sec. 7.1.4 to perform a fast online adaptation.

7.1.1 Problem setup

In this work, we consider the surface rearrangement task, Algorithm 1, where two agents work together to arrange a set of objects O into a set of locations L. In

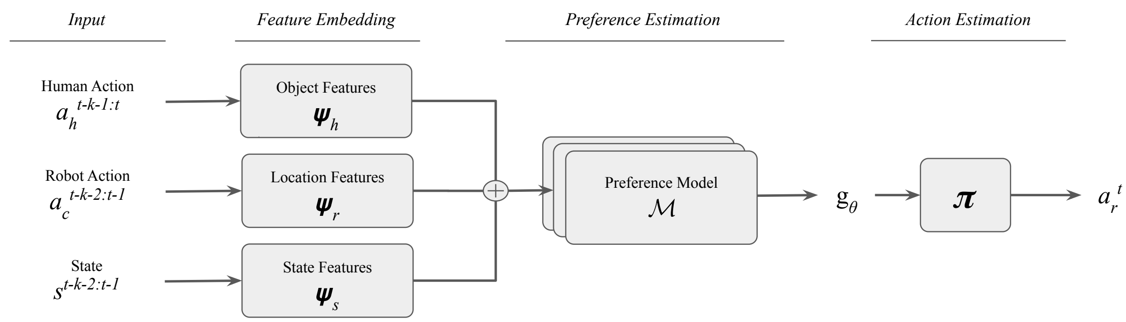


Figure 7.2: From left to right, we first embed the input state and actions using ψ. These are then concatenated and fed into the preference estimator $\mathcal{M}$. This learns to output reward parameters, g_θ which are used to initialize an online learning policy using the policy π, which determines the robot's action a_r.

this task, the assistive agent aims to help a person rearrange objects into locations. Importantly, the agent's goal is to achieve the final state that is desired by the person, which is initially unknown to the agent.

A single episode of this task consists of an object repository containing objects $o \in O$. The initial state of the episode is $|L|$ randomly chosen objects from O and L vacant locations. Each location has a capacity for a single object. Progress in the task is made by placing objects o in locations $l \in L$. A task is completed when all objects o have been placed in a location l. For simplicity, we assume that $N \leq |L|$ and that placing o in l occurs instantaneously.

Two agents interact in an episode in the following way. The human agent π_θ first picks an object given the current state s^{t-1}. Then, the robot agent $\pi_{\hat{\theta}}$ places this object into a location. The environment then returns the next state s^t and the human corrects the robot's action, returning a_c^t. An episode ξ can be represented as the following tuple: $\left(s^0, a_h^1, a_r^1, a_c^1, s^1, ...a_h^L, a_r^L, a_c^L, s^L\right)$.

7.1.2 Datasets

Modeling a Diverse User Population

The two key ideas of our method to develop assistive robots for household collaborations is that the method should be able to both effectively use a large population of preference data to pretrain good initializations and be able to quickly adapt to a

Algorithm 2 Expert Demonstration Collection

Require: $\Theta, \pi, \texttt{env}, O, L$
1: $D = [\,]$
2: **for** θ in Θ **do**
3: $\quad \xi \leftarrow \texttt{surfaceRearrangement}(\pi_h, \pi_r, \texttt{env}, \text{O}, \text{L})$
4: $\quad D\texttt{.append}(\xi)$
5: **end for**

particular preference when presented with information about that preference.

To capture these ideas in our experiments, we develop a simulated dataset of preferences. First, we sample a large set of preferences, representing a population, as encoded by g_θ. We assume that preferences within this population are drawn normally from one of several modes, each of which indicates a subpopulation of similar preferences. We sample three preference datasets: *train*, , and *test*. From each set of preferences, we sample episodes of surface rearrangement episode rollouts, thus creating three datasets: D_{train}, D_{eval}, and D_{test}. D_{train} consists of 1000 simulated preferences, sampled from four modes, with 1000 episodes per preference. D_{eval}, and D_{test} each contain 100 simulated preferences, with 20 episodes per preference.

Environments for surface rearrangement

To test the efficacy of our approach in varying difficulties, we develop three environments. Each environment scales the difficulty of the problem by increasing the size of the state space. We have a small environment, with five possible objects and five locations, a medium environment, with ten objects and ten locations, and finally a large environment, with 25 objects and 25 locations.

To collect a demonstration data set for each environment, we use Alg. 2. Importantly, to collect expert demonstrations, we set $g_\theta = \hat{g}_\theta$ and use a linear policy $\pi = \psi_h(a_h) \cdot g_\theta \cdot \psi_r(A_r)$, where all ψ are implemented as one-hot embedding layers. For each environment, we collect 100 demonstrations from each preference generated in Sec. 7.1.2.

Algorithm 3 Learning Priors for Online Linear Regression

Require: $D, \mathcal{M}, \psi_h, \psi_r, \psi_s$

```
while training do
    for (s, a_h, k) in D do
        ĝ_θ ← M(ψ_s(s), ψ_h(a_h), ψ_r(a_c))
        a_r ← argmax_{a_r ∈ A_r} ψ_h(a_h^t) · ĝ_θ · ψ_r(A_r)
        loss ← p(a_c) log q(a_r)
        training ← M.update(loss)
    end for
end while
```

7.1.3 Learning preferences in a diverse user population

The first step of our proposed algorithm aims to minimize regret by achieving good zero-shot performance. Ultimately, we want to model $p(a_r|s, a_h)$. However, this problem is ill-posed, as two policies parameterized by different preferences will correctly take two different actions a_r given the same state and human action. To account for this ambiguity, we include a history of the prior state k and action pairs taken under current preference and maximize $p(a_r|s^{t-k-2:t-1}, a_h^{t-k-1:t}, a_c^{t-k-2:t-1})$. For the sake of brevity, we will slightly abuse the notation and refer to this distribution as $p(a_r|s, a_h, k)$.

Again, when training assistive agents, achieving low zero-shot performance is not our only objective. We also need an agent that adapts online to incoming user behavior while maintaining collaborative fluency. This means developing a lightweight, low-parameter model capable of performing action inference and policy adaptation in real time.

To do this, instead of learning $p(a_r|s, a_h, k)$ directly, we first learn a latent space that corresponds to the weights of a logistic regression problem. These weights serve as input to the second step of our algorithm, Sec. 7.1.4. Thus, we train our model to maximize $M_{\psi,\psi}(s, a_h, k, t) = p(g_\theta|s, a_h, k, t)$. In this way, we place an inductive bias over the latent space of the model, enticing it to learn a matrix of size $O \times L$, which can be used as the weights of an online logistic regression problem. We treat this as a classification problem and minimize the cross entropy loss between our model's predictions and the collected expert demonstrations: $L = p(a_c) \cdot \log q(a_r)$ where $q(a_r) = \psi(a_h) \cdot M(s, a_h, k, t) \cdot \psi(A_r)$, as shown in Alg. 3.

7.1.4 Bootstrapping shallow linear models for fast, online adaptation

The second step of our proposed algorithm aims to minimize regret through online adaptation. Using the output of the model learned in Sect. 7.1.3, we can employ online logistic regression, which has been shown to work well to teach human preferences to agents through corrective feedback in robot control tasks. Importantly, since online logistic regression has a very simple update rule to estimate g_θ that operates on a much smaller number of parameters than a large, nonlinear network, we can adapt this initial estimate of the person's preference *in situ* without risking large human or robot idle time, thereby maintaining collaborative fluency.

To update our estimate of g_θ, we use a linear approximation of the QMDP solution to the DEC-POMDP in Section 2.3.2 and stochastic gradient descent, resulting in the following update rule:

$$\hat{g}_\theta = \hat{g}_\theta - \alpha \left(\psi_h(a_h) \cdot \psi_r(a_r) - \psi_h(a_h) \cdot \psi_r(a_c) \right)$$

where α is the learning rate.

	Small		Medium		Large	
	No Prior	Prior	No Prior	Prior	No Prior	Prior
ShallowLinear	0.413	0.665	0.215	0.518	0.096	0.289
DeepLinear	0.425	0.680	0.199	0.504	0.101	0.303
MLP	0.605	0.759	0.361	0.653	0.120	0.358
Transformer	0.729	**0.771**	0.603	**0.673**	0.160	**0.412**

Table 7.1: We compare zero-shot performance on the test set of each environment. We have two axes of comparison: model complexity in the rows, and inductive prior in the columns. Results are reported in terms of accuracy. We can see that the highest capacity, attention based model trained with an inductive prior outperforms all other models in every environment.

7.2 Experimental design

To test our algorithm, we designed several experiments. First, we validate the need for large, nonlinear models to learn the distribution of preferences embedded in the demonstration dataset, Sect. 7.2.1. Then we explore how our algorithm performs in its intended use case: fast, online adaptation. We tested this in two scenarios. Sect. 7.2.2 analyzes adaptation to a single preference over time, while Sect. 7.2.2 explores how well our algorithm performs when the preference generating the behavior changes without explicit communication to the robot.

7.2.1 Zero-shot coordination

We evaluated our model in each environment over the test set using Alg. 3. While we are in search of an algorithm that performs regret minimization, this metric is relative to a specific preference. To understand model performance in an absolute sense and compare across environments, we report accuracy in terms of the number of correct robot action predictions. This metric is inversely correlated with regret.

We choose our baselines to examine two key questions: 1) are high-capacity, nonlinear models necessary for disambiguation between preferences in a highly diverse preference space, and 2) how does inducing an inductive prior over the latent space affect zero-shot performance?

To answer these questions, we introduce baselines across two axes: model complexity and model bias. To determine the effect of high-capacity nonlinear models on zero-shot performance we compare four levels of model complexity in terms of how we implement ψ in Fig. 7.2:

- **ShallowLinear**. Typical online IRL settings learn a shallow model from scratch using MaxEntIRL. To bootstrap this process, one could perform the same process over the offline dataset, thereby encoding the diverse preference population in the initial model weights. Our intuition, though, is that since demonstrations are drawn from a large, diverse population of preferences, and that the relations between preferences and people are not known a priori, this disambiguation will benefit from a nonlinear function approximator. We expect nonlinear, high-capacity models to outperform this baseline.

- **DeepLinear**. Since the space of preferences is very large, it could simply be that increasing model capacity without introducing nonlinearity may capture the preference distribution. To test this, we introduce DeepLinear, which simply adds additional model parameters in both width and depth. We expect this model to outperform a ShallowLinear model but underperform nonlinear methods.
- **Multi-Layer Perceptron**. To test the importance of modeling the preference distribution with a nonlinear model, we introduce a multi-layer perceptron baseline. We expect this model to outperform both linear methods but underperform attention-based mechanisms.
- **Causal Transformer**. Finally, since we are passing a history of behavior to the model at every time step, we can infer the current preference from this sequence of behaviors. Attention-based mechanisms, specifically causal transformers, have been shown to excel at modeling sequential data. To test this we implement ψ as a transformer, and expect it to outperform all other methods.

The second axis of baselines we develop compares the importance of introducing an inductive bias over the latent space to learn g_θ. We compare an implementation of the above models in which each model minimizes $L = p(a_r) \cdot \log \mathcal{M}(s, a_h, k)$ to our proposed inductive bias, which minimizes $L = p(ar) \cdot \log \psi(a_h) \cdot \mathcal{M}(s, a_h, k) \cdot \psi(A_r)$.

Implementation details

To implement our models we make the following decisions. We perform a separate parameter sweep for each model and environment for the following parameters and ranges: learning rate $(1e^{-3}, 1e^{-6})$, the dimensionality of hidden layers $(2^5, 2^8)$, and the number of layers in ψ $(3, 5, 7, 10, 12)$. We set the size of the input history to be 50, padding when necessary. For each model, we implement all ψ as a single, one-hot embedding space of vocabulary size 208, where 0-7 are special characters, 8-107 are location indices, and 108-207 are object indices. To implement ψ, we use PyTorch [236] and base our implementation of a causal transformer on Decision Transformer [62]. We implement π as a simple linear model for inductive bias and as an MLP for no inductive bias. All models are trained using the appropriate training and evaluation sets, which do not overlap with the test set, with 10 epochs of early stopping.

7.2.2 Test-time adaptation

Developing assistive policies is not only about achieving good zero-shot performance. The space of actual human preferences is almost boundless and is likely impossible to capture in advance of an interaction. Therefore, it is important to develop algorithms that can rapidly align with the preferences associated with a person's *in-situ* behavior. We study this in two settings. First, we analyze our algorithm's ability to adapt to a stationary preference over the course of multiple episodes. Then, we analyze our algorithm's ability to adapt in scenarios where preferences are nonstationary. Here, we are interested in an algorithm's ability to 1) maintain decent performance in the face of a preference change and 2) rapidly recover after the change in preference.

Stationary Preferences

To test our algorithm's ability to adapt to stationary preferences, we averaged the performance of our bootstrapped online IRL algorithm on all preferences in the testing set over 20 episodes in each testing environment.

We compare against a linear model that learns from scratch and a method that optimizes over all transformer parameters between episodes but keeps inference computation constant. We measure computation cost in terms of FLOPS and calculate these values empirically using FVCore. We expect to see that the bootstrapped online IRL algorithm achieves a similar performance to the online transformer method, but at a fraction of the compute.

Nonstationary Preferences

Similarly to the stationary preferences experiment, we run IRL over 20 episodes. In this analysis, however, we switch to a different random objective after 10 episodes. Again, we compare against a linear model learning from scratch and an online transformer implementation. We expect to see that the linear method starts with poor performance but adapts quickly when exposed to incoming behavior. We expect to see that the transformer method starts with good performance and adapts more

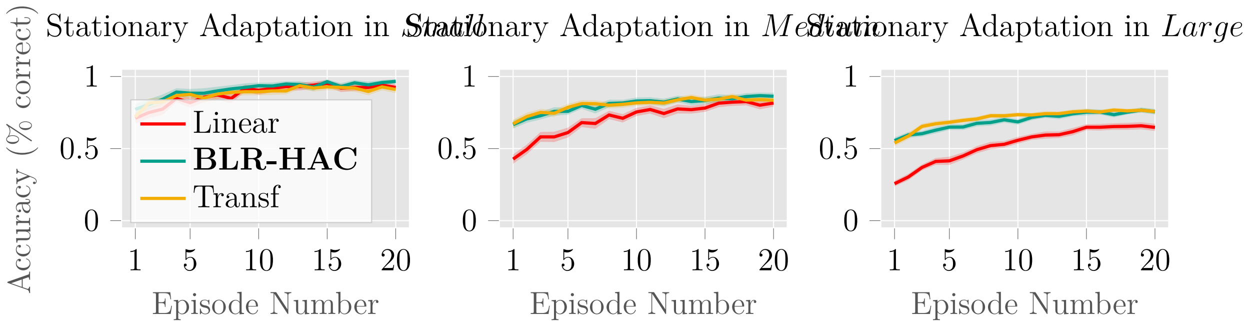


Figure 7.3: Learning curves for each test environment for each algorithm. We report the average accuracy over each episode. BLR-HAC is able to achieve the low zero-shot performance of the transformer method, and the fast adaptation of the linear method. Additionally, we can see that as the episode length increases, these differences in performance are more notable, with the linear method failing to catch up to the other two methods over the course of 20 episodes.

slowly as behavior data is accumulated. Finally, we expect our method to achieve the benefits of both the linear from scratch and the transformer methods: it should start off with reasonable performance and adapt quickly as data are aggregated.

Implementation Details

For both experiments, we do a hyperparameter sweep over the learning rate in the range $(1e^{-2}, 1e^{-5})$ for the transformer and $(1, 5, 10)$ for the linear models. In both cases, we used the maximum learning rate for all experiments. In addition, we use stochastic gradient descent for optimization in both cases. To train the transformer method, we perform five steps of gradient descent between each episode.

7.3 Results

From running the experiments outlined in Sec. 7.2, we have three main results. First, we find support for our hypothesis that nonlinear, high-capacity models trained with inductive biases can learn a diverse population of user preferences. In Tab. 7.1, we see that the attention-based method trained with an inductive prior outperforms all other methods, achieving $77.1\%, 67.3\%,$ and 41.2% accuracy on the small, medium, and

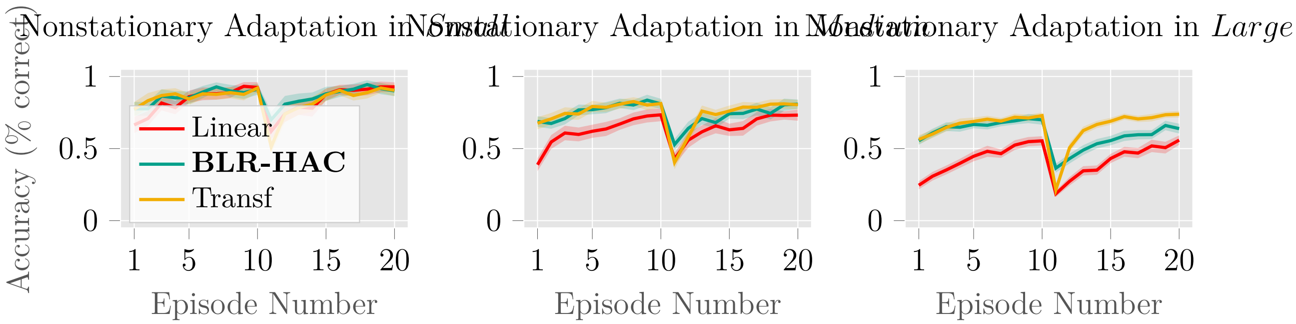


Figure 7.4: Learning curves over each test environment for each algorithm. We report average accuracy over episodes. BLR-HAC is able to perform on par with the transformer method in the small and medium environments and part of the large environment. BLR-HAC outperforms all methods in all environments immediately after the preference switch. In the large environment, though, the transformer recovers more quickly as it has access to more data.

large environments, respectively. We see that the difference in performance between models trained with and without the inductive prior increases as the difficulty of the problem increases. Additionally, we see the general trend that higher capacity nonlinear models outperform lower capacity linear models. These results empirically justify our desire to use a high-capacity nonlinear model to bootstrap a linear model in an online logistic regression problem.

Our second set of results is shown in Fig. 7.3. Here, we plot the test-time adaptation accuracy for three models: linear (in red), BLR-HAC (in green), and online transformer (in yellow). From these graphs, we can see support for our hypothesis that bootstrapped, shallow linear models trained with IRL achieve good accuracy with low computation. We can see that BLR-HAC and Transformer both start with higher accuracy than Linear in all cases and that this difference increases as the problem complexity increases. Furthermore, we see how BLR-HAC achieves a similar performance over episodes as the transformer method, but at a fraction of the computation. Although both methods have similar inference compute, of OxL FLOPS, BLR-HAC uses only $2xOxL$ FLOPS, while the Transformer method uses $\sim 400M$ FLOPS during updates.

Finally, in Fig. 7.4 results from test-time adaption with nonstationary preferences.

These results show mixed support for our hypothesis that bootstrapped, shallow linear models trained with IRL recover well from unexpected shifts in user behavior. In each graph, episodes 1-10 show results similar to the previous set of experiments. However, in episode 10, the preference shifts, and all models suffer a drop in performance. Interestingly, in all cases, BLR-HAC suffers the smallest drop in performance. Although this is a positive result, we also see that as the environment becomes more complex, BLR-HAC suffers in its adaptation rate from episodes 10-20. While it adapts on par with the linear method (though still achieves higher performance due to its better initial performance), it adapts slower than the transformer-based method. This is likely due to the fact that the transformer is able to make better use of the larger amounts of data that are being aggregated in the large environment.

7.4 Discussion, limitations, and future work

We develop policies for assistive agents that are both well-initialized and highly adaptable. Through simulated experiments, our method achieves both the good initializations of large, nonlinear models trained with behavior cloning and the fast adaptation to user behavior present in low-capacity models trained with online MaxEntIRL. Importantly, BLR-HAC initializes better than ShallowLinear on test data that are far from the initial distribution, meaning that our approach should ideally allow for faster adaptation to populations for whom it is difficult to collect data for offline pretraining.

Future work should explore applying BLR-HAC to user studies with real people to determine whether the better initializations and faster adaptations of our method hold outside of simulation and are preferred. It is also important to study the effect of the size of the surface rearrangement problem on these results.

User studies also provide an opportunity to improve our method. Collecting interaction data through interactive simulators, such as AI Habitat [273, 311], deployed on platforms such as Amazon Mechanical Turk [72] or Prolific [244] would allow us to pretrain BLR-HAC with real data.

Finally, our method also assumes a single, synchronized modality of corrective actions: direct state corrections. This makes our learning problem easier by maximizing the correlation between the leader's corrections and their reward function. We would like to extend our approach to account for other modalities of corrections issued asynchronously, such as those expressed in real time through verbal or nonverbal communication.

We show that we can use offline datasets to bootstrap assistive collaborations by pretraining assistive agents. However, this method necessitates using specific subpopulations of the larger human population, i.e. those represented by the dataset. This leads to ethical questions such as: Are the preferences present in the dataset representative of the larger population? How does this affect people who have preferences outside this subpopulation? These questions are especially pertinent in assistive settings, where agents are likely to encounter out-of-distribution phenomena at test time. Questions such as these are the motivation for this work.

We assume that a critical part of providing assistance is to reduce unnecessary burden placed on individuals while acting in alignment with their preference. When a person's preferences are well represented in the data set, pretraining necessarily minimizes a person's burden to bring the agent into alignment with their preference. When a person's preferences are not well represented in the data set, our method quickly aligns to the person's preference quickly by using their in-situ, goal-directed behavior. Thus, while the model does not have an initial representation of these out-of-domain preferences, it knows how to interpret goal-directed behaviors to learn such a representation.

We believe that there is ample opportunity for future work to continue to explore solutions to these ethical dilemmas. For example, by learning more generalizable features of preferences that allow for better representations of human preferences, or by teaching agents to learn to learn preferences, which would improve an assistive agents ability to adapt to out-of-distribution preferences.

7.5 Conclusion

In this work, we laid out an argument for why assistive agents should be both well-initialized and highly adaptable. We introduced a novel formulation of assistive human-agent collaboration as collaborative inverse reinforcement learning and introduced an algorithm BLR-HAC that takes advantage of sophisticated population-level modeling found in deep neural networks with the fast adaptation of shallow, low-capacity inverse reinforcement learning methods. Finally, we verified these claims through simulated experiments. With these experiments, we show how we can use naturalistic human behaviors in order to provide value-aligned robotic assistance.

8 Conclusion

Providing robotic assistance to people with household tasks is a challenging undertaking that would give back valuable free time and help people for whom these tasks are physically burdensome. In this thesis, we lay out an argument that robots that provide such assistance should do so according to a value alignment process: to assist a person, a robot must take actions that support a person's goals. We also identify the components of an assistive robotic system and a large literature review of recent work in assistive robotics, Chapter 3.

We then demonstrate how naturalistic behaviors can be collected in a simplified version of such a task in Chapter 4 and show how these can be used to train feature spaces that could be used for downstream rearrangement tasks in Chapter 5. We show evidence that personalization is needed in assistive robotics, by demonstrating how models trained at the population level suffer when generalizing across participants.

Seeing this evidence for the need to personalize such general feature spaces, we present a method for doing so using in-context learning in Chapter 6. We explore this idea in complex, open-vocabulary, sequentially dependent rearrangement problems, which model household tasks. Our first study in this uses in-context learning to personalize the feature spaces of large, generalized pretrained models trained on internet scale data. While we find that we can use these methods to personalize these large models for rearrangement problems, we find that the solution is not sufficient on its own and propose a method to use naturalistic human behavior to continually refine these feature spaces.

Finally, seeing that in-context prompting was not enough to personalize alone, we introduce a method for continual, shallow fine-tuning of a large model's preference representation in Chapter 7: BLR-HAC. This method finetunes large, pretrained models by using naturalistic corrective behaviors during the execution of a collaborative rearrangement task. This method works by training a small model on top of a large, pretrained model, that learns to interpret corrective actions as goal oriented, thereby allowing rapid training for individuals and continued refinement over the lifecycle of the interaction.

8.1 Generalizability

Although we presented several different methods for providing robot assistance in surface rearrangement tasks, the space of general rearrangement tasks is quite large and can vary according to many different task variables. As such, it is important to discuss the ways in which our methods can generalize to a broader space of tasks and the preferences that can be held within these tasks.

First, since our pipeline (e.g., the work done in Chapter 6) uses vision-and-language models to capture preferences, it could theoretically generalize to any preference that is able to be represented as an image, in text, or as a combination of both. Bin-picking tasks, for example, could be represented as images of the placement bins and abstract images of the items that can be placed. The bin each object belongs to can be represented as is currently, through prior examples of a task plan. More complex tasks, such as building a three-dimensional structure, such as a car or a Lego model, could require additional intermediate representations.

In practice, we only experimented with preferences that could be represented as a planar image of, for example, a table top. We do extend previous work, which typically considers a set of objects in view (e.g. Dalle-Bot [148], which rearranges only the items in view on a table top at the beginning of a task) to be able to arrange only a subset of items that are not in view in the current scene. The set of objects from which you would like to select can be passed in as an unorganized list of either images or object descriptions.

Additionally, while our methods do capture preferences over the order in which objects are placed, they do this at a coarse level, assuming access to general pick_and_place functions that can reliably pick and place any object. Our current method would not be able to account for preferences that operate over finer grained details of the task planning process, such as the exact trajectory a robot takes as it executes a pick_and_place action.

Some rearrangement tasks also require memory or maintaining some amount of knowledge about multiple states at any given task. For example, when stocking aisles in a grocery store, a person may need to remember what items are in inventory, what items have run out on the shelves, and what items other employees are in charge of stocking. While such a task could likely be represented in text or in images, it is an open question if large foundation models can reason over increasingly complex tasks such as this one. Additionally, there may be a missing social component of such a system: how does the robot execute the rearrangement task when its objectives conflict with the objectives of another person? As these memory requirements grow, as is common in many rearrangement tasks, methods that rely on models with fixed context length may struggle to perform well.

8.2 Future research directions

Given these various limitations in the generalizability of our approach, as well as the experience we gained in building the systems presented, we propose several directions for future work, as follows.

8.2.1 Real robot implementation

First, connecting DegustaBot (Chapter 6) and BLR-HAC (Chapter 7 could result in a real system that would be able to provide robotic assistance in a real surface rearrangement task, such as setting the table. By instantiating the pre-trained model in BLR-HAC with a modern foundation model, such as GPT4-o, and implementing a top-down pick and place robot, a working system could be built.

Such a set-up would be beneficial in a variety of ways. First, it would allow researchers to easily test the generalizability of new objects and new tasks. Participants could bring in real objects from home and use them to perform a novel rearrangement task. Additionally, while the images we use in Chapter 6 are drawn from real images, a realistic setup would allow us to determine the method's ability to be robust to confounding features such as lighting variation.

A realistic setup could also help researchers perform user studies which would be critical for understanding whether people actually enjoy this kind of assistance, as well as for collecting naturalistic preferences and additional naturalistic behaviors that people use to express their preferences in such a task. With such datasets, additional improvements to the method could be made to make it generalize to more types of user preferences.

Finally, a system such as this would also allow us to test whether or not goal representations (such as capture in Chapters 6 and 7 can be used by downstream action execution policies. Our current work assumes a coarse action execution policy that relies on general pick_and_place actions. Future work could consider these task plans as the goal representation input for a more fine-grained goal conditions execution policy that could account for preferences in the trajectory a robot takes. This would disambiguate preferences about what should be done from how that task should be done, which would hopefully improve sample efficiency of online preference learning methods. A real world robot system could be a test-bed for such an implementation to test such hypotheses.

8.2.2 Modulating between conflicting goals

While much of the focus of the work in this dissertation focuses on capturing and refining an individuals goal in a rearrangement task, these tasks can also have a social element when executed in the real world. People can have high-level values and low-level values that may not always be consistent with one another. Philosophical literature has suggested that people's preferences can operate at different ranges that can sometimes conflict with one another, e.g., wanting to be safe but not wanting to wear a motorcycle helmet [304]. For example, if asked, a person may proclaim a desire to stay safe while operating motor vehicles, but in practice not wear their helmet or buckle their seatbelt when they drive a motorcycle or car.

Our current methods, which rely entirely on *in situ* behaviors, likely lean heavily on aligning to low level values (e.g. what the person is doing right now) and may therefore not be preferred in tasks where higher level values are of more concern or are difficult for a person to act on. *Ex-situ* behavior, for example interviewing someone about their preferred table arrangement, may reveal higher-level value information that could be used in conjunction with *in-situ* information to accomplish tasks in a way that satisfies both sets of values. Studying such questions would also allow us to understand what a robot should do when these value systems are incompatible and help us develop robots that can still provide assistance even though it is unclear what the person actually wants.

These ideas also occur in environments with multiple people. In a household, for example, different people may have different ideas about how to optimally load a dishwasher. A robot cannot load a dishwasher to please two people, but must accomplish the task. Understanding how the robot fits into this social dynamic, and how it should align its objectives to the family unit, is critical for long-term adoption in realistic households.

8.2.3 Incorporating multiple modalities of naturalistic human behaviors

Finally the work presented in this thesis relies on a relatively small set of naturalistic behaviors through which people can communicate their goals: gaze, visual demonstrations, text demonstrations, and corrective actions. In reality, people express their content or discontent through a huge amount of different behaviors, such as the prosody of speech, facial expressions, laughter, or other non-language utterances. Researchers should explore the multi-modal nature of human behaviors and human preferences.

First, researchers could train models to combine multiple different modalities of in situ, naturalistic, and goal-directed behaviors, similar to our work in Chapter 5. Training coordinated feature spaces for multiple different modes of behavior could allow for redundancy in understanding user goals, as well as being able to infer goals or preferences that are not apparent in any single signal. Large datasets of this kind could also bootstrap alignment to novel preferences or to novel tasks, for example through retrieval-augmented generation. This would make it not necessary to perform costly data collection on every new preference or task.

Finally, even though language and vision are very general forms of representing a preference, people often express their preferences through other means. Language is particularly interesting because of its variety and capability for expression. Verbal utterances, for example, can indicate when someone is unhappy with a current action, or eye gaze can reveal what a person really wants a robot to do. Idiolects can also develop that indicate additional information, potentially through short hand (e.g., "pass me that thingy" instead of "pass me the round blue object on the rightmost side"), which can be useful for executing collaborative rearrangement tasks. With a better understanding of how to incorporate language into methods such as the ones we present here, a larger variation of preferences could likely be captured with fewer data examples.

8.3 Closing thoughts

With the work presented in this thesis, we have argued that robots can provide value-aligned assistance using naturalistic human behavior that can be used in rearrangement tasks that model household chores. Although we have presented these methods and tested them in mostly simulated environments, it is increasingly important to understand what people want a robot to do. As such, we believe that it is very important to 1) understand how people execute rearrangement tasks in their homes to capture a realistic dataset of people's preferences and how strongly these preferences are held, and 2) get robots in front of people so that we can understand how these preferences may change as a robot executes more of these types of chores. With this work, we hope that we have pushed forward the knowledge in this field, so that we can one day achieve the goal of deploying general purpose assistive robots in household environments.

Bibliography

[1] Pieter Abbeel and Andrew Y Ng. Apprenticeship learning via inverse reinforcement learning. In *Proceedings of the twenty-first international conference on Machine learning*, page 1, 2004. 2.3.1

[2] Nichola Abdo, Cyrill Stachniss, Luciano Spinello, and Wolfram Burgard. Robot, organize my shelves! tidying up objects by predicting user preferences. In *2015 IEEE international conference on robotics and automation (ICRA)*, pages 1557–1564. IEEE, 2015. 2.2.1, 6

[3] Urja Acharya, Siya Kunde, Lucas Hall, Brittany A. Duncan, and Justin M. Bradley. Inference of user qualities in shared control. In *2018 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA)*, pages 588–595. IEEE COMPUTER SOC, 2018. 3.4.1

[4] Nevena Ackovska, Vesna Kirandziska, Ana Tanevska, Liljana Bozinovska, and Adrijan Bozinovski. Robot - assisted therapy for autistic children. In *SOUTHEASTCON 2017*. IEEE, 2017. 3.4.3

[5] Henny Admoni and Siddhartha S. Srinivasa. Predicting user intent through eye gaze for shared autonomy. In *Proceedings of the AAAI Fall Symposium: Shared Autonomy in Research and Practice*, pages 298–303. AAAI Press, 2016. 5

[6] Henny Admoni, Christopher Datsikas, and Brian Scassellati. Speech and gaze conflicts in collaborative human-robot interactions. In *Annual Conference of the Cognitive Science Society (CogSci)*, pages 104–109, 2014. 4

[7] Henny Admoni, Thomas Weng, Bradley Hayes, and Brian Scassellati. Robot nonverbal behavior improves task performance in difficult collaborations. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*, pages 51–58. IEEE Press, 2016. 3.4.3, 3.5.3

[8] D. J. Agravante, A. Cherubini, A. Bussy, P. Gergondet, and A. Kheddar. Collaborative human-humanoid carrying using vision and haptic sensing. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 607–612, May 2014. doi: 10.1109/ICRA.2014.6906917. 2.1.3

[9] Roxana Agrigoroaie, Francois Ferland, and Adriana Tapus. The enrichme

project: Lessons learnt from a first interaction with the elderly. In A Agah, JJ Cabibihan, AM Howard, MA Salichs, and H He, editors, *SOCIAL ROBOTICS, (ICSR 2016)*, volume 9979, pages 735–745. SPRINGER-VERLAG BERLIN, 2016. 3.4.3

[10] Michael Ahn, Anthony Brohan, Noah Brown, Yevgen Chebotar, Omar Cortes, Byron David, Chelsea Finn, Chuyuan Fu, Keerthana Gopalakrishnan, Karol Hausman, Alex Herzog, Daniel Ho, Jasmine Hsu, Julian Ibarz, Brian Ichter, Alex Irpan, Eric Jang, Rosario Jauregui Ruano, Kyle Jeffrey, Sally Jesmonth, Nikhil Joshi, Ryan Julian, Dmitry Kalashnikov, Yuheng Kuang, Kuang-Huei Lee, Sergey Levine, Yao Lu, Linda Luu, Carolina Parada, Peter Pastor, Jornell Quiambao, Kanishka Rao, Jarek Rettinghouse, Diego Reyes, Pierre Sermanet, Nicolas Sievers, Clayton Tan, Alexander Toshev, Vincent Vanhoucke, Fei Xia, Ted Xiao, Peng Xu, Sichun Xu, Mengyuan Yan, and Andy Zeng. Do as i can and not as i say: Grounding language in robotic affordances. In *arXiv preprint arXiv:2204.01691*, 2022. 2.4.1, 2.4.3

[11] Sharath Chandra Akkaladevi, Matthias Plasch, Andreas Pichler, and Bernhard Rinner. Human robot collaboration to reach a common goal in an assembly process. In D Pearce and HS Pinto, editors, *PROCEEDINGS OF THE EIGHTH EUROPEAN STARTING AI RESEARCHER SYMPOSIUM (STAIRS 2016)*, volume 284 of *Frontiers in Artificial Intelligence and Applications*, pages 3–14. IOS PRESS, 2016. 3.4.1, 3.5.3

[12] Maryam Alimardani, Linda Kemmeren, Kazuki Okumura, and Kazuo Hiraki. Robot-assisted mindfulness practice: Analysis of neurophysiological responses and affective state change. In *2020 29TH IEEE INTERNATIONAL CONFERENCE ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 683–689. IEEE, 2020. 3.3.2, 3.4.3

[13] Rasmus S. Andersen, Simon Bogh, Thomas B. Moeslund, and Ole Madsen. Task space hri for cooperative mobile robots in fit-out operations inside ship superstructures. In *2016 25TH IEEE INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 880–887. IEEE, 2016. 3.4.1

[14] Rasmus S. Andersen, Ole Madsen, Thomas B. Moeslund, and Heni Ben Amor. Projecting robot intentions into human environments. In *2016 25TH IEEE INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 294–301. IEEE, 2016. 3.4.3, 3.5.2

[15] Antonio Andriella, Carme Torras, Carla Abdelnour, and Guillem Alenyà. Introducing caresser: A framework for in situ learning robot social assistance from expert knowledge and demonstrations. *User Modeling and User-Adapted*

Interaction, 03 2022. doi: 10.1007/s11257-021-09316-5. 2.3.1

[16] Dejanira Araiza-Illan and Alberto de San Bernabe Clemente. Dynamic regions to enhance safety in human-robot interactions. In *2018 IEEE 23RD INTERNATIONAL CONFERENCE ON EMERGING TECHNOLOGIES AND FACTORY AUTOMATION (ETFA)*, IEEE International Conference on Emerging Technologies and Factory Automation-ETFA, pages 693–698. IEEE, 2018. 3.4.3

[17] Reuben M. Aronson and Henny Admoni. Gaze for error detection during human-robot shared manipulation. In *RSS Workshop: Towards a Framework for Joint Action*, 2018. 4, 5

[18] Reuben M. Aronson and Henny Admoni. Gaze complements control input for goal prediction during assisted teleoperation. *Robotics science and systems*, 2022. URL https://par.nsf.gov/biblio/10327640. 2.3.1

[19] Reuben M. Aronson, Thiago Santini, Thomas C. Kübler, Enkelejda Kasneci, Siddhartha Srinivasa, and Henny Admoni. Eye-hand behavior in human-robot shared manipulation. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '18, page 4–13, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450349536. doi: 10.1145/3171221.3171287. URL https://doi.org/10.1145/3171221.3171287. 2.1.1, 2.1.3, 2.3.1, 3.4.3, 3.5.1, 5, 5.1.1

[20] P. K. Artemiadis and K. J. Kyriakopoulos. Emg-based control of a robot arm using low-dimensional embeddings. *IEEE Transactions on Robotics*, 26(2): 393–398, April 2010. ISSN 1552-3098. doi: 10.1109/TRO.2009.2039378. 2.1.1

[21] Pablo Azagra, Yoan Mollard, Florian Golemo, Ana Cristina Murillo, Manuel Lopes, and Javier Civera. A Multimodal Human-Robot Interaction Dataset. NIPS 2016, workshop Future of Interactive Learning Machines, December 2016. URL https://hal.inria.fr/hal-01402479. Poster. 2.1.2

[22] Chris L Baker, Joshua B Tenenbaum, and Rebecca R Saxe. Goal inference as inverse planning. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 29, 2007. 2.3, 2.3.1, 7

[23] Malek Baklouti, Eric Monacelli, Vincent Guitteny, and Serge Couvet. Intelligent assistive exoskeleton with vision based interface. In S Helal, S Mitra, J Wong, C Chang, and M Mokhtari, editors, *SMART HOMES AND HEALTH TELEMATICS*, volume 5120 of *Lecture Notes in Computer Science*, pages 123–135. SPRINGER-VERLAG BERLIN, 2008. 3.4.2

[24] Jimmy Baraglia, Maya Cakmak, Yukie Nagai, Rajesh Rao, and Minoru Asada. Initiative in robot assistance during collaborative task execution. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*, HRI '16, pages 67–74. IEEE Press, 2016. 3.4.1, 3.5.1

[25] Andrea Bartl, Stefanie Bosch, Michael Brandt, Monique Dittrich, and Birgit Lugrin. The influence of a social robot's persona on how it is perceived and accepted by elderly users. In A Agah, JJ Cabibihan, AM Howard, MA Salichs, and H He, editors, *SOCIAL ROBOTICS, (ICSR 2016)*, volume 9979 of *Lecture Notes in Artificial Intelligence*, pages 681–691. SPRINGER-VERLAG BERLIN, 2016. 3.4.3

[26] Emanuele Bastianelli, Giuseppe Castellucci, Danilo Croce, Luca Iocchi, Roberto Basili, and Daniele Nardi. Huric: a human robot interaction corpus. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, Reykjavik, Iceland, may 2014. European Language Resources Association (ELRA). ISBN 978-2-9517408-8-4. 2.1.2

[27] Dhruv Batra, Angel X Chang, Sonia Chernova, Andrew J Davison, Jia Deng, Vladlen Koltun, Sergey Levine, Jitendra Malik, Igor Mordatch, Roozbeh Mottaghi, et al. Rearrangement: A challenge for embodied ai. *arXiv preprint arXiv:2011.01975*, 2020. 2.2, 2.2.1, 2.3.2

[28] Dhruv Batra, Angel X Chang, Sonia Chernova, Andrew J Davison, Jia Deng, Vladlen Koltun, Sergey Levine, Jitendra Malik, Igor Mordatch, Roozbeh Mottaghi, et al. Rearrangement: A challenge for embodied ai. *arXiv preprint arXiv:2011.01975*, 2020. 1

[29] Jackson Beatty. Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, 91(2):276, 1982. 4

[30] Roman Bednarik, Piotr Bartczak, Hana Vrzakova, Jani Koskinen, Antti-Pekka Elomaa, Antti Huotarinen, David Gil de Gómez Pérez, and Mikael von und zu Fraunberg. Pupil size as an indicator of visual-motor workload and expertise in microsurgical training tasks. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*, ETRA '18, pages 60:1–60:5, New York, NY, USA, 2018. ACM. ISBN 978-1-4503-5706-7. doi: 10.1145/3204493.3204577. URL `http://doi.acm.org/10.1145/3204493.3204577`. 4

[31] Marc G. Bellemare, Yavar Naddaf, Joel Veness, and Michael Bowling. The arcade learning environment: An evaluation platform for general agents. *CoRR*, abs/1207.4708, 2012. URL `http://arxiv.org/abs/1207.4708`. 2.1.1

[32] H. Ben Amor, G. Neumann, S. Kamthe, O. Kroemer, and J. Peters. Interaction primitives for human-robot cooperation tasks. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2831–2837, May 2014. doi: 10.1109/ICRA.2014.6907265. 2.1.3

[33] Atef Ben-Youssef, Chloé Clavel, Slim Essid, Miriam Bilac, Marine Chamoux, and Angelica Lim. Ue-hri: A new dataset for the study of user engagement in spontaneous human-robot interactions. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, ICMI 2017, pages 464–472, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-5543-8. doi: 10.1145/3136755.3136814. URL `http://doi.acm.org/10.1145/3136755.3136814`. 2.1.2

[34] Beatriz Bernardo, Patricia Alves-Oliveira, Maria Graca Santos, Francisco S. Melo, and Ana Paiva. An interactive tangram game for children with autism. In D Traum, W Swartout, P Khooshabeh, S Kopp, S Scherer, and A Leuski, editors, *INTELLIGENT VIRTUAL AGENTS, IVA 2016*, volume 10011 of *Lecture Notes in Artificial Intelligence*, pages 500–504. SPRINGER-VERLAG BERLIN, 2016. 3.4.3

[35] Jaishankar Bharatharaj, Loulin Huang, Ahmed Al-Jumaily, Mohan Rajesh Elara, and Chris Krageloh. Investigating the effects of robot-assisted therapy among children with autism spectrum disorder using bio-markers. In *2017 INTERNATIONAL CONFERENCE ON ADVANCED TECHNOLOGIES IN DESIGN, MECHANICAL AND AERONAUTICAL ENGINEERING (ATDMAE 2017)*, volume 234 of *IOP Conference Series-Materials Science and Engineering*. IOP PUBLISHING LTD, 2017. 3.4.3

[36] Gurit E. Birnbaum, Moran Mizrahi, Guy Hoffman, Harry T. Reis, Eli J. Finkel, and Omri Sass. Machines as a source of consolation: Robot responsiveness increases human approach behavior and desire for companionship. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*, HRI '16, pages 165–171. IEEE Press, 2016. 3.4.3

[37] Han Bo, Dhanya Menoth Mohan, Muhammad Azhar, Kana Sreekanth, and Domenico Campolo. Human-robot collaboration for tooling path guidance. In *2016 6TH IEEE INTERNATIONAL CONFERENCE ON BIOMEDICAL ROBOTICS AND BIOMECHATRONICS (BIOROB)*, Proceedings of the IEEE RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics, pages 1340–1345. IEEE, 2016. 3.4.2, 3.5.3

[38] Christopher Bodden, Daniel Rakita, Bilge Mutlu, and Michael Gleicher. Evaluating intent-expressive robot arm motion. In *2016 25TH IEEE INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 658–663. IEEE, 2016. 3.4.3

[39] Mayara Bonani, Raquel Oliveira, Filipa Correia, André Rodrigues, Tiago Guerreiro, and Ana Paiva. What my eyes can't see, a robot can show me: Exploring the collaboration between blind people and robots. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '18, pages 15–27. Association for Computing Machinery,

2018. 3.4.2

[40] Jean-David Boucher, Ugo Pattacini, Amelie Lelong, Gerrard Bailly, Frederic Elisei, Sascha Fagel, Peter Ford Dominey, and Jocelyne Ventre-Dominey. I Reach Faster When I See You Look: Gaze Effects in Human-Human and Human-Robot Face-to-Face Cooperation. *Frontiers in neurorobotics*, 6(May): 1–11, January 2012. ISSN 1662-5218. doi: 10.3389/fnbot.2012.00003. 2.1.3, 4

[41] C. Braunagel, E. Kasneci, W. Stolzmann, and W. Rosenstiel. Driver-activity recognition in the context of conditionally autonomous driving. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, pages 1652–1657, Sept 2015. doi: 10.1109/ITSC.2015.268. 2.1.1

[42] C Breazeal, A Brooks, D Chilongo, J Gray, G Hoffman, C Kidd, H Lee, J Lieberman, and A Lockerd. Working collaboratively with humanoid robots. In *2004 4TH IEEE/RAS INTERNATIONAL CONFERENCE ON HUMANOID ROBOTS, VOLS 1 AND 2, PROCEEDINGS*, IEEE-RAS International Conference on Humanoid Robots, pages 253–272. IEEE, 2004. 3.4.3

[43] Priscilla Briggs, Matthias Scheutz, and Linda Tickle-Degnen. Are robots ready for administering health status surveys? first results from an hri study with subjects with parkinson's disease. In *PROCEEDINGS OF THE 2015 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI'15)*, ACM IEEE International Conference on Human-Robot Interaction, pages 327–334. ASSOC COMPUTING MACHINERY, 2015. 3.3.3, 3.4.3

[44] Alexander Broad, Jacob Arkin, Nathan Ratliff, Thomas Howard, and Brenna Argall. Real-time natural language corrections for assistive robotic manipulators. *Int. J. Rob. Res.*, 36(5-7):684–698, June 2017. ISSN 0278-3649. doi: 10.1177/0278364917706418. URL `https://doi.org/10.1177/0278364917706418`. 2.1.1

[45] Christina Broehl, Jochen Nelles, Christopher Brandl, Alexander Mertens, and Christopher M. Schlick. Tam reloaded: A technology acceptance model for human-robot cooperation in production systems. In C Stephanidis, editor, *HCI INTERNATIONAL 2016 - POSTERS' EXTENDED ABSTRACTS, PT I*, volume 617 of *Communications in Computer and Information Science*, pages 97–103. SPRINGER INTERNATIONAL PUBLISHING AG, 2016. 3.4.1

[46] S. W. Brose, D. J. Weber, B. A. Salatin, G. G. Grindle, H. Wang, J. J. Vazquez, and R. A. Cooper. The role of assistive robotics in the lives of persons with disability. *Am J Phys Med Rehabil*, 89(6):509–521, Jun 2010. 3.1.3

[47] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural*

information processing systems, 33:1877–1901, 2020. 2.4.2

[48] Ha-Duong Bui and Nak Young Chong. An integrated approach to human-robot-smart environment interaction interface for ambient assisted living. In *2018 IEEE WORKSHOP ON ADVANCED ROBOTICS AND ITS SOCIAL IMPACTS (ARSO)*, IEEE Workshop on Advanced Robotics and its Social Impacts, pages 32–37. IEEE, 2018. 3.4.3

[49] Jonathan Cacace, Riccardo Caccavale, Alberto Finzi, and Vincenzo Lippiello. Variable admittance control based on virtual fixtures for human-robot co-manipulation. In *2019 IEEE INTERNATIONAL CONFERENCE ON SYSTEMS, MAN AND CYBERNETICS (SMC)*, IEEE International Conference on Systems Man and Cybernetics Conference Proceedings, pages 1569–1574. IEEE, 2019. 3.4.2

[50] Jonathan Cacace, Alberto Finzi, and Vincenzo Lippiello. Enhancing shared control via contact force classification in human-robot cooperative task execution. In F Ficuciello, F Ruggiero, and A Finzi, editors, *HUMAN FRIENDLY ROBOTICS*, volume 7 of *Springer Proceedings in Advanced Robotics*, pages 167–179. SPRINGER INTERNATIONAL PUBLISHING AG, 2019. 3.4.2

[51] Maya Cakmak, Siddhartha S. Srinivasa, Min Kyung Lee, Jodi Forlizzi, and Sara Kiesler. Human preferences for robot-human hand-over configurations. In *2011 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS*, IEEE International Conference on Intelligent Robots and Systems. IEEE, 2011. 3.4.1

[52] Gerard Canal, Guillem Alenya, and Carme Torras. Personalization framework for adaptive robotic feeding assistance. In A Agah, JJ Cabibihan, AM Howard, MA Salichs, and H He, editors, *SOCIAL ROBOTICS, (ICSR 2016)*, volume 9979 of *Lecture Notes in Artificial Intelligence*, pages 22–31. SPRINGER-VERLAG BERLIN, 2016. 3.4.1

[53] Gerard Canal, Emmanuel Pignat, Guillem Alenya, Sylvain Calinon, and Carme Torras. Joining high-level symbolic planning with low-level motion primitives in adaptive hri: application to dressing assistance. In *2018 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA)*, IEEE International Conference on Robotics and Automation ICRA, pages 3273–3278. IEEE COMPUTER SOC, 2018. 3.4.1, 3.5.1

[54] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In *CVPR*, 2017. 4.3

[55] Marc G. Carmichael and Dikai Liu. Admittance control scheme for implementing model-based assistance-as-needed on a robot. In *2013 35TH ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE*

AND BIOLOGY SOCIETY (EMBC), IEEE Engineering in Medicine and Biology Society Conference Proceedings, pages 870–873. IEEE, 2013. 3.4.2

[56] Micah Carroll, Rohin Shah, Mark K Ho, Tom Griffiths, Sanjit Seshia, Pieter Abbeel, and Anca Dragan. On the utility of learning about humans for human-ai coordination. *Advances in neural information processing systems*, 32, 2019. 2.4.3, 7

[57] Elizabeth Cha, Naomi T. Fitter, Yunkyung Kim, Terrence Fong, and Maja J. Mataric. Effects of robot sound on auditory localization in human-robot collaboration. In *HRI '18: PROCEEDINGS OF THE 2018 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION*, ACM IEEE International Conference on Human-Robot Interaction, pages 434–442. ASSOC COMPUTING MACHINERY, 2018. 3.4.3

[58] Wan-Ling Chang and Selma Sabanovic. Interaction expands function: Social shaping of the therapeutic robot paro in a nursing home. In *PROCEEDINGS OF THE 2015 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI'15)*, ACM IEEE International Conference on Human-Robot Interaction, pages 343–350. ASSOC COMPUTING MACHINERY, 2015. 3.4.3

[59] Boyuan Chen, Fei Xia, Brian Ichter, Kanishka Rao, Keerthana Gopalakrishnan, Michael S Ryoo, Austin Stone, and Daniel Kappler. Open-vocabulary queryable scene representations for real world planning. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 11509–11522. IEEE, 2023. 2.4.1

[60] Boyuan Chen, Zhuo Xu, Sean Kirmani, Brian Ichter, Danny Driess, Pete Florence, Dorsa Sadigh, Leonidas Guibas, and Fei Xia. Spatialvlm: Endowing vision-language models with spatial reasoning capabilities. *arXiv preprint arXiv:2401.12168*, 2024. URL `https://arxiv.org/abs/2401.12168`. 6.5

[61] Kangwei Chen, Victoria Plaza-Leiva, Byung-Cheol Min, Aaron Steinfeld, and Mary Bernardine Dias. Navcue: Context immersive navigation assistance for blind travelers. In *ELEVENTH ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN ROBOT INTERACTION (HRI'16)*, ACM IEEE International Conference on Human-Robot Interaction, page 559. ASSOC COMPUTING MACHINERY, 2016. 3.4.3

[62] Lili Chen, Kevin Lu, Aravind Rajeswaran, Kimin Lee, Aditya Grover, Michael Laskin, Pieter Abbeel, Aravind Srinivas, and Igor Mordatch. Decision transformer: Reinforcement learning via sequence modeling, 2021. 7.2.1

[63] Sean Chen, Jensen Gao, Siddharth Reddy, Glen Berseth, Anca D. Dragan, and Sergey Levine. Asha: Assistive teleoperation via human-in-the-loop reinforce-

ment learning. In *2022 International Conference on Robotics and Automation (ICRA)*, pages 7505–7512, 2022. doi: 10.1109/ICRA46639.2022.9812442. 2.4.3

[64] Tiffany L. Chen and Charles C. Kemp. Lead me by the hand: Evaluation of a direct physical interface for nursing assistant robots. In *PROCEEDINGS OF THE 5TH ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI 2010)*, ACMIEEE International Conference on Human-Robot Interaction, pages 367–374. IEEE, 2010. 3.4.2

[65] Zhichao Chen, Yutaka Nakamura, and Hiroshi Ishiguro. Android as a receptionist in a shopping mall using inverse reinforcement learning. *IEEE Robotics and Automation Letters*, 7(3):7091–7098, 2022. doi: 10.1109/LRA.2022.3180042. 2.3.1

[66] Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using RNN encoder–decoder for statistical machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1724–1734, Doha, Qatar, October 2014. Association for Computational Linguistics. doi: 10.3115/v1/D14-1179. URL `https://www.aclweb.org/anthology/D14-1179`. 5.2.2

[67] Hyunjin Choi, Jangmok Lee, and Kyoungchul Kong. A human-robot interface system for walkon suit : a powered exoskeleton for complete paraplegics. In *IECON 2018 - 44TH ANNUAL CONFERENCE OF THE IEEE INDUSTRIAL ELECTRONICS SOCIETY*, IEEE Industrial Electronics Society, pages 5057–5061. IEEE, 2018. 3.4.2

[68] Paul F. Christiano, Jan Leike, Tom B. Brown, Miljan Martic, Shane Legg, and Dario Amodei. Deep reinforcement learning from human preferences. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, page 4302–4310, Red Hook, NY, USA, 2017. Curran Associates Inc. ISBN 9781510860964. 1

[69] Panagiota Christodoulou, Alecia Adelaide May Reid, Dimitrios Pnevmatikos, Carlos Rioja del Rio, and Nikolaos Fachantidis. Students participate and evaluate the design and development of a social robot. In *2020 29TH IEEE INTERNATIONAL CONFERENCE ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 739–744. IEEE, 2020. 3.3.2, 3.4.3

[70] Matei Ciocarlie, Kaijen Hsiao, Adam Leeper, and David Gossow. Mobile manipulation through an assistive home robot. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5313–5320, 2012. doi: 10.1109/IROS.2012.6385907. 2.4.3

[71] Sven Cremer, Kris Doelling, Cody L. Lundberg, Mike McNair, Jeongsik Shin, and Dan Popa. Application requirements for robotic nursing assistants in hospital environments. In D Popa and MBJ Wijesundara, editors, *SENSORS FOR NEXT-GENERATION ROBOTICS III*, volume 9859 of *Proceedings of SPIE*. SPIE-INT SOC OPTICAL ENGINEERING, 2016. 3.4.1

[72] Kevin Crowston. Amazon mechanical turk: A research tool for organizations and information systems scholars. In Anol Bhattacherjee and Brian Fitzgerald, editors, *Shaping the Future of ICT Research. Methods and Approaches*, pages 210–221, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg. 7.4

[73] Tamas Cserteg, Gabor Erdos, and Gergely Horvath. Assisted assembly process by gesture controlled robots. In L Wang, editor, *51ST CIRP CONFERENCE ON MANUFACTURING SYSTEMS*, volume 72 of *Procedia CIRP*, pages 51–56. ELSEVIER SCIENCE BV, 2018. 3.4.1

[74] Yuchen Cui, Scott Niekum, Abhinav Gupta, Vikash Kumar, and Aravind Rajeswaran. Can foundation models perform zero-shot task specification for robot manipulation? In *Learning for Dynamics and Control Conference*, pages 893–905. PMLR, 2022. 2.4.1

[75] Dima Damen, Hazel Doughty, Giovanni Maria Farinella, Sanja Fidler, Antonino Furnari, Evangelos Kazakos, Davide Moltisanti, Jonathan Munro, Toby Perrett, Will Price, and Michael Wray. Scaling egocentric vision: The EPIC-KITCHENS dataset. *CoRR*, abs/1804.02748, 2018. URL `http://arxiv.org/abs/1804.02748`. 2.1.2, 2.1.2

[76] Daniel P. Davison, Frances M. Wijnen, Vicky Charisi, Jan van der Meij, Vanessa Evers, and Dennis Reidsma. Working with a social robot in school: A long-term real-world unsupervised deployment. In *PROCEEDINGS OF THE 2020 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI '20)*, ACM IEEE International Conference on Human-Robot Interaction, pages 63–72. ASSOC COMPUTING MACHINERY, 2020. 3.3.2, 3.4.3

[77] Alessandro De Luca and Fabrizio Flacco. Integrated control for phri: Collision avoidance, detection, reaction and collaboration. In JP Desai, LPS Jay, and L Zollo, editors, *2012 4TH IEEE RAS & EMBS INTERNATIONAL CONFERENCE ON BIOMEDICAL ROBOTICS AND BIOMECHATRONICS (BIOROB)*, Proceedings of the IEEE RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics, pages 288–295. IEEE, 2012. 3.4.3

[78] Joseph DelPreto and Daniela Rus. Sharing the load: Human-robot team lifting using muscle activity. In A Howard, K Althoefer, F Arai, F Arrichiello, B Caputo, J Castellanos, K Hauser, V Isler, J Kim, H Liu, P Oh, V Santos, D Scaramuzza, A Ude, R Voyles, K Yamane, and A Okamura, editors, *2019*

INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA), IEEE International Conference on Robotics and Automation ICRA, pages 7906–7912. IEEE, 2019. 3.4.1, 3.5.3

[79] Joseph DelPreto, Andres F. Salazar-Gomez, Stephanie Gil, Ramin M. Hasani, Frank H. Guenther, and Daniela Rus. Plug-and-play supervisory control using muscle and brain signals for real-time gesture and error detection. In *Robotics: Science and Systems*, 2018. 2.1.1

[80] Sandra Devin and Rachid Alami. An implemented theory of mind to improve human-robot shared plans execution. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*, HRI '16, pages 319–326. IEEE Press, 2016. 3.4.1

[81] Francesca Dino, Rohola Zandie, Hojjat Ahdollahi, Sarah Schoeder, and Mohammad H. Mahoor. Delivering cognitive behavioral therapy using a conversational social robot. In *2019 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS)*, IEEE International Conference on Intelligent Robots and Systems, pages 2089–2095. IEEE, 2019. 3.3.2, 3.4.3

[82] Athanasios C. Dometios, Xanthi S. Papageorgiou, Antonis Arvanitakis, Costas S. Tzafestas, and Petros Maragos. Real-time end-effector motion behavior planning approach using on-line point-cloud data towards a user adaptive assistive bath robot. In A Bicchi and A Okamura, editors, *2017 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS)*, IEEE International Conference on Intelligent Robots and Systems, pages 5031–5036. IEEE, 2017. 3.4.2

[83] Vishnu Sashank Dorbala, Gunnar Sigurdsson, Robinson Piramuthu, Jesse Thomason, and Gaurav S Sukhatme. Clip-nav: Using clip for zero-shot vision-and-language navigation. *arXiv preprint arXiv:2211.16649*, 2022. 2.4.1, 6.3.2

[84] Barzin Doroodgar, Maurizio Ficocelli, Babak Mobedi, and Goldie Nejat. The search for survivors: Cooperative human-robot interaction in search and rescue environments using semi-autonomous robots. In *2010 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA)*, IEEE International Conference on Robotics and Automation ICRA, pages 2858–2863. IEEE, 2010. 3.4.1

[85] Anca Dragan and Siddhartha Srinivasa. A policy blending formalism for shared control. *The International Journal of Robotics Research*, May 2013. 5

[86] Anca D. Dragan, Kenton C.T. Lee, and Siddhartha S. Srinivasa. Legibility and predictability of robot motion. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 301–308, 2013. doi:

10.1109/HRI.2013.6483603. 3.4.3

[87] Kiana Ehsani, Winson Han, Alvaro Herrasti, Eli VanderBilt, Luca Weihs, Eric Kolve, Aniruddha Kembhavi, and Roozbeh Mottaghi. Manipulathor: A framework for visual object manipulation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4497–4506, 2021. 2.2.1

[88] Ilias El Makrini, Kelly Merckaert, Dirk Lefeber, and Bram Vanderborght. Design of a collaborative architecture for human-robot assembly tasks. In A Bicchi and A Okamura, editors, *2017 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS)*, IEEE International Conference on Intelligent Robots and Systems, pages 1624–1629. IEEE, 2017. 3.4.1, 3.5.3

[89] Anthropic Team et. al. Claude 3 haiku, 2024. 1, 6.1

[90] Gemini Team et. al. Gemini: A family of highly capable multimodal models, 2024. 1, 6.1

[91] OpenAI et. al. Gpt-4 technical report, 2023. 1, 6.1

[92] Jing Fan, Linda Beuscher, Paul A. Newhouse, Lorraine C. Mion, and Nilanjan Sarkar. A robotic coach architecture for multi-user human-robot interaction (ramu) with the elderly and cognitively impaired. In *2016 25TH IEEE INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 445–450. IEEE, 2016. 3.3.3, 3.4.3

[93] Miguel Faria, Rui Silva, Patricia Alves-Oliveira, Francisco S. Melo, and Ana Paiva. "me and you together" movement impact in multi-user collaboration tasks. In A Bicchi and A Okamura, editors, *2017 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS)*, IEEE International Conference on Intelligent Robots and Systems, pages 2793–2798. IEEE, 2017. 3.4.3

[94] Logan Farrell, Philip Strawser, Kimberly Hambuchen, William Baker, and Julia Badger. Supervisory control of a humanoid robot in microgravity for manipulation tasks. In A Bicchi and A Okamura, editors, *2017 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS)*, IEEE International Conference on Intelligent Robots and Systems, pages 3797–3802. IEEE, 2017. 3.4.1

[95] Alireza Fathi, Yin Li, and James M. Rehg. Learning to recognize daily actions using gaze. In Andrew Fitzgibbon, Svetlana Lazebnik, Pietro Perona, Yoichi Sato, and Cordelia Schmid, editors, *Computer Vision – ECCV 2012*, pages 314–327, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg. ISBN 978-3-642-

33718-5. 2.1.2, 2.1.2

[96] Beatriz Quintino Ferreira, Kelly Karipidou, Filipe Rosa, Sofia Petisca, Patricia Alves-Oliveira, and Ana Paiva. A study on trust in a robotic suitcase. In A Agah, JJ Cabibihan, AM Howard, MA Salichs, and H He, editors, *SOCIAL ROBOTICS, (ICSR 2016)*, volume 9979 of *Lecture Notes in Artificial Intelligence*, pages 179–189. SPRINGER-VERLAG BERLIN, 2016. 3.4.1

[97] Roya Firoozi, Johnathan Tucker, Stephen Tian, Anirudha Majumdar, Jiankai Sun, Weiyu Liu, Yuke Zhu, Shuran Song, Ashish Kapoor, Karol Hausman, et al. Foundation models in robotics: Applications, challenges, and the future. *arXiv preprint arXiv:2312.07843*, 2023. 2.4.1

[98] Kerstin Fischer, Lars Christian Jensen, Franziska Kirstein, Sebastian Stabinger, Oezguer Erkent, Dadhichi Shukla, and Justus Piater. The effects of social gaze in human-robot collaborative assembly. In A Tapus, E Andre, JC Martin, F Ferland, and M Ammi, editors, *SOCIAL ROBOTICS (ICSR 2015)*, volume 9388 of *Lecture Notes in Artificial Intelligence*, pages 204–213. SPRINGER-VERLAG BERLIN, 2015. 3.4.3, 3.5.3

[99] T Fong, C Thorpe, and C Baur. Collaboration, dialogue, and human-robot interaction. In RA Jarvis and A Zelinsky, editors, *ROBOTICS RESEARCH*, volume 6 of *SPRINGER TRACTS IN ADVANCED ROBOTICS*, pages 255–266. SPRINGER-VERLAG BERLIN, 2003. 3.1.3, 3.4.3

[100] Kristin S. Fuglerud and Ivar Solheim. The use of social robots for supporting language training of children kristin. In G Craddock, C Doran, L McNutt, and D Rice, editors, *TRANSFORMING OUR WORLD THROUGH DESIGN, DIVERSITY AND EDUCATION*, volume 256 of *Studies in Health Technology and Informatics*, pages 401–408. IOS PRESS, 2018. 3.4.3

[101] Samir Yitzhak Gadre, Mitchell Wortsman, Gabriel Ilharco, Ludwig Schmidt, and Shuran Song. Cows on pasture: Baselines and benchmarks for language-driven zero-shot object navigation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 23171–23181, 2023. 2.4.1

[102] Edwinn Gamborino and Li-Chen Fu. Interactive reinforcement learning based assistive robot for the emotional support of children. In *2018 18TH INTERNATIONAL CONFERENCE ON CONTROL, AUTOMATION AND SYSTEMS (ICCAS)*, International Conference on Control Automation and Systems, pages 708–713. IEEE, 2018. 3.4.3

[103] Chuang Gan, Siyuan Zhou, Jeremy Schwartz, Seth Alter, Abhishek Bhandwaldar, Dan Gutfreund, Daniel LK Yamins, James J DiCarlo, Josh McDermott, Antonio Torralba, et al. The threedworld transport challenge: A visually guided task-and-motion planning benchmark for physically realistic embodied ai. *arXiv*

preprint arXiv:2103.14025, 2021. 2.2.1

[104] Xiaofeng Gao, Ran Gong, Yizhou Zhao, Shu Wang, Tianmin Shu, and Song-Chun Zhu. Joint mind modeling for explanation generation in complex human-robot collaborative tasks. In *2020 29TH IEEE INTERNATIONAL CONFERENCE ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 1119–1126. IEEE, 2020. 3.4.3

[105] Yuan Gao, Elena Sibirtseva, Ginevra Castellano, and Danica Kragic. Fast adaptation with meta-reinforcement learning for trust modelling in human-robot interaction. In *2019 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS)*, IEEE International Conference on Intelligent Robots and Systems, pages 305–312. IEEE, 2019. 3.4.3

[106] O. Can Goeruer, Benjamin Rosman, Fikret Sivrikaya, and Sahin Albayrak. Social cobots: Anticipatory decision-making for collaborative robots incorporating unexpected human behaviors. In *HRI '18: PROCEEDINGS OF THE 2018 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION*, ACM IEEE International Conference on Human-Robot Interaction, pages 398–406. ASSOC COMPUTING MACHINERY, 2018. 3.4.1

[107] Felix Ferdinand Goldau, Tejas Kumar Shastha, Maria Kyrarini, and Axel Graeser. Autonomous multi-sensory robotic assistant for a drinking task. In *2019 IEEE 16TH INTERNATIONAL CONFERENCE ON REHABILITATION ROBOTICS (ICORR)*, International Conference on Rehabilitation Robotics ICORR, pages 210–216. IEEE, 2019. 3.4.1

[108] Michael A. Goodrich, Mark Colton, Bonnie Brinton, and Martin Fujiki. A case for low-dose robotics in autism therapy. In *PROCEEDINGS OF THE 6TH ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTIONS (HRI 2011)*, ACM IEEE International Conference on Human-Robot Interaction, pages 143–144. IEEE, 2011. 3.4.3

[109] Deepak Gopinath, Siddarth Jain, and Brenna D Argall. Human-in-the-loop optimization of shared autonomy in assistive robotics. *IEEE Robotics and Automation Letters*, 2(1):247–254, 2016. 5

[110] Eliza M. Grames, Andrew N. Stillman, Morgan W. Tingley, and Chris S. Elphick. An automated approach to identifying search terms for systematic reviews using keyword co-occurrence networks. *Methods in Ecology and Evolution*, 10(10):1645–1654, 2019. doi: https://doi.org/10.1111/2041-210X.13268. URL `https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.13268`. 3.2

[111] Elena Corina Grigore, Kerstin Eder, Anthony G. Pipe, Chris Melhuish, and Ute

Leonards. Joint action understanding improves robot-to-human object handover. In N Amato, editor, *2013 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS)*, IEEE International Conference on Intelligent Robots and Systems, pages 4622–4629. IEEE, 2013. 3.4.1

[112] Arzu Guneysu and Bert Arnrich. Socially assistive child-robot interaction in physical exercise coaching. In A Howard, K Suzuki, and L Zollo, editors, *2017 26TH IEEE INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 670–675. IEEE, 2017. 3.4.3

[113] Dylan Hadfield-Menell, Anca Dragan, Pieter Abbeel, and Stuart Russell. Cooperative inverse reinforcement learning. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*, NIPS'16, page 3916–3924, Red Hook, NY, USA, 2016. Curran Associates Inc. ISBN 9781510838819. 1, 2.3, 2.3.2

[114] Kelsey P. Hawkins, Nam Vo, Shray Bansal, and Aaron F. Bobick. Probabilistic human action prediction and wait-sensitive planning for responsive human-robot collaboration. In *2013 13TH IEEE-RAS INTERNATIONAL CONFERENCE ON HUMANOID ROBOTS (HUMANOIDS)*, IEEE-RAS International Conference on Humanoid Robots, pages 499–506. IEEE, 2013. 3.4.1

[115] Kelsey P. Hawkins, Shray Bansal, Nam N. Vo, and Aaron F. Bobick. Anticipating human actions for collaboration in the presence of task and sensor uncertainty. In *2014 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA)*, IEEE International Conference on Robotics and Automation ICRA, pages 2215–2222. IEEE, 2014. 3.4.1

[116] Mary Hayhoe and Dana Ballard. Eye movements in natural behavior. *Trends in Cognitive Sciences*, 9(4):188–194, 2005. ISSN 1364-6613. doi: 10.1016/j.tics.2005.02.009. 2.1.1, 2.1.3, 4, 5

[117] Rafi Hayne, Ruikun Luo, and Dmitry Berenson. Considering avoidance and consistency in motion planning for human-robot manipulation in a shared workspace. In A Okamura, A Menciassi, A Ude, D Burschka, D Lee, F Arrichiello, H Liu, H Moon, J Neira, K Sycara, K Yokoi, P Martinet, P Oh, P Valdastri, and V Krovi, editors, *2016 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA)*, IEEE International Conference on Robotics and Automation ICRA, pages 3948–3954. IEEE, 2016. 3.4.3

[118] Jerry Zhi-Yang He, Zackory Erickson, Daniel S. Brown, Aditi Raghunathan, and Anca Dragan. Learning representations that enable generalization in assistive tasks. In *6th Annual Conference on Robot Learning*, 2022. URL

https://openreview.net/forum?id=b88HF4vd_ej. 2.4.3

[119] Jacqueline Hemminghaus and Stefan Kopp. Towards adaptive social behavior generation for assistive robots using reinforcement learning. In *PROCEEDINGS OF THE 2017 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI'17)*, ACM IEEE International Conference on Human-Robot Interaction, pages 332–340. IEEE, 2017. 3.4.3

[120] Laura V. Herlant, Rachel M. Holladay, and Siddhartha S. Srinivasa. Assistive teleoperation of robot arms via automatic time-optimal mode switching. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*, HRI '16, pages 35–42. IEEE Press, 2016. 3.4.3, 3.5.1, 3.5.2

[121] Antti Hietanen, Alireza Changizi, Minna Lanz, Joni Kamarainen, Pallab Ganguly, Roel Pieters, and Jyrki Latokartano. Proof of concept of a projection-based safety system for human-robot collaborative engine assembly. In *2019 28TH IEEE INTERNATIONAL CONFERENCE ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN. IEEE, 2019. 3.4.3

[122] Guy Hoffman. Evaluating fluency in human–robot collaboration. *IEEE Transactions on Human-Machine Systems*, 49(3):209–218, 2019. doi: 10.1109/THMS.2019.2904558. 2.3.1, 7.1

[123] Guy Hoffman, Gurit E. Birnbaum, Keinan Vanunu, Omri Sass, and Harry T. Reis. Robot responsiveness to human disclosure affects social impression and appeal. In *HRI'14: PROCEEDINGS OF THE 2014 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION*, ACM IEEE International Conference on Human-Robot Interaction, pages 1–8. IEEE, 2014. 3.4.3, 3.5.1

[124] Y Horiguchi, T Sawaragi, and G Akashi. Naturalistic human-robot collaboration based upon mixed-initiative interactions in teleoperating environment. In *SMC 2000 CONFERENCE PROCEEDINGS: 2000 IEEE INTERNATIONAL CONFERENCE ON SYSTEMS, MAN & CYBERNETICS, VOL 1-5*, IEEE International Conference on Systems Man and Cybernetics Conference Proceedings, pages 876–881. IEEE, 2000. 3.4.3

[125] Yafei Hu, Quanting Xie, Vidhi Jain, Jonathan Francis, Jay Patrikar, Nikhil Keetha, Seungchan Kim, Yaqi Xie, Tianyi Zhang, Zhibo Zhao, et al. Toward general-purpose robots via foundation models: A survey and meta-analysis. *arXiv preprint arXiv:2312.08782*, 2023. 2.4.1

[126] Chenguang Huang, Oier Mees, Andy Zeng, and Wolfram Burgard. Visual language maps for robot navigation. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 10608–10615. IEEE, 2023. 2.4.1

[127] Chien-Ming Huang and Bilge Mutlu. Anticipatory robot control for efficient human-robot collaboration. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*, HRI '16, pages 83–90. IEEE Press, 2016. 2.1.1, 3.4.1, 3.5.1, 3.5.2

[128] Wenlong Huang, Pieter Abbeel, Deepak Pathak, and Igor Mordatch. Language models as zero-shot planners: Extracting actionable knowledge for embodied agents. In *International Conference on Machine Learning*, pages 9118–9147. PMLR, 2022. 2.4.1

[129] Wenlong Huang, Fei Xia, Ted Xiao, Harris Chan, Jacky Liang, Pete Florence, Andy Zeng, Jonathan Tompson, Igor Mordatch, Yevgen Chebotar, et al. Inner monologue: Embodied reasoning through planning with language models. *arXiv preprint arXiv:2207.05608*, 2022. 2.4.1

[130] Wenlong Huang, Chen Wang, Ruohan Zhang, Yunzhu Li, Jiajun Wu, and Li Fei-Fei. Voxposer: Composable 3d value maps for robotic manipulation with language models. *arXiv preprint arXiv:2307.05973*, 2023. 2.4.1

[131] I Iossifidis and G Schoner. Autonomous reaching and obstacle avoidance with the anthropomorphic arm of a robotic assistant using the attractor dynamics approach. In *2004 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, VOLS 1- 5, PROCEEDINGS*, IEEE International Conference on Robotics and Automation ICRA, pages 4295–4300. IEEE, 2004. 3.4.1

[132] Tariq Iqbal and Laurel D. Riek. Human-Robot Teaming: Approaches from Joint Action and Dynamical Systems. In *Humanoid Robotics: A Reference*, pages 2293–2312. Springer Netherlands, 2019. doi: 10.1007/978-94-007-6046-2_137. URL https://doi.org/10.1007/978-94-007-6046-2_137. 3.1.3

[133] Shunki Itadera, Taisuke Kobayashi, Jun Nakanishi, Tadayoshi Aoyama, and Yasuhisa Hasegawa. Impedance control based assistive mobility aid through online classification of user's state. In *2019 IEEE/SICE INTERNATIONAL SYMPOSIUM ON SYSTEM INTEGRATION (SII)*, IEEE/SICE International Symposium on System Integration, pages 243–248. IEEE, 2019. 3.4.2

[134] Ekaterina Ivanova, Axel Krause, Marie Schaelicke, Franziska Schellhardt, Natalie Jankowski, Josy Achner, Henning Schmidt, Michael Joebges, and Joerg Krueger. Let's do this together: Bi-manu-interact, a novel device for studying human haptic interactive behavior. In F Amirabdollahian, E Burdet, and L Masia, editors, *2017 INTERNATIONAL CONFERENCE ON REHABILITATION ROBOTICS (ICORR)*, International Conference on Rehabilitation Robotics ICORR, pages 708–713. IEEE, 2017. 3.3.3, 3.4.2

[135] Nathanael Jarrasse, Jamie Paik, Viviane Pasqui, and Guillaume Morel. How

can human motion prediction increase transparency? In *2008 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, VOLS 1-9*, IEEE International Conference on Robotics and Automation ICRA, pages 2134–2139. IEEE, 2008. 3.4.2

[136] Nathanaël Jarrassé, Themistoklis Charalambous, and Etienne Burdet. A Framework to Describe, Analyze and Generate Interactive Motor Behaviors. *PLoS ONE*, 7(11):e49945, nov 2012. ISSN 19326203. doi: 10.1371/journal.pone.0049945. 3.1.1, 3.1.3

[137] Shervin Javdani, Siddhartha Srinivasa, and Andrew Bagnell. Shared autonomy via hindsight optimization. In *Proceedings of Robotics: Science and Systems*, Rome, Italy, July 2015. doi: 10.15607/RSS.2015.XI.032. 2.1.1, 3.4.3, 3.5.1, 3.5.3, 4.5, 5

[138] Shervin Javdani, Henny Admoni, Stefania Pellegrinelli, Siddhartha S. Srinivasa, and J. Andrew Bagnell. Shared autonomy via hindsight optimization for teleoperation and teaming. *The International Journal of Robotics Research*, 37 (7):717–742, 2018. doi: 10.1177/0278364918776060. URL https://doi.org/10.1177/0278364918776060. 1, 2.1.1, 2.3.1, 3.5.3, 4, 4.1.3, 7

[139] D. B. Jayagopi, S. Sheiki, D. Klotz, J. Wienke, J. Odobez, S. Wrede, V. Khalidov, L. Nyugen, B. Wrede, and D. Gatica-Perez. The vernissage corpus: A conversational human-robot-interaction dataset. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 149–150, March 2013. doi: 10.1109/HRI.2013.6483545. 2.1.2

[140] Lars Christian Jensen, Kerstin Fischer, Stefan-Daniel Suvei, and Leon Bodenhagen. Timing of multimodal robot behaviors during human-robot collaboration. In A Howard, K Suzuki, and L Zollo, editors, *2017 26TH IEEE INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 1061–1066. IEEE, 2017. 3.4.1

[141] Roland S. Johansson, Göran Westling, Anders Bäckström, and J. Randall Flanagan. Eye–hand coordination in object manipulation. *Journal of Neuroscience*, 21 (17):6917–6932, 2001. ISSN 0270-6474. doi: 10.1523/JNEUROSCI.21-17-06917.2001. URL http://www.jneurosci.org/content/21/17/6917. 2.1.1, 2.1.3, 4

[142] W. Kabsch. A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A*, 32(5):922–923, Sep 1976. doi: 10.1107/S0567739476001873. URL https://doi.org/10.1107/S0567739476001873. 6.2.3

[143] Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelli-*

gence, 101(1):99 – 134, 1998. ISSN 0004-3702. doi: https://doi.org/10.1016/S0004-3702(98)00023-X. URL http://www.sciencedirect.com/science/article/pii/S000437029800023X. 4.1.3

[144] Junko Kanero, Idil Franko, Cansu Oranc, Orhun Ulusahin, Sumeyye Koskulu, Zeynep Adiguzel, Aylin C. Kuntay, and Tilbe Goksun. Who can benefit from robots? effects of individual differences in robot-assisted language learning. In *2018 JOINT IEEE 8TH INTERNATIONAL CONFERENCE ON DEVELOPMENT AND LEARNING AND EPIGENETIC ROBOTICS (ICDL-EPIROB)*, Joint IEEE International Conference on Development and Learning and Epigenetic Robotics ICDL-EpiRob, pages 212–217. IEEE, 2018. 3.4.3

[145] Mincheul Kang, Youngsun Kwon, and Sung-Eui Yoon. Automated task planning using object arrangement optimization. In *2018 15th International Conference on Ubiquitous Robots (UR)*, pages 334–341. IEEE, 2018. 2.2.1, 6

[146] Yash Kant, Arun Ramachandran, Sriram Yenamandra, Igor Gilitschenski, Dhruv Batra, Andrew Szot, and Harsh Agrawal. Housekeep: Tidying virtual households using commonsense reasoning. In *European Conference on Computer Vision*, pages 355–373. Springer, 2022. 2.2.1

[147] Ivan Kapelyukh and Edward Johns. My house, my rules: Learning tidying preferences with graph neural networks. In *Conference on Robot Learning*, pages 740–749. PMLR, 2022. 1, 2.2.1, 6, 6.5

[148] Ivan Kapelyukh, Vitalis Vosylius, and Edward Johns. Dall-e-bot: Introducing web-scale diffusion models to robotics. In *IEEE Robotics and Automation Letters (RA-L)*, 2023. 6.5, 8.1

[149] Moritz Kassner, William Patera, and Andreas Bulling. Pupil: An open source platform for pervasive eye tracking and mobile gaze-based interaction. In *Adjunct Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '14 Adjunct, pages 1151–1160, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-3047-3. doi: 10.1145/2638728.2641695. URL http://doi.acm.org/10.1145/2638728.2641695. 4.1.4

[150] James Kennedy, Paul Baxter, Emmanuel Senft, and Tony Belpaeme. Social robot tutoring for child second language learning. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*, HRI '16, pages 231–238. IEEE Press, 2016. 3.4.3

[151] Dae-Jin Kim, Rebekah Hazlett-Knudsen, Heather Culver-Godfrey, Greta Rucks, Tara Cunningham, David Portee, John Bricout, Zhao Wang, and Aman Behal. How autonomy impacts performance and satisfaction: Results from a study with spinal cord injured subjects using an assistive robot. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 42(1):2–14,

2012. 5

[152] Min-Gyu Kim, Iris Oosterling, Tino Lourens, Wouter Staal, Jan Buitelaar, Jeffrey Glennon, Iris Smeekens, and Emilia Barakova. Designing robot-assisted pivotal response training in game activity for children with autism. In *2014 IEEE INTERNATIONAL CONFERENCE ON SYSTEMS, MAN AND CYBERNETICS (SMC)*, IEEE International Conference on Systems Man and Cybernetics Conference Proceedings, pages 1101–1106. IEEE, 2014. 3.4.3

[153] Diederick P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015. 5.2.2

[154] Ross A. Knepper, Christoforos I. Mavrogiannis, Julia Proft, and Claire Liang. Implicit communication in a joint action. In *PROCEEDINGS OF THE 2017 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI'17)*, ACM IEEE International Conference on Human-Robot Interaction, pages 283–292. IEEE, 2017. 3.4.3

[155] Jonathan Kofman, Xianghai Wu, Timothy J Luu, and Siddharth Verma. Teleoperation of a robot manipulator using a vision-based human-robot interface. *IEEE Transactions on Industrial Electronics*, 52(5):1206–1219, 2005. 2.1.1

[156] Dimosthenis Kontogiorgos, Andre Pereira, Boran Sahindal, Sanne van Waveren, and Joakim Gustafson. Behavioural responses to robot conversational failures. In *PROCEEDINGS OF THE 2020 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI '20)*, ACM IEEE International Conference on Human-Robot Interaction, pages 53–62. ASSOC COMPUTING MACHINERY, 2020. 3.4.3

[157] Hema S. Koppula, Ashesh Jain, and Ashutosh Saxena. Anticipatory planning for human-robot teams. In MA Hsieh, O Khatib, and V Kumar, editors, *EXPERIMENTAL ROBOTICS*, volume 109 of *Springer Tracts in Advanced Robotics*, pages 453–470. SPRINGER INTERNATIONAL PUBLISHING AG, 2016. 3.4.1

[158] Maria Koskinopoulou, Stylimos Piperakis, and Panos Frahanias. Learning from demonstration facilitates human-robot collaborative task execution. In *ELEVENTH ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN ROBOT INTERACTION (HRI'16)*, ACM IEEE International Conference on Human-Robot Interaction, pages 59–66. ASSOC COMPUTING MACHINERY, 2016. 3.4.1

[159] Panagiotis N. Koustoumpardis, Konstantinos I. Chatzilygeroudis, Aris I. Synodinos, and Nikos A. Aspragathos. Human robot collaboration for folding fabrics based on force/rgb-d feedback. In T Borangiu, editor, *ADVANCES IN*

ROBOT DESIGN AND INTELLIGENT CONTROL, volume 371 of *Advances in Intelligent Systems and Computing*, pages 235–243. SPRINGER-VERLAG BERLIN, 2016. 3.4.1

[160] Sven Kratz and Fred Rabelo Ferriera. Immersed remotely: Evaluating the use of head mounted devices for remote collaboration in robotic telepresence. In *2016 25TH IEEE INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 638–645. IEEE, 2016. 3.4.1

[161] Krzysztof Krejtz, Andrew T. Duchowski, Anna Niedzielska, Cezary Biele, and Izabela Krejtz. Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze. *PLOS ONE*, 13(9):1–23, 09 2018. doi: 10.1371/journal.pone.0203629. URL https://doi.org/10.1371/journal.pone.0203629. 4

[162] Woo Young Kwon and Il Hong Suh. A temporal bayesian network with application to design of a proactive robotic assistant. In *2012 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA)*, IEEE International Conference on Robotics and Automation ICRA, pages 3685–3690. IEEE, 2012. 3.4.1

[163] Erik Kyrkjebo, Mads Johan Laastad, and Oyvind Stavdahl. Feasibility of the ur5 industrial robot for robotic rehabilitation of the upper limbs after stroke. In AA Maciejewski, A Okamura, A Bicchi, C Stachniss, DZ Song, DH Lee, F Chaumette, H Ding, JS Li, J Wen, J Roberts, K Masamune, NY Chong, N Amato, N Tsagwarakis, P Rocco, T Asfour, WK Chung, Y Yasuyoshi, Y Sun, T Maciekeski, K Althoefer, J AndradeCetto, WK Chung, E Demircan, J Dias, P Fraisse, R Gross, H Harada, Y Hasegawa, M Hayashibe, K Kiguchi, K Kim, T Kroeger, Y Li, S Ma, H Mochiyama, CA Monje, I Rekleitis, R Roberts, F Stulp, CHD Tsai, and L Zollo, editors, *2018 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS)*, IEEE International Conference on Intelligent Robots and Systems, pages 6124–6129. IEEE, 2018. 3.4.2

[164] Jens Lambrecht and Sebastian Nimpsch. Human prediction for the natural instruction of handovers in human robot collaboration. In *2019 28TH IEEE INTERNATIONAL CONFERENCE ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN. IEEE, 2019. 3.4.1

[165] Michael F. Land and Mary Hayhoe. In what ways do eye movements contribute to everyday activities? *Vision Research*, 41(25):3559–3565, 2001. 2.1.1, 2.1.3, 4, 5

[166] H. O. Latif, N. Sherkat, and A. Lotfi. Teleoperation through eye gaze (telegaze): A multimodal approach. In *2009 IEEE International Conference on Robotics*

and Biomimetics (ROBIO), pages 711–716, Dec 2009. doi: 10.1109/ROBIO.2009.5420585. 2.1.3

[167] Edith Law, Vicky Cai, Qi Feng Liu, Sajin Sasy, Joslin Goh, Alex Blidaru, and Dana Kulic. A wizard-of-oz study of curiosity in human-robot interaction. In A Howard, K Suzuki, and L Zollo, editors, *2017 26TH IEEE INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 607–614. IEEE, 2017. 3.4.3

[168] Hyunglae Lee and Neville Hogan. Essential considerations for design and control of human-interactive robots. In A Okamura, A Menciassi, A Ude, D Burschka, D Lee, F Arrichiello, H Liu, H Moon, J Neira, K Sycara, K Yokoi, P Martinet, P Oh, P Valdastri, and V Krovi, editors, *2016 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA)*, IEEE International Conference on Robotics and Automation ICRA, pages 3069–3074. IEEE, 2016. 3.4.2

[169] Iolanda Leite, Marissa McCoy, Monika Lohani, Daniel Ullman, Nicole Salomons, Charlene Stokes, Susan Rivers, and Brian Scassellati. Emotional storytelling in the classroom: Individual versus group interaction between children and robots. In *PROCEEDINGS OF THE 2015 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI'15)*, ACM IEEE International Conference on Human-Robot Interaction, pages 75–82. ASSOC COMPUTING MACHINERY, 2015. 3.3.2, 3.4.3

[170] Brian Lester, Rami Al-Rfou, and Noah Constant. The power of scale for parameter-efficient prompt tuning. *arXiv preprint arXiv:2104.08691*, 2021. 2.4.2

[171] Xiang Lisa Li and Percy Liang. Prefix-tuning: Optimizing continuous prompts for generation. *arXiv preprint arXiv:2101.00190*, 2021. 2.4.2

[172] Yanan Li, Keng Peng Tee, Wei Liang Chan, Rui Yan, Yuanwei Chua, and Dilip Kumar Limbu. Role adaptation of human and robot in collaborative tasks. In *2015 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA)*, IEEE International Conference on Robotics and Automation ICRA, pages 5602–5607. IEEE COMPUTER SOC, 2015. 3.4.2

[173] Zhongyu Li and Ralph Hollis. Toward a ballbot for physically leading people: A human-centered approach. In *2019 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS)*, IEEE International Conference on Intelligent Robots and Systems, pages 4827–4833. IEEE, 2019. 3.4.3

[174] Jacky Liang, Wenlong Huang, Fei Xia, Peng Xu, Karol Hausman, Brian Ichter, Pete Florence, and Andy Zeng. Code as policies: Language model programs

for embodied control. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 9493–9500. IEEE, 2023. 2.4.1

[175] Donghwan Lim, Wansoo Kim, Heedon Lee, Hojun Kim, Kyoosik Shin, Taejoon Park, JiYeong Lee, and Changsoo Han. Development of a lower extremity exoskeleton robot with a quasi-anthropomorphic design approach for load carriage. In *2015 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS)*, IEEE International Conference on Intelligent Robots and Systems, pages 5345–5350. IEEE, 2015. 3.4.2

[176] Kevin Lin, Christopher Agia, Toki Migimatsu, Marco Pavone, and Jeannette Bohg. Text2motion: From natural language instructions to feasible plans. *arXiv preprint arXiv:2303.12153*, 2023. 2.4.1

[177] Tsung-Chi Lin, Achyuthan Unni Krishnan, and Zhi Li. Physical fatigue analysis of assistive robot teleoperation via whole-body motion mapping. In *2019 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS)*, IEEE International Conference on Intelligent Robots and Systems, pages 2240–2245. IEEE, 2019. 3.4.3

[178] Michael L. Littman, Anthony R. Cassandra, and Leslie Pack Kaelbling. Learning policies for partially observable environments: Scaling up. In Armand Prieditis and Stuart Russell, editors, *Machine Learning Proceedings 1995*, pages 362 – 370. Morgan Kaufmann, San Francisco (CA), 1995. ISBN 978-1-55860-377-6. doi: https://doi.org/10.1016/B978-1-55860-377-6.50052-9. URL http://www.sciencedirect.com/science/article/pii/B9781558603776500529. 2.3.2, 4.1.3

[179] Bo Liu, Yuqian Jiang, Xiaohan Zhang, Qiang Liu, Shiqi Zhang, Joydeep Biswas, and Peter Stone. Llm+ p: Empowering large language models with optimal planning proficiency. *arXiv preprint arXiv:2304.11477*, 2023. 2.4.1

[180] Fangchen Liu, Kuan Fang, Pieter Abbeel, and Sergey Levine. Moka: Open-vocabulary robotic manipulation through mark-based visual prompting. *arXiv preprint arXiv:2403.03174*, 2024. 2.4.2

[181] Hongyi Liu, Yuquan Wang, Wei Ji, and Lihui Wang. A context-aware safety system for human-robot collaboration. In D Sormaz, G Suer, and FF Chen, editors, *28TH INTERNATIONAL CONFERENCE ON FLEXIBLE AUTOMATION AND INTELLIGENT MANUFACTURING (FAIM2018): GLOBAL INTEGRATION OF INTELLIGENT MANUFACTURING AND SMART INDUSTRY FOR GOOD OF HUMANITY*, volume 17 of *Procedia Manufacturing*, pages 238–245. ELSEVIER SCIENCE BV, 2018. 3.4.3

[182] Zeyi Liu, Arpit Bahety, and Shuran Song. Reflect: Summarizing robot experiences for failure explanation and correction. *arXiv preprint arXiv:2306.15724*,

2023. 2.4.1

[183] Dylan P. Losey, Craig G. McDonald, Edoardo Battaglia, and Marcia K. O'Malley. A review of intent detection, arbitration, and communication aspects of shared control for physical human–robot interaction, jan 2018. ISSN 00036900. URL http://asmedigitalcollection.asme.org/appliedmechanicsreviews/article-pdf/70/1/010804/5964415/amr_070_01_010804.pdf. 3.5.3

[184] Dylan P. Losey, Krishnan Srinivasan, Ajay Mandlekar, Animesh Garg, and Dorsa Sadigh. Controlling assistive robots with learned latent actions. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 378–384, 2020. doi: 10.1109/ICRA40945.2020.9197197. 1

[185] Dylan P Losey, Andrea Bajcsy, Marcia K O'Malley, and Anca D Dragan. Physical interaction as communication: Learning robot objectives online from human corrections. *The International Journal of Robotics Research*, 41(1): 20–44, 2022. 1, 2.3.1, 2.3.2, 7

[186] Wing-Yue Geoffrey Louie, Tiago Vaquero, Goldie Nejat, and J. Christopher Beck. An autonomous assistive robot for planning, scheduling and facilitating multi-user activities. In *2014 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA)*, IEEE International Conference on Robotics and Automation ICRA, pages 5292–5298. IEEE, 2014. 3.4.3

[187] Nichola Lubold, Erin Walker, and Heather Pon-Barry. Effects of voice-adaptation and social dialogue on perceptions of a robotic learning companion. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*, HRI '16, pages 255–262. IEEE Press, 2016. 3.4.3

[188] Jing Luo, Chenguang Yang, Etienne Burdet, and Yanan Li. Adaptive impedance control with trajectory adaptation for minimizing interaction force. In *2020 29TH IEEE INTERNATIONAL CONFERENCE ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 1360–1365. IEEE, 2020. 3.4.2

[189] Yecheng Jason Ma, William Liang, Guanzhi Wang, De-An Huang, Osbert Bastani, Dinesh Jayaraman, Yuke Zhu, Linxi Fan, and Anima Anandkumar. Eureka: Human-level reward design via coding large language models. *arXiv preprint arXiv:2310.12931*, 2023. 2.4.1

[190] Tamotsu Machino, Satoshi Iwaki, Hiroaki Kawata, Yoshimasa Yanagihara, Yoshito Nanjo, and Ken-ichiro Shimokura. Remote-collaboration system using mobile robot with camera and projector. In *2006 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA), VOLS 1-10*, IEEE International Conference on Robotics and Automation ICRA, pages 4063+. IEEE, 2006. 3.3.4, 3.4.3

[191] G. Maeda, M. Ewerton, R. Lioutikov, H. Ben Amor, J. Peters, and G. Neumann. Learning interaction for collaborative tasks with probabilistic movement primitives. In *2014 IEEE-RAS International Conference on Humanoid Robots*, pages 527–534, Nov 2014. doi: 10.1109/HUMANOIDS.2014.7041413. 2.1.3

[192] Guilherme Maeda, Marco Ewerton, Rudolf Lioutikov, Heni Ben Amor, Jan Peters, and Gerhard Neumann. Learning interaction for collaborative tasks with probabilistic movement primitives. In *2014 14th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, IEEE-RAS International Conference on Humanoid Robots, pages 527–534. IEEE, 2014. 3.4.1

[193] J. Mainprice, R. Hayne, and D. Berenson. Predicting human reaching motion in collaborative tasks using inverse optimal control and iterative re-planning. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 885–892, May 2015. doi: 10.1109/ICRA.2015.7139282. 4

[194] Norjasween Abdul Malik, Hanafiah Yussof, and Fazah Akhtar Hanapiah. Development of imitation learning through physical therapy using a humanoid robot. In H Yussof, editor, *MEDICAL AND REHABILITATION ROBOTICS AND INSTRUMENTATION (MRRI2013)*, pages 191–197. ELSEVIER SCIENCE BV, 2014. 3.4.3

[195] Norjasween Abdul Malik, Hanafiah Yussof, and Fazah Akhtar Hanapiah. Interactive scenario development of robot-assisted therapy for cerebral palsy: A face validation survey. In H Yussof, editor, *2016 IEEE INTERNATIONAL SYMPOSIUM ON ROBOTICS AND INTELLIGENT SENSORS (IRIS 2016)*, volume 105 of *Procedia Computer Science*, pages 322–327. ELSEVIER SCIENCE BV, 2017. 3.4.3

[196] Eleonora Mariotti, Emanuele Magrini, and Alessandro De Luca. Admittance control for human-robot interaction using an industrial robot equipped with a f/t sensor. In A Howard, K Althoefer, F Arai, F Arrichiello, B Caputo, J Castellanos, K Hauser, V Isler, J Kim, H Liu, P Oh, V Santos, D Scaramuzza, A Ude, R Voyles, K Yamane, and A Okamura, editors, *2019 INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA)*, IEEE International Conference on Robotics and Automation ICRA, pages 6130–6136. IEEE, 2019. 3.4.2

[197] Nikolas Martelaro, Victoria C. Nneji, Wendy Ju, and Pamela Hinds. Tell me more: Designing hri to encourage more trust, disclosure, and companionship. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*, HRI '16, pages 181–188. IEEE Press, 2016. 3.3.2, 3.4.3

[198] Maja Mataric, Adriana Tapus, Carolee Winstein, and Jon Eriksson. Socially assistive robotics for stroke and mild tbi rehabilitation. In A Gaggioli, EA Keshner, PL Weiss, and G Riva, editors, *ADVANCED TECHNOLOGIES IN REHA-*

BILITATION: EMPOWERING COGNITIVE, PHYSICAL, SOCIAL AND COMMUNICATIVE SKILLS THROUGH VIRTUAL REALITY, ROBOTS, WEARABLE SYSTEMS AND BRAIN-COMPUTER INTERFACES, volume 145 of *Studies in Health Technology and Informatics*, pages 249–262. IOS PRESS, 2009. 3.4.3

[199] Maja J Matarić and Brian Scassellati. Socially assistive robotics. *Springer handbook of robotics*, pages 1973–1994, 2016. 3.1.3

[200] Alexander Mertens, Ulrich Reiser, Benedikt Brenken, Mathias Luedtke, Martin Haegele, Alexander Verl, Christopher Brandl, and Christopher Schlick. Assistive robots in eldercare and daily living: Automation of individual services for senior citizens. In S Jeschke, H Liu, and D Schilberg, editors, *INTELLIGENT ROBOTICS AND APPLICATIONS, PT I: ICIRA 2011*, volume 7101 of *Lecture Notes in Artificial Intelligence*, pages 542+. SPRINGER-VERLAG BERLIN, 2011. 3.4.1

[201] S. Meyer and Ch. Fricke. Robotic companions in stroke therapy: A user study on the efficacy of assistive robotics among 30 patients in neurological rehabilitation. In A Howard, K Suzuki, and L Zollo, editors, *2017 26TH IEEE INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 135–142. IEEE, 2017. 3.4.3

[202] Grégoire Milliez, Raphaël Lallement, Michelangelo Fiore, and Rachid Alami. Using human knowledge awareness to adapt collaborative plan generation, explanation and monitoring. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*, HRI '16, pages 43–50. IEEE Press, 2016. 3.4.1

[203] Matthias Minderer, Alexey Gritsenko, Austin Stone, Maxim Neumann, Dirk Weissenborn, Alexey Dosovitskiy, Aravindh Mahendran, Anurag Arnab, Mostafa Dehghani, Zhuoran Shen, Xiao Wang, Xiaohua Zhai, Thomas Kipf, and Neil Houlsby. Simple open-vocabulary object detection. In *Computer Vision – ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part X*, page 728–755, Berlin, Heidelberg, 2022. Springer-Verlag. ISBN 978-3-031-20079-3. doi: 10.1007/978-3-031-20080-9_42. URL https://doi.org/10.1007/978-3-031-20080-9_42. 6.3.2

[204] Jun Miura, Shin Kadekawa, Kota Chikaarashi, and Junichi Sugiyama. Human-robot collaborative remote object search. In E Menegatti, N Michael, K Berns, and H Yamaguchi, editors, *INTELLIGENT AUTONOMOUS SYSTEMS 13*, volume 302 of *Advances in Intelligent Systems and Computing*, pages 979–991. SPRINGER-VERLAG BERLIN, 2016. 3.4.1

[205] Brian Mok. Effects of proactivity and expressivity on collaboration with interactive robotic drawers. In *ELEVENTH ACM/IEEE INTERNATIONAL*

CONFERENCE ON HUMAN ROBOT INTERACTION (HRI'16), ACM IEEE International Conference on Human-Robot Interaction, pages 633–634. ASSOC COMPUTING MACHINERY, 2016. 3.4.1, 3.4.3, 3.5.3

[206] Natawut Monaikul, Bahareh Abbasi, Zhanibek Rysbek, Barbara Di Eugenio, and Milos Zefran. Role switching in task-oriented multimodal human-robot collaboration. In *2020 29TH IEEE INTERNATIONAL CONFERENCE ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 1150–1156. IEEE, 2020. 3.4.3

[207] Hee-Seung Moon, Woohyun Kim, Seungjae Han, and Jiwon Seo. Observation of human trajectory in response to haptic feedback from mobile robot. In *2018 18TH INTERNATIONAL CONFERENCE ON CONTROL, AUTOMATION AND SYSTEMS (ICCAS)*, International Conference on Control Automation and Systems, pages 1530–1534. IEEE, 2018. 3.4.3

[208] Marius Mosbach, Maksym Andriushchenko, and Dietrich Klakow. On the stability of fine-tuning {bert}: Misconceptions, explanations, and strong baselines. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=nzpLWnVAyah. 2.4.3

[209] Caio Mucchiani, Suneet Sharma, Megan Johnson, Justine Sefcik, Nicholas Vivio, Justin Huang, Pamela Cacchione, Michelle Johnson, Roshan Rai, Adrian Canoso, Tessa Lau, and Mark Yim. Evaluating older adults' interaction with a mobile assistive robot. In A Bicchi and A Okamura, editors, *2017 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS)*, IEEE International Conference on Intelligent Robots and Systems, pages 840–847. IEEE, 2017. 3.4.3

[210] Sarah Luisa Mueller, Stefan Schroeder, Sabina Jeschke, and Anja Richert. Design of a robotic workmate. In VG Duffy, editor, *DIGITAL HUMAN MODELING: APPLICATIONS IN HEALTH, SAFETY, ERGONOMICS, AND RISK MANAGEMENT: ERGONOMICS AND DESIGN*, volume 10286 of *Lecture Notes in Computer Science*, pages 447–456. SPRINGER INTERNATIONAL PUBLISHING AG, 2017. 3.4.1

[211] Luca Muratore, Arturo Laurenzi, and Nikos G. Tsagarakis. A self-modulated impedance multimodal interaction framework for human-robot collaboration. In A Howard, K Althoefer, F Arai, F Arrichiello, B Caputo, J Castellanos, K Hauser, V Isler, J Kim, H Liu, P Oh, V Santos, D Scaramuzza, A Ude, R Voyles, K Yamane, and A Okamura, editors, *2019 INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA)*, IEEE International Conference on Robotics and Automation ICRA, pages 4998–5004. IEEE, 2019. 3.4.3

[212] Saida Mussakhojayeva, Nazerke Kalidolda, and Anara Sandygulova. Adaptive

strategies for multi-party interactions with robots in public spaces. In A Kheddar, E Yoshida, SS Ge, K Suzuki, JJ Cabibihan, F Eyssel, and H He, editors, *SOCIAL ROBOTICS, ICSR 2017*, volume 10652 of *Lecture Notes in Artificial Intelligence*, pages 749–758. SPRINGER INTERNATIONAL PUBLISHING AG, 2017. 3.4.3

[213] Rajkumar Muthusamy, Jose Miguel Indave, Praveen Kumar Muthusamy, Eman Fayez Hasan, Yahya Zweiri, Ville Kyrki, and Dongming Gan. Investigation and design of robotic assistance control system for cooperative manipulation. In *2019 9TH IEEE ANNUAL INTERNATIONAL CONFERENCE ON CYBER TECHNOLOGY IN AUTOMATION, CONTROL, AND INTELLIGENT SYSTEMS (IEEE-CYBER 2019)*, IEEE Annual International Conference on Cyber Technology in Automation Control and Intelligent Systems, pages 889–895. IEEE, 2019. 3.4.2

[214] Mahdi Nabipour and S. Ali A. Moosavian. Dynamics modeling and performance analysis of robowalk. In *2018 6TH RSI INTERNATIONAL CONFERENCE ON ROBOTICS AND MECHATRONICS (ICROM 2018)*, RSI International Conference on Robotics and Mechatronics ICRoM, pages 445–450. IEEE, 2018. 3.4.2

[215] Vinod Nair and Geoffrey E. Hinton. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ICML'10, pages 807–814, USA, 2010. Omnipress. ISBN 978-1-60558-907-7. URL `http://dl.acm.org/citation.cfm?id=3104322.3104425`. 5.2.2

[216] Koki Nakabayashi, Yukiko Iwasaki, Shota Takahashi, and Hiroyasu Iwata. Experimental evaluation of cooperativeness and collision safety of a wearable robot arm. In JJ Cabibihan, F Mastrogiovanni, AK Pandey, S Rossi, and M Staffa, editors, *2018 27TH IEEE INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (IEEE RO-MAN 2018)*, IEEE RO-MAN, pages 1026–1031. IEEE, 2018. 3.4.3

[217] Soroush Nasiriany, Fei Xia, Wenhao Yu, Ted Xiao, Jacky Liang, Ishita Dasgupta, Annie Xie, Danny Driess, Ayzaan Wahid, Zhuo Xu, et al. Pivot: Iterative visual prompting elicits actionable knowledge for vlms. *arXiv preprint arXiv:2402.07872*, 2024. 2.4.2, 6.1

[218] Heramb Nemlekar, Dharini Dutia, and Zhi Li. Object transfer point estimation for fluent human-robot handovers. In A Howard, K Althoefer, F Arai, F Arrichiello, B Caputo, J Castellanos, K Hauser, V Isler, J Kim, H Liu, P Oh, V Santos, D Scaramuzza, A Ude, R Voyles, K Yamane, and A Okamura, editors, *2019 INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA)*, IEEE International Conference on Robotics and Automation

ICRA, pages 2627–2633. IEEE, 2019. 3.4.1

[219] Benjamin Newman, Kevin Carlberg, and Ruta Desai. Optimal assistance for object-rearrangement tasks in augmented reality, 2020. URL https://arxiv.org/abs/2010.07358. 2.2.1, 2.4.3, 3.4.3

[220] Benjamin A. Newman, Abhijat Biswas, Sarthak Ahuja, Siddharth Girdhar, Kris K. Kitani, and Henny Admoni. Examining the effects of anticipatory robot assistance on human decision making. In Alan R. Wagner, David Feil-Seifer, Kerstin S. Haring, Silvia Rossi, Thomas Williams, Hongsheng He, and Shuzhi Sam Ge, editors, *Social Robotics*, pages 590–603. Springer International Publishing, 2020. 2.2.1, 2.3.1, 3.4.1, 3.5.4

[221] Benjamin A. Newman, Reuben M. Aronson, Kris Kitani, and Henny Admoni. Helping people through space and time: Assistance as a perspective on human-robot interaction. *Frontiers in Robotics and AI*, 8, 2022. ISSN 2296-9144. doi: 10.3389/frobt.2021.720319. URL https://www.frontiersin.org/articles/10.3389/frobt.2021.720319. 1.1.2, 6, 7

[222] Benjamin A. Newman, Reuben M. Aronson, Siddhartha S. Srinivasa, Kris Kitani, and Henny Admoni. Harmonic: A multimodal dataset of assistive human–robot collaboration. *The International Journal of Robotics Research*, 41 (1):3–11, 2022. doi: 10.1177/02783649211050677. URL https://doi.org/10.1177/02783649211050677. 1.1.3, 2.2.1, 2.3.1, 3.4.3, 3.5.1, 5, 5.1.2

[223] Benjamin A. Newman, Christopher Jason Paxton, Kris Kitani, and Henny Admoni. Towards online adaptation for autonomous household assistants. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '23, page 506–510, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450399708. doi: 10.1145/3568294.3580136. URL https://doi.org/10.1145/3568294.3580136. 2.3.1, 6

[224] Benjamin A. Newman, Pranay Gupta, Yonatan Bisk, Kris Kitani, Henny Admoni, and Chris Paxton. Leveraging vision and language models for zero-shot, personalization of household multi-object rearrangement tasks. In *Proceedings of Human-Large Language Model Interaction Workshop HRI'24*. ACM, 2024. 2.4.1

[225] Benjamin A. Newman, Pranay Gupta, Kris Kitani, Yonatan Bisk, Henny Admoni, and Chris Paxton. Degustabot: Zero-shot visual preference estimation for personalized multi-object rearrangement, 2024. URL https://arxiv.org/abs/2407.08876. 1.1.5

[226] Benjamin A. Newman, Christopher Jason Paxton, Kris Kitani, and Henny Admoni. Bootstrapping linear models for fast online adaptation in human-agent collaboration. In *Proc. of the 23rd International Conference on Autonomous*

Agents and Multiagent Systems), Auckland, New Zealand, May 6 – 10, 2024. IFAAMAS, 2024. 1.1.6, 2.2.1

[227] Hai Nguyen and Charles C. Kemp. Bio-inspired assistive robotics: Service dogs as a model for human-robot interaction and mobile manipulation. In *2008 2ND IEEE RAS & EMBS INTERNATIONAL CONFERENCE ON BIOMEDICAL ROBOTICS AND BIOMECHATRONICS (BIOROB 2008), VOLS 1 AND 2*, Proceedings of the IEEE RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics, pages 542–549. IEEE, 2008. 3.4.1

[228] Phuong D. H. Nguyen, Fabrizio Bottarel, Ugo Pattacini, Matej Hoffmann, Lorenzo Natale, and Giorgio Metta. Merging physical and social interaction for effective human-robot collaboration. In T Asfour, editor, *2018 IEEE-RAS 18TH INTERNATIONAL CONFERENCE ON HUMANOID ROBOTS (HUMANOIDS)*, IEEE-RAS International Conference on Humanoid Robots, pages 710–717. IEEE, 2018. 3.4.3

[229] Guangtao Nie, Zhi Zheng, Jazette Johnson, Amy R. Swanson, Amy S. Weitlauf, Zachary E. Warren, and Nilanjan Sarkar. Predicting response to joint attention performance in human-human interaction based on human-robot interaction for young children with autism spectrum disorder. In JJ Cabibihan, F Mastrogiovanni, AK Pandey, S Rossi, and M Staffa, editors, *2018 27TH IEEE INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (IEEE RO-MAN 2018)*, IEEE RO-MAN, pages 1069–1074. IEEE, 2018. 3.4.3

[230] Stefanos Nikolaidis, Anton Kuznetsov, David Hsu, and Siddharta Srinivasa. Formalizing human-robot mutual adaptation: A bounded memory model. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*, HRI '16, pages 75–82. IEEE Press, 2016. 3.4.1, 3.5.3

[231] Tomoyuki Noda, Jun-ichiro Furukawa, Tatsuya Teramae, Sang-Ho Hyon, and Jun Morimoto. An electromyogram based force control coordinated in assistive interaction. In *2013 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA)*, IEEE International Conference on Robotics and Automation ICRA, pages 2657–2662. IEEE, 2013. 3.4.3

[232] Hiroki Ono, Kaito Koike, Takeshi Morita, and Takahira Yamaguchi. Ontologies-based pupil robot interaction with group discussion. In IJ Rudas, C Janos, C Toro, J Botzheim, RJ Howlett, and LC Jain, editors, *KNOWLEDGE-BASED AND INTELLIGENT INFORMATION & ENGINEERING SYSTEMS (KES 2019)*, volume 159 of *Procedia Computer Science*, pages 2071–2080. ELSEVIER, 2019. 3.4.3

[233] Xanthi S. Papageorgiou, Georgia Chalvatzaki, Athanasios C. Dometios, and Costas S. Tzafestas. Human-centered service robotic systems for assisted

living. In NA Aspragathos, PN Koustoumpardis, and VC Moulianitis, editors, *ADVANCES IN SERVICE AND INDUSTRIAL ROBOTICS, RAAD 2018*, volume 67 of *Mechanisms and Machine Science*, pages 132–140. SPRINGER INTERNATIONAL PUBLISHING AG, 2019. 3.4.2

[234] Christopher Parlitz, Martin Hägele, Peter Klein, Jan Seifert, and Kerstin Dautenhahn. Care-o-bot 3 — rationale for human-robot interaction design. In *39th International Symposium on Robotics, ISR 2008*, 10 2008. 3.4.1

[235] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in PyTorch. In *NIPS Autodiff Workshop*, 2017. 5.2.2

[236] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019. URL http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf. 7.2.1

[237] L. Peternel, N. Tsagarakis, and A. Ajoudani. Towards multi-modal intention interfaces for human-robot co-manipulation. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2663–2669, Oct 2016. doi: 10.1109/IROS.2016.7759414. 2.1.3

[238] L. Peternel, N. Tsagarakis, and A. Ajoudani. A human–robot co-manipulation approach based on human sensorimotor information. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(7):811–822, July 2017. doi: 10.1109/TNSRE.2017.2694553. 2.1.3

[239] Rifca Peters, Joost Broekens, Kangqi Li, and Mark A. Neerincx. Robots expressing dominance: Effects of behaviours and modulation. In *2019 8TH INTERNATIONAL CONFERENCE ON AFFECTIVE COMPUTING AND INTELLIGENT INTERACTION (ACII)*, International Conference on Affective Computing and Intelligent Interaction. IEEE, 2019. 3.4.3

[240] T. X. N. Pham, K. Hayashi, C. Becker-Asano, S. Lacher, and I. Mizuuchi. Evaluating the usability and users' acceptance of a kitchen assistant robot in household environment. In A Howard, K Suzuki, and L Zollo, editors, *2017 26TH IEEE INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 987–992. IEEE, 2017. 3.4.3

[241] H. Pirsiavash and D. Ramanan. Detecting activities of daily living in first-person camera views. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2847–2854, June 2012. doi: 10.1109/CVPR.2012.6248010. 2.1.2, 2.1.2

[242] Alex Polishuk and Igor Verner. An elementary science class with a robot teacher. In W Lepuschitz, M Merdan, G Koppensteiner, R Balogh, and D Obdrzalek, editors, *ROBOTICS IN EDUCATION: LATEST RESULTS AND DEVELOPMENTS*, volume 630 of *Advances in Intelligent Systems and Computing*, pages 263–273. SPRINGER INTERNATIONAL PUBLISHING AG, 2018. 3.3.3, 3.4.3

[243] Juergen Pripfl, Tobias Koertner, Daliah Batko-Klein, Denise Hebesberger, Markus Weninger, Christoph Gisinger, Susanne Frennert, Hakan Eftring, Margarita Antona, Ilia Adami, Astrid Weiss, Markus Bajones, and Markus Vincze. Results of a real world trial with a mobile social service robot for older adults. In *ELEVENTH ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN ROBOT INTERACTION (HRI'16)*, ACM IEEE International Conference on Human-Robot Interaction, pages 497–498. ASSOC COMPUTING MACHINERY, 2016. 3.4.1

[244] Prolific. Prolific, 2014 Online. URL https://www.prolific.co. 7.4

[245] Reid Pryzant, Dan Iter, Jerry Li, Yin Tat Lee, Chenguang Zhu, and Michael Zeng. Automatic prompt optimization with" gradient descent" and beam search. *arXiv preprint arXiv:2305.03495*, 2023. 2.4.2

[246] Xavier Puig, Kevin Ra, Marko Boben, Jiaman Li, Tingwu Wang, Sanja Fidler, and Antonio Torralba. Virtualhome: Simulating household activities via programs. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 8494–8502, 2018. 2.2.1

[247] Pupil Labs, Inc. Pupil labs - pupil, 2017. https://pupil-labs.com/pupil/. 4.1.4

[248] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017. URL https://www.R-project.org/. 3.2

[249] Mattia Racca, Ville Kyrki, and Maya Cakmak. Interactive tuning of robot program parameters via expected divergence maximization. In *PROCEEDINGS OF THE 2020 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI '20)*, ACM IEEE International Conference on Human-Robot Interaction, pages 629–638. ASSOC COMPUTING MACHINERY, 2020. 3.4.1

[250] S. M. Mizanoor Rahman. Human features-based variable admittance control for improving hri and performance in power-assisted heavy object manipulation.

In *2019 12TH INTERNATIONAL CONFERENCE ON HUMAN SYSTEM INTERACTION (HSI)*, Conference on Human System Interaction, pages 87–92. IEEE, 2019. 3.4.1

[251] S. M. Mizanoor Rahman. Bioinspired dynamic affect-based motion control of a humanoid robot to collaborate with human in manufacturing. In *2019 12TH INTERNATIONAL CONFERENCE ON HUMAN SYSTEM INTERACTION (HSI)*, Conference on Human System Interaction, pages 76–81. IEEE, 2019. 3.4.3, 3.5.3

[252] S. M. Mizanoor Rahman. Grasp rehabilitation of stroke patients through object manipulation with an intelligent power assist robotic system. In K Arai, R Bhatia, and S Kapoor, editors, *PROCEEDINGS OF THE FUTURE TECHNOLOGIES CONFERENCE (FTC) 2019, VOL 2*, volume 1070 of *Advances in Intelligent Systems and Computing*, pages 244–259. SPRINGER INTERNATIONAL PUBLISHING AG, 2020. 3.4.1

[253] Aditi Ramachandran, Alexandru Litoiu, and Brian Scassellati. Shaping productive help-seeking behavior during robot-child tutoring interactions. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*, HRI '16, pages 247–254. IEEE Press, 2016. 3.4.3

[254] Shreyas Sundara Raman, Vanya Cohen, Eric Rosen, Ifrah Idrees, David Paulius, and Stefanie Tellex. Planning with large language models via corrective re-prompting. In *NeurIPS 2022 Foundation Models for Decision Making Workshop*, 2022. 2.4.1

[255] Robin Rasch, Dennis Sprute, Aljoscha Pörtner, Sven Battermann, and Matthias König. Tidy up my room: Multi-agent cooperation for service tasks in smart environments. *Journal of Ambient Intelligence and Smart Environments*, 11(3): 261–275, 2019. 2.2.1

[256] Siddharth Reddy, Anca D. Dragan, and Sergey Levine. Where do you think you're going?: Inferring beliefs about dynamics from behavior. *CoRR*, abs/1805.08010, 2018. URL `http://arxiv.org/abs/1805.08010`. 2.1.1

[257] Siddharth Reddy, Sergey Levine, and Anca D. Dragan. Shared autonomy via deep reinforcement learning. *CoRR*, abs/1802.01744, 2018. URL `http://arxiv.org/abs/1802.01744`. 2.1.1

[258] Mauricio Reyes, Ivan Meza, and Luis A. Pineda. The positive effect of negative feedback in hri using a facial expression robot. In JTKV Koh, BJ Dunstan, D SilveraTawil, and M Velonaki, editors, *CULTURAL ROBOTICS, CR 2015*, volume 9549 of *Lecture Notes in Computer Science*, pages 44–54. SPRINGER INTERNATIONAL PUBLISHING AG, 2016. 3.4.3, 3.5.3

[259] Jimin Rhim, Anthony Cheung, David Pham, Subin Bae, Zhitian Zhang, Trista

Townsend, and Angelica Lim. Investigating positive psychology principles in affective robotics. In *2019 8TH INTERNATIONAL CONFERENCE ON AFFECTIVE COMPUTING AND INTELLIGENT INTERACTION (ACII)*, International Conference on Affective Computing and Intelligent Interaction. IEEE, 2019. 3.4.3

[260] Iain E. Richardson. *The H.264 Advanced Video Compression Standard*. Wiley Publishing, 2nd edition, 2010. ISBN 0470516925. 4.3, 4.3.2, 4.3.3

[261] Nina Riether, Frank Hegel, Britta Wrede, and Gernot Horstmann. Social facilitation with social robots? In *HRI'12: PROCEEDINGS OF THE SEVENTH ANNUAL ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION*, ACM IEEE International Conference on Human-Robot Interaction, pages 41–47. ASSOC COMPUTING MACHINERY, 2012. 3.4.3

[262] Paul Robinette, Wenchen Li, Robert Allen, Ayanna M. Howard, and Alan R. Wagner. Overtrust of robots in emergency evacuation scenarios. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*, HRI '16, pages 101–108. IEEE Press, 2016. 3.4.3

[263] Rinat Rosenberg-Kima, Yaacov Koren, Maya Yachini, and Goren Gordon. Human-robot-collaboration (hrc): Social robots as teaching assistants for training activities in small groups. In *HRI '19: 2019 14TH ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION*, ACM IEEE International Conference on Human-Robot Interaction, pages 522–523. IEEE, 2019. 3.3.3, 3.4.3

[264] Martin J. Rosenstrauch, Tessa J. Pannen, and Joerg Krueger. Human robot collaboration - using kinect v2 for iso/ts 15066 speed and separation monitoring. In S Wang and Z Wu, editors, *7TH CIRP CONFERENCE ON ASSEMBLY TECHNOLOGIES AND SYSTEMS (CATS 2018)*, volume 76 of *Procedia CIRP*, pages 183–186. ELSEVIER SCIENCE BV, 2018. 3.4.3

[265] Alessandra Rossi, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L. Walters. How social robots influence people's trust in critical situations. In *2020 29TH IEEE INTERNATIONAL CONFERENCE ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 1020–1025. IEEE, 2020. 3.4.3

[266] Selma Sabanovic, Casey C. Bennett, Wan-Ling Chang, and Lesa Huber. Paro robot affects diverse interaction modalities in group sensory therapy for older adults with dementia. In *2013 IEEE 13TH INTERNATIONAL CONFERENCE ON REHABILITATION ROBOTICS (ICORR)*, International Conference on Rehabilitation Robotics ICORR. IEEE, 2013. 3.4.3

[267] Dorsa Sadigh, Anca D. Dragan, Shankar Sastry, and Sanjit A. Seshia. Active

preference-based learning of reward functions. In Nancy M. Amato, Siddhartha S. Srinivasa, Nora Ayanian, and Scott Kuindersma, editors, *Robotics: Science and Systems XIII, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, July 12-16, 2017*, 2017. doi: 10.15607/RSS.2017.XIII.053. URL http://www.roboticsproceedings.org/rss13/p53.html. 1

[268] Esther Salichs, Enrique Fernandez-Rodicio, Jose Carlos Castillo, Alvaro Castro-Gonzalez, Maria Malfaz, and Miguel Angel Salichs. A social robot assisting in cognitive stimulation therapy. In Y Demazeau, B An, J Bajo, and A FernandezCaballero, editors, *ADVANCES IN PRACTICAL APPLICATIONS OF AGENTS, MULTI-AGENT SYSTEMS, AND COMPLEXITY: THE PAAMS COLLECTION*, volume 10978 of *Lecture Notes in Artificial Intelligence*, pages 344–347. SPRINGER INTERNATIONAL PUBLISHING AG, 2018. 3.4.3

[269] Anara Sandygulova, Wafa Johal, Zhanel Zhexenova, Bolat Tleubayev, Aida Zhanatkyzy, Aizada Turarova, Zhansaule Telisheva, Anna CohenMiller, Thibault Asselborn, and Pierre Dillenbourg. Cowriting kazakh: Learning a new script with a robot. In *PROCEEDINGS OF THE 2020 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI '20)*, ACM IEEE International Conference on Human-Robot Interaction, pages 113–120. ASSOC COMPUTING MACHINERY, 2020. 3.4.3

[270] Thiago Santini, Wolfgang Fuhl, Thomas Kübler, and Enkelejda Kasneci. Bayesian identification of fixations, saccades, and smooth pursuits. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications*, ETRA '16, pages 163–170, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-4125-7. doi: 10.1145/2857491.2857512. URL http://doi.acm.org/10.1145/2857491.2857512. 5.2.1

[271] Gabriel Sarch, Zhaoyuan Fang, Adam W Harley, Paul Schydlo, Michael J Tarr, Saurabh Gupta, and Katerina Fragkiadaki. Tidee: Tidying up novel rooms using visuo-semantic commonsense priors. In *European conference on computer vision*, pages 480–496. Springer, 2022. 2.2.1

[272] Celal Savur, Shitij Kumar, and Ferat Sahin. A framework for monitoring human physiological response during human robot collaborative task. In *2019 IEEE INTERNATIONAL CONFERENCE ON SYSTEMS, MAN AND CYBERNETICS (SMC)*, IEEE International Conference on Systems Man and Cybernetics Conference Proceedings, pages 385–390. IEEE, 2019. 3.4.1

[273] Manolis Savva, Abhishek Kadian, Oleksandr Maksymets, Yili Zhao, Erik Wijmans, Bhavana Jain, Julian Straub, Jia Liu, Vladlen Koltun, Jitendra Malik, et al. Habitat: A platform for embodied ai research. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9339–9347, 2019. 7.4

[274] Jonas Schmidtler and Klaus Bengler. Size-weight illusion in human-robot collaboration. In *2016 25TH IEEE INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 874–879. IEEE, 2016. 3.4.1

[275] Jonas Schmidtler, Moritz Koerber, and Klaus Bengler. A trouble shared is a trouble halved - usability measures for human-robot collaboration. In *2016 IEEE INTERNATIONAL CONFERENCE ON SYSTEMS, MAN, AND CYBERNETICS (SMC)*, IEEE International Conference on Systems Man and Cybernetics Conference Proceedings, pages 217–222. IEEE, 2016. 3.4.1

[276] Sebastian Schneider and Franz Kummert. Motivational effects of acknowledging feedback from a socially assistive robot. In A Agah, JJ Cabibihan, AM Howard, MA Salichs, and H He, editors, *SOCIAL ROBOTICS, (ICSR 2016)*, volume 9979 of *Lecture Notes in Artificial Intelligence*, pages 870–879. SPRINGER-VERLAG BERLIN, 2016. 3.3.2, 3.4.3

[277] Sarah Strohkorb Sebo, Margaret Traeger, Malte Jung, and Brian Scassellati. The ripple effects of vulnerability: The effects of a robot's vulnerable behavior on trust in human-robot teams. In *HRI '18: PROCEEDINGS OF THE 2018 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION*, ACM IEEE International Conference on Human-Robot Interaction, pages 178–186. ASSOC COMPUTING MACHINERY, 2018. 3.4.3

[278] Sarah Strohkorb Sebo, Ling Liang Dong, Nicholas Chang, and Brian Scassellati. Strategies for the inclusion of human members within human-robot teams. In *PROCEEDINGS OF THE 2020 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI '20)*, ACM IEEE International Conference on Human-Robot Interaction, pages 309–317. ASSOC COMPUTING MACHINERY, 2020. 3.4.3

[279] Dhruv Shah, Michael Robert Equi, Błażej Osiński, Fei Xia, Brian Ichter, and Sergey Levine. Navigation with large language models: Semantic guesswork as a heuristic for planning. In *Conference on Robot Learning*, pages 2683–2699. PMLR, 2023. 2.4.1

[280] Dhruv Shah, Błażej Osiński, Sergey Levine, et al. Lm-nav: Robotic navigation with large pre-trained models of language, vision, and action. In *Conference on Robot Learning*, pages 492–504. PMLR, 2023. 2.4.1

[281] Kamran Shamaei, Yuhang Che, Adithyavairavan Murali, Siddarth Sen, Sachin Patil, Ken Goldberg, and Allison M. Okamura. A paced shared-control teleoperated architecture for supervised automation of multilateral surgical tasks. In *2015 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS)*, IEEE International Conference on Intelligent Robots and Systems, pages 1434–1439. IEEE, 2015. 3.4.1

[282] Syamimi Shamsuddin, Winal Zikril Zulkifli, Lim Thiam Hwee, and Hanafiah Yussof. Animal robot as augmentative strategy to elevate mood: A preliminary study for post-stroke depression. In A Ronzhin, G Rigoll, and R Meshcheryakov, editors, *INTERACTIVE COLLABORATIVE ROBOTICS (ICR 2017)*, volume 10459 of *Lecture Notes in Artificial Intelligence*, pages 209–218. SPRINGER INTERNATIONAL PUBLISHING AG, 2017. 3.4.3

[283] Mingyang Shao, Silas Franco Dos Reis Alves, Omar Ismail, Xinyi Zhang, Goldie Nejat, and Beno Benhabib. You are doing great! only one rep left: An affect-aware social robot for exercising. In *2019 IEEE INTERNATIONAL CONFERENCE ON SYSTEMS, MAN AND CYBERNETICS (SMC)*, IEEE International Conference on Systems Man and Cybernetics Conference Proceedings, pages 3811–3817. IEEE, 2019. 3.4.3

[284] Ali Sharifara, Ashwin Ramesh Babu, Akilesh Rajavenkatanarayanan, Christopher Collander, and Fillia Makedon. A robot-based cognitive assessment model based on visual working memory and attention level. In M Antona and C Stephanidis, editors, *UNIVERSAL ACCESS IN HUMAN-COMPUTER INTERACTION: METHODS, TECHNOLOGIES, AND USERS, UAHCI 2018, PT I*, volume 10907 of *Lecture Notes in Computer Science*, pages 583–597. SPRINGER INTERNATIONAL PUBLISHING AG, 2018. 3.4.3

[285] Amir Mehdi Shayan, Alireza Sarmadi, Armin Pirastehzad, Hadi Moradi, and Pegah Soleiman. Roboparrot 2.0: A multi-purpose social robot. In *2016 4TH RSI INTERNATIONAL CONFERENCE ON ROBOTICS AND MECHATRONICS (ICROM)*, RSI International Conference on Robotics and Mechatronics ICRoM, pages 422–427. IEEE, 2016. 3.3.2, 3.4.3

[286] Samira Sheikhi and Jean-Marc Odobez. Recognizing the visual focus of attention for human robot interaction. In *Proceedings of the Third International Conference on Human Behavior Understanding*, HBU'12, pages 99–112, Berlin, Heidelberg, 2012. Springer-Verlag. ISBN 978-3-642-34013-0. doi: 10.1007/978-3-642-34014-7_9. URL http://dx.doi.org/10.1007/978-3-642-34014-7_9. 2.1.2

[287] T Shibata, T Mitsui, K Wada, A Touda, T Kumasaka, K Tagami, and K Tanie. Mental commit robot and its application to therapy of children. In *2001 IEEE/ASME INTERNATIONAL CONFERENCE ON ADVANCED INTELLIGENT MECHATRONICS PROCEEDINGS, VOLS I AND II*, IEEE ASME International Conference on Advanced Intelligent Mechatronics, pages 1053–1058. IEEE, 2001. 3.4.3

[288] Mohit Shridhar, Lucas Manuelli, and Dieter Fox. Cliport: What and where pathways for robotic manipulation. In *Conference on Robot Learning*, pages 894–906. PMLR, 2022. 2.4.1

[289] Beibei Shu, Gabor Sziebig, and Sakari Pieska. Human-robot collaboration: Task sharing through virtual reality. In *IECON 2018 - 44TH ANNUAL CONFERENCE OF THE IEEE INDUSTRIAL ELECTRONICS SOCIETY*, IEEE Industrial Electronics Society, pages 6040–6044. IEEE, 2018. 3.4.3, 3.5.2

[290] Tianmin Shu, M. S. Ryoo, and Song-Chun Zhu. Learning social affordance for human-robot interaction. In *International Joint Conference on Artificial Intelligence (IJCAI)*, 2016. 2.1.2, 2.1.2

[291] Sergio D. Sierra, Mario F. Jimenez, Marcela C. Munera, Anselmo Frizera-Neto, and Carlos A. Cifuentes. Remote-operated multimodal interface for therapists during walker-assisted gait rehabilitation: A preliminary assessment. In *HRI '19: 2019 14TH ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION*, ACM IEEE International Conference on Human-Robot Interaction, pages 528–529. IEEE, 2019. 3.4.2

[292] G. A. Sigurdsson, A. Gupta, C. Schmid, A. Farhadi, and K. Alahari. Actor and observer: Joint modeling of first and third-person videos. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 7396–7404, Los Alamitos, CA, USA, jun 2018. IEEE Computer Society. doi: 10.1109/CVPR.2018.00772. URL https://doi.ieeecomputersociety.org/10.1109/CVPR.2018.00772. 2.1.2, 2.1.2

[293] Gunnar A. Sigurdsson, Abhinav Gupta, Cordelia Schmid, Ali Farhadi, and Karteek Alahari. Charades-ego: A large-scale dataset of paired third and first person videos. *CoRR*, abs/1804.09626, 2018. URL http://arxiv.org/abs/1804.09626. 2.1.2, 2.1.2

[294] Tom Silver, Varun Hariprasad, Reece S Shuttleworth, Nishanth Kumar, Tomás Lozano-Pérez, and Leslie Pack Kaelbling. Pddl planning with pretrained large language models. In *NeurIPS 2022 foundation models for decision making workshop*, 2022. 2.4.1

[295] Tomas Simon, Hanbyul Joo, Iain Matthews, and Yaser Sheikh. Hand keypoint detection in single images using multiview bootstrapping. In *CVPR*, 2017. 4.3

[296] Ishika Singh, Valts Blukis, Arsalan Mousavian, Ankit Goyal, Danfei Xu, Jonathan Tremblay, Dieter Fox, Jesse Thomason, and Animesh Garg. Progprompt: Generating situated robot task plans using large language models. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 11523–11530. IEEE, 2023. 2.4.1

[297] H. P. Chapa Sirithunge, M. A. Viraj J. Muthugala, A. G. Buddhika P. Jayasekara, and D. P. Chandima. A wizard of oz study of human interest towards robot initiated human-robot interaction. In JJ Cabibihan, F Mastrogiovanni, AK Pandey, S Rossi, and M Staffa, editors, *2018 27TH IEEE*

INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (IEEE RO-MAN 2018), IEEE RO-MAN, pages 515–521. IEEE, 2018. 3.4.3

[298] David Sirkin, Brian Mok, Stephen Yang, and Wendy Ju. Mechanical ottoman: How robotic furniture offers and withdraws support. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 11–18, 2015. 3.4.1

[299] Edward J. Sondik. The optimal control of partially observable markov processes over the infinite horizon: Discounted costs. *Operations Research*, 26(2):282–304, 1978. ISSN 0030364X, 15265463. URL http://www.jstor.org/stable/169635. 4.1.3

[300] Chan Hee Song, Jiaman Wu, Clayton Washington, Brian M Sadler, Wei-Lun Chao, and Yu Su. Llm-planner: Few-shot grounded planning for embodied agents with large language models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2998–3009, 2023. 2.4.1

[301] Siddhartha S Srinivasa, Dave Ferguson, Casey J Helfrich, Dmitry Berenson, Alvaro Collet, Rosen Diankov, Garratt Gallagher, Geoffrey Hollinger, James Kuffner, and Michael Vande Weghe. Herb: a home exploring robotic butler. *Autonomous Robots*, 28(1):5–20, 2010. 3.4.1

[302] Siddhartha S Srinivasa, Dmitry Berenson, Maya Cakmak, Alvaro Collet, Mehmet R Dogar, Anca D Dragan, Ross A Knepper, Tim Niemueller, Kyle Strabala, Mike Vande Weghe, et al. Herb 2.0: Lessons learned from developing a mobile manipulator for the home. *Proceedings of the IEEE*, 100(8):2410–2428, 2012. 3.4.1

[303] Kalin Stefanov and Jonas Beskow. A multi-party multi-modal dataset for focus of visual attention in human-human and human-robot interaction. In *LREC*, 2016. 2.1.2

[304] Bonnie Steinbock, John D. Arras, and Alex John London. *Ethical Issues in Modern Medicine*. McGraw-Hill, Boston, 2003. 8.2.2

[305] Stereolabs Inc. Stereolabs, 2018. https://www.stereolabs.com/. 4.1.4

[306] Brett Stoll, Samantha Reig, Lucy He, Ian Kaplan, Malte F. Jung, and Susan R. Fussell. "wait, can you move the robot?": Examining telepresence robot use in collaborative teams. In *HRI '18: PROCEEDINGS OF THE 2018 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION*, ACM IEEE International Conference on Human-Robot Interaction, pages 14–22. ASSOC COMPUTING MACHINERY, 2018. 3.4.1

[307] Sarah Strohkorb, Ethan Fukuto, Natalie Warren, Charles Taylor, Bobby Berry, and Brian Seassellati. Improving human-human collaboration between children

with a social robot. In *2016 25TH IEEE INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 551–556. IEEE, 2016. 3.4.3

[308] DJ Strouse, Kevin McKee, Matt Botvinick, Edward Hughes, and Richard Everett. Collaborating with humans without human data. *Advances in Neural Information Processing Systems*, 34:14502–14515, 2021. 2.4.3, 7

[309] Hang Su, Juan Sandoval, Mohatashem Makhdoomi, Giancarlo Ferrigno, and Elena De Momi. Safety-enhanced human-robot interaction control of redundant robot for teleoperated minimally invasive surgery. In *2018 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA)*, IEEE International Conference on Robotics and Automation ICRA, pages 6611–6616. IEEE COMPUTER SOC, 2018. 3.4.1

[310] Petr Svarny, Michael Tesar, Jan Kristof Behrens, and Motel Hoffmann. Safe physical hri: Toward a unified treatment of speed and separation monitoring together with power and force limiting. In *2019 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS)*, IEEE International Conference on Intelligent Robots and Systems, pages 7580–7587. IEEE, 2019. 3.4.3

[311] Andrew Szot, Alexander Clegg, Eric Undersander, Erik Wijmans, Yili Zhao, John M Turner, Noah D Maestre, Mustafa Mukadam, Devendra Singh Chaplot, Oleksandr Maksymets, Aaron Gokaslan, Vladimír Vondruš, Sameer Dharur, Franziska Meier, Wojciech Galuba, Angel X Chang, Zsolt Kira, Vladlen Koltun, Jitendra Malik, Manolis Savva, and Dhruv Batra. Habitat 2.0: Training home assistants to rearrange their habitat. In A. Beygelzimer, Y. Dauphin, P. Liang, and J. Wortman Vaughan, editors, *Advances in Neural Information Processing Systems*, 2021. URL `https://openreview.net/forum?id=DPHsCQ8OpA`. 2.2.1, 2.3.2, 7.4

[312] Aaquib Tabrez, Shivendra Agrawal, and Bradley Hayes. Explanation-based reward coaching to improve human performance via reinforcement learning. In *HRI '19: 2019 14TH ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION*, ACM IEEE International Conference on Human-Robot Interaction, pages 249–257. IEEE, 2019. 3.4.3

[313] A. R. Taheri, M. Alemi, A. Meghdari, H. R. PourEtemad, and N. Mahboob Basiri. Social robots as assistants for autism therapy in iran: Research in progress. In *2014 SECOND RSI/ISM INTERNATIONAL CONFERENCE ON ROBOTICS AND MECHATRONICS (ICROM)*, RSI International Conference on Robotics and Mechatronics ICRoM, pages 760–766. IEEE, 2014. 3.4.3

[314] Akira Taniguchi, Shota Isobe, Lotfi El Hafi, Yoshinobu Hagiwara, and Tadahiro Taniguchi. Autonomous planning based on spatial concepts to tidy up home

environments with service robots. *Advanced Robotics*, 35(8):471–489, 2021. 2.2.1

[315] Adriana Tapus, Cristian Tapus, and Maja J. Mataric. Hands-off therapist robot behavior adaptation to user personality for post-stroke rehabilitation therapy. In *PROCEEDINGS OF THE 2007 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, VOLS 1-10*, IEEE International Conference on Robotics and Automation ICRA, pages 1547+. IEEE, 2007. 3.3.2, 3.4.3

[316] Stefanie Tellex, Thomas Kollar, Steven Dickerson, Matthew R. Walter, Ashish G. Banerjee, Seth Teller, and Nicholas Roy. Understanding natural language commands for robotic navigation and mobile manipulation. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, pages 1507–1514, San Francisco, CA, August 2011. 2.1.1

[317] Yunus Terzioglu, Bilge Mutlu, and Erol Sahin. Designing social cues for collaborative robots: The role of gaze and breathing in human-robot collaboration. In *PROCEEDINGS OF THE 2020 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI '20)*, ACM IEEE International Conference on Human-Robot Interaction, pages 343–357. ASSOC COMPUTING MACHINERY, 2020. 3.4.3, 3.5.3

[318] Thalmic Labs, Inc. Myo Gesture Control Armband, 2018. https://www.myo.com/. 4.1.4

[319] John Till, Caroline E. Bryson, Scotty Chung, Andrew Orekhov, and D. Caleb Rucker. Efficient computation of multiple coupled cosserat rod models for real-time simulation and control of parallel continuum manipulators. In *2015 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA)*, IEEE International Conference on Robotics and Automation ICRA, pages 5067–5074. IEEE COMPUTER SOC, 2015. 3.4.3

[320] Inc. Tobii. Tobii pro website, 2017. 2.1.2

[321] Cristen Torrey, Susan R. Fussell, and Sara Kiesler. How a robot should give advice. In H Kuzuoka, V Evers, M Imai, and J Forlizzi, editors, *PROCEEDINGS OF THE 8TH ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI 2013)*, ACMIEEE International Conference on Human-Robot Interaction, pages 275–282. IEEE, 2013. 3.4.3

[322] Pinar Uluer, Hatice Kose, Bulent Koray Oz, Turgut Can Aydinalev, and Duygun Erol Barkana. Towards an affective robot companion for audiology rehabilitation: How does pepper feel today? In *2020 29TH IEEE INTERNATIONAL CONFERENCE ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 567–572. IEEE, 2020. 3.4.3

[323] Vaibhav V. Unhelkar, Ho Chit Siu, and Julie A. Shah. Comparative perfor-

mance of human and mobile robotic assistants in collaborative fetch-and-deliver tasks. In *HRI'14: PROCEEDINGS OF THE 2014 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION*, ACM IEEE International Conference on Human-Robot Interaction, pages 82–89. IEEE, 2014. 3.4.1

[324] Vaibhav V. Unhelkar, Shen Li, and Julie A. Shah. Decision-making for bidirectional communication in sequential human-robot collaborative tasks. In *PROCEEDINGS OF THE 2020 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI '20)*, ACM IEEE International Conference on Human-Robot Interaction, pages 329–341. ASSOC COMPUTING MACHINERY, 2020. 3.4.3

[325] Peggy van Minkelen, Carmen Gruson, Pleun van Hees, Mirle Willems, Jan de Wit, Rian Aarts, Jaap Denissen, and Paul Vogt. Using self-determination theory in social robots to increase motivation in l2 word learning. In *PROCEEDINGS OF THE 2020 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI '20)*, ACM IEEE International Conference on Human-Robot Interaction, pages 369–377. ASSOC COMPUTING MACHINERY, 2020. 3.4.3

[326] Simone Varrasi, Santo Di Nuovo, Daniela Conti, and Alessandro Di Nuovo. Social robots as psychometric tools for cognitive assessment: A pilot test. In F Ficuciello, F Ruggiero, and A Finzi, editors, *HUMAN FRIENDLY ROBOTICS*, volume 7 of *Springer Proceedings in Advanced Robotics*, pages 99–112. SPRINGER INTERNATIONAL PUBLISHING AG, 2019. 3.3.3, 3.4.3

[327] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. Attention is all you need. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf. 7

[328] Vighnesh Vatsal and Guy Hoffman. Design and analysis of a wearable robotic forearm. In *2018 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA)*, IEEE International Conference on Robotics and Automation ICRA, pages 5489–5496. IEEE COMPUTER SOC, 2018. 3.4.3

[329] Ajay Vishwanath, Aalind Singh, Yi Han Victoria Chua, Justin Dauwels, and Nadia Magnenat-Thalmann. Humanoid co-workers: How is it like to work with a robot? In *2019 28TH IEEE INTERNATIONAL CONFERENCE ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN. IEEE, 2019. 3.4.3

[330] Dinh-Son Vu, Ulysse Cote Allard, Clement Gosselin, Francois Routhier, Benoit Gosselin, and Alexandre Campeau-Lecours. Intuitive adaptive orientation control of assistive robots for people living with upper limb disabilities. In F Amirabdollahian, E Burdet, and L Masia, editors, *2017 INTERNATIONAL CONFERENCE ON REHABILITATION ROBOTICS (ICORR)*, International Conference on Rehabilitation Robotics ICORR, pages 795–800. IEEE, 2017. 3.4.3

[331] H. Wandke. Assistance in human–machine interaction: A conceptual framework and a proposal for a taxonomy. *Theoretical Issues in Ergonomics Science*, 6(2):129–155, 2005. ISSN 1464536X. doi: 10.1080/1463922042000295669. URL https://www.tandfonline.com/doi/abs/10.1080/1463922042000295669. 3.1.3

[332] Yeping Wang, Gopika Ajaykumar, and Chien-Ming Huang. See what i see: Enabling user-centric robotic assistance using first-person demonstrations. In *PROCEEDINGS OF THE 2020 ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI '20)*, ACM IEEE International Conference on Human-Robot Interaction, pages 639–648. ASSOC COMPUTING MACHINERY, 2020. 3.4.1, 3.5.3

[333] Zihao Wang, Shaofei Cai, Guanzhou Chen, Anji Liu, Xiaojian Ma, and Yitao Liang. Describe, explain, plan and select: Interactive planning with large language models enables open-world multi-task agents. *arXiv preprint arXiv:2302.01560*, 2023. 2.4.1

[334] Ziheng Wang and Ann Majewicz Fey. Human-centric predictive model of task difficulty for human-in-the-loop control tasks. *PLOS ONE*, 13(4):1–21, 04 2018. doi: 10.1371/journal.pone.0195053. URL https://doi.org/10.1371/journal.pone.0195053. 5

[335] Shih-En Wei, Varun Ramakrishna, Takeo Kanade, and Yaser Sheikh. Convolutional pose machines. In *CVPR*, 2016. 4.3

[336] Luca Weihs, Matt Deitke, Aniruddha Kembhavi, and Roozbeh Mottaghi. Visual room rearrangement. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5922–5931, 2021. 2.2.1

[337] Jacqueline Kory Westlund, Goren Gordon, Samuel Spaulding, Jin Joo Lee, Luke Plummer, Marayna Martinez, Madhurima Das, and Cynthia Breazeal. Lessons from teachers on performing hri studies with young children in schools. In *ELEVENTH ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN ROBOT INTERACTION (HRI'16)*, ACM IEEE International Conference on Human-Robot Interaction, pages 383–390. ASSOC COMPUTING MACHINERY, 2016. 3.4.3

[338] Whirlpool. How to load a dishwasher properly, 2024. URL https://www.whirlpool.com/blog/kitchen/proper-way-to-load-a-dishwasher.html. 1

[339] Iris Wieser, Sibel Toprak, Andreas Grenzing, Tobias Hinz, Sayantan Auddy, Ethem Can Karaoguz, Abhilash Chandran, Melanie Remmels, Ahmed El Shinawi, Josip Josifovski, Leena Chennuru Vankadara, Faiz Ul Wahab, Alireza M. Bahnemiri, Debasish Sahu, Stefan Heinrich, Nicolas Navarro-Guerrero, Erik Strahl, Johannes Twiefel, and Stefan Wermter. Robotic home assistant with memory aid functionality. In G Friedrich, M Helmert, and F Wotawa, editors, *KI 2016: ADVANCES IN ARTIFICIAL INTELLIGENCE*, volume 9904 of *Lecture Notes in Artificial Intelligence*, pages 102–115. SPRINGER-VERLAG BERLIN, 2016. 3.4.1

[340] Katie Winkle and Paul Bremner. Investigating the real world impact of emotion portrayal through robot voice and motion. In A Howard, K Suzuki, and L Zollo, editors, *2017 26TH IEEE INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 627–634. IEEE, 2017. 3.4.3

[341] Luke J. Wood, Abolfazl Zaraki, Michael L. Walters, Ori Novanda, Ben Robins, and Kerstin Dautenhahn. The iterative development of the humanoid robot kaspar: An assistive robot for children with autism. In A Kheddar, E Yoshida, SS Ge, K Suzuki, JJ Cabibihan, F Eyssel, and H He, editors, *SOCIAL ROBOTICS, ICSR 2017*, volume 10652 of *Lecture Notes in Artificial Intelligence*, pages 53–63. SPRINGER INTERNATIONAL PUBLISHING AG, 2017. 3.4.3

[342] Bryce Woodworth, Francesco Ferrari, Teofilo E. Zosa, and Laurel D. Riek. Preference learning in assistive robotics: Observational repeated inverse reinforcement learning. In Finale Doshi-Velez, Jim Fackler, Ken Jung, David Kale, Rajesh Ranganath, Byron Wallace, and Jenna Wiens, editors, *Proceedings of the 3rd Machine Learning for Healthcare Conference*, volume 85 of *Proceedings of Machine Learning Research*, pages 420–439. PMLR, 17–18 Aug 2018. URL https://proceedings.mlr.press/v85/woodworth18a.html. 2.3.1

[343] Jimmy Wu, Rika Antonova, Adam Kan, Marion Lepert, Andy Zeng, Shuran Song, Jeannette Bohg, Szymon Rusinkiewicz, and Thomas Funkhouser. Tidybot: Personalized robot assistance with large language models. *arXiv preprint arXiv:2305.05658*, 2023. 1, 1.1.5, 2.2.1, 6, 6.3.2

[344] Min Wu, Yanhao He, and Steven Liu. Shared impedance control based on reinforcement learning in a human-robot collaboration task. In K Berns and D Gorges, editors, *ADVANCES IN SERVICE AND INDUSTRIAL ROBOTICS*, volume 980 of *Advances in Intelligent Systems and Computing*, pages 95–103.

SPRINGER INTERNATIONAL PUBLISHING AG, 2020. 3.4.2

[345] Hanwei Xu, Yujun Chen, Yulun Du, Nan Shao, Yanggang Wang, Haiyu Li, and Zhilin Yang. Gps: Genetic prompt search for efficient few-shot learning. *arXiv preprint arXiv:2210.17041*, 2022. 2.4.2

[346] Tian Linger Xu, Hui Zhang, and Chen Yu. Cooperative gazing behaviors in human multi-robot interaction. *Interaction Studies*, 14:390–418, 01 2013. 2.1.2

[347] Yiqing Xu, Jiayuan Mao, Yilun Du, Tomas Lozáno-Pérez, Leslie Pack Kaebling, and David Hsu. "set it up!": Functional object arrangement with compositional generative models. In *Robotics: Science and Systems*, 2024. 1.1.5, 6.5

[348] Zhi Yan, Nathan Crombez, Jocelyn Buisson, Yassine Ruichck, Tomas Krajnik, and Li Sun. A quantifiable stratification strategy for tidy-up in service robotics. In *2021 IEEE international conference on advanced robotics and its social impacts (ARSO)*, pages 182–187. IEEE, 2021. 2.2.1

[349] Chengrun Yang, Xuezhi Wang, Yifeng Lu, Hanxiao Liu, Quoc V Le, Denny Zhou, and Xinyun Chen. Large language models as optimizers. *arXiv preprint arXiv:2309.03409*, 2023. 2.4.2

[350] Jianwei Yang, Hao Zhang, Feng Li, Xueyan Zou, Chunyuan Li, and Jianfeng Gao. Set-of-mark prompting unleashes extraordinary visual grounding in gpt-4v. *arXiv preprint arXiv:2310.11441*, 2023. 6.2.1

[351] Sriram Yenamandra, Arun Ramachandran, Mukul Khanna, Karmesh Yadav, Jay Vakil, Andrew Melnik, Michael Büttner, Leon Harz, Lyon Brown, Gora Chand Nandi, Arjun PS, Gaurav Kumar Yadav, Rahul Kala, Robert Haschke, Yang Luo, Jinxin Zhu, Yansen Han, Bingyi Lu, Xuan Gu, Qinyuan Liu, Yaping Zhao, Qiting Ye, Chenxiao Dou, Yansong Chua, Volodymyr Kuzma, Vladyslav Humennyy, Ruslan Partsey, Jonathan Francis, Devendra Singh Chaplot, Gunjan Chhablani, Alexander Clegg, Theophile Gervet, Vidhi Jain, Ram Ramrakhya, Andrew Szot, Austin Wang, Tsung-Yen Yang, Aaron Edsinger, Charlie Kemp, Binit Shah, Zsolt Kira, Dhruv Batra, Roozbeh Mottaghi, Yonatan Bisk, and Chris Paxton. Towards open-world mobile manipulation in homes: Lessons from the neurips 2023 homerobot open vocabulary mobile manipulation challenge, 2024. URL `https://arxiv.org/abs/2407.06939`. 1

[352] Mingxin Yu, Yingzi Lin, David Schmidt, Xiangzhou Wang, and Yu Wang. Human-robot interaction based on gaze gestures for the drone teleoperation. *Journal of Eye Movement Research*, 7(4), Sep. 2014. doi: 10.16910/jemr.7.4.4. URL `https://bop.unibe.ch/JEMR/article/view/2387`. 2.1.3

[353] Wenhao Yu, Nimrod Gileadi, Chuyuan Fu, Sean Kirmani, Kuang-Huei Lee, Montse Gonzalez Arenas, Hao-Tien Lewis Chiang, Tom Erez, Leonard Hasenclever, Jan Humplik, et al. Language to rewards for robotic skill synthesis.

arXiv preprint arXiv:2306.08647, 2023. 2.4.1

[354] Setareh Zafari, Isabel Schwaninger, Matthias Hirschmanner, Christina Schmidbauer, Astrid Weiss, and Sabine T. Koeszegi. "you are doing so great!" - the effect of a robot's interaction style on self-efficacy in hri. In *2019 28TH IEEE INTERNATIONAL CONFERENCE ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN. IEEE, 2019. 3.4.3

[355] Andy Zeng, Maria Attarian, Brian Ichter, Krzysztof Choromanski, Adrian Wong, Stefan Welker, Federico Tombari, Aveek Purohit, Michael Ryoo, Vikas Sindhwani, et al. Socratic models: Composing zero-shot multimodal reasoning with language. *arXiv preprint arXiv:2204.00598*, 2022. 2.4.1

[356] Ruohan Zhang, Zhuode Liu, Luxin Zhang, Jake A. Whritner, Karl S. Muller, Mary M. Hayhoe, and Dana H. Ballard. AGIL: learning attention from human for visuomotor tasks. In *Computer Vision - ECCV 2018 - 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part XI*, pages 692–707, 2018. doi: 10.1007/978-3-030-01252-6_41. URL https://doi.org/10.1007/978-3-030-01252-6_41. 2.1.1

[357] Zhizheng Zhang, Zhibo Chen, and Weiping Li. Automating robotic furniture with a collaborative vision-based sensing scheme. In JJ Cabibihan, F Mastrogiovanni, AK Pandey, S Rossi, and M Staffa, editors, *2018 27TH IEEE INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (IEEE RO-MAN 2018)*, IEEE RO-MAN, pages 719–725. IEEE, 2018. 3.4.1

[358] Fangyun Zhao, Curt Henrichs, and Bilge Mutlu. Task interdependence in human-robot teaming. In *2020 29TH IEEE INTERNATIONAL CONFERENCE ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 1143–1149. IEEE, 2020. 3.4.1

[359] Huiying Zhou, Liu Yang, Honghao Lv, Kang Yi, Huayong Yang, and Geng Yang. Development of a synchronized human-robot-virtuality interaction system using cooperative robot and motion capture device. In *2019 IEEE/ASME INTERNATIONAL CONFERENCE ON ADVANCED INTELLIGENT MECHATRONICS (AIM)*, IEEE ASME International Conference on Advanced Intelligent Mechatronics, pages 329–334. IEEE, 2019. 3.4.3

[360] Huaijiang Zhu, Volker Gabler, and Dirk Wollherr. Legible action selection in human-robot collaboration. In A Howard, K Suzuki, and L Zollo, editors, *2017 26TH IEEE INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION (RO-MAN)*, IEEE RO-MAN, pages 354–359. IEEE, 2017. 3.4.3

[361] Brian D. Ziebart, Andrew Maas, J. Andrew Bagnell, and Anind K. Dey. Maximum entropy inverse reinforcement learning. In *Proc. AAAI*, pages 1433–1438, 2008. 2.3.1, 4.1.3

[362] Andrea Zignoli, Francesco Biral, Kouta Yokoyama, and Tomoyuki Shimono. Including a musculoskeletal model in the control loop of an assistive robot for the design of optimal target forces. In *45TH ANNUAL CONFERENCE OF THE IEEE INDUSTRIAL ELECTRONICS SOCIETY (IECON 2019)*, IEEE Industrial Electronics Society, pages 5394–5400. IEEE, 2019. 3.4.2