

Interactive and Explainable Robot Learning

A Comprehensive Review

Suggested Citation: Esmaeil Seraj, Kin Man Lee, Zulfiqar Zaidi, Qingyu Xiao, Zhaoxin Li, Arthur Scaquetti do Nascimento, Sanne Van Waveren, Pradyumna Tambwekar, Rohan Paleja, Devleena Das and Matthew Gombolay (2024), "Interactive and Explainable Robot Learning", : Vol. 12, No. 2-3, pp 75–349. DOI: 10.1561/23000000081.

Esmail Seraj

Georgia Institute of Technology
eseraj3@gatech.edu

Kin Man Lee

Georgia Institute of Technology
klee863@gatech.edu

Zulfiqar Zaidi

Georgia Institute of Technology
zzaidi8@gatech.edu

Qingyu Xiao

Georgia Institute of Technology
qxiao33@gatech.edu

Zhaoxin Li

Georgia Institute of Technology
zli3088@gatech.edu

Arthur Scaquetti do Nascimento

Georgia Institute of Technology
anascimento7@gatech.edu

Sanne van Waveren

Georgia Institute of Technology
sanne@gatech.edu

Pradyumna Tambwekar

Georgia Institute of Technology
ptambwekar3@gatech.edu

Rohan Paleja

Georgia Institute of Technology
rpaleja3@gatech.edu

Devleena Das

Georgia Institute of Technology
ddas41@gatech.edu

Matthew Gombolay

Georgia Institute of Technology
matthew.gombolay@cc.gatech.edu

This article may be used only for the purpose of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval.

now
the essence of knowledge
Boston — Delft

Contents

1	Introduction	77
2	Background	82
2.1	Explainable AI	82
2.2	Interactive Robot Learning	85
3	Interactive Robot Learning	89
3.1	Preliminaries	91
3.2	Levels and Modalities of Interaction	93
3.3	Learned Representations in Interactive Robot Learning	98
3.4	Learning from Demonstration	99
3.5	Learning from Human Feedback	109
3.6	Leveraging Natural Language for Interactive Robot Learning	113
3.7	Adaptive and Adaptable Interactive Robot Learning Methods	117
3.8	Mixed-Initiative Interaction and Shared-Autonomy	124
3.9	User Expectations and Trust	127
4	Feedback Mechanisms Aiding Humans in Interactive and Explainable Robot Learning	131
4.1	Approaches to Interpretability and Explainability of the Robot Feedback	134
4.2	Contents and Types of Robot Feedback	137

4.3	Modalities of Robot Feedback	142
4.4	Feedback Timing and Frequency	147
4.5	Implicit vs Explicit Feedback	151
4.6	The relationship between robot feedback and the expertise and expectations of the end-user	154
4.7	Design Guidelines for Effective Feedback in Interactive Robot Learning, Based on Feedback Content, Modality, and Frequency	156
5	Long-Term Robot Learning	160
5.1	Algorithmic Needs for Long-Term Interactive Robot Learning Methods	162
5.2	Communication in Long-Term Human-Robot Interaction	165
5.3	Evolving Learning Goals and Policy Updates	172
5.4	Design Guidelines and Future Work	177
5.5	Conclusion	178
6	Challenges and Techniques for Safety in Interactive Robot Learning	180
6.1	Physical Safety in Interactive Robot Learning	181
6.2	Perceived Safety in Interactive Robot Learning	185
6.3	Safety Aspects of Explainability and Transparency	187
6.4	Formal Methods and Guarantees	188
6.5	Reliability, Robustness, Resilience, and Error Recovery	191
6.6	Conclusion	195
7	Ethical Implications of Interactive and Explainable Robot Learning	197
7.1	Accountability	198
7.2	Privacy	201
7.3	Biases	204
7.4	Transparency	207
7.5	Diversity, Equity, and Inclusion	208
7.6	Psychological Wellbeing	209
7.7	Design Guidelines for Ethical Concerns in Interactive and Explainable Robot Learning	209

7.8	Conclusion	211
8	Evaluation, Metrics, and Benchmarking of Interactive and Explainable Robot Learning	212
8.1	Task Performance Metrics	214
8.2	Metrics for Explainability	218
8.3	Metrics for Safety	222
8.4	Benchmarks	223
8.5	Conclusion	229
9	Applications	230
9.1	Healthcare	231
9.2	Domestic and Household Tasks	235
9.3	Industrial Automation	236
9.4	Autonomous Vehicles	239
9.5	Other Applications	241
9.6	Design Guidelines and Future Work	241
10	Takeaways and Conclusion	243
	Acknowledgements	247
	References	248

Interactive and Explainable Robot Learning

Esmaeil Seraj¹, Kin Man Lee¹, Zulfiqar Zaidi¹, Qingyu Xiao¹,
Zhaoxin Li¹, Arthur Scaquetti do Nascimento¹, Sanne Van Waveren¹,
Pradyumna Tambwekar¹, Rohan Paleja¹, Devleena Das¹ and
Matthew Gombolay¹

¹*School of Interactive Computing; Georgia Institute of Technology*

ABSTRACT

This review paper embarks on a comprehensive exploration of approaches, evaluation methods, and ethical considerations in explainable and interactive systems for robotic applications, distinctly focusing on intelligent systems that are specifically designed for learning automated agents. Given the increasing integration of robots in daily life, it is crucial to focus on intelligent systems that not only can learn and adapt, but also can offer clarity and comprehension for their actions. The interactive component of these systems is thoroughly examined, evaluating the algorithms, the modalities used in interaction, and the significance of mixed-initiative and shared autonomy. We spotlight adaptive and adaptable methods, emphasizing the centrality of user-inspired research and personalized approaches in interactive robotics. A rigorous examination of safety and ethical considerations of these intelligent systems anchors the discussion, including aspects of transparency, privacy, accountability, biases, and psychological well-being. The review evaluates existing metrics and benchmarking standards for such systems and

Esmaeil Seraj, Kin Man Lee, Zulfiqar Zaidi, Qingyu Xiao, Zhaoxin Li, Arthur Scaquetti do Nascimento, Sanne Van Waveren, Pradyumna Tambwekar, Rohan Paleja, Devleena Das and Matthew Gombolay (2024), “Interactive and Explainable Robot Learning”, : Vol. 12, No. 2-3, pp 75–349. DOI: 10.1561/23000000081.

©2024 M. Gombolay

explores their practical applications across domains such as healthcare, domestic tasks, and industrial automation. Concluding with key insights and directions for future research, we provide design guidelines and points of consensus for each subject in order to equip readers with a nuanced understanding of current trends and tools in explainable and interactive robotic systems, paving the way for informed research and application in this dynamic field.

1

Introduction

The field of robotics has improved human well-being in the past, ranging from robots being deployed in manufacturing assembly lines to automate repetitive or dangerous tasks (Zhang *et al.*, 2014) to more recent robots that navigate a hospital floor to quickly deliver resources (Muio, 2019). Across numerous domains, including healthcare (Morgan *et al.*, 2022), manufacturing, disaster response (Ramchurn *et al.*, 2016), space exploration (Banerjee and Padhi, 2018), and household assistance (Brose *et al.*, 2010), robotics has the potential to increase productivity, augment humans as teammates, and perform a variety of tasks. Importantly, in the future, through high-accuracy sensing and large computational capabilities, robots can handle the dull, dirty, and dangerous tasks, while humans will be able to transition to focus on a higher level of problem-solving and task support.

However, robotics is still relatively limited in that the deployment of a robot must be done within a fixed setting, where robots may be caged to avoid dynamic agents (e.g., humans), and the task the robot must accomplish must be pre-defined and can only be changed by an army of engineers (Sanneman *et al.*, 2021). This falls short of how we envision general-use robots that can actively learn new behaviors or

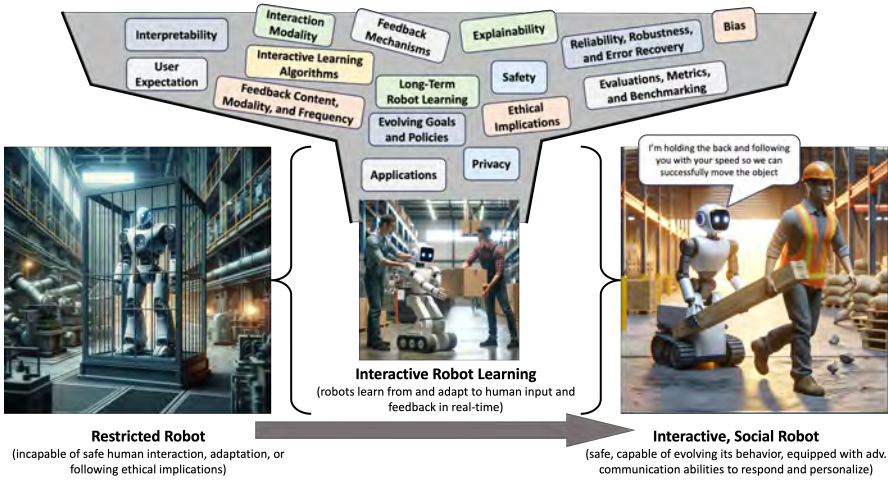


Figure 1.1: Overview of the interactive and explainable robot learning. A robot that is incapable of adapting to human styles and preferences or accounting for physical and ethical safety of its counterparts must be restricted to avoid potential harm (left-side). Through interactive robot learning, user expectations can be accounted for in long term, safety and human preferences can be communicated in an interpretable and explainable manner, all while making sure ethical implications are not violated. As a result, an interactive, social robot is trained which can freely and understandably interact and work with humans.

concepts, interact successfully with humans, and perform desired tasks correctly. To reach this vision, where we can deploy robotic agents in dynamic task settings (e.g., households, hospitals, and collaborative manufacturing), we must ask the questions of "How can we enable end-users to teach robots new behaviors?" and "How should a robot ask for help in a language end-users can understand?" We display an overview of how we see the field transitioning to this vision via explainable and interactive robot learning in Figure 1.1, requiring the combination of different techniques and characteristics across the fields of Artificial Intelligence and Robotics.

The Need for Interactive Learning – As we shift toward robotic systems that actively learn new tasks after they are deployed, we must create systems where a non-roboticist human end-user can *interact* with a robot and provide the robot with new information to improve its ability to make decisions. A successful interaction should support the transfer of

intended knowledge between the human to the robot. This is challenging for a variety of reasons. First, there are many different modalities that a human can use to communicate (e.g., by demonstration or natural language). Next, the robot can utilize a number of different techniques to derive knowledge from the communicated information. This derived knowledge must then be incorporated into a robot’s programming while understanding that the communicated knowledge may only be applicable to a limited number of task settings. Finally, end-users that vary in robotics expertise and capability require personalized systems to support this transference, leading to a need for adaptive techniques that can support a variety of users.

The Need for Explainable Systems – However, not only do we need the human to be able to communicate with the robot, but we also need the capability for the robot to *explain* its needs (i.e., request specific feedback) and justify a set of actions to ensure safety and reliability. Robot explanations should support providing the human with an understanding of what the robot understands or does not understand, and potentially why. This proves challenging for several reasons. First, a robot may utilize components that are difficult to explain (e.g., black-box modules) and may only be able to support low-fidelity explanations (i.e., explanations that are not faithful to the original model) or local explanations (i.e., explanations that only hold under a specific partition of the state space). Second, it is non-trivial to explain what an agent does not know but needs to know in order to understand a concept or behavior. Lastly, communication modalities, which depend on a robot’s physical design, may only be able to support certain types of explanations. These facets together make all-purpose targeted explanations toward human end-users difficult to generate.

Understanding how we can enable the aforementioned capabilities will be critical in helping to democratize access to robot systems (as there will be a much-lowered cost due to the removal of having automation/software engineers in the loop) and support deployment of these systems into mid-size entities, such as hospitals, small factories, and households, furthering robotic research and bringing us to a new frontier in robotic systems.

In this review paper, we present a comprehensive exploration of

approaches, evaluation methods, and ethical considerations in interactive and explainable robot learning. We place a specific emphasis on robots and avoid confounding techniques that are not amenable to the sequential decision-making problems typically encountered in robotics. There have been several other review papers, notably on Robot Learning from Demonstration (Ravichandar *et al.*, 2020; Adams *et al.*, 2022), Explainable Artificial Intelligence (Arrieta *et al.*, 2020; Anjomshoaie *et al.*, 2019; Linardatos *et al.*, 2021), Safety in Robotics (Brunke *et al.*, 2022; Gu *et al.*, 2022), Human-Robot Interaction (Wang *et al.*, 2019a; Badr and Abdul-Hassan, 2020; Halme *et al.*, 2018), and Social Robotics (Guemghar *et al.*, 2022; Koh *et al.*, 2021). These review papers each cover a set of specific techniques that are aimed toward addressing a major challenge within a specific sub-field of AI or Robotics. Furthermore, each review paper has direct or indirect implications in improving either the explainability or interactability of robots. This article, while having overlaps across the variety of aforementioned subfields and respective review papers, is unique in that we are specifically focused on building toward systems that can effectively learn through the feedback loop facilitated by the *combination of explainability and interactability*.

In this paper, we first include a background section that covers basic concepts and terminologies that are used throughout this paper (Chapter 3). Afterward, we present approaches that seek to enable robot learning from user interaction (Chapter 4), provide robotic feedback to the human to aid the human in teaching the robot using explainable AI (xAI) techniques (Chapter 5), support robot learning in the longer-term (Chapter 6), design for safety while supporting robot learning (Chapter 7), account for ethical implications such as privacy, transparency, and bias (Chapter 8), and evaluate these systems across objective and subjective metrics (Chapter 9). We conclude with a discussion of downstream applications, including healthcare, manufacturing, automation, and autonomous navigation, that will benefit from interactive and explainable robot learning systems (Chapter 10).

This article covers a wide range of concepts beneficial to creating robotic agents that can successfully interact with humans. Concluding with key insights and directions for future research, we provide design guidelines and points of consensus for each subject discussed in the paper

to equip readers with a nuanced understanding of current trends and tools in explainable and interactive robotic systems, paving the way for informed research and application in this dynamic field. This would be of interest to researchers studying Learning from Demonstration, Safety in Robotics, Explainable AI for Embodied Agents, Social Robotics, and Human-Robot Interaction.

2

Background

This article provides an extensive overview of various methods under the umbrella of interactive and explainable robot learning. Prior to discussing each of these methods, we wanted to begin by defining some key terminology and taxonomy in these fields to help ground the discourse in the rest of this article. The goal of this section is to equip the reader with the necessary context to adequately comprehend and interpret the topics discussed in this article.

2.1 Explainable AI

Explainable AI is a rapidly expanding area of research; However, due to this rise in popularity, terms within this field are often misused or overloaded. In this section, we clearly define important terms within xAI and clarify any misconceptions regarding concepts within Explainable AI.

1. What is an explanation?: Firstly, it is important to understand how to characterize an explanation in order to develop grounded approaches to Explainable Robot Learning. The function of an explanation for an autonomous agent has generally been defined as opening the “black box” of an AI system, wherein a “black box” model is a system

whose decision making processes, to transform an input from its feature space into the output space, is obfuscated from the user (Guidotti *et al.*, 2018; Adadi and Berrada, 2018). However, the characterization of what makes an explanation can not be defined so universally. A user’s goals, needs or disposition are often the deciding factors for whether or not a “statement” or “concept” presented by the robot constitutes an explanation (Hoffman *et al.*, 2023). The “who” in the human-robot interaction often dictates the effectiveness of an explanation, i.e. “How is the explanation going to be used by the user?,” or “What is the user’s knowledge of the robot and domain?,” etc (Ehsan and Riedl, 2020; Gilpin *et al.*, 2022; Tambwekar and Gombolay, 2023).

2. Interpretability: In machine learning, “interpretability” and “explainability” are closely related concepts, but they have distinct meanings ((Rudin *et al.*, 2022)). Interpretability refers to how easily a human can understand or predict the outputs of a machine learning model or algorithm. When a model is interpretable, its internal mechanics – such as how it processes input data and arrives at a decision – are clear and straightforward((Lipton, 2018; Rudin, 2019)). For instance, a simple decision tree can be considered highly interpretable because you can easily trace how inputs lead to a specific conclusion ((Paleja *et al.*, 2023a)). Explanations have been found to increase a user’s propensity to blindly accept advice from an autonomous agent, also known as “innappropriate compliance” (Silva *et al.*, 2023). Interpretable methods empower a user to truly understand the decision making process of a robot and support their decision making in order to lessen their reliance on “black-box” systems.

3. Explainability: Explainability involves the ability to explain the actions, decisions or concepts, of a model or robot, in a human-understandable manner when the underlying algorithm itself is complex and not inherently comprehensible ((Silva *et al.*, 2023; Das *et al.*, 2023; Madumal *et al.*, 2020)). Methods under the umbrella of explainability follow a “human-centered” paradigm with a stakeholder focus (Ehsan *et al.*, 2022). For complex models like deep neural networks, explainability might require additional tools or techniques to translate their processes and decisions into understandable terms (Poursabzi-Sangdeh *et al.*, 2018). Explainable methods focus on providing understandable reasons

behind a model’s output, even if the internal workings of the model are complex.

4. Explanations in Robot Learning: A fundamental stratification critical to understanding Explainable-AI for Robot Learning is explanations which explain the learned policy vs explanations which explain the learned reward. Explaining the learned policy focuses on clarifying how the robot selects actions based on its current state and available information, which is valuable for understanding decision-making in specific environments and embodiments (Paleja *et al.*, 2021; Silva and Gombolay, 2021; Sequeira and Gervasio, 2020; Hayes and Shah, 2017a; Topin and Veloso, 2019). This approach often involves visualizing decision boundaries, action distributions, or state-action trajectories, highlighting how the robot navigates and responds to different situations. In contrast, explaining the learned reward provides insight into the robot’s objectives and preferences, which offers a more abstract understanding applicable across different embodiments and dynamics (Huang *et al.*, 2019; Tabrez *et al.*, 2019b; Sanneman and Shah, 2021). Explaining the reward function enables a user to capture the underlying goals that guide the robot’s actions, independent of specific policies or dynamics.

5. Mental Models: Two integral concepts that are crucial to understand for developing Explainable Robot Learning methods are “*mental models*” and “*conceptual models*.” Mental models are defined as a user’s understanding or insight regarding the functionality of an underlying system, which in the context of this paper is an autonomous system or robot (Greca and Moreira, 2000). Measuring and improving mental models has been shown to improve team-performance and understanding (Bansal *et al.*, 2019). Conceptual models are precise and complete representations of the actual decision making process of system (Greca and Moreira, 2000). To make an autonomous system fully transparent (Belle and Papantonis, 2021), an explanation should reconcile any differences between the conceptual model and the user’s mental model of the autonomous system or robot, thereby making mental model assessment a useful measurement tool to gauge the helpfulness of an explanation or interaction with a robot (Gero *et al.*, 2020).

6. Responsible AI (Fairness & Accountability): Explainable AI

is a core component of Responsible Artificial Intelligence Research, which relates to the *fairness* and *accountability* of autonomous systems (Arrieta *et al.*, 2020). An unfair or biased system is one which makes predictions or decisions that inequitably disadvantages or excludes any individual due to their membership in some group (Hardt *et al.*, 2016; Dwork *et al.*, 2012). These unfair decisions are often caused by discriminative feature associations learnt by a predictor, or unbalanced datasets which propagate harmful stereotypes. Beyond bias, many autonomous systems are being deployed without adequate privacy protection measures or testing to minimize the negative impact of decisions in consequential circumstances. Accountability refers to the idea that an individual, in this case the designers of interactive systems, are responsible for their actions and thereby must be able to explain their aims, motivations and reasons (Novelli *et al.*, 2023). Within the framework of Responsible AI, Accountability motivates practitioners to put in place methods to accurately assess and mitigate any potential damage caused by autonomous systems which interact with humans (Intelligence, n.d.; Tsamados *et al.*, 2021). These concepts of fairness and accountability are strongly intertwined with any discussion of interactive and explainable autonomous systems, and we hope that readers keep these topics in mind while processing the content in this article and envisioning future work in this field.

2.2 Interactive Robot Learning

In this section, we define some key terms which will be referenced across multiple sections of this paper. We first define important terms relating to human-robot interaction such as Mixed-Initiative Interaction and Theory of Mind. Next, we move onto interactive robot learning and define important terms such as Long-Term Learning, Active Learning, and the Correspondence Problem.

1. **Mixed-Initiative Interaction:** Mixed-initiative interaction refers to a flexible interaction strategy in which each agent (human or machine) contributes what it is best suited at the most appropriate time (Allen *et al.*, 1999; Seraj *et al.*, 2023). The goal is to maximize the effectiveness of the team through fluid coordination between agents.

Unlike conventional computer systems, the physical embodiment and mobility of robots involves a broader range of actions and interactions with humans that requires more robust mixed-initiative strategies (Schermerhorn and Scheutz, 2011; Jiang and Arkin, 2015). Interactive robot learning is an attractive approach for devising human-robot mixed-initiative strategies, which can enable the dynamic adaptation of roles and responsibilities based on the evolving context of the task and the capabilities of each participant through physical, interactive experience.

2. Theory of Mind: Theory of Mind (ToM) refers to the cognitive ability to attribute mental states, i.e., beliefs, intentions, desires, and knowledge, to oneself and others (Carlson *et al.*, 2013; Frith and Frith, 2005). For robots engaged in interactive learning, implementing ToM can be transformative, offering more nuanced ways to understand the human partner’s expectations, actions, and future behavior (Camerer *et al.*, 2004; Sugden, 2008; Chong *et al.*, 2005; Li *et al.*, 2019a). Prior work has delved into the integration of ToM into robotic systems (Konan *et al.*, 2022), showing that robots with rudimentary ToM capabilities tend to facilitate more effective human-robot interactions, as they can anticipate human needs and actions more accurately.

3. Correspondence Problem: The correspondence problem in robotics encapsulates two primary challenges: aligning a robot’s actions with human expectations and adapting human demonstrations to fit the robot’s unique form factor.

On one front, the correspondence problem concerns the alignment between the robot’s internal states and actions and the expectations and interpretations of human users. This aspect of the correspondence problem becomes pronounced when there’s a divergence between the robot’s performed actions and what humans anticipate, leading to potential confusion or erosion of trust. Traditionally, this challenge was met with rigid, rule-based explanations that scarcely bridged the gap in understanding between humans and robots. Nevertheless, advancements, particularly the incorporation of Theory of Mind (ToM) into robotics, have ushered in a new era of solutions. Robots equipped with ToM can generate context-aware, nuanced explanations that consider the human user’s likely understanding or misconceptions, thereby offering feedback and explanations that are more effectively aligned with human

expectations (Nehaniv, Dautenhahn, *et al.*, 2002; Johnson and Demiris, 2004; Billard and Grollman, 2013; Argall *et al.*, 2009a).

Conversely, the second facet of the correspondence problem revolves around the physical embodiment of robots. It specifically addresses the challenge of translating actions performed by a human teacher during a demonstration into actions that a robot, with its distinct form factor, can perceive or execute. This translation process necessitates two pivotal mappings: the record mapping, which involves capturing or representing the teacher’s actions in a dataset, and the embodiment mapping, which ensures the recorded data aligns with the robot’s capabilities (Argall *et al.*, 2009b). For instance, a human teaching a robot to move an object on a table demonstrates actions represented by human joint angles, observed by the robot through camera images. The robot, however, must interpret these actions not directly from the images but through a process that extracts the pertinent information, exemplifying record mapping. Furthermore, given the inherent differences in physical capabilities between humans and robots, embodiment mapping becomes essential to translate the demonstration into a format congruent with the robot’s operational paradigm.

4. Long-Term Learning: Long-term learning are machine learning techniques designed for the long-term adaptation and enhancement of models with new data. These methods contrast with short-term learning strategies, which are often focused on immediate or cross-sectional data processing without the necessity for future adaptation. In the context of interactive robot learning, long-term learning serves as a foundational tool for updating policies prospectively, enabling robots to refine their actions and strategies as they encounter new information. This long-term learning perspective ensures that robots can evolve and adapt over time, as opposed to short-term learning approaches that might optimize for the present without consideration for future changes or data.

5. Active Learning: In robotics, active learning involves training a robot on an initial dataset initially, followed by the use of an acquisition function to select specific data points for labeling by an expert oracle, which could be a human or specialized algorithm. The acquisition function typically selects points for relabeling based on criteria such as the model’s uncertainty, informativeness, or expected learning gain.

By actively choosing the most informative data points or tasks for labeling or exploration, the robot can effectively reduce the amount of data needed to achieve the desired level of performance (Settles, 2009). In the context of human-robot interaction, this method enhances the robot’s ability to understand and respond to human actions and intentions more effectively, enabling more personalized and contextually appropriate responses (Bobu *et al.*, 2021; Bıyık *et al.*, 2019). This iterative interaction between human and robot not only streamlines the learning process but also ensures that the robot’s development is closely aligned with human feedback, leading to more robust and user-tailored robotic behaviors (Bobu *et al.*, 2021; Brown *et al.*, 2021; Bıyık *et al.*, 2019).

3

Interactive Robot Learning

Envision a future where a robot and a human collaborator work side by side in a factory, seamlessly communicating and learning from one another. The robot observes the human’s expert maneuvers, listens to verbal instructions, and even understands non-verbal cues to master a complex assembly task. This is the world of interactive robot learning – a dynamic and evolving field standing at the intersection of robotics, artificial intelligence, and human-computer interaction in which the collaborative and iterative process between humans and robots transcends traditional boundaries, enabling robots to acquire knowledge and skills through active engagement. In this Chapter, we delve into the intricacies of this collaborative and iterative process where humans and robots engage actively to transfer knowledge and skills, paving the way for robots to learn and perform tasks more effectively in diverse environments.

This chapter aims to provide a comprehensive understanding of the multifaceted nature of interactive robot learning, highlighting the technological advancements and challenges in creating robots that can effectively learn from and collaborate with humans. We begin with Preliminaries (section 3.1), introducing fundamental concepts such as

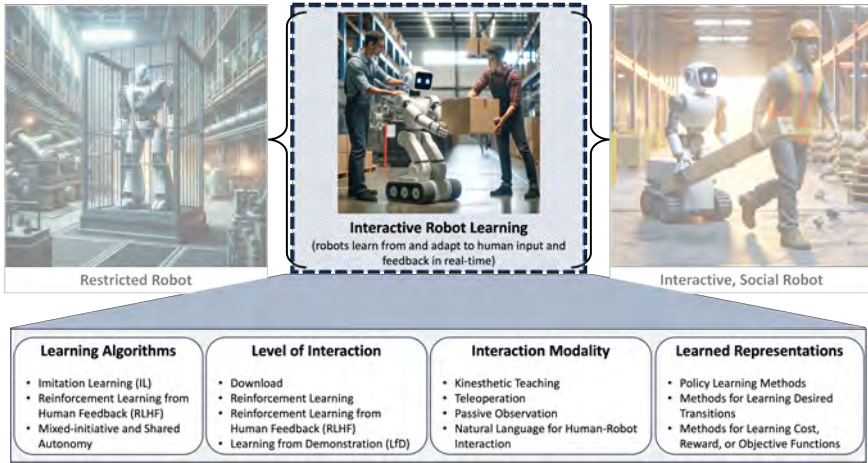


Figure 3.1: Overview of the key concepts discussed in this chapter, interactive robot learning.

Markov Decision Processes, and reinforcement learning. In Section 3.2, we address the spectrum of interactivity in robot learning, ranging from basic command-and-control interactions to more sophisticated learning from demonstration (LfD) methodologies and we examine how different modes of communication, such as Kinesthetic Teaching, Teleoperation, and Natural Language contribute to the robot’s learning process. In Section 3.3, we look at the various representations used to represent the knowledge encapsulated within a model. In Section 3.4, we delve into techniques ranging from direct imitation to sophisticated methods like Inverse Reinforcement Learning, each offering unique approaches to skill acquisition. Section 3.5 expands on how robots can learn from human feedback. Section 3.6 discusses natural language based interactive learning methods. In Section 3.7 we explore adaptive and adaptable learning methods.

In Section 3.8, we explore the dynamic of control and initiative sharing between humans and robots, crucial for effective collaboration. Lastly, in Section 3.9, we tackle the psychological and social aspects of interactive robot learning, emphasizing the importance of building trust and meeting user expectations for successful long-term interactions.

3.1 Preliminaries

Markov Decision Process (MDP): A MDP is a mathematical framework commonly used to model sequential decision-making problems (possibly stochastic) where an agent interacts with an environment over time. Standard definitions of a MDP is denoted by the 5-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{R}, \mathcal{T}, \gamma)$, where:

- **State space ($\mathcal{S}$):** The set of states representing the current situation or configuration of the system.
- **Action space ($\mathcal{A}$):** The set of possible moves or decisions that the agent can take.
- **Reward ($\mathcal{R} : \mathcal{S} \times \mathcal{A}$):** A numerical, scalar signal that indicates the immediate benefit or cost associated with an agent's action in a particular state.
- **Transition function ($\mathcal{T} : \mathcal{S} \times \mathcal{A} \times \mathcal{S}$):** The transition function which defines the probability of transitioning from one state to another given an action.
- **Discount factor (γ):** The discount factor models the present value of future rewards. γ typically lies within the interval $[0, 1]$.

The MDP is often extended with a sixth element in the tuple, the initial state distribution $\rho_0(s)$, which indicates the possible starting configurations of the system.

Reinforcement Learning (RL): Reinforcement Learning (RL) is a subfield of machine learning that focuses on enabling agents to make sequential decisions by interacting with an environment (Sutton and Barto, 2018). Unlike supervised learning, where models are trained on labeled data, and unsupervised learning, where models discover patterns without explicit guidance, RL involves learning through trial and error (Russell and Norvig, 2010).

At its core, RL involves an agent, a set of actions that the agent can take, an environment in which the agent operates, and a system

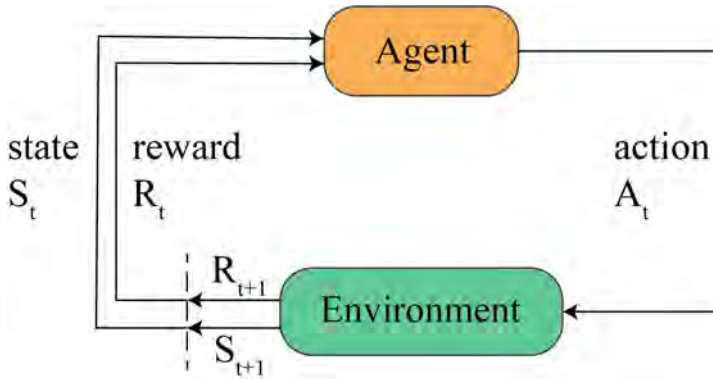


Figure 3.2: Figure depicting the learning process in reinforcement learning.

of rewards or penalties that the agent receives based on its actions as shown in Figure 3.2. The goal of the agent is to learn a policy $\pi : \mathcal{S} \rightarrow \mathcal{A}$, a strategy for selecting actions, that maximizes the cumulative reward over time. Typical RL approaches operate over some variant of the MDP formulation.

A key mathematical foundation of RL is the Bellman equation, which expresses the value ($V(s)$) of being in a particular state (s) as the sum of the immediate reward ($R(s, a)$) and the discounted value of the expected future rewards (Bellman, 1957):

$$V(s) = \max_a \left(R(s, a) + \gamma \sum_{s'} P(s' | s, a) V(s') \right) \quad (3.1)$$

Here, γ is the discount factor that determines the importance of future rewards, and $P(s' | s, a)$ is the probability of transitioning to state s' from state s by taking action a .

RL is particularly well-suited for problems where explicit training data is scarce or unavailable, and the robot must learn from experience. Therefore, RL frameworks often allow end-users to define custom reward functions tailored to specific tasks or objectives (Cabi *et al.*, 2019). Popular algorithms in RL include Q-learning, Deep Q Networks (DQN), Policy Gradient methods, and more recently, algorithms based on deep neural networks such as Proximal Policy Optimization (PPO) and Trust Region Policy Optimization (TRPO) (Arulkumaran *et al.*,

2017). The application domains of reinforcement learning are diverse, ranging from game playing (e.g., AlphaGo) to robotics, finance, health-care, and autonomous systems. RL has shown remarkable success in solving complex decision-making problems, making it a powerful tool for addressing real-world challenges where sequential decision-making is crucial (Mnih *et al.*, 2015; Schulman *et al.*, 2017).

Natural Language Processing: The field of Natural Language Processing (NLP) is rapidly advancing, witnessing significant progress. In this chapter, we also explore how NLP advancements have facilitated progress in interactive robot learning. For a comprehensive understanding of NLP concepts and developments, we recommend reading the survey paper by Zhao *et al.* (Zhao *et al.*, 2023).

3.2 Levels and Modalities of Interaction

In this section, we delve into the nuanced spectrum of interaction within the domain of interactive robot learning.

3.2.1 Levels of Interaction

The degree of interaction between humans and robots is a thoroughly examined subject in human-robot interaction studies (Onnasch and Roesler, 2021, Mukherjee *et al.*, 2022, Moorman *et al.*, 2023b). Similarly, we propose a continuum for interactive robot learning, classifying it based on the extent of interaction, ranging from minimal involvement (i.e., zero interaction) to significant user engagement (i.e., high involvement).

- **Download:** In this mode, users merely download a new behavioral capability without engaging in any interaction that contributes to the robot’s learning process. Users observe the robot’s performance without actively participating in its learning. An example of such an interaction could involve observing a newly added capability for a robot or a car, such as a Roomba or a Tesla, as part of a software update.

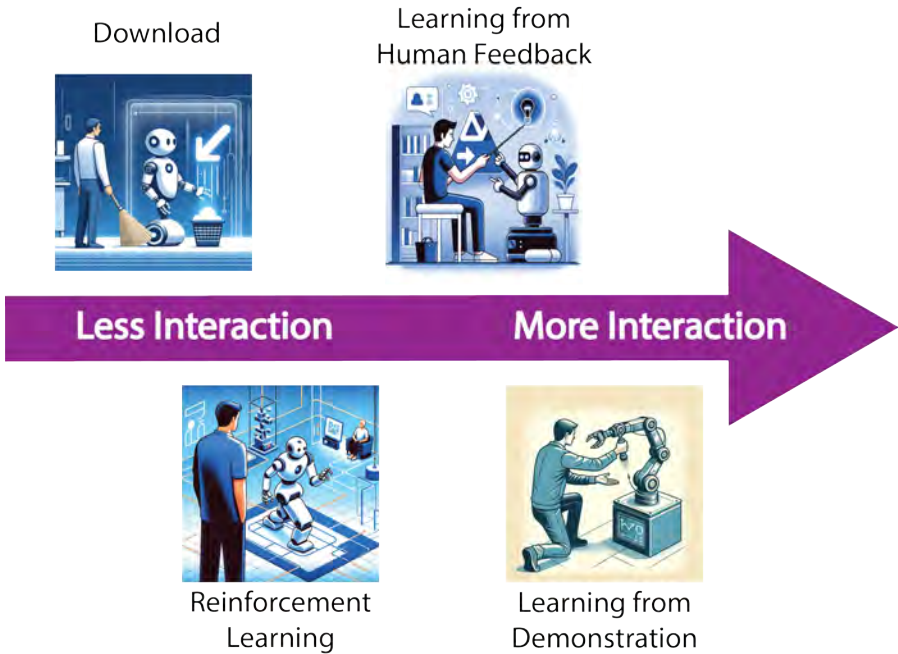


Figure 3.3: Depiction of the gradation of level of human-robot interaction in interactive robot learning.

- **Reinforcement Learning:** RL presents a low-involvement, learning from trial and error class of robot learning methods (Sutton and Barto, 2018). The interaction is limited to users observing the robot’s learning process through experiences.
- **Interactive Reinforcement Learning or Learning from Human Feedback:** This category demands increased user involvement, where users provide feedback that the robot utilizes to enhance its performance. It encompasses methods such as RLHF (Christiano *et al.*, 2017c) and PbRL (Wirth *et al.*, 2017). Users actively participate by offering feedback or corrective demonstrations to guide the robot’s learning. For a more detailed analysis of methods and algorithms using interactive reinforcement learning methods, refer to Section 3.5.
- **Learning from Demonstration:** Algorithms and methods

falling under this category necessitate the highest level of user involvement. Users are required to provide a demonstration illustrating a specific task or desired behavior. The robot then learns from this human-provided data (Ravichandar *et al.*, 2020).

Figure 3.3 visualizes this continuum depicting the various level of interactions attributed to various algorithms.

3.2.2 Modalities in Interactive Robot Learning

Learning in robots involves various modalities, each offering unique approaches to impart knowledge and skills. This section explores key modalities for interactive robot learning, including *kinesthetic teaching*, *teleoperation*, *passive observations*, and *natural language* commands.

Kinesthetic teaching involves physically moving the robot to demonstrate a behavior. During this interaction, the robot’s state parameters, such as joint positions, velocities, and torques, are continuously captured. This recorded data serves as the basis for training machine learning models (Ravichandar *et al.*, 2020). Kinesthetic teaching doesn’t necessitate additional external sensors, making it a straightforward method. Kinesthetic has been primarily applied to manipulation tasks on robotic arms due to its intuitive nature (Stavridis *et al.*, 2022). However, its application to other platforms, like legged robots or robotic hands, is limited due to the non-intuitive nature of demonstrations on such setups. Additionally, the direct recording of data on the robot using integrated sensors eliminates the correspondence problem, offering accurate and context-aware learning.

Teleoperation involves a user controlling the robot using external inputs such as a joystick, graphical user interface, or other interfaces. The recorded robot state and trajectory data are then utilized to train machine learning models (Ravichandar *et al.*, 2020). Tele-operation has been a widely used method to collect data for a range of tasks such as trajectory learning (the process of learning and reproducing specific paths or movements Abbeel *et al.*, 2010), task learning (the acquisition of skills or knowledge to perform a particular task Peters *et al.*, 2003), grasping (Wonsick and Padir, 2020) navigation (Cèsar-Tondreau *et al.*, 2021) and other high-level tasks (Mohseni-Kabir *et al.*, 2015).

Various interfaces have been developed to tele-operate the robots and collect data.

- **Joystick:** In this method, the human instructor directs the demonstration and conveys information about the actions to the robot's controller through either a wireless or wired joystick (Wang *et al.*, 2019b). This strategy has been designed for diverse LfD applications, including tasks such as robot polishing (Nagata *et al.*, 2001), robot soccer (Katagami and Yamada, 2001), and robotic welding (Nose *et al.*, 1991). Joysticks have been a very common interface for tele-operating robots, although it may introduce correspondence problems due to the mismatch in degrees of freedom (Ravichandar *et al.*, 2020).
- **Shadow Robots:** Another approach to tele-operate a robot involves the use of shadow robots, a concept demonstrated by Katsura *et al.*, 2010 and Fu *et al.*, 2024. This interface allows for the tele-operation of a robot by kinesthetically manipulating a twin shadow robot. The real-time movements of the shadow robot are replicated on the primary robot, offering a highly intuitive control mechanism that surpasses traditional joystick interfaces. This method capitalizes on the human operator's natural sense of movement, enhancing the teleoperation experience and providing a more seamless and responsive control interface for intricate robotic tasks.
- **Haptic Devices:** Haptic devices and interfaces are also an active area of research to tele-operate a robot (Ravichandar *et al.*, 2020). Su *et al.*, 2016 incorporated haptic feedback into their work to guide the robot to perform a peg-in-hole task, which involves inserting a peg into a corresponding hole with precision and accuracy. Kormushev *et al.*, 2011 also incorporated haptic feedback into their system to teach a robot to perform an ironing task and a door opening task. The inclusion of haptic interfaces enhances the tele-operation experience, providing operators with a tactile understanding of the robot's interactions with its environment.
- **Graphical User Interface (GUI):** GUIs have been developed

as well to serve as interfaces for tele-operation. In the work by Tavares *et al.*, 2023, a dynamic interface for tele-operating robots was introduced, adapting in real-time to the user's estimated attention level. Baker *et al.*, 2004 developed an efficient interface for a search and rescue robot that lowers user's cognitive workload, enhances their awareness and increase their efficiency. The incorporation of well-designed GUIs presents a practical avenue for refining the tele-operation of robots, enabling the human user to efficiently provide demonstrations or feedback to the robot.

- **Remote tele-operation:** This is also an active area of research as remote tele-operation allows to control and collect data on the robot from a considerable physical distances. This approach not only enables effective robot control but also presents an opportunity for scalable crowd-sourced demonstrations (Ravichandar *et al.*, 2020). Mandlekar *et al.*, 2018b developed an interface for mobile devices to remotely collect high quality 6-DoF trajectories through tele-operation. Through this approach, the researchers amassed 137.5 hours of manipulation data from remote workers, showcasing the potential for leveraging remote tele-operation in training robots for intricate multi-step manipulation tasks.

Passive observation demonstrations, a user performs a task, and external sensors, such as cameras or motion capture systems, capture information on how the task is accomplished. The robot remains a passive observer. Since the human performs the task themselves, minimal training is required. Additionally, this modality of teaching is highly useful for robots with high degrees of freedom or non-anthropomorphic robots (Ravichandar *et al.*, 2020). However, there is a disparity between human and robot actions, and there is a need of a translation layer that maps the human actions to the robot actions (See "Correspondence Problem" in Section 2.2). Issues such as occlusions, rapid human movements, and sensor noise also need to be addressed. Numerous studies have employed passive observation in LfD applications, including tasks like furniture assembly (Hayes and Scassellati, 2014), autonomous driving (Codevilla *et al.*, 2017), and table-top actions (Pervez *et al.*, 2017).

Recent advancements in **natural language** processing have spurred

a notable paradigm shift in human-robot interaction. Language, in its natural form, has become a potent tool for seamlessly communicating with robots across various applications. Refer to Section 3.6 for a more detailed discussion on how natural language is being employed as an input modality for interactive robot learning.

3.3 Learned Representations in Interactive Robot Learning

A method of categorizing robot learning algorithms, as proposed in previous literature (Celemin *et al.*, 2022), involves an exploration of the diverse manners in which knowledge is encapsulated within a model. Within this framework, three distinct classifications emerge, each corresponding to the unique approaches of methods to learning from data and interactions.

- **Policy Learning Methods (Actions):** This category encompasses methods that concentrate on learning to generate specific actions based on the information provided by the current state. A policy, denoted as $\pi(s_t) = a_t$, is acquired, mapping the current state to an action. In interactive robot learning, this method is widely employed when the robot is provided with demonstrations by a human within the robot’s action space. For instance, several DAGger (Ross *et al.*, 2011) methods have learned a direct policy from expert or human data.
- **Methods for Learning Desired Transitions:** This class of methods is designed to predict the subsequent desired state that the robot should visit given the current state, i.e., $\pi(s_t) = s_{t+1}^{des}$. Such approaches become particularly relevant when demonstrations or feedback from human interactions do not align directly with the action space of the agent. By anticipating desired transitions, these methods facilitate effective learning. For example, in Carvalho *et al.*, 2022, the robot learned state transitions for 3D block insertion from passive human demonstrations, and a low-level controller then achieved the end-effector movement needed to realize the desired state transitions.

- **Methods for Learning Cost, Reward, or Objective Functions:** In this category, methods aim to learn an objective function derived from interactions or data. The acquired objective function is then utilized to inform decision-making processes, guiding the agent in selecting actions aligned with the learned objectives. This paradigm provides a structured way to imbue the model with a sense of purpose. For example, in Sharma *et al.*, 2022a, a cost function was learned using natural language, and the robot’s action was based on the optimization of that cost function to achieve the desired behavior.

This categorization framework serves as a comprehensive guide for understanding and classifying various learning algorithms based on their distinct approaches to representing and utilizing learned behaviors. Additional details about this categorization can be found in Celemin *et al.* (2022).

3.4 Learning from Demonstration

Learning from demonstration (LfD) (Billard *et al.*, 2008; Argall *et al.*, 2009a; Chernova and Thomaz, 2014; Ravichandar *et al.*, 2020) — also known as imitation learning (Schaal, 1999; Hussein *et al.*, 2017) or apprenticeship learning (Abbeel and Ng, 2004) — is a paradigm in machine learning where a system learns a task policy from example demonstrations, typically provided by a human expert. This approach is particularly relevant in interactive robot learning, where robots are trained to reproduce complex, human-like behaviors. Through LfD, robots can acquire a wide range of skills without the need for explicit programming of each task. LfD is beneficial in scenarios where defining a clear reward function is difficult, which is a typical requirement in reinforcement learning. We adapt definitions on social learning mechanisms from Tomasello (1999), Call and Carpenter (2002), and Chernova and Thomaz (2014) that allow us to describe LfD succinctly.

- **Mimicry:** An observer copies the actions of others without appreciation of their purpose. The observer discovers the effects after conducting the action.

- **Emulation:** An observer witnesses others produce a particular result, then employs their own set of actions in attempt to produce the same result.
- **Imitation:** An observer reproduces similar actions of others to achieve the same results with the same goal. Note that achieving the same results does not necessitate operating under the same goal; Imitation establishes the distinction that the observer attempts to learn the underlying goal of the actions.

In a robotic system represented by a MDP, the LfD problem can then be defined as follows:

Definition (Learning from Demonstration). Given a set of demonstrations $\mathcal{D} = \{\tau_0, \tau_1, \dots, \tau_N\}$ of trajectories $\tau_i = \{(s_0, a_0), (s_1, a_1), \dots, (s_{M_i}, a_{M_i})\}_i$, find a robot policy $\hat{\pi}$ that *mimicks*, *emulates*, or *imitates* the underlying expert policy π^* ¹ represented by $\mathcal{D}$.

LfD methods typically follow one of two approaches:

- Directly learn a policy $\hat{\pi}^*$ that mimicks the trajectories in $\mathcal{D}$
- Learn a reward function $\hat{\mathcal{R}}(s, a)$ from $\mathcal{D}$ that attempts to capture the true unknown reward $\mathcal{R}$, and (optionally) synthesize a policy $\hat{\pi}^*$ that emulates or imitates π^* .

We will discuss fundamental methods within each category and present examples within the literature that build on these ideas.

3.4.1 Direct Policy Learning

A simple technique commonly applied in LfD to directly learn a policy for mimicry is **behavioral cloning** (BC) (Michie *et al.*, 1990). BC is a technique where a model is trained to replicate expert behavior by using supervised learning on the set of expert demonstrations. The objective

¹The π^* notation assumes that the expert is optimal; however, in practice, the expert is more often suboptimal than not.

is to learn a parameterized policy $\hat{\pi}^* : \mathcal{S} \rightarrow \mathcal{A}$, mapping states to expert actions by minimizing a loss function $\mathcal{L}$ over all demonstrations in $\mathcal{D}$:

$$\hat{\pi}^* = \arg \min_{\pi \in \Pi} \sum_{i=0}^N \sum_{j=0}^{M_i} \mathcal{L}(\pi^*(s_j), \pi(s_j)) \quad (3.2)$$

For discrete actions, the problem is represented as a classification problem, where the loss function takes the form of a cross-entropy loss. For continuous actions, the problem is represented as a regression problem, where the loss function typically takes the form of a L_2 norm.

Another perspective on BC is to view the problem as minimizing the difference between the action distribution of the robot and demonstrator policy. Utilizing the Kullback-Leibler (KL) divergence (Kullback and Leibler, 1951), the BC objective can alternatively be stated as:

$$\hat{\pi}^* = \arg \min_{\pi \in \Pi} \mathbb{E}_{s \sim \mathcal{S}_{\mathcal{D}}} [D_{KL}(\pi^*(\cdot|s) \parallel \pi(\cdot|s))] \quad (3.3)$$

The main advantages of BC include ease of implementation, sample-efficiency, and strong in-distribution performance. As a model-free, offline method, it does not require any interactions with the environment. In practice, however, BC typically struggles with generalization to out-of-distribution (OOD) scenarios unseen in the data. This problem is known as *covariate shift*.

A method proposed that addresses covariate shift in BC is **Dataset Aggregation (Dagger)** (Ross *et al.*, 2011). DAgger is an online learning method that queries the expert, π^* , to provide additional labels for scenerios not in the training set. DAgger iteratively performs rollouts of trajectories in the environment from the current policy, then queries the expert for corrections to the actions executed. The original set of demonstrations is then augmented with the new corrective demonstrations and the policy is retrained on the aggregated dataset.

In Algorithm 1, β_i is a tunable parameter that weighs between the pretrained policy and the new, updated policy. Typically, $\beta_1 = 1$, and decreases exponentially as $\mathcal{D}$ grows, since the policy in first few iterations

Algorithm 1 DAgger (Dataset Aggregation) (Ross *et al.*, 2011)

```

Initialize  $\mathcal{D} \leftarrow \emptyset$ .
Initialize  $\hat{\pi}^*$  with a pre-trained BC policy
Initialize  $\hat{\pi}_1$  to some  $\pi \in \Pi$ 
for  $i=1$  to  $N$  do
    Let  $\pi_i = \beta_i \hat{\pi}^* + (1 - \beta_i) \hat{\pi}_i$ . (optional)
    Sample  $T$ -step trajectories using  $\pi_i$ .
    Create  $\mathcal{D}_i = \{(s, \pi^*(s))\}$  of visited states and expert corrections.
    Aggregate datasets:  $\mathcal{D} \leftarrow \mathcal{D} \cup \mathcal{D}_i$ .
    Train classifier  $\hat{\pi}_{i+1}$  on  $\mathcal{D}$ .
end for
Return the best  $\hat{\pi}_i$  on validation.

```

is the least robust. As the number of iterations increase, the learned policy converges towards the true policy and reduces the distributional mismatch between the training data and the real environment. Although DAgger offers a simple method to reduce the problems with covariate shift, the algorithm assumes there is an expert it can query and that the policy can interact with the environment, which is not always feasible. Additionally, the iterative nature of the algorithm can be time-consuming and induce a high workload on the expert demonstrator.

Recently proposed approaches to BC leverage energy-based models and diffusion models such as Implicit Behavior Cloning (IBC) (Florence *et al.*, 2022), where the policy representation is changed to represent the action distribution using energy-based models, and Diffusion Policy (Chi *et al.*, 2023a), where a visuomotor robot policy is learned through a denoising diffusion process on the robot action space. These approaches have been shown to express multimodal behaviors, generalizing across multiple tasks on real-world robotic platforms at the cost of large-scale data requirements.

3.4.2 Inverse Reinforcement Learning

Inverse Reinforcement Learning (IRL) (Russell, 1998; Ng and Russell, 2000) is a method in machine learning where the objective is to infer a

reward function $\hat{\mathcal{R}}$ given the observed behavior of an expert agent (e.g. the demonstrator). Unlike BC methods, a reward mechanism provides a signal the robot can use to reason about the goal or results of the observed actions for emulation or imitation. The expert can be suboptimal; inferring the reward enable methods that allow the robot to perform actions that potentially outperform the expert at achieving the desired goal. Additionally, the robot may have capabilities different from the agent (see the correspondence problem in 2.2), where reasoning about the reward as opposed to direct actions allow the robot to best leverage its' action repertoire for the task at hand. A well inferred reward function, $\hat{\mathcal{R}}$, should be invariant to the system's transition dynamics, $\mathcal{T}$, to generalize across multiple systems. The formal definition of the IRL problem is given as follows:

Definition (Inverse Reinforcement Learning). Given a MDP $\mathcal{M}$ with an unknown reward function $\mathcal{R}(s, a)$ and a dataset of trajectories $\mathcal{D} = \{\tau_0, \tau_1, \dots, \tau_N\}$, $\tau_i = \{(s_0, a_0), (s_1, a_1), \dots, (s_{M_i}, a_{M_i})\}_i$, find the optimal reward function, $\hat{R}^*(s, a)$ such that $\hat{R}^*(s, a) = \max_{R(s, a)} \mathbb{E} [\sum_{t=0}^{\infty} \gamma^t R(s, a) | \pi^*]$, where π^* is the underlying expert policy derived from $\mathcal{D}$.

IRL methods employ a feature-based reward function. One way to represent the reward function is through a linear combination of weights and feature map of the form:

$$\mathcal{R}(s, a) = \mathbf{w}^\top \phi(s, a) \quad (3.4)$$

where $\mathbf{w} \in \mathbb{R}^d$ is a weight vector and $\phi(s, a) : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}^d$ is a feature map of states and actions. With this reward formulation, the value function of an arbitrary policy π in RL can be expressed as:

$$V^\pi(s) = \mathbf{w}^\top \mu(s, \pi), \quad \mu(s, \pi) = \mathbb{E} \left[\sum_{t=0}^{\infty} \gamma^t \phi(s_t, \pi(s_t)) \mid s_0 = s \right] \quad (3.5)$$

where $\mu(s, \pi)$ is known as the *feature expectation*. The goal is then to

find an optimal weight vector $\mathbf{w}^*$ that maximizes the value function under the optimal policy π^* . That is:

$$\mathbf{w}^{*\top} \mu(s, \pi^*) \geq \mathbf{w}^{*\top} \mu(s, \pi), \quad \forall s \in \mathcal{S}, \quad \forall \pi \quad (3.6)$$

However, finding $\mathbf{w}^*$ is an ill-posed problem (i.e. there are infinitely many solutions). In fact, $\mathbf{w}^* = \mathbf{0}$ is a perfectly valid (albeit useless) solution. Because there are infinitely many reward functions that can explain why demonstrations are performed a certain way, reward ambiguity is the main challenge that IRL algorithms attempt to address. Additionally, suboptimal or heterogeneous demonstrations complicates the learning process and requires additional mechanisms to differentiate between varying levels of performance and intentions among demonstrators. For complex environments, a simulator is often required in order to accurately represent the feature map, where performance is now linked to simulation fidelity.

The literature on IRL algorithms is vast (Arora and Doshi, 2021; Adams *et al.*, 2022); covering all of these methods is out of the scope of this document. Most IRL algorithms fit into one of three categories: maximum margin methods, maximum entropy methods, and Bayesian methods. We will cover the core ideas of these methods, and a few notable exceptions that do not fit into any of these categories.

- **Maximum Margin Methods:** The first category of methods proposed for IRL. Max-margin methods aim to find a reward function that maximize the margin between the optimal value function to all other value functions. The core idea is similar to support vector machines (SVMs) (Cortes and Vapnik, 1995) in essence. Notable examples of max-margin IRL are Ng and Russell, 2000; Abbeel and Ng, 2004; Ratliff *et al.*, 2006.
- **Maximum Entropy Methods:** Max-margin methods introduce a bias in the learned reward function that can skew the robot policy towards trajectories unrelated to the observed expert demonstrations. To counteract this bias, maximum entropy methods aim to obtain a distribution of trajectories that makes the

least assumptions beyond to matching the feature expectations by utilizing the principle of maximum entropy (Jaynes, 1957). Examples of maximum entropy IRL include Ziebart *et al.*, 2008; Ziebart *et al.*, 2013; Boularias *et al.*, 2011; Wulfmeier *et al.*, 2015.

- **Bayesian Methods:** In Bayesian IRL, Bayesian updates are performed to maximize the posterior distribution of a reward function. The update is performed as follows:

$$Pr(\mathcal{R} \mid \tau) \propto Pr(\tau \mid \mathcal{R})Pr(\mathcal{R}) \quad (3.7)$$

where $Pr(\tau \mid \mathcal{R})$ is the observation likelihood of the reward and $Pr(\mathcal{R})$ is the prior distribution over the reward. This approach is advantageous if expert knowledge about the reward exists and offers a probabilistic formulation to model a mixture of reward functions. Examples of Bayesian IRL are Ramachandran and Amir, 2007; Choi and Kim, 2011; Levine *et al.*, 2011.

3.4.3 Generative Learning

Advances in adversarial learning with generative adversarial networks (GAN) (Goodfellow *et al.*, 2014) have crossed over into imitation learning and IRL to form two popular state-of-the-art methods used for research in interactive robot learning: **Generative Adversarial Imitation Learning** (GAIL) (Ho and Ermon, 2016) and **Adversarial Inverse Reinforcement Learning** (AIRL) (Fu *et al.*, 2017). Both methods adversarially train a generative policy against a discriminator that aims to classify trajectories from the learned policy from expert demonstrations. While similar in nature, GAIL is not an IRL algorithm, as it does not recover a reward function, unlike AIRL.

3.4.4 Skill Learning

In the domain of robotics, learning spans a diverse array of methodologies, covering both low-level skill acquisition and high-level task learning. Low-level skills refer to fundamental capabilities such as motor control and manipulation, while high-level skills entail complex, goal-oriented

behaviors and decision-making processes. This section explores interactive learning techniques tailored for both low-level and high-level skill acquisition in robotics.

Low-level Skill Learning

Learning low-level skills in robotics through imitation learning is common, specifically through the use of movement primitives. Movement primitives are fundamental building blocks that enable robots to execute complex tasks by combining simple, pre-defined motions. This approach is a structured learning process that is robust and sample-efficient, especially in dynamic and high-speed environments. The integration of these primitives into robotic systems paves the way for more intuitive and adaptive behaviors, bridging the gap between high-level decision making and low-level motor control.

Dynamic Movement Primitives (DMP) (Schaal, 2006; Ijspeert *et al.*, 2013) is the earliest introduced and most common class of movement primitives for learning low-level skills. The core idea behind DMPs is to represent movements as a combination of a stable, nonlinear dynamical system with a learnable forcing function to modulate the shape of the movement. In its most basic form, a DMP is represented with point attractor dynamics as follows:

$$\tau \dot{z} = \alpha_z (\beta_z (g - y) - z) + f(x), \quad \tau \dot{y} = z, \quad \tau \dot{x} = \alpha_x x \quad (3.8)$$

where $y, \dot{y}$ are position and velocity data from the demonstration, g is the attractor goal, τ is a time scaling parameter that provides a time-invariance property to DMPs, z and $\dot{z}$ are the time-scaled velocity and acceleration, x is the phase variable that indicates the progression in the movement, and f is the forcing function. The forcing function is defined with a linear combination of radial basis functions. The classical DMP has spawned a variety of extensions and adaptations that are tailored to different applications (Saveriano *et al.*, 2023).

A probabilistic framework to represent and learn movement primitives was introduced by Paraschos *et al.*, 2013 called **Probabilistic Movement Primitives** (ProMP). ProMPs represent movements as probability distributions over trajectories rather than deterministic func-

tions (e.g. DMPs), which reduces uncertainty and increases robustness by modeling the variability in the demonstrated movements. Additionally, ProMPs can be conditioned to reach desired via-points and blend multiple movement primitives into one for strong adaptability in movement representation. A ProMP is represented with a state distribution parameterized by θ as follows:

$$p(\mathbf{y}_t; \theta) = \int p(\mathbf{y}_t | \omega) p(\omega | \theta_\omega) \quad (3.9)$$

$$p(\mathbf{y}_t | \omega) = \mathcal{N}(\mathbf{y}_t | \Phi_t \omega, \Sigma_y) \quad (3.10)$$

$$p(\omega | \theta_\omega) = \mathcal{N}(\omega | \theta_\omega = \{\mu_\omega, \Sigma_\omega\}) \quad (3.11)$$

where $\mathbf{y}_t$ is the robot joint state at time t , Φ_t is a time-dependent basis matrix, and ω is a weight vector that is marginalized by introducing a distribution over the weights with θ .

Other movement primitive approaches for modeling low-level movements include Kernelized Movement Primitives (KMP) (Huang *et al.*, 2018b), which is a non-parametric solution that targets domains with high-dimensional inputs and enables extrapolation. Recently, ProDMP (Li *et al.*, 2023) was proposed, providing a framework that inherits properties from both DMPs and ProMPs.

High-Level Skill Learning

High-level task learning is a crucial aspect of interactive robot learning, especially for complex, long-horizon tasks involving multiple actions. Take the example of autonomous food delivery, which encompasses various sub-tasks like planning the route to restaurant, picking up food, and delivering it to the customer.

Hierarchical Task Networks (HTN) address this complexity by decomposing tasks into trees with task decompositions. In this framework, the state comprises relevant environmental features, and actions represent primitive tasks or sequences of tasks. HTNs are advantageous as they enable reasoning over abstractions, allow task decomposition, leverage previously acquired skills, and are human-readable due to their tree structure (Figure 3.4).

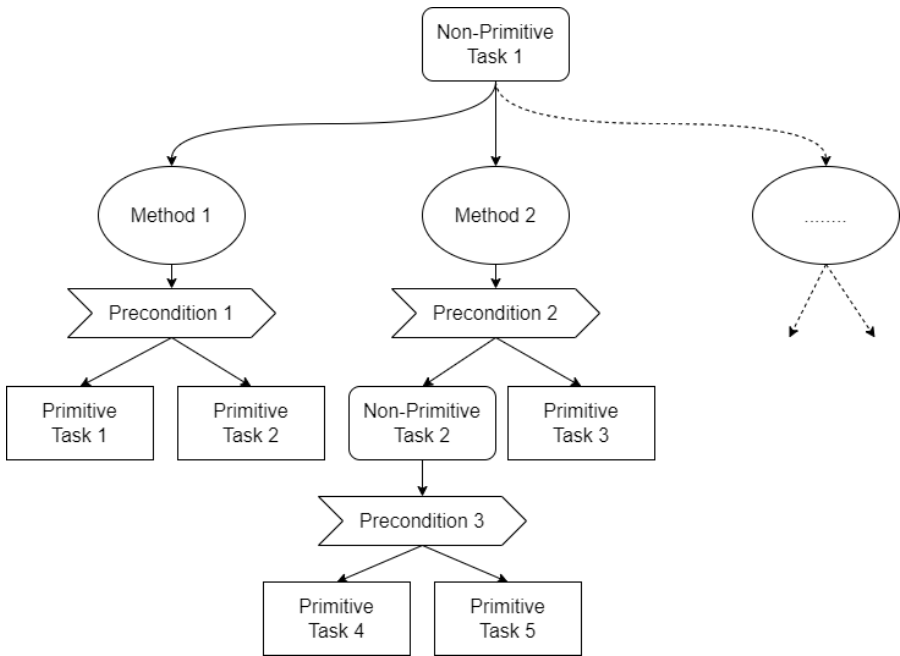


Figure 3.4: A hierarchical task learning network framework.

Mohseni-Kabir *et al.*, 2014 leverage the human-readable property of HTNs, framing the HTN specification as an interactive discourse between developers and agents. Task decomposition and composition patterns are learned through human interaction. Chen *et al.*, 2021a present an unsupervised learning approach for HTNs, learning from unannotated demonstrations. They draw an analogy between learning complex HTNs and series/parallel reduction of resistor networks in Electrical Engineering. The learned HTN can reconstruct tasks and predict necessary actions. French *et al.* French *et al.*, 2019 introduce Behavior Trees (BTs) as an alternative high-level task representation. BTs, composed of control and action nodes, offer real-time execution, modularity, and transparency. They effectively represent high-level tasks, demonstrated by instructing a Fetch robot in a household cleaning task.

In conclusion, high-level skill learning is a pivotal aspect of interactive robot learning, essential for tackling complex, multi-step tasks.

3.5 Learning from Human Feedback

Reinforcement learning has found extensive applications in the field of robotics (Kober *et al.*, 2013). However, our focus lies specifically on interactive reinforcement learning and works that involve learning from human feedback. The fundamental concept underlying these approaches is to leverage human feedback for purposes such as learning a reward function, shaping a reward function, molding a policy, directing the agent’s exploration process, or augmenting the agent’s value function (Arzate Cruz and Igarashi, 2020a).

3.5.1 Preference based Reinforcement Learning (PbRL) and Reinforcement Learning from Human Feedback (RLHF)

Preference-based reinforcement learning (PbRL) algorithms aim to address the reinforcement learning problem by utilizing preferences expressed by experts between states, actions, or trajectories. The objective is to learn a policy that aligns closely with the preferences provided by the expert. Therefore, a PbRL algorithm must discern the expert’s intentions and identify a policy that best matches them. To refine its understanding, these algorithms generate trajectories that are then assessed by the expert, allowing for exploration of the state-action space (Wirth *et al.*, 2017). Reinforcement learning from human feedback (RLHF) can be considered a specialized form of PbRL, where the preference labels are supplied directly by humans. This approach has garnered significant attention lately, demonstrating success in various domains such as robotic control and natural language processing tasks through techniques like fine-tuning pre-trained models (Ouyang *et al.*, 2022; Bai *et al.*, 2022). Such a methods allow a robot to align its behavior with the envisioned goals of the human trainer (Li *et al.*, 2021b).

Christiano *et al.*, 2017a introduced an approach involving learning a reward function from human feedback. This method generated multiple trajectories using an agent’s policy, sent pairs of trajectory segments to a human for ranking, and learned a reward function based on those preferences. The effectiveness of this method was demonstrated in solving complex RL tasks, including Atari games and simulated robot

locomotion, with feedback provided on less than 1% of the agent's interactions with the environment.

The Evaluative Reinforcement (TAMER) algorithm, introduced by Bradley Knox and Stone, 2008, allows the agent to model the human's reward function based on human feedback and choose actions predicted to yield the most reward. Warnell *et al.*, 2018 presented a modified version of the TAMER algorithm tailored for deep RL models, incorporating a deep convolutional neural network to model the human agent and a novel approach to handling user feedback. This approach minimizes a weighted difference between the human reward and the predicted value for each state-action pair individually. Furthermore, Arakawa *et al.*, 2018 introduced the DQN-TAMER algorithm, combining TAMER and Deep TAMER. This algorithm utilizes both environments and human binary feedback to shape the reward function. Experiments in a maze-like environment involved a robot receiving implicit human feedback in the form of facial expressions of the user. This mechanism was developed to address the imprecision and intermittency of explicit human feedback.

Celemin and Ruiz-del-Solar, 2015 proposed a Convergent Actor-Critic method, COACH (CORrective Advice Communicated by Humans). This framework enables non-experts to use numerical binary feedback to formulate policies through corrective suggestions. An extension of this work, named deep COACH, was introduced by Arumugam *et al.*, 2019. Deep COACH incorporates a deep neural network with a replay memory buffer and an autoencoder. Unlike the original COACH implementation, deep COACH utilizes raw pixels from the testbed as input.

To improve on suboptimal demonstrations in BC and IRL, Brown *et al.* (2019) proposed Trajectory-ranked Reward EXtrapolation (T-REX). T-REX is an PbRL method that learns to extrapolate intent of the demonstrator via suboptimal demonstrations that are ranked via preference labels. The resulting reward function can then be used in combination with deep RL to learn a policy that outperforms the demonstrator. One limitation of T-REX is that the algorithm assumes the preference labels are distributed according to the reward, which Knox *et al.* (2022) argues is a flawed assumption and suggests that users provide preferences based on regret of the behavior. Building on this idea,

works such as Hejna *et al.* (2023) and An *et al.* (2024) proposed PbRL methods without reward modeling. Both methods utilize a contrastive objective to learn a policy based on the provided preferences. Hejna *et al.* (2023) attempts to learn a regret-based model through their proposed contrastive objective. Another problem with learning reward functions through preferences is reward misidentification, where non-causal distractor features that are correlated with causal features of the true reward function influences the learned reward function. Tien *et al.* (2023) provides a series of sensitivity and ablation studies on several benchmark domains with PbRL methods to highlight the reward misidentification problem in introducing misaligned behaviors.

3.5.2 Reward Shaping

Reward shaping is a strategy to influence the behavior of a learning agent by adjusting its reward function to align with the preferences of the RL designer or the human user, expediting the learning process. Thomaz *et al.*, 2006 proposed a technique tailored for non-expert human trainers. These trainers could provide a positive or negative numerical reward to the agent, modifying its subsequent actions. In the event of negative feedback, the agent would attempt to retract its last action to improve its score. An important observation from this work was the phenomenon of users providing anticipatory feedback, assuming that their input guides the bot in future states. Understanding these teaching strategies becomes crucial in interactive RL, underscoring the necessity to comprehend user preferences in teaching agents and enhance the feedback mechanism to foster trust and improve its quality and quantity.

An alternative approach to reward shaping involves the manual creation of heuristic functions incentivizing the agent to take specific actions in particular environmental states (Rosenfeld *et al.*, 2018, Bianchi *et al.*, 2014). These heuristic functions provide automatic feedback to the learning agent. Experiments indicate that using heuristic functions as input for interactive RL is a natural method for infusing human knowledge into the algorithm. However, a drawback is that heuristic functions are challenging to construct and demand programming skills.

Deep versions of reward shaping methods, benefiting significantly

from function approximation, minimize the feedback required to achieve good results. This advantage arises from the generalization of user feedback across similar states, allowing human knowledge to be injected into multiple analogous states rather than just one. While reward-shaping methods effectively accelerate the learning process in environments with sparse rewards, challenges such as "credit allocation" and "reward hacking" are acknowledged (Arzate Cruz and Igarashi, 2020a). The former relates to mapping human rewards to corresponding actions, particularly in rapidly changing environments where human feedback may lag. The latter involves the risk of the agent achieving maximum rewards through unexpected methods not anticipated by humans.

3.5.3 Exploration Guidance

The guided exploration process aims to enhance this exploration by integrating human knowledge, directing the agent towards states with high rewards.

Thomaz and Breazeal, 2008b conducted an experiment using the game Sophie's Kitchen to assess the impact of human guidance on minimizing the agent's action space and accelerating learning efficiency. Results indicate that leveraging human prior knowledge effectively assists the agent in reducing low utility attempts, proving more efficient than relying solely on scalar reward functions. Suay and Chernova, 2011 introduced a variation where users guide exploration by identifying goal states in the environment. The bot adjusts its policy to favor actions leading to the specified goal. Experiments showed that relying solely on exploration guides from users yielded optimal results and reduced the required amount of user feedback.

While utilizing human feedback for guidance in interactive RL algorithms proves effective in terms of performance, successful implementation necessitates a model of the environment. The interactive RL algorithm must discern the actions required to reach the state proposed by the user, presenting a challenge in complex environments where precise models are challenging to create. Moreover, these algorithms usually necessitate training by humans, entailing significant domain expertise and active involvement in the training process from the human

trainer.

3.5.4 Augmented Value Functions

The value function-based method involves estimating future rewards with the aim of maximizing the final reward at the end of a task by incorporating human knowledge (Arzate Cruz and Igarashi, 2020b). These methods merge the value representing human preferences with the agent's obtained value from the environment, facilitating the learning process.

A proposed approach, Human-Agent Transfer (HAT) (Taylor *et al.*, 2011), combines human preference and agent value. The algorithm generates a strategy based on recorded human trainer preferences, shaping the Q-value function. This process ensures a stable reward for the state-action pair in the Q-learning process. Another interactive RL algorithm, RLfD (Brys *et al.*, 2015), integrates expert demonstrations as input. The demonstrations are used to create a potential-based piecewise Gaussian function, assigning high values to state-action pairs demonstrated by the expert and 0 values where no demonstrations occurred. This function serves a dual purpose in biasing the exploration process of a $Q(\lambda)$ -learning algorithm. First, the Q-function of the RL algorithm is initialized with the potential-based function values. Subsequently, the potential-based function acts as a shaping mechanism complementing the reward function from the environment. This dual biasing mechanism aims to leverage human knowledge from the expert throughout the entire learning process.

While the augmented value function method shows promise in minimizing the required human feedback, there is a limited number of studies exploring its effectiveness (Li *et al.*, 2021b).

3.6 Leveraging Natural Language for Interactive Robot Learning

As robots venture into non-expert-populated environments, providing user-friendly interfaces becomes imperative. Conventional interfaces such as keyboards, mouse, or touchscreens demand training complexity for commanding intricate robotic behavior (Zucker *et al.*, 2015). In

contrast, natural language (NL) interfaces offer intuitive ways for users to interact with and control robots, requiring minimal training (Tellex *et al.*, 2020a).

3.6.1 Language Grounding Based Methods

Language grounding refers to the process of associating linguistic symbols, such as words or phrases, with corresponding entities, concepts, or actions in the real world. It involves establishing a meaningful connection between language and the external environment, enabling robots to interpret language based on their sensory perceptions and interactions with the world. Grounding language ensures that linguistic expressions carry contextually relevant meanings and facilitate effective communication between humans and robots.

Instruction following (Mei *et al.*, 2016, Arumugam *et al.*, 2017) was one of the first areas in which progress was made. This supervised problem involves executing a sequence of actions based on a given natural language command. A common approach treats language commands and action sequences as a machine translation problem, employing RNN-based Sequence to Sequence models (Mei *et al.*, 2016). Some variations of the problem focus on learning the grounding from natural language to sub-goals or goals, simplifying the complexity of language-guided actions (Gopalan *et al.*, 2018, Andreas and Klein, 2015).

Efforts have also been made to mitigate the need for extensive supervision by transforming the instruction following problem into a reinforcement learning (RL) problem. Branavan *et al.*, 2009 initially approached this by employing classical policy gradient methods. Recent advancements extend this approach to more complex environments with visual inputs (Das *et al.*, 2018, Hermann *et al.*, 2017, Misra *et al.*, 2017). A prevalent strategy involves modeling the agent and its environment as a MDP and representing the instruction as the state of the environment. These RL-based agents demonstrate capabilities such as answering questions about object properties or navigating to objects in simulated environments. Silva *et al.*, 2022a leveraged natural language specifications, enabling users to describe new tasks for the robot and facilitating the robot's reasoning about task similarities and

differences through language-based task embedding thereby speeding up the learning process.

The task of associating object names with their corresponding visual representations is an active research domain in deep learning. A prevalent approach, initially introduced in (Karpathy and Fei-Fei, 2017), involves aligning vectorized object representations in images with vectorized sentence representations describing those objects. These methods have been applied to physical robots in scenarios such as object retrieval, with training often conducted on simulated images (Hatori *et al.*, 2018, Shridhar *et al.*, 2020). In a related study, Blukis *et al.*, 2018 focused on end-to-end learning, aiming to establish a direct mapping between navigation instructions and low-level control actions. This process leverages the robot’s sensor input and control actions, translating language instructions into corresponding control actions. The study successfully demonstrated the practical implementation of this approach on physical robots.

Recent approaches with language grounding include using natural language descriptions of various tools to generate feature representations that can be further used for language-conditioned meta-learning (Ren *et al.*, 2023b), using natural language feedback as a reward signal (Sumers *et al.*, 2021), and specifying safety constraints for RL with natural language (Yang *et al.*, 2021).

3.6.2 Foundation Models

Foundation models, characterized by their vast number of parameters in the billions, represent a significant advancement in machine learning. These models, typically based on Transformer architectures (refer to Section 3.1), undergo pre-training on extensive datasets collected from the internet. This process imbues them with a wealth of world knowledge, thereby enhancing their generalization capabilities (Firoozi *et al.*, 2023). Large language models (LLMs) constitute a subset of foundation models, specifically trained on large-scale textual data. Prominent examples include GPT (Generative Pre-trained Transformer), BERT (Bidirectional Encoder Representations from Transformers), and Gemini. On the other hand, Vision-Language-Action models (VLAs) are

foundational models that integrate information from visual inputs (such as images or videos), textual inputs (such as natural language descriptions), and action outputs (such as commands or decisions) to enable robots to perceive, comprehend, and interact with the environment.

Recent advancements in LLMs have prompted increased interest in VLA models (Zeng *et al.*, 2023). RT-1, introduced by Brohan *et al.*, 2022, was a preliminary groundbreaking model that leveraged transformer architectures for robotics, but it had limitations that impacted its performance. Specifically, RT-1 faced challenges with generalization and robustness in unstructured environments, as well as issues with handling diverse and novel tasks. The model sometimes failed to adapt effectively to new situations or demonstrated brittle behavior when faced with unexpected changes. RT-2, introduced by Brohan *et al.*, 2023b, was built on RT-1’s foundation, addressing some of these shortcomings by incorporating multi-modal inputs and enhanced learning mechanisms, leading to improved adaptability and performance across a broader range of robotic tasks. RT-2 presents an end-to-end model capable of mapping robot observations to actions, establishing a VLA model. In contrast to previous methods that involved integrating vision-language models (VLMs) into robot policies or designing novel architectures, RT-2 introduces VLA instantiation through fine-tuning. The approach fine-tunes a large-scale VLM (Chen *et al.*, 2023b, Driess *et al.*, 2023) pre-trained on an extensive network-scale dataset, transforming it into a VLA model. Actions are integrated as text tokens directly into the pre-training dataset, creating multi-modal sentences that unify robot actions and natural language responses (Driess *et al.*, 2023). These multimodal statements enable robots to respond to commands through observation, outputting corresponding actions. This processing mirrors the decoding process in LLMs, where action-related tokens are decoded and converted into robot actions. VLA models hold significant potential for enhancing the generalization capabilities of robots.

3.6.3 Reward Shaping and Refinement Methods

Another domain of exploration involves leveraging natural language for the creation and refinement of reward functions, providing an intuitive

means to articulate objectives that can be subsequently optimized to induce desired robot behaviors. In this context, MacGlashan *et al.*, 2015 introduced a framework that translates natural language commands into reward functions through demonstrations of these commands being executed in the environment. Additionally, Goyal *et al.*, 2019 presented a reward shaping technique that maps free-form natural language instructions to intermediate rewards, that can be integrated into standard reinforcement learning algorithms. Wang *et al.*, 2023 prompted a LLM to directly formulate a reward function for legged locomotion, while Yu *et al.*, 2023 directed an LLM to define reward parameters, which were subsequently optimized to achieve a variety of robotic tasks.

In summary, interactive robot learning has made remarkable progress in harnessing natural language for human-robot interaction. Through endeavors such as language grounding, reward shaping, and the adoption of foundational models, researchers have augmented robotic capabilities and improved their ability to generalize. With natural language interfaces increasingly central to human-robot interactions, these advancements hold the potential to foster more intuitive, user-friendly, and capable robotic systems.

3.7 Adaptive and Adaptable Interactive Robot Learning Methods

Adaptive methods refer to algorithms capable of autonomously acquiring necessary behavioral and cognitive skills for task performance through interactions with the environment (Nolfi, 2021). These methods prioritize continuous adaptation and modification of the robot's behavior and function in response to contextual changes, such as varying conditions or tasks, without requiring human intervention. Adaptable methods in robotics refer to approaches where humans play a central role in modifying the behavior or capabilities of robots. Unlike adaptive methods, which enable robots to autonomously learn and improve their performance over time, adaptable methods involve direct human intervention in guiding or directing the changes made to the robot's behavior or function. This intervention can take various forms, such as providing explicit instructions, adjusting parameters, or redesigning algorithms based on human expertise or preferences. Adaptable methods thus facili-

tate a more direct and controlled modification process, allowing humans to tailor the robot's behavior to specific requirements or preferences, but may result in a larger workload on the end user.

We will go over some main works that have targeted the problem of adaptive and adaptable learning method for robotics.

3.7.1 Adaptive Learning Methods

Continual or Incremental Learning

Continual learning, also known as lifelong learning or incremental learning, are adaptive methods that allows the robot to acquire and adapt to new knowledge and experiences over time without catastrophically forgetting previously learned information. In continual learning, the system incrementally learns from a stream of data or tasks, integrating new information while preserving previously acquired knowledge. The goal is to enable the system to continually improve its performance and adapt to changing environments or requirements without the need to retrain from scratch (Lesort *et al.*, 2020). Continual learning is particularly important in dynamic and evolving domains where data distributions or task requirements may change over time.

Gao *et al.*, 2021 present an approach for continual imitation learning that relies on deep generative replay (Shin *et al.*, 2017) and action-conditioned video prediction to create state and action trajectories for past tasks. These generated trajectories, or pseudo-data, are combined with current task demonstrations to train the policy network guiding the robot's actions. However, the authors identify a challenge in generating high-quality video frames for an extended sequence of tasks.

Auddy *et al.*, 2023 introduce a novel approach for continual learning from demonstrations using hypernetworks. This method generates parameters for neural ordinary differential equation solvers, and its effectiveness is validated through real-world robotic tasks that involve changes in position and orientation, demonstrating the practical applicability of this continual learning strategy in robotics.

However, continual learning algorithms may have to deal with problems such as imbalanced or scarce data (Sprechmann *et al.*, 2018),

catastrophic forgetting (French, 1999), or data distribution shifts (Gepnerth and Hammer, 2016).

Intrinsic Motivation

Intrinsic motivation, as discussed by (Schmidhuber, 2010), plays a crucial role in promoting long-term and open-ended learning in robotics. Intrinsic motivation for robot learning refers to the use of internal drives or rewards within the robot itself to guide its learning process, rather than relying solely on external rewards or supervision. These intrinsic motivators can include curiosity, novelty-seeking, competence, and the desire for autonomy. By incorporating intrinsic motivation mechanisms into robotic systems, researchers aim to promote self-directed learning, exploration, and adaptation, which can lead to more robust and flexible behaviors in diverse environments. This approach is inspired by theories from psychology and neuroscience that suggest intrinsic motivation plays a crucial role in human learning and development.

One strategy for implementing intrinsic motivation involves using a reward function comprising two scalar real-valued components: the standard external reward tied to achieving specific goals, and the intrinsic reward, which stems from the novelty of behaviors or observations. These components are typically combined using a function, such as simple addition. This approach not only promotes ongoing learning but also enriches the robots' capacity to evolve and adapt over time, as evidenced in studies by Baldassarre and Mirolli, 2013 and Cangelosi and Schlesinger, 2015.

Another approach to intrinsic motivation, proposed by Mohamed and Rezende, 2015, involves defining empowerment using mutual information as an internal drive. This method utilizes a variational lower bound on mutual information, coupled with convolutional networks for processing visual input streams, to develop a stochastic optimization algorithm. This algorithm enables scalable information maximization and empowerment-based reasoning directly from pixels to actions.

Unsupervised reinforcement learning (URL) serves as another approach for intrinsic motivation in robotic systems. Unlike traditional RL methods that rely on external reward signals, URL algorithms aim

to discover and learn policies without explicit rewards. Instead, they focus on maximizing some approximation of mutual information, akin to representation learning in supervised settings. By leveraging URL, robots can autonomously explore and discover novel skills and behaviors in their environment, thereby fostering intrinsic motivation and continuous learning. These approaches have shown promise in enabling robots to acquire diverse skills and adapt to various tasks without the need for explicit rewards, thus facilitating self-directed learning and exploration as evidenced by Eysenbach *et al.*, 2022.

In Pathak *et al.*, 2017, the authors propose a self-supervised approach grounded in curiosity-driven exploration. Here, the agent learns to forecast the outcomes of its actions via a self-supervised learning objective, thereby incentivizing exploration of the environment to unearth new skills.

Nguyen and Oudeyer, 2014 introduced a method that combines active, intrinsically motivated learning with imitation learning, enabling robots to learn high-dimensional sensorimotor inverse models efficiently. This approach, tested with a robot arm learning to use a flexible fishing line, enhances the robot's ability to produce diverse outcomes and refine control in large spaces. Nguyen and Oudeyer, 2012 developed an interactive learning system that blends internally guided learning with social interaction, allowing robots to learn motor skills with minimal reliance on human guidance and adapt flexibly in human interactions. Qureshi *et al.*, 2018 proposed a reinforcement learning framework motivated intrinsically, where the robot gains rewards through an action-conditional predictive model. This method led to robots learning social skills from real, uncontrolled human-robot interaction experiences, showing an enhanced ability in making human-like decisions.

In conclusion, intrinsic motivation enables robots to learn autonomously and adapt effectively to new challenges and environments. This approach fosters continuous skill development and enhances the robot's ability to handle complex tasks.

3.7.2 Adaptable Learning Methods

Curriculum Learning

Curriculum learning in robotics is a training strategy inspired by educational practices, where learning tasks are organized in a structured sequence of increasing complexity or difficulty. This approach involves presenting the robot with a series of tasks or learning objectives arranged in a curriculum, starting from simpler tasks and gradually progressing to more challenging ones (Mendoza Barrenechea, 2023). By following this curriculum, the robot learns foundational skills or concepts before tackling more complex problems, allowing for a more systematic and efficient learning process. Curriculum learning aims to enhance the robot's learning performance, adaptability, and generalization capabilities by providing a scaffolded learning experience that builds upon prior knowledge and experiences.

Karpathy and Panne, 2012 proposed a method that involves a two-level curriculum: at the high level, it specifies the order in which different skills should be learned, while at the low level, it defines the learning process within each skill. The study focused on developing integrated motor skills for a planar articulated figure capable of executing parameterized hops, flips, rolls, and acrobatic sequences. Moreover, the proposed curriculum could be customized to generate individualized motor skill sets for articulated figures of varying proportions. Qin *et al.*, 2019 used curriculum learning to teach a six-legged robot to walk on various terrains. In Yu *et al.*, 2018, curriculum learning is applied to train a humanoid robot to walk. The training starts in a simple environment and progressively advances to more complex settings.

Despite their benefits in speeding up the learning process, these methods come with notable drawbacks. They depend on expert knowledge to define the learning stages and introduce new hyperparameters, such as transition thresholds, which are challenging to optimize. Furthermore, they presume that problems can be broken down into sub-problems of increasing difficulty on a single axis, a presumption that may not always be applicable in complex real-world situations (Nolfi, 2021). In order to address these weaknesses, automated curriculum learning is an active area of research. This approach aims to automatically define the

curriculum without relying on human experts (Portelas *et al.*, 2020).

Active Learning

Active learning in robotics refers to a methodology where a robot intelligently selects specific data points for labeling by an expert or an oracle, such as a human (Settles, 2009). Unlike passive learning, where the robot passively receives labeled data or demonstrations, active learning involves the robot actively engaging with its environment to solicit informative feedback. The selection criteria for data points or tasks are typically based on measures of uncertainty, informativeness, or expected learning gain (Wallkötter *et al.*, 2021). By actively choosing the most informative data points or tasks for labeling or exploration, the robot can effectively reduce the amount of labeled data needed to achieve a desired level of performance, thereby improving the efficiency and effectiveness of its learning process, enabling rapid adaptation to new environments.

Chernova and Veloso, 2009 proposed Confidence-Based Autonomy (CBA) algorithm to facilitate policy learning from demonstration through interactive human-agent collaboration. CBA comprised two key components: Confident Execution, which autonomously identified states requiring demonstration, requested demonstrations from human teachers, and learned policies based on acquired data; and Corrective Demonstration, which allowed teachers to correct agent mistakes through additional demonstrations to enhance policy and task performance. Through evaluation in a complex simulated driving domain, CBA demonstrated superior learning performance compared to individual components, striking a balance between demonstration efficiency and learning accuracy.

Maeda *et al.*, 2017 introduced an approach where robots actively request information to incrementally learn movement primitives. Using a Gaussian Process to generate a trajectory, the robot can extrapolate from a library of demonstrations with defined confidence margins. This process allows the robot to generalize from minimal demonstrations and importantly, to identify when such generalizations are reliable. Real-world applications demonstrated a robotic arm requesting specific demonstrations from users to enhance its reaching capabilities, even

becoming proficient in reaching for objects without prior demonstrations, by learning from related ones.

Baranes and Oudeyer, 2013 proposed an active learning architecture for complex, high-dimensional spaces. This method efficiently learns policies through active goal exploration and has shown to surpass direct policy sampling in learning speed. It involves an autonomous exploration of increasingly complex tasks, guided by competence progress. This approach was successfully tested in various robotic contexts, such as learning inverse kinematics in a redundant robotic arm, developing omnidirectional locomotion in a quadruped robot, and controlling a fishing rod with a flexible wire.

Bullard *et al.*, 2019 introduced an active learning approach that employs imitation learning to empower an agent operating in constrained environments to simultaneously consider its internal learning objectives and externally imposed environmental constraints within its objective function. Experiments demonstrated the efficacy of environmentally-aware reasoning in outperforming established active learning baselines under constrained conditions.

Grollman and Jenkins, 2007 applied active learning in teaching robots new tasks, including mimicry and ball seeking. These various implementations demonstrate the versatility and efficiency of active learning in robotics, highlighting its potential in enhancing robot adaptability and skill acquisition.

Evolutionary Methods

Evolutionary methods in robotics, as detailed by Floreano *et al.*, 2008, involve an approach that mirrors natural evolution, a field known as evolutionary robotics. This process starts by creating a population of artificial genotypes, which are essentially blueprints for the robots' characteristics or phenotypes.

The process begins with the random generation of genotypes for the initial robot population. Each genotype corresponds to a robot, which then interacts with its environment in one or more evaluation sessions. The performance of these robots is quantified using a fitness score, typically a scalar value. During these interactions, the characteristics

of the robots might change due to learning or developmental processes. The most successful robots, in terms of fitness, are then selected to reproduce. Their offspring are created with variations, introduced through mutations or other genetic operations. This cycle continues, generation after generation, until a predetermined number of evaluations have been completed. The human involvement is primarily in crafting the fitness function, which could be, for example, measuring the distance a robot travels in a given episode. This evolutionary approach allows for the development of adaptable methods through a process that closely emulates natural selection and evolution. This concept is exemplified in studies like the one employing an evolutionary algorithm for visual tracking, which starts with large stationary targets and progresses to small moving ones (Harvey *et al.*, 1994).

3.8 Mixed-Initiative Interaction and Shared-Autonomy

This section on Mixed-Initiative Interaction and Shared-Autonomy delve into advanced robotics interaction strategies that emphasize collaboration between humans and robots. We highlight the distinction between these two concepts and discuss the role of interactive robot learning in producing adaptive collaboration strategies.

3.8.1 Mixed-Initiative Interaction

Mixed-initiative interaction (defined generally in Section 2.2), in the context of robotics, refers to an interaction strategy where both the human operator and the robot have the ability to initiate actions or decisions (Jiang and Arkin, 2015; Seraj *et al.*, 2023). The human operator typically provides high-level guidance, insights, and decisions, while the machine contributes precision, and consistency, leveraging the strengths of each agent to efficiently accomplish a shared goal. Jiang and Arkin (2015) proposed the following definition for mixed-initiative in human-robot interaction:

A collaboration strategy for human-robot teams where humans and robots opportunistically seize (relinquish) initiative from (to) each other as a mission is being executed, where initiative is an element of the mission

that can range from low-level motion control of the robot to high-level specification of mission goals, and the initiative is mixed only when each member is authorized to intervene and seize control of it.

Under this definition, a system is not considered a mixed-initiative system if there is no overlap between the initiatives of the robot and the human, or if the initiative is heavily one-sided (e.g. only human intervention). However, systems in the literature that are classified as mixed-initiative rarely meet this criteria (Murphy *et al.*, 2000; Finzi and Orlandini, 2005; Loizou and Kumar, 2007; Clare *et al.*, 2012), as true mixed-initiative systems require robust safety measures to protect the human operator during robot intervention.

Due to the complex nature and diversity of human behaviors, modeling the interactions within a mixed-initiative team is a major challenge (Natarajan *et al.*, 2023a). Suboptimality in human behavior is one such aspect of human modeling that can degrade the performance of mixed-initiative teams if not taken into consideration (Natarajan *et al.*, 2023a; Nikolaidis *et al.*, 2017a; Tabrez and Hayes, 2019; Kwon *et al.*, 2020). Interactive robot learning can alleviate these issues by iteratively improving and adapting from the human partner’s feedback data, helping the robot understand and predict human intention over time. For example, Guo *et al.* (2018) proposed a mixed-initiative task and motion planning strategy with inverse reinforcement learning to learn the operator’s task preferences and reduce task constraint violations. Natarajan *et al.* (2023b) sought to maximize performance in mixed-initiative teaming under suboptimality and proposed a human-robot model optimized through a variant of BA-POMCP (Katt *et al.*, 2017).

3.8.2 Shared-Autonomy

Similar to mixed-initiative, shared autonomy, or shared control (Abbink *et al.*, 2018; Chiou *et al.*, 2021a)), is a term used generally to describe collaborative systems with any level of interactivity between the human and the robot (Ishiguro, 2017; Schilling *et al.*, 2016). The term has additionally been used specifically to describe systems where the robot’s control policies involve mixing of both human and robot input (Dragan

and Srinivasa, 2013; Muelling *et al.*, 2017; Reddy *et al.*, 2018; Chiou *et al.*, 2021a). In this article, we will use shared control to describe the latter, as the bulk of interactive robot learning research focus on learning optimized input mixtures (Dragan and Srinivasa, 2013; Jun Jeon *et al.*, 2020; Reddy *et al.*, 2018). In that sense, shared control can be considered a subset of mixed-initiative systems that focuses on low-level motion control. Shared control systems are often deployed to assist human teleoperators (Selvaggio *et al.*, 2022; Gottardi *et al.*, 2022), with the intention to reduce the operator's cognitive load (Boessenkool *et al.*, 2012).

Adaptive control is a key component of shared control, dynamically adjusting the balance of control between the human and the robot according to the specific context to maximize overall performance (Nikolaidis *et al.*, 2017b). Important questions that arise are "how" and "when" to adjust this balance. Interactive robot learning can be an effective solution for answering these questions in shared control, with similar reasoning that were presented for mixed-initiative systems. Human idiosyncrasies lead to unique differences in how an individual performs a task with the robot. To address this, (Schrum *et al.*, 2022b) proposed a meta-learning framework with imitation learning to adapt to individual users in shared control tasks. Additionally, reinforcement learning has been used to implicitly learn the operator's goals to provide assisted control (Reddy *et al.*, 2018; Schaff and Walter, 2020). Jun Jeon *et al.* (2020) proposed an assistive manipulation framework that dynamically adjusts the user's granularity of control by learning latent actions. All of these methods adapt and correct for the user's preferences and sub-optimality in some way, highlighting the utility of interactive learning in shared control to increase overall system performance.

3.8.3 Takeaways for Mixed-Initiative and Shared Autonomy

In conclusion, the integration of interactive robot learning in mixed-initiative and shared control systems can greatly enhance the cooperative efficiency of the team. Through an iterative learning process, robots can finely tune their responses based on human feedback, continually adapting to individual user preferences and compensating for human

suboptimal behaviors. As robot capabilities continue to evolve, learning methods need to place a larger focus on human-centered approaches to maintain strong cohesion and fluidity in the human-robot team.

3.9 User Expectations and Trust

Building trust is a pivotal factor in determining the successful collaboration of interactive agents, whether its human-human, human-robot, or robot-robot collaboration (Hancock *et al.*, 2011a; Sanders *et al.*, 2011). Despite numerous studies on the topic of trust, there is no consensus on a single definition, as defining trust is heavily context-dependent (Cameron *et al.*, 2015). In the context of automation, Lee and See (2004) provided a commonly used definition of trust on an agent (human or automation) as *the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability*. The role of trust in human-robot interaction (HRI) is complex, as robots are perceived to have a degree of self-governance, in contrast to automation, where the system behavior's is pre-programmed (Lewis *et al.*, 2018). In this section, we first provide some context for why trust is necessary for interactive robot systems, platforms primed for interactive learning, then discuss how building trust can impact the effectiveness of interactive learning.

3.9.1 Effect of Trust on Interactive Robots

Widespread adoption of interactive robots requires trust of the user (Hancock *et al.*, 2011b; Pinxteren *et al.*, 2019; Chi *et al.*, 2023b), and in some cases, trust has been shown to be the main factor in robot acceptance (Parasuraman and Riley, 1997). When individuals trust a robot, they are more likely to engage with and follow the robot's instructions (Muir and Moray, 1996; Lewandowsky *et al.*, 2000). Lack of trust in a robot leads to under-reliance (Parasuraman and Riley, 1997; Lyons and Stokes, 2012), where a user disregards sound advice or actions from a robot. Conversely, misplaced trust in the robot's capabilities leads to over-compliance (Parasuraman and Riley, 1997; Parasuraman and Manzey, 2010; Lyons and Stokes, 2012). Over-compliance can lead to

disastrous outcomes if the user accepts a poor decision made by the robot (e.g. faulty autopilot (Mosier *et al.*, 1997)). Therefore, user trust in the robot should be well-calibrated to the robot’s capabilities for optimal performance of a human-robot system (Lewis *et al.*, 2018).

The importance of trust calibration has led to numerous studies investigating factors that influence user trust within human-robot teams (Okamura and Yamada, 2020; Chiou *et al.*, 2021b; Herse *et al.*, 2023; Lee *et al.*, 2023). Khavas *et al.*, 2020 presented a review of techniques for trust modeling and argued that trust must first be modeled before it can be properly calibrated. Following this line of reasoning, researchers have developed a variety of trust models in conjunction with robotic systems that adapt to these models. For example, Saeidi *et al.* (2016) proposed a two-way mathematical trust model (i.e. human-to-robot and robot-to-human) based on agent performance and a mixed-initiative control scheme for bilateral teleoperation of mobile robots. Similarly, Hannum *et al.* (2023) proposed a motion planner utilizing a two-way trust model that dynamically adjusts trust levels based on human motions, past interactions, and each agent’s performance. While the community has invested significant effort into modeling trust, these models are generally domain-specific, highlighting a need for a general trust model that can be used across multiple robotic tasks (Khavas *et al.*, 2020).

3.9.2 Utility of Trust in Interactive Robot Learning

In the field of interactive robot learning, considerable attention has been placed on adaptive robot behavior conditioned on learned trust models. Examples of learned trust models for HRI include OPTIMo (Xu and Dudek, 2015), an online, dynamic Bayesian network that infers beliefs over latent trust states, Partially Observable Markov Decision Process (POMDP) frameworks embedded with learned belief over the latent state of human trust (Chen *et al.*, 2018a; Chen *et al.*, 2020d; Natarajan *et al.*, 2023b), and a dynamic trust model utilizing recurrent neural networks (RNN) for diverse robotic tasks (Soh *et al.*, 2020). A major problem with learned systems is the lack of transparency in the agent’s decision-making process, and transparency is an important factor in

building trust (Lewis *et al.*, 2021; Setchi *et al.*, 2020a). Explainable AI is a rapidly growing field attempting to address this issue, with many recent works focused on robotics applications (Tabrez and Hayes, n.d.; Paleja *et al.*, 2023b; Das *et al.*, 2021a; Chen *et al.*, 2020b). Since trust is easily lost and hard to regain (Esterwood and Jr, 2023; Alhaji *et al.*, 2021), learning-based systems need to actively manage the user's expectations during the learning process as the robot's capabilities change.

Trust plays a critical role in both classical and learning-based interactive robot systems. Trust needs to be well-calibrated to avoid user under-reliance and over-compliance. As interactive learning methods continue to evolve, managing user expectations and maintaining an appropriate level of trust will be paramount for the successful deployment and acceptance of these systems in the real-world.

3.9.3 Design Guidelines and Conclusion

Below, we outline a series of design guidelines aimed at establishing interactive robot learning systems. These guidelines are crafted to enhance the development and implementation process, ensuring that these systems are both effective and user-centric:

1. **Selection of Interaction Modality and Level:** Choose the interaction modality based on the robotic setup, alongside user needs and preferences. Considerations such as the task at hand, the volume of data required, and the ease of providing demonstrations are crucial in determining the most suitable modality.
2. **Selection of Algorithms and Models:** Tailor the choice of algorithms and models to the specific use-case. For instance, in situations where demonstration data is limited, techniques like movement primitives may be appropriate. Conversely, for scenarios demanding high robustness under diverse conditions, leveraging large foundational models trained on extensive datasets is advisable.
3. **Adapting Robot Behaviors:** When there's a need for robots to adjust their behaviors to new users or contexts, readers are

encouraged to consult Section 3.7 for methodologies that align best with the particular scenario at hand.

4. **Ensure Robust Safety Measures for Mixed-Initiative Systems:** In mixed-initiative interactions, where both the robot and user can initiate actions, robust safety measures are essential to protect the user during robot interventions. This includes mechanisms for emergency stop, user override, and conflict resolution to ensure that the collaboration remains safe and effective.
5. **Actively Manage User Expectations:** As the robot's capabilities evolve through interactive learning, communicate these changes effectively to the user. This involves setting realistic expectations about the robot's current performance level and potential limitations. Regular feedback and updates on the learning progress can help users adjust their trust and reliance on the robot appropriately.

In conclusion, interactive robot learning stands at the confluence of robotics, artificial intelligence, and human-robot interaction, heralding a future where robots and humans collaborate more seamlessly and intelligently. This chapter has navigated through the multifaceted landscape of interactive robot learning, shedding light on the advancements, challenges, and methodologies that shape this dynamic field. From understanding the foundational principles of interaction modalities and learning representations to exploring the nuances of trust, mixed-initiative interaction, and shared autonomy, we have underscored the importance of user-centric approaches and adaptive learning mechanisms. This discussion paves the stage for the next chapter, where we will explore feedback mechanisms that aid human in interactive and explainable robot learning.

4

Feedback Mechanisms Aiding Humans in Interactive and Explainable Robot Learning

Feedback Mechanisms in interactive robot learning refer to the systems and processes through which a robot communicates back to the human user, providing information about its actions, decisions, or learning progress. These mechanisms are essential for creating a dynamic, two-way interaction where the robot not only receives input and instructions from the human but also actively contributes to the learning experience by offering insights, suggestions, or clarifications. By utilizing feedback mechanisms, robots become active participants in the learning process, facilitating a more immersive, contextual, and collaborative learning environment that aligns with the principles of situated learning theory.

This section aims to shed light on the numerous ways in which robots can actively contribute to the learning experience, by focusing on feedback mechanisms that are emitted from the robot to the human. Drawing upon the principles of situated learning theory, which emphasizes learning as an interactive process occurring within a specific context or environment, this section explores how robotic feedback systems can foster an immersive and contextual learning environment (Clancey, 1995; McLellan, 1996). We underscore the importance of designing robot learning systems that are not merely passive recipients of human

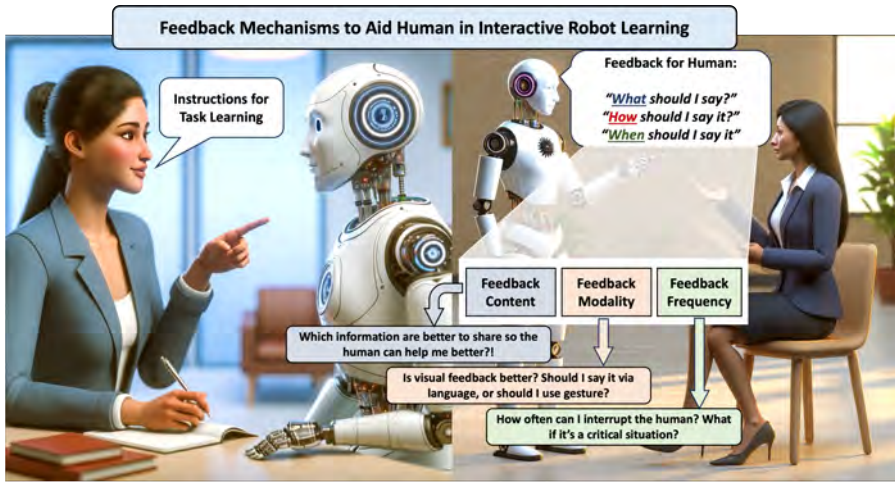


Figure 4.1: Feedback content, frequency, and modality can be considered the three major categories of the primary challenges on the topic of feedback mechanisms for aiding humans in interactive robot learning.

instruction but active participants in a collaborative learning journey, echoing the situated learning theory's emphasis on participation and collaboration in authentic contexts (Natarajan *et al.*, 2023a; Konan *et al.*, 2022).

While a significant portion of prior research is geared toward enhancing the robot's learning experience from human guidance (Chernova and Thomaz, 2014; Wang *et al.*, 2019a), an often-overlooked aspect is how robots can contribute to this collaborative learning process. In interactive robot learning scenarios, the quality and effectiveness of the learning process depend not only on the human's ability to teach but also on the robot's ability to assist and facilitate its own learning through effective communication (Sheridan, 2016; Natarajan *et al.*, 2023a). This is particularly important for establishing a shared understanding between the human and the robot, a phenomenon that is increasingly relevant in coordinated and cooperative interactions (Salas *et al.*, 1992; MacMillan *et al.*, 2004; Natarajan *et al.*, 2023a; Seraj, 2022).

Feedback mechanisms implemented in robots serve as instrumental tools to enhance this *dual-sided* learning experience (Breazeal *et al.*, 2016; Nocentini *et al.*, 2019). The term *dual-sided* in this context un-

derscores the reciprocal nature of the learning experience, where both humans and robots are actively learning from and adapting to each other. This mutual exchange is facilitated through particular and often domain-dependent feedback and communication mechanisms which are critical for effective collaboration and the ongoing development of more intuitive and capable robotic systems. These mechanisms (i.e., Figure 4.3), ranging from simple visual or auditor cues (e.g., red alarm light or a buzz sound) to complex dialogues (e.g., textual instructions, scene description, or language conversation), aim to provide the human demonstrator with information that helps optimize the teaching process (Hindriks and Liebens, 2019; Breazeal *et al.*, 2016; Breazeal *et al.*, 2004; Nocentini *et al.*, 2019). Recently, Reinforcement Learning from Human Feedback (RLHF) (Knox and Stone, 2011; Griffith *et al.*, 2013; Lin *et al.*, 2020; Christiano *et al.*, 2017b) has become increasingly prominent as an interactive tool for robot learning. RLHF in robot learning is pivotal in enhancing robot feedback mechanisms, as it allows for the fine-tuning of robotic behaviors based on direct human guidance and evaluation. When integrated with Large Language Models (LLMs) (Zhao *et al.*, 2023), RLHF can empower robots with the ability to process and interpret complex, nuanced human language, making the feedback process more sophisticated and context-aware. This integration leads to more accurate and relevant responses from the robot, aligning closely with the intentions and expectations of human users (Griffith *et al.*, 2013; Lin *et al.*, 2020). However, challenges such as language hallucinations, where the model generates plausible but incorrect or irrelevant information (Ji *et al.*, 2023), and issues with language grounding, where the robot must accurately connect language with real-world concepts and actions (Steels and Hild, 2012; Cangelosi, 2010), remain crucial areas to tackle. Addressing these issues is essential for ensuring the reliability and practical applicability of RLHF with LLMs in interactive and explainable robot learning.

The effectiveness of the feedback mechanism between human and the robot can be dependent on various factors, including the timing, modality, and the level of expertise of the human involved (Natarajan *et al.*, 2023a). Through this section, we will delve into the types, modalities, and dependencies of such feedback mechanisms, discussing how

they can be designed to improve both the robot’s learning outcomes and the overall quality of the human-robot interactive experience (i.e., Figure 4.1). We expect readers to gain a comprehensive understanding of the intricate dynamics of feedback in interactive robot learning. Readers will learn about the various feedback types and how these cater to different interaction scenarios, the effectiveness of different modalities in conveying information, and the crucial role of timing and frequency in feedback delivery. This knowledge is intended to equip readers with insights to critically analyze and potentially contribute to the development of more efficient, adaptive, and user-friendly interactive robot systems.

4.1 Approaches to Interpretability and Explainability of the Robot Feedback

In the realm of interactive robot learning, interpretability and explainability play pivotal roles, ensuring that the feedback provided by the robot is comprehensible and meaningful to the user (Arnold *et al.*, 2021b). Particularly in complex systems, where deciphering the rationale behind the robot’s actions or decisions can be daunting, these concepts become paramount. By incorporating mechanisms for interpretability and explainability, we enable users to better understand and interact with the robot, fostering more effective learning experiences. "interpretability" and "explainability" are closely related concepts, but they have distinct meanings ((Rudin *et al.*, 2022)). The definition of those two concepts are in section 2.1. By incorporating the interpretable and explainable methods in robot feedback, we enhance trust and transparency in human-robot interactions.

Within the context of interpretable robot feedback, (Hristov *et al.*, 2018) provides a method to enhance human-robot interaction in learning from human demonstrations by aligning compressed, low-dimensional representations of high-dimensional sensory input with user-defined symbols and semantics, ensuring that the robot’s interpretations are more aligned with human semantic notions. This is accomplished by introducing an auxiliary classification loss to a Variational Autoencoder. The auxiliary loss impacts the model’s latent space, inducing properties

that make it well-suited for robust symbol classification. Augmented Reality (AR) has been introduced to enhance the interpretability of a robot’s knowledge, specifically in tasks related to opening bottles (Liu *et al.*, 2018b). The proposed Temporal And-Or graph (T-AOG) representation, learned from human demonstrations, allows users to visualize the tree structure of the task on the AR interface, facilitating better understanding and supervision of the robot’s actions. The AR interface enables users to diagnose task failures, interactively teach the robot new actions, and patch its knowledge structure through intuitive visualizations. (Kulkarni *et al.*, 2020) propose a novel framework for designing environments to promote explicable behavior in robots by considering multiple tasks and time horizons, explores the balance between the cost of environmental design and the cost of maintaining interpretable robot behavior over time.

Within the realm of interpretability, mechanistic explanation has emerged as an increasingly important area of study. A mechanistic explanation aims to delve into the inner workings of the network, often by identifying the circuits within a model that drive specific behaviors. This involves examining the trained neural network’s weights to reverse engineer the algorithms learned by the model (Wang *et al.*, 2022a; Nanda *et al.*, 2023; Conmy *et al.*, 2023). For instance, in the context of predictive modeling for interactive robot learning, mechanistic interpretability may involve unraveling the primary features influencing the robot’s feedback, understanding their interactions, and elucidating the precise computations or algorithms utilized by the model for predictions. It ensures transparency and accountability in decision-making processes, which are essential for fostering trust between humans and robots. Ultimately, mechanistic interpretability facilitates informed decision-making and the seamless integration of machine learning systems into real-world interactive robot learning scenarios. (Nanda, 2022; Zimmermann *et al.*, 2023)

Explanations for Robot Learning approaches, typically those which utilize Reinforcement Learning, generally fall under two categories: (1) Explaining the reward function or (2) Explaining the policy. Reward-explanations aim to teach users the functionality of the underlying value-system of a robot (Sanneman and Shah, 2022). These methods

seek to enable a user to understand the high-level goals of a robot. Reward-based explanations typically fall under two subcategories, which are approaches which explain the features intrinsic to the reward and their respective weights (Juozapaitis *et al.*, 2019), or approaches which summarize trajectories to reveal the underlying goals of the robot (Lage *et al.*, 2019; Huang *et al.*, 2018a). Reward-based explanations are crucial in human-robot cooperative settings, to ensure that humans and robot's are working towards maximizing the same objective (Tabrez *et al.*, 2019b). Policy-explanations seek to enable users to interpret the decision making process of a robot to infer the action taken by the robot at any state (Tambwekar and Gombolay, 2023). These approaches either leverage transparent architectures, such as differentiable decision trees, to distill the exact decision making process of the robot at any given state (Tambwekar and Gombolay, 2023; Silva and Gombolay, 2021), or utilize state abstractions to generate an explanation which can empower a user to predict or verify the next action or set of actions at any given state (Topin and Veloso, 2019; Boggess *et al.*, 2022). Policy-explanations are imperative in safety-critical settings wherein users or stakeholders require the ability to correctly interpret and verify a robot's behavior and decision-making process.

Researchers have also explored ways to employ additional tools to generate explainable feedback, understandable by humans, in the domain of interactive robot learning. Methods have been proposed to address the challenge of explaining the internal workings of a robot to enhance human understanding, particularly focusing on explanations tailored for children (Tulli *et al.*, 2020). The proposed search-based approach generates contrastive explanations using optimal and sub-optimal plans in a child-robot learning scenario involving a zero-sum game. While the explainable agency did not significantly affect children's game efficiency, it positively impacted their perceived difficulty, suggesting potential improvements in the learning process through reflection on the robot's explainable agency. Various explanation types have also been study in a eXplainable Human-Robot Interaction (XHRI) framework (Kerzel *et al.*, 2022). The framework incorporates various explanation types, from non-verbal cues to comprehensive verbal and multimodal explanations. The study advocates for context-dependent interaction mediums and efficient

communication channels, highlighting the importance of combining visual and verbal elements for effective explanations. Two example approaches are provided for explaining neural modules in a neuro-robotic architecture. The authors call for future research to express neural reasoning processes faithfully and to support non-expert users with easy-to-interpret multimodal explanations in HRI scenarios. A self-explaining autonomous social robot has been developed to generate social behavior and verbal self-explanations, catering to users' information needs during interactions (Stange *et al.*, 2022). The authors outline requirements for explainable behavior generation architectures, introduce a socio-interactive framework for behavior explanations, and describe the implementation within a social robot's architecture. Evaluation results show the robot's ability to autonomously generate naturalistic social behavior and verbally self-explain its actions, addressing users' requests effectively.

In summary, understanding and explaining the actions of intelligent systems, especially in interactive robot learning, is crucial for effective human-robot interaction, with interpretability focusing on how easily one can comprehend the workings of a model and explainability involving the ability to elucidate complex model decisions. Researchers are exploring diverse techniques, such as aligning sensory input representations with semantics, leveraging augmented reality for visualizing and supervising robot actions, and employing search-based approaches for generating contrastive explanations. These efforts collectively contribute to crafting transparent and comprehensible interactive robotic systems, enhancing collaborative and learning experiences. In the following sections, we talk in detail about what feedback mechanisms have been used by the robots toward interpretable and explainable robot learning. We also talk about feedback timing, feedback penalization, and the implicitness of the feedback.

4.2 Contents and Types of Robot Feedback

The field of feedback mechanisms in interactive robot learning is as diverse as the range of tasks and environments in which robots are deployed (Hoffman and Breazeal, 2004). Understanding the taxonomy

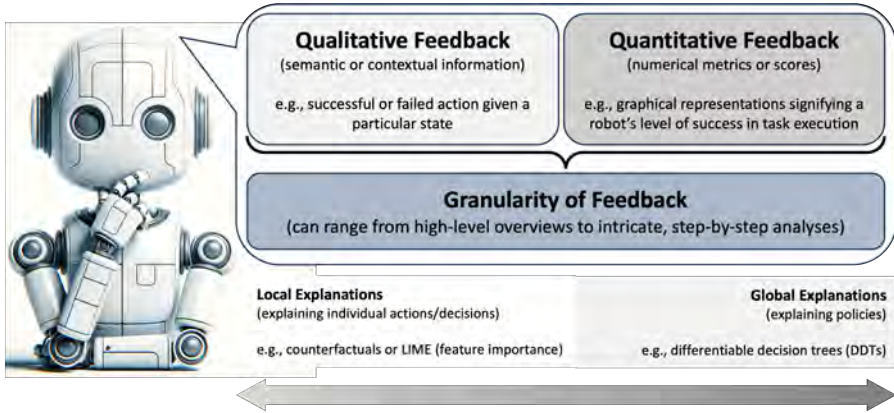


Figure 4.2: Robots can share qualitative or quantitative feedback with their human counterparts in interactive robot learning scenarios. Regardless of the feedback content and type, granularity of the feedback (e.g., high- or low-level analysis) also needs to be determined.

of feedback types is critical to robotic systems that effectively facilitate bi-directional learning (Cuayáhuitl *et al.*, 2013; Jones and Rock, 2002; Natarajan *et al.*, 2023a). Feedback can generally be categorized into two broad types (see Figure 4.2): *qualitative* and *quantitative*. Qualitative feedback often takes the form of semantic or contextual information. For example, a robot might convey why a particular action was taken or failed, using natural language (Jones and Rock, 2002; Perzanowski *et al.*, 2001) or symbolic representations (Hoffman and Breazeal, 2004; Rickel and Johnson, 2000). On the other hand, quantitative feedback provides numerical metrics or scores, possibly accompanied by graphical representations to signify how well the robot understood or executed a task (Steinfeld *et al.*, 2006; Akgun *et al.*, 2012; Goodrich, Schultz, *et al.*, 2008). Prior work has shown the efficacy of these types of feedback in different learning paradigms; qualitative feedback tends to be more useful in tasks that require interpretability and context, whereas quantitative feedback is generally more beneficial for tasks that can be clearly measured or benchmarked (Hoffman and Breazeal, 2004; Jones and Rock, 2002; Steinfeld *et al.*, 2006; Goodrich, Schultz, *et al.*, 2008).

Content-wise, feedback can cover various dimensions, including task performance, interpretability, errors, and progress metrics (Kim *et al.*,

2020c; Mirnig *et al.*, 2012). For example, Kim *et al.* (2020) explored how error-related feedback in robot learning can be used to correct and refine robot behaviors, enhancing the overall learning efficiency. Mirnig *et al.* (2012) investigated the impact of different multi-modal feedback strategies on user perceptions of robot competence and likability. Some interactive robot learning systems focus on performance-oriented feedback, conveying how successfully the robot has performed a task in real-time. This often includes quantitative metrics like completion time, success rate, or efficiency (see Chapter 8). Such metrics are particularly helpful in training scenarios where the objective is clear and quantifiable (Seraj *et al.*, 2023). Prior work has shown that real-time quantitative feedback on task performance metrics significantly improved the learning curve of users in collaborative tasks with robots (Gast *et al.*, 2009; Bray *et al.*, 2012; Skantze, 2016). In contrast, when it comes to tasks with an element of subjectivity or ambiguity, prior research has highlighted the importance of interpretability in the feedback provided (Silva *et al.*, 2023). Silva *et al.* (2023) emphasize the role of feedback in explainable AI, showing how robots that provide explanations of their decision-making processes can foster better human understanding and trust, particularly in scenarios where the robot’s actions are not immediately intuitive. Robots may employ more qualitative content like explanations or justifications, especially when decisions are made based on complex algorithms or machine learning models.

Another critical angle is the granularity of feedback, which can range from high-level overviews to intricate, step-by-step analyses (Thomaz and Breazeal, 2008a; Argall *et al.*, 2009a). For example, in teaching a robot how to assemble a complex machine, an end-user might find it helpful to get high-level feedback about each successfully completed module (Zhu and Hu, 2018). However, if the robot fails at a specific step, more granular feedback would be necessary to identify the exact cause of the failure (MacGlashan *et al.*, 2017). The need for varying granularity has been addressed in prior work by employing multi-modal feedback systems that can switch between text, audio, and visual information to fit the context (Andronas *et al.*, 2021; Lackey *et al.*, 2011; Chen *et al.*, 2022b; Shu *et al.*, 2019; Xue *et al.*, 2020).

Temporal aspects also interplay with the type and content of feed-

back (Cini *et al.*, 2021). Immediate, real-time feedback is invaluable for tasks requiring quick corrections, while delayed, summary-based feedback might be more appropriate for tasks that involve complex analyses or those that unfold over an extended period (Candon *et al.*, 2023). Timing can also be adaptive, with robots choosing the most effective time to deliver feedback based on the flow of interaction and the user’s behavior, a concept explored in several studies involving real-world human-robot interaction scenarios (Cini *et al.*, 2021; Candon *et al.*, 2023; Mitsunaga *et al.*, 2008). This adaptability in feedback timing resonates with the distinctions between Markov Decision Processes (MDPs) and Partially Observable Markov Decision Processes (POMDPs) in robotic modeling (see Section 3.1 for a formal definition of MDPs). In MDP models, where the environment’s state is fully observable and predictable, real-time feedback can be seamlessly integrated. Conversely, in POMDP scenarios, where there is uncertainty and partial observability, delayed or summary feedback might be more suitable, allowing for the accumulation of enough information to make informed feedback decisions. This parallel underscores the importance of considering the underlying robotic modeling approach when designing feedback systems.

Lastly, the modality of delivering feedback, whether it be verbal (Candon *et al.*, 2023; St. Clair and Mataric, 2015), auditory (Badr and Abdul-Hassan, 2020), visual (Halme *et al.*, 2018), or haptic (Liu and Wang, 2018; Casalino *et al.*, 2018a; Haruna *et al.*, 2021), often complements the type and content. For example, auditory feedback (Badr and Abdul-Hassan, 2020) might be more appropriate for conveying complex, qualitative information, while visual cues (Halme *et al.*, 2018) like LED indicators or screen displays might suffice for simple, quantitative metrics. Combining different modalities can lead to a richer, more informative feedback experience. Researchers have been increasingly interested in developing multi-modal robot feedback mechanisms, which can adaptively switch between or integrate various types of content, to better match the user’s needs and the situational context (Andronas *et al.*, 2021; Lackey *et al.*, 2011; Chen *et al.*, 2022b; Shu *et al.*, 2019; Xue *et al.*, 2020).

In summary, the contents and types of robot feedback are influenced by numerous factors including the nature of the task, the temporal

dynamics of the interaction, and the available modalities for communication. Innovations continue to emerge in designing feedback mechanisms that are not only informative but also adaptive, multi-modal, and context-aware, thereby promising more effective and nuanced human-robot interactive learning experiences. As we look toward the future, integrating LLMs into feedback systems presents a frontier with both challenges and opportunities. One promising direction is leveraging LLMs to provide user-centered feedback to improve the effectiveness of a user's interaction with a robot. LLMs, with their advanced natural language processing capabilities, could aid in generating more nuanced and human-like explanations of robotic actions and decisions, thereby addressing the need for feedback that is not only informative but also easily comprehensible to diverse user groups. Additionally, the potential of LLMs in understanding and even predicting user needs and emotional states could lead to more personalized and empathetic interactions in robot learning scenarios. However, this integration also brings forth challenges, particularly in ensuring that the feedback provided by LLM-augmented robots aligns accurately with the users' real-time context and does not introduce biases. Ethical considerations, especially concerning privacy and data security, will also be paramount as these models often require processing large datasets. Furthermore, LLMs can be prone to hallucinations, whereby the model produces nonsensical or unfaithful text to the source input or context (Ji *et al.*, 2023). We strongly urge any researchers who apply LLMs to provide feedback, which could detrimentally impact a user interacting with a robot, to consider incorporating counter-measures such as knowledge-augmented generation (Martino *et al.*, 2023), or post-hoc fact checking (Dhuliawala *et al.*, 2023) to counter the threat of hallucinations. While these methods do not yet provide guarantees towards fully mitigating hallucinations, they provide a step in the right direction towards utilizing LLMs in interactive, user-facing settings. Future research in this domain must navigate these complexities to fully harness the potential of LLMs in creating more dynamic, safe, and user-friendly interactive robot learning experiences.

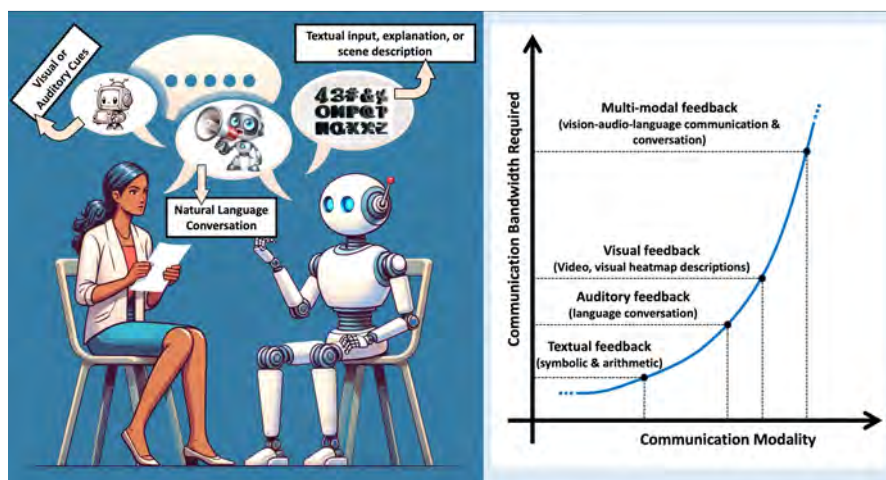


Figure 4.3: Deciding the communication and feedback modality in interactive robot learning with humans is challenging due to subjectivity, end-user dependency, and application dependency. The right-hand curve depicts the communication modality as a function of communication bandwidth (i.e., amount of bits to be transmitted). Typically, textual communication (e.g., symbolic and arithmetic) require the least amount of bits, while multi-modal feedback (e.g., vision-audio-language) necessitates a large communication bandwidth.

4.3 Modalities of Robot Feedback

After determining *what* information is necessary for communication and feedback to effectively inform the human counterpart in robot learning (i.e., feedback content), the next step is to ascertain *how* to share that information and identify the best communication modality. (Figure 4.1). The modalities through which feedback is delivered significantly influence the efficacy of the human-robot learning process (Mirnig *et al.*, 2012; Hoffman and Breazeal, 2004). As our discussion in previous section has shown, the medium can often be as impactful as the message itself. Historically, visual and auditory feedback mechanisms have been the most extensively explored. Visual modalities could range from simple LED indicators that represent states or conditions, to more sophisticated graphical user interfaces that can display a variety of quantitative and qualitative information (Silva and Gombolay, 2021). Research has found that visual feedback is particularly effective for tasks requiring

fine-grained attention to detail or those that are highly quantitative in nature (Halme *et al.*, 2018). For instance, visual cues can be employed to show performance metrics or even offer real-time graphical plots for a more nuanced understanding of the robot’s actions (Silva and Gombolay, 2021).

Auditory feedback, in contrast, is more common in scenarios where the complexity of the task demands that the human user is not tied to a screen or a display (Badr and Abdul-Hassan, 2020; Teo *et al.*, 2018; Cha *et al.*, 2018). Text-to-speech algorithms (Alonso Martin *et al.*, 2020) can articulate complex qualitative feedback, offer contextual explanations, or even engage in dialogues with human users to resolve ambiguities. Prior work has demonstrated the utility of auditory feedback in tasks that require multi-tasking or involve a level of urgency that precludes the human operator from engaging in a visual interface (Badr and Abdul-Hassan, 2020; Teo *et al.*, 2018; Cha *et al.*, 2018). Moreover, auditory cues can serve to attract immediate attention, making them especially useful in alerting systems or real-time corrective mechanisms.

In the diverse landscape of robot feedback modalities, understanding the potential accessibility issues, along with the strengths and limitations of each mode, play a crucial role in their effectiveness and applicability. Visual feedback, while highly informative, can pose accessibility challenges for individuals with visual impairments. The visual feedback modality excels in delivering detailed and spatial information, but its efficacy is limited if the user cannot see the display or interpret visual data effectively (e.g., color blindness). In contrast, auditory feedback is beneficial in these scenarios, offering an alternative means of conveying information through sound. However, its use is constrained for those with hearing impairments or in noisy environments where audio cues can be lost. We will further discuss the limitations of feedback modalities in the design guideline section 4.7.

The landscape of feedback modalities is not limited to the auditory and visual spectrum. Haptic feedback has been utilized as a vital modality, particularly in tasks that require a physical interaction between the robot and the human (Liu and Wang, 2018; Casalino *et al.*, 2018a; Haruna *et al.*, 2021; Urakami and Seaborn, 2023). Whether it is through a simple vibration alert (Scheggi *et al.*, 2012) on a wearable device or

more complex force feedback mechanisms (Mintchev *et al.*, 2019; Scheggi *et al.*, 2016), haptic cues can serve to enhance the user's situational awareness and guide physical actions in real-time. This is particularly salient in medical or industrial settings where tactile sensations can be critical (Urakami and Seaborn, 2023). Haptic feedback offers a strong alternative in situations where both visual and auditory modalities are less effective. Providing tactile information directly through touch, it can be particularly powerful for users with visual and auditory disabilities. However, its limitation lies in its inability to convey complex or detailed data, which visual and auditory feedback can achieve more efficiently. Due to physical interaction with human, the haptic feedback modality can also pose safety issues. Gesture-based feedback, on the other hand, provides a non-verbal means of communication that can be universally understood, but its interpretation may vary across different cultures and its subtlety can be missed by those not attuned to such cues.

Gesture-based and olfactory feedback are relatively less explored but promising modalities (Bonarini, 2020; Saunderson and Nejat, 2019; Urakami and Seaborn, 2023). Gesture-based feedback is gaining traction in scenarios involving non-verbal communications, such as directing a user's attention to a specific object in the environment. Recently, there has been several works dedicated to developing sign-language-based communication for human interactions with underwater robots (Chiarella *et al.*, 2015; Fulton *et al.*, 2019; Chiarella *et al.*, 2018; Xu, 2017). The underwater environment poses a set of technical challenges which are not readily surmountable thus limiting the spectrum from which possible communication modalities can be chosen. In such cases, gesture-based feedback can be of great interest. Olfactory feedback, while niche, has potential applications in tasks where smell can be a crucial factor, like detection of hazardous materials or quality control in food industries (Urakami and Seaborn, 2023). Yet, just like other modalities discussed above, olfactory feedback is limited by its need for the user to be able to perceive and interpret different scents accurately. This modality is also less developed compared to others, presenting a frontier with unexplored accessibility challenges.

Finally, multi-modal feedback (Hoffman and Breazeal, 2004; Rickel and Johnson, 2000; Urakami and Seaborn, 2023), which combines two or

more of these modalities, has been the focus of recent research. Studies have shown that multi-modal feedback is often more effective than its single-modal counterparts, especially in complex environments where information density is high and quick adaptability is required (Andronas *et al.*, 2021; Lackey *et al.*, 2011; Chen *et al.*, 2022b; Shu *et al.*, 2019; Xue *et al.*, 2020). For instance, a robot could combine visual displays with auditory explanations to facilitate both quick understanding and deep contextual awareness (Ince *et al.*, 2021). Multi-modal feedback can often mitigate the limitations of individual modes by offering complementary information. For instance, combining visual with auditory feedback can enhance understanding for a broader audience, including those with specific sensory impairments. Similarly, integrating haptic feedback with visual or auditory cues can provide a more immersive and accessible experience, ensuring that information is not missed due to the limitations of a single modality. In this way, the modality can be chosen to best match the content and type of feedback necessary for the task at hand, making the learning experience more efficient, engaging, and effective. Lastly, drawing from human factors literature, the concept of dual-encoding (Paivio, 1990; Mayer and Moreno, 2003) suggests that people process visual and auditory information through separate channels, enhancing learning and retention when both are used together. This aligns with the effectiveness of multi-modal feedback in robot learning, where leveraging dual-encoding can significantly improve comprehension and task performance by engaging multiple sensory modalities.

Table 4.1 provides an overview of recent studies that utilize different communication modalities for interactive robotics, as discussed in this section. Each modality is paired with references to relevant studies investigating feedback and transparency in interactive robotics using the respective communication channels.

In summary, the modality of feedback plays a vital role in the interactive learning process and the overall user experience. Advances in technology and research continue to broaden the palette of modalities available, providing new opportunities for more effective and nuanced communication in human-robot interaction. While each feedback modality has its strengths, they also come with inherent accessibility challenges and limitations. A summary of the feedback modalities' weaknesses

Table 4.1: Summary of Communication Modalities Utilized in Interactive Robotics

Modality	Reference
Visual	(Rossi <i>et al.</i> , 2021), (Attia <i>et al.</i> , 2018), (Hindemith <i>et al.</i> , 2021), (Struckmeier <i>et al.</i> , 2019), (Tabrez <i>et al.</i> , 2022), (Wengefeld <i>et al.</i> , 2020), (Yigitbas <i>et al.</i> , 2021), (Edmonds <i>et al.</i> , 2019a), (Matsumaru, 2007), (Shah <i>et al.</i> , 2022), (Zakershahrak <i>et al.</i> , 2021), (Han <i>et al.</i> , 2021), (Zakershahrak <i>et al.</i> , 2020), (Liu <i>et al.</i> , 2018a), (Matsumaru, 2006), (Lyons, 2013), (Moon <i>et al.</i> , 2014), (Dragan <i>et al.</i> , 2013), (Seraj and Talamonti, 2024), (Bertozzi <i>et al.</i> , 2000), (Sheikholeslami <i>et al.</i> , 2018), (Diethelm <i>et al.</i> , 2021), (Hirschmanner <i>et al.</i> , 2021), (Kwon <i>et al.</i> , 2018), (Matarese <i>et al.</i> , 2021), (Sanders <i>et al.</i> , 2014), (Tojo <i>et al.</i> , 2000), (Vitale <i>et al.</i> , 2018), (Yonezawa <i>et al.</i> , 2013), (Gross and Boehme, 2000), (Broekens and Chetouani, 2021), (Mobahi and Ansari, 2003)
Haptic	(Casalino <i>et al.</i> , 2018b), (Grushko <i>et al.</i> , 2021b), (Grushko <i>et al.</i> , 2021a), (Valdivia <i>et al.</i> , 2022), (Che <i>et al.</i> , 2020), (Yonezawa <i>et al.</i> , 2013),
Text/Speech	(Wortham <i>et al.</i> , 2017), (Lyons, 2013), (Olatunji <i>et al.</i> , 2021), (Arntz <i>et al.</i> , 2021), (Chen <i>et al.</i> , 2020e), (Hayes and Shah, 2017b), (Hindemith <i>et al.</i> , 2021), (Iucci <i>et al.</i> , 2021a), (Kallinen, 2017), (Kaptein <i>et al.</i> , 2019), (Mota and Sridharan, 2021), (Roncone <i>et al.</i> , 2017), (Rotsidis <i>et al.</i> , 2019), (Sanders <i>et al.</i> , 2014), (Struckmeier <i>et al.</i> , 2019), (Virgolin <i>et al.</i> , 2021), (Vitale <i>et al.</i> , 2018), (Wang <i>et al.</i> , 2016a), (Wang <i>et al.</i> , 2016b), (Yigitbas <i>et al.</i> , 2021), (Edmonds <i>et al.</i> , 2019a), (Gao <i>et al.</i> , 2020a), (Shah <i>et al.</i> , 2022), (Arnold <i>et al.</i> , 2021a), (Han <i>et al.</i> , 2021), (Fiazza and Fiorini, 2021), (Nesset <i>et al.</i> , 2021), (Gross and Boehme, 2000), (Aroyo <i>et al.</i> , 2021), (Cantucci and Falcone, 2020), (Diethelm <i>et al.</i> , 2021), (Hirschmanner <i>et al.</i> , 2021), (Kim and Hinds, 2006), (Straten <i>et al.</i> , 2020), (Tojo <i>et al.</i> , 2000), (Zakershahrak <i>et al.</i> , 2021), (Tabrez <i>et al.</i> , 2019a), (Zakershahrak <i>et al.</i> , 2020),

Table 4.2: Strengths and Weaknesses of Different Feedback Modalities

Modality	Strengths	Weaknesses
Visual	High information density; intuitive for spatial tasks	Accessibility challenges for visually impaired users
Auditory	Effective for sequential information; can be used without visual focus	Can be challenging in noisy environments; accessibility issues for hearing impaired users
Haptic	Provides tactile feedback; useful for hands-on tasks	Limited bandwidth for conveying complex information
Natural Language	Intuitive and flexible; high bandwidth of communication	Requires advanced processing; risk of misinterpretation and language grounding issues

and strengths are provided in Table 4.2. The choice of modality or combination thereof should be informed by the specific context of use, the nature of the task, and the individual needs and capabilities of the user. Understanding these dynamics is crucial for designing effective and inclusive human-robot interaction systems. These varied modalities can be tailored to fit the unique needs of specific tasks, learning paradigms, and human-robot collaboration scenarios.

4.4 Feedback Timing and Frequency

The timing and frequency of feedback in human-robot interactive learning are critical dimensions that influence both the learning efficacy and the user experience (Cini *et al.*, 2021; Candon *et al.*, 2023; Mitsunaga *et al.*, 2008). Research in this domain investigates how different timing strategies and feedback frequencies can optimize human comprehension, robot performance, and the overall effectiveness of the learning process (Gawron, 2008; Durso and Gronlund, 1999; Wickens, 2002; Stout *et al.*, 2017; Bui *et al.*, 2019). The first aspect to consider is the feedback schedule, i.e., whether feedback is provided in real-time, in batches, or at specific milestones (Mitsunaga *et al.*, 2008; Natarajan *et al.*, 2023a; Skantze *et al.*, 2014). Real-time feedback has been found to be most effective in scenarios where immediate corrective action is crucial, such as in surgical robotics or disaster management. In these high-stakes applications, auditory or haptic modalities, as mentioned in

the previous section, are commonly employed because of their ability to convey urgent information without requiring visual focus from the user. Additionally, auditory and haptic feedback are crucial for facilitating environments where visual cues may be obstructed or irrelevant, such as in tasks requiring attention to be directed elsewhere. Moreover, these modalities can offer a more intuitive understanding of abstract or spatial information through sound patterns or tactile sensations, enhancing cognitive processing in ways visual feedback alone cannot (Mayer and Moreno, 2003). This multimodal approach supports diverse learning styles and accessibility, ensuring all users can effectively interact with and learn from robotic systems.

Conversely, batched or milestone-based feedback has been favored in tasks that involve complex data interpretation or where the human user is expected to perform reflective analysis (Kruijff-Korbayová *et al.*, 2011; Lutkebohle *et al.*, 2009; Petersen and Stryk, 2011). In this context, visual modalities with rich graphical interfaces can offer an array of metrics and interpretive data. For example, visual modalities may display charts showing progress over time, allowing humans to evaluate longer-term patterns in robot behavior (Halme *et al.*, 2018). Prior work has indicated that this kind of delayed, aggregate feedback is useful for tasks requiring strategic planning or complex decision-making. Additionally, this timing structure can be particularly suitable for content-based feedback, where a nuanced understanding of the robot’s decision-making process is required (Dautenhahn, 2007).

The frequency of feedback is another essential aspect, closely tied to the type and modality of the feedback being provided. An important aspect to consider when optimizing communication frequency is to investigate its impact on workload, situational awareness, and decision-making (Natarajan *et al.*, 2023a). Too many messages can be overwhelming and reduce situational awareness, resulting in low-quality decision-making, while, too few messages can lead to low situational awareness, insufficient knowledge of the world state, and thus, reduced teamwork effectiveness and performance (Natarajan *et al.*, 2023a; Gawron, 2008; Durso and Gronlund, 1999; Wickens, 2002; Stout *et al.*, 2017; Bui *et al.*, 2019; Seraj *et al.*, 2023). For instance, real-time auditory feedback can quickly become overwhelming if given too frequently, leading to

cognitive overload and reduced effectiveness. The same can be true for haptic feedback, especially if used in a high-frequency manner for tasks that do not require constant vigilance. Overall, different types of tasks pose unique constraints on feedback timing and frequency. For example, in repetitive or routine tasks, feedback can be sparse, focusing on deviations or errors, thereby not overwhelming the user with redundant information (Natarajan *et al.*, 2023a; Saunderson and Nejat, 2019; Mirnig *et al.*, 2012). In contrast, tasks that involve novel or complex decision-making often necessitate more frequent and detailed feedback to aid understanding and performance (Bonarini, 2020; Saunderson and Nejat, 2019; Natarajan *et al.*, 2023a; Seraj *et al.*, 2023).

The concept of adaptive frequency has also gained attention in recent literature (Maderna *et al.*, 2022). This entails dynamically adjusting the feedback frequency based on metrics like user performance, task complexity, or user engagement levels (Maderna *et al.*, 2022; Zhang *et al.*, 2020; Kotowick and Shah, 2018). Additionally, incorporating workload and psychophysiological measures can provide a more nuanced guide for adaptive frequency (Novak *et al.*, 2015; Prinzel III *et al.*, 2003). For instance, tracking indicators such as heart rate variability or galvanic skin response could inform the system about the user's stress levels or cognitive load, enabling the feedback mechanism to adjust its frequency and intensity accordingly. This approach ensures that feedback is not only timely but also attuned to the user's current state, reducing the risk of information overload or under-stimulation (Yang *et al.*, 2022a; Bekele and Sarkar, 2014; Bethel *et al.*, 2007; Guerrero *et al.*, 2012). Incorporating physiological measures into adaptive feedback systems, however, presents unique challenges, such as accurately interpreting the data in real-time and ensuring privacy and ethical handling of sensitive biometric information. These challenges necessitate advanced algorithms and robust data protection measures to effectively utilize physiological indicators without compromising user trust or system integrity. These adaptive systems often employ machine learning algorithms to optimize the feedback schedule, aiming to deliver it at moments when it would be most impactful (Oliff *et al.*, 2020; Sheridan, 2016; Nourmohammadi *et al.*, 2022). An interesting avenue for optimizing adaptive robot frequency in human-robot interactive learning could be the com-

bination of feedback types (see Section 4.2) to adapt the frequency of feedback. For example, evaluative feedback could be frequent in the initial learning phases and under a specific feedback modality, transitioning to descriptive feedback as the user becomes more proficient. Further, during high-workload scenarios, a robot could shift to more visual feedback if psychophysiological measures indicate that auditory processing capabilities are diminished (Bekele and Sarkar, 2014; Bethel *et al.*, 2007; Guerrero *et al.*, 2012). Conversely, in lower-stress situations where cognitive capacity is less taxed, more complex, verbal feedback could be provided.

Lastly, Theory of Mind (ToM)-based feedback (see Chapter 2 for description of ToM in the context of this work), which could be complementary to, or integrated with, other types of feedback mechanisms previously discussed, can also guide the timing and frequency of the robot feedback. Prior work has delved into the integration of ToM into robotic systems (Irfan *et al.*, 2021), showing that robots with rudimentary ToM capabilities tend to facilitate more effective human-robot interactions, as they can anticipate human needs and actions more accurately. For instance, explicit and implicit feedback could be combined in novel ways when the robot has a better understanding of the human's current mental state. The frequency and timing of such feedback could also be optimized based on ToM; a robot that understands human fatigue or information overload could choose to provide feedback at more opportune moments. Another noteworthy point is the potential for ToM to be dynamic and adaptive (Irfan *et al.*, 2019; Irfan *et al.*, 2021).

In summary, the timing and frequency of feedback are multifaceted components that interact with the types and modalities of feedback in complex ways. While no one-size-fits-all strategy exists, researchers and practitioners are increasingly looking toward adaptive and context-aware feedback systems to optimize human-robot interactive learning. Such systems consider the intricacies of human cognition, task demands, and the interactive dynamics between humans and robots to tailor feedback schedules effectively. The goal is not just to optimize robot performance but also to enhance human understanding, engagement, and trust in robotic systems. As research in this domain continues to evolve, it

becomes increasingly clear that feedback timing and frequency are as crucial as the content and modality in shaping effective and meaningful human-robot interactions. Looking to the future, opportunities exist for exploring how emerging technologies like AI-driven analytics and real-time biometric feedback can further refine these adaptive systems. Open challenges remain in balancing the precision of feedback timing with the fluidity of human-robot interactions, particularly in unstructured environments. Additionally, investigating the long-term effects of varied feedback frequencies on learning efficacy and user satisfaction is a crucial area of ongoing research. Finally, investigating and integrating the ToM for better designing the robotic feedback attributes (e.g., content, modality, and frequency) is another promising area of active research. A robot with a well-developed ToM could choose the modality of feedback that is most likely to be effective for a particular individual or situation. For example, visual cues may be more effective for conveying spatial relationships, while verbal explanations may be better for articulating complex decision-making processes. These choices can be made more intelligently when the robot has some understanding of what the human is likely to find intuitive or understandable, a feature that can be facilitated through ToM. These challenges present a fertile ground for future work, promising to deepen our understanding and enhance the sophistication of feedback mechanisms in interactive robot learning.

4.5 Implicit vs Explicit Feedback

Feedback from the interactive robot can be explicit or implicit. Explicit communication involves deliberate and direct transmission of information with a clear intent, while implicit communication relies on inferred cues rather than purposeful transfer of information (Gildert *et al.*, 2018). Although implicit and explicit feedback are different concepts and have different meanings, they are not mutually exclusive (Pascher *et al.*, 2023). The robot can concurrently offer both types of feedback to improve communication with humans. (Che *et al.*, 2020) introduces a planning framework for mobile robot navigation that combines implicit (robot motion) and explicit (visual/audio/haptic feedback) communication. The planner generates communicative actions to enhance the

robot's transparency and efficiency. A user study demonstrates the robot's capability to effectively communicate its intent, resulting in easier understanding, reduced user effort, and increased trust compared to relying solely on collision avoidance, which ultimately enhances the explainability of the robot.

Implicit feedback in interactive robot learning requires interpretation by the human trainers or the learning system itself. Types of implicit feedback from robots could be the robot's movements or backchanneling (Salem *et al.*, 2012; Lala *et al.*, 2017; Park *et al.*, 2017; Hussain *et al.*, 2019; Murray *et al.*, 2022; Urakami and Seaborn, 2023). Backchanneling refers to the non-verbal and verbal cues, such as nods, facial expressions, or short verbal utterances, that listeners use to provide feedback to a speaker during a conversation (Terrell and Mutlu, 2012). Backchanneling from a robot can serve as feedback to guide and inform the human collaborator or learner. (Murray *et al.*, 2022) use data augmentation for effective nodding behavior in a robot and show case incorporating backchanneling behaviors in robots enhances natural and engaging interactions, mirroring human conversational dynamics. The physical form of communication such as robot's movement is particularly useful when verbal communication is less effective or when the interaction relies on visual or kinesthetic learning (Bethel *et al.*, 2007). Backchanneling, such as nodding or head-shaking, offers a way of communication in noisy settings or when speech is not an option. In situations requiring human assistance, a robot's raised hand or wave can effectively capture attention (McNeill, 2019). (Nikolaidis *et al.*, 2017a) propose a adaptive model using probabilistic finite-state controller which allows the robot to guide adaptable participants in collaborative tasks, demonstrating improved effectiveness and maintained human trust. Such implicit cues are essential for seamless and effective teamwork between humans and robots.

Explicit feedback, characterized by direct and clear communication from the robot, contrasts with implicit feedback, which relies on interpretation of movements or backchanneling. Typically, explicit feedback takes the form of clear messages, signals, or actions that are easily understood ((Breazeal *et al.*, 2005; Hindriks and Liebens, 2019)), leaving little room for misinterpretation. One type of explicit is feedback from robots

is the gestures or body language that have an explicit meaning without further interpretation, such as the “okay” hand sign. It is important to distinguish non-verbal explicit cues from implicit backchannel gestures. ((Gildert *et al.*, 2018)). The main difference is that the “okay” hand sign from a robot leaves little room for humans to interpret, while backchannel gestures such as thumbs-up and nods require human interpretation and could have different meanings in different contexts. For example, the thumbs-up from a robot can signify praise when the human demonstrator provides an excellent demonstration for the robot to learn. Additionally, the thumbs-up can also indicate agreement in the context where the robot clearly understands human instructions and agrees with them.

Another type of explicit feedback is question asking or querying humans. (Cakmak and Thomaz, 2012a) propose a method to enhance learning from demonstration by letting robot ask questions. The questions are categorized into labels, demonstrations and feature queries. The findings highlight that feature queries are common in human learning and perceived as intelligent when employed by the robot offers guidelines for designing question-asking behaviors in robot learners. (Rosenthal *et al.*, 2012) explore the capability of autonomous robots to independently solicit aid from humans present in their vicinity, without the need for continuous human supervision. It introduces an innovative model that considers the factor of human presence and assimilates it into the strategic navigation plans of robots. Upon encountering a situation that necessitates assistance, the robot is programmed to autonomously proceed to the offices of available occupants to request their intervention. This approach significantly improves the probability of the robot receiving the required support in a timely manner. The utilization of feedback in the form of questions can be extended to address the challenge of learning reward functions. This can lie in the category of Inverse Reinforcement learning introduced in section 3.4.2. (Bıyık *et al.*, 2019) investigate the problem of reward function learning by querying a human expert. The authors introduce a method that optimizes robot’s question selection for human experts, considering both the robot’s uncertainty and the human’s ease of answering. Results demonstrate that this approach leads to easier questions for the human

and more rapid reward learning for the robot.

In summary, implicit feedback relies on the contextual background of the interaction for human interpretation, whereas explicit feedback provides precise information about robot's intent, aiding users in the learning process. Both implicit and explicit feedback contribute to guiding the training process and enhancing the explainability of the system, enabling human trainers or the robots themselves to adapt the learning strategy accordingly. ((Breazeal *et al.*, 2016)).

4.6 The relationship between robot feedback and the expertise and expectations of the end-user

The relationship between robot feedback and the expertise and expectations of the end-user. is a crucial element in interactive robot learning. It involves not only the information being communicated but also the manner in which it is conveyed. Novice users need simpler, more digestible feedback, while experts may demand detailed, data-rich feedback ((Nocentini *et al.*, 2019; Natarajan *et al.*, 2023a)). If novice users are provided with complex robot feedback containing terminologies understandable only to experienced users, it will undermine system's interpretability and explainability. Even with in the same group of users, different individuals may have different preferences to the feedback. For the most effective human-robot collaboration, feedback mechanisms must be flexible and adaptable to different users' preferences and expertise levels.

Methods have been proposed to help robot provide personalized feedback. (Lee *et al.*, 2020) introduce an interactive robotic exercise coaching system that uses machine learning and rule-based models for personalized feedback, significantly improving exercise assessment accuracy with an ensemble voting method as evidenced by higher F1-scores in post-stroke rehabilitation exercises. (Leyzberg *et al.*, 2018) evaluates a personalized social robot tutoring system using an adaptive Hidden Markov Model for curriculum personalization with Spanish-speaking first graders learning English. The study reveals that students receiving personalized robotic tutoring based on their individual skill levels significantly outperform those with non-personalized instruction,

highlighting the importance of tailoring robot feedback in education. (Maroto-Gómez *et al.*, 2023) introduce an autonomous decision-making robot aimed at personalized interactions based on user preferences. The study evaluates two algorithms, Top Label as Class and Ranking by Pairwise Comparison, selecting the latter for superior preference predictions. The method enables the robot’s decision-making to balance activity exploration with the selection of favored activities. In user studies, participants favor the personalized activity selection, positively impacting their perception of the robot’s likability and intelligence. The above works collectively emphasize the effectiveness of personalized approaches across diverse domains, from policy summarization and education to decision-making for interactive robots.

With respect to the expertise of the participants, two types of participant expertise should be taken into account: mastery in the task domain and understanding in robotics and machine learning. This expertise can be classified into two levels: novice and expert ((Chernova and Thomaz, 2014)). A novice user is someone who lacks extensive experience or proficiency in a particular domain, whereas an expert is skilled and knowledgeable in that domain. For instance, participants with a high level of expertise in machine learning are generally quicker in making predictions about the system’s learning process compared to those with low expertise. ((Saunders *et al.*, 2015; Buchina *et al.*, 2016; Faccio *et al.*, 2023; Natarajan *et al.*, 2023a)) As a result, for novice users, simplified feedback helps them understand basic concepts without being overwhelmed by technical details. While experts may require data-driven insights, statistical information, or advanced error analysis to refine their interaction with the robot. (Carreno-Medrano *et al.*, 2022) also points out that the frequency of robot’s feedback should also be affected by the human partner’s expertise. However, one challenge here is to infer the users’ expertise level. (Enayati *et al.*, 2018) assess an operator’s expertise in tele-operation by analyzing their performance using machine learning methods with motion-based and task-related features. And (Carreno-Medrano *et al.*, 2022) use observations from potentially non-optimal human demonstrations to assess both the human’s objectives and their expertise level simultaneously. Further, (Milliken and Hollinger, 2017) use a POMDP to infer likelihood of a user being an expert after each

observed task execution. In summary, to provide users with personalized feedback, we need to jointly infer users' preferences or expertise levels and generate corresponding feedback.

In summary, successful human-robot collaboration hinges on tailoring feedback to align with users' expertise and preferences. Recent work on recommending gameplay strategies to users, showed that the alignment of a recommended strategy significantly impacted the system's perception and usability (Dodeja *et al.*, 2024). Novice users benefit from simplified feedback, while experts prefer detailed information. Crucially, the ability to infer users' preferences and expertise becomes pivotal in providing personalized feedback. This tailored robot feedback not only influences users' trust but also significantly impacts the performance of robots across diverse tasks.

4.7 Design Guidelines for Effective Feedback in Interactive Robot Learning, Based on Feedback Content, Modality, and Frequency

The design of feedback mechanisms is pivotal to the success of human-robot interaction. The efficacy of these systems hinges not only on their technical capabilities but also on how intuitively and effectively they can communicate with users. Effective feedback design enhances learning, ensures user engagement, and bolsters the overall interaction experience. This section outlines a set of comprehensive design guidelines derived from extensive research in feedback content, timing, frequency, and modalities, as provided in sections 4.2–4.4. These guidelines aim to provide a foundational framework for developers and designers to create more responsive, accessible, and user-friendly robotic systems.

4.7.1 Design Guidelines

1. **Enhancing Interpretability and Explainability in Model Operations:** Ensure that the workings of machine learning models or algorithms are easily understandable by humans. Models should be interpretable, allowing users to trace how inputs lead to specific output. For example, in a domain where a human teaches a robot to cook, the robot's learning model could utilize a decision

tree algorithm, which visually represents the steps taken during the cooking process and the factors influencing each decision to enhance interpretability.

2. **Multimodal Feedback Integration:** Utilize a combination of visual, auditory, and haptic feedback. For instance, a robot performing a demonstration task could show visual cues on a screen, provide verbal instructions, and offer haptic feedback for hands-on guidance. This approach ensures that information is accessible to all users, regardless of sensory abilities, and can adapt to various learning environments (Hoffman and Breazeal, 2004; Rickel and Johnson, 2000; Urakami and Seaborn, 2023). By leveraging the principle of dual-encoding, which posits that individuals process visual and auditory information through separate cognitive channels, this multimodal feedback strategy enhances learning and retention, making the instructional content more memorable and easier to understand.
3. **Adaptive Feedback Timing:** Align feedback delivery with the task's demands. Real-time feedback can be crucial in high-stakes environments like surgical settings, where immediate corrective action is vital. Conversely, in educational or strategic planning contexts, delayed feedback encourages reflective thinking. For example, a robot in a learning environment might provide immediate feedback during a task and follow up with a comprehensive summary at the end of a session (Maderna *et al.*, 2022; Zhang *et al.*, 2020; Kotowick and Shah, 2018).
4. **Granular Feedback Customization:** Offer feedback that matches the user's expertise and the task's complexity. Beginners might benefit from detailed, step-wise feedback, while experienced users may prefer an overview. For instance, a robot teaching a complex assembly task could provide detailed guidance initially, gradually reducing the level of detail as the user gains proficiency (MacGlashan *et al.*, 2017).
5. **User-Centric Feedback Content:** Design feedback that resonates with the user's understanding. Qualitative feedback, rich

in explanations and contextual information, is vital for tasks requiring deeper understanding, while quantitative feedback suits scenarios with clear performance metrics. For example, in tasks involving decision-making, a robot might explain the reasoning behind its choices to ensure the user understands the basis of its actions (Natarajan *et al.*, 2023a).

6. **Balanced Feedback Frequency:** Maintain an optimal frequency that keeps users engaged without overwhelming them. For instance, in a training scenario, a robot could provide more frequent feedback in the early stages of learning, gradually reducing the frequency as the user's skill level increases (Saunderson and Nejat, 2019; Mirnig *et al.*, 2012).
7. **Accessibility and Inclusivity in Feedback Design:** Ensure feedback is accessible, considering users with diverse needs. For example, a robot interacting with users who have hearing impairments could rely more on visual and haptic feedback, whereas for visually impaired users, auditory and haptic feedback would be more effective.
8. **Iterative Improvement Based on Evaluation:** Regularly assess and refine feedback mechanisms. This can involve user feedback sessions, performance analytics, and adaptability tests. Continuous evaluation helps in fine-tuning the system to user needs and evolving task requirements.
9. **Integration of Implicit and Explicit Feedback:** Design systems that allow robots to concurrently offer both implicit and explicit feedback to enhance communication with humans. This integration should aim to improve transparency, efficiency, and overall effectiveness of the interaction (Gildert *et al.*, 2018). For example: In a hospital setting, a medication-delivery robot combines implicit cues like smooth movements and visual status indicators with explicit feedback such as verbal instructions and on-screen prompts.

In conclusion, the design of feedback mechanisms in interactive

robot learning systems is a multifaceted challenge that requires careful consideration of various factors, including modality, timing, content, and user needs. By adhering to these guidelines, developers and designers can create systems that not only facilitate effective learning and cooperation but also foster a positive and inclusive user experience. Future advancements in technology and further research in human-robot interaction will continue to shape and refine these guidelines, driving towards increasingly sophisticated and user-friendly robotic systems.

5

Long-Term Robot Learning

Through this chapter, we provide a comprehensive understanding of the complexities and nuances of long-term robot learning, highlighting the innovative approaches and methodologies that are shaping the future of human-robot interaction.

Imagine a robot, initially trained by human demonstrators, now seamlessly integrated into a dynamic environment – be it a bustling factory floor, a busy household, or a collaborative educational setting. Unlike traditional robotic systems confined to repetitive tasks, this robot adapts, evolves, and continues to learn, fostering a robust, long-term collaboration with its human counterparts (see Figure 5.1). This vision encapsulates the essence of long-term robot learning, a paradigm shift from viewing robot learning as a one-off cycle to a continuous, interactive process. In this Chapter, we delve into the multifaceted nature of long-term robot learning, exploring the intricate blend of interactive learning methods and advanced feedback mechanisms necessary for sustaining and enriching human-robot collaboration over extended periods.

In Section 5.1, we explore the motivations and technological key-contributions for sustaining long-term learning. This exploration includes a discussion on Personalized Robot Learning (see Subsection 5.1.1)

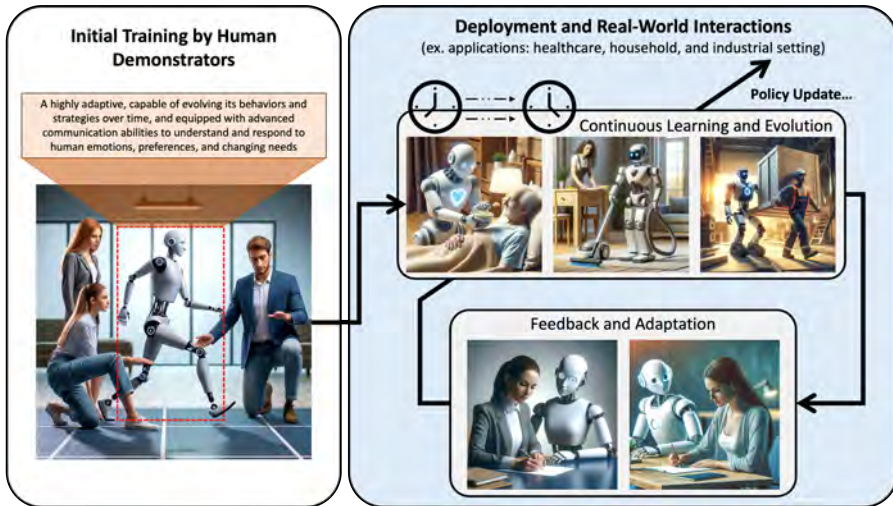


Figure 5.1: This diagram depicts the overall long-term robot learning cycle, starting from initial training of a robot by human demonstrators. An idea robot for long-term interaction and learning is highly adaptive, capable of evolving its behaviors and strategies over time, and is equipped with advanced communication abilities to understand and respond to human emotions, preferences, and changing needs. Upon deployment and experiencing real-world interactions, the robot can leverage feedback and adaptation method to continuously learning and evolve its policy.

and Use-Inspired Robot Learning (see Subsection 5.1.2), highlighting the critical role of customizing the learning experience to meet the needs of individual users or specific target groups. These sections underscore the significance of adopting both user-centric and use-centric approaches in the context of long-term learning in robotics.

Section 5.2, shifts the focus to the social and communicative aspects of sustained interactions. Here, we discuss the role of Theory of Mind (ToM) for Long-Term Interactions in helping robots understand and respond to human mental states and intentions in long term. The Correspondence Problem (section 5.2.1) examines the challenge of aligning robot actions with human expectations, and Feedback Loops in Long-Term Learning (section 5.2.2) emphasizes the importance of continuous and adaptive feedback in maintaining effective, long-term learning relationships in human-robot interaction.

In Section 5.3, we discuss how the objectives and targets of Inter-

active Robot Learning can change and develop over time (Subsection 5.3.1), necessitating flexible and responsive learning strategies. Moreover we cover how to address those changes through periodic updates and modifications in the robot’s learning policies, so it can stay relevant and effective in the face of new challenges and information.

5.1 Algorithmic Needs for Long-Term Interactive Robot Learning Methods

In this section, we explore the reasons for long-term and open-ended learning algorithms in robotics. Long-term robot learning refers to the process by which robots continuously acquire and adapt skills, knowledge, and behaviors over extended periods of time through interaction with their environment (Nolfi, 2021). Unlike traditional learning approaches that focus on short-term tasks or specific datasets, long-term learning aims to enable robots to accumulate experience and improve their performance over time across various tasks, environments, and scenarios. This involves mechanisms for continual learning, adaptation, memory retention, and skill transfer, allowing robots to maintain relevance and effectiveness in dynamic and evolving real-world settings.

The need for long-term robot learning arises from the inherent complexity and variability of real-world environments, tasks, and user interactions. Traditional approaches to robot learning often rely on static datasets or predefined tasks, limiting their applicability to dynamic and evolving scenarios. In contrast, long-term learning enables robots to continuously adapt and improve their capabilities over time, addressing challenges such as changing user preferences, environmental conditions, and task requirements (Yan *et al.*, 2023). Moreover, long-term learning facilitates the development of lifelong companions and assistants that can effectively collaborate with humans over extended periods, supporting diverse applications ranging from household chores and healthcare assistance to industrial automation and autonomous transportation, as shown in 9. As such, the pursuit of long-term robot learning is essential for advancing the capabilities and practical utility of robotic systems in real-world contexts (Irfan *et al.*, 2023).

Readers can refer to the discussion in Sections 3.7.1 and 3.7.2, which

delve into adaptive and adaptable learning methods that adjust the robot's behavior and function in response to changing conditions, environments, and goals. Particularly relevant to long-term robot learning is the exploration of intrinsic motivation methods as discussed in Section 3.7.1. Intrinsic motivators such as curiosity, novelty-seeking, competence, and the desire for autonomy play a crucial role in promoting self-directed learning, exploration, and adaptation within robotic systems. By incorporating intrinsic motivation mechanisms, researchers aim to cultivate more robust and flexible behaviors in diverse environments, contributing to the long-term learning and adaptation capabilities of robots.

5.1.1 Personalized Robot Learning

Personalized Robot Learning focuses on enabling robots to adapt their skills to the unique needs and preferences of individual users or specific groups. This field is driven by the goal of creating robots that can provide tailored assistance, drawing upon a user's preferences, personality traits, experiences, attitudes, and habits to display contextually relevant behaviors (Hellou *et al.*, 2021).

Central to this research is the development of user-adaptive robotic systems. These systems are designed to manage a variety of situations based on user-specific data such as identity, preferences, and level of expertise (Martins *et al.*, 2019). For example, Sekmen and Challa, 2013 developed a mobile robot that learns and adapts to the behaviors and preferences of the people it interacts with. Another contribution is by (Gross *et al.*, 2009), who designed an autonomous shopping assistant robot with capabilities in autonomous navigation and interactive communication, helping users find specific products.

A key focus in recent research is adapting robots to diverse user demonstrations (Ziebart, 2010; Paleja *et al.*, 2020a). Techniques like grouping demonstrations into similar clusters and applying Inverse Reinforcement Learning (more details on IRL in Subsection 3.4.2) have been explored (Nikolaidis *et al.*, 2015a). The Expectation-Maximization (EM) algorithm, which clusters demonstrations and solves the IRL problem, follows a similar approach (Babes-Vroman *et al.*, 2011; Ramponi *et al.*, 2020). Addressing the challenge of learning from varied demonstrations,

(Chen *et al.*, 2020c) introduced a method that optimizes both task goals and human strategic preferences. To overcome the need for knowing the number of different strategies beforehand, (Chen *et al.*, 2023a) developed FLAIR, a system that quickly adapts to new demonstrations for rapid user personalization, expanding its model only as needed for efficiency.

RL has also been an extensively used tool for personalizing robot learning (Hengst *et al.*, 2020). For instance, (He *et al.*, 2024) presented a constrained multi-objective reinforcement learning algorithm for personalized end-to-end robotic control, enabling a single model to approximate user-specific optimal policies. In virtual reality, (Heuvel *et al.*, 2022) used reinforcement learning for personalized navigation control. The study conducted by (Keizer *et al.*, 2014) employed distinct Markov Decision Processes (MDPs) to simulate the actions of a bartender robot tasked with serving beverages to a multitude of users. This approach empowered the robot to tailor its interactions to accommodate the diverse preferences and requirements of individual users. A study by (Görür *et al.*, 2017) utilized a partially observable Markov decision process (POMDP) to adjust robot actions based on human intentions, and (Qureshi *et al.*, 2017) focused on using reinforcement learning for a robot to select appropriate greeting actions based on individual preferences. Additionally, PbRL and learning from human feedback have been extensively utilized to align a robot's behavior with user expectations, as discussed in Section 3.5.

The utilization of latent embeddings to represent types of behaviors or groups of individuals is vastly used in works that touch on personalization. Commonly, those embeddings are learned through some network, and refined via backpropagation in gradient descent-based methods, e.g., (Schrum *et al.*, 2022b; Schrum *et al.*, 2022a). A drawback is that this process is very time-consuming. To overcome that, (Schrum *et al.*, 2023a) and (Silva *et al.*, 2022c) introduced a method that constitutes doing consecutive forward passes in the network that is responsible to capture the embeddings information, by smartly designing networks with some memory nature to it, such as Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997) and transformers.

5.1.2 Use-inspired Robot Learning

Long-term, use-inspired robot learning focuses on developing lifelong learning strategies that are tailored to meet the specific requirements of distinct users or application areas, such as autonomous vehicles, surgical robotics, and home assistance robots. This approach emphasizes the importance of robots continuously adapting and improving over time to cater to the evolving needs of their designated tasks and environments. For a comprehensive exploration of this subject, please refer to Chapter 9.

5.2 Communication in Long-Term Human-Robot Interaction

Traditionally, the focus of communication in robot learning has been centered around basic forms of feedback, be it through visual cues, audio signals, or textual messages (see Section 4.3). While these modalities have their merits, they often fall short in long-term interactions where the context and complexities can evolve. In long-term interactions, bi-directional feedback loops are crucial for aligning human and robot values, facilitating a continuous exchange of preferences and objectives that ensure both parties adjust and refine their behaviors and expectations collaboratively (Yuan *et al.*, 2022). This iterative process not only enhances mutual understanding but also ensures that the robot's actions remain consistently aligned with human ethical standards and goals over extended periods (Mahmud *et al.*, 2022). Achieving such seamless human-robot value alignment leads to an increasing need for more nuanced and adaptive communication methods, including those rooted in the ToM, shared mental models, and feedback loops, to sustain a productive human-robot relationship over an extended period.

Designing the communication mechanisms for long-term interactive robot learning is pivotal, as it delves into the complexities of ensuring that both the robot and the human end-user are in alignment and share a synchronized understanding (Natarajan *et al.*, 2023a; Bonarini, 2020; Irfan *et al.*, 2021; Irfan *et al.*, 2019). To understand the communication needs in long-term interactive robot learning, we first need to understand the concept of *mental models* in human-robot interaction.

A mental model in the context of human-robot interaction refers to the internal representation or cognitive framework that a human user or a robot possesses about each other's capabilities, intentions, and ways of functioning (Chong *et al.*, 2005; Konan *et al.*, 2022). This model guides how they predict and interpret each other's behaviors and actions. In human-robot interaction, a person's mental model of the robot includes assumptions about how the robot works, its abilities, limitations, and how it might respond in different situations. Simultaneously, advanced robots might also develop a form of mental model about their human users, which includes understanding their preferences, behaviors, and likely responses. A *shared mental model* in long-term human-robot interaction is a mutual, often evolving understanding between the human and the robot about their respective roles, capabilities, and strategies to achieve common goals (Ramaraj, 2021; Powers and Kiesler, 2006; Rueben *et al.*, 2021). This shared model is crucial for effective collaboration, especially over extended periods. In long-term interactions, as both the human and the robot learn from and adapt to each other, their shared mental model is continuously refined and updated. In this context, the shared mental model is integral to achieving a harmonious and productive human-robot partnership, especially in complex tasks requiring teamwork and mutual reliance. When there is a misalignment in mental models, ineffective or even counterproductive interactions can arise (Colman *et al.*, 2014; Chong *et al.*, 2005; Seraj, 2023; Seraj, 2022; Seraj *et al.*, 2022). This misalignment can be costly in long-term learning scenarios where the cumulative effect of misunderstandings can lead to poor learning outcomes or reduced user trust (Mutlu *et al.*, 2016). Therefore, mechanisms that allow a robot to convey crucial information and explanations become indispensable for harmonizing the differing mental models.

Prior work in the field has often emphasized the critical role of communication in establishing shared team mental model and achieving successful team coordination, including human-robot interactions (Seraj *et al.*, 2022; Bonarini, 2020; Kidd and Breazeal, 2008; Baxter *et al.*, 2011). These studies suggest that effective communication can facilitate smoother interactions, more rapid learning, and a more reliable collaboration. However, the need for communication extends beyond

initial interactions and becomes even more pressing as the robot and human engage in long-term collaborative tasks (Butchibabu *et al.*, 2016; Butchibabu, 2016). Over time, as both the robot and the human learn and adapt, their respective mental models will also change, thereby creating a moving target for alignment. Thus, the robot’s ability to clearly and adaptively communicate is not just a one-off requirement but a continuous need.

In the following sub-sub-sections, we will explore two critical avenues for effective communication in long-term robot learning: (1) the ToM and the correspondence problem and, (2) feedback loops in long-term learning. These topics go beyond the simplistic feedback mechanisms, offering insights into how robots can be designed to better understand human mental states, beliefs, and intentions, and how ongoing, reciprocal communication can be managed and optimized for the benefit of both parties involved.

5.2.1 Theory of Mind and The Correspondence Problem in Long-Term Robot Learning

One of the most intriguing applications of ToM in robotics is in solving what is known as the Correspondence Problem (Carlson *et al.*, 2013). In a long-term interaction, both the robot’s and the human’s understanding and expectations can change. A static model of the human’s mental state would quickly become obsolete. However, ToM allows for adaptive understanding, which can be continuously updated as more data becomes available. In the realm of long-term human-robot interaction, the correspondence problem, introduced in Chapter 2, emerges as a critical challenge, particularly from a communication perspective (Del Duchetto and Hanheide, 2022; Datteri, 2013). Over extended periods of interaction, ensuring consistent and accurate understanding between humans and robots becomes increasingly complex (Del Duchetto and Hanheide, 2022; Datteri, 2013). As robots are often programmed to follow certain rules or algorithms, their actions might seem opaque or undesired to humans over time as humans develop new habits or values, leading to a disparity in the shared mental model. This misalignment can result in communication breakdowns, reduced cooperation

efficiency, and even diminished trust in the robotic system over extended human-robot interaction periods (Irfan *et al.*, 2019; Chen *et al.*, 2018a). Tackling this problem requires addressing the correspondence problem through sophisticated communication strategies, where the robot not only conveys its actions but also provides context and rationale in a way that is easily interpretable by the human user. Advanced methods, such as employing natural language processing or interactive feedback loops, can aid in this process by making the robot's decision-making process more transparent and understandable, thereby bridging the gap in correspondence and fostering a more harmonious and effective long-term partnership (Nehaniv, Dautenhahn, *et al.*, 2002; Del Duchetto and Hanheide, 2022; Carlson *et al.*, 2013; Datteri, 2013).

In summary, ToM offers an advanced and increasingly necessary layer to robotic systems, especially in contexts that require long-term human-robot interaction. The capabilities extend from offering more context-aware feedback to solving complex cognitive challenges like the correspondence problem. While much has been achieved, it remains a vibrant field of study, and its full integration into robotic learning systems holds the promise of significantly enhancing the depth and effectiveness of human-robot collaboration over extended periods. To effectively address the correspondence problem through ToM-based frameworks, specific design guidelines should be considered. Firstly, it is crucial to develop ToM models that can accurately gauge and respond to the evolving mental states and intentions of human users. This requires the integration of sophisticated cognitive modeling techniques that can process and interpret a wide range of human behaviors and cues. Additionally, communication mechanisms need to be designed in a way that they can provide timely and context-sensitive responses that align with the user's current mental state and understanding. This might involve adaptive feedback that changes in complexity and form based on the user's demonstrated comprehension and interaction history. Furthermore, to enhance long-term learning and cooperation, these systems should be capable of updating and refining their ToM models continuously, based on ongoing interactions. This adaptive capability ensures that the robot remains attuned to the user's changing needs and expectations over time. In practical terms, designers should focus

on creating interfaces and interaction protocols that facilitate clear, intuitive, and meaningful exchanges of feedback, ensuring that the robotic system can communicate its understanding and intentions effectively, thereby maintaining alignment with the human user's perspective. Such tailored approaches in designing ToM-integrated systems promise not only to mitigate the correspondence problem but also to elevate the overall quality of long-term human-robot interaction, fostering a more seamless and effective collaborative experience.

Future research in this area presents several challenges and opportunities. One key area is the development of more sophisticated ToM models that can accurately interpret and respond to a broader range of human emotions and intentions. This requires advancements in areas like affective computing and natural language understanding. Additionally, ensuring that ToM-enabled robots can function effectively in diverse and dynamic real-world environments is a significant challenge, necessitating research into more robust and adaptable systems. Another important consideration is the ethical dimension of ToM in robotics, particularly in terms of privacy, autonomy, and the implications of robots making assumptions about human mental states. Establishing guidelines for the responsible use of ToM in robotics will be crucial as these technologies continue to evolve. Overall, addressing these challenges will require interdisciplinary collaboration, bringing together expertise from robotics, psychology, ethics, and other fields to realize the full potential of ToM in long-term human-robot interaction.

5.2.2 Feedback Loops in Long-Term Learning

In the domain of long-term robot learning, the significance of effective feedback loops between the human and the robot is paramount. (Irfan *et al.*, 2021; Irfan *et al.*, 2019). Feedback loops serve as a critical mechanism for iterative learning and adaptation, helping both parties to optimize their performance and actions over time. In prior work, much emphasis has been laid on short-term or single-session interactions, but the dynamics of feedback change significantly when we extend the interaction horizon (Irfan *et al.*, 2021; Natarajan *et al.*, 2023a). When interaction stretches over days, weeks, or even longer, continuous feedback

plays a crucial role in maintaining optimal performance, facilitating mutual understanding, and building trust (Kanda *et al.*, 2007; Vinanzi *et al.*, 2019). It allows for real-time adjustments and learning, ensuring both parties remain aligned and effective over extended periods. This ongoing exchange is essential not just for operational efficiency but also for fostering a reliable and transparent relationship between humans and robots (Kanda *et al.*, 2007; Natarajan *et al.*, 2023a; Vinanzi *et al.*, 2019; Irfan *et al.*, 2021).

A key question in this context is how feedback loops adapt over time to suit the evolving needs and competencies of both the robot and the human. Previous research has demonstrated the benefits of adaptive feedback systems in robotics which can modulate the amount and type of feedback based on the task at hand, the learning state of the robot, and the observed or inferred state of the human user (Irfan *et al.*, 2021; Kidd and Breazeal, 2008; Chisari *et al.*, 2022). In long-term settings, adaptive feedback loops must be designed to evolve in a way that accounts for progress in learning, changes in user preferences, or shifts in the environment. These adaptive systems may employ machine learning algorithms to continually update their models, much like in the previously discussed ToM frameworks.

Additionally, feedback loops are also closely tied to the problem of user engagement in long-term interactions (Kidd, 2007; Irfan *et al.*, 2019). Maintaining user interest and motivation over an extended period is challenging, and well-designed feedback mechanisms can contribute significantly in this regard. Kidd (2007) conducted experiments comparing human interactions with a physical robot versus an animated character and revealed that the physical robots significantly enhanced social presence, credibility, and user engagement. This finding emphasizes the impact of a robot's tangible presence in fostering more engaging and effective long-term human-robot interactions. The research also suggests that the incorporation of nuanced expressions and gestures in robots can further amplify user engagement, crucial for sustained interactions. These insights are particularly relevant in designing feedback mechanisms in sociable robot systems, where continuous and engaging feedback loops are essential for maintaining user interest and effective communication over extended periods. Additionally, personalized,

context-aware feedback, likely incorporating ToM, can make the interaction more engaging for the user, thus promoting a more effective and enduring learning experience for both the robot and the human (Irfan *et al.*, 2019). Irfan *et al.* (Irfan *et al.*, 2019) highlight the importance of personalization in maintaining long-term interactions between humans and robots through organizing a workshop, emphasizing on the need for customizing robotic behavior to individual user requirements for their successful long-term adoption in everyday and professional environments.

Another consideration in long-term learning is the issue of feedback decay (i.e., Forgetfulness) (Biswas and Murray, 2015). Just as skills and knowledge can degrade over time if not reinforced, the effectiveness of a given feedback mechanism may also diminish if it becomes too predictable or if the human or robot becomes desensitized to it. Some approaches for combating this include feedback randomization, where the timing and type of feedback are varied to maintain engagement, and meta-feedback systems that not only provide task-specific feedback but also give feedback on how well the feedback process itself is functioning (Biswas and Murray, 2015; Mirnig *et al.*, 2012).

In conclusion, the management and optimization of feedback loops in long-term robot learning are complex and multi-dimensional challenges. While much progress has been made, significant research opportunities remain in developing systems that can adapt their feedback strategies over time, manage the complexities of feedback frequency and timing, and maintain user engagement and effectiveness in long-term interactions. As robots become increasingly integrated into various aspects of human life, the design of effective long-term feedback loops will likely emerge as a critical area of focus. In terms of design guidelines, future systems should incorporate adaptive learning algorithms that can personalize feedback based on user preferences and performance metrics. This approach would ensure that feedback remains relevant, engaging, and effective over extended periods. Additionally, designers should consider implementing meta-feedback systems that monitor and adjust the feedback process itself to prevent feedback decay and maintain user interest.

Future research should also address the challenges of integrating

feedback loops with the evolving capabilities of robots, such as enhanced autonomy and decision-making. Investigating the long-term psychological impacts of continuous interaction with robots and how feedback mechanisms can be optimized to foster positive human-robot relationships is another important area. Open problems include developing feedback systems that are sensitive to the nuanced changes in human behavior and preferences over time, and ensuring these systems are ethical, respecting user privacy and autonomy. As the field advances, exploring these areas will be crucial in realizing the full potential of long-term human-robot interactions.

5.3 Evolving Learning Goals and Policy Updates

In the general case of long-term robot learning, the evolution of the human's and the robot's goals imply on the need for updating the general decision-making process for the robot, i.e., policy update. This section aims to distill the literature for both the evolving learning goals and policy updates, while making some links between the two.

5.3.1 Evolving Learning Goals

After a robot is deployed and ready to interact with humans, it is natural that the objectives change over time. For illustration purposes, let us consider a robot responsible to screw bolts in walls in a construction site. The robot knows specific locations where the bolts should be screwed and what amount of torque to apply. After some time, new bolts need to be screwed on that same wall. Given that, the human workers should be able to easily convey this goal expansion to the robot and have it successfully perform the new task. Similarly, there could be changes in the torque requirements, which would lead to the same need for humans to teach the robot this other new task (screwing the same bolts, but applying a different torque). The change of task is only one of many factors could cause changes of the learning goals. Additionally, as the human-robot teaching-learning process can be time consuming, often intimidating to some users (Lasota and Shah, 2015; Lasota *et al.*, 2017; Akalin *et al.*, 2022) and the cobots can be deployed to users who are not

familiar with programming, ideally the robot will be able to adjust for changes as seamlessly as possible to ensure a healthy long-term relationship.

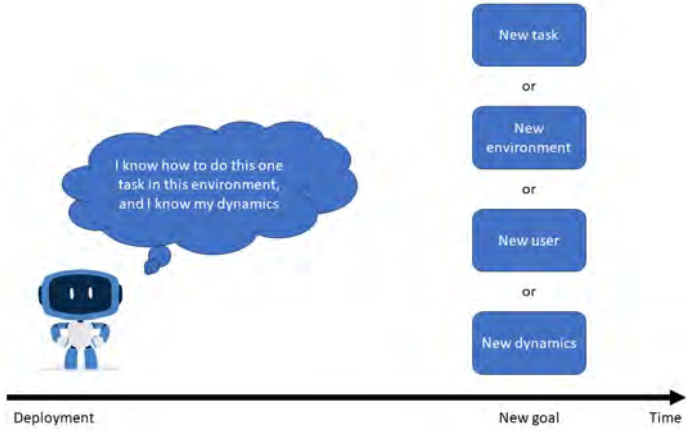


Figure 5.2: Overview of causes for evolving learning goals covered in this section.

Wu *et al.*, 2022 explore the change of task scenario, where robots are deployed in construction sites knowing how to perform just one specific chore, even though they are physically capable of performing a wide variety of tasks and have access to experienced human teachers. For those cases, the tools to employ to account for this change in goal are based on asking the human to teach the robot the new task, and are grounded in methods such as LfD 3.4, Imitation Learning (Celemin *et al.*, 2022) and interactive cases of reward learning (Adams *et al.*, 2022), better explored in 3.3, and 3.4.2.

Another example of evolving learning goals (or factors that could cause it) is changes in the dynamics or environment. Let us consider a household robot, which gets taught by the user how to load/unload their dishwasher. Some time down the line the dishwasher gets replaced with a model that has a slightly different internal configuration, or even get moved to a different location where both the dishwasher’s internal configuration and the kitchen arrangement are considerably different. However, the overall user preference for how utensils should be placed

in the dishwasher remains the same.

Non-interactive ways of teaching a robot, such as by Song *et al.*, 2020 tackle the issue of environmental changes that affects a robot’s ability to perform a task. However, this strategy falls short because it is not possible to anticipate every possible change prior to deployment. Song *et al.*, 2020 utilizes Evolutionary Strategies (discussed in Subsection 3.7.2) combined with Meta-Learning to quickly adapt a previously learned policy to the changes that the robot is experiencing, with no human input. Conversely, Ghadirzadeh *et al.*, 2021 propose an interactive robot learning framework that is able to adapt for changes in the dynamics by asking the user for demonstrations in the new environment.

An alternative to address similar issues is using perception frameworks that captures the dynamics of an environment in a contextual manner, enhancing the awareness of a teaming robot with respect to environmental changes and allowing for self-adaptation (Yasar *et al.*, 2023; Bozhinoski *et al.*, 2022).

Another possible cause for an evolving learning goal is the change of the human. Tasks that are sensitive to individual factors bring the need to adapt when there is any change in the interaction, e.g., the human partner changes how they perform their part of the task, or even is replaced by a different person. Other than the techniques for personalized robot learning shown in Subsection 5.1.1, Variational Autoencoders (defined in 2.2) along with demonstrations from human-human interactions are shown to be an useful tool to teach a robot how to perceive those changes and generate the complementary action from observing the human partner, since it can capture multiple spatio-temporal features of an interaction, as seen in (Bütepage *et al.*, 2020; Prasad *et al.*, 2022).

Addressing the evolution of learning goals for interactive robot learning in the long-term horizon is crucial for the democratization of collaborative robots, while meeting the evolving interaction with their human users and ensuring healthy collaboration over time. Researchers must be aware of this phenomenon during the development of better robots, especially those aimed to novice users (i.e., without making assumptions that the users are familiar with programming). Investigating the topics covered in this section will help researchers to develop

and deploy robotic agents capable to perform adaptation and continual improvement as it interacts with humans without overburdening the user.

5.3.2 Policy Updates Due to Evolving Learning Goals

In long-term human-robot collaboration, there are countless scenarios where the robot's policy may need updates, many of them motivated by the evolution of learning goals. For instance, the robot might need recalibration due to physical wear, because one (or both) of the parties might have changed their usual behavior (e.g. the human has new preferences for how to complete a task), or because a new requirement is in place. This Subsection is dedicated to delve into the literature of how human teachers interact with cobots when the learning goals have evolved and a policy update is due.

While Subsection 3.7 explain some methods used in long-term robot learning, this Section delves into how those methods can be used to address the evolution of goals – mainly by direct updates in the policy or updates in auxiliary tasks. As examples of continual learning, described in Subsection 3.7.1, being used to improve tasks that aid the overall quality of a human-robot interaction, Azagra *et al.*, 2020 focus on tackling interactions between a human with a robot querying for the placement of potentially uncatalogued objects using speech and vision for objects and gestures recognition. This multi-modal incremental learning pipeline is able to improve recognition accuracy over time, which is crucial to update policies that rely on features such as vision in the long-term.

Similar to continual learning (Subsection 3.7.1), few-shot learning in the context of long-term interactive robot learning is a class of techniques aimed to learn different capabilities (i.e., update the policy) using very limited amount of *new* data, and is often paired with meta-learning (Parnami and Lee, 2022). Ghadirzadeh *et al.*, 2021 takes advantage of that pairing to update a manipulator policy, which is trained for a set of hardware models instead of a specific one, extending it to other models by collecting the least possible amount of demonstrations. In this work, the authors use a generative meta-model, i.e.,

a model that captures the common structure shared across different robotic manipulators. A subpolicy is then trained using that meta-model, along with valid motor actions for a few known robots, using an off-the-shelf motion planner. A latent variable captures which robot was used. That latent variable is passed to then train a second policy using few-shot learning. This allows for extending the capacity of a deployed policy in cases where the agent experiences hardware malfunctions, or even undergoes drastic changes in the way it physically interacts with the environment and human user.

While still borrowing key-ideas from Section 3.7, we explore other ways to deal with policy updates within the context of interaction robot learning for the long-term. Active learning (detailed in Subsection 3.7.2) is present in approaches where the robot asks the human teacher to provide *specific* corrections or demonstrations, since the robotic agent requests the human collaborator for the most informative action (Silver *et al.*, 2012; Yu and Wang, 2013), or only the necessary ones (Del Duchetto *et al.*, 2018).

Specifically, Krishna *et al.*, 2022 demonstrates active learning in practice for policy updates in interactive robot learning, by not assuming that the humans will be willing to answer all asked questions. For that, an RL-based generalization of active learning is proposed, using the insight of inferring the human’s preferences through contextual information and tailoring the questions according to what the user is more likely to answer, which is crucial for long-term learning. In contrast, (Cakmak and Thomaz, 2012b) proposes guidelines for what types of questions to ask when using active learning for policy updates, by analysing humans’ perceptions of a robot’s proficiency according to the request, which is imperative for the success of a human-robot team (as discussed in Section 3.9).

In conclusion, addressing the need to continuously update a robot’s capacities is imperative to healthy long-term human-robot interactions. As previously pointed, many factors bring the need for a robot to adapt, some of them being so drastic to the point they require learning a completely new skill. Even though there are still many challenges to face, the near future looks promising for the community.

5.4 Design Guidelines and Future Work

When aiming for reaching seamless interaction with cobots, researchers and practitioners should be aware of the challenges inherent to long-term robot learning, such as mutual understanding, the evolution of goals, and a potential demand of new sets of tasks so long they are within the robot's physical capabilities. We propose a set of guidelines to assist the development of future research:

- **Long-Term Learning:** Robots require the ability to adapt and refine their behaviors and strategies over time through continuous learning. This entails utilizing adaptive and adaptable methods, thoroughly described in 3.7. These techniques enable robots to dynamically adjust to new information, changing environments, and evolving user preferences, ensuring their effectiveness and relevance over extended periods.
- **Personalization and User-Centric Approaches:** Emphasize the importance of personalizing the learning process to individual users or specific target groups. Robots should be able to extract and utilize user-specific data (e.g., preferences, needs, personality traits) to tailor their actions and responses, thereby enhancing the user experience and effectiveness of human-robot interaction.
- **Effective Communication:** Develop advanced communication mechanisms, including the use of ToM and shared mental models (Chapter 2), to ensure both the robot and the human user are in alignment when human preferences potentially alter over time. Communication strategies should be capable of conveying the robot's understanding and intentions in a way that is easily interpretable by humans, facilitating a productive and harmonious relationship over extended periods of time.
- **Feedback Mechanisms:** Implement continuous and adaptive feedback loops that allow for real-time adjustments and learning. Feedback should be tailored to the evolving competencies of both the robot and the human, ensuring that it remains relevant, engaging, and effective over extended periods.

- **Handling Evolving Learning Goals:** Design robots with the flexibility to accommodate changing learning objectives, which may arise from new task requirements, environmental changes, variations in the dynamics, or shifts in user preferences. This could involve leveraging techniques like LfD, Imitation Learning, interactive reward learning, and frameworks that capture dynamic environmental changes, such as meta learning (Beck *et al.*, 2023). In particular, meta learning can enhance the adaptability and robustness of interactive robot learning systems, as it focuses on learning how to learn, enabling robots to quickly adapt to new scenarios based on prior experience. Meta-Learning techniques, such as meta-reinforcement learning, can leverage mixed supervision, rewards, and imitation learning across diverse tasks, facilitating effective adaptation in both few-shot and multi-shot settings (Kaushik *et al.*, 2020; Anne *et al.*, 2021).

5.5 Conclusion

In deploying robots for long-term interaction, a pivotal focus lies in adaptive and adaptable learning methods, coupled with personalized robot learning. This approach is essential in tailoring robotic behavior to meet the specific needs and preferences of different user groups. Central to this endeavor is the integration of advanced communication strategies, particularly those grounded in ToM and shared mental models, which are instrumental in aligning mutual understanding and addressing the correspondence problem in human-robot interactions.

When deploying a robot, some concerns that emerge is on how to account for scenarios that the robot might face and can not be predicted when first being programmed. Those scenarios could be caused by changes intrinsic or extrinsic to the agent: changes in the agent's physical placement in the world or hardware malfunction are causes of the former, while changes in the environment and/or in the task to be performed are causes of the latter. Moreover, incorporating human preferences and behavior is critical when having machines working alongside humans.

To ensure that the human-robot teaming will be as harmoniously

and seamless as possible, researchers in the field of long-term interactive robot learning need to have those concerns always mapped, and address them by ensuring that the robot will always have its decision-making policy up to date with respect to all factors, finding the middle ground between leveraging human feedback and not overwhelming the user.

There has been exciting progress in this field in the last decades, but key challenges arise to tackle more sophisticated scenarios, where everything might change in a very small window of time – which requires spotless and efficient solutions for refining policy due to evolution of goals in the long-term context.

6

Challenges and Techniques for Safety in Interactive Robot Learning

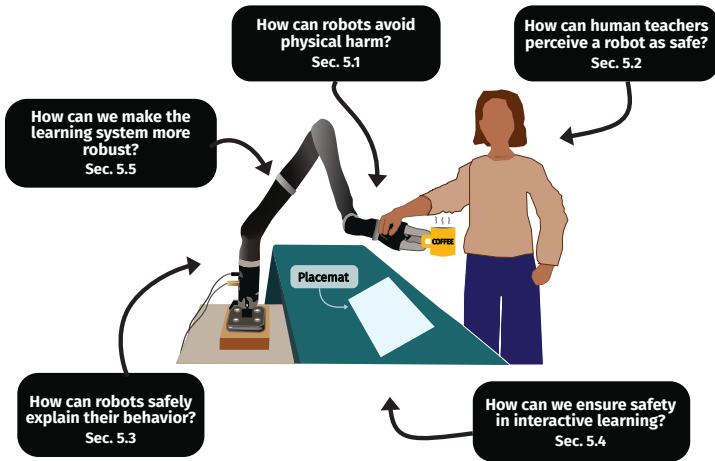


Figure 6.1: Overview figure depicting the different topics covered in this chapter.

Safety is a major challenge in realizing the full potential of robots that can interactively learn and explain their reasoning and behaviors. Safety concerns include the risk for physical injuries when the robot and the human closely interact or the loss of trust in the robot when the robot fails or behaves unreliably. In interactive robot learning, the

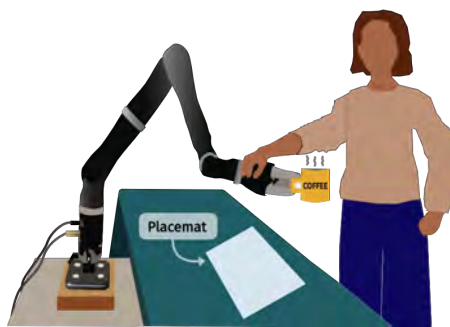


Figure 6.2: Interactive robot learning example.

robot has the ability to iteratively learn from human input and it might be hard for robot designers to anticipate all possible behaviors and outcomes. This context raises questions about what safety means and whether a distinction needs to be made between safe hardware and safe interactions (Salem and Dautenhahn, 2015). This chapter covers various safety challenges around interactive and explainable robot learning. First, we focus on physical and perceived safety questions that arise when people interactively teach robots. Physical and perceived safety refer to that robots should avoid collisions and physical injuries and that robots should elicit feelings of safety and comfort in people, respectively. Second, we consider algorithmic safety, which pertains to the transparency concerns in the development of explainable robot learning systems. We then give an overview of formal methods in robotics to ensure safety and provide guarantees on the robot’s behavior and discuss the integration of formal methods-based approaches into interactive robot learning. Lastly, we discuss different reliability, robustness, resilience and failure recovery strategies of interactive learning algorithms.

6.1 Physical Safety in Interactive Robot Learning

Physical safety refers to avoiding physical harm or injuries when the robot is interacting with its environment and users (Lasota *et al.*, 2017). To date, robot safety approaches have primarily focused on ensuring the physical safety of the robot (Akalın *et al.*, 2022), contributing methods

to avoid collisions between the robot and its environment (e.g., Kulić and Croft, 2006; Haddadin *et al.*, 2008). For instance, in the illustrated typical interactive robot learning setup in Fig. 6.2, the robot is operating in a physical space that is occupied with various objects, such as the table, and shares this space with a human. In such a situation, we must ensure that the robot is never colliding with itself, objects in the environment, and the user. Classically, physical safety has been encoded as a set of hard constraints that the robot needs to adhere to (Brunke *et al.*, 2022; Lasota *et al.*, 2017; Gu *et al.*, 2022). To this end, obstacles are modeled as stay-out-regions and the robot’s occupied space should never intersect with such regions (Kulić and Croft, 2005; Ekvall and Kragic, 2008; Lee *et al.*, 2018; Wang *et al.*, 2017; Koptev *et al.*, 2021; Lacevic *et al.*, 2022). While recent approaches can model those regions with high accuracy, over approximations of those regions can be used for higher computational efficiency.

Interactive robot learning requires us to look beyond classical *avoid-collisions* formulations, since many interactive robot learning scenarios require the human to closely interact and potentially actively be in contact with the robot (Ravichandar *et al.*, 2020). Many industrial robots are deployed in safety cages or operate with laser-triggered stay-out-zones that automatically stop any robot movement if a human enters the zone. Unlike these industrial robots that operate in isolation, robots that interact with humans and operate in human-populated environments do not have such safety precautions, which creates more risks and safety concerns (Winfield *et al.*, 2021). Popular interactive learning approaches often require the human to touch the robot, e.g., to kinesthetically guide the robot through a task (see Chapter 3 for more details). For instance, in the scenario in Fig. 6.2, the human is guiding the robot arm so that the robot learns that it should place the coffee mug on the placemat. Similarly, there are interactive robot learning setups, in which the human has indirect contact with the robot, e.g., in handover tasks where the robot is in contact with the user through objects (Kupcsik *et al.*, 2018; Quispe *et al.*, 2017; Ortenzi *et al.*, 2021; Hedlund *et al.*, 2021). For example, the robot may hand over the coffee mug to the human in Fig. 6.2, resulting in indirect contact via the mug. These interactive robot learning scenarios require new definitions of

physical safety that explicitly consider contact between the robot and its users (Haddadin, 2015; Zacharaki *et al.*, 2020).

Industrial safety definitions, such as the ISO norm 15066, specify additional requirements for collaborative robots that are operating alongside and together with humans (Barattini *et al.*, 2019). These requirements include reduced velocity limits, constraining the robot’s force limits, and compliant control modes, e.g., where humans can override the robots motion while in contact or the robot avoids excessive forces by adjusting its contact forces throughout the interaction. The implementation of such safety definitions requires a tight integration of robot sensors (e.g., forces and torques) with the robot’s controllers (Ferraguti *et al.*, 2020) and the interactive learning approach. While maximum limits on interaction forces can be encoded through hard constraints, soft constraints (Gaurav *et al.*, 2022) are highly suitable to model safe interactions. To this end, we can specify desired low operating forces that the robot should adhere to, but the robot is allowed to violate such constraints when necessary. However, such violation of the soft constraints can be punished through extra costs so that the robot is encouraged to satisfy the soft constraints (Karlsson *et al.*, 2020). Moreover, by designing appropriate cost functions, we can shape the interaction of the robot with the human, e.g., quadratically punishing larger forces in the interaction. Incorporating such constraints into interactive robot learning poses new challenges because the human teacher would need to be able to specify the type of constraint and define the consequences, i.e., costs, of violating them. In the illustrative example in Fig. 6.2, imagine that other objects are blocking the placemat. If the mug placement is modeled as a hard constraint, this blockage would make the robot’s plan to place the mug on the placemat infeasible. However, the person might be fine with the robot placing the mug *close to* the placemat instead if the placemat is blocked, which would be feasible if it was modeled as a soft constraint.

In addition to the robot’s occupied space, we may also consider that the robot should safely handle potentially dangerous objects (Mueller *et al.*, 2018; Haddadin *et al.*, 2011). For instance, the robot should not spill hot coffee close to the human in Fig. 6.2. Similarly, a robot should consider object characteristics when manipulating objects alongside peo-

ple. Some examples include: the robot should not drop fragile objects from a height, should present the handle of a cup if its content is hot, should orient the sharp end of objects away from the user. Unique to interactive robot learning, we must consider potential safety challenges that arise when we ask people to demonstrate to a robot how to manipulate potentially dangerous objects, because such requests must never cause harm to a person. For instance, asking a person touch and move a robot's arm to place a cup of hot coffee or cut food with a knife raises several potential safety concerns. First, the person might feel unsafe closely interacting with the robot when it holds these objects. Second, the person might not trust the robot to perform well and might feel burdened with the responsibility of ensuring the robot's safety. Third, the robot might not be able to learn safety constraints from demonstrations only. Interactive robot learning raises the question of how we can ensure safety of the teacher if we cannot yet ensure the safety of the robot learner. Robots need new ways of incorporating safety constraints in their behavior and research needs to be done to understand when the responsibility of teaching those constraints on the user is acceptable and when it is not.

To account for the probabilistic nature of human-robot interaction, physical safety has also been encoded through probabilistic constraints. In contrast to hard constraints, probabilistic constraints can express probabilistic relationships between variables or parameters. A popular type of probabilistic constraints are chance constraints which specify that the probability of violating the desired constraint should be below a user-defined threshold. For example, they can be used to model that the robot should avoid colliding with at least 95% probability (Du Toit and Burdick, 2011; Shah *et al.*, 2012; Zhu and Alonso-Mora, 2019). Such formulations are helpful to ensure that the robot is able to find feasible actions and behave in a less conservative manner. Moreover, it allows the robot to operate with a residual risk, defined by the user (Xiao *et al.*, 2020; Nyberg *et al.*, 2021; Majumdar and Pavone, 2020). Yet, ensuring that the robot always avoids physical contact with objects and humans at all times is not necessarily suitable for interactive robot learning setups.

So far, interactive robot learning research has primarily focused

on realizing the interactive learning setup in various applications and improving the learning loop itself. Only recently, there have been works that address physical safety aspects of interactive robot learning setups. These works already show that safe interactive robot learning has great potential. Yet, we require more research to realize interactive robot learning techniques that are safe to use by humans at all times.

6.2 Perceived Safety in Interactive Robot Learning

Although physical safety is crucial in interactive robot learning, solely ensuring physical safety is not enough to provide safe interactions with users (Salem *et al.*, 2015). While physical safety primarily focuses on avoiding physically harmful situations with users, the robot should also be *perceived* as safe (Rubagotti *et al.*, 2022; Lasota *et al.*, 2017). Perceived safety refers to the user’s feeling of safety and comfort when interacting with the robot (Bartneck *et al.*, 2009). When users perceive the robot as safe, they generally tend to have higher trust in the robot, increasing acceptability and decreasing psychological stress (Rubagotti *et al.*, 2022). There are various factors that influence the perceived safety of a robot (Akalın *et al.*, 2023), ranging from its appearance or ergonomic design when interacting with it to how the robot behaves during and after the interaction. In interactive robot learning, we must consider how the robot might influence how safe and comfortable people feel. The fact that the robot is learning makes it inherently prone to failures or undesired behavior. The robot might behave poorly or fail in the beginning of the learning process which can affect people’s trust in themselves and the robot (Hedlund *et al.*, 2021; Moorman *et al.*, 2023a). This initial teaching experience might also shape people’s perceptions of the robot in further interactions. Additionally, some interactive learning techniques require the human to touch the robot, which might cause discomfort if the robot is a big industrial robot that is more likely to cause physical harm compared to a small, table-top robot.

Ensuring physical safety does not necessarily lead to perceived safety, because the robot’s behavior might make people feel unsafe even though the robot is guaranteed to never collide with them (Salem *et al.*, 2015). For example, in our interactive robot learning scenario in Fig. 6.2, the

robot arm might approach a person with an objectively safe speed, which allows the arm to always stop in time and never collide with the person, but the person might still feel that the robot moves faster and comes closer than they would like.

In interactive learning setups, humans will work closely with robots, potentially over longer periods of time. However, adaptation to or familiarization with the robot can influence people's perceived safety (Akalın *et al.*, 2022), because people will better understand what to expect from the robot and how their teaching affects the robot's behavior. For instance, a fast-moving robot might be perceived as unsafe by a person in their first encounter but over time, the person might start perceiving the robot as safe if it never causes accidents. As a result, for perceived safety, the robot needs to adapt its behavior between users and also over the course of the interaction (Lasota *et al.*, 2017). These human factors make it difficult to ensure perceived safety in most interactive robot learning settings, unlike physical safety, where it is possible to provide guarantees.

Human-aware planning that adapts the robot's behavior in the proximity of humans has shown to increase perceived safety. For instance, Lasota and Shah (2015) showed that a human-aware planning for a collaborative robot improved team efficiency and humans felt safe during the interaction. Other works have used constraints in motion planning to improve passenger comfort in autonomous transportation (Morales *et al.*, 2015; Karlsson *et al.*, 2021). The work by Norouzzadeh *et al.* (2012) presents a model predictive control scheme to keep the robot in a perceived safe zone when the robot and human are in permanent contact. The perceived safety zone is derived in a study, e.g., which forces users perceived as safe. The reinforcement learning setup in Marta *et al.* (2021) uses human feedback during training to update the robot's behavior for perceived safety. To date, physical safety and perceived safety have often not been studied together. Recent work presents how to combine physical and perceived safety for human-drone interaction by parameterizing the drone's speed and distance to humans and adapting it to their feedback on perceived safety (Van Waveren *et al.*, 2023b).

An increasing body of work is developing new, promising techniques to increase the perceived safety of robots. Yet, there does not exist

work to improve perceived safety in interactive robot learning while ensuring physical safety. To address this gap, we require new advances in interactive robot learning in which perceived safety is another learning criterion. Similarly, the robot learning system might be perceived as more predictable if it explains its decision-making, what its intentions are, and describes what it will do next (Sado *et al.*, 2023). These aspects can ultimately increase perceived safety of the robot and are crucial both, during as well as after learning (Akalın *et al.*, 2022; Akalın *et al.*, 2023). The next section will discuss the safety aspects of explainability and transparency in interactive robot learning.

6.3 Safety Aspects of Explainability and Transparency

Explainability and transparency are vital to build trust in interactive robot learning and the robot itself (Edmonds *et al.*, 2019b). For example, during learning in our interactive learning scenario in Fig. 6.2, we want the robot to explain to the user that it requires input on where to place the coffee mug. Moreover, we want the robot to be transparent about its capabilities and limitations, including joint and torque limits, sensor obstructions, and singularities. For instance, the robot arm may not allow the human to push it into joint limits or singularities. After learning, we want the robot to explain to its user what it has learned and communicate the learned behavior in a transparent manner (Broekens and Chetouani, 2019), e.g., as a visualization of the policy distribution in the workspace. However, similar to most machine learning approaches (Linardatos *et al.*, 2021), explainability and transparency are usually not built into interactive robot learning approaches.

By design, most interactive robot learning approaches use complex deep learning components, which are essentially black-boxes (Castelvecchi, 2016). As a result, these approaches are difficult to analyze for learned behavior. Interpretable Continuous Control Trees (ICCTs) as introduced by Paleja *et al.* (2022) are a novel tree-based model technique designed to produce interpretable policies. ICCTs decision-tree-like representation that can be optimized during training for sparseness and other optimization criteria. Iucci *et al.* (2021b) generate explanations

in reinforcement learning settings for human-robot collaboration by decomposing reward functions and generating natural language outputs. As shown in Li *et al.* (2019b), one can leverage formal specifications to encode interpretable reward functions and robot behaviors. For more details on formal specifications, see Sec. 6.4. The modality of explanations also influences how well explanations are perceived by users as investigated by Tambwekar *et al.* (2023b) with experiments using natural language, saliency maps, and decision trees. The use of explanations is promising to calibrate trust and avoid under- or overreliance on the robot learning system (Sanneman and Shah, 2020).

6.4 Formal Methods and Guarantees

Ensuring that interactive robot learning approaches adhere to desired properties at all times is crucial for safe operation and their certification. The trustworthiness of the robot learning systems should be certifiable, i.e. the robot should do what the human has instructed it to do *at all times*. Due to the black-box nature of many learning approaches, it is difficult to encode constraints such as for physical safety into them (Brunke *et al.*, 2022). For instance, constrained learning approaches aim to check constraints during each learning update and computing gradients to not violate the constraints (Achiam *et al.*, 2017). Yet, these approaches cannot guarantee that the robot adheres to the constraints at all times, resulting in remaining risks. Formal methods are mathematical techniques that rely on formal logic and mathematics to model systems, define their properties, and reason about their behavior. Recently, formal methods-based techniques (Könighofer *et al.*, 2022; Fulton and Platzer, 2018) have increasingly been applied to human-robot interactions, to ensure correct and safe robot behavior.

Formal methods in robotics is a term to denote mathematically rigorous techniques that can be used to formally specify and verify system properties to build trustworthy and reliable robots (Luckcuck *et al.*, 2019). First, desired system properties, e.g., physical safety, need to be encoded in a formal specification language. The term *formal* specification language refers to the fact that the language has an unambiguous syntax and semantics for the desired domain, allowing one to

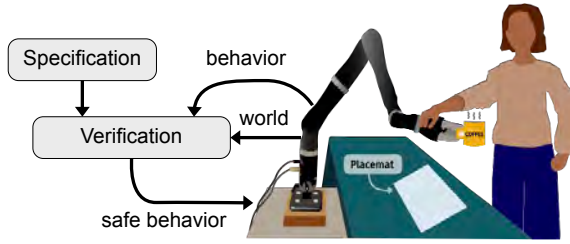


Figure 6.3: Behavior verification overview.

reason about system properties. With the specification of the desired robot behavior at hand, verification tools then check whether a given model, e.g., the robot’s behavior and the evolution of the environment over time, adheres to the specification. If the verification is successful, the model is guaranteed to adhere to the specification. In the opposite case, the verification tool provides a counterexample where the model is not adhering to the specification.

In our example robot learning scenario in Fig. 6.3, we may specify that the robot should always keep sufficient distance between the cup of hot coffee and the human. To avoid verifying the machine learning model itself, we can verify the resulting outputs during the robot’s operation instead, i.e., the robots behavior. Whenever the robot computes a new behavior, the verification technique checks whether the robot’s behavior over time adheres to the specification. To this end, the verification considers the world state, which includes information about the human. For instance, it might include the current position of the human and other auxiliary information, e.g., a prediction of the human’s motion over time. If the behavior is verified as safe, the robot is allowed to execute the safe behavior. In case the behavior is deemed unsafe, i.e., the cup might potentially come too close to the human, we can trigger the execution of a fail-safe behavior instead, e.g., the robot arm retracts and moves the coffee away from the human.

Temporal logics have increasingly become a popular specification language in robotics. Users can specify desired safety properties of the robot, often without additional knowledge about the robot’s dynamics. For instance, Linear Temporal Logic (LTL) (Pnueli, 1977) allows

one to create temporal specifications of robot behaviors with Boolean propositions that are interpreted over time (Kress-Gazit *et al.*, 2009; Kress-Gazit *et al.*, 2018; Schillinger *et al.*, 2018; Finucane *et al.*, 2010). In interactive robot learning, LTL has been used to specify constraints in imitation learning setups (Wang *et al.*, 2022d), specify rewards (Li *et al.*, 2017), learn interpretable models (Camacho and McIlraith, 2019), and to interactively train robots (Shah *et al.*, 2020b). Metric Interval Temporal Logics (MITL) (Alur *et al.*, 1996) allows users to additionally define time intervals in temporal operators, providing more fine-grained control (Lin and Baras, 2020; Barbosa *et al.*, 2021). Signal Temporal Logics (STL) (Donzé and Maler, 2010) consider continuous domains and provide quantitative semantics for specifications, i.e., a continuous measure to what extent a scenario satisfies or violates a specification. STL has successfully been used to specify desired robot behaviors in learning from demonstration settings (Puranic *et al.*, 2021), to optimize demonstrated robot skills (Dhonthi *et al.*, 2022), to learn interpretable models (Camacho and McIlraith, 2019), to specify different robot behaviors such as driving styles (Karlsson *et al.*, 2021), and to guide robot learning (Van Waveren *et al.*, 2023a)

There are various ways to use formal specifications to constrain robot learning while providing guarantees on the learned behavior. For instance, specifications can be turned into discrete automata that encode allowed behavior transitions of the robot (Fainekos *et al.*, 2005; Li *et al.*, 2021a; Kantaros *et al.*, 2020). During learning or execution of the learned behavior, the robot is only allowed to perform a behavior if it is in the set of allowed transitions. The continuous nature of STL specifications is well suited to specify sets of safe states (Pek *et al.*, 2023). With techniques, such as reachability analysis (Bansal *et al.*, 2020; Shao *et al.*, 2021), one can verify that the robot never leaves the set of safe states (Chen *et al.*, 2018b; Selim *et al.*, 2022). Control Barrier Functions (CBFs) are another way of representing safe sets (Ames *et al.*, 2019). STL specifications can be used to (automatically) synthesize CBFs for guaranteed robot behavior (Yang *et al.*, 2020a; Safaoui *et al.*, 2020; Lindemann *et al.*, 2020). The created CBFs can then be used to constrain the robot's behavior. Moreover, specifications can also be used to directly synthesize safe controllers (Raman *et al.*, 2015; Gundana

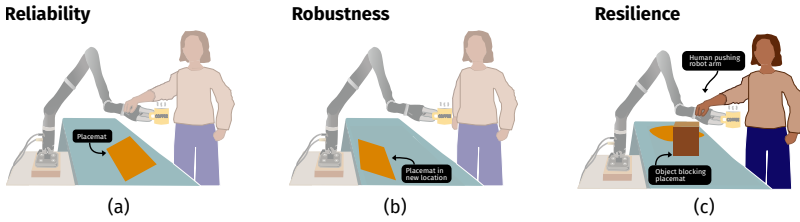


Figure 6.4: Examples for reliable, robust, and resilient interactive robot learning.

and Kress-Gazit, 2021; Finucane *et al.*, 2010; Wongpiromsarn *et al.*, 2012). Alternatively, one can also formulate staying in the safe set in a probabilistic setting. For example, we can formulate a stochastic optimization problem in which we optimize an objective function, e.g., encoding the desired task, subject to probabilistic constraints, such as chance-constraints, to stay in a predefined safety envelope during the task (Schrump and Gombolay, 2019; Schrump *et al.*, 2022a).

Formal-methods based approaches have been used to ensure safety in robotic systems, and are promising to apply in interactive robot learning settings for trustworthy operation. Verification techniques can verify whether a model adheres to a specification. While it is hard to verify the machine learning models directly, we can verify the robot’s behavior and trigger fail-safe behavior when the robot violates our constraints. Similarly, specifications can be used to synthesize safe controllers or functions that constrain the robot’s controller to safe behavior.

6.5 Reliability, Robustness, Resilience, and Error Recovery

Reliability of a learning system is defined as the system’s ability to consistently and successfully perform a given task on any input data and initial conditions as long as this input data and conditions come from the same training distribution. For instance, when we teach our robot to successfully place the coffee mug on the placemat in the example scenario in Fig. 6.4a, the robot acts reliably if it successfully performs this task for various initial poses of the coffee mug. Reliability is crucial for the adoption of interactive robot learning systems and for the successful deployment of interactive learning-based robots (Ravichandrar *et al.*,

2020).

Robustness takes it a step further and is defined as the robot’s ability to successfully perform the learned task even when the data is slightly perturbed or out of the initial training distribution. In other words, the robot should not only succeed at its task in the original environment or scenario but also perform well in variations. In our example, the robot is robust if it successfully places the coffee mug even if the placemat has slightly different appearance (e.g., size or color) and is moved across the table (Fig. 6.4b). Several techniques exist to achieve robustness in robot learning, including but not limited to domain randomization (Mehta *et al.*, 2020; Muratore *et al.*, 2022), active learning (Kroemer *et al.*, 2010; Schrum and Gombolay, 2019; Fitzgerald *et al.*, 2022; Guo *et al.*, 2022), and constrained policy learning (Mueller *et al.*, 2018).

Recently, language and the use of Large Language Models (LLMs) have increasingly received attention in robotics and interactive robot learning (Tellex *et al.*, 2020b; Liu *et al.*, 2023a). Hence, it is important to also consider robustness against natural language confounds and ambiguities (Deits *et al.*, 2013; Karli and Fitzgerald, 2023). LLMs might support the use of natural language to instruct robots; however, these instructions might be highly context-dependent. Hence, interactive learning approaches that use natural language instructions need to have some level of context-awareness, which can help resolve ambiguous instructions. At the same time, LLMs can be used to provide natural language explanations of the robot’s behavior or ask for clarifications to the user. Importantly, using LLMs trained on large datasets also introduces challenges with being mindful and mitigating potential biases, ethical implications, and undesired instances where the model outputs incorrect text, or false or misleading information (hallucinations). Recent works have already explored how reliability and robustness of LLMs can be improved, e.g., the recent KnowNo framework to provide statistical coverage guarantees on LLM-based planners’ uncertainty (Ren *et al.*, 2023a). However, such works require accurate calibration datasets and human interventions to resolve ambiguities during runtime.

Although these works increase the robustness of the robot, the robot is not necessarily resilient to large disturbances. Resilience in the context of interactive robot learning is defined as the robot’s ability to

successfully perform the learned task even under heavy disturbances and to adapt to unexpected inputs and changes in the environments (e.g., an object blocking the location where the robot wants to place the objects as illustrated in Fig. 6.4c). Training robots to cope with such situations is challenging since the disturbances can be potentially arbitrarily large. However, an advantage of interactive robot learning is that there is a human in the loop. Whenever the robot fails, or is close to failing, it is possible to use human feedback to correct the robot and potentially prevent it from failing.

An increasing body of work has explored human feedback for error recovery and policy correction in interactive robot learning (Bajcsy *et al.*, 2018; Fitzgerald *et al.*, 2019; Celemin *et al.*, 2019; Sharma *et al.*, 2022b). Van Waveren *et al.* (2022) propose to generate correction mechanisms, or shields, from non-roboticist (i.e., non-expert) human feedback to correct high-level robot policy failures. To help non-experts identify and correct robot failures, Das *et al.* (2021b) propose an approach to automatically generate natural language explanations for robot failures found that explanations which capture both the context about the environment in which the failure happened and previously successful actions were effective.

The effectiveness of the corrections is highly correlated to the quality of the human feedback, which is highly heterogeneous. Moreover, human feedback is generally sub-optimal, which often leads to additional effort required when correcting robots. Recent works address the heterogeneity of human teachers through the use of personalized embeddings that implicitly capture the quality of the feedback during the pre-deployment training phase, which can be thought of making the robots aware that there will be several different ways of correcting its faulty behavior (Schrum *et al.*, 2022b; Paleja *et al.*, 2020b). Work by Chen *et al.* (2021b) proposes a new approach to learn noise-robust reward functions which improves performance when learning from sub-optimal demonstrations. Celemin and Kober (2023) propose ICREATE, a framework that is an extension of DAgger which explicitly considers inconsistent or ambiguous human feedback.

As described in Sec. 6.3, providing more information about the robot's behavior and decision making can help users understand the

learning system better and adapt how they behave as a teacher. Similarly, enabling the learning system to explain possible causes of failure can help users resolve the failures or refine the learned behavior accordingly. For example, Wu *et al.*, 2021b introduce a failure recovery pipeline that makes use of a structure of failure modes and their effects of the application and a decision network to select failure recovery strategies. The authors demonstrate this in a pick-and-place task, where a robot arm might fail (fault mode) because it is unable to find the object it wants to pick up (error cause) and the recovery module might suggest that the robot re-positions itself so it can detect the object (recovery solution). Yet, this failure recovery pipeline still requires a high degree of technical expertise to understand the failure mode information, solutions, and error reports. The role of the user is to select fault modes, error causes, and preferred recovery solutions in a web-based interface, which is used to further train the robot's recovery pipeline.

Porfirio *et al.*, 2018 leverage automated formal-verification methods to provide users with some technical expertise (e.g., at least one semester of programming courses) with feedback when designed robot behavior violates task or social norms. The application of formal verification can be used to provide feedback in interactive robot learning if the constraints can be explicitly defined, for instance in temporal logics (for more details on temporal logics, see Sec. 6.4). However, challenges remain to make formal methods-based failure detection and recovery pipelines amenable to interactive robot learning with people who are not necessarily programmers or roboticists. Alternatively, and depending on the depth to which the robot understands its limitations, researchers might be inclined to adopt the strategy of simply having the robot ask the human for something they can not do by itself (Knepper *et al.*, 2015).

In interactive robot learning, reliability, robustness, and resilience are crucial for successful and safe task performance. Reliability ensures task success in an initial scenario, robustness enables the robot to adapt to variations, and resilience means that the robot can successfully operate under more significant disturbances. While human feedback is promising to correct robot behavior and refine learned policies, it gives rise to new safety concerns like ensuring safety of newly learned behaviors,

resolving ambiguous instructions, and transparently explaining what the robot has learned. More work is needed to fully leverage the benefits of automated failure recovery and human feedback in interactive robot learning.

6.6 Conclusion

Interactive robot learning has great potential to build successfully operating robots in various situations and conditions. It can leverage advances in the machine learning domain to boost the robustness of the learned behavior. Moreover, it can directly leverage the human in the loop to query corrections and additional information to optimize the resilience of the robot's behavior.

At the same time, interactive robot learning presents new challenges and calls for research in the several directions related to safety. Firstly, as robots might closely interact with humans during interactively robot learning, it becomes crucial to go beyond the *avoid-collision* notion of physical safety, and include aspects such as force, speed, and handling of potentially dangerous objects during the learning process. Secondly, ensuring physical safety does not necessarily lead to robots that are perceived as safe by their teachers. In interactive robot learning, it becomes crucial to investigate how people perceive robot learning systems, how this perception develops over longer periods of time, especially since the system itself learns and adapts over time.

A unique challenge becomes to optimize for perceived safety in interactive learning systems while the system is learning its planner or policy at the same time. Furthermore, due to the inherent black-box nature of learning systems and the fact that many teachers might not necessarily possess programming experience or technical expertise, it becomes increasingly pressing to endow robots with the ability to be transparent about what input is needed from teachers, how the teachers' input was understood by the system, and what behaviors were learned from the input. Additionally, it is hard to provide guarantees in interactive robot learning because we do not necessarily explicitly model and reason over the interaction.

Future work should investigate how we can formally define different

interactive robot learning setups and how we can apply formal methods-based techniques to these settings so that we can give some guarantees on the interaction and learned behaviors. Finally, interactive robot learning needs to deal with ambiguous and suboptimal human input, which will require some level of context-awareness.

7

Ethical Implications of Interactive and Explainable Robot Learning

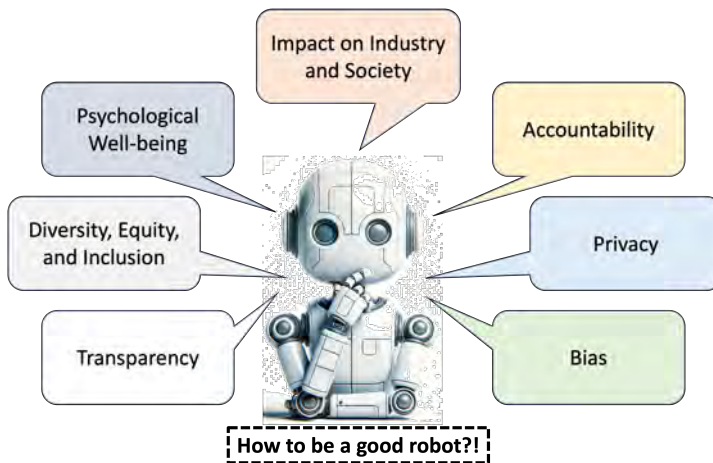


Figure 7.1: Overview of key ethical implication and factors discussed in this chapter.

Ethical issues are also important aspects to consider when developing interactive robot learning systems. For example, consider a scenario where a hospital introduces an interactive robot to enhance patient care. The robot is designed to learn from its interactions with healthcare

professionals, adapting its behavior to provide personalized assistance. However, as the robot becomes an integral part of the healthcare team, questions about accountability, privacy, and potential biases in its learning algorithms emerge. And patients may raise concerns about the privacy of their medical information, wondering how the robot accesses and uses sensitive data. This sparks a broader discussion about the ethical dimensions of using interactive robot learning in healthcare settings.

In this section, we embark on an exploration of the ethical dimensions surrounding explainable and interactive robot learning. Central to our inquiry are the themes of *accountability*, *privacy*, and the inherent *biases* that may arise within learning algorithms. Accountability involves assigning responsibility in the decision-making process of a robot as it learns and interacts with humans (Raji *et al.*, 2020). Privacy encompasses the protection of individuals' personal information and the confidentiality of their interactions with robots (Stahl and Wright, 2018). Bias, on the other hand, pertains to the presence of systematic and unfair favoritism or discrimination in how a robot makes decisions or responds to inputs during the learning process (Friedman and Nissenbaum, 1996; Ferrer *et al.*, 2021). Our analysis will extend to examining current regulatory frameworks and tools designed to mitigate risks, including privacy violations, varying levels of transparency, and biases ((Friesen *et al.*, 2021; Galli *et al.*, 2021)). Additionally, we will scrutinize how interactive robot learning intersects with and impacts critical issues of diversity, equity, and inclusion ((Hagendorff, 2020b; Cachat-Rosset and Klarsfeld, 2023)). This investigation aims to provide a comprehensive understanding of the ethical landscape in which these advanced technologies operate.

7.1 Accountability

In this section, we will investigate the moral and ethical responsibility and accountability assignments related to decision-making AI. Ethics refers to a set of principles governing right and wrong conduct within a particular context or society. It is concerned with determining what is morally acceptable behavior and often involves the application of

philosophical principles to real-world situations. Morality, on the other hand, relates to personal beliefs and principles regarding what is right and wrong conduct. It encompasses individual values, principles, and judgments about what constitutes good or bad behavior. ((Anderson and Anderson, 2011; Wang and Siau, 2018; Wallach and Vallor, 2020)). In the context of interactive robot learning, moral considerations may involve ethical principles that govern the actions and interactions of robots with humans. Accountability remains a controversial topic in learning-based systems, and there has been considerable debate regarding which party should be held responsible when an AI system causes damage to users, properties, or violates social norms ((Smith, 2021)). The crux of this debate centers on whether the designers and developers or the users and operators should bear this responsibility. (Smith, 2021) explore the use of opaque AI systems in clinical decision-making, pointing out that the vagueness of AI systems complicates accountability, making it hard to assign responsibility. A legal framework could guide stakeholders and address potential patient harm scenarios. (Lima *et al.*, 2021) examine perceptions of moral responsibility in AI and human agents in bail decision-making contexts. Results show AI agents are held causally responsible, similar to humans, but humans are attributed more present-looking and forward-looking responsibility. The study highlights the need for explainable AI in decision-making systems.

"Responsibility gaps" refer to situations in the context of advanced technology and AI where it becomes unclear who (or what) should be held accountable for the actions or decisions made by these systems (Amoroso and Tamburrini, 2019). This gap arises particularly with technologies like autonomous systems, robots, and AI algorithms that can make decisions or take actions independently of human control. (Santoni de Sio and Mecacci, 2021) analyzes the "responsibility gap" in AI through philosophical and legal perspectives. It critiques current simplistic approaches and proposes an integrated method based on "meaningful human control" to address these gaps.

(Gunkel, 2020) explore three approaches to address responsibility gaps. The first is using the strict instrumental theory, maintains that all responsibility should be attributed to human beings. It views technology, including sophisticated robots, as mere tools or instruments of

human decision-making and action. Under this theory, humans are fully accountable for whatever the technology does or fails to do. Second is assigning quasi-responsibility to machines. This perspective suggests that machines, particularly advanced robots, can be assigned a level of functional responsibility, implying that while robots are not full moral agents, they can hold a certain degree of accountability for their actions. The last is the distributed responsibility. This approach proposes a more nuanced distribution of responsibility across a network of interacting components, which includes both human and non-human elements. It acknowledges the complexity of modern socio-technical systems and suggests that responsibility should be shared among various contributors to the system, rather than being assigned to a single agent.

Value alignment has been adopted to address the accountability issue in interactive robot learning. Value alignment refers to the process of ensuring that the values of individuals, entities, or systems are consistent or compatible with each other ((Sierra *et al.*, 2021)). In interactive robot learning, value alignment helps establish clear lines of responsibility between humans and robots. When robots are designed to align with human values, it becomes evident who is responsible for the robot's behavior – whether it's the designers, programmers, or the robot itself ((Gabriel, 2020; Nyholm, 2023)). (Sierra *et al.*, 2021) introduce a formal model representing values through preferences and outlines methods for computing value aggregations, both in terms of preferences among groups of agents and preferences related to sets of values. The paper defines and computes value alignment concerning a given norm with respect to a particular value. The focus is on norms as they govern behavior, and the alignment of a system with a specific value is influenced by the norms it follows.

Explainable models in interactive robot learning can also help alleviate the responsibility gaps arising from the increasing complexity of models used in this field. With the improvement in computing performance and the growing popularity of GPT, Large Language Models (LLMs) are being increasingly employed in interactive robot learning ((Liu *et al.*, 2023a; Park *et al.*, 2023; Koubaa, 2023; Tambwekar *et al.*, 2023a)). However, LLMs often function as complex models with opaque decision-making processes. This lack of transparency can pose challenges

in understanding how the model reaches specific decisions or responses, complicating accountability in cases of misleading outputs or errors. Explainable Artificial Intelligence (XAI) enhances transparency by providing insights into the decision-making processes of LLMs ((Ali and Kostakos, 2023; Datta and Dickerson, 2023)). This allows developers, users, and stakeholders to understand and interpret model outputs, ultimately reducing the responsibility gaps.

In conclusion, ethical considerations and accountability in interactive robot learning pose complex challenges. Debates over responsibility highlight the need for clear attributions, with legal frameworks, explainable AI, and value alignment proposed as solutions. The concept of "responsibility gaps" emphasizes the need for distinct accountability in advanced technologies. Various approaches, such as strict instrumental theory, quasi-responsibility, and distributed responsibility, are explored. Additionally, the use of explainable models in interactive robot learning is crucial for transparency, reducing accountability challenges linked to complex models like Large Language Models.

7.2 Privacy

Privacy concerns in interactive robot learning stem from their data collection capabilities, mobility in real-world environments, and the perception of robots as social entities by humans. This section delves into mechanisms to bolster resilience against data leakage and explores governmental inclinations to regulate this evolving field. The intersection of robotics and privacy necessitates a nuanced examination of tools and regulatory frameworks to ensure responsible and secure deployment of interactive robots in diverse settings.

One side of the privacy in interactive robots is to constrain the navigation and manipulation of the robot before it get access to the privacy data. For example, patients' units could be considered as private if the interactive robot is not designed for nursing. And in government or military buildings, there could be areas that are classified that should be forbidden for the interactive service robots to enter. The most intuitive method entails delimiting the robot's permissible planning space by explicitly defining its configuration space ((LaValle, 2006)). Yet,

the manual articulation of this "configuration space" proves to be a formidable task for humans, as their apprehension of the world transcends mere numerical abstraction, in contrast to the robotic perspective. Human cognition is inherently imbued with semantic nuances, wherein the world is perceived through layers of meaning.

To address this challenge, researchers have sought to utilize semantic maps to create an environment representation that is consistent with human comprehension. This approach goes beyond the numerical constraints and offers a more intuitive means for humans to articulate and specify the configuration space ((Galindo *et al.*, 2005), (Galindo *et al.*, 2008), (Kantaros and Pappas, 2019), (Cartillier *et al.*, 2021)). In a healthcare setting, a semantic map might include information about the locations of medical equipment, patient rooms, and safety zones. People can manually specify which areas the robot shouldn't access, which offers a more intuitive way for specifying the robot's configuration space. By incorporating semantic meanings, these maps not only enhance interpretability but also contribute to a more seamless integration of human insights into the robotic planning process.

Restrict the camera of the robot from collecting sensitive contents is another side of the privacy concerns. In interactive robot learning, this pertains to eliminating privacy-related information before it is stored and utilized in the training procedure. (Templeman *et al.*, 2014) propose a framework to allows users to mark private areas, like bathrooms, to prevent unintended image sharing. It identifies and flags images from "blacklisted" private areas using detailed image analysis, combining fine-grained and coarse-grained features. (Fleck *et al.*, 1996) present an automatic method to detect nudity in images without manual intervention. The technique combines color and texture analysis to create a skin mask effective across various skin tones. It then uses geometric constraints to identify human figures. Pixelating, blurring, redacting, and replacing are also adopted to obscure the privacy-sensitive parts of the image or video ((Zhao and Stasko, 1998), (Boyle *et al.*, 2000), (Rother *et al.*, 2004), (Gerstner *et al.*, 2012), (Korshunov and Ebrahimi, 2013), (Raval *et al.*, 2014), (Hubers *et al.*, 2015), (Hubers *et al.*, 2015)).

In the realm of interactive and explainable robot learning, the integration of federated learning alongside techniques like differential

privacy has become pivotal for addressing privacy-related concerns (Triastcyn and Faltings, 2019; Wei *et al.*, 2020; Hu *et al.*, 2020; Wei *et al.*, 2023). Unlike traditional centralized machine learning, where data is collected and stored in a central server for training, federated learning allows training models directly on decentralized devices or servers where the data resides, without exchanging the data itself (McMahan *et al.*, 2017; Li *et al.*, 2020b; Chen *et al.*, 2023a). This approach helps to preserve user privacy by keeping sensitive data localized and reducing the risk of data breaches or unauthorized access. Differential privacy, a key component of federated learning, adds noise to the training process, preventing the extraction of sensitive information from individual data samples and further enhancing privacy protection (Dwork, 2006; Dwork, 2008). By combining federated learning with differential privacy techniques, interactive robot learning systems can effectively mitigate privacy concerns associated with centralized data collection and model training. This integration fosters an environment of confidentiality and user trust, essential for the successful deployment and acceptance of interactive robot learning systems.

The last defence to the privacy is to protect the collected data against leakage. Multiple countries have passed laws to regulate privacy issues in machine learning and robot community. The United States proposed the California Consumer Privacy Act ((State of California - Department of Justice - Office of the Attorney General, 2018)), granting Californian residents rights to know and delete their personal data, applying to companies, including those handling data collected by robots. The European Union enacted the General Data Protection Regulation ((*General Data Protection Regulation (GDPR) – Official Legal Text* 2016)), a comprehensive global data protection law with strict rules impacting how robots and AI systems manage data. Canada's Personal Information Protection and Electronic Documents Act ((Act, 2000)) sets rules for handling personal information in commercial activities, applying to data collected by robots. Japan's Act on the Protection of Personal Information ((Ministry of Justice, Japan, 2020)) governs personal data processing, requiring businesses to specify data usage, obtain consent, and ensure security.

In summary, in tackling privacy concerns within interactive robot

learning, the approach involves three main facets. First, there's a need to control the robot's movements and actions concerning privacy-sensitive data. Techniques like defining permissible planning spaces or using semantic maps are used to establish these constraints. Second, privacy concerns extend to the data collected by robots, especially content captured by cameras. Additionally, integrating techniques like federated learning and differential privacy further strengthens these efforts, ensuring robust privacy protection in interactive robot learning systems. The last is to protect the collected data against leakage. This mainly depends on the legislative efforts from governments, like the California Consumer Privacy Act and the General Data Protection Regulation. Those laws play a critical role in governing how robots handle and manage data, addressing legal challenges related to privacy in the realm of robot learning and human-robot interaction systems.

7.3 Biases

The issue of bias in learning systems is a significant concern in interactive and explainable robots. The bias often stems from the data used to train the systems, which may reflect historical or societal biases, and from the subjective decisions made during their design and development by humans.

There are two types of bias in interactive robot learning: data-level bias and model-level bias ((Londoño *et al.*, 2022)). Bias at the data level in robot learning occurs when the data includes human biases or lacks fair representation of diverse samples. The learning algorithms can substantially reproduce and potentially magnify the unfair scenarios that are either explicitly or implicitly present in the data used (Silva *et al.*, 2021). For example, in an interactive robot learning scenario, human teachers might conduct biased demonstrations and provide prejudiced feedback, leading the robot to learn and exhibit biased behavior. Additionally, interactions occurring in specific contexts could result in contextual bias, limiting the applicability of the trained robot to other environments (Leslie, 2020). Model-level bias refers to discriminatory behavior or reinforcement of harmful stereotypes. Stemming from inherent characteristics of machine learning algorithms

and the design of objective functions during training, this bias may cause robots to associate certain features, such as race, age, or gender, with specific characteristics in their predictions or interactions (Cavazos *et al.*, 2020). Addressing this issue requires prioritizing ethical considerations, including the use of diverse datasets, mitigation strategies to identify and correct biases, and establishment of policies to ensure adherence to ethical standards and mitigate potential harm (Hundt *et al.*, 2022b).

Before addressing bias, it is essential to detect bias in the model. Explainable models play a significant role in bias identification. Tools such as SHAP (SHapley Additive exPlanations, (Lundberg and Lee, 2017)) and LIME (Local Interpretable Model-agnostic Explanations, (Ribeiro *et al.*, 2016)) are commonly employed for understanding the importance of different features in the model's decision-making process. These tools allow for the analysis of individual features and predictions, enabling the identification of any disproportionate contributions to biased predictions ((Teso and Kersting, 2018; Teso and Kersting, 2019; Sano *et al.*, 2021)).

Addressing data-level bias can be approached in various ways within the context of interactive robot learning. One approach involves preprocessing the training data to obtain a cleaner or restored version of the dataset. Methodologies from the broader field of machine learning can be leveraged for this purpose. For instance, a common but naive strategy is to remove bias-sensitive entries such as gender and age ((Grgic-Hlaca *et al.*, 2016)). However, this approach may inadvertently eliminate correlations between these sensitive attributes and other relevant factors ((Londoño *et al.*, 2022)). Instead of directly removing bias-sensitive attributes, innovative methods have been proposed. For example, (Hagendorff, 2020a) suggest a novel approach that selects training data based on ethical assessments of behavioral data, considering social and psychological backgrounds. Additionally, researchers have developed algorithms to repair contextual bias in Deep Neural Networks (DNNs). (Agarwal *et al.*, 2022) introduced a data repair algorithm focusing on curating a balanced dataset by selecting samples that are fair in terms of co-occurrence with various classes for a protected attribute. This approach has been shown to train fair models across different tasks, architectures, and methodologies, without compromising overall per-

formance. Similarly, (Calmon *et al.*, 2017) presented a probabilistic approach to data preprocessing aimed at reducing discrimination in algorithmic decision-making. Another strategy to mitigate data-level bias is to create a diverse dataset. (Padalkar *et al.*, 2023) investigated this methodology by developing a policy for a versatile robotic system capable of effectively adjusting to various robotic platforms, tasks, and surroundings. They offered datasets in standardized layouts along with models for robotic manipulation, showcasing beneficial knowledge transfer and enhanced performance across a range of robotic systems. The model trained on data from 22 different robotic configurations displayed a success rate 50% higher than that of current state-of-the-art techniques.

Methods of adding a additional reward or penalty term to the reward function during training are widely adopted to mitigate the model-level bias ((Kamishima *et al.*, 2012), (Iosifidis and Ntoutsi, 2018), (Agarwal *et al.*, 2018), (Wilson *et al.*, 2019), (Tartaglione *et al.*, 2021)). This suggests that a negative reward could be applied if the interactive robot learning algorithm is favoring a certain category of policy. Transfer learning is also use to enhance fairness. (Schumann *et al.*, 2019) explore the challenge of ensuring fair predictions by models in new or unexpected cases, highlighting the gap in fairness literature regarding this issue. It treats such scenarios as domain adaptation problems, proposing a method to transfer debiasing strategies from a well-researched source domain to a less understood target domain. This approach provides theoretical guarantees for improving fairness across various domains and demonstrates empirically that these strategies can enhance fairness metrics even in data-sparse situations. (Zhang *et al.*, 2018) use adversarial learning technique to mitigate biases in learning systems. The approach aims to enhance prediction accuracy while minimizing the adversary's success in predicting the protected attribute, thus reducing bias. Drawing inspiration from these works, addressing model-level bias in interactive robot learning systems could benefit from similar ideas. (Hundt *et al.*, 2022b) reveal extensive bias and discrimination in robot learning. The study shows robots perpetuating toxic stereotypes related to gender, race, and physiognomy. The interdisciplinary analysis calls for a pause on robot learning methods manifesting harmful stereotypes

until safety and fairness are ensured. The authors also advocate for comprehensive policy changes and interdisciplinary research.

In addition, explainability methods are another approach to reduce the bias in robot learning. Explainable AI (XAI) provides transparency in how the robot makes decisions((Kim *et al.*, 2018), (Cultrera *et al.*, 2020), (Paleja *et al.*, 2022), (Omeiza *et al.*, 2021b)). By understanding the rationale behind decisions, developers and users can identify and address potential biases in the robot’s learning process or its decision-making algorithms. For instance, if a robot consistently makes decisions that favor one group over another, the underlying reasoning provided by XAI can help pinpoint the source of this bias. Here, explainability in interactive robot learning is not just about making the robot’s decisions understandable; it’s also a critical tool for identifying, understanding, and reducing biases in AI systems.

In conclusion, mitigating bias in interactive robot learning involves addressing both data-level and model-level biases. Data-level bias can be reduced through preprocessing techniques or using more diverse dataset to ensure fair representation. Model-level bias mitigation includes adding reward or penalty terms during training, employing transfer learning for domain adaptation, and using adversarial learning techniques. Additionally, Explainable AI (XAI) plays a crucial role in both revealing bias and reducing biases by making the learning process and its outcomes understandable and accountable. This multifaceted approach is essential for developing fair and unbiased AI systems.

7.4 Transparency

Transparency in the context of intelligent agents, particularly robots, refers to the capacity of the system to provide meaningful information to its end-users. A transparent robot or AI algorithm should include mechanisms that allow end-users to obtain precise interpretations of various aspects, such as the system’s capabilities, objectives, ongoing progress, sensory inputs, reliability, and comprehensive reports of any unexpected events. This information should be presented in a format that is readily understandable to humans, enabling users to comprehend and assess the system’s operations effectively (Andreas Theodorou and

Bryson, 2017).

As robots transition into real-world applications, there is a growing emphasis on transparency in regulatory frameworks. General Data Protection Regulation (GDPR) by (European Parliament and Council of the European Union, 2016) and the “Civil Law Rules on Robotics” by (European Parliament, 2017) specifically underscore transparency principles. These regulations necessitate that autonomous or artificially intelligent robots provide rationales for every decision, emphasizing the need to make system computations understandable and explainable to humans. The resolution also advocates for a “black box” in advanced robots, recording data on transactions and the logic behind each decision. Therefore, the resolution essentially calls for interactive robots to be both explainable and interpretable, highlighting the importance of providing rationales for every decision made by autonomous or artificially intelligent robots, and emphasizing the need to make system computations understandable to humans.

7.5 Diversity, Equity, and Inclusion

As robots increasingly engage with individuals in diverse social settings, it has become imperative to design robots that cater to the information needs and values of a broad user base. This research area is crucial for interactive robotics, where users differ in terms of race, ethnicity, and gender. Diversity, within this context, encompasses variations among individuals, including race, nationality, gender, sexual identity, disability, and other distinctive characteristics (Grubbs, 2020). Equity involves providing fair opportunities based on individual needs, striving for an equitable playing field (Walster *et al.*, 1973). Inclusion aims to establish an environment where individuals, regardless of differences, can actively engage and access opportunities for participation (Goodin, 1996).

Recent discourse within the robotics community has brought attention to potential shortcomings, uncovering biases present in current robot learning algorithms (Londoño *et al.*, 2022), emphasizing how these biases can manifest as racial and gender stereotypes (Hundt *et al.*, 2022a). The deployment of pre-trained machine learning models in robots can result in the inheritance and enactment of biases, including

racial, gender, and physiognomic stereotypes (Hundt *et al.*, 2022a). In the era of foundational robotics models (Firoozi *et al.*, 2023), where extensive data trains general-purpose robots, it becomes imperative to scrutinize the impact of these models on diversity, equity, and inclusion. Bommasani *et al.*, 2022 provide a comprehensive exploration of the potential harms caused by foundational models, underscoring the disproportionate impact on marginalized communities. In their work, Birhane and Dijk, 2020 delves into the question of accountability regarding the harm caused by these biases. Addressing these inequities necessitates a multifaceted approach, encompassing proactive interventions (e.g., data-centric and model-centric changes) and reactive measures (e.g., feedback mechanisms and accountability). More detailed discussion on the algorithms and methods being used to mitigate such effects can be found in Section 7.3.

7.6 Psychological Wellbeing

Interactive and explainable robotics offer promising avenues for enhancing psychological well-being by providing personalized and empathetic support in various contexts. These technologies enable robots to engage with individuals in meaningful interactions, offering companionship, assistance, and therapeutic interventions tailored to their unique needs (Feil-Seifer and Matarić, 2009; Broadbent *et al.*, 2011). For an in-depth examination of methods and techniques used in interactive and explainable robotics aimed at enhancing mental well-being, readers are encouraged to consult Section 9.1.

7.7 Design Guidelines for Ethical Concerns in Interactive and Explainable Robot Learning

Ethical concerns in interactive and explainable robot learning are crucial aspects to consider in the development and deployment of interactive and explainable robot learning systems. We propose a set of guidelines to assist the development of future research:

1. **Ethical Decision-making Frameworks:** Equip interactive robots with ethical decision-making frameworks to guide their

behavior in morally ambiguous situations. Incorporate ethical principles such as beneficence, non-maleficence, autonomy, and justice into the robot's decision-making processes.

2. **Accountability Frameworks:** Establish clear lines of accountability within the interactive robot learning ecosystem. Define roles and responsibilities for developers, operators, and users to ensure accountability for the robot's actions and decisions.
3. **Privacy-preserving Measures:** Implement robust privacy-preserving measures to safeguard users' sensitive information. Minimize data collection and processing, and employ techniques such as federated learning and differential privacy to mitigate privacy risks associated with interactive robot learning.
4. **Informed Consent:** Obtain informed consent from users before collecting or processing personal data or engaging in potentially intrusive activities. Clearly communicate the purposes and implications of robot interaction to enable meaningful consent.
5. **Fairness-aware Design:** Mitigate biases in interactive robot learning algorithms to promote fair treatment of all individuals. Implement fairness-aware techniques during model training and evaluation to prevent the propagation of discriminatory outcomes.
6. **Transparency:** Ensure transparency by enabling robots to provide clear explanations for their decisions, making system computations understandable to humans, and allowing users to interpret the reasoning behind each decision.
7. **Uphold Principles of Diversity, Equity, and Inclusion:** Ensure diversity, equity, and inclusion in interactive robotics by designing robots that cater to the information needs and values of a broad user base, encompassing variations in race, ethnicity, gender, sexual identity, disability, and other distinctive characteristics. Strive for fairness and equity by mitigating biases in robot learning algorithms, addressing potential manifestations of racial and gender stereotypes. Additionally, promote transparency and

accountability regarding the impact of biases in robotics, ensuring proactive interventions and reactive measures to address inequities and foster diversity and inclusion.

7.8 Conclusion

In conclusion, the integration of interactive and explainable robot learning stands as a transformative force, with profound impacts on industry and society. The amalgamation of these technologies not only enhances production efficiency and safety but also facilitates rapid skill acquisition, fostering collaboration and acceptance. Nonetheless, challenges persist, particularly in building trust through explainable robots amidst concerns of manipulation and ethical dilemmas. Moreover, ethical considerations, accountability frameworks, privacy safeguards, bias mitigation strategies, transparency mandates, and efforts toward diversity, equity, and inclusion are vital aspects that must be addressed to navigate the complexities of interactive robot learning responsibly. Moving forward, a multidisciplinary approach, informed by ongoing research and ethical deliberations, is crucial to harnessing the potential benefits while mitigating potential risks, particularly in domains like healthcare where the impact on psychological wellbeing holds significant promise.

8

Evaluation, Metrics, and Benchmarking of Interactive and Explainable Robot Learning

Metrics and benchmarks play a pivotal role in evaluating the performance of robotic systems, with the choice of metrics varying according to the specific application. For instance, in autonomous navigation tasks, metrics such as the success rate Rosenberger *et al.* (2020) and collision rate (Nishimura and Yonetani, 2020) are of paramount importance. Conversely, for a robotic tour guides, tour accuracy and user satisfaction with the experience should be focused Cantucci and Falcone (2022). Although literature reviews have been conducted on metrics within the domains of robot handover (Castro *et al.*, 2021), Human-Robot Interaction (HRI) (Coronado *et al.*, 2022), and XAI (Hoffman *et al.*, 2018), these reviews are either very specific in certain application or incorporate excessive metrics, which, within the authors' knowledge, have not been utilized in the papers pertaining to interactive and explainable robot learning. Therefore, the existing surveys do not align precisely with the focus of our topic. In this chapter, we exclusively examine papers related to interactive and explainable robot learning. We aim to elucidate the definition, differentiate between metrics that appear similar, and provide design guidelines for the application of these metrics. Additionally, we present examples illustrating the application

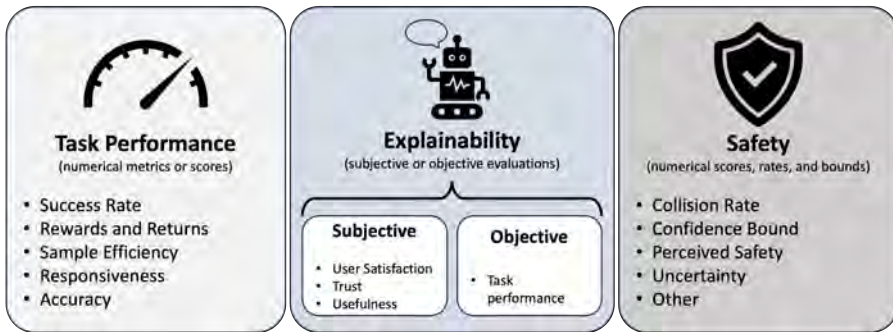


Figure 8.1: Overview for metrics in interactive robot and explainable learning.

of these metrics in relevant research.

Figure 8.1 summarizes the metrics which are discussed in the subsequent sections. In section 8.1, we discuss task performance metrics such as success rate, rewards, sample efficiency, responsiveness, and accuracy, their relevance extends beyond mere performance evaluation in interactive robot learning. In section 8.2, we discuss how task performance metrics, serving as objective measures, also complement subjective metrics such as user satisfaction, trust, and usefulness to provide a comprehensive evaluation of interactive and explainable robot system. In section 8.3, we discuss safety in interactive and explainable robot learning. We summarize physical safety metrics such as collision rate or confidence, the psychological metrics, arising from humans teaming or collaboration with robots, and the perceived safety metric. Although safety have a broad spectrum of concerns, we also include some rare yet insightful metrics for reference. In Section 8.4, we summarize the benchmarks utilized by research papers focusing on interactive and explainable robot learning. Furthermore, we include a list of papers associated with these benchmarks, providing detailed insights into how these benchmarks are applied within the scope of their respective studies.

8.1 Task Performance Metrics

Success rate is a crucial metric in evaluating the performance of interactive robot learning. It is often defined by the probability of the robot that can successfully complete a task. This metric is widely used in many interactive robot learning applications, such as objects manipulation (Rosenberger *et al.*, 2020; Mittal *et al.*, 2023; Jang *et al.*, 2022), object handovers between humans and robots, (Koene *et al.*, 2014; Yang *et al.*, 2020b; Grigore *et al.*, 2013), food preparation (Tai *et al.*, 2023; Grannen and Wu, 2022), autonomous navigation (Li *et al.*, 2020a; Eysenbach *et al.*, 2019), surgical assistance (Long *et al.*, 2023; Scheikl *et al.*, 2022; Chiu *et al.*, 2021), social assistance (Satake *et al.*, 2009; Yousuf *et al.*, 2019), etc.

Success rate is very useful in scenarios where the objectives are clearly defined and success can be unambiguously determined. The objectives could be completion of a physical task, achieving a specific threshold or target value, or meeting any other criteria that is clearly defined. This could include physical tasks like passing the correct tools for the worker repairing a vehicle (Rosenberger *et al.*, 2020) or social interactions like approaching and assisting visitors in a shopping mall (Satake *et al.*, 2009). Thus, the proportion of the successful interactions between robot and human can be clearly annotated. Success rate are sometimes used together with rewards or sample efficiency, which will be discussed later in this section. The design guidelines in these studies emphasize the establishment of specific and measurable objectives for robot interactions and tasks, ensuring a clear definition of what constitutes success.

Accuracy is a crucial metric in interactive robot learning. Unlike the success rate, which solely captures the completion of the final goal, accuracy is extensively used to evaluate performance in a range of scenarios, especially for subtasks. For instance, Sun *et al.* (2020) apply accuracy to assess the performance of predicted assembly sequences for individual components. Wang *et al.* (2021) employ this metric to measure the effectiveness of their proposed model for predicting human hand-over intentions. Similarly, Chao *et al.* (2010) utilize accuracy to quantify the learning performance of robots in recognizing symbols from human instructors. Another study Kim *et al.* (2017) focus on using

accuracy to evaluate a robot’s capability in recognizing gestures and accurately mapping these gestures to corresponding robot actions in parallel. This widespread adoption of accuracy as a measure underscores its significance in designing interactive robot learning systems. It suggests a design guideline where accuracy should be prioritized alongside success rate, ensuring that each subtask is finely tuned for optimal overall performance.

Accumulated rewards or *returns* stand out as fundamental metrics in the realm of reinforcement learning (RL) as shown in Figure 3.2. They typically refer to the total reward an agent or robot accumulates over a duration and are usually calculated as the aggregate of the rewards multiplied by a discount factor over time. In many scenarios, there’s a strong link between accumulated rewards and success rates, typically because the highest rewards are often given upon task completion. Nonetheless, this correlation is not always consistent. For instance, in meta-reinforcement learning (meta-RL), policies may struggle to adjust to new tasks for improved performance, despite higher rewards in smaller tasks (Jiang *et al.*, 2021). Similarly, in inverse reinforcement learning (IRL), the ambiguity can result in multiple paths to high rewards, not necessarily translating to higher success rates (Chen *et al.*, 2020c). However, unlike the success rate, which only evaluates the final outcome, the accumulated rewards or returns offer a more comprehensive insight into the process of the motion of the agent or robot, capturing the quality of actions from start to finish. Moreover, accumulated rewards sometimes serve as the constant metric (achieving the same rewards with less demonstrations or samples) to measure the sample efficiency (Mueller *et al.*, 2018), which will be discussed later in this section. Accumulated rewards or returns are also widely used in interactive robot learning with RL framework. Moreira *et al.* (2020a) use returns to prove their proposed algorithm was better than the baseline algorithm in object manipulation. Cruz *et al.* (2016) employ accumulated rewards to assess their interactive RL method against baselines, validating the method’s performance. Hoque *et al.* (2023) used accumulated reward as a key metric to compare their Fleet-Dagger algorithm with other baseline algorithms. Li *et al.* (2020a) apply accumulated reward to evaluate their HRL4IN algorithm in robot navigation against baseline

models. Lee *et al.* (2021) measure the performance of their proposed unsupervised PEBBLE algorithm against baselines using returns. The design guidelines in these studies focus on ensuring the reward shaping within the learning framework is representative and consistent, which is crucial for effectively comparing different RL algorithms.

Sample efficiency refers to an algorithm’s ability to achieve a desired performance level with fewer training samples (Rajeswaran *et al.*, 2018). In interactive robot learning, sample efficiency can also be reflected by achieving a desired performance with fewer number of demonstrations or corrections made by human (Mueller *et al.*, 2018). Sample efficiency is crucial because it is time consuming and expensive to obtain human demonstrations. This metric is usually used for the scenarios where robot learns from human teacher, as the involvement of human expertise can significantly accelerate the robot’s learning process and reduce the number of learning epochs required to meet specific performance benchmarks (Argall *et al.*, 2009a). Note that sample efficiency differs from temporal efficiency. The former emphasizes learning effectively from a limited number of interactions, whereas the latter focuses on the amount of wall-clock time to finish a task with or without human in the loop. We delve into the implications of temporal efficiency on *responsiveness* in subsequent discussions.

To evaluate sample efficiency, it is essential to measure other metrics as constant metrics, such as success rate (Long *et al.*, 2023), accumulated rewards (Subramanian *et al.*, 2016), accuracy (Nair *et al.*, 2020), or other metrics within the interest. Many works in interactive robot learning proposed methods to improve the sample efficiency with human in the loop. Zhang *et al.* (2023) show human feedback on robot learning can significantly improve the sample efficiency for learning object moving task compared to pure RL frameworks such as soft actor-critic (SAC) algorithm. Mueller *et al.* (2018) illustrate the sample efficiency can be significantly improved for placing and pouring task with human corrections and annotated constraints. Karamcheti *et al.* (2022) introduce language-informed learning to improve the sample efficiency across multiple tasks on a robot arm. The design guidelines for employing sample efficiency as a metric are particularly pertinent in scenarios where robots are challenging to train using pure Reinforcement Learning (RL)

algorithms or when the availability of human demonstrations is limited.

Responsiveness in the context of interactive robot learning and operation is defined by several key factors, including execution time, human idle time, and robot idle time. It encapsulates the system's ability to react and adapt to inputs or changes in the environment promptly. This refers to the duration a robot takes to complete a given task or respond to a command. Shorter execution times indicate higher responsiveness, as the robot can quickly carry out actions or adapt to new instructions. Long execution time can impair a robot's performance, particularly in fast-paced, dynamic environments such as autonomous driving (Schrum *et al.*, 2023b) or playing sports (Zaidi *et al.*, 2023). Utilizing large deep learning models can achieve high performance (Mahler *et al.*, 2017; Mahler *et al.*, 2018), but their long execution time poses significant challenges on the robot's ability to frequently update its planning to a dynamic environment. Many works used execution time as one of the metrics to evaluate the robot system. Morrison *et al.* (2020) use this metric to assess their real-time reactive grasping system, while Kivrak *et al.* (2021) employ it to measure path planning efficiency in dynamic environments for social robots. Moreover, robot and human idle time are also crucial in human-robot collaboration, serving as measures of team fluency (Hoffman, 2019). Casalino *et al.* (2019) use both human and robot idle times to evaluate the efficiency of their scheduling method in human-robot assembly tasks. Lasota and Shah (2015) and Nikolaidis *et al.* (2015b) also utilize these metrics to assess their human-aware motion planning algorithms and improve human-robot teaming efficiency, respectively. The design guideline for execution time hinges on whether the robot operates in a dynamic environment. For human and robot idle times, the key consideration is the involvement of human-robot collaboration and ensuring that users or robots do not experience significant delays or inefficiencies during the interaction.

In conclusion, we've examined key metrics essential for assessing interactive robot learning performance. Success rate is a core metric evaluating the robot's task completion capability, from physical to social interactions. Accumulated rewards or returns, particularly relevant in Reinforcement Learning, offer a holistic view of performance over time.

Accuracy, important for evaluating robot perception, provides nuanced performance insights. Sample efficiency highlights the algorithm's ability to achieve outcomes efficiently with minimal training or human intervention, underlining the importance of optimizing human input. Responsiveness, which includes execution, robot idle, and human idle times, is vital for seamless human-robot interactions in changing environments. Collectively, these metrics create a comprehensive framework for evaluating and improving interactive robot learning.

8.2 Metrics for Explainability

Evaluating an explainable robot system involves two types of metrics: objective and subjective. Objective metrics focus on quantifiable aspects of robot task performance as discussed. These metrics offer clear, measurable data on how well the robot performs tasks alongside its explanation component. Subjective metrics assess user perceptions, satisfaction, trust, and situation awareness. These metrics capture the human experience, including how users understand, interact with, and feel about the robot, which objective metrics cannot quantify. This section discusses both objective and subjective metrics and their application in recent interactive and explainable robot learning studies.

8.2.1 Objective Metrics

Task performance metrics, which discussed in section 8.1, are usually adopted as objective metrics in explainable robot learning. These metrics are not directly used to assess the explainability, but provide a objective way to measure the impact of the explainability on the performance of a robotic system. For instance, success rate is adopted by Mitrevski *et al.* (2020) as a measure of impact of explainability on object manipulation tasks. Cruz *et al.* (2022) use success rate as a metric to measure the impact of explainability on non-expert end-users. Sample efficiency and accuracy are both employed by Riley and Sridharan (2019) to assessed the explainability impact of their framework, which aids in understanding objects, relationships, and outcomes in a scene.

However, we also observe other performance metrics, which are not widely used in interactive robot learning, but adopted in explainable robot learning. F1 score, a variation of accuracy, are largely used in explainable robot learning to evaluate the performance impact from explainability where the data has false positive and false negative labels. For instance, Kulesza *et al.* (2015) exams the changes in F1 score to assess explanatory debugging in human-robot interaction, addressing challenges in building mental models by users. Das *et al.* (2021b) utilize F1 score as metrics to evaluate explainability for failure identification. Recall is another variation of accuracy metric used in explainable robot learning. For instance, Das and Chernova (2021) use recall to evaluate the explainability of a system. Recall is chosen as a metric in this paper to measure the failure identification. Compliance ratio is also a task performance in explainable robot learning. Pynadath *et al.* (2022) use compliance ratios to assess the impact of explainability, providing insights into user adherence and correct implementation of robot guidance or explanations.

In conclusion, task performance in interactive robot learning can be used as objective metrics in explainable robot learning. But we also observe the nuanced difference in task performance metrics between explainable robot learning and interactive robot learning lies in the requirement for the robot to provide explanations to humans. F1 score has a better balance between the precision and recall. Additionally, the compliance ratio, measuring the human willingness to accept the explanation, is also a significant factor for explainable robot learning.

8.2.2 Subjective Metrics

Subjective metrics play a pivotal role in interactive robot learning, particularly in assessing the explainability of robot systems. Explainability, the ability of a robot to clearly articulate its actions and decisions to human users, is vital for user comprehension and effective human-robot interaction. To measure the quality of explainability, it is essential to obtain direct feedback from users, typically obtained through well-structured user studies, questionnaires, or interviews. This feedback provides valuable insights into user perceptions and experiences, making

subjective metrics a direct and impactful means of measuring explainability in human-robot interactions.

The Likert scale, as noted by (Joshi *et al.*, 2015), stands out as a predominant method for assessing explainability in user studies. For instance, Hoffman *et al.* (2023) craft questions such as "The explanation of how the [software, algorithm, tool] works is satisfying" to gauge user satisfaction. Similarly, to evaluate human trust in robot systems, Hoffman *et al.* (2023) utilize inquiries like "I am confident in the [software, algorithm, tool]. I feel that it works well." Despite its widespread application, recent research has highlighted a significant drawback of relying on single-prompt Likert scales. Such an approach, though straightforward, is problematic due to the nature that a high degree of measurement error arises when a participant is asked to respond only to a single prompt Schrum *et al.* (2020). A more effective strategy for capturing subjective metrics involves averaging feedback across a series of prompts. For example, to measure human trust, a diverse set of statements could be used, including "Most robots make poor teammates," "Most robots possess adequate decision-making capability," "Most robots are pleasant towards people," and "Most robots are not precise in their actions," as suggested by (Mittu *et al.*, 2016). Consequently, the construction of a Likert scale represents a critical consideration in the evaluation of subjective metrics, necessitating meticulous attention to ensure reliability and validity.

User satisfaction emerges as a key subjective metric in the realm of interactive and explainable robot learning. This metric has been widely employed to evaluate the impact of explainability on user perceptions. For example, Cantucci and Falcone (2022) use user satisfaction to compare the effectiveness of a robotic tour guide with and without explainability features in an art museum setting. Similarly, Adamik *et al.* (2022) apply this metric to assess user satisfaction regarding the explainability of a robotic system in image processing tasks.

Trust in robotic systems is another crucial subjective metric. Research such as Gao *et al.* (2020b) use trust as a measure for evaluating the quality of automatically generated explanations by robots. This includes assessing how the clarity and relevance of these explanations influence the trust users place in the systems. Liu *et al.* (2021a) also

emphasize trust in their human study, demonstrating that robotic systems with enhanced explainability yield better performance and greater user trust. These studies highlight the significant interplay between explainability and trust in shaping user experiences with AI and robotic systems.

Usefulness, as a subjective metric, encapsulates a broad range of aspects and is often defined by amalgamating multiple metrics. For instance, Tambwekar and Gombolay (2023) combine ease of use Belanche *et al.* (2012), attitude, trust, and intention Belanche *et al.* (2012) to use as metrics for measuring explainability, subsequently using overall usefulness as a gauge for performance. Similarly, Olivares-Alarcos *et al.* (2023) employ relevance, understandability, and interpretability to assess explainability, with usefulness serving as an overarching metric for performance evaluation. These approaches underscore the multifaceted nature of usefulness as a metric, integrating various dimensions to assess the overall performance of robotic systems.

Situation awareness (SA) is a subjective metric to evaluate the individual's perception and understanding of the situation (Endsley, 1995). Situation Awareness Global Assessment Technique (SAGAT) is often employed to collect subjective evaluations of situation awareness (Endsley, 1995). Paleja *et al.* (2021) utilize SAGAT to measure human machining teaming. Endsley (2023) adapt SA to measure the team situation awareness and showed improvements in AI transparency for Tesla autopilot. These applications demonstrate SA can be used for evaluating how well the system's model of the environment matches the human operator's understanding or expectations.

In this section, we explore the metrics for evaluating the explainability of interactive robot learning systems. We distinguish between objective and subjective metrics, underscoring their respective roles in quantifying task performance and capturing human perceptions of explainability. Objective metrics, including F1 scores, accuracy, recall, and success rates, provide tangible benchmarks for assessing the effectiveness of explainability in improving robotic task performance. These measures, highlighted through various studies, underscore the hypothesis that enhanced explainability leads to better user engagement and task efficiency. Subjective metrics underline the significance of user-centric

evaluations in interactive robot learning, particularly in the realm of system explainability. Metrics such as user satisfaction, trust, and usefulness, predominantly derived from user feedback and perceptions, offer invaluable insights into how humans interact with and comprehend robotic systems. Likert scale is one of the most common way to measure the subjective metrics, but the sufficient and valid prompts are needed to measure the metrics accurately.

8.3 Metrics for Safety

In the field of interactive robotic learning, ensuring the safety of both humans and the robots themselves is paramount as human will closely interact with robots. This involves various measures to enhance the physical safety or the perceived safety. As discussed in 6.1, physical safety refers to avoiding physical harm or injuries when the robot is interacting with its environment and users. One typical metric to measure the physical safety is collision rate. Collision rate is often used to measure the safety of a robot's navigation system, with lower rates indicating better avoidance of obstacles and safer interaction with the environment (Nishimura and Yonetani, 2020; Spencer *et al.*, 2022). Shao *et al.* (2021) use collision rate to measure their reachability approach for safe navigation of self-driving vehicle and UAVs. Yang *et al.* (2022b) utilize collision rate to evaluate their proposed control barrier function method, which achieved safe inverse reinforcement learning.

Additionally, the quantification of uncertainty or confidence in a robot's decision-making can serve as a measurement for the safety of its policies, especially in learning scenarios (Kelly *et al.*, 2019; Hu *et al.*, 2023). The probabilistic nature of this approach not only enables users to set specific confidence thresholds (such as a 95% confidence level) for the robot's actions but also facilitates the measurement of uncertainties during interactive robot training (Menda *et al.*, 2019). This, in turn, aids in determining the optimal number of demonstrations required for effective robot learning (Brown and Niekum, 2018; Brown and Niekum, 2017).

In contexts where robots and humans cooperate, risk assessments consider factors such as human states, transitions, and proximity to

ensure safety Hayashi *et al.* (2022). Moreover, achieving perceived safety is crucial in human-robot interactions, even when robots are physically safe. The psychological impact on humans when cooperating with robots warrants significant consideration. For example, a robot's rapid movements and proximity to a human can be perceived as a threat of danger or physical harm. This aspect of perceived safety can be assessed through user feedback, as indicated by studies from Akalin *et al.* (2022), Marta *et al.* (2021), and Van Waveren *et al.* (2023b). Akalin *et al.* (2022) also highlight a strong correlation between trust, sense of control, comfort, and perceived safety, emphasizing the multifaceted nature of safety perceptions in human-robot collaborations. Furthermore, indices such as the Social Individual Index (SII) and Relative Motion Index (RMI), measure the psychological safety of humans around moving robots, indicating how these interactions affect human perceptions and reactions Kivrak *et al.* (2021).

In other scenarios, specific metrics were adopted due to their relevance and utility. For instance, in autonomous driving, a key metric is the rate of human intervention, which is a critical indicator of the system's safety and reliability (Wang *et al.*, 2022b). The more humans intervene with the autonomous car, the less reliable the system is. In the context of surgical robotics, emphasis is placed on operative safety, with significant metrics including the frequency of tool-to-tool collisions and the number of usage of controls like clutch pedals Raza *et al.* (2014). These specialized metrics, adapted to their unique applications and contexts, are vital for ensuring the safe and harmonious integration of robots into human-centric environments.

8.4 Benchmarks

Interactive robot learning is a dynamic field where standardized benchmarks play a critical role in shaping advancements. These benchmarks are designed to test robotic systems in diverse scenarios, focusing on their ability to learn and adapt through interaction. From simulated environments to physical manipulation tasks, each benchmark addresses specific challenges and contributes uniquely to the development of intelligent, adaptable robotic systems. This section summarizes the benchmarks

used in our topic of interactive and explainable robot learning. We also list relevant papers and provide inside how the benchmarks are applied to these studies.

OpenAI Gym (Brockman *et al.*, 2016) is a comprehensive toolkit designed specifically for research in reinforcement learning. It features an expanding array of benchmark problems unified by a common interface, facilitating ease of use and consistency in testing. The OpenAI Gym environment is also adaptable for use in interactive robot learning benchmarks. For example, Diaz *et al.* (2021) utilize Duckietown (Chevalier-Boisvert *et al.*, 2018), a simulator for self-driving cars within the OpenAI Gym framework, to evaluate the performance of interactive imitation learning. Similarly, Juan *et al.* (2021) employ OpenAI Gym environments to test their Reinforcement Learning (RL) method, which incorporates human feedback, on tasks like CartPole and Mountain-Car. These instances demonstrate the versatility of OpenAI Gym in supporting a wide range of interactive robot learning research.

The Open X-Embodiment dataset (Padalkar *et al.*, 2023) is one of the most extensive resources available for robot learning, collected with 527 skills across 160,266 tasks. This dataset is particularly useful for training generalized models in robotics that can handle a vast number of tasks. Data collection methods in this dataset are diverse, including human VR, expert policy, human kinesthetic, human spacemouse, joystick, keyboard, and puppeteering techniques. Additionally, much of the data includes annotations in either templated or natural language. Both RT-1 (Brohan *et al.*, 2022) and RT-2 (Brohan *et al.*, 2023a) utilize this dataset for training purposes. These attributes make Open X-Embodiment a vital tool for advancing the capabilities of robotic systems.

Habitat 3.0 (Puig *et al.*, 2023) is a photo-realistic simulator platform designed to study two typical tasks in human-robot interaction (HRI): social navigation and social rearrangement in home-like environments. This platform features accurate humanoid avatar simulation and provides human interaction interfaces such as mouse, keyboard, and VR. Lidard *et al.* (2024) utilize this platform to test their algorithm, which involves a Boston Dynamics Spot robot and a human collaborating to clean a house. Raychaudhuri *et al.* (2024) conduct experiments in Habitat to validate their approach for the object-goal navigation task.

These studies show the effectiveness of Habitat 3.0 in advancing research in HRI.

RLBench (James *et al.*, 2020) is an extensive robot learning benchmark suite designed for research in reinforcement learning and imitation learning. It features 100 unique, hand-designed tasks, ranging from simple single-objective tasks to complex, multi-stage tasks. Although primarily intended for reinforcement learning and imitation learning research, RLBench is also widely used in interactive robot learning studies to assess performance. Chu *et al.* (2023) use this benchmark to validate their proposed language feedback interactive RL. Liu *et al.* (2022) employ RLBench to evaluate their instructRL method, which integrates both vision and language representations to predict robot actions. These applications show RLBench’s utility in diverse research scenarios.

The DeepMind Control Suite (Tassa *et al.*, 2018) presents a collection of continuous control tasks, designed with standardized structures and clear, interpretable rewards. It is also suitable for benchmarks in interactive robot learning. Huang *et al.* (2023) employ this suite to evaluate their interactive, vision-based Reinforcement Learning (RL) framework for self-driving cars. Additionally, Liu *et al.* (2023c) carry out comprehensive experiments using the DeepMind Control Suite benchmark to assess their algorithms designed for learning from multiple oracles. These uses underscore the flexibility of DeepMind Control in accommodating a variety of research needs, especially in the realms of interactive and explainable robot learning.

AI2-THOR, or The House Of inteRActions, (Kolve *et al.*, 2017) is a comprehensive framework developed for visual AI research. AI2-THOR supports a broad spectrum of research areas, extending from deep reinforcement learning and imitation learning to interaction-based learning, strategic planning, visual question answering, unsupervised learning for representation, object recognition and segmentation, as well as cognitive model development. In the AI2-THOR environment, Lu *et al.* (2022) implement their Atomic Skill Completion method to train agents for four interactive tasks simultaneously, achieving remarkable effectiveness. Similarly, Wu *et al.* (2021a) create a Virtual Reality (VR)-based 3D simulation environment leveraging the AI2-

THOR platform. This was used to explore the communicative learning of natural gestures in robotics. These applications demonstrate its versatility and effectiveness as a platform for advanced research in interactive task training and communicative learning in robotics.

RoboTurk (Mandlekar *et al.*, 2018a) emerges as a solution to the limitations of Imitation Learning in robotic manipulation. This innovative crowdsourcing platform enables high-quality 6-DoF trajectory-based teleoperation using common mobile devices like iPhones. These interactive features make RoboTurk a popular platform for data collection and benchmarking in the realm of interactive robot learning. For instance, Clever *et al.* (2021) investigate transformer model predictions for seven different manipulation tasks, including pick-and-place operations in industrial, household, and caregiving robotics, using a substantial dataset of over 500 demonstrations from the open-source RoboTurk. Additionally, Mandlekar *et al.* (2020) enhance the RoboTurk system to allow remote users the capability to oversee policy execution and intervene as necessary through a web browser and smartphone. These studies illustrate that RoboTurk serves as an effective platform for the interactive collection of data and build benchmarks.

Mujoco haptix (Kumar and Todorov, 2015) have created a virtual reality setup that integrates real-time motion capture, physics-based simulation, and stereoscopic visualization. This system allows users, equipped with a CyberGlove, to virtually 'reach into' the simulation and interact with objects using a remotely operated virtual hand. The effectiveness of this system is assessed using selected tasks from the Southampton Hand Assessment Procedure, a clinically recognized hand function evaluation. Additionally, this system is employed in the DARPA Hand Proprioception and Touch Interfaces program by teams developing neural control interfaces within a simulated environment. The ease of tracking hand motion makes the Mujoco Haptix a popular choice for benchmarking in studies related to dexterous hand control. Rajeswaran *et al.* (2017) utilize the Mujoco HAPTIX system to capture finger movements to convert real-world demonstrations into virtual reality. Similarly, Nemec *et al.* (2017) implement the MuJoCo HAPTIX simulation environment to model a KUKA robot's interaction for the task of opening doors. These implementations of the Mujoco HAPTix

in varied research contexts highlight its effectiveness and precision in simulating complex hand movements and interactions within virtual environments.

Robosuite (Zhu *et al.*, 2020) serves as an advanced simulation framework for robot learning. It provides a collection of benchmark environments designed for reproducible research in this field. Lin *et al.* (2023b) assess their interactive hierarchical policy learning system by testing it on four complex manipulation tasks within the robosuite framework. Meanwhile, Liu *et al.* (2023b) apply their approach in an interactive imitation learning setting using the robosuite simulator. Furthermore, Hoque *et al.* (2021) conduct an evaluation of their ThriftyDagger method on a challenging long-horizon peg insertion task, also within the Robosuite simulation environment. These studies collectively demonstrate the robustness and versatility of the robosuite simulation environment in facilitating the development and testing of advanced robotic learning methods across a range of complex tasks.

The Interactive Gibson Benchmark (Xia *et al.*, 2020) is a pioneering platform in the field of robotics, specifically designed for evaluating solutions in Interactive Navigation. This benchmark is an extension of the original Gibson Environment, offering a comprehensive setup for training and testing robotic navigation systems. What sets it apart is its focus on interaction: robots are not just navigating through static environments, but are also engaging with and manipulating objects within these spaces. Li *et al.* (2022) showcase their transformer-based memory technique in the iGibson environment. Pérez-D’Arpino *et al.* (2021) deploy their navigation strategy in the interactive Gibson simulator, training it to steer a simulated non-holonomic wheeled robot across various settings. Additionally, Schmalstieg *et al.* (2023) utilize eight training scenes from the iGibson challenge to test their hierarchical interactive multi-object search method. These varied applications in the iGibson environment illustrate its capacity to effectively test and refine innovative robotic navigation and interaction methods, contributing significantly to advancements in the field of robot learning.

Meta-World (Yu *et al.*, 2020) is an innovative, open-source benchmark specifically designed for meta-reinforcement learning and multitask learning in robotics. The primary aim of Meta-World is to facilitate the

development of algorithms that are capable of generalizing their learning experience, thereby accelerating the acquisition of new, unseen tasks. Goyal *et al.* (2021) use Meta-World to test their language reward in robot learning. Furthermore, Hansen *et al.* (2022) utilize Meta-World as a testing ground to verify the effectiveness of their framework designed to speed up model-based RL algorithms through demonstrations. These experiments in the Meta-World environment underscore its utility as a versatile and challenging platform for testing and refining various aspects of robotic manipulation and learning strategies.

CALVIN (Composing Actions from Language and Vision) (Mees *et al.*, 2022b) is a cutting-edge, open-source simulated benchmark aimed at advancing the field of robotic manipulation. It focuses on teaching robots to perform long-horizon tasks that are specified solely through human language and interpreted via onboard sensors. This benchmark is designed to bridge the gap between natural language processing and physical task execution, enabling the development of AI agents capable of understanding and executing complex instructions in a variety of settings. Many works (Mees *et al.*, 2022a; Sheikh *et al.*, 2023) utilize CALVIN to develop language conditioned policy for robot learning, and these works highlight the CALVIN benchmark’s effectiveness in facilitating innovative research and development in language-driven, long-term policy learning in robotics.

While there are benchmarks for human-agent interaction involving various interfaces like smartphones, hand sensors, and language, specific benchmarks for explainable robotics techniques are not as prevalent. In the broader field of machine learning, some benchmarks have been established for Explainable Artificial Intelligence (XAI), such as those by Liu *et al.* (2021b) and Holmberg (2022). However, these benchmarks primarily focus on static AI models and do not fully capture the dynamic interactions inherent in robotics, where explainability must encompass the robot’s physical actions and decisions in real-time and in complex, often unpredictable environments. The unique requirements of interactive and explainable robot learning call for the development of specialized datasets and benchmarks that can effectively evaluate how well a robotic system communicates its intentions, decisions, and actions to human users. Moreover, the development and implementation

of subjective metrics for benchmarks in explainable robotics require a uniform and standardized approach to evaluation. The subjective nature of explainability—how users perceive, understand, and trust the explanations provided by robots—requires careful consideration in how these aspects are measured. Without a consistent methodology, comparisons across different systems and studies can become challenging, leading to potentially misleading conclusions about the effectiveness of explainability approaches.

8.5 Conclusion

In this chapter, we delve into the crucial aspects of evaluation, metrics, and benchmarking within the realm of interactive and explainable robot learning. It underscores the significance of metrics and benchmarks in validating the performance, safety, and explainability of robotic systems, which is critical for their effective integration across various sectors like healthcare, manufacturing, and service industries. Key metrics such as success rate, accuracy, accumulated rewards, sample efficiency, and responsiveness are highlighted for their roles in ensuring that robots perform tasks accurately, reliably, and in ways that are understandable and trustworthy to humans.

Looking towards the future, we suggest a continued focus on developing comprehensive and versatile benchmarks that can adapt to the evolving needs of interactive and explainable robot learning. The need for benchmarks that can rigorously test and validate new approaches in real-world scenarios is evident.

In conclusion, we suggest that advancing the field of interactive and explainable robot learning requires a concerted effort in refining evaluation strategies, expanding upon existing metrics, and creating dynamic benchmarks. Such endeavors will not only foster trust and reliability in robotic systems but also improve the field towards innovative discoveries, enhancing human-robot collaboration across all sectors.

9

Applications

Interactive and explainable robot learning are rapidly growing fields that are increasingly being deployed in a multitude of industries. From cleaning robots in the home to heavy-duty manipulators in the warehouse, interactive robot learning has the potential to revolutionize automation by endowing robots with skills difficult to achieve through traditional learning or control methods by leveraging human expertise from the domain where these robots are deployed and improving using human feedback. The vast majority of robots in the real-world are deployed with fixed capabilities and often fail to adapt to new environments and users. Furthermore, there is a need for explainable robot learning methods to build trust and increase user acceptance of robots into everyday life (Section 3.9). Thus far, explainable robot learning methods have been relegated to frameworks limited to specific domains in simulation, but have yet to be rigorously evaluated on any real-world robotic systems (Sakai and Nagai, 2022).

To highlight the potential of interactive robot learning for real-world impact, we examine existing literature that utilize interactive robot learning on a variety of domains in this section, which include healthcare, industrial automation, domestic and household environments, and robot

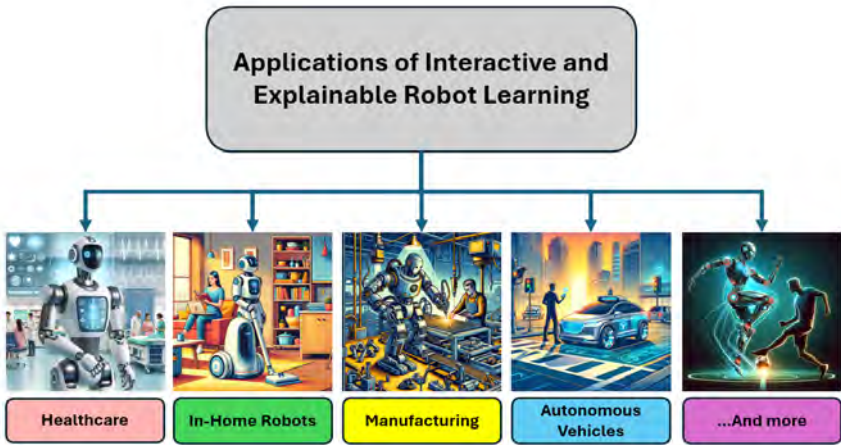


Figure 9.1: Overview of the applications of interactive robot learning covered in this section.

navigation. For each of these application areas, we discuss domain-specific challenges that hinder the large-scale deployment of robots with interactive learning capabilities. We additionally provide some insights as to where explainable robot learning is needed to improve robot adoption rates for these real-world domains. Lastly, we conclude by introducing unique applications of interactive robot learning that do not necessarily fit into any of these categories, which can provide novel perspectives and methods towards solving application challenges of interactive robot learning methods.

9.1 Healthcare

Within the healthcare industry, interactive robot learning is an attractive method for assistive and medical robots to acquire new skills towards the treatment of patients (Okamura *et al.*, 2010; Kyrarini *et al.*, 2021). Robotic-assisted surgery can provide surgeons with extreme precision on the operating table using a robotic system, such as the da Vinci Surgical System, a tool that provides demonstrations for learning-based applications. Several interactive robot learning methods have been proposed for surgical applications, for instance, in soft tissue

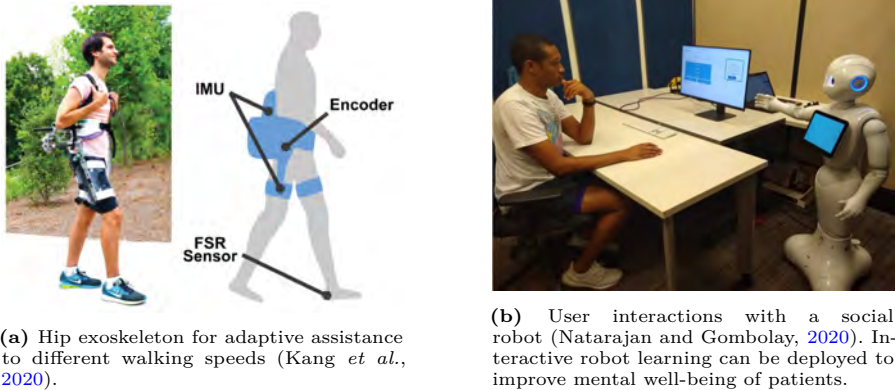


Figure 9.2: Example applications of interactive robot learning for healthcare

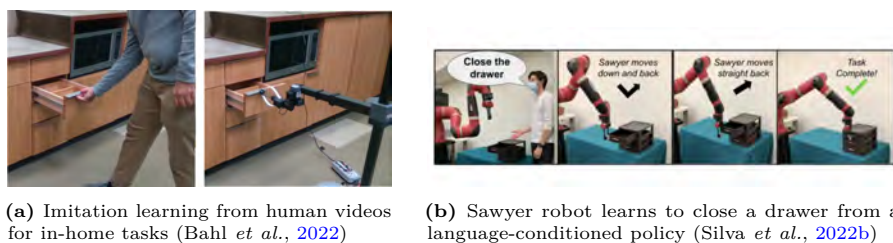
manipulation (Pore *et al.*, 2021), using a novel variant of GAIL (Section 3.4.3) and ophthalmic microsurgery (Keller *et al.*, 2020), addressing the sensory and dexterity limitations of human surgeons by utilizing an RL policy warm-started with behavioral cloning. LfD has also been proposed for navigation of a surgical tool in eye surgeries (Kim *et al.*, 2020a).

Beyond their emerging role in surgical aid, hospitals are increasingly integrating robotics for logistical support and patient care. In particular, interactive robot learning has been harnessed to enhance the coordination of patient care, with robots learning scheduling policies directly from nurse-led demonstrations of resource distribution, ultimately providing a decision support system through (Gombolay *et al.*, 2018a). Furthermore, interactive robot learning has shown promise in physical rehabilitation, aiming to assist patients in the restoration of mobility and the enhancement of their motor skills. Interactive robot learning methods were applied to improve gait rehabilitation through exoskeletons (Kang *et al.*, 2020; Shepherd *et al.*, 2022), examples including an RL policy which is able to adapt its physical parameters (e.g. stiffness, damping, and inertia coefficients) to cater to individual needs (Bingjing *et al.*, 2019), a imitation learning system for robot assisted rehabilitation exercises using dynamic movement primitives to model normal limb movement and identify abnormal forces resulting

from the patient's mobility impairment (Escarabajal *et al.*, 2023), and a POMDP framework in a robotic system for adaptive resistance in upper limb stroke rehabilitation exercises (Kan *et al.*, 2011). Each of these methods provide an automated, adaptable framework tailored to individual patient needs, facilitating a personalized approach to mobility restoration.

Recently, socially assistive robots (i.e. robots that assist people through social interactions (Feil-Seifer and Mataric, 2005)) have emerged as an avenue for improving mental health (Guemghar *et al.*, 2022; Koh *et al.*, 2021) and act as an alternative for social interaction. While such robots are not a complete substitute for human contact (Moyle *et al.*, 2016; Casey *et al.*, 2020; Fiske *et al.*, 2019), social robots have been shown to reduce depression and loneliness for patients with major depressive disorder in some cases (Chen *et al.*, 2020f). Interactive learning methods have been proposed and evaluated on social robot platforms to act as a companion robot for the elderly by utilizing RL with human feedback to have conversations in different situations (Yang *et al.*, 2017), adapting robot responses to varying empathetic levels, which are captured from the user through visual and task-related features, and transformed into a measure of valence of feeling (Leite *et al.*, 2012), and deployed as a personalized social robot to aid the educational development of children with autism spectrum disorders (Clabaugh *et al.*, 2019). Research involving assistive robots with ToM (Section 5.2.1) capabilities in healthcare settings has demonstrated that when robots could make educated guesses about the patients' mental states or intentions, the level of user satisfaction and compliance increased substantially (Cifuentes *et al.*, 2020; Broadbent *et al.*, 2009; Tapus *et al.*, 2009). These methods highlight the potential of socially assistive robots to cater to diverse individual needs in the patients' mental well-being. However, Parks, 2010 cautions against the negative consequences of substituting the "human touch" with robots, highlighting the potential for increased isolation.

Despite the advancements within interactive robot learning in the healthcare domain, few explainable robot learning techniques have been developed for healthcare. Thus far, explainable AI in the healthcare domain have been focused on imaging and diagnosis systems (Chaddad



(a) Imitation learning from human videos for in-home tasks (Bahl *et al.*, 2022)

(b) Sawyer robot learns to close a drawer from a language-conditioned policy (Silva *et al.*, 2022b)

Figure 9.3: Example applications of interactive robot learning for household tasks

et al., 2023; Hulsén, 2023), but not on robots. Setchi *et al.* (2020b) presents a hypothetical scenario of how explainable robot learning can be used to help assist an elderly person with dementia. In this example, the patient with slurred speech is confused about why the robot decided to fetch their dentures, where the robot then explains that it has detected a change in the patients' speech patterns and needs the dentures to speak clearly. Setchi *et al.* (2020b) additionally motivates the need for more research on explainable robot learning methods to understand how to construct human-friendly and trustworthy explanations using analogies, past experiences and real-time information.

While interactive robot learning has shown great potential in improving healthcare robots, there remain several unique challenges that must be addressed to achieve democratization of these robots in the real-world. Notably, one unique challenge of interactive learning in the healthcare sector is the level of data security and privacy measures that must be deployed when collecting and storing sensitive data from the interactions between patients and robots. Interactive robot learning practitioners targeting healthcare applications should thoroughly understand the ethics implications presented in Chapter 7. Safety of the learning methods is another significant challenge to overcome. Interactive learning methods require interaction data to improve, thus requiring robust safety measures while the robots are learning to avoid harming patients, a population that is already in a vulnerable state. Practitioners in the healthcare domain are advised to be familiar with the safety concepts in Chapter 6.

9.2 Domestic and Household Tasks

Interactive robot learning has the potential to greatly enhance the functionality and adaptability of robots to serve personalized needs in the home. When properly applied, these learning methods enable household robots to adapt to the habits of the inhabitants, as well as the specific configurations of different homes by learning individual tendencies through interactions, such as how a person prefers their plates to be stacked, or when a person prefers to have their meals. Much of interactive robot learning in the home has focused on various forms of human activity recognition to learn robust robot policies for imitation (Bahl *et al.*, 2022; Shah *et al.*, 2020a). For example, Bahl *et al.* (2022) introduced a method that attempts to extract human intent from human videos, then use this information to improve the policy with a sampling-based policy optimization approach, demonstrating one-shot generalization over 20 different everyday household tasks; while in Shah *et al.* (2020a) a robot learned a policy with off-policy RL to set up a dinner table by inferring task specifications in linear temporal logic (LTL) from human demonstrations of the task.

Several methods have been developed that leverage human feedback to enhance policy learning and improve adaptation to varying user requirements. For instance, a deep reinforcement learning framework that incorporates feedback from a human advisor during training for policy-shaping (Moreira *et al.*, 2020b). This approach has demonstrated that the integration of interactive feedback can expedite the learning process in a domestic object organization task by allowing the advisor to impart prior knowledge related to the task at hand. Shek *et al.* (2023) proposed a framework that takes an object-centric approach that utilizes a single instance of human feedback (i.e. one-shot adaptation) and presented good adaptive performance in a can fetching task, showing the value of incorporating human feedback to improve sample-efficiency. Furthermore, in Tseng *et al.* (2018), a reinforcement learning framework was introduced that employs reward shaping based on human feedback observations. This framework was successfully implemented in an in-home service robot, demonstrating an improved capacity for adapting to individual user needs and preferences. Robots equipped with ToM can

better assist in household tasks by understanding and adapting to the preferences and habits of different family members, thereby becoming more integrated and accepted as part of the household (Kidd and Breazeal, 2008; Sung *et al.*, 2010; De Graaf *et al.*, 2016; De Graaf *et al.*, 2016). Such methods underscore the potential of human-augmented learning processes in improving the adaptability and efficiency of in-home robotics systems.

Learning methods leveraging natural language (Section 3.6) became ubiquitous with the advent of large language models (LLM) (Section 3.6.2), and interactive robot learning is no exception. Natural language is an important data modality used in interactive robot learning for in-home tasks, used to train robot policies that exhibit diverse behavior required for performing a wide-range of tasks in an unstructured home. Silva *et al.* (2022b) incorporates the contextual structure of natural language as prior knowledge of a task to a manipulator to learn multi-task skills, achieving high generalization to new, unseen tasks. In Lin *et al.* (2023a), a LLM is used as a framework to process both human gesture and language instruction to facilitate human-robot collaboration in tool fetching tasks, demonstrating the utility of language in learning policies that can adapt to humans in home environments. While foundation models offer great performance in generalization, LLMs have the potential to generate harmful content or "hallucinate" (Chapter 4). Due to the complex nature and diversity of tasks in an unstructured home environment, it is imperative that interactive LLMs utilized in robots are transparent and clearly communicate their intent to the user.

9.3 Industrial Automation

With an estimated 3.5 million industrial robots deployed across the world as of 2021 (International Federation of Robotics, 2022), warehouse robots stand to gain the most from improvements to interactive and explainable robot learning methods that can better streamline and automate tasks relevant to this domain. Many of these robots are deployed in e-commerce warehouses to sort, pack, and move goods. Some are used in manufacturing on the assembly line to build cars, aircraft, and electronics. Interactive robot learning can benefit these

types of settings in several ways, such as increasing efficiency and productivity, adapting to new tasks and environments, and enhancing human-robot collaboration in the workplace.

One form of interactive robot learning that has improved industrial automation is LfD (Section 3.4). LfD methods have been commonly applied for robotic task automation in manufacturing (Zhu and Hu, 2018; Ravichandar *et al.*, 2020; Kyrarini *et al.*, 2019). Tasks in this domain tend to be repetitive, making learning from demonstration particularly attractive as a quick way to teach a robot to perform a new skill without the need for reprogramming. While industrial environments are relatively static compared to a home environment, unexpected scenarios still occur regularly, especially in manipulation tasks with complex dynamics. Rethink Robotics, a robotics company once at the forefront of collaborative robots that aimed to revolutionize robotics in manufacturing with its simple-to-program Baxter and Sawyer product lines, was not able to survive due to precision and adaptation issues, highlighting the need for manufacturing robots to have strong capabilities of adaptation against perturbations and new tasks. Recent popular approaches tackle these challenges by improving on an LfD policy with reinforcement learning (Wang *et al.*, 2022c; Luo *et al.*, 2021) to increase robustness against new tasks. Some methods leverage human feedback online for additional training data when the robot does not know how to perform a task (Mandlekar *et al.*, 2023) or fails to perform the task (Luo *et al.*, 2021).

One unique element of integrating robots into manufacturing and warehouse domains is mixed-initiative and shared autonomy (Section 3.8). While robots excel at repetitive and precise tasks, they still struggle in scenarios requiring nuanced decision-making. Mixed human-robot teams can alleviate these issues by offering adaptability through real-time human input. Robots must then learn to collaborate efficiently with a human-in-the-loop. LfD has been explored extensively for robots to learn skills that are compatible for human collaboration in manufacturing (Nikolaidis *et al.*, 2015b; Maeda *et al.*, 2014; Rozo *et al.*, 2013). Due to the strength and speed of industrial manipulators, interactive learning approaches have focused on enhancing safety by predicting human motion and intention (Liu *et al.*, 2019), and planning with collision



(a) Human demonstrator teaches robot a peg-in-hole task via kinesthetic teaching (Abu-Dakka *et al.*, 2014)



(b) Collaborative robot working together with a human operator (Sauppé and Mutlu, 2015). Interactive robot learning can improve teaming efficiency by adapting to the human collaborator.

Figure 9.4: Example applications of interactive robot learning for manufacturing tasks

avoidance (Aivaliotis *et al.*, 2019) while maintaining good collaborative performance. Robots with ToM capabilities have improved efficiency and safety by anticipating human co-worker actions and adjusting their behavior accordingly (Gombolay *et al.*, 2015; Gombolay *et al.*, 2018b). Another aspect to mixed-initiative collaboration are the social capabilities of the robot. Results from Sauppé and Mutlu (2015) suggests that a warehouse robot's social capabilities affect how workers adapt to and interact with the robot, indicating that interactive robot learning methods adapting for social preferences of the workers may improve coordination effectiveness.

Furthermore, explainable robot learning is another avenue for potentially improve the collaborative effectiveness of human-robot teams. While few works have evaluated explainable frameworks on real-world manufacturing robots, explainable AI frameworks have been developed for improving manufacturing processes such as quality inspection systems (Goldman *et al.*, 2021; Lee *et al.*, 2022) and a decision-support system for supply chain management (Zajec *et al.*, 2021). Baum *et al.* (2023) suggested that lack of information and understanding on the human is the main reason for poor adoption of intelligent support systems in factory environments, motivating the need for explanatory information in AI decision-making. Similar types of evaluations on collaborative

robots are needed to access the efficacy of explainable frameworks on improving rates of robot adoption in manufacturing.

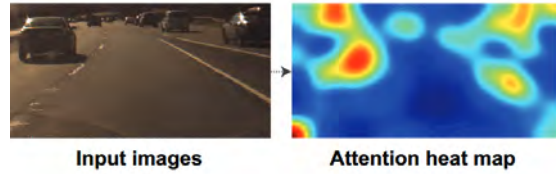
The greatest hurdle for interactive and explainable robot learning in industrial automation is the safety of the deployed methods. While the problem is also present for general robot learning, interactive robot learning in manufacturing forces frequent proximate interactions between the robot and human to learn, where the combination of this domain and learning approach has the greatest risk of physical human injury when taking into consideration the increased size, speed, and strength of industrial robots. Thus, interactive robot learning methods in the manufacturing domain need to incorporate safety mechanisms that specifically mitigate high-speed collision risks with users. Practitioners in industrial automation should thoroughly understand the concepts presented in Chapter 6. As a result of the immense safety concerns, there is a lack of data from human-robot teams in manufacturing, which further complicates the development of new interactive and explainable learning methods in this domain. To address these challenges, researchers should emphasize investigating human-centered approaches that provide robot systems with capabilities to better predict and adapt to its human collaborators with proactive safety measures as opposed to reactive safety measures (Lasota *et al.*, 2017), which in turn can improve the robustness and safety of these systems.

9.4 Autonomous Vehicles

Autonomous driving technology has emerged as a transformative force in the transportation sector, driven by the technology's potential to transform how people commute, reduce traffic accidents, and improve traffic efficiency. This has resulted in a dramatic increase in the interest and attention of autonomous vehicle (AV) research in recent years (Parekh *et al.*, 2022), with numerous research advances in path planning (Dai and Lee, 2020; Chiang *et al.*, 2020; Qiao and Wu, 2019), control (Yin *et al.*, 2015; Wasala *et al.*, 2020; Chebly *et al.*, 2019), and object detection (Gupta *et al.*, 2021; Chen *et al.*, 2020a; Jung *et al.*, 2020) in the self-driving domain. Similarly, interactive robot learning has emerged as a field that can enhance the capabilities of AVs. For example,



(a) Human demonstrators provide demonstrations in a driving simulator (Schrum *et al.*, 2022b).



(b) Visual explanations for an AV via an attention heat map. Figure adapted from Kim and Canny (2017).

Figure 9.5: Example applications of interactive and explainable robot learning for autonomous vehicles

interactive robot learning techniques applied in the self-driving domain can provide enhanced personalization of vehicle behavior (Schrum *et al.*, 2023b; Schrum *et al.*, 2022b; Schrum *et al.*, 2023a) by anticipating the needs of the passengers to offer a more comfortable driving experience. Other interactive learning techniques for AVs have focused on adapting to road users (Hu *et al.*, 2023; Chen *et al.*, 2023c; Tian *et al.*, 2022) – such as other vehicles, pedestrians, or cyclists – to increase safety and efficiency when operating in crowded scenarios where these interactions are common.

Recent accidents caused by AVs (Tesla Deaths 2024; Stanton *et al.*, 2019) have increased public scrutiny and hampered trust (Yurtsever *et al.*, 2020) on self-driving technology. Providing explanations is considered as one way to build confidence and trust in AV technology (Koo *et al.*, 2015; Kenesei *et al.*, 2022). Explainable robot learning techniques in AVs utilized vision-based explanations (Kim and Canny, 2017; Zeng *et al.*, 2019; Kim *et al.*, 2020b) to provide the user an opportunity to better understand how the AV senses its surroundings, as well as decision-tree based explanations (Brewitt *et al.*, 2021; Omeiza *et al.*, 2021a; Mankodiya *et al.*, 2021; Paleja *et al.*, 2022) for interpretability in AV decision-making. Atakishiyev *et al.* (2023) provides a comprehensive overview of explainable learning methods for AVs.

9.5 Other Applications

Other applications of interactive robot learning include, but are not limited to, sports (Lee *et al.*, 2023; Mülling *et al.*, 2013; Gomez-Gonzalez *et al.*, 2016; Abeyruwan *et al.*, 2022; Chen *et al.*, 2023a; Aşık *et al.*, 2019), dancing (Meng *et al.*, 2014; Tidemann and Öztürk, 2007), and art (Chen *et al.*, 2022a; Bidgoli *et al.*, 2020; Berio *et al.*, 2016) robots. Many of these applications utilize LfD for robots to imitate human behavior for performing tasks specific to each of these domains. Sports robots such as the table tennis robots (Lee *et al.*, 2023; Mülling *et al.*, 2013; Gomez-Gonzalez *et al.*, 2016; Abeyruwan *et al.*, 2022; Chen *et al.*, 2023a; Aşık *et al.*, 2019) require precise, agile movements, where learning of movement primitives is valuable for ensuring real-time execution of robot movement. Dancing robots require replication of complex human movements that are difficult to program by hand, thus researchers have explored imitating human dance motions from images (Tidemann and Öztürk, 2007), and human feedback in interactive RL to create new dance movements (Meng *et al.*, 2014). In the art domain, systems were developed to reproduce human-like calligraphic styles in graffiti (Chen *et al.*, 2022a; Berio *et al.*, 2016), and robotic painting through brushstroke imitation (Bidgoli *et al.*, 2020).

9.6 Design Guidelines and Future Work

Interactive and explainable robot learning represents a paradigm shift in how robots are programmed, interact, and learn from humans across various real-world domains. The advancement of this field is not only about improving the technical capabilities of robots but also about ensuring these systems can be effectively and safely integrated into human environments. Trust, safety, user adaptability, and ethical considerations form a foundation to which the future of interactive and explainable robot learning should be built. Drawing from the discussions across healthcare, domestic and household tasks, industrial automation, autonomous vehicles, and other innovative applications, we distill a set of key design guidelines for the development and deployment of future robot learning systems:

1. **User-Centered Design:** Prioritize the development of interactive learning systems that are intuitive, user-friendly, and tailored to the specific needs and preferences of diverse user groups. This approach ensures that robots can adapt to the individual habits and environments of users, enhancing user experience and satisfaction. Adaptivity to the user has been a recurring theme across all domains and incorporating feedback mechanisms (Chapter 4) with adaptive learning methods (Section 3.7.1) is key to enable user adaptivity.
2. **Safety and Ethical Considerations:** Implement robust safety protocols (Chapter 6) and ethical guidelines (Chapter 7) to protect users from potential harm and ensure privacy, especially in sensitive settings such as healthcare, and dangerous settings such as manufacturing and autonomous driving. This involves incorporating fail-safe mechanisms, secure data handling practices, and ethical decision-making algorithms.
3. **Transparency:** Design explainable robot learning systems with a focus on transparency (Section 7.4), allowing users to understand the rationale behind robot actions and decisions. This is crucial for users to trust robots (Section 3.9) and increase robot acceptance in all domains.

By adhering to these guidelines, the field of interactive and explainable robot learning can move forward in a manner that maximizes its potential to be deployed on real-world robots. Future work should focus on advancing technical capabilities while considering these design guidelines, and increasingly evaluate these learning algorithms in real-world domains.

10

Takeaways and Conclusion

In the rapidly evolving field of robotics, the emergence of Interactive and Explainable Robot Learning (IERL) marks a significant leap forward, harmonizing the dynamic capabilities of robots with the critical need for transparency. This fusion not only elevates the potential for robots to become more autonomous and intelligent but also ensures they remain aligned with human values and expectations. By delving into the intricacies of IERL, this review explores the synergistic relationship between interactive learning methods, feedback systems, long-term adaptation, and ethical guidelines, all while highlighting explainability as a cornerstone for fostering deeper human-robot connections. Through this lens, we investigate how these interconnected facets contribute to the development of robots that can learn from, adapt to, and explain their actions to human users, ultimately leading to more trustworthy and effective robotic systems.

Interactive Robot Learning (IRL) (Chapter 3) forms the bedrock of IERL, offering a paradigm where humans and robots engage in a symbiotic relationship to enhance the robot's learning curve. This interaction is not just about task execution but also about the robot's ability to explain its actions and learning process. Explainability in this con-

text aids in demystifying the robot's decision-making process, thereby enhancing human trust and the efficacy of the learning interaction. The review expands on various IRL methodologies, such as imitation learning, reinforcement learning, and their hybrids, emphasizing how explainability can be integrated, for instance, through understandable feedback loops and decision-making processes that are transparent to the user. The evolution of IRL necessitates advancements in robots' ability to process and interpret complex human instructions, gestures, and feedback, aligning with the overarching goal of making robots more relatable and their actions more interpretable to humans.

Feedback Mechanisms (Chapter 4) within Interactive and Explainable Robot Learning (IERL) are essential, acting as vital channels for bidirectional communication between humans and robots. These mechanisms are not just pathways for information and corrections but also foundational elements that enhance the interactivity and learning efficiency of robots. By integrating explainability into feedback systems, robots become capable of not only receiving instructions and corrections but also of providing insights into their decision-making processes. This level of transparency allows users to understand the rationale behind a robot's actions or errors, enabling them to offer more precise and constructive feedback. This review explores the richness of feedback modalities — including visual signals, auditory cues, and tactile feedback — each offering unique avenues for embedding explainable AI (XAI) techniques that make robot behaviors and learning processes more interpretable to human users. Future research directions are discussed, focusing on developing adaptive feedback systems that are personalized to the user's needs and preferences. These systems aim to not only convey information effectively but also elucidate the reasoning behind the robot's actions, fostering a deeper understanding and engagement. In the context of interactive learning, this enhanced feedback loop is instrumental in creating an environment where humans and robots can effectively collaborate and learn from each other, paving the way for more nuanced and effective human-robot interactions.

Long-term Learning (Chapter 5) is essential for robots operating in human environments, requiring not only the capacity to adapt and interact over extended periods but also to do so in a way that

fosters understanding and trust. The integration of explainability into long-term learning strategies is paramount, enabling robots to not only evolve in response to new information and environments but also to communicate the basis of their adaptation in a clear and understandable manner. This aspect is crucial for maintaining a healthy, trust-based relationship between humans and robots. Additionally, the development of Theory of Mind (ToM) capabilities is emphasized, allowing robots to better infer and articulate the intentions and emotions behind human actions, which is vital for effective and nuanced interaction. Looking forward, the presented review highlights the need for enhancing robots' adaptive learning algorithms, ensuring that as these systems become more sophisticated and autonomous, they retain the ability to explain their evolving thought processes and actions. This dual focus on interactive learning and explainability ensures that long-term robot learning remains aligned with human needs and expectations, fostering a collaborative and dynamic human-robot partnership.

Safety and Ethical Considerations (Chapter 6 - 7) in the realm of Interactive and Explainable Robot Learning (IERL) are paramount, ensuring that robots operate within the boundaries of both physical safety and moral responsibility. The commitment to safety involves employing formal methods to validate and verify the actions and decisions of robots, ensuring they pose no harm to human users or their environments. Ethical considerations extend beyond mere safety, encompassing the principles of accountability, privacy, and the avoidance of bias, thereby upholding the dignity and rights of all involved parties. In the context of interactive robot learning, these considerations become even more critical as the dynamics of human-robot interaction introduce complex ethical scenarios. Explainability serves as a crucial bridge in this context, not just as a means to make robot actions understandable, but as a mechanism to ensure that those actions are justifiable and aligned with ethical norms. By enabling robots to explain their decision-making processes, stakeholders can assess not only the appropriateness of the actions in terms of safety and ethics but also the alignment with societal values and individual expectations. This dual focus on interactive learning and explainability underlines the importance of developing robotic systems that are not only intelligent and adaptive

but also principled and accountable.

Evaluation Metrics and Benchmarking (Chapter 8) are fundamental in assessing the progress and effectiveness of Interactive and Explainable Robot Learning (IERL) systems. Chapter 8 advocates for the creation of advanced metrics and benchmarks designed to not only evaluate the performance of robots in interactive and explainable contexts but also to deeply understand the subtleties of human-robot interactions. Such metrics are crucial for determining the robots' ability to articulate their learning processes and decision-making rationales comprehensively, which, in turn, influences their acceptance and practicality in real-life settings. Enhanced evaluation strategies are essential for providing a holistic view of a robot's capabilities, ensuring that they meet the diverse expectations and needs of users across various scenarios. This comprehensive approach to evaluation underpins the continuous improvement of robots, making them more adept, trustworthy, and valuable as partners in human environments.

In conclusion, the harmonious fusion of interactive learning, robust feedback mechanisms, sustained learning over time, rigorous safety protocols, comprehensive ethical guidelines, and detailed evaluation metrics stands at the forefront of innovation. Central to this multidimensional approach is explainability, a crucial element that not only strengthens the functionality and reliability of robots but also bolsters their acceptance within society. By prioritizing the development of robots that can learn from human interaction and effectively communicate their reasoning, this integrated strategy ensures that advancements in robotics technology are in concert with human principles and ethics. Such a commitment to explainability and interaction paves the way for a future marked by trustful, effective, and enriching human-robot collaborations, reinforcing the essential alignment between robotic advancements and societal values.

Acknowledgements

The authors also would like to greatly thank Prof. Sonia Chernova for her constructive comments and feedback and being a sounding board, guiding our efforts throughout the process of preparing this work.

References

- Abbeel, P., A. Coates, and A. Y. Ng. (2010). “Autonomous Helicopter Aerobatics through Apprenticeship Learning”. *The International Journal of Robotics Research*. 29(13): 1608–1639. DOI: [10.1177/0278364910371999](https://doi.org/10.1177/0278364910371999). eprint: <https://doi.org/10.1177/0278364910371999>. URL: <https://doi.org/10.1177/0278364910371999>.
- Abbeel, P. and A. Y. Ng. (2004). “Apprenticeship learning via inverse reinforcement learning”. en. In: *Twenty-first international conference on Machine learning - ICML '04*. Banff, Alberta, Canada: ACM Press. 1. DOI: [10.1145/1015330.1015430](https://doi.org/10.1145/1015330.1015430). URL: <http://portal.acm.org/citation.cfm?doid=1015330.1015430>.
- Abbink, D. A., T. Carlson, M. Mulder, J. C. F. de Winter, F. Aminravan, T. L. Gibo, and E. R. Boer. (2018). “A Topology of Shared Control Systems—Finding Common Ground in Diversity”. *IEEE Transactions on Human-Machine Systems*. 48(5): 509–525. ISSN: 2168-2305. DOI: [10.1109/THMS.2018.2791570](https://doi.org/10.1109/THMS.2018.2791570). URL: <https://ieeexplore.ieee.org/document/8353134>.
- Abeyruwan, S., L. Graesser, D. B. D’Ambrosio, A. Singh, A. Shankar, A. Bewley, D. Jain, K. Choromanski, and P. R. Sanketi. (2022). “i-Sim2Real: Reinforcement Learning of Robotic Policies in Tight Human-Robot Interaction Loops”. en. URL: <http://arxiv.org/abs/2207.06572>.

- Abu-Dakka, F. J., B. Nemec, A. Kramberger, A. G. Buch, N. Krüger, and A. Ude. (2014). “Solving peg-in-hole tasks by human demonstration and exception strategies”. *Industrial Robot: An International Journal*. 41(6): 575–584. ISSN: 0143-991X. DOI: [10.1108/IR-07-2014-0363](https://doi.org/10.1108/IR-07-2014-0363). URL: <https://doi.org/10.1108/IR-07-2014-0363> (accessed on 03/13/2024).
- Achiam, J., D. Held, A. Tamar, and P. Abbeel. (2017). “Constrained policy optimization”. In: *International conference on machine learning*. PMLR. 22–31.
- Act, P. (2000). “Personal Information Protection and Electronic Documents Act”. *Department of Justice, Canada. Full text available at* <http://laws.justice.gc.ca/en/P-8.6/text.html>.
- Adadi, A. and M. Berrada. (2018). “Peeking inside the black-box: a survey on explainable artificial intelligence (XAI)”. *IEEE access*. 6: 52138–52160.
- Adamik, M., A. P. Madsen, and M. Rehm. (2022). “Explainability in Collaborative Robotics: The Effect of Informing the User on Task Performance and Trust”. In: *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 1252–1257.
- Adams, S., T. Cody, and P. A. Beling. (2022). “A survey of inverse reinforcement learning”. *Artificial Intelligence Review*. 55(6): 4307–4346.
- Agarwal, A., A. Beygelzimer, M. Dudík, J. Langford, and H. Wallach. (2018). “A reductions approach to fair classification”. In: *International conference on machine learning*. PMLR. 60–69.
- Agarwal, S., S. Muku, S. Anand, and C. Arora. (2022). “Does data repair lead to fair models? curating contextually fair data to reduce model bias”. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 3298–3307.
- Aivaliotis, P., S. Aivaliotis, C. Gkournelos, K. Kokkalis, G. Michalos, and S. Makris. (2019). “Power and force limiting on industrial robots for human-robot collaboration”. *Robotics and Computer-Integrated Manufacturing*. 59(Oct.): 346–360. ISSN: 0736-5845. DOI: [10.1016/j.rcim.2019.05.001](https://doi.org/10.1016/j.rcim.2019.05.001). URL: <https://www.sciencedirect.com/science/article/pii/S0736584518302928>.

- Akalin, N., A. Kiselev, A. Kristoffersson, and A. Loutfi. (2023). “A Taxonomy of Factors Influencing Perceived Safety in Human–Robot Interaction”. *International Journal of Social Robotics*: 1–12.
- Akalin, N., A. Kristoffersson, and A. Loutfi. (2022). “Do you feel safe with your robot? Factors influencing perceived safety in human-robot interaction based on subjective and objective measures”. *International journal of human-computer studies*. 158: 102744.
- Akgun, B., M. Cakmak, J. W. Yoo, and A. L. Thomaz. (2012). “Trajectories and keyframes for kinesthetic teaching: A human-robot interaction perspective”. In: *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. 391–398.
- Alhaji, B., M. Prilla, and A. Rausch. (2021). “Trust Dynamics and Verbal Assurances in Human Robot Physical Collaboration”. *Frontiers in Artificial Intelligence*. 4. ISSN: 2624-8212. URL: <https://www.frontiersin.org/articles/10.3389/frai.2021.703504>.
- Ali, T. and P. Kostakos. (2023). “HuntGPT: Integrating Machine Learning-Based Anomaly Detection and Explainable AI with Large Language Models (LLMs)”. *arXiv preprint arXiv:2309.16021*.
- Allen, J., C. Guinn, and E. Horvitz. (1999). “Mixed-initiative interaction”. *IEEE Intelligent Systems and their Applications*. 14(5): 14–23. DOI: [10.1109/5254.796083](https://doi.org/10.1109/5254.796083).
- Alonso Martin, F., M. Malfaz, Á. Castro-González, J. C. Castillo, and M. Á. Salichs. (2020). “Four-features evaluation of text to speech systems for three social robots”. *Electronics*. 9(2): 267.
- Alur, R., T. Feder, and T. A. Henzinger. (1996). “The benefits of relaxing punctuality”. *Journal of the ACM*. 43(1): 116–146.
- Ames, A. D., S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada. (2019). “Control barrier functions: Theory and applications”. In: *2019 18th European control conference (ECC)*. IEEE. 3420–3431.
- Amoroso, D. and G. Tamburrini. (2019). “What makes human control over weapon systems “meaningful””. *Report submitted by the International Committee for Robot Arms Control (ICRAC) to the GGE on lethal autonomous weapons systems of the CCW, Geneva*: 20–21.

- An, G., J. Lee, X. Zuo, N. Kosaka, K.-M. Kim, and H. O. Song. (2024). “Direct Preference-based Policy Optimization without Reward Modeling”. *Advances in Neural Information Processing Systems*. 36.
- Anderson, M. and S. L. Anderson. (2011). *Machine ethics*. Cambridge University Press.
- Andreas, J. and D. Klein. (2015). “Alignment-Based Compositional Semantics for Instruction Following”. In: *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Ed. by L. Màrquez, C. Callison-Burch, and J. Su. Lisbon, Portugal: Association for Computational Linguistics. 1165–1174. DOI: [10.18653/v1/D15-1138](https://doi.org/10.18653/v1/D15-1138). URL: <https://aclanthology.org/D15-1138>.
- Andreas Theodorou, R. H. W. and J. J. Bryson. (2017). “Designing and implementing transparency for real time inspection of autonomous robots”. *Connection Science*. 29(3): 230–241. DOI: [10.1080/09540091.2017.1310182](https://doi.org/10.1080/09540091.2017.1310182). eprint: <https://doi.org/10.1080/09540091.2017.1310182>. URL: <https://doi.org/10.1080/09540091.2017.1310182>.
- Andronas, D., G. Apostolopoulos, N. Fourtakas, and S. Makris. (2021). “Multi-modal interfaces for natural Human-Robot Interaction”. *Procedia Manufacturing*. 54: 197–202.
- Anjomshoe, S., A. Najjar, D. Calvaresi, and K. Främling. (2019). “Explainable agents and robots: Results from a systematic literature review”. In: *18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019*. International Foundation for Autonomous Agents and Multiagent Systems. 1078–1088.
- Anne, T., J. Wilkinson, and Z. Li. (2021). “Meta-learning for fast adaptive locomotion with uncertainties in environments and robot dynamics”. In: *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 4568–4575.
- Arakawa, R., S. Kobayashi, Y. Unno, Y. Tsuboi, and S.-i. Maeda. (2018). “DQN-TAMER: Human-in-the-Loop Reinforcement Learning with Intractable Feedback”. *ArXiv*. abs/1810.11748. URL: <https://api.semanticscholar.org/CorpusID:53105957>.
- Argall, B. D., S. Chernova, M. Veloso, and B. Browning. (2009a). “A survey of robot learning from demonstration”. *Robotics and autonomous systems*. 57(5): 469–483.

- Argall, B. D., S. Chernova, M. Veloso, and B. Browning. (2009b). “A survey of robot learning from demonstration”. *Robotics and Autonomous Systems*. 57(5): 469–483. ISSN: 0921-8890. DOI: <https://doi.org/10.1016/j.robot.2008.10.024>. URL: <https://www.sciencedirect.com/science/article/pii/S0921889008001772>.
- Arnold, T., D. Kasenberg, and M. Scheutz. (2021a). “Explaining in Time: Meeting Interactive Standards of Explanation for Robotic Systems”. *J. Hum.-Robot Interact.* 10(3). DOI: [10.1145/3457183](https://doi.org/10.1145/3457183). URL: <https://doi.org/10.1145/3457183>.
- Arnold, T., D. Kasenberg, and M. Scheutz. (2021b). “Explaining in time: Meeting interactive standards of explanation for robotic systems”. *ACM Transactions on Human-Robot Interaction (THRI)*. 10(3): 1–23.
- Arntz, A., S. C. Eimler, C. Straßmann, and H. U. Hoppe. (2021). “On the Influence of Autonomy and Transparency on Blame and Credit in Flawed Human-Robot Collaboration”. In: *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction. HRI '21 Companion*. Boulder, CO, USA: Association for Computing Machinery. 377–381. ISBN: 9781450382908. DOI: [10.1145/3434074.3447196](https://doi.org/10.1145/3434074.3447196). URL: <https://doi.org/10.1145/3434074.3447196>.
- Arora, S. and P. Doshi. (2021). “A survey of inverse reinforcement learning: Challenges, methods and progress”. *Artificial Intelligence*. 297: 103500. ISSN: 0004-3702. DOI: <https://doi.org/10.1016/j.artint.2021.103500>. URL: <https://www.sciencedirect.com/science/article/pii/S0004370221000515>.
- Aroyo, A. M., D. Pasquali, A. Kothig, F. Rea, G. Sandini, and A. Sciutti. (2021). “Expectations Vs. Reality: Unreliability and Transparency in a Treasure Hunt Game With Icube”. *IEEE Robotics and Automation Letters*. 6(3): 5681–5688. DOI: [10.1109/LRA.2021.3083465](https://doi.org/10.1109/LRA.2021.3083465).
- Arrieta, A. B., N. Diaz-Rodriguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-López, D. Molina, R. Benjamins, et al. (2020). “Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI”. *Information fusion*. 58: 82–115.

- Arulkumaran, K., M. P. Deisenroth, M. Brundage, and A. A. Bharath. (2017). “Deep Reinforcement Learning: A Brief Survey”. *IEEE Signal Processing Magazine*. 34(6): 26–38. DOI: [10.1109/MSP.2017.2743240](https://doi.org/10.1109/MSP.2017.2743240).
- Arumugam, D., S. Karamcheti, N. Gopalan, L. Wong, and S. Tellex. (2017). “Accurately and Efficiently Interpreting Human-Robot Instructions of Varying Granularities”. In: *Proceedings of Robotics: Science and Systems*. Cambridge, Massachusetts. DOI: [10.15607/RSS.2017.XIII.056](https://doi.org/10.15607/RSS.2017.XIII.056).
- Arumugam, D., J. K. Lee, S. Saskin, and M. L. Littman. (2019). “Deep Reinforcement Learning from Policy-Dependent Human Feedback”. *ArXiv*. abs/1902.04257. URL: <https://api.semanticscholar.org/CorpusID:21704492>.
- Arzate Cruz, C. and T. Igarashi. (2020a). “A Survey on Interactive Reinforcement Learning: Design Principles and Open Challenges”. In: *DIS '20*. <conf-loc>, <city>Eindhoven</city>, <country>Netherlands</country>: Association for Computing Machinery. 1195–1209. ISBN: 9781450369749. DOI: [10.1145/3357236.3395525](https://doi.org/10.1145/3357236.3395525). URL: <https://doi.org/10.1145/3357236.3395525>.
- Arzate Cruz, C. and T. Igarashi. (2020b). “A survey on interactive reinforcement learning: Design principles and open challenges”. In: *Proceedings of the 2020 ACM designing interactive systems conference*. 1195–1209.
- Aşık, O., B. Görür, and H. L. Akın. (2019). “End-to-End Deep Imitation Learning: Robot Soccer Case Study”. en. In: *RoboCup 2018: Robot World Cup XXII*. Ed. by D. Holz, K. Genter, M. Saad, and O. von Stryk. *Lecture Notes in Computer Science*. Cham: Springer International Publishing. 137–149. ISBN: 978-3-030-27544-0. DOI: [10.1007/978-3-030-27544-0_11](https://doi.org/10.1007/978-3-030-27544-0_11).
- Atakishiyev, S., M. Salameh, H. Yao, and R. Goebel. (2023). “Explainable Artificial Intelligence for Autonomous Driving: A Comprehensive Overview and Field Guide for Future Research Directions”. en. URL: <http://arxiv.org/abs/2112.11561> (accessed on 03/07/2024).
- Attia, M., M. Hossny, S. Nahavandi, M. Dalvand, and H. Asadi. (2018). “Towards Trusted Autonomous Surgical Robots”. In: *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 4083–4088. DOI: [10.1109/SMC.2018.00692](https://doi.org/10.1109/SMC.2018.00692).

- Auddy, S., J. Hollenstein, M. Saveriano, A. Rodríguez-Sánchez, and J. Piater. (2023). “Continual learning from demonstration of robotics skills”. *Robotics and Autonomous Systems*. 165: 104427. ISSN: 0921-8890. DOI: <https://doi.org/10.1016/j.robot.2023.104427>. URL: <https://www.sciencedirect.com/science/article/pii/S0921889023000660>.
- Azagra, P., J. Civera, and A. C. Murillo. (2020). “Incremental Learning of Object Models From Natural Human–Robot Interactions”. *IEEE Transactions on Automation Science and Engineering*. 17(4): 1883–1900. DOI: [10.1109/TASE.2020.2980246](https://doi.org/10.1109/TASE.2020.2980246).
- Babes-Vroman, M., V. Marivate, K. Subramanian, and M. L. Littman. (2011). “Apprenticeship Learning About Multiple Intentions”. In: *International Conference on Machine Learning*. URL: <https://api.semanticscholar.org/CorpusID:9737656>.
- Badr, A. A. and A. K. Abdul-Hassan. (2020). “A review on voice-based interface for human-robot interaction”. *Iraqi Journal for Electrical And Electronic Engineering*. 16(2): 91–102.
- Bahl, S., A. Gupta, and D. Pathak. (2022). “Human-to-Robot Imitation in the Wild”. en. In: *Robotics: Science and Systems XVIII*. Robotics: Science and Systems Foundation. ISBN: 978-0-9923747-8-5. DOI: [10.15607/RSS.2022.XVIII.026](https://doi.org/10.15607/RSS.2022.XVIII.026). URL: <http://www.roboticsproceedings.org/rss18/p026.pdf> (accessed on 11/04/2023).
- Bai, Y., A. Jones, K. Ndousse, A. Askell, A. Chen, N. DasSarma, D. Drain, S. Fort, D. Ganguli, T. Henighan, *et al.* (2022). “Training a helpful and harmless assistant with reinforcement learning from human feedback”. *arXiv preprint arXiv:2204.05862*.
- Bajcsy, A., D. P. Losey, M. K. O’Malley, and A. D. Dragan. (2018). “Learning from physical human corrections, one feature at a time”. In: *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. 141–149.
- Baker, M., R. Casey, B. Keyes, and H. Yanco. (2004). “Improved interfaces for human-robot interaction in urban search and rescue”. In: *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No.04CH37583)*. Vol. 3. 2960–2965 vol.3. DOI: [10.1109/ICSMC.2004.1400783](https://doi.org/10.1109/ICSMC.2004.1400783).
- Baldassarre, G. and M. Mirolli, eds. (2013). *Intrinsically Motivated Learning in Natural and Artificial Systems*. Berlin: Springer.

- Banerjee, A. and R. Padhi. (2018). “Nonlinear Guidance and Autopilot Design for Lunar Soft Landing”. In: *2018 AIAA Guidance, Navigation, and Control Conference*. 1872.
- Bansal, G., B. Nushi, E. Kamar, W. S. Lasecki, D. S. Weld, and E. Horvitz. (2019). “Beyond accuracy: The role of mental models in human-AI team performance”. In: *Proceedings of the AAAI conference on human computation and crowdsourcing*. Vol. 7. No. 1. 2–11.
- Bansal, S., A. Bajcsy, E. Ratner, A. D. Dragan, and C. J. Tomlin. (2020). “A hamilton-jacobi reachability-based framework for predicting and analyzing human motion for safe planning”. In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 7149–7155.
- Baranes, A. and P.-Y. Oudeyer. (2013). “Active learning of inverse models with intrinsically motivated goal exploration in robots”. *Robotics and Autonomous Systems*. 61(1): 49–73. ISSN: 0921-8890. DOI: <https://doi.org/10.1016/j.robot.2012.05.008>. URL: <https://www.sciencedirect.com/science/article/pii/S0921889012000644>.
- Barattini, P., F. Vicentini, G. S. Virk, and T. Haidegger. (2019). *Human-robot interaction: safety, standardization, and benchmarking*. CRC Press.
- Barbosa, F. S., J. Karlsson, P. Tajvar, and J. Tumova. (2021). “Formal Methods for Robot Motion Planning with Time and Space Constraints”. In: *Formal Modeling and Analysis of Timed Systems: 19th International Conference, FORMATS 2021, Paris, France, August 24–26, 2021, Proceedings 19*. Springer. 1–14.
- Bartneck, C., D. Kulić, E. Croft, and S. Zoghbi. (2009). “Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots”. *Int. Journal of Social Robotics*. 1(1): 71–81.
- Baum, D., K. Baum, T. P. Gros, and V. Wolf. (2023). “XAI Requirements in Smart Production Processes: A Case Study”. en. In: *Explainable Artificial Intelligence*. Ed. by L. Longo. *Communications in Computer and Information Science*. Cham: Springer Nature Switzerland. 3–24. ISBN: 978-3-031-44064-9. DOI: [10.1007/978-3-031-44064-9_1](https://doi.org/10.1007/978-3-031-44064-9_1).

- Baxter, P., T. Belpaeme, L. Canamero, P. Cosi, Y. Demiris, V. Enescu, A. Hiolle, I. Kruijff-Korabayova, R. Looije, M. Nalin, *et al.* (2011). “Long-term human-robot interaction with young users”. In: *IEEE/ACM human-robot interaction 2011 conference (robots with children workshop)*. Vol. 80. IEEE/ACM.
- Beck, J., R. Vuorio, E. Z. Liu, Z. Xiong, L. Zintgraf, C. Finn, and S. Whiteson. (2023). “A survey of meta-reinforcement learning”. *arXiv preprint arXiv:2301.08028*.
- Bekele, E. and N. Sarkar. (2014). “Psychophysiological feedback for adaptive human–robot interaction (HRI)”. In: *Advances in physiological computing*. Springer. 141–167.
- Belanche, D., L. V. Casalo, and C. Flavián. (2012). “Integrating trust and personal values into the Technology Acceptance Model: The case of e-government services adoption”. *Cuadernos de Economía y Dirección de la Empresa*. 15(4): 192–204.
- Belle, V. and I. Papantonis. (2021). “Principles and practice of explainable machine learning”. *Frontiers in big Data*: 39.
- Bellman, R. (1957). *Dynamic Programming*. 1st ed. Princeton, NJ, USA: Princeton University Press.
- Berio, D., S. Calinon, and F. F. Leymarie. (2016). “Learning dynamic graffiti strokes with a compliant robot”. en. In: *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Daejeon, South Korea: IEEE. 3981–3986. ISBN: 978-1-5090-3762-9. DOI: [10.1109/IROS.2016.7759586](https://doi.org/10.1109/IROS.2016.7759586). URL: <http://ieeexplore.ieee.org/document/7759586/>.
- Bertozzi, M., A. Broggi, and A. Fascioli. (2000). “Vision-based intelligent vehicles: State of the art and perspectives”. *Robotics and Autonomous systems*. 32(1): 1–16.
- Bethel, C. L., K. Salomon, R. R. Murphy, and J. L. Burke. (2007). “Survey of psychophysiology measurements applied to human-robot interaction”. In: *RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE. 732–737.

- Bianchi, R. A. C., M. F. Martins, C. H. C. Ribeiro, and A. H. R. Costa. (2014). “Heuristically-Accelerated Multiagent Reinforcement Learning”. *IEEE Transactions on Cybernetics*. 44(2): 252–265. DOI: [10.1109/TCYB.2013.2253094](https://doi.org/10.1109/TCYB.2013.2253094).
- Bidgoli, A., M. L. De Guevara, C. Hsiung, J. Oh, and E. Kang. (2020). “Artistic Style in Robotic Painting; a Machine Learning Approach to Learning Brushstroke from Human Artists”. en. URL: <https://arxiv.org/abs/2007.03647v2>.
- Billard, A., S. Calinon, R. Dillmann, and S. Schaal. (2008). “Robot Programming by Demonstration”. en. In: *Springer Handbook of Robotics*. Ed. by B. Siciliano and O. Khatib. Berlin, Heidelberg: Springer. 1371–1394. ISBN: 978-3-540-30301-5. DOI: [10.1007/978-3-540-30301-5_60](https://doi.org/10.1007/978-3-540-30301-5_60). URL: https://doi.org/10.1007/978-3-540-30301-5_60.
- Billard, A. and D. Grollman. (2013). “Robot learning by demonstration”. *Scholarpedia*. 8(12): 3824.
- Bingjing, G., H. Jianhai, L. Xiangpan, and Y. Lin. (2019). “Human–robot interactive control based on reinforcement learning for gait rehabilitation training robot”. en. *International Journal of Advanced Robotic Systems*. 16(2): 1729881419839584. ISSN: 1729-8806. DOI: [10.1177/1729881419839584](https://doi.org/10.1177/1729881419839584). URL: <https://doi.org/10.1177/1729881419839584> (accessed on 11/03/2023).
- Birhane, A. and J. van Dijk. (2020). “Robot Rights? Let’s Talk about Human Welfare Instead”. In: *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society. AIES ’20*. New York, NY, USA: Association for Computing Machinery. 207–213. ISBN: 9781450371100. DOI: [10.1145/3375627.3375855](https://doi.org/10.1145/3375627.3375855). URL: <https://doi.org/10.1145/3375627.3375855>.
- Biswas, M. and J. Murray. (2015). “Robotic companionship: how forgetfulness affects long-term human-robot interaction”. In: *Intelligent Robotics and Applications: 8th International Conference, ICIRA 2015, Portsmouth, UK, August 24–27, 2015, Proceedings, Part II 8*. Springer. 37–48.
- Biyyik, E., M. Palan, N. C. Landolfi, D. P. Losey, and D. Sadigh. (2019). “Asking easy questions: A user-friendly approach to active reward learning”. *arXiv preprint arXiv:1910.04365*.

- Blukis, V., D. Misra, R. A. Knepper, and Y. Artzi. (2018). “Mapping Navigation Instructions to Continuous Control Actions with Position-Visitation Prediction”. In: *Proceedings of The 2nd Conference on Robot Learning*. Ed. by A. Billard, A. Dragan, J. Peters, and J. Morimoto. Vol. 87. *Proceedings of Machine Learning Research*. PMLR. 505–518. URL: <https://proceedings.mlr.press/v87/blukis18a.html>.
- Bobu, A., M. Wiggert, C. Tomlin, and A. D. Dragan. (2021). “Feature expansive reward learning: Rethinking human input”. In: *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 216–224.
- Boessenkool, H., D. A. Abbink, C. J. Heemskerk, F. C. van der Helm, and J. G. Wildenbeest. (2012). “A task-specific analysis of the benefit of haptic shared control during telemanipulation”. *IEEE Transactions on Haptics*. 6(1): 2–12.
- Bogges, K., S. Kraus, and L. Feng. (2022). “Toward policy explanations for multi-agent reinforcement learning”. *arXiv preprint arXiv:2204.12568*.
- Bommasani, R. *et al.* (2022). “On the Opportunities and Risks of Foundation Models”. arXiv: [2108.07258](https://arxiv.org/abs/2108.07258) [cs.LG].
- Bonarini, A. (2020). “Communication in human-robot interaction”. *Current Robotics Reports*. 1: 279–285.
- Boularias, A., J. Kober, and J. Peters. (2011). “Relative Entropy Inverse Reinforcement Learning”. In: *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*. Ed. by G. Gordon, D. Dunson, and M. Dudík. Vol. 15. *Proceedings of Machine Learning Research*. Fort Lauderdale, FL, USA: PMLR. 182–189. URL: <https://proceedings.mlr.press/v15/boularias11a.html>.
- Boyle, M., C. Edwards, and S. Greenberg. (2000). “The effects of filtered video on awareness and privacy”. In: *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. 1–10.
- Bozhinoski, D., M. Garzon Oviedo, N. Hammoudeh Garcia, H. Deshpande, G. van der Hoorn, J. Tjerngren, A. Wąsowski, and C. Hernandez Corbato. (2022). “MROS: runtime adaptation for robot control architectures”. *Advanced Robotics*. 36(11): 502–518.

- Bradley Knox, W. and P. Stone. (2008). “TAMER: Training an Agent Manually via Evaluative Reinforcement”. In: *2008 7th IEEE International Conference on Development and Learning*. 292–297. DOI: [10.1109/DEVLRN.2008.4640845](https://doi.org/10.1109/DEVLRN.2008.4640845).
- Branavan, S., H. Chen, L. Zettlemoyer, and R. Barzilay. (2009). “Reinforcement Learning for Mapping Instructions to Actions”. In: *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*. Ed. by K.-Y. Su, J. Su, J. Wiebe, and H. Li. Suntec, Singapore: Association for Computational Linguistics. 82–90. URL: <https://aclanthology.org/P09-1010>.
- Bray, L. C. J., S. R. Anumandla, C. M. Thibeault, R. V. Hoang, P. H. Goodman, S. M. Dascalu, B. D. Bryant, and F. C. Harris Jr. (2012). “Real-time human–robot interaction underlying neurorobotic trust and intent recognition”. *Neural networks*. 32: 130–137.
- Breazeal, C., K. Dautenhahn, and T. Kanda. (2016). “Social robotics”. *Springer handbook of robotics*: 1935–1972.
- Breazeal, C., J. Gray, G. Hoffman, and M. Berlin. (2004). “Social robots: Beyond tools to partners”. In: *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759)*. IEEE. 551–556.
- Breazeal, C., C. D. Kidd, A. L. Thomaz, G. Hoffman, and M. Berlin. (2005). “Effects of nonverbal communication on efficiency and robustness in human-robot teamwork”. In: *2005 IEEE/RSJ international conference on intelligent robots and systems*. IEEE. 708–713.
- Brewitt, C., B. Gyevnar, S. Garcin, and S. V. Albrecht. (2021). “GRIT: Fast, Interpretable, and Verifiable Goal Recognition with Learned Decision Trees for Autonomous Driving”. en. In: *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Prague, Czech Republic: IEEE. 1023–1030. ISBN: 978-1-66541-714-3. DOI: [10.1109/IROS51168.2021.9636279](https://doi.org/10.1109/IROS51168.2021.9636279). URL: <https://ieeexplore.ieee.org/document/9636279/>.
- Broadbent, E., R. Stafford, and B. MacDonald. (2009). “Acceptance of healthcare robots for the older population: Review and future directions”. *International journal of social robotics*. 1: 319–330.

- Broadbent, E., R. Tamagawa, A. Patience, B. Knock, N. Kerse, K. Day, and B. A. MacDonald. (2011). “Attitudes towards health-care robots in a retirement village”. en. *Australas J Ageing*. 31(2): 115–120.
- Brockman, G., V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba. (2016). “Openai gym”. *arXiv preprint arXiv:1606.01540*.
- Broekens, J. and M. Chetouani. (2019). “Towards transparent robot learning through TDRL-based emotional expressions”. *IEEE Transactions on Affective Computing*. 12(2): 352–362.
- Broekens, J. and M. Chetouani. (2021). “Towards Transparent Robot Learning Through TDRL-Based Emotional Expressions”. *IEEE Transactions on Affective Computing*. 12(2): 352–362. DOI: [10.1109/TAFFC.2019.2893348](https://doi.org/10.1109/TAFFC.2019.2893348).
- Brohan, A., N. Brown, J. Carbajal, Y. Chebotar, X. Chen, K. Chormanski, T. Ding, D. Driess, A. Dubey, C. Finn, *et al.* (2023a). “Rt-2: Vision-language-action models transfer web knowledge to robotic control”. *arXiv preprint arXiv:2307.15818*.
- Brohan, A., N. Brown, J. Carbajal, Y. Chebotar, X. Chen, K. Chormanski, T. Ding, D. Driess, A. Dubey, C. Finn, P. Florence, C. Fu, M. G. Arenas, K. Gopalakrishnan, K. Han, K. Hausman, A. Herzog, J. Hsu, B. Ichter, A. Irpan, N. J. Joshi, R. Julian, D. Kalashnikov, Y. Kuang, I. Leal, L. Lee, T. E. Lee, S. Levine, Y. Lu, H. Michalewski, I. Mordatch, K. Pertsch, K. Rao, K. Reymann, M. S. Ryoo, G. Salazar, P. Sanketi, P. Sermanet, J. Singh, A. Singh, R. Soricut, H. T. Tran, V. Vanhoucke, Q. Vuong, A. Wahid, S. Welker, P. Wohlhart, J. Wu, F. Xia, T. Xiao, P. Xu, S. Xu, T. Yu, and B. Zitkovich. (2023b). “RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control”. *CoRR*. abs/2307.15818. DOI: [10.48550/ARXIV.2307.15818](https://doi.org/10.48550/ARXIV.2307.15818). arXiv: [2307.15818](https://arxiv.org/abs/2307.15818). URL: <https://doi.org/10.48550/arXiv.2307.15818>.
- Brohan, A., N. Brown, J. Carbajal, Y. Chebotar, J. Dabis, C. Finn, K. Gopalakrishnan, K. Hausman, A. Herzog, J. Hsu, *et al.* (2022). “Rt-1: Robotics transformer for real-world control at scale”. *arXiv preprint arXiv:2212.06817*.

- Brose, S. W., D. J. Weber, B. A. Salatin, G. G. Grindle, H. Wang, J. J. Vazquez, and R. A. Cooper. (2010). “The role of assistive robotics in the lives of persons with disability”. *American Journal of Physical Medicine & Rehabilitation*. 89(6): 509–521.
- Brown, D., W. Goo, P. Nagarajan, and S. Niekum. (2019). “Extrapolating beyond suboptimal demonstrations via inverse reinforcement learning from observations”. In: *International conference on machine learning*. PMLR. 783–792.
- Brown, D. and S. Niekum. (2018). “Efficient probabilistic performance bounds for inverse reinforcement learning”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 32. No. 1.
- Brown, D. S. and S. Niekum. (2017). “Toward probabilistic safety bounds for robot learning from demonstration”. In: *2017 AAAI Fall Symposium Series*.
- Brown, D. S., J. Schneider, A. Dragan, and S. Niekum. (2021). “Value alignment verification”. In: *International Conference on Machine Learning*. PMLR. 1105–1115.
- Brunke, L., M. Greeff, A. W. Hall, Z. Yuan, S. Zhou, J. Panerati, and A. P. Schoellig. (2022). “Safe learning in robotics: From learning-based control to safe reinforcement learning”. *Annual Review of Control, Robotics, and Autonomous Systems*. 5: 411–444.
- Brys, T., A. Harutyunyan, H. B. Suay, S. Chernova, M. E. Taylor, and A. Nowé. (2015). “Reinforcement Learning from Demonstration through Shaping”. In: *Proceedings of the 24th International Conference on Artificial Intelligence. IJCAI’15*. Buenos Aires, Argentina: AAAI Press. 3352–3358. ISBN: 9781577357384.
- Buchina, N., S. Kamel, and E. Barakova. (2016). “Design and evaluation of an end-user friendly tool for robot programming”. In: *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 185–191.
- Bui, H., V. S. Chau, M. Degl’Innocenti, L. Leone, and F. Vicentini. (2019). “The resilient organisation: A meta-analysis of the effect of communication on team diversity and team performance”. *Applied Psychology*. 68(4): 621–657.

- Bullard, K., Y. Schroecker, and S. Chernova. (2019). “Active Learning within Constrained Environments through Imitation of an Expert Questioner”. In: *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*. International Joint Conferences on Artificial Intelligence Organization. 2045–2052. DOI: [10.24963/ijcai.2019/283](https://doi.org/10.24963/ijcai.2019/283). URL: <https://doi.org/10.24963/ijcai.2019/283>.
- Butchibabu, A. (2016). “Anticipatory communication strategies for human robot team coordination”. *PhD thesis*. Massachusetts Institute of Technology.
- Butchibabu, A., C. Sparano-Huiban, L. Sonenberg, and J. Shah. (2016). “Implicit coordination strategies for effective team communication”. *Human factors*. 58(4): 595–610.
- Bütepage, J., A. Ghadirzadeh, Ö. Öztimur Karadağ, M. Björkman, and D. Kragic. (2020). “Imitating by generating: Deep generative models for imitation of interactive tasks”. *Frontiers in Robotics and AI*. 7.
- Cabi, S., S. G. Colmenarejo, A. Novikov, K. Konyushkova, S. E. Reed, R. Jeong, K. Zolna, Y. Aytar, D. Budden, M. Vecerík, O. O. Sushkov, D. Barker, J. Scholz, M. Denil, N. de Freitas, and Z. Wang. (2019). “Scaling data-driven robotics with reward sketching and batch reinforcement learning”. *Robotics: Science and Systems XVI*. URL: <https://api.semanticscholar.org/CorpusID:211032272>.
- Cachat-Rosset, G. and A. Klarsfeld. (2023). “Diversity, Equity, and Inclusion in Artificial Intelligence: An Evaluation of Guidelines”. *Applied Artificial Intelligence*. 37(1): 2176618.
- Cakmak, M. and A. L. Thomaz. (2012a). “Designing robot learners that ask good questions”. In: *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. 17–24.
- Cakmak, M. and A. L. Thomaz. (2012b). “Designing Robot Learners That Ask Good Questions”. In: *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction. HRI '12*. Association for Computing Machinery. 17–24. DOI: [10.1145/2157689.2157693](https://doi.org/10.1145/2157689.2157693).

- Call, J. and M. Carpenter. (2002). “Three sources of information in social learning.” In: *Imitation in animals and artifacts. Complex adaptive systems*. Cambridge, MA, US: Boston Review. 211–228. ISBN: 0-262-04203-7 (Hardcover). DOI: [10.7551/mitpress/3676.001.0001](https://doi.org/10.7551/mitpress/3676.001.0001).
- Calmon, F., D. Wei, B. Vinzamuri, K. Natesan Ramamurthy, and K. R. Varshney. (2017). “Optimized pre-processing for discrimination prevention”. *Advances in neural information processing systems*. 30.
- Camacho, A. and S. A. McIlraith. (2019). “Learning interpretable models expressed in linear temporal logic”. In: *Proceedings of the International Conference on Automated Planning and Scheduling*. Vol. 29. 621–630.
- Camerer, C. F., T.-H. Ho, and J.-K. Chong. (2004). “A cognitive hierarchy model of games”. *The Quarterly Journal of Economics*. 119(3): 861–898.
- Cameron, D., J. Aitken, E. Collins, L. Boorman, A. Chua, S. Fernando, O. McAree, U. Martinez-Hernandez, and J. Law. (2015). “Framing Factors: The Importance of Context and the Individual in Understanding Trust in Human-Robot Interaction”. In:
- Candon, K., H. Zhou, S. Gillet, and M. Vázquez. (2023). “Verbally Soliciting Human Feedback in Continuous Human-Robot Collaboration: Effects of the Framing and Timing of Reminders”. In: *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 290–300.
- Cangelosi, A. (2010). “Grounding language in action and perception: From cognitive agents to humanoid robots”. *Physics of life reviews*. 7(2): 139–151.
- Cangelosi, A. and M. Schlesinger. (2015). *Developmental Robotics: From Babies to Robots*. The MIT Press. ISBN: 9780262325295. DOI: [10.7551/mitpress/9320.001.0001](https://doi.org/10.7551/mitpress/9320.001.0001). URL: <https://doi.org/10.7551/mitpress/9320.001.0001>.
- Cantucci, F. and R. Falcone. (2020). “Towards trustworthiness and transparency in social human-robot interaction”. In: *2020 IEEE International Conference on Human-Machine Systems (ICHMS)*. 1–6. DOI: [10.1109/ICHMS49158.2020.9209397](https://doi.org/10.1109/ICHMS49158.2020.9209397).

- Cantucci, F. and R. Falcone. (2022). “Collaborative Autonomy: Human–Robot Interaction to the Test of Intelligent Help”. *Electronics*. 11(19): 3065.
- Carlson, S. M., M. A. Koenig, and M. B. Harms. (2013). “Theory of mind”. *Wiley Interdisciplinary Reviews: Cognitive Science*. 4(4): 391–402.
- Carreno-Medrano, P., S. L. Smith, and D. Kulić. (2022). “Joint estimation of expertise and reward preferences from human demonstrations”. *IEEE Transactions on Robotics*. 39(1): 681–698.
- Cartillier, V., Z. Ren, N. Jain, S. Lee, I. Essa, and D. Batra. (2021). “Semantic mapnet: Building allocentric semantic maps and representations from egocentric views”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. No. 2. 964–972.
- Carvalho, J., D. Koert, M. Daniv, and J. Peters. (2022). “Residual Robot Learning for Object-Centric Probabilistic Movement Primitives”. arXiv: [2203.03918 \[cs.RO\]](https://arxiv.org/abs/2203.03918).
- Casalino, A., C. Messeri, M. Pozzi, A. M. Zanchettin, P. Rocco, and D. Prattichizzo. (2018a). “Operator awareness in human–robot collaboration through wearable vibrotactile feedback”. *IEEE Robotics and Automation Letters*. 3(4): 4289–4296.
- Casalino, A., C. Messeri, M. Pozzi, A. M. Zanchettin, P. Rocco, and D. Prattichizzo. (2018b). “Operator Awareness in Human–Robot Collaboration Through Wearable Vibrotactile Feedback”. *IEEE Robotics and Automation Letters*. 3(4): 4289–4296. DOI: [10.1109/LRA.2018.2865034](https://doi.org/10.1109/LRA.2018.2865034).
- Casalino, A., A. M. Zanchettin, L. Piroddi, and P. Rocco. (2019). “Optimal scheduling of human–robot collaborative assembly operations with time petri nets”. *IEEE Transactions on Automation Science and Engineering*. 18(1): 70–84.
- Casey, D., E. Barrett, T. Kovacic, D. Sancarlo, F. Ricciardi, K. Murphy, A. Koumpis, A. Santorelli, N. Gallagher, and S. Whelan. (2020). “The Perceptions of People with Dementia and Key Stakeholders Regarding the Use and Impact of the Social Robot MARIO”. *International Journal of Environmental Research and Public Health*. 17(22). ISSN: 1660-4601. DOI: [10.3390/ijerph17228621](https://doi.org/10.3390/ijerph17228621). URL: <https://www.mdpi.com/1660-4601/17/22/8621>.

- Castelvecchi, D. (2016). “Can we open the black box of AI?” *Nature News*. 538(7623): 20.
- Castro, A., F. Silva, and V. Santos. (2021). “Trends of human-robot collaboration in industry contexts: Handover, learning, and metrics”. *Sensors*. 21(12): 4113.
- Cavazos, J. G., P. J. Phillips, C. D. Castillo, and A. J. O’Toole. (2020). “Accuracy comparison across face recognition algorithms: Where are we on measuring race bias?” *IEEE transactions on biometrics, behavior, and identity science*. 3(1): 101–111.
- Celemin, C. and J. Kober. (2023). “Knowledge-and ambiguity-aware robot learning from corrective and evaluative feedback”. *Neural Computing and Applications*: 1–19.
- Celemin, C., R. Pérez-Dattari, E. Chisari, G. Franzese, L. de Souza Rosa, R. Prakash, Z. Ajanović, M. Ferraz, A. Valada, J. Kober, et al. (2022). “Interactive imitation learning in robotics: A survey”. *Foundations and Trends® in Robotics*. 10(1-2): 1–197.
- Celemin, C. and J. Ruiz-del-Solar. (2015). “COACH: Learning continuous actions from COrrective Advice Communicated by Humans”. In: *2015 International Conference on Advanced Robotics (ICAR)*. 581–586. DOI: [10.1109/ICAR.2015.7251514](https://doi.org/10.1109/ICAR.2015.7251514).
- Celemin, C., J. Ruiz-del-Solar, and J. Kober. (2019). “A fast hybrid reinforcement learning framework with human corrective feedback”. *Autonomous Robots*. 43: 1173–1186.
- Cèsar-Tondreau, B., G. Warnell, E. Stump, K. Kochersberger, and N. R. Waytowich. (2021). “Improving Autonomous Robotic Navigation Using Imitation Learning”. en. *Front Robot AI*. 8(June): 627730.
- Cha, E., N. T. Fitter, Y. Kim, T. Fong, and M. J. Matarić. (2018). “Effects of robot sound on auditory localization in human-robot collaboration”. In: *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction*. 434–442.
- Chaddad, A., J. Peng, J. Xu, and A. Bouridane. (2023). “Survey of Explainable AI Techniques in Healthcare”. *Sensors (Basel, Switzerland)*. 23(2): 634. ISSN: 1424-8220. DOI: [10.3390/s23020634](https://doi.org/10.3390/s23020634). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9862413/>.

- Chao, C., M. Cakmak, and A. L. Thomaz. (2010). “Transparent active learning for robots”. In: *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 317–324.
- Che, Y., A. M. Okamura, and D. Sadigh. (2020). “Efficient and trustworthy social navigation via explicit and implicit robot–human communication”. *IEEE Transactions on Robotics*. 36(3): 692–707.
- Chebly, A., R. Talj, and A. Charara. (2019). “Coupled longitudinal/lateral controllers for autonomous vehicles navigation, with experimental validation”. *Control Engineering Practice*. 88(July): 79–96. ISSN: 0967-0661. DOI: [10.1016/j.conengprac.2019.05.001](https://doi.org/10.1016/j.conengprac.2019.05.001). URL: <https://www.sciencedirect.com/science/article/pii/S0967066118301680>.
- Chen, G., S. Baek, J.-D. Florez, W. Qian, S.-W. Leigh, S. Hutchinson, and F. Dellaert. (2022a). “GTGraffiti: Spray Painting Graffiti Art from Human Painting Motions with a Cable Driven Parallel Robot”. In: *2022 International Conference on Robotics and Automation (ICRA)*. 4065–4072. DOI: [10.1109/ICRA46639.2022.9812008](https://doi.org/10.1109/ICRA46639.2022.9812008). URL: <https://ieeexplore.ieee.org/document/9812008>.
- Chen, G., Z. Mao, H. Yi, X. Li, B. Bai, M. Liu, and H. Zhou. (2020a). “Pedestrian Detection Based on Panoramic Depth Map Transformed from 3D-LiDAR Data”. en. *Periodica Polytechnica Electrical Engineering and Computer Science*. 64(3): 274–285. ISSN: 2064-5279. DOI: [10.3311/PPee.14960](https://doi.org/10.3311/PPee.14960). URL: <https://pp.bme.hu/eecs/article/view/14960>.
- Chen, H., M. C. Leu, and Z. Yin. (2022b). “Real-time multi-modal human–robot collaboration using gestures and speech”. *Journal of Manufacturing Science and Engineering*. 144(10): 101007.
- Chen, K., N. S. Srikanth, D. Kent, H. Ravichandar, and S. Chernova. (2021a). “Learning Hierarchical Task Networks with Preferences from Unannotated Demonstrations”. In: *Proceedings of the 2020 Conference on Robot Learning*. Ed. by J. Kober, F. Ramos, and C. Tomlin. Vol. 155. *Proceedings of Machine Learning Research*. PMLR. 1572–1581. URL: <https://proceedings.mlr.press/v155/chen21d.html>.

- Chen, K., T. Hwu, H. J. Kashyap, J. L. Krichmar, K. Stewart, J. Xing, and X. Zou. (2020b). “Neurorobots as a Means Toward Neuroethology and Explainable AI”. *Frontiers in Neurorobotics*. 14. ISSN: 1662-5218. URL: <https://www.frontiersin.org/articles/10.3389/fnbot.2020.570308>.
- Chen, L., S. Jayanthi, R. R. Paleja, D. Martin, V. Zakharov, and M. Gombolay. (2023a). “Fast Lifelong Adaptive Inverse Reinforcement Learning from Demonstrations”. en. In: *Proceedings of The 6th Conference on Robot Learning*. PMLR. 2083–2094. URL: <https://proceedings.mlr.press/v205/chen23e.html>.
- Chen, L., R. Paleja, M. Ghuy, and M. Gombolay. (2020c). “Joint Goal and Strategy Inference across Heterogeneous Demonstrators via Reward Network Distillation”. In: *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction. HRI '20*. Cambridge, United Kingdom: Association for Computing Machinery. 659–668. ISBN: 9781450367462. DOI: [10.1145/3319502.3374791](https://doi.org/10.1145/3319502.3374791). URL: <https://doi.org/10.1145/3319502.3374791>.
- Chen, L., R. Paleja, and M. Gombolay. (2021b). “Learning from sub-optimal demonstration via self-supervised reward regression”. In: *Conference on robot learning*. PMLR. 1262–1277.
- Chen, M., S. Nikolaidis, H. Soh, D. Hsu, and S. Srinivasa. (2018a). “Planning with Trust for Human-Robot Collaboration”. In: *2018 13th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 307–315. URL: <https://ieeexplore.ieee.org/document/9473525>.
- Chen, M., S. Nikolaidis, H. Soh, D. Hsu, and S. Srinivasa. (2020d). “Trust-Aware Decision Making for Human-Robot Collaboration: Model Learning and Planning”. en. *ACM Transactions on Human-Robot Interaction*. 9(2): 1–23. ISSN: 2573-9522. DOI: [10.1145/3359616](https://doi.org/10.1145/3359616). URL: <https://dl.acm.org/doi/10.1145/3359616>.
- Chen, M., Q. Tam, S. C. Livingston, and M. Pavone. (2018b). “Signal temporal logic meets reachability: Connections and applications”. In: *International Workshop on the Algorithmic Foundations of Robotics*. Springer. 581–601.

- Chen, S., K. Boggess, and L. Feng. (2020e). “Towards Transparent Robotic Planning via Contrastive Explanations”. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 6593–6598. DOI: [10.1109/IROS45743.2020.9341773](https://doi.org/10.1109/IROS45743.2020.9341773).
- Chen, S.-C., W. Moyle, C. Jones, and H. Petsky. (2020f). “A social robot intervention on depression, loneliness, and quality of life for Taiwanese older adults in long-term care”. *International Psychogeriatrics*. 32(8): 981–991. DOI: [10.1017/S1041610220000459](https://doi.org/10.1017/S1041610220000459).
- Chen, X., J. Djolonga, P. Padlewski, B. Mustafa, S. Changpinyo, J. Wu, C. R. Ruiz, S. Goodman, X. Wang, Y. Tay, S. Shakeri, M. Dehghani, D. Salz, M. Lucic, M. Tschannen, A. Nagrani, H. Hu, M. Joshi, B. Pang, C. Montgomery, P. Pietrzyk, M. Ritter, A. Piergiovanni, M. Minderer, F. Pavetic, A. Waters, G. Li, I. Alabdulmohsin, L. Beyer, J. Amelot, K. Lee, A. P. Steiner, Y. Li, D. Keysers, A. Arnab, Y. Xu, K. Rong, A. Kolesnikov, M. Seyedhosseini, A. Angelova, X. Zhai, N. Houlsby, and R. Soricut. (2023b). “PaLI-X: On Scaling up a Multilingual Vision and Language Model”. arXiv: [2305.18565](https://arxiv.org/abs/2305.18565) [cs.CV].
- Chen, Y., S. Li, X. Tang, K. Yang, D. Cao, and X. Lin. (2023c). “Interaction-Aware Decision-Making for Autonomous Vehicles”. *IEEE Transactions on Transportation Electrification*. 9(Sept.): 4704–4715. ISSN: 2332-7782. DOI: [10.1109/TTE.2023.3240454](https://doi.org/10.1109/TTE.2023.3240454). URL: <https://ieeexplore.ieee.org/abstract/document/10029923>.
- Chernova, S. and A. L. Thomaz. (2014). *Robot learning from human teachers*. Morgan & Claypool Publishers.
- Chernova, S. and M. Veloso. (2009). “Interactive policy learning through confidence-based autonomy”. *J. Artif. Int. Res.* 34(1): 1–25. ISSN: 1076-9757.
- Chevalier-Boisvert, M., F. Golemo, Y. Cao, B. Mehta, and L. Paull. (2018). “Duckietown Environments for OpenAI Gym”. <https://github.com/duckietown/gym-duckietown>.
- Chi, C., S. Feng, Y. Du, Z. Xu, E. Cousineau, B. Burchfiel, and S. Song. (2023a). “Diffusion Policy: Visuomotor Policy Learning via Action Diffusion”. In: *Proceedings of Robotics: Science and Systems (RSS)*.

- Chi, O. H., C. G. Chi, D. Gursoy, and R. Nunkoo. (2023b). "Customers' acceptance of artificially intelligent service robots: The influence of trust and culture". *International Journal of Information Management*. 70(June): 102623. ISSN: 0268-4012. DOI: [10.1016/j.ijinfomgt.2023.102623](https://doi.org/10.1016/j.ijinfomgt.2023.102623). URL: <https://www.sciencedirect.com/science/article/pii/S026840122300004X>.
- Chiang, K.-W., G.-J. Tsai, Y.-H. Li, Y. Li, and N. El-Sheimy. (2020). "Navigation Engine Design for Automated Driving Using INS/GNSS/3D LiDAR-SLAM and Integrity Assessment". en. *Remote Sensing*. 12(10): 1564. ISSN: 2072-4292. DOI: [10.3390/rs12101564](https://doi.org/10.3390/rs12101564). URL: <https://www.mdpi.com/2072-4292/12/10/1564>.
- Chiarella, D., M. Bibuli, G. Bruzzone, M. Caccia, A. Ranieri, E. Zereik, L. Marconi, and P. Cutugno. (2015). "Gesture-based language for diver-robot underwater interaction". In: *Oceans 2015-genova*. IEEE. 1–9.
- Chiarella, D., M. Bibuli, G. Bruzzone, M. Caccia, A. Ranieri, E. Zereik, L. Marconi, and P. Cutugno. (2018). "A novel gesture-based language for underwater human–robot interaction". *Journal of Marine Science and Engineering*. 6(3): 91.
- Chiou, M., N. Hawes, and R. Stolkin. (2021a). "Mixed-initiative Variable Autonomy for Remotely Operated Mobile Robots". en. *ACM Transactions on Human-Robot Interaction*. 10(4): 1–34. ISSN: 2573-9522. DOI: [10.1145/3472206](https://doi.org/10.1145/3472206). URL: <https://dl.acm.org/doi/10.1145/3472206> (accessed on 12/19/2023).
- Chiou, M., F. McCabe, M. Grigoriou, and R. Stolkin. (2021b). "Trust, Shared Understanding and Locus of Control in Mixed-Initiative Robotic Systems". en. In: *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*. 684–691. DOI: [10.1109/RO-MAN50785.2021.9515476](https://doi.org/10.1109/RO-MAN50785.2021.9515476). URL: <http://arxiv.org/abs/2107.00690>.
- Chisari, E., T. Welschhold, J. Boedecker, W. Burgard, and A. Valada. (2022). "Correct me if i am wrong: Interactive learning for robotic manipulation". *IEEE Robotics and Automation Letters*. 7(2): 3695–3702.

- Chiu, Z.-Y., F. Richter, E. K. Funk, R. K. Orosco, and M. C. Yip. (2021). “Bimanual regrasping for suture needles using reinforcement learning for rapid motion planning”. In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 7737–7743.
- Choi, J. and K.-e. Kim. (2011). “MAP Inference for Bayesian Inverse Reinforcement Learning”. In: *Advances in Neural Information Processing Systems*. Ed. by J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, and K. Weinberger. Vol. 24. Curran Associates, Inc. URL: https://proceedings.neurips.cc/paper_files/paper/2011/file/3a15c7d0bbe60300a39f76f8a5ba6896-Paper.pdf.
- Chong, J.-K., C. F. Camerer, and T.-H. Ho. (2005). “Cognitive hierarchy: A limited thinking theory in games”. In: *Experimental Business Research: Marketing, Accounting and Cognitive Perspectives Volume III*. Springer. 203–228.
- Christiano, P. F., J. Leike, T. Brown, M. Martic, S. Legg, and D. Amodei. (2017a). “Deep Reinforcement Learning from Human Preferences”. In: *Advances in Neural Information Processing Systems*. Ed. by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett. Vol. 30. Curran Associates, Inc. URL: https://proceedings.neurips.cc/paper_files/paper/2017/file/d5e2c0adad503c91f91df240d0cd4e49-Paper.pdf.
- Christiano, P. F., J. Leike, T. Brown, M. Martic, S. Legg, and D. Amodei. (2017b). “Deep reinforcement learning from human preferences”. *Advances in neural information processing systems*. 30.
- Christiano, P. F., J. Leike, T. B. Brown, M. Martic, S. Legg, and D. Amodei. (2017c). “Deep Reinforcement Learning from Human Preferences”. *ArXiv*. abs/1706.03741. URL: <https://api.semanticscholar.org/CorpusID:4787508>.
- Chu, K., X. Zhao, C. Weber, M. Li, and S. Wermter. (2023). “Accelerating Reinforcement Learning of Robotic Manipulations via Feedback from Large Language Models”. *arXiv preprint arXiv:2311.02379*.
- Cifuentes, C. A., M. J. Pinto, N. Céspedes, and M. Múnera. (2020). “Social robots in therapy and care”. *Current Robotics Reports*. 1: 59–74.

- Cini, F., T. Banfi, G. Ciuti, L. Craighero, and M. Controzzi. (2021). “The relevance of signal timing in human-robot collaborative manipulation”. *Science Robotics*. 6(58): eabg1308.
- Clabaugh, C., K. Mahajan, S. Jain, R. Pakkar, D. Becerra, Z. Shi, E. Deng, R. Lee, G. Ragusa, and M. Matarić. (2019). “Long-Term Personalization of an In-Home Socially Assistive Robot for Children With Autism Spectrum Disorders”. *Frontiers in Robotics and AI*. 6(Nov.): 110. ISSN: 2296-9144. DOI: [10.3389 / frobt .2019 .00110](https://doi.org/10.3389/frobt.2019.00110). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7805891/> (accessed on 11/04/2023).
- Clancey, W. J. (1995). “A tutorial on situated learning”. In: *Proceedings of the international conference on computers and education (Taiwan)*. Charlottesville, VA: AACE. 49–70.
- Clare, A. S., J. C. Macbeth, and M. L. Cummings. (2012). “Mixed-initiative strategies for real-time scheduling of multiple unmanned vehicles”. In: *2012 American Control Conference (ACC)*. IEEE. 676–682.
- Clever, H. M., A. Handa, H. Mazhar, K. Parker, O. Shapira, Q. Wan, Y. Narang, I. Akinola, M. Cakmak, and D. Fox. (2021). “Assistive teleop: Leveraging transformers to collect robotic task demonstrations”. *arXiv preprint arXiv:2112.05129*.
- Codevilla, F., M. Mueller, A. Dosovitskiy, A. López, and V. Koltun. (2017). “End-to-end Driving via Conditional Imitation Learning”. Oct.
- Colman, A. M., B. D. Pulford, and C. L. Lawrence. (2014). “Explaining strategic coordination: Cognitive hierarchy theory, strong Stackelberg reasoning, and team reasoning.” *Decision*. 1(1): 35.
- Conmy, A., A. N. Mavor-Parker, A. Lynch, S. Heimersheim, and A. Garriga-Alonso. (2023). “Towards automated circuit discovery for mechanistic interpretability”. *arXiv preprint arXiv:2304.14997*.
- Coronado, E., T. Kiyokawa, G. A. G. Ricardez, I. G. Ramirez-Alpizar, G. Venture, and N. Yamanobe. (2022). “Evaluating quality in human-robot interaction: A systematic search and classification of performance and human-centered factors, measures and metrics towards an industry 5.0”. *Journal of Manufacturing Systems*. 63: 392–410.

- Cortes, C. and V. Vapnik. (1995). “Support-vector networks”. en. *Machine Learning*. 20(3): 273–297. ISSN: 1573-0565. DOI: [10.1007/BF00994018](https://doi.org/10.1007/BF00994018). URL: <https://doi.org/10.1007/BF00994018>.
- Cruz, F., S. Magg, C. Weber, and S. Wermter. (2016). “Training agents with interactive reinforcement learning and contextual affordances”. *IEEE Transactions on Cognitive and Developmental Systems*. 8(4): 271–284.
- Cruz, F., C. Young, R. Dazeley, and P. Vamplew. (2022). “Evaluating human-like explanations for robot actions in reinforcement learning scenarios”. In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 894–901.
- Cuayáhuitl, H., M. Van Otterlo, N. Dethlefs, and L. Frommberger. (2013). “Machine learning for interactive systems and robots: a brief introduction”. In: *Proceedings of the 2nd Workshop on Machine Learning for Interactive Systems: Bridging the Gap Between Perception, Action and Communication*. 19–28.
- Cultrera, L., L. Seidenari, F. Becattini, P. Pala, and A. Del Bimbo. (2020). “Explaining autonomous driving by learning end-to-end visual attention”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. 340–341.
- Dai, Y. and S.-G. Lee. (2020). “Perception, Planning and Control for Self-Driving System Based on On-board Sensors”. en. *Advances in Mechanical Engineering*. 12(9): 1687814020956494. ISSN: 1687-8132. DOI: [10.1177/1687814020956494](https://doi.org/10.1177/1687814020956494). URL: <https://doi.org/10.1177/1687814020956494>.
- Das, A., S. Datta, G. Gkioxari, S. Lee, D. Parikh, and D. Batra. (2018). “Embodied Question Answering”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Das, D., S. Banerjee, and S. Chernova. (2021a). “Explainable AI for Robot Failures: Generating Explanations that Improve User Assistance in Fault Recovery”. In: *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction. HRI '21*. New York, NY, USA: Association for Computing Machinery. 351–360. ISBN: 978-1-4503-8289-2. DOI: [10.1145/3434073.3444657](https://doi.org/10.1145/3434073.3444657). URL: <https://dl.acm.org/doi/10.1145/3434073.3444657>.

- Das, D., S. Banerjee, and S. Chernova. (2021b). “Explainable AI for robot failures: Generating explanations that improve user assistance in fault recovery”. In: *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 351–360.
- Das, D. and S. Chernova. (2021). “Semantic-based explainable ai: Leveraging semantic scene graphs and pairwise ranking to explain robot failures”. In: *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 3034–3041.
- Das, D., S. Chernova, and B. Kim. (2023). “State2explanation: Concept-based explanations to benefit agent learning and user understanding”. *arXiv preprint arXiv:2309.12482*.
- Datta, T. and J. P. Dickerson. (2023). “Who’s Thinking? A Push for Human-Centered Evaluation of LLMs using the XAI Playbook”. *arXiv preprint arXiv:2303.06223*.
- Datteri, E. (2013). “Predicting the long-term effects of human-robot interaction: A reflection on responsibility in medical robotics”. *Science and engineering ethics*. 19: 139–160.
- Dautenhahn, K. (2007). “Socially intelligent robots: dimensions of human–robot interaction”. *Philosophical transactions of the royal society B: Biological sciences*. 362(1480): 679–704.
- De Graaf, M. M., S. Ben Allouch, and J. A. van Dijk. (2016). “Long-term evaluation of a social robot in real homes”. *Interaction studies*. 17(3): 462–491.
- Deits, R., S. Tellex, P. Thaker, D. Simeonov, T. Kollar, and N. Roy. (2013). “Clarifying commands with information-theoretic human-robot dialog”. *Journal of Human-Robot Interaction*. 2(2): 58–79.
- Del Duchetto, F. and M. Hanheide. (2022). “Learning on the Job: Long-Term Behavioural Adaptation in Human-Robot Interactions”. *IEEE Robotics and Automation Letters*. 7(3): 6934–6941.
- Del Duchetto, F., A. Kucukyilmaz, L. Iocchi, and M. Hanheide. (2018). “Do Not Make the Same Mistakes Again and Again: Learning Local Recovery Policies for Navigation From Human Demonstrations”. *IEEE Robotics and Automation Letters*. 3(4): 4084–4091. DOI: [10.1109/LRA.2018.2861080](https://doi.org/10.1109/LRA.2018.2861080).

- Dhonthi, A., P. Schillinger, L. Rozo, and D. Nardi. (2022). “Optimizing demonstrated robot manipulation skills for temporal logic constraints”. In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 1255–1262.
- Dhuliawala, S., M. Komeili, J. Xu, R. Raileanu, X. Li, A. Celikyilmaz, and J. Weston. (2023). “Chain-of-verification reduces hallucination in large language models”. *arXiv preprint arXiv:2309.11495*.
- Diaz, M., T. Fevens, and L. Paull. (2021). “Uncertainty-Aware Policy Sampling and Mixing for Safe Interactive Imitation Learning”. In: *2021 18th Conference on Robots and Vision (CRV)*. IEEE. 72–78.
- Diethelm, I. G., S. S. Hansen, F. B. Leth, K. Fischer, and O. Palinko. (2021). “Effects of Gaze and Speech in Human-Robot Medical Interactions”. In: *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction. HRI '21 Companion*. Boulder, CO, USA: Association for Computing Machinery. 349–353. ISBN: 9781450382908. DOI: [10.1145/3434074.3447190](https://doi.org/10.1145/3434074.3447190). URL: <https://doi.org/10.1145/3434074.3447190>.
- Dodeja, L., P. Tambwekar, E. Hedlund-Botti, and M. Gombolay. (2024). “Towards the design of user-centric strategy recommendation systems for collaborative Human-AI tasks”. *International Journal of Human-Computer Studies*: 103216.
- Donzé, A. and O. Maler. (2010). “Robust satisfaction of temporal logic over real-valued signals”. In: *Int. Conf. on Formal Modeling and Analysis of Timed Systems*. 92–106.
- Dragan, A. D. and S. S. Srinivasa. (2013). “A policy-blending formalism for shared control”. en. *The International Journal of Robotics Research*. 32(7): 790–805. ISSN: 0278-3649, 1741-3176. DOI: [10.1177/0278364913490324](http://journals.sagepub.com/doi/10.1177/0278364913490324). URL: <http://journals.sagepub.com/doi/10.1177/0278364913490324>.
- Dragan, A. D., K. C. Lee, and S. S. Srinivasa. (2013). “Legibility and predictability of robot motion”. In: *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 301–308. DOI: [10.1109/HRI.2013.6483603](https://doi.org/10.1109/HRI.2013.6483603).

- Driess, D., F. Xia, M. S. M. Sajjadi, C. Lynch, A. Chowdhery, B. Ichter, A. Wahid, J. Tompson, Q. Vuong, T. Yu, W. Huang, Y. Chebotar, P. Sermanet, D. Duckworth, S. Levine, V. Vanhoucke, K. Hausman, M. Toussaint, K. Greff, A. Zeng, I. Mordatch, and P. Florence. (2023). “PaLM-E: An Embodied Multimodal Language Model”. In: *Proceedings of the 40th International Conference on Machine Learning*. Ed. by A. Krause, E. Brunskill, K. Cho, B. Engelhardt, S. Sabato, and J. Scarlett. Vol. 202. *Proceedings of Machine Learning Research*. PMLR. 8469–8488. URL: <https://proceedings.mlr.press/v202/driess23a.html>.
- Du Toit, N. E. and J. W. Burdick. (2011). “Probabilistic collision checking with chance constraints”. *IEEE Transactions on Robotics*. 27(4): 809–815.
- Durso, F. T. and S. D. Gronlund. (1999). “Situation awareness”. *Handbook of applied cognition*. 283: 314.
- Dwork, C. (2006). “Differential privacy”. In: *International colloquium on automata, languages, and programming*. Springer. 1–12.
- Dwork, C. (2008). “Differential privacy: A survey of results”. In: *International conference on theory and applications of models of computation*. Springer. 1–19.
- Dwork, C., M. Hardt, T. Pitassi, O. Reingold, and R. Zemel. (2012). “Fairness through awareness”. In: *Proceedings of the 3rd innovations in theoretical computer science conference*. 214–226.
- Edmonds, M., F. Gao, H. Liu, X. Xie, S. Qi, B. Rothrock, Y. Zhu, Y. N. Wu, H. Lu, and S.-C. Zhu. (2019a). “A tale of two explanations: Enhancing human trust by explaining robot behavior”. *Science Robotics*. 4(37): eaay4663. DOI: [10.1126/scirobotics.aay4663](https://doi.org/10.1126/scirobotics.aay4663). eprint: <https://www.science.org/doi/pdf/10.1126/scirobotics.aay4663>. URL: <https://www.science.org/doi/abs/10.1126/scirobotics.aay4663>.
- Edmonds, M., F. Gao, H. Liu, X. Xie, S. Qi, B. Rothrock, Y. Zhu, Y. N. Wu, H. Lu, and S.-C. Zhu. (2019b). “A tale of two explanations: Enhancing human trust by explaining robot behavior”. *Science Robotics*. 4(37): eaay4663. DOI: [10.1126/scirobotics.aay4663](https://doi.org/10.1126/scirobotics.aay4663). eprint: <https://www.science.org/doi/pdf/10.1126/scirobotics.aay4663>. URL: <https://www.science.org/doi/abs/10.1126/scirobotics.aay4663>.

- Ehsan, U. and M. O. Riedl. (2020). “Human-centered explainable ai: Towards a reflective sociotechnical approach”. In: *HCI International 2020-Late Breaking Papers: Multimodality and Intelligence: 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings 22*. Springer. 449–466.
- Ehsan, U., P. Wintersberger, Q. V. Liao, E. A. Watkins, C. Manger, H. Daumé III, A. Riener, and M. O. Riedl. (2022). “Human-Centered Explainable AI (HCXAI): beyond opening the black-box of AI”. In: *CHI conference on human factors in computing systems extended abstracts*. 1–7.
- Ekvall, S. and D. Kragic. (2008). “Robot learning from demonstration: a task-level planning approach”. *International Journal of Advanced Robotic Systems*. 5(3): 33.
- Enayati, N., G. Ferrigno, and E. De Momi. (2018). “Skill-based human–robot cooperation in tele-operated path tracking”. *Autonomous Robots*. 42: 997–1009.
- Endsley, M. R. (1995). “Measurement of situation awareness in dynamic systems”. *Human factors*. 37(1): 65–84.
- Endsley, M. R. (2023). “Supporting Human-AI Teams: Transparency, explainability, and situation awareness”. *Computers in Human Behavior*. 140: 107574.
- Escarabajal, R. J., J. L. Pulloquina, P. Zamora-Ortiz, Á. Valera, V. Mata, and M. Vallés. (2023). “Imitation Learning-Based System for the Execution of Self-Paced Robotic-Assisted Passive Rehabilitation Exercises”. en. *IEEE Robotics and Automation Letters*. 8(7): 4283–4290. ISSN: 2377-3766, 2377-3774. DOI: [10.1109/LRA.2023.3281884](https://doi.org/10.1109/LRA.2023.3281884). URL: <https://ieeexplore.ieee.org/document/10141665/> (accessed on 11/03/2023).
- Esterwood, C. and L. P. R. Jr. (2023). “Three Strikes and you are out!: The impacts of multiple human–robot trust violations and repairs on robot trustworthiness”. *Computers in Human Behavior*. 142(May): 107658. ISSN: 0747-5632. DOI: [10.1016/j.chb.2023.107658](https://doi.org/10.1016/j.chb.2023.107658). URL: <https://www.sciencedirect.com/science/article/pii/S0747563223000092>.

- European Parliament. (2017). “European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL))”. URL: <https://op.europa.eu/en/publication-detail/-/publication/13fd56d0-8a65-11e8-ac6a-01aa75ed71a1/language-en/format-PDF>.
- European Parliament and Council of the European Union. (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council*. of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). OJ L 119, 4.5.2016, p. 1–88. URL: <https://data.europa.eu/eli/reg/2016/679/oj> (accessed on 04/13/2023).
- Eysenbach, B., R. R. Salakhutdinov, and S. Levine. (2019). “Search on the replay buffer: Bridging planning and reinforcement learning”. *Advances in Neural Information Processing Systems*. 32.
- Eysenbach, B., R. Salakhutdinov, and S. Levine. (2022). “The Information Geometry of Unsupervised Reinforcement Learning”. In: *International Conference on Learning Representations*. URL: <https://openreview.net/forum?id=3wU2UX0voE>.
- Faccio, M., I. Granata, A. Menini, M. Milanese, C. Rossato, M. Bottin, R. Minto, P. Pluchino, L. Gamberini, G. Boschetti, *et al.* (2023). “Human factors in cobot era: A review of modern production systems features”. *Journal of Intelligent Manufacturing*. 34(1): 85–106.
- Fainekos, G. E., H. Kress-Gazit, and G. J. Pappas. (2005). “Temporal logic motion planning for mobile robots”. In: *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*. IEEE. 2020–2025.
- Feil-Seifer, D. and M. Mataric. (2005). “Defining socially assistive robotics”. In: *9th International Conference on Rehabilitation Robotics, 2005. ICORR 2005*. 465–468. DOI: [10.1109/ICORR.2005.1501143](https://doi.org/10.1109/ICORR.2005.1501143).
- Feil-Seifer, D. and M. J. Matarić. (2009). “Toward Socially Assistive Robotics for Augmenting Interventions for Children with Autism Spectrum Disorders”. In: *Experimental Robotics*. Ed. by O. Khatib, V. Kumar, and G. J. Pappas. Berlin, Heidelberg: Springer Berlin Heidelberg. 201–210. ISBN: 978-3-642-00196-3.

- Ferraguti, F., M. Bertuletti, C. T. Landi, M. Bonfè, C. Fantuzzi, and C. Secchi. (2020). “A control barrier function approach for maximizing performance while fulfilling to ISO/TS 15066 regulations”. *IEEE Robotics and Automation Letters*. 5(4): 5921–5928.
- Ferrer, X., T. van Nuenen, J. M. Such, M. Coté, and N. Criado. (2021). “Bias and discrimination in AI: a cross-disciplinary perspective”. *IEEE Technology and Society Magazine*. 40(2): 72–80.
- Fiazza, M.-C. and P. Fiorini. (2021). “Design for Interpretability: Meeting the Certification Challenge for Surgical Robots”. In: *2021 IEEE International Conference on Intelligence and Safety for Robotics (ISR)*. 264–267. DOI: [10.1109/ISR50024.2021.9419378](https://doi.org/10.1109/ISR50024.2021.9419378).
- Finucane, C., G. Jing, and H. Kress-Gazit. (2010). “LTLMoP: Experimenting with language, temporal logic and robot control”. In: *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*. 1988–1993.
- Finzi, A. and A. Orlandini. (2005). “A mixed-initiative approach to human-robot interaction in rescue scenarios”. In: *American Association for Artificial Intelligence*.
- Firoozi, R., J. Tucker, S. Tian, A. Majumdar, J. Sun, W. Liu, Y. Zhu, S. Song, A. Kapoor, K. Hausman, B. Ichter, D. Driess, J. Wu, C. Lu, and M. Schwager. (2023). “Foundation Models in Robotics: Applications, Challenges, and the Future”. arXiv: [2312.07843](https://arxiv.org/abs/2312.07843) [cs.R0].
- Fiske, A., P. Henningsen, and A. Buyx. (2019). “Your Robot Therapist Will See You Now: Ethical Implications of Embodied Artificial Intelligence in Psychiatry, Psychology, and Psychotherapy”. *J Med Internet Res*. 21(5): e13216. ISSN: 1438-8871. DOI: [10.2196/13216](https://doi.org/10.2196/13216). URL: <http://www.ncbi.nlm.nih.gov/pubmed/31094356>.
- Fitzgerald, T., P. Koppol, P. Callaghan, R. Q. J. H. Wong, R. Simmons, O. Kroemer, and H. Admoni. (2022). “INQUIRE: INteractive Querying for User-aware Informative REasoning”. In: *6th Annual Conference on Robot Learning*.
- Fitzgerald, T., E. Short, A. Goel, and A. Thomaz. (2019). “Human-guided trajectory adaptation for tool transfer”. In: *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*. 1350–1358.

- Fleck, M. M., D. A. Forsyth, and C. Bregler. (1996). “Finding naked people”. In: *Computer Vision—ECCV’96: 4th European Conference on Computer Vision Cambridge, UK, April 15–18, 1996 Proceedings Volume II 4*. Springer. 593–602.
- Floreano, D., P. Husbands, and S. Nolfi. (2008). “Evolutionary Robotics”. In: 1423–1451. DOI: [10.1007/978-3-540-30301-5_62](https://doi.org/10.1007/978-3-540-30301-5_62).
- Florence, P., C. Lynch, A. Zeng, O. A. Ramirez, A. Wahid, L. Downs, A. Wong, J. Lee, I. Mordatch, and J. Tompson. (2022). “Implicit behavioral cloning”. In: *Conference on Robot Learning*. PMLR. 158–168.
- French, K., S. Wu, T. Pan, Z. Zhou, and O. Jenkins. (2019). “Learning Behavior Trees From Demonstration”. In: 7791–7797. DOI: [10.1109/ICRA.2019.8794104](https://doi.org/10.1109/ICRA.2019.8794104).
- French, R. M. (1999). “Catastrophic forgetting in connectionist networks”. *Trends in Cognitive Sciences*. 3(4): 128–135. ISSN: 1364-6613. DOI: [https://doi.org/10.1016/S1364-6613\(99\)01294-2](https://doi.org/10.1016/S1364-6613(99)01294-2). URL: <https://www.sciencedirect.com/science/article/pii/S1364661399012942>.
- Friedman, B. and H. Nissenbaum. (1996). “Bias in computer systems”. *ACM Transactions on information systems (TOIS)*. 14(3): 330–347.
- Friesen, P., R. Douglas-Jones, M. Marks, R. Pierce, K. Fletcher, A. Mishra, J. Lorimer, C. Véliz, N. Hallowell, M. Graham, *et al.* (2021). “Governing AI-Driven Health Research: Are IRBs Up to the Task?” *Ethics & Human Research*. 43(2): 35–42.
- Frith, C. and U. Frith. (2005). “Theory of mind”. *Current biology*. 15(17): R644–R645.
- Fu, J., K. Luo, and S. Levine. (2017). “Learning Robust Rewards with Adversarial Inverse Reinforcement Learning”. *ArXiv*. abs/1710.11248.
- Fu, Z., T. Z. Zhao, and C. Finn. (2024). “Mobile ALOHA: Learning Bimanual Mobile Manipulation with Low-Cost Whole-Body Teleoperation”. In: *arXiv*.
- Fulton, M., C. Edge, and J. Sattar. (2019). “Robot communication via motion: Closing the underwater human-robot interaction loop”. In: *2019 International Conference on Robotics and Automation (ICRA)*. IEEE. 4660–4666.

- Fulton, N. and A. Platzer. (2018). “Safe reinforcement learning via formal methods: Toward safe control through proof and learning”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 32. No. 1.
- Gabriel, I. (2020). “Artificial intelligence, values, and alignment”. *Minds and machines*. 30(3): 411–437.
- Galindo, C., J.-A. Fernández-Madrigal, J. González, and A. Saffiotti. (2008). “Robot task planning using semantic maps”. *Robotics and autonomous systems*. 56(11): 955–966.
- Galindo, C., A. Saffiotti, S. Coradeschi, P. Buschka, J.-A. Fernandez-Madrigal, and J. González. (2005). “Multi-hierarchical semantic maps for mobile robotics”. In: *2005 IEEE/RSJ international conference on intelligent robots and systems*. IEEE. 2278–2283.
- Galli, A., S. Marrone, V. Moscato, and C. Sansone. (2021). “Reliability of explainable artificial intelligence in adversarial perturbation scenarios”. In: *International Conference on Pattern Recognition*. Springer. 243–256.
- Gao, C., H. Gao, S. Guo, T. Zhang, and F. Chen. (2021). “CRIL: Continual Robot Imitation Learning via Generative and Prediction Model”. In: *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 6747–5754. DOI: [10.1109/IROS51168.2021.9636069](https://doi.org/10.1109/IROS51168.2021.9636069).
- Gao, X., R. Gong, Y. Zhao, S. Wang, T. Shu, and S.-C. Zhu. (2020a). “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)*. 1119–1126. DOI: [10.1109/RO-MAN47096.2020.9223595](https://doi.org/10.1109/RO-MAN47096.2020.9223595).
- Gao, X., R. Gong, Y. Zhao, S. Wang, T. Shu, and S.-C. Zhu. (2020b). “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. 1119–1126.
- Gast, J., A. Bannat, T. Rehrl, F. Wallhoff, G. Rigoll, C. Wendt, S. Schmidt, M. Popp, and B. Farber. (2009). “Real-time framework for multimodal human-robot interaction”. In: *2009 2nd Conference on Human System Interactions*. IEEE. 276–283.

- Gaurav, A., K. Rezaee, G. Liu, and P. Poupart. (2022). “Learning soft constraints from constrained expert demonstrations”. *arXiv preprint arXiv:2206.01311*.
- Gawron, V. J. (2008). *Human performance, workload, and situational awareness measures handbook*. Crc Press.
- “General Data Protection Regulation (GDPR) – Official Legal Text”. (2016). <https://gdpr-info.eu/>.
- Gepperth, A. and B. Hammer. (2016). “Incremental learning algorithms and applications”. In: *European Symposium on Artificial Neural Networks (ESANN)*. Bruges, Belgium.
- Gero, K. I., Z. Ashktorab, C. Dugan, Q. Pan, J. Johnson, W. Geyer, M. Ruiz, S. Miller, D. R. Millen, M. Campbell, *et al.* (2020). “Mental models of AI agents in a cooperative game setting”. In: *Proceedings of the 2020 chi conference on human factors in computing systems*. 1–12.
- Gerstner, T., D. DeCarlo, M. Alexa, A. Finkelstein, Y. Gingold, and A. Nealen. (2012). “Pixelated image abstraction”. In: *Proceedings of the Symposium on Non-Photorealistic Animation and Rendering*. Citeseer. 29–36.
- Ghadirzadeh, A., X. Chen, P. Poklukar, C. Finn, M. Björkman, and D. Kragic. (2021). “Bayesian meta-learning for few-shot policy adaptation across robotic platforms”. In: *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 1274–1280.
- Gildert, N., A. G. Millard, A. Pomfret, and J. Timmis. (2018). “The need for combining implicit and explicit communication in cooperative robotic systems”. *Frontiers in Robotics and AI*. 5: 65.
- Gilpin, L. H., A. R. Paley, M. A. Alam, S. Spurlock, and K. J. Hammond. (2022). ““Explanation” is Not a Technical Term: The Problem of Ambiguity in XAI”. *arXiv preprint arXiv:2207.00007*.
- Goldman, C. V., M. Baltaxe, D. Chakraborty, and J. Arinez. (2021). “Explaining learning models in manufacturing processes”. *Procedia Computer Science*. 180: 259–268.

- Gombolay, M., X. J. Yang, B. Hayes, N. Seo, Z. Liu, S. Wadhwa-
nia, T. Yu, N. Shah, T. Golen, and J. Shah. (2018a). “Robotic
assistance in the coordination of patient care”. *The International
Journal of Robotics Research*. 37(10): 1300–1316. DOI: [10.1177/
0278364918778344](https://doi.org/10.1177/0278364918778344). eprint: <https://doi.org/10.1177/0278364918778344>.
URL: <https://doi.org/10.1177/0278364918778344>.
- Gombolay, M. C., R. A. Gutierrez, S. G. Clarke, G. F. Sturla, and J. A.
Shah. (2015). “Decision-making authority, team efficiency and hu-
man worker satisfaction in mixed human–robot teams”. *Autonomous
Robots*. 39: 293–312.
- Gombolay, M. C., R. J. Wilcox, and J. A. Shah. (2018b). “Fast schedul-
ing of robot teams performing tasks with temporospatial constraints”.
IEEE Transactions on Robotics. 34(1): 220–239.
- Gomez-Gonzalez, S., G. Neumann, B. Schölkopf, and J. Peters. (2016).
“Using probabilistic movement primitives for striking movements”.
In: *2016 IEEE-RAS 16th International Conference on Humanoid
Robots (Humanoids)*. 502–508. DOI: [10.1109/HUMANOIDS.2016.
7803322](https://doi.org/10.1109/HUMANOIDS.2016.7803322).
- Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley,
S. Ozair, A. Courville, and Y. Bengio. (2014). “Generative Adversar-
ial Nets”. In: *Advances in Neural Information Processing Systems*.
Ed. by Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and
K. Q. Weinberger. Vol. 27. Curran Associates, Inc. URL: [https://
proceedings.neurips.cc/paper_files/paper/2014/file/5ca3e9b122f61f8f06494c97b
Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2014/file/5ca3e9b122f61f8f06494c97bPaper.pdf).
- Goodin, R. E. (1996). “Inclusion and exclusion”. *European Journal of
Sociology*. 37(2): 343–371. DOI: [10.1017/S0003975600007219](https://doi.org/10.1017/S0003975600007219).
- Goodrich, M. A., A. C. Schultz, *et al.* (2008). “Human–robot inter-
action: a survey”. *Foundations and Trends® in Human–Computer
Interaction*. 1(3): 203–275.
- Gopalan, N., D. Arumugam, L. Wong, and S. Tellex. (2018). “Sequence-
to-Sequence Language Grounding of Non-Markovian Task Specifica-
tions”. In: *Proceedings of Robotics: Science and Systems*. Pittsburgh,
Pennsylvania. DOI: [10.15607/RSS.2018.XIV.067](https://doi.org/10.15607/RSS.2018.XIV.067).

- Görür, O., B. Rosman, G. Hoffman, and S. Albayrak. (2017). “Toward Integrating Theory of Mind into Adaptive Decision- Making of Social Robots to Understand Human Intention”. In:
- Gottardi, A., S. Tortora, E. Tosello, and E. Menegatti. (2022). “Shared Control in Robot Teleoperation With Improved Potential Fields”. *IEEE Transactions on Human-Machine Systems*. 52(3): 410–422. ISSN: 2168-2305. DOI: [10.1109/THMS.2022.3155716](https://doi.org/10.1109/THMS.2022.3155716). URL: <https://ieeexplore.ieee.org/document/9734752>.
- Goyal, P., S. Niekum, and R. Mooney. (2021). “Pixl2r: Guiding reinforcement learning using natural language by mapping pixels to rewards”. In: *Conference on Robot Learning*. PMLR. 485–497.
- Goyal, P., S. Niekum, and R. J. Mooney. (2019). “Using Natural Language for Reward Shaping in Reinforcement Learning”. In: *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*. International Joint Conferences on Artificial Intelligence Organization. 2385–2391. DOI: [10.24963/ijcai.2019/331](https://doi.org/10.24963/ijcai.2019/331). URL: <https://doi.org/10.24963/ijcai.2019/331>.
- Grannen, J. and Y. Wu. (2022). “Learning Bimanual Scooping Policies for Food Acquisition”. In: *Proceedings of the 6th Conference on Robot Learning (CoRL)*.
- Greca, I. M. and M. A. Moreira. (2000). “Mental models, conceptual models, and modelling”. *International journal of science education*. 22(1): 1–11.
- Grgic-Hlaca, N., M. B. Zafar, K. P. Gummadi, and A. Weller. (2016). “The case for process fairness in learning: Feature selection for fair decision making”. In: *NIPS symposium on machine learning and the law*. Vol. 1. No. 2. Barcelona, Spain. 11.
- Griffith, S., K. Subramanian, J. Scholz, C. L. Isbell, and A. L. Thomaz. (2013). “Policy shaping: Integrating human feedback with reinforcement learning”. *Advances in neural information processing systems*. 26.
- Grigore, E. C., K. Eder, A. G. Pipe, C. Melhuish, and U. Leonards. (2013). “Joint action understanding improves robot-to-human object handover”. In: *2013 IEEE/RSJ international conference on intelligent robots and systems*. IEEE. 4622–4629.

- Grollman, D. H. and O. C. Jenkins. (2007). “Dogged Learning for Robots”. In: *Proceedings 2007 IEEE International Conference on Robotics and Automation*. 2483–2488. DOI: [10.1109/ROBOT.2007.363692](https://doi.org/10.1109/ROBOT.2007.363692).
- Gross, H.-M., H. Boehme, C. Schroeter, S. Mueller, A. Koenig, E. Einhorn, C. Martin, M. Merten, and A. Bley. (2009). “TOOMAS: Interactive Shopping Guide Robots in Everyday Use - Final Implementation and Experiences from Long-Term Field Trials”. In: *Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS'09*. St. Louis, MO, USA: IEEE Press. 2005–2012. ISBN: 9781424438037.
- Gross, H.-M. and H.-J. Boehme. (2000). “PERSES-a vision-based interactive mobile shopping assistant”. In: *Smc 2000 conference proceedings. 2000 ieee international conference on systems, man and cybernetics. 'cybernetics evolving to systems, humans, organizations, and their complex interactions' (cat. no.0*. Vol. 1. 80–85 vol.1. DOI: [10.1109/ICSMC.2000.884968](https://doi.org/10.1109/ICSMC.2000.884968).
- Grubbs, V. (2020). “Diversity, Equity, and Inclusion That Matter”. en. *N Engl J Med*. 383(4): e25.
- Grushko, S., A. Vysocký, D. Heczko, and Z. Bobovský. (2021a). “Intuitive Spatial Tactile Feedback for Better Awareness about Robot Trajectory during Human–Robot Collaboration”. *Sensors*. 21(17). ISSN: 1424-8220. DOI: [10.3390/s21175748](https://doi.org/10.3390/s21175748). URL: <https://www.mdpi.com/1424-8220/21/17/5748>.
- Grushko, S., A. Vysocký, P. Oščádal, M. Vocetka, P. Novák, and Z. Bobovský. (2021b). “Improved Mutual Understanding for Human–Robot Collaboration: Combining Human-Aware Motion Planning with Haptic Feedback Devices for Communicating Planned Trajectory”. *Sensors*. 21(11). ISSN: 1424-8220. DOI: [10.3390/s21113673](https://doi.org/10.3390/s21113673). URL: <https://www.mdpi.com/1424-8220/21/11/3673>.
- Gu, S., L. Yang, Y. Du, G. Chen, F. Walter, J. Wang, Y. Yang, and A. Knoll. (2022). “A review of safe reinforcement learning: Methods, theory and applications”. *arXiv preprint arXiv:2205.10330*.

- Guemghar, I., P. Pires de Oliveira Padilha, A. Abdel-Baki, D. Jutras-Aswad, J. Paquette, and M.-P. Pomey. (2022). “Social Robot Interventions in Mental Health Care and Their Outcomes, Barriers, and Facilitators: Scoping Review”. *JMIR Mental Health*. 9(4): e36094. ISSN: 2368-7959. DOI: [10.2196/36094](https://doi.org/10.2196/36094). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9066335/> (accessed on 11/05/2023).
- Guerrero, C. R., J. F. Marinero, J. P. Turiel, and P. R. Farina. (2012). “Using psychophysiological feedback to enhance physical human robot interaction in a cooperative scenario”. In: *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*. IEEE. 1176–1181.
- Guidotti, R., A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, and D. Pedreschi. (2018). “A survey of methods for explaining black box models”. *ACM computing surveys (CSUR)*. 51(5): 1–42.
- Gundana, D. and H. Kress-Gazit. (2021). “Event-based signal temporal logic synthesis for single and multi-robot tasks”. *IEEE Robotics and Automation Letters*. 6(2): 3687–3694.
- Gunkel, D. J. (2020). “Mind the gap: responsible robotics and the problem of responsibility”. *Ethics and Information Technology*. 22(4): 307–320.
- Guo, C., T. Bai, X. Wang, X. Zhang, Y. Lu, X. Dai, and F.-Y. Wang. (2022). “ShadowPainter: Active learning enabled robotic painting through visual measurement and reproduction of the artistic creation process”. *Journal of Intelligent & Robotic Systems*. 105(3).
- Guo, M., S. Andersson, and D. V. Dimarogonas. (2018). “Human-in-the-Loop Mixed-Initiative Control Under Temporal Tasks”. In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. 6395–6400. DOI: [10.1109/ICRA.2018.8460793](https://doi.org/10.1109/ICRA.2018.8460793).
- Gupta, A., A. Anpalagan, L. Guan, and A. S. Khwaja. (2021). “Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues”. *Array*. 10(July): 100057. ISSN: 2590-0056. DOI: [10.1016/j.array.2021.100057](https://doi.org/10.1016/j.array.2021.100057). URL: <https://www.sciencedirect.com/science/article/pii/S2590005621000059>.

- Haddadin, S. (2015). “Physical safety in robotics”. *Formal Modeling and Verification of Cyber-Physical Systems: 1st International Summer School on Methods and Tools for the Design of Digital Systems, Bremen, Germany, September 2015*: 249–271.
- Haddadin, S., A. Albu-Schaeffer, F. Haddadin, J. Roßmann, and G. Hirzinger. (2011). “An experimental safety study for stab/puncture and incised wounds”. *IEEE Robotics & Automation Magazine*. 12: 20–34.
- Haddadin, S., A. Albu-Schaeffer, A. De Luca, and G. Hirzinger. (2008). “Collision detection and reaction: A contribution to safe physical human-robot interaction”. In: *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 3356–3363.
- Hagendorff, T. (2020a). “Ethical behavior in humans and machines—Evaluating training data quality for beneficial machine learning”. *arXiv preprint arXiv:2008.11463*.
- Hagendorff, T. (2020b). “The ethics of AI ethics: An evaluation of guidelines”. *Minds and machines*. 30(1): 99–120.
- Halme, R.-J., M. Lanz, J. Kämäräinen, R. Pieters, J. Latokartano, and A. Hietanen. (2018). “Review of vision-based safety systems for human-robot collaboration”. *Procedia Cirp*. 72: 111–116.
- Han, Z., D. Giger, J. Allspaw, M. S. Lee, H. Admoni, and H. A. Yanco. (2021). “Building the Foundation of Robot Explanation Generation Using Behavior Trees”. *J. Hum.-Robot Interact*. 10(3). DOI: [10.1145/3457185](https://doi.org/10.1145/3457185). URL: <https://doi.org/10.1145/3457185>.
- Hancock, P. A., D. R. Billings, and K. E. Schaefer. (2011a). “Can You Trust Your Robot?” en. *Ergonomics in Design*. 19(3): 24–29. ISSN: 1064-8046. DOI: [10.1177/1064804611415045](https://doi.org/10.1177/1064804611415045). URL: <https://doi.org/10.1177/1064804611415045>.
- Hancock, P. A., D. R. Billings, K. E. Schaefer, J. Y. C. Chen, E. J. de Visser, and R. Parasuraman. (2011b). “A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction”. en. *Human Factors*. 53(5): 517–527. ISSN: 0018-7208. DOI: [10.1177/0018720811417254](https://doi.org/10.1177/0018720811417254). URL: <https://doi.org/10.1177/0018720811417254>.

- Hannum, C., R. Li, and W. Wang. (2023). “A Trust-Assist Framework for Human–Robot Co-Carry Tasks”. en. *Robotics*. 12(2): 30. ISSN: 2218-6581. DOI: [10.3390/robotics12020030](https://doi.org/10.3390/robotics12020030). URL: <https://www.mdpi.com/2218-6581/12/2/30>.
- Hansen, N., Y. Lin, H. Su, X. Wang, V. Kumar, and A. Rajeswaran. (2022). “Modem: Accelerating visual model-based reinforcement learning with demonstrations”. *arXiv preprint arXiv:2212.05698*.
- Hardt, M., E. Price, and N. Srebro. (2016). “Equality of opportunity in supervised learning”. *Advances in neural information processing systems*. 29.
- Haruna, M., N. Kawaguchi, M. Ogino, and T. Koike-Akino. (2021). “Comparison of three feedback modalities for haptics sensation in remote machine manipulation”. *IEEE Robotics and Automation Letters*. 6(3): 5040–5047.
- Harvey, I., P. Husbands, and D. Cliff. (1994). “Seeing the Light: Artificial Evolution, Real Vision”. In: *Proceedings of the Third International Conference on Simulation of Adaptive Behavior: From Animals to Animats 3: From Animals to Animats 3. SAB94*. Brighton, United Kingdom: MIT Press. 392–401. ISBN: 0262531224.
- Hatori, J., Y. Kikuchi, S. Kobayashi, K. Takahashi, Y. Tsuboi, Y. Unno, W. Ko, and J. Tan. (2018). “Interactively Picking Real-World Objects with Unconstrained Spoken Language Instructions”. In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. 3774–3781. DOI: [10.1109/ICRA.2018.8460699](https://doi.org/10.1109/ICRA.2018.8460699).
- Hayashi, M., T. Miyake, M. Kamezaki, J. Yamato, K. Saito, T. Hamada, E. Sakurai, S. Sugano, and J. Ohya. (2022). “Preliminary investigation of collision risk assessment with vision for selecting targets paid attention to by mobile robot”. In: *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 624–629.
- Hayes, B. and B. Scassellati. (2014). “Discovering Task Constraints Through Observation and Active Learning”. In: DOI: [10.1109/IROS.2014.6943191](https://doi.org/10.1109/IROS.2014.6943191).

- Hayes, B. and J. A. Shah. (2017a). “Improving robot controller transparency through autonomous policy explanation”. In: *Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction*. 303–312.
- Hayes, B. and J. A. Shah. (2017b). “Improving Robot Controller Transparency Through Autonomous Policy Explanation”. In: *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction. HRI '17*. Vienna, Austria: Association for Computing Machinery. 303–312. ISBN: 9781450343367. DOI: [10.1145/2909824.3020233](https://doi.org/10.1145/2909824.3020233). URL: <https://doi.org/10.1145/2909824.3020233>.
- He, X., Z. Hu, H. Yang, and C. Lv. (2024). “Personalized robotic control via constrained multi-objective reinforcement learning”. *Neurocomputing*. 565: 126986. ISSN: 0925-2312. DOI: <https://doi.org/10.1016/j.neucom.2023.126986>. URL: <https://www.sciencedirect.com/science/article/pii/S0925231223011098>.
- Hedlund, E., M. Johnson, and M. Gombolay. (2021). “The Effects of a Robot’s Performance on Human Teachers for Learning from Demonstration Tasks”. In: *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 207–215.
- Hejna, J., R. Rafailov, H. Sikchi, C. Finn, S. Niekum, W. B. Knox, and D. Sadigh. (2023). “Contrastive preference learning: Learning from human feedback without rl”. *arXiv preprint arXiv:2310.13639*.
- Hellou, M., N. Gasteiger, J. Y. Lim, M. Jang, and H. S. Ahn. (2021). “Personalization and Localization in Human-Robot Interaction: A Review of Technical Methods”. *Robotics*. 10(4). ISSN: 2218-6581. DOI: [10.3390/robotics10040120](https://doi.org/10.3390/robotics10040120). URL: <https://www.mdpi.com/2218-6581/10/4/120>.
- Hengst, F. den, E. M. Grua, A. el Hassouni, and M. Hoogendoorn. (2020). “Reinforcement learning for personalization: A systematic literature review”. *Data Science*. 3(2): 107–147. ISSN: 2451-8492. DOI: [10.3233/DS-200028](https://doi.org/10.3233/DS-200028).

- Hermann, K. M., F. Hill, S. Green, F. Wang, R. Faulkner, H. Soyer, D. Szepesvari, W. M. Czarnecki, M. Jaderberg, D. Teplyashin, M. Wainwright, C. Apps, D. Hassabis, and P. Blunsom. (2017). “Grounded Language Learning in a Simulated 3D World”. *CoRR*. abs/1706.06551. arXiv: [1706.06551](https://arxiv.org/abs/1706.06551). URL: <http://arxiv.org/abs/1706.06551>.
- Herse, S., J. Vitale, and M.-A. Williams. (2023). “Simulation Evidence of Trust Calibration: Using POMDP with Signal Detection Theory to Adapt Agent Features for Optimised Task Outcome During Human-Agent Collaboration”. en. *International Journal of Social Robotics*. Aug. ISSN: 1875-4805. DOI: [10.1007/s12369-023-01041-w](https://doi.org/10.1007/s12369-023-01041-w). URL: <https://doi.org/10.1007/s12369-023-01041-w>.
- Heuvel, J. de, N. Corral, L. Bruckschen, and M. Bennewitz. (2022). “Learning Personalized Human-Aware Robot Navigation Using Virtual Reality Demonstrations from a User Study”. In: *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE. DOI: [10.1109/ro-man53752.2022.9900554](https://doi.org/10.1109/ro-man53752.2022.9900554). URL: <http://dx.doi.org/10.1109/RO-MAN53752.2022.9900554>.
- Hindemith, L., A. Vollmer, C. B. Wiebel-Herboth, and B. Wrede. (2021). “Improving HRI through robot architecture transparency”. *CoRR*. abs/2108.11608. arXiv: [2108.11608](https://arxiv.org/abs/2108.11608). URL: <https://arxiv.org/abs/2108.11608>.
- Hindriks, K. V. and S. Liebens. (2019). “A robot math tutor that gives feedback”. In: *International Conference on Social Robotics*. Springer. 601–610.
- Hirschmanner, M., S. Gross, S. Zafari, B. Krenn, F. Neubarth, and M. Vincze. (2021). “Investigating Transparency Methods in a Robot Word-Learning System and Their Effects on Human Teaching Behaviors”. In: *2021 30th IEEE International Conference on Robot Human Interactive Communication (RO-MAN)*. 175–182. DOI: [10.1109/RO-MAN50785.2021.9515518](https://doi.org/10.1109/RO-MAN50785.2021.9515518).
- Ho, J. and S. Ermon. (2016). “Generative Adversarial Imitation Learning”. In: *Neural Information Processing Systems*.
- Hochreiter, S. and J. Schmidhuber. (1997). “Long short-term memory”. *Neural computation*. 9(8): 1735–1780.

- Hoffman, G. (2019). “Evaluating fluency in human–robot collaboration”. *IEEE Transactions on Human-Machine Systems*. 49(3): 209–218.
- Hoffman, G. and C. Breazeal. (2004). “Collaboration in human-robot teams”. In: *AIAA 1st intelligent systems technical conference*. 6434.
- Hoffman, R. R., S. T. Mueller, G. Klein, and J. Litman. (2018). “Metrics for explainable AI: Challenges and prospects”. *arXiv preprint arXiv:1812.04608*.
- Hoffman, R. R., S. T. Mueller, G. Klein, and J. Litman. (2023). “Measures for explainable AI: Explanation goodness, user satisfaction, mental models, curiosity, trust, and human-AI performance”. *Frontiers in Computer Science*. 5: 1096257.
- Holmberg, L. (2022). “Towards Benchmarking Explainable Artificial Intelligence Methods”. *arXiv preprint arXiv:2208.12120*.
- Hoque, R., A. Balakrishna, E. Novoseller, A. Wilcox, D. S. Brown, and K. Goldberg. (2021). “ThriftyDagger: Budget-aware novelty and risk gating for interactive imitation learning”. *arXiv preprint arXiv:2109.08273*.
- Hoque, R., L. Y. Chen, S. Sharma, K. Dharmarajan, B. Thananjeyan, P. Abbeel, and K. Goldberg. (2023). “Fleet-dagger: Interactive robot fleet learning with scalable human supervision”. In: *Conference on Robot Learning*. PMLR. 368–380.
- Hristov, Y., A. Lascarides, and S. Ramamoorthy. (2018). “Interpretable latent spaces for learning from demonstration”. In: *Conference on Robot Learning*. PMLR. 957–968.
- Hu, H., Z. Zhang, K. Nakamura, A. Bajcsy, and J. F. Fisac. (2023). “Deception Game: Closing the Safety-Learning Loop in Interactive Robot Autonomy”. In: *7th Annual Conference on Robot Learning*.
- Hu, R., Y. Guo, H. Li, Q. Pei, and Y. Gong. (2020). “Personalized federated learning with differential privacy”. *IEEE Internet of Things Journal*. 7(10): 9530–9539.
- Huang, S. H., K. Bhatia, P. Abbeel, and A. D. Dragan. (2018a). “Establishing appropriate trust via critical states”. In: *2018 IEEE/RSJ international conference on intelligent robots and systems (IROS)*. IEEE. 3929–3936.

- Huang, S. H., D. Held, P. Abbeel, and A. D. Dragan. (2019). “Enabling robots to communicate their objectives”. *Autonomous Robots*. 43: 309–326.
- Huang, Y., P. Peng, Y. Zhao, Y. Zhai, H. Xu, and Y. Tian. (2023). “Simoun: Synergizing Interactive Motion-appearance Understanding for Vision-based Reinforcement Learning”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 176–185.
- Huang, Y., L. Rozo, J. Silvério, and D. G. Caldwell. (2018b). “Kernelized Movement Primitives”. en. URL: <http://arxiv.org/abs/1708.08638>.
- Hubers, A., E. Andrulis, W. D. Smart, L. Scott, T. Stirrat, D. Tran, R. Zhang, R. Sowell, and C. Grimm. (2015). “Video manipulation techniques for the protection of privacy in remote presence systems”. In: *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts*. 59–60.
- Hulsén, T. (2023). “Explainable Artificial Intelligence (XAI): Concepts and Challenges in Healthcare”. en. *AI*. 4(3): 652–666. ISSN: 2673-2688. DOI: [10.3390/ai4030034](https://doi.org/10.3390/ai4030034). URL: <https://www.mdpi.com/2673-2688/4/3/34>.
- Hundt, A., W. Agnew, V. Zeng, S. Kacianka, and M. Gombolay. (2022a). “Robots Enact Malignant Stereotypes”. In: *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency. FAccT ’22*. Seoul, Republic of Korea: Association for Computing Machinery. 743–756. ISBN: 9781450393522. DOI: [10.1145/3531146.3533138](https://doi.org/10.1145/3531146.3533138). URL: <https://doi.org/10.1145/3531146.3533138>.
- Hundt, A., W. Agnew, V. Zeng, S. Kacianka, and M. Gombolay. (2022b). “Robots enact malignant stereotypes”. In: *2022 ACM Conference on Fairness, Accountability, and Transparency*. 743–756.
- Hussain, N., E. Erzin, T. M. Sezgin, and Y. Yemez. (2019). “Speech driven backchannel generation using deep Q-network for enhancing engagement in human-robot interaction”. *arXiv preprint arXiv:1908.01618*.
- Hussein, A., M. M. Gaber, E. Elyan, and C. Jayne. (2017). “Imitation Learning: A Survey of Learning Methods”. *ACM Computing Surveys*. 50(2): 21:1–21:35. ISSN: 0360-0300. DOI: [10.1145/3054912](https://doi.org/10.1145/3054912). URL: <https://doi.org/10.1145/3054912>.

- Ijspeert, A. J., J. Nakanishi, H. Hoffmann, P. Pastor, and S. Schaal. (2013). “Dynamical Movement Primitives: Learning Attractor Models for Motor Behaviors”. en. *Neural Computation*. 25(2): 328–373. ISSN: 0899-7667, 1530-888X. DOI: [10.1162/NECO_a_00393](https://doi.org/10.1162/NECO_a_00393). URL: <https://direct.mit.edu/neco/article/25/2/328-373/7850>.
- Ince, G., R. Yorganci, A. Ozkul, T. B. Duman, and H. Köse. (2021). “An audiovisual interface-based drumming system for multimodal human–robot interaction”. *Journal on Multimodal User Interfaces*. 15: 413–428.
- Intelligence, A. “High-Level Expert Group on (AI HLEG),(2018). Ethics Guidelines for Trustworthy AI”.
- International Federation of Robotics, I. (2022). “World Robotics Report: “All-Time High” with Half a Million Robots Installed in one Year — ifr.org”. <https://ifr.org/ifr-press-releases/news/wr-report-all-time-high-with-half-a-million-robots-installed>.
- Iosifidis, V. and E. Ntoutsis. (2018). “Dealing with bias via data augmentation in supervised learning scenarios”. *Jo Bates Paul D. Clough Robert Jäschke*. 24(11).
- Irfan, B., A. Ramachandran, S. Spaulding, D. F. Glas, I. Leite, and K. L. Koay. (2019). “Personalization in long-term human-robot interaction”. In: *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 685–686.
- Irfan, B., A. Ramachandran, S. Spaulding, S. Kalkan, G. I. Parisi, and H. Gunes. (2021). “Lifelong learning and personalization in long-term human-robot interaction (leap-hri)”. In: *Companion of the 2021 ACM/IEEE international conference on human-robot interaction*. 724–727.
- Irfan, B., A. Ramachandran, M. Staffa, and H. Gunes. (2023). “Lifelong Learning and Personalization in Long-Term Human-Robot Interaction (LEAP-HRI): Adaptivity for All”. In: *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction. HRI '23*. Stockholm, Sweden: Association for Computing Machinery. 929–931. ISBN: 9781450399708. DOI: [10.1145/3568294.3579956](https://doi.org/10.1145/3568294.3579956). URL: <https://doi.org/10.1145/3568294.3579956>.

- Ishiguro Hiroshi, T. K. (2017). *Human-Robot Interaction in Social Robotics*. Boca Raton: CRC Press. ISBN: 978-1-315-21652-2. DOI: [10.1201/b13004](https://doi.org/10.1201/b13004).
- Iucci, A., A. Hata, A. Terra, R. Inam, and I. Leite. (2021a). “Explainable Reinforcement Learning for Human-Robot Collaboration”. In: *2021 20th International Conference on Advanced Robotics (ICAR)*. 927–934. DOI: [10.1109/ICAR53236.2021.9659472](https://doi.org/10.1109/ICAR53236.2021.9659472).
- Iucci, A., A. Hata, A. Terra, R. Inam, and I. Leite. (2021b). “Explainable Reinforcement Learning for Human-Robot Collaboration”. In: *2021 20th International Conference on Advanced Robotics (ICAR)*. 927–934. DOI: [10.1109/ICAR53236.2021.9659472](https://doi.org/10.1109/ICAR53236.2021.9659472).
- James, S., Z. Ma, D. R. Arrojo, and A. J. Davison. (2020). “Rlbench: The robot learning benchmark & learning environment”. *IEEE Robotics and Automation Letters*. 5(2): 3019–3026.
- Jang, E., A. Irpan, M. Khansari, D. Kappler, F. Ebert, C. Lynch, S. Levine, and C. Finn. (2022). “Bc-z: Zero-shot task generalization with robotic imitation learning”. In: *Conference on Robot Learning*. PMLR. 991–1002.
- Jaynes, E. T. (1957). “Information Theory and Statistical Mechanics”. *Phys. Rev.* 106(4): 620–630. DOI: [10.1103/PhysRev.106.620](https://doi.org/10.1103/PhysRev.106.620). URL: http://prola.aps.org/abstract/PR/v106/i4/p620_1.
- Ji, Z., N. Lee, R. Frieske, T. Yu, D. Su, Y. Xu, E. Ishii, Y. J. Bang, A. Madotto, and P. Fung. (2023). “Survey of hallucination in natural language generation”. *ACM Computing Surveys*. 55(12): 1–38.
- Jiang, P., S. Song, and G. Huang. (2021). “Exploration with task information for meta reinforcement learning”. *IEEE Transactions on Neural Networks and Learning Systems*.
- Jiang, S. and R. C. Arkin. (2015). “Mixed-initiative human-robot interaction: definition, taxonomy, and survey”. In: *2015 IEEE International conference on systems, man, and cybernetics*. IEEE. 954–961.
- Johnson, M. and Y. Demiris. (2004). “Abstraction in recognition to solve the correspondence problem for robot imitation”.
- Jones, H. and S. Rock. (2002). “Dialogue-based human-robot interaction for space construction teams”. In: *Proceedings, IEEE Aerospace Conference*. Vol. 7. IEEE. 7–7.

- Joshi, A., S. Kale, S. Chandel, and D. K. Pal. (2015). “Likert Scale: Explored and Explained”. *Current Journal of Applied Science and Technology*. 7(4): 396–403. DOI: [10.9734/BJAST/2015/14975](https://doi.org/10.9734/BJAST/2015/14975). URL: <https://journalcjast.com/index.php/CJAST/article/view/381>.
- Juan, R., J. Huang, R. Gomez, K. Nakamura, Q. Sha, B. He, and G. Li. (2021). “Shaping progressive net of reinforcement learning for policy transfer with human evaluative feedback”. In: *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 1281–1288.
- Jun Jeon, H., D. Losey, and D. Sadigh. (2020). “Shared Autonomy with Learned Latent Actions”. en. In: *Robotics: Science and Systems XVI*. Robotics: Science and Systems Foundation. ISBN: 978-0-9923747-6-1. DOI: [10.15607/RSS.2020.XVI.011](https://doi.org/10.15607/RSS.2020.XVI.011). URL: <http://www.roboticsproceedings.org/rss16/p011.pdf>.
- Jung, Y., S.-W. Seo, and S.-W. Kim. (2020). “Curb Detection and Tracking in Low-Resolution 3D Point Clouds Based on Optimization Framework”. *IEEE Transactions on Intelligent Transportation Systems*. 21(9): 3893–3908. ISSN: 1558-0016. DOI: [10.1109/TITS.2019.2938498](https://doi.org/10.1109/TITS.2019.2938498). URL: <https://ieeexplore.ieee.org/document/8826330>.
- Juozapaitis, Z., A. Koul, A. Fern, M. Erwig, and F. Doshi-Velez. (2019). “Explainable reinforcement learning via reward decomposition”. In: *IJCAI/ECAI Workshop on explainable artificial intelligence*.
- Kallinen, K. (2017). “The Effects of Transparency and Task Type on Trust, Stress, Quality of Work, and Co-Worker Preference During Human-Autonomous System Collaborative Work”. In: *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction. HRI '17*. Vienna, Austria: Association for Computing Machinery. 153–154. ISBN: 9781450348850. DOI: [10.1145/3029798.3038386](https://doi.org/10.1145/3029798.3038386). URL: <https://doi.org/10.1145/3029798.3038386>.
- Kamishima, T., S. Akaho, H. Asoh, and J. Sakuma. (2012). “Fairness-aware classifier with prejudice remover regularizer”. In: *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2012, Bristol, UK, September 24–28, 2012. Proceedings, Part II 23*. Springer. 35–50.

- Kan, P., R. Huq, J. Hoey, R. Goetschalckx, and A. Mihailidis. (2011). “The development of an adaptive upper-limb stroke rehabilitation robotic system”. *Journal of NeuroEngineering and Rehabilitation*. 8(1): 33. ISSN: 1743-0003. DOI: [10.1186 / 1743- 0003- 8- 33](https://doi.org/10.1186/1743-0003-8-33). URL: <https://doi.org/10.1186/1743-0003-8-33> (accessed on 11/03/2023).
- Kanda, T., R. Sato, N. Saiwaki, and H. Ishiguro. (2007). “A two-month field trial in an elementary school for long-term human–robot interaction”. *IEEE Transactions on robotics*. 23(5): 962–971.
- Kang, I., P. Kunapuli, and A. J. Young. (2020). “Real-Time Neural Network-Based Gait Phase Estimation Using a Robotic Hip Exoskeleton”. en. *IEEE Transactions on Medical Robotics and Bionics*. 2(1): 28–37. ISSN: 2576-3202. DOI: [10.1109/TMRB.2019.2961749](https://doi.org/10.1109/TMRB.2019.2961749). URL: <https://ieeexplore.ieee.org/document/8941004/>.
- Kantaros, Y., M. Malencia, V. Kumar, and G. J. Pappas. (2020). “Reactive temporal logic planning for multiple robots in unknown environments”. In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 11479–11485.
- Kantaros, Y. and G. J. Pappas. (2019). “Optimal temporal logic planning for multi-robot systems in uncertain semantic maps”. In: *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 4127–4132.
- Kaptein, F., J. Broekens, K. Hindriks, and M. Neerinx. (2019). “Evaluating Cognitive and Affective Intelligent Agent Explanations in a Long-Term Health-Support Application for Children with Type 1 Diabetes”. In: *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. 1–7. DOI: [10.1109/ACII.2019.8925526](https://doi.org/10.1109/ACII.2019.8925526).
- Karamcheti, S., M. Srivastava, P. Liang, and D. Sadigh. (2022). “Lila: Language-informed latent actions”. In: *Conference on Robot Learning*. PMLR. 1379–1390.
- Karli, U. B. and T. Fitzgerald. (2023). “Resolving Ambiguities in LLM-enabled Human-Robot Collaboration”. In: *2nd Workshop on Language and Robot Learning: Language as Grounding*.
- Karlsson, J., F. S. Barbosa, and J. Tumova. (2020). “Sampling-based motion planning with temporal logic missions and spatial preferences”. *IFAC-PapersOnLine*. 53(2): 15537–15543.

- Karlsson, J., S. van Waveren, C. Pek, I. Torre, I. Leite, and J. Tumova. (2021). “Encoding human driving styles in motion planning for autonomous vehicles”. In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 1050–1056.
- Karpathy, A. and L. Fei-Fei. (2017). “Deep Visual-Semantic Alignments for Generating Image Descriptions”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 39(4): 664–676. DOI: [10.1109/TPAMI.2016.2598339](https://doi.org/10.1109/TPAMI.2016.2598339).
- Karpathy, A. and M. van de Panne. (2012). “Curriculum Learning for Motor Skills”. In: *Advances in Artificial Intelligence*. Ed. by L. Kosseim and D. Inkpen. Berlin, Heidelberg: Springer Berlin Heidelberg. 325–330. ISBN: 978-3-642-30353-1.
- Katagami, D. and S. Yamada. (2001). “Real robot learning with human teaching”. In: *4th Japan-Australia Joint Workshop on Intelligent and Evolutionary Systems*. 263–270.
- Katsura, S., Y. Matsumoto, and K. Ohnishi. (2010). “Shadow robot for teaching motion”. *Robotics and Autonomous Systems*. 58(7): 840–846. ISSN: 0921-8890. DOI: <https://doi.org/10.1016/j.robot.2010.03.004>. URL: <https://www.sciencedirect.com/science/article/pii/S092188901000062X>.
- Katt, S., F. A. Oliehoek, and C. Amato. (2017). “Learning in POMDPs with Monte Carlo Tree Search”. In: *Proceedings of the 34th International Conference on Machine Learning - Volume 70. ICML’17*. Sydney, NSW, Australia: JMLR.org. 1819–1827.
- Kaushik, R., T. Anne, and J.-B. Mouret. (2020). “Fast online adaptation in robotics through meta-learning embeddings of simulated priors”. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 5269–5276.
- Keizer, S., M. Ellen Foster, Z. Wang, and O. Lemon. (2014). “Machine Learning for Social Multiparty Human–Robot Interaction”. *ACM Trans. Interact. Intell. Syst.* 4(3). ISSN: 2160-6455. DOI: [10.1145/2600021](https://doi.org/10.1145/2600021). URL: <https://doi.org/10.1145/2600021>.

- Keller, B., M. Draelos, K. Zhou, R. Qian, A. N. Kuo, G. Konidaris, K. Hauser, and J. A. Izatt. (2020). "Optical Coherence Tomography-Guided Robotic Ophthalmic Microsurgery via Reinforcement Learning from Demonstration". en. *IEEE TRANSACTIONS ON ROBOTICS*. 36(4).
- Kelly, M., C. Sidrane, K. Driggs-Campbell, and M. J. Kochenderfer. (2019). "Hg-dagger: Interactive imitation learning with human experts". In: *2019 International Conference on Robotics and Automation (ICRA)*. IEEE. 8077–8083.
- Kenesei, Z., K. Ásványi, L. Kökény, M. Jászberényi, M. Miskolczi, T. Gyulavári, and J. Syahrivar. (2022). "Trust and perceived risk: How different manifestations affect the adoption of autonomous vehicles". *Transportation Research Part A: Policy and Practice*. 164: 379–393. ISSN: 0965-8564. DOI: <https://doi.org/10.1016/j.tra.2022.08.022>. URL: <https://www.sciencedirect.com/science/article/pii/S0965856422002269>.
- Kerzel, M., J. Ambsdorf, D. Becker, W. Lu, E. Strahl, J. Spisak, C. Gäde, T. Weber, and S. Wermter. (2022). "What's on Your Mind, NICO? XHRI: A Framework for eXplainable Human-Robot Interaction". *KI-Künstliche Intelligenz*. 36(3-4): 237–254.
- Khavas, Z. R., S. R. Ahmadzadeh, and P. Robinette. (2020). "Modeling Trust in Human-Robot Interaction: A Survey". en. In: *Social Robotics*. Ed. by A. R. Wagner, D. Feil-Seifer, K. S. Haring, S. Rossi, T. Williams, H. He, and S. Sam Ge. Vol. 12483. Cham: Springer International Publishing. 529–541. ISBN: 978-3-030-62055-4 978-3-030-62056-1. DOI: [10.1007/978-3-030-62056-1_44](https://doi.org/10.1007/978-3-030-62056-1_44). URL: http://link.springer.com/10.1007/978-3-030-62056-1_44.
- Kidd, C. D. and C. Breazeal. (2008). "Robots at home: Understanding long-term human-robot interaction". In: *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 3230–3235.
- Kidd, C. D. (2007). "Engagement in long-term human-robot interaction". *PhD thesis*. Citeseer.

- Kim, J. W., C. He, M. Urias, P. Gehlbach, G. D. Hager, I. Iordachita, and M. Kobilarov. (2020a). “Autonomously Navigating a Surgical Tool Inside the Eye by Learning from Demonstration”. In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. 7351–7357. DOI: [10.1109/ICRA40945.2020.9196537](https://doi.org/10.1109/ICRA40945.2020.9196537).
- Kim, J. and J. Canny. (2017). “Interpretable Learning for Self-Driving Cars by Visualizing Causal Attention”. en. In: *2017 IEEE International Conference on Computer Vision (ICCV)*. Venice: IEEE. 2961–2969. ISBN: 978-1-5386-1032-9. DOI: [10.1109/ICCV.2017.320](https://doi.org/10.1109/ICCV.2017.320). URL: <http://ieeexplore.ieee.org/document/8237582/>.
- Kim, J., S. Moon, A. Rohrbach, T. Darrell, and J. Canny. (2020b). “Advisable Learning for Self-Driving Vehicles by Internalizing Observation-to-Action Rules”. en. In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA: IEEE. 9658–9667. ISBN: 978-1-72817-168-5. DOI: [10.1109/CVPR42600.2020.00968](https://doi.org/10.1109/CVPR42600.2020.00968). URL: <https://ieeexplore.ieee.org/document/9157238/>.
- Kim, J., A. Rohrbach, T. Darrell, J. Canny, and Z. Akata. (2018). “Textual explanations for self-driving vehicles”. In: *Proceedings of the European conference on computer vision (ECCV)*. 563–578.
- Kim, S. K., E. A. Kirchner, L. Schloßmüller, and F. Kirchner. (2020c). “Errors in human-robot interactions and their effects on robot learning”. *Frontiers in Robotics and AI*. 7: 558531.
- Kim, S. K., E. A. Kirchner, A. Stefes, and F. Kirchner. (2017). “Intrinsic interactive reinforcement learning—Using error-related potentials for real world human-robot interaction”. *Scientific reports*. 7(1): 1–16.
- Kim, T. and P. Hinds. (2006). “Who Should I Blame? Effects of Autonomy and Transparency on Attributions in Human-Robot Interaction”. In: *ROMAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication*. 80–85. DOI: [10.1109/ROMAN.2006.314398](https://doi.org/10.1109/ROMAN.2006.314398).
- Kivrak, H., F. Cakmak, H. Kose, and S. Yavuz. (2021). “Social navigation framework for assistive robots in human inhabited unknown environments”. *Engineering Science and Technology, an International Journal*. 24(2): 284–298.
- Knepper, R. A., S. Tellex, A. Li, N. Roy, and D. Rus. (2015). “Recovering from failure by asking for help”. *Autonomous Robots*. 39(3): 347–362.

- Knox, W. B., S. Hatgis-Kessell, S. Booth, S. Niekum, P. Stone, and A. Allievi. (2022). “Models of human preference for learning reward functions”. *arXiv preprint arXiv:2206.02231*.
- Knox, W. B. and P. Stone. (2011). “Augmenting reinforcement learning with human feedback”. In: *ICML 2011 Workshop on New Developments in Imitation Learning (July 2011)*. Vol. 855. 3.
- Kober, J., J. A. Bagnell, and J. Peters. (2013). “Reinforcement learning in robotics: A survey”. *The International Journal of Robotics Research*. 32(11): 1238–1274. DOI: [10.1177/0278364913495721](https://doi.org/10.1177/0278364913495721). eprint: <https://doi.org/10.1177/0278364913495721>. URL: <https://doi.org/10.1177/0278364913495721>.
- Koene, A., S. Endo, A. Remazeilles, M. Prada, and A. M. Wing. (2014). “Experimental testing of the coglaboration prototype system for fluent human-robot object handover interactions”. In: *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*. IEEE. 249–254.
- Koh, W. Q., S. A. Felding, K. B. Budak, E. Toomey, and D. Casey. (2021). “Barriers and facilitators to the implementation of social robots for older adults and people with dementia: a scoping review”. *eng. BMC geriatrics*. 21(1): 351. ISSN: 1471-2318. DOI: [10.1186/s12877-021-02277-9](https://doi.org/10.1186/s12877-021-02277-9).
- Kolve, E., R. Mottaghi, W. Han, E. VanderBilt, L. Weihs, A. Herrasti, M. Deitke, K. Ehsani, D. Gordon, Y. Zhu, *et al.* (2017). “Ai2-thor: An interactive 3d environment for visual ai”. *arXiv preprint arXiv:1712.05474*.
- Konan, S. G., E. Seraj, and M. Gombolay. (2022). “Iterated Reasoning with Mutual Information in Cooperative and Byzantine Decentralized Teaming”. In: *International Conference on Learning Representations*.
- Könighofer, B., R. Bloem, R. Ehlers, and C. Pek. (2022). “Correct-by-Construction Runtime Enforcement in AI–A Survey”. In: *Principles of Systems Design: Essays Dedicated to Thomas A. Henzinger on the Occasion of His 60th Birthday*. Springer. 650–663.

- Koo, J., J. Kwac, W. Ju, M. Steinert, L. Leifer, and C. Nass. (2015). “Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance”. en. *International Journal on Interactive Design and Manufacturing (IJIDeM)*. 9(4): 269–275. ISSN: 1955-2505. DOI: [10.1007/s12008-014-0227-2](https://doi.org/10.1007/s12008-014-0227-2). URL: <https://doi.org/10.1007/s12008-014-0227-2>.
- Koptev, M., N. Figueroa, and A. Billard. (2021). “Real-time self-collision avoidance in joint space for humanoid robots”. *IEEE Robotics and Automation Letters*. 6(2): 1240–1247.
- Kormushev, P., S. Calinon, and D. G. Caldwell. (2011). “Imitation Learning of Positional and Force Skills Demonstrated via Kinesthetic Teaching and Haptic Input”. *Advanced Robotics*. 25: 581–603. URL: <https://api.semanticscholar.org/CorpusID:17439462>.
- Korshunov, P. and T. Ebrahimi. (2013). “Using face morphing to protect privacy”. In: *2013 10th IEEE International Conference on Advanced Video and Signal Based Surveillance*. IEEE. 208–213.
- Kotowick, K. and J. Shah. (2018). “Effects of an Adaptive Modality Selection Algorithm for Navigation Systems”. In: *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*. 543–556.
- Koubaa, A. (2023). “Rosgpt: Next-generation human-robot interaction with chatgpt and ros”.
- Kress-Gazit, H., G. E. Fainekos, and G. J. Pappas. (2009). “Temporal-logic-based reactive mission and motion planning”. *IEEE transactions on robotics*. 25(6): 1370–1381.
- Kress-Gazit, H., M. Lahijanian, and V. Raman. (2018). “Synthesis for robots: Guarantees and feedback for robot behavior”. *Annual Review of Control, Robotics, and Autonomous Systems*. 1: 211–236.
- Krishna, R., D. Lee, L. Fei-Fei, and M. S. Bernstein. (2022). “Socially situated artificial intelligence enables learning from human interaction”. *Proceedings of the National Academy of Sciences*. 119(39). DOI: [10.1073/pnas.2115730119](https://doi.org/10.1073/pnas.2115730119).
- Kroemer, O. B., R. Detry, J. Piater, and J. Peters. (2010). “Combining active learning and reactive control for robot grasping”. *Robotics and Autonomous systems*. 58(9): 1105–1116.

- Kruijff-Korbayová, I., G. Athanasopoulos, A. Beck, P. Cosi, H. Cuayáhuitl, T. Dekens, V. Enescu, A. Hiolle, B. Kiefer, H. Sahli, *et al.* (2011). “An event-based conversational system for the nao robot”. In: *Proceedings of the Paralinguistic Information and its Integration in Spoken Dialogue Systems Workshop*. Springer. 125–132.
- Kulesza, T., M. Burnett, W.-K. Wong, and S. Stumpf. (2015). “Principles of explanatory debugging to personalize interactive machine learning”. In: *Proceedings of the 20th international conference on intelligent user interfaces*. 126–137.
- Kulić, D. and E. A. Croft. (2005). “Safe planning for human-robot interaction”. *Journal of Robotic Systems*. 22(7): 383–396.
- Kulić, D. and E. A. Croft. (2006). “Real-time safety for human–robot interaction”. *Robotics and Autonomous Systems*. 54(1): 1–12.
- Kulkarni, A., S. Sreedharan, S. Keren, T. Chakraborti, D. E. Smith, and S. Kambhampati. (2020). “Designing environments conducive to interpretable robot behavior”. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 10982–10989.
- Kullback, S. and R. A. Leibler. (1951). “On Information and Sufficiency”. *The Annals of Mathematical Statistics*. 22(1): 79–86. DOI: [10.1214/aoms/1177729694](https://doi.org/10.1214/aoms/1177729694). URL: <https://doi.org/10.1214/aoms/1177729694>.
- Kumar, V. and E. Todorov. (2015). “Mujoco haptix: A virtual reality system for hand manipulation”. In: *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*. IEEE. 657–663.
- Kupcsik, A., D. Hsu, and W. S. Lee. (2018). “Learning dynamic robot-to-human object handover from human feedback”. *Robotics Research: Volume 1*: 161–176.
- Kwon, M., E. Biyik, A. Talati, K. Bhasin, D. P. Losey, and D. Sadigh. (2020). “When Humans Aren’t Optimal: Robots that Collaborate with Risk-Aware Humans”. en. In: *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. Cambridge United Kingdom: ACM. 43–52. ISBN: 978-1-4503-6746-2. DOI: [10.1145/3319502.3374832](https://doi.org/10.1145/3319502.3374832). URL: <https://dl.acm.org/doi/10.1145/3319502.3374832>.

- Kwon, M., S. H. Huang, and A. D. Dragan. (2018). “Expressing Robot Incapability”. In: *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction. HRI '18*. Chicago, IL, USA: Association for Computing Machinery. 87–95. ISBN: 9781450349536. DOI: [10.1145/3171221.3171276](https://doi.org/10.1145/3171221.3171276). URL: <https://doi.org/10.1145/3171221.3171276>.
- Kyrarini, M., M. A. Haseeb, D. Ristić-Durrant, and A. Gräser. (2019). “Robot learning of industrial assembly task via human demonstrations”. en. *Autonomous Robots*. 43(1): 239–257. ISSN: 1573-7527. DOI: [10.1007/s10514-018-9725-6](https://doi.org/10.1007/s10514-018-9725-6). URL: <https://doi.org/10.1007/s10514-018-9725-6>.
- Kyrarini, M., F. Lygerakis, A. Rajavenkatanarayanan, C. Sevastopoulos, H. R. Nambiappan, K. K. Chaitanya, A. R. Babu, J. Mathew, and F. Makedon. (2021). “A survey of robots in healthcare”. *Technologies*. 9(1): 8.
- Lacevic, B., A. M. Zanchettin, and P. Rocco. (2022). “Safe Human-robot collaboration via collision checking and explicit representation of danger zones”. *IEEE Transactions on Automation Science and Engineering*. 20(2): 846–861.
- Lackey, S., D. Barber, L. Reinerman, N. I. Badler, and I. Hudson. (2011). “Defining next-generation multi-modal communication in human robot interaction”. In: *Proceedings of the human factors and ergonomics society annual meeting*. Vol. 55. No. 1. SAGE Publications Sage CA: Los Angeles, CA. 461–464.
- Lage, I., D. Lifschitz, F. Doshi-Velez, and O. Amir. (2019). “Exploring computational user models for agent policy summarization”. In: *IJCAI: proceedings of the conference*. Vol. 28. NIH Public Access. 1401.
- Lala, D., P. Milhorat, K. Inoue, M. Ishida, K. Takanashi, and T. Kawahara. (2017). “Attentive listening system with backchanneling, response generation and flexible turn-taking”. In: *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*. 127–136.
- Lasota, P. A., T. Fong, J. A. Shah, et al. (2017). “A survey of methods for safe human-robot interaction”. *Foundations and Trends® in Robotics*. 5(4): 261–349.

- Lasota, P. A. and J. A. Shah. (2015). “Analyzing the effects of human-aware motion planning on close-proximity human–robot collaboration”. *Human factors*. 57(1): 21–33.
- LaValle, S. M. (2006). *Planning algorithms*. Cambridge university press.
- Lee, J. D. and K. A. See. (2004). “Trust in automation: designing for appropriate reliance”. eng. *Human Factors*. 46(1): 50–80. ISSN: 0018-7208. DOI: [10.1518/hfes.46.1.50_30392](https://doi.org/10.1518/hfes.46.1.50_30392).
- Lee, J., M. Laskey, R. Fox, and K. Goldberg. (2018). “Constraint Estimation and Derivative-Free Recovery for Robot Learning from Demonstrations”. In: *2018 IEEE 14th International Conference on Automation Science and Engineering (CASE)*. IEEE. 270–277.
- Lee, K., L. Smith, and P. Abbeel. (2021). “PEBBLE: Feedback-Efficient Interactive Reinforcement Learning via Relabeling Experience and Unsupervised Pre-training”. In: *International Conference on Machine Learning*.
- Lee, K. M., A. Krishna, Z. Zaidi, R. Paleja, L. Chen, E. Hedlund-Botti, M. Schrum, and M. Gombolay. (2023). “The Effect of Robot Skill Level and Communication in Rapid, Proximate Human-Robot Collaboration”. In: *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction. HRI '23*. Stockholm, Sweden: Association for Computing Machinery. 261–270. ISBN: 9781450399647. DOI: [10.1145/3568162.3577002](https://doi.org/10.1145/3568162.3577002). URL: <https://doi.org/10.1145/3568162.3577002>.
- Lee, M. H., D. P. Siewiorek, A. Smailagic, A. Bernardino, and S. B. Badia. (2020). “Towards personalized interaction and corrective feedback of a socially assistive robot for post-stroke rehabilitation therapy”. In: *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 1366–1373.
- Lee, M., J. Jeon, and H. Lee. (2022). “Explainable AI for domain experts: a post Hoc analysis of deep learning for defect classification of TFT–LCD panels”. *Journal of Intelligent Manufacturing*. 33(6): 1747–1759. DOI: [10.1007/s10845-021-01758-3](https://doi.org/10.1007/s10845-021-01758-3). URL: https://ideas.repec.org/a/spr/joinma/v33y2022i6d10.1007_s10845-021-01758-3.html.

- Leite, I., A. Pereira, G. Castellano, S. Mascarenhas, C. Martinho, and A. Paiva. (2012). “Modelling Empathy in Social Robotic Companions”. en. In: *Advances in User Modeling*. Ed. by L. Ardissono and T. Kuflik. *Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer. 135–147. ISBN: 978-3-642-28509-7. DOI: [10.1007/978-3-642-28509-7_14](https://doi.org/10.1007/978-3-642-28509-7_14).
- Leslie, D. (2020). “Understanding bias in facial recognition technologies”. *arXiv preprint arXiv:2010.07023*.
- Lesort, T., V. Lomonaco, A. Stoian, D. Maltoni, D. Filliat, and N. Díaz-Rodríguez. (2020). “Continual learning for robotics: Definition, framework, learning strategies, opportunities and challenges”. *Information Fusion*. 58: 52–68. ISSN: 1566-2535. DOI: <https://doi.org/10.1016/j.inffus.2019.12.004>. URL: <https://www.sciencedirect.com/science/article/pii/S1566253519307377>.
- Levine, S., Z. Popovic, and V. Koltun. (2011). “Nonlinear Inverse Reinforcement Learning with Gaussian Processes”. In: *Advances in Neural Information Processing Systems*. Vol. 24. Curran Associates, Inc. URL: https://papers.nips.cc/paper_files/paper/2011/hash/c51ce410c124a10e0db5e4b97fc2af39-Abstract.html.
- Lewandowsky, S., M. Mundy, and G. P. Tan. (2000). “The dynamics of trust: comparing humans to automation”. eng. *Journal of Experimental Psychology. Applied*. 6(2): 104–123. ISSN: 1076-898X. DOI: [10.1037/1076-898x.6.2.104](https://doi.org/10.1037/1076-898x.6.2.104).
- Lewis, M., H. Li, and K. Sycara. (2021). “Chapter 14 - Deep learning, transparency, and trust in human robot teamwork”. In: *Trust in Human-Robot Interaction*. Ed. by C. S. Nam and J. B. Lyons. Academic Press. 321–352. ISBN: 978-0-12-819472-0. DOI: [10.1016/B978-0-12-819472-0.00014-9](https://doi.org/10.1016/B978-0-12-819472-0.00014-9).
- Lewis, M., K. Sycara, and P. Walker. (2018). “The Role of Trust in Human-Robot Interaction”. en. In: *Foundations of Trusted Autonomy*. Ed. by H. A. Abbass, J. Scholz, and D. J. Reid. *Studies in Systems, Decision and Control*. Cham: Springer International Publishing. 135–159. ISBN: 978-3-319-64816-3. DOI: [10.1007/978-3-319-64816-3_8](https://doi.org/10.1007/978-3-319-64816-3_8). URL: https://doi.org/10.1007/978-3-319-64816-3_8.

- Leyzberg, D., A. Ramachandran, and B. Scassellati. (2018). “The effect of personalization in longer-term robot tutoring”. *ACM Transactions on Human-Robot Interaction (THRI)*. 7(3): 1–19.
- Li, C., F. Xia, R. Martin-Martin, and S. Savarese. (2020a). “Hrl4in: Hierarchical reinforcement learning for interactive navigation with mobile manipulators”. In: *Conference on Robot Learning*. PMLR. 603–616.
- Li, G., Z. Jin, M. Volpp, F. Otto, R. Lioutikov, and G. Neumann. (2023). “ProDMP: A Unified Perspective on Dynamic and Probabilistic Movement Primitives”. *IEEE Robotics and Automation Letters*. 8(4): 2325–2332. DOI: [10.1109/LRA.2023.3248443](https://doi.org/10.1109/LRA.2023.3248443).
- Li, S., N. Li, A. Girard, and I. Kolmanovsky. (2019a). “Decision making in dynamic and interactive environments based on cognitive hierarchy theory, Bayesian inference, and predictive control”. In: *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE. 2181–2187.
- Li, T., A. K. Sahu, A. Talwalkar, and V. Smith. (2020b). “Federated learning: Challenges, methods, and future directions”. *IEEE signal processing magazine*. 37(3): 50–60.
- Li, W., R. Hong, J. Shen, and Y. Lu. (2022). “Learning to Navigate in Interactive Environments with the Transformer-based Memory”.
- Li, X., Z. Serlin, G. Yang, and C. Belta. (2019b). “A formal methods approach to interpretable reinforcement learning for robotic planning”. *Science Robotics*. 4(37): eaay6276.
- Li, X., C.-I. Vasile, and C. Belta. (2017). “Reinforcement learning with temporal logic rewards”. In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 3834–3839.
- Li, Y., E. M. Shahrivar, and J. Liu. (2021a). “Safe linear temporal logic motion planning in dynamic environments”. In: *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 9818–9825.

- Li, Z., L. Shi, A. I. Cristea, and Y. Zhou. (2021b). “A Survey of Collaborative Reinforcement Learning: Interactive Methods and Design Patterns”. In: *Proceedings of the 2021 ACM Designing Interactive Systems Conference. DIS '21*. Virtual Event, USA: Association for Computing Machinery. 1579–1590. ISBN: 9781450384766. DOI: [10.1145/3461778.3462135](https://doi.org/10.1145/3461778.3462135). URL: <https://doi.org/10.1145/3461778.3462135>.
- Lidard, J., H. Pham, A. Bachman, B. Boateng, and A. Majumdar. (2024). “Risk-Calibrated Human-Robot Interaction via Set-Valued Intent Prediction”. *arXiv preprint arXiv:2403.15959*.
- Lima, G., N. Grgić-Hlača, and M. Cha. (2021). “Human perceptions on moral responsibility of AI: A case study in AI-assisted bail decision-making”. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–17.
- Lin, L.-H., Y. Cui, Y. Hao, F. Xia, and D. Sadigh. (2023a). “Gesture-informed robot assistance via foundation models”. In: *Conference on Robot Learning*. PMLR. 3061–3082.
- Lin, J., Z. Ma, R. Gomez, K. Nakamura, B. He, and G. Li. (2020). “A review on interactive reinforcement learning from human social feedback”. *IEEE Access*. 8: 120757–120765.
- Lin, Z., Y. Chen, and Z. Liu. (2023b). “Sketch RL: Interactive Sketch Generation for Long-Horizon Tasks via Vision-Based Skill Predictor”. *IEEE Robotics and Automation Letters*. 9(1): 867–874.
- Lin, Z. and J. S. Baras. (2020). “Metric interval temporal logic based reinforcement learning with runtime monitoring and self-correction”. In: *2020 American Control Conference (ACC)*. IEEE. 5400–5406.
- Linardatos, P., V. Papastefanopoulos, and S. Kotsiantis. (2021). “Explainable AI: A Review of Machine Learning Interpretability Methods”. *Entropy*. 23(1). ISSN: 1099-4300. URL: <https://www.mdpi.com/1099-4300/23/1/18>.
- Lindemann, L., G. J. Pappas, and D. V. Dimarogonas. (2020). “Control barrier functions for nonholonomic systems under risk signal temporal logic specifications”. In: *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE. 1422–1428.

- Lipton, Z. C. (2018). “The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery.” *Queue*. 16(3): 31–57.
- Liu, H., Y. Zhang, W. Si, X. Xie, Y. Zhu, and S.-C. Zhu. (2018a). “Interactive Robot Knowledge Patching Using Augmented Reality”. In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. 1947–1954. DOI: [10.1109/ICRA.2018.8462837](https://doi.org/10.1109/ICRA.2018.8462837).
- Liu, H., Y. Zhang, W. Si, X. Xie, Y. Zhu, and S.-C. Zhu. (2018b). “Interactive robot knowledge patching using augmented reality”. In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 1947–1954.
- Liu, H., Y. Zhu, and S.-C. Zhu. (2021a). “Patching interpretable And-Or-Graph knowledge representation using augmented reality”. *Applied AI Letters*. 2(4): e43.
- Liu, H., L. Lee, K. Lee, and P. Abbeel. (2022). “Instruction-following agents with jointly pre-trained vision-language models”.
- Liu, H. and L. Wang. (2018). “Gesture recognition for human-robot collaboration: A review”. *International Journal of Industrial Ergonomics*. 68: 355–367.
- Liu, H., A. Chen, Y. Zhu, A. Swaminathan, A. Kolobov, and C.-A. Cheng. (2023a). “Interactive Robot Learning from Verbal Correction”. *arXiv preprint arXiv:2310.17555*.
- Liu, H., S. Dass, R. Martin-Martin, and Y. Zhu. (2023b). “Model-Based Runtime Monitoring with Interactive Imitation Learning”. *arXiv preprint arXiv:2310.17552*.
- Liu, X., T. Yoneda, C. Wang, M. Walter, and Y. Chen. (2023c). “Active policy improvement from multiple black-box oracles”. In: *International Conference on Machine Learning*. PMLR. 22320–22337.
- Liu, Y., S. Khandagale, C. White, and W. Neiswanger. (2021b). “Synthetic benchmarks for scientific research in explainable machine learning”. *arXiv preprint arXiv:2106.12543*.

- Liu, Z., Q. Liu, W. Xu, Z. Liu, Z. Zhou, and J. Chen. (2019). “Deep Learning-based Human Motion Prediction considering Context Awareness for Human-Robot Collaboration in Manufacturing”. *Procedia CIRP*. 11th CIRP Conference on Industrial Product-Service Systems 83(Jan.): 272–278. ISSN: 2212-8271. DOI: [10.1016/j.procir.2019.04.080](https://doi.org/10.1016/j.procir.2019.04.080). URL: <https://www.sciencedirect.com/science/article/pii/S2212827119306948>.
- Loizou, S. G. and V. Kumar. (2007). “Mixed initiative control of autonomous vehicles”. In: *Proceedings 2007 IEEE International Conference on Robotics and Automation*. IEEE. 1431–1436.
- Londoño, L., J. V. Hurtado, N. Hertz, P. Kellmeyer, S. Voenekey, and A. Valada. (2022). “Fairness and Bias in Robot Learning”. *arXiv preprint arXiv:2207.03444*.
- Long, Y., W. Wei, T. Huang, Y. Wang, and Q. Dou. (2023). “Human-in-the-loop embodied intelligence with interactive simulation environment for surgical robot learning”. *IEEE Robotics and Automation Letters*.
- Lu, J., J. Salvador, R. Mottaghi, and A. Kembhavi. (2022). “ASC me to Do Anything: Multi-task Training for Embodied AI”. *arXiv preprint arXiv:2202.06987*.
- Luckcuck, M., M. Farrell, L. A. Dennis, C. Dixon, and M. Fisher. (2019). “Formal specification and verification of autonomous robotic systems: A survey”. *ACM Computing Surveys (CSUR)*. 52(5): 1–41.
- Lundberg, S. M. and S.-I. Lee. (2017). “A unified approach to interpreting model predictions”. *Advances in neural information processing systems*. 30.
- Luo, J., O. Sushkov, R. Pevcevičute, W. Lian, C. Su, M. Vecerik, N. Ye, S. Schaal, and J. Scholz. (2021). “Robust Multi-Modal Policies for Industrial Assembly via Reinforcement Learning and Demonstrations: A Large-Scale Study”. en. URL: <http://arxiv.org/abs/2103.11512>.
- Lutkebohle, I., J. Peltason, L. Schillingmann, B. Wrede, S. Wachsmuth, C. Elbrechter, and R. Haschke. (2009). “The curious robot-structuring interactive robot learning”. In: *2009 IEEE International Conference on Robotics and Automation*. IEEE. 4156–4162.
- Lyons, J. B. (2013). “Being Transparent about Transparency”. In: URL: <https://api.semanticscholar.org/CorpusID:18712525>.

- Lyons, J. B. and C. K. Stokes. (2012). “Human-human reliance in the context of automation”. eng. *Human Factors*. 54(1): 112–121. ISSN: 0018-7208. DOI: [10.1177/0018720811427034](https://doi.org/10.1177/0018720811427034).
- MacGlashan, J., M. Babes-Vroman, M. desJardins, M. Littman, S. Muresan, S. Squire, S. Tellex, D. Arumugam, and L. Yang. (2015). “Grounding English Commands to Reward Functions”. In: *Proceedings of Robotics: Science and Systems*. Rome, Italy. DOI: [10.15607/RSS.2015.XI.018](https://doi.org/10.15607/RSS.2015.XI.018).
- MacGlashan, J., M. K. Ho, R. Loftin, B. Peng, G. Wang, D. L. Roberts, M. E. Taylor, and M. L. Littman. (2017). “Interactive learning from policy-dependent human feedback”. In: *International conference on machine learning*. PMLR. 2285–2294.
- MacMillan, J., E. E. Entin, and D. Serfaty. (2004). “Communication overhead: The hidden cost of team cognition.”
- Maderna, R., M. Pozzi, A. M. Zanchettin, P. Rocco, and D. Praticchizzo. (2022). “Flexible scheduling and tactile communication for human–robot collaboration”. *Robotics and Computer-Integrated Manufacturing*. 73: 102233.
- Madumal, P., T. Miller, L. Sonenberg, and F. Vetere. (2020). “Explainable reinforcement learning through a causal lens”. In: *Proceedings of the AAAI conference on artificial intelligence*. Vol. 34. No. 03. 2493–2500.
- Maeda, G., M. Ewerton, R. Lioutikov, H. Ben Amor, J. Peters, and G. Neumann. (2014). “Learning interaction for collaborative tasks with probabilistic movement primitives”. en. In: *2014 IEEE-RAS International Conference on Humanoid Robots*. Madrid, Spain: IEEE. 527–534. ISBN: 978-1-4799-7174-9. DOI: [10.1109/HUMANOIDS.2014.7041413](https://doi.org/10.1109/HUMANOIDS.2014.7041413). URL: <http://ieeexplore.ieee.org/document/7041413/>.
- Maeda, G., M. Ewerton, T. Osa, B. Busch, and J. Peters. (2017). “Active Incremental Learning of Robot Movement Primitives”. In: *Proceedings of the 1st Annual Conference on Robot Learning*. Ed. by S. Levine, V. Vanhoucke, and K. Goldberg. Vol. 78. *Proceedings of Machine Learning Research*. PMLR. 37–46. URL: <https://proceedings.mlr.press/v78/maeda17a.html>.

- Mahler, J., J. Liang, S. Niyaz, M. Laskey, R. Doan, X. Liu, J. A. Ojea, and K. Goldberg. (2017). “Dex-net 2.0: Deep learning to plan robust grasps with synthetic point clouds and analytic grasp metrics”. *arXiv preprint arXiv:1703.09312*.
- Mahler, J., M. Matl, X. Liu, A. Li, D. Gealy, and K. Goldberg. (2018). “Dex-net 3.0: Computing robust vacuum suction grasp targets in point clouds using a new analytic model and deep learning”. In: *2018 IEEE International Conference on robotics and automation (ICRA)*. IEEE. 5620–5627.
- Mahmud, S., S. Saisubramanian, and S. Zilberstein. (2022). “Reveale: Reward verification and learning using explanations”.
- Majumdar, A. and M. Pavone. (2020). “How should a robot assess risk? towards an axiomatic theory of risk in robotics”. In: *Robotics Research: The 18th International Symposium ISRR*. Springer. 75–84.
- Mandlekar, A., C. Garrett, D. Xu, and D. Fox. (2023). “Human-in-the-Loop Task and Motion Planning for Imitation Learning”. en. URL: <http://arxiv.org/abs/2310.16014>.
- Mandlekar, A., D. Xu, R. Martín-Martín, Y. Zhu, L. Fei-Fei, and S. Savarese. (2020). “Human-in-the-loop imitation learning using remote teleoperation”. *arXiv preprint arXiv:2012.06733*.
- Mandlekar, A., Y. Zhu, A. Garg, J. Booher, M. Spero, A. Tung, J. Gao, J. Emmons, A. Gupta, E. Orbay, *et al.* (2018a). “Roboturk: A crowdsourcing platform for robotic skill learning through imitation”. In: *Conference on Robot Learning*. PMLR. 879–893.
- Mandlekar, A., Y. Zhu, A. Garg, J. Booher, M. Spero, A. Tung, J. Gao, J. Emmons, A. Gupta, E. Orbay, S. Savarese, and L. Fei-Fei. (2018b). “ROBOTURK: A Crowdsourcing Platform for Robotic Skill Learning through Imitation”. In: *Proceedings of The 2nd Conference on Robot Learning*. Ed. by A. Billard, A. Dragan, J. Peters, and J. Morimoto. Vol. 87. *Proceedings of Machine Learning Research*. PMLR. 879–893. URL: <https://proceedings.mlr.press/v87/mandlekar18a.html>.
- Mankodiya, H., M. S. Obaidat, R. Gupta, and S. Tanwar. (2021). “XAI-AV: Explainable Artificial Intelligence for Trust Management in Autonomous Vehicles”. In: *2021 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI)*. 1–5. DOI: [10.1109/CCCI52664.2021.9583190](https://doi.org/10.1109/CCCI52664.2021.9583190).

- Maroto-Gómez, M., Á. Castro-González, J. C. Castillo, M. Malfaz, and M. Á. Salichs. (2023). “An adaptive decision-making system supported on user preference predictions for human–robot interactive communication”. *User Modeling and User-Adapted Interaction*. 33(2): 359–403.
- Marta, D., C. Pek, G. I. Melsión, J. Tumova, and I. Leite. (2021). “Human-feedback shield synthesis for perceived safety in deep reinforcement learning”. *IEEE Robotics and Automation Letters*. 7(1): 406–413.
- Martino, A., M. Iannelli, and C. Truong. (2023). “Knowledge injection to counter large language model (llm) hallucination”. In: *European Semantic Web Conference*. Springer. 182–185.
- Martins, G. S., L. Santos, and J. Dias. (2019). “User-Adaptive Interaction in Social Robots: A Survey Focusing on Non-physical Interaction”. *International Journal of Social Robotics*. 11(1): 185–205. ISSN: 1875-4805. DOI: [10.1007/s12369-018-0485-4](https://doi.org/10.1007/s12369-018-0485-4). URL: <https://doi.org/10.1007/s12369-018-0485-4>.
- Matarese, M., A. Sciutti, F. Rea, and S. Rossi. (2021). “Toward Robots’ Behavioral Transparency of Temporal Difference Reinforcement Learning With a Human Teacher”. *IEEE Transactions on Human-Machine Systems*. 51(6): 578–589. DOI: [10.1109/THMS.2021.3116119](https://doi.org/10.1109/THMS.2021.3116119).
- Matsumaru, T. (2006). “Mobile Robot with Preliminary-announcement and Display Function of Forthcoming Motion using Projection Equipment”. In: *ROMAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication*. 443–450. DOI: [10.1109/ROMAN.2006.314368](https://doi.org/10.1109/ROMAN.2006.314368).
- Matsumaru, T. (2007). “Mobile Robot with Preliminary-announcement and Indication Function of Forthcoming Operation using Flat-panel Display”. In: *Proceedings 2007 IEEE International Conference on Robotics and Automation*. 1774–1781. DOI: [10.1109/ROBOT.2007.363579](https://doi.org/10.1109/ROBOT.2007.363579).
- Mayer, R. E. and R. Moreno. (2003). “Nine ways to reduce cognitive load in multimedia learning”. *Educational psychologist*. 38(1): 43–52.
- McLellan, H. (1996). *Situated learning perspectives*. Educational Technology.

- McMahan, B., E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas. (2017). “Communication-efficient learning of deep networks from decentralized data”. In: *Artificial intelligence and statistics*. PMLR. 1273–1282.
- McNeill, D. (2019). *Gesture and thought*. University of Chicago press.
- Mees, O., L. Hermann, and W. Burgard. (2022a). “What matters in language conditioned robotic imitation learning over unstructured data”. *IEEE Robotics and Automation Letters*. 7(4): 11205–11212.
- Mees, O., L. Hermann, E. Rosete-Beas, and W. Burgard. (2022b). “Calvin: A benchmark for language-conditioned policy learning for long-horizon robot manipulation tasks”. *IEEE Robotics and Automation Letters*. 7(3): 7327–7334.
- Mehta, B., M. Diaz, F. Golemo, C. J. Pal, and L. Paull. (2020). “Active domain randomization”. In: *Conference on Robot Learning*. PMLR. 1162–1176.
- Mei, H., M. Bansal, and M. R. Walter. (2016). “Listen, Attend, and Walk: Neural Mapping of Navigational Instructions to Action Sequences”. In: *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*. AAAI’16. Phoenix, Arizona: AAAI Press. 2772–2778.
- Menda, K., K. Driggs-Campbell, and M. J. Kochenderfer. (2019). “Ensembledagger: A bayesian approach to safe imitation learning”. In: *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 5041–5048.
- Mendoza Barrenechea, D. (2023). “Curriculum Learning for Robot Manipulation using Deep Reinforcement Learning”. *MA thesis*. Imperial College London.
- Meng, Q., I. Tholley, and P. W. H. Chung. (2014). “Robots learn to dance through interaction with humans”. en. *Neural Computing and Applications*. 24(1): 117–124. ISSN: 1433-3058. DOI: [10.1007/s00521-013-1504-x](https://doi.org/10.1007/s00521-013-1504-x). URL: <https://doi.org/10.1007/s00521-013-1504-x>.
- Michie, D., M. Bain, and J. Hayes-Michie. (1990). “Cognitive models from subcognitive skills”. *IEE control engineering series*. 44: 71–99.
- Milliken, L. and G. A. Hollinger. (2017). “Modeling user expertise for choosing levels of shared autonomy”. In: *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2285–2291.

- Ministry of Justice, Japan. (2020). “Act on the Protection of Personal Information”. <https://www.japaneselawtranslation.go.jp/en/laws/view/4241/en>.
- Mintchev, S., M. Salerno, A. Cherpillod, S. Scaduto, and J. Paik. (2019). “A portable three-degrees-of-freedom force feedback origami robot for human–robot interactions”. *Nature Machine Intelligence*. 1(12): 584–593.
- Mirnig, N., B. Gonsior, S. Sosnowski, C. Landsiedel, D. Wollherr, A. Weiss, and M. Tscheligi. (2012). “Feedback guidelines for multi-modal human-robot interaction: How should a robot give feedback when asking for directions?” In: *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. IEEE. 533–538.
- Misra, D., J. Langford, and Y. Artzi. (2017). “Mapping Instructions and Visual Observations to Actions with Reinforcement Learning”. In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. Ed. by M. Palmer, R. Hwa, and S. Riedel. Copenhagen, Denmark: Association for Computational Linguistics. 1004–1015. DOI: [10.18653/v1/D17-1106](https://doi.org/10.18653/v1/D17-1106). URL: <https://aclanthology.org/D17-1106>.
- Mitrevski, A., P. G. Plöger, and G. Lakemeyer. (2020). “Representation and experience-based learning of explainable models for robot action execution”. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 5641–5647.
- Mitsunaga, N., C. Smith, T. Kanda, H. Ishiguro, and N. Hagita. (2008). “Adapting robot behavior for human–robot interaction”. *IEEE Transactions on Robotics*. 24(4): 911–916.
- Mittal, M., C. Yu, Q. Yu, J. Liu, N. Rudin, D. Hoeller, J. L. Yuan, R. Singh, Y. Guo, H. Mazhar, *et al.* (2023). “Orbit: A unified simulation framework for interactive robot learning environments”. *IEEE Robotics and Automation Letters*.
- Mittu, R., D. Sofge, A. Wagner, and W. F. Lawless. (2016). *Robust intelligence and trust in autonomous systems*. Springer.

- Mnih, V., K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis. (2015). “Human-level control through deep reinforcement learning”. *Nature*. 518(7540): 529–533. ISSN: 1476-4687. DOI: [10.1038/nature14236](https://doi.org/10.1038/nature14236). URL: <https://doi.org/10.1038/nature14236>.
- Mobahi, H. and S. Ansari. (2003). “Fuzzy perception, emotion and expression for interactive robots”. In: *SMC’03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme - System Security and Assurance (Cat. No.03CH37483)*. Vol. 4. 3918–3923 vol.4. DOI: [10.1109/ICSMC.2003.1244500](https://doi.org/10.1109/ICSMC.2003.1244500).
- Mohamed, S. and D. J. Rezende. (2015). “Variational Information Maximisation for Intrinsically Motivated Reinforcement Learning”. arXiv: [1509.08731 \[stat.ML\]](https://arxiv.org/abs/1509.08731).
- Mohseni-Kabir, A., S. Chernova, and C. Rich. (2014). “Collaborative Learning of Hierarchical Task Networks from Demonstration and Instruction”. In: *AAAI Fall Symposium on Artificial Intelligence and Human-Robot Interaction*.
- Mohseni-Kabir, A., C. Rich, S. Chernova, C. L. Sidner, and D. Miller. (2015). “Interactive Hierarchical Task Learning from a Single Demonstration”. In: *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 205–212.
- Moon, A., D. M. Troniak, B. Gleeson, M. K. Pan, M. Zheng, B. A. Blumer, K. MacLean, and E. A. Croft. (2014). “Meet Me Where i’m Gazing: How Shared Attention Gaze Affects Human-Robot Handover Timing”. In: *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction. HRI ’14*. Bielefeld, Germany: Association for Computing Machinery. 334–341. ISBN: 9781450326582. DOI: [10.1145/2559636.2559656](https://doi.org/10.1145/2559636.2559656). URL: <https://doi.org/10.1145/2559636.2559656>.
- Moorman, N., E. Hedlund-Botti, M. Schrum, M. Natarajan, and M. C. Gombolay. (2023a). “Impacts of robot learning on user attitude and behavior”. In: *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 534–543.

- Moorman, N., E. Hedlund-Botti, M. Schrum, M. Natarajan, and M. C. Gombolay. (2023b). “Impacts of Robot Learning on User Attitude and Behavior”. In: *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction. HRI '23*. Stockholm, Sweden: Association for Computing Machinery. 534–543. ISBN: 9781450399647. DOI: [10.1145/3568162.3576996](https://doi.org/10.1145/3568162.3576996). URL: <https://doi.org/10.1145/3568162.3576996>.
- Morales, Y., A. Watanabe, F. Ferreri, J. Even, T. Ikeda, K. Shinozawa, T. Miyashita, and N. Hagita. (2015). “Including human factors for planning comfortable paths”. In: *Proc. of the IEEE Int. Conf. on Robotics and Automation*. IEEE. 6153–6159.
- Moreira, I., J. Rivas, F. Cruz, R. Dazeley, A. Ayala, and B. Fernandes. (2020a). “Deep reinforcement learning with interactive feedback in a human–robot environment”. *Applied Sciences*. 10(16): 5574.
- Moreira, I., J. Rivas, F. Cruz, R. Dazeley, A. Ayala, and B. Fernandes. (2020b). “Deep Reinforcement Learning with Interactive Feedback in a Human-Robot Environment”. *Applied Sciences*. 10(16): 5574. ISSN: 2076-3417. DOI: [10.3390/app10165574](https://doi.org/10.3390/app10165574). URL: <http://arxiv.org/abs/2007.03363> (accessed on 11/10/2023).
- Morgan, A. A., J. Abdi, M. A. Syed, G. E. Kohen, P. Barlow, and M. P. Vizcaychipi. (2022). “Robots in healthcare: a scoping review”. *Current Robotics Reports*. 3(4): 271–280.
- Morrison, D., P. Corke, and J. Leitner. (2020). “Learning robust, real-time, reactive robotic grasping”. *The International journal of robotics research*. 39(2-3): 183–201.
- Mosier, K. L., L. J. Skitka, S. Heers, and M. Burdick. (1997). “Automation bias: decision making and performance in high-tech cockpits”. eng. *The International Journal of Aviation Psychology*. 8(1): 47–63. ISSN: 1050-8414. DOI: [10.1207/s15327108ijap0801_3](https://doi.org/10.1207/s15327108ijap0801_3).
- Mota, T. and M. Sridharan. (2021). “Answer me this: Constructing Disambiguation Queries for Explanation Generation in Robotics”. In: *2021 IEEE International Conference on Development and Learning (ICDL)*. 1–8. DOI: [10.1109/ICDL49984.2021.9515605](https://doi.org/10.1109/ICDL49984.2021.9515605).

- Moyle, W., C. Jones, B. Sung, M. Bramble, S. O'Dwyer, M. Blumenstein, and V. Estivill-Castro. (2016). "What Effect Does an Animal Robot Called CuDDler Have on the Engagement and Emotional Response of Older People with Dementia? A Pilot Feasibility Study". en. *International Journal of Social Robotics*. 8(1): 145–156. ISSN: 1875-4805. DOI: [10.1007/s12369-015-0326-7](https://doi.org/10.1007/s12369-015-0326-7). URL: <https://doi.org/10.1007/s12369-015-0326-7> (accessed on 11/05/2023).
- Mueller, C., J. Venicx, and B. Hayes. (2018). "Robust robot learning from demonstration and skill repair using conceptual constraints". In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 6029–6036.
- Muelling, K., A. Venkatraman, J.-S. Valois, J. E. Downey, J. Weiss, S. Javdani, M. Hebert, A. B. Schwartz, J. L. Collinger, and J. A. Bagnell. (2017). "Autonomy infused teleoperation with application to brain computer interface controlled manipulation". en. *Autonomous Robots*. 41(6): 1401–1422. ISSN: 1573-7527. DOI: [10.1007/s10514-017-9622-4](https://doi.org/10.1007/s10514-017-9622-4). URL: <https://doi.org/10.1007/s10514-017-9622-4>.
- Muir, B. M. and N. Moray. (1996). "Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation". eng. *Ergonomics*. 39(3): 429–460. ISSN: 0014-0139. DOI: [10.1080/00140139608964474](https://doi.org/10.1080/00140139608964474).
- Mukherjee, D., K. Gupta, L. H. Chang, and H. Najjaran. (2022). "A Survey of Robot Learning Strategies for Human-Robot Collaboration in Industrial Settings". *Robotics and Computer-Integrated Manufacturing*. 73: 102231. ISSN: 0736-5845. DOI: <https://doi.org/10.1016/j.rcim.2021.102231>. URL: <https://www.sciencedirect.com/science/article/pii/S0736584521001137>.
- Mülling, K., J. Kober, O. Kroemer, and J. Peters. (2013). "Learning to select and generalize striking movements in robot table tennis". en. *The International Journal of Robotics Research*. 32(3): 263–279. ISSN: 0278-3649, 1741-3176. DOI: [10.1177/0278364912472380](https://doi.org/10.1177/0278364912472380). URL: <http://journals.sagepub.com/doi/10.1177/0278364912472380>.
- Muoio, D. (2019). "Diligent robotics collects \$3M seed funding, Launches Autonomous Robot assistants for Hospitals". URL: <https://www.mobihealthnews.com/news/north-america/diligent-robotics-collects-3m-seed-funding-launches-autonomous-robot-assistants>.

- Muratore, F., F. Ramos, G. Turk, W. Yu, M. Gienger, and J. Peters. (2022). “Robot learning from randomized simulations: A review”. *Frontiers in Robotics and AI*: 31.
- Murphy, R. R., J. Casper, M. Micire, J. Hyams, *et al.* (2000). “Mixed-initiative control of multiple heterogeneous robots for urban search and rescue”. *submitted to IEEE Transactions on Robotics and Automation*.
- Murray, M., N. Walker, A. Nanavati, P. Alves-Oliveira, N. Filippov, A. Sauppe, B. Mutlu, and M. Cakmak. (2022). “Learning backchanneling behaviors for a social robot via data augmentation from human-human conversations”. In: *Conference on robot learning*. PMLR. 513–525.
- Mutlu, B., N. Roy, and S. Šabanović. (2016). “Cognitive human–robot interaction”. *Springer handbook of robotics*: 1907–1934.
- Nagata, F., K. Watanabe, K. Kiguchi, K. Tsuda, S. Kawaguchi, Y. Noda, and M. Komino. (2001). “Joystick teaching system for polishing robots using fuzzy compliance control”. In: *Proceedings 2001 IEEE International Symposium on Computational Intelligence in Robotics and Automation (Cat. No.01EX515)*. 362–367. DOI: [10.1109/CIRA.2001.1013226](https://doi.org/10.1109/CIRA.2001.1013226).
- Nair, A., S. Bahl, A. Khazatsky, V. Pong, G. Berseth, and S. Levine. (2020). “Contextual imagined goals for self-supervised robotic learning”. In: *Conference on Robot Learning*. PMLR. 530–539.
- Nanda, N. (2022). “A Comprehensive Mechanistic Interpretability Explainer & Glossary”.
- Nanda, N., L. Chan, T. Lieberum, J. Smith, and J. Steinhardt. (2023). “Progress measures for grokking via mechanistic interpretability”. *arXiv preprint arXiv:2301.05217*.
- Natarajan, M. and M. Gombolay. (2020). “Effects of Anthropomorphism and Accountability on Trust in Human Robot Interaction”. In: *2020 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 33–42. URL: <https://ieeexplore.ieee.org/document/9484201> (accessed on 11/22/2023).
- Natarajan, M., E. Seraj, B. Altundas, R. Paleja, S. Ye, L. Chen, R. Jensen, K. C. Chang, and M. Gombolay. (2023a). “Human-Robot Teaming: Grand Challenges”. *Current Robotics Reports*: 1–20.

- Natarajan, M., C. Xue, and M. Gombolay. (2023b). “Mixed-Initiative Human-Robot Teaming under Suboptimality”.
- Nehaniv, C. L., K. Dautenhahn, *et al.* (2002). “The correspondence problem”. *Imitation in animals and artifacts*. 41.
- Nemec, B., L. Žlajpah, and A. Ude. (2017). “Door opening by joining reinforcement learning and intelligent control”. In: *2017 18th International Conference on Advanced Robotics (ICAR)*. IEEE. 222–228.
- Nesset, B., D. A. Robb, J. Lopes, and H. Hastie. (2021). “Transparency in HRI: Trust and Decision Making in the Face of Robot Errors”. In: *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction. HRI '21 Companion*. Boulder, CO, USA: Association for Computing Machinery. 313–317. ISBN: 9781450382908. DOI: [10.1145/3434074.3447183](https://doi.org/10.1145/3434074.3447183). URL: <https://doi.org/10.1145/3434074.3447183>.
- Ng, A. Y. and S. J. Russell. (2000). “Algorithms for Inverse Reinforcement Learning”. In: *Proceedings of the Seventeenth International Conference on Machine Learning. ICML '00*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. 663–670. ISBN: 978-1-55860-707-1.
- Nguyen, S. M. and P.-Y. Oudeyer. (2012). “Interactive learning gives the tempo to an intrinsically motivated robot learner”. In: *2012 12th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2012)*. 645–652. DOI: [10.1109/HUMANOIDS.2012.6651588](https://doi.org/10.1109/HUMANOIDS.2012.6651588).
- Nguyen, S. M. and P.-Y. Oudeyer. (2014). “Socially guided intrinsic motivation for robot learning of motor skills”. *Autonomous Robots*. 36(3): 273–294. ISSN: 1573-7527. DOI: [10.1007/s10514-013-9339-y](https://doi.org/10.1007/s10514-013-9339-y). URL: <https://doi.org/10.1007/s10514-013-9339-y>.
- Nikolaidis, S., D. Hsu, and S. Srinivasa. (2017a). “Human-robot mutual adaptation in collaborative tasks: Models and experiments”. *The International Journal of Robotics Research*. 36(5-7): 618–634.

- Nikolaidis, S., R. Ramakrishnan, K. Gu, and J. Shah. (2015a). “Efficient Model Learning from Joint-Action Demonstrations for Human-Robot Collaborative Tasks”. In: *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction. HRI '15*. Portland, Oregon, USA: Association for Computing Machinery. 189–196. ISBN: 9781450328838. DOI: [10.1145/2696454.2696455](https://doi.org/10.1145/2696454.2696455). URL: <https://doi.org/10.1145/2696454.2696455>.
- Nikolaidis, S., R. Ramakrishnan, K. Gu, and J. Shah. (2015b). “Efficient model learning from joint-action demonstrations for human-robot collaborative tasks”. In: *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*. 189–196.
- Nikolaidis, S., Y. X. Zhu, D. Hsu, and S. Srinivasa. (2017b). “Human-Robot Mutual Adaptation in Shared Autonomy”. In: *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction. HRI '17*. New York, NY, USA: Association for Computing Machinery. 294–302. ISBN: 978-1-4503-4336-7. DOI: [10.1145/2909824.3020252](https://doi.org/10.1145/2909824.3020252). URL: <https://dl.acm.org/doi/10.1145/2909824.3020252>.
- Nishimura, M. and R. Yonetani. (2020). “L2b: Learning to balance the safety-efficiency trade-off in interactive crowd-aware robot navigation”. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 11004–11010.
- Nocentini, O., L. Fiorini, G. Acerbi, A. Sorrentino, G. Mancipopi, and F. Cavallo. (2019). “A survey of behavioral models for social robots”. *Robotics*. 8(3): 54.
- Nolfi, S. (2021). *Behavioral and Cognitive Robotics: An Adaptive Perspective*.
- Norouzzadeh, S., T. Lorenz, and S. Hirche. (2012). “Towards safe physical human-robot interaction: an online optimal control scheme”. In: *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. IEEE. 503–508.
- Nose, H., K. Kawabata, and Y. Suzuki. (1991). “Method of teaching a robot”.
- Nourmohammadi, A., M. Fathi, and A. H. Ng. (2022). “Balancing and scheduling assembly lines with human-robot collaboration tasks”. *Computers & Operations Research*. 140: 105674.

- Novak, D., B. Beyeler, X. Omlin, and R. Riener. (2015). “Workload estimation in physical human–robot interaction using physiological measurements”. *Interacting with computers*. 27(6): 616–629.
- Novelli, C., M. Taddeo, and L. Floridi. (2023). “Accountability in artificial intelligence: what it is and how it works”. *AI & SOCIETY*: 1–12.
- Nyberg, T., C. Pek, L. Dal Col, C. Norén, and J. Tumova. (2021). “Risk-aware motion planning for autonomous vehicles with safety specifications”. In: *2021 IEEE Intelligent Vehicles Symposium (IV)*. IEEE. 1016–1023.
- Nyholm, S. (2023). “Responsibility gaps, value alignment, and meaningful human control over artificial intelligence”. In: *Risk and responsibility in context*. Routledge. 191–213.
- Okamura, A. M., M. J. Matarić, and H. I. Christensen. (2010). “Medical and health-care robotics”. *IEEE Robotics & Automation Magazine*. 17(3): 26–37.
- Okamura, K. and S. Yamada. (2020). “Adaptive trust calibration for human-AI collaboration”. *PLoS ONE*. 15(2): e0229132. ISSN: 1932-6203. DOI: [10.1371/journal.pone.0229132](https://doi.org/10.1371/journal.pone.0229132). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7034851/>.
- Olatunji, S. A., T. Oron-Gilad, N. Markfeld, D. Gutman, V. Sarne-Fleischmann, and Y. Edan. (2021). “Levels of Automation and Transparency: Interaction Design Considerations in Assistive Robots for Older Adults”. *IEEE Transactions on Human-Machine Systems*. 51(6): 673–683. DOI: [10.1109/THMS.2021.3107516](https://doi.org/10.1109/THMS.2021.3107516).
- Oliff, H., Y. Liu, M. Kumar, M. Williams, and M. Ryan. (2020). “Reinforcement learning for facilitating human-robot-interaction in manufacturing”. *Journal of Manufacturing Systems*. 56: 326–340.
- Olivares-Alarcos, A., A. Andriella, S. Foix, and G. Alenyà. (2023). “Robot explanatory narratives of collaborative and adaptive experiences”. In: *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 11964–11971.

- Omeiza, D., H. Web, M. Jirotko, and L. Kunze. (2021a). “Towards Accountability: Providing Intelligible Explanations in Autonomous Driving”. en. In: *2021 IEEE Intelligent Vehicles Symposium (IV)*. Nagoya, Japan: IEEE. 231–237. ISBN: 978-1-72815-394-0. DOI: [10.1109/IV48863.2021.9575917](https://doi.org/10.1109/IV48863.2021.9575917). URL: <https://ieeexplore.ieee.org/document/9575917/>.
- Omeiza, D., H. Webb, M. Jirotko, and L. Kunze. (2021b). “Explanations in autonomous driving: A survey”. *IEEE Transactions on Intelligent Transportation Systems*. 23(8): 10142–10162.
- Onnasch, L. and E. Roesler. (2021). “A Taxonomy to Structure and Analyze Human–Robot Interaction”. *International Journal of Social Robotics*. 13(4): 833–849. ISSN: 1875-4805. DOI: [10.1007/s12369-020-00666-5](https://doi.org/10.1007/s12369-020-00666-5). URL: <https://doi.org/10.1007/s12369-020-00666-5>.
- Ortenzi, V., A. Cosgun, T. Pardi, W. P. Chan, E. Croft, and D. Kulić. (2021). “Object handovers: a review for robotics”. *IEEE Transactions on Robotics*. 37(6): 1855–1873.
- Ouyang, L., J. Wu, X. Jiang, D. Almeida, C. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, *et al.* (2022). “Training language models to follow instructions with human feedback”. *Advances in Neural Information Processing Systems*. 35: 27730–27744.
- Padalkar, A., A. Pooley, A. Jain, A. Bewley, A. Herzog, A. Irpan, A. Khazatsky, A. Rai, A. Singh, A. Brohan, *et al.* (2023). “Open x-embodiment: Robotic learning datasets and rt-x models”. *arXiv preprint arXiv:2310.08864*.
- Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford university press.
- Paleja, R., L. Chen, Y. Niu, A. Silva, Z. Li, S. Zhang, C. Ritchie, S. Choi, K. C. Chang, H. E. Tseng, *et al.* (2023a). “Interpretable Reinforcement Learning for Robotics and Continuous Control”. *arXiv preprint arXiv:2311.10041*.
- Paleja, R., M. Ghuy, N. Ranawaka Arachchige, R. Jensen, and M. Gombolay. (2021). “The utility of explainable ai in ad hoc human-machine teaming”. *Advances in neural information processing systems*. 34: 610–623.

- Paleja, R., Y. Niu, A. Silva, C. Ritchie, S. Choi, and M. Gombolay. (2022). “Learning Interpretable, High-Performing Policies for Continuous Control”.
- Paleja, R., Y. Niu, A. Silva, C. Ritchie, S. Choi, and M. Gombolay. (2023b). “Learning Interpretable, High-Performing Policies for Autonomous Driving”. DOI: [10.48550/arXiv.2202.02352](https://doi.org/10.48550/arXiv.2202.02352). URL: <http://arxiv.org/abs/2202.02352>.
- Paleja, R., A. Silva, L. Chen, and M. Gombolay. (2020a). “Interpretable and Personalized Apprenticeship Scheduling: Learning Interpretable Scheduling Policies from Heterogeneous User Demonstrations”. In: *Proceedings of the 34th International Conference on Neural Information Processing Systems. NIPS’20*. Vancouver, BC, Canada: Curran Associates Inc. ISBN: 9781713829546.
- Paleja, R., A. Silva, L. Chen, and M. Gombolay. (2020b). “Interpretable and Personalized Apprenticeship Scheduling: Learning Interpretable Scheduling Policies from Heterogeneous User Demonstrations”. In: *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*. Vol. 33. 6417–6428.
- Paraschos, A., C. Daniel, J. R. Peters, and G. Neumann. (2013). “Probabilistic Movement Primitives”. In: *Advances in Neural Information Processing Systems*. Ed. by C. J. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger. Vol. 26. Curran Associates, Inc. URL: https://proceedings.neurips.cc/paper_files/paper/2013/file/e53a0a2978c28872a4505bdb51db06dc-Paper.pdf.
- Parasuraman, R. and D. H. Manzey. (2010). “Complacency and bias in human use of automation: an attentional integration”. eng. *Human Factors*. 52(3): 381–410. ISSN: 0018-7208. DOI: [10.1177/0018720810376055](https://doi.org/10.1177/0018720810376055).
- Parasuraman, R. and V. Riley. (1997). “Humans and Automation: Use, Misuse, Disuse, Abuse”. en. *Human Factors*. 39(2): 230–253. ISSN: 0018-7208. DOI: [10.1518/001872097778543886](https://doi.org/10.1518/001872097778543886). URL: <https://doi.org/10.1518/001872097778543886>.
- Parekh, D., N. Poddar, A. Rajpurkar, M. Chahal, N. Kumar, G. P. Joshi, and W. Cho. (2022). “A Review on Autonomous Vehicles: Progress, Methods and Challenges”. en. *Electronics*. 11(14): 2162. ISSN: 2079-9292. DOI: [10.3390/electronics11142162](https://doi.org/10.3390/electronics11142162). URL: <https://www.mdpi.com/2079-9292/11/14/2162>.

- Park, H. W., M. Gelsomini, J. J. Lee, and C. Breazeal. (2017). “Telling stories to robots: The effect of backchanneling on a child’s storytelling”. In: *Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction*. 100–108.
- Park, J., S. Lim, J. Lee, S. Park, M. Chang, Y. Yu, and S. Choi. (2023). “Clara: Classifying and disambiguating user commands for reliable interactive robotic agents”. *IEEE Robotics and Automation Letters*.
- Parks, J. A. (2010). “Lifting the Burden of Women’s Care Work: Should Robots Replace the “Human Touch”?” *Hypatia*. 25(1): 100–120. DOI: [10.1111/j.1527-2001.2009.01086.x](https://doi.org/10.1111/j.1527-2001.2009.01086.x).
- Parnami, A. and M. Lee. (2022). “Learning from few examples: A summary of approaches to few-Shot Learning”.
- Pascher, M., U. Gruenefeld, S. Schneegass, and J. Gerken. (2023). “How to Communicate Robot Motion Intent: A Scoping Review”. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.
- Pathak, D., P. Agrawal, A. A. Efros, and T. Darrell. (2017). “Curiosity-driven Exploration by Self-supervised Prediction”. *CoRR*. abs/1705.05363. arXiv: [1705.05363](https://arxiv.org/abs/1705.05363). URL: <http://arxiv.org/abs/1705.05363>.
- Pek, C., G. F. Schuppe, F. Esposito, J. Tumova, and D. Kragic. (2023). “SpaTiaL: Monitoring and Planning of Robotic Tasks Using Spatio-Temporal Logic Specifications”. *Autonomous Robots*.
- Pérez-D’Arpino, C., C. Liu, P. Goebel, R. Martín-Martín, and S. Savarese. (2021). “Robot navigation in constrained pedestrian environments using reinforcement learning”. In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 1140–1146.
- Pervez, A., Y. Mao, and D. Lee. (2017). “Learning Deep Movement Primitives using Convolutional Neural Networks”. In: DOI: [10.1109/HUMANOIDS.2017.8246874](https://doi.org/10.1109/HUMANOIDS.2017.8246874).
- Perzanowski, D., A. C. Schultz, W. Adams, E. Marsh, and M. Bugajska. (2001). “Building a multimodal human-robot interface”. *IEEE intelligent systems*. 16(1): 16–21.

- Peters, R. A., C. L. Campbell, W. Bluethmann, and E. L. Huber. (2003). “Robonaut task learning through teleoperation”. *2003 IEEE International Conference on Robotics and Automation (Cat. No.03CH37422)*. 2: 2806–2811 vol.2. URL: <https://api.semanticscholar.org/CorpusID:5568701>.
- Petersen, K. and O. von Stryk. (2011). “An event-based communication concept for human supervision of autonomous robot teams”. *International Journal on Advances in Intelligent Systems*. 4(3&4): 357–369.
- Pinxteren, M. M. van, R. W. Wetzels, J. Rüger, M. Pluymaekers, and M. Wetzels. (2019). “Trust in humanoid robots: implications for services marketing”. *Journal of Services Marketing*. 33(4): 507–518. ISSN: 0887-6045. DOI: [10.1108/JSM-01-2018-0045](https://doi.org/10.1108/JSM-01-2018-0045). URL: <https://doi.org/10.1108/JSM-01-2018-0045>.
- Pnueli, A. (1977). “The temporal logic of programs”. In: *Proc. of the Annual Symposium on Foundations of Computer Science*. 46–57.
- Pore, A., E. Tagliabue, M. Piccinelli, D. Dall’Alba, A. Casals, and P. Fiorini. (2021). “Learning from Demonstrations for Autonomous Soft-tissue Retraction^{*}”. In: *International Symposium on Medical Robotics, ISMR 2021, Atlanta, GA, USA, November 17-19, 2021*. IEEE. 1–7. DOI: [10.1109/ISMR48346.2021.9661514](https://doi.org/10.1109/ISMR48346.2021.9661514). URL: <https://doi.org/10.1109/ISMR48346.2021.9661514>.
- Porfirio, D., A. Saupé, A. Albarghouthi, and B. Mutlu. (2018). “Authoring and verifying human-robot interactions”. In: *Proceedings of the 31st annual acm symposium on user interface software and technology*. 75–86.
- Portelas, R., C. Colas, L. Weng, K. Hofmann, and P.-Y. Oudeyer. (2020). “Automatic Curriculum Learning For Deep RL: A Short Survey”. In: *International Joint Conference on Artificial Intelligence*. URL: <https://api.semanticscholar.org/CorpusID:212644660>.
- Poursabzi-Sangdeh, F., D. G. Goldstein, J. M. Hofman, J. W. Vaughan, and H. Wallach. (2018). “Manipulating and measuring model interpretability.” (2018).

- Powers, A. and S. Kiesler. (2006). “The advisor robot: tracing people’s mental model from a robot’s physical attributes”. In: *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. 218–225.
- Prasad, V., D. Koert, R. Stock-Homburg, J. Peters, and G. Chaltatzaki. (2022). “MILD: Multimodal Interactive Latent Dynamics for Learning Human-Robot Interaction”. In: *2022 IEEE-RAS 21st International Conference on Humanoid Robots (Humanoids)*. 472–479. DOI: [10.1109/Humanoids53995.2022.10000239](https://doi.org/10.1109/Humanoids53995.2022.10000239).
- Prinzel III, L. J., F. G. Freeman, M. W. Scerbo, P. J. Mikulka, and A. T. Pope. (2003). “Effects of a psychophysiological system for adaptive automation on performance, workload, and the event-related potential P300 component”. *Human factors*. 45(4): 601–614.
- Puig, X., E. Undersander, A. Szot, M. D. Cote, T.-Y. Yang, R. Partsey, R. Desai, A. W. Clegg, M. Hlavac, S. Y. Min, *et al.* (2023). “Habitat 3.0: A co-habitat for humans, avatars and robots”. *arXiv preprint arXiv:2310.13724*.
- Puranic, A., J. Deshmukh, and S. Nikolaidis. (2021). “Learning from demonstrations using signal temporal logic”. In: *Conference on Robot Learning*. PMLR. 2228–2242.
- Pynadath, D. V., N. Gurney, and N. Wang. (2022). “Explainable reinforcement learning in human-robot teams: the impact of decision-tree explanations on transparency”. In: *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 749–756.
- Qiao, B. and X. Wu. (2019). “Real Time Trajectory Re-planning for Autonomous Vehicle Lane Changing in Uncertain Traffic”. en. In: *2019 IEEE 28th International Symposium on Industrial Electronics (ISIE)*. Vancouver, BC, Canada: IEEE. 1524–1529. ISBN: 978-1-72813-666-0. DOI: [10.1109/ISIE.2019.8781322](https://doi.org/10.1109/ISIE.2019.8781322). URL: <https://ieeexplore.ieee.org/document/8781322/>.

- Qin, B., Y. Gao, and Y. Bai. (2019). “Sim-to-real: Six-legged Robot Control with Deep Reinforcement Learning and Curriculum Learning”. In: *2019 4th International Conference on Robotics and Automation Engineering (ICRAE)*. 1–5. DOI: [10.1109/ICRAE48301.2019.9043822](https://doi.org/10.1109/ICRAE48301.2019.9043822).
- Quispe, A. C. H., E. Martinson, and K. Oguchi. (2017). “Learning user preferences for robot-human handovers”. In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 834–839.
- Qureshi, A., Y. Nakamura, Y. Yoshikawa, and H. Ishiguro. (2017). “Robot gains Social Intelligence through Multimodal Deep Reinforcement Learning”. Feb.
- Qureshi, A. H., Y. Nakamura, Y. Yoshikawa, and H. Ishiguro. (2018). “Intrinsically motivated reinforcement learning for human–robot interaction in the real-world”. *Neural Networks*. 107: 23–33. ISSN: 0893-6080. DOI: <https://doi.org/10.1016/j.neunet.2018.03.014>. URL: <https://www.sciencedirect.com/science/article/pii/S0893608018301072>.
- Rajeswaran, A., V. Kumar, A. Gupta, G. Vezzani, J. Schulman, E. Todorov, and S. Levine. (2017). “Learning complex dexterous manipulation with deep reinforcement learning and demonstrations”. *arXiv preprint arXiv:1709.10087*.
- Rajeswaran, A., V. Kumar, A. Gupta, G. Vezzani, J. Schulman, E. Todorov, and S. Levine. (2018). “Learning Complex Dexterous Manipulation with Deep Reinforcement Learning and Demonstrations”. *Robotics: Science and Systems XIV*.
- Raji, I. D., A. Smart, R. N. White, M. Mitchell, T. Gebru, B. Hutchinson, J. Smith-Loud, D. Theron, and P. Barnes. (2020). “Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing”. In: *Proceedings of the 2020 conference on fairness, accountability, and transparency*. 33–44.
- Ramachandran, D. and E. Amir. (2007). “Bayesian Inverse Reinforcement Learning”. In: *Proceedings of the 20th International Joint Conference on Artificial Intelligence. IJCAI’07*. Hyderabad, India: Morgan Kaufmann Publishers Inc. 2586–2591.

- Raman, V., A. Donzé, D. Sadigh, R. M. Murray, and S. A. Seshia. (2015). “Reactive synthesis from signal temporal logic specifications”. In: *Proceedings of the 18th international conference on hybrid systems: Computation and control*. 239–248.
- Ramaraj, P. (2021). “Robots that help humans build better mental models of robots”. In: *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 595–597.
- Ramchurn, S. D., F. Wu, W. Jiang, J. E. Fischer, S. Reece, S. Roberts, T. Rodden, C. Greenhalgh, and N. R. Jennings. (2016). “Human-agent collaboration for disaster response”. *Autonomous Agents and Multi-Agent Systems*. 30: 82–111.
- Ramponi, G., A. Likmeta, A. M. Metelli, A. Tirinzoni, and M. Restelli. (2020). “Truly Batch Model-Free Inverse Reinforcement Learning about Multiple Intentions”. In: *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*. Ed. by S. Chiappa and R. Calandra. Vol. 108. *Proceedings of Machine Learning Research*. PMLR. 2359–2369. URL: <https://proceedings.mlr.press/v108/ramponi20a.html>.
- Ratliff, N. D., J. A. Bagnell, and M. A. Zinkevich. (2006). “Maximum margin planning”. en. In: *Proceedings of the 23rd international conference on Machine learning - ICML '06*. Pittsburgh, Pennsylvania: ACM Press. 729–736. ISBN: 978-1-59593-383-6. DOI: [10.1145/1143844.1143936](https://doi.org/10.1145/1143844.1143936). URL: <http://portal.acm.org/citation.cfm?doid=1143844.1143936>.
- Raval, N., A. Srivastava, K. Lebeck, L. Cox, and A. Machanavajjhala. (2014). “Markit: Privacy markers for protecting visual secrets”. In: *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: Adjunct publication*. 1289–1295.
- Ravichandar, H., A. S. Polydoros, S. Chernova, and A. Billard. (2020). “Recent advances in robot learning from demonstration”. *Annual review of control, robotics, and autonomous systems*. 3: 297–330.
- Raychaudhuri, S., T. Campari, U. Jain, M. Savva, and A. X. Chang. (2024). “MOPA: Modular Object Navigation With PointGoal Agents”. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 5763–5773.

- Raza, S. J., S. Froghi, A. Chowriappa, K. Ahmed, E. Field, A. P. Stegemann, S. Rehman, M. Sharif, Y. Shi, G. E. Wilding, *et al.* (2014). “Construct validation of the key components of Fundamental Skills of Robotic Surgery (FSRS) curriculum—a multi-institution prospective study”. *Journal of surgical education*. 71(3): 316–324.
- Reddy, S., A. Dragan, and S. Levine. (2018). “Shared Autonomy via Deep Reinforcement Learning”. en. In: *Robotics: Science and Systems XIV*. Robotics: Science and Systems Foundation. ISBN: 978-0-9923747-4-7. DOI: [10.15607/RSS.2018.XIV.005](https://doi.org/10.15607/RSS.2018.XIV.005). URL: <http://www.roboticsproceedings.org/rss14/p05.pdf>.
- Ren, A. Z., A. Dixit, A. Bodrova, S. Singh, S. Tu, N. Brown, P. Xu, L. Takayama, F. Xia, J. Varley, *et al.* (2023a). “Robots that ask for help: Uncertainty alignment for large language model planners”. *arXiv preprint arXiv:2307.01928*.
- Ren, A. Z., B. Govil, T.-Y. Yang, K. R. Narasimhan, and A. Majumdar. (2023b). “Leveraging language for accelerated learning of tool manipulation”. In: *Conference on Robot Learning*. PMLR. 1531–1541.
- Ribeiro, M. T., S. Singh, and C. Guestrin. (2016). “" Why should i trust you?" Explaining the predictions of any classifier”. In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. 1135–1144.
- Rickel, J. and W. L. Johnson. (2000). “Task-oriented collaboration with embodied agents in virtual worlds”. *Embodied conversational agents*: 95–122.
- Riley, H. and M. Sridharan. (2019). “Integrating non-monotonic logical reasoning and inductive learning with deep learning for explainable visual question answering”. *Frontiers in Robotics and AI*. 6: 125.
- Roncone, A., O. Mangin, and B. Scassellati. (2017). “Transparent role assignment and task allocation in human robot collaboration”. In: *2017 IEEE International Conference on Robotics and Automation (ICRA)*. 1014–1021. DOI: [10.1109/ICRA.2017.7989122](https://doi.org/10.1109/ICRA.2017.7989122).
- Rosenberger, P., A. Cosgun, R. Newbury, J. Kwan, V. Ortenzi, P. Corke, and M. Grafinger. (2020). “Object-independent human-to-robot handovers using real time robotic vision”. *IEEE Robotics and Automation Letters*. 6(1): 17–23.

- Rosenfeld, A., M. Cohen, M. E. Taylor, and S. Kraus. (2018). “Leveraging human knowledge in tabular reinforcement learning: a study of human subjects”. *The Knowledge Engineering Review*. 33: e14. DOI: [10.1017/S0269888918000206](https://doi.org/10.1017/S0269888918000206).
- Rosenthal, S., M. Veloso, and A. K. Dey. (2012). “Is someone in this office available to help me? Proactively seeking help from spatially-situated humans”. *Journal of Intelligent & Robotic Systems*. 66: 205–221.
- Ross, S., G. Gordon, and D. Bagnell. (2011). “A reduction of imitation learning and structured prediction to no-regret online learning”. In: *Proceedings of the fourteenth international conference on artificial intelligence and statistics*. JMLR Workshop and Conference Proceedings. 627–635.
- Rossi, A., M. M. Scheunemann, G. L’Arco, and S. Rossi. (2021). “Evaluation of a Humanoid Robot’s Emotional Gestures for Transparent Interaction”. In: *Social Robotics*. Ed. by H. Li, S. S. Ge, Y. Wu, A. Wykowska, H. He, X. Liu, D. Li, and J. Perez-Osorio. Cham: Springer International Publishing. 397–407. ISBN: 978-3-030-90525-5.
- Rother, C., V. Kolmogorov, and A. Blake. (2004). “" GrabCut" interactive foreground extraction using iterated graph cuts”. *ACM transactions on graphics (TOG)*. 23(3): 309–314.
- Rotsidis, A., A. Theodorou, J. J. Bryson, and R. H. Wortham. (2019). “Improving Robot Transparency: An Investigation With Mobile Augmented Reality”. In: *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 1–8. DOI: [10.1109/RO-MAN46459.2019.8956390](https://doi.org/10.1109/RO-MAN46459.2019.8956390).
- Rozo, L., S. Calinon, D. Caldwell, P. Jimenez, and C. Torras. (2013). “Learning Collaborative Impedance-Based Robot Behaviors”. en. *Proceedings of the AAAI Conference on Artificial Intelligence*. 27(1): 1422–1428. ISSN: 2374-3468. DOI: [10.1609/aaai.v27i1.8543](https://doi.org/10.1609/aaai.v27i1.8543). URL: <https://ojs.aaai.org/index.php/AAAI/article/view/8543>.
- Rubagotti, M., I. Tusseyeva, S. Baltabayeva, D. Summers, and A. Sandygulova. (2022). “Perceived safety in physical human–robot interaction—A survey”. *Robotics and Autonomous Systems*. 151: 104047.

- Rudin, C. (2019). “Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell* 1: 206–215”. DOI: <https://doi.org/10.1038/s42256-019-0048-x>.
- Rudin, C., C. Chen, Z. Chen, H. Huang, L. Semenova, and C. Zhong. (2022). “Interpretable machine learning: Fundamental principles and 10 grand challenges”. *Statistic Surveys*. 16: 1–85.
- Rueben, M., J. Klow, M. Duer, E. Zimmerman, J. Piacentini, M. Browning, F. J. Bernieri, C. M. Grimm, and W. D. Smart. (2021). “Mental models of a mobile shoe rack: Exploratory findings from a long-term in-the-wild study”. *ACM Transactions on Human-Robot Interaction (THRI)*. 10(2): 1–36.
- Russell, S. (1998). “Learning agents for uncertain environments (extended abstract)”. en. In: *Proceedings of the eleventh annual conference on Computational learning theory*. Madison Wisconsin USA: ACM. 101–103. ISBN: 978-1-58113-057-7. DOI: [10.1145/279943.279964](https://doi.org/10.1145/279943.279964). URL: <https://dl.acm.org/doi/10.1145/279943.279964> (accessed on 12/24/2023).
- Russell, S. and P. Norvig. (2010). *Artificial Intelligence: A Modern Approach*. 3rd ed. Prentice Hall.
- Sado, F., C. K. Loo, W. S. Liew, M. Kerzel, and S. Wermter. (2023). “Explainable Goal-Driven Agents and Robots - A Comprehensive Review”. *ACM Comput. Surv.* 55(10). ISSN: 0360-0300. DOI: [10.1145/3564240](https://doi.org/10.1145/3564240). URL: <https://doi.org/10.1145/3564240>.
- Saeidi, H., F. McLane, B. Sadrfaidpour, E. Sand, S. Fu, J. Rodriguez, J. R. Wagner, and Y. Wang. (2016). “Trust-based mixed-initiative teleoperation of mobile robots”. In: *2016 American Control Conference (ACC)*. 6177–6182. DOI: [10.1109/ACC.2016.7526640](https://doi.org/10.1109/ACC.2016.7526640). URL: <https://ieeexplore.ieee.org/abstract/document/7526640>.
- Safaoui, S., L. Lindemann, D. V. Dimarogonas, I. Shames, and T. H. Summers. (2020). “Control design for risk-based signal temporal logic specifications”. *IEEE Control Systems Letters*. 4(4): 1000–1005.
- Sakai, T. and T. Nagai. (2022). “Explainable autonomous robots: a survey and perspective”. *Advanced Robotics*. 36(5-6): 219–238. DOI: [10.1080/01691864.2022.2029720](https://doi.org/10.1080/01691864.2022.2029720). URL: <https://doi.org/10.1080/01691864.2022.2029720>.

- Salas, E., T. L. Dickinson, S. A. Converse, and S. I. Tannenbaum. (1992). "Toward an understanding of team performance and training."
- Salem, M. and K. Dautenhahn. (2015). "Evaluating trust and safety in HRI: Practical issues and ethical challenges". *Emerging Policy and Ethics of Human-Robot Interaction*.
- Salem, M., S. Kopp, I. Wachsmuth, K. Rohlfing, and F. Joublin. (2012). "Generation and evaluation of communicative robot gesture". *International Journal of Social Robotics*. 4: 201–217.
- Salem, M., G. Lakatos, F. Amirabdollahian, and K. Dautenhahn. (2015). "Towards safe and trustworthy social robots: ethical challenges and practical issues". In: *Social Robotics: 7th International Conference, ICSR 2015, Paris, France, October 26-30, 2015, Proceedings 7*. Springer. 584–593.
- Sanders, T., K. E. Oleson, D. Billings, J. Y. Chen, and P. Hancock. (2011). "A Model of Human-Robot Trust: Theoretical Model Development". en. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 55(1): 1432–1436. ISSN: 2169-5067. DOI: [10.1177/1071181311551298](https://doi.org/10.1177/1071181311551298). URL: <https://journals.sagepub.com/doi/abs/10.1177/1071181311551298>.
- Sanders, T. L., T. Wixon, K. E. Schafer, J. Y. C. Chen, and P. A. Hancock. (2014). "The influence of modality and transparency on trust in human-robot interaction". In: *2014 IEEE International Inter-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*. 156–159. DOI: [10.1109/CogSIMA.2014.6816556](https://doi.org/10.1109/CogSIMA.2014.6816556).
- Sanneman, L., C. Fourie, J. A. Shah, *et al.* (2021). "The state of industrial robotics: Emerging technologies, challenges, and key research directions". *Foundations and Trends® in Robotics*. 8(3): 225–306.
- Sanneman, L. and J. Shah. (2021). "Explaining reward functions to humans for better human-robot collaboration". *arXiv preprint arXiv:2110.04192*.
- Sanneman, L. and J. A. Shah. (2020). "Trust considerations for explainable robots: A human factors perspective". *arXiv preprint arXiv:2005.05940*.
- Sanneman, L. and J. A. Shah. (2022). "An empirical study of reward explanations with human-robot interaction applications". *IEEE Robotics and Automation Letters*. 7(4): 8956–8963.

- Sano, T., T. Horii, K. Abe, and T. Nagai. (2021). “Temperament estimation of toddlers from child–robot interaction with explainable artificial intelligence”. *Advanced Robotics*. 35(17): 1068–1077.
- Santoni de Sio, F. and G. Mecacci. (2021). “Four responsibility gaps with artificial intelligence: Why they matter and how to address them”. *Philosophy & Technology*. 34: 1057–1084.
- Satake, S., T. Kanda, D. F. Glas, M. Imai, H. Ishiguro, and N. Hagita. (2009). “How to approach humans? Strategies for social robots to initiate interaction”. In: *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*. 109–116.
- Saunders, J., D. S. Syrdal, K. L. Koay, N. Burke, and K. Dautenhahn. (2015). ““teach me–show me”—end-user personalization of a smart home and companion robot”. *IEEE Transactions on Human-Machine Systems*. 46(1): 27–40.
- Saunderson, S. and G. Nejat. (2019). “How robots influence humans: A survey of nonverbal communication in social human–robot interaction”. *International Journal of Social Robotics*. 11: 575–608.
- Sauppe, A. and B. Mutlu. (2015). “The Social Impact of a Robot Co-Worker in Industrial Settings”. en. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Seoul Republic of Korea: ACM. 3613–3622. ISBN: 978-1-4503-3145-6. DOI: [10.1145/2702123.2702181](https://doi.org/10.1145/2702123.2702181). URL: <https://dl.acm.org/doi/10.1145/2702123.2702181>.
- Saveriano, M., F. J. Abu-Dakka, A. Kramberger, and L. Peternel. (2023). “Dynamic Movement Primitives in Robotics: A Tutorial Survey”. en. *The International Journal of Robotics Research*. 42(13): 1133–1184. ISSN: 0278-3649, 1741-3176. DOI: [10.1177/02783649231201196](https://doi.org/10.1177/02783649231201196). URL: <http://arxiv.org/abs/2102.03861>.
- Schaal, S. (1999). “Is imitation learning the route to humanoid robots?” eng. *Trends in Cognitive Sciences*. 3(6): 233–242. ISSN: 1879-307X. DOI: [10.1016/s1364-6613\(99\)01327-3](https://doi.org/10.1016/s1364-6613(99)01327-3).

- Schaal, S. (2006). “Dynamic Movement Primitives -A Framework for Motor Control in Humans and Humanoid Robotics”. In: *Adaptive Motion of Animals and Machines*. Ed. by H. Kimura, K. Tsuchiya, A. Ishiguro, and H. Witte. Tokyo: Springer Tokyo. 261–280. ISBN: 978-4-431-31381-6. DOI: [10.1007/4-431-31381-8_23](https://doi.org/10.1007/4-431-31381-8_23). URL: https://doi.org/10.1007/4-431-31381-8_23.
- Schaff, C. and M. R. Walter. (2020). “Residual policy learning for shared autonomy”. *arXiv preprint arXiv:2004.05097*.
- Scheggi, S., M. Aggravi, and D. Prattichizzo. (2016). “Cooperative navigation for mixed human–robot teams using haptic feedback”. *IEEE Transactions on Human-Machine Systems*. 47(4): 462–473.
- Scheggi, S., F. Chinello, and D. Prattichizzo. (2012). “Vibrotactile haptic feedback for human-robot interaction in leader-follower tasks”. In: *Proceedings of the 5th International Conference on Pervasive Technologies Related to Assistive Environments*. 1–4.
- Scheikl, P. M., E. Tagliabue, B. Gyenes, M. Wagner, D. Dall’Alba, P. Fiorini, and F. Mathis-Ullrich. (2022). “Sim-to-real transfer for visual reinforcement learning of deformable object manipulation for robot-assisted surgery”. *IEEE Robotics and Automation Letters*. 8(2): 560–567.
- Schermerhorn, P. W. and M. Scheutz. (2011). “Disentangling the Effects of Robot Affect, Embodiment, and Autonomy on Human Team Members in a Mixed-Initiative Task”. In: *International Conference on Advances in Computer-Human Interaction*. URL: <https://api.semanticscholar.org/CorpusID:913462>.
- Schilling, M., S. Kopp, S. Wachsmuth, B. Wrede, H. J. Ritter, T. Brox, B. Nebel, and W. Burgard. (2016). “Towards A Multidimensional Perspective on Shared Autonomy”. In: *AAAI Fall Symposia*.
- Schillinger, P., M. Bürger, and D. V. Dimarogonas. (2018). “Decomposition of finite LTL specifications for efficient multi-agent planning”. In: *Distributed Autonomous Robotic Systems*. 253–267.
- Schmalstieg, F., D. Honerkamp, T. Welschehold, and A. Valada. (2023). “Learning Hierarchical Interactive Multi-Object Search for Mobile Manipulation”. *IEEE Robotics and Automation Letters*.

- Schmidhuber, J. (2010). “Formal Theory of Creativity, Fun, and Intrinsic Motivation (1990–2010)”. *IEEE Transactions on Autonomous Mental Development*. 2(3): 230–247. DOI: [10.1109/TAMD.2010.2056368](https://doi.org/10.1109/TAMD.2010.2056368).
- Schrum, M., M. J. Connolly, E. Cole, M. Ghetiya, R. Gross, and M. C. Gombolay. (2022a). “Meta-Active Learning in Probabilistically Safe Optimization”. *IEEE Robotics and Automation Letters*. 7(4): 10713–10720.
- Schrum, M. L. and M. C. Gombolay. (2019). “When your robot breaks: Active learning during plant failure”. *IEEE Robotics and Automation Letters*. 5(2): 438–445.
- Schrum, M. L., E. Hedlund-Botti, and M. Gombolay. (2023a). “Reciprocal MIND MELD: Improving Learning From Demonstration via Personalized, Reciprocal Teaching”. In: *Proceedings of The 6th Conference on Robot Learning*. Ed. by K. Liu, D. Kulic, and J. Ichnowski. Vol. 205. *Proceedings of Machine Learning Research*. PMLR. 956–966. URL: <https://proceedings.mlr.press/v205/schrum23a.html>.
- Schrum, M. L., M. Johnson, M. Ghuy, and M. C. Gombolay. (2020). “Four years in review: Statistical practices of likert scales in human-robot interaction studies”. In: *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. 43–52.
- Schrum, M. L., E. Hedlund-Botti, N. Moorman, and M. C. Gombolay. (2022b). “MIND MELD: Personalized Meta-Learning for Robot-Centric Imitation Learning”. In: *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 157–165. DOI: [10.1109/HRI53351.2022.9889616](https://doi.org/10.1109/HRI53351.2022.9889616).
- Schrum, M. L., E. Sumner, M. C. Gombolay, and A. Best. (2023b). “MAVERIC: A Data-Driven Approach to Personalized Autonomous Driving”. en. URL: <http://arxiv.org/abs/2301.08595>.
- Schulman, J., F. Wolski, P. Dhariwal, A. Radford, and O. Klimov. (2017). “Proximal Policy Optimization Algorithms.” *CoRR*. abs/1707.06347. URL: <http://dblp.uni-trier.de/db/journals/corr/corr1707.html#SchulmanWDRK17>.
- Schumann, C., X. Wang, A. Beutel, J. Chen, H. Qian, and E. H. Chi. (2019). “Transfer of machine learning fairness across domains”. *arXiv preprint arXiv:1906.09688*.

- Sekmen, A. and P. Challa. (2013). “Assessment of adaptive human–robot interactions”. *Knowledge-Based Systems*. 42: 49–59. ISSN: 0950-7051. DOI: <https://doi.org/10.1016/j.knosys.2013.01.003>. URL: <https://www.sciencedirect.com/science/article/pii/S0950705113000075>.
- Selim, M., A. Alanwar, S. Kousik, G. Gao, M. Pavone, and K. H. Johansson. (2022). “Safe reinforcement learning using black-box reachability analysis”. *IEEE Robotics and Automation Letters*. 7(4): 10665–10672.
- Selvaggio, M., J. Cacace, C. Pacchierotti, F. Ruggiero, and P. R. Giordano. (2022). “A Shared-Control Teleoperation Architecture for Non-prehensile Object Transportation”. *IEEE Transactions on Robotics*. 38(1): 569–583. ISSN: 1941-0468. DOI: [10.1109/TRO.2021.3086773](https://doi.org/10.1109/TRO.2021.3086773). URL: <https://ieeexplore.ieee.org/document/9477290>.
- Sequeira, P. and M. Gervasio. (2020). “Interestingness elements for explainable reinforcement learning: Understanding agents’ capabilities and limitations”. *Artificial Intelligence*. 288: 103367.
- Seraj, E. (2022). “Embodied Team Intelligence in Multi-Robot Systems.” In: *AAMAS*. 1869–1871.
- Seraj, E. (2023). “Embodied, intelligent communication for multi-agent cooperation”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 37. No. 13. 16135–16136.
- Seraj, E. and W. Talamonti. (2024). “Towards In-Vehicle Multi-Task Facial Attribute Recognition: Investigating Synthetic Data and Vision Foundation Models”. *arXiv preprint arXiv:2403.06088*.
- Seraj, E., Z. Wang, R. Paleja, D. Martin, M. Sklar, A. Patel, and M. Gombolay. (2022). “Learning efficient diverse communication for cooperative heterogeneous teaming”. In: *Proceedings of the 21st international conference on autonomous agents and multiagent systems*. 1173–1182.
- Seraj, E., J. Y. Xiong, M. L. Schrum, and M. Gombolay. (2023). “Mixed-Initiative Multiagent Apprenticeship Learning for Human Training of Robot Teams”. In: *Thirty-seventh Conference on Neural Information Processing Systems*.

- Setchi, R., M. B. Dehkordi, and J. S. Khan. (2020a). “Explainable Robotics in Human-Robot Interactions”. *Procedia Computer Science*. Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 24th International Conference KES2020 176(Jan.): 3057–3066. ISSN: 1877-0509. DOI: [10.1016/j.procs.2020.09.198](https://doi.org/10.1016/j.procs.2020.09.198). (Accessed on 12/10/2023).
- Setchi, R., M. B. Dehkordi, and J. S. Khan. (2020b). “Explainable robotics in human-robot interactions”. *Procedia Computer Science*. 176: 3057–3066.
- Settles, B. (2009). “Active Learning Literature Survey”. *Computer Sciences Technical Report* No. 1648. University of Wisconsin–Madison. URL: <http://axon.cs.byu.edu/~martinez/classes/778/Papers/settles.activelearning.pdf>.
- Shah, A., S. Li, and J. Shah. (2020a). “Planning With Uncertain Specifications (PUnS)”. en. *IEEE Robotics and Automation Letters*. 5(2): 3414–3421. ISSN: 2377-3766, 2377-3774. DOI: [10.1109/LRA.2020.2977217](https://doi.org/10.1109/LRA.2020.2977217). URL: <https://ieeexplore.ieee.org/document/9018240/> (accessed on 11/05/2023).
- Shah, A., S. Wadhwan, and J. Shah. (2020b). “Interactive robot training for non-markov tasks”. *arXiv preprint arXiv:2003.02232*.
- Shah, N., P. Verma, T. Angle, and S. Srivastavam. (2022). “JEDAI: A System for Skill-Aligned Explainable Robot Plannin”. In: *2022 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2022)*. 1–8.
- Shah, S. K., C. D. Pahlajani, N. A. Lacock, and H. G. Tanner. (2012). “Stochastic receding horizon control for robots with probabilistic state constraints”. In: *2012 IEEE International Conference on Robotics and Automation*. IEEE. 2893–2898.
- Shao, Y. S., C. Chen, S. Kousik, and R. Vasudevan. (2021). “Reachability-based trajectory safeguard (RTS): A safe and fast reinforcement learning safety layer for continuous control”. *IEEE Robotics and Automation Letters*. 6(2): 3663–3670.
- Sharma, P., B. Sundaralingam, V. Blukis, C. Paxton, T. Hermans, A. Torralba, J. Andreas, and D. Fox. (2022a). “Correcting Robot Plans with Natural Language Feedback”. *ArXiv*. abs/2204.05186. URL: <https://api.semanticscholar.org/CorpusID:248085271>.

- Sharma, P., B. Sundaralingam, V. Blukis, C. Paxton, T. Hermans, A. Torralba, J. Andreas, and D. Fox. (2022b). “Correcting robot plans with natural language feedback”. *arXiv preprint arXiv:2204.05186*.
- Sheikh, J., A. Melnik, G. C. Nandi, and R. Haschke. (2023). “Language-Conditioned Semantic Search-Based Policy for Robotic Manipulation Tasks”. *arXiv preprint arXiv:2312.05925*.
- Sheikholeslami, S., J. W. Hart, W. P. Chan, C. P. Quintero, and E. A. Croft. (2018). “Prediction and Production of Human Reaching Trajectories for Human-Robot Interaction”. In: *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction. HRI '18*. Chicago, IL, USA: Association for Computing Machinery. 321–322. ISBN: 9781450356152. DOI: [10.1145/3173386.3176924](https://doi.org/10.1145/3173386.3176924). URL: <https://doi.org/10.1145/3173386.3176924>.
- Shek, A., B. Y. Su, R. Chen, and C. Liu. (2023). “Learning from Physical Human Feedback: An Object-Centric One-Shot Adaptation Method”. In: *2023 IEEE International Conference on Robotics and Automation (ICRA)*. 9910–9916. DOI: [10.1109/ICRA48891.2023.10161416](https://ieeexplore.ieee.org/document/10161416). URL: <https://ieeexplore.ieee.org/document/10161416>.
- Shepherd, M. K., D. D. Molinaro, G. S. Sawicki, and A. J. Young. (2022). “Deep Learning Enables Exoboot Control to Augment Variable-Speed Walking”. *IEEE Robotics and Automation Letters*. 7(2): 3571–3577. ISSN: 2377-3766. DOI: [10.1109/LRA.2022.3147565](https://ieeexplore.ieee.org/abstract/document/9699101). URL: <https://ieeexplore.ieee.org/abstract/document/9699101>.
- Sheridan, T. B. (2016). “Human–robot interaction: status and challenges”. *Human factors*. 58(4): 525–532.
- Shin, H., J. K. Lee, J. Kim, and J. Kim. (2017). “Continual Learning with Deep Generative Replay”. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems. NIPS'17*. Long Beach, California, USA: Curran Associates Inc. 2994–3003. ISBN: 9781510860964.
- Shridhar, M., D. Mittal, and D. Hsu. (2020). “INGRESS: Interactive visual grounding of referring expressions”. *The International Journal of Robotics Research*. 39(2-3): 217–232. DOI: [10.1177/0278364919897133](https://doi.org/10.1177/0278364919897133). eprint: <https://doi.org/10.1177/0278364919897133>. URL: <https://doi.org/10.1177/0278364919897133>.

- Shu, B., G. Sziebig, and R. Pieters. (2019). “Architecture for safe human-robot collaboration: Multi-modal communication in virtual reality for efficient task execution”. In: *2019 IEEE 28th International Symposium on Industrial Electronics (ISIE)*. IEEE. 2297–2302.
- Sierra, C., N. Osman, P. Noriega, J. Sabater-Mir, and A. Perelló. (2021). “Value alignment: a formal approach”. *arXiv preprint arXiv:2110.09240*.
- Silva, A. and M. Gombolay. (2021). “Encoding human domain knowledge to warm start reinforcement learning”. In: *Proceedings of the AAAI conference on artificial intelligence*. Vol. 35. No. 6. 5042–5050.
- Silva, A., N. Moorman, W. Silva, Z. Zaidi, N. Gopalan, and M. Gombolay. (2022a). “LanCon-Learn: Learning With Language to Enable Generalization in Multi-Task Manipulation”. *IEEE Robotics and Automation Letters*. 7(2): 1635–1642. DOI: [10.1109/LRA.2021.3139667](https://doi.org/10.1109/LRA.2021.3139667).
- Silva, A., N. Moorman, W. Silva, Z. Zaidi, N. Gopalan, and M. Gombolay. (2022b). “LanCon-Learn: Learning With Language to Enable Generalization in Multi-Task Manipulation”. *IEEE Robotics and Automation Letters*. 7(2): 1635–1642. DOI: [10.1109/LRA.2021.3139667](https://doi.org/10.1109/LRA.2021.3139667).
- Silva, A., M. Schrum, E. Hedlund-Botti, N. Gopalan, and M. Gombolay. (2023). “Explainable artificial intelligence: Evaluating the objective and subjective impacts of xai on human-agent interaction”. *International Journal of Human–Computer Interaction*. 39(7): 1390–1404.
- Silva, A., P. Tambwekar, and M. Gombolay. (2021). “Towards a comprehensive understanding and accurate evaluation of societal biases in pre-trained transformers”. In: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 2383–2389.
- Silva, A., P. Tambwekar, and M. Gombolay. (2022c). “FedPC: Federated Learning for Language Generation with Personal and Context Preference Embeddings”. *arXiv preprint arXiv:2210.03766*.
- Silver, D., J. A. Bagnell, and A. Stentz. (2012). “Active learning from demonstration for robust autonomous navigation”. In: *2012 IEEE International Conference on Robotics and Automation*. 200–207. DOI: [10.1109/ICRA.2012.6224757](https://doi.org/10.1109/ICRA.2012.6224757).
- Skantze, G. (2016). “Real-time coordination in human-robot interaction using face and voice”. *Ai Magazine*. 37(4): 19–31.

- Skantze, G., A. Hjalmarsson, and C. Oertel. (2014). “Turn-taking, feedback and joint attention in situated human–robot interaction”. *Speech Communication*. 65: 50–66.
- Smith, H. (2021). “Clinical AI: opacity, accountability, responsibility and liability”. *Ai & Society*. 36(2): 535–545.
- Soh, H., Y. Xie, M. Chen, and D. Hsu. (2020). “Multi-Task Trust Transfer for Human-Robot Interaction”. *The International Journal of Robotics Research*. 39(2-3): 233–249. ISSN: 0278-3649, 1741-3176. DOI: [10.1177/0278364919866905](https://doi.org/10.1177/0278364919866905). URL: <http://arxiv.org/abs/1807.01866>.
- Song, X., Y. Yang, K. Choromanski, K. Caluwaerts, W. Gao, C. Finn, and J. Tan. (2020). “Rapidly Adaptable Legged Robots via Evolutionary Meta-Learning”. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 3769–3776. DOI: [10.1109/IROS45743.2020.9341571](https://doi.org/10.1109/IROS45743.2020.9341571).
- Spencer, J., S. Choudhury, M. Barnes, M. Schmittle, M. Chiang, P. Ramadge, and S. Srinivasa. (2022). “Expert Intervention Learning: An online framework for robot learning from explicit and implicit human feedback”. *Autonomous Robots*: 1–15.
- Sprechmann, P., S. Jayakumar, J. Rae, A. Pritzel, A. P. Badia, B. Uria, O. Vinyals, D. Hassabis, R. Pascanu, and C. Blundell. (2018). “Memory-based Parameter Adaptation”. In: *International Conference on Learning Representations*. URL: <https://openreview.net/forum?id=rkfOvGbcW>.
- St. Clair, A. and M. Mataric. (2015). “How robot verbal feedback can improve team performance in human-robot task collaborations”. In: *Proceedings of the tenth annual acm/ieee international conference on human-robot interaction*. 213–220.
- Stahl, B. C. and D. Wright. (2018). “Ethics and privacy in AI and big data: Implementing responsible research and innovation”. *IEEE Security & Privacy*. 16(3): 26–33.
- Stange, S., T. Hassan, F. Schröder, J. Konkol, and S. Kopp. (2022). “Self-explaining social robots: An explainable behavior generation architecture for human-robot interaction”. *Frontiers in Artificial Intelligence*. 5: 87.

- Stanton, N. A., P. M. Salmon, G. H. Walker, and M. Stanton. (2019). “Models and methods for collision analysis: A comparison study based on the Uber collision with a pedestrian”. *Safety Science*. 120: 117–128. ISSN: 0925-7535. DOI: <https://doi.org/10.1016/j.ssci.2019.06.008>. URL: <https://www.sciencedirect.com/science/article/pii/S0925753519303704>.
- State of California - Department of Justice - Office of the Attorney General. (2018). “California Consumer Privacy Act (CCPA)”. <https://oag.ca.gov/privacy/ccpa>.
- Stavridis, S., D. Papageorgiou, and Z. Doulgeri. (2022). “Kinesthetic teaching of bi-manual tasks with known relative constraints”. In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 11796–11801. DOI: [10.1109/IROS47612.2022.9981196](https://doi.org/10.1109/IROS47612.2022.9981196).
- Steels, L. and M. Hild. (2012). *Language grounding in robots*. Springer Science & Business Media.
- Steinfeld, A., T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, and M. Goodrich. (2006). “Common metrics for human-robot interaction”. In: *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. 33–40.
- Stout, R. J., J. A. Cannon-Bowers, and E. Salas. (2017). “The role of shared mental models in developing team situational awareness: Implications for training”. In: *Situational awareness*. Routledge. 287–318.
- Straten, C. L. v., J. Peter, R. Kühne, and A. Barco. (2020). “Transparency about a Robot’s Lack of Human Psychological Capacities: Effects on Child-Robot Perception and Relationship Formation”. *J. Hum.-Robot Interact.* 9(2). DOI: [10.1145/3365668](https://doi.org/10.1145/3365668). URL: <https://doi.org/10.1145/3365668>.
- Struckmeier, O., M. Racca, and V. Kyrki. (2019). “Autonomous Generation of Robust and Focused Explanations for Robot Policies”. In: *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 1–8. DOI: [10.1109/RO-MAN46459.2019.8956323](https://doi.org/10.1109/RO-MAN46459.2019.8956323).

- Su, Z., O. Kroemer, G. Loeb, G. Sukhatme, and S. Schaal. (2016). “Learning to Switch Between Sensorimotor Primitives Using Multimodal Haptic Signals”. In: vol. 9825. 170–182. ISBN: 978-3-319-43487-2. DOI: [10.1007/978-3-319-43488-9_16](https://doi.org/10.1007/978-3-319-43488-9_16).
- Suay, H. B. and S. Chernova. (2011). “Effect of human guidance and state space size on Interactive Reinforcement Learning”. In: *2011 RO-MAN*. 1–6. DOI: [10.1109/ROMAN.2011.6005223](https://doi.org/10.1109/ROMAN.2011.6005223).
- Subramanian, K., C. L. Isbell Jr, and A. L. Thomaz. (2016). “Exploration from demonstration for interactive reinforcement learning”. In: *Proceedings of the 2016 international conference on autonomous agents & multiagent systems*. 447–456.
- Sugden, R. (2008). “Nash equilibrium, team reasoning and cognitive hierarchy theory”. *Acta Psychologica*. 128(2): 402–404.
- Sumers, T. R., M. K. Ho, R. D. Hawkins, K. Narasimhan, and T. L. Griffiths. (2021). “Learning rewards from linguistic feedback”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. No. 7. 6002–6010.
- Sun, Y., W. Wang, Y. Chen, and Y. Jia. (2020). “Learn how to assist humans through human teaching and robot learning in human–robot collaborative assembly”. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*. 52(2): 728–738.
- Sung, J., R. E. Grinter, and H. I. Christensen. (2010). “Domestic robot ecology: An initial framework to unpack long-term acceptance of robots at home”. *International Journal of Social Robotics*. 2: 417–429.
- Sutton, R. S. and A. G. Barto. (2018). *Reinforcement learning: An introduction*. MIT press.
- Tabrez, A., S. Agrawal, and B. Hayes. (2019a). “Explanation-Based Reward Coaching to Improve Human Performance via Reinforcement Learning”. In: *Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI '19)*. IEEE Press. Daegu, Republic of Korea. 249–257.
- Tabrez, A., S. Agrawal, and B. Hayes. (2019b). “Explanation-based reward coaching to improve human performance via reinforcement learning”. In: *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 249–257.

- Tabrez, A. and B. Hayes. (2019). “Improving Human-Robot Interaction Through Explainable Reinforcement Learning”. In: *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 751–753. DOI: [10.1109/HRI.2019.8673198](https://doi.org/10.1109/HRI.2019.8673198).
- Tabrez, A. and B. Hayes. “Mediating Trust and Influence in Human-Robot Interaction via Explainable AI”. en.
- Tabrez, A., M. B. Luebbers, and B. Hayes. (2022). “Descriptive and Prescriptive Visual Guidance to Improve Shared Situational Awareness in Human-Robot Teaming”. In: *2022 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2022)*. 1–8.
- Tai, Y.-L., Y. C. Chiu, Y.-W. Chao, and Y.-T. Chen. (2023). “SCONE: A Food Scooping Robot Learning Framework with Active Perception”. In: *7th Annual Conference on Robot Learning*.
- Tambwekar, P., L. Dodeja, N. Vaska, W. Xu, and M. Gombolay. (2023a). “A Computational Interface to Translate Strategic Intent from Unstructured Language in a Low-Data Setting”. In: *Findings of the Association for Computational Linguistics: EMNLP 2023*. Ed. by H. Bouamor, J. Pino, and K. Bali. Singapore: Association for Computational Linguistics. 12801–12819. DOI: [10.18653/v1/2023.findings-emnlp.853](https://doi.org/10.18653/v1/2023.findings-emnlp.853). URL: <https://aclanthology.org/2023.findings-emnlp.853>.
- Tambwekar, P. and M. Gombolay. (2023). “Towards Reconciling Usability and Usefulness of Explainable AI Methodologies”. *arXiv preprint arXiv:2301.05347*.
- Tambwekar, P., A. Silva, and M. Gombolay. (2023b). “The Design and Preliminary Results of a User Study Measuring Diverse Explainability Preferences”.
- Tapus, A., C. Tapus, and M. Mataric. (2009). “The role of physical embodiment of a therapist robot for individuals with cognitive impairments”. In: *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE. 103–107.
- Tartaglione, E., C. A. Barbano, and M. Grangetto. (2021). “End: Entangling and disentangling deep representations for bias correction”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 13508–13517.

- Tassa, Y., Y. Doron, A. Muldal, T. Erez, Y. Li, D. d. L. Casas, D. Budden, A. Abdolmaleki, J. Merel, A. Lefrancq, *et al.* (2018). “Deepmind control suite”. *arXiv preprint arXiv:1801.00690*.
- Tavares, A., J. L. Silva, and R. Ventura. (2023). “Physiologically Attentive User Interface for Improved Robot Teleoperation”. In: *Proceedings of the 28th International Conference on Intelligent User Interfaces. IUI '23*. Sydney, NSW, Australia: Association for Computing Machinery. 776–789. ISBN: 9798400701061. DOI: [10.1145/3581641.3584084](https://doi.org/10.1145/3581641.3584084). URL: <https://doi.org/10.1145/3581641.3584084>.
- Taylor, M. E., H. B. Suay, and S. Chernova. (2011). “Integrating Reinforcement Learning with Human Demonstrations of Varying Ability”. In: *The 10th International Conference on Autonomous Agents and Multiagent Systems - Volume 2. AAMAS '11*. Taipei, Taiwan: International Foundation for Autonomous Agents and Multiagent Systems. 617–624. ISBN: 0982657161.
- Tellex, S., N. Gopalan, H. Kress-Gazit, and C. Matuszek. (2020a). “Robots That Use Language”. *Annual Review of Control, Robotics, and Autonomous Systems*. 3(1): 25–55. DOI: [10.1146/annurev-control-101119-071628](https://doi.org/10.1146/annurev-control-101119-071628). eprint: <https://doi.org/10.1146/annurev-control-101119-071628>. URL: <https://doi.org/10.1146/annurev-control-101119-071628>.
- Tellex, S., N. Gopalan, H. Kress-Gazit, and C. Matuszek. (2020b). “Robots that use language”. *Annual Review of Control, Robotics, and Autonomous Systems*. 3: 25–55.
- Templeman, R., M. Korayem, D. J. Crandall, and A. Kapadia. (2014). “PlaceAavoider: Steering First-Person Cameras away from Sensitive Spaces.” In: *NDSS*. Vol. 14. Citeseer. 23–26.
- Teo, G., L. Reinerman-Jones, G. Matthews, J. Szalma, F. Jentsch, and P. Hancock. (2018). “Enhancing the effectiveness of human-robot teaming with a closed-loop system”. *Applied ergonomics*. 67: 91–103.
- Terrell, A. and B. Mutlu. (2012). “A regression-based approach to modeling addressee backchannels”. In: *Proceedings of the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue*. 280–289.
- “Tesla Deaths”. (2024). <https://www.tesladeaths.com/>.

- Teso, S. and K. Kersting. (2018). " " Why Should I Trust Interactive Learners?" Explaining Interactive Queries of Classifiers to Users". *arXiv preprint arXiv:1805.08578*.
- Teso, S. and K. Kersting. (2019). "Explanatory interactive machine learning". In: *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. 239–245.
- Thomaz, A. L. and C. Breazeal. (2008a). "Teachable robots: Understanding human teaching behavior to build more effective robot learners". *Artificial Intelligence*. 172(6-7): 716–737.
- Thomaz, A. L. and C. Breazeal. (2008b). "Teachable robots: Understanding human teaching behavior to build more effective robot learners". *Artificial Intelligence*. 172(6): 716–737. ISSN: 0004-3702. DOI: <https://doi.org/10.1016/j.artint.2007.09.009>. URL: <https://www.sciencedirect.com/science/article/pii/S000437020700135X>.
- Thomaz, A. L., G. Hoffman, and C. Breazeal. (2006). "Reinforcement Learning with Human Teachers: Understanding How People Want to Teach Robots". In: *ROMAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication*. 352–357. DOI: [10.1109/ROMAN.2006.314459](https://doi.org/10.1109/ROMAN.2006.314459).
- Tian, R., L. Sun, A. Bajcsy, M. Tomizuka, and A. D. Dragan. (2022). "Safety Assurances for Human-Robot Interaction via Confidence-aware Game-theoretic Human Models". en. In: *2022 International Conference on Robotics and Automation (ICRA)*. Philadelphia, PA, USA: IEEE. 11229–11235. ISBN: 978-1-72819-681-7. DOI: [10.1109/ICRA46639.2022.9812048](https://doi.org/10.1109/ICRA46639.2022.9812048). URL: <https://ieeexplore.ieee.org/document/9812048/>.
- Tidemann, A. and P. Öztürk. (2007). "Self-organizing Multiple Models for Imitation: Teaching a Robot to Dance the YMCA". en. In: *New Trends in Applied Artificial Intelligence*. Ed. by H. G. Okuno and M. Ali. *Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer. 291–302. ISBN: 978-3-540-73325-6. DOI: [10.1007/978-3-540-73325-6_29](https://doi.org/10.1007/978-3-540-73325-6_29).

- Tien, J., J. Z.-Y. He, Z. Erickson, A. Dragan, and D. S. Brown. (2023). “Causal Confusion and Reward Misidentification in Preference-Based Reward Learning”. In: *The Eleventh International Conference on Learning Representations*. URL: https://openreview.net/forum?id=R0Xxvr_X3ZA.
- Tojo, T., Y. Matsusaka, T. Ishii, and T. Kobayashi. (2000). “A conversational robot utilizing facial and body expressions”. In: *Smc 2000 conference proceedings. 2000 ieee international conference on systems, man and cybernetics. 'cybernetics evolving to systems, humans, organizations, and their complex interactions' (cat. no.0. Vol. 2. 858–863 vol.2*. DOI: [10.1109/ICSMC.2000.885957](https://doi.org/10.1109/ICSMC.2000.885957).
- Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Harvard University Press. ISBN: 978-0-674-00070-4. DOI: [10.2307/j.ctvj4jc](https://doi.org/10.2307/j.ctvj4jc). URL: <http://www.jstor.org/stable/j.ctvj4jc>.
- Topin, N. and M. Veloso. (2019). “Generation of Policy-Level Explanations for Reinforcement Learning”. *Proceedings of the AAAI Conference on Artificial Intelligence*. 33(01): 2514–2521. DOI: [10.1609/aaai.v33i01.33012514](https://doi.org/10.1609/aaai.v33i01.33012514). URL: <https://ojs.aaai.org/index.php/AAAI/article/view/4097>.
- Triastcyn, A. and B. Faltings. (2019). “Federated learning with bayesian differential privacy”. In: *2019 IEEE International Conference on Big Data (Big Data)*. IEEE. 2587–2596.
- Tsamados, A., N. Aggarwal, J. Cowls, J. Morley, H. Roberts, M. Taddeo, and L. Floridi. (2021). “The ethics of algorithms: key problems and solutions”. *Ethics, Governance, and Policies in Artificial Intelligence*: 97–123.
- Tseng, S.-H., F.-C. Liu, and L.-C. Fu. (2018). “Active Learning on Service Providing Model: Adjustment of Robot Behaviors Through Human Feedback”. en. *IEEE Transactions on Cognitive and Developmental Systems*. 10(3): 701–711. ISSN: 2379-8920, 2379-8939. DOI: [10.1109/TCDS.2017.2775621](https://doi.org/10.1109/TCDS.2017.2775621). URL: <https://ieeexplore.ieee.org/document/8119555/> (accessed on 11/10/2023).
- Tulli, S., M. Couto, M. Vasco, E. Yadollahi, F. Melo, and A. Paiva. (2020). “Explainable agency by revealing suboptimality in child-robot learning scenarios”. In: *International Conference on Social Robotics*. Springer. 23–35.

- Urakami, J. and K. Seaborn. (2023). “Nonverbal Cues in Human–Robot Interaction: A Communication Studies Perspective”. *ACM Transactions on Human-Robot Interaction*. 12(2): 1–21.
- Valdivia, A. A., R. Shailly, N. Seth, F. Fuentes, D. P. Losey, and L. H. Blumenschein. (2022). “Wrapped Haptic Display for Communicating Physical Robot Learning”. In: *2022 IEEE 5th International Conference on Soft Robotics (RoboSoft)*. 823–830. DOI: [10.1109/RoboSoft54090.2022.9762210](https://doi.org/10.1109/RoboSoft54090.2022.9762210).
- Van Waveren, S., C. Pek, I. Leite, J. Tumova, and D. Kragic. (2023a). “Generating Scenarios from High-Level Specifications for Object Rearrangement Tasks”. In: *IEEE Int. Conf. on Intelligent Robots and Systems*.
- Van Waveren, S., C. Pek, J. Tumova, and I. Leite. (2022). “Correct Me If I’m Wrong: Using Non-Experts to Repair Reinforcement Learning Policies”. In: *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 493–501. DOI: [10.1109/HRI53351.2022.9889604](https://doi.org/10.1109/HRI53351.2022.9889604).
- Van Waveren, S., R. Rudling, I. Leite, P. Jensfelt, and C. Pek. (2023b). “Increasing perceived safety in motion planning for human-drone interaction”. In: *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 446–455.
- Vinanzi, S., M. Patacchiola, A. Chella, and A. Cangelosi. (2019). “Would a robot trust you? Developmental robotics model of trust and theory of mind”. *Philosophical Transactions of the Royal Society B*. 374(1771): 20180032.
- Virgolin, M., M. Bellone, K. Wolff, and M. Wahde. (2021). “A Mobile Interactive Robot for Social Distancing in Hospitals”. In: *2021 Fifth IEEE International Conference on Robotic Computing (IRC)*. 87–91. DOI: [10.1109/IRC52146.2021.00020](https://doi.org/10.1109/IRC52146.2021.00020).

- Vitale, J., M. Tonkin, S. Herse, S. Ojha, J. Clark, M.-A. Williams, X. Wang, and W. Judge. (2018). “Be More Transparent and Users Will Like You: A Robot Privacy and User Experience Design Experiment”. In: *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction. HRI '18*. Chicago, IL, USA: Association for Computing Machinery. 379–387. ISBN: 9781450349536. DOI: [10.1145/3171221.3171269](https://doi.org/10.1145/3171221.3171269). URL: <https://doi.org/10.1145/3171221.3171269>.
- Wallach, W. and S. Vallor. (2020). “Moral machines”. *Ethics of Artificial Intelligence*. Oxford University Press: 383–412.
- Wallkötter, S., S. Tulli, G. Castellano, A. Paiva, and M. Chetouani. (2021). “Explainable Embodied Agents Through Social Cues: A Review”. *J. Hum.-Robot Interact.* 10(3). DOI: [10.1145/3457188](https://doi.org/10.1145/3457188). URL: <https://doi.org/10.1145/3457188>.
- Walster, E., E. Berscheid, and G. Walster. (1973). “New Directions in Equity Research”. *Journal of Personality and Social Psychology*. 25(Feb.): 151–176. DOI: [10.1037/h0033967](https://doi.org/10.1037/h0033967).
- Wang, K., A. Variengien, A. Conmy, B. Shlegeris, and J. Steinhardt. (2022a). “Interpretability in the wild: a circuit for indirect object identification in gpt-2 small”. *arXiv preprint arXiv:2211.00593*.
- Wang, L., A. D. Ames, and M. Egerstedt. (2017). “Safety barrier certificates for collisions-free multirobot systems”. *IEEE Transactions on Robotics*. 33(3): 661–674.
- Wang, N., D. Pynadath, and S. Hill. (2016a). “The Impact of POMDP-Generated Explanations on Trust and Performance in Human-Robot Teams”. In: *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems (AAMAS '16)*. International Foundation for Autonomous Agents and Multiagent Systems. Sao Paulo, Brazil. 997–1005.
- Wang, N., D. Pynadath, and S. Hill. (2016b). “Trust Calibration within a Human-Robot Team: Comparing Automatically Generated Explanations”. In: *Proceedings of the Eleventh ACM/IEEE International Conference on Human Robot Interaction (HRI '16)*. IEEE Press. Christchurch, New Zealand. 109–116.

- Wang, T.-H., A. Amini, W. Schwarting, I. Gilitschenski, S. Karaman, and D. Rus. (2022b). “Learning interactive driving policies via data-driven simulation”. In: *2022 International Conference on Robotics and Automation (ICRA)*. IEEE. 7745–7752.
- Wang, W., Y. Chen, R. Li, and Y. Jia. (2019a). “Learning and comfort in human–robot interaction: A review”. *Applied Sciences*. 9(23): 5152.
- Wang, W., Y. Chen, R. Li, and Y. Jia. (2019b). “Learning and Comfort in Human–Robot Interaction: A Review”. *Applied Sciences*. 9(23). ISSN: 2076-3417. DOI: [10.3390/app9235152](https://doi.org/10.3390/app9235152). URL: <https://www.mdpi.com/2076-3417/9/23/5152>.
- Wang, W., R. Li, Y. Chen, Y. Sun, and Y. Jia. (2021). “Predicting human intentions in human–robot hand-over tasks through multi-modal learning”. *IEEE Transactions on Automation Science and Engineering*. 19(3): 2339–2353.
- Wang, W. and K. Siau. (2018). “Ethical and moral issues with AI”.
- Wang, Y., C. C. Beltran-Hernandez, W. Wan, and K. Harada. (2022c). “An Adaptive Imitation Learning Framework for Robotic Complex Contact-Rich Insertion Tasks”. *Frontiers in Robotics and AI*. 8. ISSN: 2296-9144. URL: <https://www.frontiersin.org/articles/10.3389/frobt.2021.777363>.
- Wang, Y., N. Figueroa, S. Li, A. Shah, and J. Shah. (2022d). “Temporal Logic Imitation: Learning Plan-Satisficing Motion Policies from Demonstrations”. *arXiv preprint arXiv:2206.04632*.
- Wang, Y.-J., B. Zhang, J. Chen, and K. Sreenath. (2023). “Prompt a Robot to Walk with Large Language Models”. arXiv: [2309.09969](https://arxiv.org/abs/2309.09969) [cs.LG].
- Warnell, G., N. Waytowich, V. Lawhern, and P. Stone. (2018). “Deep TAMER: Interactive Agent Shaping in High-Dimensional State Spaces”. In: *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence*. AAAI’18/IAAI’18/EAAI’18. New Orleans, Louisiana, USA: AAAI Press. ISBN: 978-1-57735-800-8.

- Wasala, A., D. Byrne, P. Miesbauer, J. O'Hanlon, P. Heraty, and P. Barry. (2020). "Trajectory based lateral control: A Reinforcement Learning case study". *Engineering Applications of Artificial Intelligence*. 94(Sept.): 103799. ISSN: 0952-1976. DOI: [10.1016/j.engappai.2020.103799](https://doi.org/10.1016/j.engappai.2020.103799). URL: <https://www.sciencedirect.com/science/article/pii/S0952197620301858>.
- Wei, K., J. Li, M. Ding, C. Ma, H. H. Yang, F. Farokhi, S. Jin, T. Q. Quek, and H. V. Poor. (2020). "Federated learning with differential privacy: Algorithms and performance analysis". *IEEE Transactions on Information Forensics and Security*. 15: 3454–3469.
- Wei, K., J. Li, C. Ma, M. Ding, W. Chen, J. Wu, M. Tao, and H. V. Poor. (2023). "Personalized federated learning with differential privacy and convergence guarantee". *IEEE Transactions on Information Forensics and Security*.
- Wengefeld, T., D. Höchemer, B. Lewandowski, M. Köhler, M. Beer, and H.-M. Gross. (2020). "A Laser Projection System for Robot Intention Communication and Human Robot Interaction". In: *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 259–265. DOI: [10.1109/RO-MAN47096.2020.9223517](https://doi.org/10.1109/RO-MAN47096.2020.9223517).
- Wickens, C. D. (2002). "Situation awareness and workload in aviation". *Current directions in psychological science*. 11(4): 128–133.
- Wilson, B., J. Hoffman, and J. Morgenstern. (2019). "Predictive inequity in object detection". *arXiv preprint arXiv:1902.11097*.
- Winfield, A. F., K. Winkle, H. Webb, U. Lyngs, M. Jirotko, and C. Macrae. (2021). "Robot accident investigation: a case study in responsible robotics". *Software engineering for robotics*: 165–187.
- Wirth, C., R. Akrou, G. Neumann, and J. Fürnkranz. (2017). "A Survey of Preference-Based Reinforcement Learning Methods". *Journal of Machine Learning Research*. 18(136): 1–46. URL: <http://jmlr.org/papers/v18/16-634.html>.
- Wongpiromsarn, T., A. Ulusoy, C. Belta, E. Frazzoli, and D. Rus. (2012). "Incremental temporal logic synthesis of control policies for robots interacting with dynamic agents". In: *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 229–236.

- Wonsick, M. and T. Padir. (2020). “A Systematic Review of Virtual Reality Interfaces for Controlling and Interacting with Robots”. *Applied Sciences*. 10(24). ISSN: 2076-3417. DOI: [10.3390/app10249051](https://doi.org/10.3390/app10249051). URL: <https://www.mdpi.com/2076-3417/10/24/9051>.
- Wortham, R. H., A. Theodorou, and J. J. Bryson. (2017). “Improving robot transparency: Real-time visualisation of robot AI substantially improves understanding in naive observers”. In: *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 1424–1431. DOI: [10.1109/ROMAN.2017.8172491](https://doi.org/10.1109/ROMAN.2017.8172491).
- Wu, H., H. Li, X. Fang, and X. Luo. (2022). “A survey on teaching workplace skills to construction robots”. *Expert Systems with Applications*. 205: 117658.
- Wu, Q., C.-J. Wu, Y. Zhu, and J. Joo. (2021a). “Communicative learning with natural gestures for embodied navigation agents with human-in-the-scene”. In: *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 4095–4102.
- Wu, R., S. Kortik, and C. H. Santos. (2021b). “Automated Behavior Tree Error Recovery Framework for Robotic Systems”. In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. 6898–6904. DOI: [10.1109/ICRA48506.2021.9561002](https://doi.org/10.1109/ICRA48506.2021.9561002).
- Wulfmeier, M., P. Ondruska, and I. Posner. (2015). “Maximum entropy deep inverse reinforcement learning”. *arXiv preprint arXiv:1507.04888*.
- Xia, F., W. B. Shen, C. Li, P. Kasimbeg, M. E. Tchammi, A. Tshhev, R. Martin-Martin, and S. Savarese. (2020). “Interactive gibbon benchmark: A benchmark for interactive navigation in cluttered environments”. *IEEE Robotics and Automation Letters*. 5(2): 713–720.
- Xiao, X., J. Dufek, and R. R. Murphy. (2020). “Robot risk-awareness by formal risk reasoning and planning”. *IEEE Robotics and Automation Letters*. 5(2): 2856–2863.
- Xu, A. and G. Dudek. (2015). “OPTIMo: Online Probabilistic Trust Inference Model for Asymmetric Human-Robot Collaborations”. en. In: *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. Portland Oregon USA: ACM. 221–228. ISBN: 978-1-4503-2883-8. DOI: [10.1145/2696454.2696492](https://doi.org/10.1145/2696454.2696492). URL: <https://dl.acm.org/doi/10.1145/2696454.2696492>.

- Xu, P. (2017). “Gesture-based Human-robot Interaction for Field Programmable Autonomous Underwater Robots”. *arXiv preprint arXiv:1709.08945*.
- Xue, T., W. Wang, J. Ma, W. Liu, Z. Pan, and M. Han. (2020). “Progress and prospects of multimodal fusion methods in physical human–robot interaction: A review”. *IEEE Sensors Journal*. 20(18): 10355–10370.
- Yan, Z., L. Sun, T. Krajník, T. Duckett, and N. Bellotto. (2023). “Towards Long-term Autonomy: A Perspective from Robot Learning”. *arXiv: 2212.12798 [cs.R0]*.
- Yang, C.-Y., M.-J. Lu, S.-H. Tseng, and L.-C. Fu. (2017). “A companion robot for daily care of elders based on homeostasis”. en. In: *2017 56th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*. Kanazawa: IEEE. 1401–1406. ISBN: 978-4-907764-57-9. DOI: [10.23919/SICE.2017.8105748](https://doi.org/10.23919/SICE.2017.8105748). URL: <http://ieeexplore.ieee.org/document/8105748/> (accessed on 11/03/2023).
- Yang, G., C. Belta, and R. Tron. (2020a). “Continuous-time signal temporal logic planning with control barrier functions”. In: *2020 American Control Conference (ACC)*. IEEE. 4612–4618.
- Yang, J., J. A. Barragan, J. M. Farrow, C. P. Sundaram, J. P. Wachs, and D. Yu. (2022a). “An Adaptive Human-Robotic Interaction Architecture for Augmenting Surgery Performance Using Real-Time Workload Sensing—Demonstration of a Semi-autonomous Suction Tool”. *Human Factors*: 00187208221129940.
- Yang, T.-Y., M. Y. Hu, Y. Chow, P. J. Ramadge, and K. Narasimhan. (2021). “Safe reinforcement learning with natural language constraints”. *Advances in Neural Information Processing Systems*. 34: 13794–13808.
- Yang, W., C. Paxton, M. Cakmak, and D. Fox. (2020b). “Human grasp classification for reactive human-to-robot handovers”. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 11123–11130.
- Yang, Y., L. Chen, and M. Gombolay. (2022b). “Safe Inverse Reinforcement Learning via Control Barrier Function”. *arXiv preprint arXiv:2212.02753*.

- Yasar, M. S., M. M. Islam, and T. Iqbal. (2023). “IMPRINT: Interactional Dynamics-Aware Motion Prediction in Teams Using Multi-modal Context”. *ACM Transactions on Human-Robot Interaction*. URL: <https://doi.org/10.1145/3626954>.
- Yigitbas, E., K. Karakaya, I. Jovanovikj, and G. Engels. (2021). “Enhancing Human-in-the-Loop Adaptive Systems through Digital Twins and VR Interfaces”. In: *2021 International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS)*. 30–40. DOI: [10.1109/SEAMS51251.2021.00015](https://doi.org/10.1109/SEAMS51251.2021.00015).
- Yin, G., J. Li, X. Jin, C. Bian, and N. Chen. (2015). “Integration of motion planning and model-predictive-control-based control system for autonomous electric vehicles”. *Transport*. 30(3): 353–360. ISSN: 1648-4142. DOI: [10.3846/16484142.2015.1089322](https://doi.org/10.3846/16484142.2015.1089322). URL: <https://doi.org/10.3846/16484142.2015.1089322>.
- Yonezawa, T., H. Yamazoe, and S. Abe. (2013). “Physical contact using haptic and gestural expressions for ubiquitous partner robot”. In: *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 5680–5685. DOI: [10.1109/IROS.2013.6697179](https://doi.org/10.1109/IROS.2013.6697179).
- Yousuf, M. A., Y. Kobayashi, Y. Kuno, K. Yamazaki, and A. Yamazaki. (2019). “Social interaction with visitors: mobile guide robots capable of offering a museum tour”. *IEEJ Transactions on Electrical and Electronic Engineering*. 14(12): 1823–1835.
- Yu, C.-C. and C.-C. Wang. (2013). “Interactive Learning from Demonstration with a Multilevel Mechanism for Collision-Free Navigation in Dynamic Environments”. In: *2013 Conference on Technologies and Applications of Artificial Intelligence*. 240–245. DOI: [10.1109/TAAI.2013.55](https://doi.org/10.1109/TAAI.2013.55).
- Yu, T., D. Quillen, Z. He, R. Julian, K. Hausman, C. Finn, and S. Levine. (2020). “Meta-world: A benchmark and evaluation for multi-task and meta reinforcement learning”. In: *Conference on robot learning*. PMLR. 1094–1100.
- Yu, W., N. Gileadi, C. Fu, S. Kirmani, K.-H. Lee, M. G. Arenas, H.-T. L. Chiang, T. Erez, L. Hasenclever, J. Humplik, B. Ichter, T. Xiao, P. Xu, A. Zeng, T. Zhang, N. Heess, D. Sadigh, J. Tan, Y. Tassa, and F. Xia. (2023). “Language to Rewards for Robotic Skill Synthesis”. arXiv: [2306.08647 \[cs.R0\]](https://arxiv.org/abs/2306.08647).

- Yu, W., G. Turk, and C. K. Liu. (2018). “Learning Symmetric and Low-Energy Locomotion”. *ACM Trans. Graph.* 37(4). ISSN: 0730-0301. DOI: [10.1145/3197517.3201397](https://doi.org/10.1145/3197517.3201397). URL: <https://doi.org/10.1145/3197517.3201397>.
- Yuan, L., X. Gao, Z. Zheng, M. Edmonds, Y. N. Wu, F. Rossano, H. Lu, Y. Zhu, and S.-C. Zhu. (2022). “In situ bidirectional human-robot value alignment”. *Science robotics*. 7(68): eabm4183.
- Yurtsever, E., J. Lambert, A. Carballo, and K. Takeda. (2020). “A Survey of Autonomous Driving: Common Practices and Emerging Technologies”. en. *IEEE Access*. 8: 58443–58469. ISSN: 2169-3536. DOI: [10.1109/ACCESS.2020.2983149](https://doi.org/10.1109/ACCESS.2020.2983149). URL: <http://arxiv.org/abs/1906.05113>.
- Zacharaki, A., I. Kostavelis, A. Gasteratos, and I. Dokas. (2020). “Safety bounds in human robot interaction: A survey”. *Safety science*. 127: 104667.
- Zaidi, Z., D. Martin, N. Belles, V. Zakharov, A. Krishna, K. M. Lee, P. Wagstaff, S. Naik, M. Sklar, S. Choi, *et al.* (2023). “Athletic mobile manipulator system for robotic wheelchair tennis”. *IEEE Robotics and Automation Letters*. 8(4): 2245–2252.
- Zajec, P., J. M. Rožanec, I. Novalija, B. Fortuna, D. Mladenčić, and K. Kenda. (2021). “Towards Active Learning Based Smart Assistant for Manufacturing”. en. In: *Advances in Production Management Systems. Artificial Intelligence for Sustainable and Resilient Production Systems*. Ed. by A. Dolgui, A. Bernard, D. Lemoine, G. von Cieminski, and D. Romero. *IFIP Advances in Information and Communication Technology*. Cham: Springer International Publishing. 295–302. ISBN: 978-3-030-85910-7. DOI: [10.1007/978-3-030-85910-7_31](https://doi.org/10.1007/978-3-030-85910-7_31).
- Zakershahra, M., Z. Gong, N. Sadassivam, and Y. Zhang. (2020). “Online Explanation Generation for Planning Tasks in Human-Robot Teaming”. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 6304–6310. DOI: [10.1109/IROS45743.2020.9341792](https://doi.org/10.1109/IROS45743.2020.9341792).

- Zakershahrak, M., S. R. Marpally, A. Sharma, Z. Gong, and Y. Zhang. (2021). “Order Matters: Generating Progressive Explanations for Planning Tasks in Human-Robot Teaming”. In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. 3751–3757. DOI: [10.1109/ICRA48506.2021.9561762](https://doi.org/10.1109/ICRA48506.2021.9561762).
- Zeng, F., W. Gan, Y. Wang, N. Liu, and P. S. Yu. (2023). “Large Language Models for Robotics: A Survey”. arXiv: [2311.07226 \[cs.R0\]](https://arxiv.org/abs/2311.07226).
- Zeng, W., W. Luo, S. Suo, A. Sadat, B. Yang, S. Casas, and R. Urtasun. (2019). “End-To-End Interpretable Neural Motion Planner”. en. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, CA, USA: IEEE. 8652–8661. ISBN: 978-1-72813-293-8. DOI: [10.1109/CVPR.2019.00886](https://doi.org/10.1109/CVPR.2019.00886). URL: <https://ieeexplore.ieee.org/document/8954347/>.
- Zhang, B. H., B. Lemoine, and M. Mitchell. (2018). “Mitigating unwanted biases with adversarial learning”. In: *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*. 335–340.
- Zhang, G. Q., X. Li, R. Boca, J. Newkirk, B. Zhang, T. A. Fuhlbrigge, H. K. Feng, and N. J. Hunt. (2014). “Use of industrial robots in additive manufacturing—a survey and feasibility study”. In: *ISR/Robotik 2014; 41st International Symposium on Robotics*. VDE. 1–6.
- Zhang, R., D. Bansal, Y. Hao, A. Hiranaka, J. Gao, C. Wang, R. Martin-Martin, L. Fei-Fei, and J. Wu. (2023). “A Dual Representation Framework for Robot Learning with Human Guidance”. In: *Conference on Robot Learning*. PMLR. 738–750.
- Zhang, S., Y. Chen, J. Zhang, and Y. Jia. (2020). “Real-time adaptive assembly scheduling in human-multi-robot collaboration according to human capability”. In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 3860–3866.
- Zhao, Q. A. and J. T. Stasko. (1998). “Evaluating image filtering based techniques in media space applications”. In: *Proceedings of the 1998 ACM conference on Computer supported cooperative work*. 11–18.
- Zhao, W. X., K. Zhou, J. Li, T. Tang, X. Wang, Y. Hou, Y. Min, B. Zhang, J. Zhang, Z. Dong, *et al.* (2023). “A survey of large language models”. *arXiv preprint arXiv:2303.18223*.

- Zhu, H. and J. Alonso-Mora. (2019). “Chance-constrained collision avoidance for mavs in dynamic environments”. *IEEE Robotics and Automation Letters*. 4(2): 776–783.
- Zhu, Y., J. Wong, A. Mandlekar, R. Martin-Martin, A. Joshi, S. Nasiriany, and Y. Zhu. (2020). “robosuite: A modular simulation framework and benchmark for robot learning”. *arXiv preprint arXiv:2009.12293*.
- Zhu, Z. and H. Hu. (2018). “Robot Learning from Demonstration in Robotic Assembly: A Survey”. en. *Robotics*. 7(2): 17. ISSN: 2218-6581. DOI: [10.3390/robotics7020017](https://doi.org/10.3390/robotics7020017). URL: <https://www.mdpi.com/2218-6581/7/2/17>.
- Ziebart, B. D. (2010). “Modeling Purposeful Adaptive Behavior with the Principle of Maximum Causal Entropy”. In: URL: <https://api.semanticscholar.org/CorpusID:11919065>.
- Ziebart, B. D., J. A. Bagnell, and A. K. Dey. (2013). “The Principle of Maximum Causal Entropy for Estimating Interacting Processes”. *IEEE Transactions on Information Theory*. 59(4): 1966–1980. DOI: [10.1109/TIT.2012.2234824](https://doi.org/10.1109/TIT.2012.2234824).
- Ziebart, B. D., A. Maas, J. A. Bagnell, and A. K. Dey. (2008). “Maximum Entropy Inverse Reinforcement Learning”. In: *Proceedings of the 23rd National Conference on Artificial Intelligence - Volume 3. AAAI’08*. Chicago, Illinois: AAAI Press. 1433–1438. ISBN: 9781577353683.
- Zimmermann, R. S., T. Klein, and W. Brendel. (2023). “Scale alone does not improve mechanistic interpretability in vision models”. *arXiv preprint arXiv:2307.05471*.
- Zucker, M., S. Joo, M. X. Grey, C. Rasmussen, E. Huang, M. Stilman, and A. Bobick. (2015). “A General-purpose System for Teleoperation of the DRC-HUBO Humanoid Robot”. *Journal of Field Robotics*. 32(3): 336–351. DOI: <https://doi.org/10.1002/rob.21570>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/rob.21570>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/rob.21570>.