Human-in-the-loop Robot Learning

Huihan Liu
Why Robotics?

World-changing Power of Automation
Why Robotics?

Intelligence is actualized in the Actions

Visual Perception

Action

Policy
Why Robotics?

Intelligence is actualized in the Actions
Why Robotics?

Intelligence is actualized in the Actions

“an agent's ability to achieve goals in a wide range of environments”
Legg, S; Hutter, M (2007). A Collection of Definitions of Intelligence

“the capacity to act purposefully, … and to deal effectively with its environment”
Wechsler, D (1944). The measurement of adult intelligence
Building Robot Autonomy
Great Advances in Robot Learning

Learning-based methods hold great promise of robot autonomy

BC-Z [Jang et al. 2021]
Implicit BC [Florence et al. 2021]
Robomimic [Mandlekar et al. 2021]
R3M [Nair et al. 2022]
Are we ready for full robot autonomy?

So many failure cases in real world…
Hard challenges for real-world deployment

unstructured real-world environments

diverse objects and scenes
Alternative Paradigm

Bringing Human in the Loop!

\( \pi : O \rightarrow A \)

- \( O \): sensory data
- \( A \): motor actions
- Environment

Diagram depicting the interaction between sensory data, actions, and the environment with human involvement.
Human-in-the-loop Robot Learning System

human agent

robot agent

human-robot collaboration

reducing effort and increasing trust

experiences

never ending data engine

improving autonomy and robustness
Human-in-the-loop Robot Learning System

Dataset → Policy Update

π₀ → π₁ → ... → πᵢ

Deployment

Never-ending Loop

Learning
Human-in-the-loop Robot Learning System

Novel Algorithms for Continual Learning

human-agent

robot-agent

human-robot collaboration

improving autonomy and robustness

experiences

never ending data engine
Iterative Learning with Human samples

IWR: Intervention Weighted Resampling

Data Collection

$\mathcal{D} \rightarrow \mathcal{D}

Training with IWR

$\mathcal{D}_R \sim \mathcal{D}_I \sim \pi_\theta

Human-in-the-Loop Imitation Learning using Remote Teleoperation [Mandlekar et al. 2021]
Learning to incorporate human cost

Learn an intervention cost function to learn expected accumulation of intervention cost

Minimize human cost for policy learning
Interpreting Human Task Specifications

Reward Sketching: Learning a reward model from human sketching of rewards

Scaling data-driven robotics with reward sketching and batch reinforcement learning [Cabi et al. 2019]
Human-in-the-loop Robot Learning System

New mechanisms for human-robot teaming

human-agent

robot-agent

human-robot collaboration

experiences

never-ending data engine

improving autonomy and robustness
Effective Human Shared Control

Teleoperation Interface: Shared human robot control, Intervene easily

Human-in-the-Loop Policy Learning with RoboTurk

Effective Human Shared Control

Teleoperation Interface: Shared human-robot control, Intervene easily

https://www.youtube.com/watch?v=wQfOKoQf4J0
Effective Human Shared Control

Crowdsourcing System: RoboTurk

RoboTurk allows several simultaneous users to teleoperate robotic arms
Actively Asking for Human Feedback

“I’m not sure what to do here - Can you help?”

Knowing when to ask for help

ThriftyDAgger: Budget-Aware Novelty and Risk Gating for Interactive Imitation Learning [Hoque et al. 2021]
Human-in-the-loop Robot Learning System

Easy and Intuitive Human Feedback

Reducing effort and increasing trust

Improving autonomy and robustness

Never ending data engine

Human-agent

Robot-agent

Human-robot collaboration experiences
“Hey Robot! Stay away from the yellow bottle.”

“Now go from under the bottle of bleach.”
“I spilled my coke. Can you help?”
Preference, Ranking, Scoring

PEBBLE: Feedback-Efficient Interactive Reinforcement Learning via Relabeling Experience and Unsupervised Pre-training [Lee et al. 2021]
Skill Preferences: Learning to Extract and Execute Skills from Human Feedback [Wang et al. 2021]
Preference, Ranking, Scoring

Correct Me if I am Wrong: Interactive Learning for Robotic Manipulation [Chisari et al. 2021]
Learning Reward Functions from Scale Feedback [Wilde et al. 2021]
Challenges

How to design better human-robot interfaces that balances precise control and easy input form?

How to enable Continuous Integration and Continuous Deployment of HITL System?

How to learn under sparsity of human data?

How to have robot augment human intelligence?