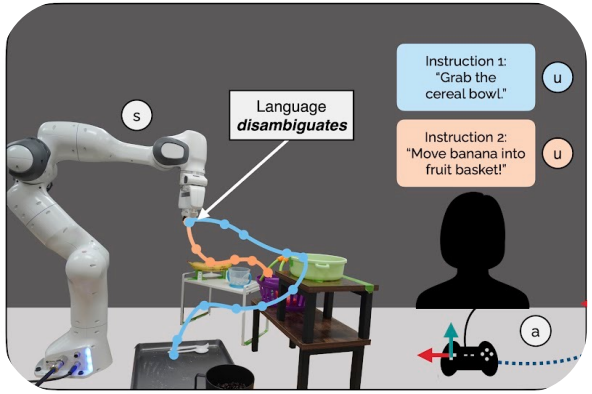
Bridging Language and Robotics for Interaction, Learning, and Teaching

Sidd Karamcheti, Dorsa Sadigh

Stanford University

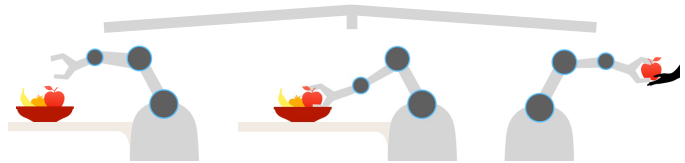
iliad
intelligent and interactive autonomous systems

Interaction



Learning

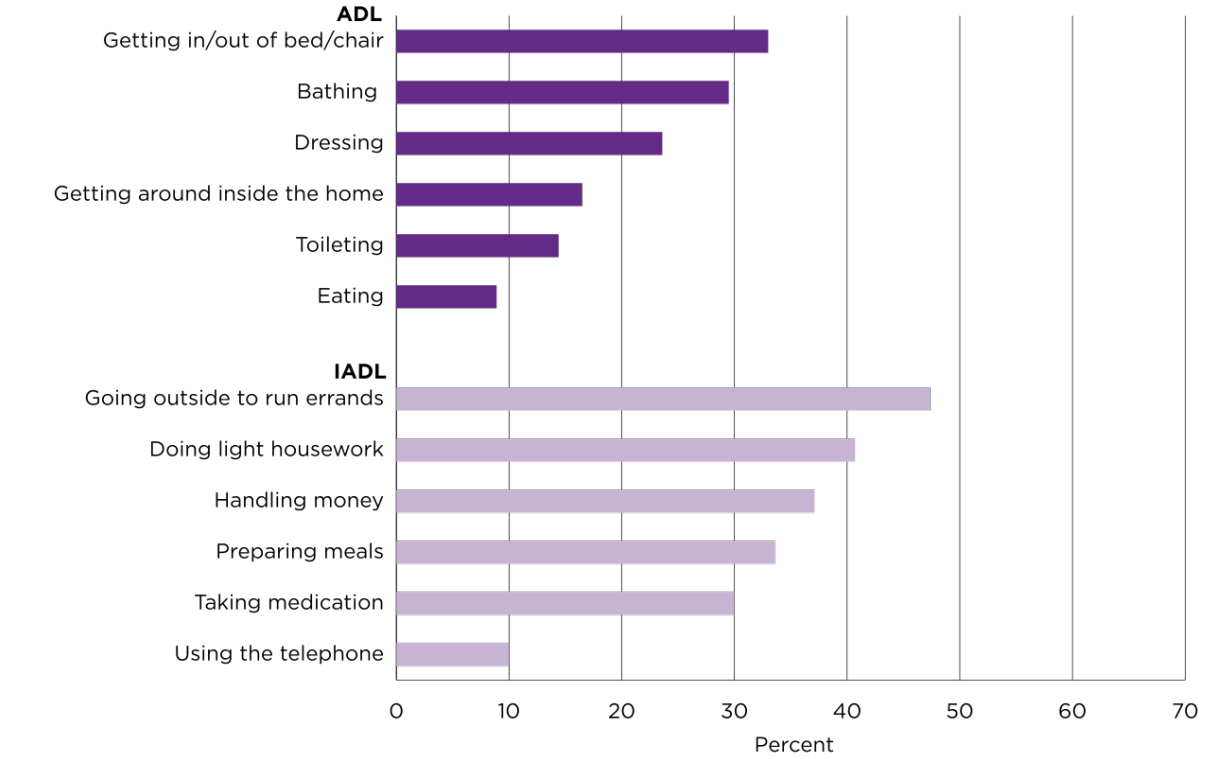
"Please get me an apple"



"Go to the table" "Pick up an apple" "Bring it to me"

Teaching

Prevalence of Difficulty Performing ADLs and IADLs in Adults 18 Years and Older With One or More Selected Symptoms That Interfere With Everyday Activities: 2014



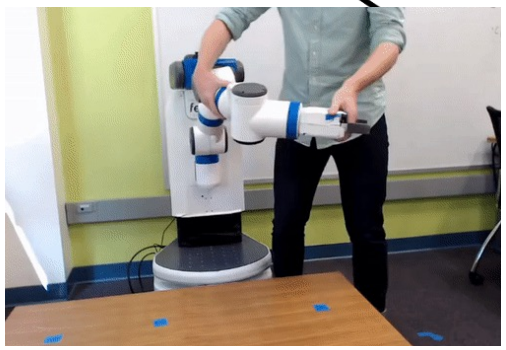
Source: U.S. Census Bureau, Social Security Administration Supplement to the 2014 Panel of the Survey of Income and Program Participation, September–November 2014.

Learning Latent Actions

Dylan Losey

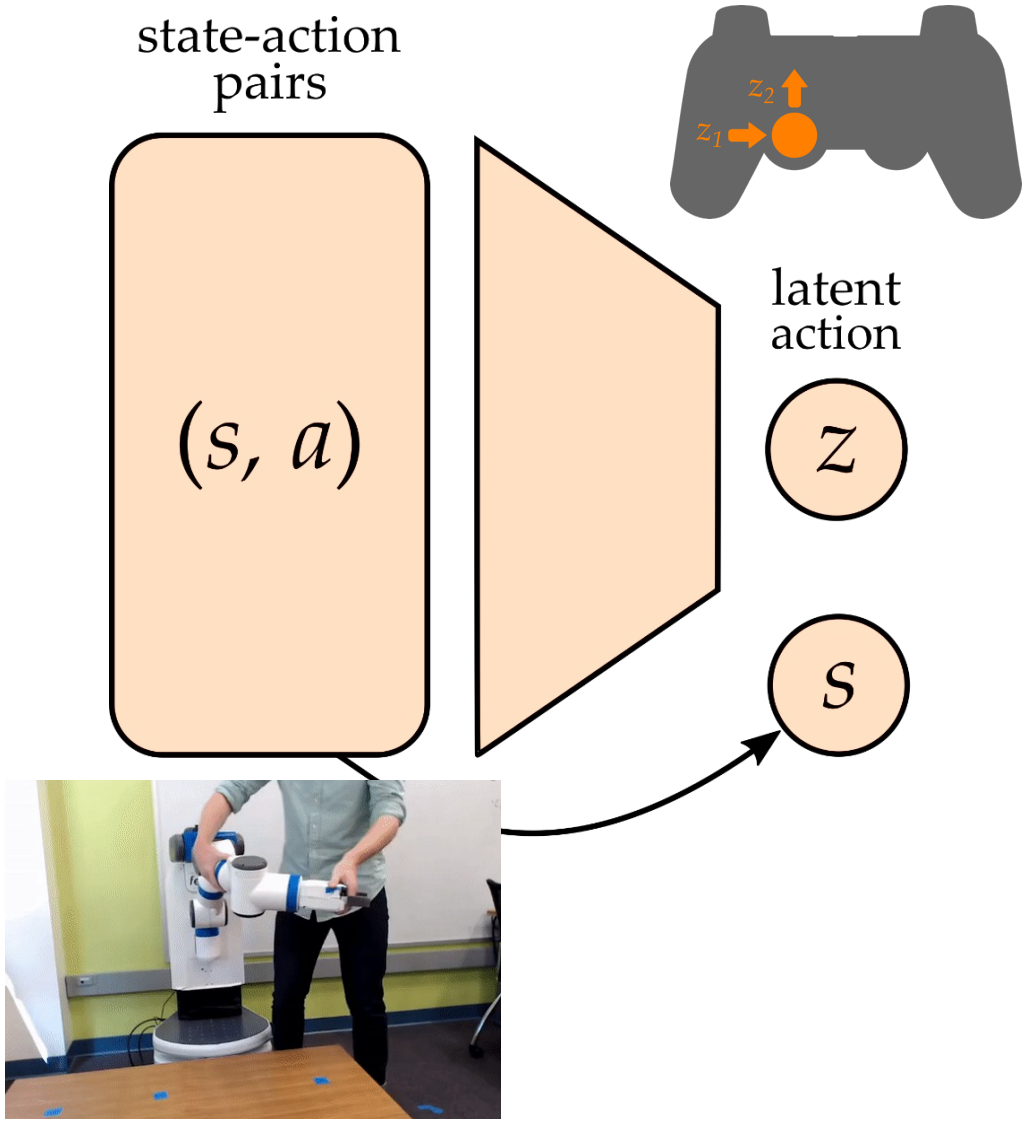
state-action pairs

(s, a)



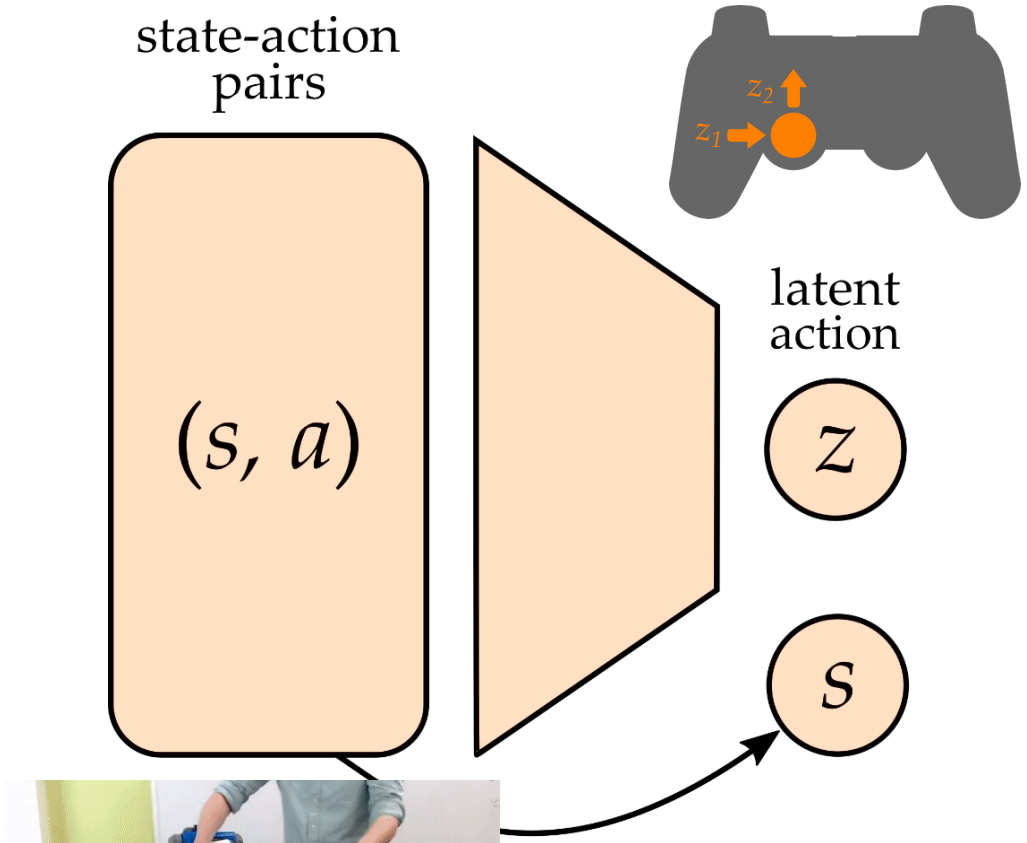
[*Losey, Srinivasan, Mandlekar, Garg, Sadigh*. ICRA 2020]

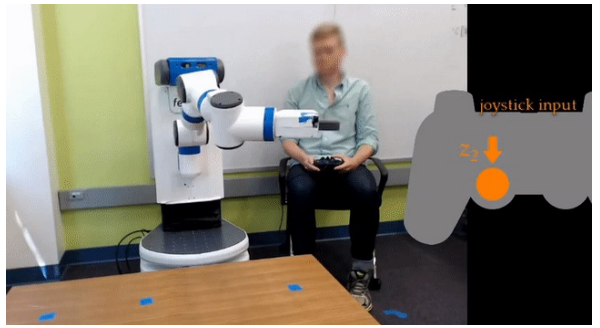
Learning Latent Actions

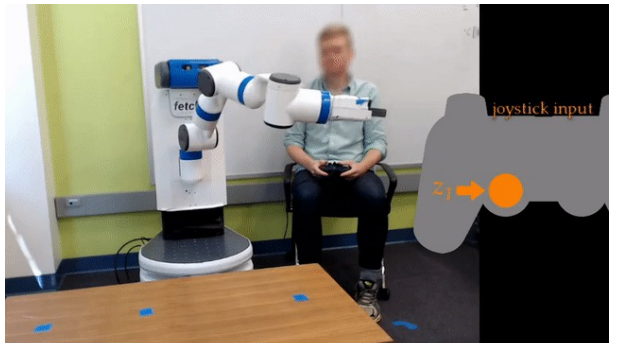


[*Losey, Srinivasan, Mandlekar, Garg, Sadigh*. ICRA 2020]

Learning Latent Actions

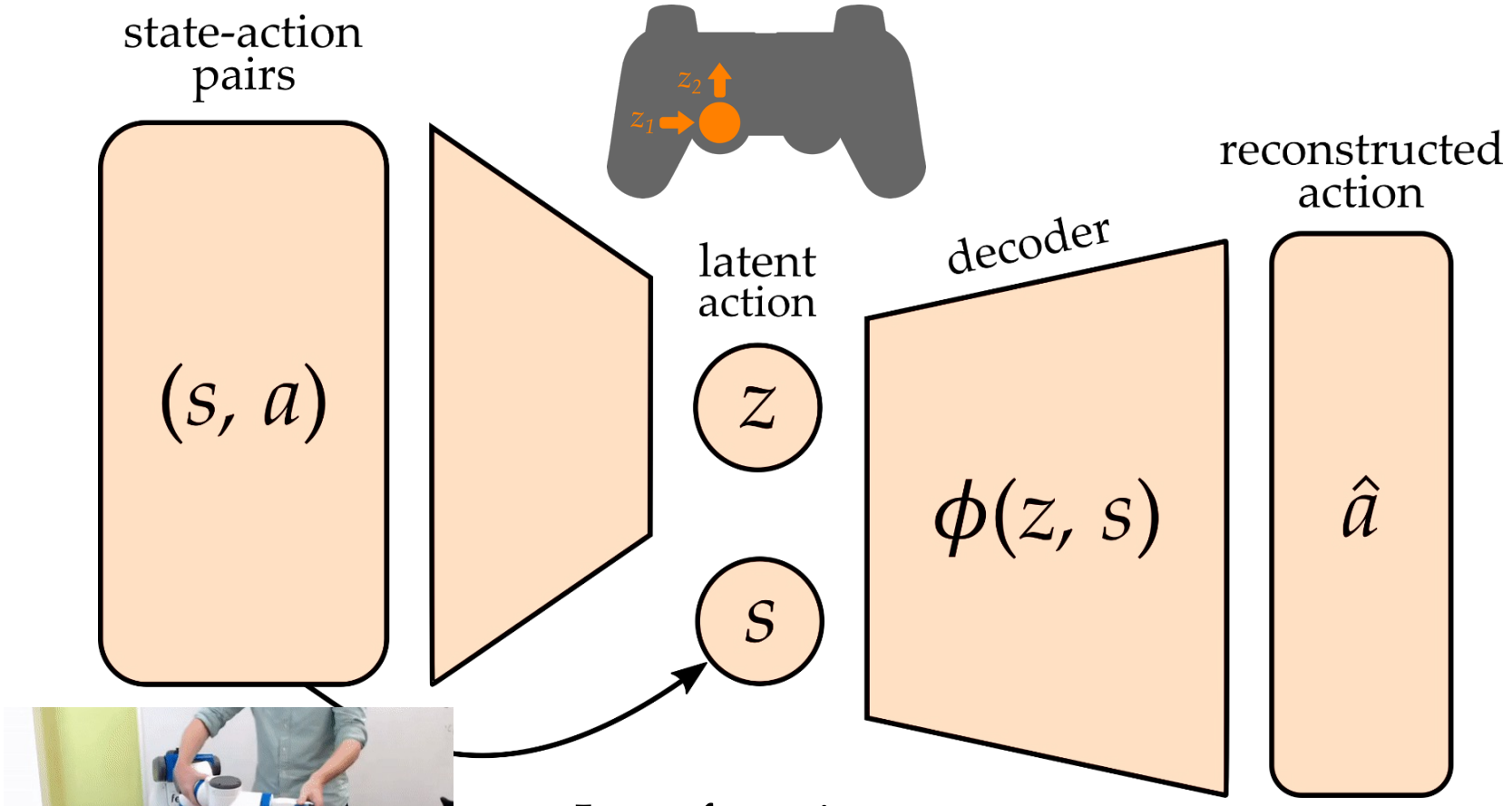


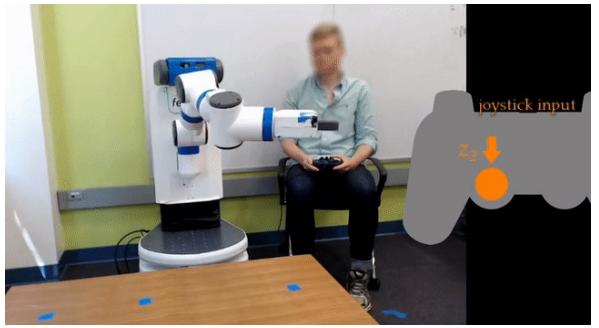


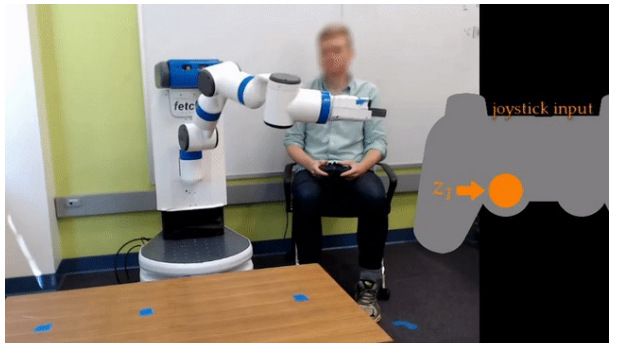


Loss function:
reachability, controllability,
temporal consistency, linearity [*Losey, Srinivasan, Mandlekar, Garg, Sadigh*. ICRA 2020]

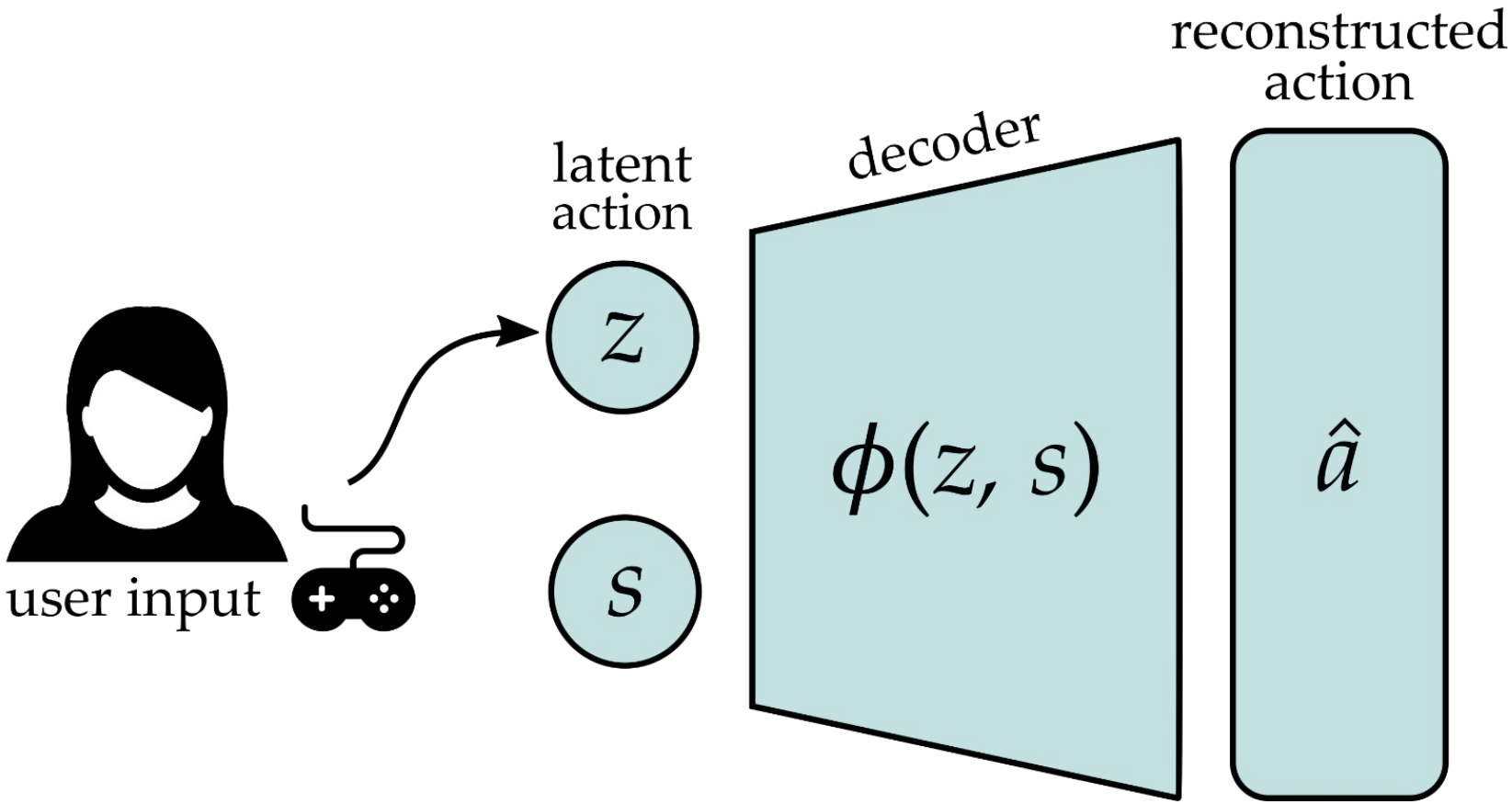
Learning Latent Actions







Loss function:
reachability, controllability,
temporal consistency, linearity [*Losey, Srinivasan, Mandlekar, Garg, Sadigh*. ICRA 2020]



[*Losey, Srinivasan, Mandlekar, Garg, Sadigh*. ICRA 2020]

4x Speed

(1) add eggs

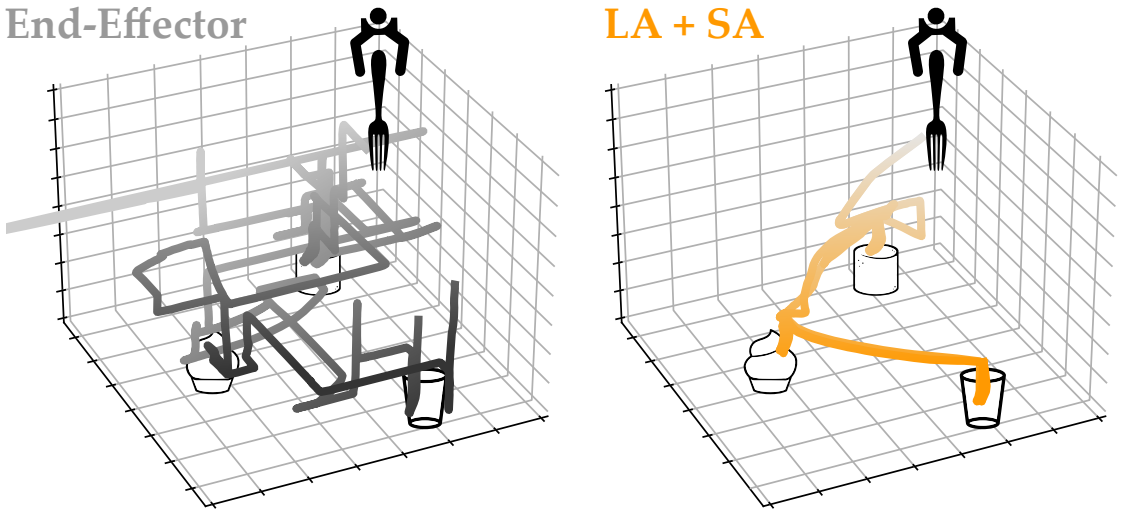
End-Effector

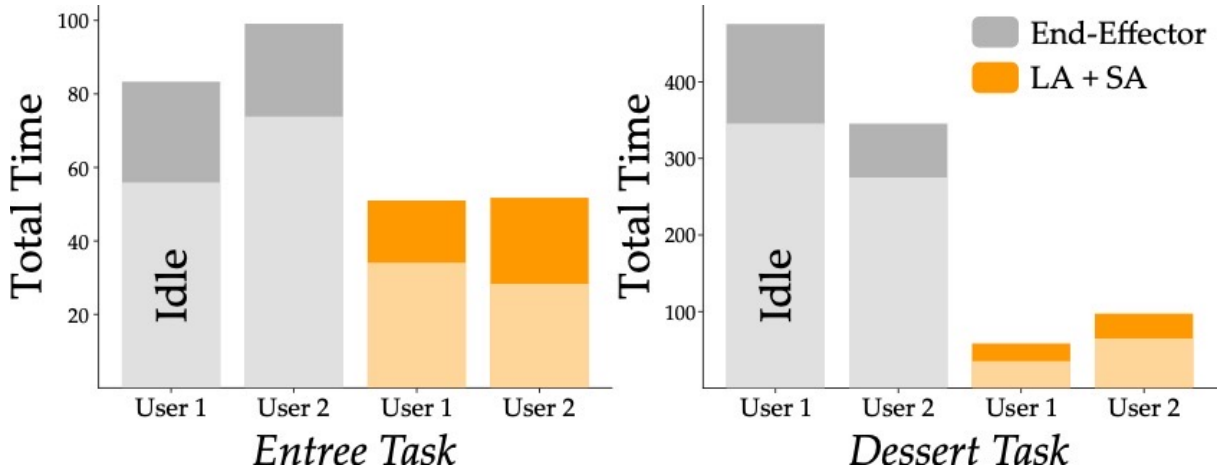
(1) add eggs

Latent Action

Entree Task

Dessert Task

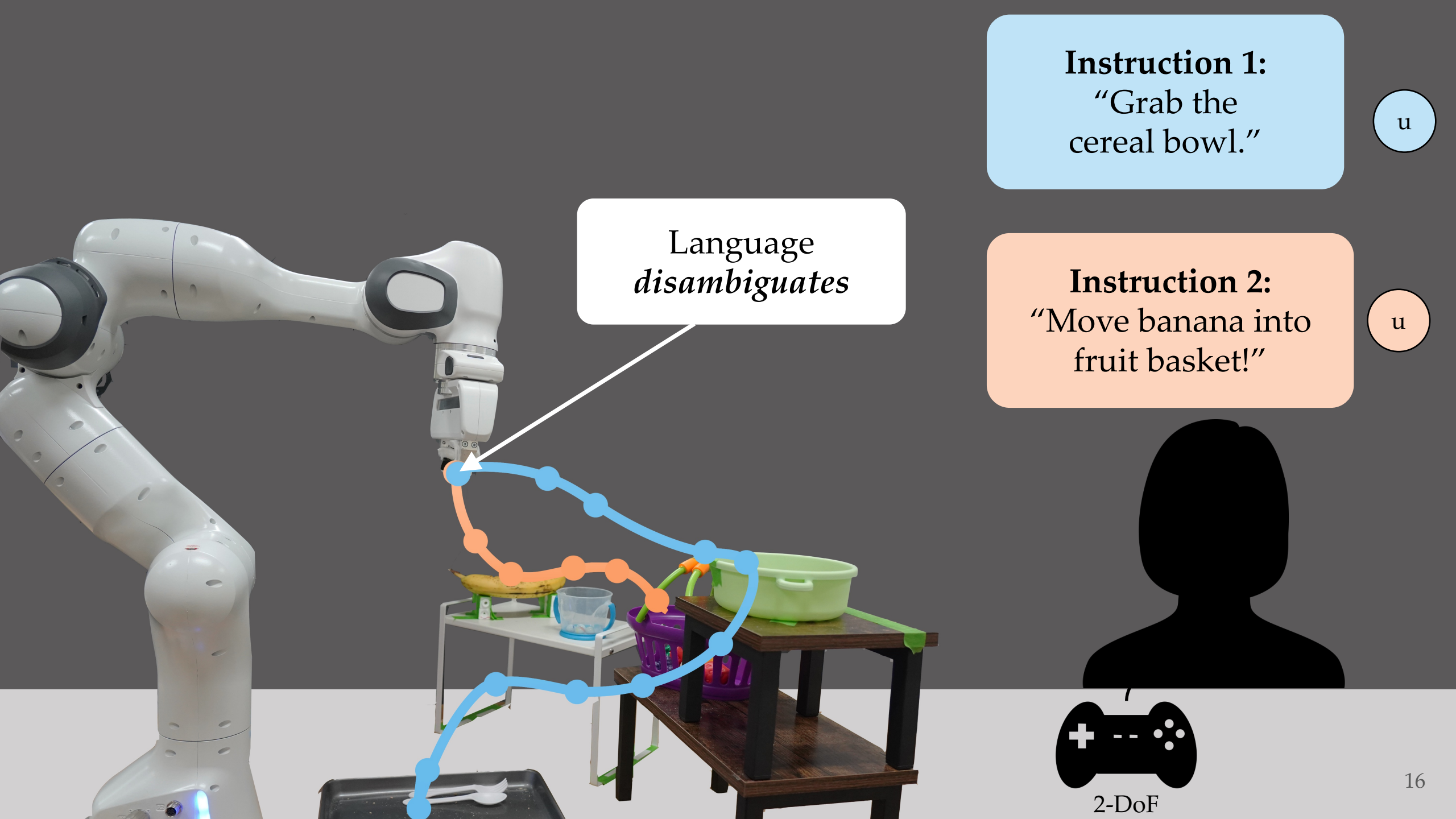




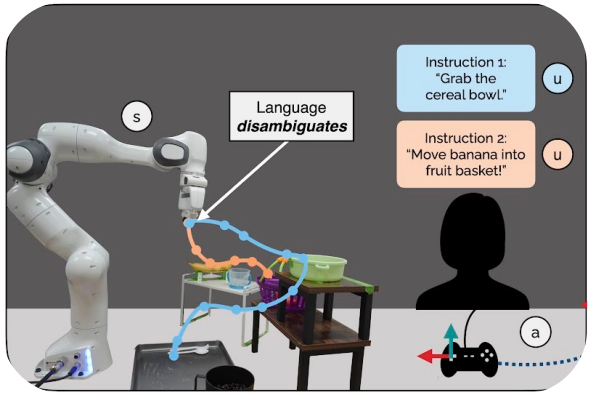
[*Losey, Jeon, Li, Srinivasan, Mandlekar, Garg, Bohg, Sadigh*. AURO 2021]

Latent actions enable intuitive low-dimensional control…

…but how far can we go without any other source of data?


Instruction 1:
"Grab the cereal bowl."
u
Language *disambiguates*
Instruction 2:
"Move banana into fruit basket!"
u
2-DoF

Interaction



LILA:
Language-Informed Latent Actions
Karamcheti*, Srivastava*, Liang, Sadigh
CoRL 2021

Learning

"Please get me an apple"

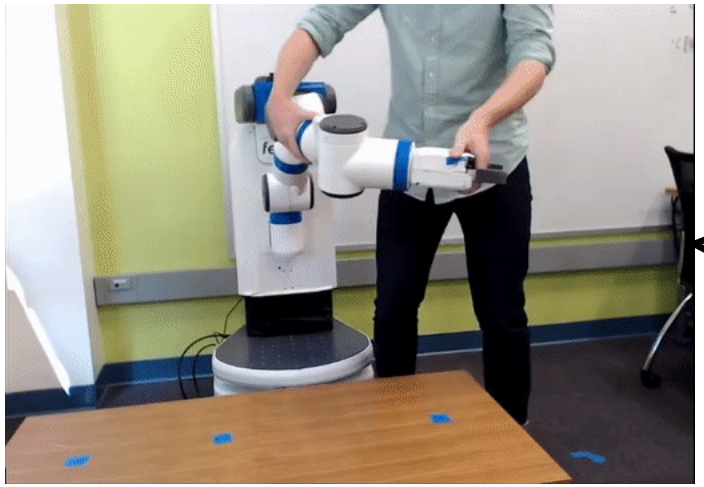
"Go to the table" "Pick up an apple" "Bring it to me"

Teaching

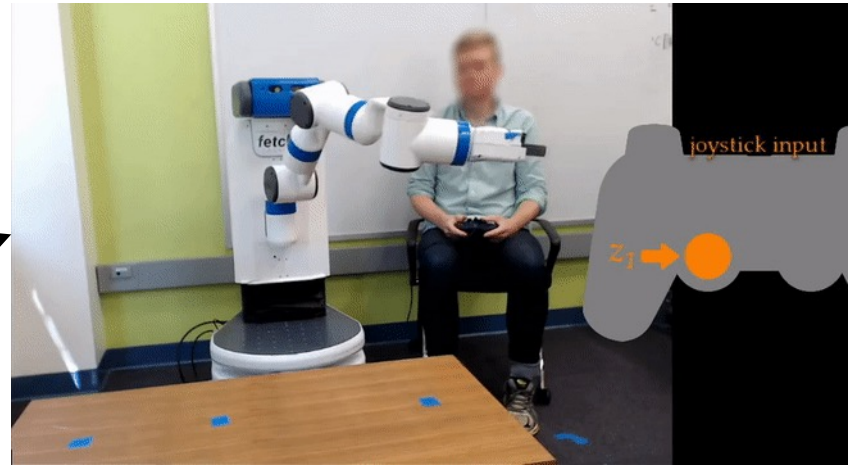
LILA: Language-Informed Latent Actions

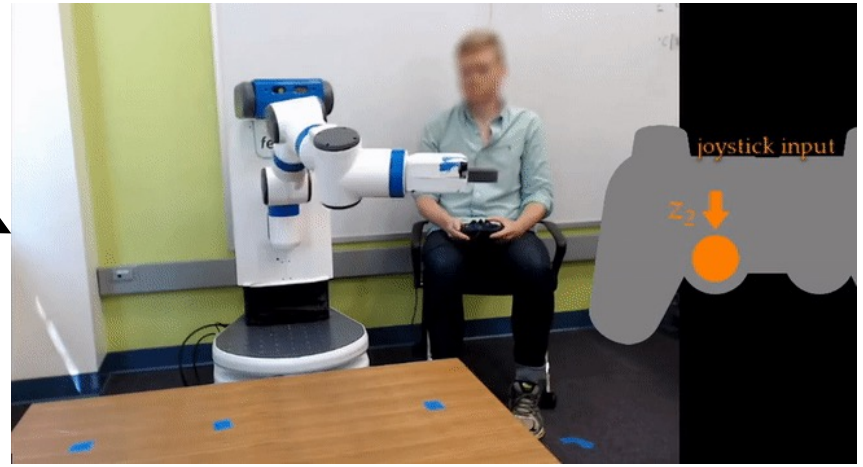
"Alone we can do so little; together we can do so much."
— Helen Keller

Revisiting Learned Latent Actions



Demonstrations for a *Single Task*

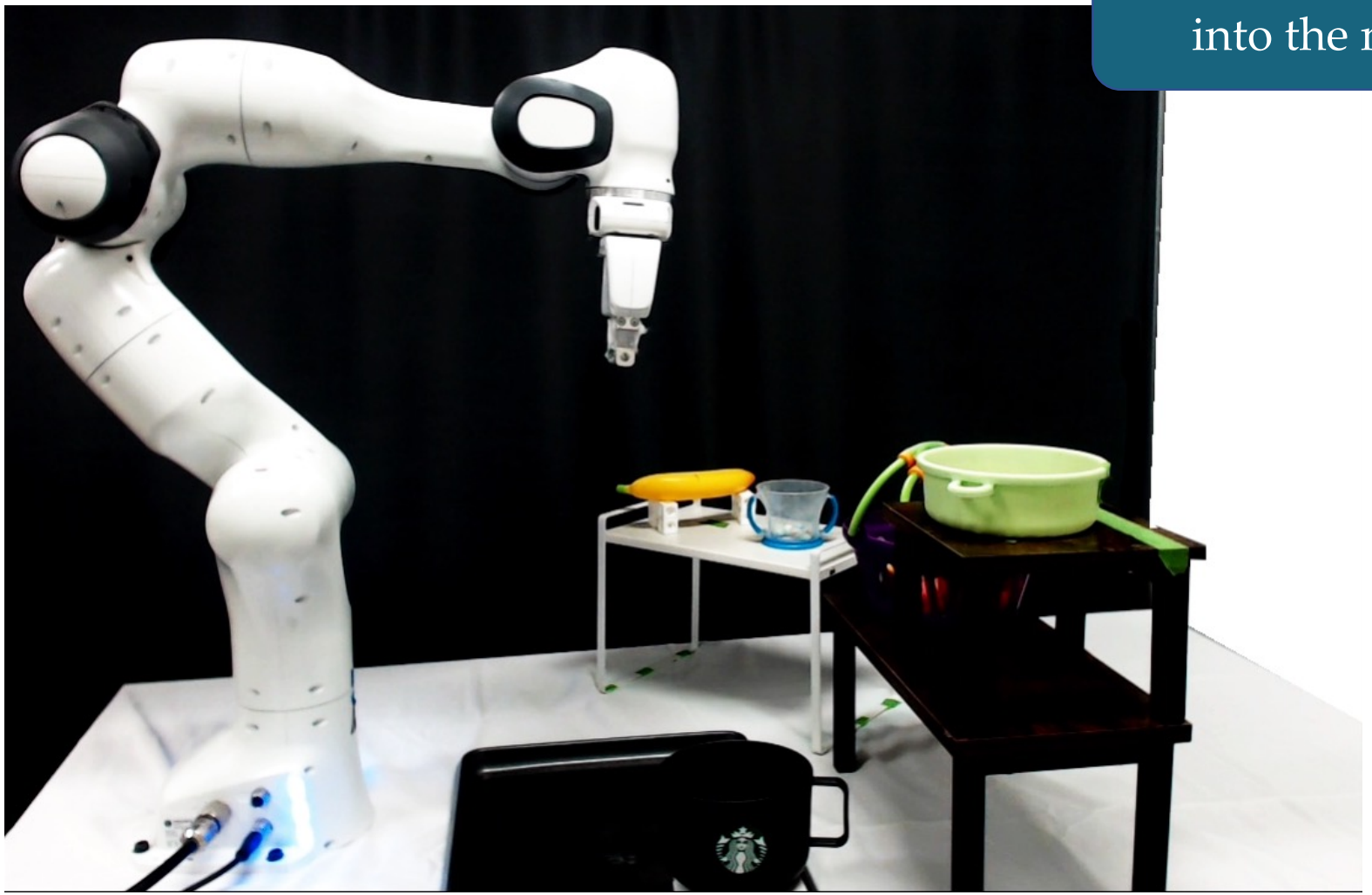




Key Idea: Language to *index & disambiguate* tasks!

[*Losey, Srinivasan, Mandlekar, Garg, Sadigh*. ICRA 2020]

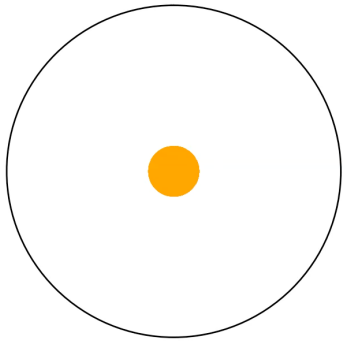
LILA in Action



"Pour the blue cup into the mug"

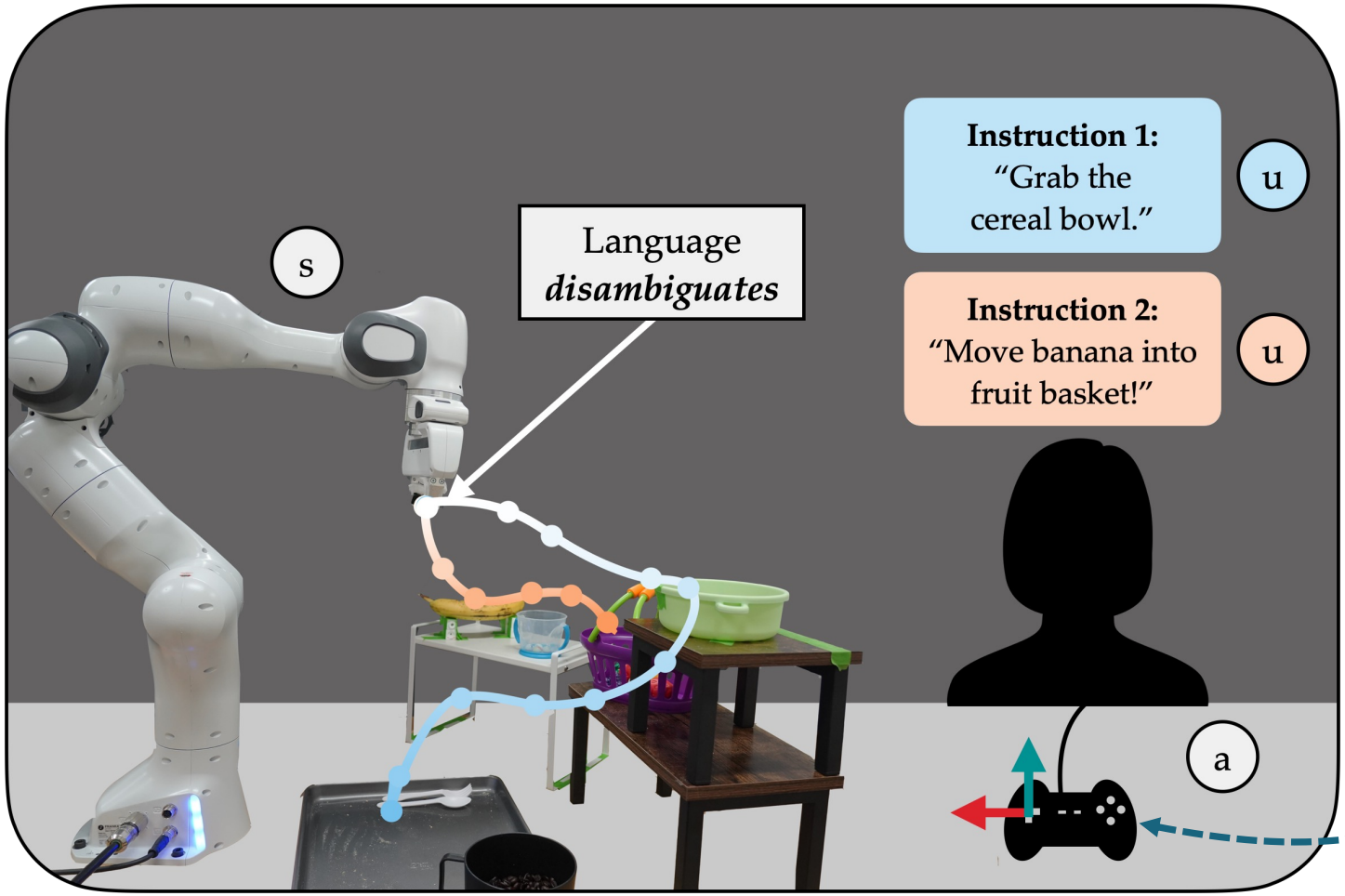
LILA: Generates *language-conditioned* control manifolds.

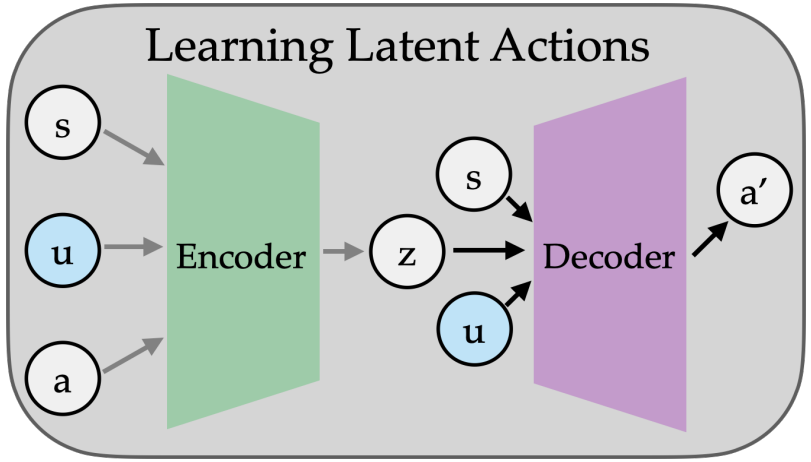
Learned from *10 demonstrations!*

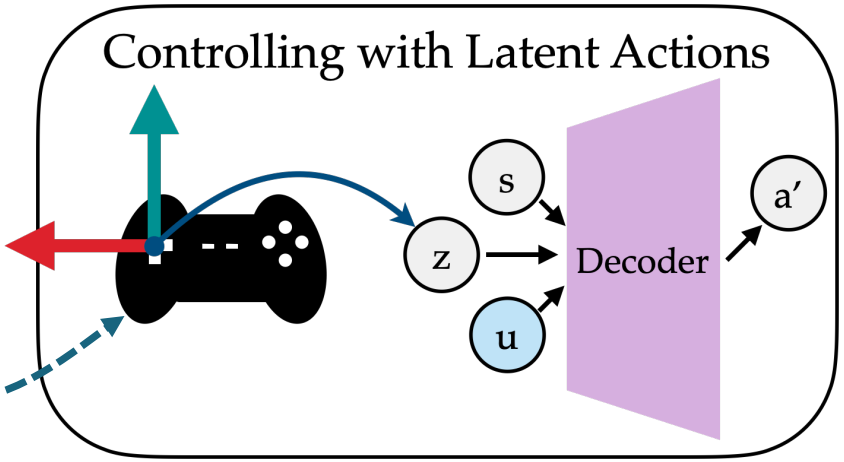


[*Karamcheti*, Srivastava*, Liang, Sadigh*. CoRL 2021]

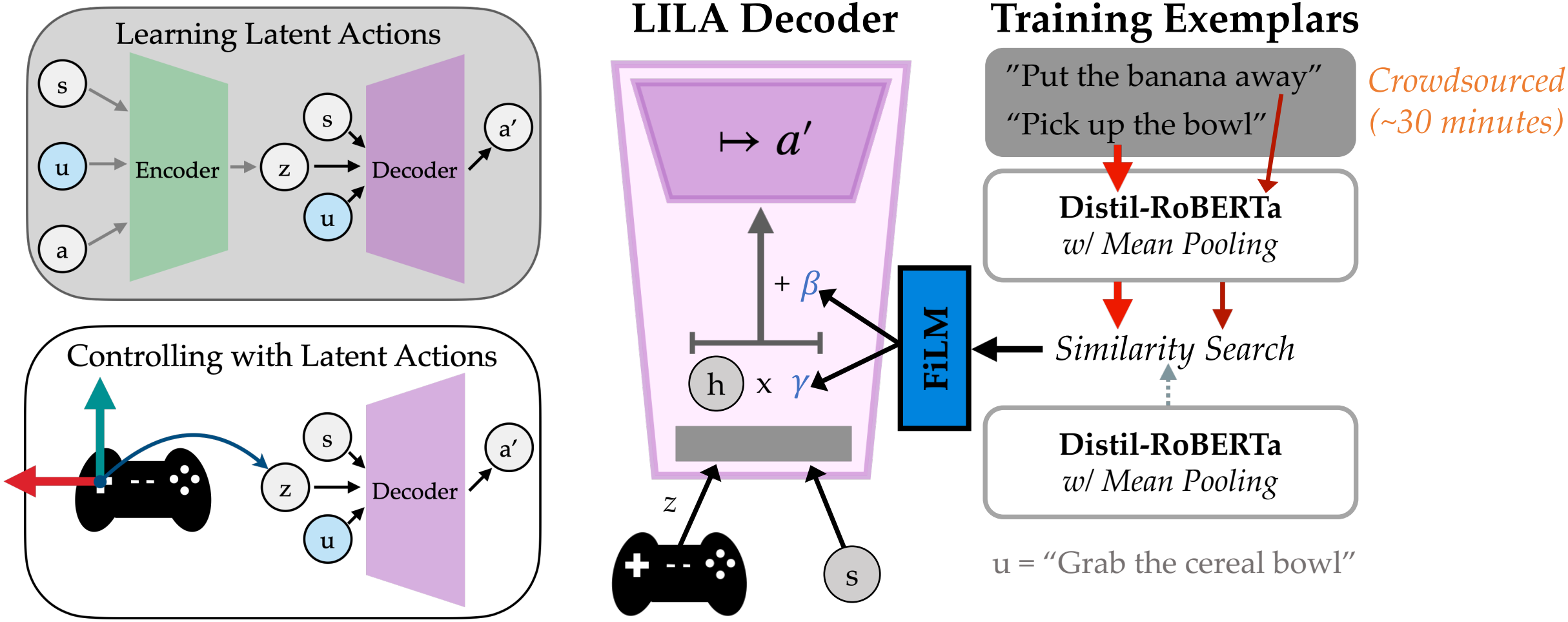
Fusing Language & Latent Actions







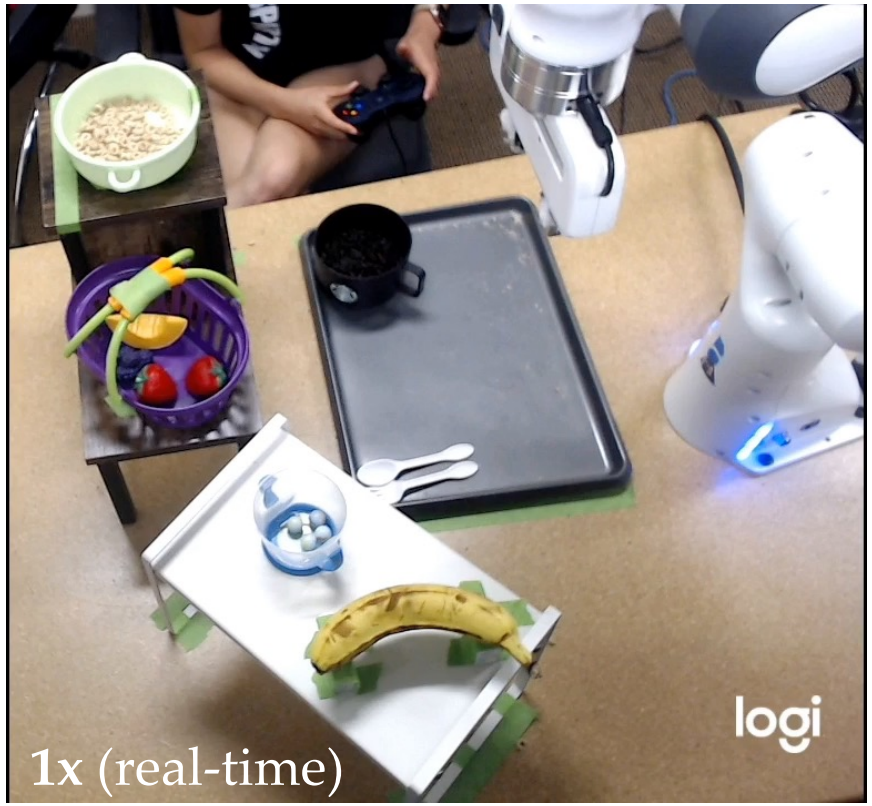
Integrating Language



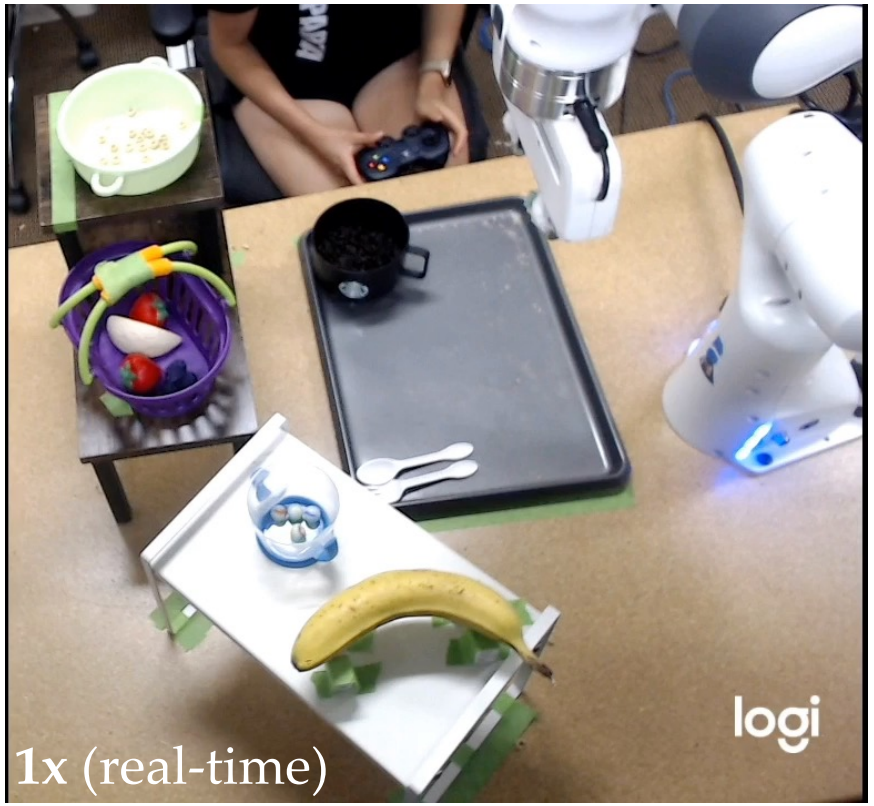
["Unnatural Language Processing: Bridging the Gap Between Synthetic & Natural Language Data," *Marzoev et. al.* arXiv 2020]
["FiLM: Visual Reasoning with a General Conditioning Layer," *Perez et. al.* AAAI 2017]

User Study – Intuitive & Effective

"Place the banana in the fruit basket."



End-Effector Control (w/ 2-DoF Joystick)
Constant mode switching & "jerky" motion…



LILA
*Smooth and intuitive – **2-3x faster!***

Language *indexes* control; utterances specify how the human and robot should *share autonomy*…

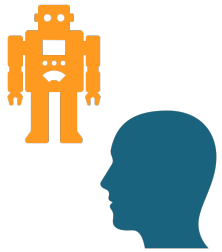
… this allows for the robot & human to play to their strengths – **symbiosis & efficiency**!

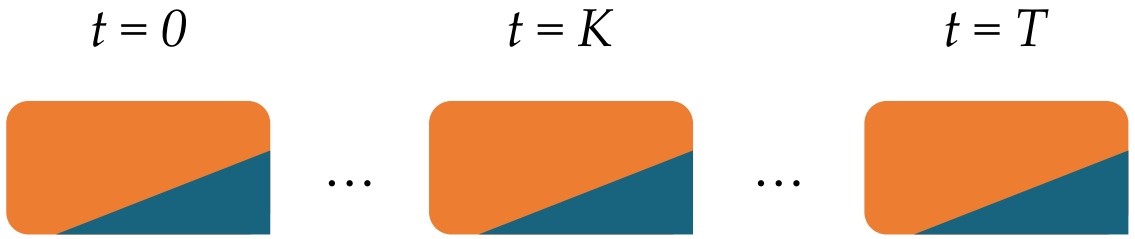
Interlude – Language & Abstraction

"We learn by rearranging what we know."
— Ludwig Wittgenstein

Abstraction: From Interaction to Learning

"Place the banana in the fruit basket."

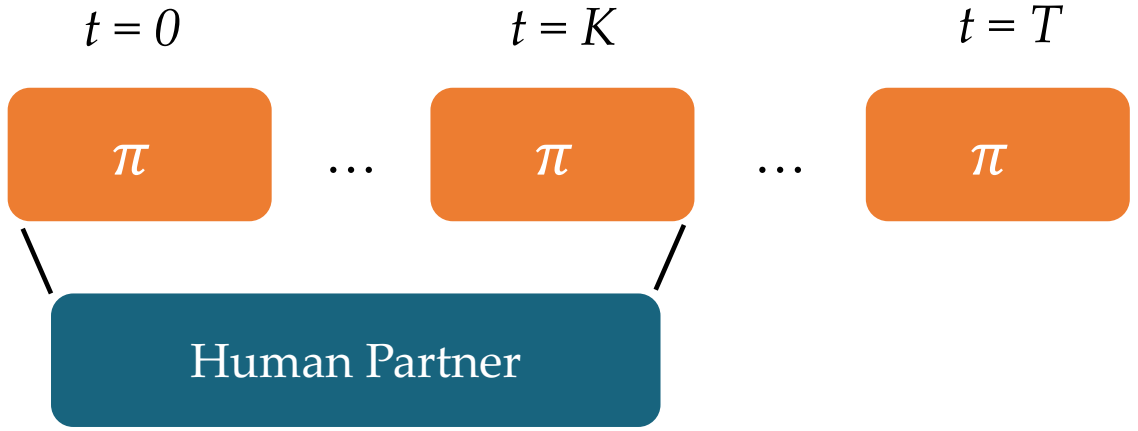




LILA: Robot & human *work together.*

Language dictates division of labor.

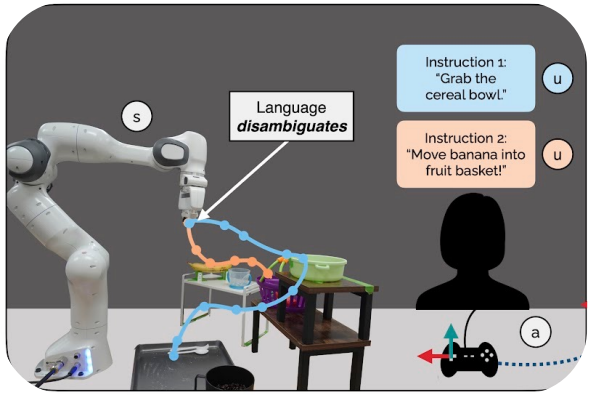
Q: Can we use this idea for **learning**?



Leverage *language abstraction* as supervision!

How?

Interaction



LILA:
Language-Informed Latent Actions
Karamcheti*, Srivastava*, Liang, Sadigh
CoRL 2021

Learning

ELLA:
Exploration through Learned Language Abstraction
Mirchandani, Karamcheti, Sadigh
NeurIPS 2021

Teaching

ELLA: Exploration through Learned Language Abstraction

"If you can't explain it simply, you don't understand it well enough."
— Albert Einstein

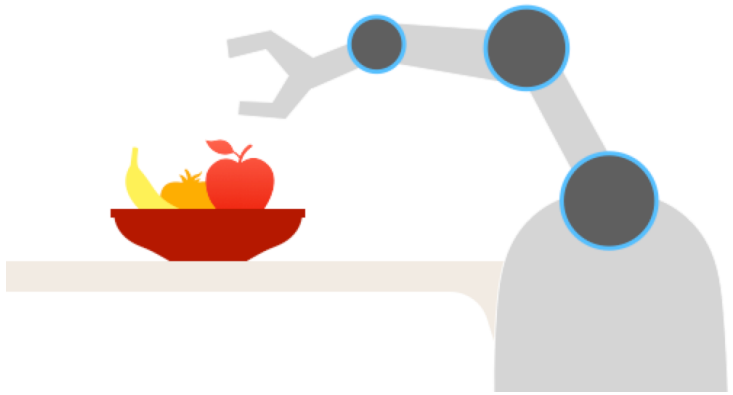
Reinforcement Learning from sparse rewards for complex tasks is **not sample efficient…**

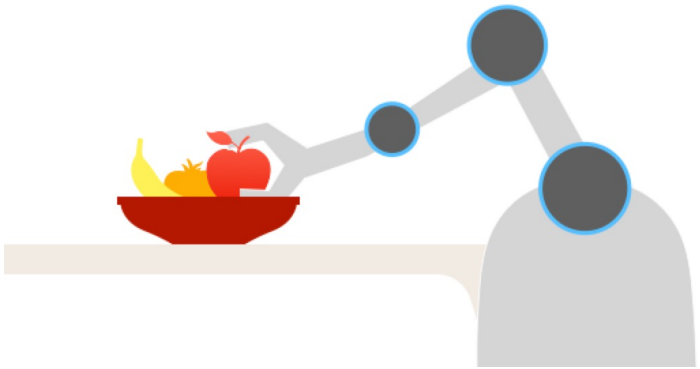
Can we use **language abstraction** to guide **exploration**?

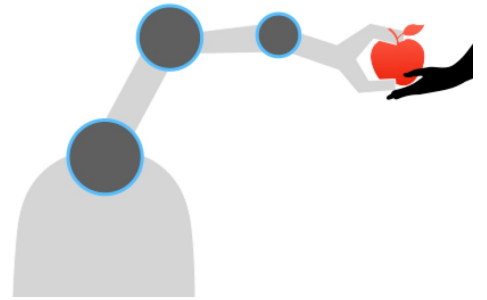
Suvir Mirchandani

[Mirchandani, Karamcheti, Sadigh, *NeurIPS 2021*]

"Bring me an apple"







"Go to the table"

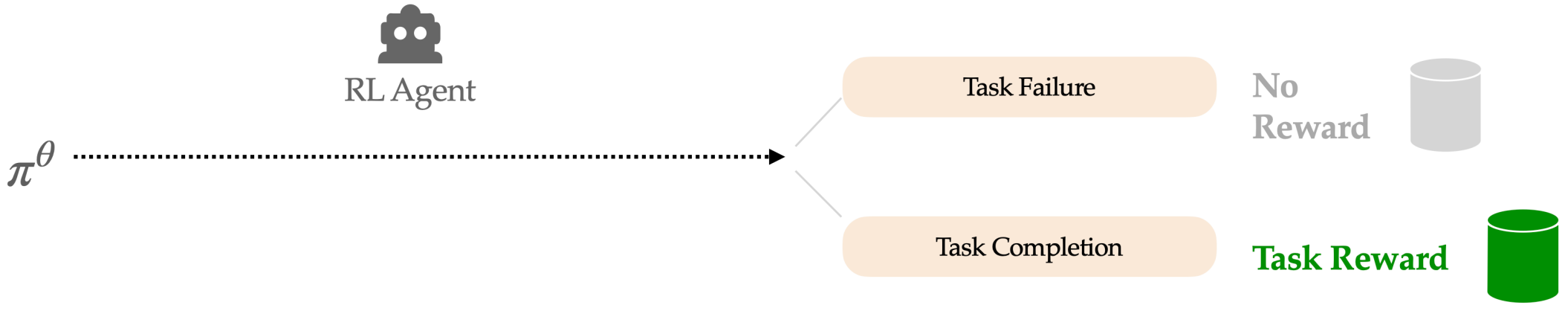
"Pick up an apple"

<turn right> …

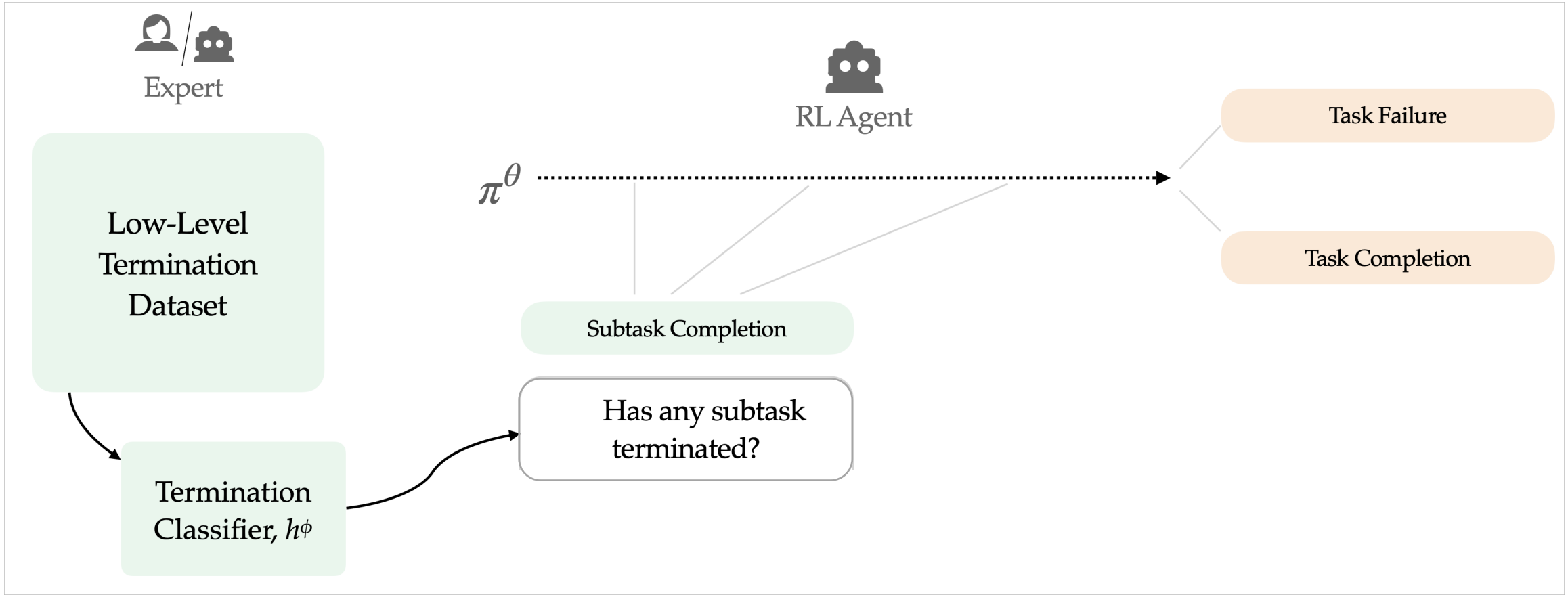
<turn left> …

<lift up> …

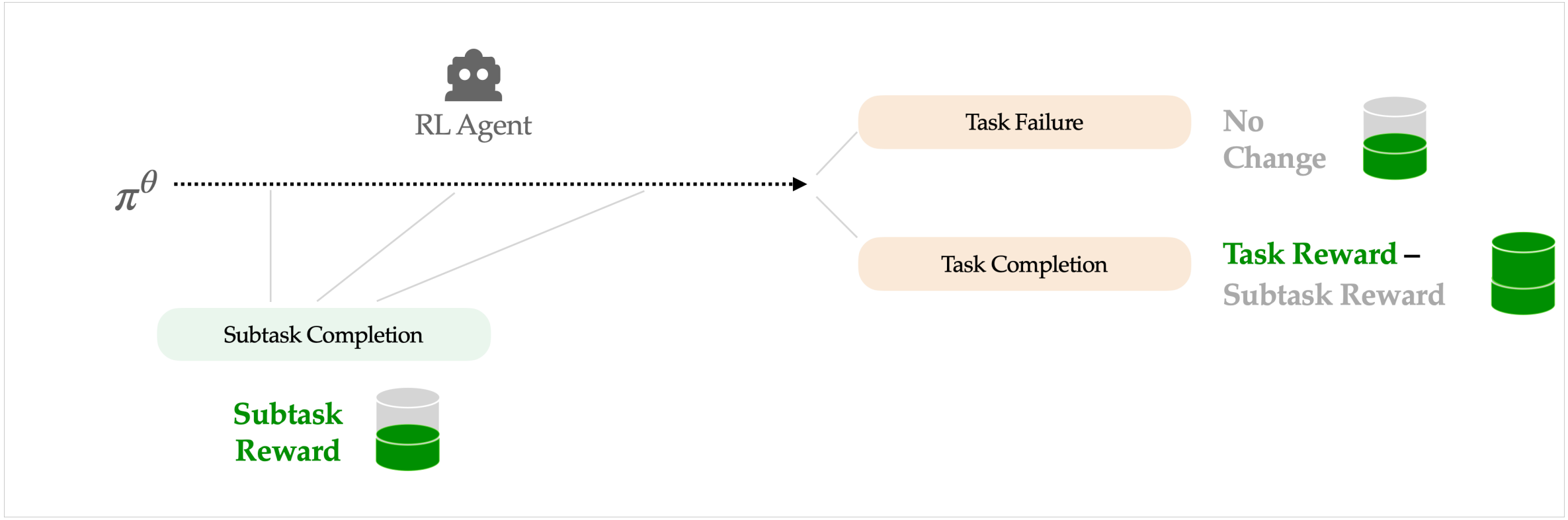
ELLA – An Overview



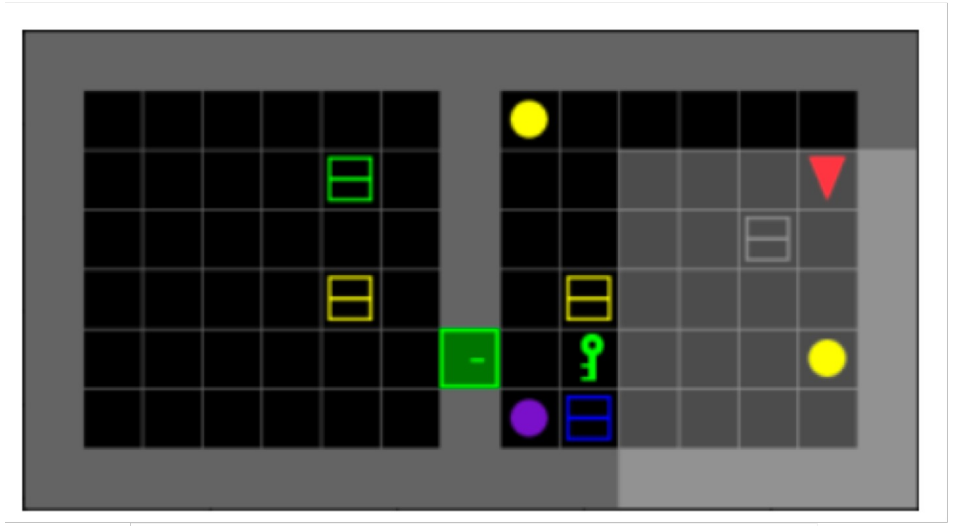
ELLA – Leveraging Language Subtasks



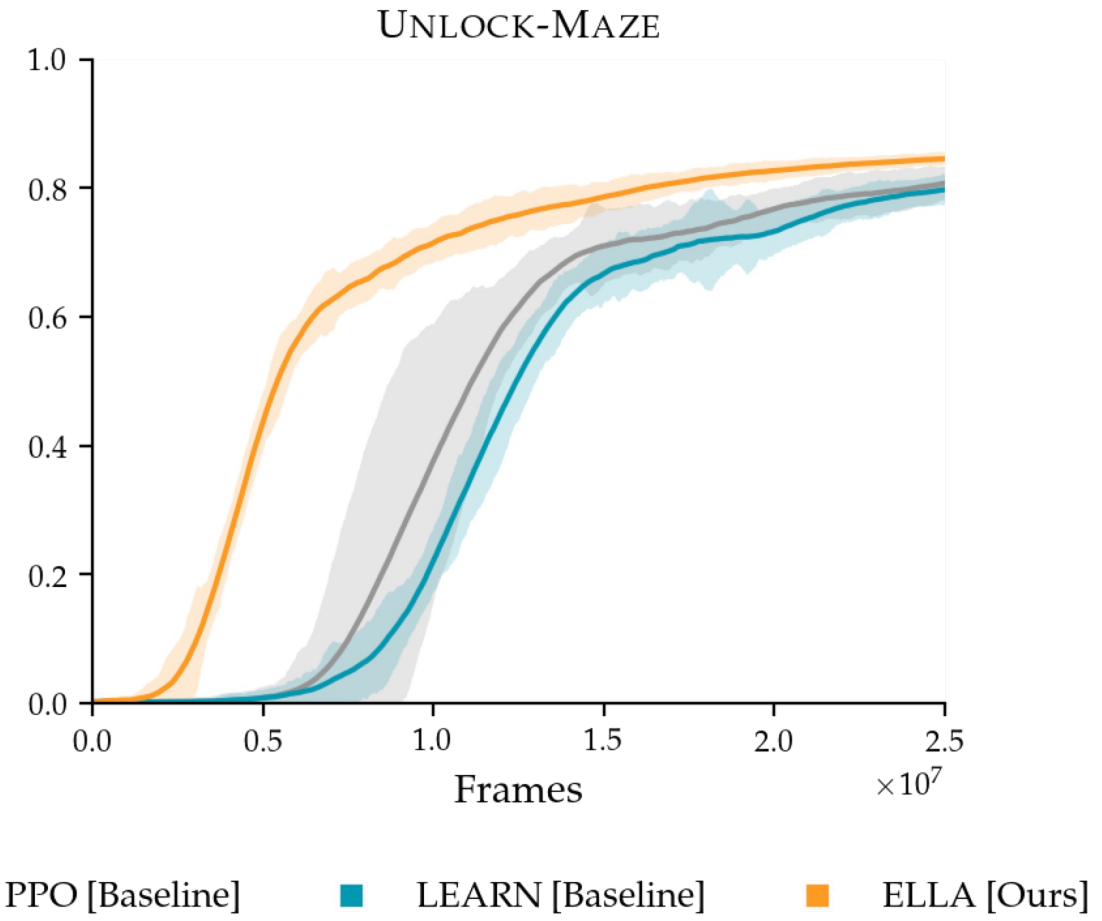
ELLA – Reward Shaping



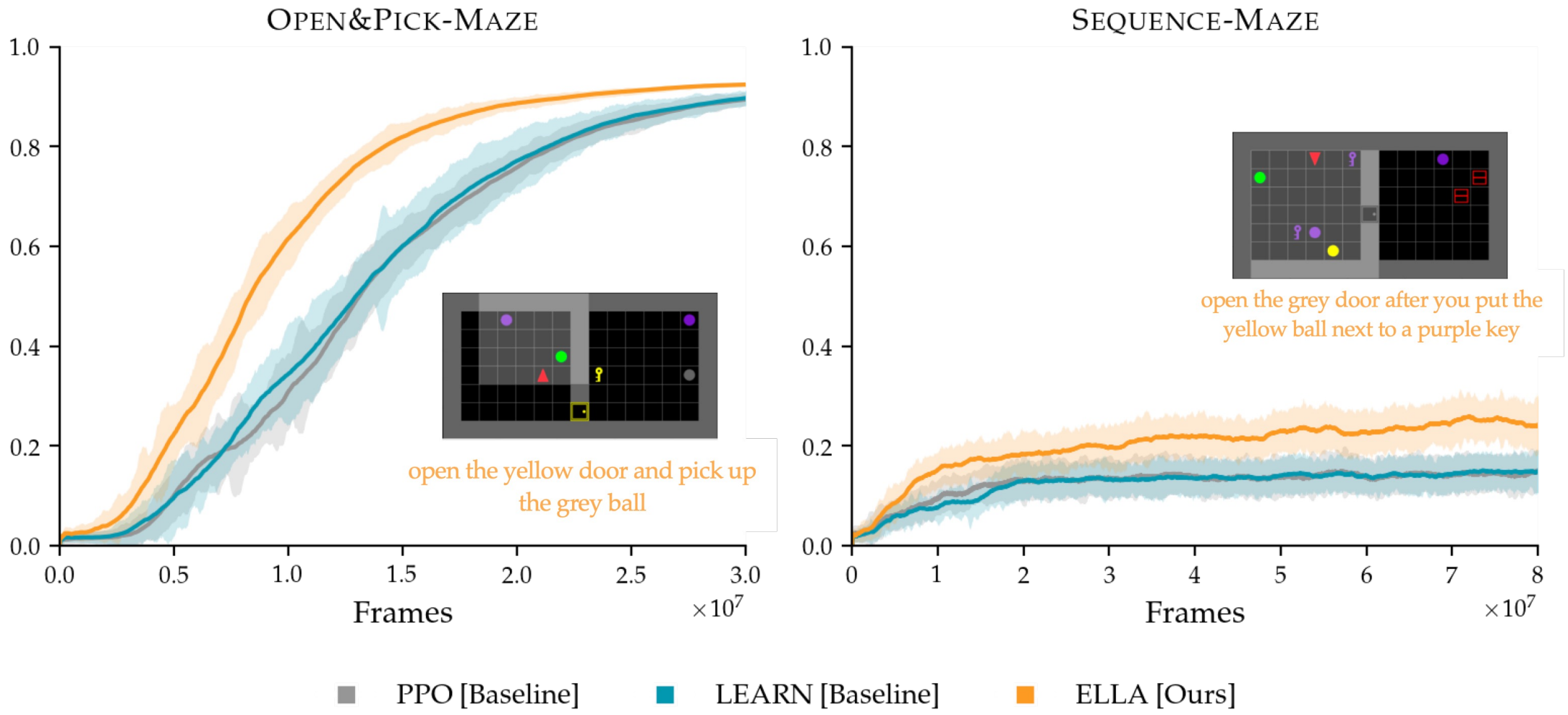
Results – Sparsity



open the green door



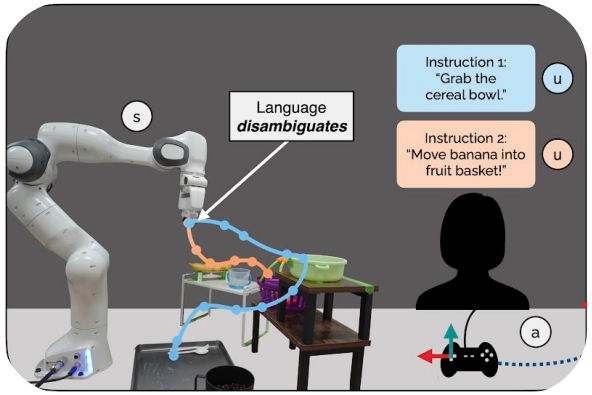
Results – Compositionality



The ability to describe our behavior relative to known abstractions *provides rich supervision…*

… *but can we go the other way?* Can we teach our hard-won abstractions to users?

Interaction



LILA:
Language-Informed Latent Actions
Karamcheti*, Srivastava*, Liang, Sadigh
CoRL 2021

Learning

ELLA:
Exploration through Learned Language Abstraction
Mirchandani, Karamcheti, Sadigh
NeurIPS 2021

Teaching

Assistive Teaching of Motor Control Tasks to Humans
Srivastava, Bıyık, Mirchandani, Goodman, Sadigh
NeurIPS 2022

Requiem — Teaching & Active Language

"The art of teaching is the art of assisting discovery."
— Mark Van Doren

Motor control tasks are everywhere…

Motor control tasks are everywhere…

and are challenging to learn!

Motor control tasks are everywhere…

and are challenging to **teach others**!

What makes teaching motor control tasks challenging?

What makes teaching motor control tasks challenging?

Requires specialized instructors

What makes teaching motor control tasks challenging?

Requires specialized instructors

Individual student variations

What makes teaching motor control tasks challenging?

Requires specialized instructors

Individual student variations

Diverse physical conditions

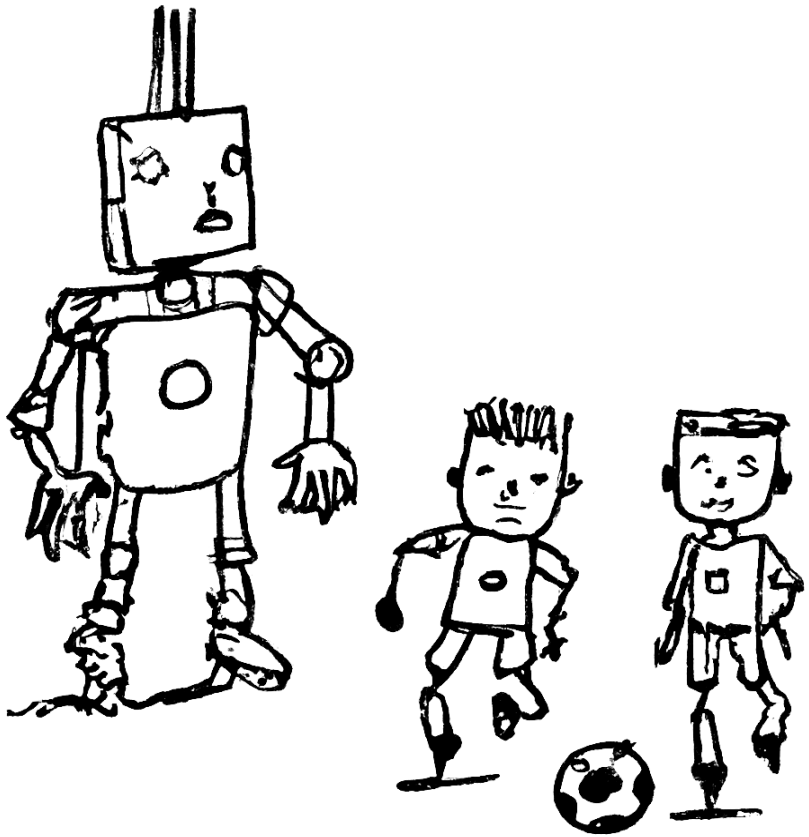
How do we teach motor control tasks?

Language to refer to chords and arpeggios

Use language to refer to the motion of pitching a ball

Use language to refer to the motion of suturing

Can AI-assistance help teach humans motor control tasks?



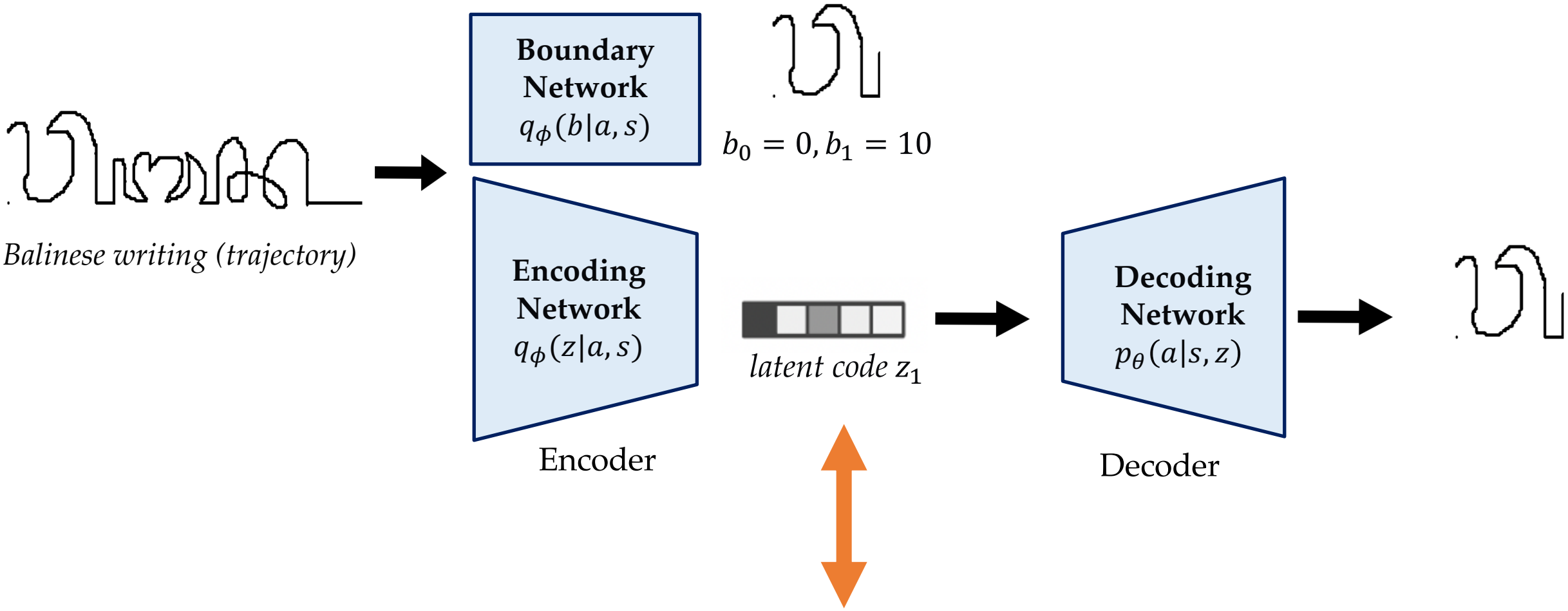
Can we leverage expert knowledge of a motor control task

to help any human learn the task themselves?

Megha Srivastava

Assistive Teaching of Motor Control Tasks to Humans
[Srivastava, Bıyık, Mirchandani, Goodman, Sadigh, *NeurIPS 2022*]

Unsupervised Skill Discovery: CompILE [Kipf et. al. '19]



There is a language utterance corresponding to each skill

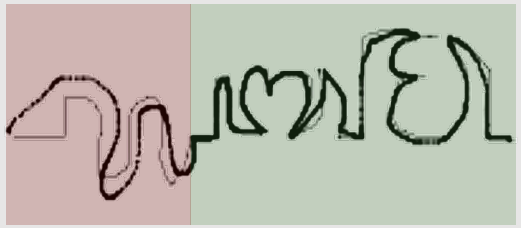
1. Extract Skills from Expert Demonstrations

τ_1^e:
τ_2^e:
τ_3^e:
τ_4^e:
⋮

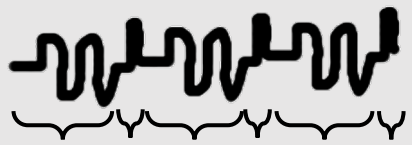
2. Diverse Skill Selection

τ_1^e:
τ_2^e:

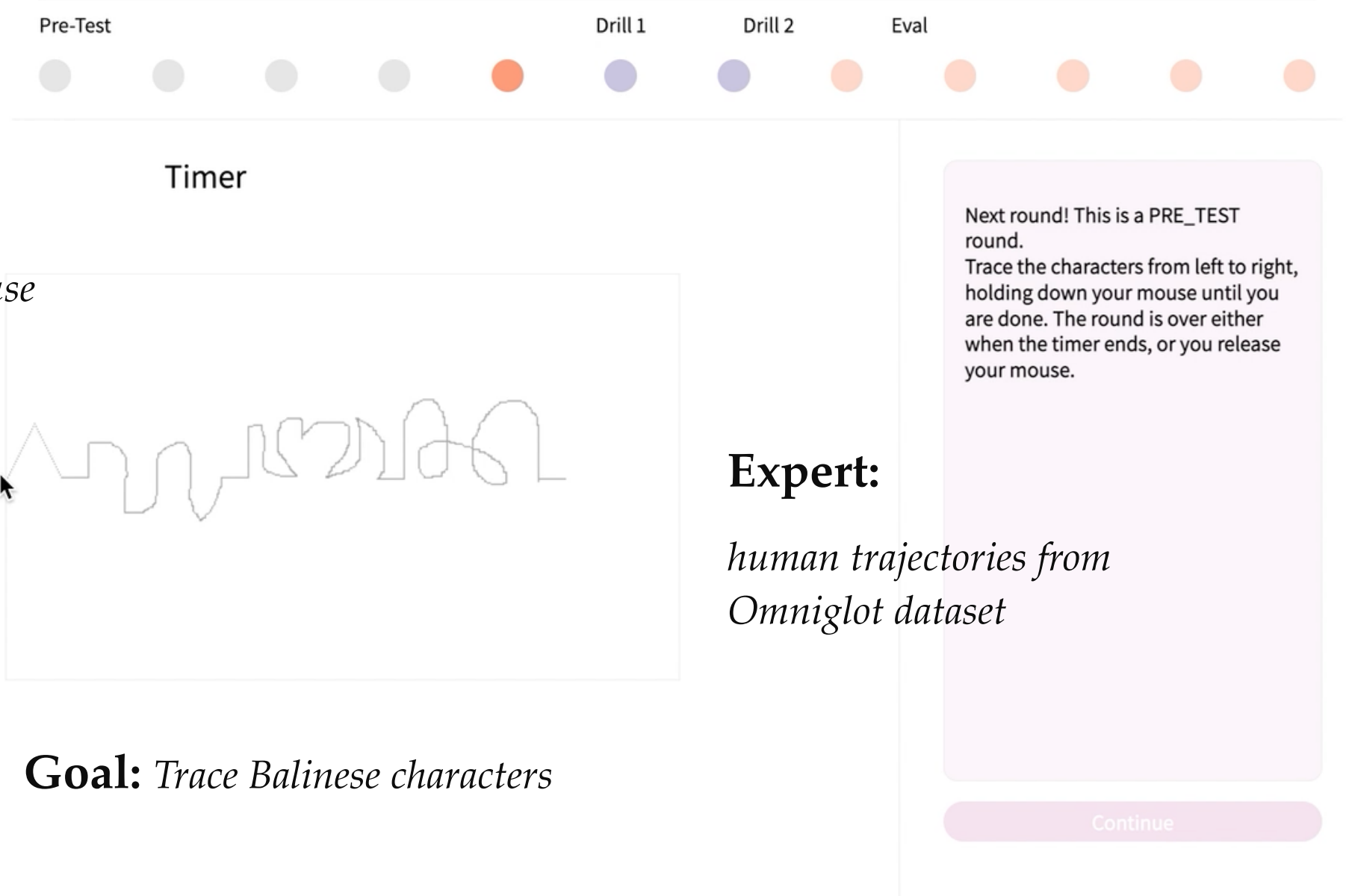
3. Individualized Skills



4. Individualized Drills



(1) Writing Task



Control:

continuous mouse control

Expert:

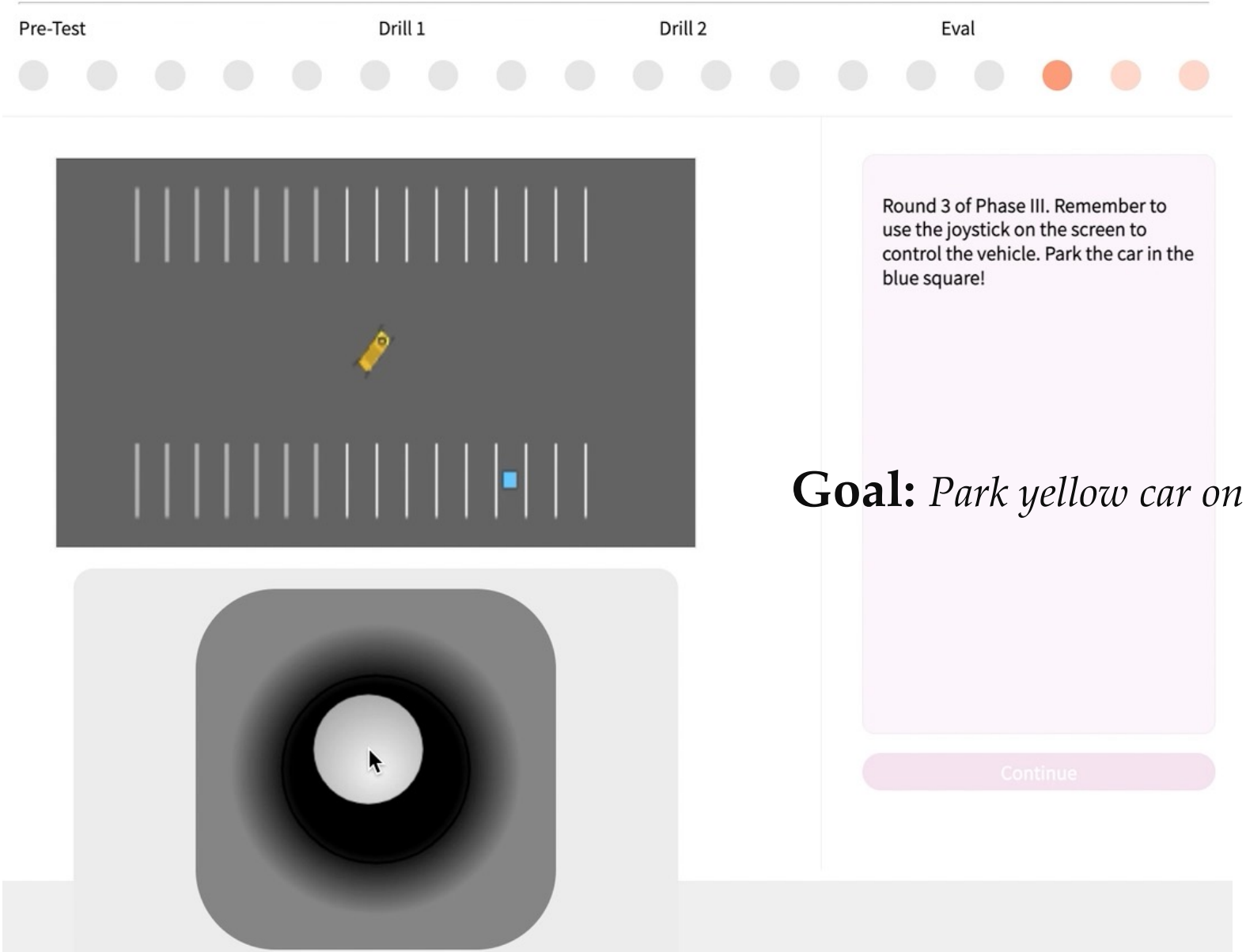
human trajectories from Omniglot dataset

Goal: *Trace Balinese characters*

(2) Parking Task

Expert:

Optimal Soft-Actor Critic Agent

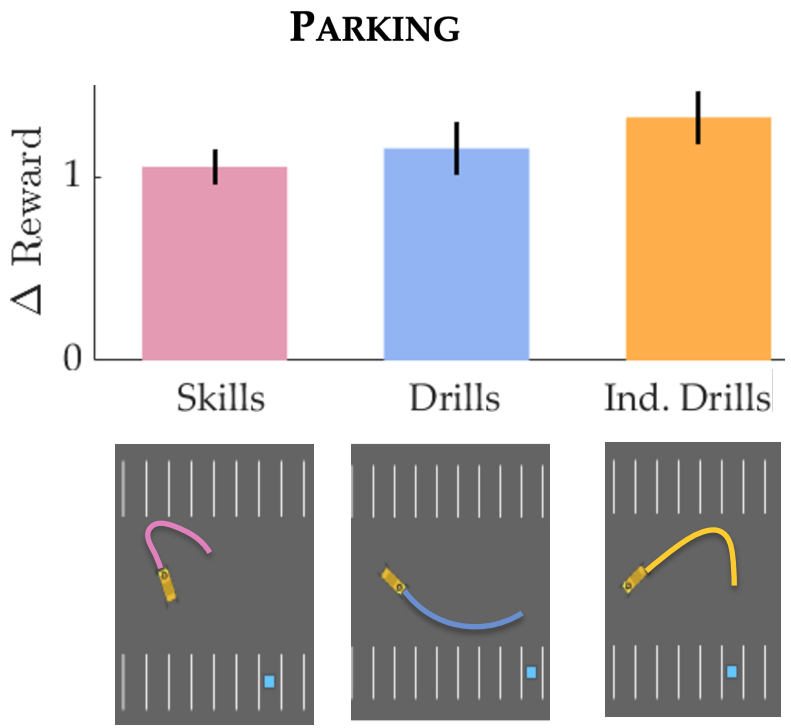


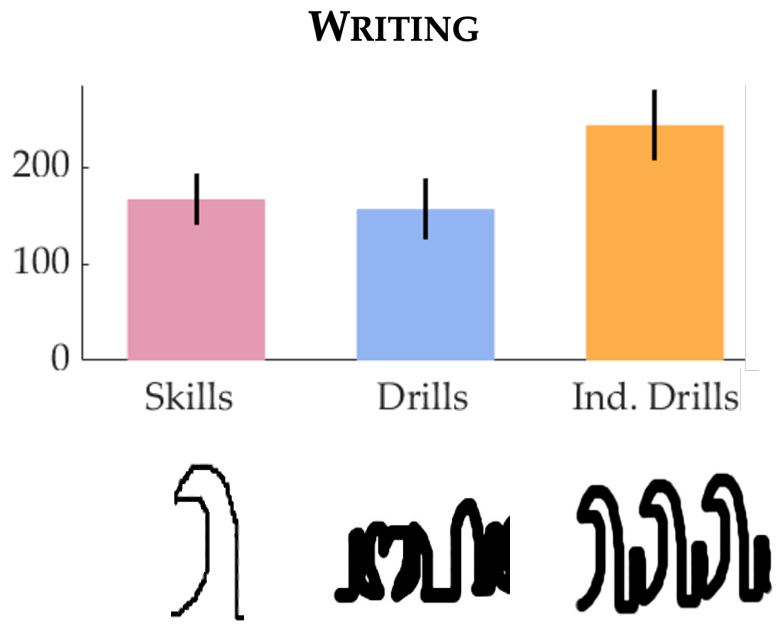
Goal: *Park yellow car on blue spot*

Control:

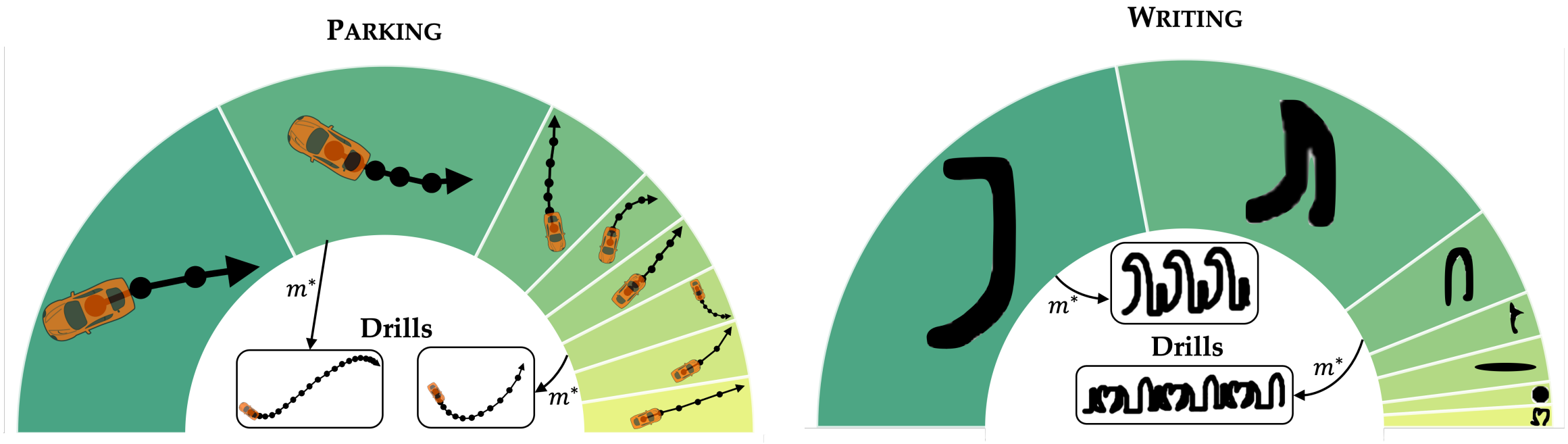
continuous mouse control

Individualized Drills Help with Student Learning

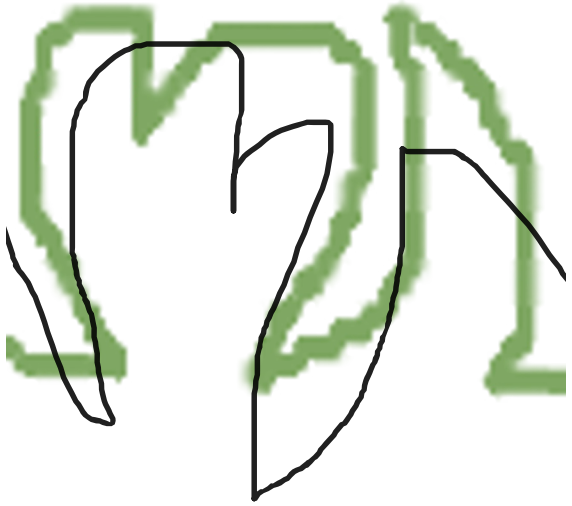




Distribution of hardest skills across individuals



Users referred to each skill using language: “heart shaped”, “Reversed C-shape”, …

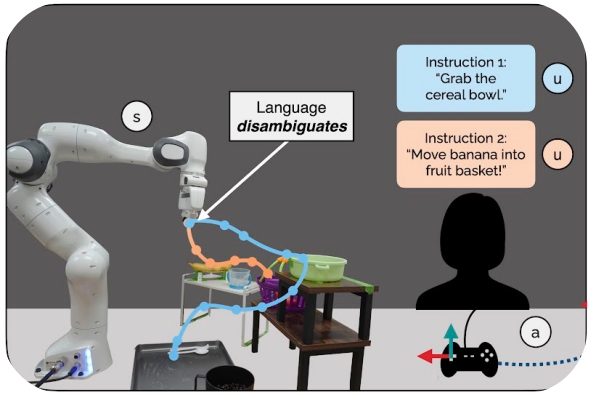


Use language to refer to *skills* or *corrections of skills*:

Do a “Heart-Shaped” character

Make it a bit smoother and symmetric

Interaction



LILA:
Language-Informed Latent Actions
Karamcheti*, Srivastava*, Liang, Sadigh
CoRL 2021

Learning

ELLA:
Exploration through Learned Language Abstraction
Mirchandani, Karamcheti, Sadigh
NeurIPS 2021

Teaching

Assistive Teaching of Motor Control Tasks to Humans
Srivastava, Bıyık, Mirchandani, Goodman, Sadigh
NeurIPS 2022