

Teacher-Learner Interaction for Robot Active Learning

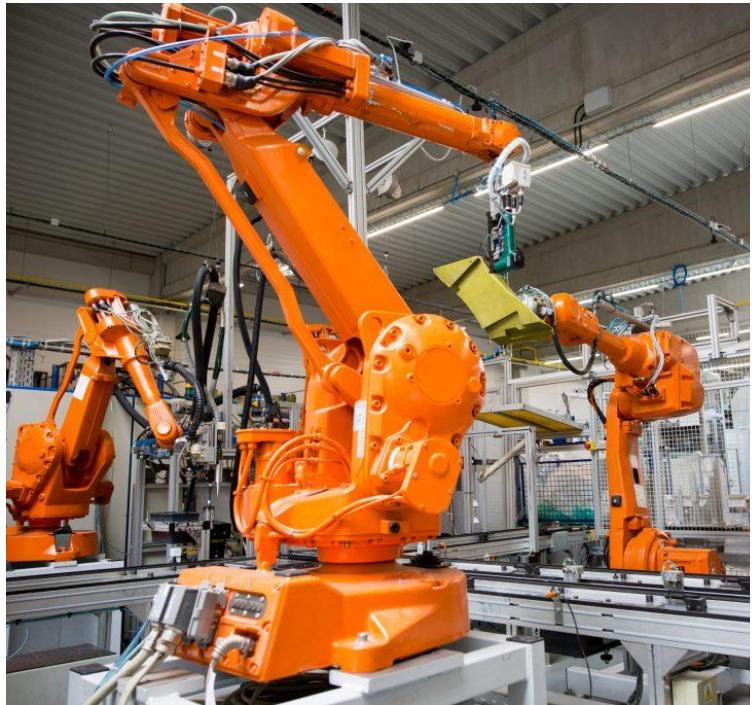
Mattia Racca

Lectio Praecursoria - 30/10/2020
The defence will start at 14.05 EET

ROBOTS: BEYOND INDUSTRIAL SETTINGS



ROBOTS: BEYOND INDUSTRIAL SETTINGS



NEW USERS, NEW CHALLENGES

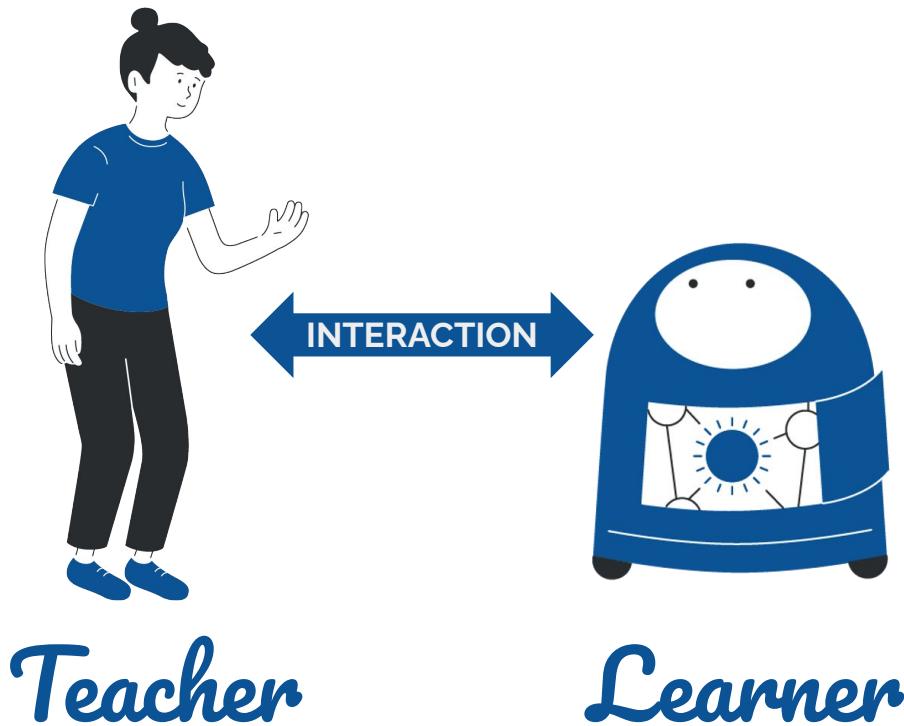


NEW USERS, NEW CHALLENGES



```
1 #!/usr/bin/env python
2
3 import rospy
4 import copy
5 import panda_primitive as pp
6 import program_interpreter as interpreter
7 from panda_pbd.srv import EnableTeaching, EnableTeachingRequest
8 from panda_pbd.msg import UserSyncGoal, MoveToContactGoal, MoveToEEGoal
9 from panda_pbd.srv import MoveFingersRequest, ApplyForceFingersRequest
10 from sensor_msgs.msg import JointState
11
12
13 class PandaPBDInterface(object):
14     def __init__(self):
15         self.program = pp.PandaProgram('A Panda Program')
16
17         self.last_pose = None
18         self.last_gripper_width = None
19         self.relaxed = False
20
21         self.default_parameters = {'kinesthetic_ft_threshold': 5.0,
22                                   'move_to_ee_default_position_speed': 0.07,
23                                   'move_to_ee_default_rotation_speed': -1.0,
24                                   'user_sync_default_force_threshold': 10.0,
25                                   'apply_force_fingers_default_force': 20.0,
26                                   'move_to_contact_default_force_threshold': 10.0,
27                                   'move_to_contact_default_torque_threshold': 10.0,
28                                   'move_to_contact_default_position_speed': 0.07,
29                                   'move_to_contact_default_rotation_speed': -1.0}
30
31     for parameter_name in self.default_parameters.keys():
32         if not rospy.has_param('~' + parameter_name):
```

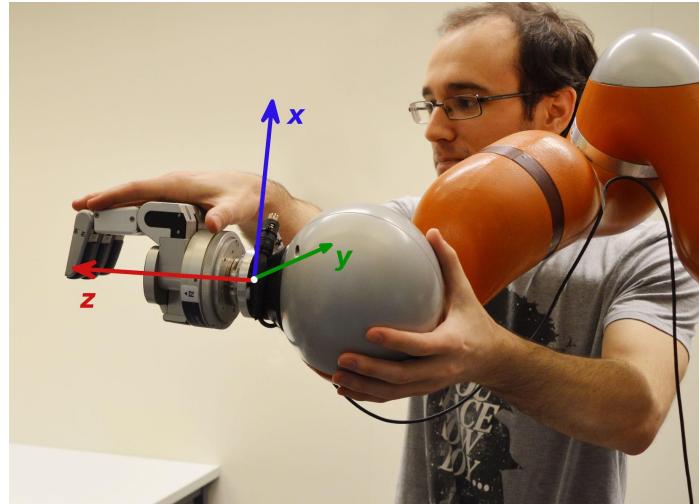
NEW USERS, NEW CHALLENGES



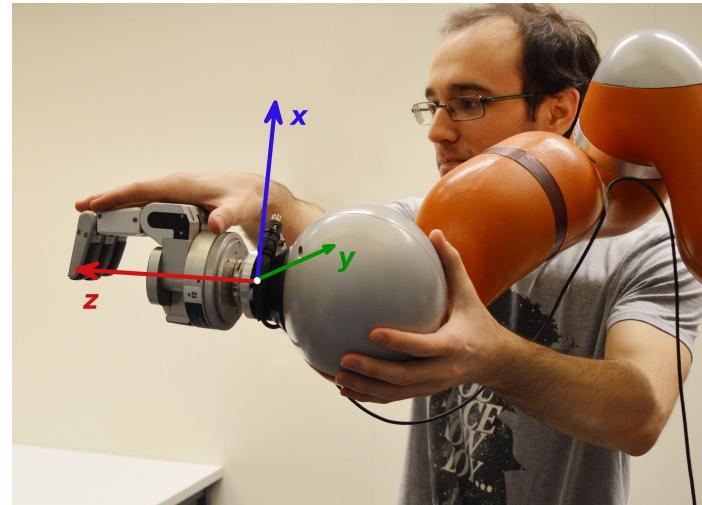
LEARNING FROM DEMONSTRATIONS



LEARNING FROM DEMONSTRATIONS

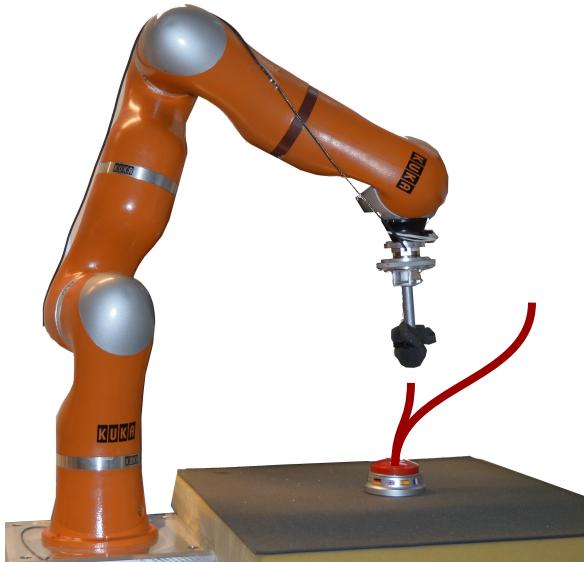


LEARNING FROM DEMONSTRATIONS

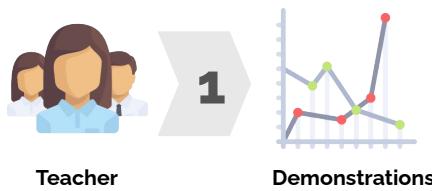
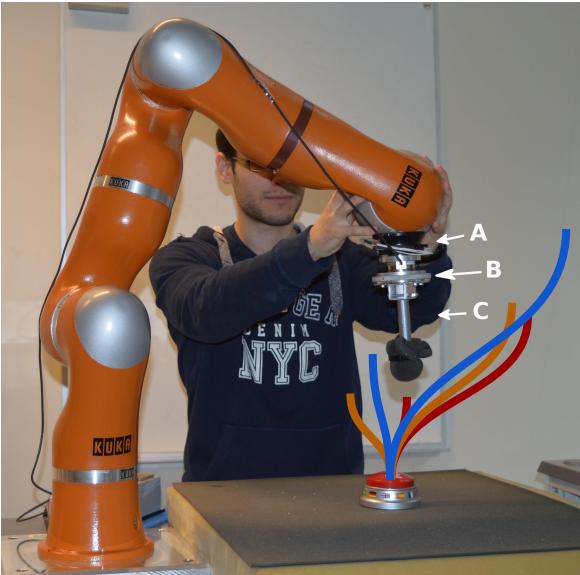


Effective teaching interface when the task is difficult to encode in a declarative way

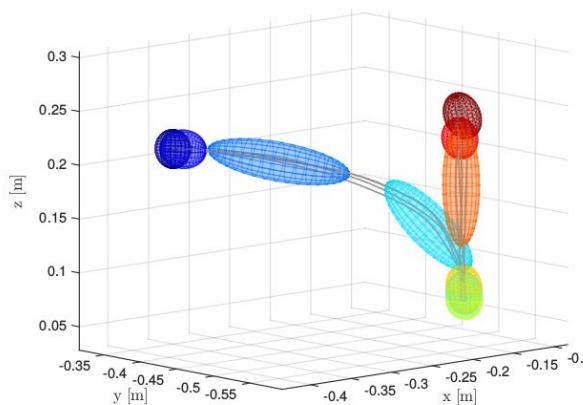
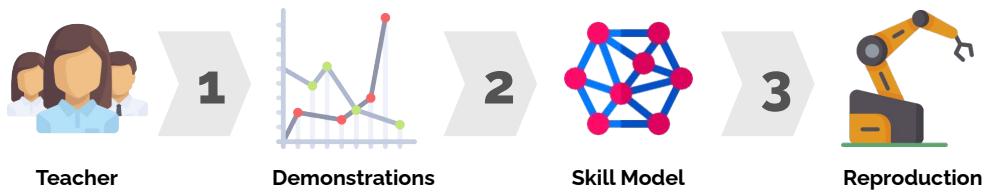
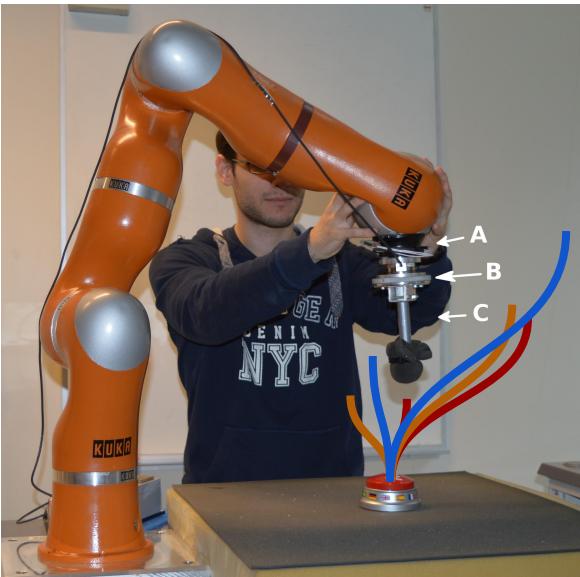
LEARNING IN-CONTACT TASKS



LEARNING IN-CONTACT TASKS



LEARNING IN-CONTACT TASKS



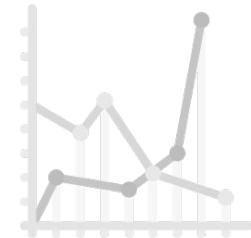
- Kinesthetic Teaching
- Hidden Semi-Markov Models (HSMM)
- Gaussian Mixture Regression (GMR)

WEAKNESSES of DEMONSTRATIONS



Teachers

1



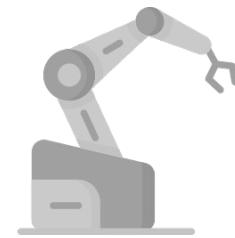
Demonstrations

2



Skill Model

3



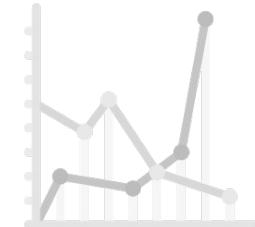
Reproduction

WEAKNESSES of DEMONSTRATIONS



Teachers

1



Demonstrations

2



Skill Model

3



Reproduction

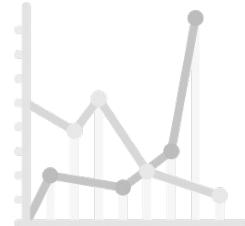
Informative Demonstrations:

demonstrations that allow the robot to learn a reliable model of the taught task

WEAKNESSES of DEMONSTRATIONS



1

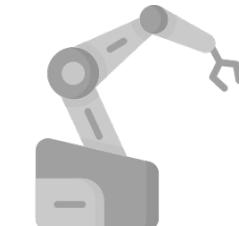


Teachers

2



Skill Model



3

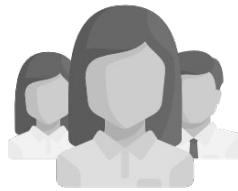
Reproduction

Informative Demonstrations:

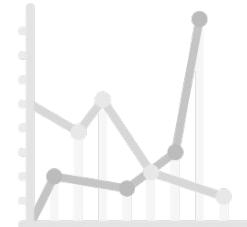
demonstrations that allow the robot to learn a reliable model of the taught task

Require some understanding of the underlying Machine Learning process!

WEAKNESSES of DEMONSTRATIONS



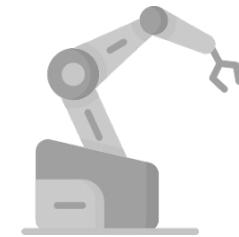
Teachers



Demonstrations

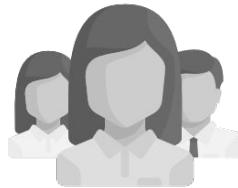


Skill Model

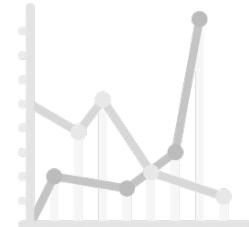


Reproduction

WEAKNESSES of DEMONSTRATIONS



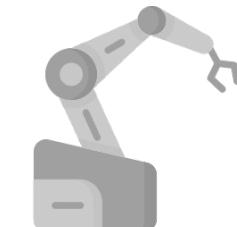
Teachers



Demonstrations



Skill Model

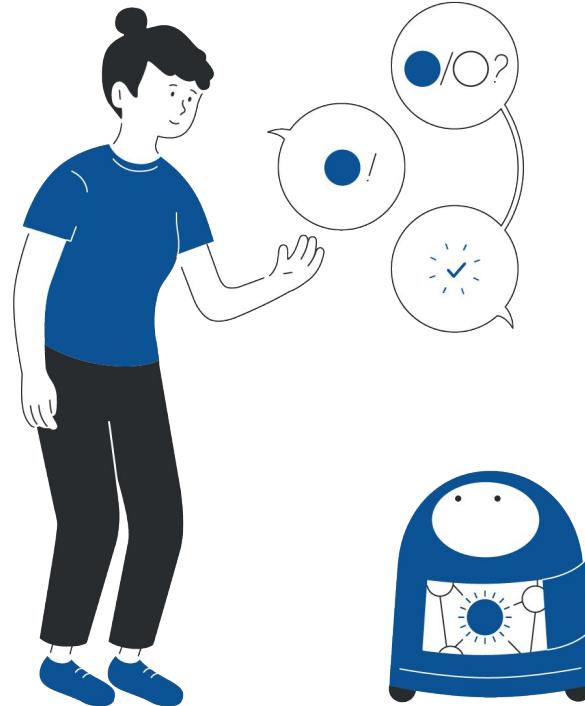


Reproduction

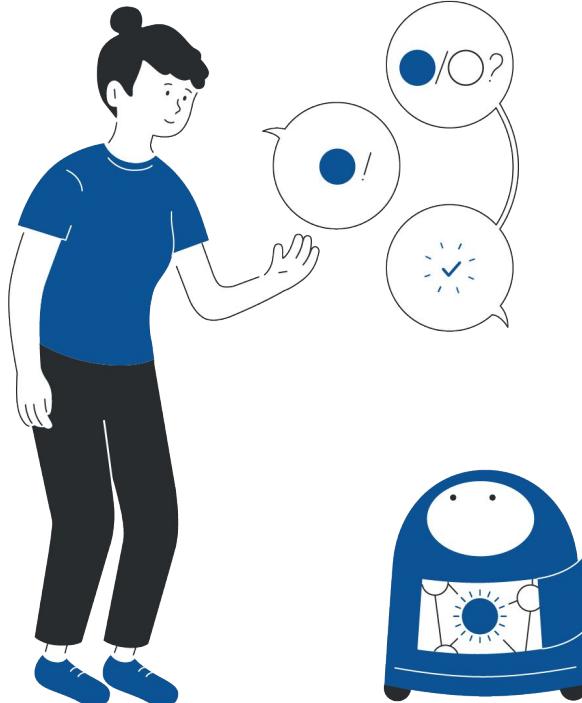
G.I.G.O.

Garbage in, Garbage out

WEAKNESSES of DEMONSTRATIONS



WEAKNESSES of DEMONSTRATIONS



ACTIVE LEARNING

The image block contains four sub-images illustrating active learning:

- (a)** A person stands at a desk with a laptop, interacting with a white humanoid robot. A computer monitor shows a user interface with a sandwich icon and text: "After getting the rye bread, do you prefer to add the ham or add the cucumber?" with options: Neither, add the ham, add the cucumber, Either, and I don't know.
- (b)** A person stands at a whiteboard with a robot arm. The whiteboard has various food icons.
- (c)** A person at a desk interacts with a white humanoid robot. A computer monitor shows a user interface asking "Do these animals have HORNs? What about lion?" with buttons for left, right, and center, and a "I don't know" option.
- (d)** A person at a desk interacts with a white humanoid robot. A computer monitor shows a user interface for a "Panda Program" with sections for "Query response buttons", "Current Primitive's parameters: Motion Speed 0.228 m/s, How was it?", and "Robot State". Buttons include "I want it slower", "I was okay like this", "I want it faster", "Go to start state", and "Execute one step".

ACTIVE LEARNING FROM QUESTIONS



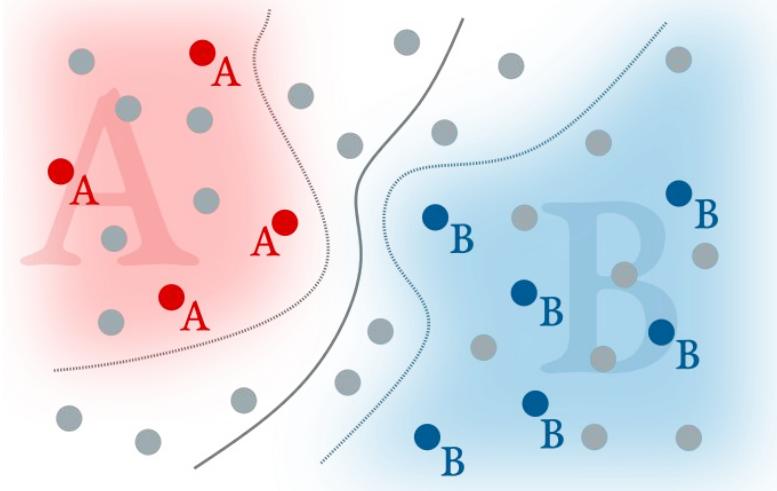
ACTIVE LEARNING FROM QUESTIONS



Queries:

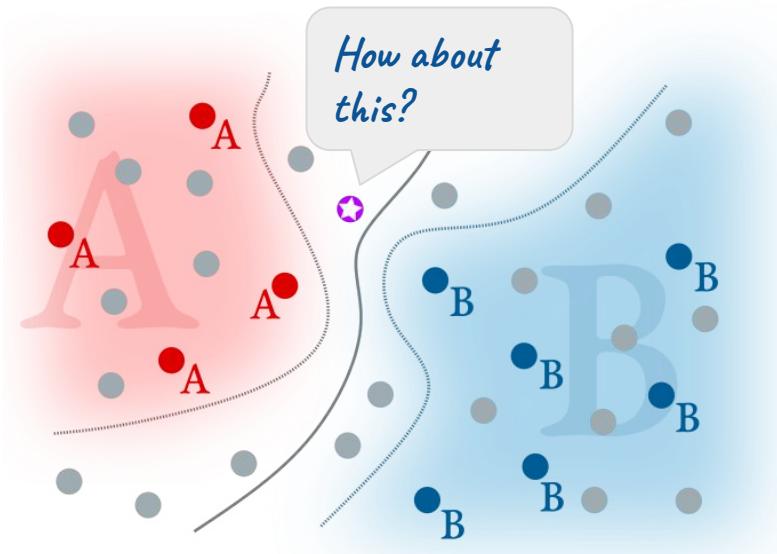
requests of information aimed at **steering the training process** to cover the current **knowledge gaps** of the learner.

ACTIVE LEARNING 101



Queries:
requests of information aimed at
steering the training process to
cover the current knowledge gaps of the learner.

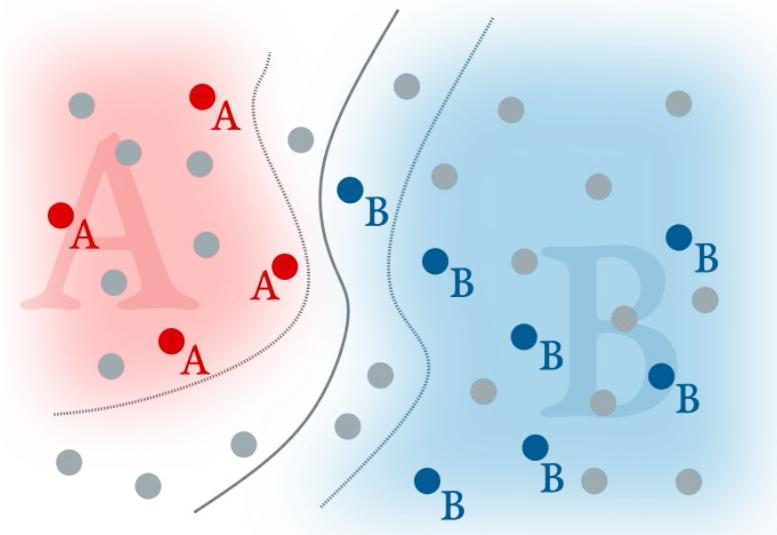
ACTIVE LEARNING 101



Queries:

requests of information aimed at
steering the training process to
cover the current **knowledge gaps** of the learner.

ACTIVE LEARNING 101



Queries:
requests of information aimed at
steering the training process to
cover the current knowledge gaps of the learner.

DEMONSTRATIONS and QUESTIONS

After getting the rye bread, do you prefer to add the ham or add the cucumber?

Neither	add the ham
add the cucumber	Either
I don't know	



DEMONSTRATIONS and QUESTIONS

After getting the rye bread, do you prefer to add the ham or add the cucumber?

Neither	add the ham
add the cucumber	Either
I don't know	



- **Demos + Queries:** the model must support both!

DEMONSTRATIONS and QUESTIONS

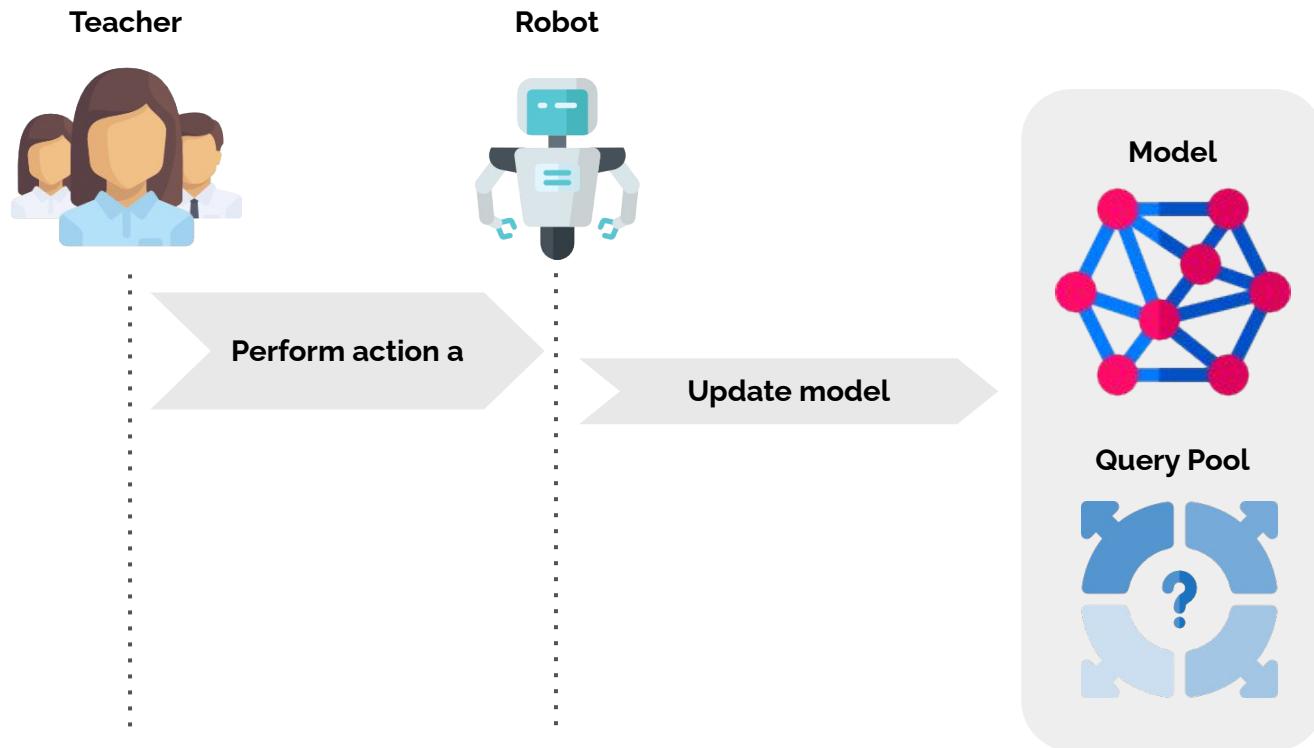
After getting the rye bread, do you prefer to add the ham or add the cucumber?

Neither	add the ham
add the cucumber	Either
I don't know	

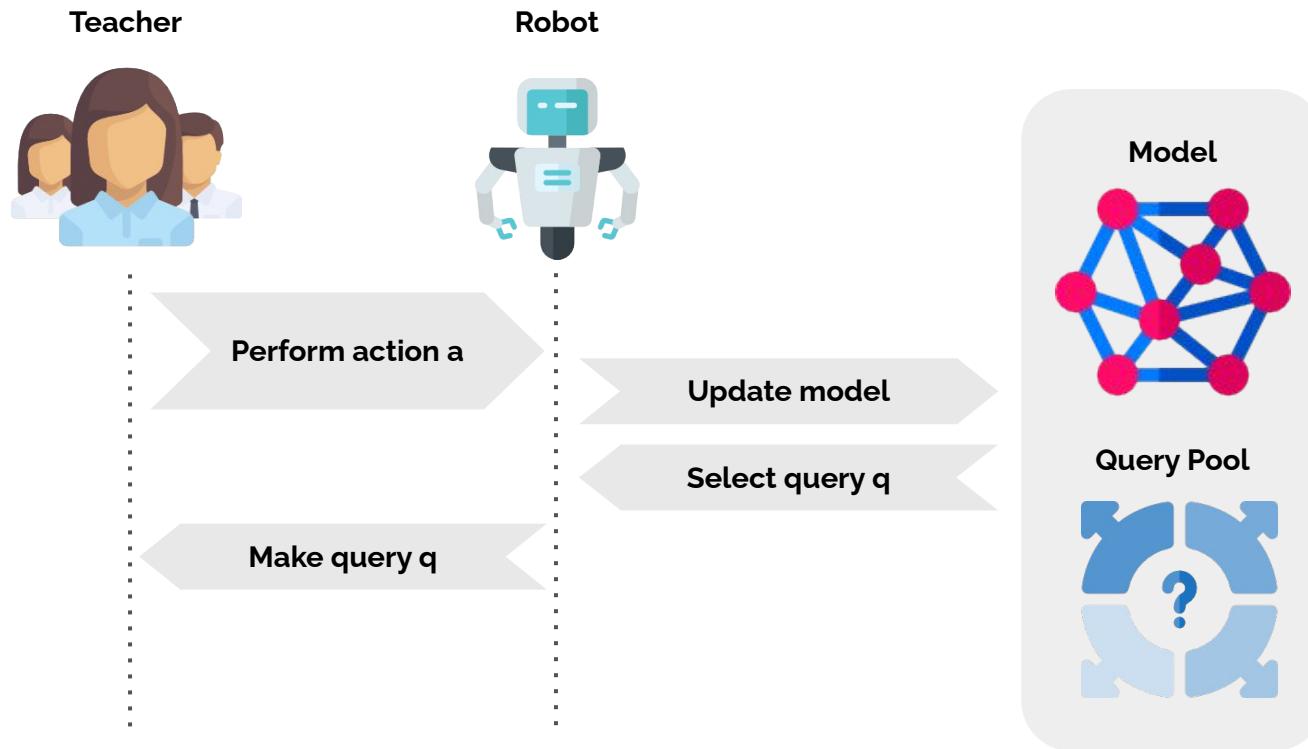


- **Demos + Queries:** the model must support both!
- **Temporal aspect:** careful design of queries

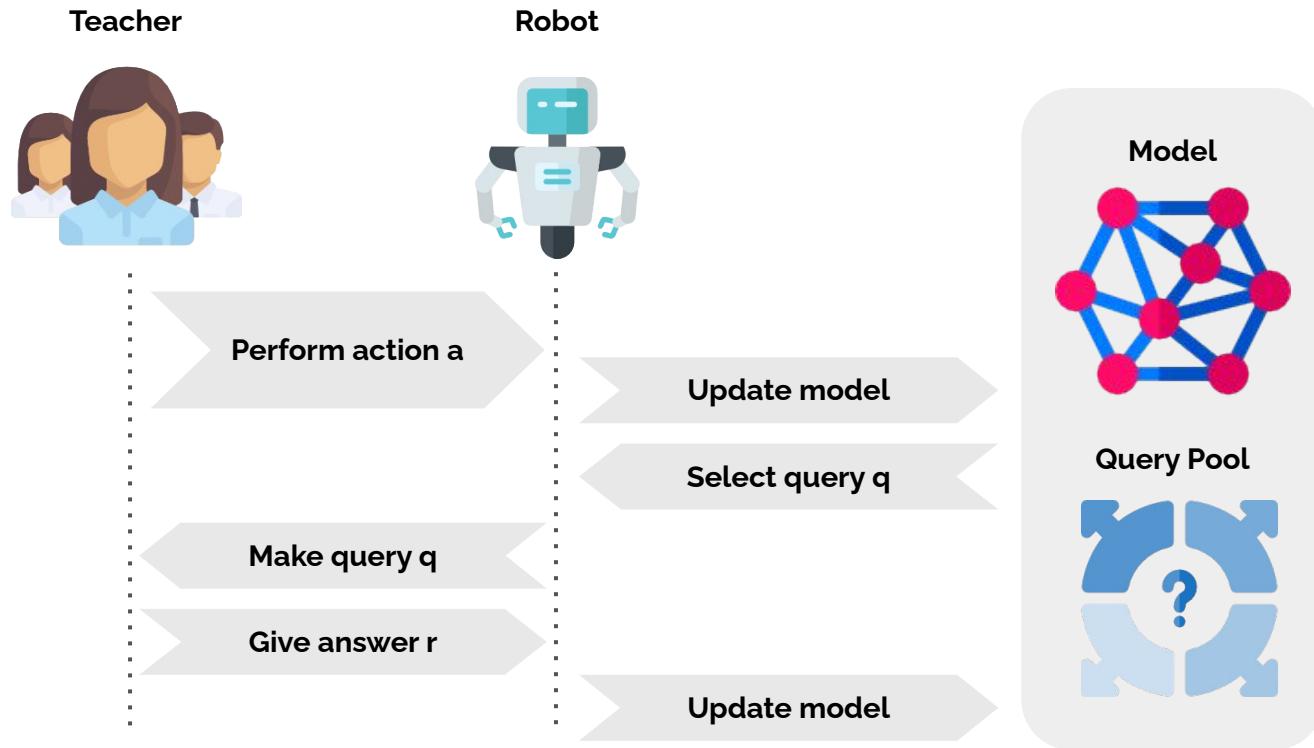
DEMONSTRATIONS and QUESTIONS



DEMONSTRATIONS and QUESTIONS



DEMONSTRATIONS and QUESTIONS



QUERY DESIGN

QUERY DESIGN

“With what probability
you do action A after
action B?”

“Do you do action A
after action B with
probability 0.3?”

QUERY DESIGN

“With what probability
you do action A after
action B?”

“Do you do action A
after action B with
probability 0.3?”

Frequency Queries

“Do you **always**
/often/never do
action A after B?”

Disambiguation Queries

“After action B, **do**
you prefer to do
action A or C?”

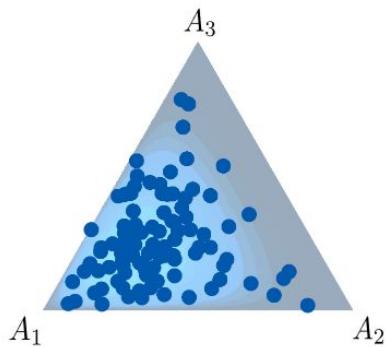
Model Friendly

User Friendly



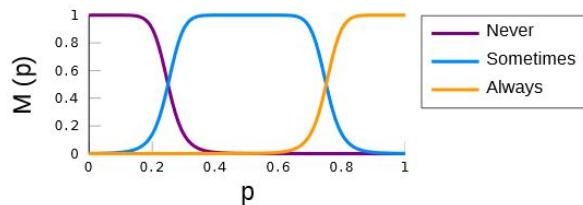
MODEL UPDATE

Question q



1. Sample the pre-query Dirichlet

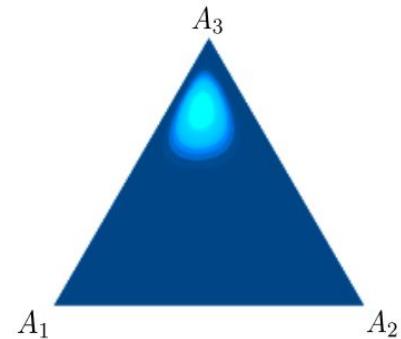
Answer a



$$w_s(q^*, r^*) = \begin{cases} M_f(s(a_{post})) & \text{if } r^* = \text{'yes'} \\ 1 - M_f(s(a_{post})) & \text{if } r^* = \text{'no'} \end{cases}$$
$$w_s(q^*, r^*) = M_{r^*}(s(a_1), s(a_2))$$

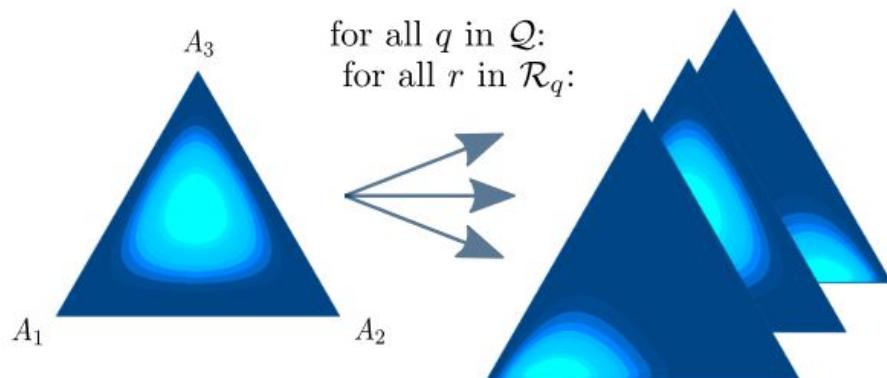
2. Filter/Weight samples based on answer

Updated model



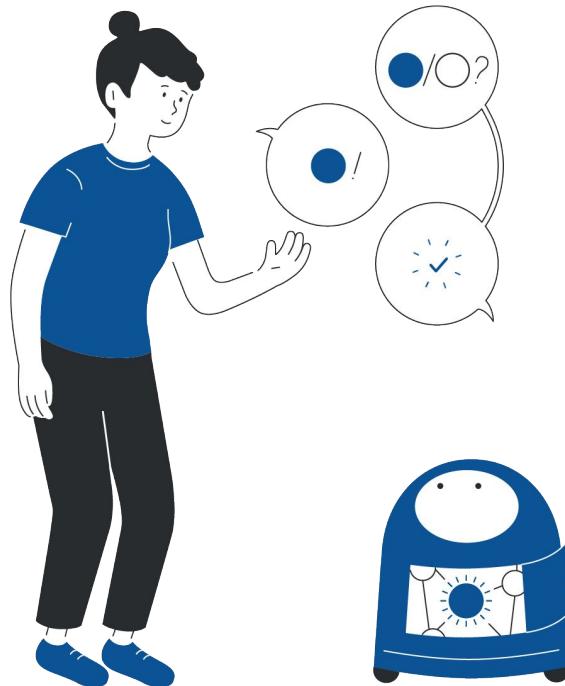
3. Fit the post-query Dirichlet

QUERY SELECTION

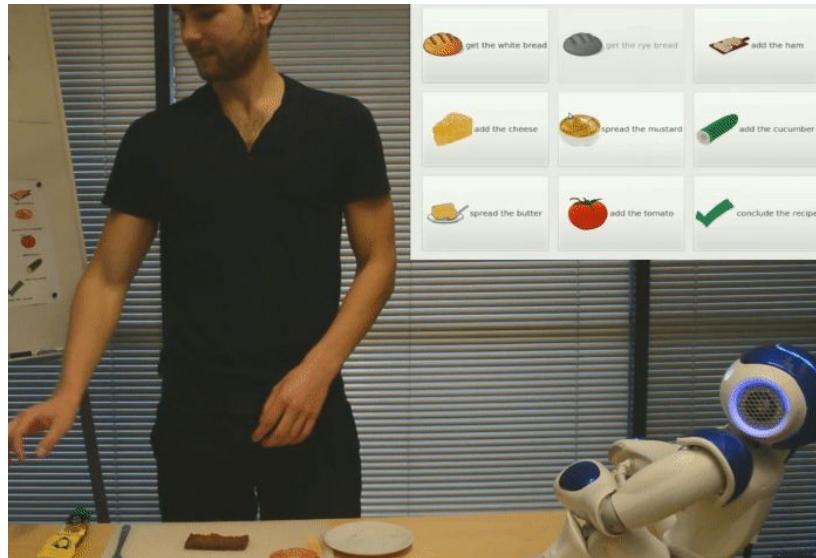


$$\begin{aligned}\Delta \mathbb{H}_q &= \underbrace{\mathbb{E}_{r}[\mathbb{H}(\text{Dir}(\cdot | \boldsymbol{\alpha}, q, r))]}_{\text{post-query}} - \underbrace{\mathbb{H}(\text{Dir}(\cdot | \boldsymbol{\alpha}))}_{\text{pre-query}} \\ &= \sum_r p(r|q) \mathbb{H}(\text{Dir}(\cdot | \boldsymbol{\alpha}, q, r)) - \mathbb{H}(\text{Dir}(\cdot | \boldsymbol{\alpha})),\end{aligned}$$

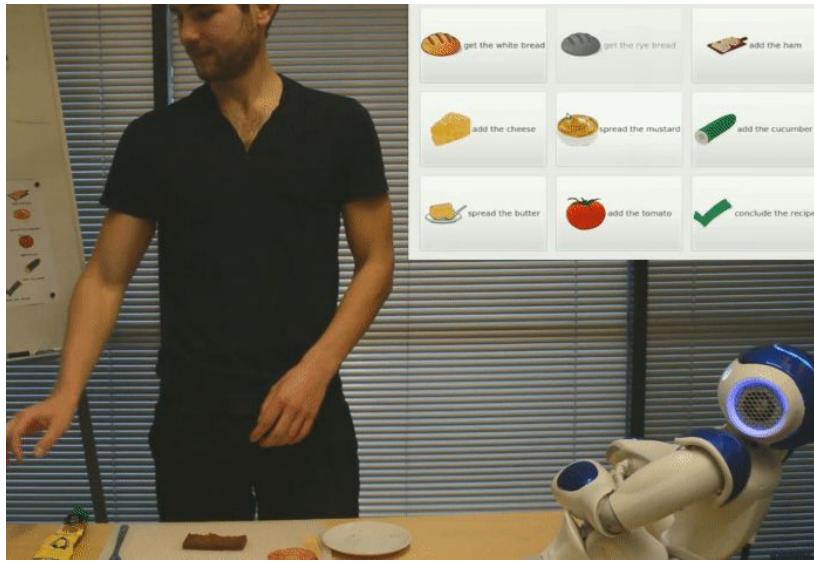
INTERACTING with an ACTIVE LEARNING ROBOT



INTERACTING with an ACTIVE LEARNING ROBOT



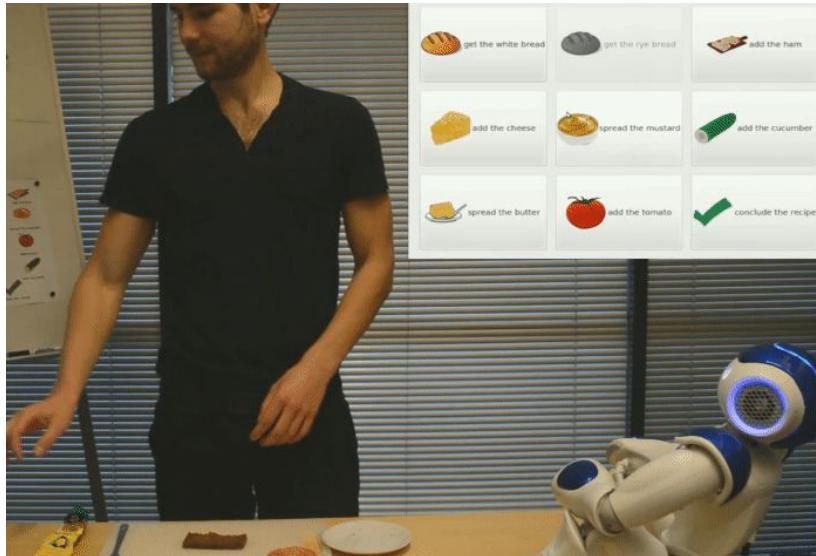
INTERACTING with an ACTIVE LEARNING ROBOT



User study:

- perception of robots using different selection strategies and effects on the teacher

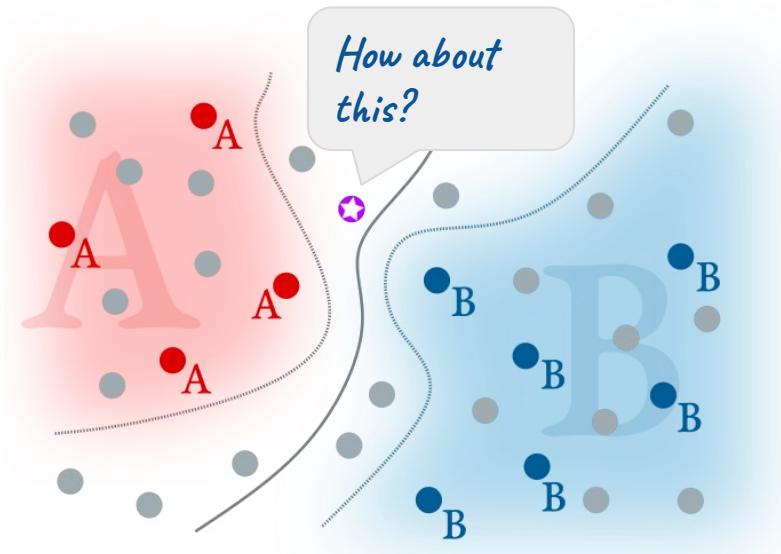
INTERACTING with an ACTIVE LEARNING ROBOT



User study:

- perception of robots using different selection strategies and effects on the teacher
- interpretation of learning behaviours, with frequent mismatches!

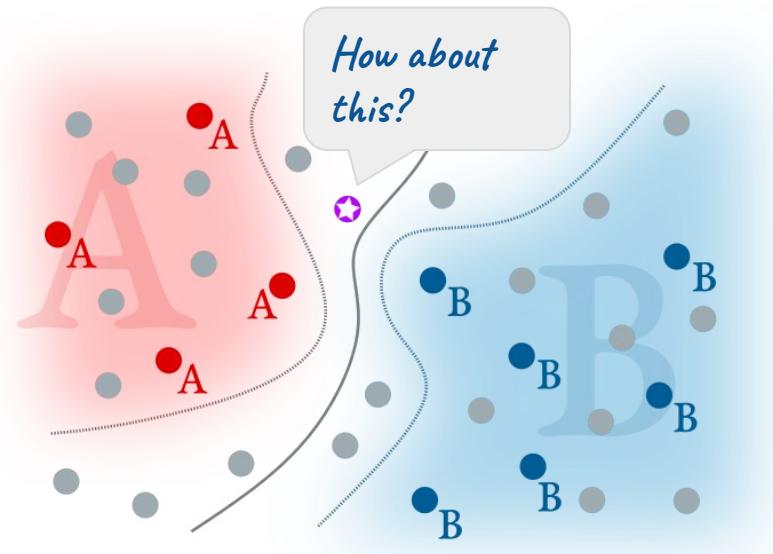
ROBOT QUERIES and REAL TEACHERS



Sample Efficiency:

- learning faster and/or with less data

ROBOT QUERIES and REAL TEACHERS



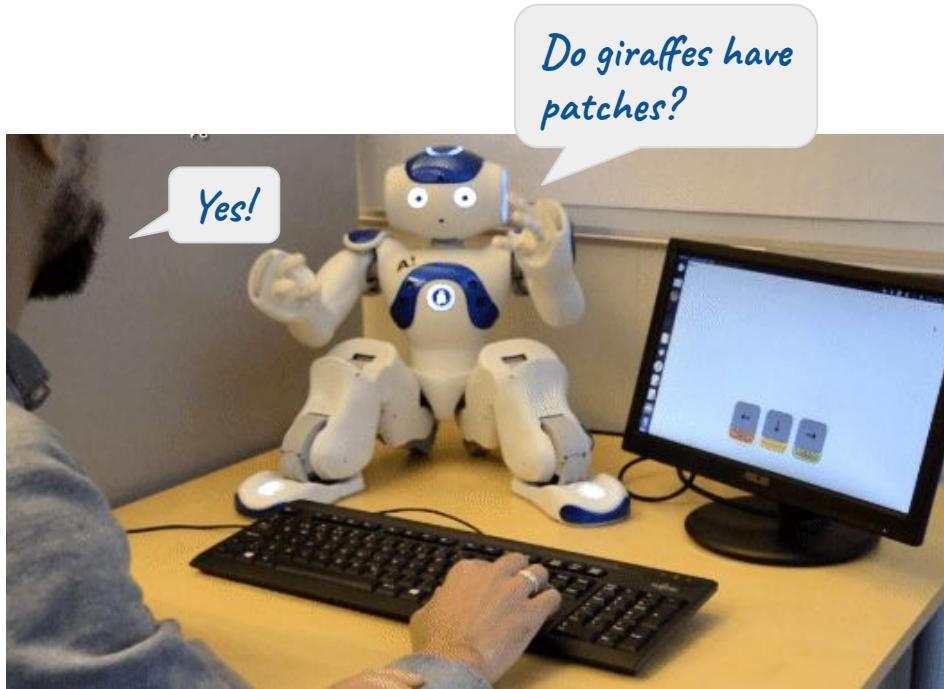
Sample Efficiency:

- learning faster and/or with less data

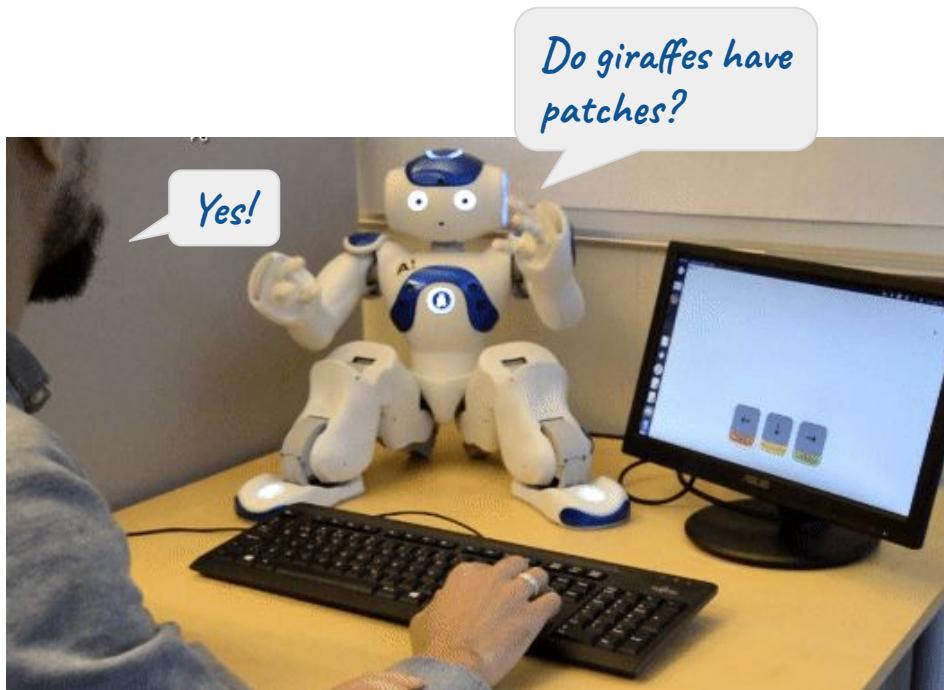
What if the efficient query selection is not the best for the teacher?

- cause errors and delays

ROBOT QUERIES and REAL TEACHERS

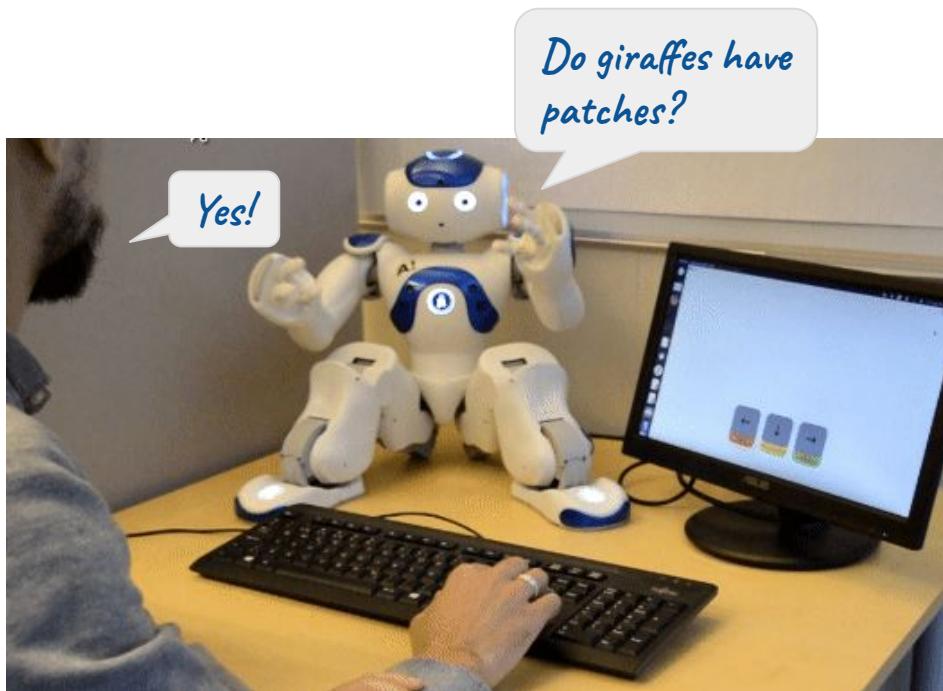


ROBOT QUERIES and REAL TEACHERS



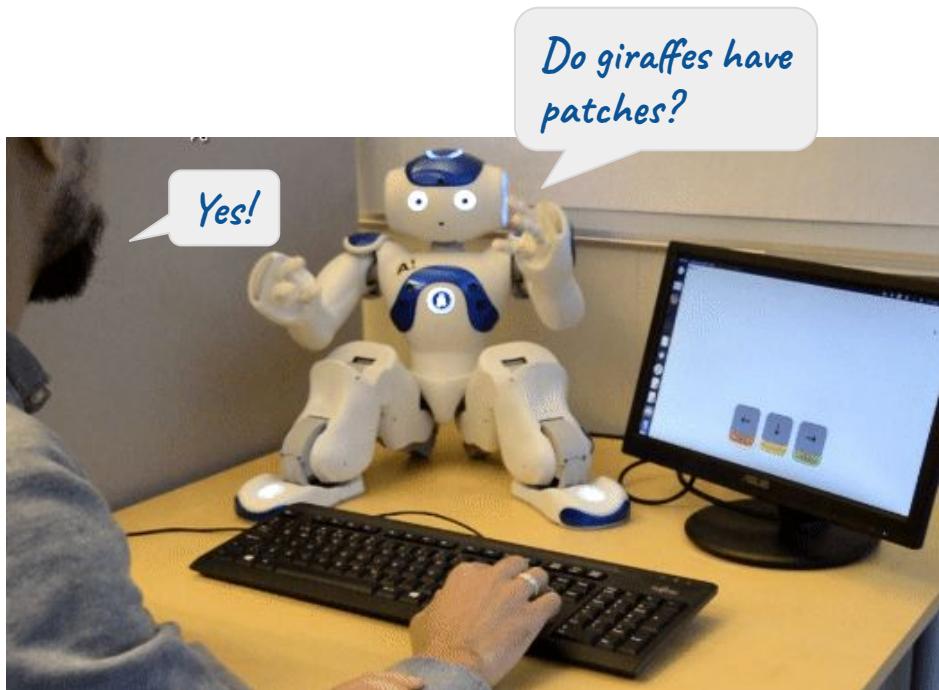
- ACTIVE LEARNING STRATEGY

ROBOT QUERIES and REAL TEACHERS



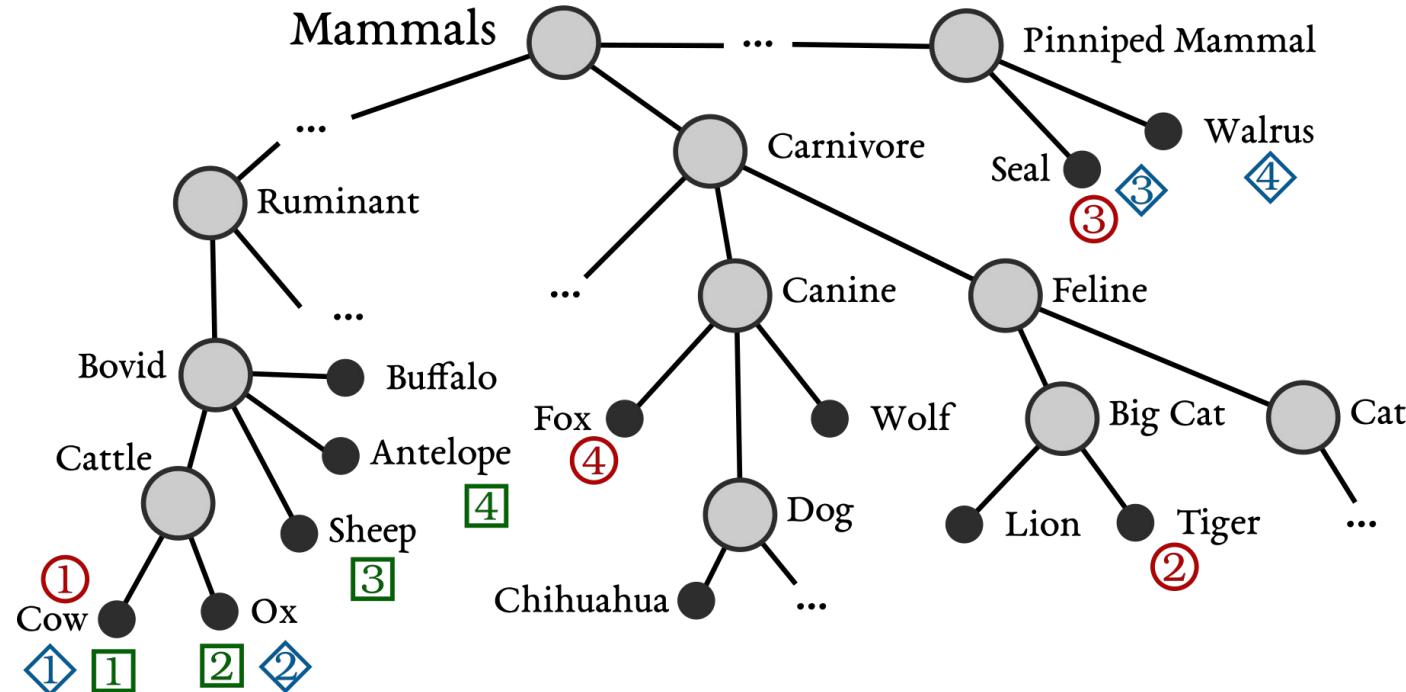
- **ACTIVE LEARNING STRATEGY**
- **MEMORY-AWARE STRATEGY**

ROBOT QUERIES and REAL TEACHERS



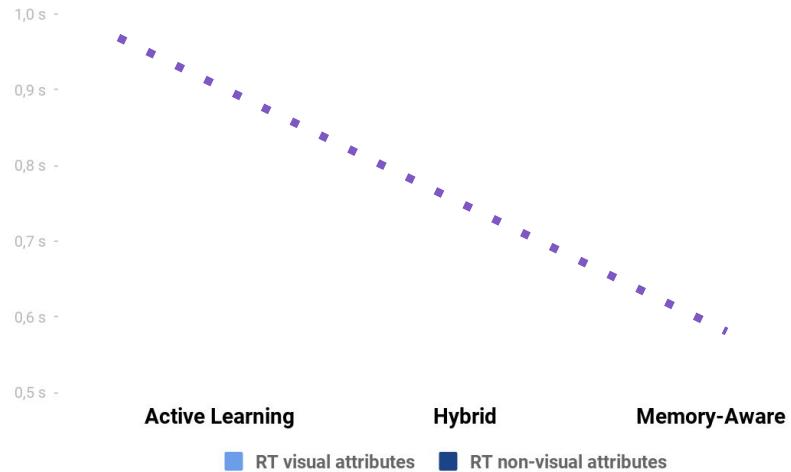
- **ACTIVE LEARNING STRATEGY**
- **MEMORY-AWARE STRATEGY**
- **HYBRID STRATEGY**

ROBOT QUERIES and REAL TEACHERS



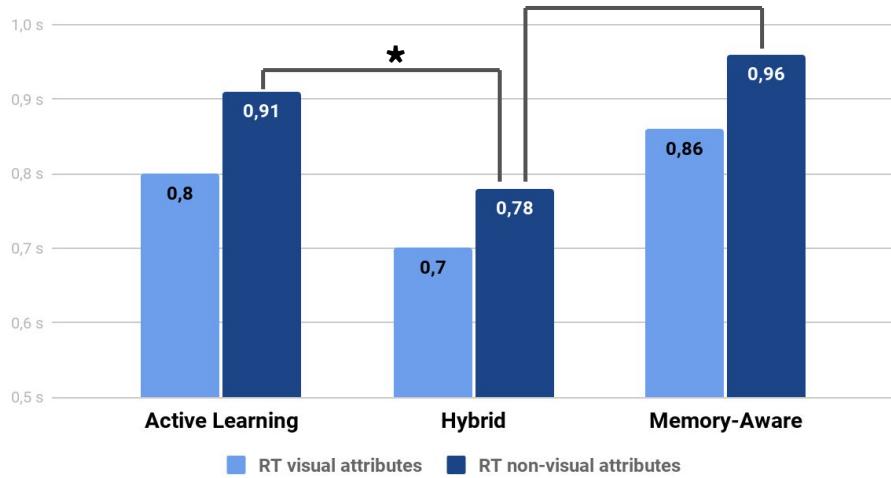
HYPOTHESES

Response Time

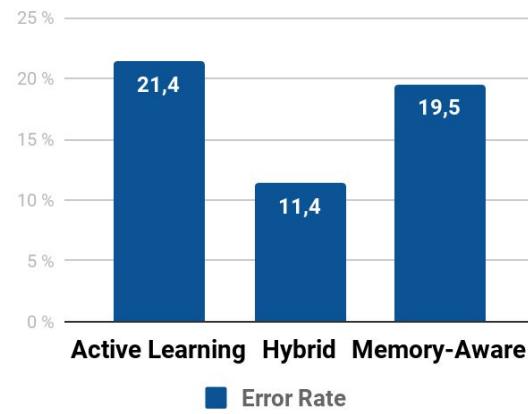


UNEXPECTED RESULTS

Response Time

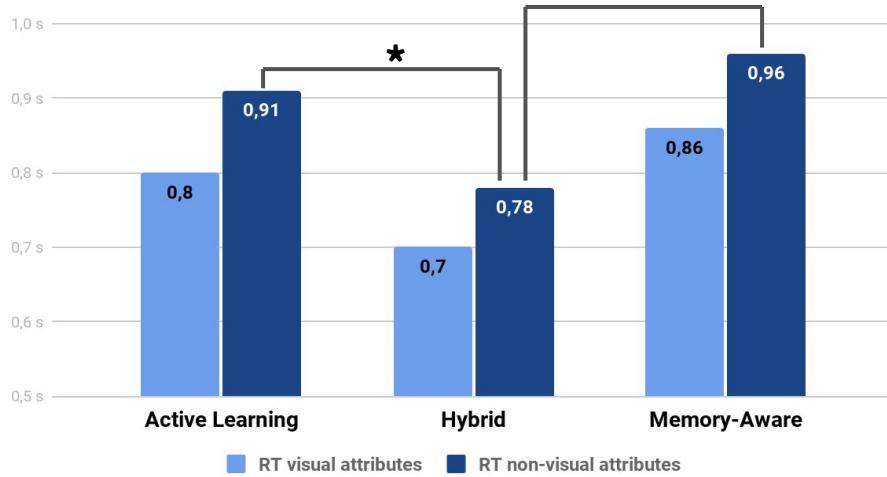


Error Rate

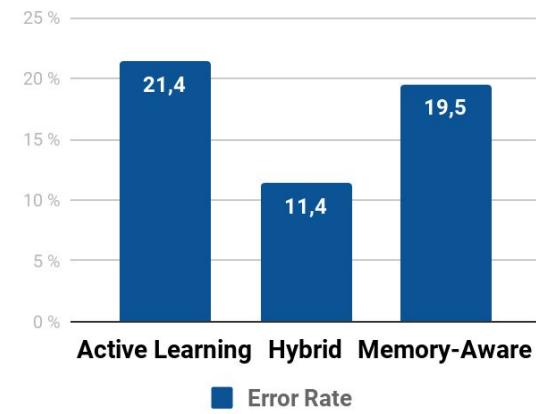


UNEXPECTED RESULTS

Response Time



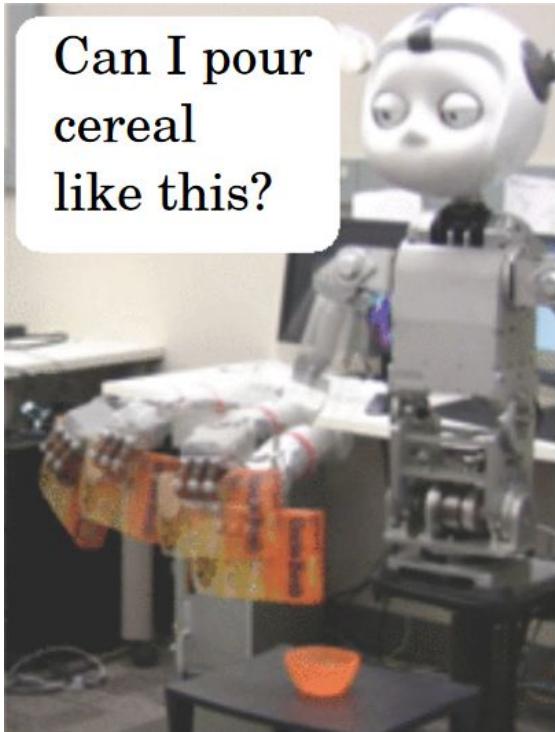
Error Rate



Different strategies drastically impact the human teacher!

EMBODIED QUESTIONS

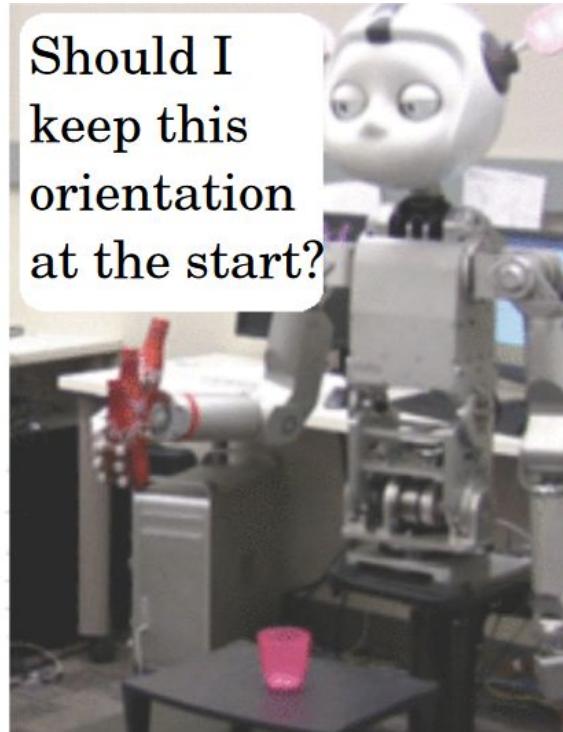
Can I pour
cereal
like this?



Can you
show me
how to add
salt
from here?



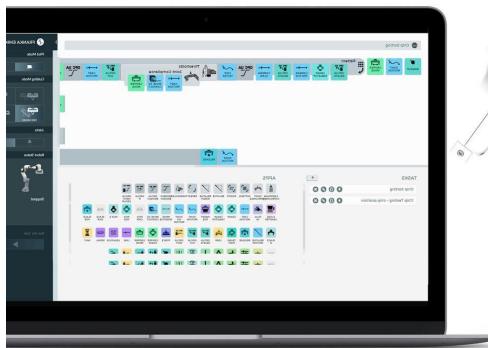
Should I
keep this
orientation
at the start?



EMBODIED QUESTIONS for ROBOT PROGRAMMING



EMBODIED QUESTIONS for ROBOT PROGRAMMING

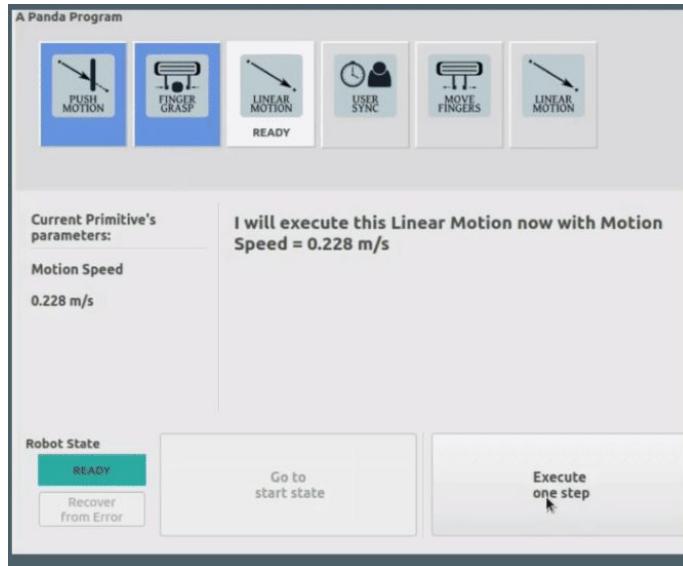


PARAMETERS

- Goal Pose
- Translational Speed
- Collision Threshold (move until you sense a contact)



TUNING ROBOT PROGRAMS



TUNING ROBOT PROGRAMS

User study:

- tune parameters faster and closer to how experts would tune them



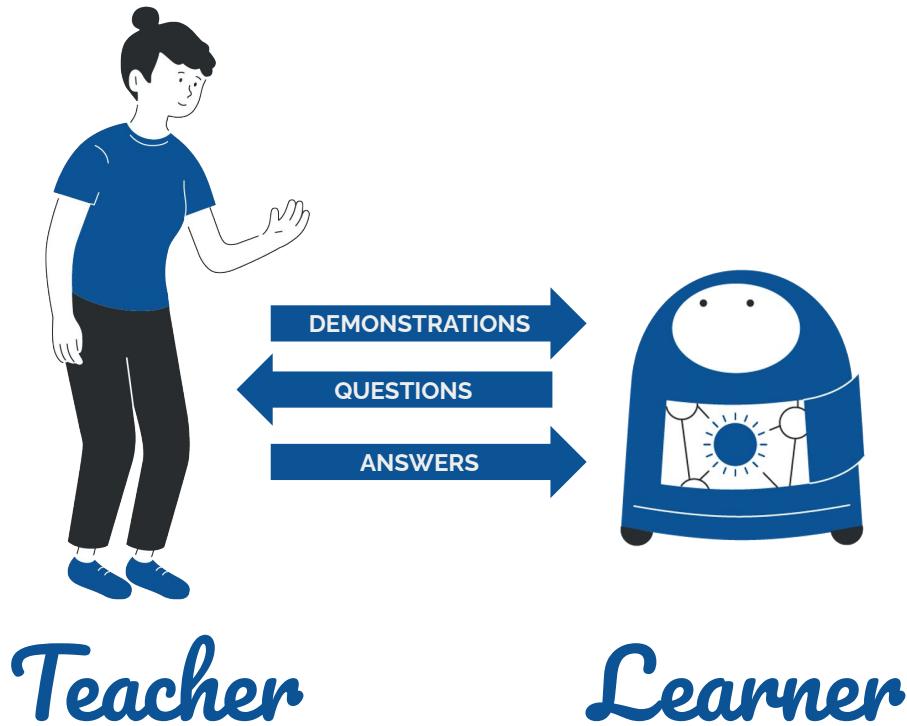
TUNING ROBOT PROGRAMS

User study:

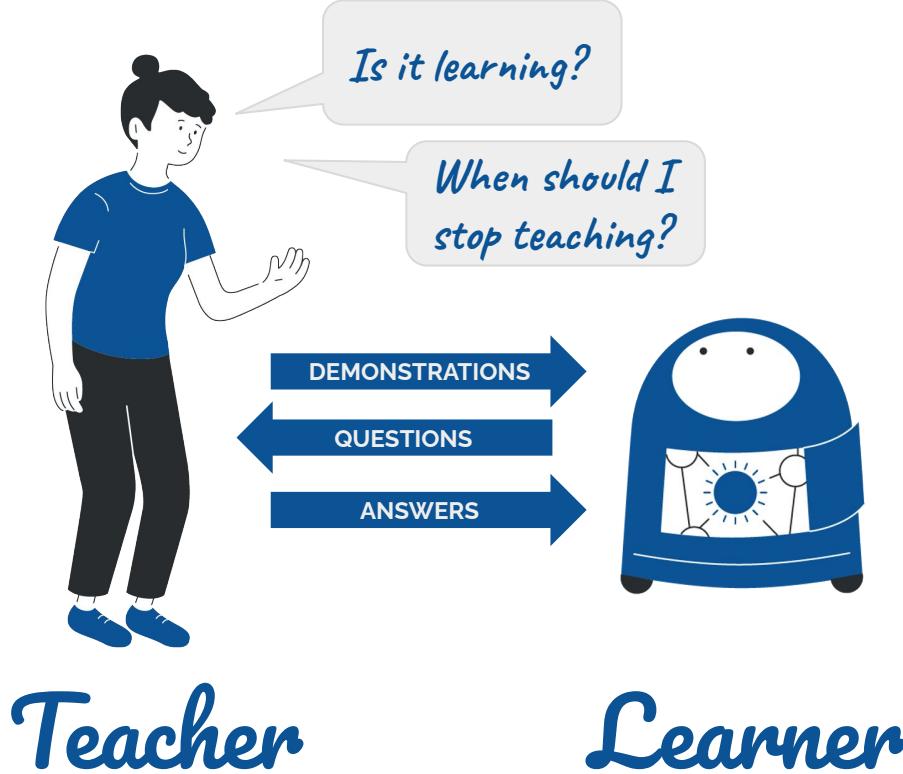
- tune parameters faster and closer to how experts would tune them
- integration of Active Learning in commercially available robot programming interface



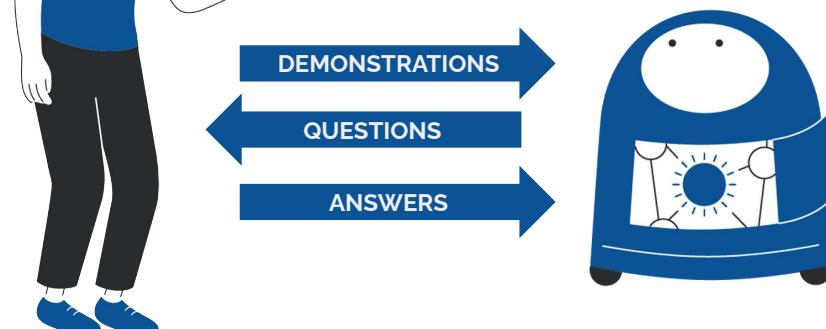
TEACHER-LEARNER TRANSPARENCY



TEACHER-LEARNER TRANSPARENCY



TEACHER-LEARNER TRANSPARENCY

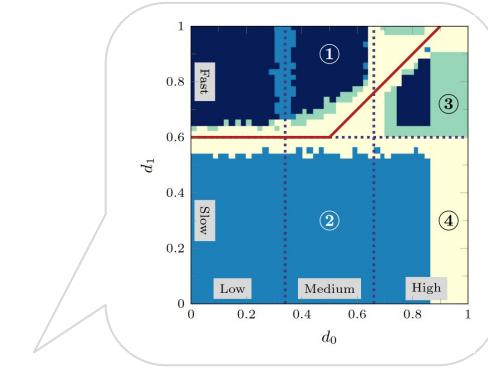


Teacher Learner

TEACHER-LEARNER TRANSPARENCY

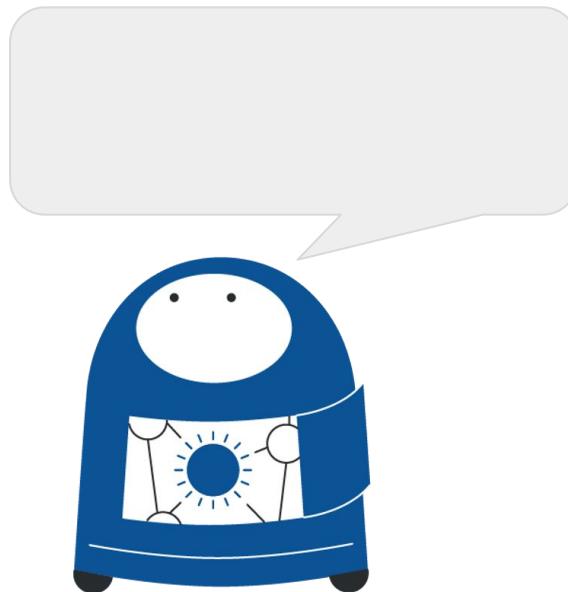
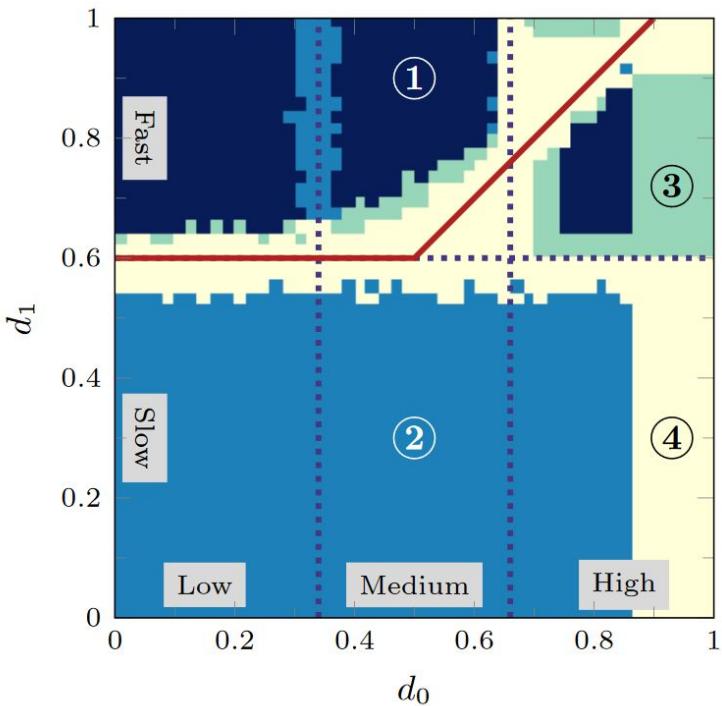
I think it's enough!

SPOILER: It was not.

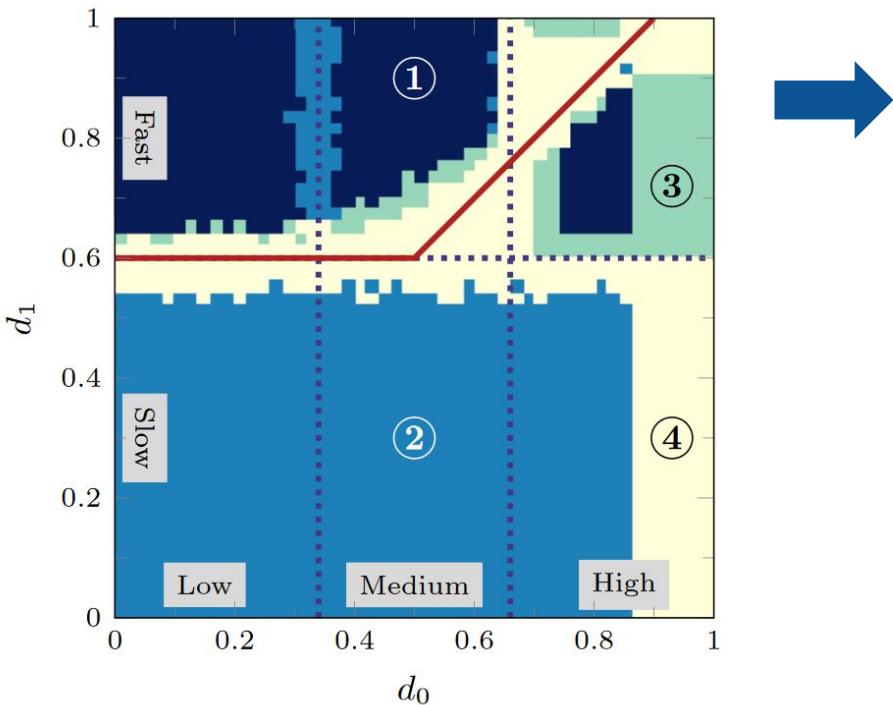


Teacher Learner

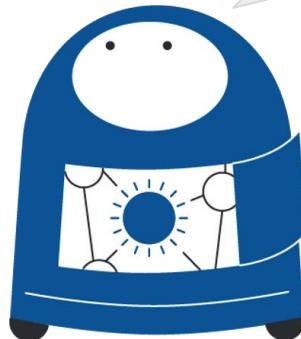
POLICY EXPLANATION



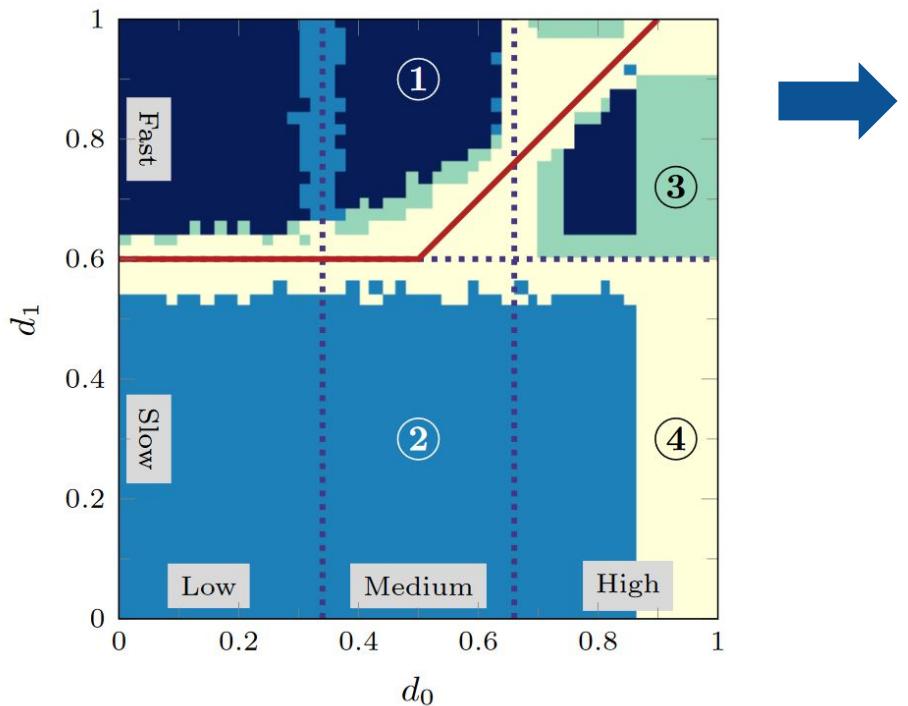
POLICY EXPLANATION



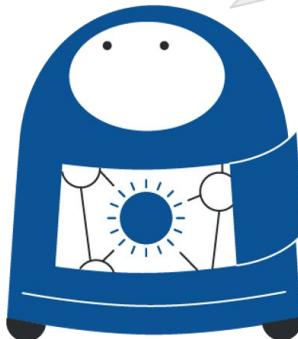
*I go to charge because my
battery is low and today
it's Friday and ...*



FOCUSED POLICY EXPLANATION



*I go to charge because my
battery is low ~~and today~~
~~it's Friday and ...~~*



FOCUSED POLICY EXPLANATION

1 Mars Rover - learning phase

Battery Level	Ground Quality	Signal Strength	Storage	Temperature
low middle high 92.0	low high 0.31	low middle high 0.83	low high 25.5	low middle high 9.3

2

Possible Actions:

- 1: move
- 2: stop and charge
- 3: stop and collect ground sample
- 4: send data to earth
- 5: return and unload the collected samples



3

Explanations:
The action was: move

**"I did move because Battery Level was high
and Ground Quality was low"**

FOCUSED POLICY EXPLANATION

1 Mars Rover - learning phase

Battery Level	Ground Quality	Signal Strength	Storage	Temperature
low middle high 92.0	low high 0.31	low middle high 0.83	low high 25.5	low middle high 9.3

2

Possible Actions:

- 1: move
- 2: stop and charge
- 3: stop and collect ground sample
- 4: send data to earth
- 5: return and unload the collected samples



3

Explanations:
The action was: move

"I did move because Battery Level was high
and Ground Quality was low"

Better understanding of
robot policies!

but we have a long way to go
before we have **Interpretable
Machine Learning!**

Interactive Robot Learning with human-in-the-loop

Interactive Robot Learning with human-in-the-loop

**Active Learning
Learning from Demonstration**

Interactive Robot Learning with human-in-the-loop

**Active Learning
Learning from Demonstration**

Focus on the Human-Robot Interaction

Teacher-Learner Interaction for Robot Active Learning

