

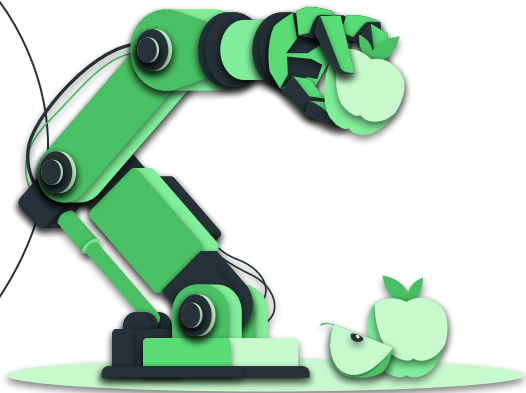
Interactive Policy Shaping for Human-Robot Collaboration with Transparent Matrix Overlays

Jake Brawer, Debasmita Ghose, Kate Candon, Meiyang Qin, Alessandro Roncone, Marynel Vázquez, Brian Scassellati

ACM/IEEE International Conference on
Human-Robot Interaction 2023 Stockholm, SE
(p. 525-533)

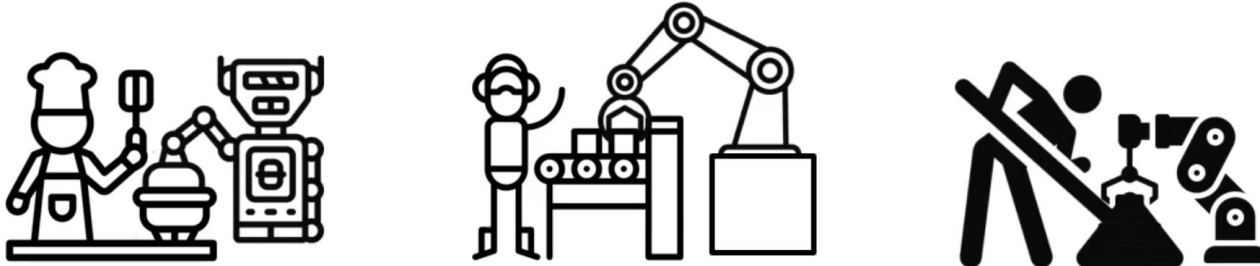
Best Technical Paper 

Guilherme Santos – 113893
Pedro Pinto – 115304



Motivation

- Human-Robot Collaboration (HRC)
- Reinforcement Learning (RL)
- Adaptation to user preferences
- Learning via shielding

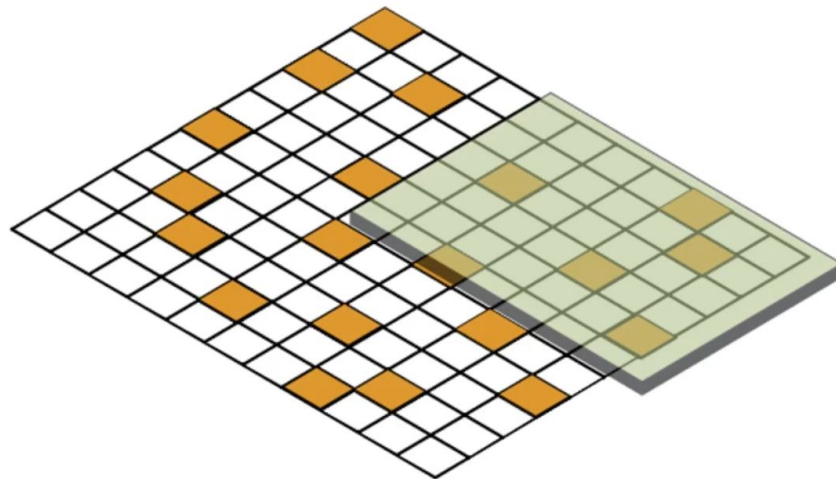


Learning Policies via shielding: <https://doi.org/10.1109/HRI53351.2022.9889604>

Transparent Matrix Overlays

Overview

- Q-value matrix
- Policy



Transparent Matrix Overlays

Types of overlays

Prohibitory Overlay

rules that prohibit actions

Transfer Overlay

transfer probability density

Permissive Overlay

rules that permit actions

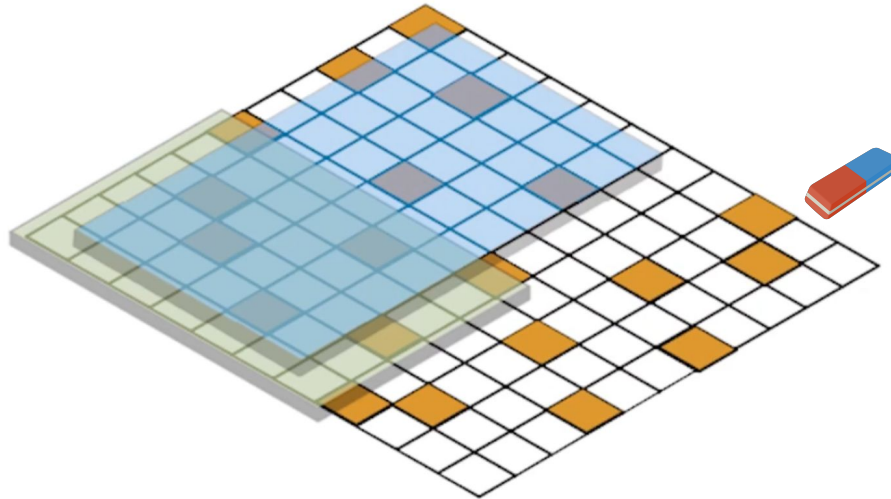
Algorithm 1: Modifying a policy by adding overlays

```
1 Input: policy  $\pi$ , state  $s$ , actions  $A$ , directive  $d$ , overlay list  $O$ 
2  $\tau \leftarrow$  Confidence Threshold
3  $l_{pre}, l_{post}, type \leftarrow processDirective(d)$ 
4  $O.append((l_{pre}, l_{post}, type))$ 
5  $O_{sorted} \leftarrow sortByType(O)$ 
6  $\hat{\pi} \leftarrow \pi$ 
7 for  $a_i \in A$  do
8   for  $l_{pre}, l_{post}, type \in O_{sorted}$  do
9     if  $type == Prohibit \ \& \ l_{pre} > \tau$  then
10       $\hat{\pi}(a_i|s) = (1 - l_{post}(a_i, s))\hat{\pi}(a_i|s)$ 
11     else if  $type == Transfer \ \& \ l_{pre} > \tau$  then
12       $l_{post}^* \leftarrow \max_{a_j \in A \setminus a_i} l_{post}(s, a_i, a_j)$ 
13       $\hat{\pi}(a_i|s) = (1 - l_{post}^*)\hat{\pi}(a_i|s) + l_{post}^*\hat{\pi}(a_j|s)$ 
14     else if  $type == Permit \ \& \ l_{pre} > \tau$  then
15       $\hat{\pi}(a_i|s) = l_{post}(a_i, s)\hat{\pi}(a_i|s)$ 
16   end
17 end
18  $\hat{\pi}(A|s) \leftarrow \frac{\hat{\pi}(A|s)}{\sum_{a \in A} \hat{\pi}(a|s)}$ 
19 Return:  $\hat{\pi}(A|s)$ 
```

Transparent Matrix Overlays

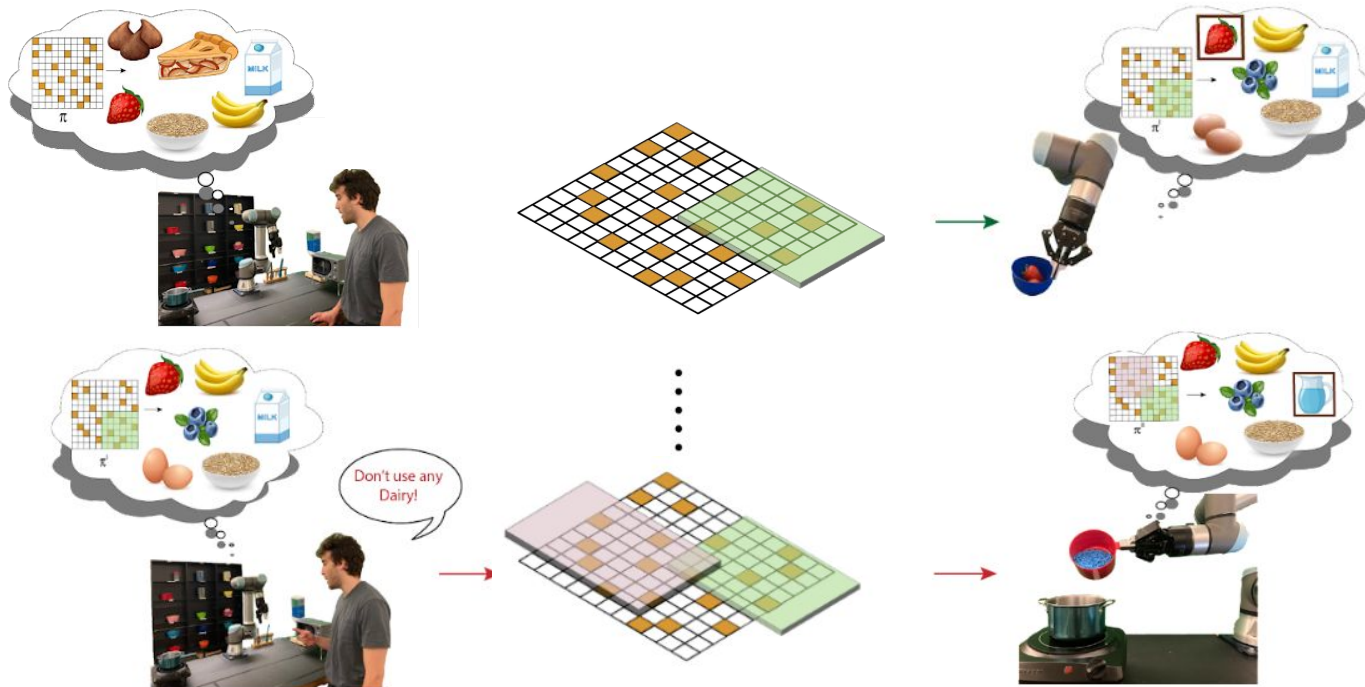
Composability and Removing

- Overlay List



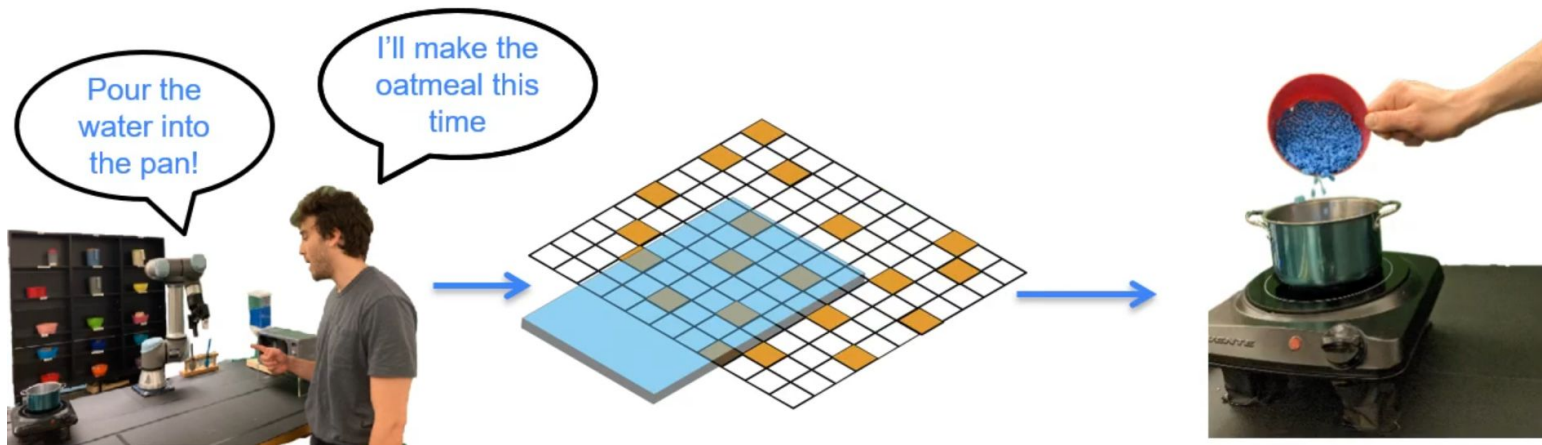
Transparent Matrix Overlays

Prohibitory and Permissive Overlays



Transparent Matrix Overlays

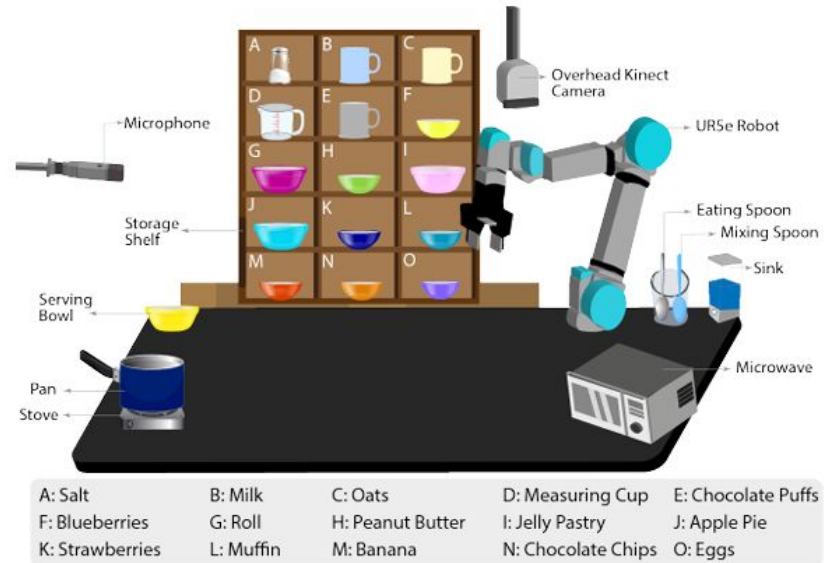
Transfer Overlays



Experiments

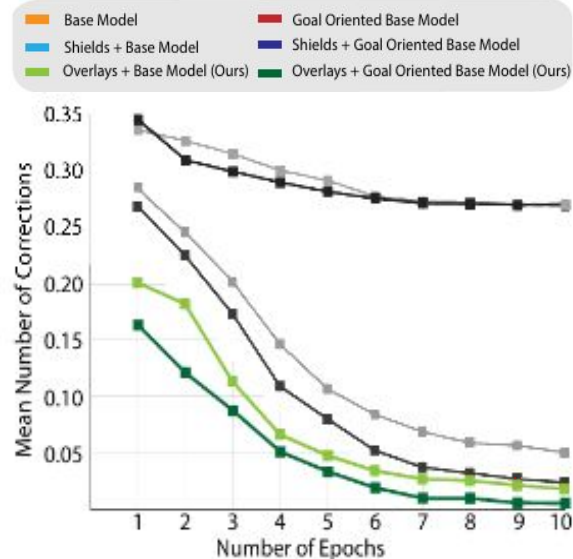
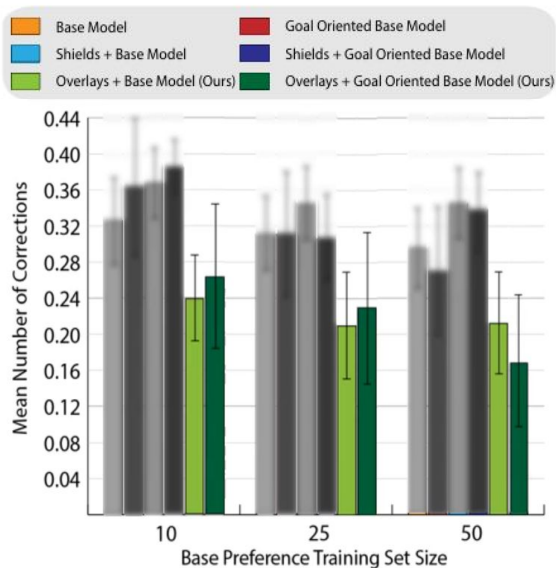
- Assistive cooking
- Rapid and temporary changes
- Adaptive and flexible policies
- Environment

Simulated vs Physical



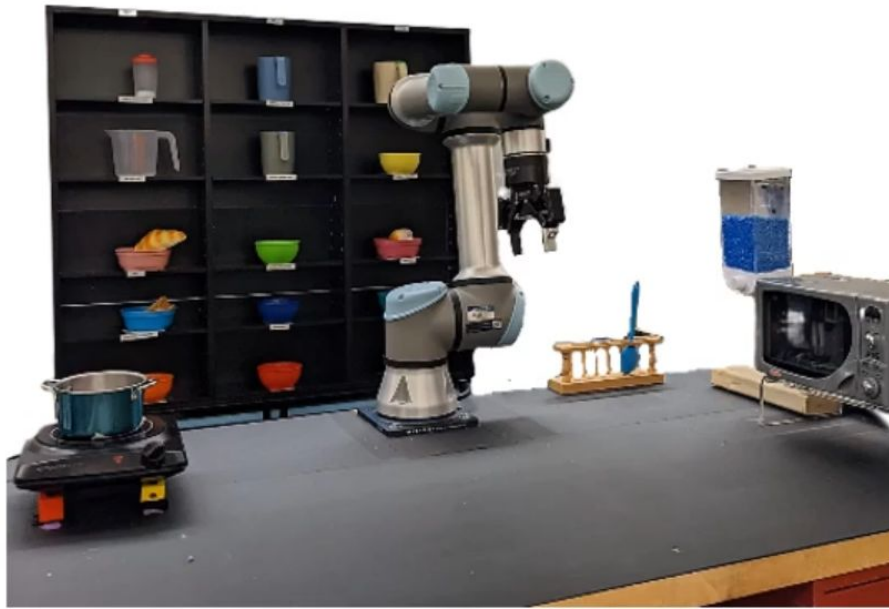
Simulated Tests

No Retraining Test vs Retraining Test



Physical Robot Tests

Proof-of-concept experiments



(video)

Physical Robot Tests

Results of the Case Study

Number of Errors	Base Model	Overlay-assisted robot	
All Types	14	2	-700%
(previous slide)	6	2	-300%
only Permissive	3	0	-300%

- 3 experienced users in each test
- same base policy form simulated tests

Conclusion

- Immediate adaptation to user preferences
- Physical Tests with few users
- Limitations:
 - Hand-crafted predicates
 - Scalability

What's next?

Interactive Policy Shaping for Human-Robot Collaboration with Transparent Matrix Overlays

Jake Brawer, Debasmita Ghose, Kate Candon, Meiying Qin, Alessandro Roncone, Marynel Vázquez, Brian Scassellati

ACM/IEEE International Conference on
Human-Robot Interaction 2023 Stockholm, SE
(p. 525-533)

Best Technical Paper 

Questions?

Guilherme Santos – 113893
Pedro Pinto – 115304

