Learning Representations that Enable Generalizations in Assistive Tasks

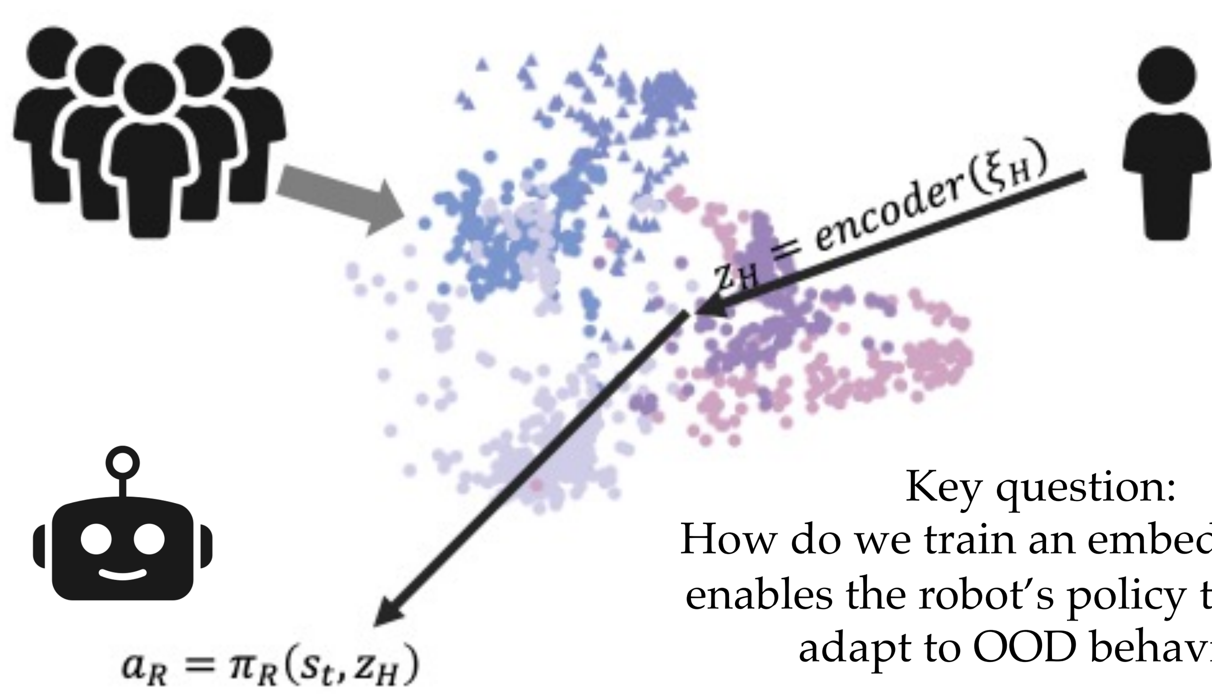
Jerry Zhi-Yang He, Aditi Raghunathan, Daniel S. Brown, Zackory Erickson, Anca D. Dragan

Goal: adapt assistance to new humans

Given: a population of human models

Find: a robot policy that achieves good *zero-shot* and *few-shot* performance on ***out-of-distribution*** human behavior

Solution: learn a human behavior embedding

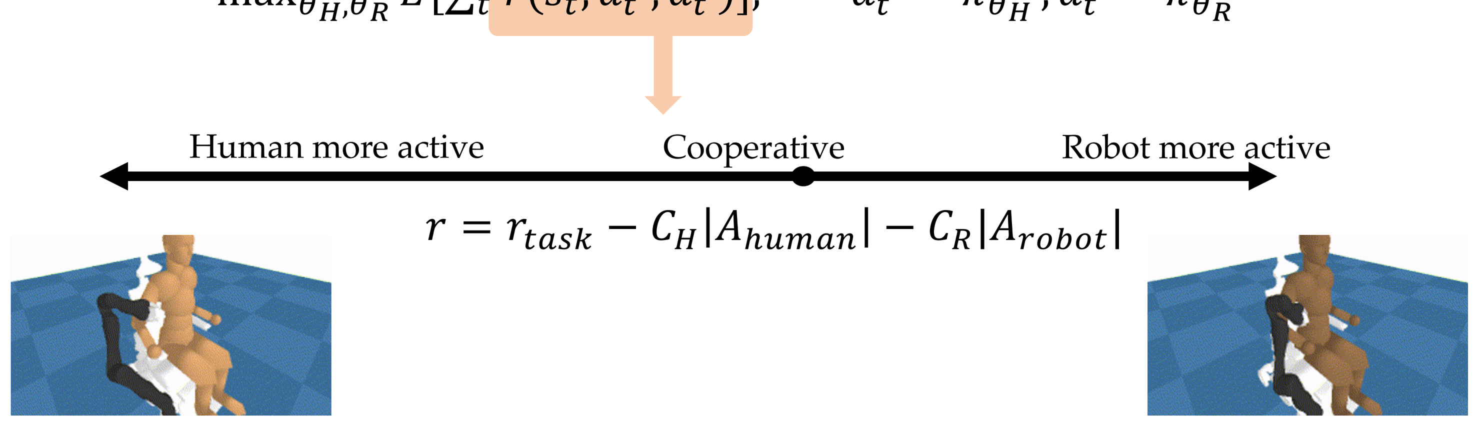


Key question:
How do we train an embedding that enables the robot's policy to quickly adapt to OOD behavior?

Human Models: synthetic generation

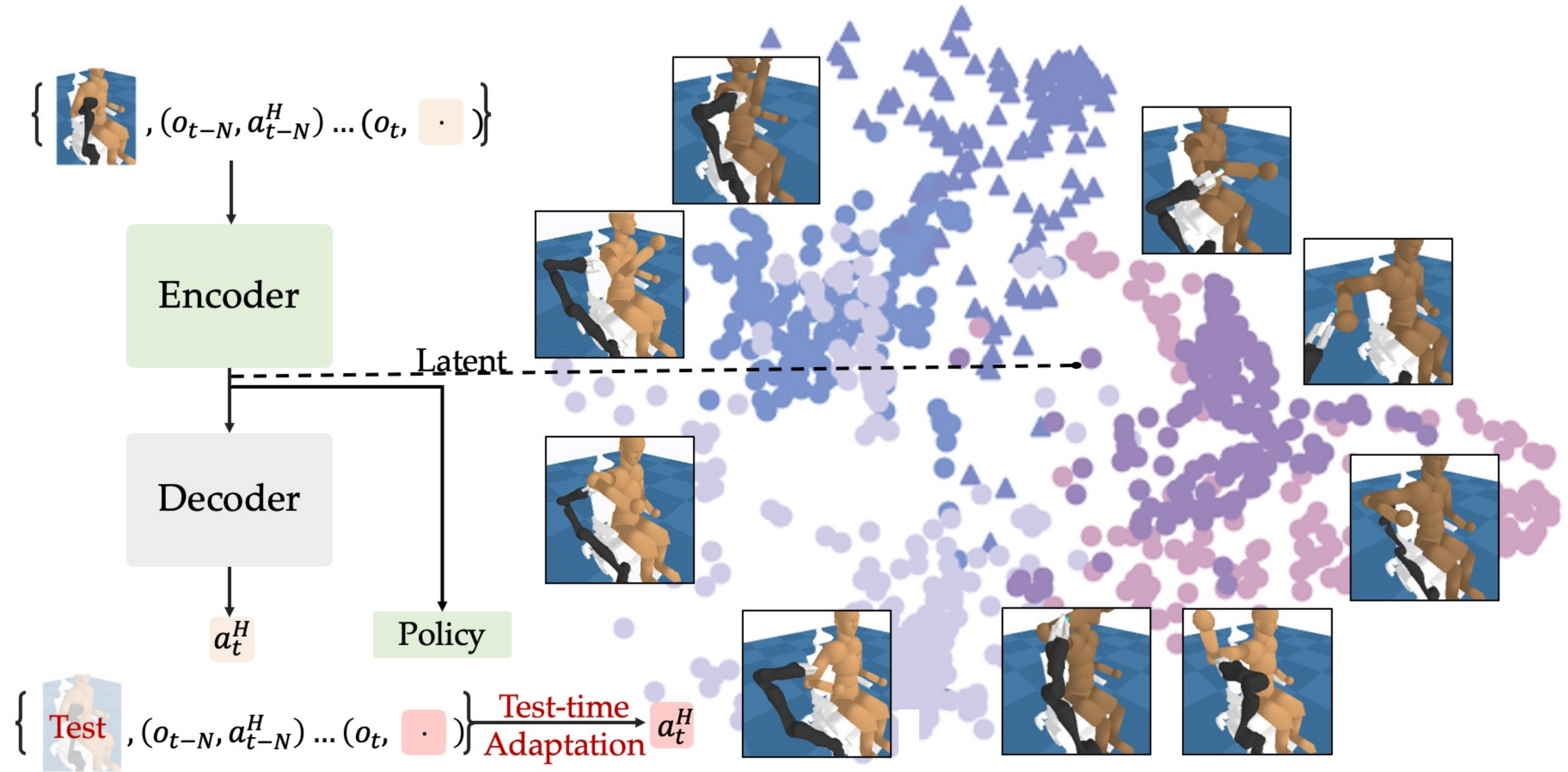
Human-robot joint optimization
Reward manipulation generates a diverse range of motions

$$\max_{\theta_H, \theta_R} E[\textstyle\sum_t r(s_t, a_t^H, a_t^R)], \qquad a_t^H \sim \pi_{\theta_H}, a_t^R \sim \pi_{\theta_R}$$



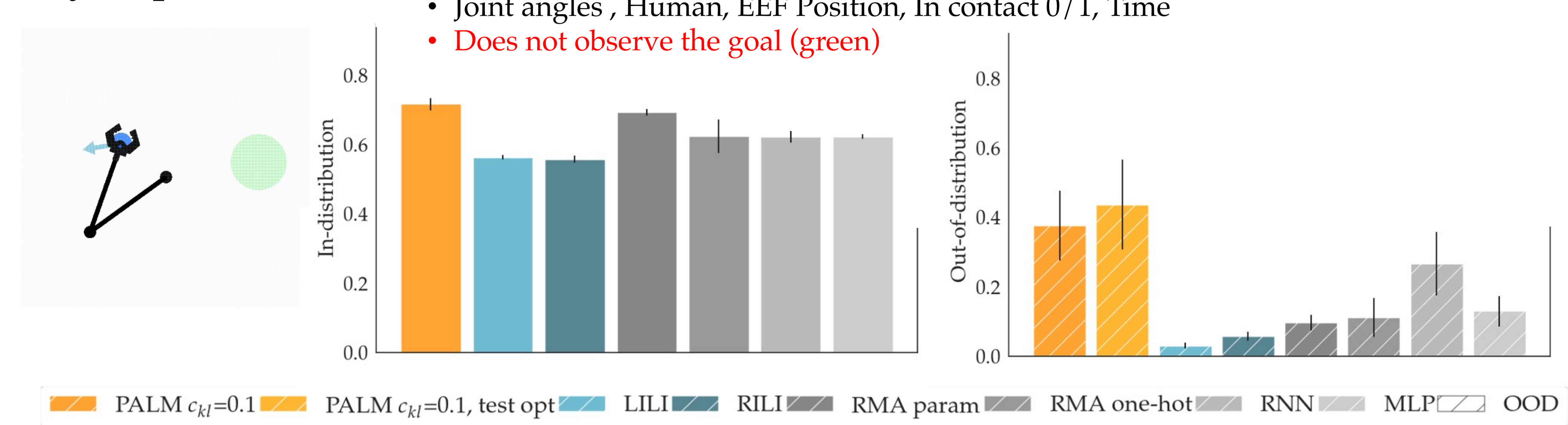
PALM: Prediction-based Assistive Latent eMbedding

- Key Loss: action prediction loss
- We train the latent space jointly with robot policy π_R

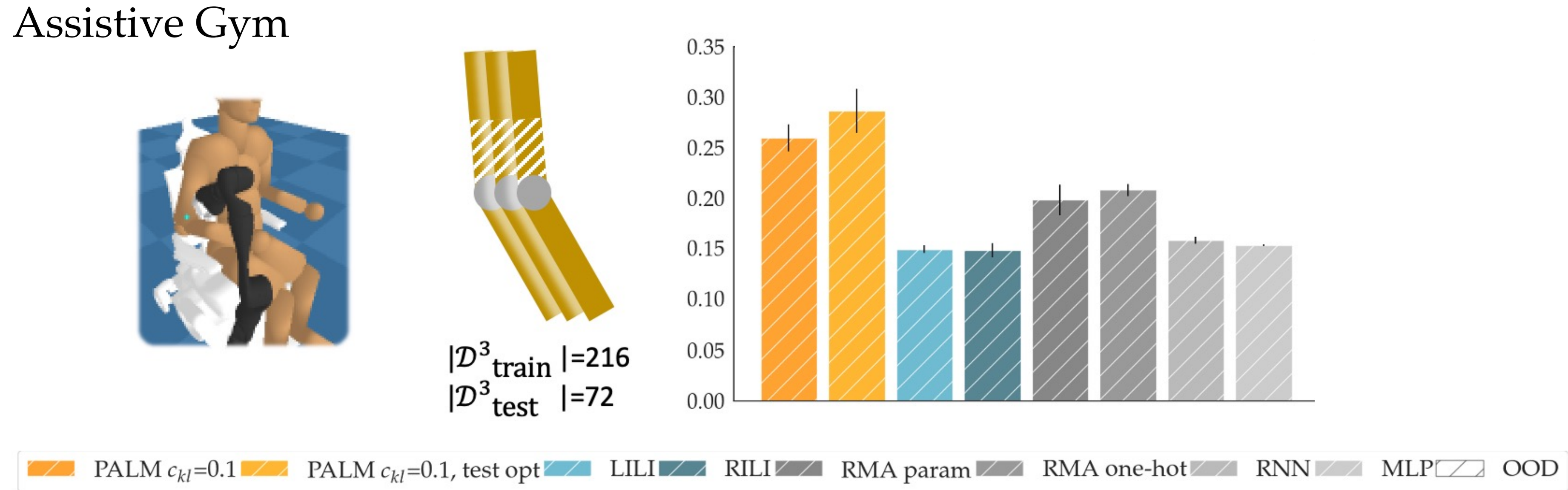


Toy Experiment

Robot observes:
- Joint angles , Human, EEF Position, In contact 0/1, Time
- Does not observe the goal (green)

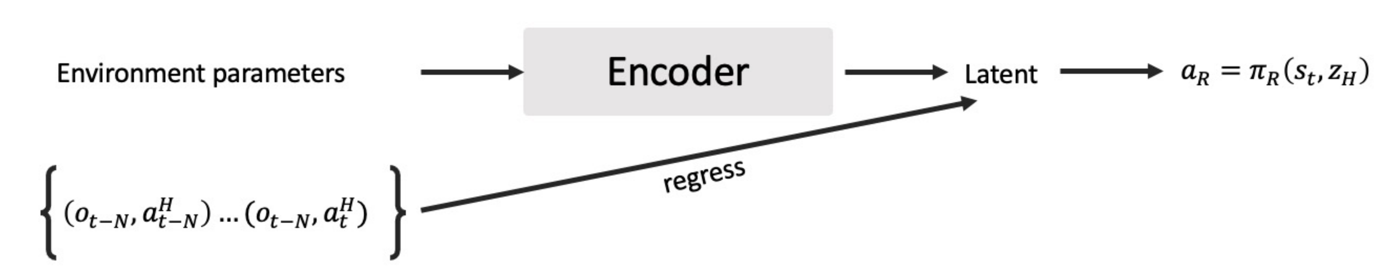


Assistive Gym



Alternatives

- RMA (*Kumar et al*): difficult to hand-code environment parameters to describe human policies



- LILI (*Xie et al*), RILI (*Parekh et al*): decodes reward, not available when assisting unknown patients w/ unknown preferences
- RNN, MLP: no latent space, worse generalization

Key Takeaways

The action-prediction objective offers a latent space that captures the structure in human policies and leads to the best performance on out-of-distribution humans.

Visualizations