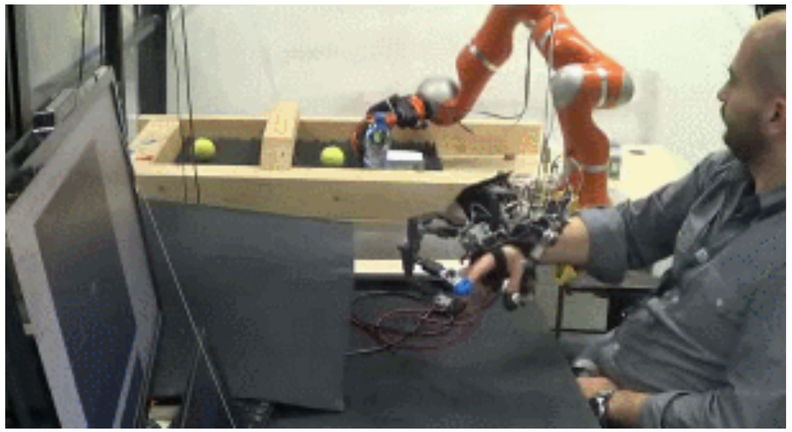
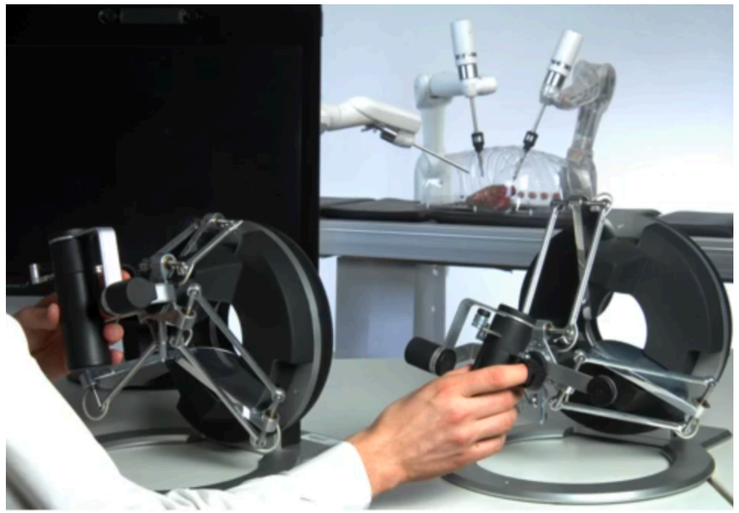
Shared Autonomy via Hindsight Optimization

Presented by Bin Yang

22 Feb. 2019

Teleoperation





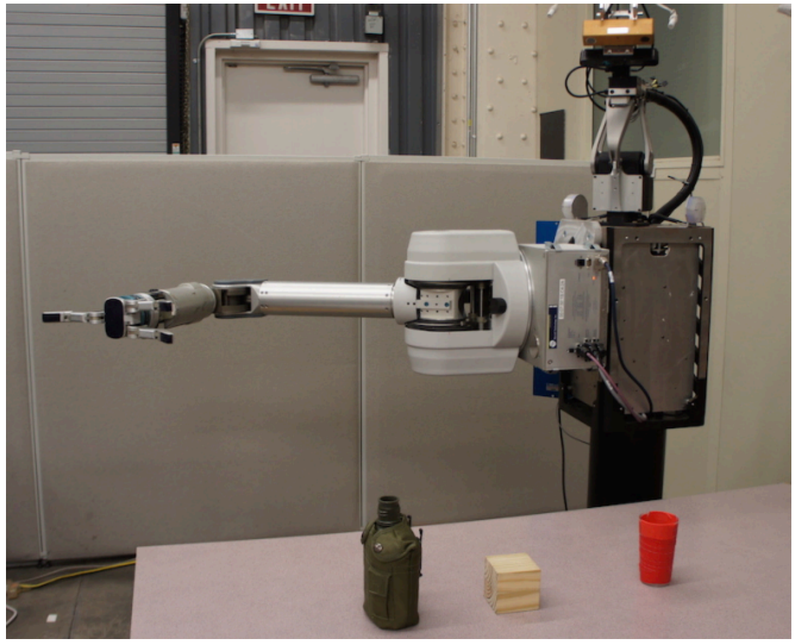
Noisy, insufficient degrees of freedom, tedious

Image credit: Javdani RSS2015 talk

Shared Autonomy

+

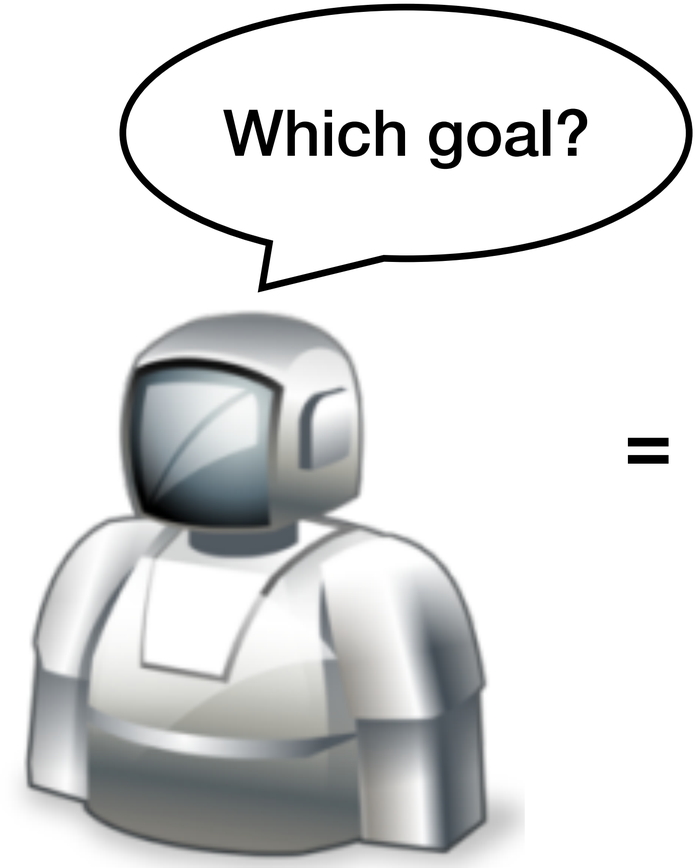
=



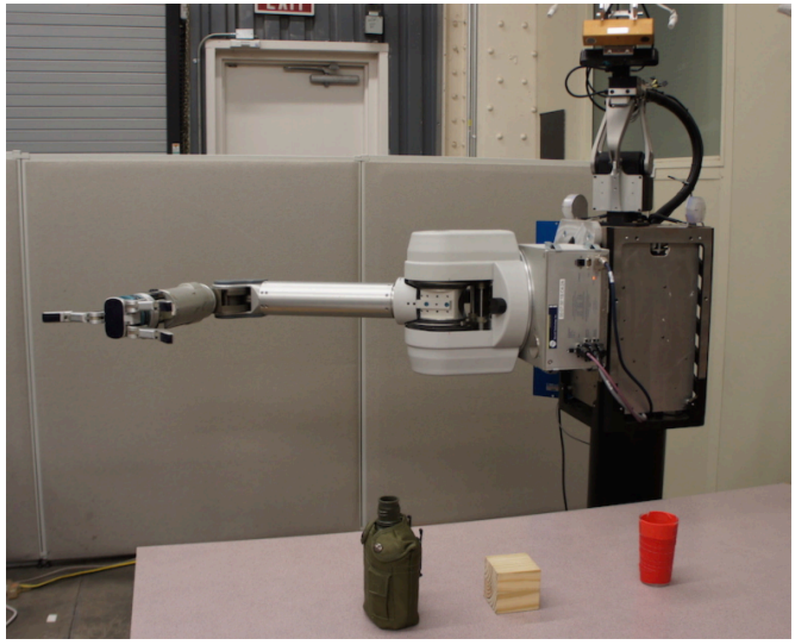
User Input **Autonomous Assistance** **Achieve Goal**

Shared Autonomy

+



=



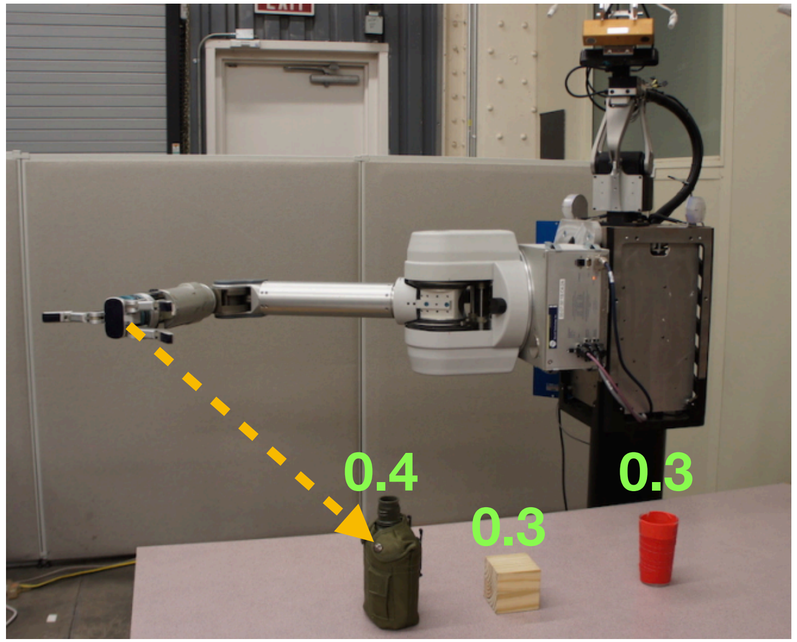
User Input **Autonomous Assistance** **Achieve Goal**

Shared Autonomy

Predict goal
Assist for single goal

[Dragan and Srinivasa 13]
[Kofman et al. 05]
[Kragic et al. 05]
[Yu et al. 05]
[McMullen et al. 14]
…

Autonomous Assistance



Achieve Goal

Shared Autonomy

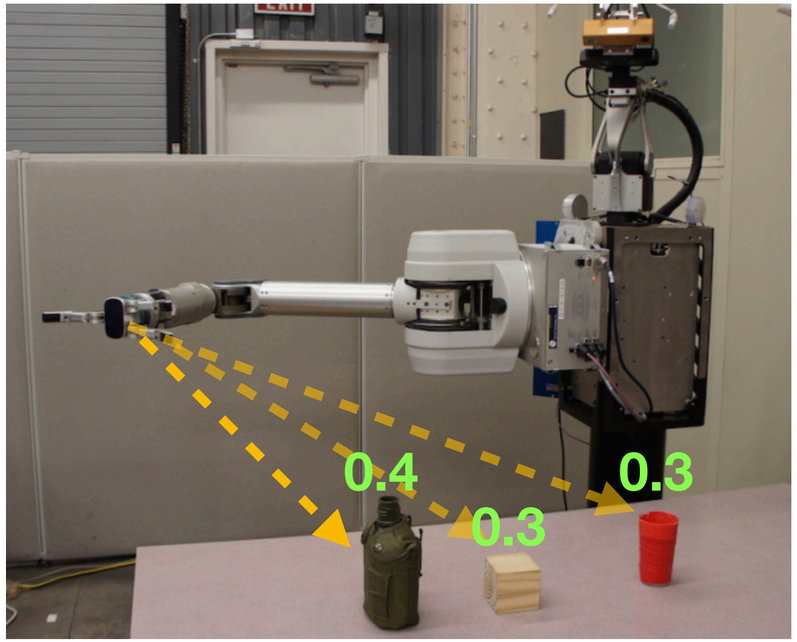
Predict goal
Assist for single goal

[Dragan and Srinivasa 13]
[Kofman et al. 05]
[Kragic et al. 05]
[Yu et al. 05]
[McMullen et al. 14]
…

Predict goal distribution
Assist for distribution

[Hauser 13]
This work!

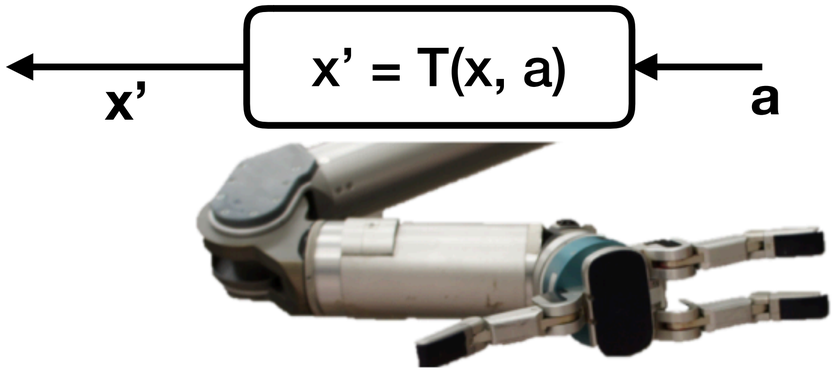
Autonomous Assistance



Achieve Goal

Method

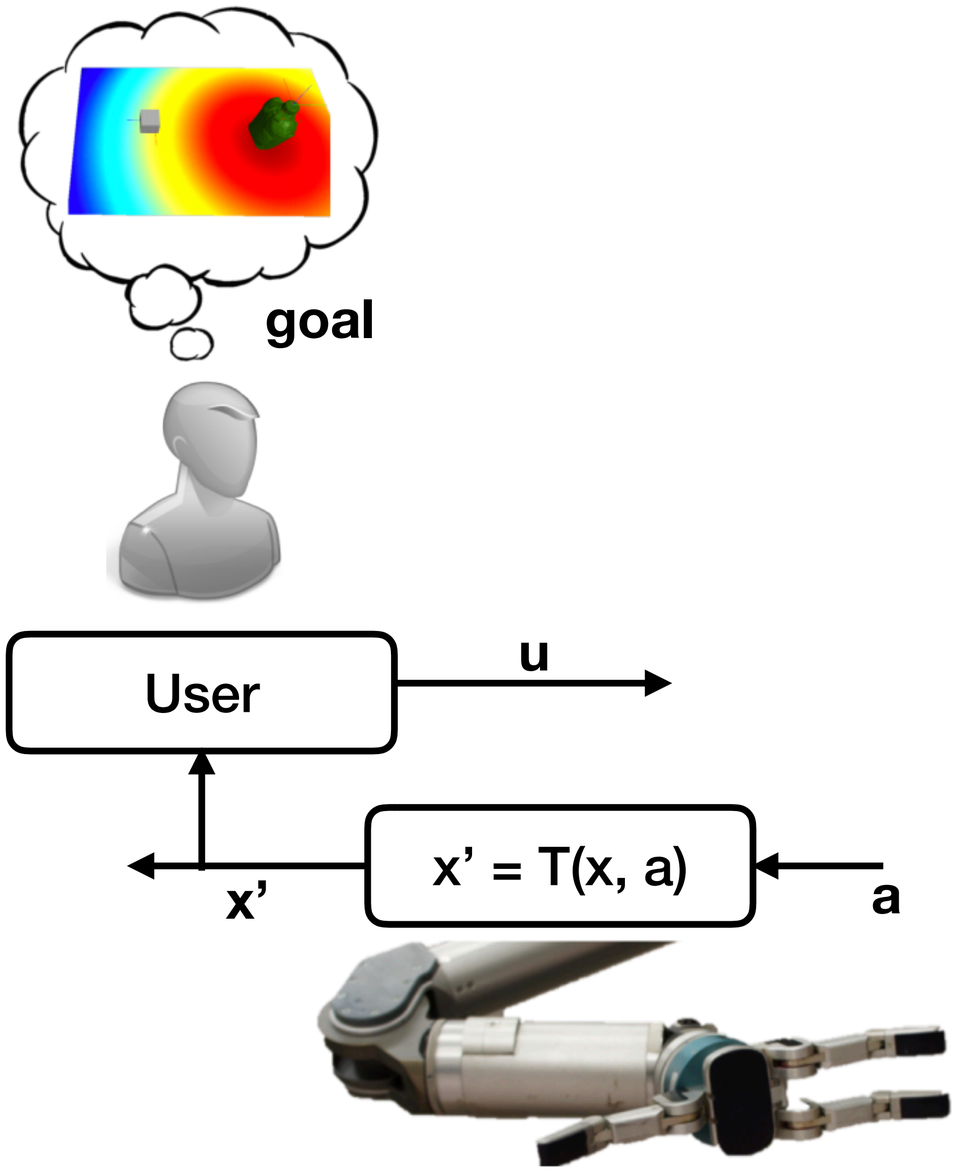
- System dynamics: x’ = T(x, a)



Method

- System dynamics: x’ = T(x, a)

- User (MDP) as $(X, U, T, C_g^{\text{usr}})$
  - User policy: $\pi_g^{\text{usr}}(x) = p(u|x, g)$
  - MaxEnt IOC: $C_g^{\text{usr}} : X \times U \to \mathcal{R}$



Method

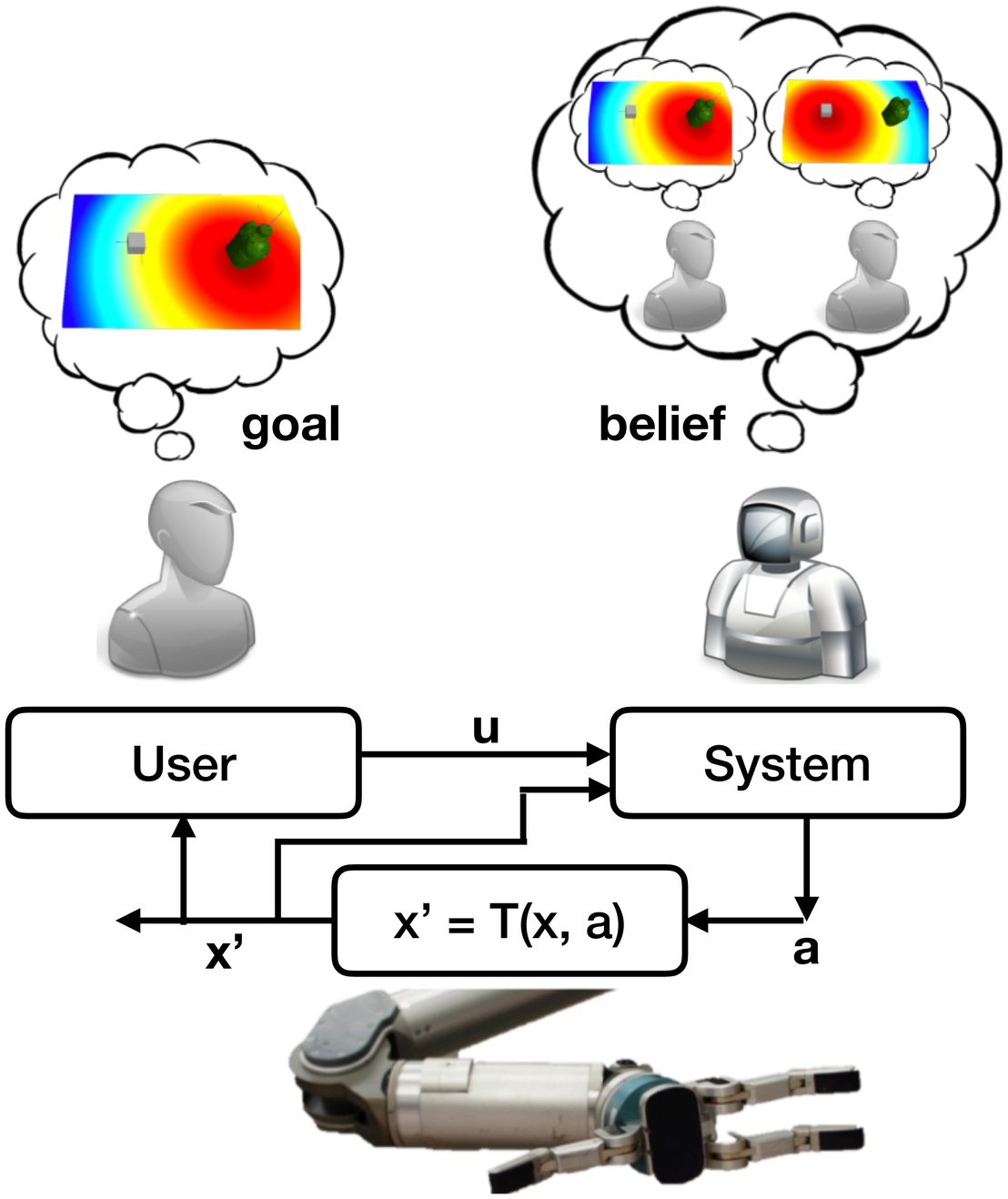
- System dynamics: x’ = T(x, a)

- User (MDP) as $(X, U, T, C_g^{\text{usr}})$
  - User policy: $\pi_g^{\text{usr}}(x) = p(u|x, g)$
  - MaxEnt IOC: $C_g^{\text{usr}} : X \times U \to \mathcal{R}$

- System (POMDP) as $(S, A, T, C^{\text{rob}}, U, \Omega)$
  - Uncertainty over user’s goal
  - System state: $S = X \times G$
  - Observation: user inputs U
  - Observation model Ω

$$p(g|\xi^{0\to t}) = \frac{p(\xi^{0\to t}|g)p(g)}{\sum_{g'} p(\xi^{0\to t}|g')p(g')}$$

  - Cost function $C^{\text{rob}} : S \times A \times U \to \mathcal{R}$



Hindsight Optimization

- MDP solution:

$$V^{\pi^{\mathrm{r}}}(s) = \mathbb{E}\left[\sum_t C^{\mathrm{r}}(s_t, u_t, a_t) \mid s_0 = s\right]$$

$$V^*(s) = \min_{\pi^{\mathrm{r}}} V^{\pi^{\mathrm{r}}}(s)$$

Hindsight Optimization

- MDP solution:

$$V^{\pi^{\mathrm{r}}}(s) = \mathbb{E}\left[\sum_t C^{\mathrm{r}}(s_t, u_t, a_t) \mid s_0 = s\right]$$

$$V^*(s) = \min_{\pi^{\mathrm{r}}} V^{\pi^{\mathrm{r}}}(s)$$

- POMDP solution:

$$V^{\pi^{\mathrm{r}}}(b) = \mathbb{E}\left[\sum_t C^{\mathrm{r}}(s_t, u_t, a_t) \mid b_0 = b\right]$$

$$V^*(b) = \min_{\pi^{\mathrm{r}}} V^{\pi^{\mathrm{r}}}(b)$$

Hindsight Optimization

- MDP solution:

$$V^{\pi^{\mathrm{r}}}(s) = \mathbb{E}\left[\sum_t C^{\mathrm{r}}(s_t, u_t, a_t) \mid s_0 = s\right]$$

$$V^*(s) = \min_{\pi^{\mathrm{r}}} V^{\pi^{\mathrm{r}}}(s)$$

- POMDP solution:

$$V^{\pi^{\mathrm{r}}}(b) = \mathbb{E}\left[\sum_t C^{\mathrm{r}}(s_t, u_t, a_t) \mid b_0 = b\right]$$

$$V^*(b) = \min_{\pi^{\mathrm{r}}} V^{\pi^{\mathrm{r}}}(b)$$

- HOP approximation:

$$\begin{aligned} V^{\mathrm{HS}}(b) &= \mathbb{E}_b\left[\min_{\pi^{\mathrm{r}}} V^{\pi^{\mathrm{r}}}(s)\right] \\ &= \mathbb{E}_g[V_g(x)] \end{aligned}$$

Hindsight Optimization

- MDP solution:

$$V^{\pi^r}(s) = \mathbb{E}\left[\sum_t C^r(s_t, u_t, a_t) \mid s_0 = s\right]$$

$$V^*(s) = \min_{\pi^r} V^{\pi^r}(s)$$

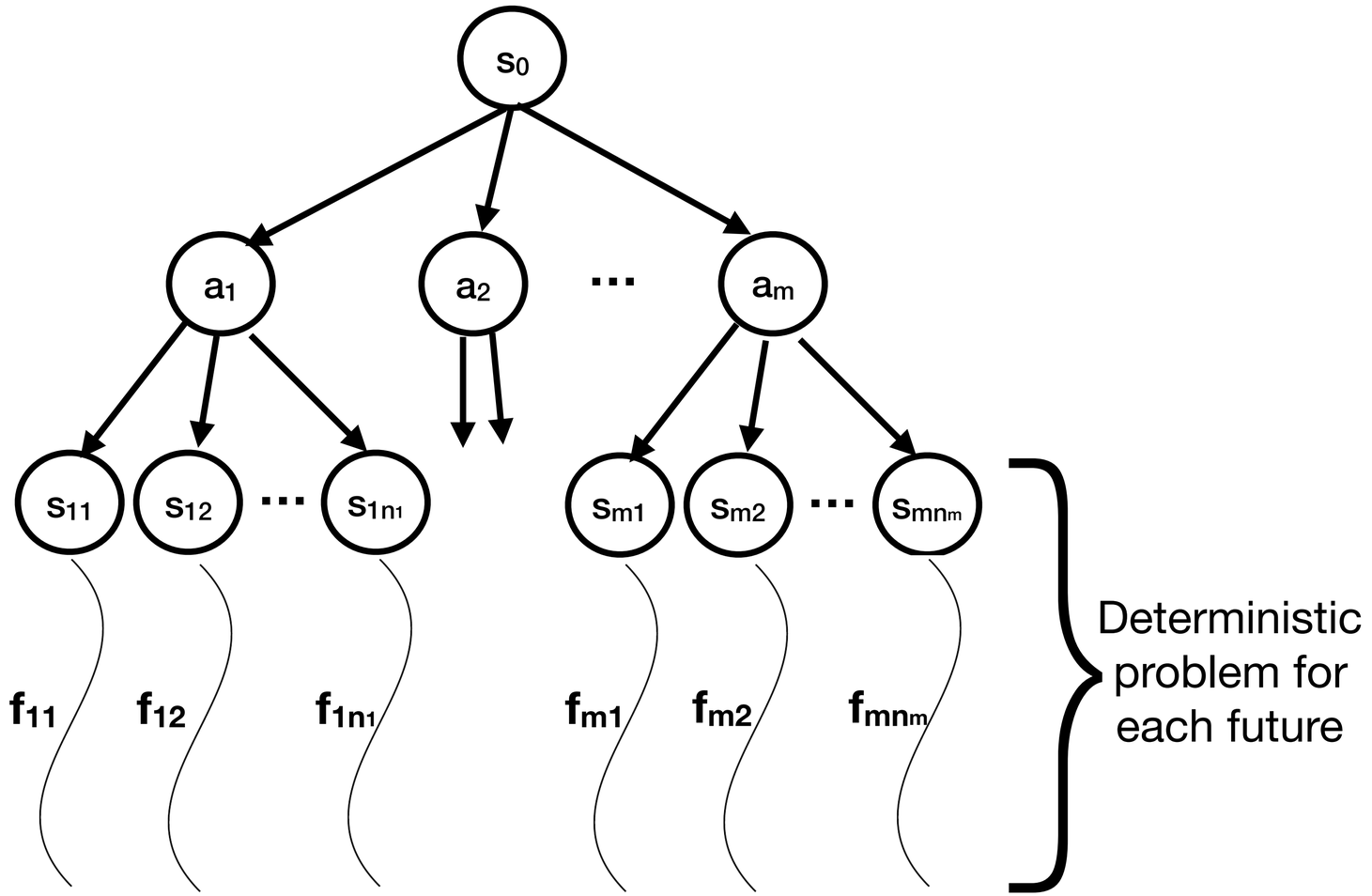
- POMDP solution:

$$V^{\pi^r}(b) = \mathbb{E}\left[\sum_t C^r(s_t, u_t, a_t) \mid b_0 = b\right]$$

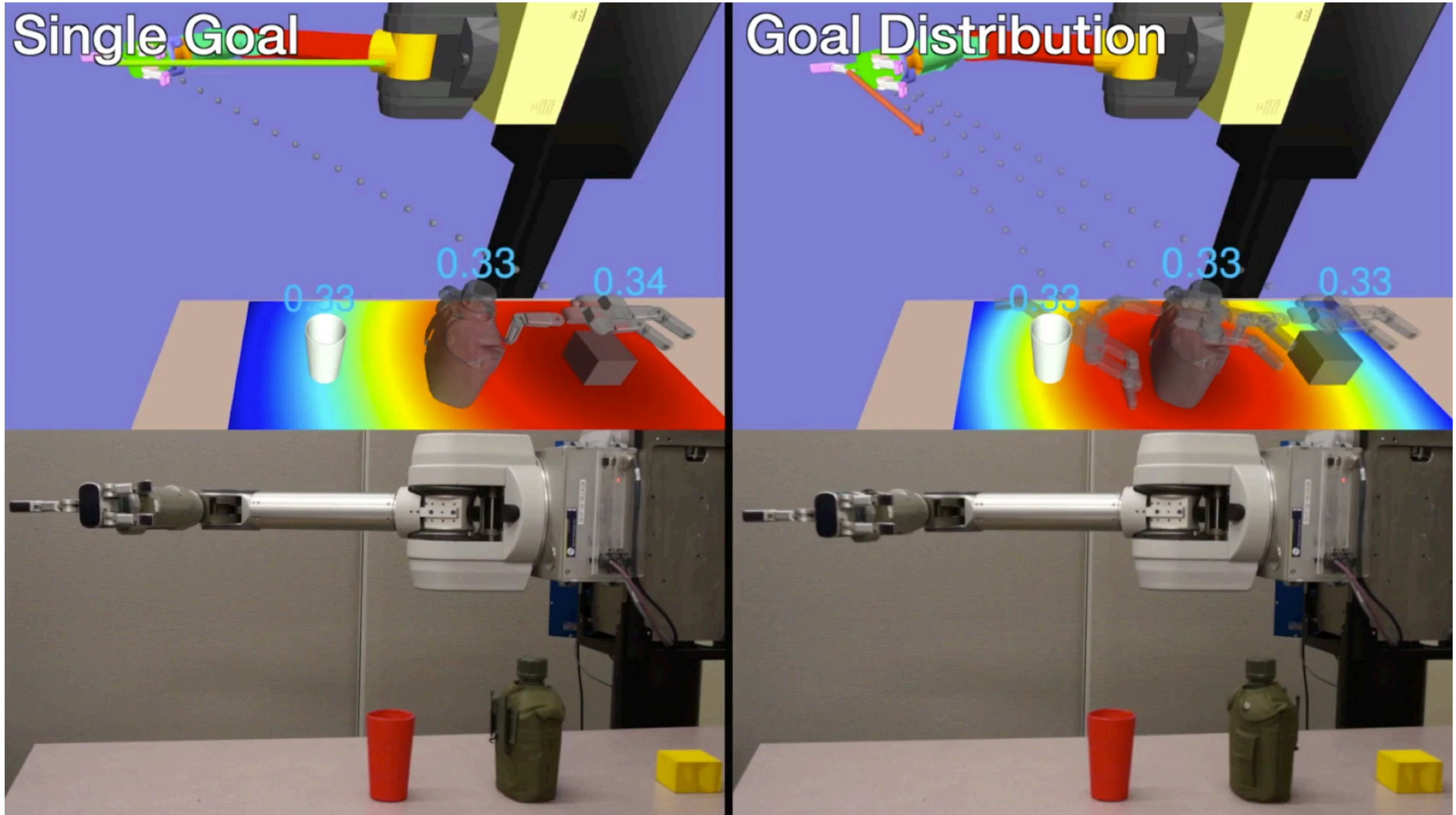
$$V^*(b) = \min_{\pi^r} V^{\pi^r}(b)$$

- HOP approximation:

$$V^{\mathrm{HS}}(b) = \mathbb{E}_b\left[\min_{\pi^r} V^{\pi^r}(s)\right]$$
$$= \mathbb{E}_g[V_g(x)]$$



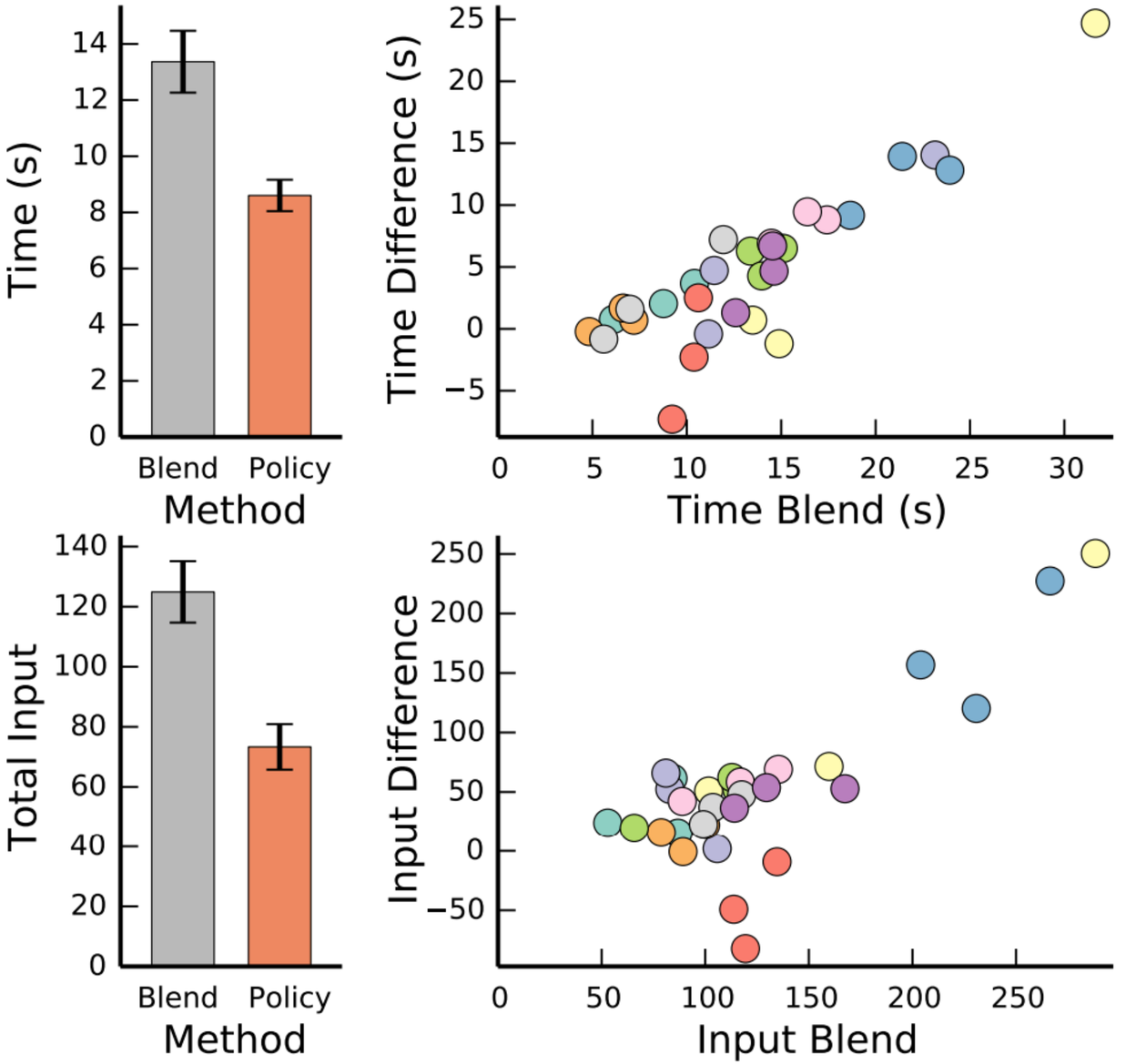
Results (video)



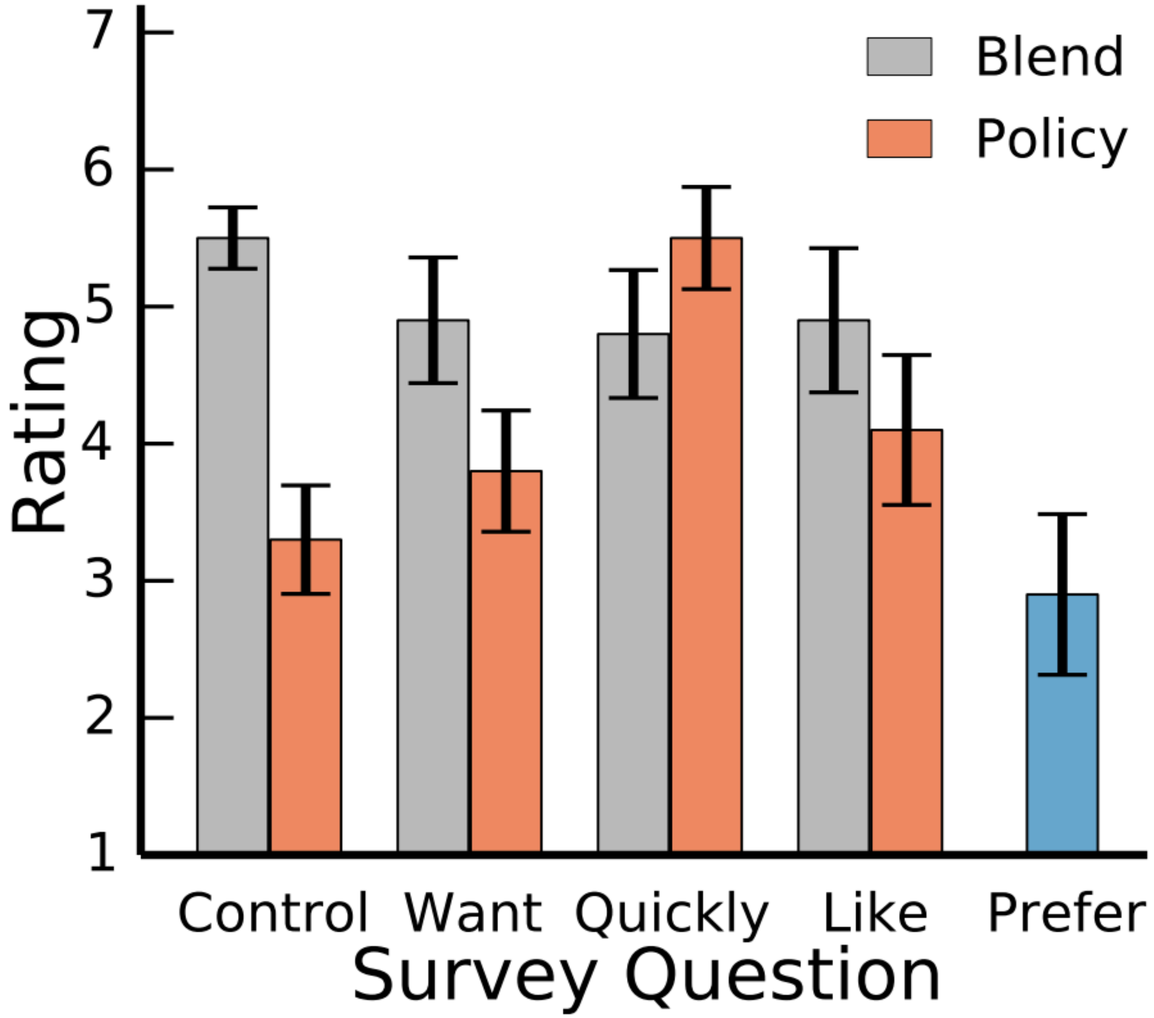
Results

Compare with method that predicts one goal, the proposed method has:

- Faster execution time
- Fewer user inputs



User Study


Blend
Policy
Rating
7
6
5
4
3
2
1
Control
Want
Quickly
Like
Prefer
Survey Question

Limitations

- Requires prior knowledge about the world:
  - a dynamics model that predicts the consequences of taking a given action in a given state of the environment;
  - the set of possible goals for the user;
  - the user's control policy given their goal.

- Suitable in constrained domains where where this knowledge can be directly hard-coded or learned.

- Unsuitable for unstructured environments with ill-defined goals and unpredictable user behavior.

References

- Javdani, S., Srinivasa, S. S., & Bagnell, J. A. (2015). Shared autonomy via hindsight optimization. *Robotics science and systems: online proceedings*, *2015*.
- RSS2015 talk: “Shared autonomy via hindsight optimization”
- Javdani, S., Admoni, H., Pellegrinelli, S., Srinivasa, S. S., & Bagnell, J. A. (2018). Shared autonomy via hindsight optimization for teleoperation and teaming. *The International Journal of Robotics Research*, *37*(7), 717-742.
- ICAPS 2015 talk: "Hindsight Optimization for Probabilistic Planning with Factored Actions”