EAIS Guest Lecture

HJI - VI

$$0 = \min \left\{ \ell(x) - V(x,t),\ D_t V + \max_{u \in \mathcal{U}} \min_{d \in D} D_x V \cdot \overbrace{f(x,u,d)}^{\dot{x} = f(x,u,d)} \right\}$$

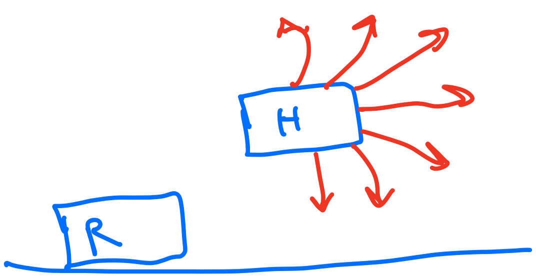
$$V(x) = \min \left\{ \ell(x),\ \max_{u \in \mathcal{U}} \min_{d \in D} V\big(\overbrace{f(x,u,d)}^{x^+ = f(x,u,d)}\big) \right\}$$

→ Q: What assumptions does this make?
What makes this hard IRL?

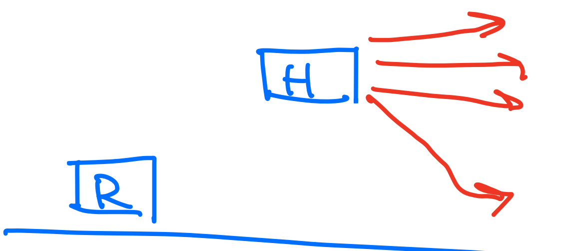
- → assume we have $f(x,u,d)$ - We know how to model disturbance set D ← This first!
- → x is observable
- → We know how to design $\ell(x)$

Interactions w/ people are hard to model!

Option 1. Be robust to anything the human can do.



Option 2: Be robust to sufficiently likely human behavior



Very Conservative!

Issue: What is "likely"?
need to know human intent!
Not observable

Partial Observability

Prev: x is fully observable

Now: get $o_t \sim P(o_t \mid x_t)$ ← observations of x

$b_t(x_t)$ → distribution over unobservable x

Prior ↑

$$b_{t+1}(x_t) = \frac{P(o_t \mid x_t)\, b(x_t)}{P(o_t)}$$

Posterior after seeing evidence o_t

This update is just Bayes' Rule

$$P(A \mid B) = \frac{P(B \mid A) P(A)}{P(B)}$$

Idea: Using stream of observations, you can reduce uncertainty about an unobservable quantity

Q: How can we design safe ctrl policies that account for robot's evolving uncertainty??

Very nuanced

Deception Game CoRL 2023

$x_{t+1} = f(x_t, u_t, d_t)$ ⇒ assume x is observable

$\theta \in \Theta$ human "type". Θ is a discrete set

θ is unobservable. θ could represent human intent, semantic class etc.

$b(\theta)$ belief over human type θ

$o_t = h(x_t, d_t)$ Observation depend on physical state x and human action d

$b_{t+1} = f_L(b_t, o_t)$ "Learning dynamics", e.g. Bayesian update rule

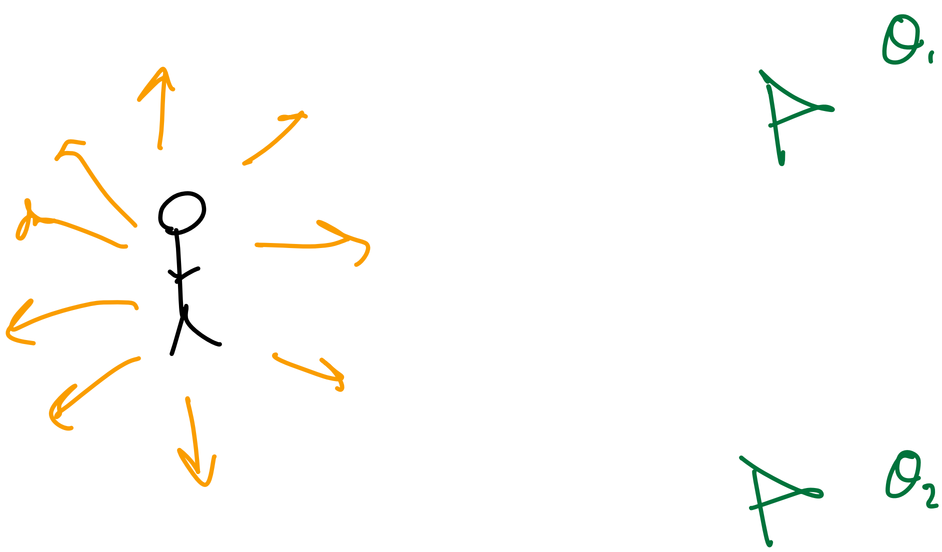
Define $z := (x_t, b_t)$ joint phys-belief state

$$F(z,u,d) = \begin{bmatrix} f(x_t, u_t, d_t) \\ f_L(b_t, o_t) \end{bmatrix}$$

$$V(x) = \min \left\{ \ell(x), \max_{u \in \mathcal{U}} \min_{d \in D} V(f(x,u,d)) \right\}$$??

Now, let's make a modeling assumption about the human

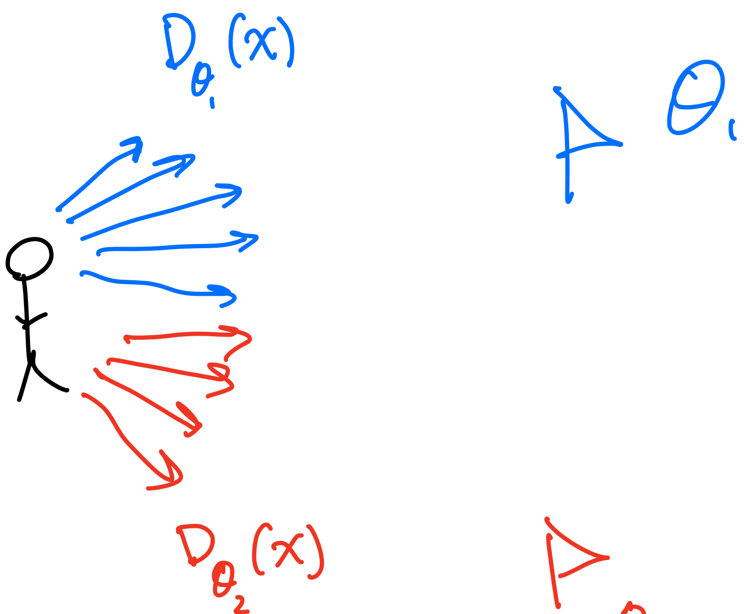
Set of all human actions $\mathcal{D}$



Type-dependent Control Set $\mathcal{D}_\theta(x)$

↓

Set of controls that we deem likely <u>if</u> the human's type is θ



Left, θ represents goal locations.

But we don't know θ, only have belief $b(\theta)$

Inference Hypothesis

One way to use $b(\theta)$ to modulate allowable human actions

$$\hat{D}(z) = \bigcup_{\theta \in \Theta} \hat{D}_\theta(z)$$

↑ Union over all types θ

$$\hat{D}_\theta(z) = \begin{cases} D_\theta(x) & \text{if } b(\theta) \geq \varepsilon \\ \emptyset & \text{otherwise} \end{cases}$$

ε → tuneable parameter

↑ only consider $D_\theta(x)$ if $b(\theta)$ is sufficiently high

Note: $b(\theta)$ evolves with time, so $\hat{D}(z)$ will also evolve w/ time, subject to learning dynamics $f_L(x, \theta)$

Belief-Space HJ

$$V(z) = \min \left\{ \ell(z), \max_{u \in \mathcal{U}} \min_{d \in \hat{D}(z)} V(F(z, u, d)) \right\}$$

↑ belief influences dyn

$$F := \{ z \mid \ell(x) < 0 \}$$

↑ This paper: only depends on physical state

Note:

- Solved via adversarial RL
- Humans can act /deceptively/