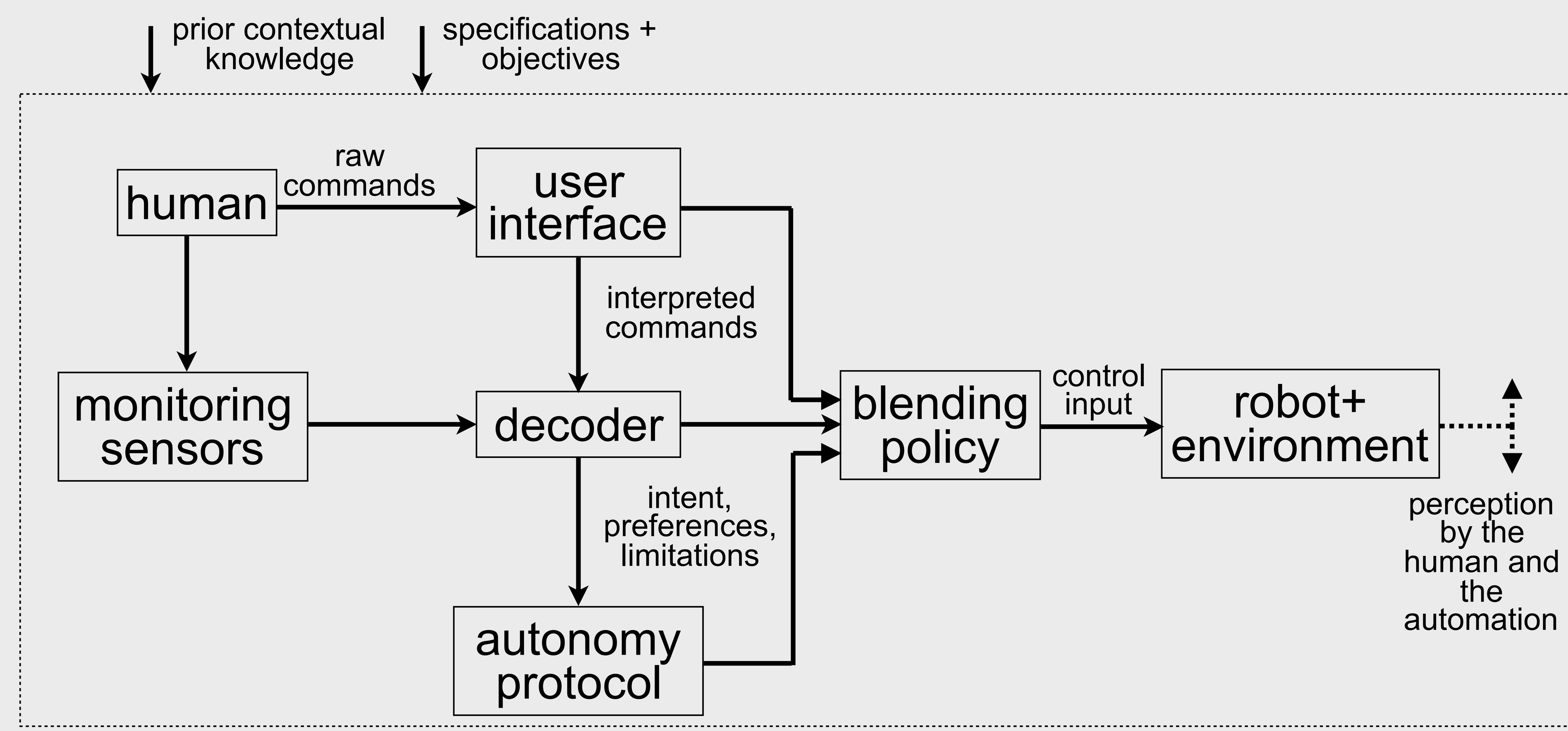
Provably Correct Shared Control for Human-Embedded Autonomous Systems

Ufuk Topcu — The University of Texas at Austin

Human-embedded systems: Humans and autonomy are responsible for collective information acquisition, perception, cognition and decision-making at multiple and varying levels of abstraction.

Objective of the project: Develop languages, algorithms and demonstrations for the formal specification and automated synthesis of *shared control* protocols.



A shared control architecture for human-embedded autonomous systems

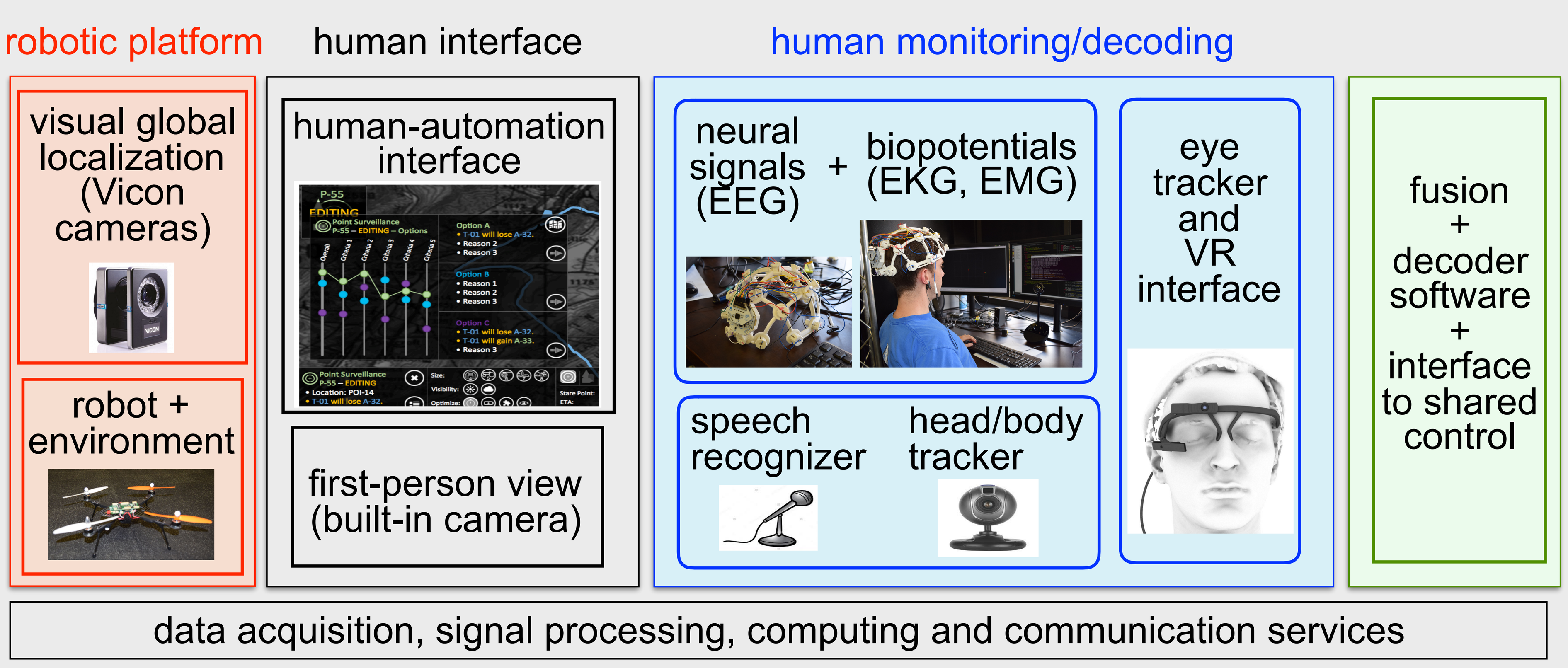
Research thrusts

Specifications and modeling for shared control: What does it mean to be provably correct in human-embedded autonomous systems, and how can we represent correctness in formal specifications?

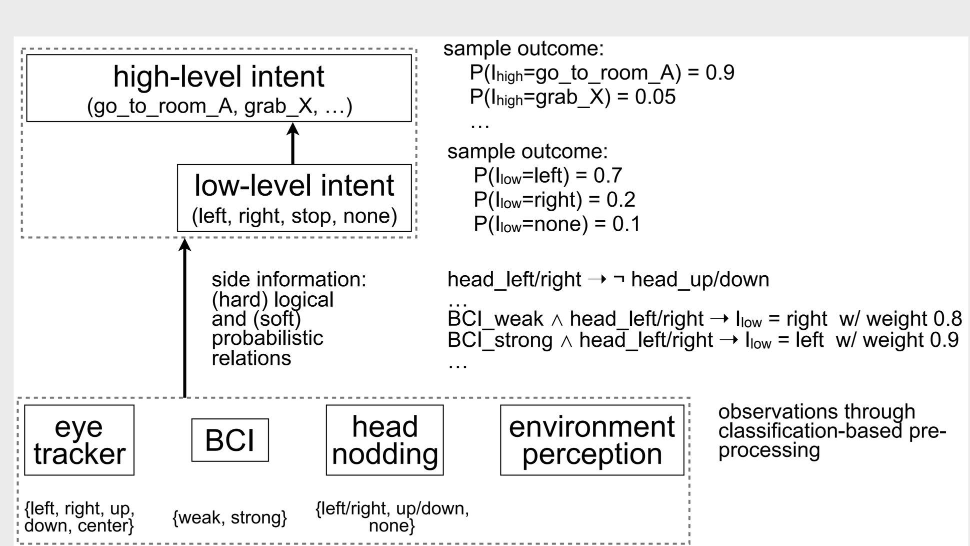
Automated synthesis of shared control protocols: How can we mathematically abstract shared control, and automatically synthesize shared control protocols from formal specifications?

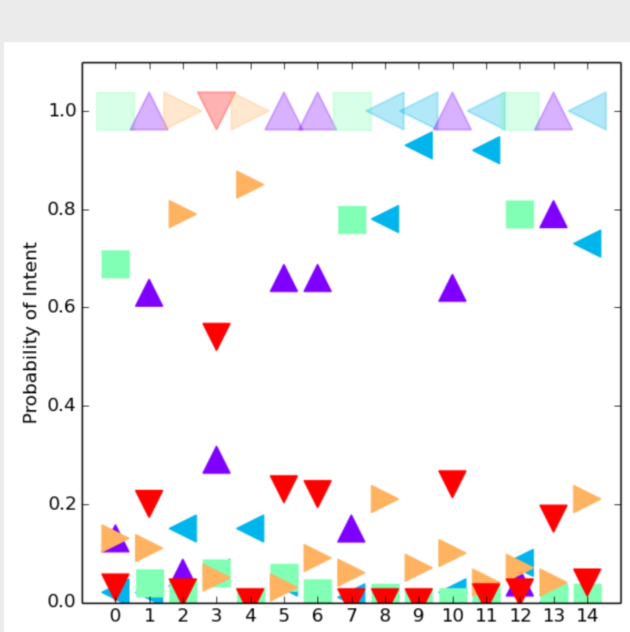
Shared control through human-autonomy interfaces: How can we account for the limitations in expressivity, precision and bandwidth of human-autonomy interfaces, and co-design controllers and interfaces?

An implementation and preliminary results



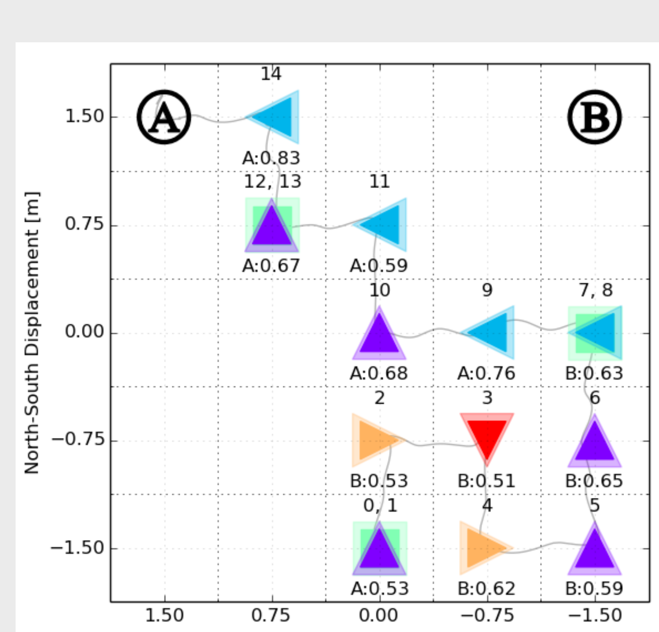
Human intent decoding—using Markov logic networks—through imperfect interfaces





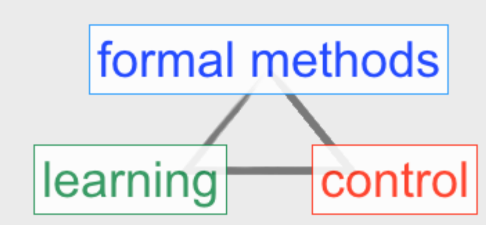
Left: motion-level intent prediction

Right: Sample trajectory, level intents and their probabilities



An enabling factor

Convergence between learning, formal methods, and control



Inverse reinforcement learning with high-level task as side information

$$\min_{\theta} \quad J^{\mathrm{mle}}(\pi_\theta | M \times \mathcal{A}_\varphi, D) - \mu g(\bar{y})$$

$$\text{s.t.} \quad \text{Bellman equation}(\theta, M \times \mathcal{A}_\varphi, D)$$

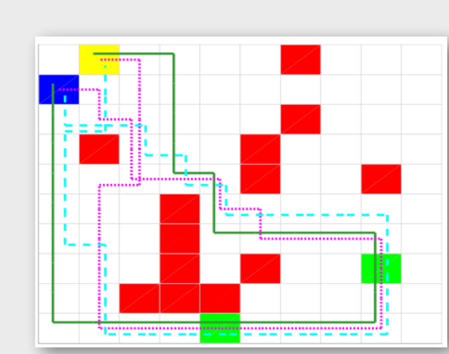
$\bar{y}(s_0, \pi_\theta)$ probability of satisfying the specification by taking policy π_θ from initial state s_0

Reward-free MDP $M = (S, A, T, \gamma)$

modified inverse reinforcement learning

Reward function R

Policy that balances between mimicking the demonstrations and probability of satisfying φ



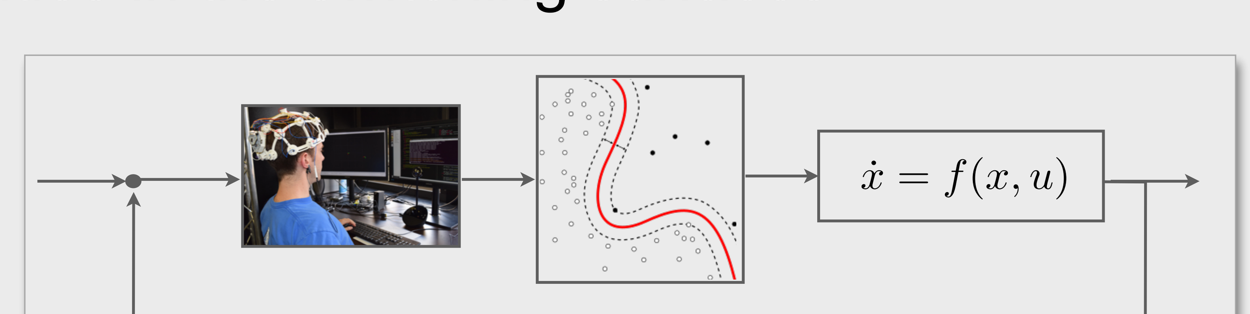
Expert demonstrations

$$\varphi_{\mathrm{cs}} = \varphi_{\mathrm{init}} \rightarrow (\varphi_{\mathrm{safe}} \wedge \varphi_{\mathrm{goal}})$$
$$\varphi_{\mathrm{init}} = \neg r \ \wedge \neg y$$
$$\varphi_{\mathrm{safe}} = \neg r \ \mathcal{U} \ y$$
$$\varphi_{\mathrm{goal}} = ((\neg y) \ \mathcal{U} \ (\Diamond g_1 \wedge \Diamond g_2)) \wedge (\Diamond y)$$

Task knowledge as side information (encoded as a temporal logic formula φ)

Stability of “classifier-in-the-loop” control systems

Translated into and studied as hybrid systems potentially with uncertainties in the switching surfaces



http://u-t-autonomous.info — *utopcu@utexas.edu*