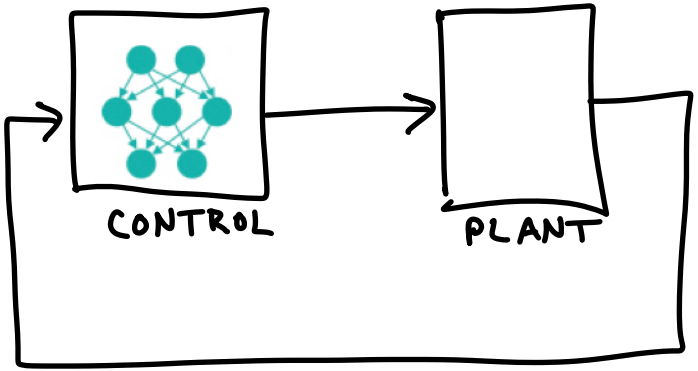
SOCIALLY RESPONSIBLE H-CPS: UNCERTAINTY REPRESENTATION

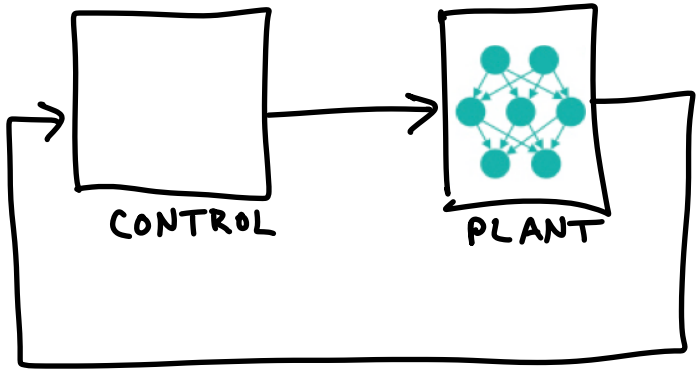
Todd Murphey
Northwestern University

Northwestern | CENTER FOR ROBOTICS AND BIOSYSTEMS

Control for Learning

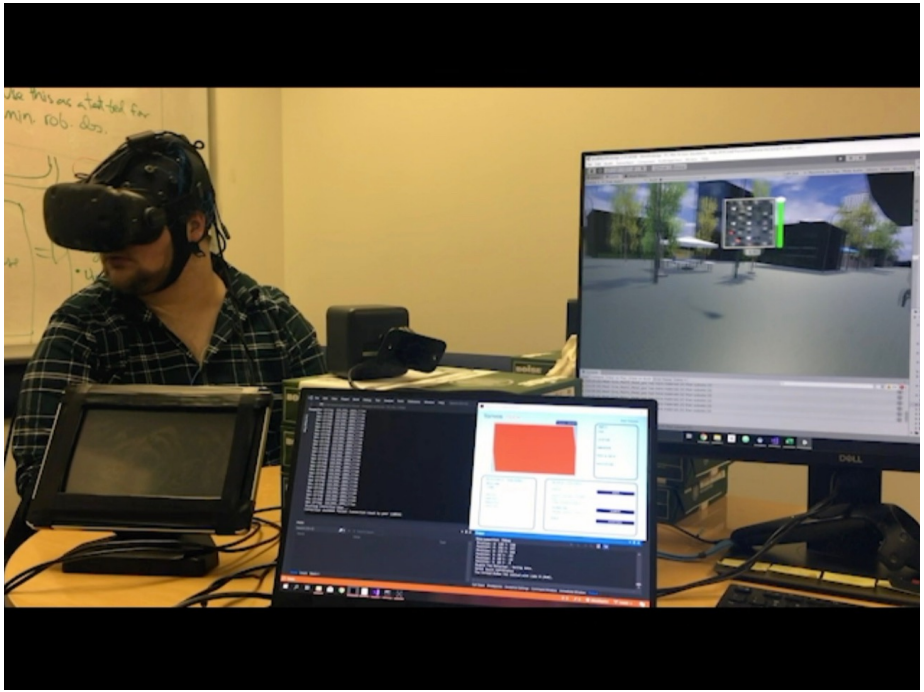


Learning for Control



Control for Learning

Research Overview



How should Control-for-Learning impact an operator adapting to a complex, high pressure environment?

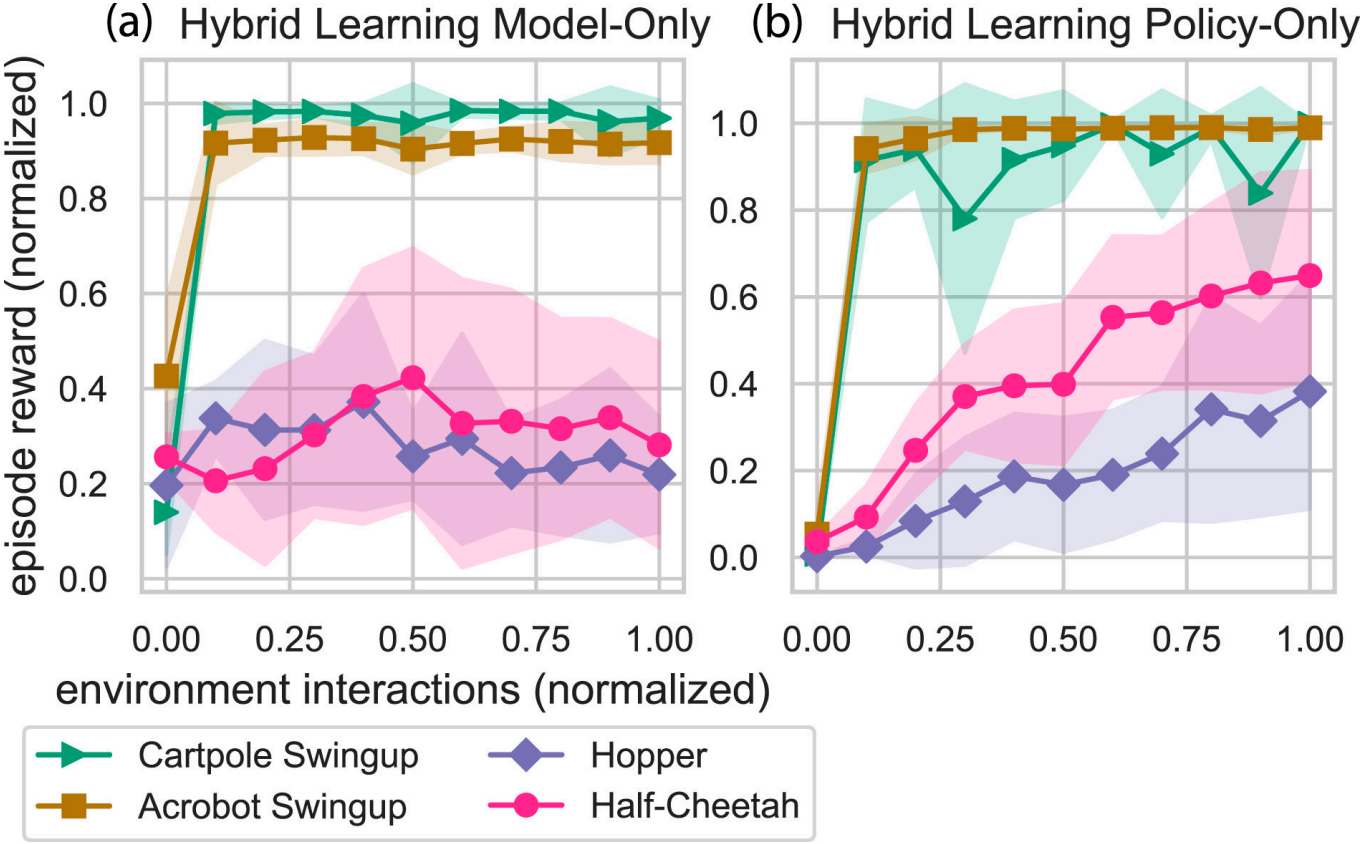
K. Fitzsimons, A. M. Acosta, J. Dewald, and T. D. Murphey, "Ergodicity reveals assistance and learning in physical human robot interaction," *Science Robotics*, 2019.

M. Schlafly, A. Prabhakar, K. Popovic, G. Schlafly, C. Kim, and T. D. Murphey, "Collaborative AI augments human cognition," In Preparation.

Responsible Representations of Algorithms

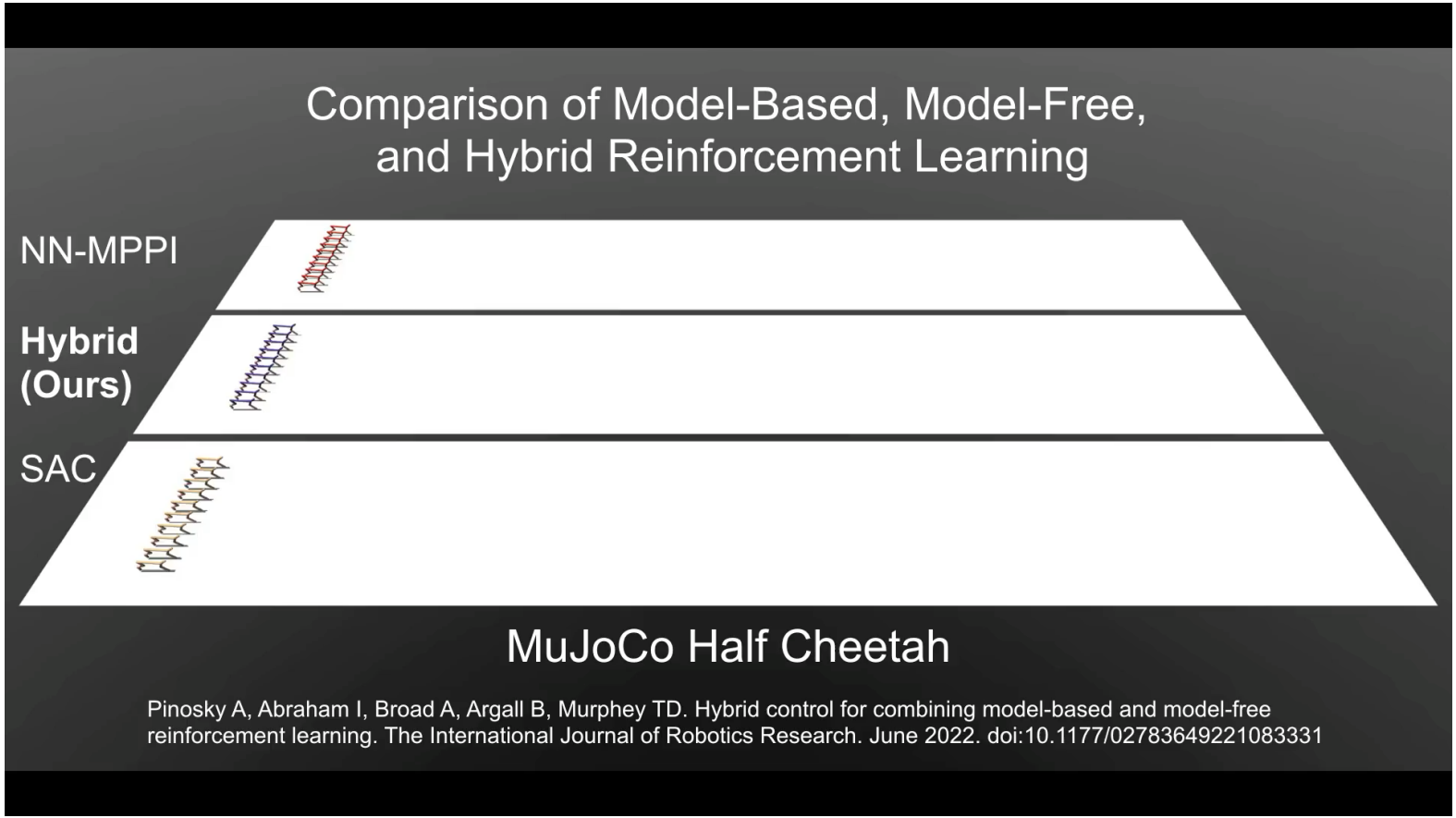
- Uncertainty quantification matters, even (or particularly) to untrained operators
- Words like ‘intelligence’ are disingenuous when taken out of context
    - The word ‘intelligence’ in computer science is only vaguely related to daily experience
- We can, and should, continue making systems more reliable/safer/socially-aware but we can also calibrate user expectations

Representations of Uncertain Performance



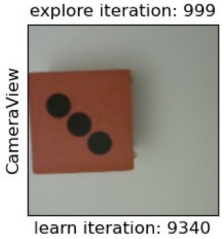
Pinosky, I. Abraham, A. Broad, B. Argall, T. Murphey, "Hybrid control for combining model-based and model-free reinforcement learning," *Int. J. of Robotics Research*, 2022.

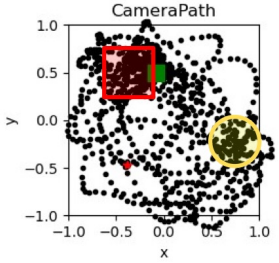
Representations of Uncertain Performance

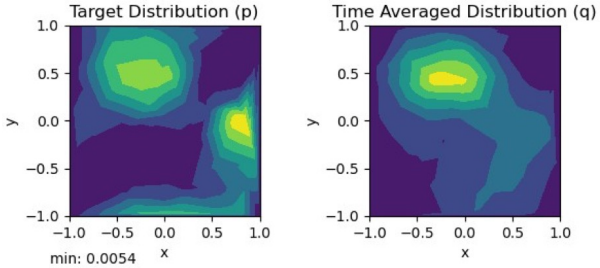


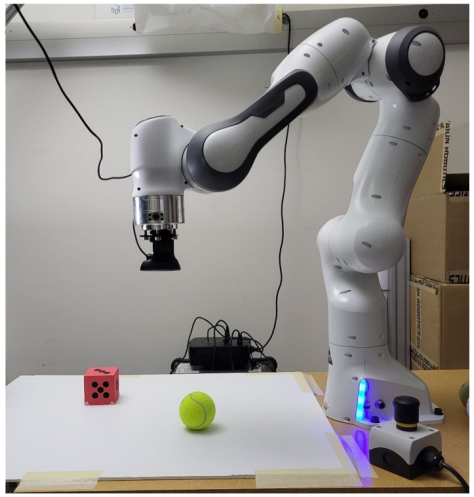
Pinosky, I. Abraham, A. Broad, B. Argall, T. Murphey, "Hybrid control for combining model-based and model-free reinforcement learning," *Int. J. of Robotics Research*, 2022.

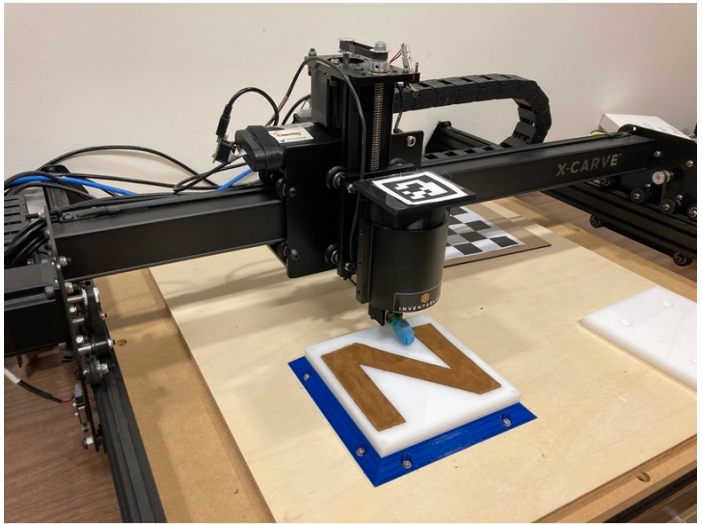
Physical Uncertainty and Prediction

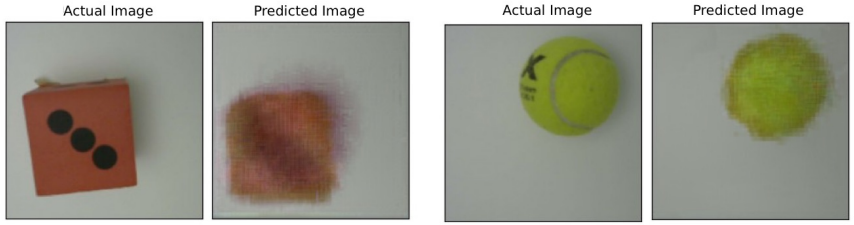












Real-time / Online Deep Learning involves compute and thoughtful, automated data collection

People might be able to understand the visual system, but they will not be able to understand the tactile system.

Key Point:

- We want H-CPS to be responsive, socially aware, unbiased, ethical, and more. **We will achieve some of those things some of the time to some degree.**
- The rest of the time, we need H-CPS to accurately represent their performance characteristics to their users, particularly H-CPS limitations.
- Technical challenges—how should a system reason about its limitations and automatically communicate them to operators on a spectrum of training backgrounds?

Needed Advances/Opportunities:

- Measures, datasets, and benchmarks for helping people create accurate mental models of robots
- Open source protocols with ‘real world’ pressure
    - virtual reality can be an enabler
- How do we make physical experimentation cheap and scalable? Pop-up facilities, field experiments
- Data-driven techniques to create interfaces that present representations of outcomes
- Models of operators as learning agents

Thanks!