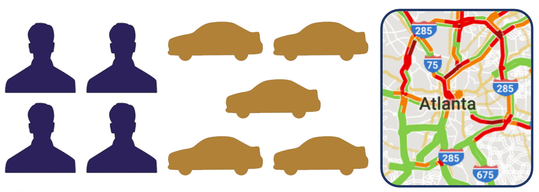
CPS:Medium:Collaborative Research:

Certifiable reinforcement learning for cyber-physical systems

Sam Coogan (Georgia Tech), **Sam Burden** (University of Washington), **Lillian Ratliff** (University of Washington)

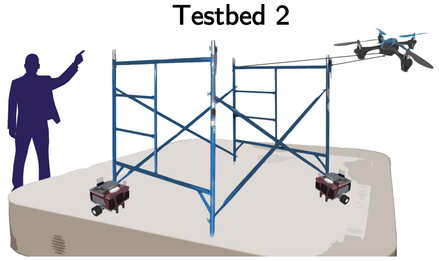
Classical optimal control problem: $\min_u c(x,u)$ such that $\dot{x} = f(x,u)$

Testbed 1



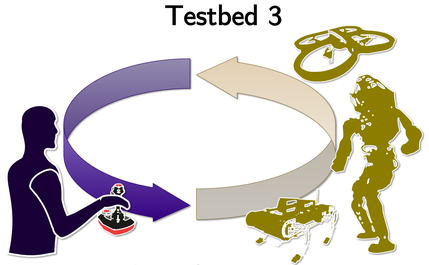
Societal-scale systems as multi-agent games

Testbed 2



Construction with collaborative robotics

Testbed 3



Teleoperation of dynamic robots

CPS have non-classical costs/dynamics:

- Hybrid or nonsmooth dynamics F
- Multiple decision-making agents with incomplete information (humans and autonomy)

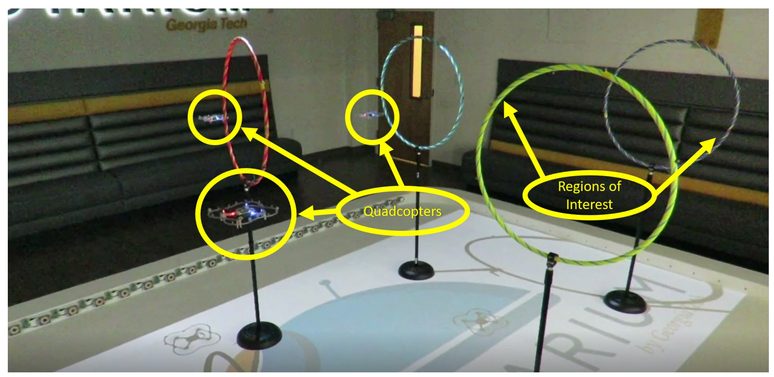


Figure. A multiagent team of quadrotors learns how to maneuver through suspected hoops to achieve complex objectives (e.g., green hoop always before blue hoop)

Thread A: Consistent credit assignment.

Challenge: RL requires a solution to the credit assignment problem in which the value of states and actions must be determined based on future rewards. In CPS, nearby trajectories may behave differently.

Recent work: Extending contraction theory to hybrid systems enables comparison of nearby states [1]. Multiagents systems can learn complex tasks with the right cost functions [2,3].

Broader Impact: Robotarium quadrotors

We are developing a quadrotor component of the open-access Robotarium at Georgia Tech. This component will include a simulator and the ability to run experiments remotely on the Robotarium.

Thread B: Metrics and measures for assessing convergence.

Challenge: RL algorithms must satisfy: 1) for fixed approximation scheme, convergence to some limit; 2) as parameters increase, this limit converges to optimum. Convergence is currently guaranteed for a very limited class of dynamics.

Recent work: Convergence guarantees for multi-agent systems including human/machine interaction [4].

Broader Impact: Learning modules for undergrad courses

Guided by the outcome of the proposed workshop, we will develop and deploy learning modules for undergraduate capstone design courses

Thread C: Scaling to high dimensions with value and policy gradients.

Challenge: Differentiability is key to scale RL algorithms to high dimensions by computing value or policy gradients. However, trajectory outcomes generally vary non-smoothly with respect to states, parameters, and control inputs in CPS.

Recent work: Offline state estimation for hybrid systems [5].

References

[1] Burden, Coogan "On infinitesimal contraction analysis for hybrid systems" IEEE CDC, 2022

[2] S. Jafarpour, M. Abate, A. Davydov, F. Bullo, S. Coogan, "Robustness certificates for implicit neural networks: a mixed monotone contractive approach," Learning For Decision and Control (L4DC), 2022.

[3] Cao, Michael E. and Bloch, Matthieu and Coogan, Samuel "Efficient Learning of Hyperrectangular Invariant Sets using Gaussian Processes " IEEE Open Journal of Control Systems, 2022, to appear.

[4] Chasnov, Yamagami, Parsa, Ratliff, Burden, "Experiments with sensorimotor games in dynamic human/machine interaction" Micro- and Nanotechnology Sensors, Systems, and Applications XI , 2019

[5] Zhang, Pace, Burden, Aravkin "Offline state estimation for hybrid systems via nonsmooth variable projection" Automatica , 2020

Georgia Tech | UNIVERSITY of WASHINGTON

Award ID#: 1836932 (Georgia Tech)
1836819 (University of Washington)