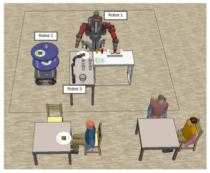
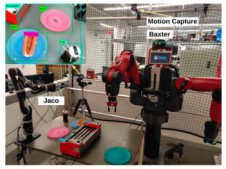
NRI: FND: A Formal Methods Approach to Safe, Composable, and Distributed Reinforcement Learning for co-Robots

PI: Calin Belta, Boston University

Problem Formulation





Given spec, e.g., "Always prepare hotdogs and drinks. If a customer requests drinks or hotdogs, deliver whenever possible. If no customer is requesting or food / drinks are not available, then wait. Always avoid collisions."

Given robot action and sensing capabilities, e.g., Robot 1 can make hotdogs, can place hotdogs on Robot2, can sense when a hotdog is ready, can sense when Robot2 is near, can sense obstacles, Robot 2 can deliver hotdogs and drinks, etc.

Find robot control strategies that satisfy the specs and are provably safe

Challenge

Model-based (optimal) control not appropriate due to complex dynamics and large observation spaces. Reinforcement learning (RL) usually used for such problems. **In RL, it is difficult to**

- Incorporate **complex task** in objective functions,
- Ensure that the learned policy satisfies the **safety requirements**,
- **Transfer** learned policies to unseen tasks,
- Effectively **distribute** a complex task among a robot team that allows each member to learn in a decentralized fashion.

Scientific Impact

Interpretable, provably safe, distributed RL algorithms developed in this projects are relevant to a wide range of **safety-critical CPS**, including autonomous driving, power networks, air traffic controllers, agriculture, military surveillance, search and rescue

Solution: Technical Approach: *Machine (reinforcement) learning + formal methods + optimal control*

- Rich, easy-to-understand, **temporal logic (TL) specification language**; can specify prior knowledge; has a satisfaction metric that can be used to guide the learning process

> (always eventually **MakeHotdog**) and (always eventually **PourDrinks**) and (if *IsNear*(Robot2) and *Ready*(drink) then **PlaceOnRobot2**(drink)) and (if *IsNear*(Robot2) and *Ready*(hotdog) then **PlaceOnRobot2**(hotdog)) and (if *CustomerRequest*(drink) and *OnRobot2*(drink) then **Deliver**(drink) else **TravelTo**(Robot3) followed by **Wait** and (if *CustomerRequest*(hotdog) and *OnRobot2*(hotdog) then **Deliver**(hotdog) else **TravelTo**(Robot1) followed by **Wait**) and always not **Collision**

- **Control Barrier Functions (CBF)** are used to ensure safety
- **Temporal logic guided policy composition method** where policies for new tasks can be constructed from a library of learned policies with little to no additional exploration
- Method that projects a **global specification to a set of local specifications** that each robot can learn to execute.

Scientific Impact

Interpretable, provably safe, distributed RL algorithms developed in this projects are relevant to a wide range of **safety-critical CPS**, including autonomous driving, power networks, air traffic controllers, agriculture, military surveillance, search and rescue

Education and outreach plans

- new courses and sections at the undergraduate and graduate level,
- the involvement of the PI in Research Internships in Science and Engineering (RISE), Discovery Internships, the BU Upward Bound Math and Science (UBMS) program, and the Technology Innovation Scholars Program (TISP) at BU.

Broader Impact:

- Avoiding reward hacking, guaranteeing safety during learning and deployment, use of formal, unambiguous specification languages can potentially encode **robotic safety standards** (ISO)
- Interpretable nature of the formal language has the potential to promote **public trust** in robots

Wenliang Liu, Noushin Mehdipour, Calin Belta, Recurrent Neural Network Controllers for Signal Temporal Logic Specifications Subject to Safety Constraints, IEEE Control Systems Letters (L-CSS), 2021 (in print)

Max Cohen and Calin Belta, Model-Based Reinforcement Learning for Approximate Optimal Control with Temporal Logic Specifications, 24th ACM International Conference on Hybrid Systems: Computation and Control (HSCC), 2021

2021 NRI & FRR Principal Investigators' Meeting
March 10-12, 2021

Award ID#: 2024606