CAREER: A Skill-Driven Cooperative Learning Framework for Cyber-Physical Autonomy

PI: Prof. Xiangnan Zhong (xzhong@fau.edu)

Department of Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL, USA

https://www.nsf.gov/awardsearch/showAward?AWD_ID=2047010&HistoricalAwards=false

Goal: The *goal* of this project is to advance foundational knowledge and scientific methodologies of reinforcement learning for generalization and scalability in cyber-physical systems (CPS).

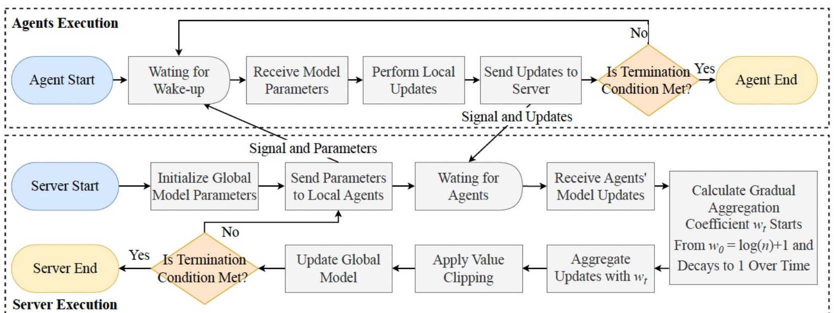
Challenge:

- The nature of many CPS is heterogeneous and high-dimensional, making the hand-coded functions and task-specific information hard to design.
- Large amount of training data is often required for achieving the desired performance which limits the generalization to other tasks.

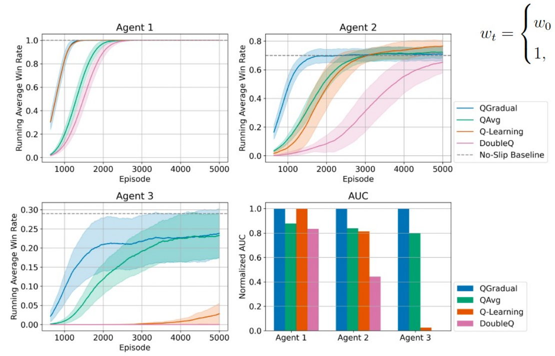
Scientific Impact: This project advances the scientific foundations and methodologies of intelligent control design for CPS in high-dimensional and heterogeneous environment. The developed algorithms and associated architectures have provided critical insights and guidelines to foster autonomous learning and generalization in CPS.

Solution:

Design Federated Reinforcement Learning (FRL) methods for CPS



Flowchart of the developed FRL-QGradual algorithm: to increase the learning efficiency while preserving the agent privacy.



Running average win rates for 3 agents in a heterogeneous stochastic Frozen Lake with varying starting points over 5000 episodes.

Phased weight-adjustment mechanism

$$w_t = \begin{cases} w_0 - \dfrac{t\,(w_0-1)}{\delta}, & \text{if } 0 < t < t_\delta, \\ 1, & \text{if } t \geq t_\delta, \end{cases}$$

Value clipping strategy to prevent overflow

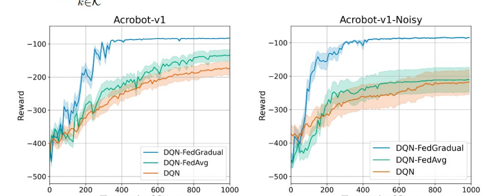
$$\Delta\bar{Q}_t(s,a) = \begin{cases} V_{min}, & \Delta\bar{Q}_t(s,a) < V_{min}, \\ \Delta\bar{Q}_t(s,a), & V_{min} \leq \Delta\bar{Q}_t(s,a) \leq V_{max}, \\ V_{max}, & \Delta\bar{Q}_t(s,a) > V_{max}, \end{cases}$$

Aggregation updates

$$\Delta\bar{Q}_t(s,a) = \frac{w_t}{n}\sum_{k\in\mathcal{K}} \Delta Q_t^k(s,a),\ \forall s,a.$$

Model updates

$$\bar{Q}_{t+1}(s,a) \leftarrow \bar{Q}_t(s,a) + \Delta\bar{Q}_t(s,a),\ \forall s,a.$$



Performance of deep FRL in the heterogeneous Acrobot-v1 and Acrobot-Noisy environments.

Design RL-based methods for hierarchical multiplayer systems

Facilitates the development of more sophisticated and adaptive control approaches and enables AI players to efficiently navigate hierarchical multiplayer environments.

Design Computationally Efficient IRL Approach

A new featurization network to accommodate trajectories and output aligned feature vectors automatically without human intervention.

Broader impact

- Provide critical insights and guidelines to foster autonomous learning and generalization in CPS.
- Integrate research and education plan to enrich the participation of students with different backgrounds.
- Foster student engagement in CPS research and equip the future workforce with expertise in CPS, artificial intelligence, learning, and control.
- Establish learning models into the current curriculum and also integrate the project's cutting-edge research into new courses.