Collaborative Research: Physics-Model-Based Neural Networks Redesign for CPS Learning and Control

Yanbing Mao[1], Lui Sha[2], Naira Hovakimyan[3], Huajie Shao[4], and Weisong Shi[5]

[1]Engineering Technology Division, Wayne State University

[2]Department of Computer Science, [3]Department of Mechanical Science and Engineering, University of Illinois at Urbana—Champaign

[4]Department of Computer Science, William and Mary College

[5]Department of Computer and Information Sciences, University of Delaware

Challenge:

- The current generation of DNNs cannot provide analyzable behaviors and verifiable properties necessary for safety assurance.
- Physics-model-based engineering methods provide analyzable behaviors and verifiable properties but do not match the performance of DNN systems.
- Purely data-driven DNNs are sensitive to spurious correlations and can violate physics laws.

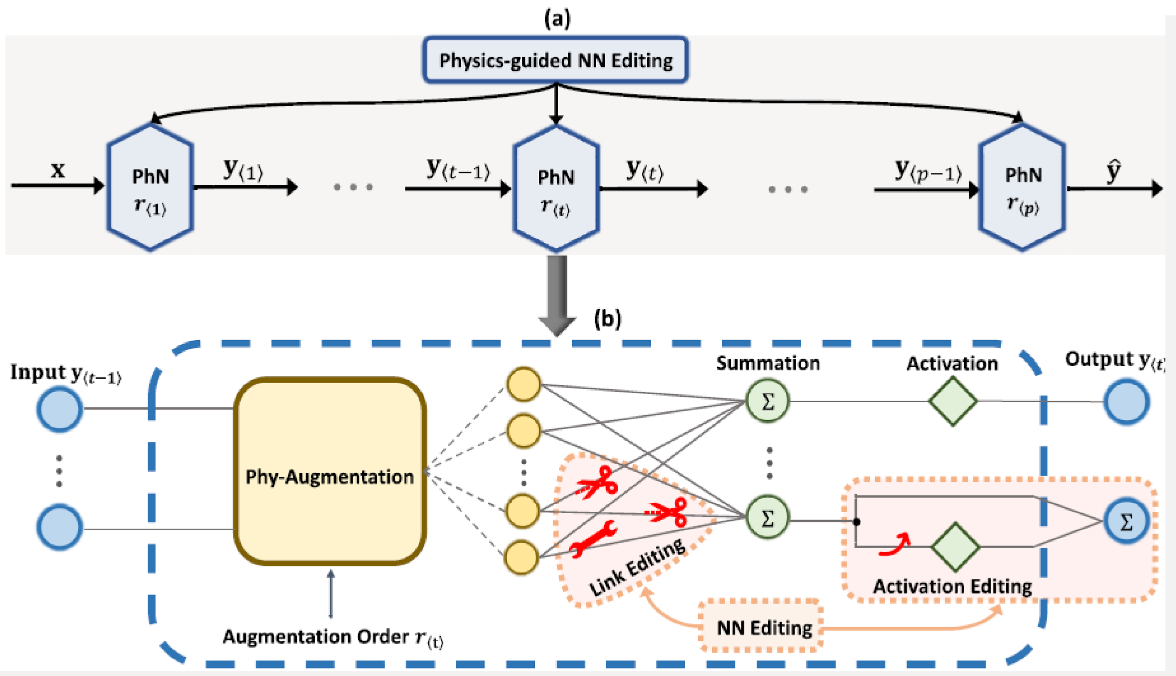
Scientific Impact:

Open the door to embedding scientific knowledge into DNNs that will feature interpretability while offering enhanced model robustness and accuracy.

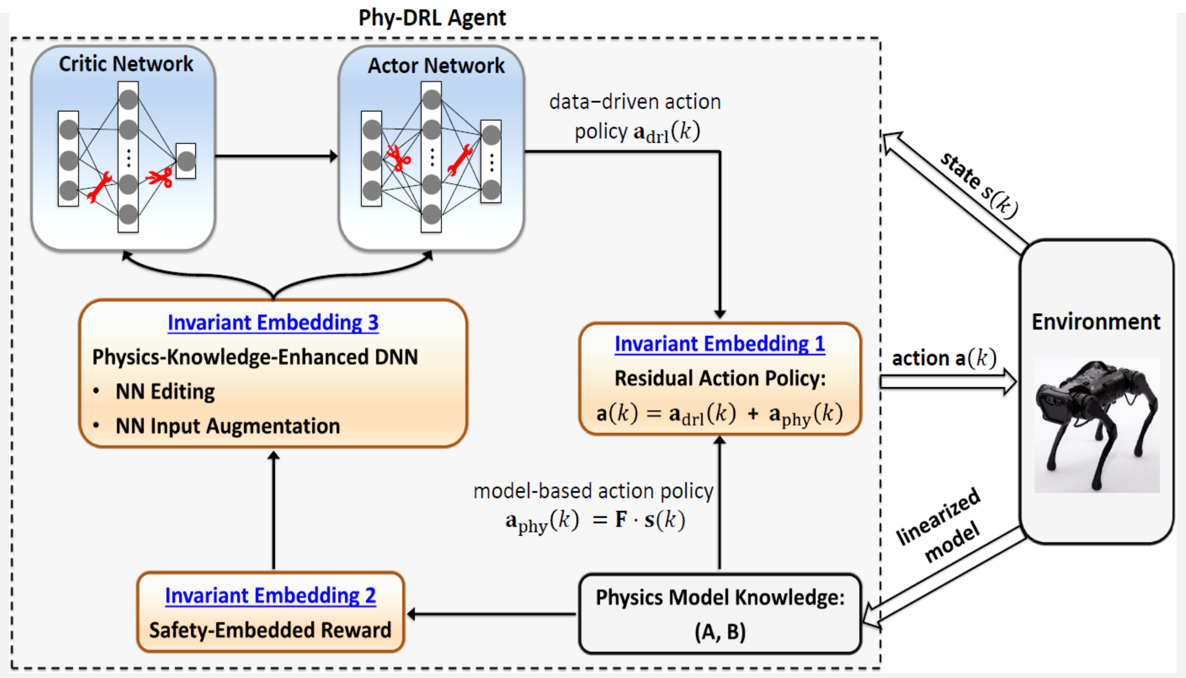
Advance fundamentally integrating deep learning and robust control to enable non-stationary, safety- and time-critical CPS to operate safely with high performance.

Solution:

- **Phy-Taylor: Partially Physics-Knowledge-Enhanced DNN via NN Editing,** aiming at achieving the DNN's strict compliance with available physics laws when applied to safety-critical physical engineering systems.
- **Phy-DRL: Physics-regulated Deep Reinforcement Learning**, aiming at provable safety guarantees for safety-critical robotic systems. It features
  - automatic construction of safety-embedded rewards,
  - mathematically provable safety guarantee,
  - fast training towards safety guarantee.



Phy-Taylor Architecture



Phy-DRL Architecture

Broad Impact: Academic and Technical Impact

- Lead to a new paradigm in how DNNs can be redesigned trustworthily and applied safely to non-stationary, safety- and time-critical CPS.
- Will be readily transferred to industrial use cases to build new CPS infrastructure in modern society, greatly promoting industrial automation and boosting the economy.

Broad Impact: Education and Outreach

- Organize several educational and outreach activities for women and minorities, such as G-BAM at Urbana-Champaign, IL.
- The vehicle platforms will be used to construct participatory demonstrations to engage K-12 students.
- etc.

Broad Impact: Potential Impact

The project will be a milestone toward the goals of DNN with explainable behavior and provable properties. The results will be broadly disseminated to the CPS, machine learning, and control communities through conference presentations, journal publications, and invited seminars.

References:

[1] Hongpeng Cao, Yanbing Mao, Lui Sha, and Marco Caccamo, “Physics-Regulated Deep Reinforcement Learning: Invariant Embeddings,” The International Conference on Learning Representations, Spotlight, pp. 1-45, 2024.

[2] Yanbing Mao, Lui Sha, Huajie Shao, Yuliang Gu, Qixin Wang, and Tarek Abdelzaher, “Phy-Taylor: Partially Physics-Knowledge-Enhanced Deep Neural Networks via NN Editing,” IEEE Transactions on Neural Networks and Learning Systems, DOI: 10.1109/TNNLS.2023.3325432.

[3]. Hongpeng Cao, Yanbing Mao, Lui Sha, and Marco Caccamo, “Physics-Model-Regulated Deep Reinforcement Learning Towards Safety & Stability Guarantees,” IEEE Conference on Decision and Control, Singapore, Singapore, pp. 8306-8311, 2023.

Award ID#: ECCS-2311084 and ECCS-2311085