2021 NSF CYBER-PHYSICAL SYSTEMS PRINCIPAL INVESTIGATORS' MEETING

CAREER: Robustifying Machine Learning for Cyber-Physical Systems

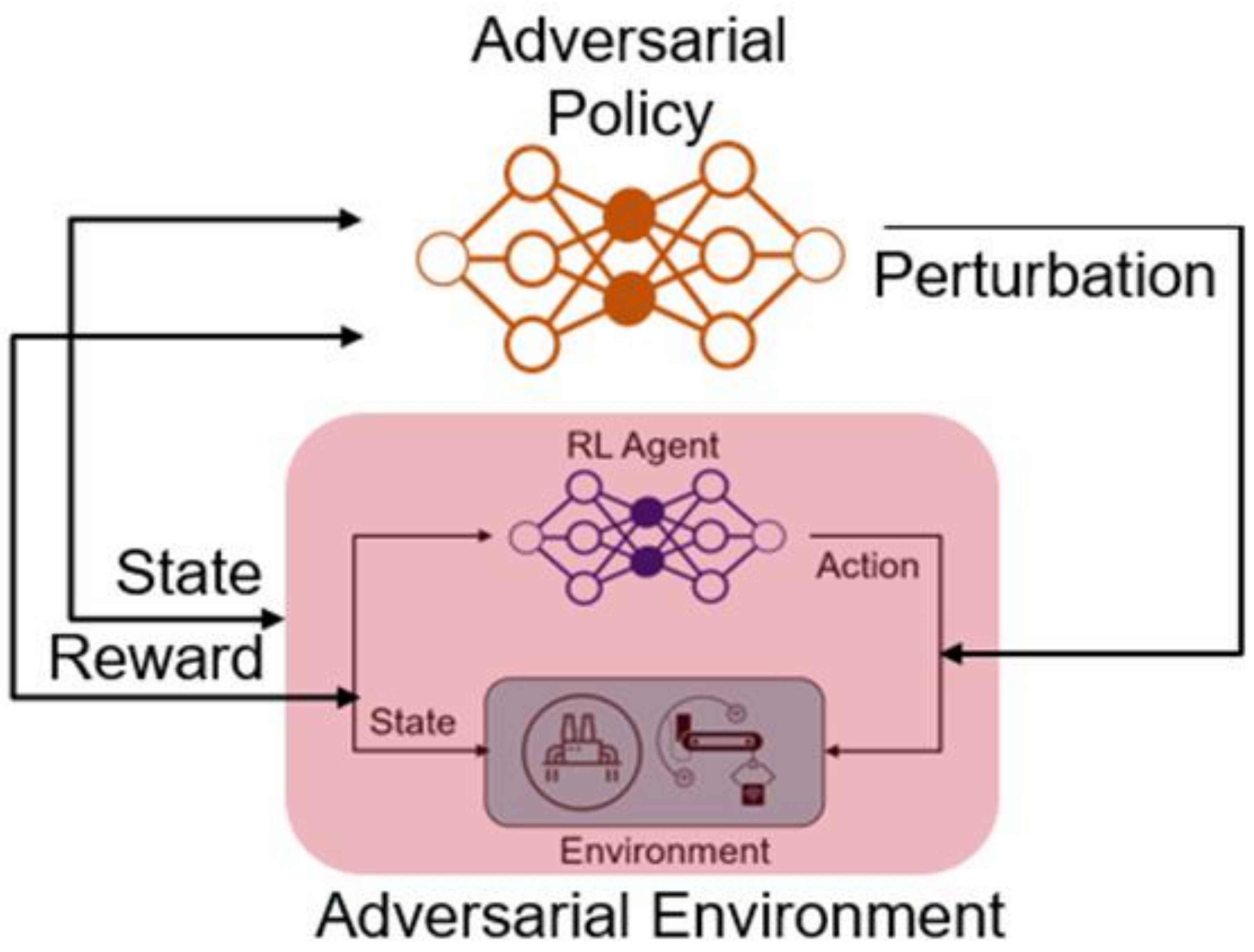
CNS-1845969, Mar 1, 2019 – Feb 29, 2024, PI: Soumik Sarkar, Iowa State University

Challenge:

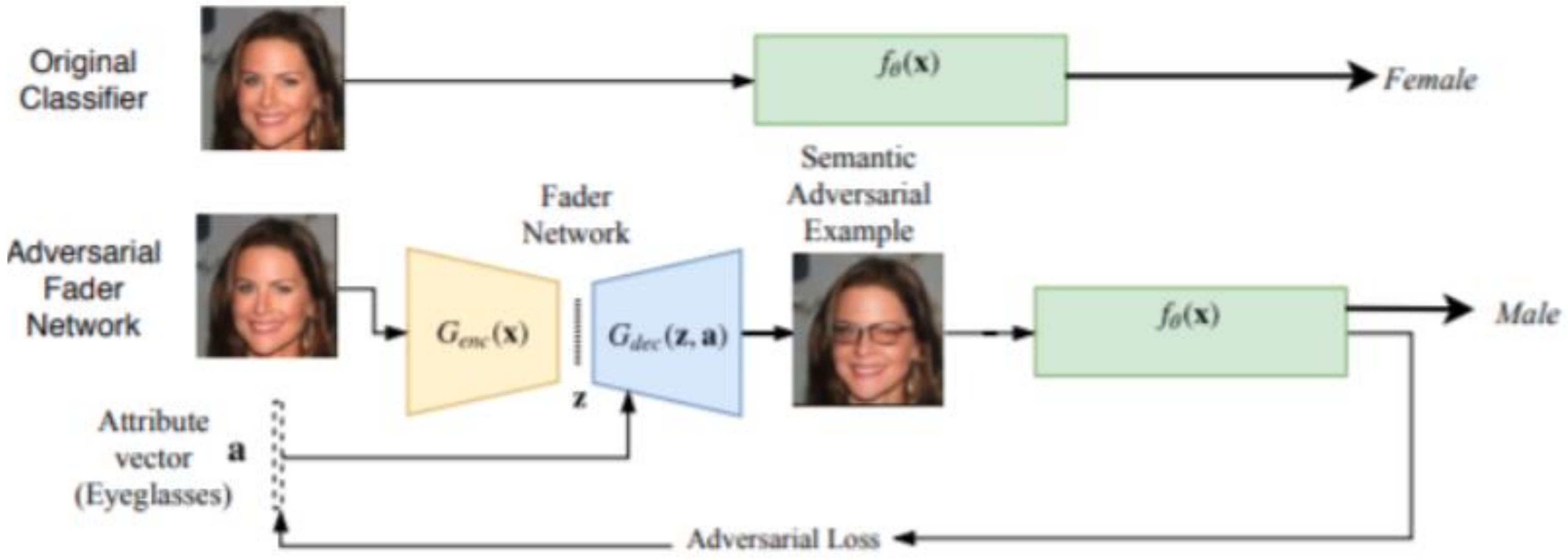
- Understand adversarial attacks on deep learning models for perception (e.g., CNN) & decision-making (e.g., RL)
- Investigate adversarial training schemes to robustify machine learning (ML) algorithms

Solution:

- Optimization based spatio-temporal attacks on RL agents
- Semantic robustness of perception models via generative modeling approach
- Robust optimization for robust learning



Robust Reinforcement Learning (RL)



Semantic Robustness for ML

Scientific Impact:

- Attack models studied are generic and applicable to any commonly used vision and RL-frameworks
- Robust models can be deployed in various CPS applications that leverage ML modules

Broader Impact:

- Robust ML frameworks will be critical for certification and adoption of ML in CPS
- Undergraduate CPS minor at Iowa State
- Supporting PhD study of 2 female and 1 URM students

Contact: soumiks@iastate.edu