Formal Methods and Machine Learning for Safe Control of Autonomous Systems

Morteza Lahijanian, Eric Frew, Majid Zamani

https://www.ariasystems.group/projects/safe-synthesis.html

University of Colorado, Boulder

Ann and H.J. Smead Department of Aerospace Engineering Sciences

Research Aim

Safety-critical autonomous systems should guarantee safe behavior, react to changes to itself and the environment, and accomplish its mission. We are incorporating machine learning with formal methods to verify and control systems with unmodelled dynamics.

Machine Learning is powerful, yet fallible;

Formal methods are rigorous, yet inflexible.

Unmodelled Dynamics

$$x^+ = f(x, a, w)$$

Non-Gaussian Noise

Using data, **efficiently** generate a control policy that is **safe**, and **reacts** to changes in the closed-loop dynamics.

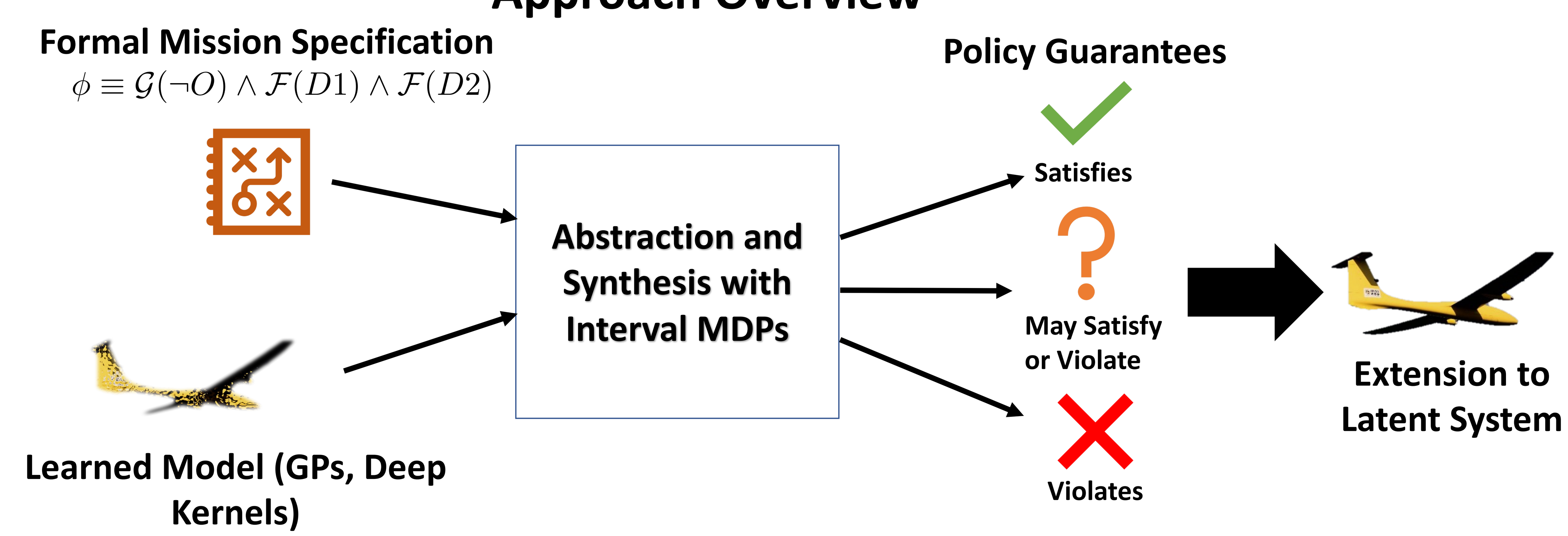
Contributions

- Fusion of machine learning with formal methods;
- Automatically generate safe, mission-accomplishing controllers;
- Address scalability and feasibility challenges;
- Develop toolboxes for end-to-end verification and synthesis.

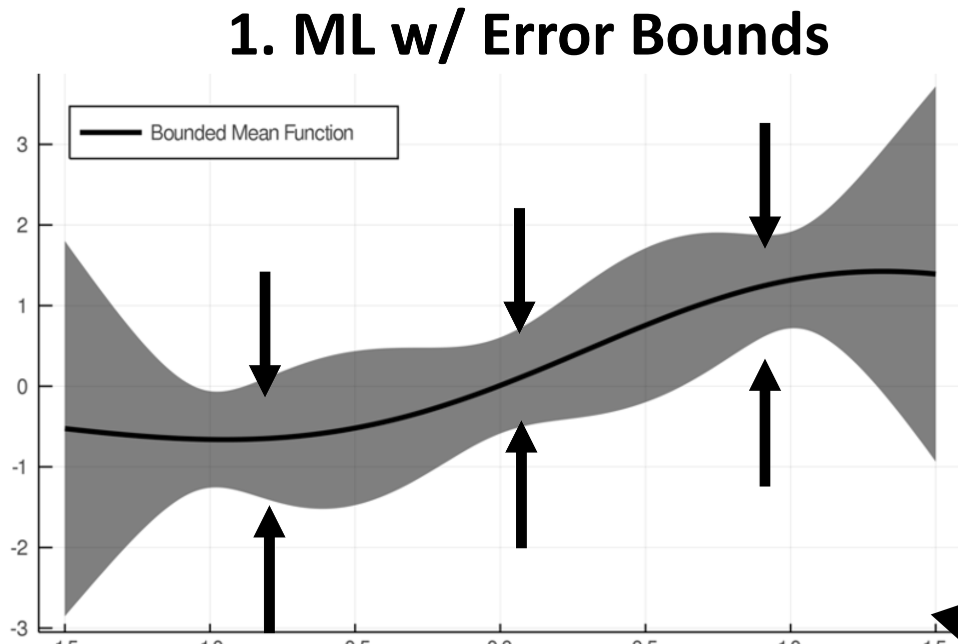
Impact Areas

- Autonomous ground vehicles, advanced air mobility applications
- Energy infrastructure – robust load balancing and forecasting
- Space exploration, extraterrestrial operations and construction

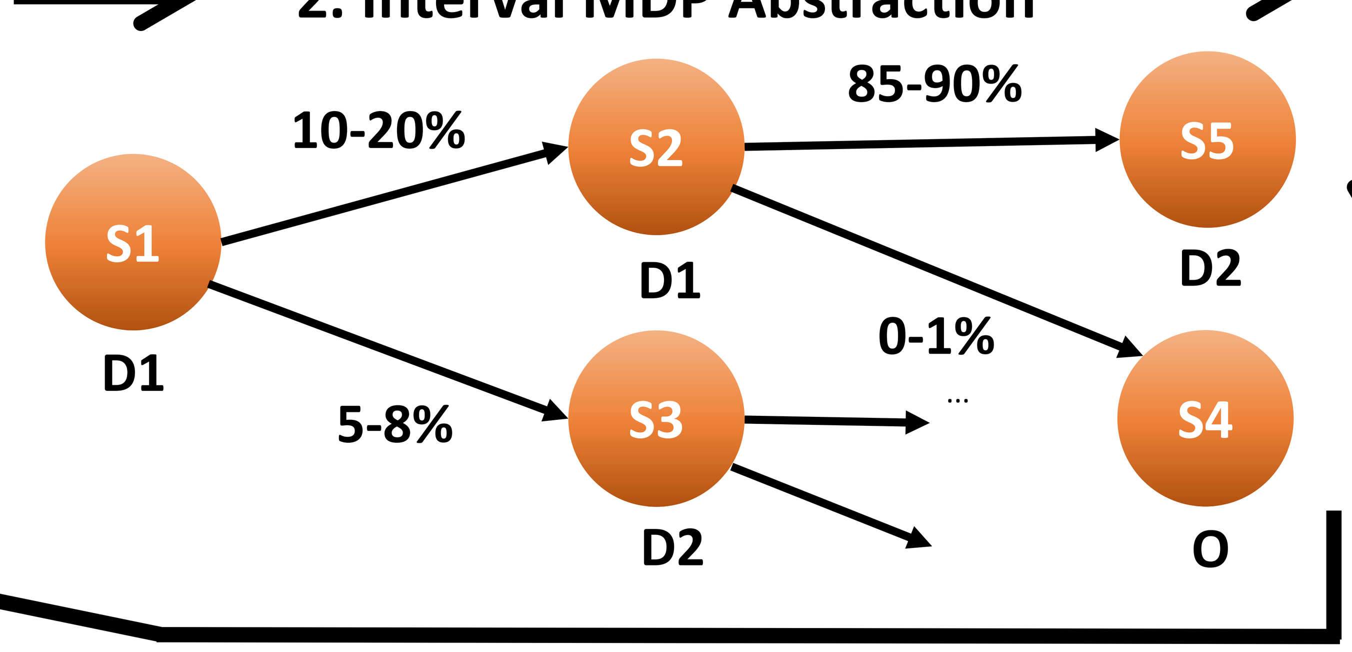
Approach Overview



1. ML w/ Error Bounds



2. Interval MDP Abstraction



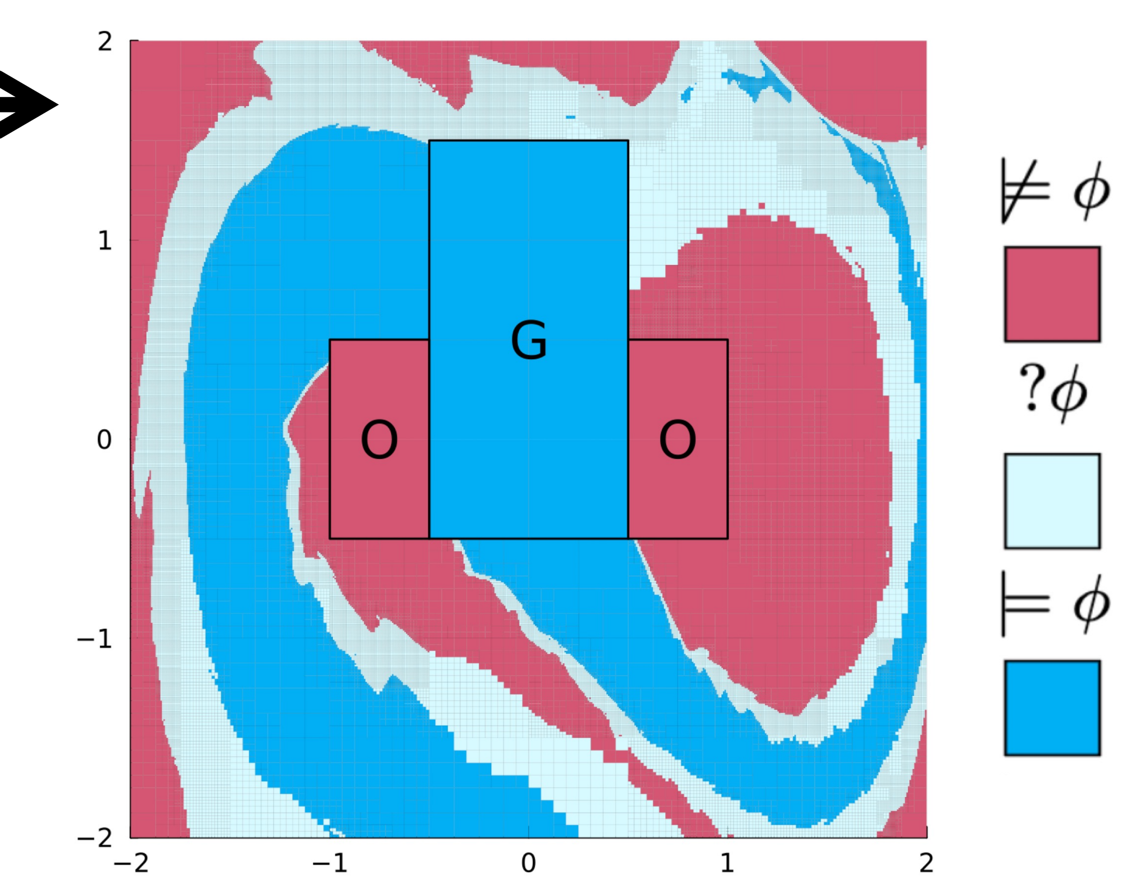
3. Policy Synthesis & Verification

Policy synthesis is a value-iteration problem that is solved with convergence guarantees.

4. Execution and Feedback

Data collected online to reduce error bounds, improve the abstraction and policy.

Ex.



Potential Stakeholders

- Autonomous systems practitioners
- Government and regulatory bodies (FAA, NHSTA, FDA)
- Commercial interests

Outreach

- Undergraduate student research mentoring
- VEX robotics competition volunteering

Broader Impacts

- Robotics to grow to $9.3B market in US by 2028
- Rigorous safety and capability necessary for public and private safety-critical applications (e.g. service robotics)

2024 NSF Cyber-Physical Systems Principal Investigators' Meeting

March 20-21, 2024

Award ID#: 2039062