High-Assurance Design of Learning-Enabled Cyber-Physical Systems with Deep Contracts

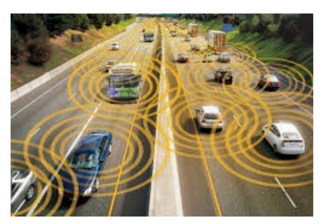
Pierluigi Nuzzo, Department of Electrical and Computer Engineering, University of Southern California

nuzzo@usc.edu

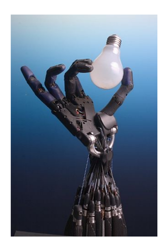
https://descyphy.usc.edu/research/cyber-physical-system-design/

Award ID#: 1846524

Learning-Enabled Cyber-Physical Systems

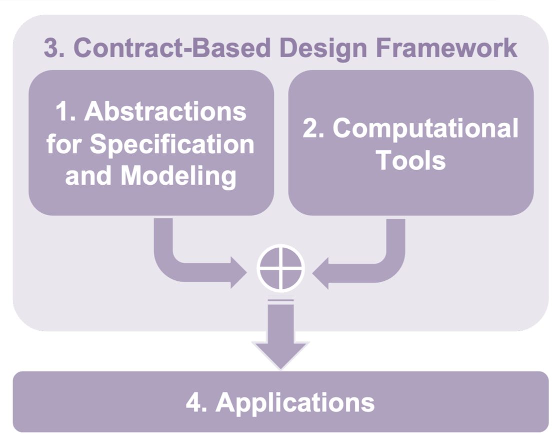


- Modern AI techniques enable ***adaptiveness*** *and* ***resilience*** *of cyber-physical systems, but also bring more* ***complexity, heterogeneity, approximations*** *and* ***uncertainty*** *in the design.*
- ***Requirements*** *are* ***not rigidly defined:*** *How to relate component-level robustness to system-level objectives, such as safety, reliability, performance, cost?*



Source: Wikipedia

Goal: *A holistic framework including modeling techniques, specification formalisms, and scalable algorithms for the design and analysis of intelligent, autonomous, cyber-physical systems including AI-enabled components with high guarantees of correctness in a modular way*



Research Organization

Reasoning with Deep Contracts

Contract $C=(V,A,G)$:
Set $V=I \cup O$ of **variables**
Set A of **assumptions**
Set G of **guarantees**

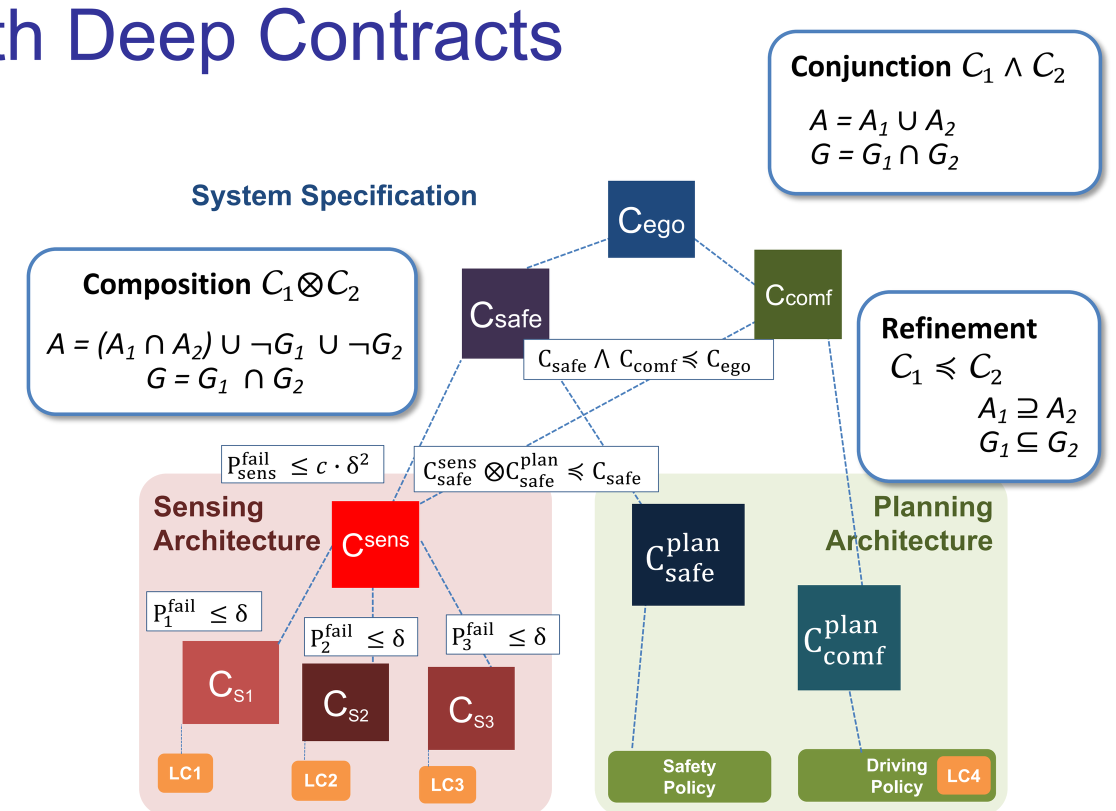
A, G: **behaviors** over V

An **implementation** M satisfies a contract if $M \cap A \subseteq G$

An **environment** E satisfies a contract if $E \subseteq A$

(A, G) is **compatible** iff $A \neq \emptyset$

(A, G) is **consistent** iff $G \neq \emptyset$

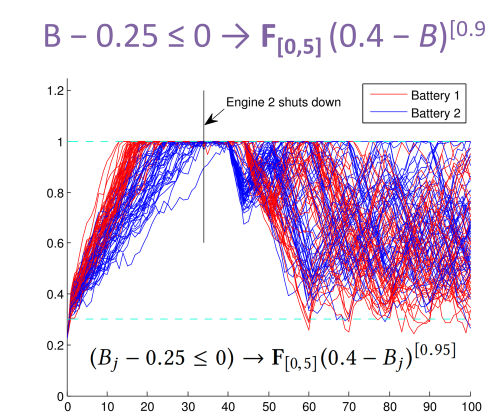


Existing contract frameworks (e.g., [Benveniste et al. '12, Nuzzo et al. '15, '18, '19]) enable **modular verification**, **hierarchical refinement**, and **design reuse** based on a rigorous calculus, but fall short of *effectively capturing uncertainty*, often leading to *pessimistic solutions (over-design)* or *intractable representations*

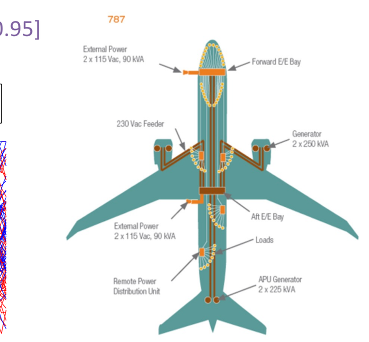
Deep Contracts *for* ***compositional reasoning*** *about* ***probabilistic system behaviors****:*
- ***Context-aware:*** *describe components conditioned to their environment and overall system goals*
- ***Stochastic:*** *express and propagate uncertainty at different abstraction layers*
- ***Vertically-integrated:*** *bridge heterogeneous models and architectures across the design hierarchy*
- ***Pervasive:*** *offers mechanisms to monitor requirements for continual assurance*

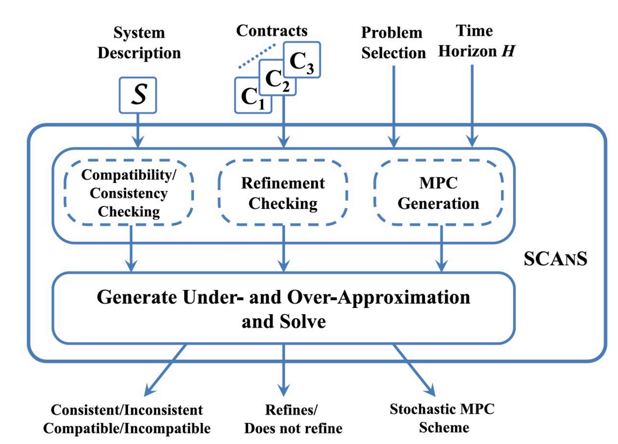
Contract Framework for Stochastic Systems

Leverage *Stochastic Signal Temporal Logic (StSTL)* to express assumptions and guarantees on real-time, real-valued, stochastic signals and formulate verification and synthesis problems as StSTL satisfiability problems



Battery Charge Levels





Enable efficient automatic generation of power management systems for richer specifications than previous solutions [Maasoumy et al. 2013, Shahsavari et al. 2015]

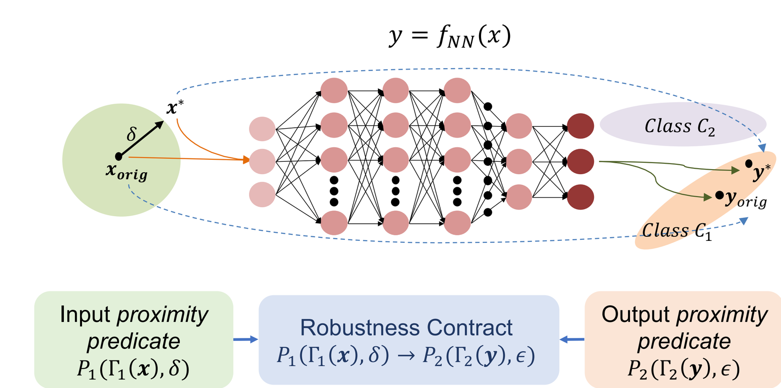
SCAnS (Stochastic Contract-Based Analysis and Synthesis)

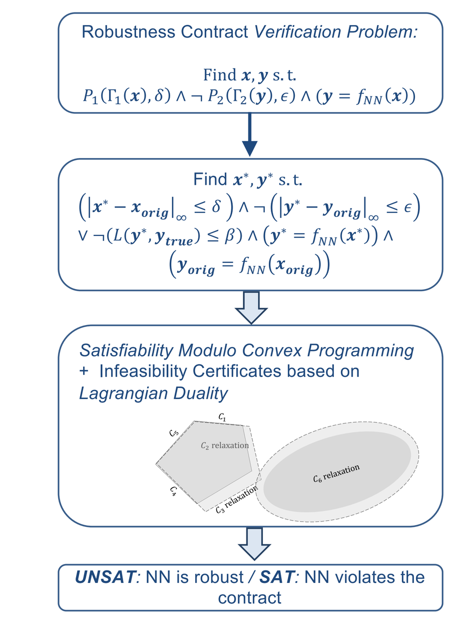
Extension: Optimizing assume-guarantee contracts to deal with performance/cost objectives and rewards in cooperating or non-cooperating multi-agent systems (e.g., connected autonomous cars)

"Stochastic Assume-Guarantee Contracts for Cyber-Physical System Design," *Trans. Embedded Computing Systems*, 2019

"Optimizing Assume-Guarantee Contracts for Cyber-Physical System Design," *Design Automation and Testing In Europe Conf.*, 2019

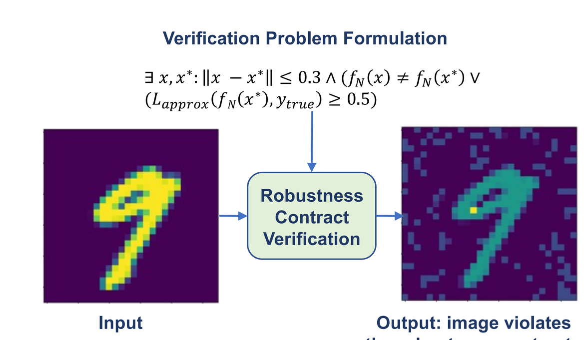
Robustness Contracts for AI-Enabled Components



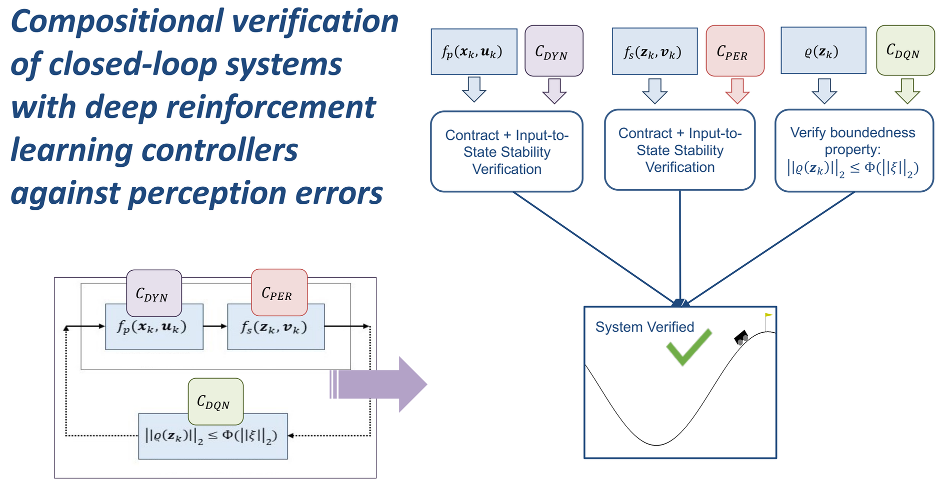


- Generalize many notions of robustness proposed in the literature
- Support sound and complete algorithms based on the coordination of Boolean satisfiability (SAT) solving and convex programming for efficient verification

Verification of neural network-based components against multiple robustness criteria



Compositional verification of closed-loop systems with deep reinforcement learning controllers against perception errors



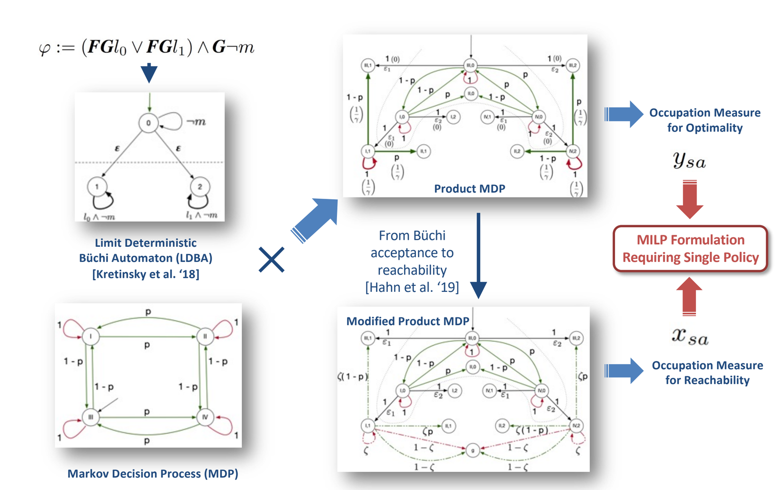
N. Naik and P. Nuzzo, *Int. Conf. Formal Methods and Models for System Design*, 2020, Best Paper Award

Synthesis of Optimal Control and Reinforcement Learning Policies from Rich Contracts

Optimal Control of Markov Decision Processes (MDPs) Under Temporal Logic Specifications

- "Soft" objective: Optimize discounted reward optimality over infinite horizon
- "Hard" constraint: Mission-critical task expressed in general linear temporal logic (LTL) must hold with probability 1

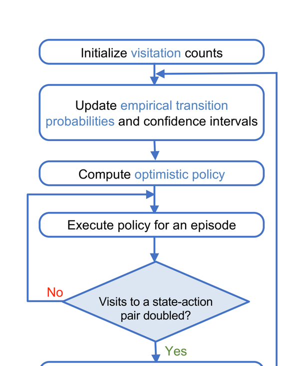
Key Insight: Optimality and LTL satisfaction can be both expressed via occupation measures that can be matched to the same deterministic policy



Sample-Efficient Reinforcement Learning for Finite-Horizon Constrained MDPs

- Uncertain environments and unknown dynamics
- Multiple reward objectives and constraints

Key Insight: Express optimal control of constrained MDPs as a linear program via occupation measures and exploit optimism in the face of uncertainty principle for learning efficiency



Probably Approximate Correctness (PAC)-Based Complexity Bound

$$\tilde{O}\left(\frac{|S||A|C^2H^2}{\epsilon^2}\log\frac{1}{\delta}\right)$$

PAC sample complexity has quadratic dependence on the horizon length

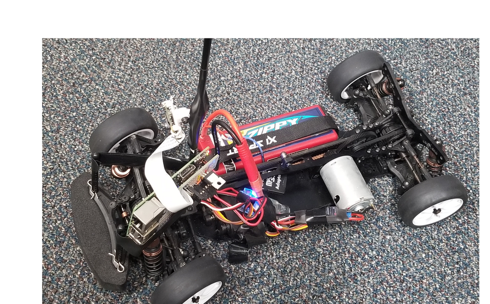
"A Sample-Efficient Algorithm for Episodic Finite-Horizon MDP with Constraints", *AAAI Conf. Artificial Intelligence*, 2021

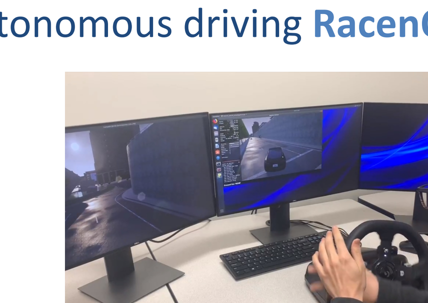
"Optimal Control of Discounted-Reward Markov Decision Processes Under Linear Temporal Logic Specifications," *American Control Conf.*, 2021

Impact on Society and Education

- Provide the foundations for rapid, compositional, certified design and operation of adaptive and resilient learning-enabled cyber-physical systems for a broad range of applications: autonomous vehicles, robotics, industrial automation, medical devices, …
- Research outcomes are part of an educational program focusing on systems engineering concepts and multidisciplinary methods to realize safe and cost-effective intelligent systems interacting with people
  - Pre-college: via the USC Viterbi SHINE Program
  - Undergraduate and graduate: via new labs and collateral initiatives such as the USC AutoDRIVE Lab, the USC Autonomous Vehicles Club, and the USC autonomous driving RacenOn! competition

SHINE
Summer High School Intensive in Next-Generation Engineering





AutoDRIVE LAB