Research Statement

Sicong Guo (stevengu@umich.edu)

Vision: Safe and Trustworthy Autonomy

The paradigm of Cyber-Physical Systems (CPS) has fundamentally transitioned from the analysis of isolated and individual units to the **compositional design** of large-scale, interconnected networks operating in highly dynamic environments. From fleets of connected automated vehicles (CAVs) regulating dense urban traffic to distributed multi-agent systems coordinating complex logistics (*e.g.*, modular reconfigurable robots capable of autonomous adaptation for disaster response on Earth or resilient infrastructure assembly in space), these emerging examples exhibit immense potential for addressing critical CPS demands in sustainable urban mobility, resilient critical infrastructure, and energy-efficient distribution.

However, the deployment of such systems introduces a set of stringent, interdependent requirements that are difficult to satisfy simultaneously: rigorous safety guarantees, operational performance optimization, and the ability to scale to higher-dimensional complexity. Standing at this critical juncture, the research community faces a methodological divide. On one hand, classical control theory, rooted in robust stability and invariant set analysis, provides the necessary **mathematical certifiability** for safety-critical infrastructure. Yet, these methods often struggle to scale to the high-dimensional and nonlinear complexity of global system-level interactions. On the other hand, the surge in data-driven techniques, particularly Reinforcement Learning (RL), has expanded the operational boundaries of CPS. RL offers unprecedented scalability and the capability to solve complex, multi-stage tasks in unstructured environments where analytical models are often intractable. Yet, these methods lack the rigorous, **correct-by-construction** safety certifications required for deployment. The "black-box" nature of deep policy networks means that statistical success in training does not necessarily imply provable safety in novel scenarios. Consequently, a critical gap remains: the development of a unified methodology that leverages the adaptivity and scalability of learning-enabled systems while enforcing the provable safety and stability guarantees (via formal control theory) within a compositional framework.

I aim to study the rigorous synthesis of these distinct fields: Control Theory, Optimization, and Formal Methods. My vision is to establish a *component-symbolic scheme* via compositional perspectives: the components (specifically *learning-enabled subsystems*) enable advanced control synthesis building upon the high-dimensional perception and adaptive decision-making required for complex environments; simultaneously, the symbolic layer (comprising automatas, temporal logic and contracts) provides the structured guardrails and interpretability necessary for *safety* and *certified trust*. In such framework, learning-enabled subsystems are not merely tasked with realizing effective decision-making, but are constrained and guided by *formal specifications*, ensuring safety and trustworthiness in autonomous operations upon *system-level design*.

My past research spans **intelligent transportation systems**, **wearable robots**, and **distributed manufacturing systems**. To address network-induced degradations in connected engineering testbeds, I developed a model-free delay compensation framework that predicts future system states from delayed remote signals without requiring an explicit plant model [2]. This shaped my view of connectivity as a coordination mechanism whose value depends on robustness to network-induced *latency* and *asynchrony* in the closed loop. Extending to mixed autonomy, I designed *fleet-level* traffic control by coordinating a *paired* set of CAVs via V2X/connected cruise control to regulate the intervening human-driven platoon, with the objective of attenuating disturbances and suppressing stop-and-go waves even under sparse CAV penetration [3]. This progression motivates my focus on **specification-driven compositional** design: formalizing temporal requirements, enforcing them with runtime monitors, and diagnosing assumption violations to trigger contract-aware fallbacks. In the domain of wearable robots, I used convex optimization to synthesize parallel elastic actuators for energy-efficient actuation [1]. By modeling motor energy (including Joule-heating-driven losses) as a convex function of spring stiffness, I obtained *globally optimal* elasticity designs under actuator constraints that enables significant energy consumption reduction, demonstrating how *constraint-aware* convex formulations enable globally optimal and *certifiably feasible* component-level designs. In distributed manufacturing domain, I developed a collaborative process parameter recommender system formulated as a *sequential matrix completion* problem, enabling networked manufacturing systems to collaboratively optimizing their performances from sparse data, without exhaustive testing [5].

Building on this foundation, my current research bridges learning with formal specifications to engineer compositional autonomy that is *amenable to runtime assurance* through monitorable requirements and diagnosable assumption violations. A central challenge in learning-based autonomy is that many safety and rule-compliance requirements are *temporal*: satisfaction depends on the *history* of events, not solely the current state. At unsignalized urban intersections, traffic conventions such as "stop-before-go" and "yield-until-clear" induce *non-Markovian* reward structure that violates standard deep RL assumptions. I therefore leverage *reward machines* [4] to encode these requirements as finite-state task specifications and integrate them into RL framework, reducing reliance on ad-hoc reward shaping while yielding interpretable, auditable decision logic. In parallel, I develop *contract-based* reward machines for

multi-agent learning in settings such as highway merging and lane-changing: *inductive* assume–guarantee contracts are compiled into local reward machines that serve as monitorable interfaces among agents, avoiding vacuous satisfaction and enabling compositional safety reasoning. Together, these threads support a *specification-to-assurance pipeline* in which I formalize temporal requirements and assume–guarantee interfaces, synthesize locally-learned policies that compose with these abstractions, monitor reward-machine progress and contract satisfaction online, and when an assumption invalidation is indicated, localize the breach and invoke targeted runtime-assurance supervisors (e.g., conservative car following).

Looking ahead, my future research aims to formalize a specification-to-assurance pipeline for *system-level safe and trustworthy autonomy*: learning interpretable task structure (e.g., reward machines and assume–guarantee contracts) from observed data and demonstrations; enabling *compositional design* via algebraic contract operations such as *refinement* and *quotient*; and coupling the resulting architectures with *runtime verification*, *statistical model invalidation*, and *diagnosable, supervisor-enforced fallbacks* for when assumptions are violated. I will pursue *inferential specification learning* to synthesize policy-equivalent reward machines and assume–guarantee interfaces for non-Markovian, multi-agent tasks, and develop *repair-oriented contract operations* (potentially grounded in refinement and quotient) that enable principled and localized redesign when global guarantees fail by updating only the responsible component while keeping the rest fixed. To support scalable verification and runtime assurance, I will leverage *lifted hybrid abstractions* with adaptive refinement to map nonlinear dynamics and multi-agent interaction into tractable representations that enable efficient runtime verification, statistical model invalidation, and diagnosis under tight compute and latency constraints.

Past Research: Foundations in Control and Optimization

Intelligent Transportation Systems My research began with the core cyber-physical challenge of integrating *geographically-separated* engineering testbeds over delayed networks [2]. To address network-induced degradations without relying on an explicit plant model, I developed a model-free delay compensation framework that predicts future states from delayed remote signals, *achieving up to* 30% *higher integration fidelity.* I then characterized the predictor's delay-dependent stability region, yielding stability boundaries in the design parameters as a function of communication delay. Building on this foundation, I moved from single-system integration to *fleet-level* mixed-autonomy traffic regulation [3]. I proposed a connected cruise and traffic control strategy in which *pairs* of CAVs coordinate via V2X/connected cruise control to stabilize the human-driven platoon between them; aligning with head-to-tail string-stability analysis, the simulation studies show that coordinated CAV pairs can guarantee stability even at *lean penetration (as low as* 10%*)*, enabling system-level congestion mitigation.

Wearable Robots I sought the optimal design of parallel elastic actuators in wearable robots like prostheses [1]. The core challenge was to minimize energy consumption (including Joule-heating-driven losses) while satisfying torque and velocity constraints. I formulated the design problem as a *convex optimization* problem by modeling the motor energy consumption as a convex function of the spring stiffness, which guarantees the feasibility of global optimal. This results in optimal designs that *reduced energy consumption by up to 63%* compared to rigid actuators. This experience reflects my commitment to pursuing globally optimal solutions not only in design, but also in ensuring system safety.

Distributed Manufacturing Systems I addressed the scalability of process parameter optimization across fleets of networked manufacturing systems [5]. A major challenge in such systems related to *machine-to-machine variability*, where identical models exhibit distinct behaviors due to assembly tolerances and wear. To address this, I developed a collaborative process parameter recommender system, formulated as a Sequential Matrix Completion task. By stacking the partially-observed process-parameter operating condition data into a low-rank matrix, I utilized *spectral clustering* to identify latent similarities between machines, and *alternating least squares* to recover missing performance entries. This enabled peer-to-peer collaboration across networked manufacturing systems, allowing machines to learn optimal local control parameter configurations with substantially faster convergence. Experimental validation on a fleet of 3D printers demonstrated up to a *40% reduction* in the amount of trials required to optimize process parameter compared to the non-collaborative baseline.

Current Research: Bridging Learning and Formal Specifications

Traffic Rule Compliant Autonomous Driving Urban driving scenarios are governed by *temporal* traffic rules, *e.g.*, stopping before proceeding, yielding until the intersection is clear, and resolving right-of-way based on arrival order. At unsignalized intersections, these requirements must be satisfied under partial observability (due to limited sensing) and strategic multi-agent interactions (with pedestrians, cyclists, and non-cooperative drivers). A key difficulty is that such traffic-rule compliance is inherently *non-Markovian*: the permissibility of "go" depends on prior events (*e.g.*, whether the vehicle has stopped, who arrived first, whether yielding obligations were met), which violates the Markov reward

assumption in standard deep RL and makes naive reward shaping brittle. I therefore leverage *reward machines* [4] to encode traffic rules as finite-state specifications over logical events and integrate them into RL framework via a product Markov Decision Process; therefore, the policy reasons jointly over physical state and specification progress. This reduces reliance on ad-hoc heuristics while yielding interpretable, auditable decision logic and a natural substrate for runtime monitoring of rule satisfaction.

Contract-Based Reward Machines In multi-agent settings (*e.g.*, highway merging and lane-changing), learning safety-critical behavior requires more than local optimality: agents must coordinate under partial information while ensuring that local decisions compose into global safety. I develop *contract-based* reward machines that compile assume–guarantee contracts formulation into local reward machines. Building on *inductive* assume–guarantee contracts, which enforce causality by requiring guarantees at the *next time step* $(t+1)$ based on assumptions up to time t, I validate a circular composition principle under compatibility conditions and use the resulting contracts as monitorable interfaces among agents. Operationally, each agent's local reward machine both (i) shapes incentives toward satisfying its guarantee and (ii) exposes when its assumptions about others are violated, enabling blame assignment and targeted safety fallbacks. I evaluate this framework in simulation on prototypical highway scenarios using `highway-env`, demonstrating decentralized training where robust global behavior emerges from locally-checkable specifications.

Future Research: Towards System-Level Safe and Trustworthy Autonomy

My long-term goal is to make learning-enabled autonomy *engineerable*: requirements are explicit, components compose predictably, and safety can be monitored, diagnosed, and repaired at runtime. Building on reward machines and inductive assume–guarantee reasoning, my future work will develop a unified specification-to-assurance workflow spanning: (i) *inferential specification learning*, (ii) *compositional contract operations for repair*, and (iii) *scalable runtime assurance*.

Inferential specification learning for non-Markovian autonomy. I plan to study when a finite-state reward machine is *policy-equivalent* to an underlying non-Markovian task, characterize minimality and identifiability relative to the available propositional event set, and connect reward placement (incremental vs. terminal) to exploration (including maximum-entropy objectives) to derive learnability-aware design principles. Algorithmically, I plan to synthesize reward machines from expert demonstrations and closed-loop execution traces using *SAT-based automata inference* with *counterexample-guided refinement*: when held-out traces or runtime monitors expose ambiguous event predicates, misspecified transitions, or unintended temporal behavior, the resulting counterexample traces are compiled into additional constraints to iteratively update the event predicates, automaton structure, and reward shaping, while explicitly handling occlusions and partial observability via uncertainty-aware event abstraction.

Compositional contracts for multi-agent learning and targeted repair. To scale from single-agent policies to *collaborating* agents, I plan to develop *repair-oriented contract operations* grounded in inductive assume–guarantee interfaces, compiling each interface into a *contract-compiled reward machine* (a local finite-state monitor) that both shapes incentives and exposes monitorable assumptions and guarantees. Beyond establishing compatibility and circular composition conditions, I aim to develop *targeted repair*: when a system-level guarantee fails, I will localize the breach to a specific agent and contract clause, compute a quotient-style "missing responsibility" specification under the fixed context of the other agents, and refine only the responsible agent's contract and local reward machine to restore the global guarantee. I also plan to evaluate and validate such scheme across mixed-autonomy driving tasks, including multi-lane highways (merges and lane changes) and urban unsignalized intersections in hierarchical/multi-task settings, measuring task completion and throughput efficiency, and the ability to recover guarantees after localized refinement.

Scalable runtime assurance via lifted hybrid abstractions and invalidation-driven refinement. To keep verification and monitoring tractable in nonlinear CPS with tight real-time constraints, I intend to use *lifted hybrid abstractions* with adaptive refinement: lift dynamics and constraints into representations that admit efficient runtime verification while preserving a refinement order so the model is tightened only where needed. These abstractions will be coupled with statistical model invalidation and runtime supervision: when data contradicts abstraction assumptions, the system either refines locally to restore soundness or switches to a conservative supervisor while collecting informative data. This closes the loop between learning, monitoring, diagnosis, and repair, enabling system-level guarantees that remain meaningful under deployment-time uncertainty.

To make reward machine and contract guarantees tractable in nonlinear CPS under real-time constraints, I intend to use *lifted hybrid abstractions* with adaptive refinement, lifting dynamics and constraints into representations that support efficient runtime verification while preserving a refinement order so tightness is increased only where needed. Coupled

with statistical model invalidation and runtime supervision, this layer detects when deployment data contradicts abstraction assumptions, then either refines locally to restore soundness or triggers a conservative supervisor while collecting informative data. In doing so, it implements the specification-to-assurance workflow by grounding learned specifications and contracts in runtime-verifiable abstractions of the physical dynamics and enabling diagnosis and repair under deployment-time uncertainty.

References

[1] Sicong Guo, Robert D. Gregg, and Edgar Bolívar-Nieto. Convex optimization for spring design of parallel elastic actuators. In *2022 American Control Conference (ACC)*, pages 3688–3694, 2022.

[2] Sicong Guo, Yuzhang Liu, Yingshi Zheng, and Tulga Ersal. A delay compensation framework for connected testbeds. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 52(7):4163–4176, 2022.

[3] Sicong Guo, Gábor Orosz, and Tamas G. Molnar. Connected cruise and traffic control for pairs of connected automated vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 24(11):12648–12658, 2023.

[4] Rodrigo Toro Icarte, Toryn Klassen, Richard Valenzano, and Sheila McIlraith. Using reward machines for high-level task specification and decomposition in reinforcement learning. In *International Conference on Machine Learning*, pages 2107–2116. PMLR, 2018.

[5] Weishi Wang, Sicong Guo, Chenhuan Jiang, Mohamed Elidrisi, Myungjin Lee, Harsha V. Madhyastha, Raed Al Kontar, and Chinedum E. Okwudire. A collaborative process parameter recommender system for fleets of networked manufacturing machines – with application to 3d printing, 2025.