STEVENS
INSTITUTE OF TECHNOLOGY
1870

Ph.D. DISSERTATION DEFENSE

Candidate:	Balaji Rao
Degree:	Doctor of Philosophy
School/Department:	Charles V. Schaefer, Jr. School of Engineering and Science/Department of Systems Engineering
Date:	8th December 2026
Time/Location:	3:30 pm / Babbio 503
Title:	Enhancing Reliability of LLM Generated Content Through Structure and Machine Checks
Chairperson:	Dr. Carlo Lipizzi, Department of Systems Engineering, School of Engineering & Sciences
Committee Members:	Dr. Mo Mansouri, Department of Systems Engineering, School of Engineering & Sciences
	Dr. William Eiers, School of Computing
	Dr. Juneyoung Lee, Amazon Web Services
	Dr. David Darian Muresan, Department of Systems Engineering, School of Engineering & Sciences

ABSTRACT

Large Language Models (LLMs) continue to make remarkable progress in the field of natural language processing; they have shown great potential in generating human-like text and assisting with complex tasks that range from Natural Language Understanding (NLU) to code generation. However, their propensity to hallucinate makes them brittle in high-stakes settings because they generate confident but unverifiable content. This dissertation advances a structure-and-check paradigm: constrain generation into machine-interpretable artifacts and subject those artifacts to external verification and (or) validation. Our study on the Identification of Variables Impacting Cascading Failures in Aerospace Systems looks at "LLMs as a crutch" for data-science decision-making, which surfaces the successes of using LLMs within strict bounds of an analytical framework and motivates structural safeguards in end-user workflows. Our study of a course-specific tutoring assistant (SSE-Edubot) provides qualitative examples of more course-relevant answers after domain adaptation through retrieval and supervised fine-tuning. Our study on a systems thinking view of an AI Economy provided the socio-technical rationale (governance, auditability, and accountability) for demanding evidence-carrying AI outputs in real deployments. Finally, ProofSeek studies coupling LLMs with formal and symbolic backends and reports proof-success and execution-time comparisons on mathematical and access-policy tasks.

Building on this foundation, this dissertation contributes three technical pieces: (1) Applied reliability (Steve): a schema-first, ontology-driven career-coaching system that turns resumes and interviews into inspectable JSON files, aligns skills to a configurable career tree, and retrieves courses via structured queries

from a grounded knowledge base. Evaluated studies show reasonably consistent outputs under repeated runs, with transparent, inspectable artifacts. (2) Neural Theorem Proving and s2n-bignum-bench: a benchmark for machine-checkable tactic synthesis over low-level cryptographic assembly proofs in HOL Light, together with supervised and proof-reward post-training of an interactive tactic policy. (3) MBSE autoformalization: a typed, solver-guided workflow for producing initial SysML requirements models from contextualized natural-language requirements. Its contribution is an inspectable connection between source interpretation, generated constraints, bounded revision, and automated checking evidence.