Neuro-Symbolic Bridge: From Perception to Estimation & Control

PIs: *Radoslav Ivanov (RPI), Ivan Ruchkin (UF)*

Students: Thomas Waite, Shuhang Tan (RPI), Yuang Geng, Trevor Turnquist (UF)

Rensselaer Polytechnic Institute, University of Florida

Problem:

CPS rely on neural networks (NNs) for perception and symbolic techniques for estimation/control. A **fundamental mismatch** exists between the uncertainties in NN outputs and the assumptions of symbolic tasks, compromising **system safety and performance**.

Key idea:

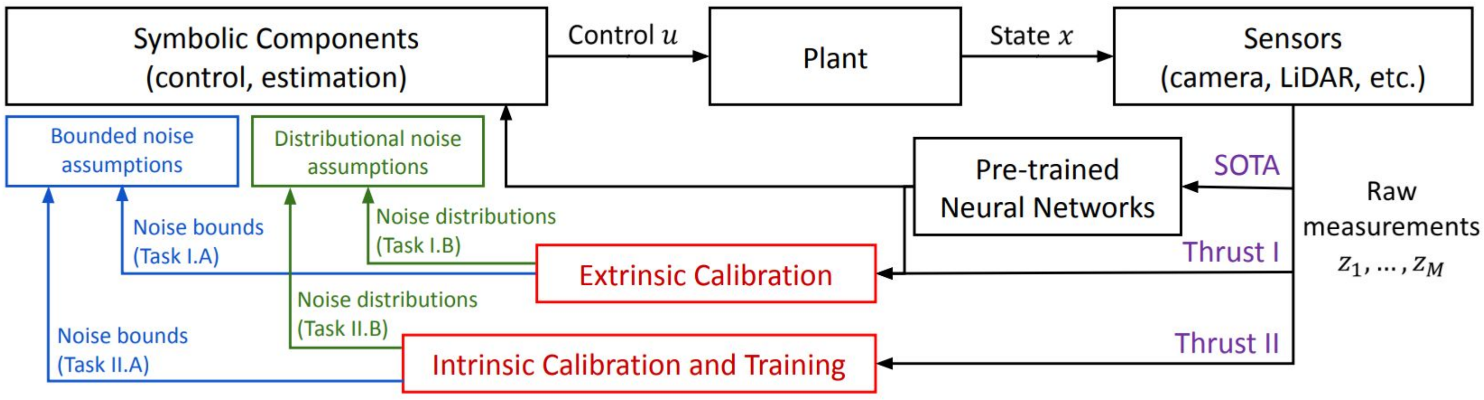
A *neuro-symbolic calibration and training* to bridge the gap between NN perception and downstream symbolic tasks:

- **Be robust** to sensor faults and perception errors.
- **Align NN uncertainties** with symbolic task assumptions (e.g., bounded or Gaussian noise).
- **Ensure temporal consistency** of calibrated outputs with system dynamics.

Broader Impacts:

- **Mentoring:** PI Ivanov is advising a high school student. PI Ruchkin is mentoring several undergraduates.
- **Workshop on neuro-symbolic CPS:** PI Ruchkin is co-chairing a CAV'25 workshop on neuro-symbolic CPS https://www.tacps.org

Approach Overview:



- *Pre-trained NN perception* processes raw sensor data.
- *Thrust I*: Extrinsic Neuro-Symbolic Calibration (post-training)
- *Thrust II*: Intrinsic Neuro-Symbolic Calibration (training & calibration)
- *Symbolic estimation*: calibrated NN outputs for improved safety and performance.

Results:

1. State-Dependent Conformal Perception Bounds for Neuro-Symbolic Verification of Autonomous Systems

Problem: given system dynamics and a dataset of IID trajectories, construct a sequence of reachable sets such that:

$$\mathbb{P}[x_k \in \mathcal{X}_k, \forall k = 1..T] \geq 1 - \alpha$$

- **Intuition**: perception errors vary across state space; can we design a state-dependent conformal prediction method?

Solution:

- Split state space into regions to minimize conformal error
  - Defined weighted loss of accumulated error

$$\min_{\mathcal{X}_1,\dots,\mathcal{X}_M} \sum_{i=1}^{M} \sum_{x_{j,k} \in \mathcal{D}} \mathbf{1}_{\{x_{j,k} \in \mathcal{X}_i\}} e_i$$

$$\text{s.t. } \mathcal{N}_i = \{n_j \mid n_j = \max_{x_{j,k} \in \mathcal{X}_i} \|g(x_{j,k}) - y_{j,k}\|_\infty \text{ for all } j = 1..N, k = 1..T\}, i = 1, \dots, M,$$

$$e_i = \inf\left\{\hat{q} \mid \frac{\|\{n \in \mathcal{N}_i \mid n \leq \hat{q}\}\|}{\|\mathcal{N}_i\|} \geq \lceil \|\mathcal{N}_i\| + 1 \rceil (1 - \frac{\alpha}{M})\right\}, i = 1, \dots, M.$$

- Synthesize state regions $\mathcal{X}_1, \dots, \mathcal{X}_M$ with **gradient-free optimization** (genetic search and simulated annealing)

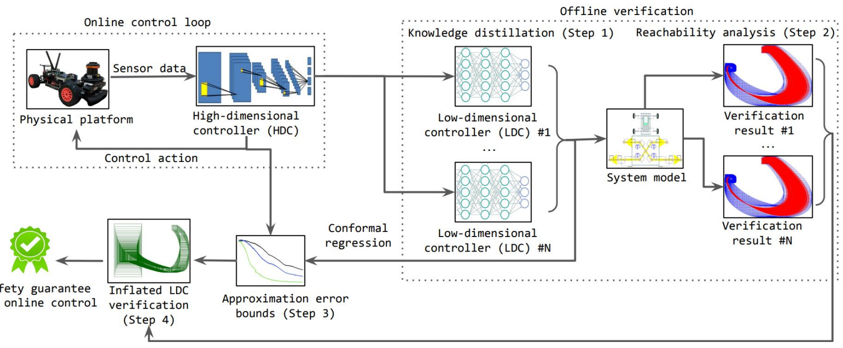
	Average Time to Verify (seconds)							Max Reachable Set Size						
Algo\M	1	2	3	4	5	6	7	1	2	3	4	5	6	7
Uniform	Failed	Failed	3,487	2,656	Failed	4,968	5,381	N/A	N/A	0.513	0.309	N/A	0.330	0.308
SA+P	-	Failed	2,051	2,851	3,902	3,396	5,737	-	N/A	0.377	0.356	0.337	0.323	0.308
SA+TD	-	Failed	**1,571**	1,898	**1,810**	**2,037**	**2,787**	-	N/A	0.378	**0.277**	**0.286**	0.275	0.233
G+P	-	Failed	1,789	2,487	4,151	4,556	3,859	-	N/A	0.377	0.334	0.289	0.264	0.238
G+TD	-	Failed	1,956	**1,736**	1,940	2,272	3,266	_	N/A	**0.353**	0.301	0.294	**0.275**	**0.197**

T. Waite, Y. Geng, T. Turnquist, I. Ruchkin, R. Ivanov. *State-Dependent Conformal Perception Bounds for Neuro-Symbolic Verification of Autonomous Systems*. ArXiv preprint, in submission.

2. Verifiable Neural Approximation of Vision Control

TEA Lab

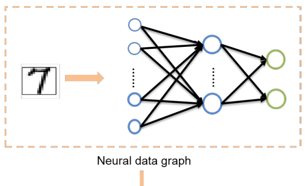
- Approximate a *high-dimensional controller (HDC)* with multiple *low-dimensional controllers (LDCs)*
- *Three statistical discrepancies* between HDC and LDCs

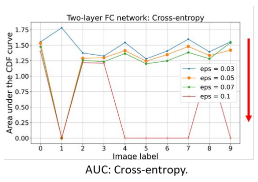


Y. Geng, J. Baldauf, S. Dutta, C. Huang, I. Ruchkin. *Bridging Dimensions: Confident Reachability for High-Dimensional Controllers*. In Proc. of FM'24.

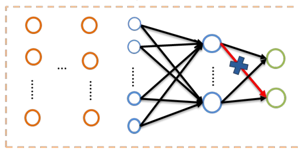
3. Analyzing Neural Network Robustness Using Graph Curvature

- A fresh look at neural network (NN) robustness from a graph theory point of view
- Use Ricci Curvature to identify bottleneck NN edges that contribute to low robustness





NN setup	$\varepsilon = 0.03$	$\varepsilon = 0.07$	$\varepsilon = 0.1$	$\varepsilon = 0.2$
[15,20], CE	1.48	1.30	1.28	N/A
[15,20], WD	1.47	1.45	1.37	1.24
[15,20], AT	1.15	1.15	1.14	1.12
[15,25,20,15], CE	2.04	1.81	1.66	N/A
[15,25,20,15], WD	1.93	1.92	1.86	N/A
[15,25,20,15], AT	2.10	2.09	2.10	2.08
CNN, CE	3.43	3.42	3.44	3.31



S. Tan, J. Sia, P. Bogdan, R. Ivanov. *Analyzing neural network robustness using graph curvature.* In 2024 International Conference on Assured Autonomy (ICAA) (pp. 110-113). IEEE.

Project info: CPS Small, $499k
06/2024-05/2027

Awards ID#: 2403615, 2403616