DeSimplex: Data Enabled Simplex for Safe Operation of Autonomous Systems

Naira Hovakimyan (PI), **Shenlong Wang** (co-PI)
University of Illinois Urbana-Champaign

Xiaofeng Wang (PI)
University of South Carolina

NSF Workshop on Safe AI

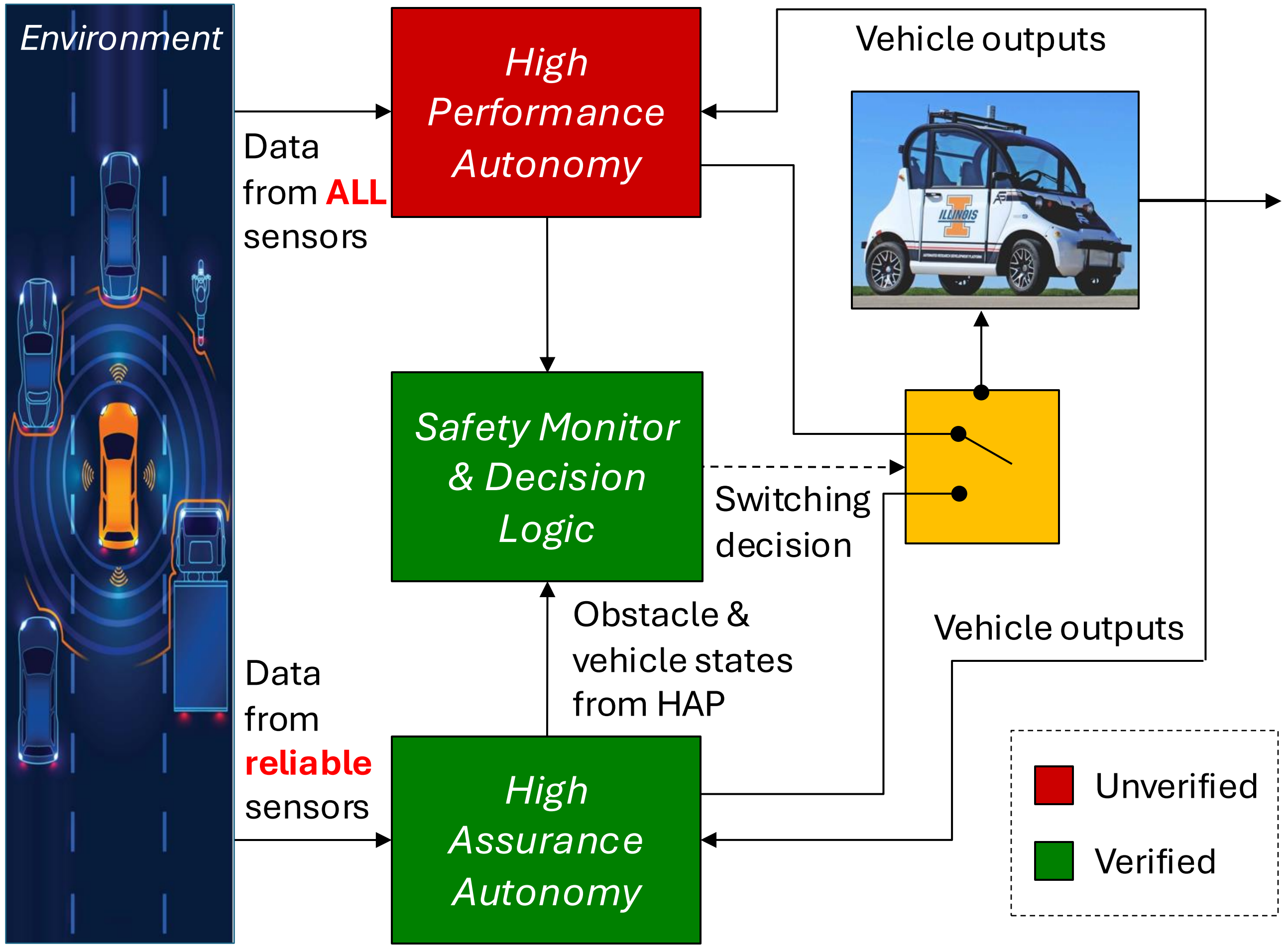
02/26/2025

DeSimplex: Data-Enabled Simplex

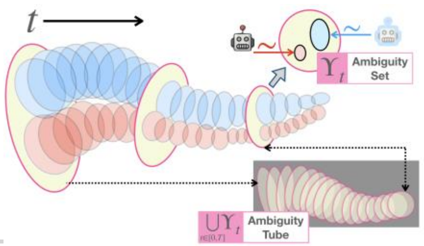
Primary Goal: Establish a verifiably safe framework that enables learning-enabled systems to adapt and perform effectively amidst unforeseen changes in system dynamics, extreme events, environmental hazards, and irregular system behaviors.

Core Principle: Transferring control authority from high-performance learning-based components to high-assurance solutions (**Simplex**) whenever safety boundaries are at risk of being violated, thus ensuring end-to-end safety.

Target Applications: Autonomous vehicles (aerial and ground) with a focus on collision avoidance and safety.



Key Components



High Performance Autonomy

HP Perception → HP Planning → HP Control

<u>Goal</u>: Optimize system performance, despite lack of verification.

- *Data collection and learning are continuously operating even if the HPA does not have the control authority.*

<u>Supporting Technologies</u>:

- HP perception: Uncertainty-aware neural semantic scene understanding
- HP planning: Optimization with chance constraints
- HP control: DiffTune and TPN

High Assurance Autonomy

HA Perception → HA Planning → HA Control

<u>Goal</u>: Deliver reliable control of the system, prioritizing safety guarantees over optimal performance.

- *Allowing for additional time to recover the learning process in HPA.*

<u>Supporting Technologies</u>:

- HA perception: [Placeholder for Shenlong]
- HA planning: Min-max optimization to handle worst-case uncertainties.
- HA control: DRAC, L1MPC

Safety monitor & Decision Logic

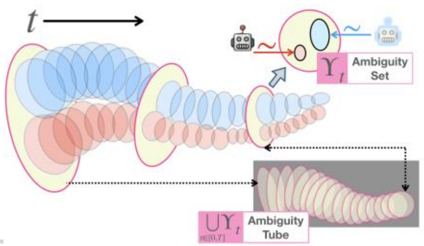
Safety Monitor → Decision Logic

<u>Safety Monitor</u>: Online assess the uncertainties and reliability of the HP learning components.

<u>Decision Logic</u>: Transition from HPA to HSA when the obstacle state and vehicle state approach the boundaries of a pre-calculated safety tube.

<u>Supporting Tech</u>: Stochastic reachability analysis, Envelope-based switching law

Key Components



High Performance Autonomy

HP Perception → HP Planning → HP Control

High Assurance Autonomy

HA Perception → HA Planning → HA Control

Safety monitor & Decision Logic

Safety Monitor → Decision Logic

Goal: Optimize system performance, despite lack of verification.

- *Data collection and learning are continuously operating even if the HPA does not have the control authority.*

Supporting Technologies:

- HP perception: Uncertainty-aware neural semantic scene understanding
- HP planning: Optimization with chance constraints
- HP control: DiffTune and TPN

Goal: Deliver reliable control of the system, prioritizing safety guarantees over optimal performance.

- *Allowing for additional time to recover the learning process in HPA.*

Supporting Technologies:

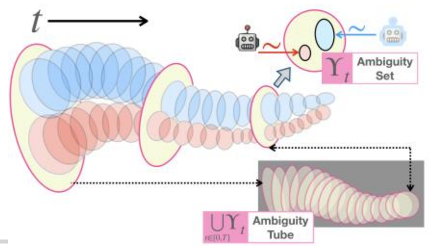
- HA perception: Uncertainty-calibrated state-estimation and mapping
- HA planning: Min-max optimization to handle worst-case uncertainties.
- HA control: DRAC, L1MPC

Safety Monitor: Online assess the uncertainties and reliability of the HP learning components.

Decision Logic: Transition from HPA to HSA when the obstacle state and vehicle state approach the boundaries of a pre-calculated safety tube.

Supporting Tech: Stochastic reachability analysis, Envelope-based switching law

Key Components



High Performance Autonomy

HP Perception → HP Planning → HP Control

High Assurance Autonomy

HA Perception → HA Planning → HA Control

Safety monitor & Decision Logic

Safety Monitor → Decision Logic

Goal: Optimize system performance, despite lack of verification.

- *Data collection and learning are continuously operating even if the HPA does not have the control authority.*

Supporting Technologies:

- HP perception: Uncertainty-aware neural semantic scene understanding
- HP planning: Optimization with chance constraints
- HP control: DiffTune and TPN

Goal: Deliver reliable control of the system, prioritizing safety guarantees over optimal performance.

- *Allowing for additional time to recover the learning process in HPA.*

Supporting Technologies:

- HA perception: Uncertainty-calibrated state-estimation and mapping
- HA planning: Min-max optimization to handle worst-case uncertainties.
- HA control: DRAC, L1-MPC

Safety Monitor: Online assess the uncertainties and reliability of the HP learning components.

Decision Logic: Transition from HPA to HAA when the obstacle state and vehicle state approach the boundaries of a pre-calculated safety tube.

Supporting Tech: Stochastic reachability analysis, Envelope-based switching law

Design Philosophy: The safety region is determined by the capacities of the high-assurance autonomy.

Publications

- Tao, Ran, et al. "DiffTune-MPC: Closed-loop learning for model predictive control." IEEE Robotics and Automation Letters (2024).
- Cheng, Sheng, et al. “Task-Parameter Nexus for Learning Task-Specific Parameters in Model-Based Control.” accepted for presentation at IEEE ICRA@40 (2024).
- Tao, Chuyuan, et al. "An Optimization-Based Planner with B-spline Parameterized Continuous-Time Reference Signals.“ accepted for presentation at 2024 IROS (2024).
- Gu, Yuliang, et al. "Proto-MPC: An Encoder-Prototype-Decoder Approach for Quadrotor Control in Challenging Winds." in Proceedings of 6th L4DC (2024).
- Xia, Hongchi, et al. "Video2Game: Real-time Interactive Realistic and Browser-Compatible Environment from a Single Video." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.
- Marques, Joao M. C., et al. "On the Overconfidence Problem in Semantic 3D Mapping." in International Conference on Robotics and Automation (ICRA). 2024.
- Jiang, Hanxiao, et al. "RoboEXP: Action-Conditioned Scene Graph via Interactive Exploration for Robotic Manipulation." in Conference on Robot Learning (CoRL). 2024.
- Tao, Ran, et al. "Guaranteed safety across autonomy architectures using safety tubes,“ in preparation.
- Kim, Minkyung , et al. "Learning Reliable Perceptual Uncertainty through Multimodal Sequential Modeling," in preparation.