CPS Safety Verification: What's Next

Sayan Mitra
mitras@illinois.edu
ECE, University of Illinois Urbana-Champaign

20th NSF CPS Program Anniversary
21 October 2026



ILLINOIS
Electrical & Computer Engineering
GRAINGER COLLEGE OF ENGINEERING

Formal verification is having its ChatGPT moment

Formal verification is having its ChatGPT moment

OpenAI's claimed Navier–Stokes blow-up proof in September and the 722 manuscripts released in October. Both came with Lean formalisations; machine-checked proofs.

Formal verification is having its ChatGPT moment

OpenAI's claimed Navier–Stokes blow-up proof in September and the 722 manuscripts released in October. Both came with Lean formalisations; machine-checked proofs.

"Math and software are readily verifiable."



All-In Podcast

Formal verification is having its ChatGPT moment

OpenAI's claimed Navier–Stokes blow-up proof in September and the 722 manuscripts released in October. Both came with Lean formalisations; machine-checked proofs.

"Math and software are readily verifiable."



All-In Podcast

Fifty years of research made "readily" possible.

Enablers of “Readily” Verifiable

1. **Some statements cannot be verified at all.** Collatz; A hybrid automaton with skewed two clocks: undecidable.

Enablers of “Readily” Verifiable

1. **Some statements cannot be verified at all.** Collatz; A hybrid automaton with skewed two clocks: undecidable.
2. **Mapping of the Decidability boundary**

Enablers of “Readily” Verifiable

1. **Some statements cannot be verified at all.** Collatz; A hybrid automaton with skewed two clocks: undecidable.
2. **Mapping of the Decidability boundary**
3. **For decidable problems, proofs are still not free.** We need a *certificate* that is checkable and implies the property: e.g., ranking functions, barrier certificates, Lyapunov functions.

Enablers of “Readily” Verifiable

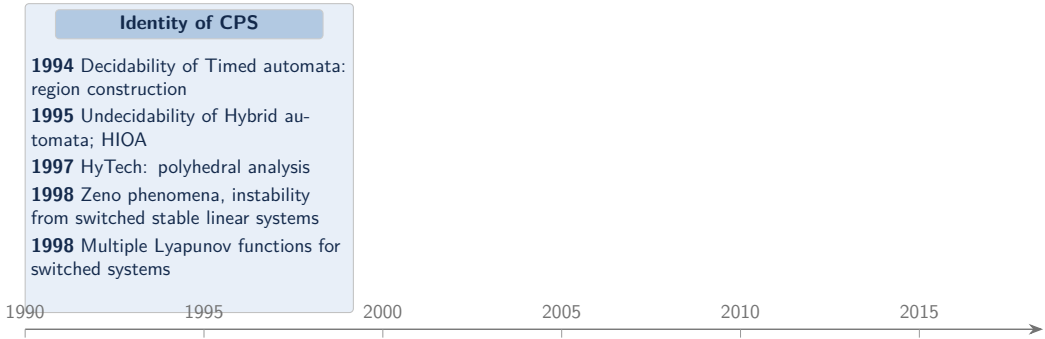
1. **Some statements cannot be verified at all.** Collatz; A hybrid automaton with skewed two clocks: undecidable.
2. **Mapping of the Decidability boundary**
3. **For decidable problems, proofs are still not free.** We need a *certificate* that is checkable and implies the property: e.g., ranking functions, barrier certificates, Lyapunov functions.
4. **Finding the certificate is a search problem.** Neural certificates, and now AI-assisted loops, Lean.

Enablers of "Readily" Verifiable

1. **Some statements cannot be verified at all.** Collatz; A hybrid automaton with skewed two clocks: undecidable.
2. **Mapping of the Decidability boundary**
3. **For decidable problems, proofs are still not free.** We need a *certificate* that is checkable and implies the property: e.g., ranking functions, barrier certificates, Lyapunov functions.
4. **Finding the certificate is a search problem.** Neural certificates, and now AI-assisted loops, Lean.

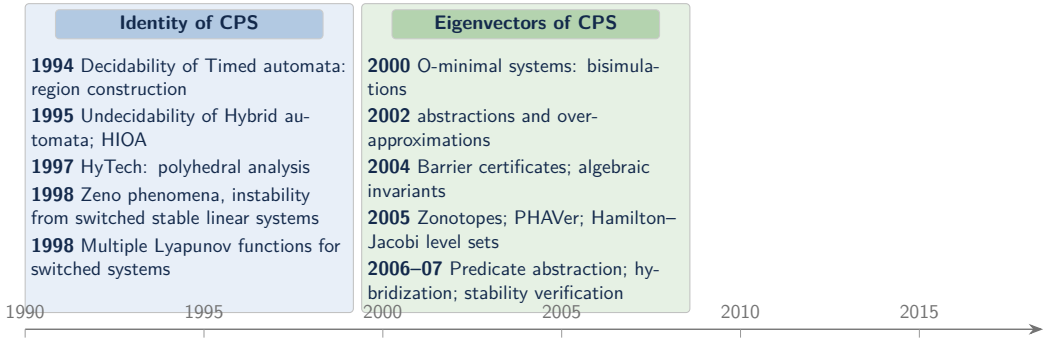
20 years of CPS research has loaded the verification gun with the first 3, now AI can empty the clips.

Three eras leading up to verifiable physical AI



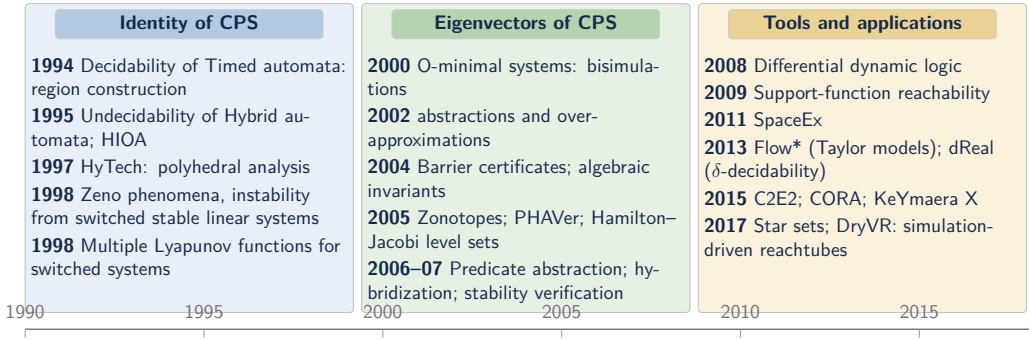
Twenty Years of Formal Verification for CPS. Habeeb P, Prabhakar, Fainekos, M, ACM TCPS, 2026.

Three eras leading up to verifiable physical AI



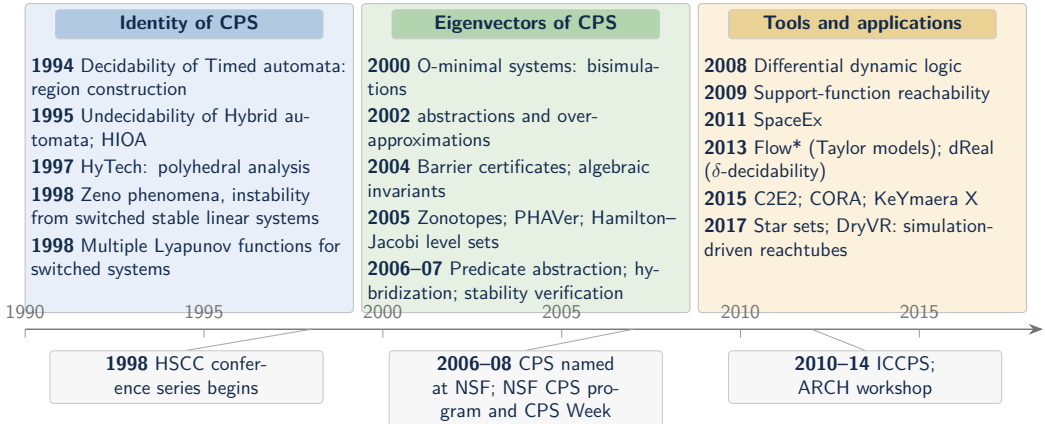
Twenty Years of Formal Verification for CPS. Habeeb P, Prabhakar, Fainekos, M, ACM TCPS, 2026.

Three eras leading up to verifiable physical AI



Twenty Years of Formal Verification for CPS. Habeeb P, Prabhakar, Fainekos, M, ACM TCPS, 2026.

Three eras leading up to verifiable physical AI

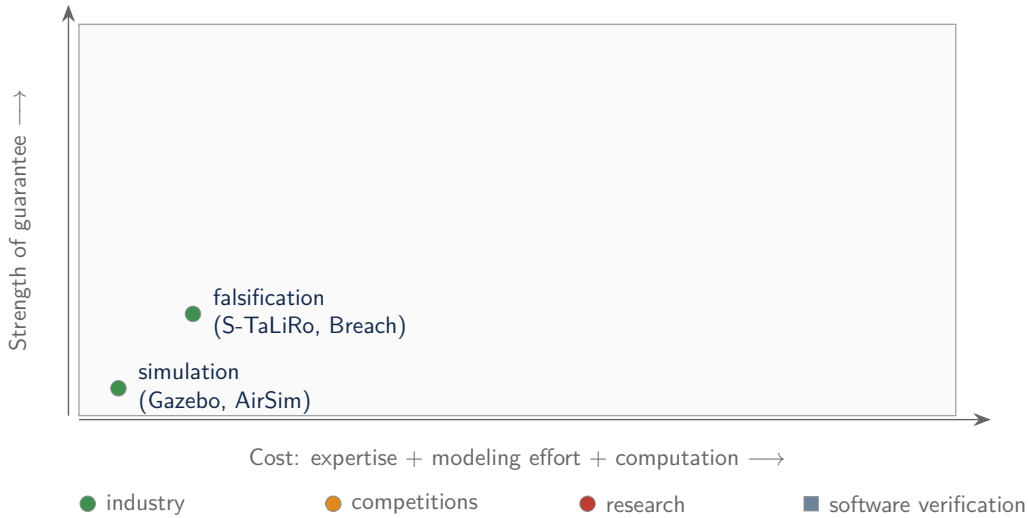


Twenty Years of Formal Verification for CPS. Habeeb P, Prabhakar, Fainekos, M, ACM TCPS, 2026.

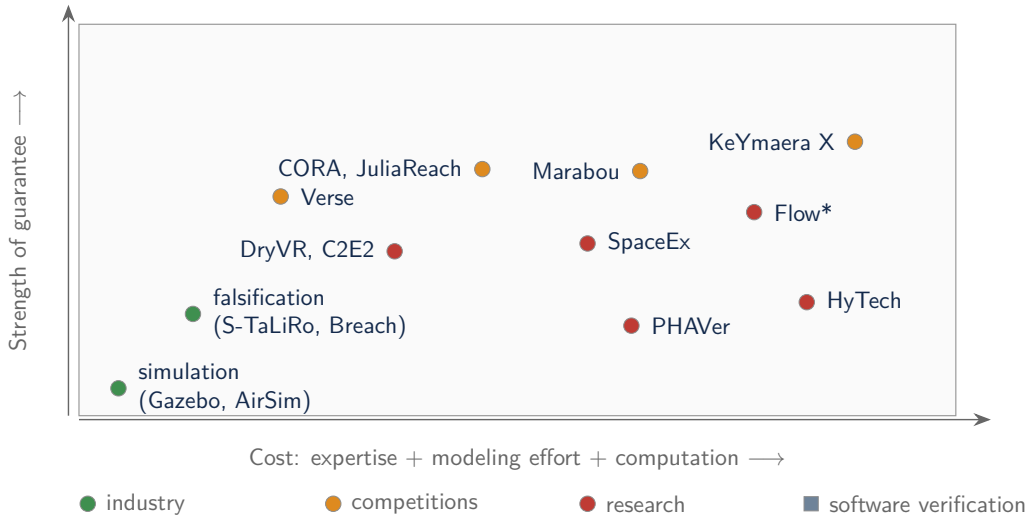
Verification cost and benefits (pre-AI)



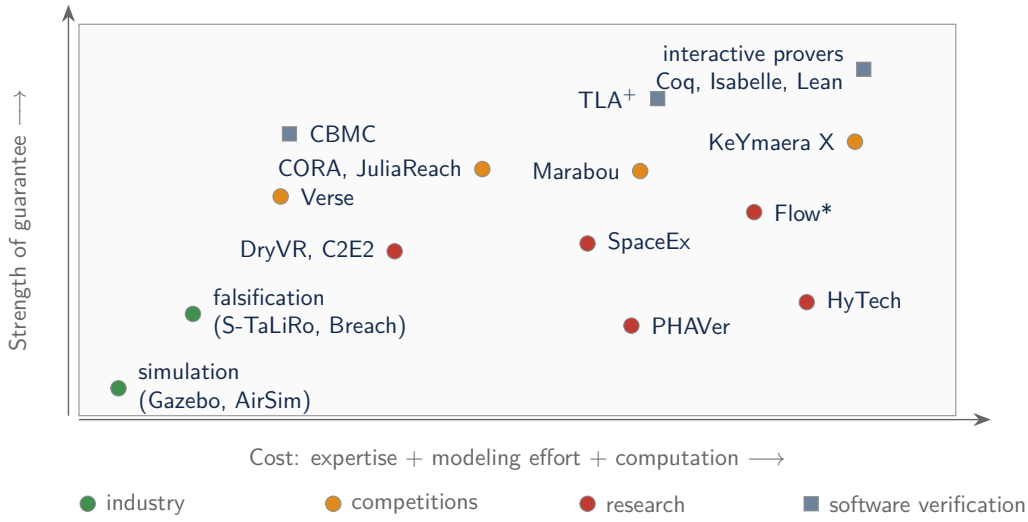
Verification cost and benefits (pre-AI)



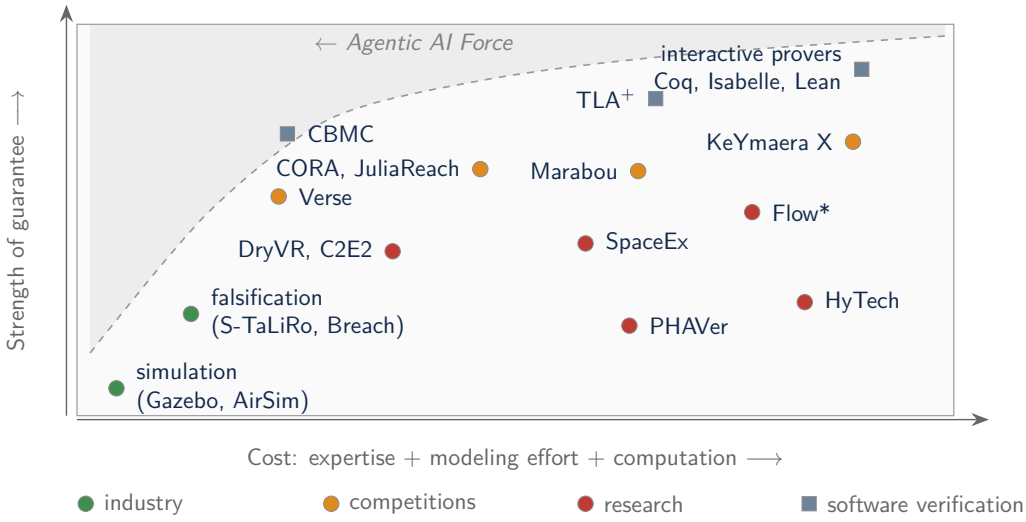
Verification cost and benefits (pre-AI)



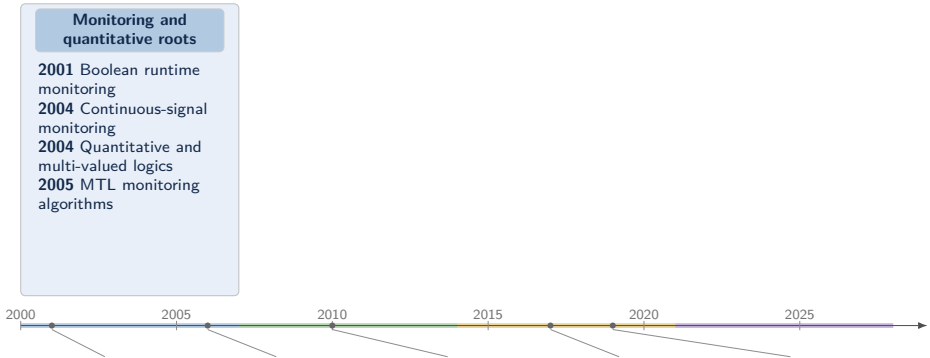
Verification cost and benefits (pre-AI)



Verification cost and benefits (pre-AI)

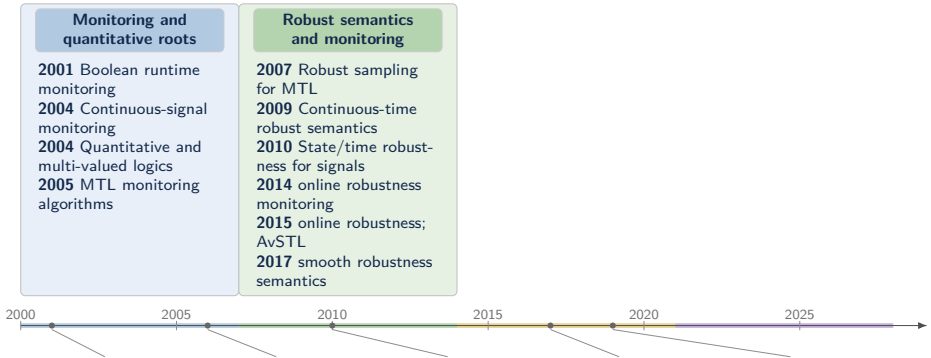


Four eras of falsification and monitoring



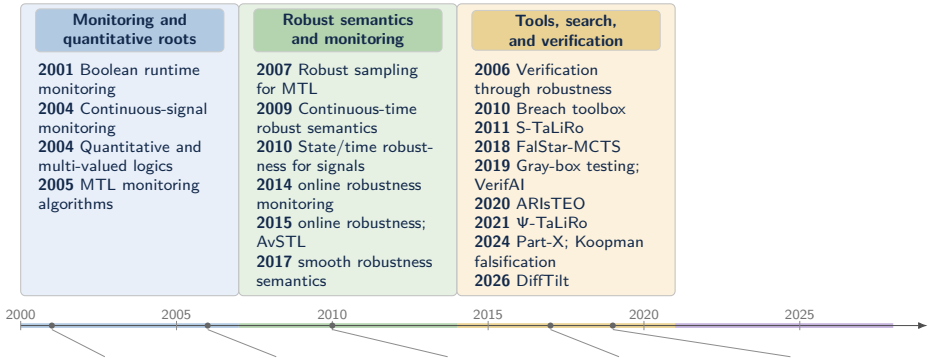
Twenty Years of Formal Verification for CPS. Habeeb P, Prabhakar, Fainekos, M, ACM TCPS, 2026.

Four eras of falsification and monitoring



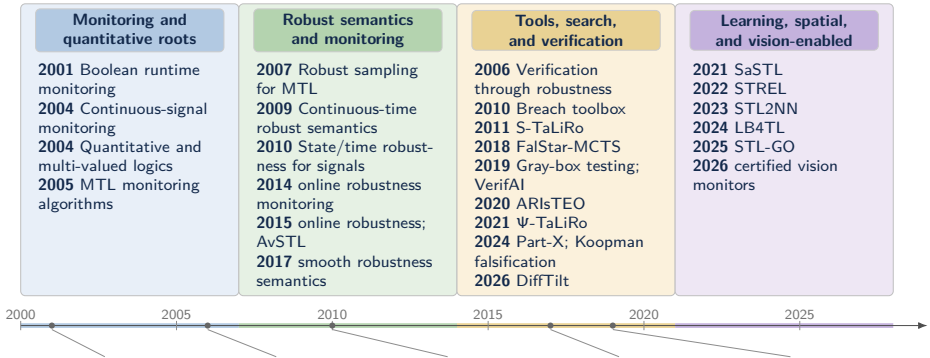
Twenty Years of Formal Verification for CPS. Habeeb P, Prabhakar, Fainekos, M, ACM TCPS, 2026.

Four eras of falsification and monitoring



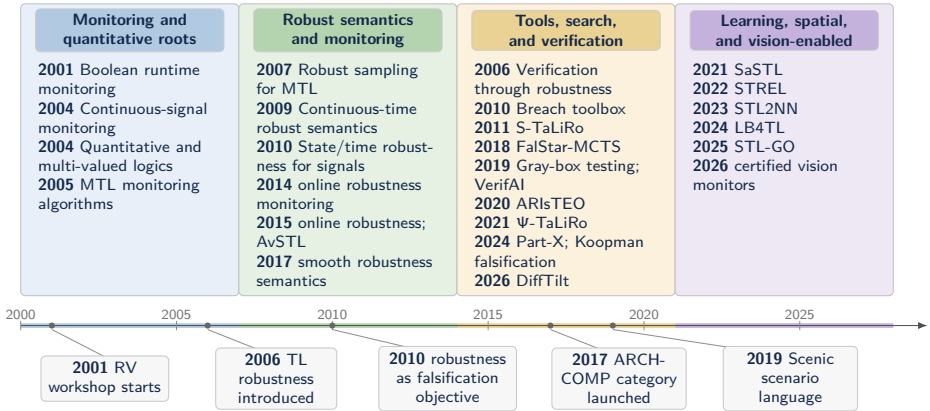
Twenty Years of Formal Verification for CPS. Habeeb P, Prabhakar, Fainekos, M, ACM TCPS, 2026.

Four eras of falsification and monitoring



Twenty Years of Formal Verification for CPS. Habeeb P, Prabhakar, Fainekos, M, ACM TCPS, 2026.

Four eras of falsification and monitoring



Twenty Years of Formal Verification for CPS. Habeeb P, Prabhakar, Fainekos, M, ACM TCPS, 2026.

Industrial adoption and remaining bottlenecks

- **Falsification, into automotive development.** The 2014 powertrain benchmark; S-TaLiRo and Breach on production Simulink models; C2E2's first verification.

Industrial adoption and remaining bottlenecks

- **Falsification, into automotive development.** The 2014 powertrain benchmark; S-TaLiRo and Breach on production Simulink models; C2E2's first verification.
- **Runtime assurance, into a standard.** Simplex codified in **ASTM F3269**, which adopts the architecture without requiring the proof.

Industrial adoption and remaining bottlenecks

- **Falsification, into automotive development.** The 2014 powertrain benchmark; S-TaLiRo and Breach on production Simulink models; C2E2's first verification.
- **Runtime assurance, into a standard.** Simplex codified in **ASTM F3269**, which adopts the architecture without requiring the proof.
- **Demonstrations (but not production).** DARPA ARCOS; ACCELERATE, where S-TaLiRo exposed a geofence false-alarm defeater; Scenic and VerifAI found edge-cases in physical vehicle; CORA, DryVR tested on vehicles

Industrial adoption and remaining bottlenecks

- **Falsification, into automotive development.** The 2014 powertrain benchmark; S-TaLiRo and Breach on production Simulink models; C2E2's first verification.
- **Runtime assurance, into a standard.** Simplex codified in **ASTM F3269**, which adopts the architecture without requiring the proof.
- **Demonstrations (but not production).** DARPA ARCOS; ACCELERATE, where S-TaLiRo exposed a geofence false-alarm defeater; Scenic and VerifAI found edge-cases in physical vehicle; CORA, DryVR tested on vehicles

FDIV vs. MCAS

Industrial adoption and remaining bottlenecks

- **Falsification, into automotive development.** The 2014 powertrain benchmark; S-TaLiRo and Breach on production Simulink models; C2E2's first verification.
- **Runtime assurance, into a standard.** Simplex codified in **ASTM F3269**, which adopts the architecture without requiring the proof.
- **Demonstrations (but not production).** DARPA ARCOS; ACCELERATE, where S-TaLiRo exposed a geofence false-alarm defeater; Scenic and VerifAI found edge-cases in physical vehicle; CORA, DryVR tested on vehicles

FDIV vs. MCAS
CPS = open; different physics, time scales and environmental assumptions

Industrial adoption and remaining bottlenecks

- **Falsification, into automotive development.** The 2014 powertrain benchmark; S-TaLiRo and Breach on production Simulink models; C2E2's first verification.
- **Runtime assurance, into a standard.** Simplex codified in **ASTM F3269**, which adopts the architecture without requiring the proof.
- **Demonstrations (but not production).** DARPA ARCOS; ACCELERATE, where S-TaLiRo exposed a geofence false-alarm defeater; Scenic and VerifAI found edge-cases in physical vehicle; CORA, DryVR tested on vehicles

FDIV vs. MCAS
CPS = open; different physics, time scales and environmental assumptions
$\Rightarrow$ Assumptions and guarantees flow cross organizational boundaries

Industrial adoption and remaining bottlenecks

- **Falsification, into automotive development.** The 2014 powertrain benchmark; S-TaLiRo and Breach on production Simulink models; C2E2's first verification.
- **Runtime assurance, into a standard.** Simplex codified in **ASTM F3269**, which adopts the architecture without requiring the proof.
- **Demonstrations (but not production).** DARPA ARCOS; ACCELERATE, where S-TaLiRo exposed a geofence false-alarm defeater; Scenic and VerifAI found edge-cases in physical vehicle; CORA, DryVR tested on vehicles

FDIV vs. MCAS
CPS = open; different physics, time scales and environmental assumptions
$\Rightarrow$ Assumptions and guarantees flow cross organizational boundaries
$\Rightarrow$ Bottlenecked by diseconomies of scale in verification.

Intellectual exports: robotics and control

- **Barrier certificates and control barrier functions.** Used as a real-time safety filter that minimally edits a nominal controller, running on quadrupeds and in adaptive cruise control.

Intellectual exports: robotics and control

- **Barrier certificates and control barrier functions.** Used as a real-time safety filter that minimally edits a nominal controller, running on quadrupeds and in adaptive cruise control.
  - **Inverted the certificate's role: static evidence became an active component.**

Intellectual exports: robotics and control

- **Barrier certificates and control barrier functions.** Used as a real-time safety filter that minimally edits a nominal controller, running on quadrupeds and in adaptive cruise control.
  - **Inverted the certificate's role: static evidence became an active component.**
- **Temporal logic became a specification interface.** Robustness turned a verdict into an objective that can be optimized against.

Intellectual exports: robotics and control

- **Barrier certificates and control barrier functions.** Used as a real-time safety filter that minimally edits a nominal controller, running on quadrupeds and in adaptive cruise control.
  - **Inverted the certificate's role: static evidence became an active component.**
- **Temporal logic became a specification interface.** Robustness turned a verdict into an objective that can be optimized against.
  - **STL as specification for synthesis, RL, MPC**

Intellectual exports: AI-enabled systems

- **Neural-network verification inherited algorithms and data-structures.** Star sets from hybrid reachability drive NNV; Taylor-model flowpipes extended into Verisig and POLAR; nonlinear arithmetic solvers in Reluplex and Marabou.

Intellectual exports: AI-enabled systems

- **Neural-network verification inherited algorithms and data-structures.** Star sets from hybrid reachability drive NNV; Taylor-model flowpipes extended into Verisig and POLAR; nonlinear arithmetic solvers in Reluplex and Marabou.
- **Verification of Physical AI.** Bound propagation through the NN controller, *observer*, and the plant. Verification ACAS Xu quantization uncovered bugs.

Learning-enabled CPS

2017 Reluplex: SMT for ReLU networks

2018–19 AI^2, DeepPoly, CROWN: abstract domains, bound propagation

2019–20 Verisig, NNV, Marabou: neural control verification

2021–22 DeepReach, NeuReach, POLAR

2023 Verse: multi-agent scenarios

Intellectual exports: AI-enabled systems

- **Neural-network verification inherited algorithms and data-structures.** Star sets from hybrid reachability drive NNV; Taylor-model flowpipes extended into Verisig and POLAR; nonlinear arithmetic solvers in Reluplex and Marabou.
- **Verification of Physical AI.** Bound propagation through the NN controller, *observer*, and the plant. Verification ACAS Xu quantization uncovered bugs.
  - Missing: Formal analysis of the observer beyond $\dot{x} = f(x, g(y)); y = h(x)$.

Learning-enabled CPS

2017 Reluplex: SMT for ReLU networks

2018–19 AI^2, DeepPoly, CROWN: abstract domains, bound propagation

2019–20 Verisig, NNV, Marabou: neural control verification

2021–22 DeepReach, NeuReach, POLAR

2023 Verse: multi-agent scenarios

Generative AI meets CPS Verification

Camera as NN. Bound the pose recovered from one image by reachability through the network.



Ladner, Shoukry and Althoff, CAV 2026

Generative AI meets CPS Verification

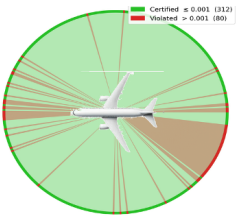
Camera as NN. Bound the pose recovered from one image by reachability through the network.



Ladner, Shoukry and Althoff, CAV 2026 ·

Abstraction of rendering. Compute set of all images that arise from a range of camera poses by approximating rendering primitives.





Li et al., 2025

Two steps towards readily verifiable & generative physical AI

1. **RSI and real-sim-to-real.** Simultaneously improving the simulator and verifying the controller, with a CEGAR-like loop IRL.

Two steps towards readily verifiable & generative physical AI

1. **RSI and real-sim-to-real.** Simultaneously improving the simulator and verifying the controller, with a CEGAR-like loop IRL.
2. **Theory of information and energy for physical AI**: Task-specific, mutual information (or topological entropy) per watt.

CPS Torch Burning Bright



HSCC and ICCPS merge into one conference and one proceedings.

Program chairs. Foundations: Taylor Johnson (Vanderbilt) · AI-enabled Autonomy: Mykel Kochenderfer (Stanford) · Systems and Applications: Jonathan Sprinkle (Vanderbilt)

Deadline **5 November 2026** · `hscc-iccps2027.taylortjohnson.com`

Thank you CPS Community, past and present

Helen Gill and David Corman

Thank you CPS Community, past and present

Helen Gill and David Corman

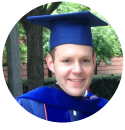
People of NSF CPS Program

Thank you CPS Community, past and present

Helen Gill and David Corman

People of NSF CPS Program

Former PhD students



Taylor Johnson
'13 · Vanderbilt



Sridhar Duggirala
'15 · UNC



Zhenqi Huang
'16 · Waymo



Chuchu Fan
'19 · MIT



Ritwika Ghosh
'20 · Bodo.ai



Hussein Sibai
'21 · WashU



Chiao Hsieh
'23 · U Kyoto



Dawei Sun
'23 · Meta



Kristina Miller
'24 · AFRL



Yangge Li
'26 · ByteDance