Simulations to Proofs through Discrepancy

for cyber-physical systems

Sayan Mitra

Electrical & Computer Engineering

University of Illinois at Urbana Champaign

UTC Institute for Advanced System Engineering

University of Connecticut

September 19th 2016

Cyberphysical systems

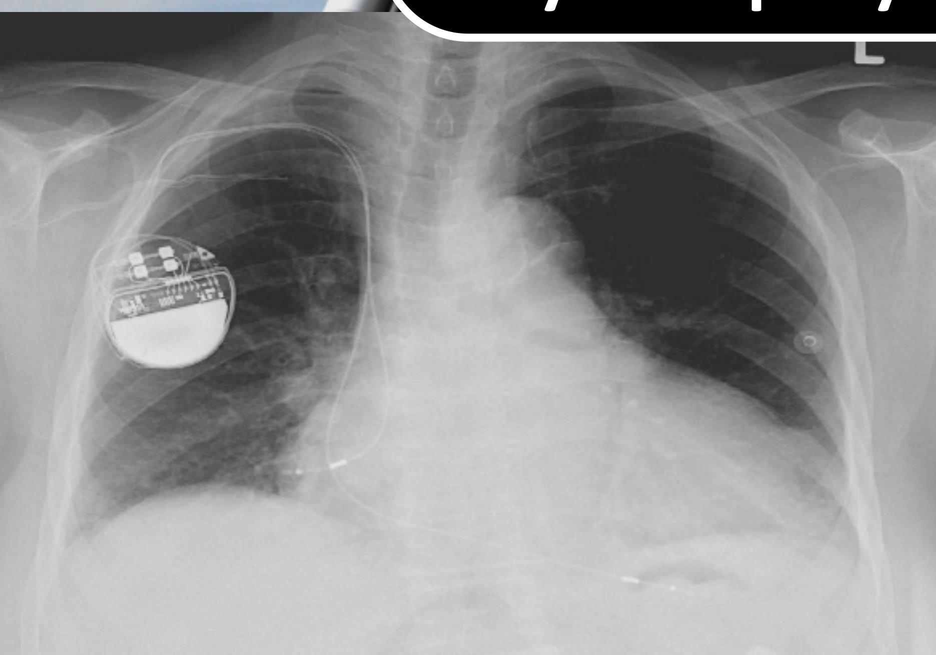


photo from [Mathews15]

photo from "Stroke and Vascular Neurology"


Number of fatal “autonomous” crashes: 1
% cost of 787 attributed to software: 50
Cars recalled in 2013: 22 M
Medical devices recalled over the decade: 2 M
% owing to software bugs: 24

“How can we design cyber-physical systems that we can bet our lives on?”

- Jeannette M. Wing

VP of Microsoft Research
Professor of Computer Science, CMU

Rigorous system engineering & Correctness properties

Invariance

Nothing "bad" ever happens

Safe separation between vehicles is maintained in adaptive cruise control

Privacy

No information leakage

Location privacy is preserved in a crowd-sourced smart navigation system

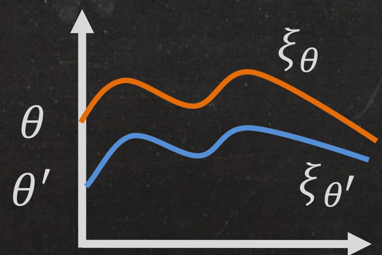
sensitivity

Quantifying sensitivity

Trajectory (or execution): evolution of states over time A model can be viewed as a mapping from a parameter d to a trajectory ξ_d . E.g., d could be initial state, private data, etc.

Sensitivity bounds the distance between trajectories as a function of the changes in parameters, that is $|\xi_d - \xi_{d'}|$

$$\dot{x} = f(x)$$



$O = ab \ldots$

$\xi_{D,O} = q_0 q_1 \ldots q_n$

$\xi_{D',O} = q_0 {q'}_1 \ldots {q_n}'$

Talk outline

Invariance

Nothing “bad” ever happens

- From Simulations to Proofs
- Tool and applications
- Compositional analysis

Privacy

conclusion

Talk outline

Invariance

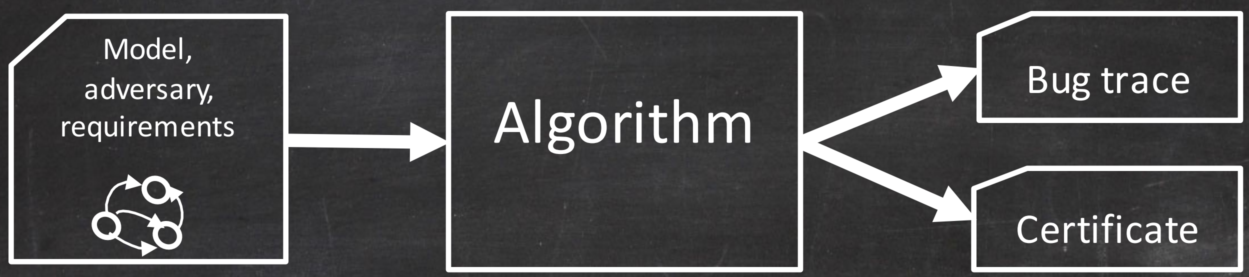
Nothing “bad” ever happens

- From Simulations to Proofs
- Tool and applications
- Compositional analysis

Privacy

conclusion

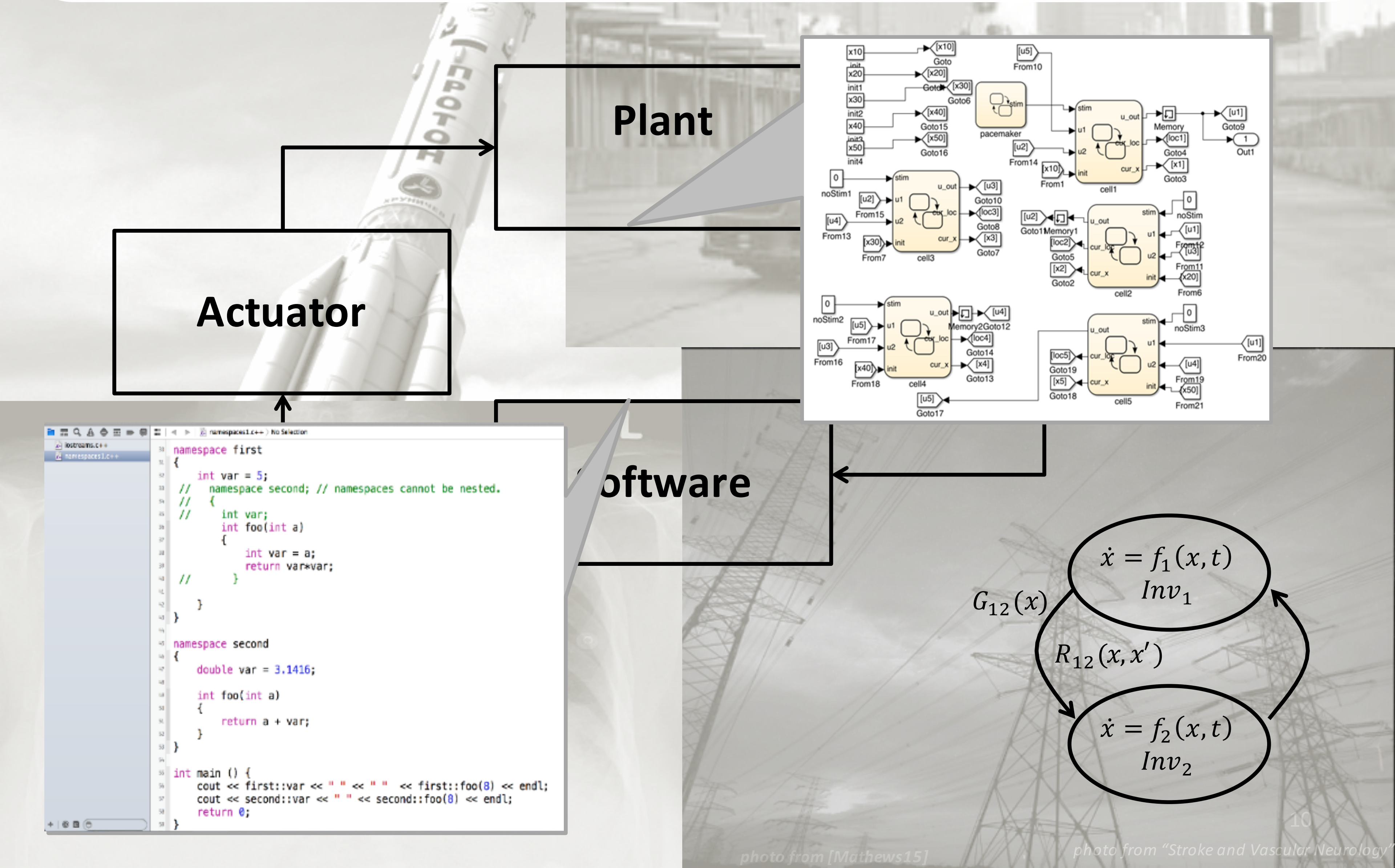
Verification problem



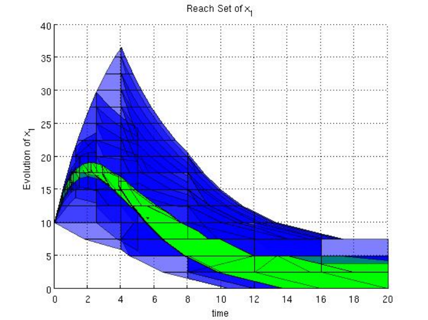
$\exists\ x_0 \in Init, u \in U, a \in A, t \in [0, T]$,
such that trajectory $\xi(x_0, a, u, t) \in U$?

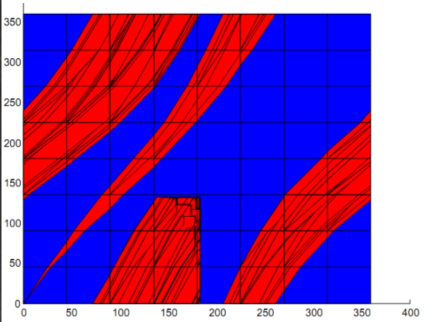
Yes (Bug-trace) / No (Safety certificate)

Hybrid automata: A model for cyberphysical systems



Brief history





Early 90's: Exactly compute unbounded time reach set

Decidable for timed automata [Alur Dill 92]

Undecidable even for rectangular dynamics [Henzinger 95]

Late 90'-00': Approximate bounded time reach set

Hamilton-Jacobi-Bellman approach [Tomlin et al. 02]

Polytopes [Henzinger 97], ellipsoids [Kurzhanski] zonotopes [Girard 05], support functions [Frehse 08]

Predicate abstraction [Alur 03], CEGAR [Clarke 03] [Mitra 13]

Today: Scalability for realistic models

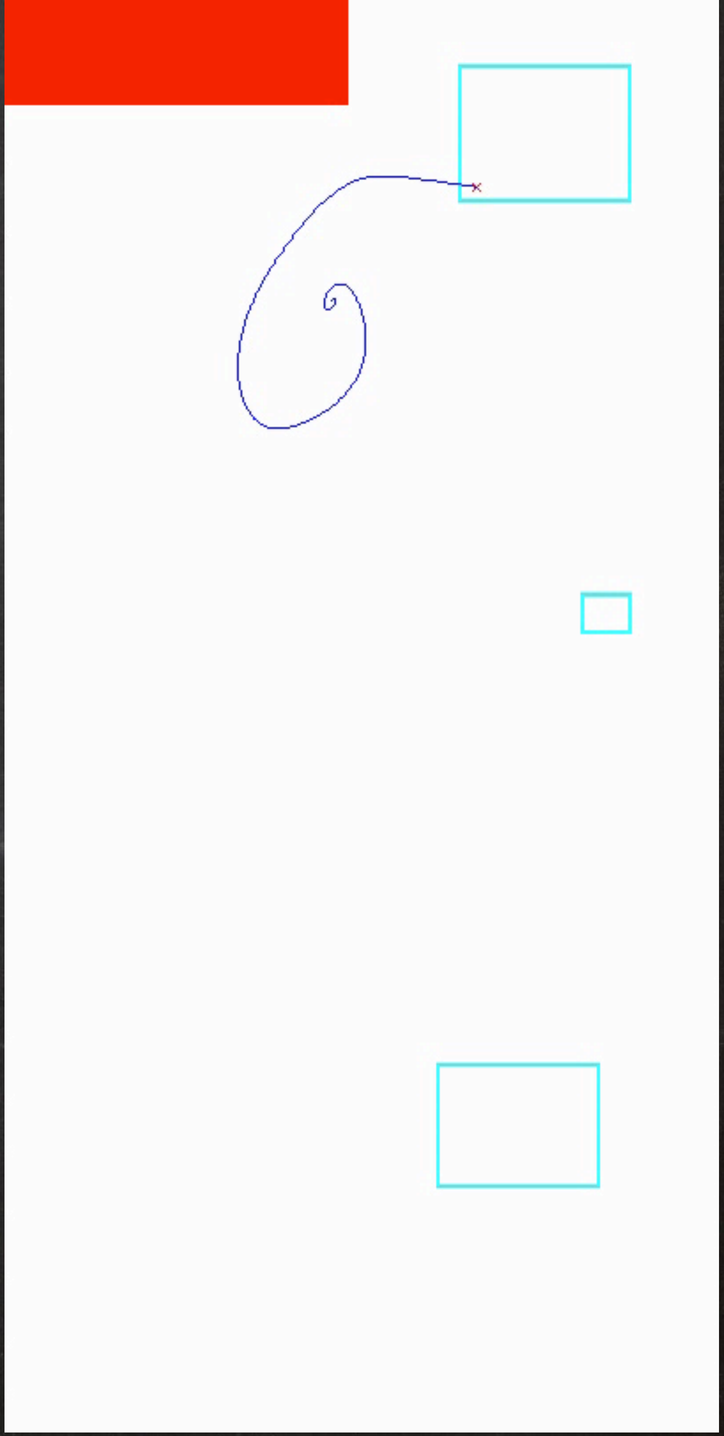
Simulation-driven algorithms [Julius 02] [Mitra 10-13][Donze 07]

Simulations to proofs

- Given start S and target U
- Compute finite cover of initial set
- Simulate from the center x_0 of each cover
- **Bloat/generalize** simulation to contain all trajectories from the cover
- Check intersection/containment with U
- Refine if needed and repeat

How to bloat or generalize simulations?

How to handle mode switches?



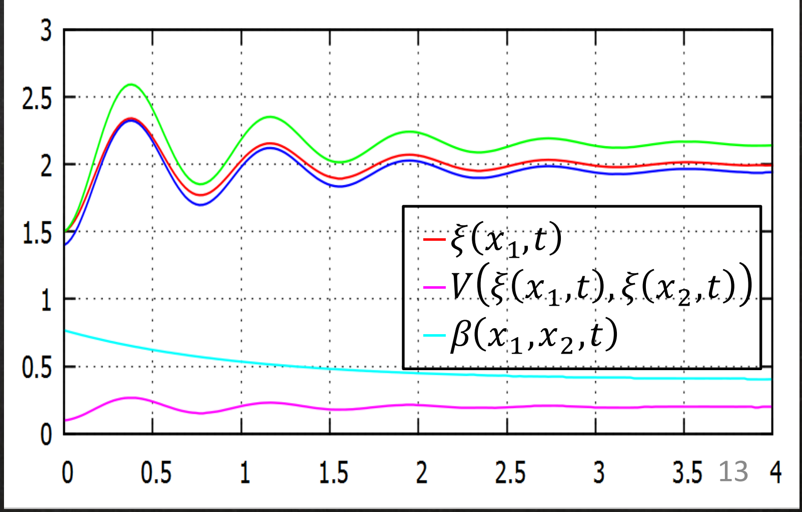
Discrepancy quantifies sensitivity

Definition. $\beta: \mathbb{R}^{2n} \times \mathbb{R}^{\geq 0} \rightarrow \mathbb{R}^{\geq 0}$ defines a discrepancy of the system if for any two states x_1 and $x_2 \in X$, For any t,

- $|\xi(x_1, t) - \xi(x_2, t)| \leq \beta(x_1, x_2, t)$ and
- $\beta \rightarrow 0$ as $x_1 \rightarrow x_2$

[EMSOFT 2013] Duggirala, Mitra & Viswanathan: Verification of annotated models from executions. EMSOFT 2013, 1-26, ACM

If L is a Lipschitz constant for *f(x,t)* then $|\xi(x_1, t) - \xi(x_2, t)| \leq e^{Lt}|x_1 - x_2|$



Guarantees for bounded invariance verification using discreapancy

Theorem. (Soundness). If Algorithm returns safe or unsafe, then A is safe or unsafe.

Definition Given HA $A = \langle V, Loc, A, D, T \rangle$, an **$\epsilon$-perturbation** of A is a new HA A' that is identical except, $\Theta' = B_\epsilon(\Theta)$, $\forall \ell \in Loc, Inv' = B_\epsilon(Inv)$ (b) a $\in$ A, $Guard_a = B_\epsilon(Guard_a)$.

A is **robustly safe** iff $\exists \epsilon > 0$, such that A' is safe for U_ϵ upto time bound T, and transition bound N. Robustly unsafe iff $\exists\ \epsilon < 0$ such that A' is safe for U_ϵ.

Theorem. (Relative Completeness) Algorithm always terminates whenever the A is either robustly safe or robustly unsafe.

Computing discrepancy functions

[ATVA 15] Fan & Mitra, Bounded verification with on-the-Fly Discrepancy Computation. ATVA 2015: 446-463, LNCS.

[HSCC 14] Huang & Mitra, Proofs from simulations and modular annotations. HSCC 2014: 183-192, ACM.

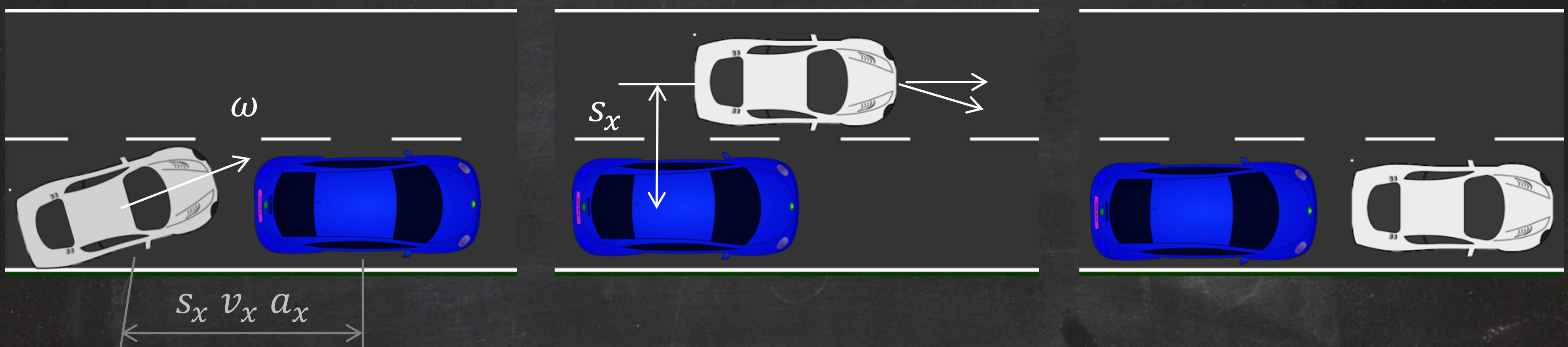
[CAV 14] Huang, Fan, Mereacre, Mitra & Kwiatkowska: Invariant Verification of Nonlinear Hybrid Automata Networks of Cardiac Cells. CAV 2014: 373-390, LNCS.

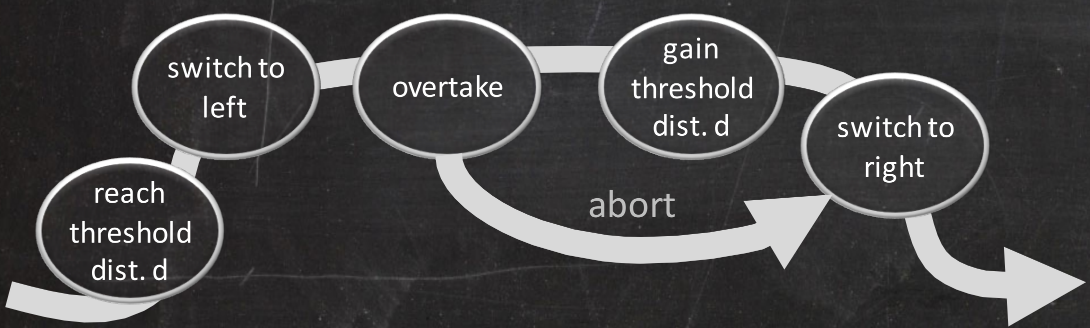
[TACAS 15] Duggirala, Mitra, Viswanathan, Potok: C2E2: A Verification Tool for Stateflow Models. TACAS 2015: 68-82, LNCS.

[CAV 15] Duggirala, Fan, Mitra, Viswanathan: Meeting a Powertrain Verification Challenge. CAV 2015, 536-543, LNCS.

[CAV 16] Fan, Qi, Mitra, Viswanathan, Duggirala: Automatic reachability analysis for nonlinear hybrid models with C2E2. CAV 2016: 531-538, LNCS.

Verification in action: an auto-pass controller





Given a controller and a safe separation requirement, we would like to check that the system is safe with respect to

a) range of initial relative positions
b) range of possible speeds
c) range road friction conditions
d) possible behaviors of “other” car
e) range of design parameters

Talk outline

Invariance

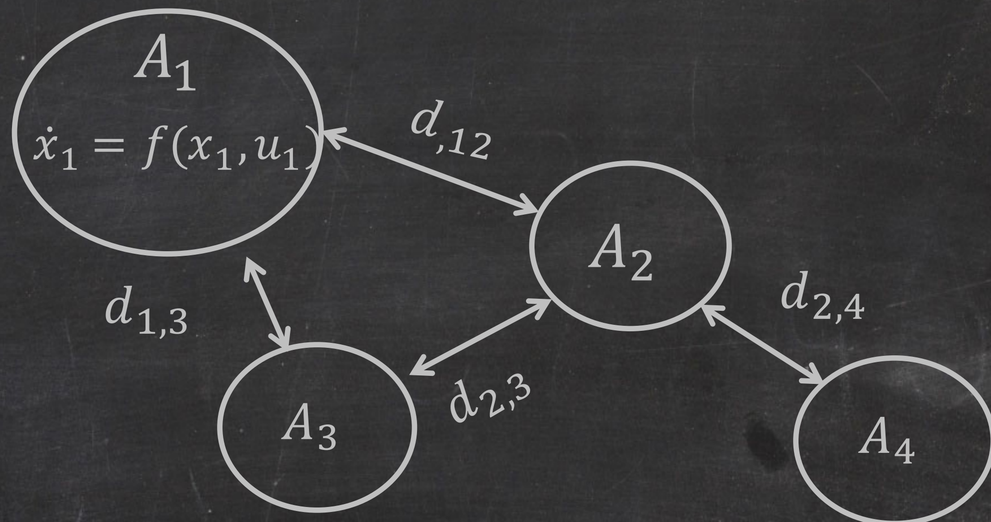
Nothing "bad" ever happens

- From Simulations to Proofs
- Tool and applications
- Compositional analysis

Privacy

conclusion

Networked cyberphysical system



- Local state vector $x_i \in \Re^n$, input $u_i \in \Re^m$
- Dynamic function f_i
- Communication possibly with delays $u_i(t) = x_j(t - d_{i,j})$

Individual dynamics

$$\dot{x}_1(t) = f_1(x_1(t), x_2(t - d_{2,1}), x_3(t - d_{3,1}))$$

Challenge: quantifying sensitivity of large networks with only node-level analysis

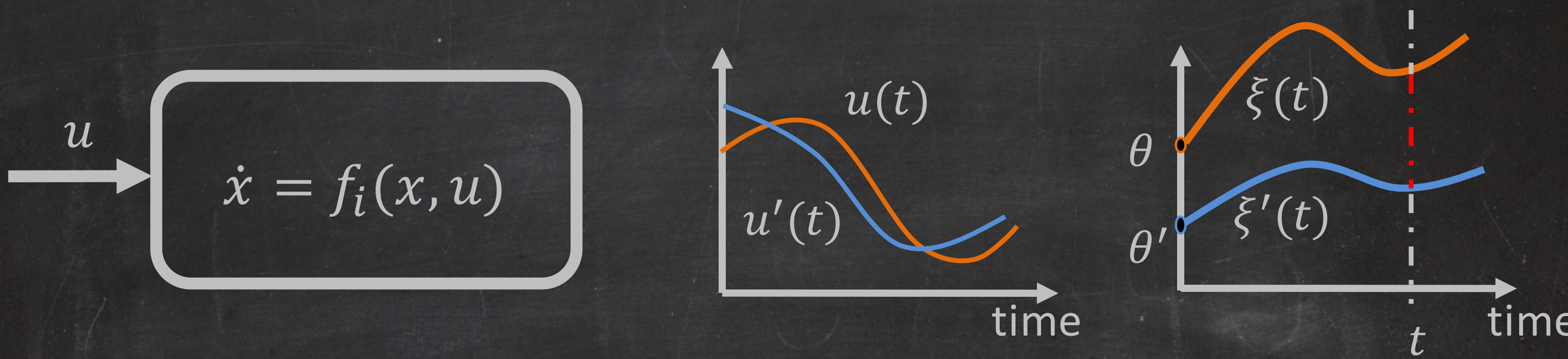
Definition. A discrepancy is a function $D: \Re_{\geq 0} \times \Re_{\geq 0} \to \Re_{\geq 0}$, such that for any $\delta \geq 0$, any pair of initial states $|\theta - \theta'| \leq \delta$, any t: $|\xi_\theta(t) - \xi_{\theta'}(t)| \leq D(\delta, t)$ and as $\delta \to 0$, $D \to 0$.

Goal: compute D only using static analysis of nodes (f_i), but not the dynamics of the entire network f.

Nodes are easier to analyze compare to the network, especially when the network has communication delays

Analysis can be applied to different topologies and delays

Input-to-State (IS) Discrepancy

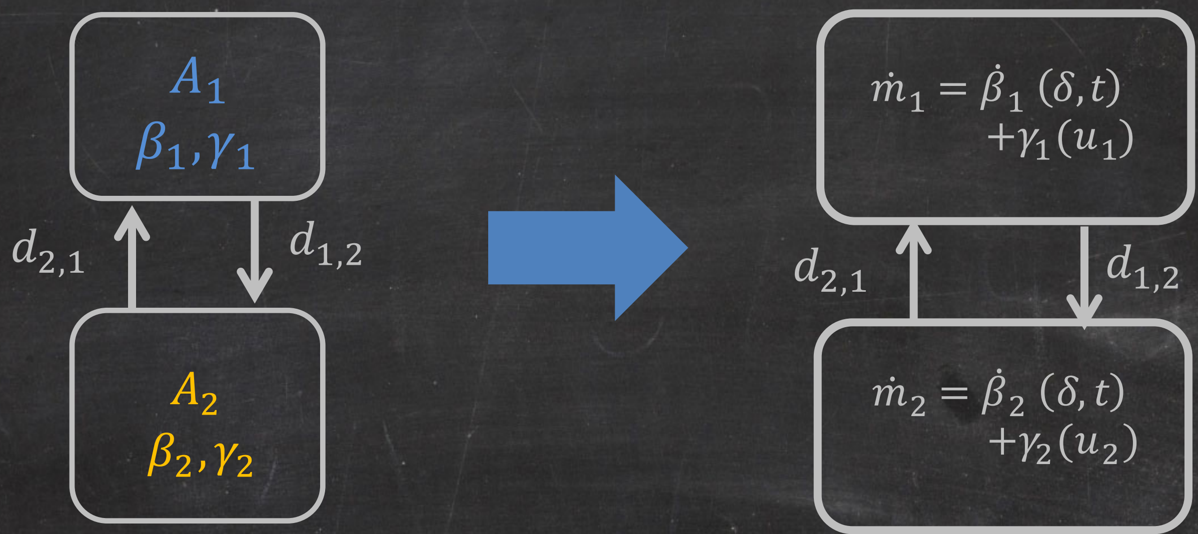


Definition. IS discrepancy of f_i is defined by two functions β and γ such that for any initial states θ, θ' and any inputs u, u',

$$|\xi(t) - \xi'(t)| \leq \beta(|\theta - \theta'|, t) + \int_0^t \gamma(|u(s) - u'(s)|)ds .$$

Also, $\beta \to 0$ as $\theta \to \theta'$, and $\gamma \to 0$ as $u \to u'$

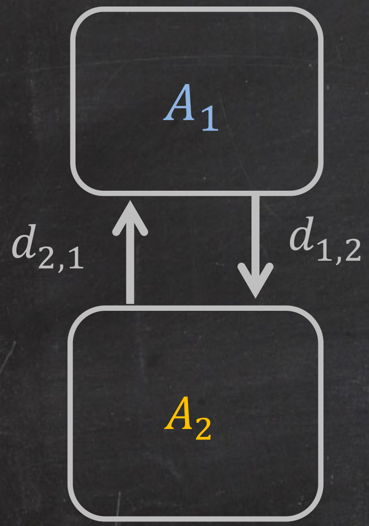
Reduced model from IS discrepancy

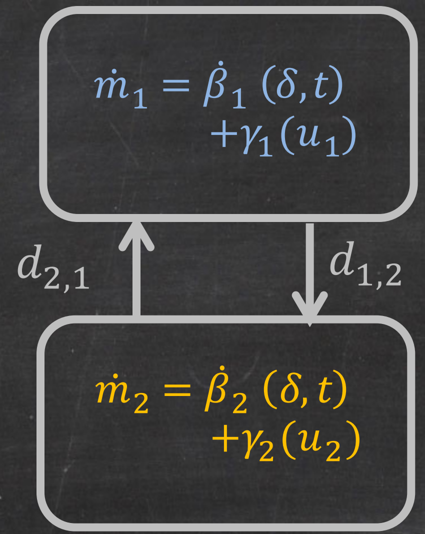


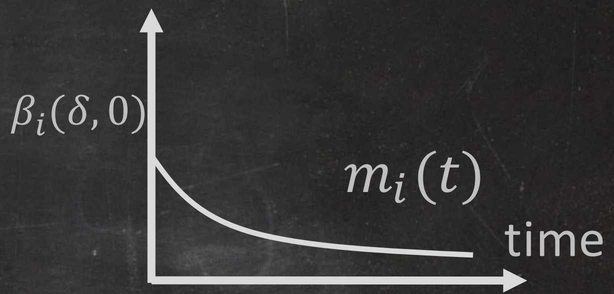
- Constructed using IS discrepancies and $\delta \geq 0$
- Identical topology and delay as the original system
- Unique initial state $[\beta_1(\delta, 0), \beta_2(\delta, 0)]$

- Easy to construct for different topologies and delays

Trajectory of reduced model gives discrepancy of original

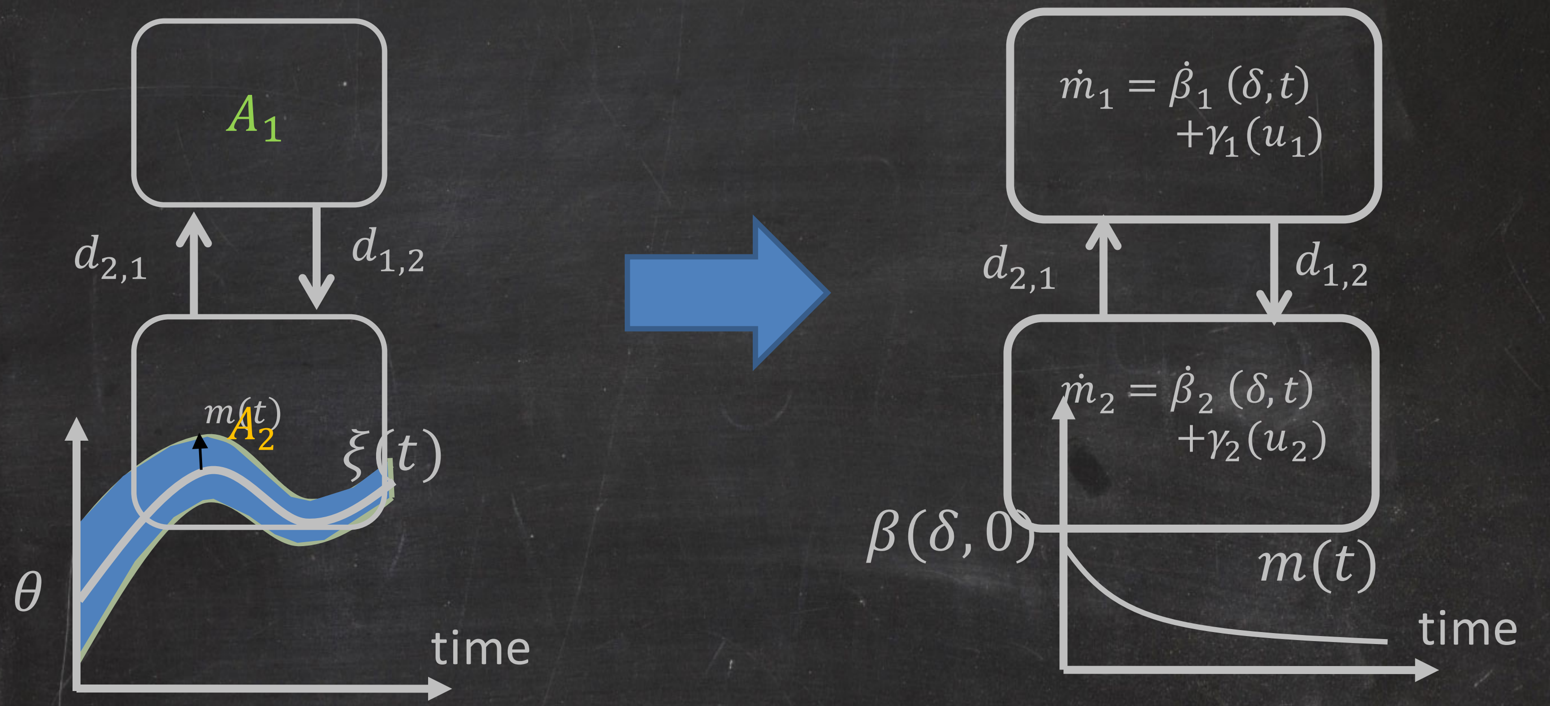






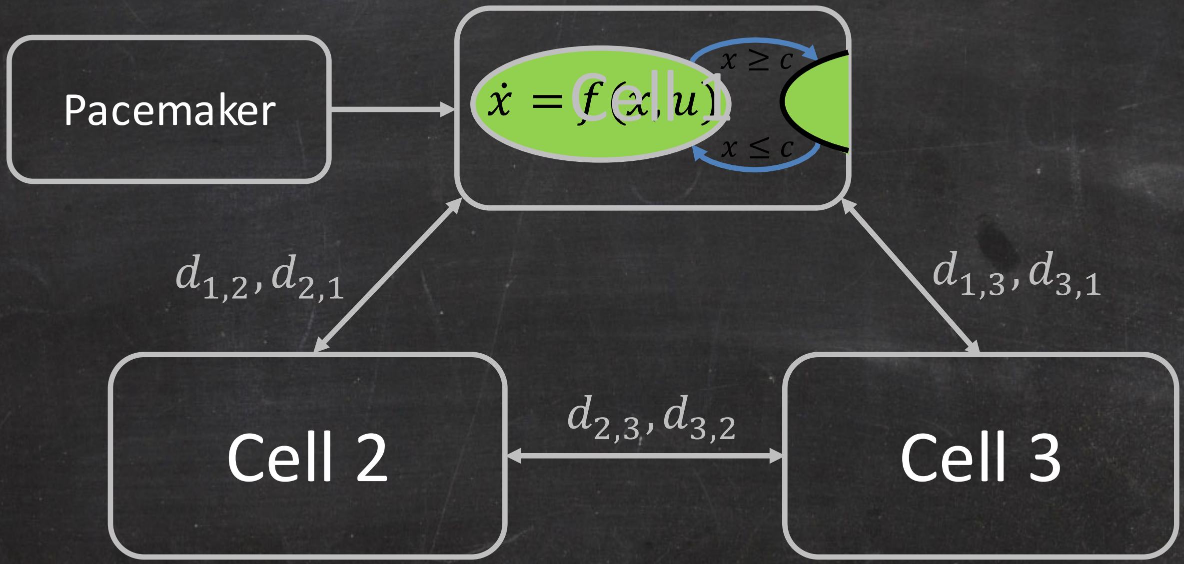
Theorem. For any pair of initial states of the network θ, θ' with $|\theta - \theta'| \leq \delta$, for all t: $\left|\xi_{\theta,i}(t) - \xi_{\theta',i}(t)\right| \leq m_i(t)$, and as $\delta \to 0$ the error bound $m(t) \to 0$.

Putting it all together gives reach set

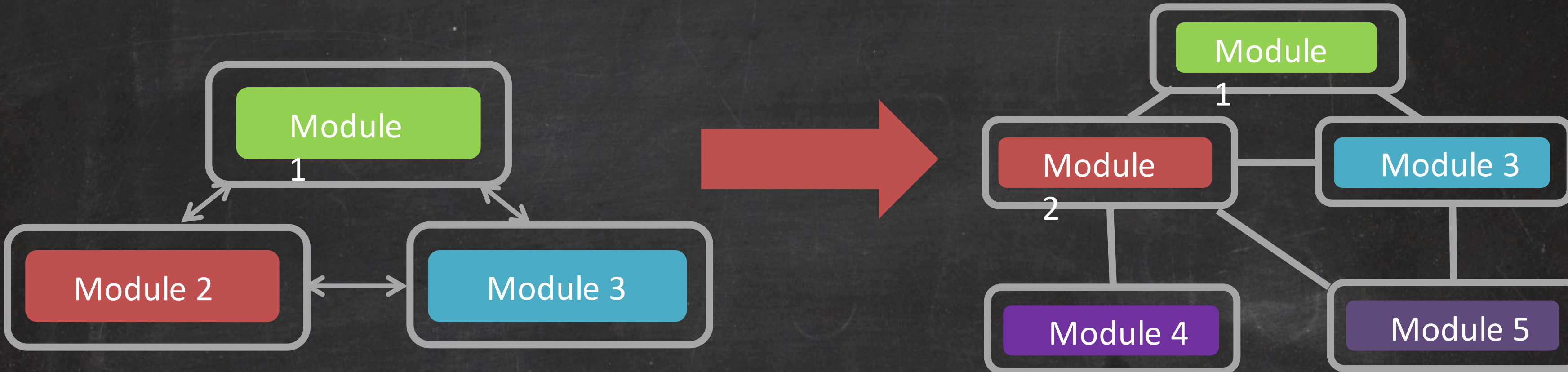


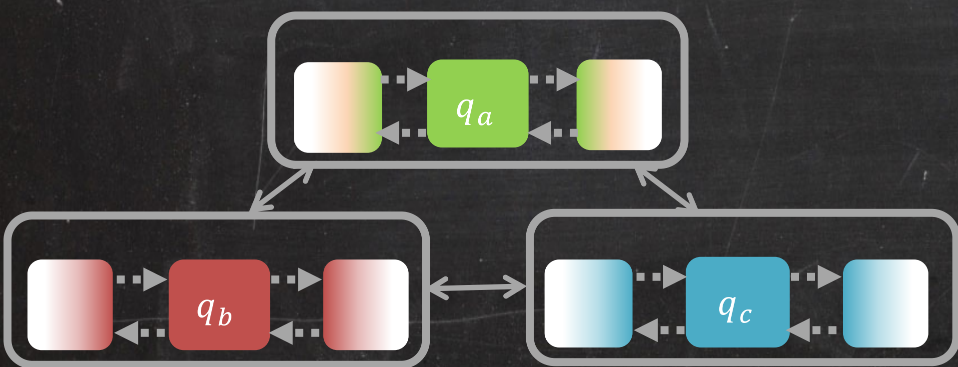
- $\xi(t) \oplus m(t)$ over-approximates reach set from the δ-neighborhood of θ
- Over-approximation can be made arbitrarily precise

Scaling to challenging benchmarks: pacemaker-heart [CAV 2014]



Exploiting modularity

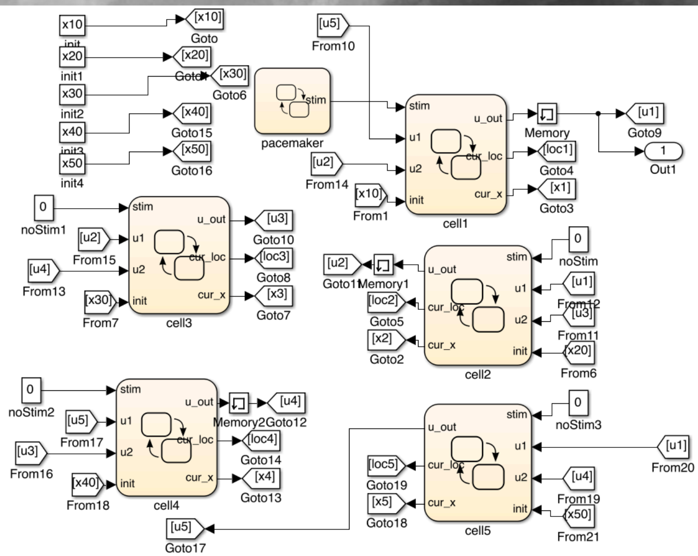


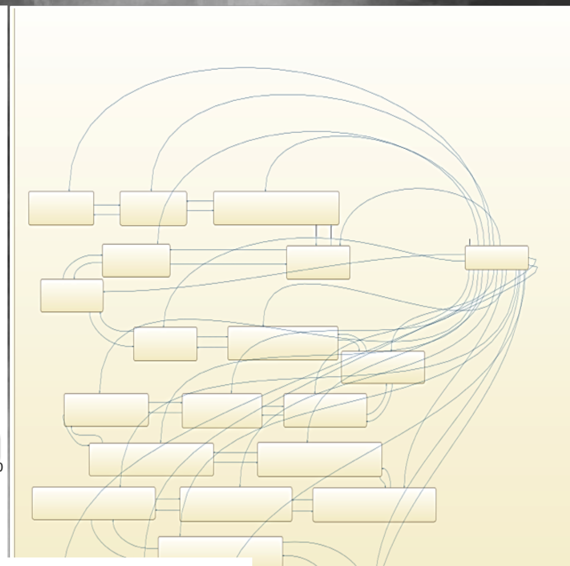


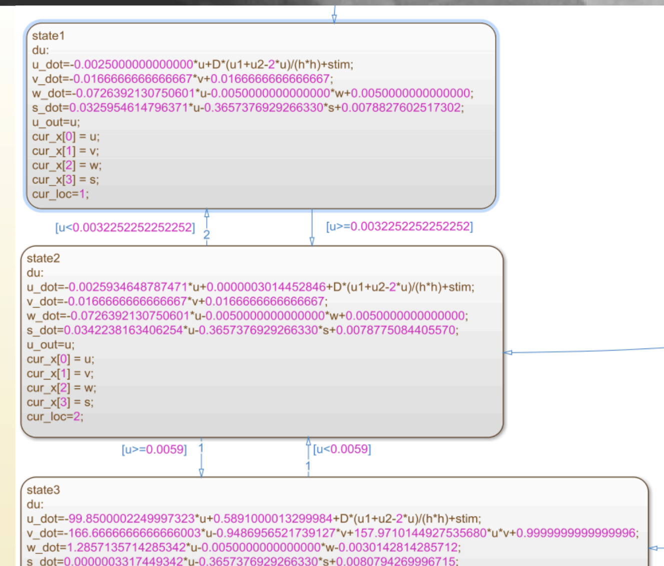
$$\dot{x}_1 = f_a(x_1, x_2, x_3)$$
$$\dot{x}_2 = f_b(x_2, x_1, x_3)$$
$$\dot{x}_3 = f_c(x_3, x_1, x_2)$$

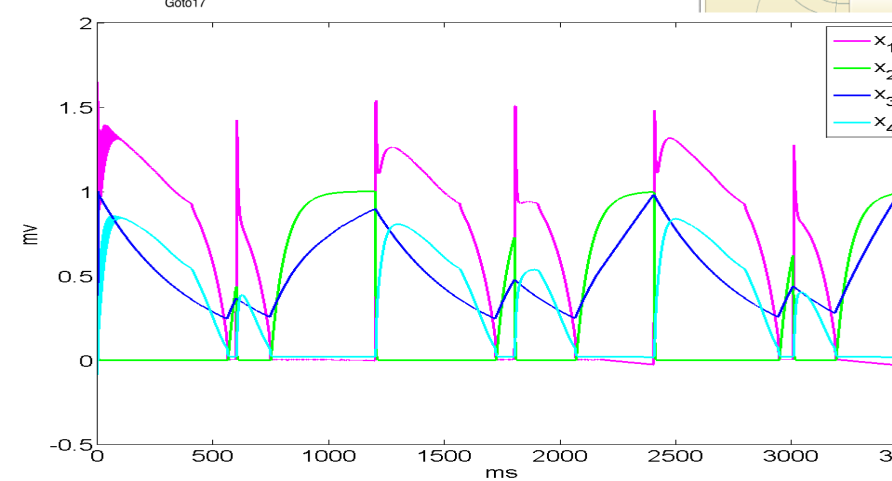
$$\times L^N$$

Pacemaker + cardiac cell network









Network	# Variables	# ODE	# Sims	Run Time (s)
8 cells (FH)	16	1	24	33
3 cells	12	2.4×10^{4}	16	105
5 cells	20	2.1×10^{7}	170	945
8 cells	32	5.0×10^{10}	73	2377
3 cells delay	6	1	16	22
8 cells delay	16	1	24	52

Talk outline

Invariance

Nothing “bad” ever happens

- From Simulations to Proofs
- Tool and applications
- Compositional analysis

Privacy

No information leakage

- Privacy in control systems
- Sensitivity to private control
- Cost of privacy

conclusion

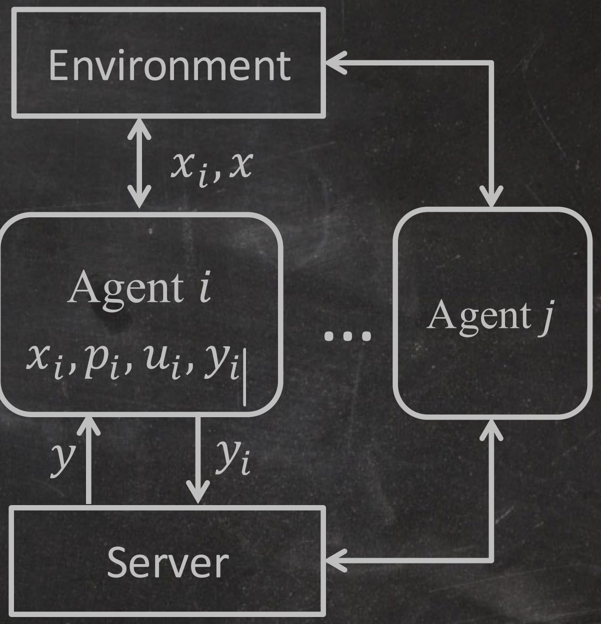
Routing delay vs. Location privacy

Publications on privacy

- [WPES12]: Z. Huang, S. Mitra and G. Dullerud, Differentially Private Iterative Synchronous Consensus.
- [IEEE Trans. CNS]: Z. Huang, Y. Wang, S. Mitra and G. Dullerud, On the Cost of Differential Privacy in Distributed Control Systems
- [CDC14]: Y. Wang, Z. Huang, S. Mitra and G. Dullerud, Entropy-minimizing Mechanism for Differential Privacy of Discrete-time Linear Feedback Systems.
- [ICDCN15]: Z. Huang, S. Mitra and N. Vaidya, Differentially Private Distributed Optimization.

Network control with randomized communication

- N agents evolve for time horizon T
- State (position) x_i
  - Affected by the environment (congestion)
  - Trajectory: $\xi = \{x(t)\}_{t \in [T]}$
- Private data (waypoints) p_i
  - Data set $D = \{p_i\}_{i \in [N]}$
- Noisy report y_i
  - $y_i = x_i + noise$
  - Observation sequence $O = \{y(t)\}_{t \in [T]} \in Obs$
- Control decision u_i computed using y, x_i, p_i

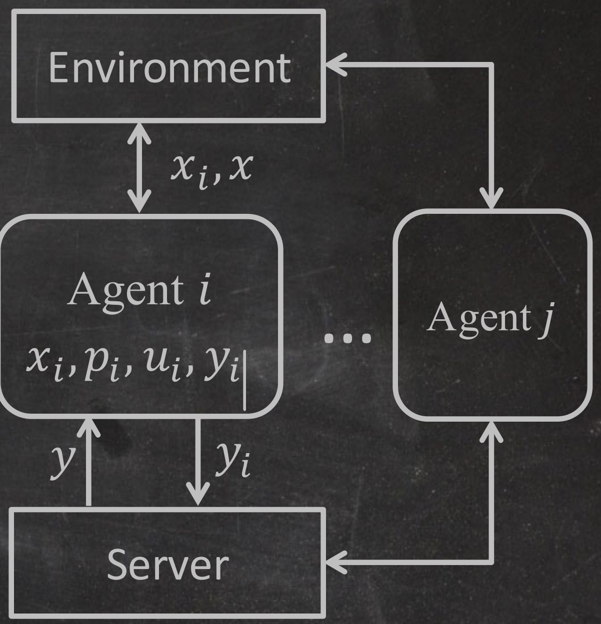


Problem: design noise mechanism for privacy

$y_i = x_i + n_i$ (random noise)

$u_i = g(x_i, y, p_i)$

$x_i^+ = f(x_i, x, u_i)$



Proposition. Fixing a data set $D = \{p_i\}_{i \in [N]}$ and an observation sequence $O = \{y(t)\}_{t \in [T]}$ uniquely determines a trajectory, denoted $\xi_{D,O}$.

Differential privacy [Dwork06]

Definition. Data sets D and D' are adjacent if D and D' differ only in agent *i*'s data, and $|p_i - p_i'| \leq \delta$ for some $\delta > 0$.

Definition. The system is ϵ-differentially private with $\epsilon > 0$, if for any adjacent D, D' and all subset of observations $S \subseteq Obs, \Pr[O_D \in S] \leq e^{\epsilon} \Pr[O_{D'} \in S]$

$\epsilon \downarrow$, privacy $\uparrow$; $\epsilon = 0$, no communication

$\epsilon \rightarrow \infty$, no privacy

Sensitivity with respect to private data

Definition. Sensitivity is a function S satisfies: for any time $t = 1,2, \ldots T$, for any observation $O \in Obs$, for any $adj(D, D')$, for any agent i:

$$| \xi_{D,O}(t) - \xi_{D',O}(t)|_1 \leq S(t)$$

- S(t) depends on dynamics f and control g
- For linear f and g, S(t) can be found analytically; general systems we use techniques from verification

Laplace Mechanism for distributed control

Theorem. The following distributed control system is ϵ-differentially private up to time T if at each time t, each agent adds an vector of independent Laplace noise $Lap(\frac{S(t)T}{\epsilon})$ to its actual state: $y_i(t) = x_i(t) + Lap\left(\frac{S(t)T}{\epsilon}\right)$, where

$Lap(\lambda)$ has the pdf $f(x) = \frac{1}{2\lambda} e^{-\frac{|x|_1}{\lambda}}$

Time horizon↑, privacy level ↑, sensitivity ↑ ⇒ noise ↑

Cost of Privacy

- Average Cost: $Cost_D = \sum_{t=0}^{T} \mathbf{E}|x_i(t) - p_i(t)|^2$
- Baseline cost $\overline{Cost_D}$: the cost when $y_i(t) = x_i(t)$
- The Cost of Privacy of a DP mechanism M is:

$$\mathbf{CoP} = \sup_{D} \mathbf{E}[Cost_D - \overline{Cost_D}]$$

Theorem. For stable system $\text{CoP} \sim O(\frac{T^3}{N^2 \epsilon^2})$, otherwise grows exponential in T

Summary of privacy work

- We introduced a notion of privacy for systems with feedback, developed privacy-preserving Laplace mechanism for dynamical systems using sensitivity

- Framework for analyzing cost of privacy
  – Linear stable dynamics $O(\frac{T^3}{N^2 \epsilon^2})$

Talk outline

Invariance

Nothing “bad” ever happens

- From Simulations to Proofs
- Tool and applications
- Compositional analysis

Privacy

No information leakage

- Privacy in control systems
- Sensitivity to private control
- Cost of privacy

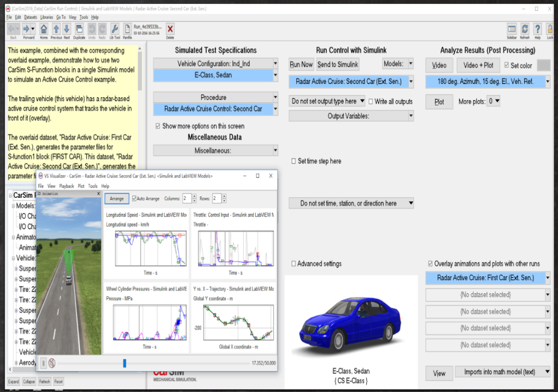
conclusion

Future: Formal methods ⇔ Data

Analysis:

Simulation data + discrepancy => algorithms => sound &complete invariance verification

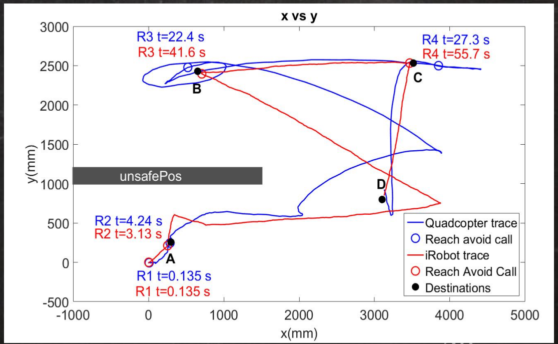
- Learn discrepancy from simulations (CarSim)
- Entropy and minimum data-rate needed for state estimation and model detection (HSCC 16)



Synthesis:

Sensitivity => privacy-preserving algorithms => trade-off between privacy and performance

- Controller synthesis with system ID [CDC15]
- Distributed optimization, learning, and fairness



Collaborators in work presented

Zhenqi Huang

Chuchu Fan

Mahesh Viswanathan

Funding support from

National Science Foundation

Science of Security Lablet of National Security Agency

Air Force Office of Scientific Research

Q & A