FORCES: Modeling Cyber Human Systems

Shankar Sastry, PI, Berkeley

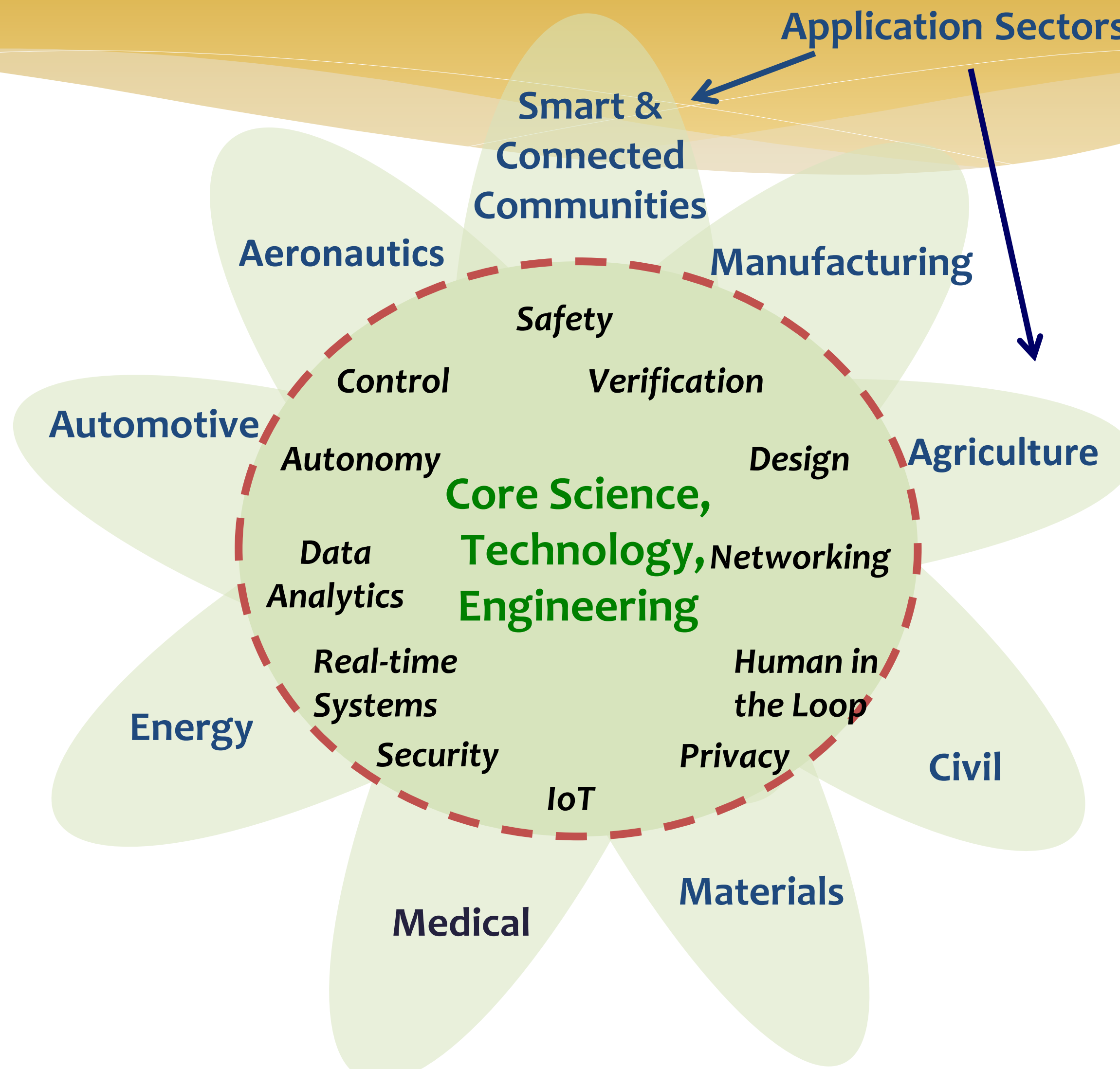
Joint work with Dorsa Sadigh, Claire Tomlin, Sanjit Seshia and Anca Dragan.

PI Meeting, January 2017

FORCES & NSF CPS Research Model

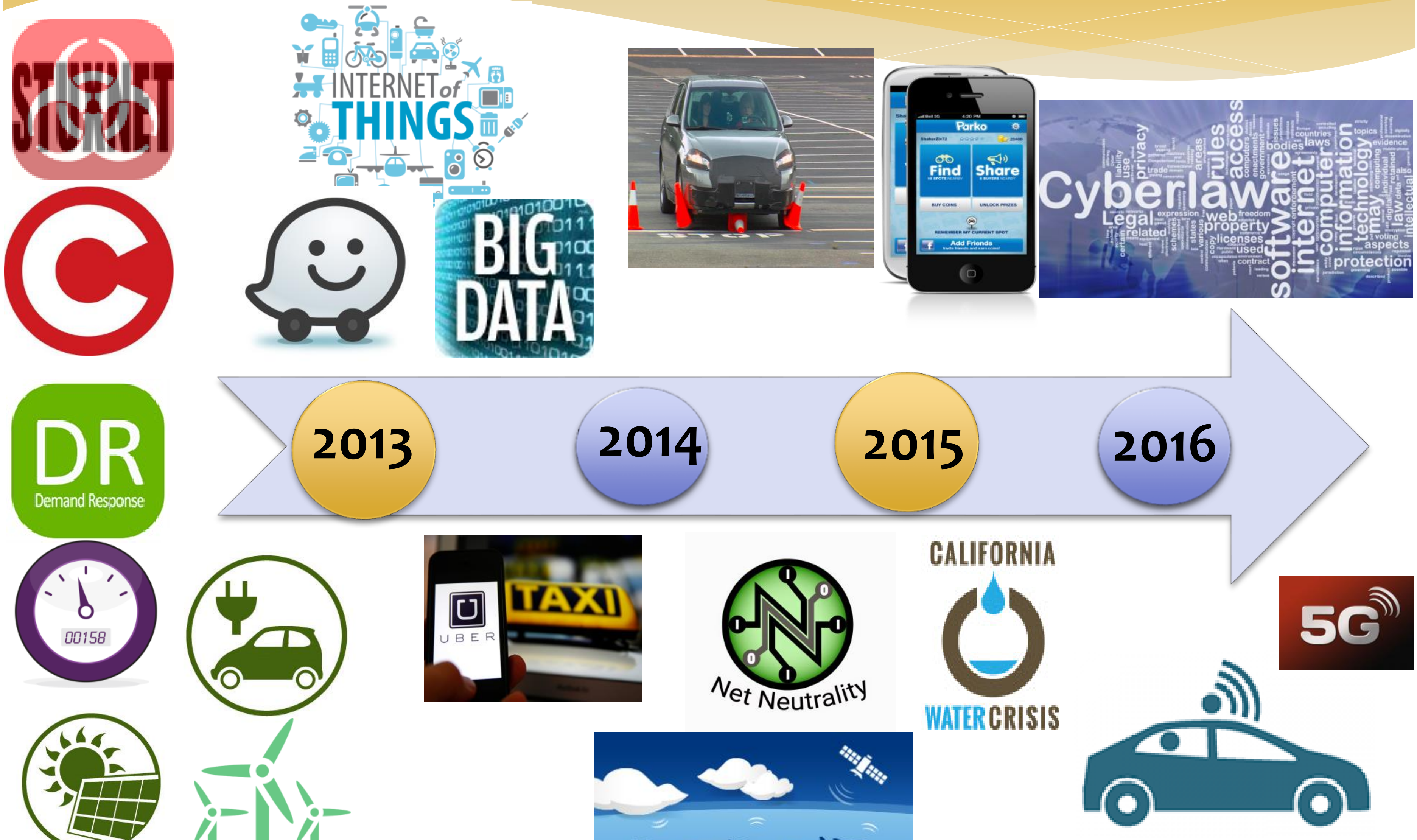
* FORCES Domains
  * Energy
  * Ground transportation
  * Air transportation
  * Smart cities

* FORCES Science
  * Robust control
  * Reliability & safety
  * Human-CPS
  * Security & privacy

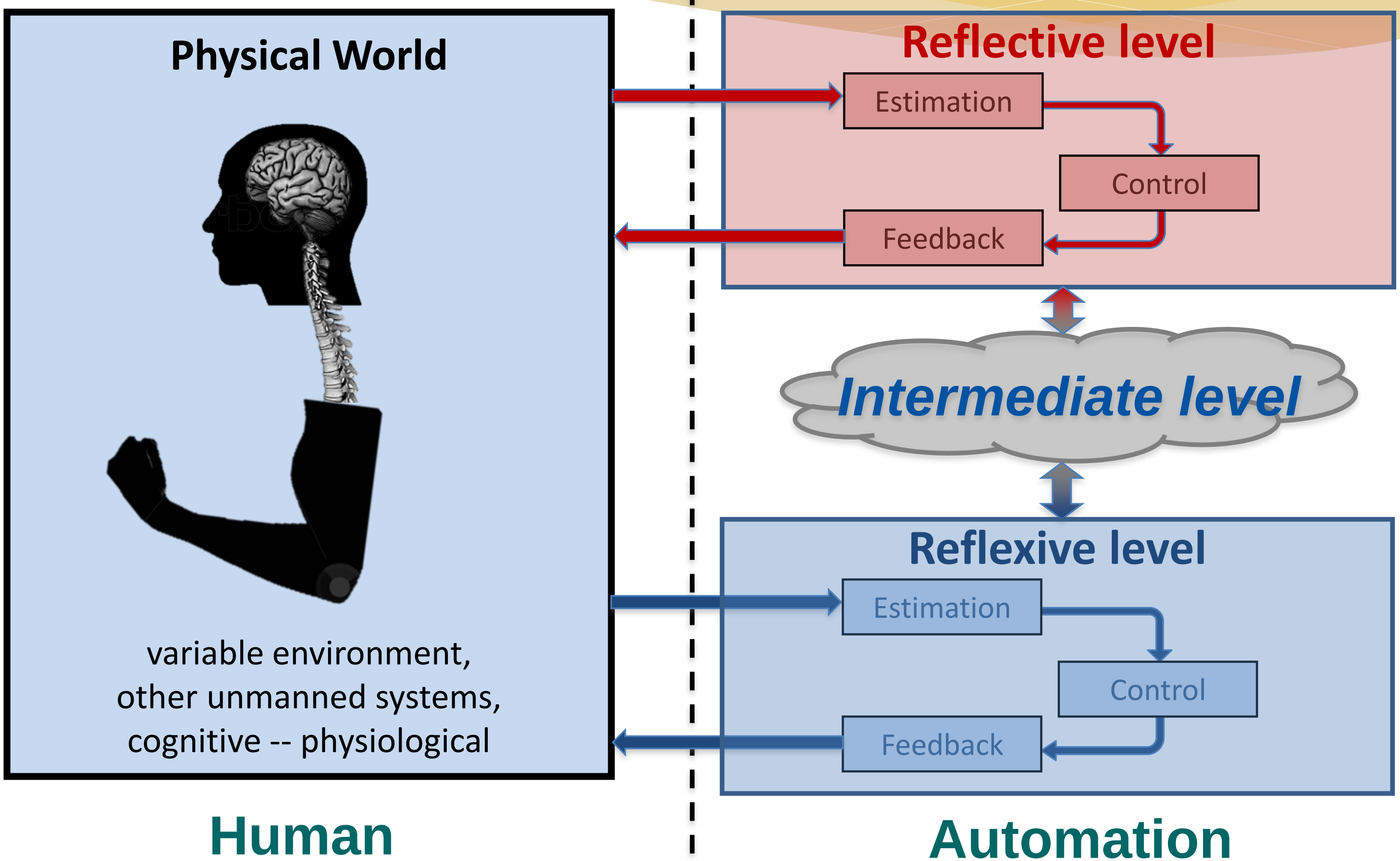


FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

CPS-FORCES continue to be on [**rapid**] ascent!



Reflexive Interacts with Reflective

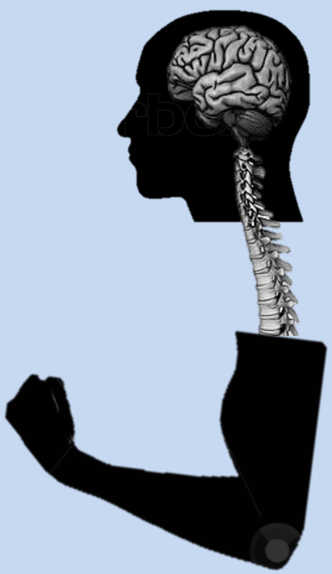


Interaction as a Stochastic Hybrid System

Physical World

safety-critical dynamical system, control by human is stochastic

$$x_{k+1} = f_k(q_k, x_k, u_k)$$



Estimation

human variability, cognitive state

$$\widehat{q}_k$$

Control

assumption mining enables safety verification

$$\exists \psi : \psi \rightarrow \varphi$$

Feedback

alter system dynamics, task specification

$$\psi \rightarrow \varphi$$

Human | **Automation**

T4Provably Correct Mixed Initiative Systems

- **Proofs** of correctness, **tools** for synthesis
- Hierarchical Decision Making and Controller Synthesis: Scaling Up
  - reinforcement learning operates on-line but often makes *myopic* decisions
  - model-based planning leverages known structure to ensure high-quality decisions
- Learning by Doing
  - learn from rich instruction; provide advice & reward to human
  - robust to inconsistency; respects neuronal learning speed in human sensori-motor loop

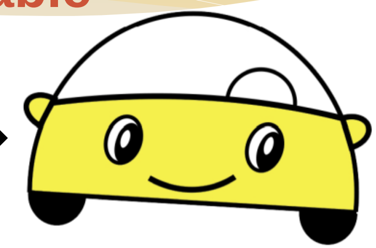
Controller Synthesis from Logic Specifications

LOGIC SPECIFICATION

Goal
Environment Model
Human Model

→ Controller Synthesis →

Realizable



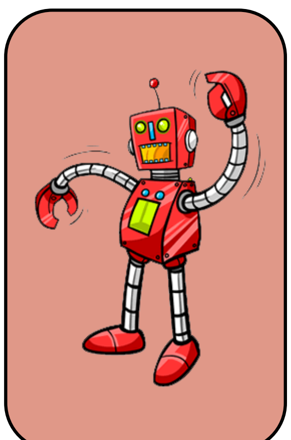
Given a **formal specification**, encoding the

- objective,
- environment model,
- **human model,**

synthesize a controller that is **guaranteed** to satisfy the **specification**.

Human-Aware Control

- Closed-Loop Human Modeling
- Planning to Leverage Effects on Humans



Controller Synthesis from Formal Specifications

- Systematic Human-Intervention
- Control under Uncertainty

Interaction with Humans

Google's Driverless Cars Run Into Problem: Cars With Drivers

By MATT RICHTEL and CONOR DOUGHERTY SEPT. 1, 2015

Email

Share

Tweet

Save

More

MOUNTAIN VIEW, Calif. — Google, a leader in efforts to create driverless cars, has run into an odd safety conundrum: Its cars don't make enough mistakes.

Last m
approa
suppos
pedestr
to appl
not so r
behind

"One of the biggest challenges facing automated cars is blending them into a world in which humans don't behave by the book."

Google's fleet of autonomous test cars is programmed to follow the letter of the law. But it can be tough to get around if you are a stickler for the rules. One Google car, in a test in 2009, couldn't get through a four-way stop because its sensors kept waiting for other (human) drivers

The Google self-driving car, with Eric Schmidt, left, the company's executive chairman, and Transportation Secretary Anthony Foxx. Justin Sullivan/Getty Images

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

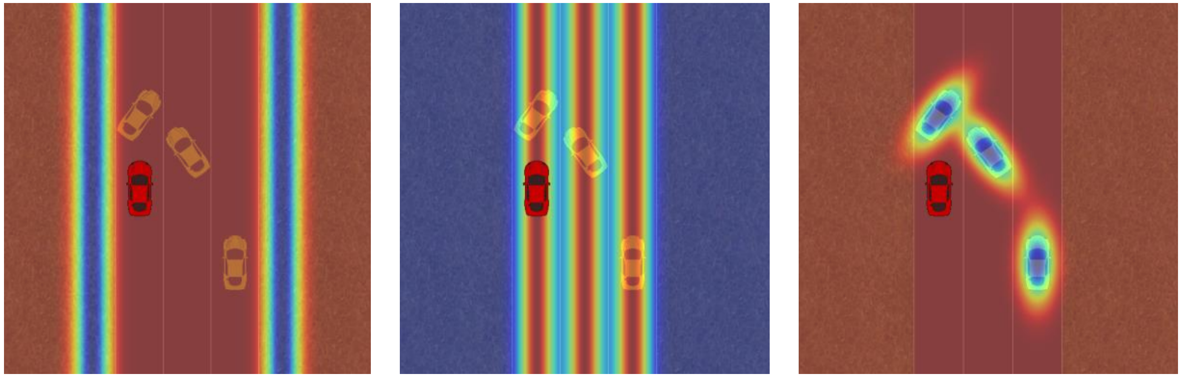
It is difficult to deal with humans, even if we eliminate the driver.

Learning Driver Models

Learn Human's reward function based on Inverse Reinforcement Learning:

$$P(\boldsymbol{u}_H | x_0, w) = \frac{\exp(R_H(x_0, \boldsymbol{u}_R, \boldsymbol{u}_H))}{\int \exp\left(R_H(x_0, \boldsymbol{u}_R, \tilde{\boldsymbol{u}}_H)\right) d\, \tilde{\boldsymbol{u}}_H}$$

$$r_H\left(x^t, u_R^t, u_H^t\right) = w^\top \phi(x^t, u_R^t, u_H^t)$$



(a) Features for the boundaries of the road

(b) Feature for staying inside the lanes.

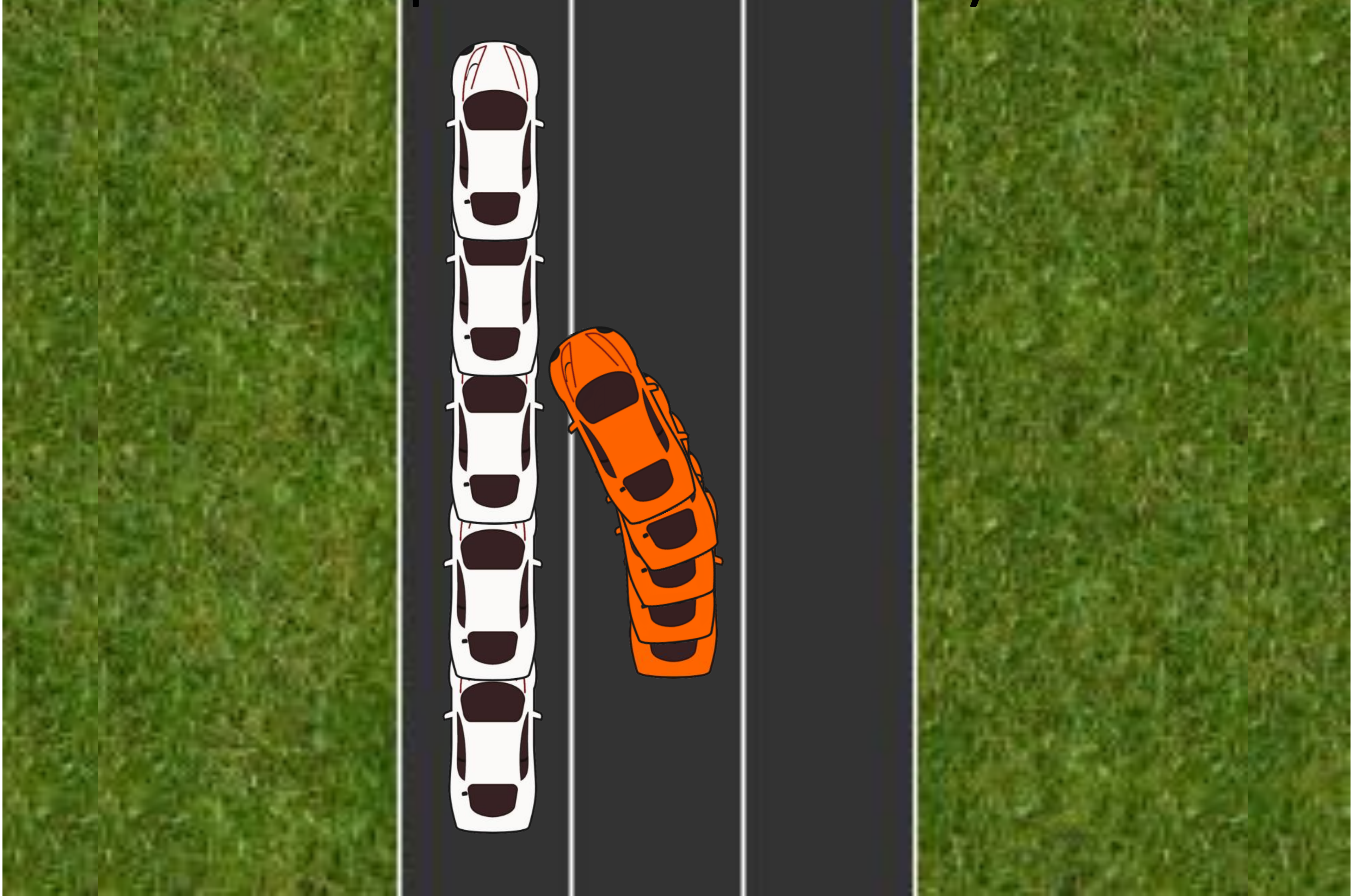
(c) Features for avoiding other vehicles.

B. Ziebart, A. Maas, J. A. Bagnell, and A. K. Dey. Maximum entropy inverse reinforcement learning. In AAAI, 2008.

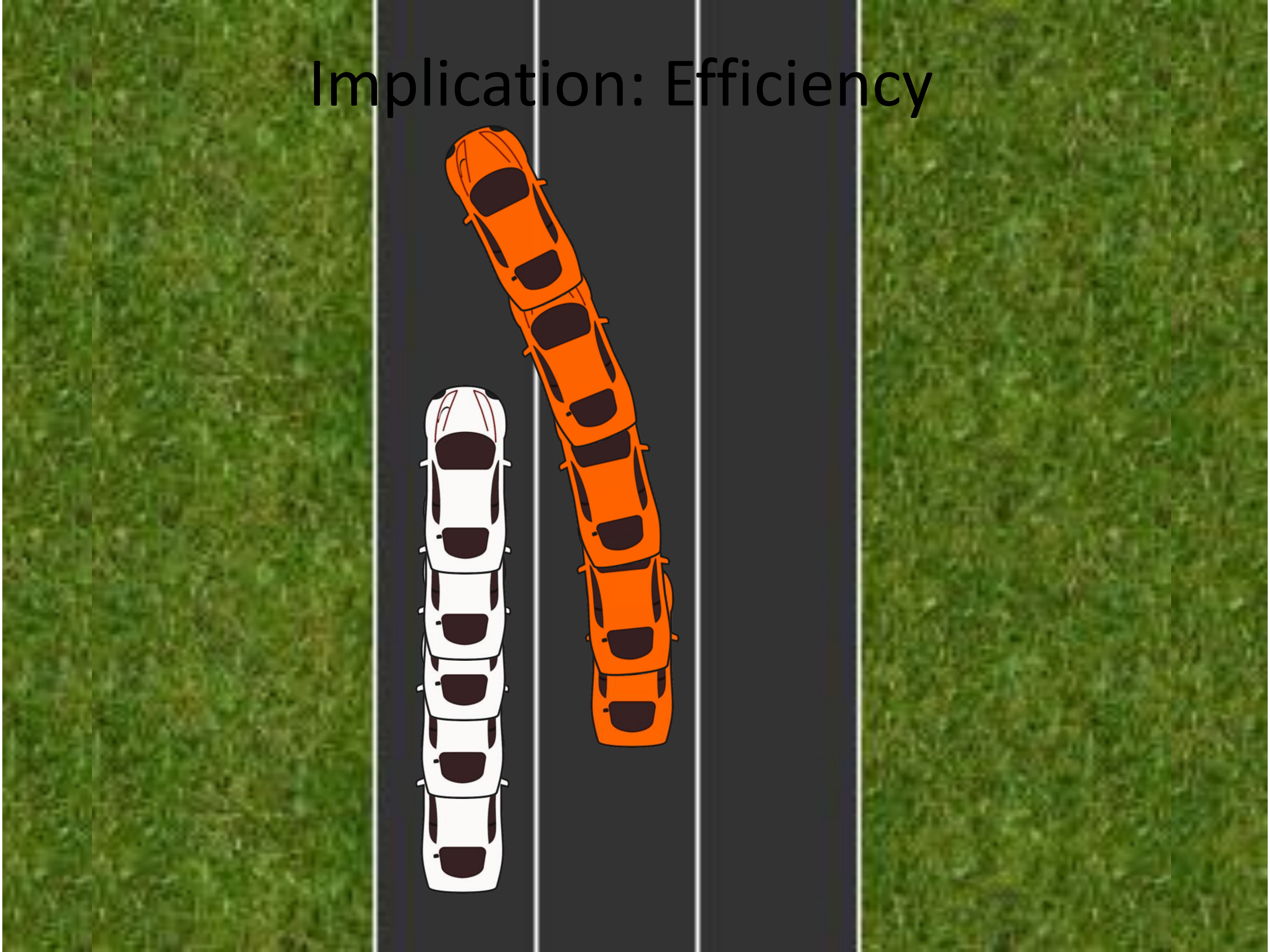
S. Levine, V. Koltun. Continuous inverse optimal control with locally optimal examples. arXiv , 2012.

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Implication: Efficiency

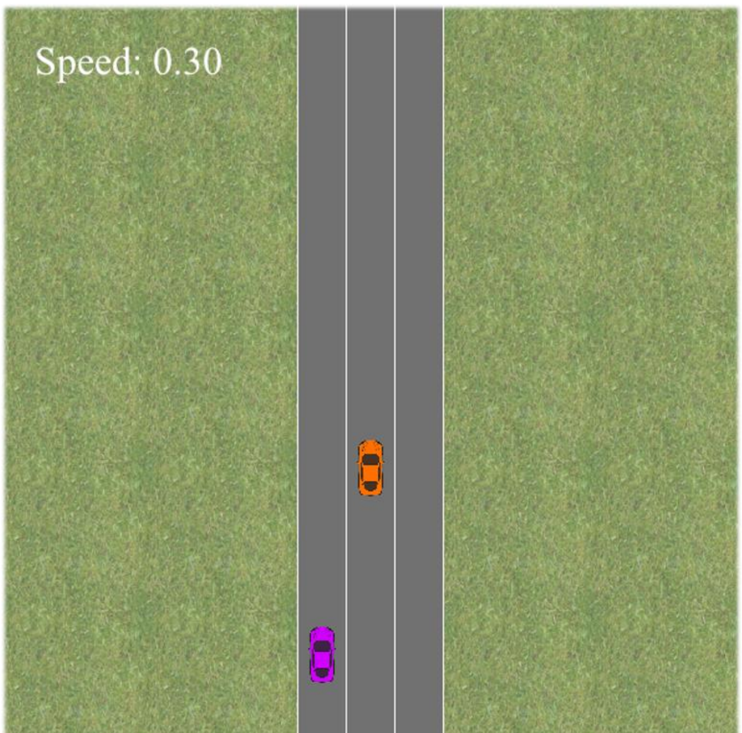


Implication: Efficiency

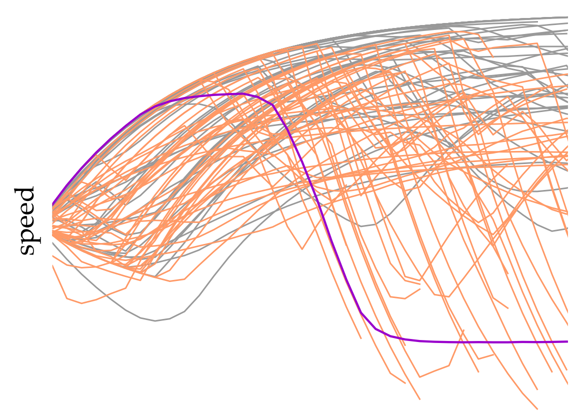


Make Human Slow Down

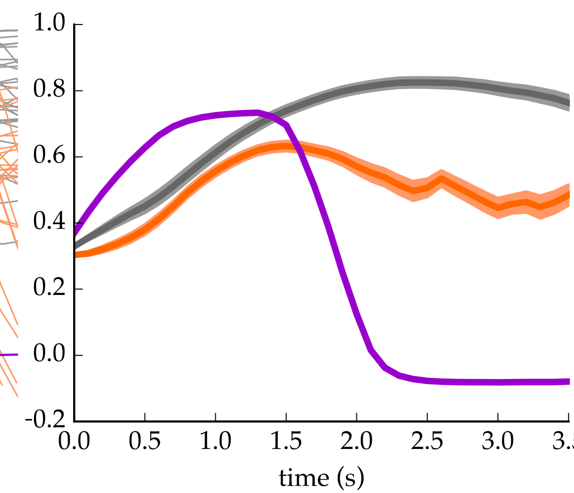
Autonomous vehicle optimizes for efficiency, and leverages affects on the human.



Affect Human
Avoid Human
Trained Human Mo

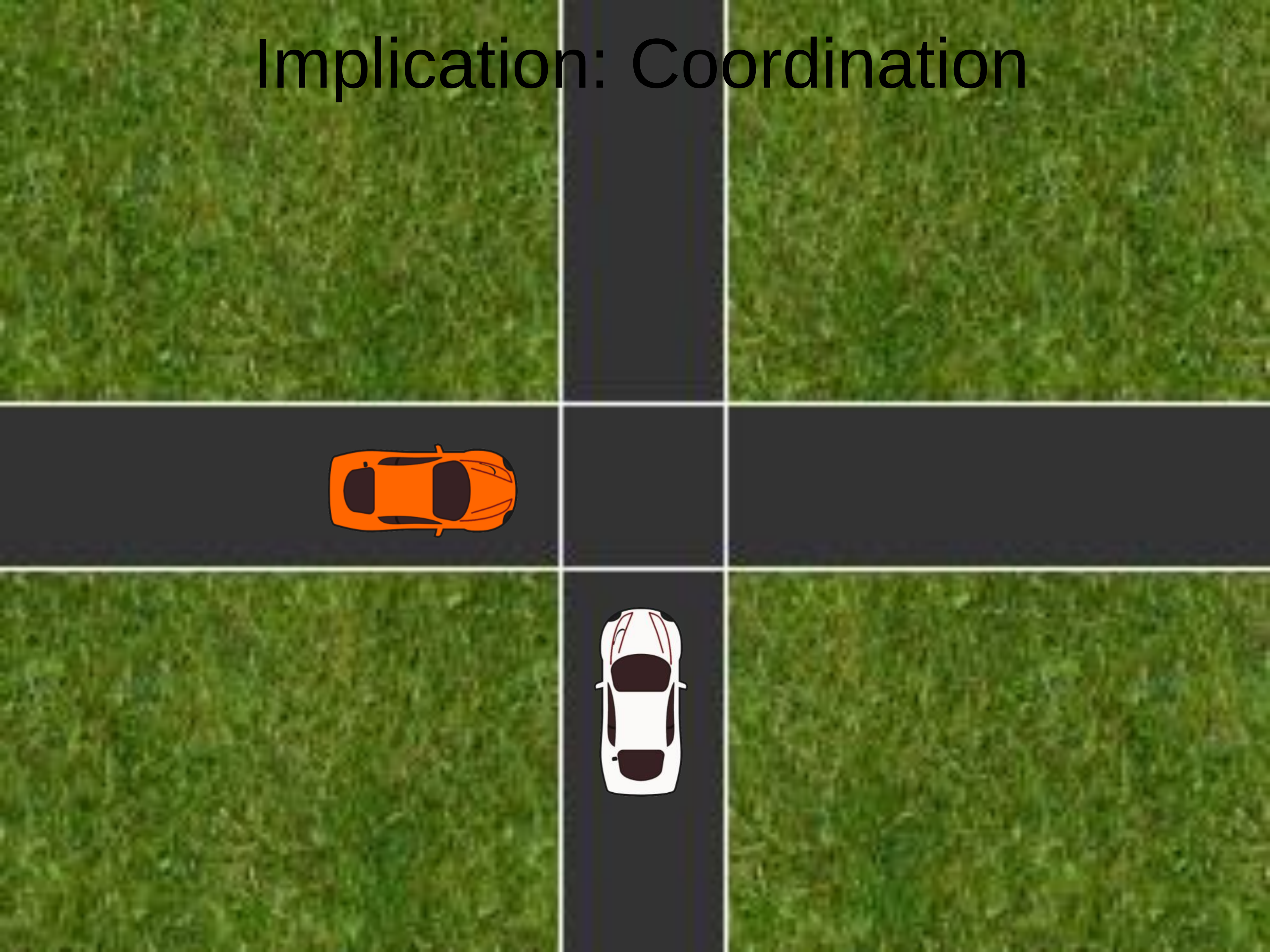


(a)

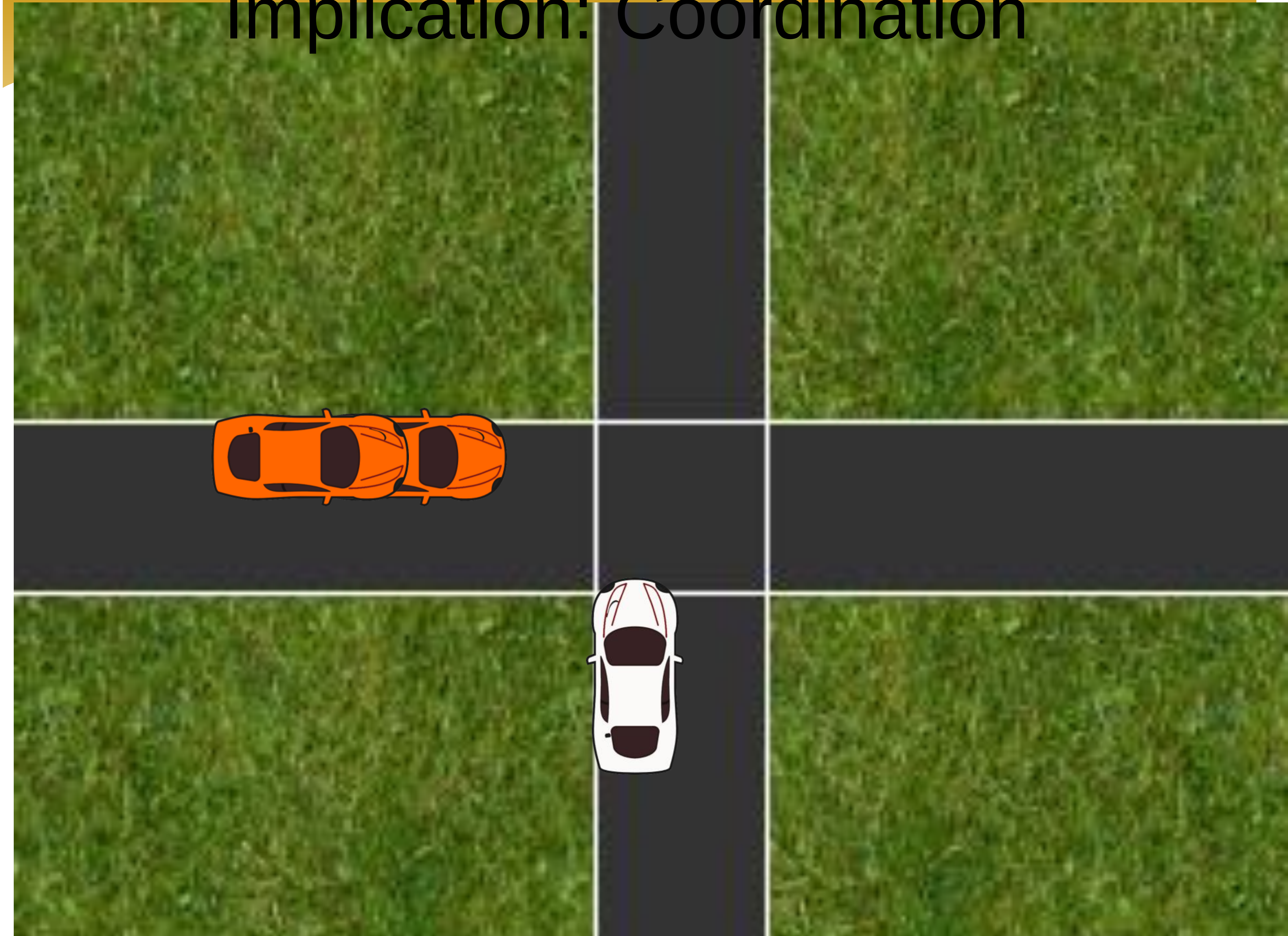


(b)

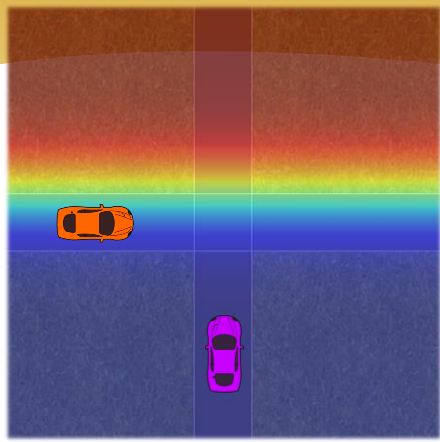
Implication: Coordination



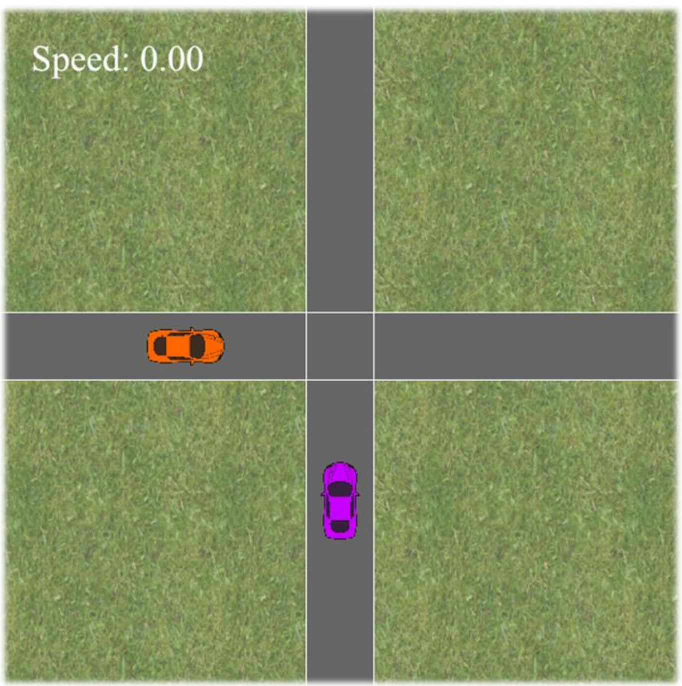
Implication: Coordination



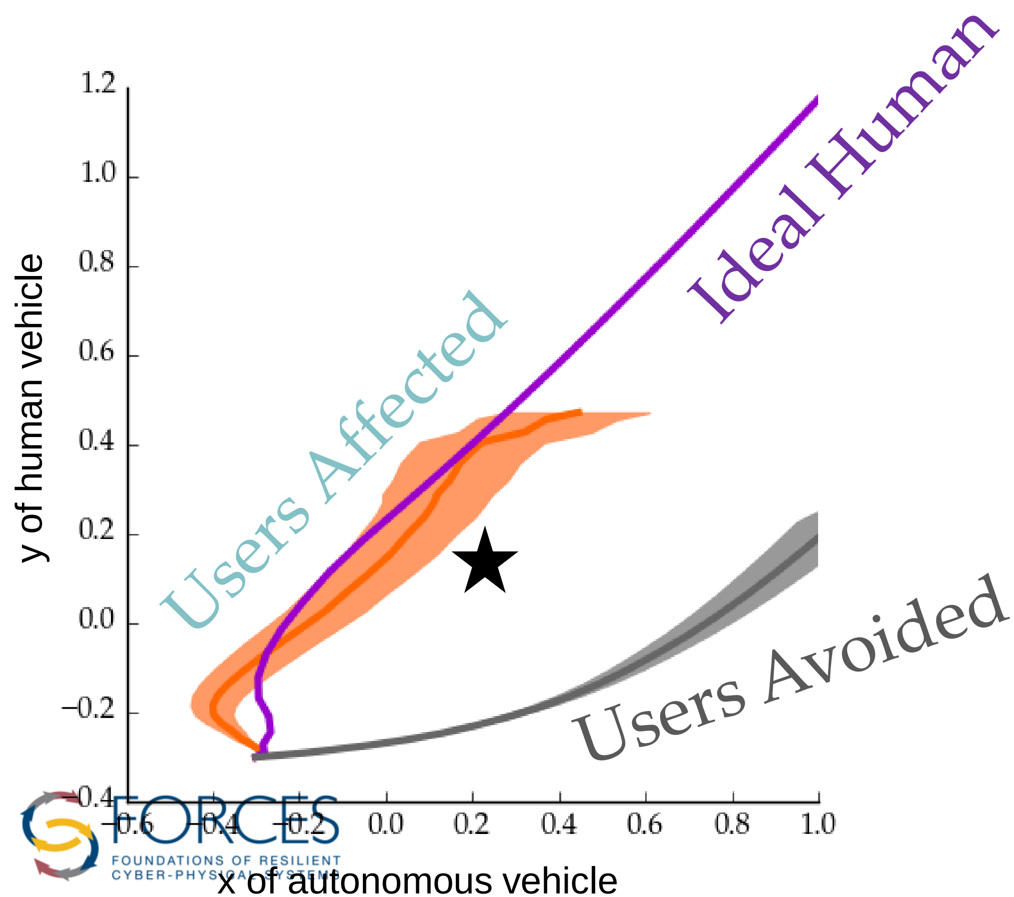
Make Human Cross First



Reward for making the human cross first.



Autonomous vehicle backs up to communicate with the human, and make her cross first.

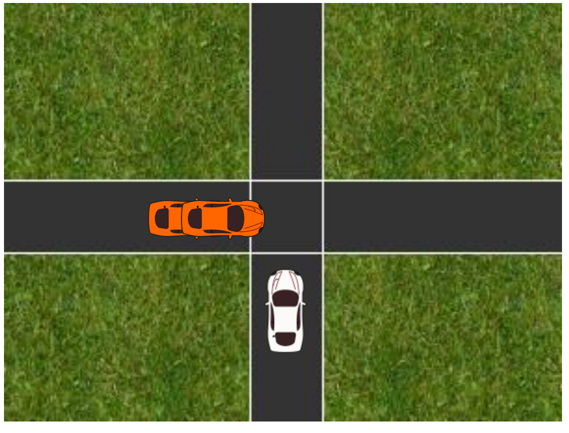


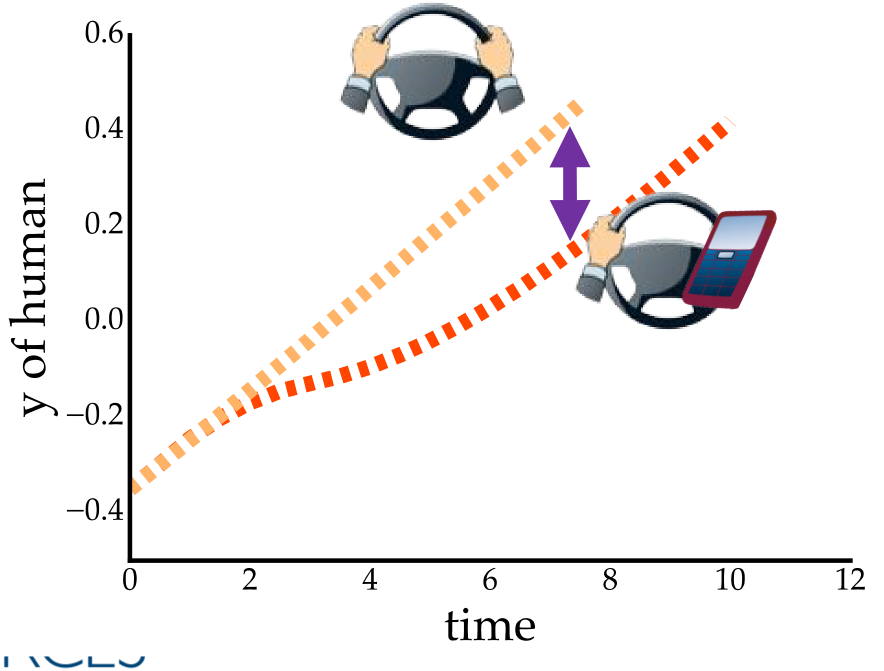
FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Active Information Gathering

Active information gathering over human's internal state.

Human's internal state:
φ : Aggressive vs Timid
Attentive vs Distracted



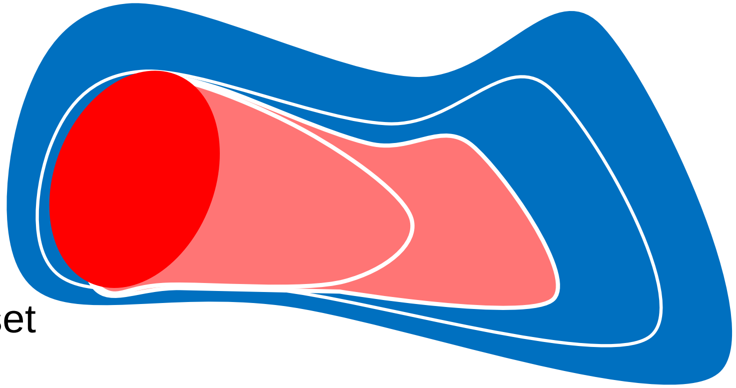


D. Sadigh, S. S. Sastry, S. Seshia, A. Dragan. Information Gathering Actions over Human Internal State. IROS 2016.

Safe Learning

* RL can not be used in a safety critical environment!
  * Machine learning algorithms converge asymptotically
  * Some natural parameterization can behave poorly during training
* Developed framework for combining arbitrary ML methods with safety analysis techniques
  * How can we use reinforcement learning to improve performance online, while still guaranteeing system safety?
  * Guaranteed-safe online learning via reachability [Gillula, Tomlin ’14]
  * Safe exploration and model validation [Akametalu, Fisac, Tomlin ’14, ’15]

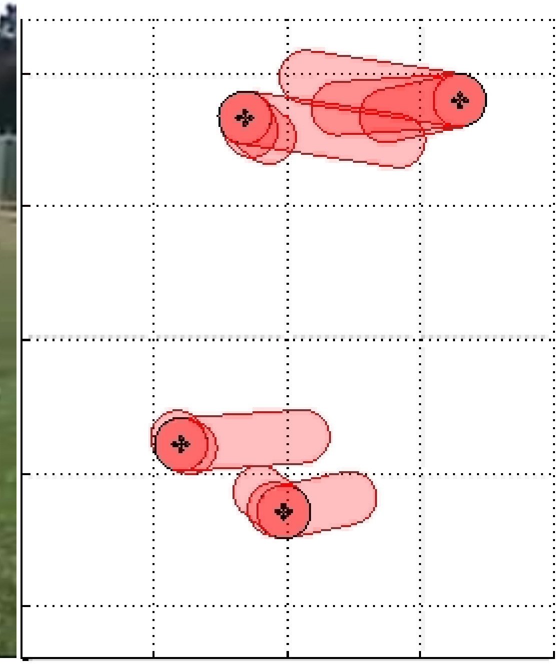
- Initialize active unsafe set = smallest candidate set
- Repeat:
  - Measure disturbance
  - Validate measured disturbance at visited states against model
  - If model inaccuracy is detected, expand unsafe set
  - Update disturbance model



FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

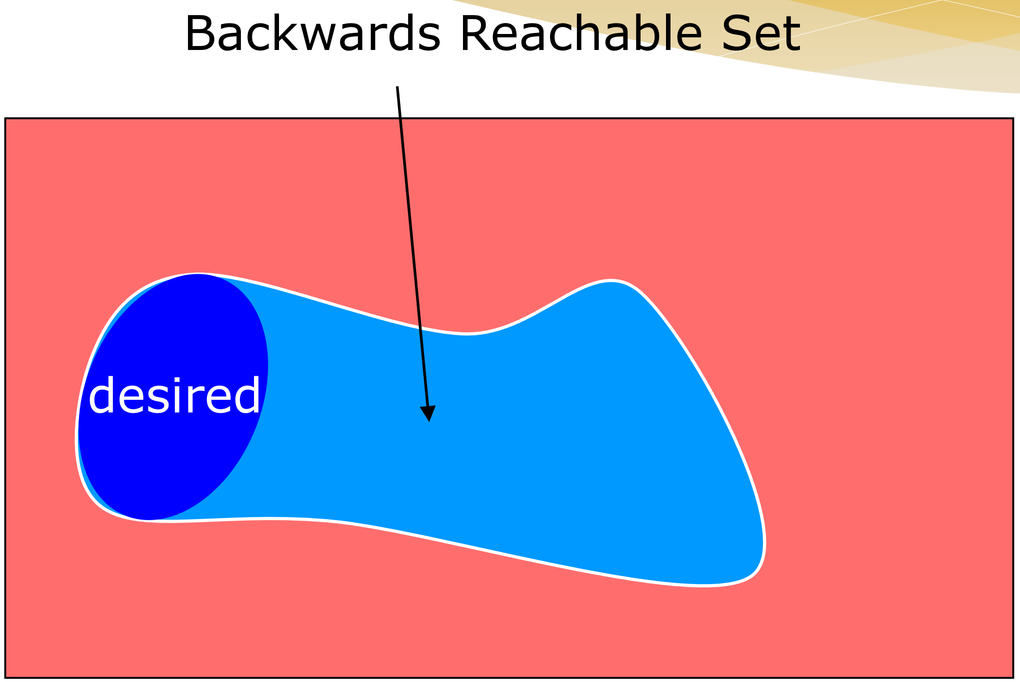
Example 1: Collision Avoidance

Pilots instructed to attempt to collide vehicles



[STARMAC: Stanford Testbed of Autonomous Rotorcraft for MultiAgent Control]

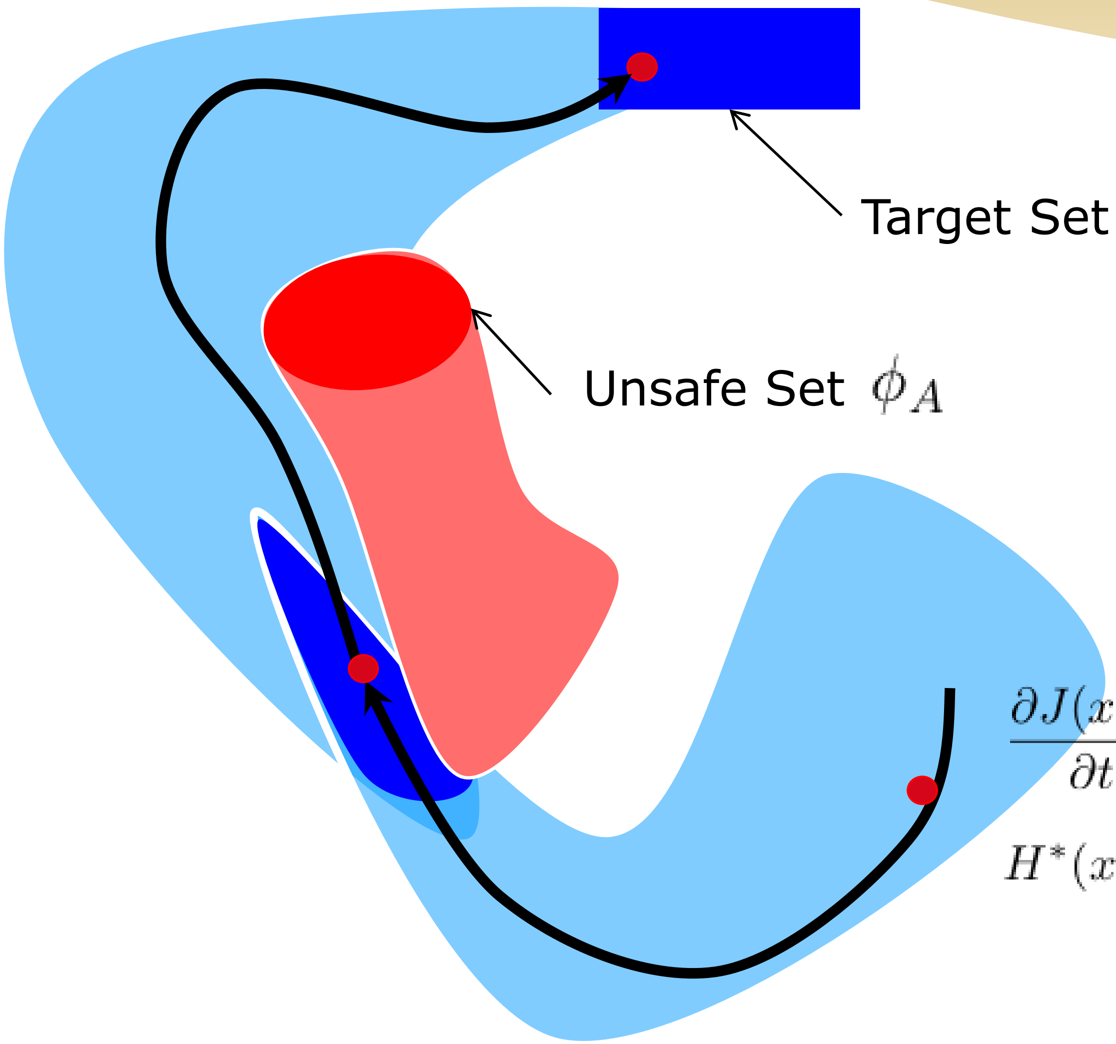
Backwards Reachable Set: Capture



Capture property can also be encoded as a condition on the system's reachable set of states

$$-\frac{\partial J(x,t)}{\partial t} = \min\{0, \min_u \max_d \frac{\partial J(x,t)}{\partial x} f(x,u,d)\}$$

Mode sequencing and reach-avoid



String together capture sets, starting from the target set and working backwards

Avoid sets can be combined with capture sets to guarantee safety

$$\frac{\partial J(x,t)}{\partial t} + \min[0, H^*(x, \frac{\partial J(x,t)}{\partial x})] = 0$$

$$H^*(x, \frac{\partial J(x,t)}{\partial x}) =$$

$$\min_u \max_d \frac{\partial J(x,t)^T}{\partial x} \cdot f(x,u,d,t)$$

Subject to $J(x,t) \geq -\phi_A$

Dealing with the curse of dimensionality

* **Impose practical constraints**
  * **Protocols, additional problem structure**

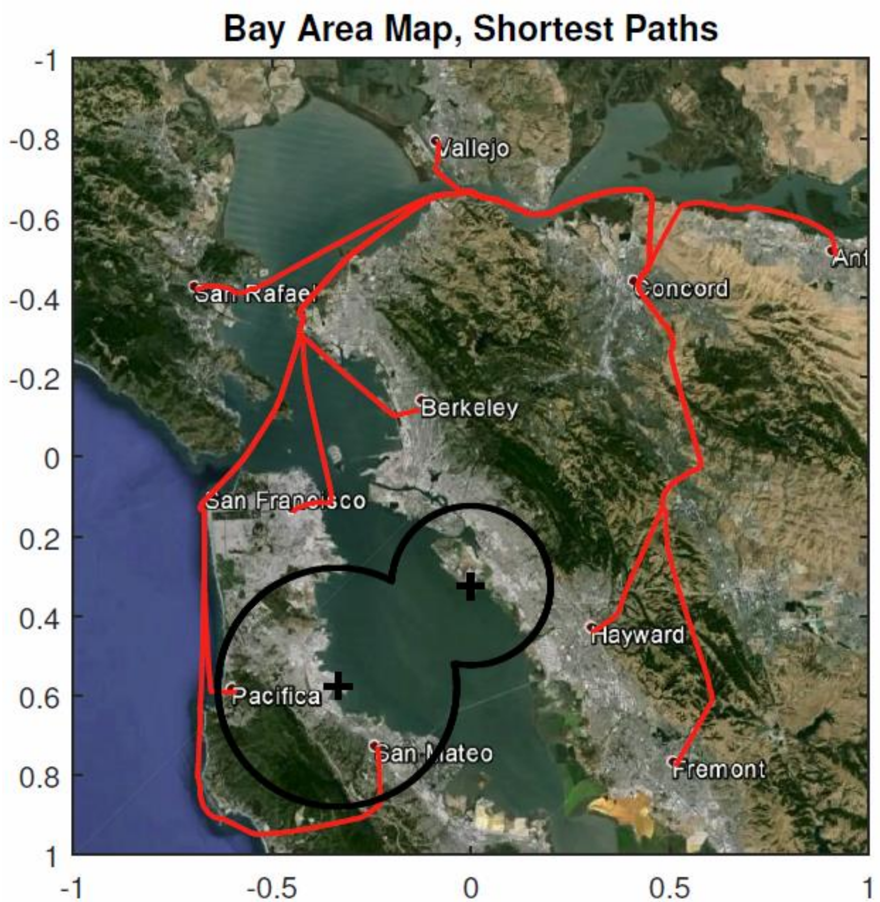
* Approximations
  * Bisimulations (Girard, Pappas, Tabuada)
  * Piecewise and multi-affine systems (Morari, Borrelli, Krogh, Johansson, Rantzer, Belta)
  * Ellipsoidal and polyhedral sets (Kurzhanski , Varaiya, Stipanovic)
  * Funnels and barrier certificates (Parillo, Majumdar, Tedrake, Papachristodoulou, Julius, Lall, Topcu)
  * Decoupling disturbances (Chen, Herbert)

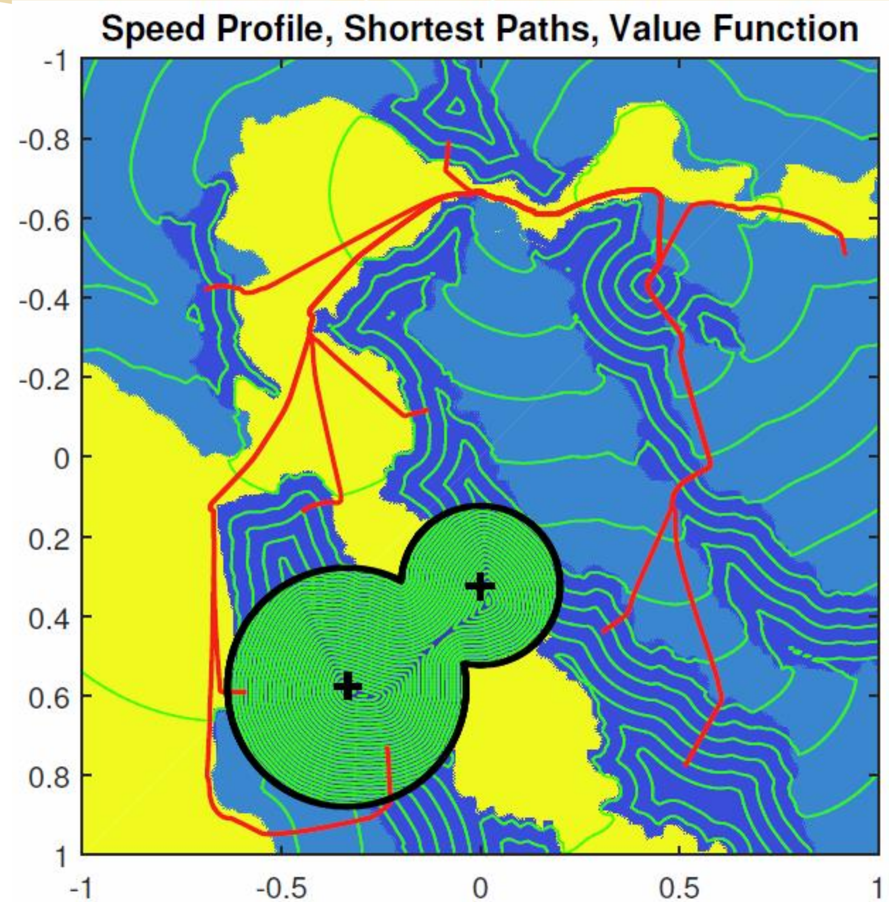
* Mathematical structure
  * Monotone systems (Sontag, Del Vecchio, Arcak, Coogan)
  * LTL specifications (Kress-Gazit, Raman, Murray, Wongpiromsarn, Belta)

* Decompositions (Mitchell, Del Vecchio, Chen, Herbert, Grizzle, Ames, Tabuada)

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

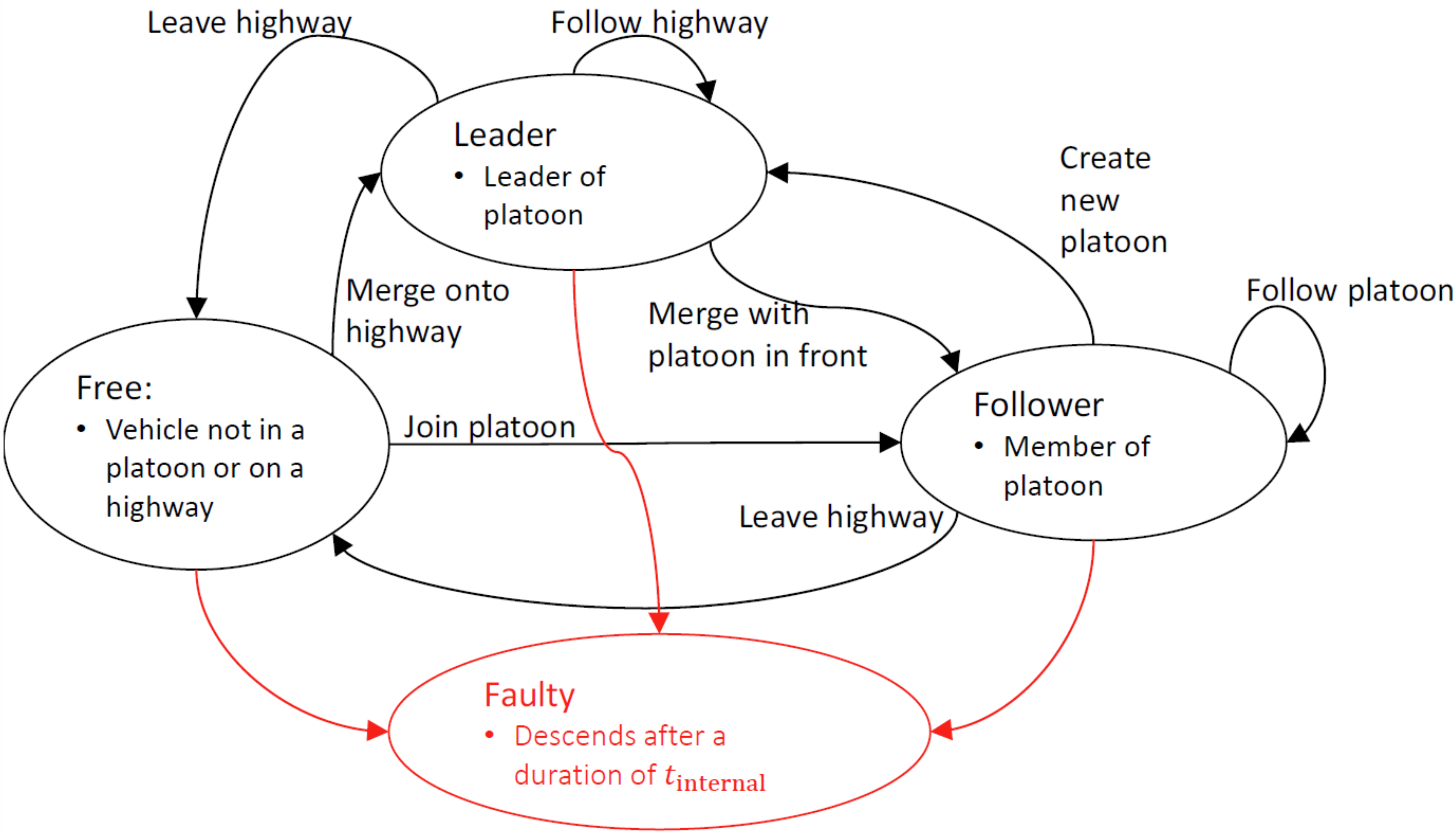
Example 2: Platooning UAVs





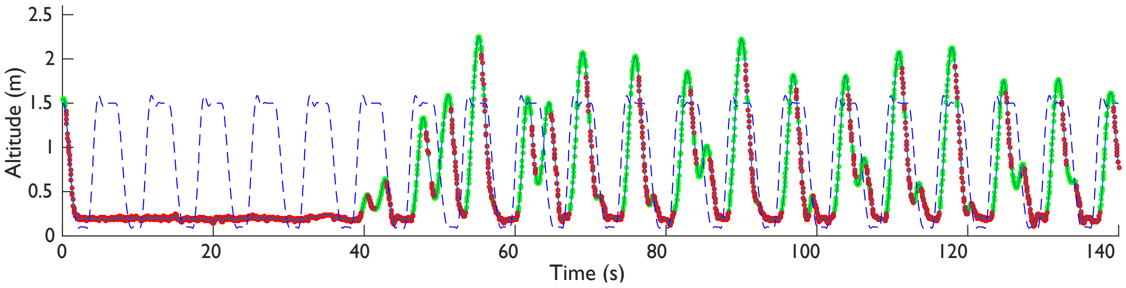
FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Example 2: Platooning UAVs



Example 3: Safe Policy Gradient Reinforcement Learning

The quadrotor first: drops



After about 1 minute,
it can roughly track the trajectory

Soon, it starts experimenting

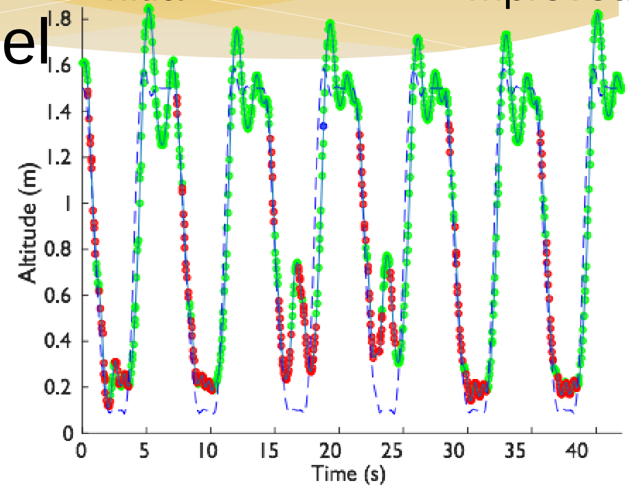
…but the safe controller steps in

[PGSD: Kolter and Ng, 2009]

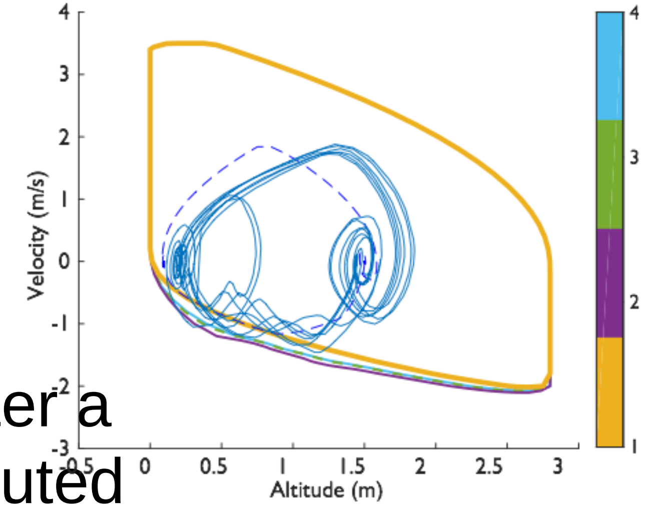
FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Example 4: Safe Learning

First computed model is locally inaccurate



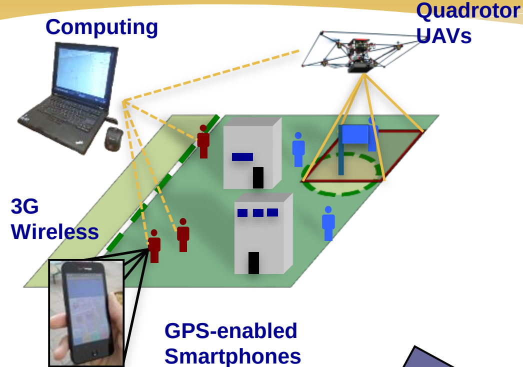
System detects inconsistency, slightly contracts safe set



Tracking resumes after a better model is computed

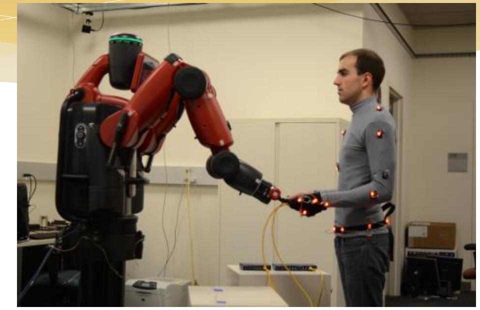
FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Integrative Experiments



Vehicle Control (Tomlin)

3D SLAM (Yang, Sastry)



Human Collaboration (Bajcsy)

Berkeley AR Headset

Robot OS