VeHICaL: Verified Human Interfaces, Control, and Learning for Semi-Autonomous Systems

S. A. Seshia, R. Bajcsy, T. Griffiths, B. Hartmann, S. Sastry, C. Tomlin (Berkeley), R. Murray (Caltech), C. Sturton (UNC Chapel Hill)

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

NSF CPS PI Meeting
October 31, 2016

Human Cyber-Physical Systems (h-CPS)

CPS that operate in concert with humans

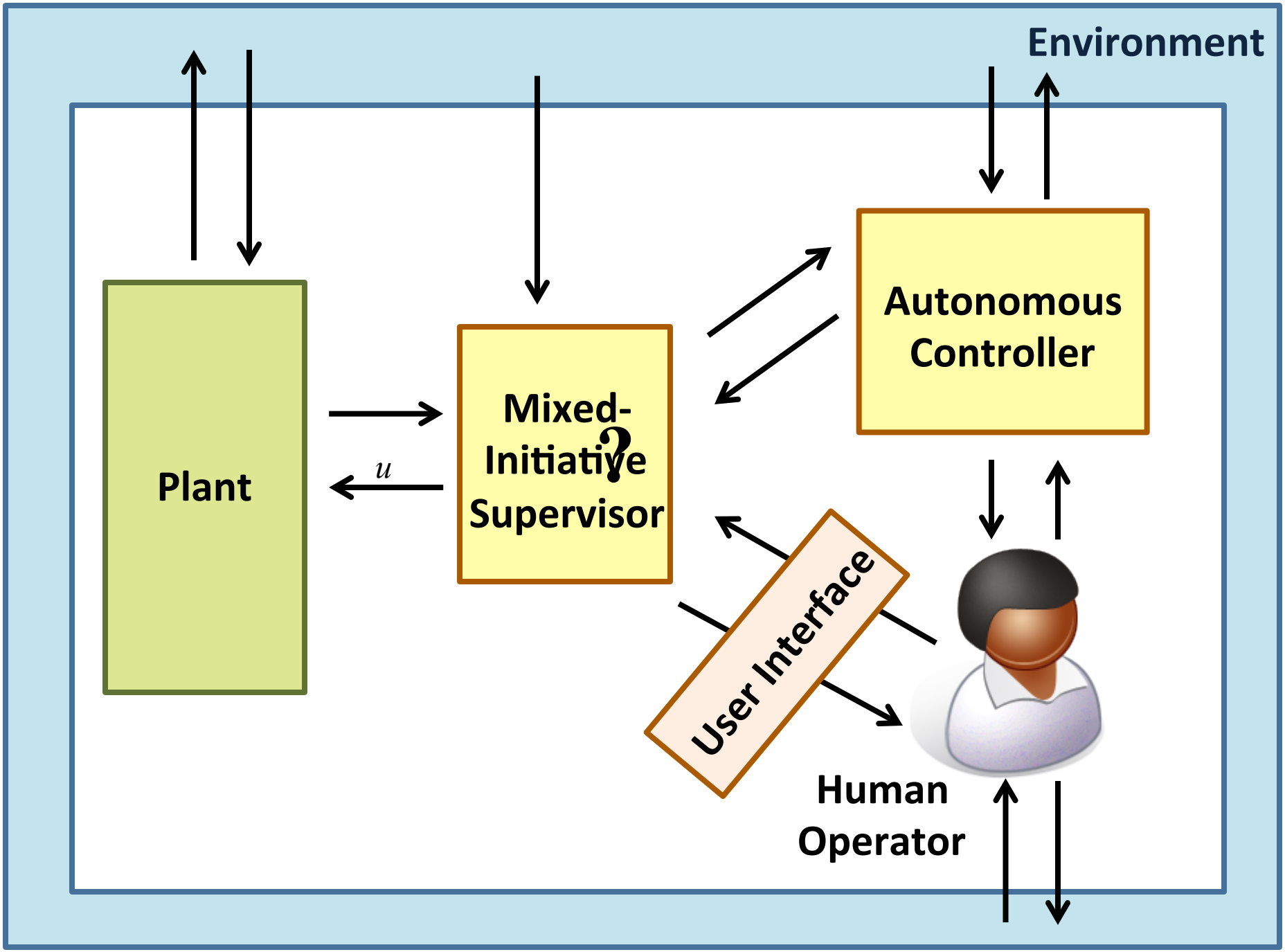
Driver Assistance in Cars

Fly-by-wire Cockpit Interfaces

UAVs with Human Operators

Robotic Surgery & Medicine

Semi-Autonomous Manufacturing


Environment
Plant
u
Mixed-Initiative Supervisor
?
Autonomous Controller
User Interface
Human Operator

Overall Project Objective of VeHICaL

*To develop a **science of verified co-design** of controllers for semi-autonomous cyber-physical systems and interfaces between humans and cyber-physical components*

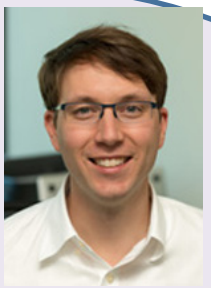
Team

Bajcsy

Robotics & Perception

Griffiths

Psychology & Cognitive Science



Hartmann

Human-Computer Interaction

Sastry

Tomlin

Control Systems

Murray

Seshia

Formal Methods

Sturton

Security & Privacy

Motivating Applications

- **Semi-Autonomous Automobiles**

TECHNOLOGY

The New York Times

The 15-Point Federal Checklist for Self-Driving Cars

By CECILIA KANG SEPT. 20, 2016

A Lesson of Tesla Crashes? Computer Vision Can't Do It All Yet

- **Semi-Autonomous UAVs**

FAA Expects 600,000 Commercial Drones In The Air Within A Year

August 29, 2016 · 3:10 PM ET npr

AARIAN MARSHALL TRANSPORTATION 08.30.16 1:28 PM WIRED

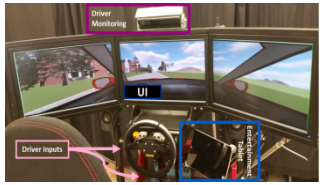
HOW TO ACE THE FAA'S NEW TEST AND BECOME A PRO DRONE PILOT

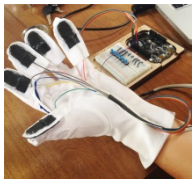
Project Goal: Develop Science of Design of h-CPS with Provable Guarantees on Behavior

Some Major Unsolved Problems to be Tackled:

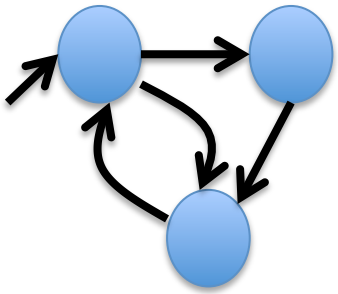
- Effective Sharing of Control between Human and Autonomous Controller
- Modeling Human Cognition, Perception, and Action, and developing Sensing Strategies to infer and evaluate such models
- Uncertainty in h-CPS Dynamics *and* in Modeling: need new techniques blending Verification, Learning, and Control
- Verifiable Human-Computer Interface Design *integrated with* Control System Design
- Privacy-preserving Learning of Models of h-CPS

Sensing & HMI Design



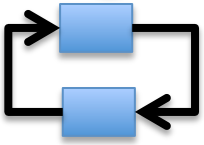


Formal Modeling & Specification



Verification

Control

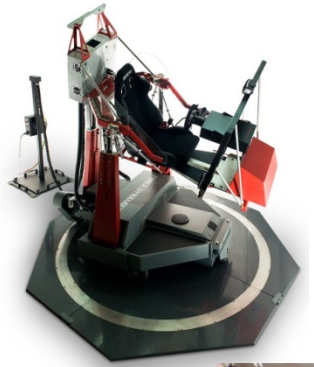


Security & Privacy

VeHiCal

System-level Integration & Validation

Learning Models from Data



Evaluation & User Studies

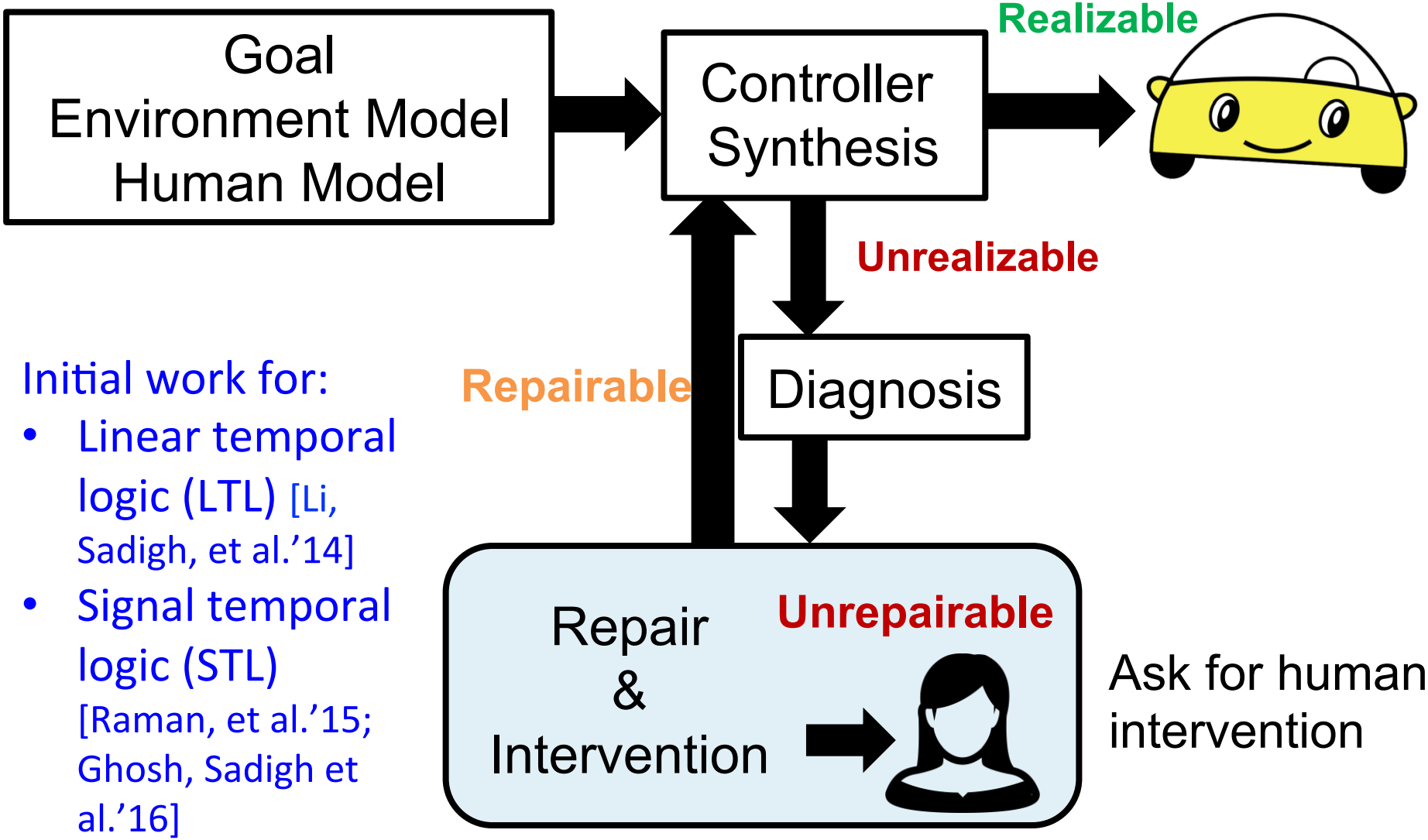
Four Project Thrusts

1. Specification and Modeling
2. Learning, Verification, and Control
3. Human-Machine Interface Design & Verification
4. Testbed and Evaluation

Specification for h-CPS: What does "correctness" mean? [Li, Sadigh, Sastry, Seshia, TACAS'14]

- Safe & Correct Autonomy
  - Autonomous mode should be proved correct
- Effective Monitoring
  - Must monitor all relevant environment variables
- Prescient Switching
  - Sufficient lead time to switch between human & auto
- Minimally Intervening
  - Should *rarely* request human intervention
- Other Properties:
  - Security & Privacy
  - HMI Properties: Absence of Mode Confusion, ...
  - ...

Correct-by-Construction (Semi-Autonomous) Control



Initial work for:

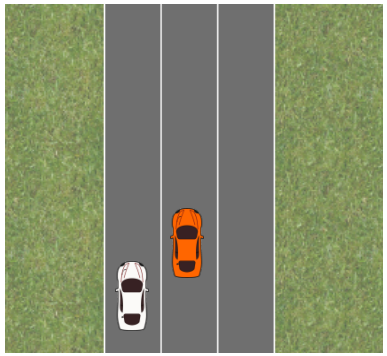
- Linear temporal logic (LTL) [Li, Sadigh, et al.'14]
- Signal temporal logic (STL) [Raman, et al.'15; Ghosh, Sadigh et al.'16]

Human-Aware Control of Vehicles

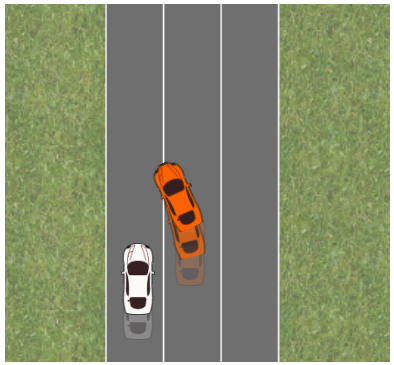
[Sadigh, Sastry, Seshia, Dragan; RSS, IROS '16]

Leverage human *responses* to estimate human internal state.

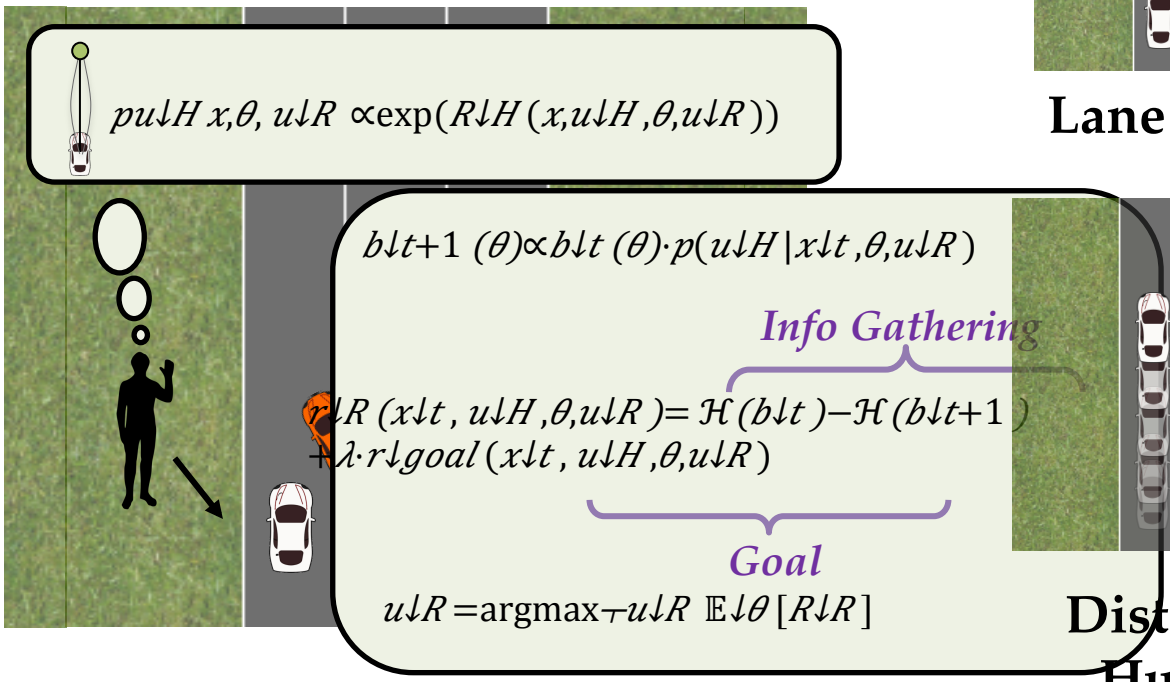
SAMPLE RESULT

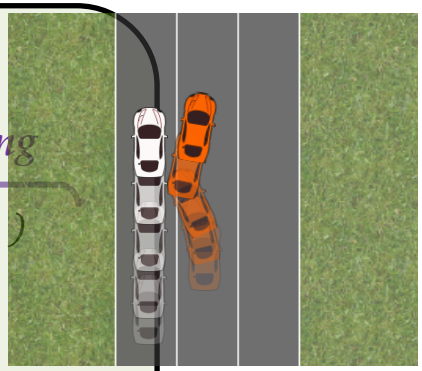


Lane Change

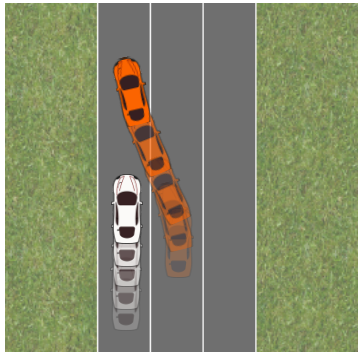


Nudging In





Distracted Human



Attentive Human

Human Modeling: Data-Driven Modeling and the "Bounded Optimal" Model

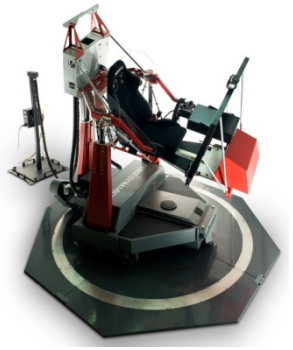
- Uncertainty in human behavior and variation (person to person) are central aspects of h-CPS
  - Probabilistic formalisms, learned from data
  - Consistent with formalisms using in verification and control
- Model human behavior based on "Resource Rationality" [Griffiths, Lieder, & Goodman, 2015]
  - Trade off error with the cost of computation, assuming an abstract computational architecture
  - an instance of "bounded optimality": choosing the best algorithm to solve a problem (Russell, 1995)

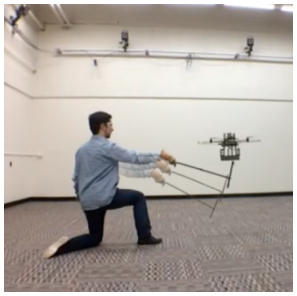
Resource rationality in VeHICal: Directions

- **Prediction:** Modeling the allocation of cognitive resources in dynamic environments
  - e.g., do we make rational choices about when to check our texts when driving?
- **Inference:** Developing models that can be used to infer when/how humans are resource-limited
  - e.g., behavior patterns that result from having less resources for tracking, relevant to prescient switching

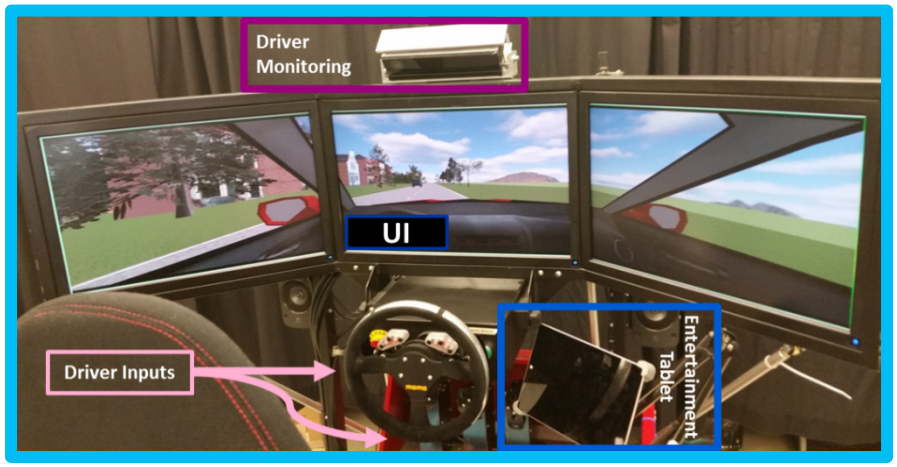
Experimentation on CPS Testbeds

- **Driving Simulator Testbed:** Force Dynamics 401CR (industrial simulator with motion feedback)
- **Automotive:** Partnerships with automotive companies
- **UAVs:** Tomlin lab, partnerships with drone companies
- **Jacobs Hall**: State-of-the-art labs for rapid prototyping of interfaces and hardware



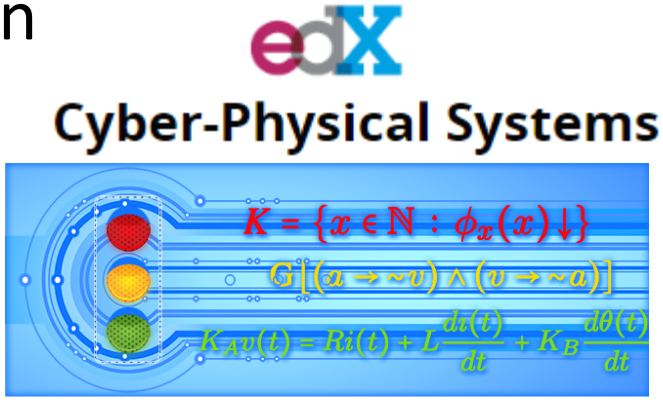


Initial work [Rezvani et al., ITSC'16]: Impact of user interface design on transfer of control



Education and Outreach activities

- Joint development and dissemination of educational materials on h-CPS
  - *Online course on h-CPS* jointly offered by PIs
- Joint mentoring of postdoctoral researchers
- Summer programs for undergraduate and high/middle school students
  - Projects conceived of jointly and offered in a coordinated manner to amplify the overall impact



Seventh Annual
Cyber-Physical Systems Principal Investigators' Meeting
Arlington, VA | October 31 – November 1, 2016

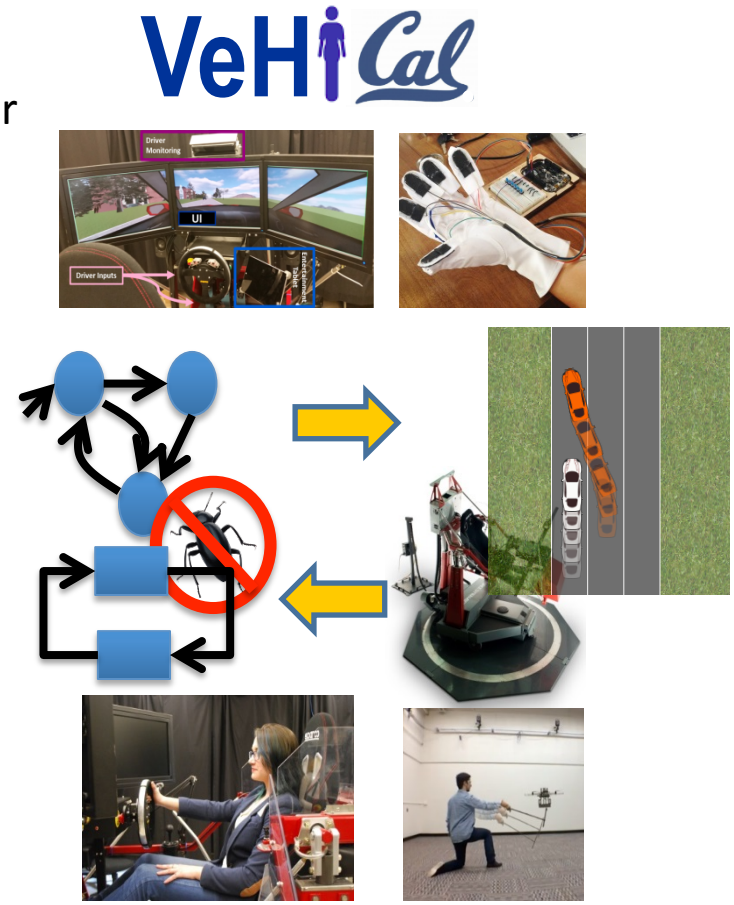
VeHICaL: Verified Human Interfaces, Control, and Learning for Semi-Autonomous Systems

Challenge:

- *Co-design human interfaces and control* for human-cyber-physical systems with *provable guarantees*
- Apply to semi-autonomous vehicles (ground and air)

Solution:

- Integrate Learning, Verification and Control
- Bounded Resource Rational Human Modeling
- Prototype Controllers & Interfaces, Evaluate on Testbed



Scientific Impact:

- Developing a Science of Co-Design of Human Interfaces and Control
- Bridging Model-Based and Data-Driven Design of CPS

Broader Impact:

- Significantly improve safety, security, and performance of systems where humans interact closely with automation
- Involve middle/high-school and undergraduate students in VeHICaL activities

CPS Awards 1545126, 1544714, 1544924

S. A. Seshia, R. Bajcsy, T. Griffiths, B. Hartmann, S. S. Sastry, C. Tomlin (UC Berkeley), R. Murray (Caltech), C. Sturton (UNC).