CIS 700/002 : Topics in Safe Autonomy

Overview

Insup Lee
CIS 700/002: Topics in Safe Autonomy
Department of Computer and Information Science
School of Engineering and Applied Science
University of Pennsylvania

22 January 2019

Course Info

- Instructor: Insup Lee
  - Office: 602 Levine
  - Office Hours: 4-5 Mon, 2-3 Thr,
- Co-Instructors/Guest Lecturers
  - Justin Gottschlich, Intel
  - James Weimer (?)
  - Rado Ivanov (?)

- Class: TTh at 10:30am – 11:45 am, 313 Towne
- Course website:
  - https://rtg.cis.upenn.edu/cis700-2019

What can you expect?

- Course is to learn about topics and approaches relevant for the development of safe autonomous systems

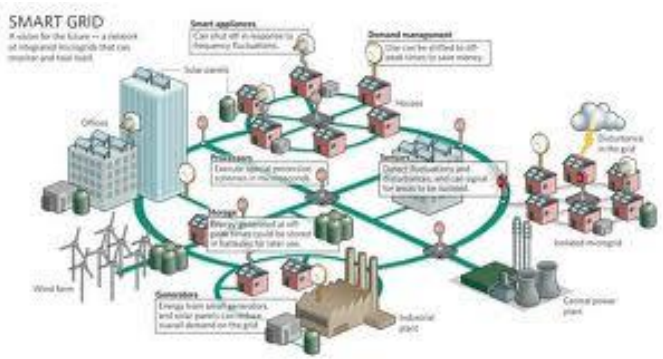
- "More than a reading group" and open ended
  - Read papers and participate in class
  - Select a few topics/papers, present them and lead discussions
  - Identify and carry out a term project
    - Individually or in group of two people
    - Implementation and evaluation: tool, application, case study, etc.
    - Write a high-quality research and/or survey paper
- No Exams
- Grade:
  - In-class presentation/participation: 40%
  - Term projects: 60%

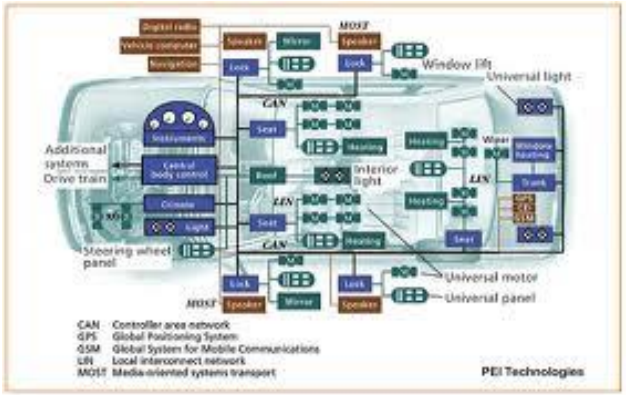
Action Items

- Pick a date to present
- Pick a topic and papers to cover
- Come to my office hour or make an appointment to discuss your topic/presentation
- Send me a copy of your ppt slides at least two days before your presentation date for feedback

Cyber-Physical Systems

- Computation + physical world + networking
- Scale + diversity + ubiquity





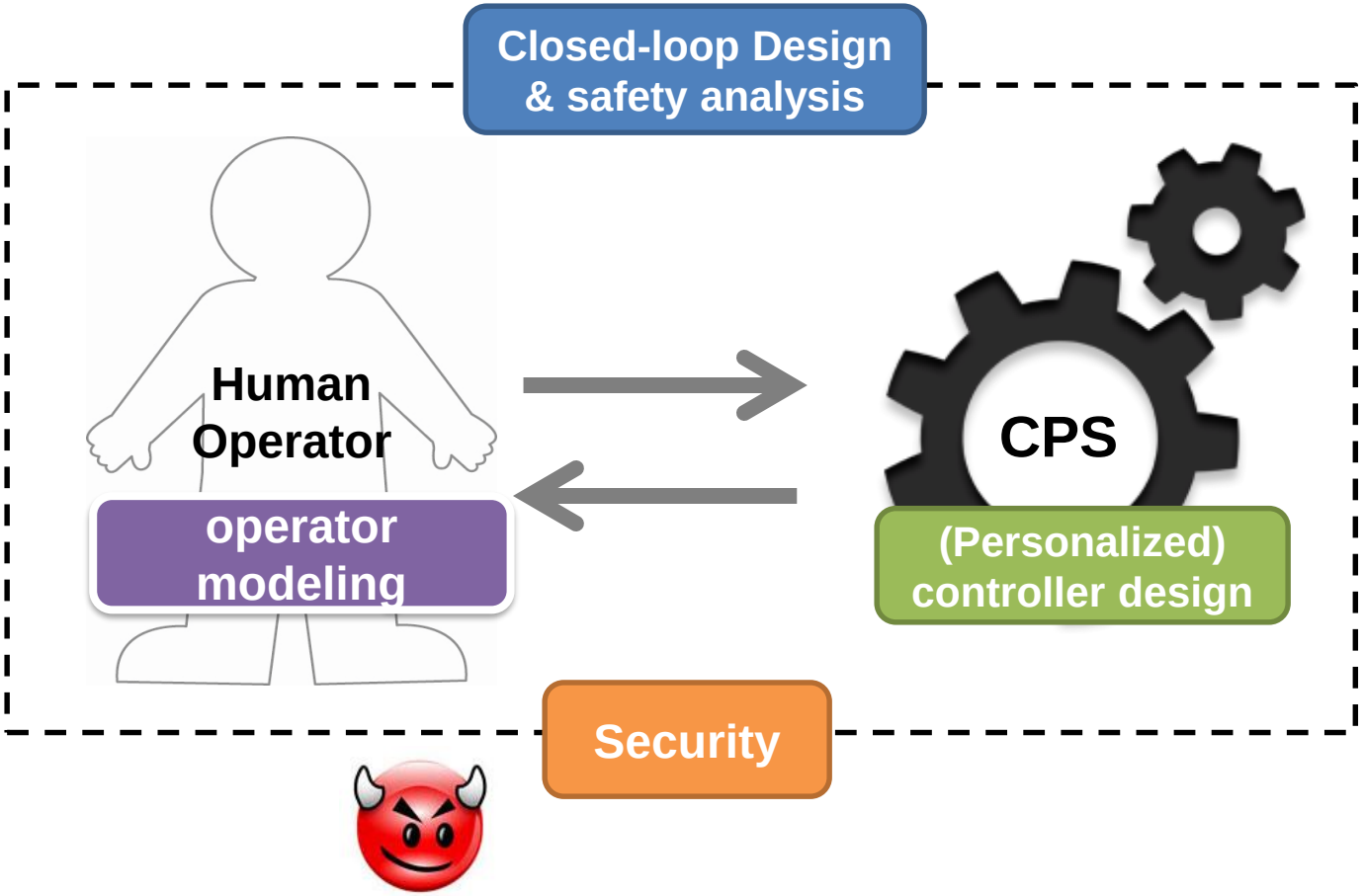
Smart City

Autonomous Systems with Operators

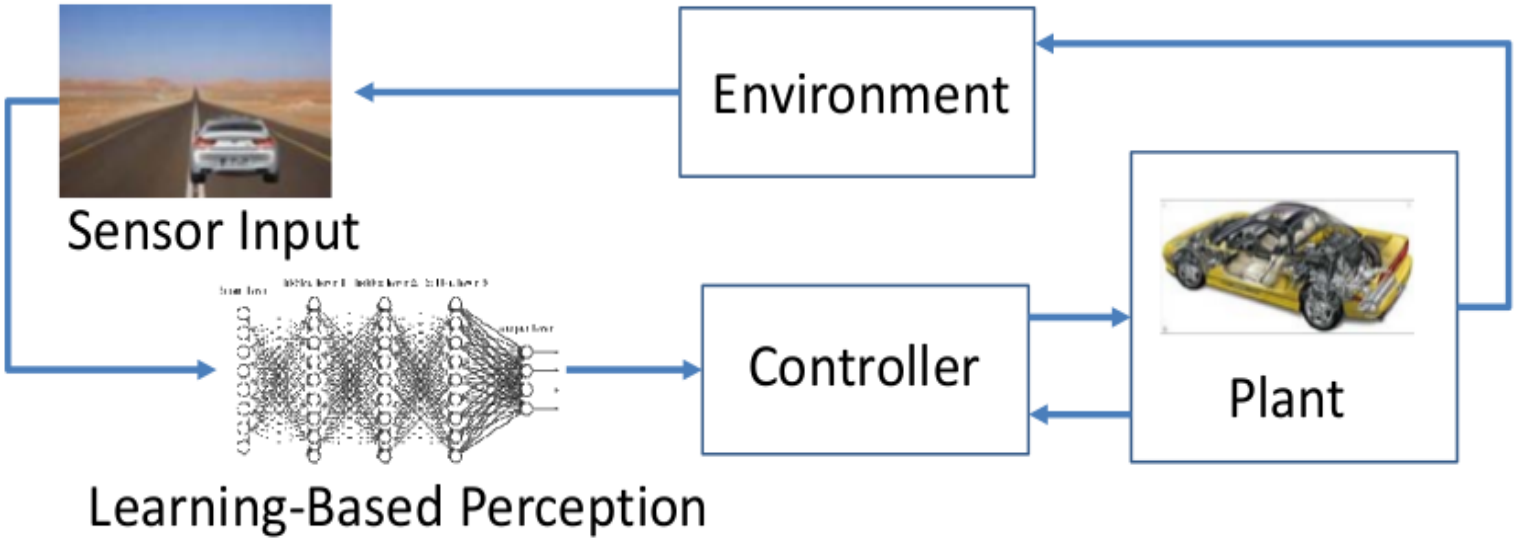
Many **autonomous CPS** interact with **human operators** who act as supervisors or collaborators

Design Time

Human-in/on-the-loop CPS: Modeling, design and analysis

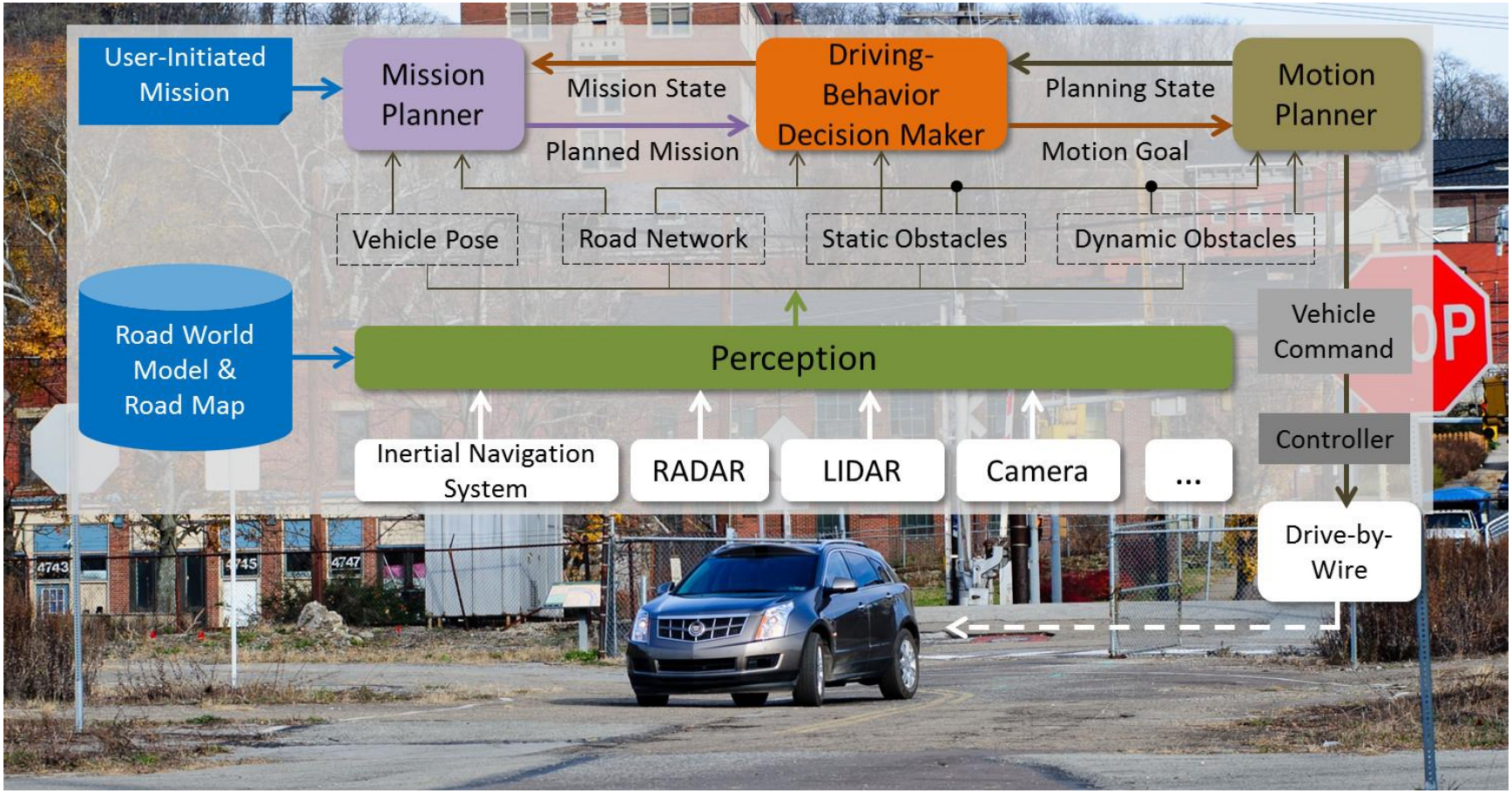


Runtime



- Assuring safety/correctness
  - Gap between design-time and runtime
- Runtime verification
- Anomaly detection
  - Failures/attacks
- Dynamic adaptation

Example Architecture



[Rajkumar]

Typical architecture/components

- Perception
- Reflection/state estimator
- Goal management/decision support/planner
- Self-adaptor
- Actuation
- Human operator

Challenges

- Design time assurance
  - Building blocks/components
  - Closed-loop systems
- Runtime assurance (Gap between models and implemented systems)
- Dealing with anomaly
  - Failures, design defects, security attacks
- Network and system support
  - Edge computing
  - Middleware

Topics for the course

- Modeling and Analysis techniques and tools
  - Verification (e.g., hybrid-system based, abstraction interpretation, falsification)
  - Simulation
  - Testing
- Verification of LEC (Leaning-Enabled Components)
- Gap between design time and runtime
  - Assumptions, coverage, etc.
- Anomaly Detection
- Runtime Verification
- Runtime Assurance
- Human-in-the-loop
  - Interpretable AI
  - Human behavior modeling
- Edge computing (and safety)

Possible Projects

- Develop an anomaly detector and evaluate using data set from Intel
- Develop a self-driving model and implement using AirSim, CARLA, Gazebo, or ROS
- Devise attacks/modules to AirSim or CARLA and evaluate detectors

Summary

- Safe autonomy is an emerging area with significant challenges.
- This course will study foundations, techniques, tools, systems support, and methods for assuring and improving safety of autonomous systems

Assignments

- Presentation signup
  - Looking for early Feb presentations on tools/simulators
- Reading
- Project ideas

End

Approaches

- Modeling and Verification Tools
  - dReach, Flow*, …
- Simulation Tools
  - Airsim, CARLA, Gazebo
- Data-driving modeling
  - Machine learning