Towards Socially Responsible Human-CPS

Lu Feng

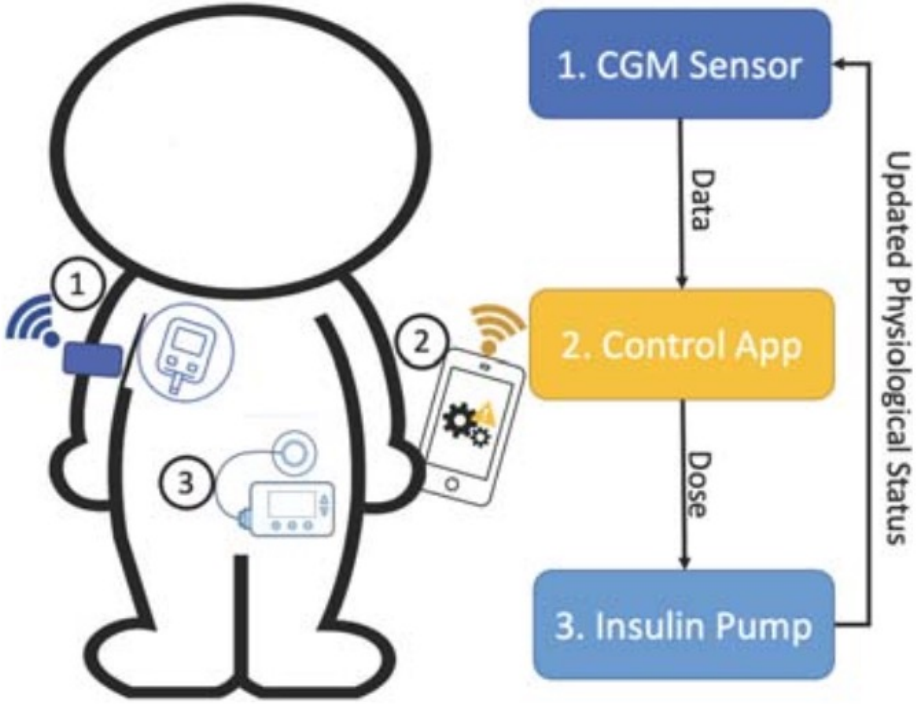
University of Virginia

NSF CPS PI Meeting, 26 January 2023

Research Summary

- Planning for Automated Vehicles with Human Trust [TCPS'22]

- NSF DASS project: Accountable Software Systems for Safety-Critical Applications



What are socially responsible human-CPS?

ACM Statement on Principles for Responsible Algorithmic Systems
(released on Oct 26, 2022)

It is imperative that algorithmic systems comply fully with established legal, ethical, and scientific norms and that the risks of their use be proportional to the specific problems being addressed.

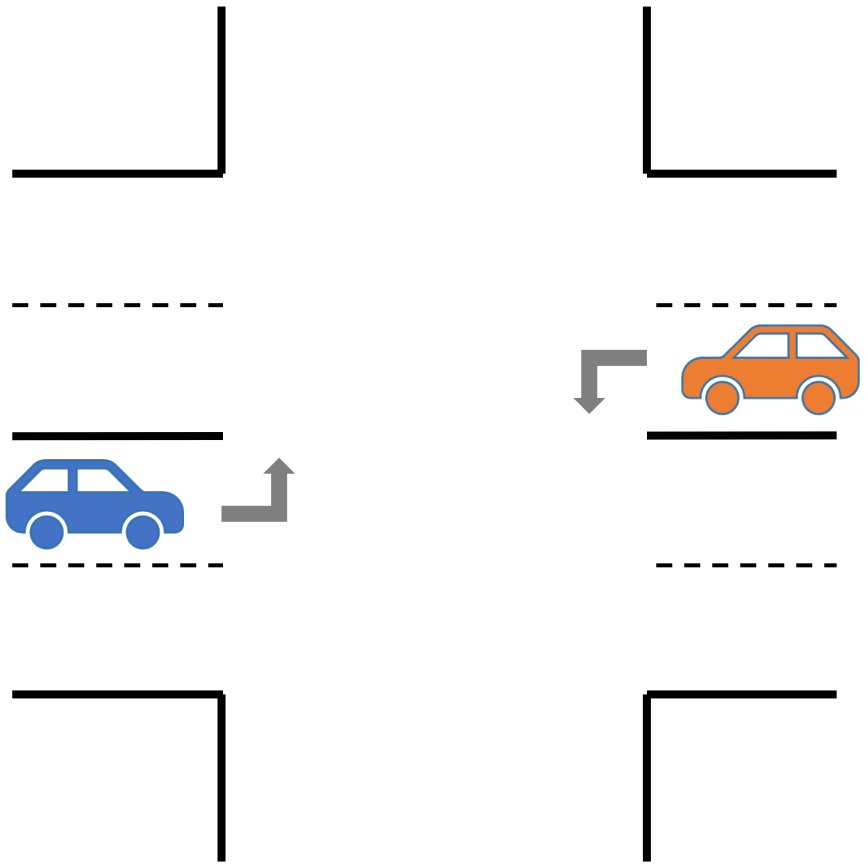
What are socially responsible human-CPS?

[ACM Statement on Principles for Responsible Algorithmic Systems]

1. Legitimacy and competency
2. Minimizing harm
3. Security and privacy
4. Transparency
5. Interpretability and explainability
6. Maintainability
7. Contestability and auditability
8. Accountability and responsibility
9. Limiting environmental impacts

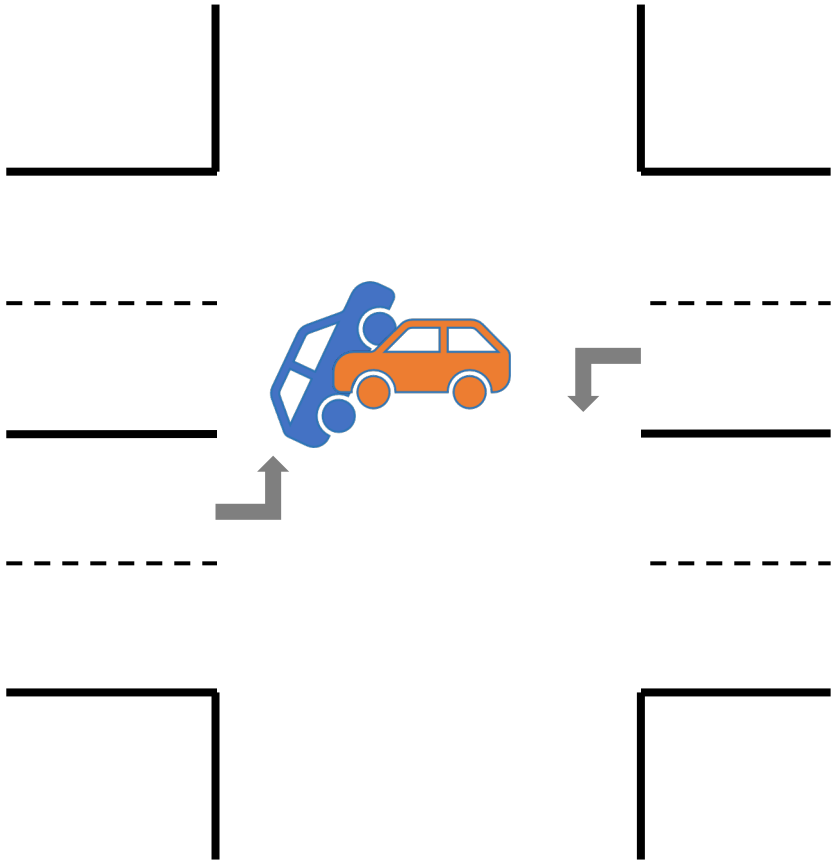
Public and private bodies should be held accountable for decisions made by algorithms they use, even if it is not feasible to explain in detail how those algorithms produced their results. Such bodies should be responsible for entire systems as deployed in their specific contexts, not just for the individual parts that make up a given system.

Example: Which car to blame?



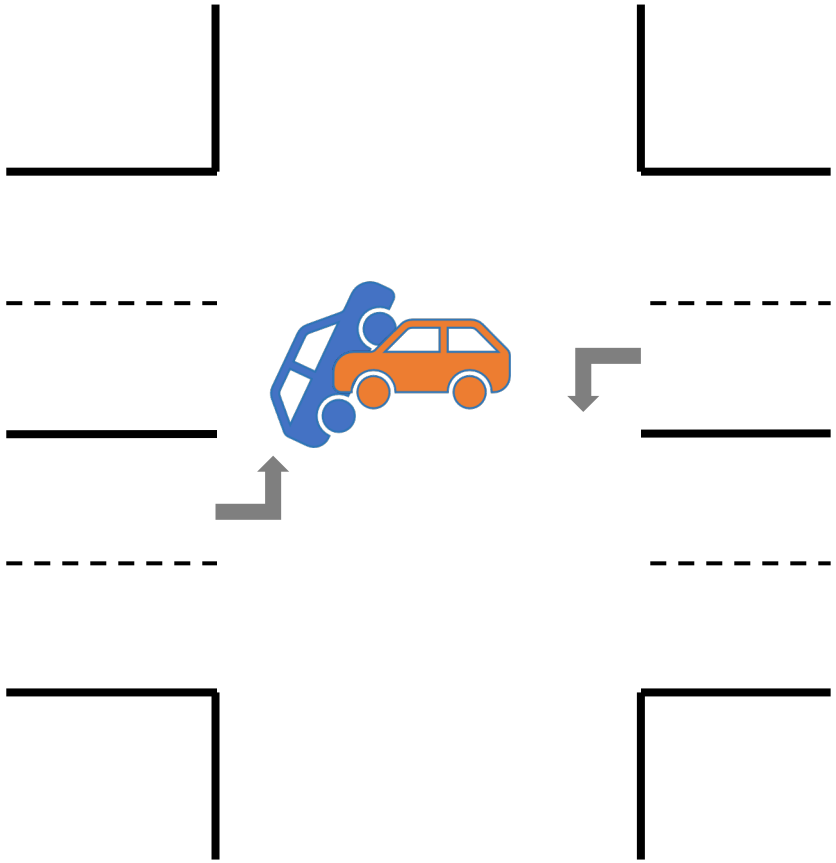
- The blue car entered the left hand turn lane, turned the left turn signal on, and initiated a left turn on a green light.
- At the same time, the orange car approached the intersection in the right turn lane. It was traveling approximately 40 mph in a 25 mph speed zone.

Example: Which car to blame?



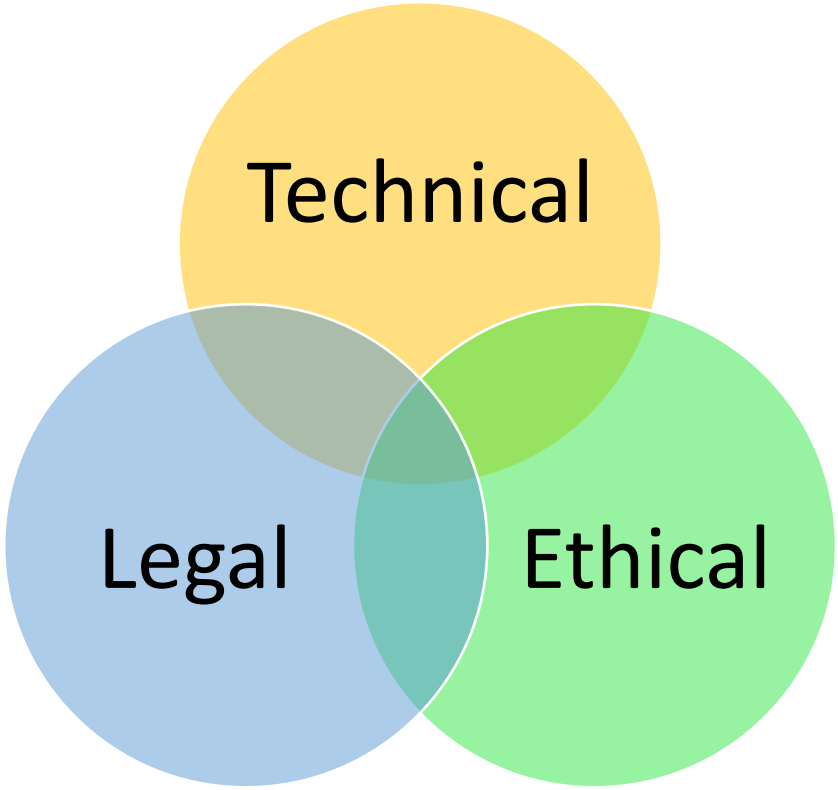
- The blue car entered the left hand turn lane, turned the left turn signal on, and initiated a left turn on a green light.
- At the same time, the orange car approached the intersection in the right turn lane. It was traveling approximately 40 mph in a 25 mph speed zone.
- The blue car came to a stop before fully completing its turn due to the oncoming orange car, and the orange car entered the intersection traveling straight from the turn lane instead of turning.
- Shortly thereafter, the orange car made contact with the rear passenger side of the blue car. The impact caused damage to the right rear door, panel, and wheel of the blue car. The blue car was towed from the scene.

Example: Which car to blame?



- A real accident happened in San Francisco on June 3
  - Blue car: General Motor Cruise AV
  - Orange car: Toyota Pirus
- San Francisco police determined that the Pirus was at fault since it was speeding and in the wrong lane.
- The NHTSA hasn't fully excused Cruise from blame.
  - "[The software could] incorrectly predict another vehicle's path or be insufficiently reactive to the sudden path change of a road user."
- General Motor has recalled self-driving cars after this crash and stated that they've made software changes to avoid this type of accident during future drives.

What advances are necessary to achieve socially responsible human CPS?



AAAI-23 Bridge: AI and Law

Feb 8, Walter E. Washington Convention Center

- Interdisciplinary dialogues between AI and legal researchers
- Discussion Topics:
  - Accountability and Safety
  - Explainability and Transparency
  - Fairness and Non-discrimination
  - Privacy
- Organizers:
  - Bryan Choi, Ohio State University
  - Lu Feng, University of Virginia
  - Sarit Kraus, Bar Ilan University
  - Christopher Yoo, University of Pennsylvania

ai-law-bridge.github.io