

# Principles of Safe Autonomy ECE 484

## Spring 2022 Lecture 1

Professors: Sayan Mitra

Jan 20, 2022

<https://publish.illinois.edu/safe-autonomy/>

<https://mitras.ece.illinois.edu/>

mitras@Illinois.edu



# Welcome from Safe Autonomy team!

## Professor

Sayan Mitra (mitras) (CSL 266)

## TAs

Tianchen Ji (tj12)

Pranav Sriram (psriram2)

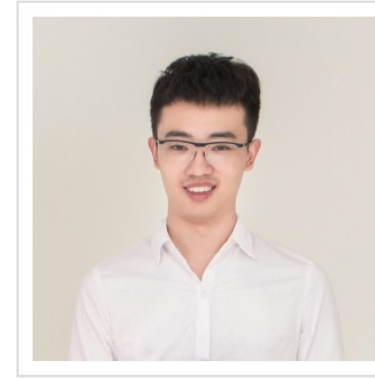
## Lab assistants

Haoyuan You (hy19)

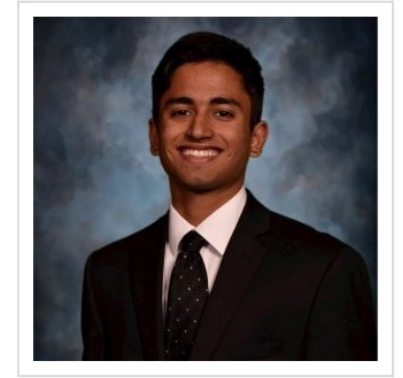
Ninghan Zhong (innghan2)



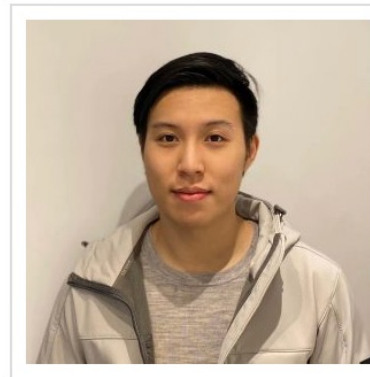
Sayan Mitra (Instructor)



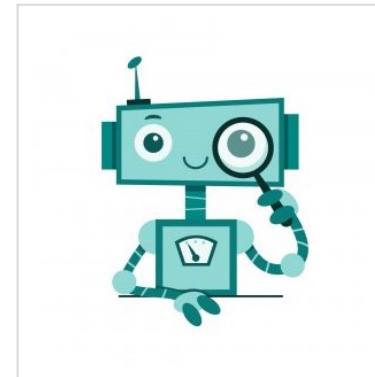
Tianchen Ji (tj12)



Pranav Sriram (psriram2)



Ninghan Zhong (ninghan2)



Haoyuan You (hy19)



# Plan for today

- ▶ What is this course about?
- ▶ Why *Principles* of Safe Autonomy?
- ▶ How will this course work?





## 100 years of *progress in safer roads*

1. Traffic infrastructure (e.g., lane markings, traffic signals, etc)
2. Police enforcement and traffic regulations
3. Driver training
4. Passenger safety (e.g., seatbelts, airbags)
5. Improved vehicle design (e.g., crumple zones)
6. Rear-view and blind spot sensors (e.g., camera)
7. Advanced Driver Assistance Systems (e.g., ABS, ACC, etc.)





# Autonomous systems will be awesome!

Driverless cars will make us more productive

- Average American drives 13,474 miles (300 hrs) per year

Our cities will be greener

- 40% of city surface is parking

Travel and deliveries will be safer

- 32K+ fatalities and 3M+ injuries every year



# PREDICTIONS SCORECARD, 2022 JANUARY 01 by Rodney Brooks

<http://rodneybrooks.com/predictions-scorecard-2022-january-01/>

The years in blue indicate when the industry leaders thought these predictions would come to pass. I have highlighted all the dates up through 2021, now numbering 17 of the 23 predictions. Not one of them has happened or is even close to happening.

## FORECASTS: [http://www.driverless-future.com/?page\\_id=384](http://www.driverless-future.com/?page_id=384) March 27, 2017

NVIDIA to introduce level-4 enabling system by **2018** (2017)

NuTonomy to provide self-driving taxi services in Singapore by **2018**, expand to 10 cities around world by **2020** (2016)

Delphi and MobilEye to provide off-the-shelf self-driving system by **2019** (2016)

Ford CEO announces fully autonomous vehicles for mobility services by **2021** (2016) ←

Volkswagen expects first self driving cars on the market by **2019** (2016)

GM: Autonomous cars could be deployed by **2020** or sooner (2016) ←

BMW to launch autonomous iNext in **2021** (2016) ←

Ford's head of product development: autonomous vehicle on the market by **2020** (2016) ←

Baidu's Chief Scientist expects large number of self-driving cars on the road by **2019** (2016)

First autonomous Toyota to be available in **2020** (2015) ←

Elon Musk now expects first fully autonomous Tesla by **2018**, approved by **2021** (2015)

US Sec Trans: Driverless cars will be in use all over the world by **2025** (2015)

Uber fleet to be driverless by **2030** (2015) ←

Ford CEO expects fully autonomous cars by **2020** (2015) ←

Next generation Audi A8 capable of fully autonomous driving in **2017** (2014)

Jaguar and Land-Rover to provide fully autonomous cars by **2024** says Director of Research and Technology (2014)

Fully autonomous vehicles could be ready by **2025**, predicts Daimler chairman (2014) ←

Nissan to provide fully autonomous vehicles by **2020** (2013) ←

Truly autonomous cars to populate roads by **2028-2032** estimates insurance think tank executive (2013)

Continental to make fully autonomous driving a reality by **2025** (2012)

Hubris on a mass delusion scale. Audi fully autonomous by 2017? That is Teslan in its delusion level.



# Building autonomous cars will be harder than going to the Moon

**Science**  
Google promises autonomous

**The Atlantic**  
TECHNOLOGY  
Google's Self-Driving Cars: 300,000 Miles a Single Accident Under Computer Control  
By REBECCA J. ROSEN AUG 9, 2012  
The automated cars are slowly building a driving record that's better than that of your average American.

**WIRED**  
12.19.17  
After P...  
Enter t...

**ars TECHNICA**  
BIZ & IT TECH SCIENCE POLICY CARS GAMING & CULT  
HYPE CYCLE—  
The hype around self-driving cars is crashing down

**Bloomberg Businessweek**  
December 31, 2018, 6:08 AM CST  
Emergent Tech ▶ Artificial Intelligence  
Self-Driving Tapping  
GM Cruise holds off on self-driving taxis for this year, says it needs more testing time to be safe  
Really? Johnny Cabs aren't that safe yet, eh? Quelle surprise  
By Katyanna Quach 24 Jul 2019 at 23:57 23 SHARE

**MORE STORIES**  
The Quiet Ways Automation Is Remaking Service Work  
SIDNEY FUSSELL  
The Instagram-Husband Revolution  
TAYLOR LORENZ  
Radio Atlantic: How to Fix Social Media  
KEVIN TOWNSEND  
A Border Is Not a Wall  
ALEXIS C. MADRIGAL

Google self-driving car

Some of you will be working on this problem for decades!





# Plan for today

- ▶ What is this course about?
- ▶ *Why Principles of Safe Autonomy?*
- ▶ How will this course work?



Another transportation revolution from a century ago

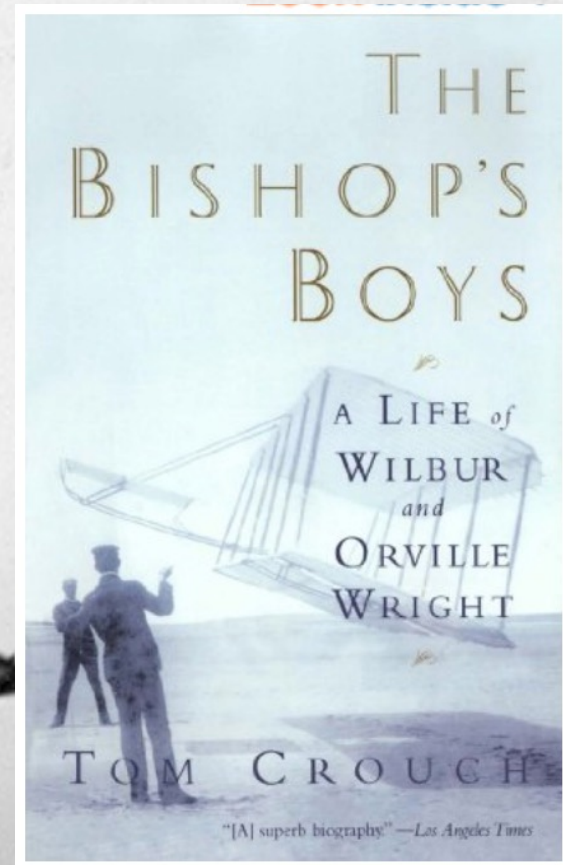
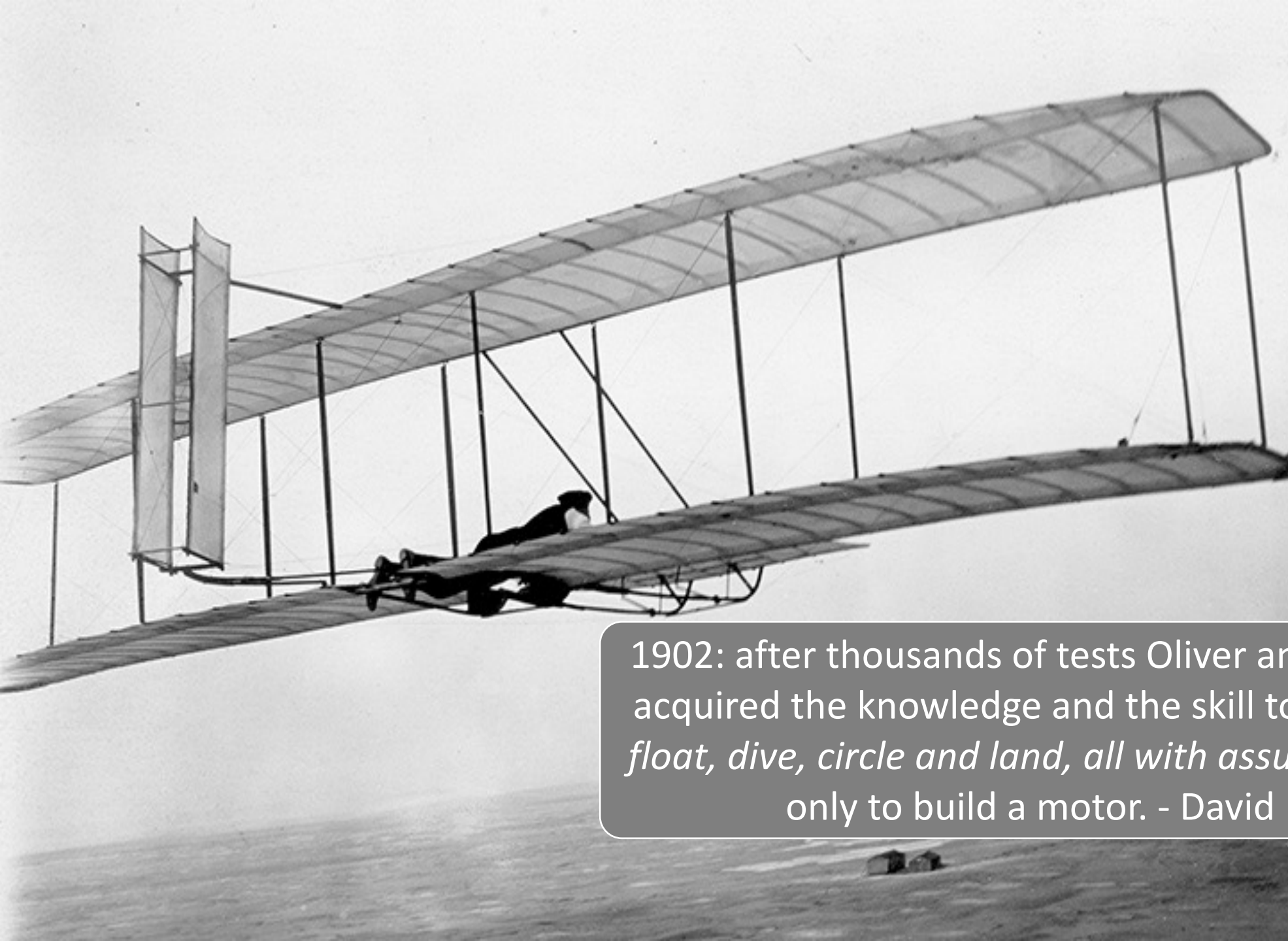


TheFilmGate@aol.com

FIRST  
ROCKET  
MAIL

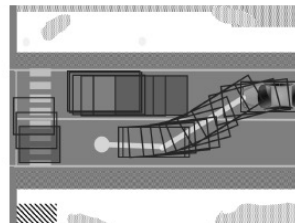
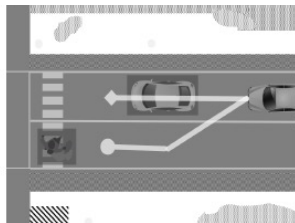
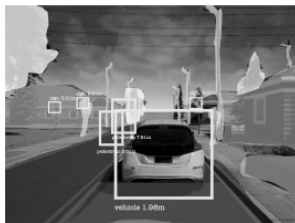
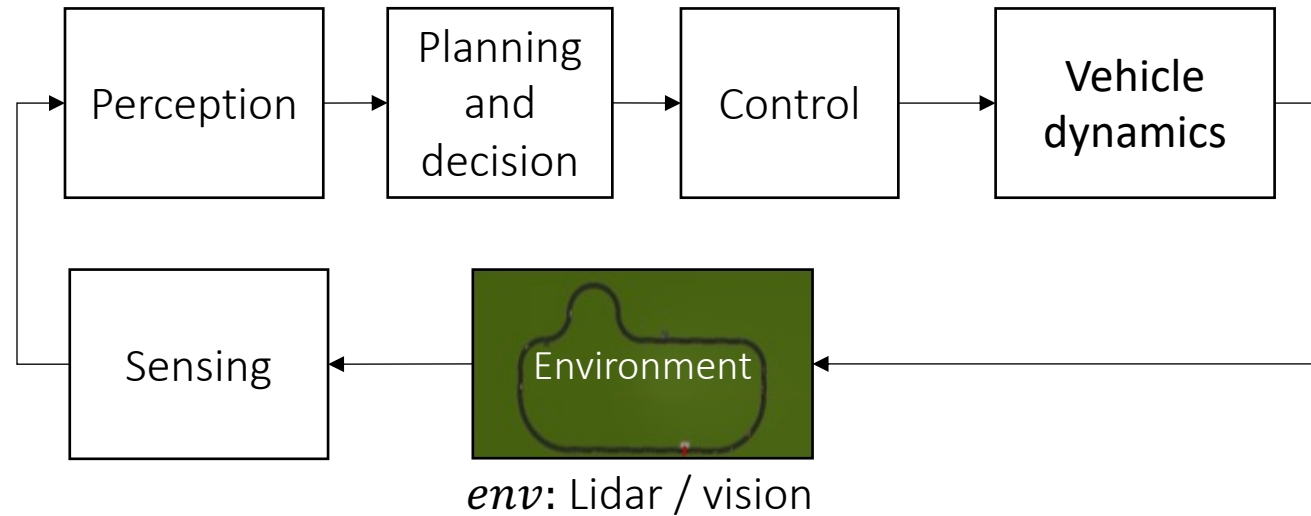
02:47:16:22





1902: after thousands of tests Oliver and Wilbur Wright had acquired the knowledge and the skill to fly. *They could soar, float, dive, circle and land, all with assurance.* Now they had only to build a motor. - David McCullough

# Autonomous GEM vehicle: An example CPS



## Sensing

Physics-based models of cameras, LIDAR, radar, GPS, and so on.

## Perception

Programs for object tracking, scene understanding, and so on.

## Decisions and planning

Programs and multi-agent models of pedestrians, cars, and so on.

## Control

Dynamical models of vehicle engine, powertrain, steering, tires, and so on.





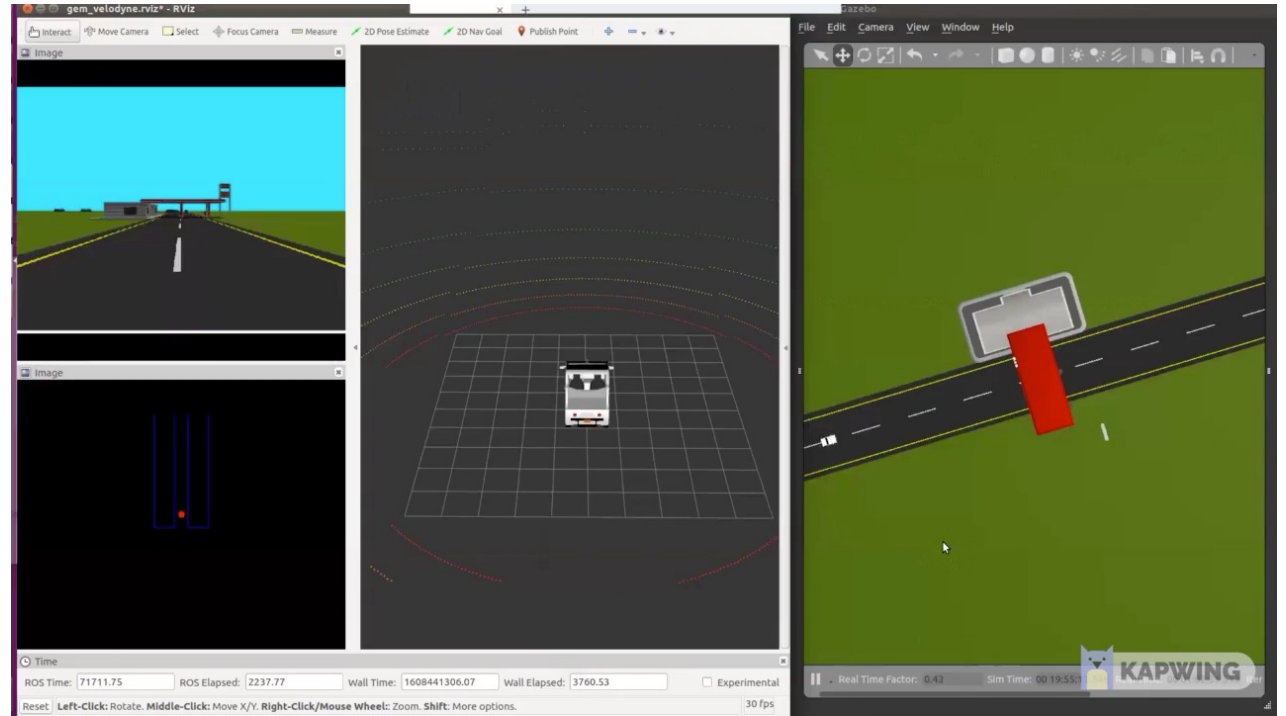




# Open problem (project B)

Simulated race car following a track with Lidar-based perception and control.

**Problem:** For a given track and initial conditions check that the *trajectory* of the car does not collide and stays in lane.

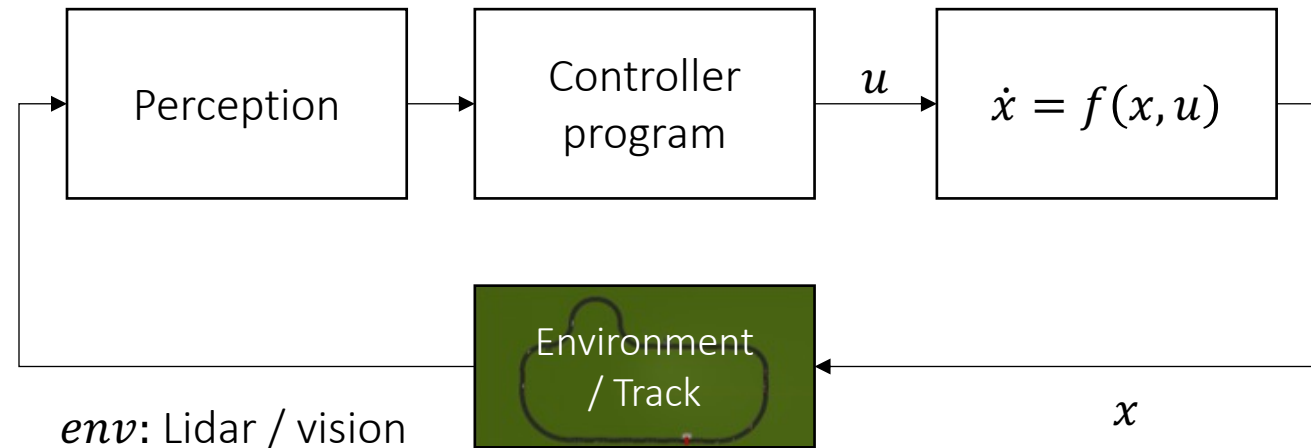


Can we check *efficiently*?

Can we *generalize to similar* tracks?

What should we assume about perception, accuracy of the vehicle model?

What should we assume about the execution of the controller?



# A disagreeable position about autonomy

“Collect lots of data, build faster computers, and train larger neural networks, and safe autonomous vehicles will follow” --- A founder

How to assure safety of an AV ? Run tests

In an absolute sense:

“Testing can be used to show the presence of bugs, but never to show their absence!” --- Edsger W. Dijkstra

Because there are infinitely many *executions* and we can only test finitely many of those in any testing algorithm

In a probabilistic sense also, purely using data to gain safety assurance is not practical



# Naively collecting test driving miles is also not going to work

Probability of a fatality caused by an accident per one hour of human driving is known to be  $10^{-6}$

Assume\* that for AV this has to be  $10^{-9}$

Data required to guarantee a probability of  $10^{-9}$  fatality per hour of driving is proportional to its inverse,  $10^9$  hours, 30 billion miles

Multi-agent, open system, with human interactions => cannot be simulated offline to generate data

Any change in software means tests have to be rerun

To learn or extrapolate about all---infinitely many---executions from a finite sampling of executions, we need to make some assumptions about the system. A collection of these assumptions defines a model

Different types of model (and data) for sensing, control, planning, and we need to understand how to analyze and compose them

*On a Formal Model of Safe and Scalable Self-driving Cars by  
Shai Shalev-Shwartz, Shaked Shammah, Amnon Shashua, 2017  
(Responsibility Sensitive Safety)*



# Plan for today

- ▶ What is this course about?
- ▶ Why *Principles* of Safe Autonomy?
- ▶ How will this course work?



# Why are we here? Course goals

Know



Components of an autonomous system , safety standards, ...

How to use software modules for perception, planning, control, ROS, Yolo, OpenCV, Z3, ...

Do



Code and analyze algorithms for perception clustering, convolution, filtering, edge detection, filtering, localization, planning, formal verification

Plan, propose, organize and execute a team project

Understand



Models, algorithms, data, biases, assumptions for building trustworthy autonomous systems

Theoretical properties of algorithms and their limitations

Get inspired







Become the Isaac Newton of Autonomy

“To do things right, first you need love, then technique.” – Antoni Gaudí



# Why are we here? Course goals

Know		Lectures, MPs, Homework
Do		MPs, Homework, Project, Participate
Understand		Lectures, MP, Homework, Exams
Get inspired		Project

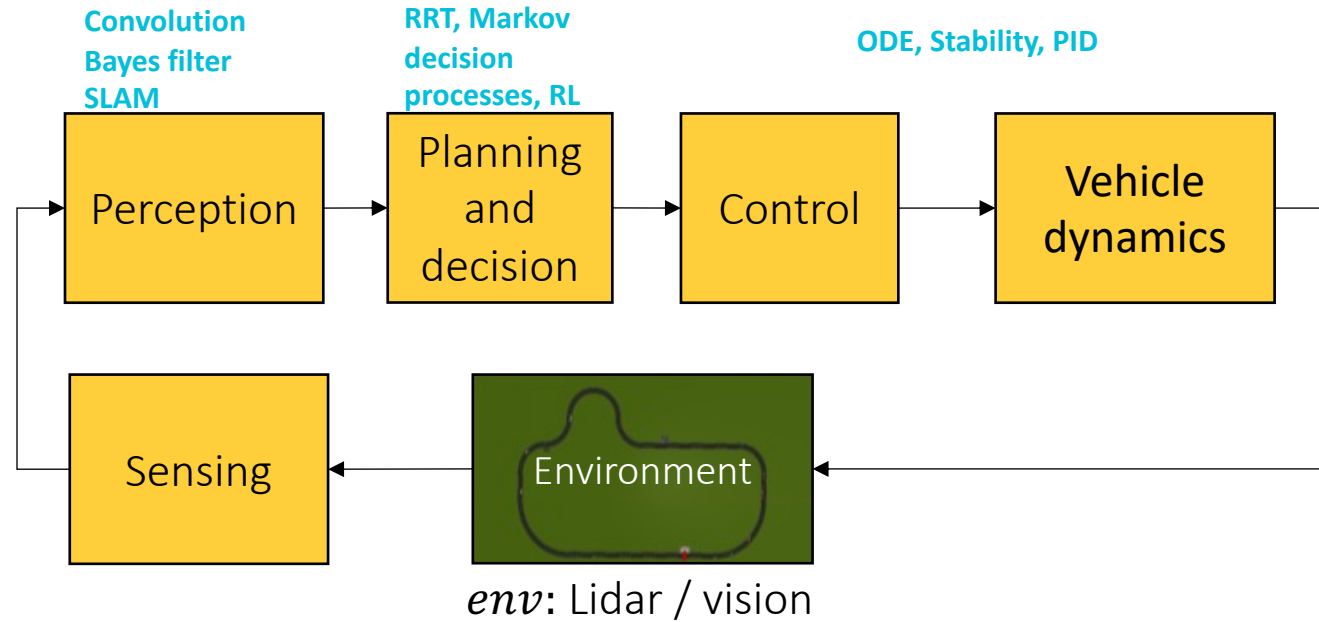




# Course structure



Safety, end-to-end testing, simulation, system integration  
 MPO, MP5, **Project**  
 State machines, model checking, hypothesis testing, ROS



Sensing	Perception	Decisions and planning	Control
MP3 Project	MP1 MP3 Project	MP4 Project	MP2 Project



# About the course

Everything starts here: <https://publish.illinois.edu/safe-autonomy/>

- ▶ Schedule, lab, resources, papers, homework, MP, code, project, gitlab links

Campuswire for announcements, but no SLA

- ▶ Discussions, forming teams, occasional polls, feedback
- ▶ Do not expect to get answers to HWs, MPs, and exam related questions in the last minute.

Canvass for MP release and grades



# Schedule: <https://publish.illinois.edu/safe-autonomy/schedule-spring-2022/>

- Simple safety (MP0, 1 week)
  - Perception (MP1 lane detection, 2 weeks)
  - Modeling and control (MP2 vehicle control, 2 weeks)
  - Filtering: localization particle filtering (MP3 localization, 2 weeks)
  - March 1: Midterm 1
  - Planning (MP4 planning, 2 weeks)
  - Spring Break
  - Decision making RL (MP5 racing, end-to-end safety)
  - End-to-end system and safety analysis
  - April 14: Midterm 2
- Group formation
  - Lab safety training
  - Labs, MPs
  - Spring Break
  - Project pitch
  - Intermediate checking
  - Practice presentation
  - Final presentation
  - Report

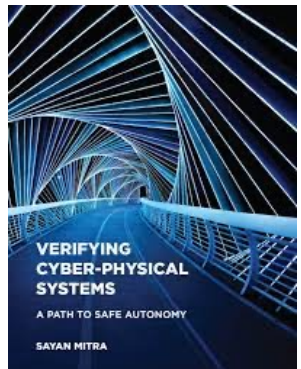
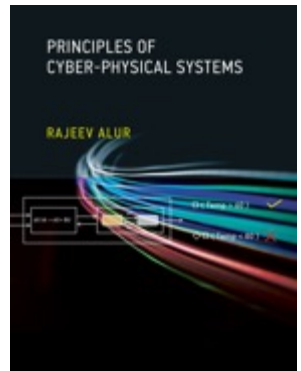
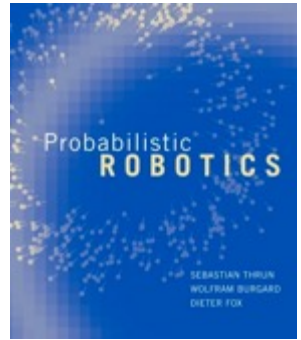


# Course materials

Lecture notes, slides, code, video lectures, lab manuals created and curated from recent research publications

## Reference books:

- ▶ *Probabilistic robotics*, By Sebastian Thrun, Wolfram Burgard and Dieter Fox, 2005
- ▶ *Principles of Cyber-Physical Systems*, Rajeev Alur, MIT Press, 2015
- ▶ *Verifying Cyber-physical Systems*, Sayan Mitra, MIT Press 2021\*  
(if you are interested in safety verification)





# Course: components and (tentative) weights

- ▶ 5-6 programming assignments or MPs 45% (group)
  - ▶ ROS + Python, Ubuntu, VM BYOD or use lab workstations
  - ▶ labs (Friday 4-8 pm starting tomorrow)
  - ▶ Office hours
- ▶ Homework assignments 10% (individual)
  - ▶ math, analysis, critical reasoning; preparation for midterms
- ▶ Midterms x2 20% (individual)
- ▶ Mini project 20% (group): more on this later, 2 tracks:
  - ▶ A. Dev and test concepts on GEM
  - ▶ B. [GRAIC autonomous racing competition / testing](#)
- ▶ Participation 5%
  - ▶ Lab and class attendance and participation, early feedback on notes, solve exercise problems in class notes, class participation, MP beta testers

Tentative grade boundaries	
A	>90
B	>80
C	>65
D	>55



# Homework, participation, & exam: Individual work 35%

- ▶ Testing principles, concepts
- ▶ Read course notes and slides routinely; exercises are provided
- ▶ Homework sets (synchronized with MPs)
- ▶ 2 in-class midterms **March 1, April 14**
- ▶ No final exam



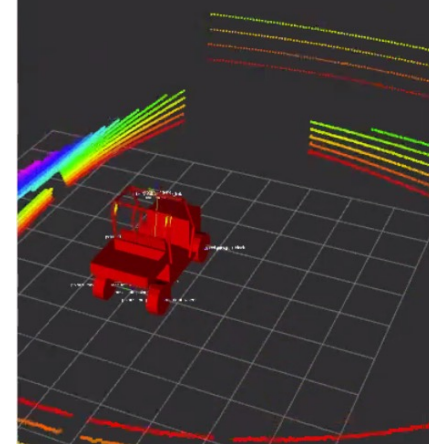
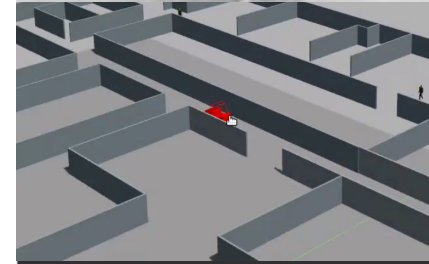
# Teamwork: MP, labs, and mini project

- ▶ In groups: Form your group of 3-4 now! [Make your group](#) (form), make new friends
  - ▶ If you do not have a group by **Sunday (12 midnight AoE)** we will assign you a group
- ▶ Each MP will build a significant component of an autonomous system over **2 weeks**
- ▶ Use our VM or your own computer with Ubuntu 16.04 or broadband internet connection to remote login to lab computers
- ▶ TAs and LAs will run live labs in ECEB 5072 (first one on zoom)
  - ▶ MP walkthrough, setup, bridge the lecture and the assignments
- ▶ **MPO+HW0 will be release this Friday (1/21), labs starts this Friday**
  - ▶ Your entire group has to attend 1 lab after each new MP is released,
  - ▶ And 1 lab after the MP is due (to demo your work). For the other weeks, at least one representative from the group should discuss progress with MPs and projects



# Mini projects: explore, inspire, and impress

- ▶ GEM Track. Build on existing SW to live demo at Highbay
  - ▶ Reliable parallel parking, lane following and pedestrian avoidance outdoor track, biker intent estimation and reaction
  - ▶ **COVID related uncertainty about highbay access; Frontload the work**
- ▶ GRAIC Track. Participate in an open simulation-based autonomous racing competition
  - ▶ Build your own driver software, test it in a number of environments
  - ▶ Test, verify, find bugs in other people's driver software
  - ▶ Automatically create complex scenarios
- ▶ Outcomes: Write research papers, jumpstart grad research, career in autonomy, incubate startup ideas, sharpen presentation skills
- ▶ We provide: Polaris GEM vehicle (camera, LIDAR, RADAR, IMU, GPS, and drive-by-wire system) modules for pedestrian detection, lane tracking, and vehicle control, a vehicle simulator, and testing facility (highbay) with indoor positioning system. GRAIC autonomy software stack
- ▶ Expertise (TA, lab and office hours, TBD)
- ▶ Timeline: [Get started, be a member of IRL from this link](#)
  - ▶ High-bay virtual site visit and training (in next 2 weeks)
  - ▶ Project pitch after Spring break
  - ▶ **Public presentation, demo, awards (End of April or May 6)**



[Spring 2020 projects](#)

[Fall 2020 projects](#)



# Summary

- ▶ Form your team. Decide track. Sign-up to be member of IRL
- ▶ Careful modeling and reasoning can expose flawed assumptions, bad design bugs, make the system explainable
- ▶ Discrete time model: states, initial states, transition function
- ▶ Requirements, e.g., safety

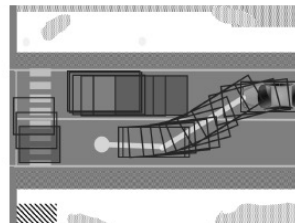
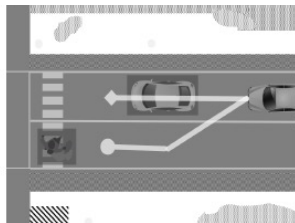
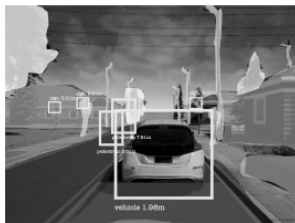
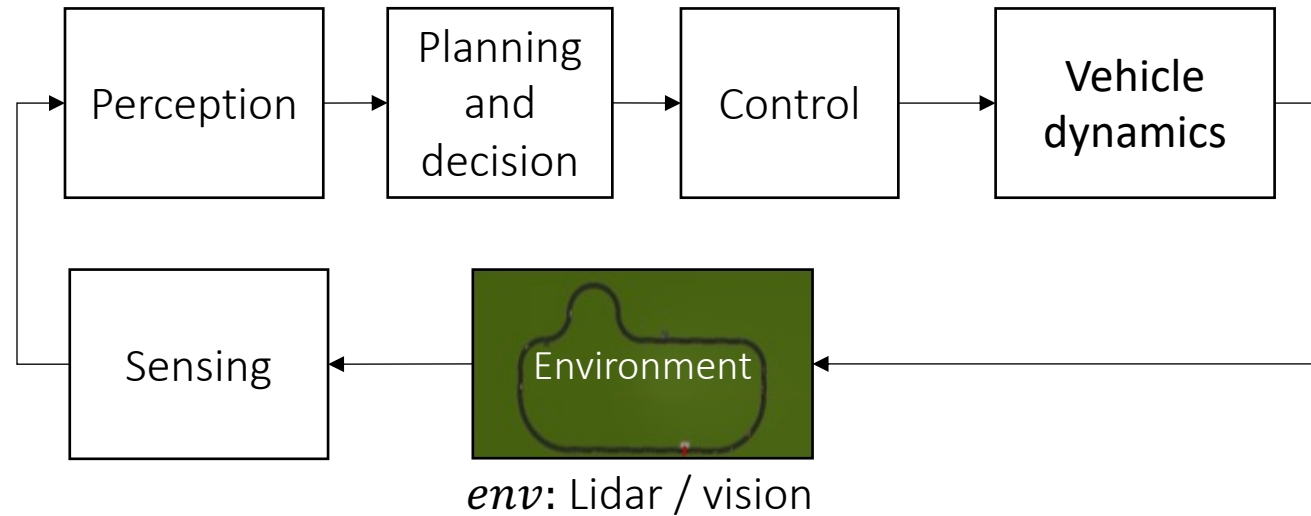




# Absolute safety analysis



# Autonomous GEM vehicle: An example CPS



## Sensing

Physics-based models of cameras, LIDAR, radar, GPS, and so on.

## Perception

Programs for object tracking, scene understanding, and so on.

## Decisions and planning

Programs and multi-agent models of pedestrians, cars, and so on.

## Control

Dynamical models of vehicle engine, powertrain, steering, tires, and so on.

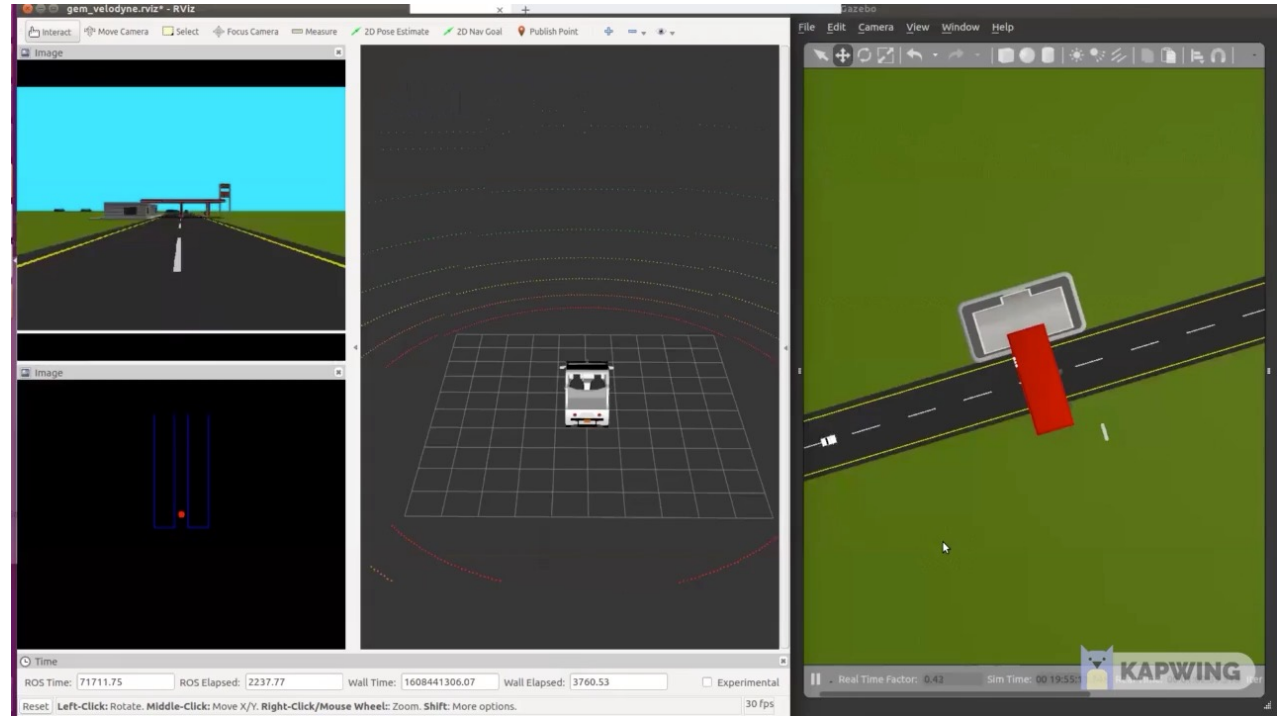




# Open problem (project B)

Simulated race car following a track with Lidar-based perception and control.

**Problem:** For a given track and initial conditions check that the *trajectory* of the car does not collide and stays in lane.

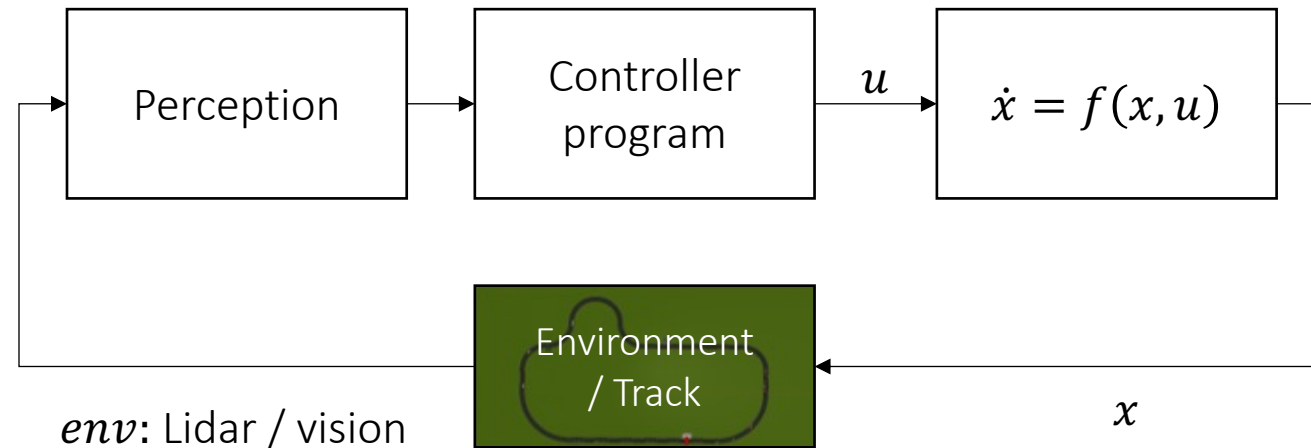


Can we check *efficiently*?

Can we *generalize to similar* tracks?

What should we assume about perception, accuracy of the vehicle model?

What should we assume about the execution of the controller?



# A disagreeable position about autonomy

“Collect lots of data, build faster computers, and train larger neural networks, and safe autonomous vehicles will follow” --- A founder

How to assure safety of an AV ? Run tests

In an absolute sense:

“Testing can be used to show the presence of bugs, but never to show their absence!” --- Edsger W. Dijkstra

Because there are infinitely many *executions* and we can only test finitely many of those in any testing algorithm

In a probabilistic sense also, purely using data to gain safety assurance is not practical





# Naively collecting test driving miles is also not going to work

Probability of a fatality caused by an accident per one hour of human driving is known to be  $10^{-6}$

Assume\* that for AV this has to be  $10^{-9}$

Data required to guarantee a probability of  $10^{-9}$  fatality per hour of driving is proportional to its inverse,  $10^9$  hours, 30 billion miles

Multi-agent, open system, with human interactions => cannot be simulated offline to generate data

Any change in software means tests have to be rerun

To learn or extrapolate about all---infinitely many---executions from a finite sampling of executions, we need to make some assumptions about the system. A collection of these assumptions defines a *model*

*On a Formal Model of Safe and Scalable Self-driving Cars by  
Shai Shalev-Shwartz, Shaked Shammah, Amnon Shashua, 2017  
(Responsibility Sensitive Safety)*

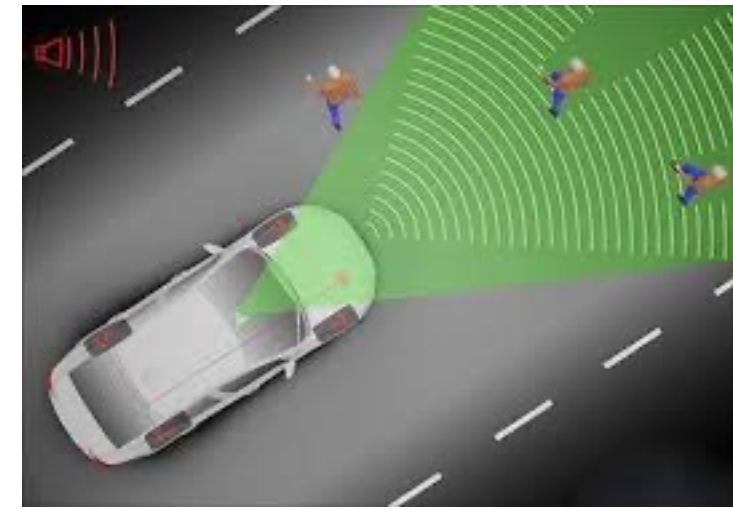
# Today: How can we use a simple mode to get absolute safety guarantees

- ▶ A simple class of models: automata
- ▶ What are executions of automata?
- ▶ What are safety requirements?
- ▶ Reachable states, Invariants for safety guarantees



# A “simple” safety scenario

A car moving down a straight road has to detect any pedestrian (or another car) in front of it and stop before it collides.



## Automatic Emergency Braking

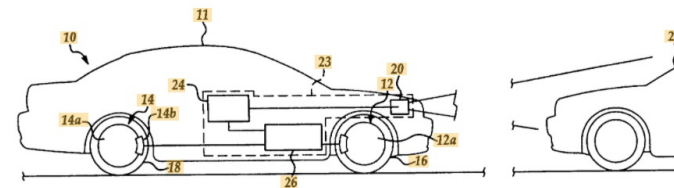


Figure 1

Not a trivial requirement

Today: There is no enforced standard for testing AEB

[www.google.com > patents](#)

[US20110168504A1 - Emergency braking system - Google ...](#)

Jump to **Patent citations** (18) - US4053026A \* 1975-12-09 1977-10-11 Nissan Motor Co., Ltd. Logic circuit for an automatic braking system for a motor ...

[www.google.com > patents](#)

[US5170858A - Automatic braking apparatus with ultrasonic ...](#)

An automatic braking apparatus includes: an ultrasonic wave emitter provided in a ... Info: Patent citations (13); Cited by (7); Legal events; Similar documents; Priority and ... US6523912B1 2003-02-25 Autonomous emergency braking system.

[www.google.com > patents](#)

[DE102004030994A1 - Brake assistant for motor vehicles ...](#)

B60T7/22 Brake-action initiating means for automatic initiation; for initiation not ... Info: Patent citations (3); Cited by (9); Legal events; Similar documents ... data from the environment sensor and then automatically initiates emergency braking.

[www.google.com.pg > patents](#)

[Braking control system for vehicle - Google Patents](#)

An automatic emergency braking system for a vehicle includes a forward viewing camera and a control. At least in part responsive to processing of captured ...

[www.automotiveworld.com > news-releases > toyota-ip... >](#)

[Toyota IP Solutions and IUPUI issue first commercial license ...](#)

Jul 22, 2020 - ... and validation of automotive automatic emergency braking (AEB) ... and Director of Patent Licensing for Toyota Motor North America. "We are ...

[insurancenewsnet.com > oarticle > patent-application-tit... >](#)

[Patent Application Titled "Multiple-Stage Collision Avoidance ...](#)

Apr 3, 2019 - No assignee for this patent application has been made. ... Automatic emergency braking systems will similarly, also, soon be required for tractor ...



“simple” ≠ Easy



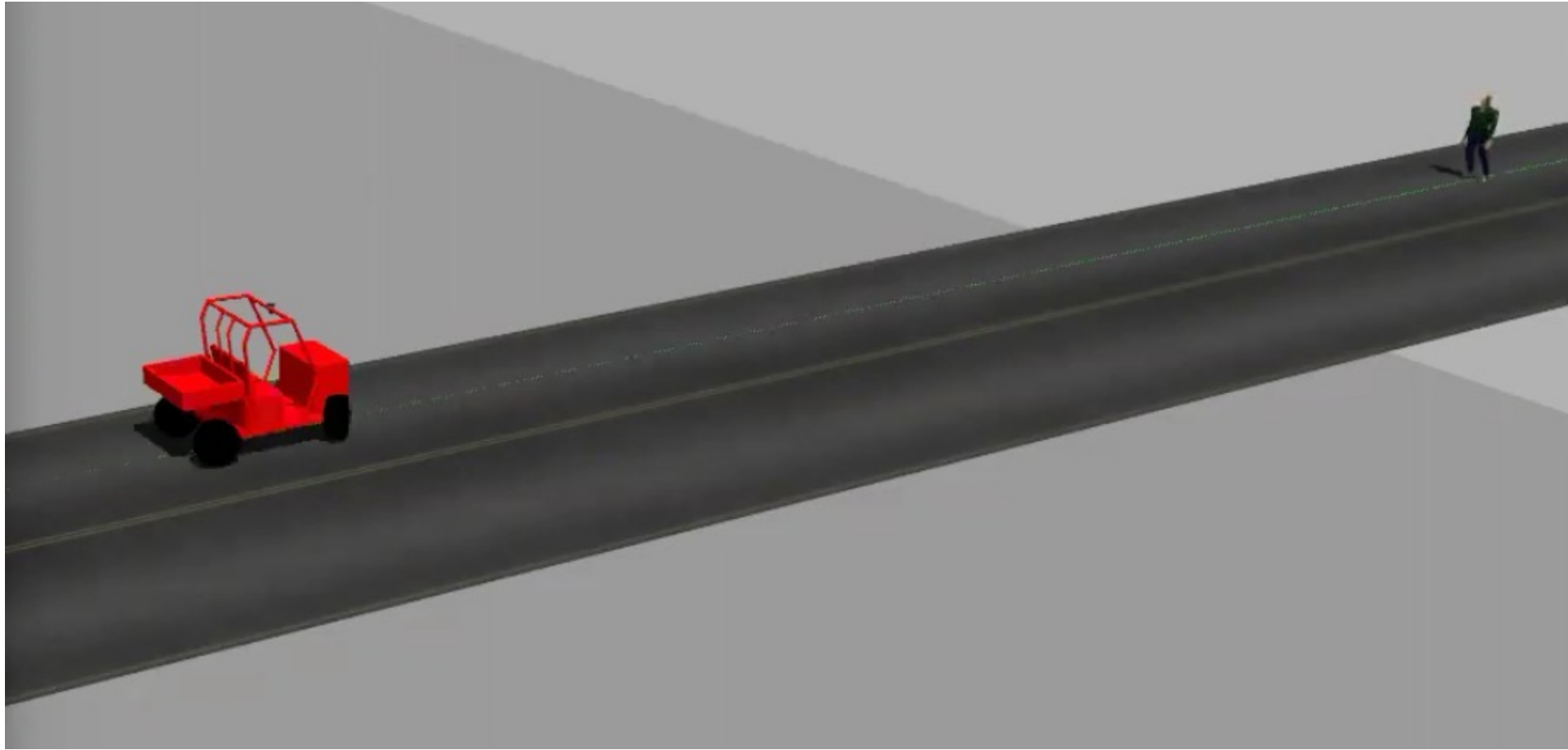
# Modeling the scenario

- ▶ What is a model of a system?
- ▶ A *mathematical model* describes how a system behaves.
  - ▶ What are the key parameters and states?
  - ▶ How are the parameters selected by nature?
  - ▶ What are the initial conditions of the state?
  - ▶ How do the state change over time? ...
  - ▶ What parts of the model are available for observation/analysis?
- ▶ Models include the implicit and explicit assumptions (biases) we are making about the system





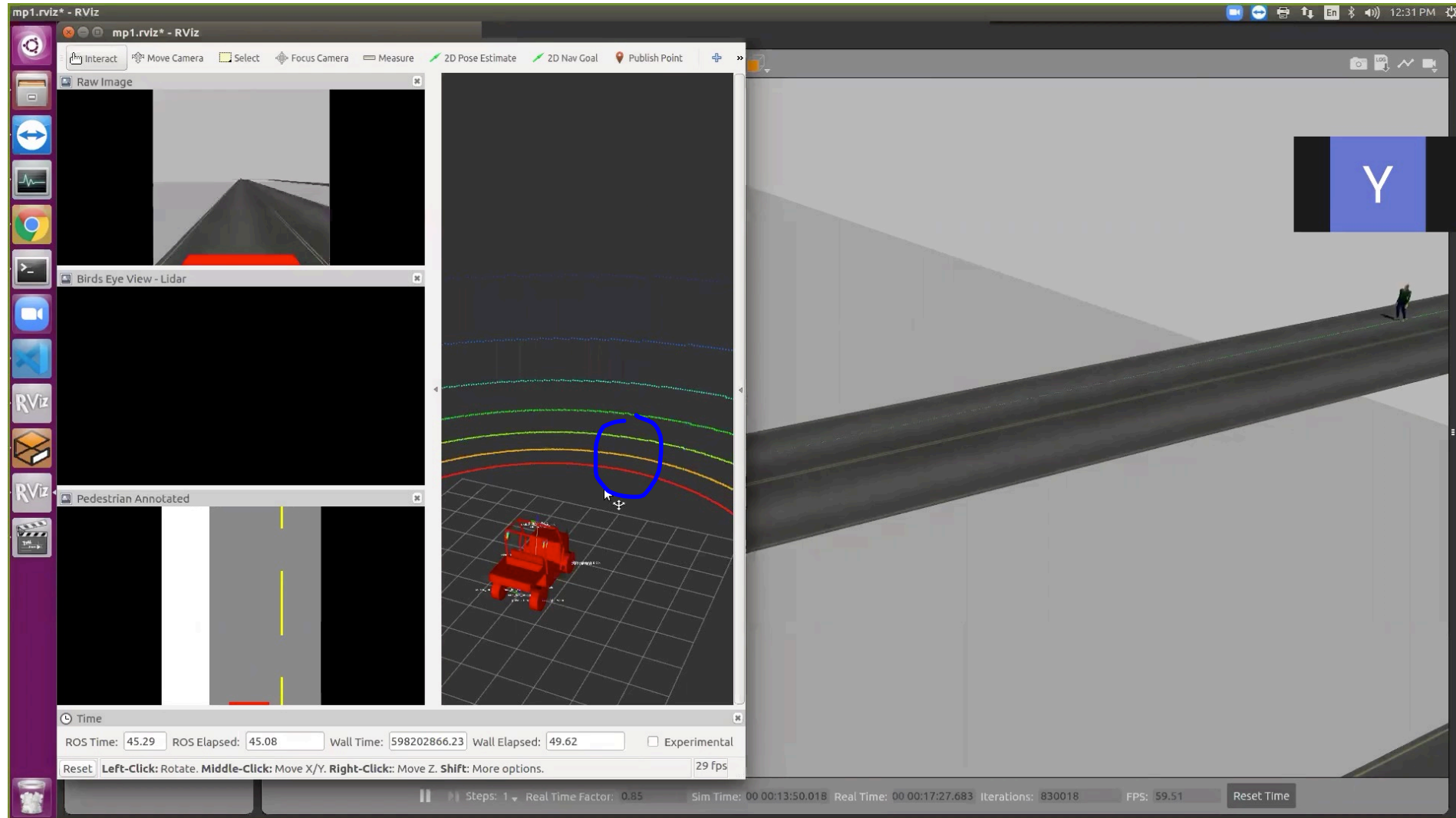
# Model (switch to notes)



“All models are wrong, some are useful.”



# MPO: Simulate model for testing



# An honest scientific approach

1. Create detailed *mathematical models* of the autonomous systems and its environment
  2. Enumerate the precise *requirements* of the system and the conditions on the environment under which it is supposed to work
  3. Analyze the system to either
    - ▶ *prove* that all behaviors meet the requirement (perhaps with high probability)
    - ▶ find counter-examples, corner cases, etc., debug and repeat
- ▶ Currently there are fundamental flaws in making this work for autonomous systems
  - ▶ **Why study this approach?**
    - ▶ Careful reasoning can expose flawed assumptions, bad design choices
    - ▶ The approach has been successful in other industries: microprocessors, aviation, cloud computing, nuclear, ..
    - ▶ Big strides in the last few years
    - ▶ Working deliberately towards a more perfect understanding is a worthwhile intellectual struggle



# Plan for today

- ▶ What is this course about?
- ▶ Why is this approach to autonomous systems important?
- ▶ How will this course work? (Administrivia)

