Goal of the Course

ADS (autonomous driving system)

Challenges of Safe ML

ML

Safety

Assuring Safety of ML components through testing, verification or synthesis (safe by construction)

What is Safety of ML components?
The Dream of Self-Driving
Lecture plan

• Levels of automation of ADS
• Current big players in ADS
• Functional reference architecture for ADS
  • And a look at some sensing technology
• A1 vs A2 autonomy
• A look at several ADS accidents
• Safety assurance of ADS
• ADS challenges
SAE J3016 Levels of Automation

**Driver**
- **0**: Warning, Emergency
- **1**: Driver Assistance
- **2**: Partial Automation
- **3**: Conditional Automation
- **4**: High Automation
- **5**: Full Automation

**Example**
- **Lane Departure Warning**
- **Adaptive Cruise Control**
- **Tesla’s Autopilot**
- **ADS for stop-and-go**
- **Shuttle in geo fenced area**
- **Robo Taxi anywhere**

**Steering & Accel**
- N/A
- N/A
- N/A

**OEDR**
- Limited
- Limited
- Any time, anywhere
SAE J3016 Levels of Automation

Object Event Detection and Recognition

<table>
<thead>
<tr>
<th>Level</th>
<th>Driver Assistance</th>
<th>Adaptive Cruise Control</th>
<th>Autopilot</th>
<th>High Automation</th>
<th>Full Automation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Warning, Emergency</td>
<td>Driver Assistance</td>
<td>Warning</td>
<td>High Automation</td>
<td>Full Automation</td>
</tr>
<tr>
<td>1</td>
<td>Lane Departure Warning</td>
<td>Adaptive Cruise Control</td>
<td>Autopilot</td>
<td>High Automation</td>
<td>Full Automation</td>
</tr>
<tr>
<td>2</td>
<td>Steering &amp; Acceleration</td>
<td>Adaptive Cruise Control</td>
<td>Autopilot</td>
<td>High Automation</td>
<td>Full Automation</td>
</tr>
<tr>
<td>3</td>
<td>Object Event Detection and Recognition</td>
<td>Autopilot</td>
<td>Autopilot</td>
<td>High Automation</td>
<td>Full Automation</td>
</tr>
<tr>
<td>4</td>
<td>Partial Automation</td>
<td>Autopilot</td>
<td>stop-and-go</td>
<td>High Automation</td>
<td>Full Automation</td>
</tr>
<tr>
<td>5</td>
<td>Conditional Automation</td>
<td>Autopilot</td>
<td>stop-and-go</td>
<td>High Automation</td>
<td>Full Automation</td>
</tr>
</tbody>
</table>

Example

- Steering & Acceleration
- Object Event Detection and Recognition
- High Automation
- Full Automation

ODR

- Limited
- Limited
- Any time, anywhere

ODD

- Limited
- Limited
- Any time, anywhere

University of Toronto, CSC2125, Lecture 1: ADS
SAE J3016 Levels of Automation vs. Operational Design Domain (ODD)

Some ODD parameters:
- Speed
- Geography
- Roadway
- Environment

Level 2 example:
- Roadway == expressway
- Speed <= 35mph
- Daytime only

Level 4 example:
- Roadway == campus roads
- Speed <= 25mph
- Daytime only

Unlimited ODD
Beyond Traditional Levels: Two types of AI

- **Starting point:**
  - All cars are manually controlled until the AI system shows itself to be available and is elected to be turned on by the human.

- **A1: Human-Centered Autonomy**
  - **Definition:** AI is not fully responsible
  - Feature axis:
    - Where/how often is it “available”? (traffic, highway, sensor-based, etc.)
    - How many seconds for take-over? (0, 1, 10, etc)
    - Teleoperation support

- **A2: Full Autonomy**
  - **Definition:** AI is fully responsible
  - Notes:
    - No teleoperation
    - No 10-second rule: It’s allowed to ask for human help, but not guaranteed to ever receive it.
    - Arrive to a safe destination or safe harbor.
    - Allow the human to take over when they choose to.
Beyond Traditional Levels: Two types of AI

L0  • Starting point:
   • All cars are manually controlled until the AI system shows itself to be available and is elected to be turned on by the human.

L1, L2, L3  • A1: Human-Centered Autonomy
   • Definition: AI is not fully responsible

L4, L5  • A2: Full Autonomy
   • Definition: AI is fully responsible
Example Players
Notable:

- April 2017: Exits testing: first rider in Phoenix
- November 2017: 4 million miles driven autonomously
- December 2017: No safety driver in Phoenix
Uber

Notable:

• December 2017: 2 million miles driven autonomously
Notable:

• Sep 2014: Released Autopilot
• Oct 2016: Started Autopilot 2 from scratch.
• Jan 2018: ~1 billion miles driven in Autopilot
• Jan 2018: ~300,000 Autopilot equipped vehicles
Audi A8 (released end of 2018)

- Thorsten Leonhardt, head of Automated Driving, Audi: “When the function is operated as intended, if the customer turns the traffic jam pilot on and uses it as intended, and the car was in control at the time of the accident, the driver goes to his insurance company and the insurance company will compensate the victims of the accident and in the aftermath they come to us and we have to pay them,” he said.
Notable Progress

- Full autonomy (A2)
  - Waymo
  - Uber
  - GM Cruise
  - nuTonomy
  - OptimusRide
  - Zenuity
  - Voyage
  - …

- Human-centered autonomy (A1)
  - Tesla Autopilot - Model S/3/X
  - Volvo PilotAssist - S90/XC90/XC60/V90
  - Audi Traffic Jam Assist - A8
  - Mercedes-Benz Drive Pilot Assist - E-Class
  - Cadillac Super Cruise - CT6
  - Comma.ai openpilot
  - …
Paths to Autonomous Future

**A1:** Human-Centered Autonomy

- **Localization and Mapping:** Where am I?
- **Scene Understanding:** Where/who/what/why of everyone else?
- **Movement Planning:** How do I get from A to B?
- **Human-Robot Interaction:** What is the physical and mental state of the driver?
- **Communicate:** How do I convey intent to the driver and to the world?

**A2:** Full Autonomy

- **Localization and Mapping:** Where am I?
- **Scene Understanding:** Where/who/what/why of everyone else?
- **Movement Planning:** How do I get from A to B?
- **Human-Robot Interaction:** What is the physical and mental state of the driver?
- **Communicate:** How do I convey intent to the driver and to the world?
What does it take to drive a car?

1. Perception

2. Decision making

3. Control
Self-Driving Car Tasks

• Localization and Mapping – Where am I
• Scene Understanding – Where is Everyone Else?
• Movement Planning – How to get from Point A to Point B
• Driver State – What is the Driver Up to?
  • Essential if driver is part of the loop!
• Safety Monitoring
Functional Reference Architecture

Source: Krzysztof Czarnecki, Waterloo
Functional Reference Architecture

- Sensor Input
- Ego Perception
- Static Environmental Perception
- Dynamic Environmental Perception
- Mission Execution
- Cloud Data

University of Toronto, CSC2125, Lecture 1: ADS
1. Perception

- Sensor Input
- Cloud Data
- Ego Perception
- Static Environmental Perception
- Dynamic Environmental Perception
- Mission Execution
- System Supervisor
- Vehicle Actuator Command
1. Perception

2. Decision making

3. Control

Functional Reference Architecture

System Supervisor

Sensor Input

Ego Perception

Static Environmental Perception

Dynamic Environmental Perception

Mission

Cloud Data

Vehicle Actuator Command

University of Toronto, CSC2125, Lecture 1: ADS
Functional Reference Architecture

System Supervisor

Sensor Input
- GPS
- IMU
- Wheel Odometry
- Cameras
- LIDAR
- RADAR

Ego Perception
- Static Environmental Perception
- Dynamic Environmental Perception

Cloud Data
- Maps
- Traffic
- Weather

Mission Execution
- Vehicle
- Actuator Command

University of Toronto, CSC2125, Lecture 1: ADS
Functional Reference Architecture

System Supervisor

GPS
IMU
Wheel Odometry
Cameras
LIDAR
RADAR
Cloud Data
Maps
Traffic
Weather

Ego Perception
Static Environmental Perception
Dynamic Environmental Perception

Mission Execution

Vehicle Actuator Command

University of Toronto, CSC2125, Lecture 1: ADS
Functional Reference Architecture

- GPS
- IMU
- Wheel Odometry
- Cameras
- LIDAR
- RADAR
- Cloud Data
  - Maps
  - Traffic
  - Weather
- Ego Perception
  - Static Environmental Perception
  - Dynamic Environmental Perception
- Mission Execution
  - Vehicle Actuator Command

- Velodyne HDL-32
Functional Reference Architecture

- Sensor Input
  - GPS
  - IMU
  - Wheel Odometry
  - Cameras
  - LIDAR
  - RADAR

- Ego Perception
  - Static
  - Dynamic

- Cloud Data
  - Maps
  - Traffic
  - Weather

- Mission
  - Execution
  - Dynamic Environmental Perception
  - Static Environmental Perception

System Supervisor

Graphs:
- (a) pedestrian, object #1, object #2
- (b) pedestrian @ -6.59 km/h
- (c) self-interference, object #1, object #2

http://www.mdpi.com/1424-8220/16/1/124/htm

University of Toronto, CSC2125, Lecture 1: ADS
Radar

- Cheap
- Does well in extreme weather
- Low resolution
- Most used automotive sensor for object detection and tracking
LIDAR

- Expensive
- Extremely accurate depth information
- Resolution much higher than radar
- 360 degrees of visibility
Camera

- Cheap
- Highest resolution
- Huge data = deep learning
- Human brains use similar sensor technology for driving
- Bad at depth estimation
- Not good in extreme weather
Comparisons

Source: https://cleantechnica.com/2016/07/29/tesla-google-disagree-lidar-right/
Lidar

- Proximity
- Detection
- Sensor Cost
- Sensor size
- Detects speed
- Provides Colour / Contrast
- Works in snow / fog / rain
- Works in dark
- Works in bright
- Range
- Resolution
Ultrasonic

- Proximity Detection
- Sensor Cost
- Sensor size
- Detects speed
- Provides Colour / Contrast
- Works in snow / fog / rain
- Works in bright
- Works in dark
- Range

University of Toronto, CSC2125, Lecture 1: ADS
Radar

- Proximity Detection
- Sensor Cost
- Sensor size
- Detects speed
- Provides Colour / Contrast
- Works in snow/ fog / rain
- Range
- Resolution
- Works in dark
- Works in bright
Passive Visual
Sensor Fusion

Sensor types:
- Ultrasonic
- Passive Visual
- Radar

Comparative attributes:
- Sensor Cost
- Sensor size
- Detects speed
- Provides Colour / Contrast
- Proximity Detection
- Range
- Resolution
- Works in dark
- Works in bright
- Works in snow / fog / rain

Graphs depicting the performance of each sensor type against these attributes.
Future of Sensor Technology: Camera vs Lidar

- **Radar and Ultrasonic:**
  - Always there to help

- **Camera:**
  - Annotated driving data grows
  - Deep learning algorithms improve

- **LIDAR:**
  - Range increases
  - Cost drops (solid-state LIDAR)
Functional Reference Architecture

System Supervisor

- GPS
- IMU
- Wheel Odometry
- Cameras
- LIDAR
- RADAR

Cloud Data
- Maps
- Traffic
- Weather

HD Map by HERE

Mission Execution

Dynamic Environmental Perception

Vehicle Actuator Command

University of Toronto, CSC2125, Lecture 1: ADS
Functional Reference Architecture

- Sensor Input:
  - GPS
  - IMU
  - Wheel Odometry
  - Cameras

- Static Environmental Perception
- Cloud Data
  - Maps
  - Traffic
  - Weather

- Dynamic Environmental Perception

- Mission Execution
- Vehicle Actuator Command

Vehicle to infrastructure (V2I) and vehicle to vehicle (V2V) communication

University of Toronto, CSC2125, Lecture 1: ADS
**Functional Reference Architecture**

**Inputs**
- Wheel speeds
- GPS/IMU
- Visual Odometry
- Laser scans
- ...

**Outputs**
- Position
- Speed
- Accel
- Yaw
- Yaw rate
- ...

System Supervisor

Vehicle Model

dx.doi.org/10.1631/jzus.A1400101

University of Toronto, CSC2125, Lecture 1: ADS
Functional Reference Architecture

Ego Perception

Road Network Pose

Lane Level Pose

Precise State

Time to collision with Static Objects

Moving object & situation recognition, and prediction

Time to Collision with Moving Objects

Behavioral Planning

Motion Control

Vehicle Actuator Command

System Supervisor

GPS

IMU

Wheel Odometry

Cameras

LIDAR

RADAR

Cloud Data

Maps

Traffic

Weather

University of Toronto, CSC2125, Lecture 1: ADS
**Functional Reference Architecture**

- **System Supervisor**
  - Ego Perception
    - Road Network Pose
    - Lane Level Pose
    - Precise State
  - Static Environmental
    - Global mapping
    - Road config, lane marking, signs, ...
  - Dynamic Environmental
    - Time to collision with Static Objects
    - Time to Collision with Moving Objects
  - Mission Execution
    - Route Level Navigation
    - Behavioral Planning
    - Motion Control

- **Cloud Data**
  - Maps
  - Traffic
  - Weather

- **Vehicle Actuator Command**

University of Toronto, CSC2125, Lecture 1: ADS
Sensor Input
- GPS
- IMU
- Wheel
- Odometry
- Cameras
- LIDAR
- RADAR

Ego Perception
- Static Environmental Perception
- Dynamic Environmental Perception
  - Moving object & situation recognition, and prediction

Mission Execution
- Route Level Navigation
- Behavioral Planning
- Motion Control

Precise State
- Time to collision with Static Objects
- Time to collision with Moving Objects

System Supervisor

Cloud Data
- Maps
- Traffic
- Weather

See https://www.cityscapes-dataset.com
Functional Reference Architecture

System Supervisor

Sensor Input
- GPS
- IMU
- Wheel Odometry
- Cameras
- LIDAR
- RADAR

Ego Perception
- Static Environmental Perception
- Cloud Data
- Maps
- Traffic
- Weather

Mission Execution
- Road Network
- Pose
- Lane Level Pose
- Precise State
- Global Mapping
- Road config, lane marking, signs, ...

Time to collision with Static Objects
- Traffic Summary
- Moving object & situation recognition, and prediction

Time to Collision with Moving Objects

Dynamic Environmental Perception

Motion Control
- Behavioral Planning
- Route Level Navigation

Vehicle Actuator Command

University of Toronto, CSC2125, Lecture 1: ADS
Sensor Input

- GPS
- IMU
- Wheel Odometry
- Cameras
- LIDAR
- RADAR

Ego Perception

- Static Environmental Perception
- Dynamic Environmental Perception

Mission Execution

- Route Level Navigation
- Behavioral Planning
- Motion Control
- Vehicle Actuator Command

Cloud Data

- Maps
- Traffic
- Weather

Precise State

- Pose
- Traffic signs, ...

Time to Collision with Static Objects

Time to Collision with Moving Objects

Route Network

- Pose
- Lane Level Pose

Continuous


University of Toronto, CSC2125, Lecture 1: ADS
Functional Reference Architecture

- Sensor Input
  - GPS
  - IMU
  - Wheel Odometry
  - Cameras
  - LIDAR
  - RADAR

- Ego Perception
  - Static Environmental Perception
  - Dynamic Environmental Perception

- Mission Execution
  - Route Level Navigation
  - Behavioral Planning
  - Motion Control

- System Supervisor

- Multiplayer games at a roundabout

- Reinforcement Learning

- Traffic
- Weather

- Shalev-Shwartz et al., 2016
Functional Reference Architecture

Sensor Input
- GPS
- IMU
- Wheel Odometry
- Cameras
- LIDAR
- RADAR

Ego Perception
- Static
- Dynamic
  - Environmental
    - Road Network
      - Pose
      - Lane Level Pose
    - Precise State
      - Global mapping
      - Road config,
        lane marking,
      - signs, …
  - Time to collision with Static Objects

Mission Execution
- Route Level Navigation
- Behavioral Planning
- Motion Control

System Supervisor

Continuous

Steering, braking, throttle

Cloud Data
- Maps
- Traffic
- Weather

Vehicle Actuator Command

University of Toronto, CSC2125, Lecture 1: ADS
Functional Reference Architecture

System Supervisor

**Ego Perception**
- Road Network Pose
- Lane Level Pose
- Precise State

**Static Environmental**
- Global mapping
- Road config, lane marking, signs, ...
- Time to collision with Static Objects

**Dynamic Environmental**
- Traffic Summary
- Moving object & situation recognition, and prediction
- Time to Collision with Moving Objects

**Mission Execution**
- Route Level Navigation
- Behavioral Planning
- Motion Control

**Cloud Data**
- Maps
- Traffic
- Weather

**Mission**
- Execution

---

University of Toronto, CSC2125, Lecture 1: ADS
Traffic Data

- Naturalistic driving
- AV sensors & perception
- Infrastructure mounted
- Birds-eye view
A1 vs A2 Autonomy

- **Starting point:**
  - All cars are manually controlled until the AI system shows itself to be available and is elected to be turned on by the human.

- **A1: Human-Centered Autonomy**
  - **Definition:** AI is not fully responsible
  - Feature axis:
    - Where/how often is it “available”? (traffic, highway, sensor-based, etc.)
    - How many seconds for take-over? (0, 1, 10, etc)
    - Teleoperation support

- **A2: Full Autonomy**
  - **Definition:** AI is fully responsible
  - Notes:
    - No teleoperation
    - No 10-second rule: It’s allowed to ask for human help, but not guaranteed to ever receive it.
    - Arrive to a safe destination or safe harbor.
    - Allow the human to take over when they choose to.
A1 vs A2 Autonomy

- **L0** • **Starting point:**
  - All cars are manually controlled until the AI system shows itself to be available and is elected to be turned on by the human.

- **L1, L2, L3** • **A1: Human-Centered Autonomy**
  - **Definition:** AI is not fully responsible

- **L4, L5** • **A2: Full Autonomy**
  - **Definition:** AI is fully responsible
Human-Centric Approach to AI (also see Safety)

90% \[\text{No} \quad \text{Human Needed} \quad \text{Yes} \quad 10\%\]

Solve the perception-control problem where possible:

And where not possible: involve the human

Perception / control (via Deep-Learning)
Paths to Autonomous Future

**A1: Human-Centered Autonomy**

- **Localization and Mapping:** Where am I?
- **Scene Understanding:** Where/who/what/why of everyone else?
- **Movement Planning:** How do I get from A to B?
- **Human-Robot Interaction:** What is the physical and mental state of the driver?
- **Communicate:** How do I convey intent to the driver and to the world?

**A2: Full Autonomy**

- **Localization and Mapping:** Where am I?
- **Scene Understanding:** Where/who/what/why of everyone else?
- **Movement Planning:** How do I get from A to B?
- **Human-Robot Interaction:** What is the physical and mental state of the driver?
- **Communicate:** How do I convey intent to the driver and to the world?
Is partially automated driving a bad idea? Observations from an on-road study

Yes, with nothing to do, drivers quickly stop paying attention, get distracted, fall asleep
Public Perception of What Drivers Do in Semi-Automated Vehicles
What Does Data Say?

A look at several autonomous driving accidents

Based on work of Prof. Mark Lawford, McMaster University
1\textsuperscript{st} Fatal Tesla Autopilot Crash
2016 - January 20th - Fatal - Tesla Model S(China)
1st Fatal Tesla Autopilot Crash Analysis

- Model S was equipped with
  - a single forward facing radar,
  - a single forward facing camera,
  - a set of 12 ultrasonic sensors.

- Camera was used by MobileEye's EyeQ3 computing platform implementing a Deep Neural Network (DNN) for its object identification and detection.

- Vehicle was also equipped with Tesla's Automatic Emergency Braking (AEB) system
  - AEB system required agreement between both the camera and the radar before any action was taken.

- Driver monitoring system consisted of a torque sensor in the steering wheel.
2nd Fatal Tesla Autopilot Crash
2017 - May 7th - Fatal - Tesla Model S (Florida)

1. Trailer turns left in front of the Tesla
2. Tesla doesn't stop, hitting the trailer and traveling under it
3. Tesla veers off road and strikes two fences and a power pole
2nd Fatal Tesla Autopilot Crash
Analysis

• Similar Model S sensors and features to 1st Tesla Autopilot crash
• No braking or avoidance action prior to collision
• Tesla commented that the camera failed to detect the truck due to white color of the trailer against a brightly lit sky and a high ride height.
• They further commented that the radar filtered out the truck as an overhead road sign to prevent false braking.
• In both cases MobileEye commented that:
  • MobileEye's system was not designed to cover all accident scenarios and that Tesla was using it outside of its intended purpose.
3rd Tesla Autopilot Crash
2018 - January 22nd - Non-Fatal – Tesla Model S (California)
Tesla Collision with Fire Truck

• **Tesla** Model S in Autopilot mode was following a **pickup** truck in left lane
• **Pickup** changed lanes to avoid a stationary **firetruck**
• **Tesla** accelerated into the back of the **firetruck** at 65 m.p.h
Similar Autopilot, lane changing lead vehicle & stationary vehicle failure
Autopilot, lane changing lead vehicle & stationary vehicle

• Tesla Model S Handbook states:

“Traffic-Aware Cruise Control cannot detect all objects and may not brake/decelerate for stationary vehicles, especially in situations when you are driving over 50 mph (80 km/h) and a vehicle you are following moves out of your driving path and a stationary vehicle or object is in front of you instead.”
Why the acceleration?

- ACC is part of Autopilot
- Set max speed (normal cruising speed) & time gap (headway) when following a lead vehicle @speed < max speed

Hypothesis:
- When pickup changed lane distance to new lead vehicle (firetruck) increased
- ACC commanded acceleration to close the gap
Another Tesla Autopilot Crash show what this might be like at full speed
Uber Autonomous Vehicle Crash
2018 - March 18th - Fatal – Uber Volvo XC90 (Arizona)
Uber Accident Details

• Uber
  • Switched off Volvo’s standard Aptiva/Intel Mobile Eye collision avoidance/mitigation system
    • Initially detected unknown object 6 seconds before impact
    • It decided it was a bicycle 1.3 second before impact and would have started braking
  • Why?
    • To reduce interference with their software? Avoid false positives?
    • Think of trying to making a right turn @Yonge & Dundas in Toronto
  • Also switched off Volvo’s Driver Distraction Detection System

• What’s a poor autonomous vehicle to do?
  • Maybe requiring having these features turned on by an industry standard assurance case would help!
4th Tesla Autopilot Crash
2018 - March 23rd - Fatal - Tesla Model X (California)
4th Tesla Autopilot Crash
Analysis

NTSB preliminary report summary states:
• During the 60 seconds prior to the crash, the driver's hands were detected on the steering wheel on three separate occasions, for a total of 34 seconds;
• for the last 6 seconds prior to the crash, the vehicle did not detect the driver’s hands on the steering wheel.
• At 8 seconds prior to the crash, the Tesla was following a lead vehicle and was traveling about 65 mph.
• At 7 seconds prior to the crash, the Tesla began a left steering movement while following a lead vehicle.
• At 4 seconds prior to the crash, the Tesla was no longer following a lead vehicle.
• At 3 seconds prior to the crash and up to the time of impact with the crash attenuator, the Tesla’s speed increased from 62 to 70.8 mph, with no pre-crash braking or evasive steering movement detected.
4th Tesla Autopilot Crash

Analysis

• Tesla stated after the accident:
  • “The driver had about five seconds and 150 meters of unobstructed view of the concrete divider with the crushed crash attenuator, but the vehicle logs show that no action was taken.”

• Oddly enough, Tesla failed to mention that the Tesla sensors and AEB had the exact same opportunity to see the concrete divider and react in a timely fashion to mitigate the outcome
A similar Tesla crash
Following lane marks – to an accident

1. Location of Police vehicle
2. Right hand lane marker as road starts to widen for turn lane
   Probably during “rush hour” no vehicles park there
Main Fallacy in existing (implicit) Assurance Cases for ADAS

• The driver is going to catch the Machine Learning (ML) failures . . . without driver attentiveness monitoring!
Getting too (artificially) intelligent with safety

• Object identification is very useful
• Can help predict and plan in addition to help partially meet some safety goals
• Pedestrian detection is an example of how ML fails badly with the key safety requirement: “Don’t hit things!”

AI/ML Version:

“I don’t know what it is so it’s not there.”

vs

Safety Version:

“I don’t know what it BUT IT’S THERE!”
If ML Doesn’t Recognize It, It’s Not There
The trouble with AI in safety critical situations

- Using ML to deal with cross walks:
  - AI does a good job with this but not ...
Lessons learned

• Production is currently taking precedence over safety and that is resulting in accidents

• The driver is not a sufficient mitigation without *real* driver attentiveness monitoring

• Interactions with other systems requirements is compromising safety (ACC acceleration in stopped vehicle accidents, interactions between control loops at different time scales)

• Current systems are not providing confidence information from ML components resulting in unsafe behaviour
  • When in doubt, slow down!

• New failure modes not discussed here – maintenance
  • replacing your windshield can now cause accidents due to sensor calibration errors!
Proper Monitoring of Driver Attentiveness

**Super Cruise**
Tested on Cadillac CT6

Super Cruise uses a camera to watch where the driver’s eyes are looking.

- Capability & Performance
- Ease of Use
- Clear When Safe to Use
- Keeping Driver Engaged
- Unresponsive Driver

**Autopilot**
Tested on Tesla X/S/3

Autopilot performed well and is easiest to use in stop-and-go traffic.

- Capability & Performance
- Ease of Use
- Clear When Safe to Use
- Keeping Driver Engaged
- Unresponsive Driver

Also work from MIT (see Lecture 2 of MIT course on Deep Learning and Self-Driving)
Self-Driving Car Tasks

• Localization and Mapping – Where am I
• Scene Understanding – Where is Everyone Else?
• Movement Planning – How to get from Point A to Point B
• Driver State – What is the Driver Up to?
  • Essential if driver is part of the loop!
• Safety Monitoring
Safety Assurance of ADS

Source: Krzysztof Czarnecki, Waterloo
Operational Design Domain (ODD)

SAE J3016 Levels of Driving Automation

A set of **conditions** under which the driving automation can operate a vehicle

<table>
<thead>
<tr>
<th>Time of day</th>
<th>Types of roads</th>
<th>Geographic area</th>
<th>Traffic conditions</th>
<th>Weather conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>day</td>
<td>residential</td>
<td>area</td>
<td>stop-and-go</td>
<td>clear</td>
</tr>
<tr>
<td>night</td>
<td>residential</td>
<td>area</td>
<td>free flowing</td>
<td>raining</td>
</tr>
<tr>
<td></td>
<td>urban</td>
<td></td>
<td></td>
<td>snowing</td>
</tr>
<tr>
<td></td>
<td>highway</td>
<td></td>
<td></td>
<td>icy</td>
</tr>
</tbody>
</table>
Dynamic Driving Task (DDT) Fallback

Who performs the DDT in the case of *system malfunction* or when *leaving the ODD*?
Automated Driving Systems (ADS)

SAE J3016 Levels of Automation

- **3**: Conditional Automation
  - Example: ADS for stop-and-go
  - ODD: limited
  - Fallback: Driver

- **4**: High Automation
  - Example: Shuttle in geo fenced area
  - ODD: unlimited
  - Fallback: ADS

- **5**: Full Automation
  - Example: Robo Taxi anywhere
ADS Hazard Sources

Mechanical faults

Electrical faults

Computer HW faults

Computer SW faults

Sensor noise & limitations

Machine learning errors

Inadequate driving behavior

DDT fallback failures

Cyber attacks

Mature best practices

ISO 26262

(ISO / PAS 21448)

SAE J3061

University of Toronto, CSC2125, Lecture 1: ADS
Assurance: ISO 26262

- HARA
- Safety Concept
- Architecture
- Vehicle Validation Test
- Verification Test
- Safety Requirements

Safety Case
ADS Hazard Sources

Mature best practices

- Mechanical faults
- Electrical faults

ISO 26262

- Computer HW faults
- Computer SW faults

(ISO / PAS 21448)

- Sensor noise & limitations
- Machine learning errors

SAE J3061

- Inadequate driving behavior
- DDT fallback failures
- Cyber attacks
Safety

Absence of unreasonable risk of mishap

- Risk
  - Severity
  - Likelihood
Driving Behavior Safety

Absence of unreasonable crash risk due to ADS driving behavior

Noncollisions

Collisions
Factors Influencing Risk Acceptability

- Risk level
- Risk reduction cost
- Benefit of the risky functionality (risk taking)
- Best practice (state of technology)
- Replacement risk
- Who controls risk
- Perception/public opinion
Assurance Target

- ODD boundary
- Acceptable risk of unknown unsafe scenarios
- Acceptable risk of known unsafe scenarios
Responsibility-Driven Safety

• Normal driving scenarios
  • Must not cause unacceptable risk increase
  • Low/high demand (incl. other road user errors)

• Emergency scenarios
  • Near-crash
    • Must avoid crash if it can
  • Crash
    • Must mitigate if it can
    • Dilemmas often addressed by blame assignment
  • Fallback
    • Must minimize overall risk

Blame vs. Injury Risk

GM Cruise Chevy vs. motorcycle crash
https://www.dmv.ca.gov/portal/wcm/connect/1877d019-d5f0-4c46-b472-78cfe289787d/GMCruise_120717.pdf?MOD=AJPERES
Blame vs Injury Risk (from the Accident Report)

A Cruise autonomous vehicle ("Cruise AV"), operating in autonomous mode in heavy traffic, was involved in a collision while traveling east on Oak Street just past the intersection with Fillmore Street. The Cruise AV was traveling in the center of three one-way lanes. Identifying a space between two vehicles (a minivan in front and a sedan behind) in the left lane, the Cruise AV began to merge into that lane. At the same time, the minivan decelerated. Sensing that its gap was closing, the Cruise AV stopped making its lane change and returned fully to the center lane. As the Cruise AV was re-centering itself in the lane, a motorcycle that had just lane-split between two vehicles in the center and right lanes moved into the center lane, glanced the side of the Cruise AV, wobbled, and fell over. At the time of the collision, the Cruise AV was traveling with the flow of traffic at 12mph, while the motorcycle was traveling at approximately 17mph. The motorcyclist got up and walked his vehicle to the side of the road, where the parties exchanged information. 911 was called pursuant to Cruise policy. The motorcyclist reported shoulder pain and was taken to receive medical care, and a police report was taken. As reported in Traffic Collision Report#l70989746, the motorcyclist was determined to be at fault for attempting to overtake and pass another vehicle on the right under conditions that did not permit that movement in safety in violation of CVC 21755(a).
High-Level Behavior Safety Requirements (Normal Driving)

1. Vehicle stability
2. Assured clear distance ahead
3. Minimum separation
4. Traffic regulations
5. Informal traffic rules (best practices)
Behavioral Safety: 1. Vehicle Stability

Skid stability

Roll stability

Friction ellipses

\[ e + \mu_y = \frac{v^2}{12/R} \]
Behavioral Safety: 2. Assured Clear Distance Ahead (ACDA)

Stopping sight distance
(Perception-reaction time and braking distance)

Perception distance
(Range + road geometry)

Limits safe speed
Behavioral Safety: 2. ACDA Perception Distance

Crests

Object height $h_2 = 2.0'$

Eye height $h_1 = 3.5'$

Curves

Sight distance ($S$)

Obstruction or backslope

Center of lane

Radius

Intersections

Overtaking

FIRST PHASE

SECOND PHASE

Opposing vehicle appears when passing vehicle reaches Point A.

Passing vehicle
Behavioral Safety:
3. Minimum Separation

Separation in terms of **distance gap**, **time gap**, and **time-to-collision**

... and various maneuver-specific gaps, including following, overtaking, turning
Behavioral Safety:  
4. Traffic Regulations

Safe speed (ACDA)  
Yielding to other road users rules  
Obeying regulatory traffic signs & signals  
Where to drive  
Reacting to emergency vehicles & school buses  
U-turn prohibitions

Safe following gap  
Passing rules  
Signaling stops & turns  
Parking restrictions  
Use of passing beam  
Required behavior at railway crossings  
...
Behavioral Safety:
5. Informal Traffic Rules

2/3 – second rule

Responding to tailgaters

How early to signal turns

Delayed acceleration at signalized intersections

Lane selection

Anticipating aberrant behaviors of other road users

Responding to animals on the roadway
WISE Drive Documentation

WISE Drive comes with comprehensive documentation (over 350 pages) available from this page.

All eight documents in two zip archives: zip1, zip2

Driving Task Specification

Maneuver Catalog


Basic Motion Control Task Catalog


Road Environment Specification

ODD Taxonomy

ADS Hazard Sources

Mature best practices

Mechanical faults

Electrical faults

ISO 26262

Computer HW faults

Computer SW faults

(ISO / PAS 21448)

Sensor noise & limitations

Machine learning errors

Inadequate driving behavior

DDT fallback failures

SAE J3061

Cyber attacks

University of Toronto, CSC2125, Lecture 1: ADS
Fail-Operational ADS Architecture

- Primary channel
  - High-performance AD function (gracefully degradable)
  - Monitor A
  - Enable Out

- Secondary channel
  - Minimal AD fallback function
  - Monitor B
  - Enable Out

Vehicle Platform

- Dependability patterns:
  - Redundancy
  - Diversity
  - Simplex
  - Graceful degradation
  - Monitoring of monitoring
  - Minimized cost

No single-point failures

University of Toronto, CSC2125, Lecture 1: ADS
ODD vs. ROD

Operational Design Domain

Restricted Operational Domain

I Colwell, B Phan, S Saleem, R Salay, K Czarnecki. An Automated Vehicle Safety Concept Based on Runtime Restriction of the Operational Design Domain. IEEE Intelligent Vehicles Symposium (IV), 2018
ROD Monitoring for Graceful Degradation

System supervisor

ROD Monitor

System Health Monitor

System functions

ROD Manager

System configurator
ADS Hazard Sources

- Mechanical faults
- Electrical faults
- Computer HW faults
- Computer SW faults
- Sensor noise & limitations
- Machine learning errors
- Inadequate driving behavior
- DDT fallback failures
- Cyber attacks

Mature best practices

ISO 26262

(ISO / PAS 21448)

SAE J3061

University of Toronto, CSC2125, Lecture 1: ADS
Challenges of Assuring Machine Learned Components

Lack of specification

Lack of inspectability

Lack of Complete Spec Affects Verification and Testing (see Lecture 4 by R. Salay)

**Best practices**
- Spec notations
- Design guidelines
- Coding guidelines

**Fault tolerance**
- Error detection & handling

**Verification methods**
- Walkthroughs
- Inspections
- Formal verification
- Static code analysis

**Testing methods**
- Requirements-based testing
- Error guessing
- Interface test
- Fault injection test
- Resource usage test
- Structural coverage

ISO 26262 Part 6

University of Toronto, CSC2125, Lecture 1: ADS
Key Recommendations (see Lecture 4)

• Partial specifications
  • Assumptions, necessary/sufficient conditions, in- and equivariants
  • Runtime monitoring, test generation, regularization

• Data requirements
  • Domain coverage (e.g., ontology)
  • Risk profiling
ADS Challenges

(an unsorted list)
Road User Intension
Will she cross the street?
Will she cross the street?
Bad Weather Driving
“Plastic Bag” Problem
Edge Cases
Driving into a Tornado
Autonomous Trap 101
Crossing Double Yellow Lines
Place Charles de Gaulle, Paris
Busy City Traffic
Vehicle To Pedestrian Communication

Clamann et al. 2016
Daimler Prototype
Unexpected Road Incursion by Pedestrians

Sudden Emergency Doctrine for human drivers

What is the expected standard for AVs?
Moral Machines

http://moralmachine.mit.edu/
Safety of Sensors and AI
Testing Challenges

• 100 million miles driven between deadly crashes (US)
  • Crashes are rare events
  • Human drivers are extremely good, when they pay attention

• Showing equal performance by an AV with 95% confidence requires demonstrating 300 million miles driven without a deadly crash
California DMV Disengagement Reports

- Google (miles driven between disengagements):
  - 2015: 2000 miles
  - 2016: 5000 miles
Tesla Autopilot Data Collection and Testing

- In 2016, on average, 1 million miles per 10h data collected
  - Object lists
  - Driver inputs
  - Vehicle state
- Since May 5, 2017, Tesla asks for permission to gather video clips from their customers
- OtA Update staging
  - Dormant mode
  - Gradual release
Testing in Virtual World

Vehicle Physics and 3D Photo-Realistic Simulation
by Aleksandar Pocuc & Igor Ilic
V2X: Major Infrastructure Requirements