Applications of Machine Learning for Autonomous Driving & Challenges in Testing & Verifications

Ching-Yao Chan

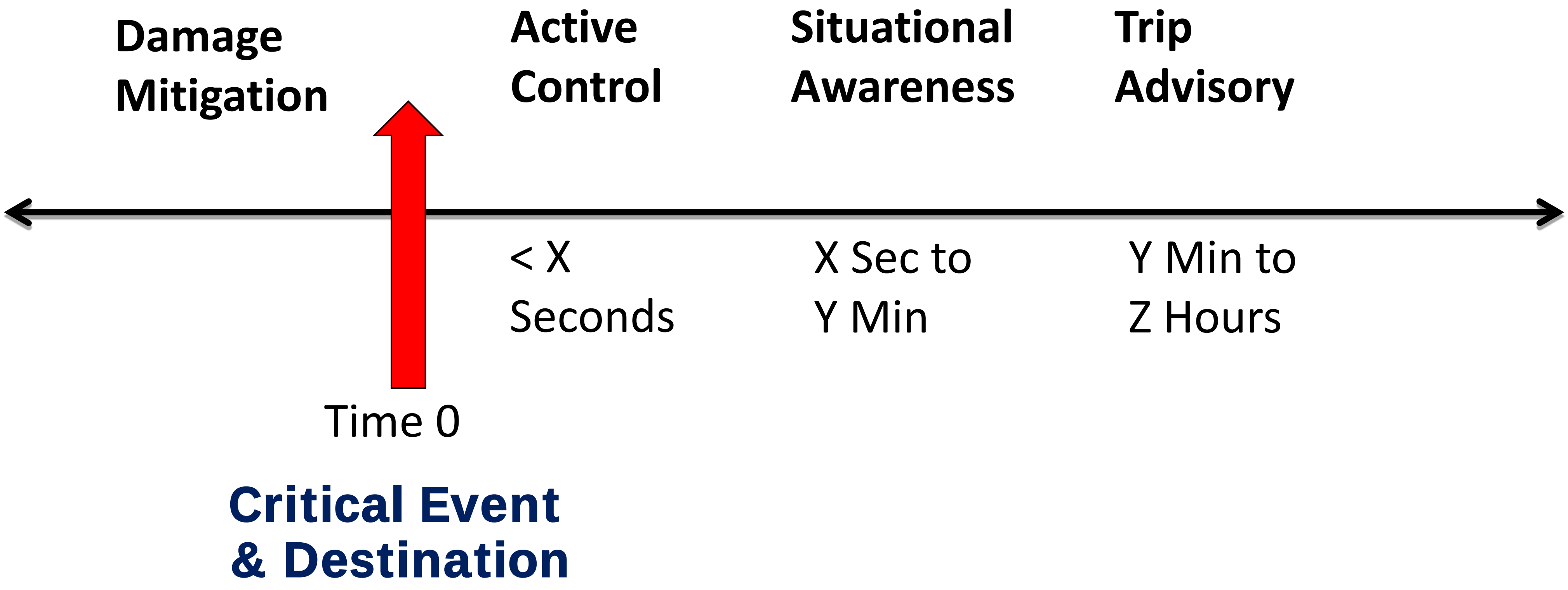
Nokia Workshop

January 11, 2018

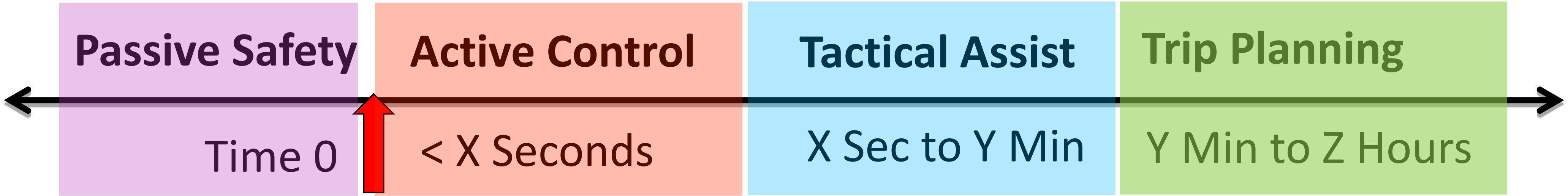
Berkeley DeepDrive

Taxonomy of A Driving Trip

- *Driving Experience Taxonomy – Classification by Timeline*



Taxonomy of Driving (by Critical Event)



Exemplar Automated & Assist Functions

- *Strategic Advisory (Route Planning)*
- *Tactical Assist (Maneuver Planning)*
- *Control and Assist (Collision Avoidance)*
- *Damage Mitigation (Minimum Risk Actions)*

Berkeley DeepDrive

Driving Tasks & Automation

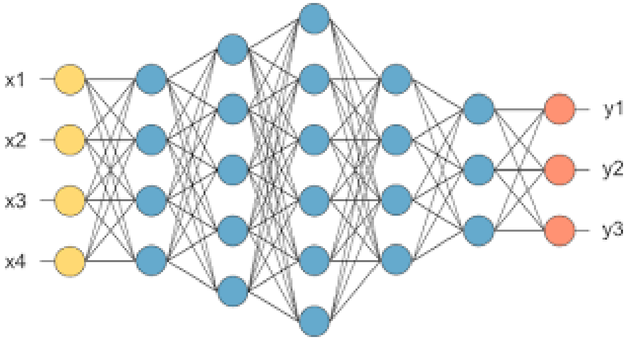
Hierarchical Level of Driver Tasks	Tasks and Considerations	Automation & Assist Functionality
Strategic (Long time intervals)	Trip planning, Route selection, Risks and costs, Satisfaction & comfort	Long-term planning, Route optimization Risk evaluation, Comfort/security assurance
Tactical (seconds)	**Situation Awareness & Negotiating maneuvers,** Obstacle avoidance, Gap acceptance, Turning and overtaking	Partially or fully support drivers for **cognition, perception, and decision making**
Control (centi- or milli-seconds)	Motor actions, Physical response	Enhance or substitute driver actions to overcome degradation of physical and mental abilities

Driving Tasks & ML / A.I.

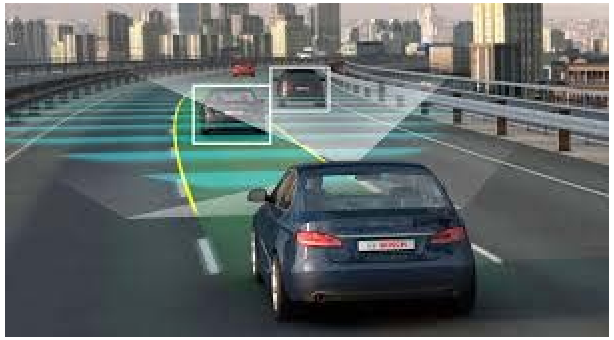
Hierarchical Level of Driver Tasks	Tasks and Considerations	Human and Artificial Intelligence
Strategic (Long time intervals)	Trip planning, Route selection	**Trip optimization (time, cost, risk, comfort, etc.)**
Tactical (seconds)	**Situation Awareness & Negotiating maneuvers,** Obstacle avoidance, Gap acceptance, Turning and overtaking	**Inference, Anticipation, Adaptation, Planning, etc. (General A.I.)**
Control (milliseconds)	Motor actions, Physical response	**Control optimization (safety, efficiency, smoothness, etc.)**

Machine Learning/A.I. & Automated Driving

A Great Enabler



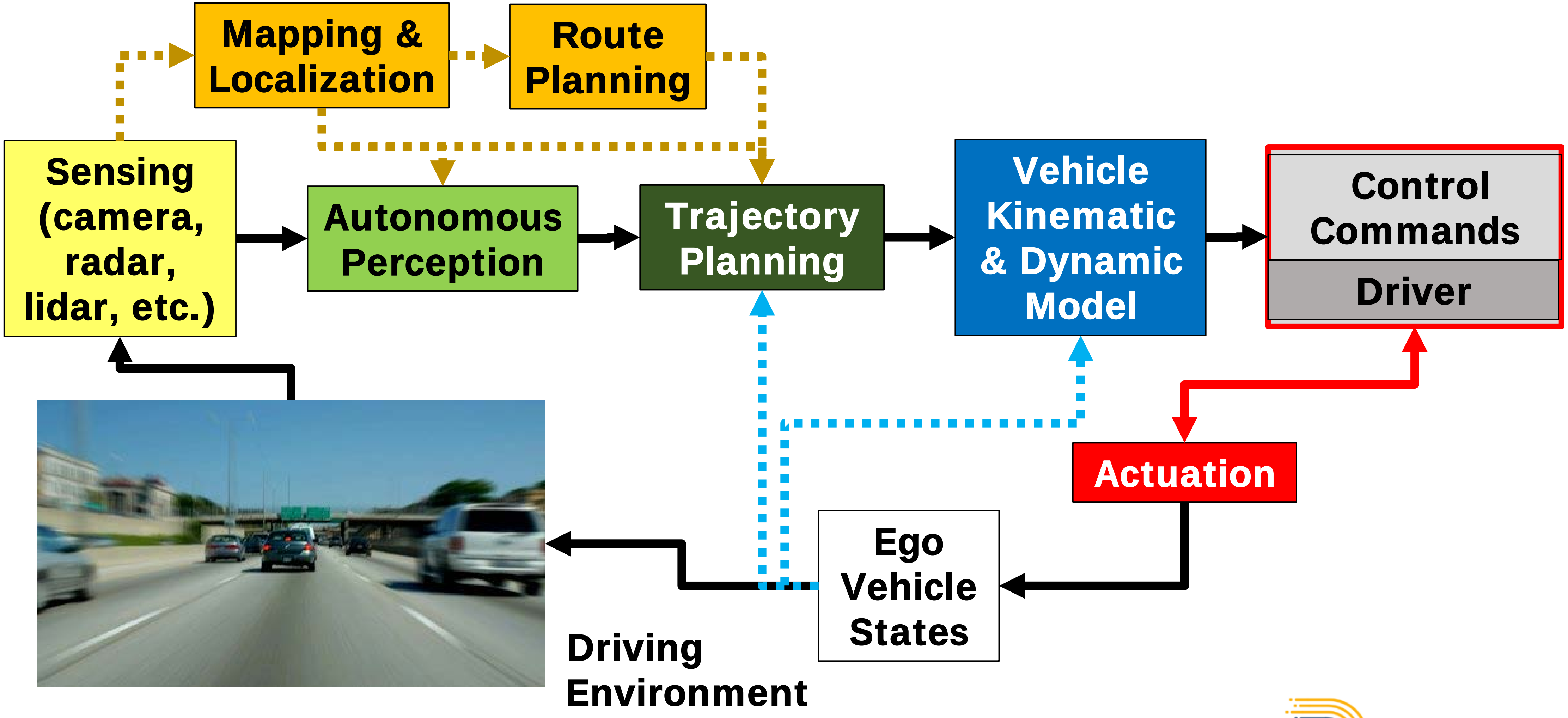
A Fitting Challenge



Where and How Best to Utilize?

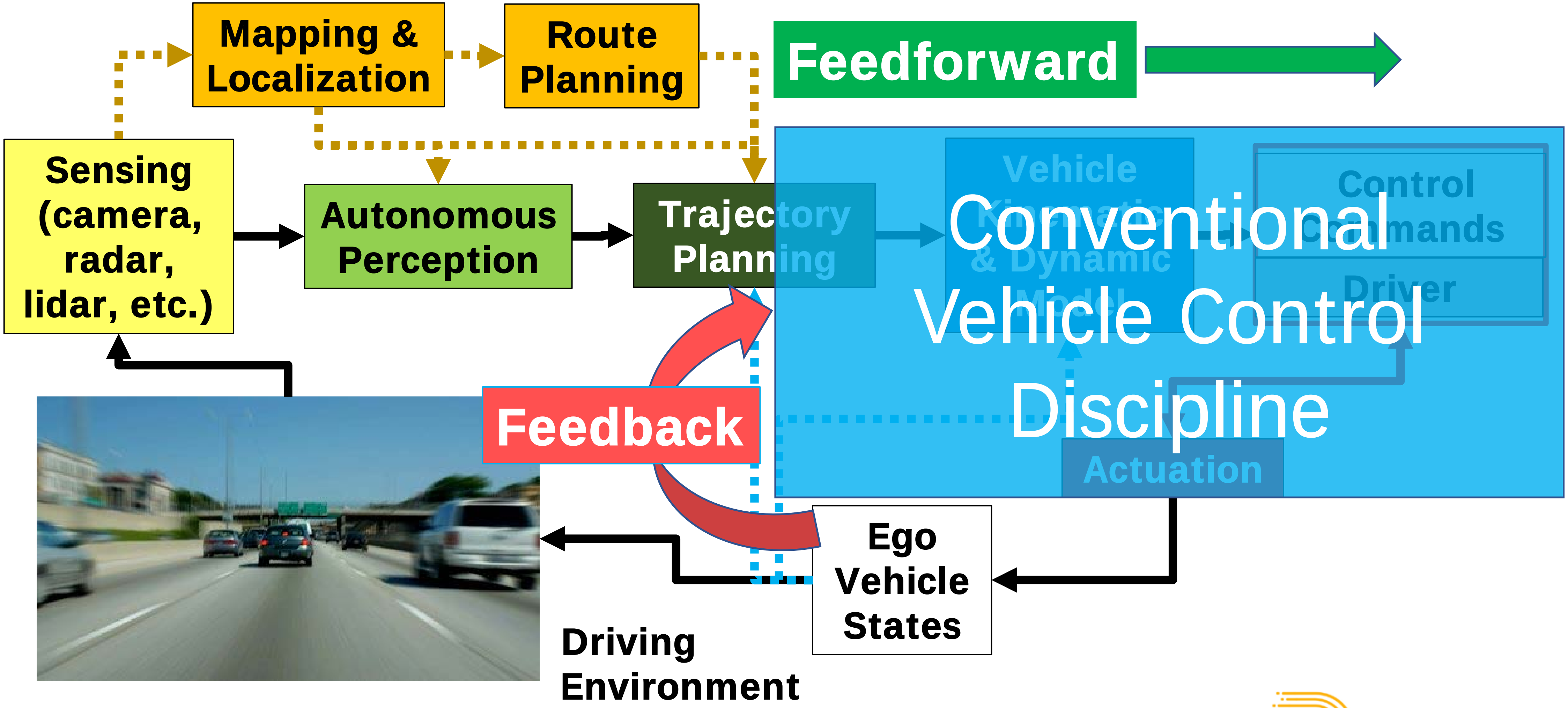
Berkeley DeepDrive

Automated Driving Systems (ADS)
- Functional Block Diagram



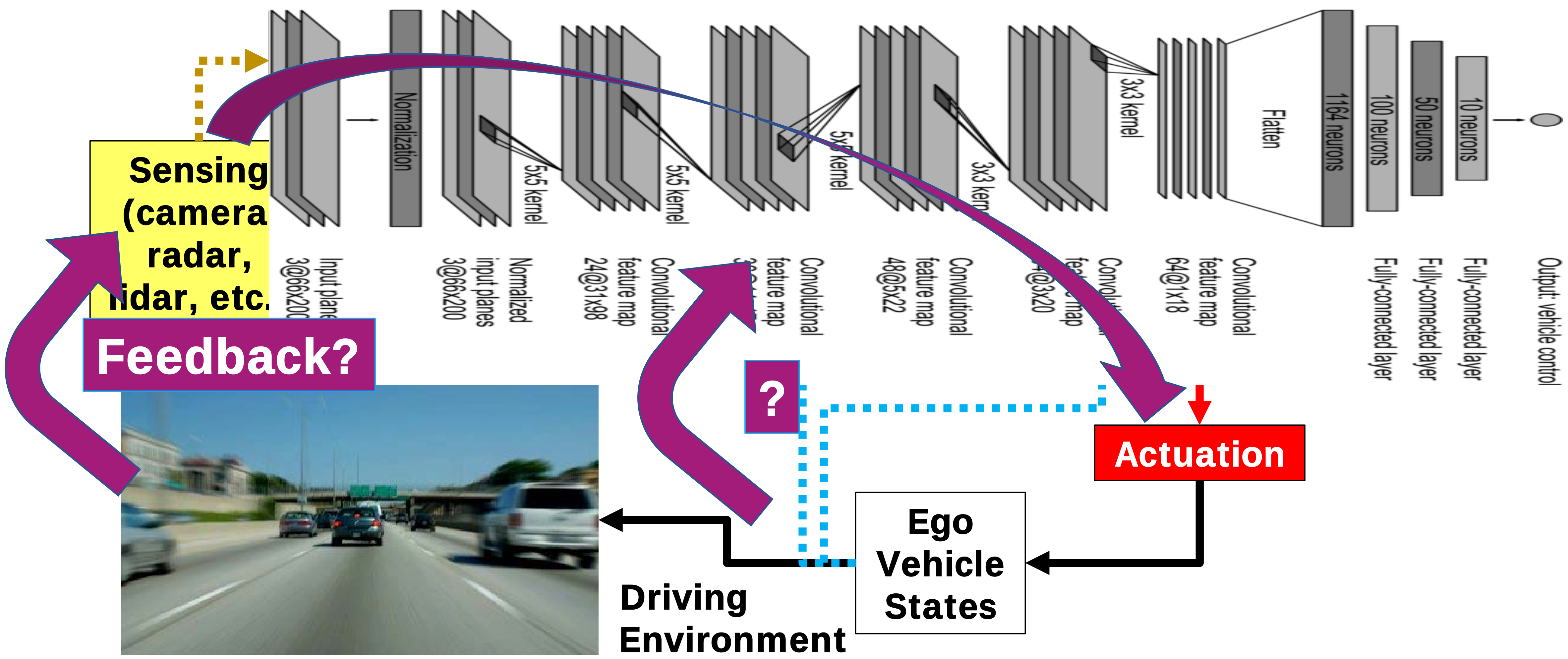
Berkeley DeepDrive

Automated Driving Systems (ADS)
- Feedforward and Feedback in Control Systems



Berkeley DeepDrive

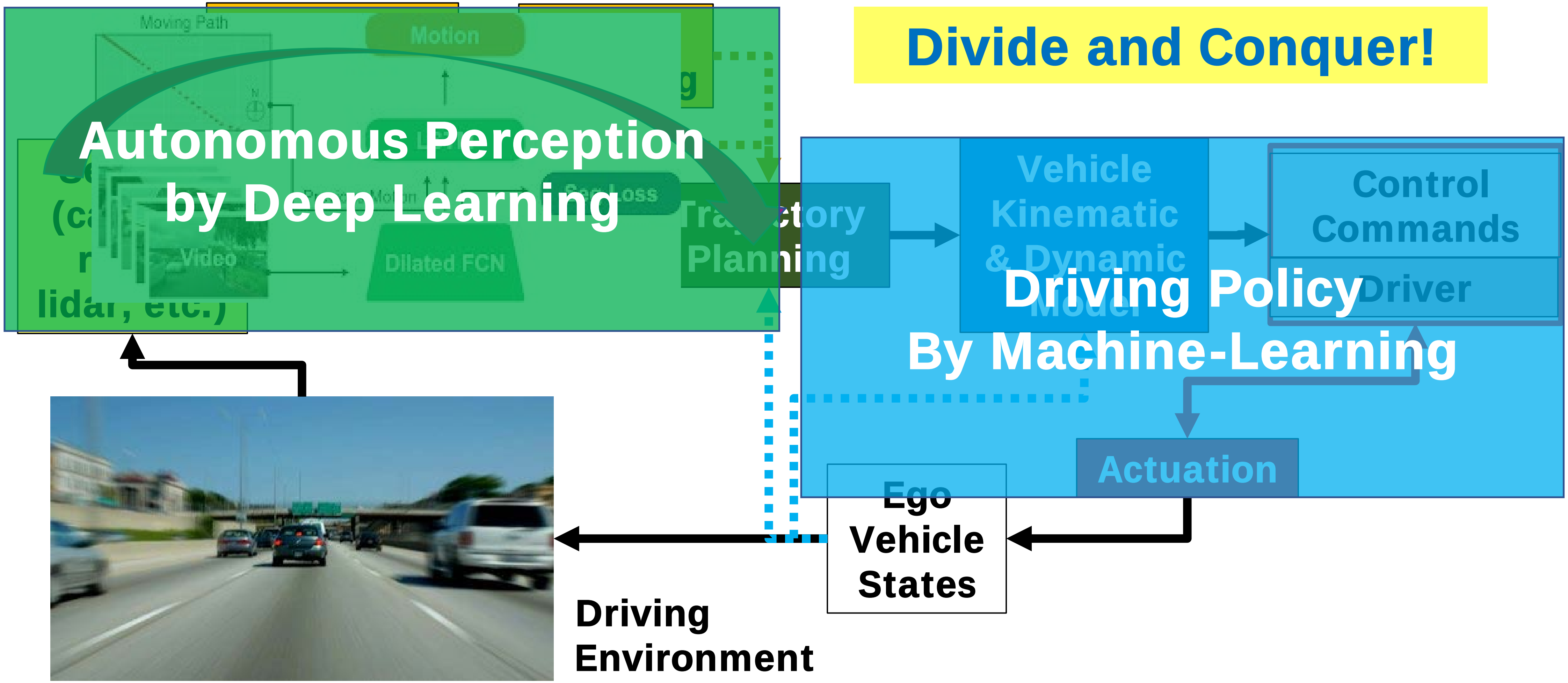
Automated Driving Systems (ADS)
- DNN End-to-End and Feedback in ADS



* End to End Learning for Self-Driving Cars, https://arxiv.org/abs/1604.07316

Berkeley DeepDrive

Automated Driving Systems (ADS)
AI / ML Application

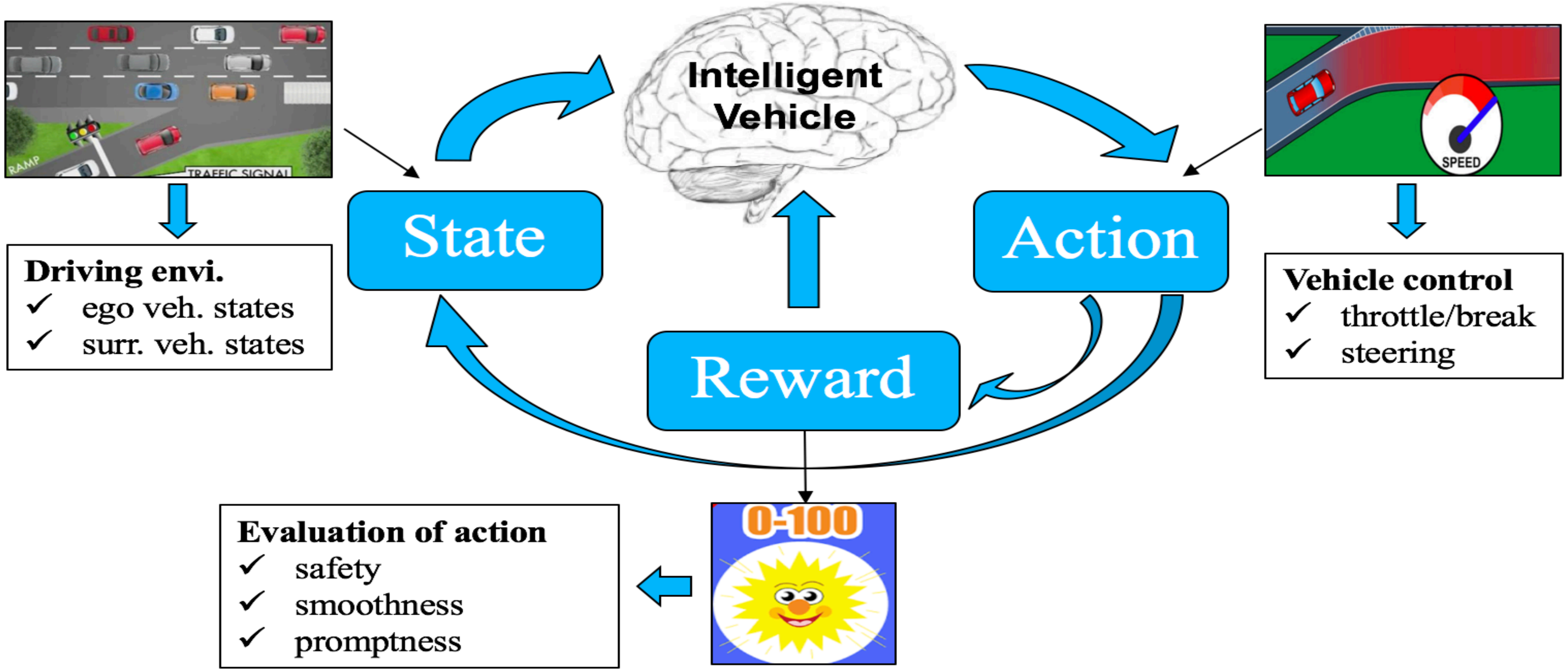


* End-To-End Learning of Driving Models From Large-Scale Video Datasets, https://arxiv.org/pdf/1612.01079.pdf

Berkeley DeepDrive

Reinforcement Learning for Automated Driving
A Use Case of Ramp Merge Automation

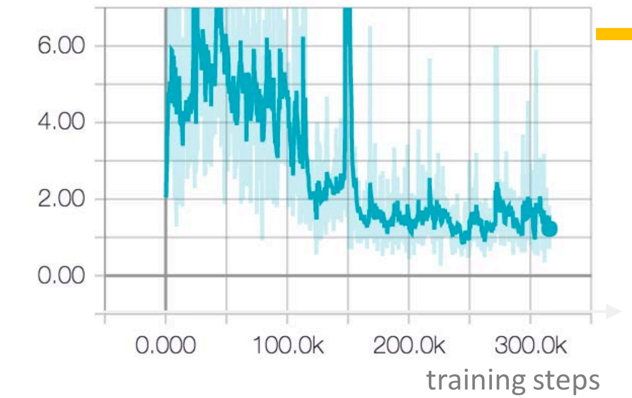
Pin Wang, Ching-Yao Chan



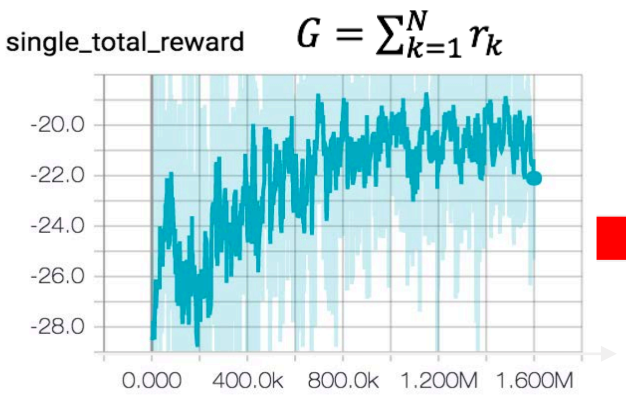
Berkeley DeepDrive

RL Training Results

Loss: $L(\theta) = \sum_{i=1}^{N} (Q^T - Q^P)_i^2$



$G = \sum_{k=1}^{N} r_k$



Hardware

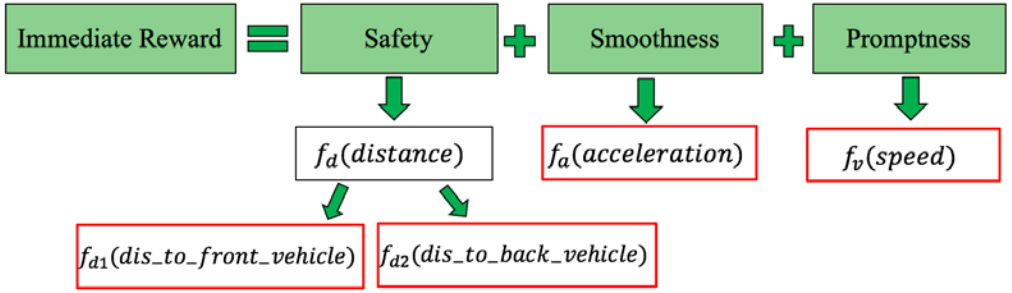
Processor: 2.5 GHz Intel Core i7 CPU

Memory: 16 GB

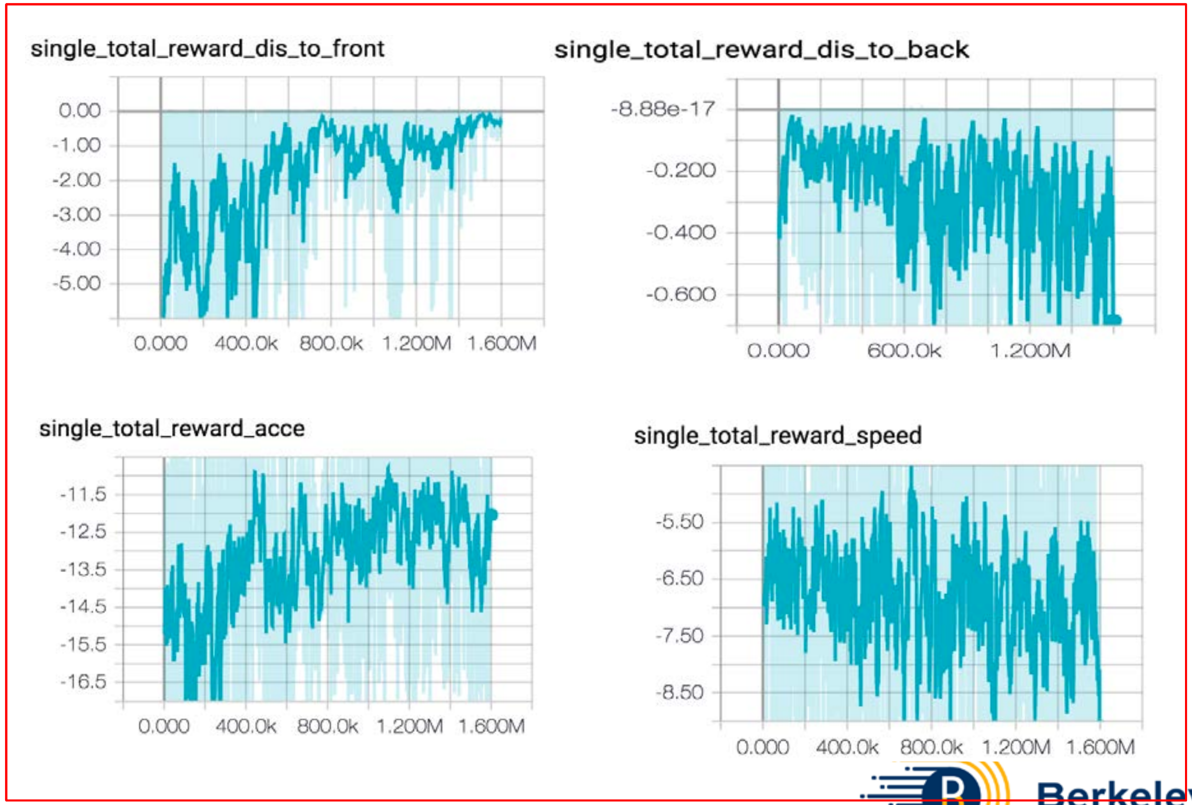
Training Time:

About 100 mins.

About 15,000 merging vehicles.



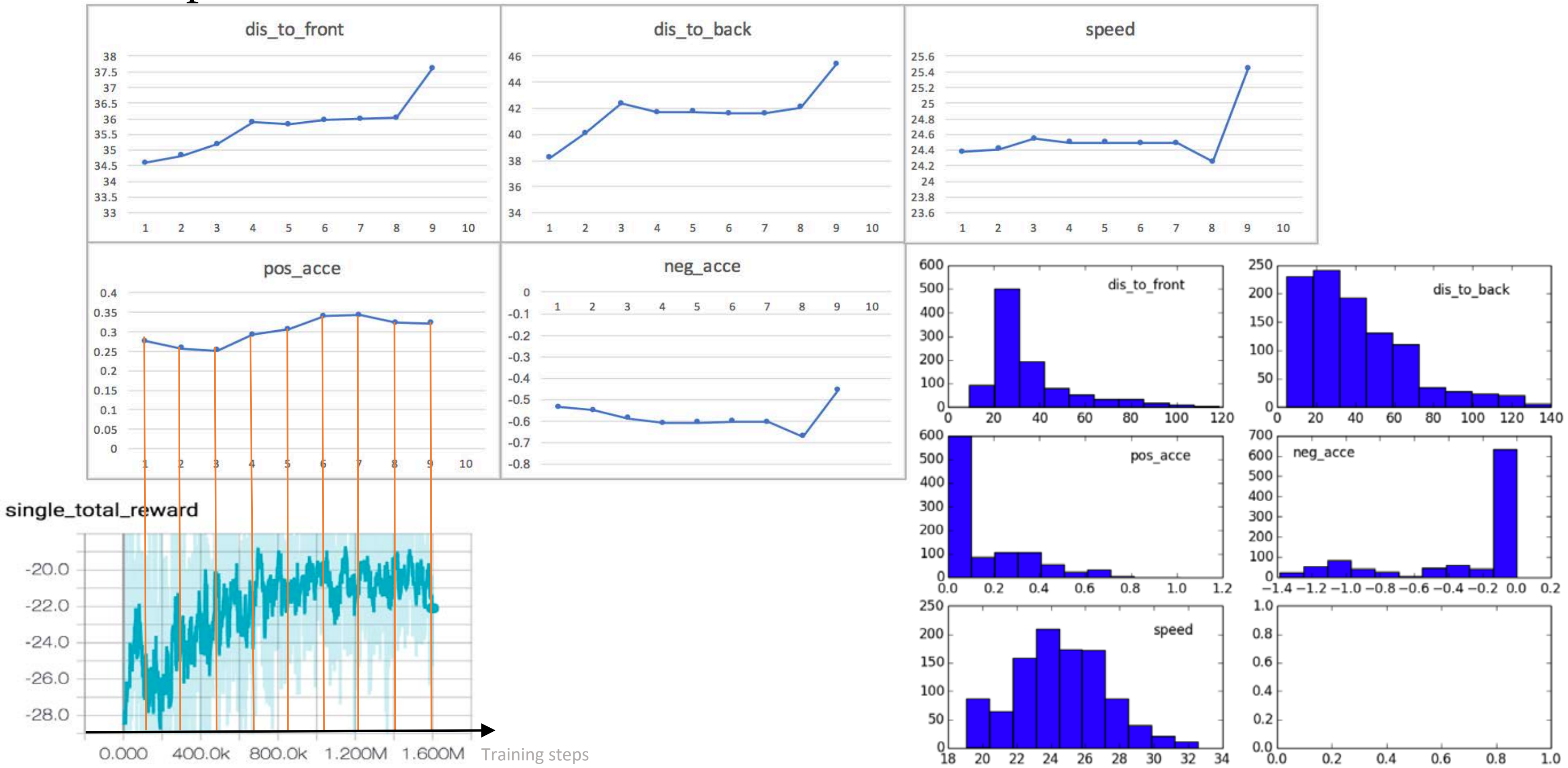
Balanced weightings



Berkeley DeepDrive

Driving Performance Metrics vs. Reward Function

- ✓ distance to gap front vehicle
- ✓ distance to gap back vehicle
- ✓ speed
- ✓ positive max. acceleration
- ✓ negative max. acceleration



Autonomous Vehicle Proposition

- Driver errors cause >90% of accidents
- Automated Driving Systems (ADS) will replace drivers to perform driving tasks
- Including drivers in the loop is not sensible, and drivers can’t be expected to take over safely or effectively.
- Do without Drivers!

Berkeley DeepDrive

Things Could Still Go Wrong, Even If Vehicles are Automated?

Tesla Fatality Incident (May 2016, Florida)

- Neither human nor the machine hits the brake.

- Driver's hands on wheel for only 25 seconds during 37-minute period
- Driver ignored 7 visual warnings and 6 audible warnings during the trip

- Tesla cruise control set at 74 mph.
- Driver has at least 3.4 seconds to react.

Berkeley DeepDrive

Questions and Comments

- Supposedly, the driver should still be monitoring the Driving environment. (SAE Level 2)
  - Can we expect drivers to be continuously vigilant?

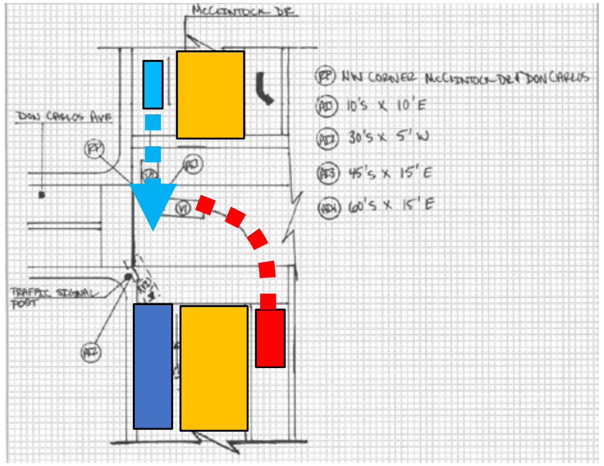
- Apparently, the “detection/perception” function failed for the ADS to timely react to the situation.
  - Why did the Tesla do not slow down?
  - Radar was the only sensor that could have detected the tractor-trailer, allegedly.

- Can AI/ML help?

Berkeley DeepDrive

UBER Accident, 03/2017

- **Recent UBER Incident (March 2017, Arizona)**



An Uber SUV crashed in Tempe March 24, 2017. (Photo: Mark Beach)

- **UBER car has the right of way, per police report.**
- **The left-turn car was "at fault," and cited.**

- **Two inside lanes were grid-locked. (orange)**
- **The outside lane was clear to proceed. (green)**

- **Did the UBER car try to "rush through the intersection"?**
- **Did it make a good judgment?**

Berkeley DeepDrive

Questions and Comments

- **How will a (conservative) human driver behave if he is in the UBER car?**
  - **Slow down as it approaches, given that the left-lane traffic is congested and partially blocking the view?**

- **Is this a failure in decision making and driving behaviors?**
  - **Defensive driving in anticipation of other road users**

- **Can AI / ML help?**

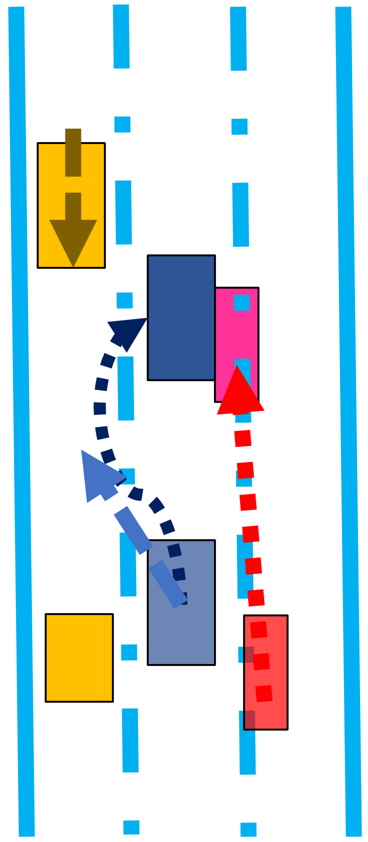
Berkeley DeepDrive

Cruise Automation Accident, 12/2017

- **Recent Cruise Incident (December 2017, SF)**

- Cruise AV intends to change lane, but a van in front slows down
- As Cruise AV aborted a lane change and was re-centering itself in the lane,
- **A motorcycle** that that had just lane-split between two vehicles moved into the center lane, **glanced the side of the Cruise AV, wobbled, and fell over**.
- At the time of the collision,
  - Cruise AV was traveling with the flow of traffic at 12 mph
  - Motorcycle at approximately 17 miles per hour.

- **Can AI/ML help?**



Berkeley DeepDrive

Safety Challenges in Real World

- These accidents may be the first to draw attention,
  - But, they won’t be the last
- There is usually a prime culprit of functional failure in the system,
  - But, multiple causes/factors are often involved
- The real world is very complicated
  - How much testing is needed?
  - How do we verify safety?

Challenges

in

AV Testing and Safety Verification

How Much Testing Is Needed?
Compared to Benchmark (Human) Performance

- The National Safety Council of US reports a rate (including deaths of pedestrians and cyclists killed in motor vehicle accidents for all roads) of
  - 1.25 deaths per 100 million vehicle miles, or
  - 12.5 deaths per billion vehicle miles) traveled in 2016.

- 80 million vehicle-miles per fatality

  ~= 20,000 miles/per year X 50 years X 80 life-times, or

  ~= 4,000 cars X 20,000 miles/per car per year (production)

  ~= 10 cars X 160,000 miles/per car per year X 50 years (prototype)

How Confident Are We about the Validity of On-Road Testing?

- Even if 10^9 km miles were driven for testing, what do we learn?

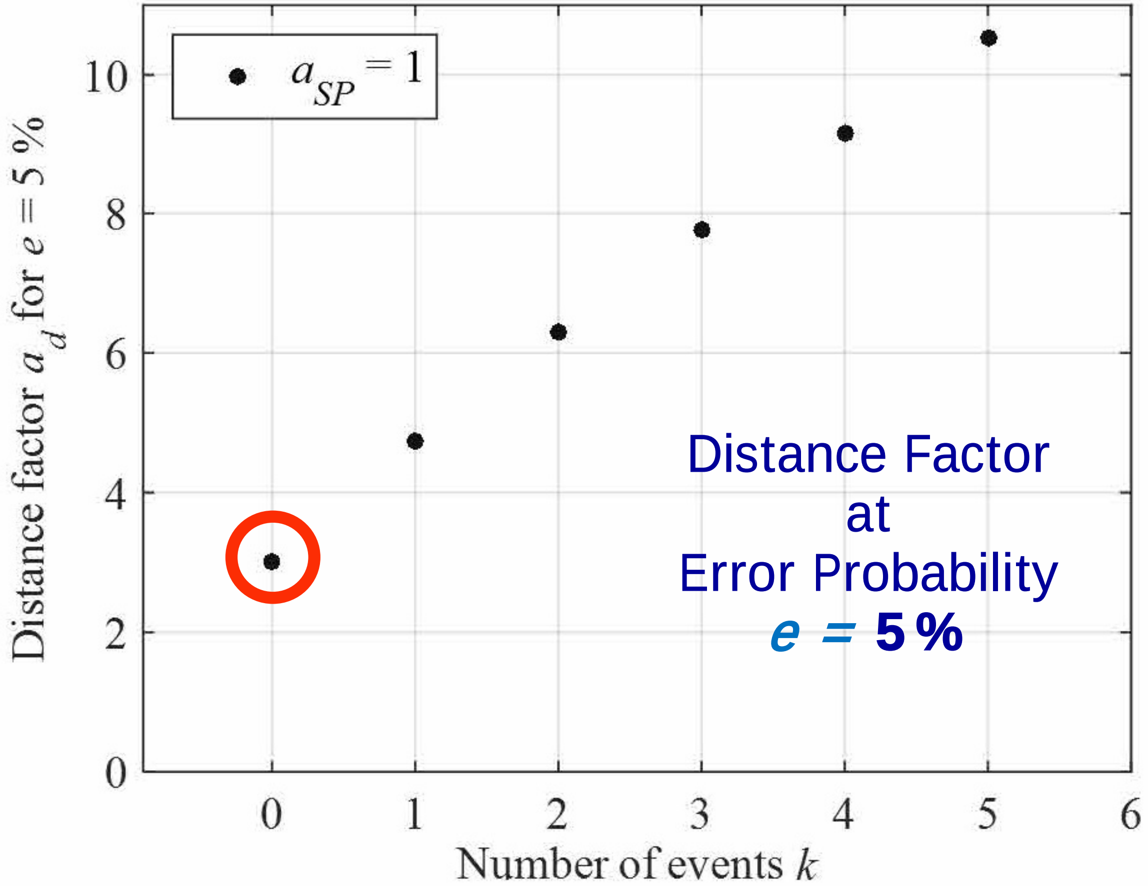
- As of 02/2016 (after the Google car-bus incident), Google cars had a total of *17 crashes over 1.3 million miles* of on-road testing, since 2009. (13 by other vehicles rear-ending Google cars)
  - Twice as high as typical statistics
  - Is Google safety performance inferior?

- At the time of the first fatal Tesla crash in 05/2016, Tesla Auto-Pilot fleet has accumulated over *130 million miles on the road*.
  - Mileage relatively higher than typical statistics
  - Is Tesla safety performance superior?

- Accidents are random events, and they must be given in the context of probability. (Topic for another day!)

Berkeley DeepDrive

A Probabilistic Model for Accident Occurrence



Poisson Distribution: $P_\lambda(k) = \frac{\lambda^{k}}{k!} e^{-\lambda}$

- Premise: The occurrence of an accident is an **independent and non-exhaustive random process**.

- The data point at zero events in the figure* means that, with a distance factor $a_d \approx 3$, the probability e is less than 5% that a vehicle performing worse than the comparison group is not involved in an event.

*** Doctoral Dissertation by Walther Wachenfeld, advised by Prof. Hermann Winner, Technischen Universität Darmstadt**

Berkeley DeepDrive

Safety Assurance of ADS

- The consensus is that it is too resource-consuming and not feasible to conduct ADS testing by physical cases “completely.”

- Safe validation must include a structured combination of the following methodologies:
  - Proving ground testing (especially corner cases)
  - On-road testing
  - Simulation

- Safety assurance is a major challenge to be addressed.

- Efforts are underway, e.g. PEGASUS project in Germany.

Berkeley DeepDrive

How to Expedite Learning and Testing?

- **Practices of Safety Assurance Testing:**
  - **Learn from database of "corner cases"**
    - **Collection of challenging scenarios and probable test cases for specifications**
  - **"Fleet" Learning**
    - **Tesla, e.g.**
  - **"Simulated" Learning**
    - **Waymo, e.g.**

Berkeley DeepDrive

Learning Corner Cases

- Testing to ensure that they can operate reliably under infrequently encountered situations
  - Strange and extreme weather
  - Emergency vehicles, fire trucks, police cars, motorcycles
  - Other road users' behaviors (e.g. pedestrians and bicyclists)
  - Unusual traffic, construction zones, etc.

- Situations that humans find understandable may not be easily recognizable to software*
  - A data set comprehensive to a human may be insufficient for a machine

- Creating Corner Cases is a topic for AI/ML methods

* Edge Cases and Unknown Unknowns, M. Wagner, 07/2017

Berkeley DeepDrive

Tesla “Fleet Learning”

- As of November 2016,
  - The Autopilot first generation fleet is over 100,000 vehicles strong.
  - Tesla has accumulated 1.3 billion miles of Autopilot data from its first generation sensor suite.
  - The actual number of miles driven with the Autopilot active is closer to 300 million miles at that point in time.
- Tesla “learn” from the data even when the Autopilot is not active to improve its Autopilot.
- Distributed and Crowd-Based Learning is a research topic for AI / ML.

Berkeley DeepDrive

Waymo "CarCraft"

Google's Virtual World:

- **At any time, there are now 25,000 virtual self-driving cars making their way through fully modeled versions of cities and test-track scenarios.**
- **Collectively, they now drive 8 million miles per day in the virtual world.**
- **In 2016, they logged 2.5 billion virtual miles versus a little over 3 million miles by Google's self-driving cars that run on public roads.**

How to create meaningful simulation scenarios is an active topic.

Berkeley DeepDrive

How Does Waymo Do it? (1)

Waymo Safety Report, 10/2017, https://waymo.com/safetyreport/

- **Behavioral Safety** (driving decisions and behavior of our vehicles on the road),
- **Functional Safety** (operate safely even when there is a system fault or failure; including backups and redundancies),
- **Crash Safety** (ability to protect people within the car),
- **Operational Safety** (safety and comfort in interaction between passenger and car), and
- **Non-collision Safety** (safety for anyone interacting with the vehicle in any capacity).

Berkeley DeepDrive

How Does Waymo Do it? (2)

Waymo Safety Report, 10/2017, https://waymo.com/safetyreport/

- Waymo's Self-Driving Systems, Safety Measures
  - Operational Design Domain (ODD)
  - Minimum Risk Condition (fallback)

- Test and Verification Methods
  - Base Vehicle Safety
  - Self-Driving Hardware Testing
  - Self-Driving Software Testing
    - Simulation, Closed-Course, Real-World Driving
  - Testing the Fully Integrated Vehicles
    - Testing on Public Road
    - Testing Crash Avoidance Capabilities
    - Hardware reliability and durability testing

Challenges
in
Safety Testing & Verification

ML/AI in Autonomous Driving

Safety Challenges for A.I.

- “Safe” means* (for ML and AI)
  - Doing what they are designed to do properly
  - Dealing with non-routine hazards
  - Providing resilience in likely gaps
  - Being adaptive by on-line learning
  - Planning to be fail-operational and fail-safe
  - Monitoring themselves with confidence
  - Maximizing controllability

* Koopman and Wagner, IEEE ITSC Magazine, Spring 2017

Meeting the Challenges

Desirable Safety Features and Requirements	Needed Research in ML/AI to Address Concerns

Berkeley DeepDrive

Safety Risks

- Safety Risk =
  Exposure X Severity X Controllability

- Fixed-route, slow-moving (driverless) shuttles minimize safety risks.

- Partially automated systems (and ADAS) have drivers as backup, to increase controllability and reduce risks

- Driverless Automated Mobility Services (Robot-Taxi) must perform well to avoid risks.

Berkeley DeepDrive

A.I. Evolution toward Safe ADS

- Higher Intelligence Beings
  - Adaptive to new domains
  - Robust to deal with adversarial inputs
  - Meta-Learning – Learning to Learn
  - General A.I.

- Continuing AI/ML research will evolve toward helping achieve ADS-Safety.

Berkeley DeepDrive

Thank you.

Ching-Yao Chan
cychan@berkeley.edu

Berkeley DeepDrive