Integrating Autonomy into Urban Systems

A Reinforcement Learning Perspective

Cathy Wu | Assistant Professor

CEE, IDSS, LIDS

NSF Workshop Control for Networked Transportation Systems, Philadelphia, 2019

Year 2050: The hope of self-driving cars

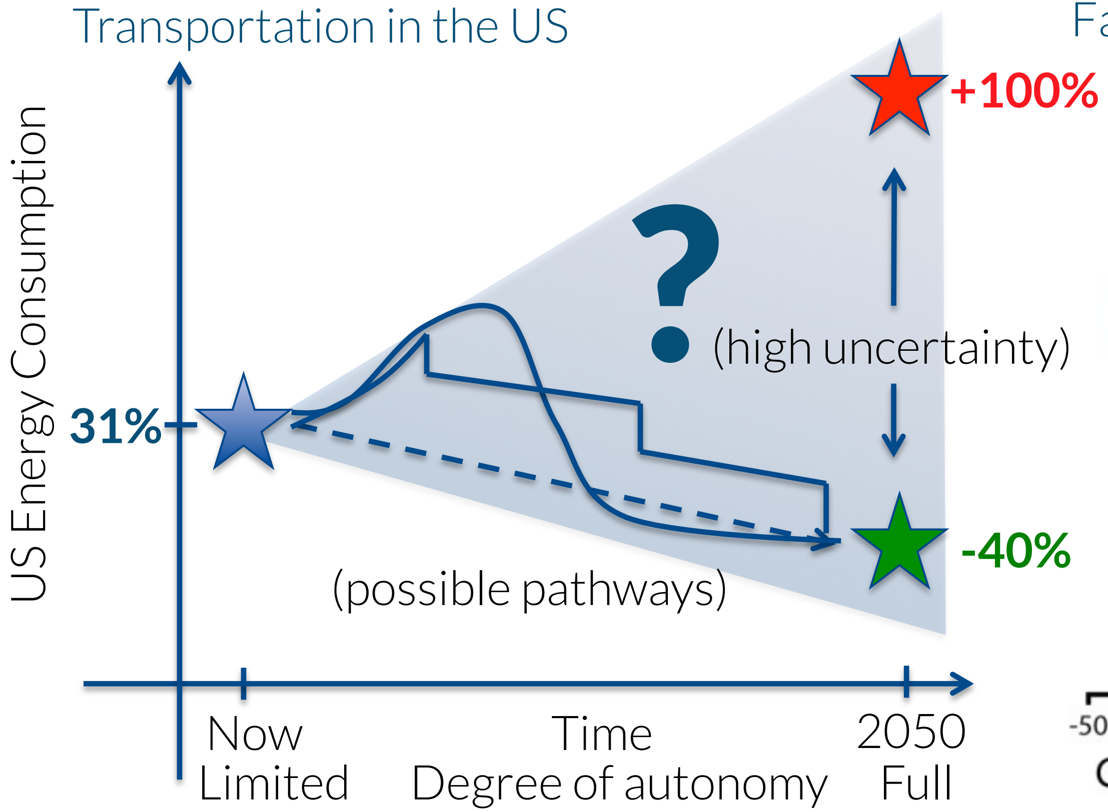
- **Traffic accidents**:
  - 37,000 fatalities
  - 41% deaths of young adults (ages 15-24)
  - 94% of serious crashes caused by human error
- **Congestion**:
  - 6.9 billion hours wasted
  - 3.1 billion gallons of fuel wasted (160$B)

- **Greenhouse gas emissions**:
  - 28% from transportation
- **Access to mobility**: 30% of population
  - 20% youth or elderly
  - 10% disabled (ages 18-64)

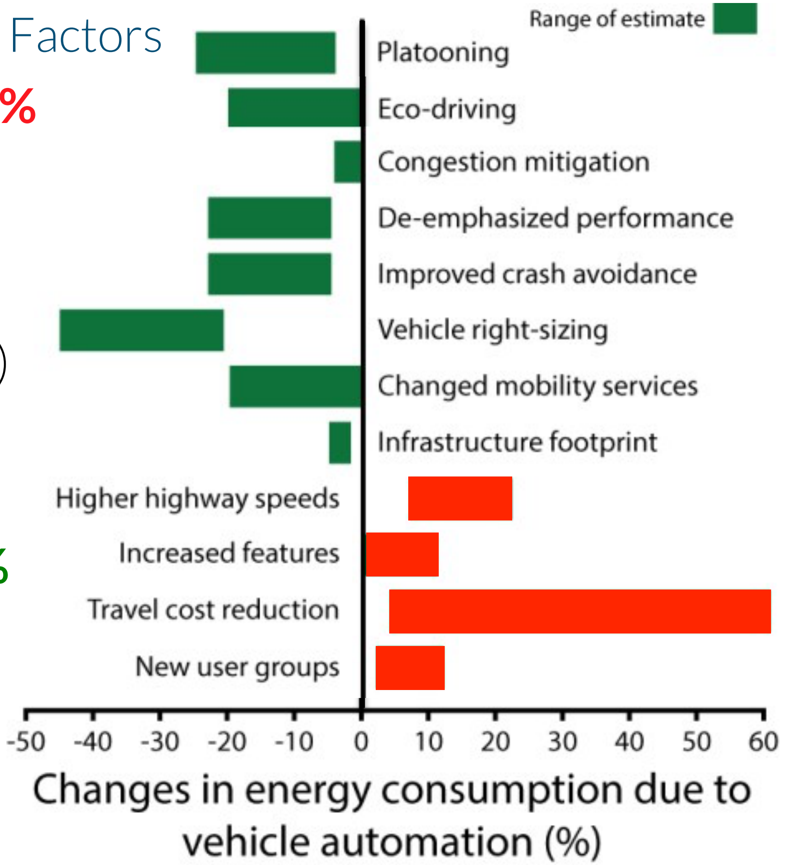
U.S. Energy Information Administration, 2017; U.S. Census Bureau, 2017.

Years 2019 to 2049: Integrating autonomy

Transportation in the US

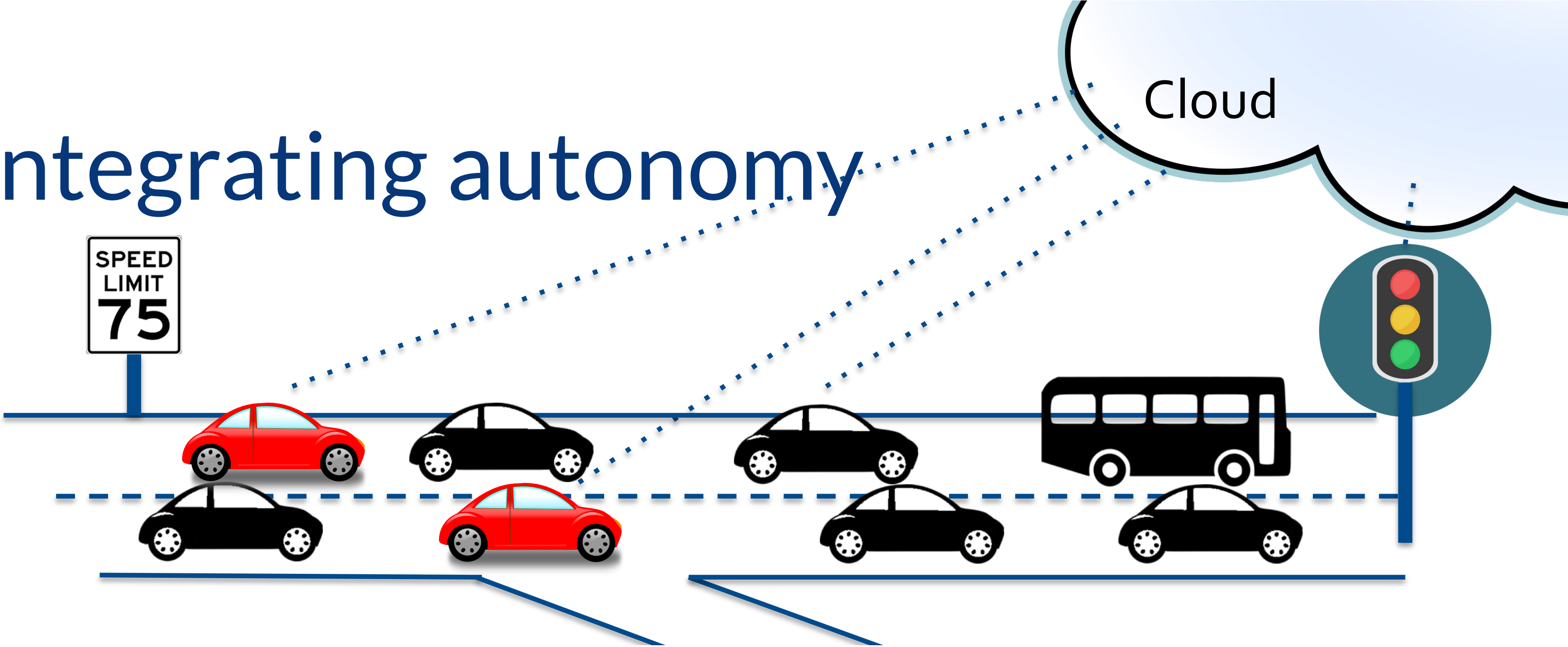


Factors



Wadud, et al. TR-A, 2016. TR-A; U.S. Energy Information Administration, 2017.

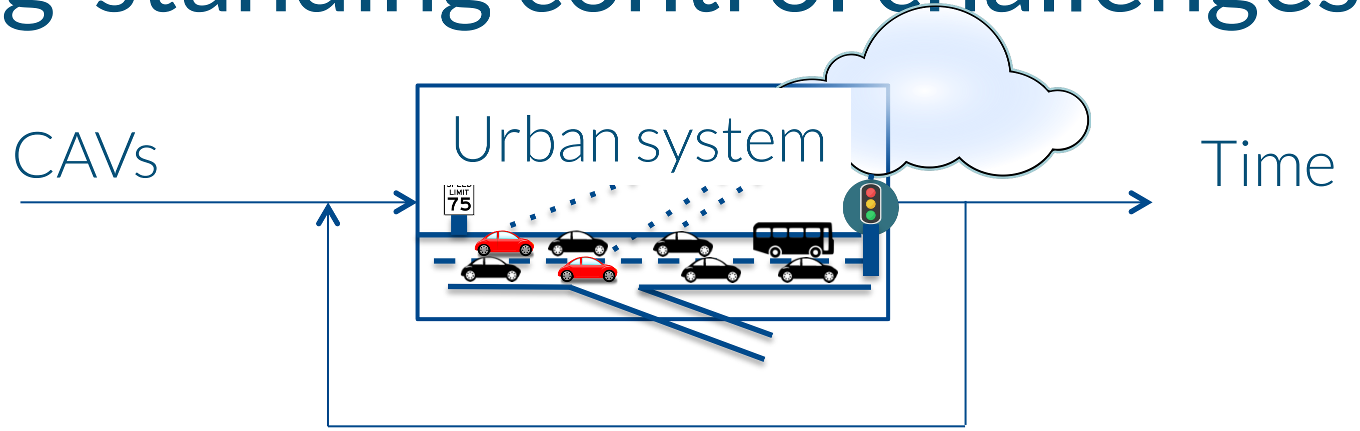
Integrating autonomy



How can we gain understanding for integrating autonomy into complex systems?

In particular: traffic congestion.

Long-standing control challenges



System complexity

- Highly complex non-linear delayed dynamics
- Human behavior modeling
- Large-scale, heterogeneity
- Computational cost

Data restrictions

- Expensive to collect data
- No data on the future
- Expensive to test / deploy
- Limited benchmarks

Integrating autonomy as control

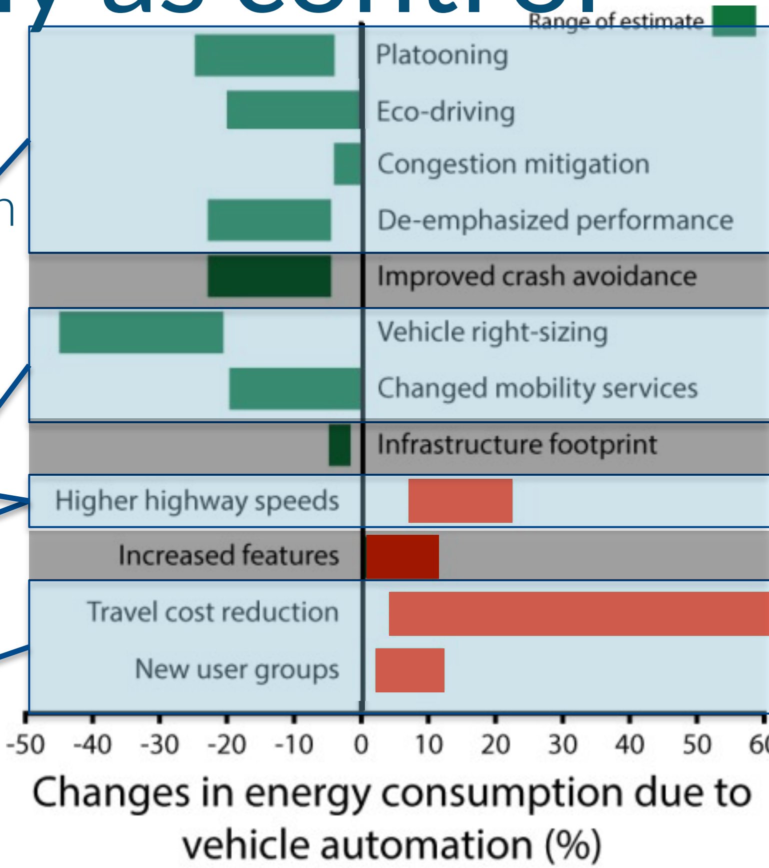
Approach for deriving insights:

- Trade precision for complexity
- Leverage compute and system decomposition
- Leverage simulation to overcome data restrictions

Deep reinforcement learning
[ICLR18, T-RO18, CoRL18, ICRA18, ITSC17, ITSC17]

Convex optimization
[ISTTT15, TR-C15, CDC15, T-ITS18]

Combinatorial optimization
[ITSC16, ITSC16, T-ASE18]



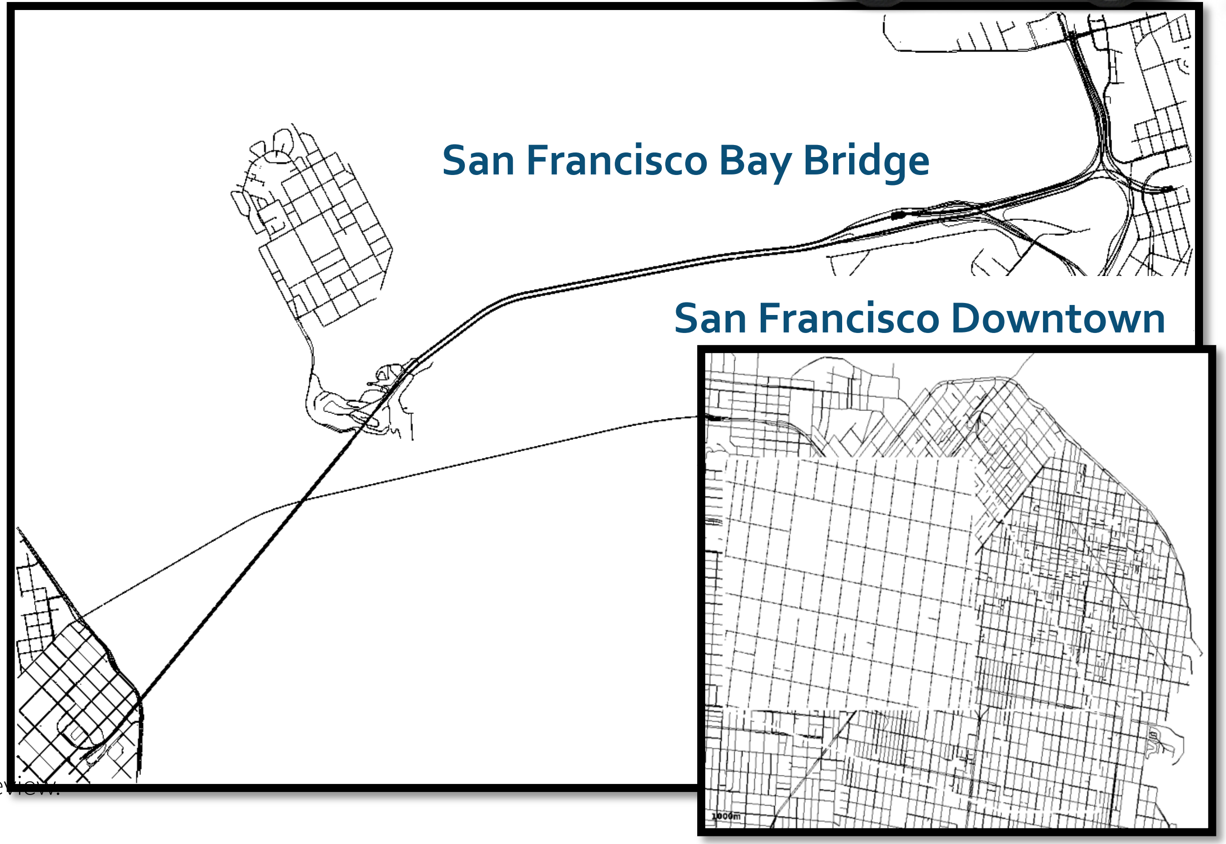
Wadud, et al. TR-A, 2016. TR-A; U.S. Energy Information Administration, 2017.

Urban networks

Setting: ~2000 vehicles

Dynamics:

- cascaded nonlinear systems
- bottlenecks
- multi-lane merges
- toll plaza dynamics

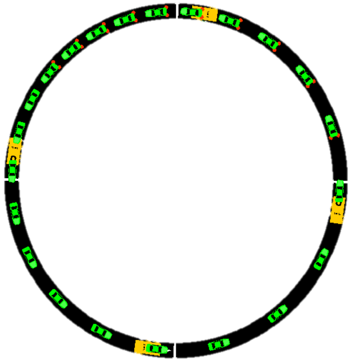


Wu, et al. IEEE ITSC, 2017; **Wu**, et al. IEEE T-RO, in review.

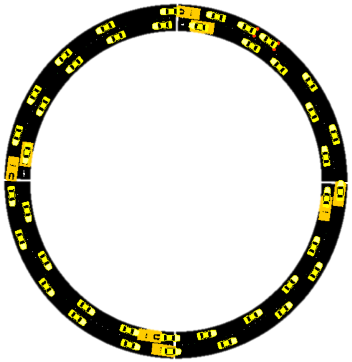
Traffic LEGO blocks

Benchmarks for autonomy in transportation

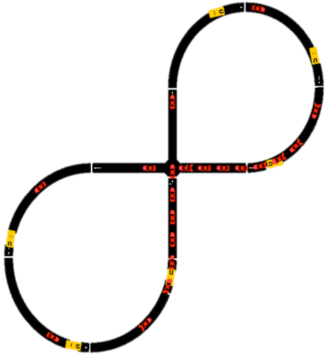
Single-lane



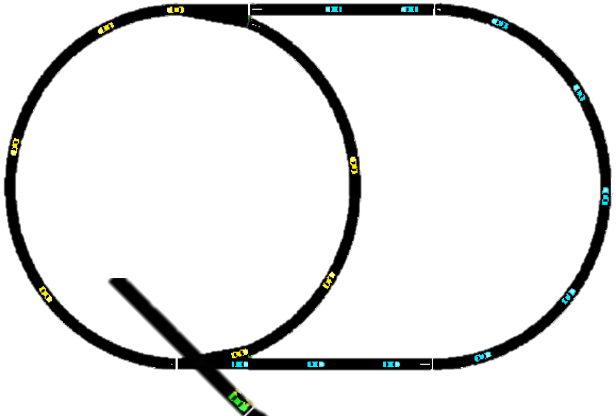
Multi-lane



Intersection



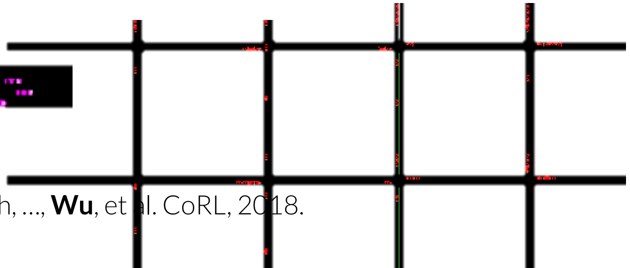
On/off-ramp



Straight highway

Bottleneck

Grid network



Signalized intersection

Wu, et al. IEEE ITSC, 2017; **Wu**, et al. IEEE T-RO, in review; Vinitsky, Kreidieh, ..., **Wu**, et al. CoRL, 2018.

Traffic jams

Sugiyama, et al.

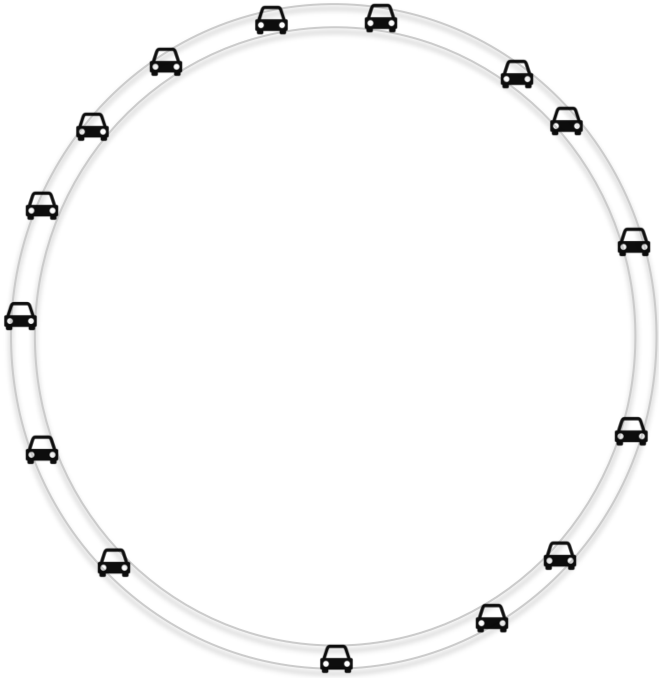
1955 — 900 papers on PDEs for traffic — **2008** — 2019

Partial differential equations (PDE)

Setting: 22 human drivers

Instructions: drive at 19 mph.

No traffic lights, stop signs, lane changes.



Traffic jams

Sugiyama, et al.

1955 — 900 papers on PDEs for traffic — **2008** — 2019

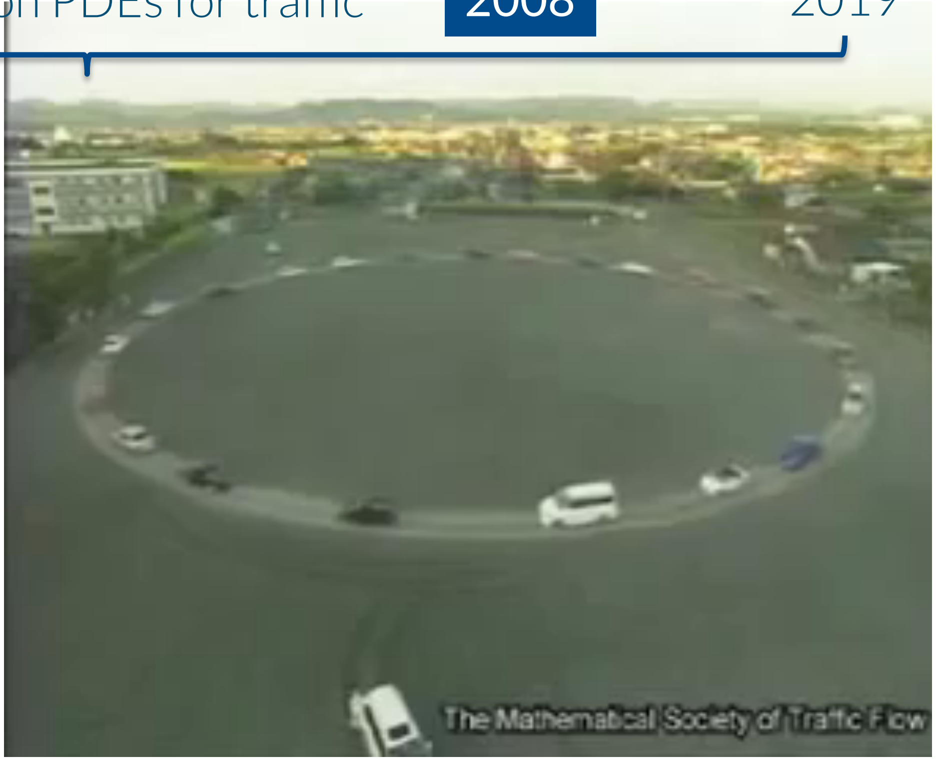
Partial differential equations (PDE)

Setting: 22 human drivers

Instructions: drive at 19 mph.

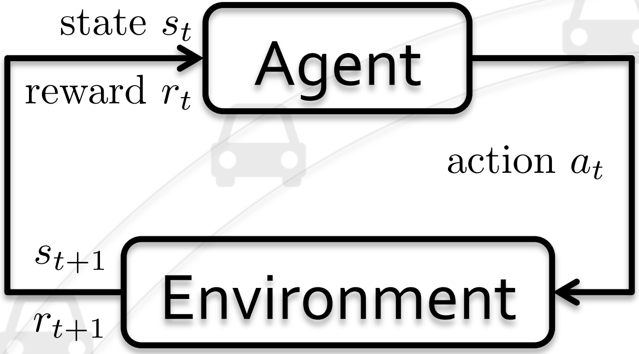
No traffic lights, stop signs, lane changes.

Traffic jams still form.



Video credits: NewScientist.com

Deep reinforcement learning (RL)



Decisions in urban systems:

Vehicle accelerations
Tactical maneuvers
Transit schedules
Traffic lights
Land use
Parking
Tolling
...

Goal:

learn policy $\pi : S \to A$
to maximize reward

$$\max_\theta \mathbb{E}\left[\sum_{t=0}^{H} r(s_t, a_t) \,\middle|\, \pi_\theta\right]$$

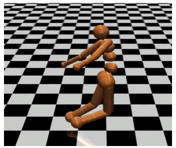
Cumulative rewards, returns

Policy parameters (deep neural network)

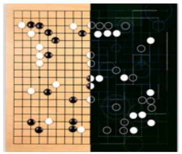
Global rewards

Average velocity
Energy consumption
Travel time
Safety, comfort

DQN (2015)



TRPO (2015)



AlphaGo (2016)

Single-lane control with RL

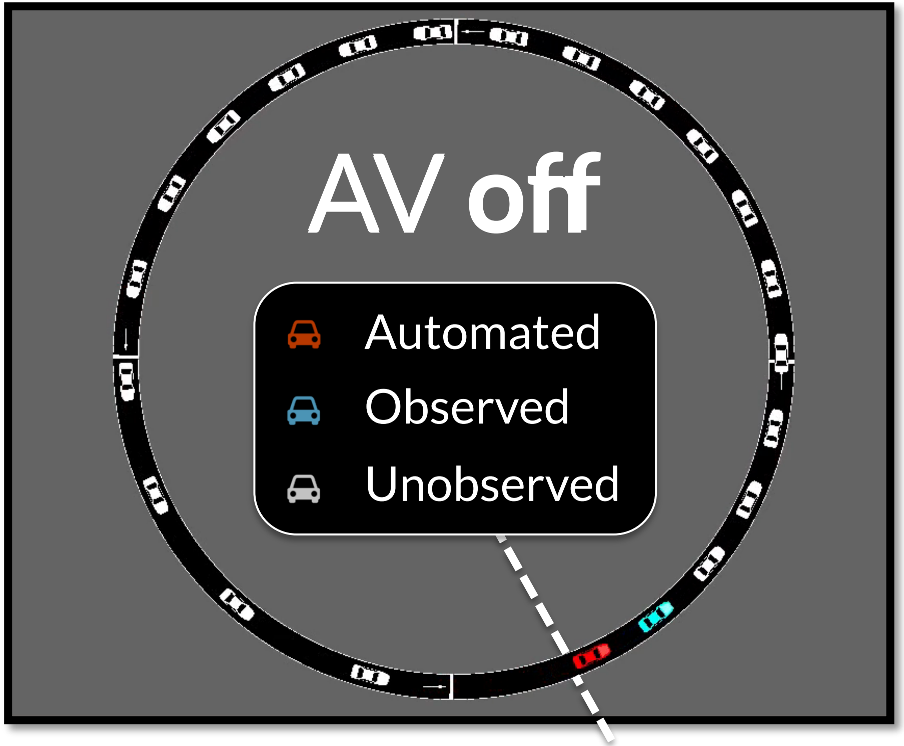
Setting: 1 AV, 21 human

Experiment

- **Goal**: maximize average velocity
- **Observation**: relative vel and headway
- **Action**: acceleration
- **Policy**: multi-layer perceptron (MLP)
- **Learning algorithm**: policy gradient

Results

- 1 AV: **+49%** average velocity
- **First near-optimal controller for single-lane**
- Uniform flow at **near-optimal velocity**
- **Generalizes** to out-of-distribution densities



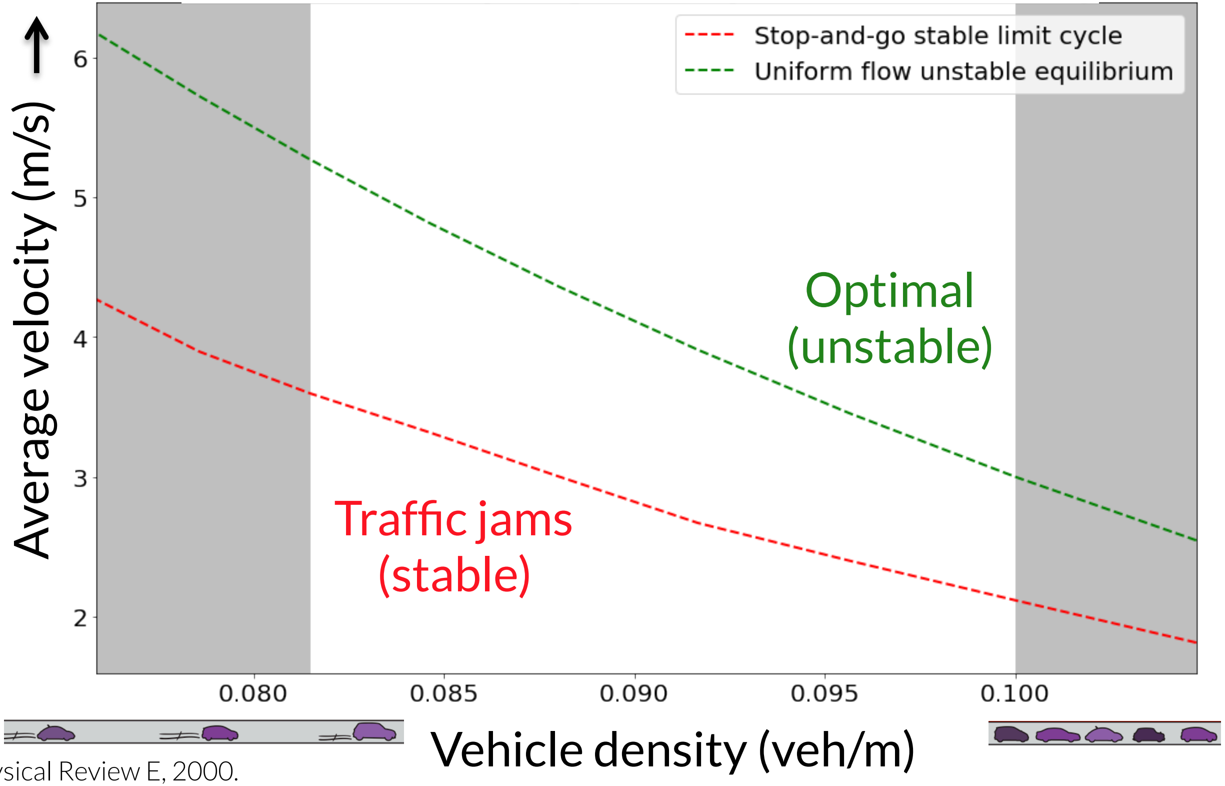
Wu, et al. CoRL, 2017; **Wu**, et al. IEEE T-RO, 2018

Single-lane: dynamical system equilibria

Human driver model

Intelligent Driver Model (IDM) [Treiber, et al. 2000]

Average velocity vs traffic density

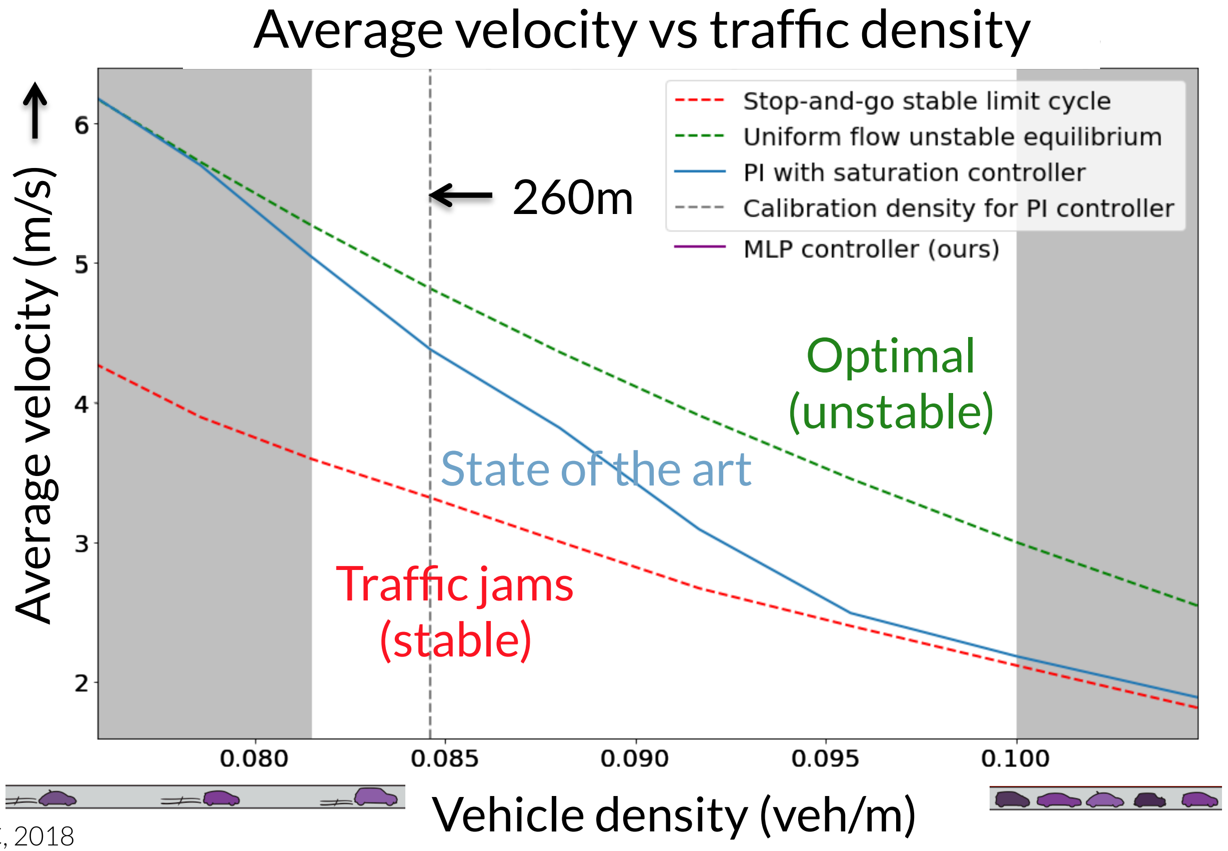


Wu, et al. CoRL, 2017;
Wu, et al. IEEE T-RO, in review; Treiber, et al. Physical Review E, 2000.

Single-lane: state of the art policy

State of the art

Proportional-integral (PI) controller with saturation [Stern, et al. 2018]



Wu, et al. CoRL, 2017;
Wu, et al. IEEE T-RO, in review; Stern, et al. TR-C, 2018

Single-lane: learned policy via deep RL

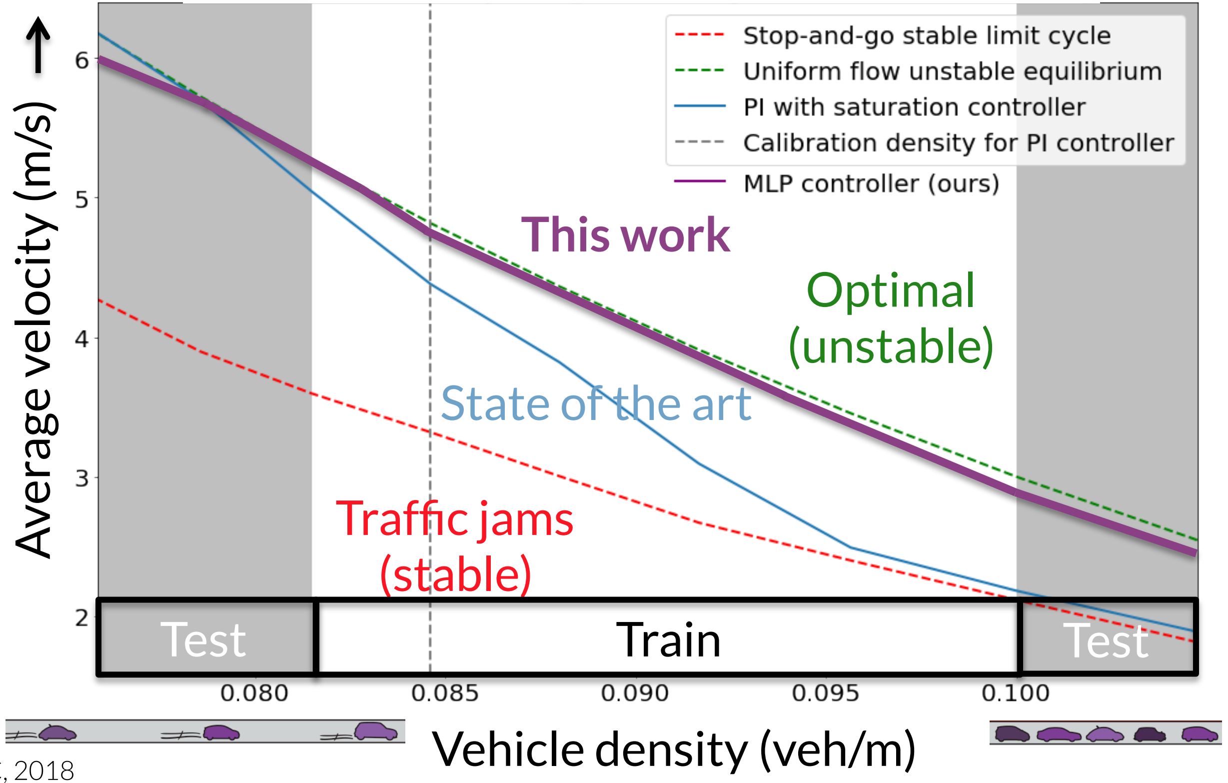
State of the art

Proportional-integral (PI) controller with saturation [Stern, et al. 2017]

Our results

- Near-optimal
- Generalizes to out-of-distribution traffic densities

Average velocity vs traffic density



Wu, et al. CoRL, 2017;
Wu, et al. IEEE T-RO, in review; Stern, et al. TR-C, 2018

Multi-lane traffic

Dynamics: mixed discrete-continuous cascaded nonlinear systems

Techniques:

- Partial differential equations
- Hybrid systems
- Formal methods
- Model predictive control

Lane-changing in traffic streams.
Laval, Daganzo. TR-B, 2006.

General lane-changing model MOBIL for car-following models.
Kesting, et al. TRR, 2007.

Multi-lane Reduction: A Stochastic Single-lane Model for Lane Changing.
Wu, et al. ITSC, 2017.

Wu, et al. IEEE T-RO, 2018

Multi-lane: mixed autonomy

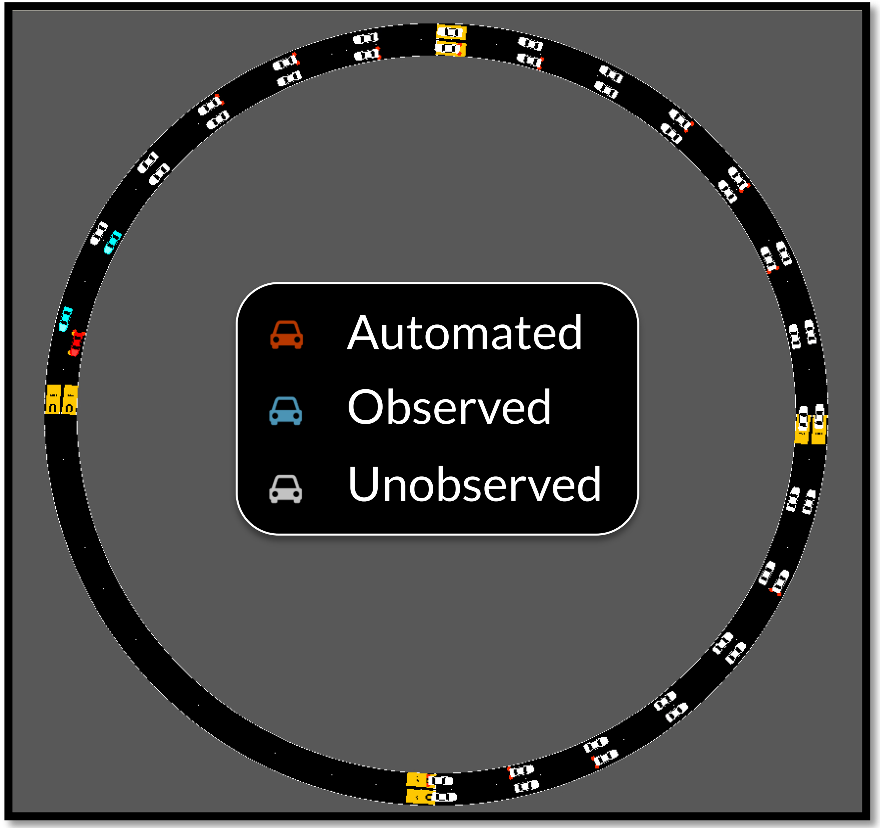
Setup: 1 AV, 41 human

Experiment

- Goal: maximize average velocity
- Observation:
  following headways, velocity
- Action: **acceleration and lane change**

Results

- **First stabilizing controller for multi-lane traffic**
- **Insight**: A single AV can stabilize multiple lanes of traffic
- **Emergent traffic break**



Wu, et al. IEEE T-RO, 2018

Multi-lane: traffic break

Setup: 1 AV, 41 human

Experiment

- Goal: maximize average velocity
- Observation:
  following headways, velocity
- Action: acceleration and lane change

Results

- **First stabilizing controller for multi-lane traffic**
- **Insight**: A single AV can stabilize multiple lanes of traffic
- **Emergent traffic break**

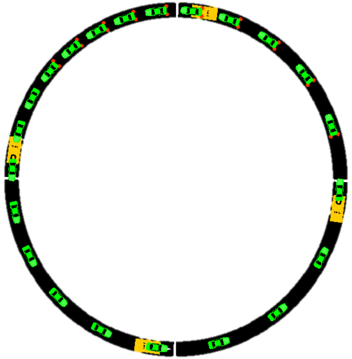
A traffic break found in the wild
(California Interstate Highway 8)

Wu, et al. IEEE T-RO, 2018

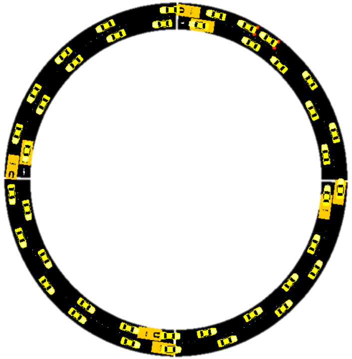
Traffic LEGO blocks

Benchmarks for autonomy in transportation

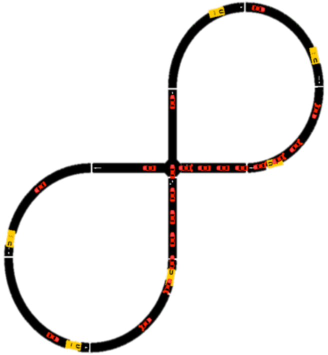
Single-lane



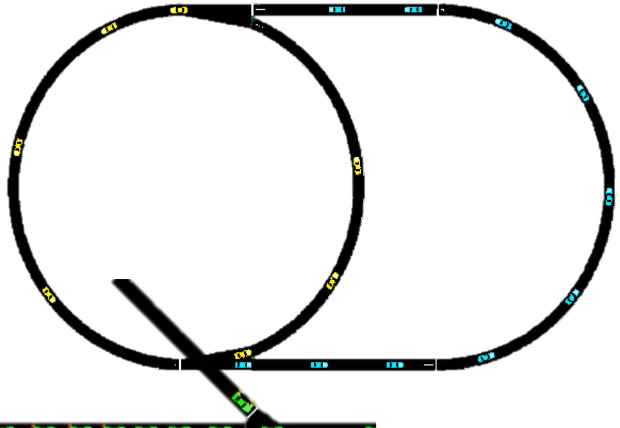
Multi-lane



Intersection



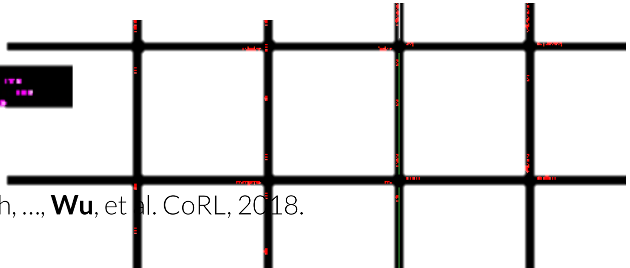
On/off-ramp



Straight highway

Bottleneck

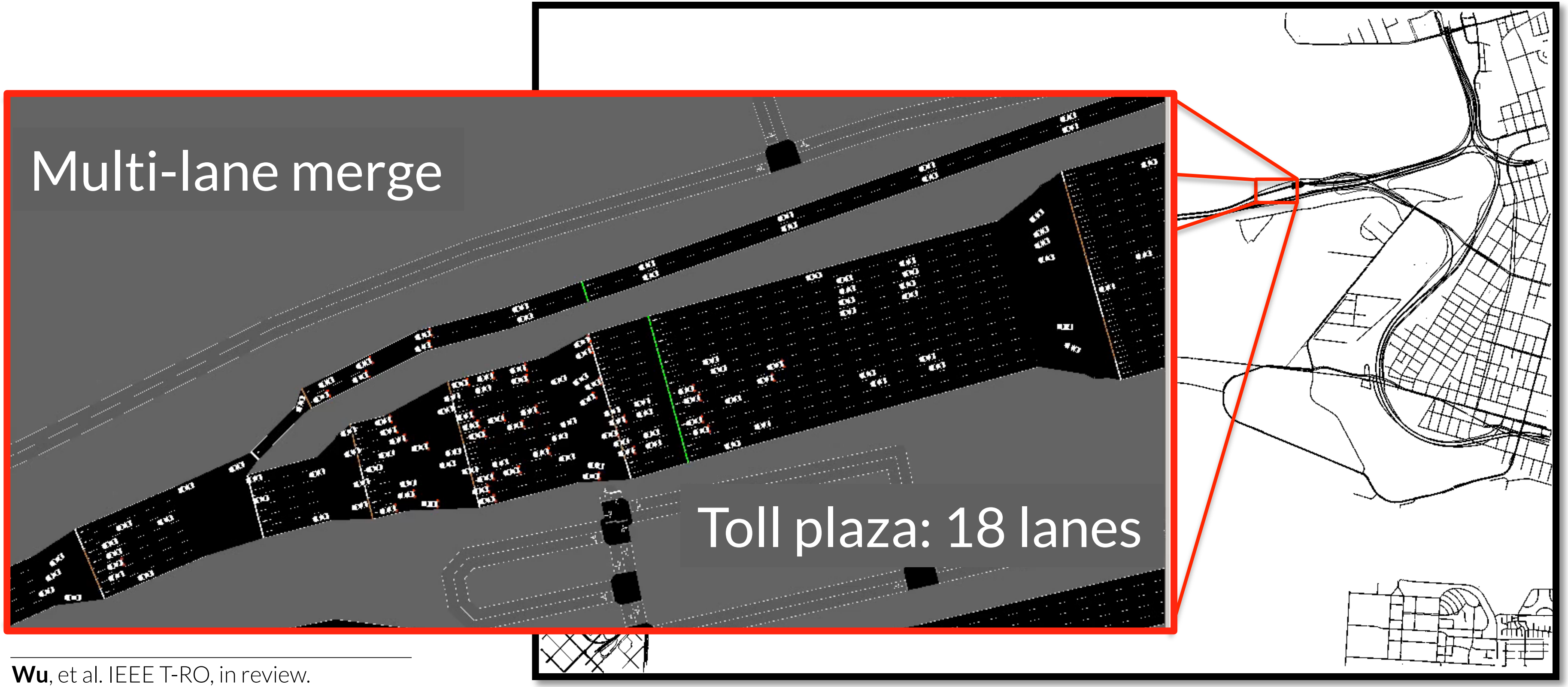
Grid network



Signalized intersection

Wu, et al. IEEE ITSC, 2017; **Wu**, et al. IEEE T-RO, in review; Vinitsky, Kreidieh, ..., **Wu**, et al. CoRL, 2018.

San Francisco Bay Bridge



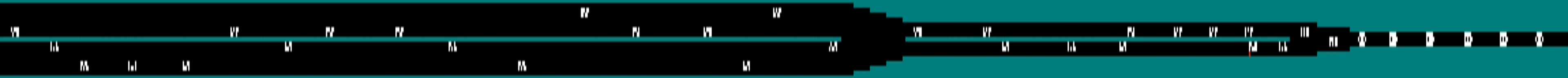
Wu, et al. IEEE T-RO, in review.

Core problem: traffic bottleneck

Eugene Vinitsky

Setting: No AVs

720 veh/hr



Phenomenon: capacity drop

Setting: 10% AVs

1020 veh/hr

40% improvement
Avoids capacity drop

Dynamics:

- Four lanes → Two lanes → One
- Cascaded nonlinear systems with right-of-way dynamics model, **merge conflicts**, and **excessive, fluctuating inflow**

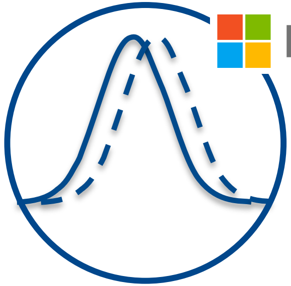
Vinitsky, Parvate, Kreidieh, **Wu**, Bayen. IEEE ITSC, 2018

Integrating autonomy: current & future

Operationalizing insights for control

Scalable RL for networked systems

Understanding adversarial driving



Microsoft

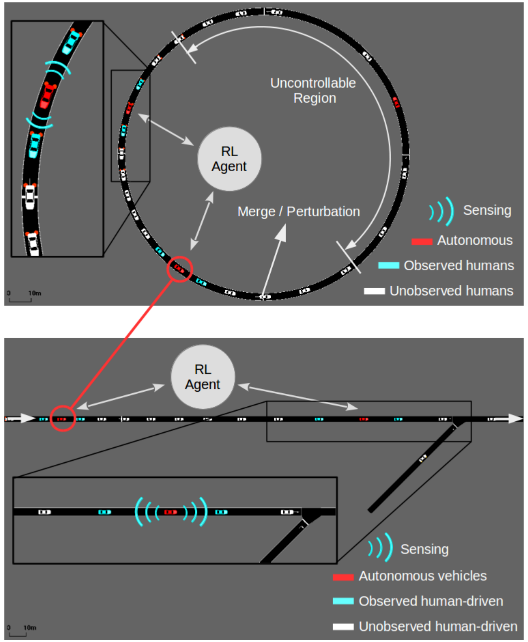
Coping with distribution shift

Scalable behavior modeling

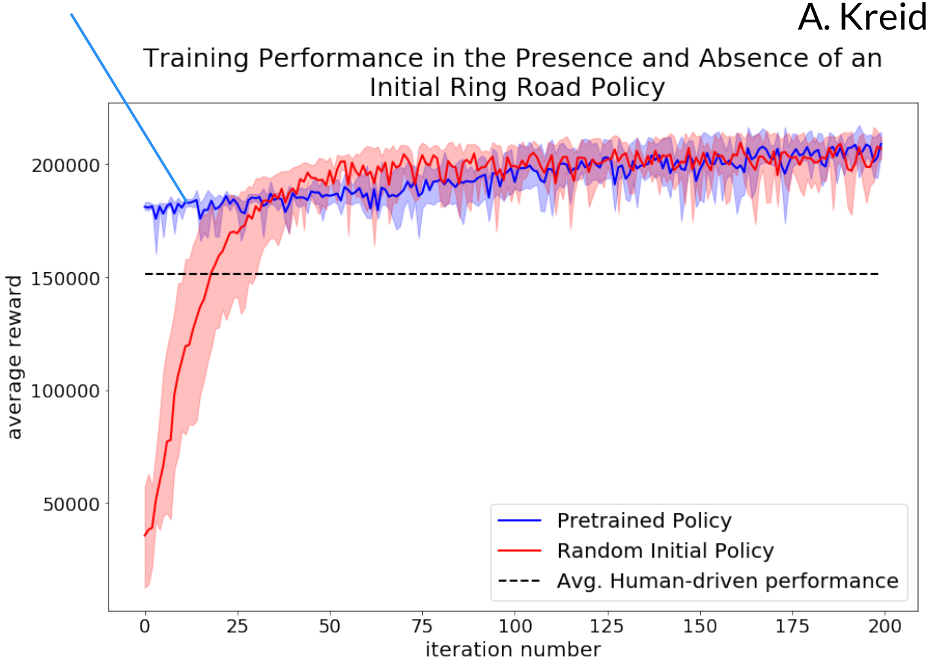
Urban decision support systems

Policy transfer

- Policies trained on ring roads, then deployed on straight roads



Initial performance boost



- Successful direct transfer!
- Closed → open networks

Kreidieh, **Wu**, Bayen, ITSC 2018.

Aravind Rajeswaren

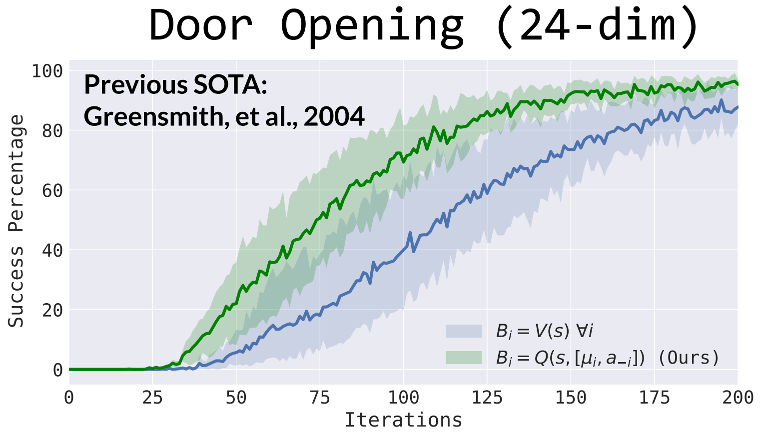
Uncertainty quantification and mitigation

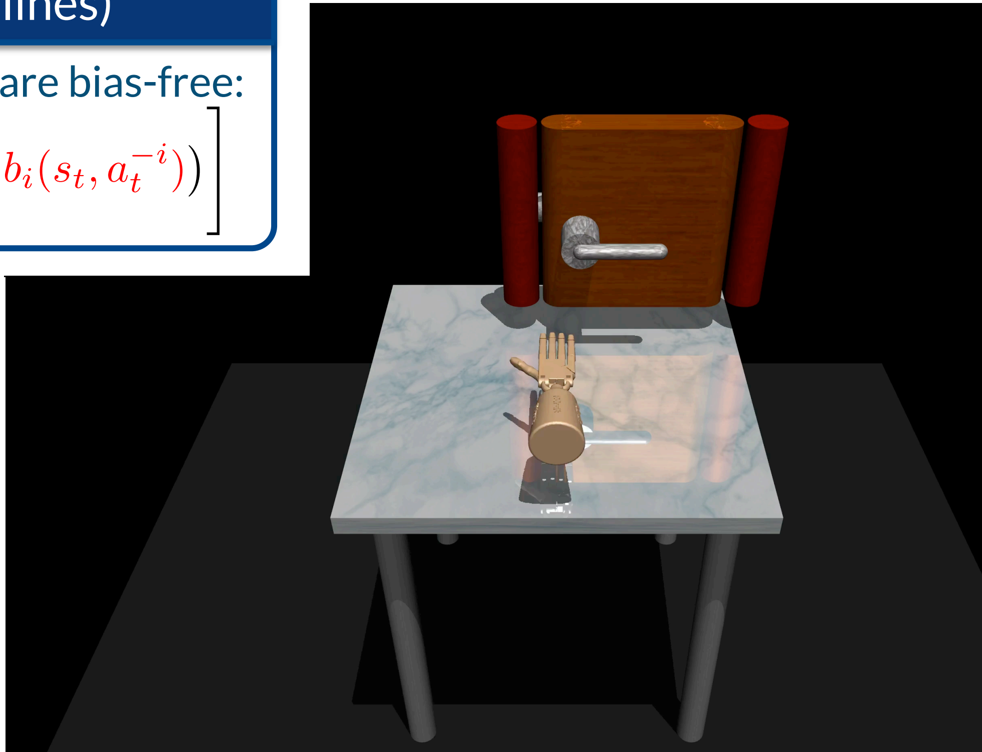
High-dimensional control: variance reduction for policy gradient via action-dependent baselines

Theorem (bias-free state-action baselines)

State-action baselines of the form $b_i(s_t, a_t^{-i})$ are bias-free:

$$g = \mathbb{E}\left[\sum_{i=1}^{m} \nabla_\theta \log \pi_\theta(a_t^i | s_t) \left(R(s_t, a_t) - b_i(s_t, a_t^{-i})\right)\right]$$





Wu, Rajeswaren, et al. ICLR, 2018; Rajeswaran, et al. arXiv, 2017.

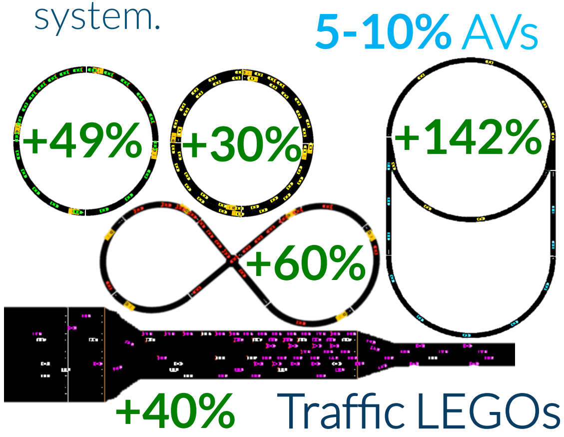
Integrating autonomy into urban systems

Challenge:
Vast uncertainty in future urban systems due to autonomy.

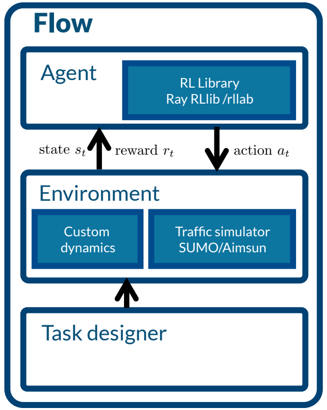
Approach:
Deep reinforcement learning (RL) provides **understanding** for integration of autonomy.

Findings:

- **Automatically discovered traffic controllers**
- **Small % of AVs** greatly affect **traffic dynamics**, which in turn, affects all parts of the urban system.



Flow: open source project to enable RL for traffic control
`flow-project.github.io`



Cathy Wu
cathywu@mit.edu

Cathy Wu - cathywu@mit.edu

Collaborators & Partners


Alexandre Bayen


Pieter Abbeel


Igor Mordatch

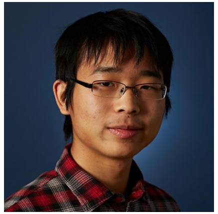

Alekh Agarwal


Adith Swaminathan


Eugene Vinitsky

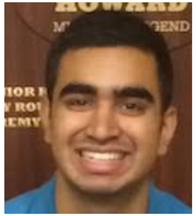

Aboudy Kreidieh


Kanaad Parvate

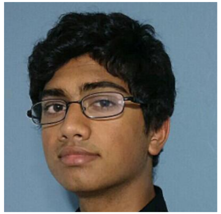

Rocky Duan


Aravind Rajeswaren


Kathy Jang


Nishant Kheterpal


Leah Dickstein


Ananth Kuchibhotla


Nathan Mandi

NSF

FLOW

Microsoft

aws

OpenAI

riselab UC Berkeley

Berkeley DeepDrive

BAIR BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

CALIFORNIA PATH