Mixed Autonomy Traffic

A Reinforcement Learning Perspective

Cathy Wu | Assistant Professor

LIDS, CEE, IDSS

IPAM Safe Operation of Connected and Autonomous Vehicle Fleets Workshop, 2020

Counterfactual reasoning with reinforcement learning?

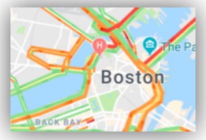
- *Motivation*: Quantify impact of technology on societal systems
- Pace of change & complexity is increasing

Environment

City

Agent





Uber



Reward

Impacts on:

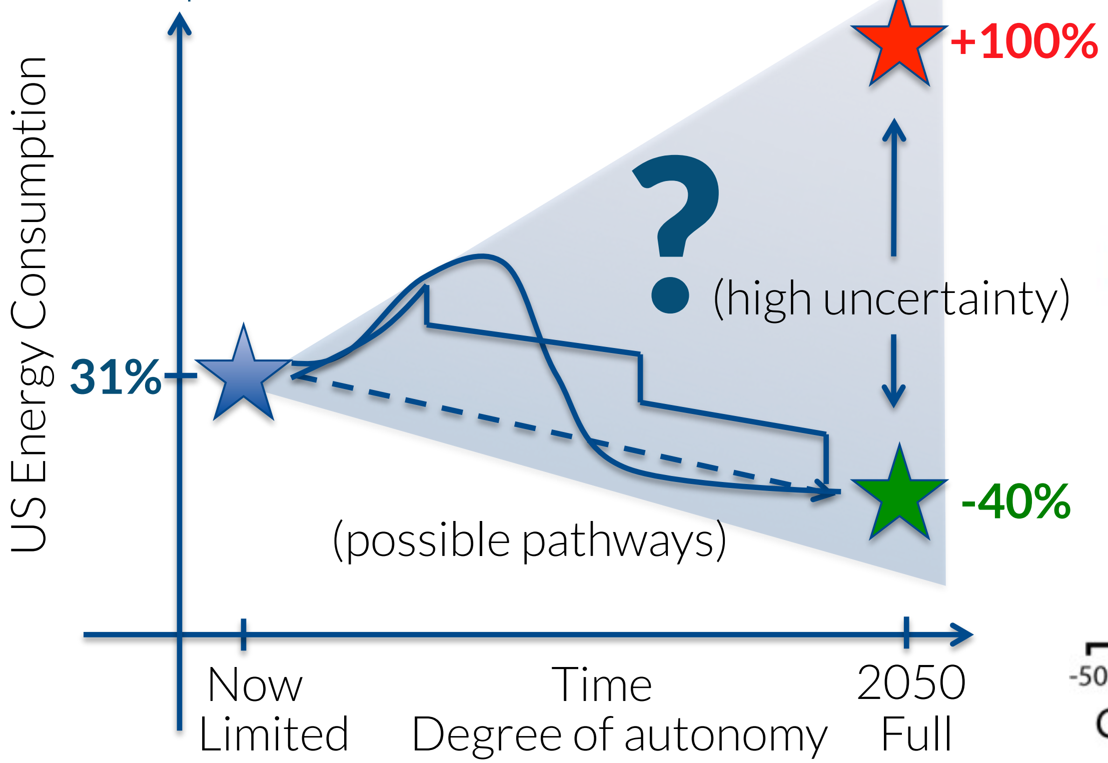
- Public safety & health
- Economic wellbeing
- Sustainability
- Resilience
- Equity & fairness

Policy evaluation → quantify impact

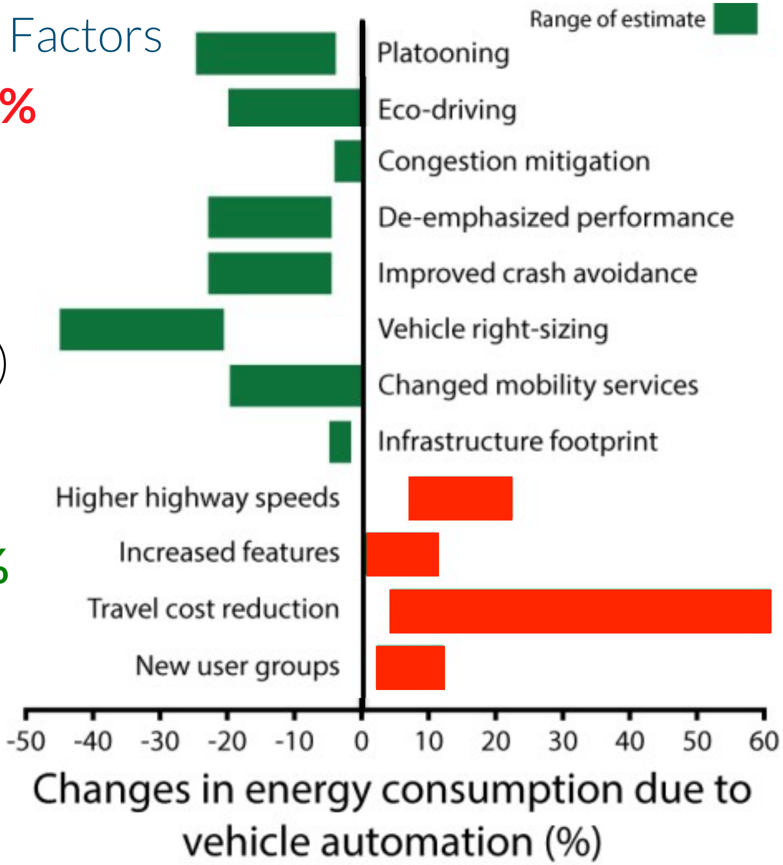
Policy learning → improve impact

Years 2020 to 2049: Mixed autonomy

Transportation in the US



Factors



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

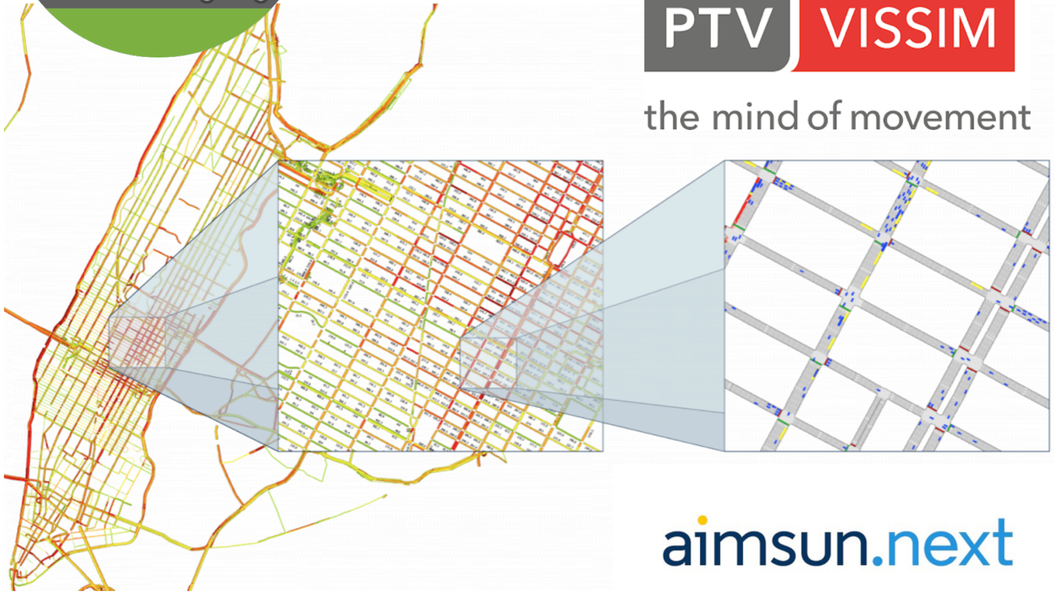
Urban simulation

PARAMICS MICROSIMULATION

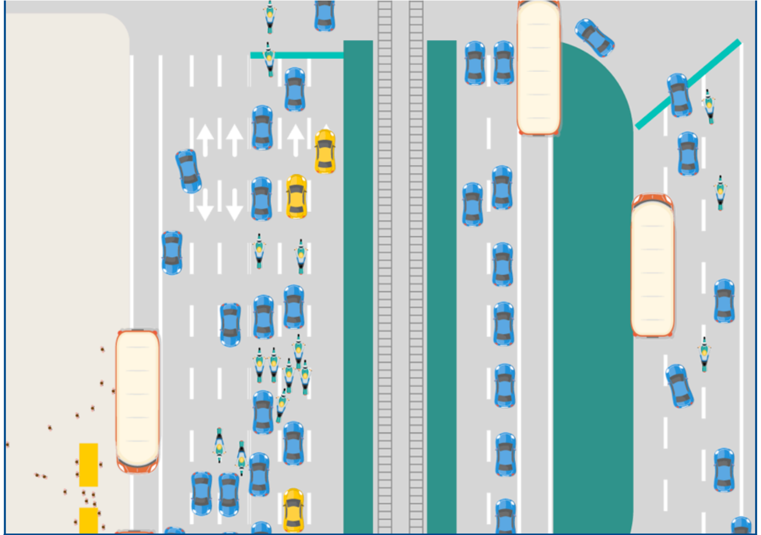
SUMO

PTV VISSIM

the mind of movement

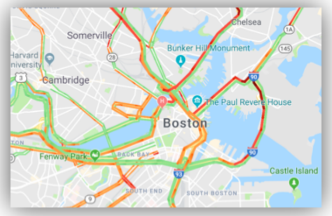


aimsun.next



Axes of difficulty in mixed autonomy traffic

500K vehicles,
10K roads



Scope
(# vehicles,
road network)

1-20
vehicles,
1 road

Curse of
dimensionality

Partial differential
equations

Model predictive
control

PID control

Feedback
linearization

Part 3: Then what?

Jackson networks

Part 2: RL?

Reservation
systems

Part 1: RL?

Polling systems

0%

Degree of autonomy

100%

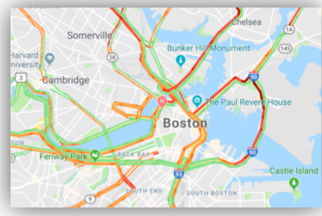
Low
uncertainty
in **rewards**

High
uncertainty
in rewards &
transitions

Low
uncertainty in
transitions

Axes of difficulty in mixed autonomy traffic

500K vehicles,
10K roads



Scope
(# vehicles,
road network)

1-20
vehicles,
1 road

Curse of
dimensionality

Part 3: Then what?

Jackson networks

Partial differential
equations

Model predictive
control

PID control

Feedback
linearization

Part 2: RL?

Reservation
systems

<u>Part 1: RL?</u>

Polling systems

0%

Degree of autonomy

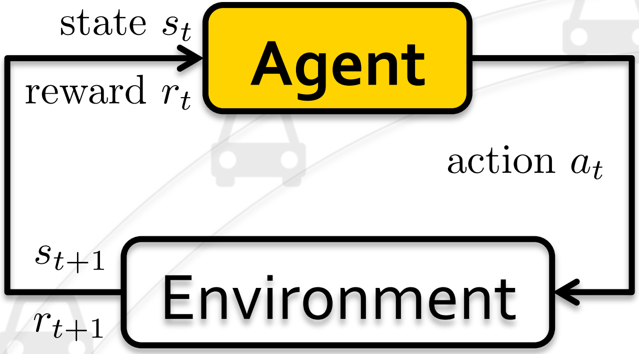
100%

Low
uncertainty
in **rewards**

High
uncertainty
in rewards &
transitions

Low
uncertainty in
transitions

Deep reinforcement learning (RL)



Goal:

learn policy $\pi : S \rightarrow A$
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

Decisions in transportation:

Vehicle accelerations

Tactical maneuvers

Transit schedules

Traffic lights

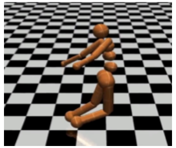
Land use

Parking

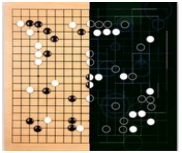
Tolling

...

DQN (2015)



TRPO (2015)

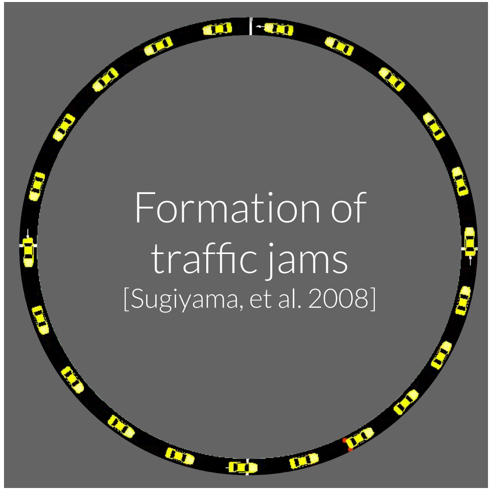


AlphaGo (2016)

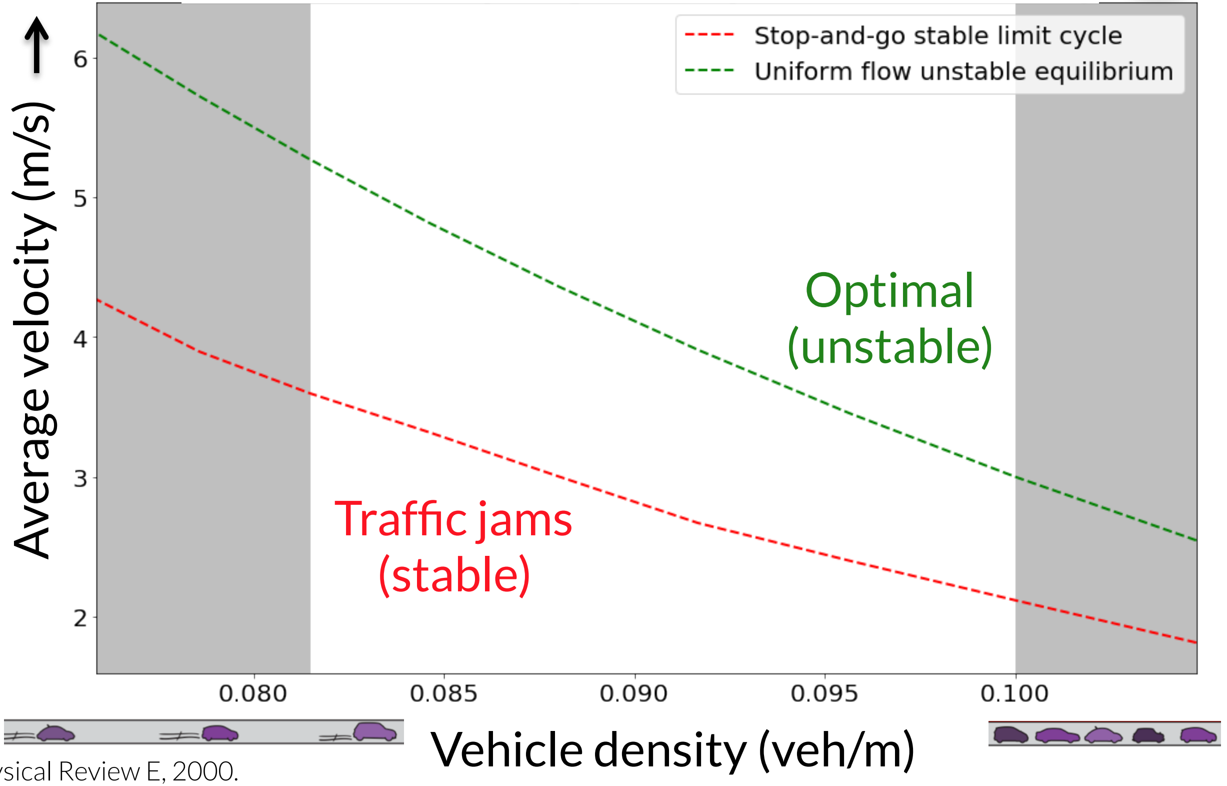
Single-lane: dynamical system equilibria

Human driver model

- Car-following model
- $\ddot{x}_i = f(v_i, v_i - v_{i-1}, x_i - x_{i-1})$
- Intelligent Driver Model [Treiber, et al. 2000]



Average velocity vs traffic density

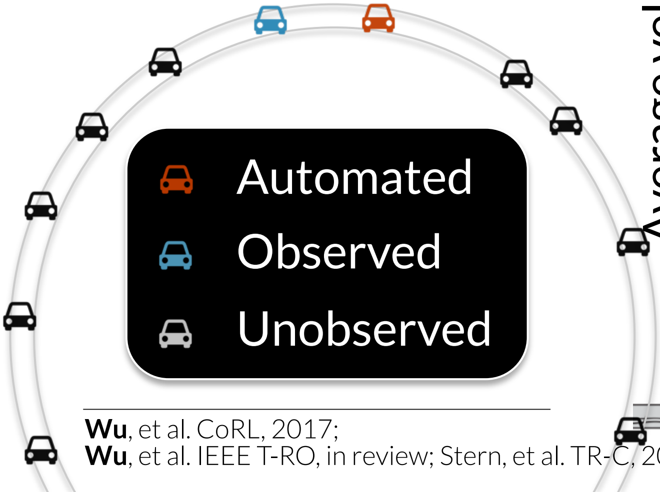


Wu, et al. CoRL, 2017;
Wu, et al. IEEE T-RO, in review; Treiber, et al. Physical Review E, 2000.

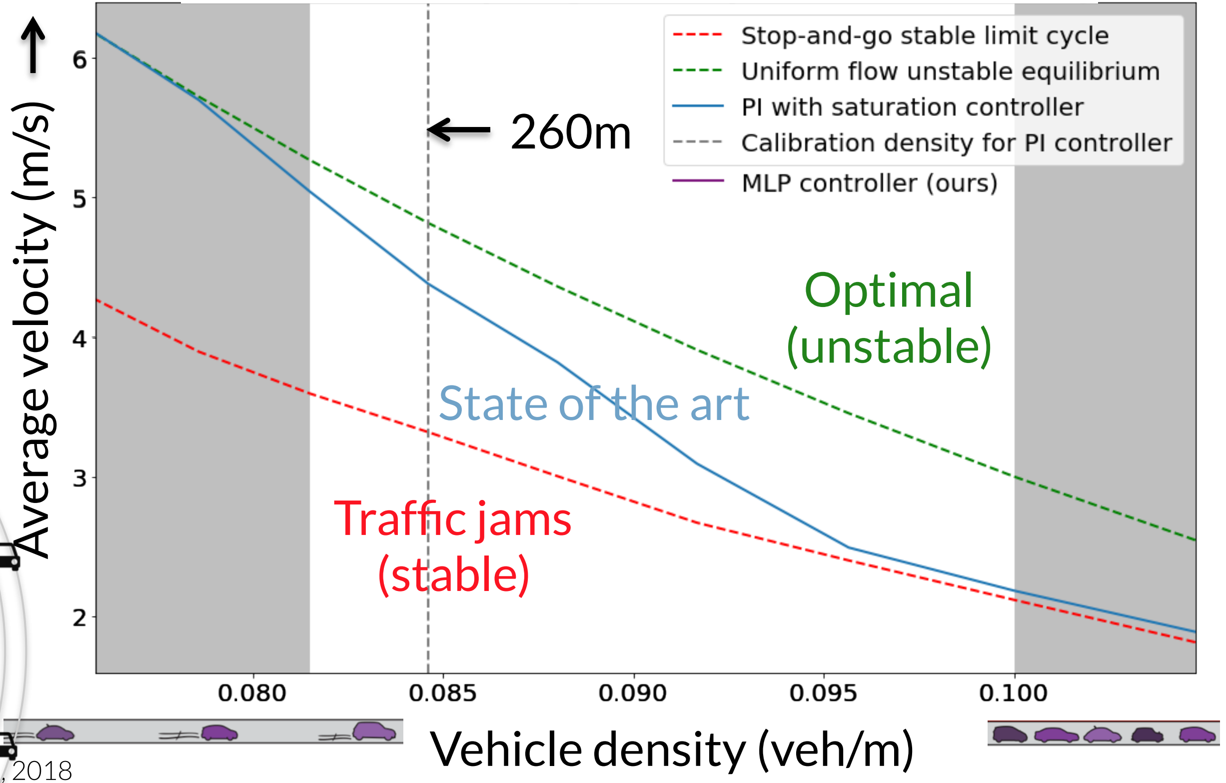
Single-lane: state of the art policy

State of the art

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



Average velocity vs traffic density



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

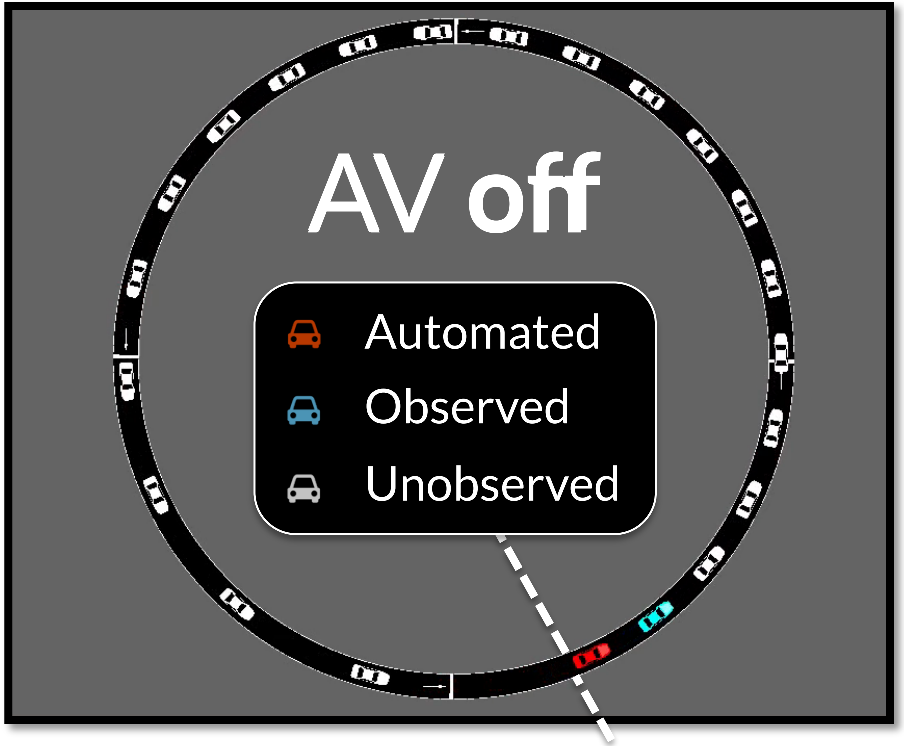
Mixed autonomy traffic (single-lane)

Wu, et al.
Stern, et al. 2017
1955 Sugiyama, et al. 2008 2019

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, in review.

Single-lane: learned policy via deep RL

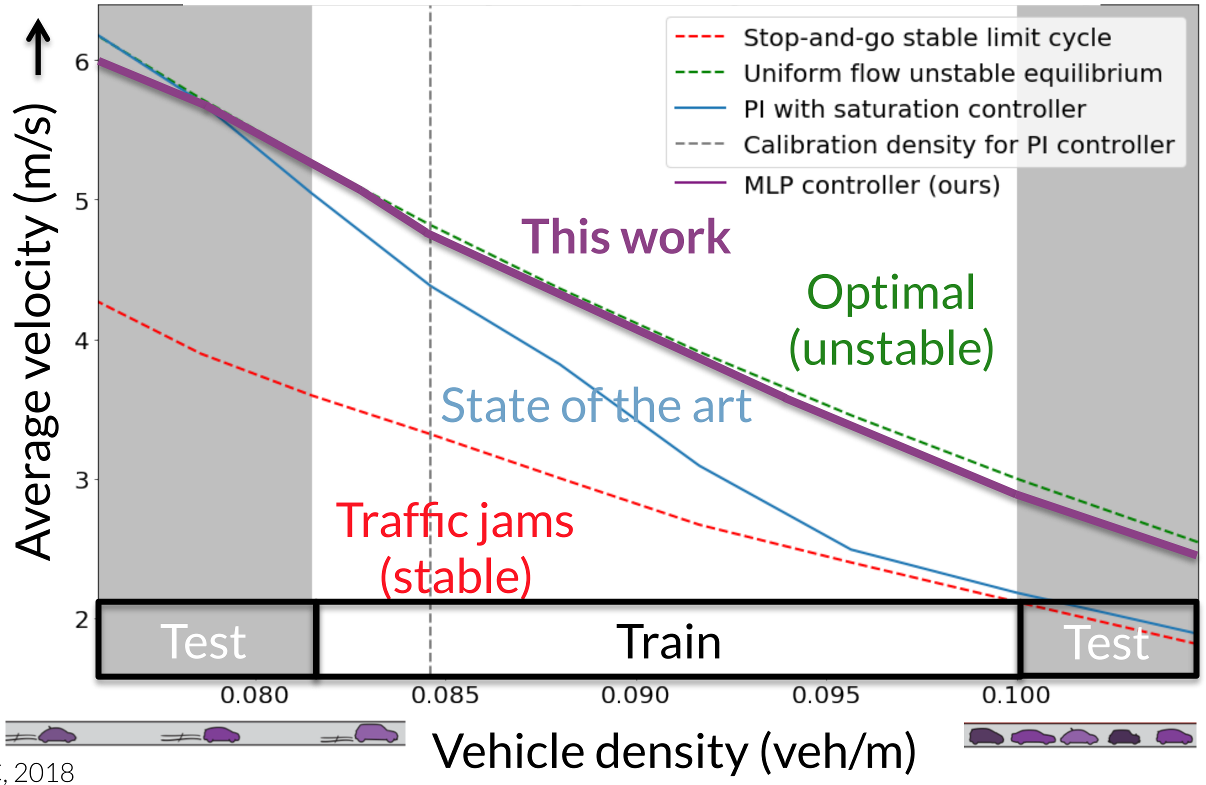
State of the art

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

Our results

- Near-optimal
- Generalizes to out-of-distribution traffic densities
- Memory not needed

Average velocity vs traffic density

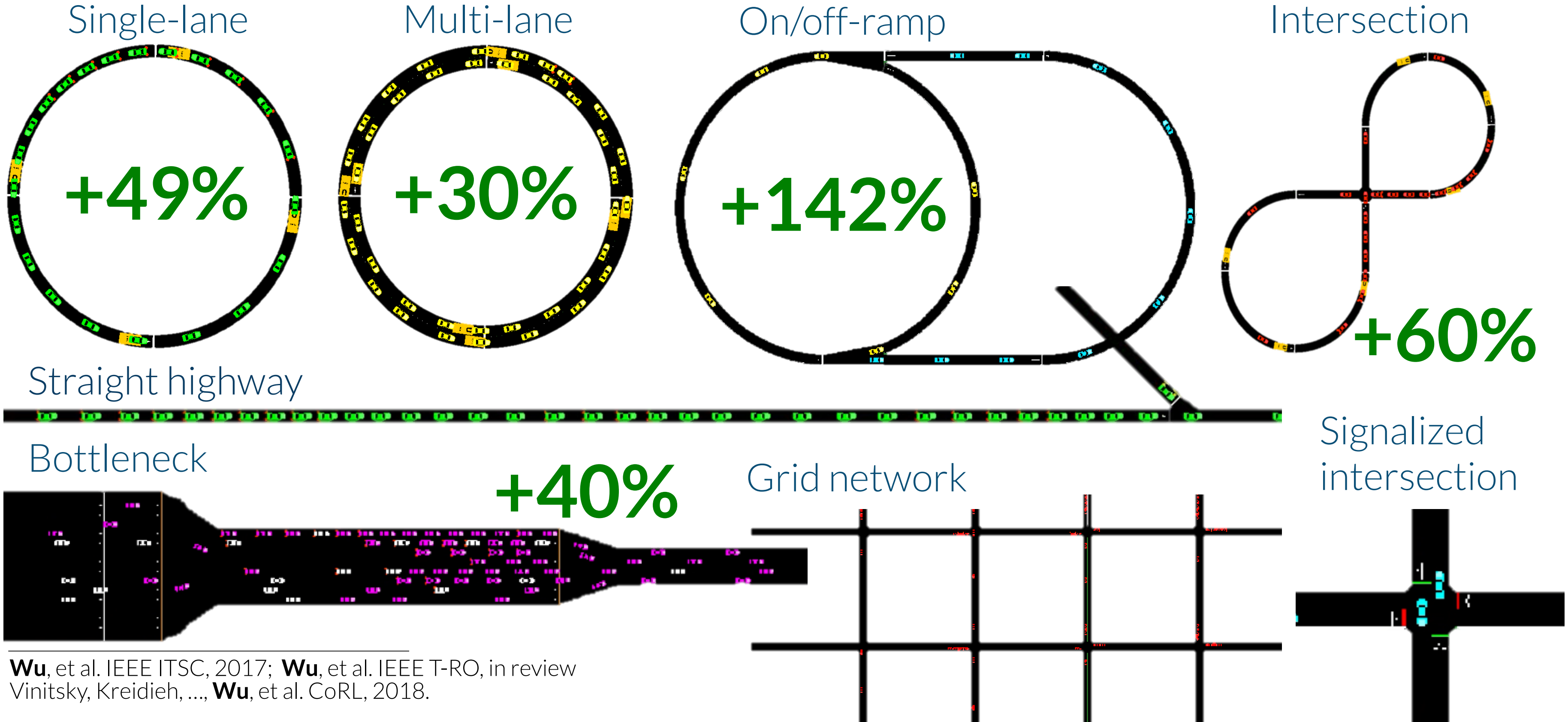


Wu, et al. CoRL, 2017;
Wu, et al. IEEE T-RO, in review; Stern, et al. TR-C, 2018

Traffic LEGO blocks

5-10% AVs

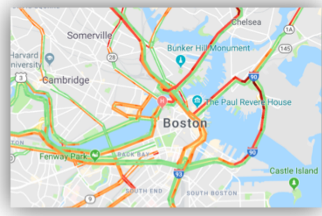
Benchmarks for autonomy in transportation



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

Axes of difficulty in mixed autonomy traffic

500K vehicles,
10K roads



Scope
(# vehicles,
road network)

Curse of
dimensionality

Partial differential
equations

Model predictive
control

PID control

Feedback
linearization

Part 3: Then what?

Jackson networks

<u>Part 2: RL?</u>

Reservation
systems

Polling systems

Part 1: RL?

0%

Degree of autonomy

100%

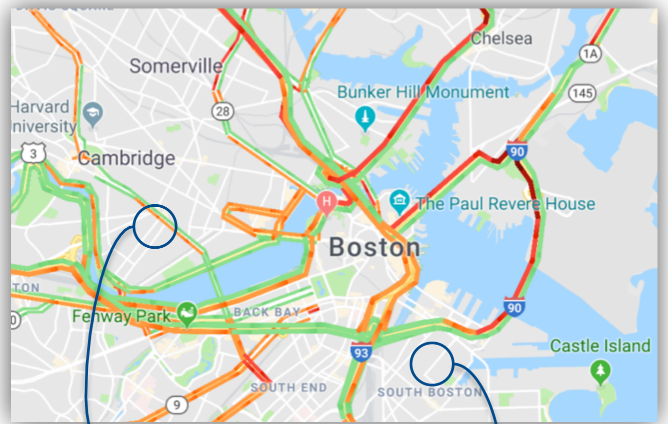
Low
uncertainty
in **rewards**

High
uncertainty
in rewards &
transitions

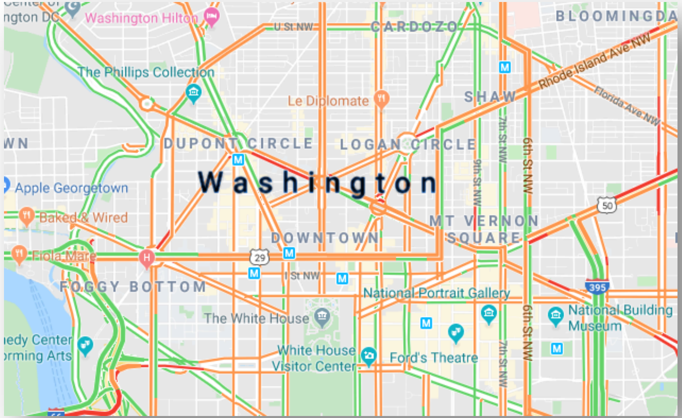
Low
uncertainty in
transitions

Challenge: combinatorial number of environments

A critical challenge to scaling deep reinforcement learning



≠

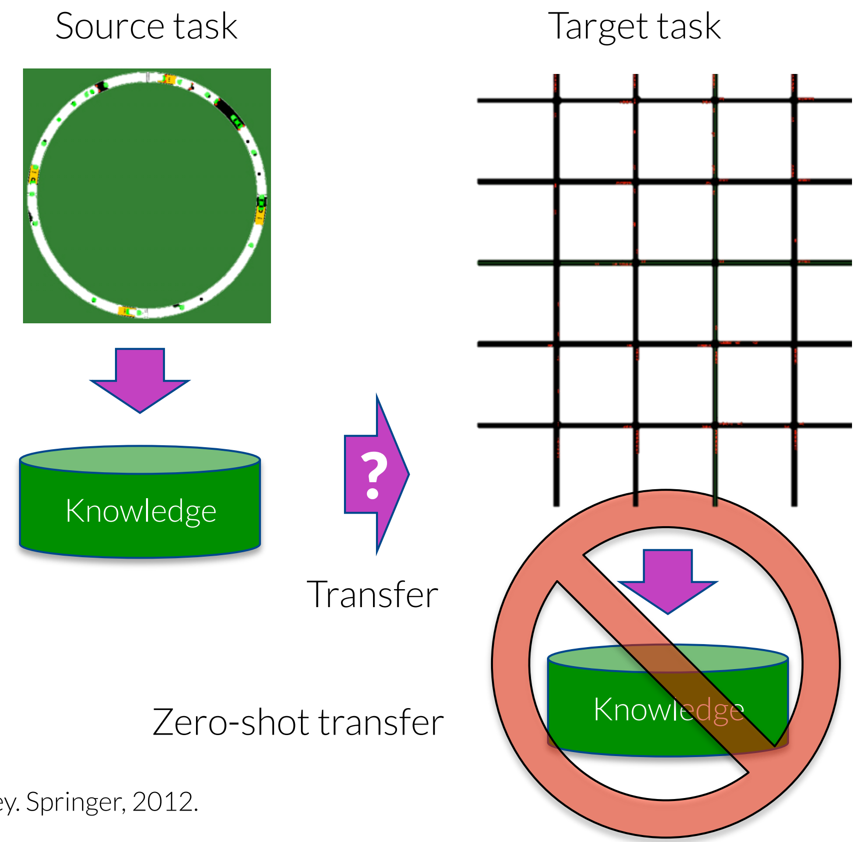


≠

Elements: Road network, Roadway signage, Rules of the road, Types of vehicles, Speed limits, Traffic lights, # Lanes, Driver behavior, …

Transfer learning across networks

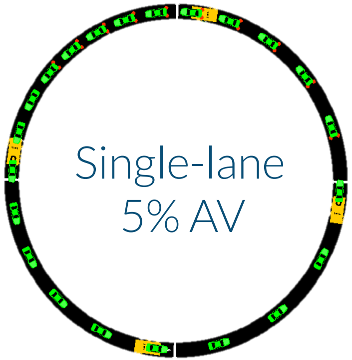
- *Research question*:
  Can knowledge be transferred across traffic scenarios?
- *Transfer learning*: The use of knowledge gained from a **source task** to bias the learning process on a **target task** towards a set of good hypotheses.
- *Zero-shot transfer*: Extreme setting where no learning is done on the target task. Out-of-distribution generalization.



Lazaric, Alessandro. Transfer in Reinforcement Learning: a Framework and a Survey. Springer, 2012.

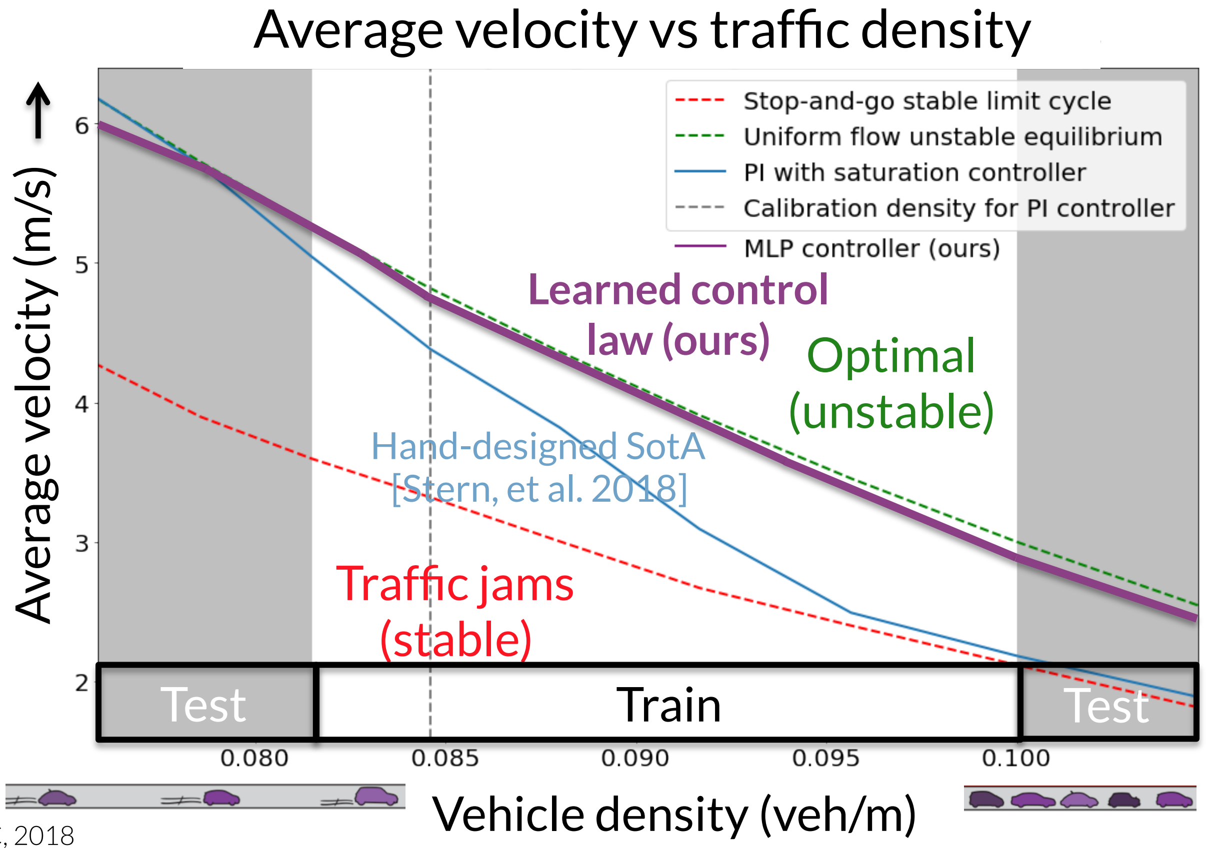
Zero-shot transfer

Circular track → More/less dense circular track



Our results

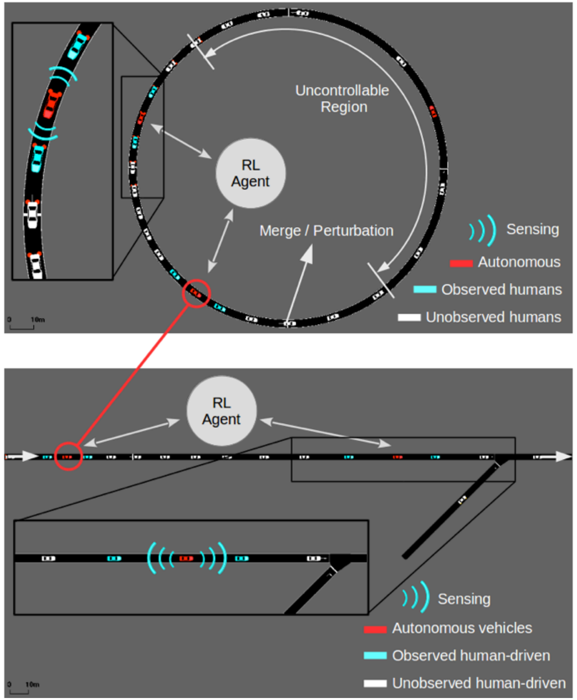
- Near-optimal
- Zero-shot transfer: Generalizes to out-of-distribution traffic densities



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

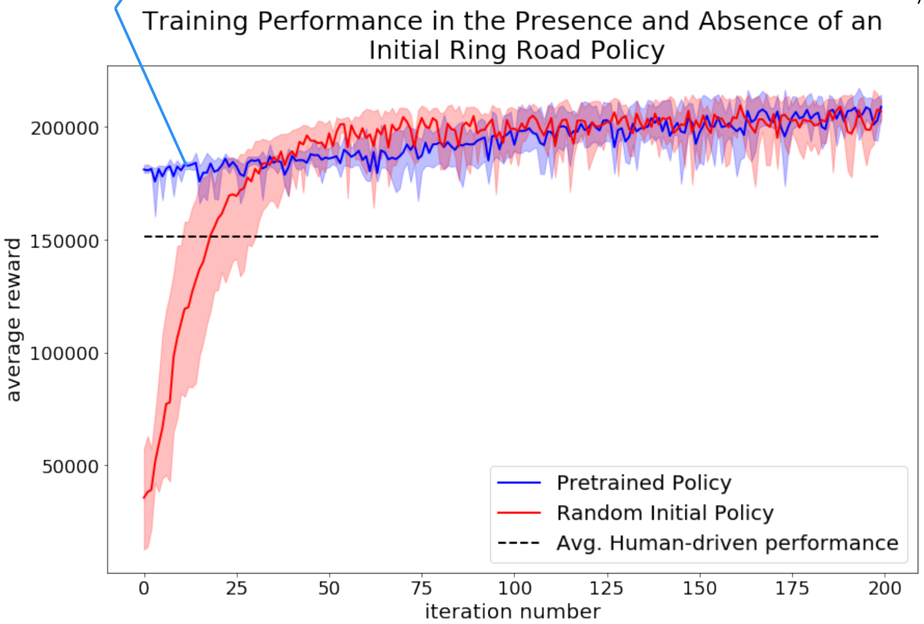
Transfer learning

Circular roads → Straight roads



Initial performance boost

Single-lane
5% AV



Aboudy Kreidieh
(UC Berkeley)

- Successful direct transfer!
- Closed → open networks

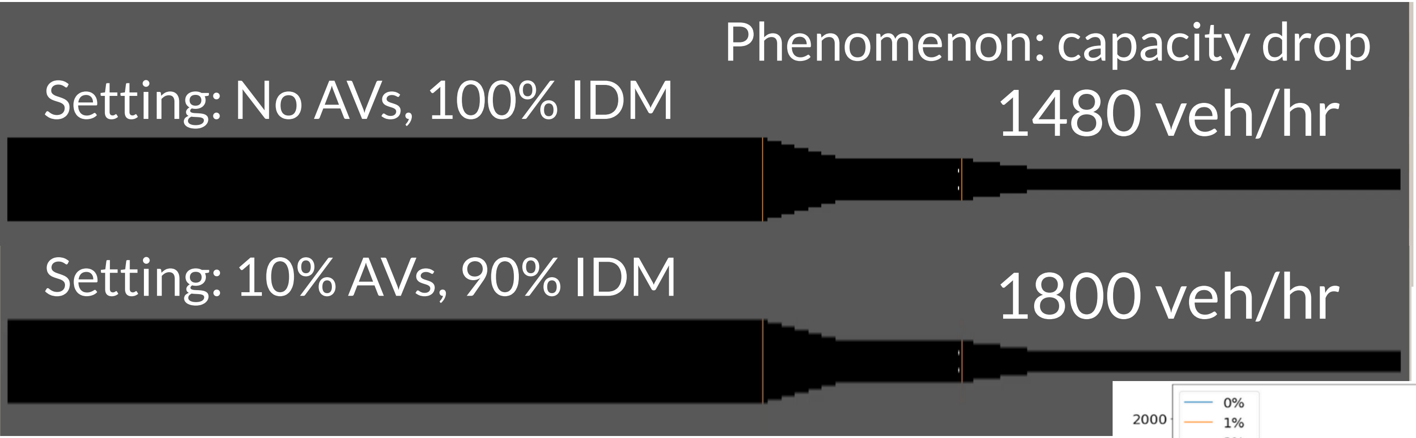
Kreidieh, **Wu**, Bayen, ITSC 2018.

Zero-shot transfer

One bottleneck → Many different bottlenecks

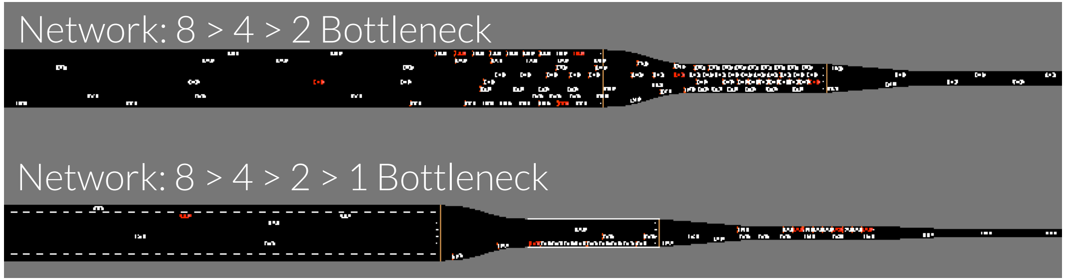
[Ongoing research]

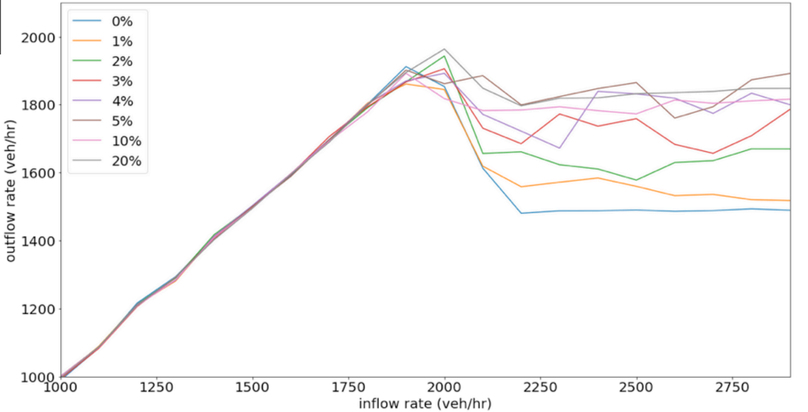
Zee Yan



Results:
- 22% improvement
- Avoids capacity drop
- Learned policy transfers to different inflow rates, number of lanes, and percent of automated vehicles

Successful transfer:





Scenario: varying inflow rates, varying % AVs.

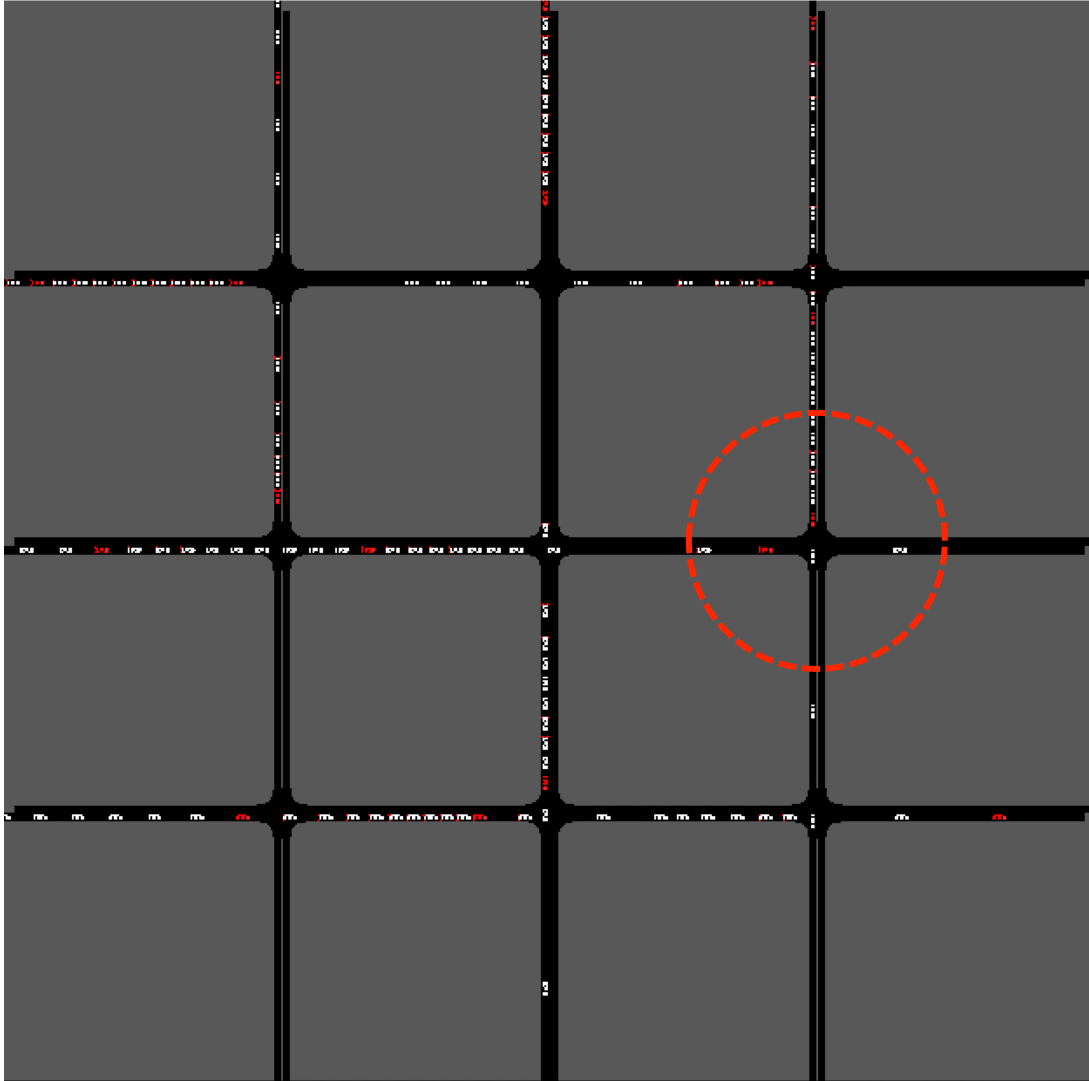
*Capacity drop experiment is a variation of: Vinitsky, Parvate, Kreidieh, **Wu**, Bayen. IEEE ITSC, 2018

Zero-shot transfer

Bottleneck → Grid?!

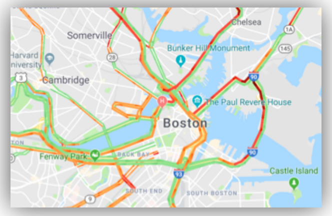
26% improvement over human baseline (IDM)

No fine-tuning!



Axes of difficulty in mixed autonomy traffic

500K vehicles,
10K roads



Scope
(# vehicles,
road network)

1-20
vehicles,
1 road

Curse of
dimensionality

Partial differential
equations

Model predictive
control

PID control

Feedback
linearization

Part 3: Then what?

Jackson networks

<u>Part 2: RL?</u>

Reservation
systems

Polling systems

Part 1: RL?

0%

Degree of autonomy

100%

Low
uncertainty
in **rewards**

**High
uncertainty**
in rewards &
transitions

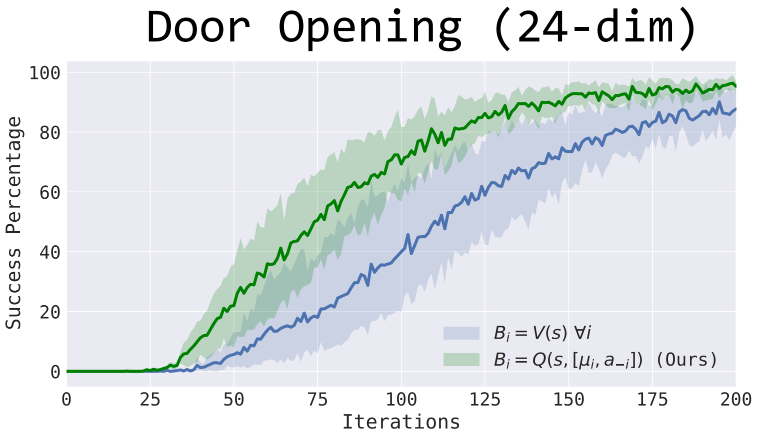
Low
uncertainty in
transitions

High-dimensional control

Key challenge in cooperative multi-agent systems: curse of dimensionality

Variance reduction in policy gradient methods

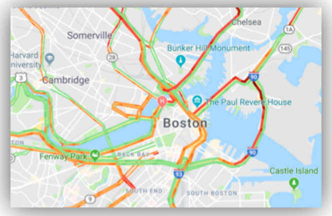
Key idea: factorizing stochastic policies → opportunity for improved control variates



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

Axes of difficulty in mixed autonomy traffic

500K vehicles,
10K roads



Scope
(# vehicles,
road network)

Curse of
dimensionality

Part 3: Then what?

Jackson networks

Partial differential
equations

Model predictive
control

PID control

Feedback
linearization

Part 2: RL?

Part 1: RL?

Reservation
systems

Polling systems

0%

Degree of autonomy

100%

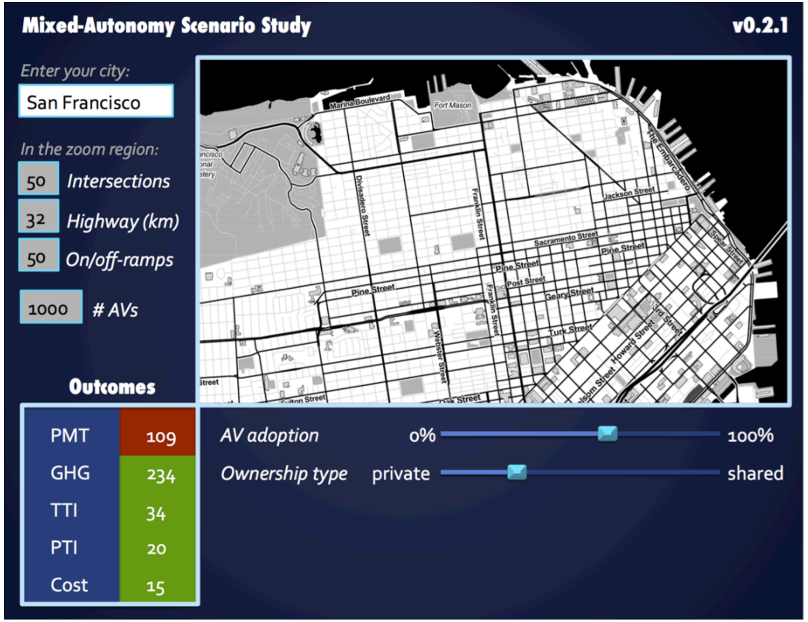
Low
uncertainty
in **rewards**

High
uncertainty
in rewards &
transitions

Low
uncertainty in
transitions

Pathways toward reality

Physical tests & deployment



Insights for urban planning & industry

[Ongoing research]

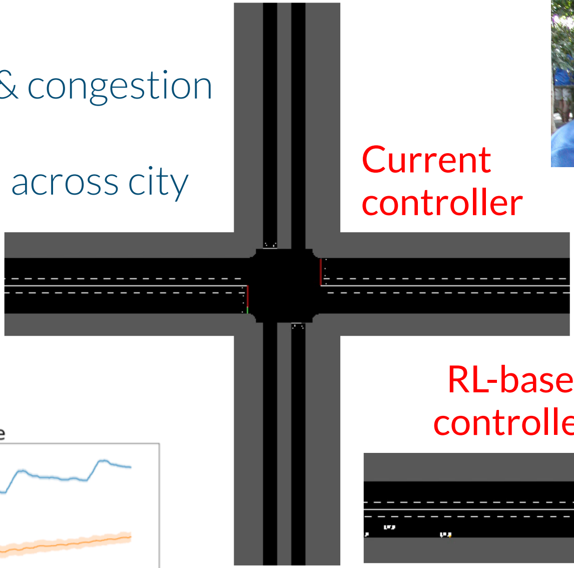
Automatic traffic signal optimization

Not just for futuristic automated vehicles!

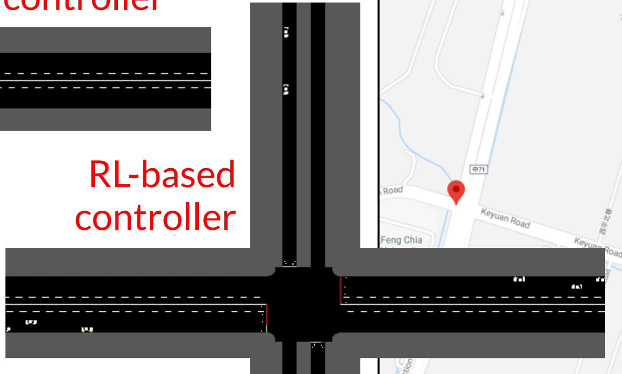
Dr. Fang-Pang Lin, Schumi Huang (NCHC)

Vindula Jayawardana

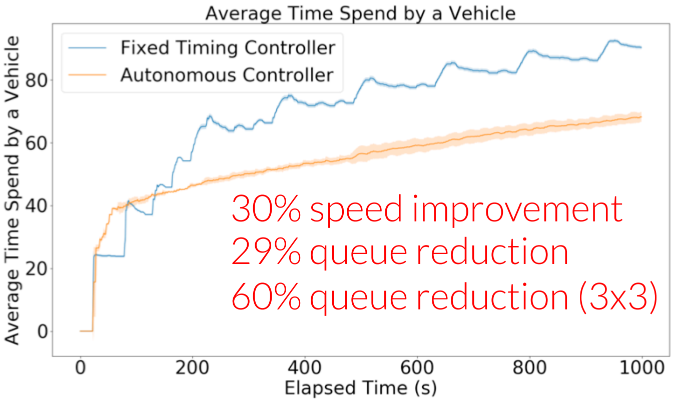
- Motivation:
  - Improve travel times & congestion
  - Mitigate air pollution
  - Improve coordination across city boundaries
- Same approach
- Potential for near-term benefits



Current controller



RL-based controller



MOST 科技部 Ministry of Science and Technology

臺中市政府 TAICHUNG CITY GOVERNMENT

NARLabs 財團法人國家實驗研究院

NCHC 國家高速網路與計算中心 National Center for High-performance Computing

Collaborators & Partners


Prof. Alexandre Bayen$^+$

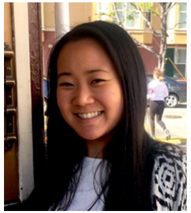

Eugene Vinitsky$^+$


Aboudy Kreidieh$^+$


Zee Yan

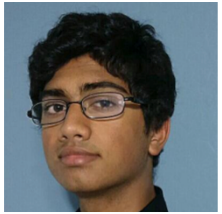

Vindula Jayawardana


Kanaad Parvate$^+$


Kathy Jang$^+$


Nishant Kheterpal$^+$


Leah Dickstein$^+$


Ananth Kuchibhotla$^+$


Nathan Mandi$^+$

IDSS
MIT INSTITUTE FOR DATA, SYSTEMS, AND SOCIETY

MIT

LIDS
LABORATORY FOR INFORMATION AND DECISION SYSTEMS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Civil and Environmental Engineering

Microsoft

NSF

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

FLOW

NARLabs 財團法人國家實驗研究院
NCHC 國家高速網路與計算中心
National Center for High-performance Computing

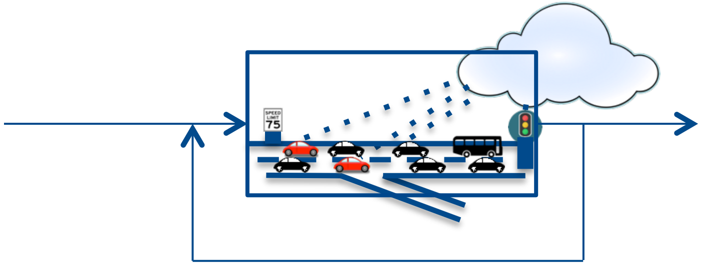
aws

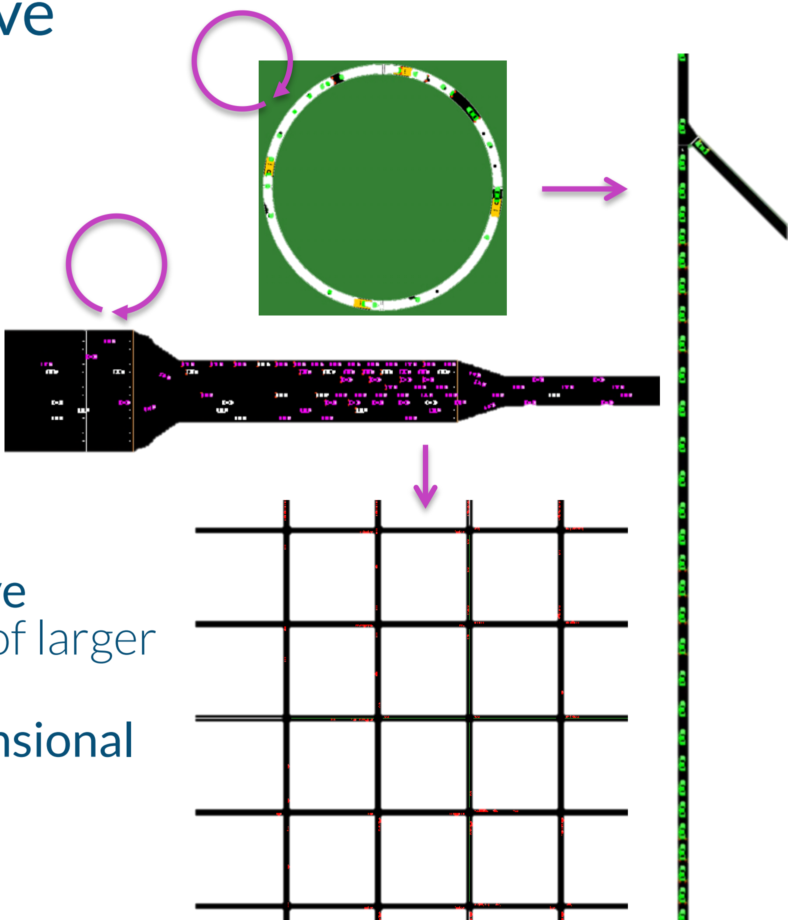
臺中市政府
TAICHUNG CITY GOVERNMENT

新竹市政府
HSINCHU CITY GOVERNMENT

$^+$ = UC Berkeley

Mixed Autonomy Traffic:
A Reinforcement Learning Perspective





- **Deep reinforcement learning** provides a pathway for understanding the impacts of mixed autonomy in urban systems
- **Transfer learning across networks can improve sample efficiency of RL** to enable the analysis of larger and more complex traffic scenarios
- Numerous opportunities to scale to **high dimensional control**
- There's everything left to do