Integrating Autonomy into Urban Systems

Cathy Wu | Assistant Professor
CEE, IDSS, LIDS
2011: The self-driving dream

- 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 ($160B)

- Greenhouse gas emissions:
  - 28% from transportation

- Access to mobility:
  - 30% of population
  - 20% youth or elderly
  - 10% disabled (ages 18-64)

2011-16: Driving the dream

2007: DARPA Urban Challenge
2011: MIT Maslab tournament (Mobile Autonomous Systems LABoratory)
2013: Cellpath demand inference
2015: Clustering for ridesharing
2016: Routing for Waymo service

2050: How will self-driving cars change urban mobility?

Short answer: it is **highly uncertain**.

Transportation today: **31% US energy consumption**

100% self-driving cars: **-40% to +100% energy**


---

Integrating autonomy

How can we gain understanding for integrating autonomy into complex systems?
In particular: traffic congestion.
Long-standing challenges

- Highly complex non-linear delayed dynamics
- Severe data limitations
- Human behavior modeling
- Large-scale, heterogeneity
- Computational cost
- Limited benchmarks

- Search possibilities
- Simulation
- Leverage mature models
- Seek insights in small settings
- Create some
The Flow Team

Thanks to:

Eugene Vinitsky
UCB MechE, PhD

Aboudy Kreidieh
UCB CEE, PhD

Kanaad Parvate, Nishant Kheterpal, Kathy Jang, Ananth Kuchibhotla, Leah Dickstein, Nathan Mandi
UCB EECS, undergraduate researchers

Alexandre Bayen
Principle Investigator
UCB EECS/CEE

Cathy Wu
Founder & Advisor
MIT CEE/IDSS

Thanks to:

Cathy Wu | UC Berkeley
Deep reinforcement learning (RL)

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

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

Global rewards
Average velocity
Energy consumption
Travel time
Safety, comfort

Policy parameters
(deep neural network)
Deep reinforcement learning (RL)

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

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

Agent
- RL Library
- Ray RLlib / rllab

Environment
- Custom dynamics
- Traffic simulator SUMO/Aimsun

Task designer
- Markov Decision Process
- Traffic network
- Traffic dynamics
- Vehicle types

Flow
flow-project.github.io

DQN (2015)
TRPO (2015)
AlphaGo (2016)

Flow: full networks
OpenStreetMaps

**Setting:** ~2000 vehicles

**Dynamics:**
- cascaded nonlinear systems
- bottlenecks
- multi-lane merges
- toll plaza dynamics

Flow: traffic LEGO blocks
Benchmarks for autonomy in transportation

Single-lane | Multi-lane | Intersection | On/off-ramp

Straight highway
Bottleneck
Grid network
Signalized intersection

Traffic jams

Partial differential equations (PDE)

**Setting:** 22 human drivers

**Instructions:** drive at 19 mph.

No traffic lights, stop signs, lane changes.

Sugiyama, et al. 2008

900 papers on PDEs for traffic 2008 2018
Traffic jams

Partial differential equations (PDE)

Setting: 22 human drivers

Instructions: drive at 19 mph.

No traffic lights, stop signs, lane changes.

Traffic jams still form.

Sugiyama, et al.

1955

900 papers on PDEs for traffic

2008

2018
Mixed autonomy traffic

1955

Partial differential equations (PDE)

Setting: 22 human drivers

Instructions: drive at 19 mph.

No traffic lights, stop signs, lane changes.

Traffic jams still form.

Sugiyama, et al.

2008

2018

900 papers on PDEs for traffic
No traffic jams!

1955: Setting 1 AV, 21 human drivers

Instructions: follow the vehicle in front, and close gaps. No tail-gaiting!

AV: Proportional-integral (PI) saturation controller

Traffic jams diminished.

1 AV: +14% average velocity.

Sugiyama, et al. 2008

Stern, et al. 2017

Prof. Daniel Work, Vanderbilt
Prof. Benedetto Piccoli, Rutgers
Prof. Benjamin Seibold, Temple
Prof. Jonathan Sprinkle, UoA
Prof. Raphael Stern, UMN
Maria Laura Delle Monache, Inria

Stern, et al. TR-C, 2018
**Setting:** 1 AV, 21 human

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

**Results**
- **1 AVs:** +49% average velocity
- **Stabilization at near-optimal velocity**

---

Single-lane: system equilibria

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

Average velocity vs traffic density

Optimal (unstable)
Traffic jams (stable)

Single-lane: state of the art

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

Average velocity vs traffic density

- Stop-and-go stable limit cycle
- Uniform flow unstable equilibrium
- PI with saturation controller
- Calibration density for PI controller
- MLP controller (ours)

Average velocity (m/s)
Vehicle density (veh/m)

State of the art
Traffic jams (stable)
Optimal (unstable)

260m

Single-lane: mixed autonomy

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

Results
• Near-optimal
• Generalized to out-of-distribution traffic densities

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.

General lane-changing model MOBIL for car-following models.


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
• 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
- **Insight**: A single AV can stabilize multiple lanes of traffic
- **Emergent traffic break**

---

Wu, et al. IEEE T-RO, 2018
Intersection: fully automated

Queuing theory
Reservation systems
Model predictive control

A multiagent approach to autonomous intersection management.

Polling-systems-based control of high-performance provably-safe autonomous intersections.
Miculescu, Karaman. CDC, 2014.

What if even one of these vehicles is not automated?
Intersection: no autonomy

Setting: 0 AV, 14 human

Dynamics: cascaded nonlinear systems with right-of-way dynamics model
**Intersection: mixed autonomy**

**Setting:** 1 AV, 13 human

**Experiment**
- Goal: maximize average velocity
- Observation: fully observed
- Action: acceleration

**Results**
- Emergent mixed-autonomy platoon
- **Insight:** A single AV can slow or stop ALL vehicles behind it
- **1 AV:** +60% average velocity
- **14 AVs:** +170% average velocity

---

San Francisco Bay Bridge

Multi-lane merge

Toll plaza: 18 lanes

Core problem: traffic bottleneck

Setting: No AVs
720 veh/hr

Phenomenon: capacity drop

Setting: 10% AVs
1020 veh/hr

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

40% improvement
Avoids capacity drop

Vinitsky, Parvate, Kreidieh, Wu, Bayen. IEEE ITSC, 2018
Onwards and upwards

Challenges:
• High-dimensional control
• Policy transfer to similar tasks
• Policy transfer to physical system
• Interpretability
• Hierarchy
• ...

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:

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

Door Opening (24-dim)

Previous SOTA: Greensmith, et al., 2004

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:

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

Door Opening (24-dim)

Policy transfer

• Ring roads $\rightarrow$ Straight roads

Initial performance boost

Successful direct transfer!

Closed $\rightarrow$ open networks

Policy transfer

- Ring roads $\rightarrow$ Straight roads

Setting: No AVs

Setting: 5% AVs

Ring-road pre-training

Straight-road fine-tuning

Flow: architecture & features

Control signals
- Longitudinal, lateral control
- Traffic light control, ramp meters

Large-scale reinforcement learning
- Hierarchical policy
- Multi-agent environments
- Distributed simulation and sampling

Scenarios and networks
- Parameterized python scenario creation
- A variety of open and closed networks
- OSM network import

Libraries
- Rich models via SUMO/Aimsun
- OpenAI gym interface
- Supports rllab and RLlib

Flow: open-source RL + microsim
Team: flow-dev@googlegroups.com
Docs: flow.readthedocs.io
Website: flow-project.github.io

A Flow Community

Controller design for AVs

Multi-agent RL

Understanding adversarial driving

Urban decision support systems

System verification

IEEE ITSC 2018 Workshop & Tutorial
Deep RL for Intelligent Transportation Systems
Available: flow-project.github.io
Exercise 02: Running RLlib Experiments

This tutorial walks you through the process of running traffic simulations in Flow with trainable RLlib-powered agents. Autonomous

Exercise 03: Running rllab Experiments

This tutorial walks you through the process of running traffic simulations in Flow with trainable rllab-powered agents. Autonomous

Exercise 04: Visualizing Experiment Results

This tutorial describes the process of visualizing and replaying the results of Flow experiments run using RL. The process of

Exercise 05: Creating Custom Scenarios

This tutorial walks you through the process of generating custom scenarios. Scenarios define the network geometry of the problem, as well as the constituents of the network, e.g. vehicles, traffic lights, etc... Various scenarios are available in Flow, depicting a diverse set of open and closed traffic networks such as ring roads, intersections/grids, straight highway merges, and more.

In this exercise, we will recreate the ring road network, seen in the figure below.
Integrating autonomy into urban systems

- **Deep reinforcement learning** can provide understanding for integration of autonomy into urban systems.
- **Small % of AVs** can greatly affect traffic dynamics, which in turn, affects all parts of the urban system.
- **Flow**: open source framework for benchmarking reinforcement learning and traffic microsimulation.