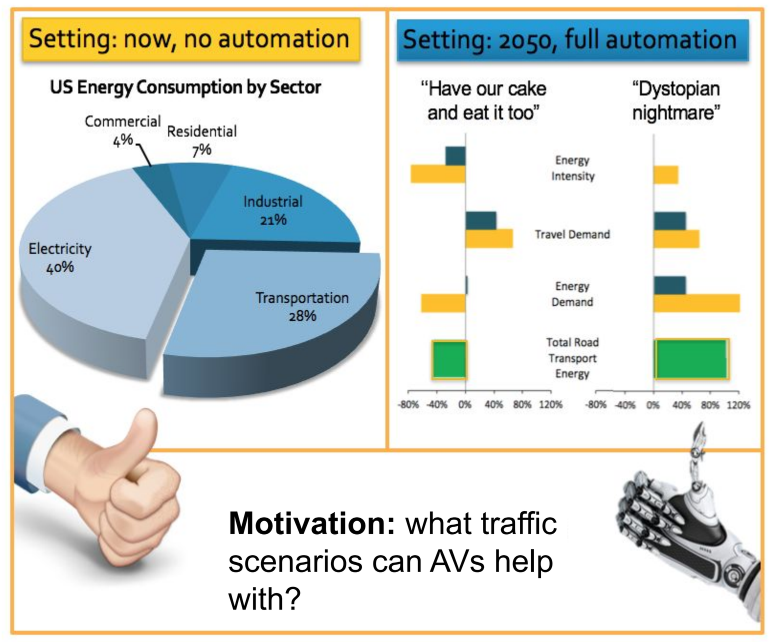
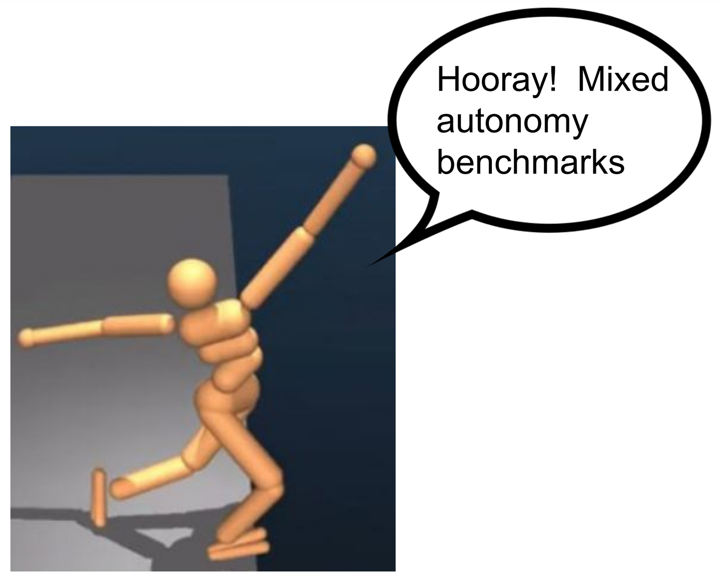
Benchmarks for Reinforcement Learning in Mixed Autonomy Traffic

Eugene Vinitsky[1*], **Aboudy Kreidieh**[2*], Luc Le Flem[3], Nishant Kheterpal[3], Kathy Jang[3], Cathy Wu[3], Fangyu Wu[3], Richard Liaw[3], Eric Liang[3], Alexandre Bayen[2,3,4] | UC Berkeley

Motivation

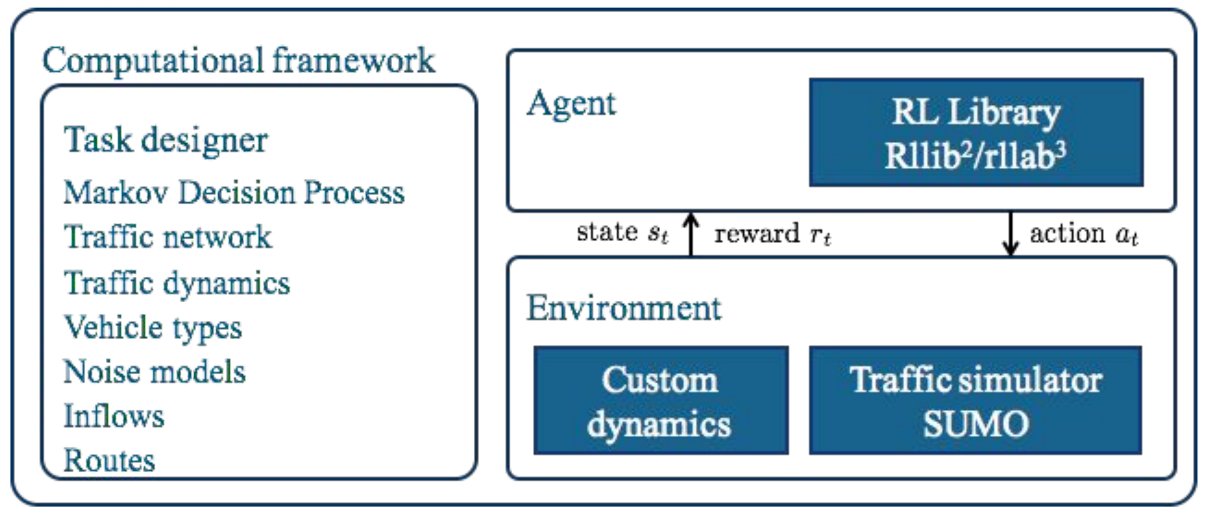


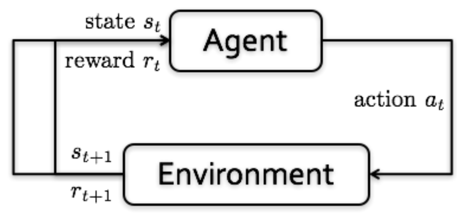
- 2019: Every car company rolls out **Automated Cruise Control**
- Steady appearance of **mixed autonomy-traffic**
- Opportunity to improve the roadways
- Mujoco/Atari benchmarks have hugely advanced RL research
- **No existing benchmarks in mixed-autonomy traffic!**
  - Time wasted rebuilding traffic scenarios
  - Impossible to compare control strategies



Problem Setup

Flow[1]:





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

Example:
max –(deviation from speed limit) + –λ(total fuel consumption)
s.t. no collisions

Control Structure

Autonomous vehicles:

$$a_t = \pi_\theta(s_t)$$

Humans: Intelligent Driver Model (IDM)

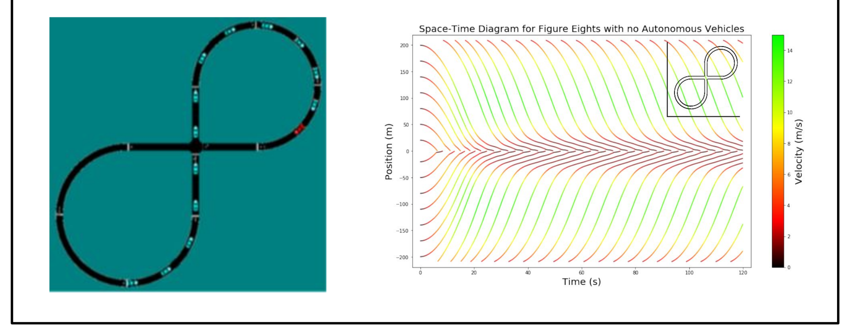
$$a_{\mathrm{IDM}} = \frac{dv_\alpha}{dt} = a\left[1 - \left(\frac{v_\alpha}{v_0}\right)^\delta - \left(\frac{s^*(v_\alpha, \Delta v_\alpha)}{s_\alpha}\right)^2\right]$$

$$s^*(v_\alpha, \Delta v_\alpha) = s_0 + \max\left(0, v_\alpha T + \frac{v_\alpha \Delta v_\alpha}{2\sqrt{ab}}\right)$$

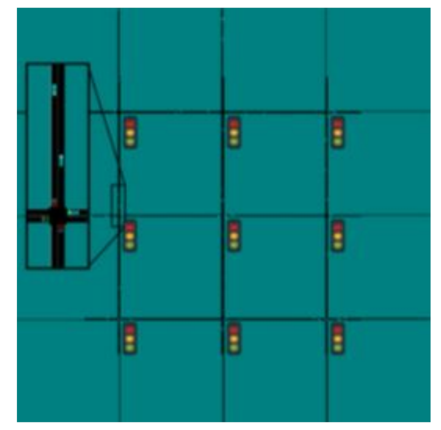
Benchmarks-v0 in *Flow*[1]

Image Key

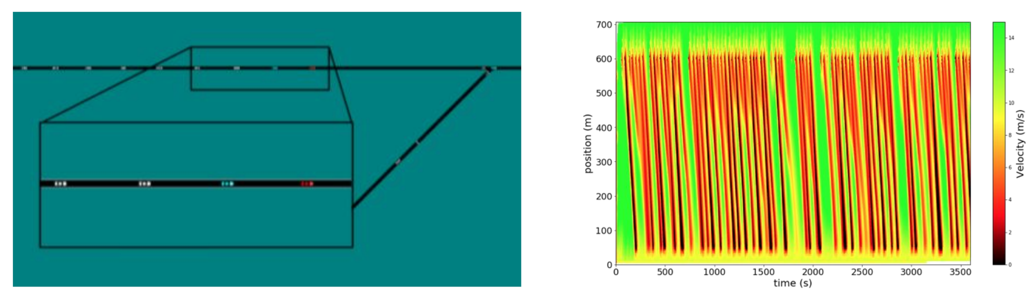
Figure Eight: action-space scaling



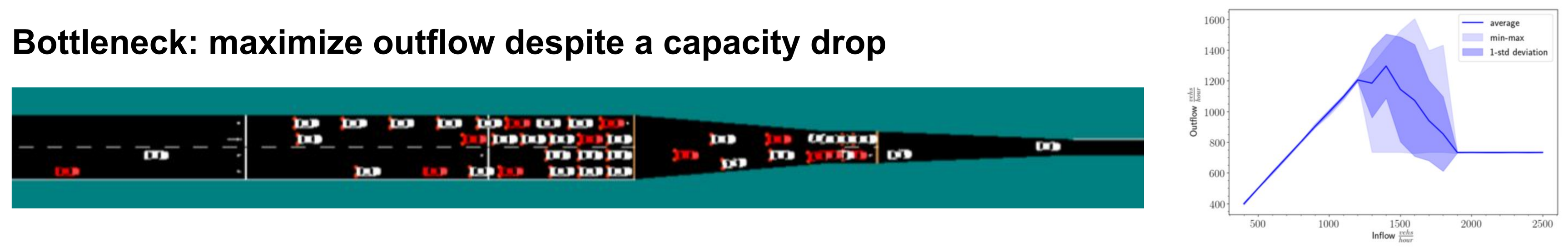
Grid: minimize delay



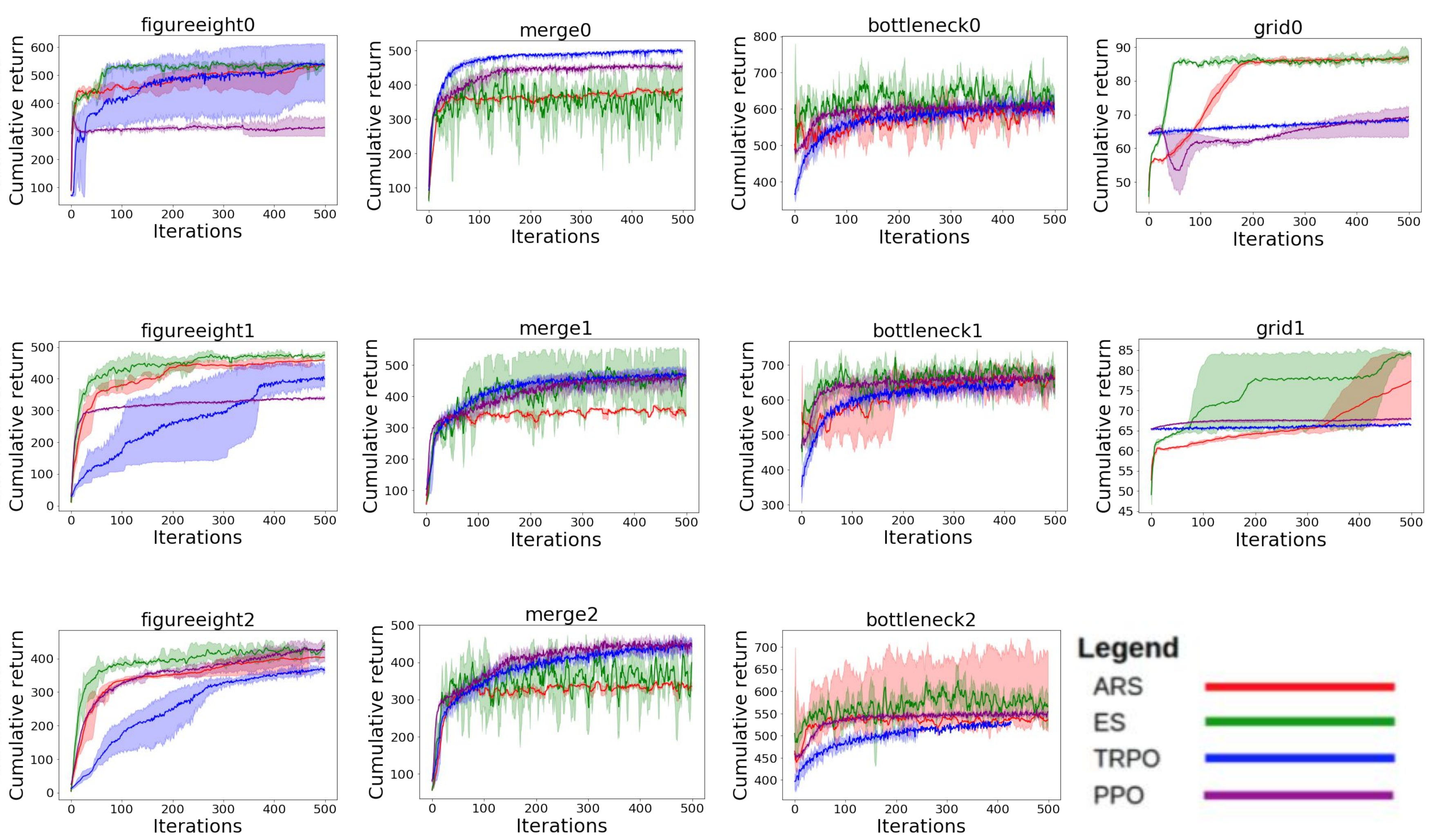
Merge: optimize shockwave dissipation



Bottleneck: maximize outflow despite a capacity drop



Results



Average and standard deviation over 40 rollout. Top scores in **bold**, human baseline in *italics*. Green/red indicates higher scores are better/worse.

Benchmark—	ARS	ES	TRPO	PPO	Human	Details	Metric
Figure Eight 0	7.3 ± .5	6.9 ± .1	**8.2 ± .1**	N/A	*4.2 ± .1*	1 AV, 13 Humans	Avg. Speed (m/s) ↑ (green)
Figure Eight 1	**6.4 ± .1**	N/A	5.6 ± .6	N/A	*4.2 ± .1*	7 AVs, 7 Humans	Avg. Speed (m/s) ↑ (green)
Figure Eight 2	5.7 ± .1	**6.0 ± .1**	5.0 ± .2	N/a	*4.2 ± .1*	14 AVs	Avg. Speed (m/s) ↑ (green)
Merge 0	11.3 ± .3	13.3 ± .5	**15.0 ± .1**	13.7 ± .4	*7.4 ± .6*	10% AVs	Avg. Speed (m/s) ↑ (green)
Merge 1	11.1 ± 3	**17.3 ± .4**	13.7 ± .2	14.6 ± .5	*7.4 ± .6*	25% AVs	Avg. Speed (m/s) ↑ (green)
Merge 2	11.5 ± .5	**17.3 ± .5**	14.1 ± .2	14.5 ± .3	*7.4 ± .6*	33% AVs	Avg. Speed (m/s) ↑ (green)
Grid 0	**270 ± 1**	271 ± 1	296 ± 3	296 ± 5	*280 ± 2*	3x3 Grid	Avg. Delay (s) ↑ (red)
Grid 1	275 ± 1	**274 ± 1**	296 ± 2	296 ± 2	*276 ± 2*	5x5 Grid	Avg. Delay (s) ↑ (red)
Bottleneck 0	1265 ± 263	**1360 ± 200**	1298 ± 268	1167 ± 264	*1023 ± 263*	10% AVs, 4x2x1 lanes, No lane changing	Outflow (vehs/hr) ↑ (green)
Bottleneck 1	1350 ± 162	**1378 ± 192**	1375 ± 61	1258 ± 200	*1135 ± 319*	10% AVs, 4x2x1 lanes, lane changing enabled	Outflow (veh/hr) ↑ (green)
Bottleneck 2	2284 ± 231	**2324 ± 264**	2131 ± 190	2143 ± 208	*1889 ± 252*	10% AVs, 8x4x2 lanes, no lane changing	Outflow (veh/hr) ↑ (green)

Conclusions/Future Work

- Using deep reinforcement learning, we can train autonomous vehicles to improve traffic
- **We open source four new benchmarks in mixed autonomy traffic**
  - Three benchmarks correspond to common traffic situations
  - AVs improve traffic metrics up to 100% in some benchmarks
  - Increasing the fraction of AVs does not necessarily improve outcomes
- **Future work: Benchmarks-v1**
  - **Generalization:** Can we find one AV controller for many scenarios?
  - **Decentralization:** how well can we do with multi-agent RL?
  - **Scaling:** Can AVs optimize traffic at the city scale?
  - **Fairness:** Can RL find the social optimum without unfair penalties?
  - **Human comfort:** Optimize w/ regard for passenger satisfaction
- For more info:
  - Website: https://flow-project.github.io
  - Github repo: https://github.com/flow-project/flow
  - RL Library: https://github.com/ray-project/ray
  - Lab twitter: https://twitter.com/BerkeleyMsl

References:

1. Wu, Cathy, et al. "Flow: Architecture and benchmarking for reinforcement learning in traffic control." *arXiv preprint arXiv:1710.05465* (2017).
2. Liang, Eric, et al. "Ray RLLib: A Composable and Scalable Reinforcement Learning Library." *arXiv preprint arXiv:1712.09381* (2017).
3. Duan, Yan, et al. "Benchmarking deep reinforcement learning for continuous control." *International Conference on Machine Learning*. 2016.

*indicates equal authorship, evinitsky@berkeley.edu, {aboudy, bayen, nkheterpal}@berkeley.edu; 1. Mechanical Engineering; 2. Civil and Environmental Engineering; 3. Electrical Engineering and Computer Science; 4. Institute of Transportation Studies