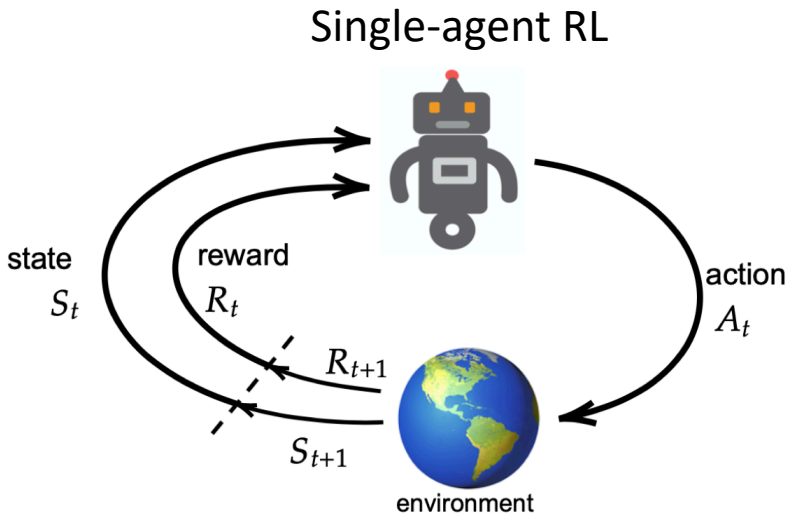
SMARTS: Scalable Multi-Agent Reinforcement Learning Training School for Autonomous Driving

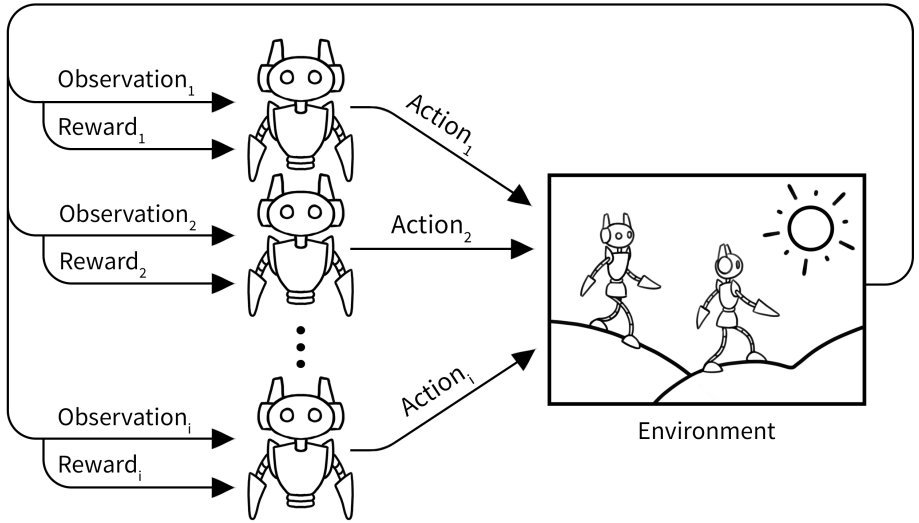
by Ming Zhou et al.

CoRL 2020 (Best System Paper Award)

Presented by Ismail Geles

Multi-agent Reinforcement Learning (MARL)





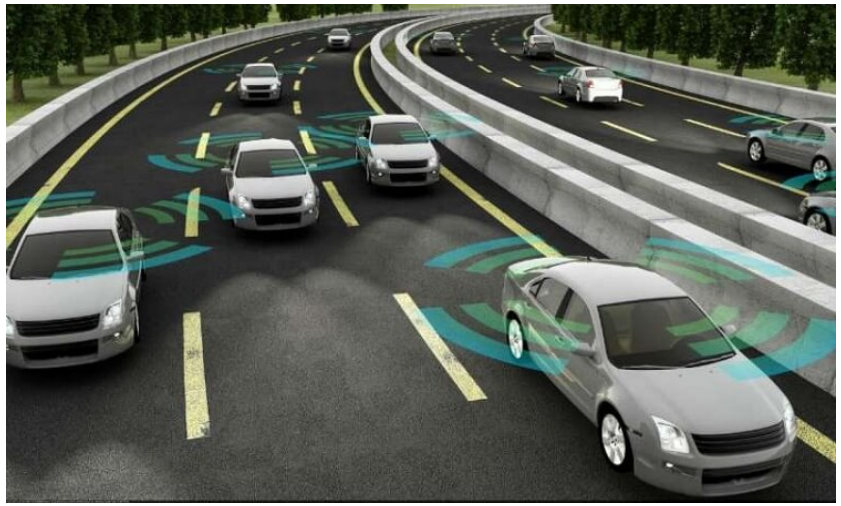
$s \in \mathcal{S} \ldots$ state space (set of states)

$a \in \mathcal{A} \ldots$ action space (set of actions)

$r \in \mathcal{R} \ldots$ set of rewards

$p(s_o) \quad \ldots$ distribution of initial states

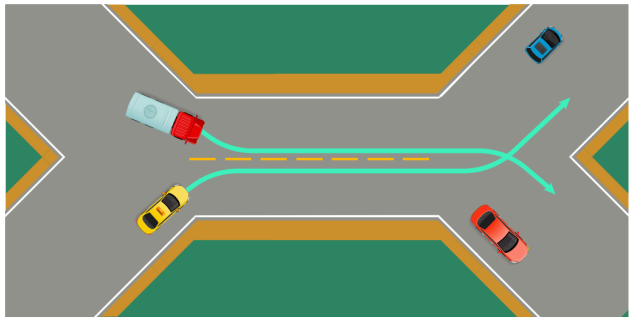
$$\mathbb{E}\left[\sum_{t\geq 0}\gamma^t R(s_t, a_t, s_{t+1}) \,\middle|\, a_t \sim \pi(\cdot|s_t), s_0\right]$$



M-levels, simulators and why SMARTS?

Level	Description
M0	Rule-based planning and control without learning
M1	Single-agent learning without coordination with other learning agents
M2	Decentralized multi-agent learning with opponent modeling
M3	Coordinated multi-agent learning and independent execution
M4	Local (Nash) equilibrium oriented multi-agent learning
M5	Social welfare oriented multi-agent learning

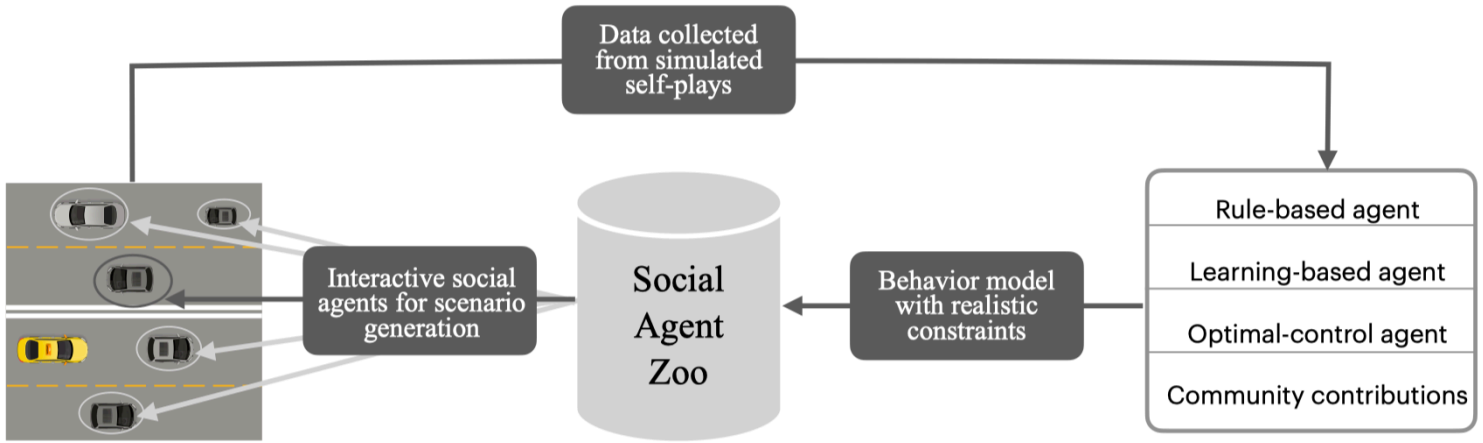
double merge scenario



Other AD RL simulators:

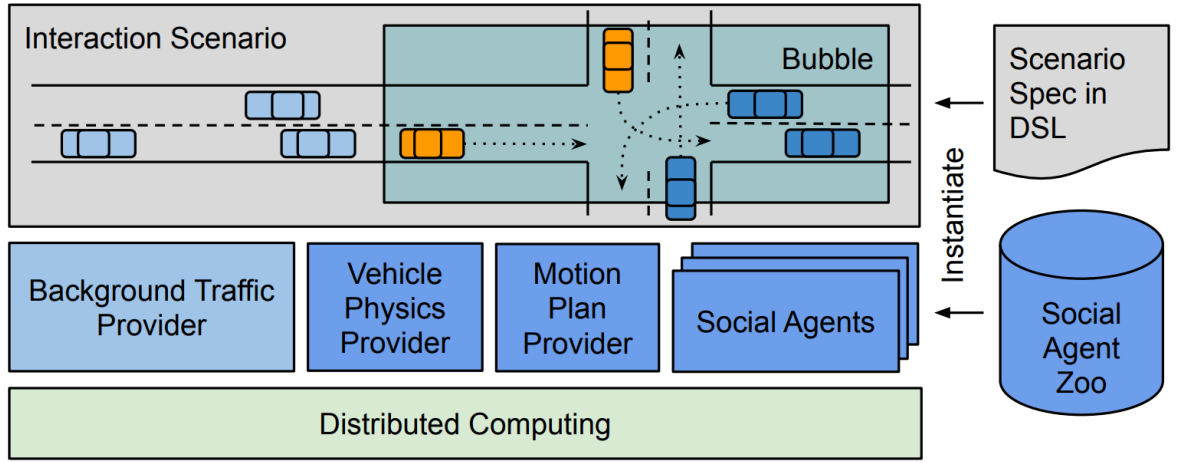
AIM4, BARK, CarRacing, CityFlow, Duckietown, Flow, Gym-TORCS, highway-env, MACAD-Gym, MADRaS,…

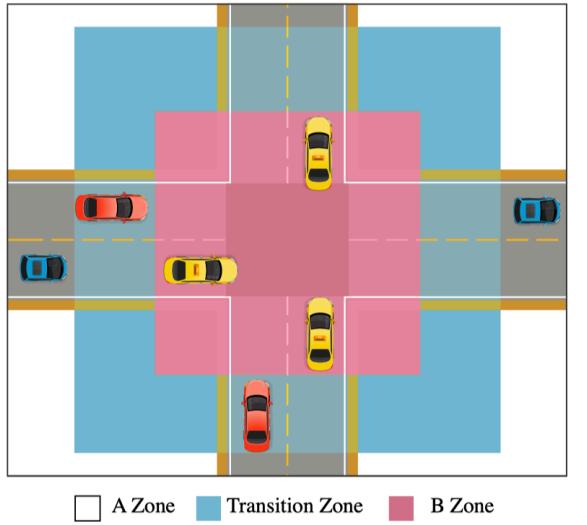
What should be the benchmark?
Diverse and realistic simulations?



Design and Architecture of SMARTS

Area	Features
Realistic Interaction	Realistic physics Heterogeneous ego and social agents Handwritten social agents Social agents trained with real-world data Social agents trained with RL Social agent zoo for crowd sourcing
Platform Integration	Multi-agent distribution Multi-simulation distribution RLlib integration for RL training SUMO integration for traffic simulation Built-in scenario composition
Research Friendliness	Gym APIs Headless mode Web-based visualization with recording Comprehensive observation & action options Multi-agent RL algorithm libraries Realistic benchmark suits

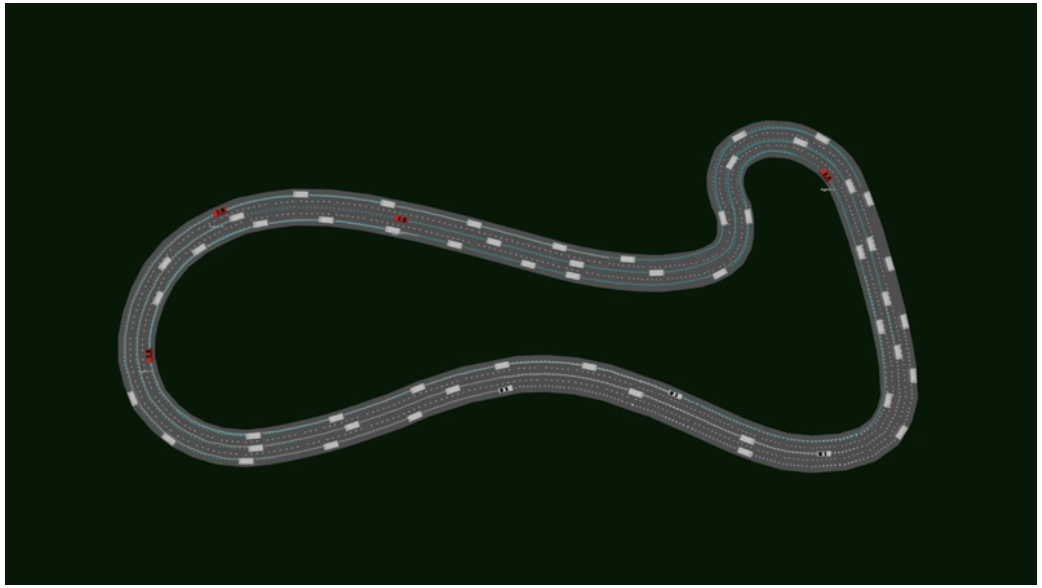


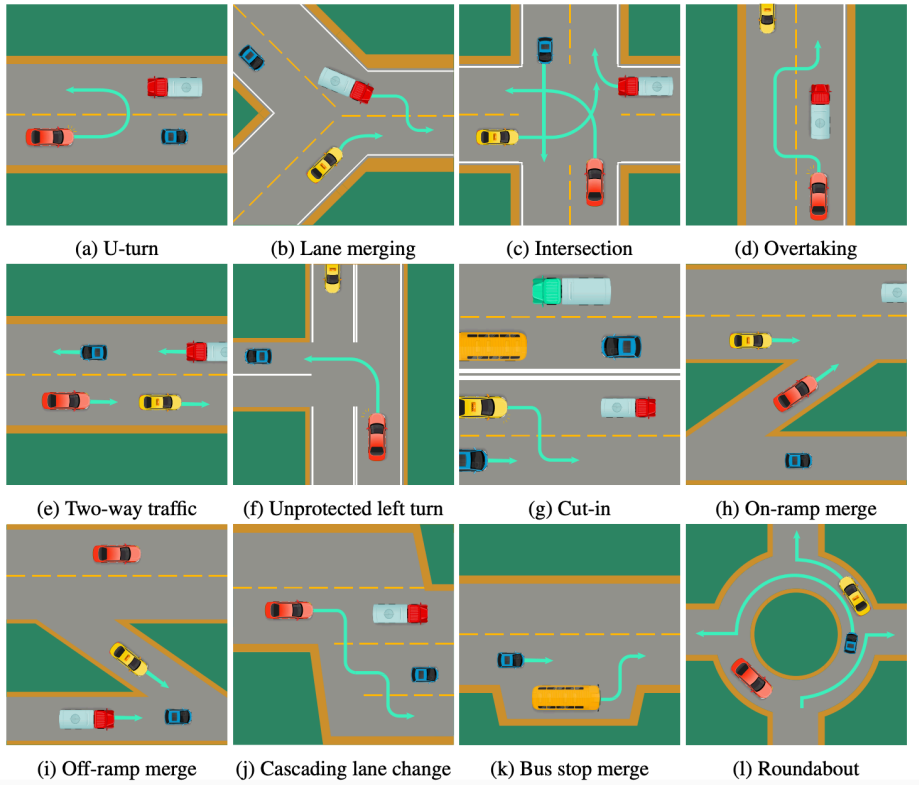


Background traffic transitions to social agent inside of "Bubble"

Setup and Experiments

- Docker Image provided
- Scenarios, Agents and Metrics
- Web-visualisation
- RL benchmarks

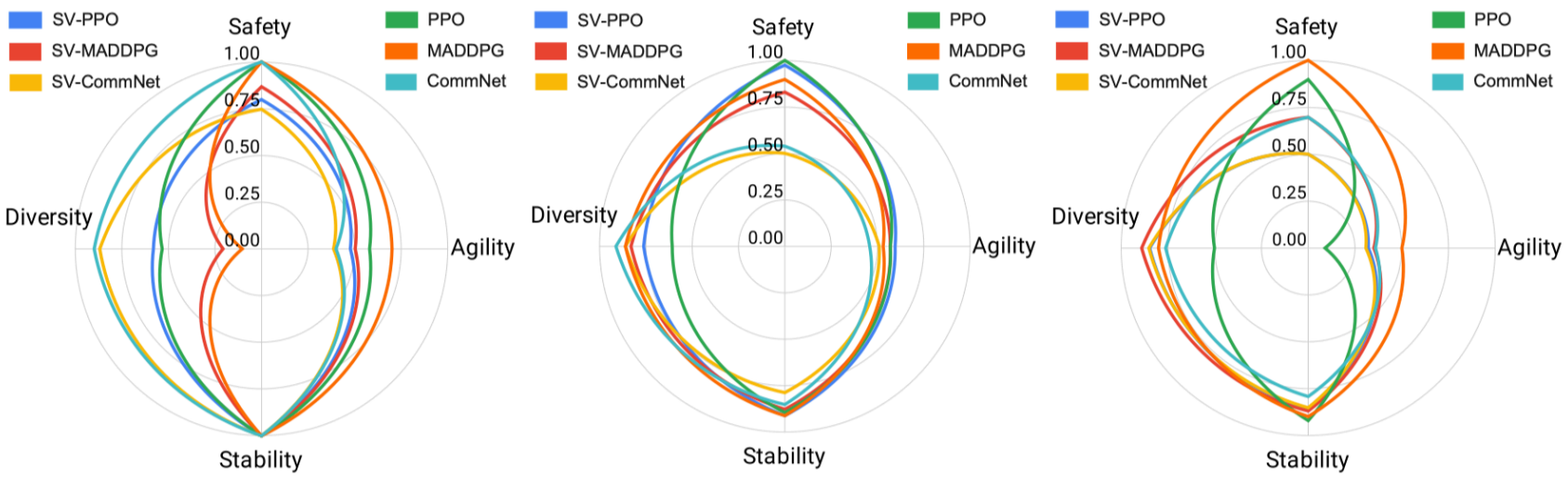




Metric	Type	Description
Collision Rate	Performance	# of collisions over # of episodes.
Completion Rate	Performance	# of missions completed over # of episodes.
Generalization	Performance	Robustness of algorithms to scenario variation.
Safety	Behavior	Integrated metrics, e.g. non-collision rate.
Agility	Behavior	Integrated metrics, e.g. speed.
Stability	Behavior	Integrated metrics for driving smoothness.
Control Diversity	Behavior	Preference for longitudinal or lateral control.
Cut-in Ratio	Behavior	Probability of cut-in in traffic flow.
Stochasticity	Behavior	Stochasticity of decision making.
Collaborative	Game theory	Compatible interests, e.g. ratio of giving way.
Competitive	Game theory	Conflicting interests, e.g. ratio of overtaking.

RL Algorithms and Benchmark Results

- Independent learning: DQN, PPO
- Centralised training: MAAC, MF-AC, MADDPG, Networked Fitted-Q
- Fully centralised training: CommNet
- Radar plots over four behaviour metrics comparing different RL agents

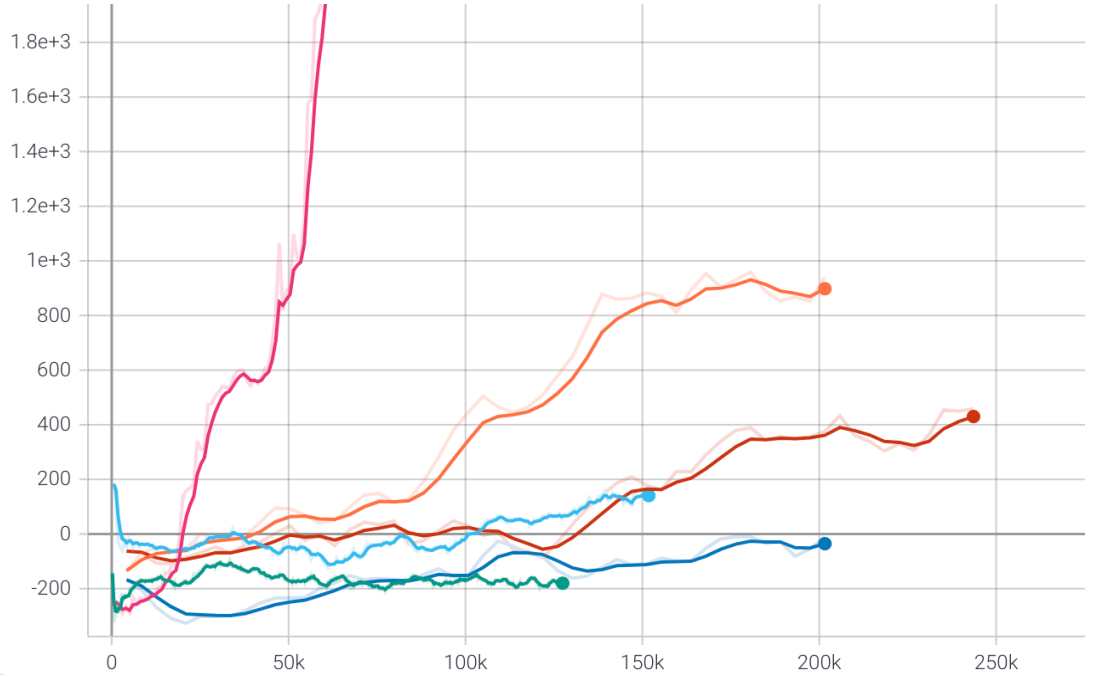


(a) Two-Way (b) Double Merge (c) Intersection

Own Experiments

- due to the number of bugs and errors in the repository, only “double merge” was tested with a few algorithms and parameters
- release version 0.4 --> might improve in future releases

Mean reward of 4 agents over time-steps



Pink: DQN, default

Light blue: MF-AC, default

Green: A2C, default

Orange: PPO, lr=0.0001, SGD_batch_size=32

Red: PPO, lr=0.0003 SGD_batch_size=64

Blue: PPO, lr=0.001, SGD_batch_size=64

Performance of various algorithms on ”double merge” scenario

Quiz

- How does an agent learn in reinforcement learning?
- How can reinforcement learning influence autonomous driving?
- What are some features SMARTS offers?
- What is a “Bubble” in the SMARTS library?

Appendix: Other available AD MA/RL Libraries

MA/RL	AIM4 [15]	Multi-agent framework for managing autonomous vehicles at intersections.	Focused on intersection only and assuming all vehicles are autonomous.
	BARK [25]	Multi-agent envs with extensible social agent models.	Similar to SMARTS in emphasizing importance of social agents; no explicit support for multi-agent research.
	CarRacing [63]	An RL env in OpenAI Gym.	Car racing for testing RL.
	CityFlow [17]	Streamlined simulator for traffic optimization with RL.	Could provide background traffic in low-fidelity setting.
	Duckietown [64]	AD simulator for education.	For the Duckietown project.
	Flow [18]	Microscopic traffic simulation for deep RL.	Wraps SUMO and Aimsun Next.
	Gym-TORCS [65]	Gym-style env for TORCS	Shows need for standard envs.
	highway-env [24]	Hand-coded interaction envs.	Shows need for interaction envs.
	MACAD-Gym [66]	Multi-agent connected car simulation using CARLA.	Links V2V communication with multi-agent coordination.
	MADRaS [67]	TORCS multi-agent wrapper.	Limited to racing.