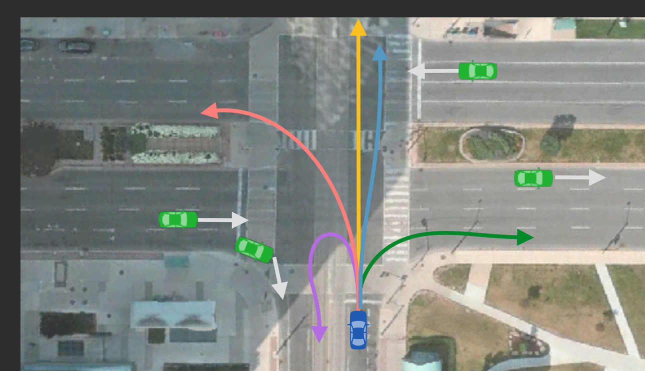
Multiple Futures Prediction

Yichuan Charlie Tang, Ruslan Salakhutdinov

Challenge

How to make multi-agent interactive multimodal future predicitions?



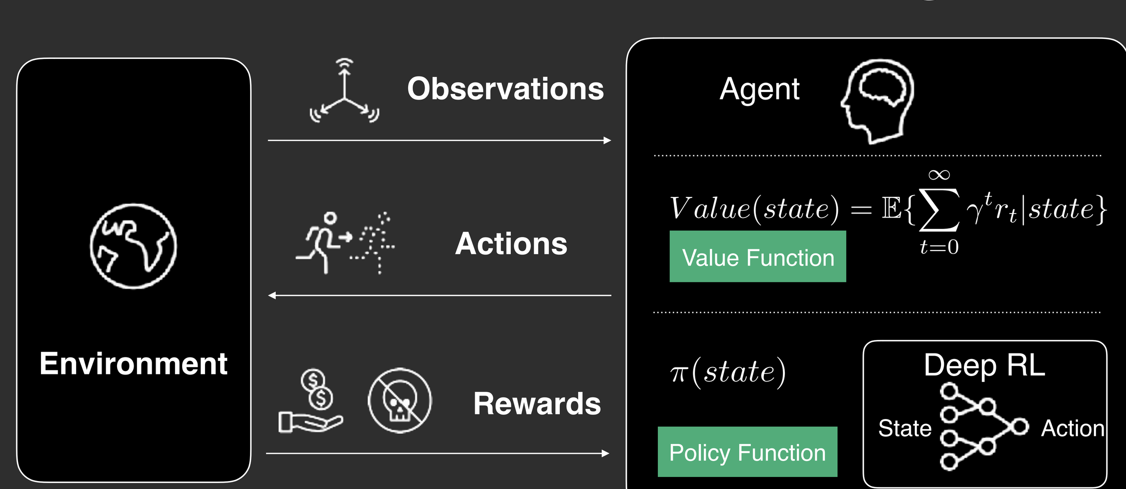
Deep RL + Self-play

- Learn policies automatically from simulation
- As training progresses, iteratively add more capable agents to the training environment
- Increases diversity and complexity of the simulation
- Leads to the learning of more robust policies

Our Contributions

- 2D sim env with topology and motion model
- Rule-based (IDM) agents with adaptive cruise control and simple lane change capabilities
- RL learning from rasterized + low dimensional inputs
- Qualitatively: learns human-like policies
- Quantitatively: improves success rate from 63% to over 98% via self-play training

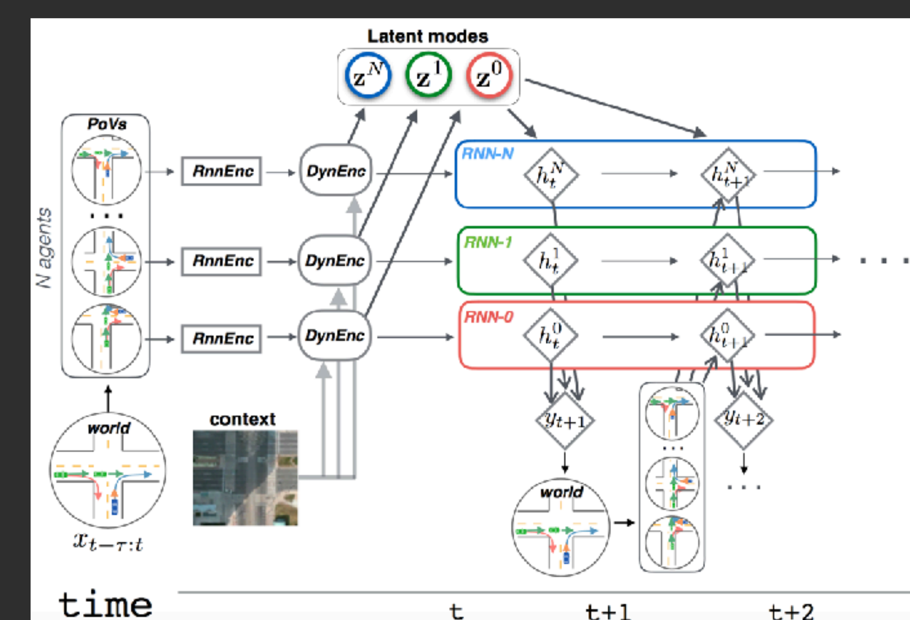
Reinforcement Learning



RL training

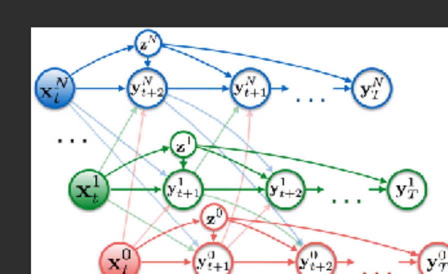
- Goal: **(A,B, or C)** -> **(D,E, or F)**
- Model-free training with policy gradient methods
- Input: rasterization + relative features
- Rewards: Success: +100. Collision: -500
- Initially trained with rule-based agents
- Reward shaping is critical

MFP Predictive Model

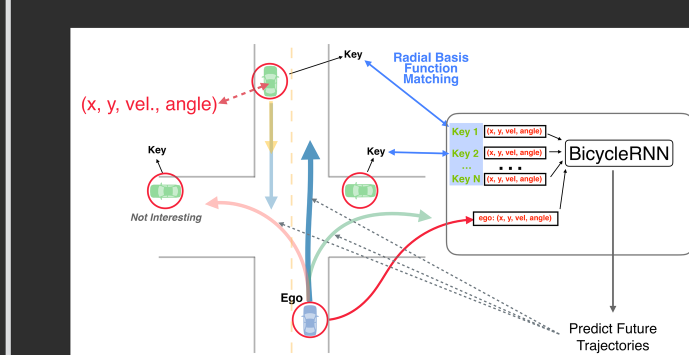


Parameters	Values
Scale	340m
Target Velocity	[10,20] m/s
# Other Agents	[0, 10]
Spawn %	1%
Reward: Success	+100,0

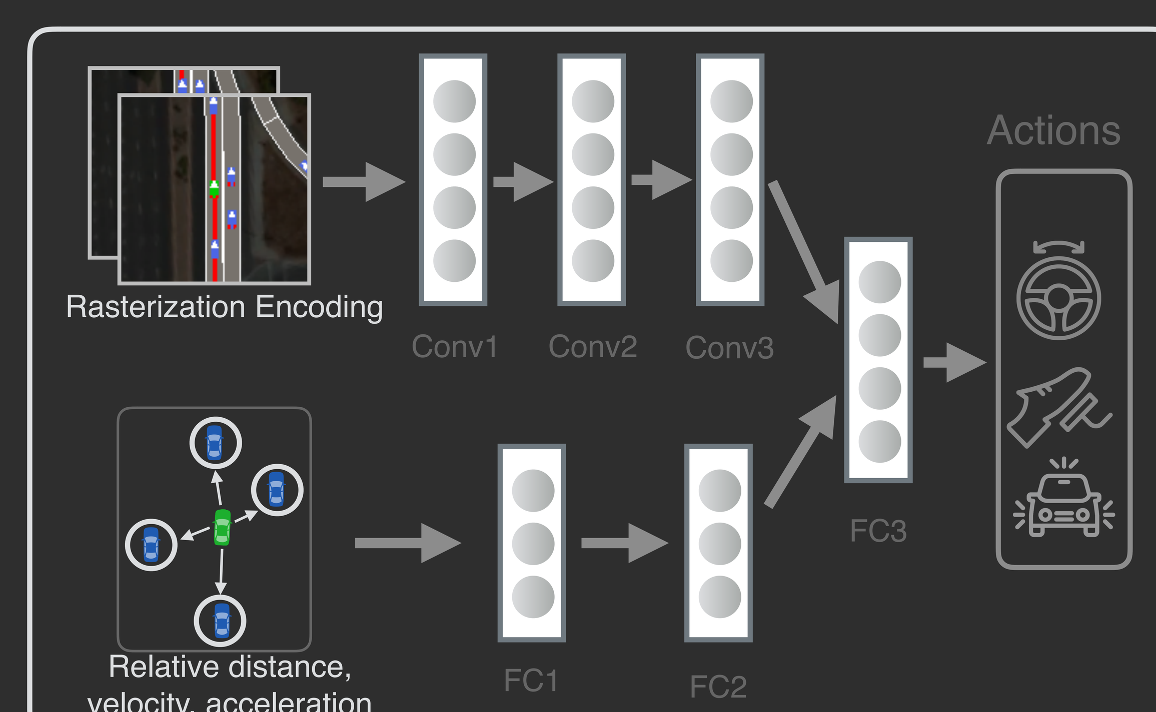
Graphical Model



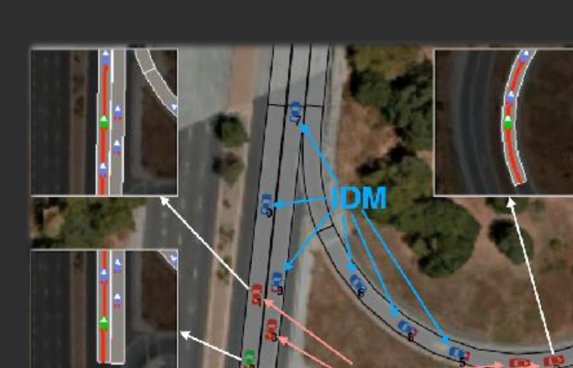
Dynamic Attention Encoding



Network Architecture



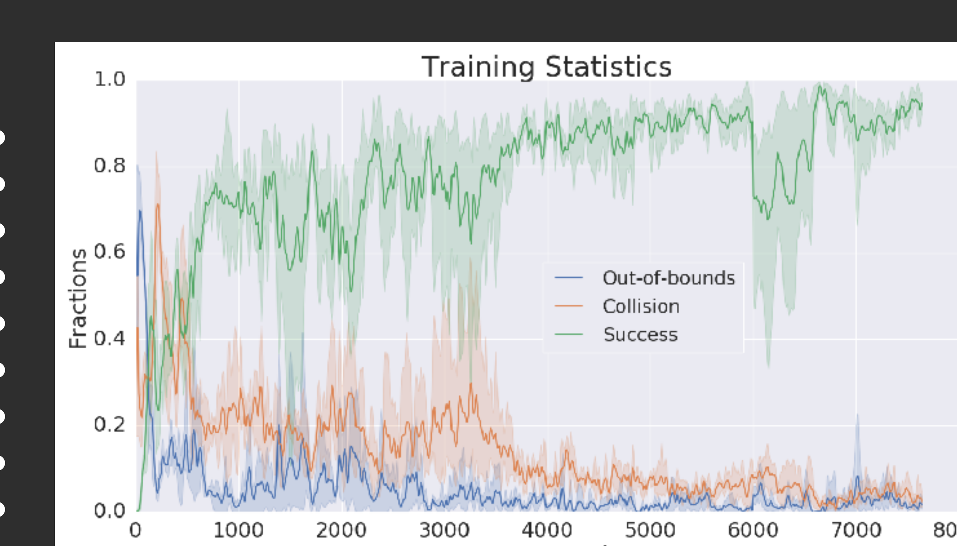
Self-play Training



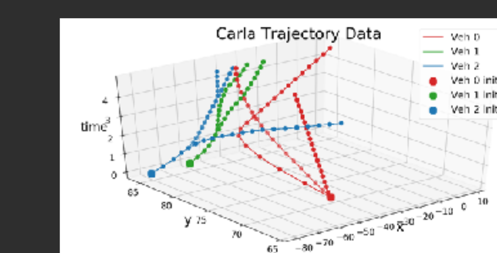
- RL overfits to environments
- While training, add RL-trained agents back to simulation: ***sparring partners***
- Self-play increases complexity
- Learns interesting merge behaviors

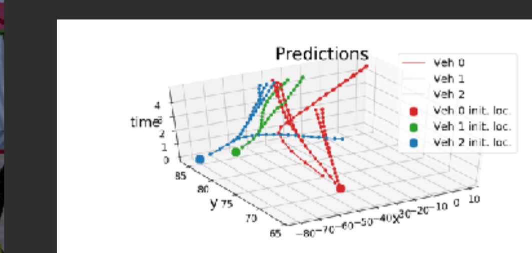
Training Populations

	Population Agent Types			
Populations	IDM	RL	SP1	SP2
Popul. 1	100%	0%	0%	0%
Popul. 2	50%	50%	0%	0%
Popul. 3	30%	30%	40%	0%
Popul. 4	10%	20%	30%	40%

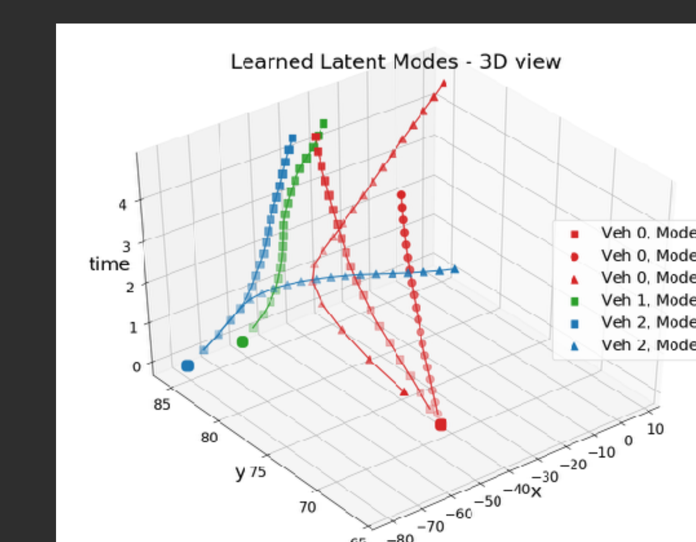


Experimental Results - Qualitative

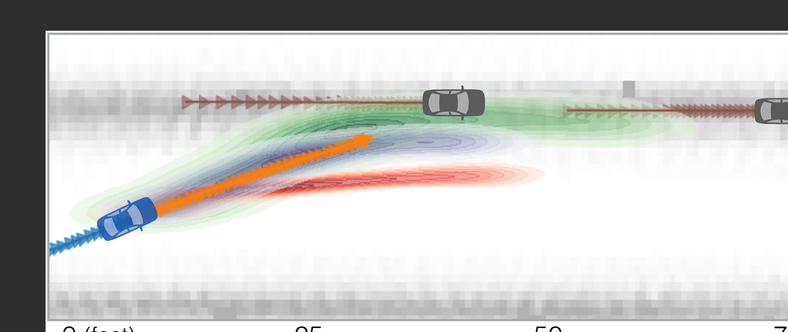


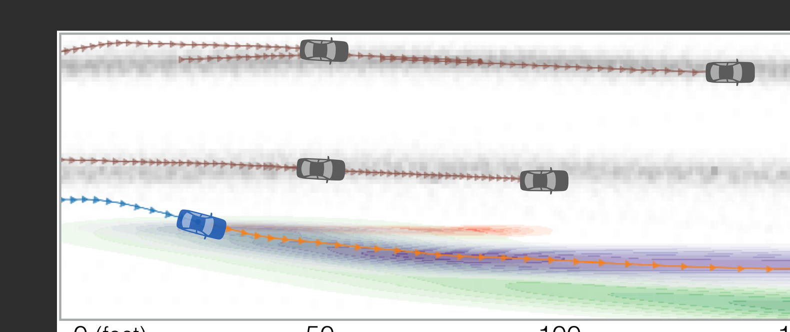


	Population Agent Types			
Populations	IDM	RL	SP1	SP2
Popul. 1	100%	0%	0%	0%
Popul. 2	50%	50%	0%	0%
Popul. 3	30%	30%	40%	0%
Popul. 4	10%	20%	30%	40%



NGSIM Results





Experimental Results - Quantitative

		Training	Testing - Success rate % (Collision rate %)			
Stage	**Agents**	**Training with Populations**	**Test against Population 1**	**Population 2**	**Population 3**	**Population 4**
0	IDM	N/A	62.8% (0.8%)			
1	RL	IDM (Population 1)	94.8% (4.0%)	77.2% (12.4%)		
2	Self-play 1	IDM+RL+SP1 (Population 3)	96.0% (3.6%)	91.2% (4.4%)	95.2% (3.2%)	
3	Self-play 2	IDM+RL+SP1+SP2 (Population4)	93.2% (6.0%)	94.8% (4.8%)	95.2% (4.4%)	**98.2%** (1.4%)