Optimizing driving entity switching of semi-automated vehicles under automation degradation

Author: Csanád Bakos, c.bakos@student.tudelft.nl | Supervisor: Yang Li | Responsible Professor: Matthijs Spaan | April-July 2021 | CSE3000 Research Project

1. Motivation

- Autonomous driving systems (ADS) can bring safety & efficiency
- Until full automation, control mediation between person & ADS needed
- Vehicle leaves operational design domain (ODD) or ADS malfunctions
- Can't crash cars just to find optimal control mediation policies

Policy (WandB id)	Accident ratio	HR level 2 ratio	A4R level 2 ratio	Unsafe action ratio	Unnecessary actions ratio	Complete episode used ratio
Baseline (333-eval)	**0.0128**	**0**	**0.0082**	**0.04632**	**0**	**0.1762**
SB3 DQN (167-train)	0.1788	0.03868	0.01535	0.1737	0.0851	0.064
SB1 DQN (184-train)	0.026	0.005363	0.02005	0.09648	0.04752	0.0008

Policy (WandB id)	Needed & approved EL0 action ratio	Needed & approved EL4 action ratio	Correct ES ratio
Baseline (333-eval)	**1**	**1**	**1**
SB3 DQN (167-train)	0.1982	0	0.8067
SB1 DQN (184-train)	0.5798	0.1222	0.5688

Figure 2: evaluation results

2. Question

How can an optimal driving entity switching strategy be found under automation degradation?

3. Method

- Markov Decision Process (MDP) formulation
- Simulate state space, autonomous system failures & leaving/entering ODD
- Baseline decision tree policy
- Deep reinforcement learning approach with Deep Q-Network (DQN)
- Safety and comfort evaluation

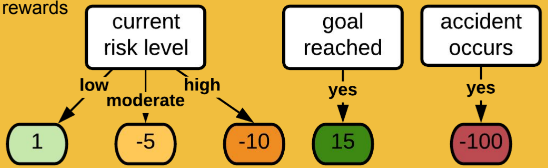

Figure 3: basic rewards

Actions:
DN: do nothing
ES: emergency stop
SL0/4: suggest shift to L0/4
EL0/4: enforce shift to L0/4

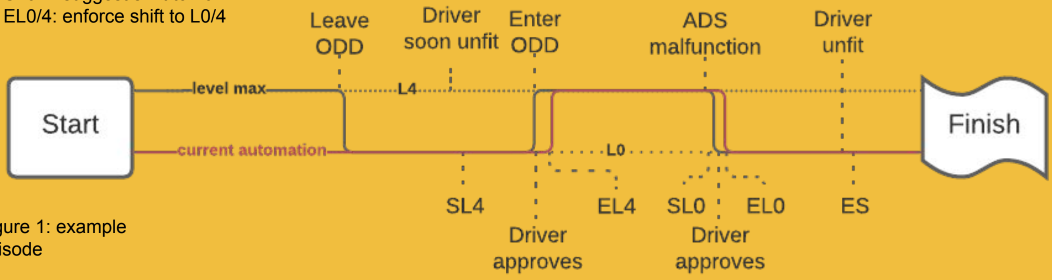

Figure 1: example episode

4. Results & Conclusions

- Baseline outperforms DQN policy, explainability is easy
- No simple reward suitable because of multi-objective optimization
- Learning complex, multi-step decision chains is problematic
- Dilemma about final say in decisions
- Baseline limitations: noisy input, need to map out all possible scenarios