Reinforcement Learning for Switching of Semi-automated Vehicles under Uncomfortable Driving Situations

CSE 3000 – Author: Ceren Uğurlu (C.Ugurlu@student.tudelft.nl) - Supervisor: Yang Li - Professor: Matthijs Spaan

Background

Automated vehicle technology developing with huge safety potential

MEDIATOR, a mediating system that enables a safe and real-time switching

Problem

Use Case: Mediator initiates shift of control when uncomfortable driving situations are detected

Goal: modelling the problem and evaluating RL algorithms compared to baseline policy

Uncomfortable Events

Manual driving

Traffic Jam | Poor Visibility | Long Trips

Automated driving

Time Pressure

Method

- Build an **MDP model** to provide a mathematical formulation of the decision problem
- Create **route simulation**
- Apply **RL algorithms** on created MDP model

MDP Formulation

- State Space
- Action Space

Do Nothing | Suggest Shift | Shift

- Transition Probability
- Reward Function

RL Algorithms

OpenAI Baselines algorithms:
- Deep Q-Learning (DQN)
- Advantage Actor Critic (A2C)
- Trust Region Policy Optimization (TRPO)

Evaluation

Reward
- Training performance: mean reward versus timesteps

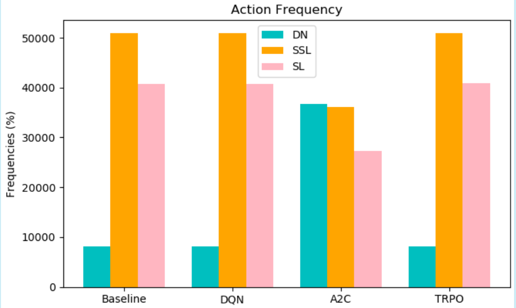
Driving Safety:
- Number of unsafe actions

Driver Comfort:
- Time to fix discomfort scenario

Decision Efficiency:
- Number of fixed scenarios
- Time to correct action
- Action distribution

Results

Algorithms	Driver Safety	Driver Comfort	Decision Efficiency	Mean Reward
Baseline	32.63	2.11	67.99	5.11
DQN	32.66	1.11	91.65	5.10
A2C	21.81	2.01	81.65	2.87
TRPO	32.66	2.11	69.32	5.10



Conclusion

- **DQN** outperforms other RL algorithms and the baseline
- **A2C** performs worst
- A learned reinforcement learning policy can be used to solve complex decision-making scenarios