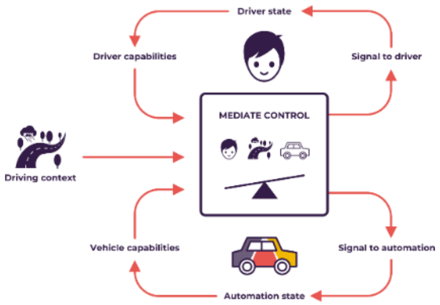
REINFORCEMENT LEARNING FOR SWITCHING CONTROL OF SEMI-AUTOMATED VEHICLES WITH DRIVER FATIGUE

Irem Ugurlu (Ugurlu@student.tudelft.nl) - Supervisors: Yang Li & Prof. Matthijs Spaan - CSE3000 Research Project

BACKGROUND

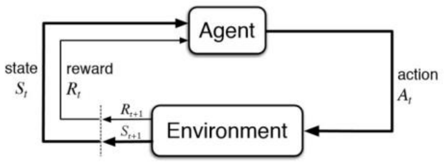
Automated driving technology is growing rapidly

MEDIATOR will develop a safe mediating system for drivers in automated vehicles



Markov Decision Process (MDP): discrete-time stochastic control process to model decision-making problems

Reinforcement Learning (RL): training of machine learning models to make a sequence of decisions



GOAL

Determine the optimal actions to guarantee driving safety and driver comfort in decision logic for specific use case of fatigue

METHOD

- Build an MDP model for the problem
- Use existing RL algorithms to evaluate the performance
- Compare with the baseline policy

MDP FORMULATION

1 - FATIGUE REPRESENTATION

KSS is used to quantify the sleepiness level
- Alert (KSS = 1~4)
- Neither alert nor sleepy (KSS = 5)
- Signs of sleepiness (KSS = 6~7)
- Sleepy (KSS = 8~9)

2 - ACTIONS

Do Nothing | Correct Fatigue | Suggest Shift Automation | Shift to Automation | Emergency Stop

EVALUATION

- OpenAI Baselines algorithms: DQN, A2C, TRPO
- 3 different metrics to evaluate
  - Driving safety
  - Driving comfort
  - Decision efficiency

RESULTS

Algorithms	Unsafe Actions (%)	DHFL (s)	Fixed Scenarios (%)
baseline	3.80	1.34	36.53
DQN	3.22	**1.29**	**87.53**
A2C	*23.06*	*2.58*	*16.97*
TRPO	**3.21**	1.30	67.16

Table 1: Driving Safety Metrics

Algorithms	Duration of Being in Fatigue (%)
baseline	2.74
DQN	1.58
A2C	*5.89*
TRPO	**1.49**

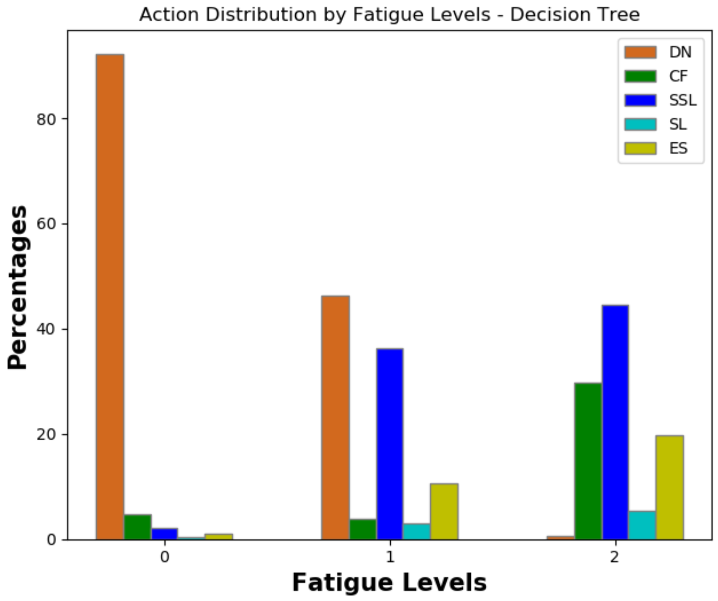
Table 2: Driving Comfort Metrics

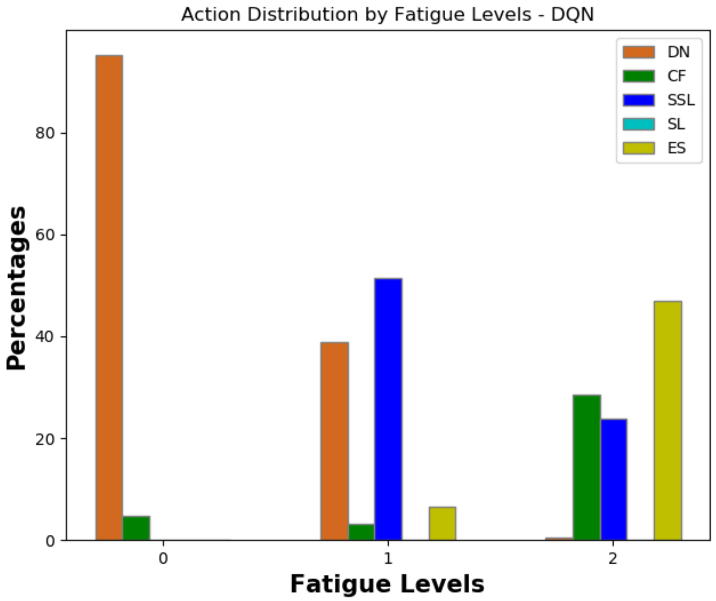
Algorithms	Time to Solve Critical Scenario (s)
baseline	*2.80*
DQN	2.67
A2C	**1.63**
TRPO	2.64

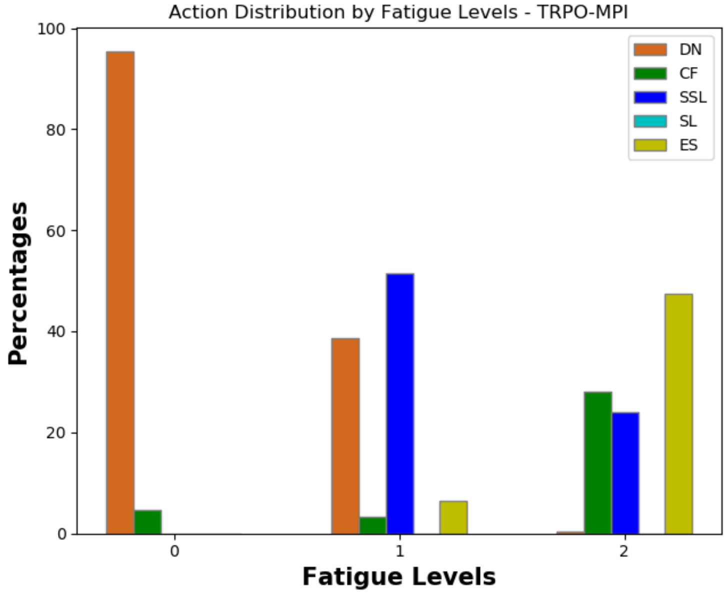
Table 3: Decision Efficiency Metrics

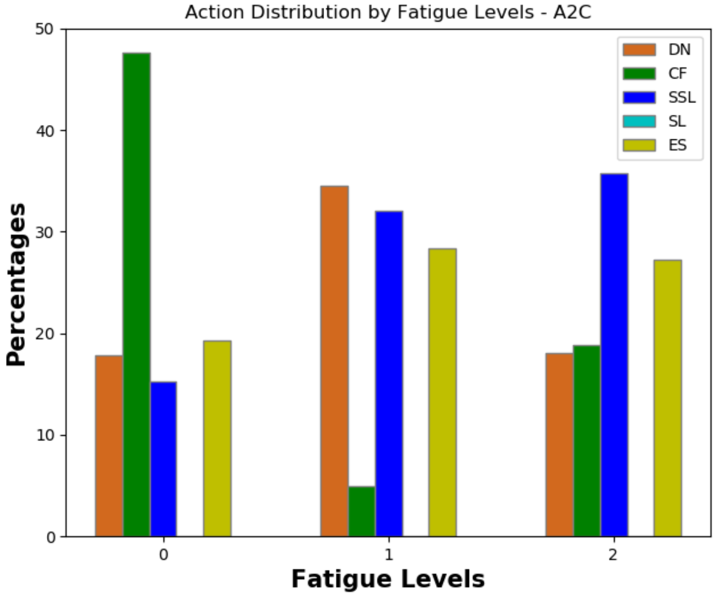
CONCLUSION

- DQN and TRPO outperforms other algorithms
- A2C is the worst algorithm
- More successful algorithms than baseline policy
- Further improvements by trying more algorithms


Action Distribution by Fatigue Levels - Decision Tree
DN
CF
SSL
SL
ES
Percentages
0
20
40
60
80
0
1
2
Fatigue Levels


Action Distribution by Fatigue Levels - DQN
DN
CF
SSL
SL
ES
Percentages
0
20
40
60
80
0
1
2
Fatigue Levels


Action Distribution by Fatigue Levels - TRPO-MPI
DN
CF
SSL
SL
ES
Percentages
0
20
40
60
80
100
0
1
2
Fatigue Levels


Action Distribution by Fatigue Levels - A2C
DN
CF
SSL
SL
ES
Percentages
0
10
20
30
40
50
0
1
2
Fatigue Levels

REFERENCES

- https://mediatorproject.eu/about/about-mediator
- https://deepsense.ai/what-is-reinforcement-learning-the-complete-guide/
- https://www.kdnuggets.com/2018/03/5-things-reinforcement-learning.html
- https://www.flaticon.com/