



Reinforcement Learning for Obstacle Navigation in Low Control Situations

David Brooks, Michael Villordon, Vincent Viray



Introduction/Objective

- An average of nearly 1 million accidents occur each year due to wet pavement [1]
- As of now, there is little research in this area; especially automated control of the car in a situation where the driver may be panicking
- As such, we introduce a proof-of-concept system for an autonomous car system in low control situations
- The primary goal of this system is to keep the occupants of the car safe, minimizing harm to the passengers if a crash is unavoidable. The secondary goal is reaching the end of the road safely, so long as a crash was avoided.
- Our project is designed around this reality, from rewards and model to experimental design

Methodology

- Our goal was for the agent to learn to avoid obstacles/walls and to stop
- As such, the rewards for our reinforcement learning model were as follows:
- High impulse when in contact with a wall or obstacle:
 - High impulse indicates a drastic crash.
 - The penalty is proportional to the impact of the car
 - If a crash is drastic enough, the episode ends
 - Penalty is offset by reduction in speed from initial
 - Penalty for "love tapping"
- Successfully stopping:
 - Successfully stopping for 1 second counts as a success and a big reward is given
- Reaching the end of the simulation
 - If the car successfully reaches the goal at the end of the road, a big reward is given.

Methodology - Continued

Agent Actions

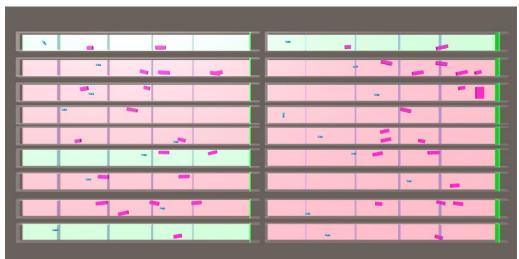
- (floats) turn left, turn right, no turn
- o (floats) move forward, move backwards, don't move
- o (boolean) break or don't break

• RL State

- LiDAR sensors
- Linear Velocity of agent
- Angular Velocity of Agent

Experimental Design

- For the car:
 - Car is spawned randomly on left or right
 - Car initially either has a slow speed or fast speed (although 90% of the time it's fast)
 - Car can have some rotation to it and angular velocity
 - Tire friction set to .1 or .4
- For obstacles:
 - There are a variable number of obstacles (1-5)
 - Obstacles are placed randomly on the left or right side of the road and with differing lengths
 - Obstacles can have a rotation to them
 - Obstacles are most of the time spread apart but sometimes jumbled together (pileup)
- Randomly there are unwinnable scenarios (the goal is blocked by a large car)



C unity

Technology

- Unity
 - Game Engine / Simulation Environment
 - o ML Agents
 - Open source
 - Reinforcement Learning Proximal Policy Optimization (PPO)
 - Neural Network
 - TensorFlow
 - Policy Gradient Method
 - Gradient descent to search space of parametrized policies
 - On-Policy algorithm
 - Policy used to make decisions is the policy being improved upon

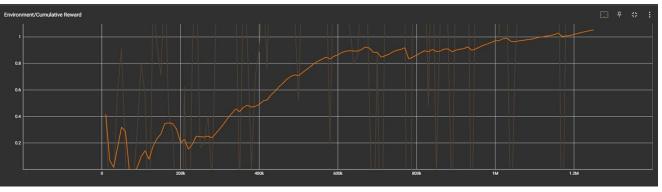


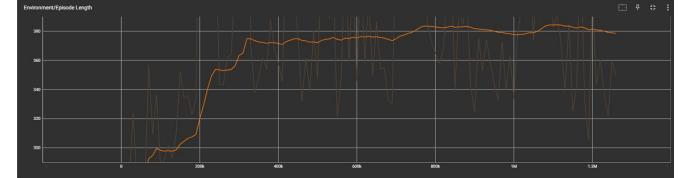
Technology - Demo

A				DEPENDENCE DEP
Tangana and	na 🕇 Bankal Thurshand Thurshand Thurshand The 🗰 Bankal			
(main stress		La Collar		
Company of the		to Trape Manufa		
Conneg Concentrations		Name of Street o		
.322		12	Sin 1927	
- 4				
1 Constants		· D · Contractor Scott		
Q-instants		1 Transford Want		
Contraction of the local division of the loc		- Contraction		
Detected Target		1 first (all lites)		
		1 has light these		
		COM - Back Fire Particles		
		- Deck Free Partners - Car Donly		
		the lating		
		No. 1 Kei Tage		
		Engine and power satisfying		
		Eren Tana		
		tina type Admenia line film		
		Mini Mutor Tangan		
		March angle Annotation Age (1914) March March	Test I	
		No. 274		
		Carlot New		
		Carl DF Ofeen MPM		
		Cardiffield Posselly Section		
		Ren To Next Dear		
		Ren 1 a Francisco		
		Rynningen für April Hands Lang Sper + Genera Maria		
		The second second		
		Rading cartings Was Dated Torque		
		Me Tour Argo Multiple		
		Opporte Angola Valocity Halp Form Postore Regular Valocity Halp Form Machingshir Valocity Halp Regis		
		March space Transition (1) March Regio Respire Transition (1) In March Regio		
		Impair Intering 2 Mile Imple		
		T B + Bay To (Base)		
	🖿 🔜 🖪 🕥 🖉 🧶 🔂 😯 🛅 🧶 🖉 🖉 🖉 🕅 🏠 😚			
	The second secon			
		1 D - De bane Gemele (Bolge)		

Results

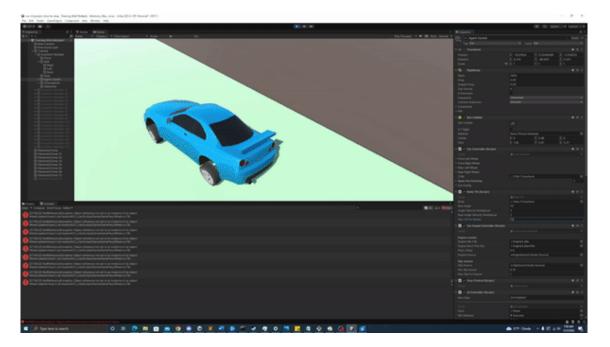
Rewards





Episode Length

Test Environment



Conclusion

- Hydroplaning and low friction/high speed environments are dangerous and difficult to navigate
- Here we provide a proof-of-concept reinforcement based machine learning model that is capable of reducing harm to the passengers of a car in a low-friction situation
- Some limitations:
 - This simulation is not immediately applicable to a real-world scenario, as the authors focused on the training, rather than how accurate the simulation environment is to a real world wet road.
 - Similarly, there was no training for curved roads, different elevations, or other moving vehicles, to simplify training (however, most hydroplaning situations are on highways, which are straight roads with high speeds)
- This system, in theory, after more improvement and simulation, could be used to safely navigate a car through obstacles in a low friction environment

References

- [1] L. Borrelli, "Hydroplaning: What it is and how to avoid it," Bankrate, 2021.
- L. Orgov an, T. B ecsi, and S. Aradi, "Autonomous drifting using reinforcement learning," 2021. [Online]. Available: https://pp.bme.hu/tr/article/view/18581/9167
- D. Schnieders, S. Bhattacharjee, K. D. Kabara, and R. Jain, "Autonomous drifting rc car with simulation-aided reinforcement learning," 2017. [Online]. Available: https://i.cs.hku.hk/fyp/2017/ fyp17014/docs/FinalProjectPlan.pdf
- A. Stocker and J. Lewis, "Variables associated with automobile tire hydroplaningg," 1972. [Online]. Available: <u>https://static.tti.tamu.edu/tti.tamu.edu/documents/147-2.pdf</u>
- R. Nave, "Friction and automobile tires." [Online]. Available: http://hyperphysics.phy-astr.gsu.edu/hbase/Mechanics/frictire.html
- https://i.cs.hku.hk/fyp/2017/fyp17014/
- https://github.com/gzrjzcx/ML-agents/blob/master/docs/Training-PPO.md
- https://github.com/Unity-Technologies/ml-agents