Reinforcement Learning for Collision Avoidance in Autonomous Driving

Nazmus Sakib
Department of Computer Science
University of Colorado Colorado Springs
nsakib@uccs.edu

Abstract

Reinforcement learning (RL) for dynamic collision avoidance in autonomous driving has gained popularity in recent years. This research intends to evaluate various RL algorithms for dynamic collision avoidance in autonomous driving. After reviewing several papers, we discovered three common reinforcement learning methods: Q-Learning, SARSA, and MLP SARSA. Therefore, the main goal of this research is to assess how well these three reinforcement learning methods in autonomous driving across a busy roadway. Additionally, we plan to train these three models in a dynamic collision avoidance urban simulation environment to evaluate their performances.

Introduction

A fundamental concept of ensuring the safety of drivers and their vehicles is collision avoidance (Yurtsever *et al.*, 2020). Due to rising vehicle density, autonomous vehicle navigation is becoming a difficult issue. There are two types of obstacles that can be found on roads: static obstacles, such as cars parked there, and dynamic obstacles, which include moving objects like animals that move randomly (Almazrouei *et al.*, 2023). Autonomous driving requires the ability to detect and avoid both static and dynamic obstacles. To ensure safety, various sensors, including vision, ultrasonic radar, and lidar, need to be used. Vision sensors can distinguish and classify obstacles, but they cannot precisely determine their distance for action. Combining ultrasonic radar or lidar data with a vision sensor improves obstacle detection accuracy in various settings. Supervised learning outperforms traditional approaches in detecting objects using vision and ultrasonic radar. However, this strategy requires extensive ground truth data from various situations to train the machine learning model (Arvind and Senthilnath, 2019). To overcome this limitation, combining supervised and reinforcement learning may eliminate the need to manually provide ground truth for obstacles (Babu *et al.*, 2016). Reinforcement learning (RL) is a machine learning technique that allows an agent, such as an autonomous car, to learn its surroundings based on prior actions, states, and rewards (Kaelbling *et al.*, 1996). In recent years, there has been an increase in the use of RL for dynamic collision avoidance in autonomous driving (Kim *et al.*, 2020).

The goal of this research is to assess several reinforcement learning algorithms for dynamic collision avoidance in autonomous driving. We identified three widely used reinforcement learning techniques: Q-Learning, SARSA, and MLP SARSA. Thus, the primary objective of this study is to evaluate the effectiveness of these three reinforcement learning techniques for autonomous driving on a busy road.

Background

Reinforcement Learning

Reinforcement learning (RL) is a branch of machine learning in which an agent functions in a given environment and tries to find a policy (π) that maximizes a cumulative reward function. The policy specifies what action should be done, a, in a state, s. The environment will thereafter shift to a new state, s', and yield a reward, r.

Q-Learning

The agent in Q-Learning, a reinforcement learning algorithm, tries to learn $Q^*(s, a)$, the optimal action-value function. The maximum expected return that results from being in a state (s), acting (a), and then according to the best course of action (π) is the definition of this function.

State Action Reward State Action (SARSA)

The State-Action-Reward-State-Action (SARSA) algorithm is a fundamental approach in reinforcement learning that operates on the principle of learning an action-value function to guide the actions of an agent within an environment. Unlike Q-Learning, which is based on the principle of maximizing future rewards, SARSA takes a more conservative approach by considering the actual next action the agent will take according to its current policy. In essence, SARSA updates its action-value function, $Q(s, a)$, based on the observed transition from the current state-action pair (s, a) to the next state-action pair (s', a') and the reward received in the process.

Multi-Layer Perceptron (MLP) based State-Action-Reward-State-Action (SARSA)

The Multi-Layer Perceptron (MLP) based State-Action-Reward-State-Action (SARSA) reinforcement learning method integrates the classic SARSA algorithm with the power of neural networks, specifically MLPs, to handle environments with high-dimensional state spaces. In this approach, the MLP is used to approximate the action-value function $Q(s, a)$, enabling the agent to learn and generalize across a vast number of states and actions. By inputting the state (and possibly the action) into the MLP, the network outputs the estimated values of taking each action in the given state, effectively learning the optimal policy through iterative updates. The SARSA update rule is applied in a way that the target for the MLP's output is adjusted based on the reward received and the estimated value of the next state-action pair, $Q(s', a')$, according to the policy being followed.

Related Work

We examine several studies on the application of reinforcement learning methods to autonomous driving. A couple of related works are listed below.

The authors of (Babu *et al.*, 2016) have developed an autonomous agent that leverages Q-learning methods to find the shortest path from a current state to a desired state in a fully dynamic environment based on camera data. Similarly, the authors of (Hong *et al.*, 2017) have used a fuzzy Q-learning method to identify obstacles in a dynamic environment utilizing input from ultrasonic sensors. The State Action Reward State Action (SARSA) method was employed by the authors of (Rais *et al.*, 2023) to avoid collisions in autonomous vehicles on highways. In an urban dynamic environment scenario, the authors of (Arvind and Senthilnath, 2019) have employed the State Action Reward State Action (SARSA) method using Multi Layer Perception (MLP) for autonomous vehicle obstacle detection and avoidance. The authors of (Arvind and Senthilnath, 2020) employed a policy-free, model-free Q-learning based reinforcement learning algorithm with a multi-layer perceptron neural network (MLP-NN) to forecast the best course of action for the vehicle in the future, given its present state.

In contrast to earlier studies, the objective of this research is to evaluate the performance of these three reinforcement learning techniques for driving autonomously over a busy road. In addition, we intend to assess the performance of these popular reinforcement learning methods by training them in a dynamic collision avoidance urban simulation scenario.

Proposed Methodology

The self-driving car used in this experiment as a reinforcement agent is fitted with sensors that must understand its environment, which is its immediate surroundings. With reinforcement learning, the agent must understand its surroundings by considering its future state and the rewards it will receive for its current actions. More detailed details on how we want to implement each of the different pieces of the reinforcement learning algorithm are likely to be given below.

States

The car has only three states: Normalized car speed, Normalized relative position of car, and Normalized relative position of the obstacle.

Actions

Autonomous cars are quite simple to control and therefore the actions are quite simple as well. The car has only four actions: Move forward, Stop, Turn Left, and Turn Right.

Reward

A straightforward reward function will be used to assist the car's agent in learning which activities result in positive and bad results. Two distinct reward functions are what we wish to incorporate. In the first case, if the car avoids a collision, a positive reward of +100 will be given. The agent will then receive a -100 reward if they hit an obstacle.

Environment

Using the pygame package, we want to train our models in a complex urban simulation environment with dynamic obstacles (Gym and Sanghi, 2021).

Conclusion

To achieve dynamic collision avoidance in autonomous vehicles, this study aims to examine three reinforcement learning (RL) algorithms: Q-Learning, SARSA, and MLP-based SARSA. The main goal is to assess critically how well these algorithms work for safely driving autonomous cars across busy highways. Through the use of these three RL methods in a simulated urban collision avoidance scenario, the research aims to provide significant insights into how well they perform about one another.

Timeline

This is a general timeline that shows how we think the project will advance.

Table 1: Project Milestones and Timelines

Date	Milestone
3/1/2024	Environment Setup for simulation.
3/10/2024	First implementation of RL agents.
3/18/2024	Complete working on implementing RL agents and prepare demo and midterm report.
4/15/2024	Update and improve RL agents.
4/25/2024	Update and improve RL agents for Final version.
5/1/2024	Complete Final version of project and report.

References

Khawla Almazrouei, Ibrahim Kamel, and Tamer Rabie. Dynamic obstacle avoidance and path planning through reinforcement learning. *Applied Sciences*, 13(14):8174, 2023.

CS Arvind and J Senthilnath. Autonomous rl: Autonomous vehicle obstacle avoidance in a dynamic environment using mlp-sarsa reinforcement learning. In *2019 IEEE 5th International Conference on Mechatronics System and Robots (ICMSR)*, pages 120–124. IEEE, 2019.

CS Arvind and J Senthilnath. Autonomous vehicle for obstacle detection and avoidance using reinforcement learning. In *Soft Computing for Problem Solving: SocProS 2018, Volume 1*, pages 55–66. Springer, 2020.

V Madhu Babu, U Vamshi Krishna, and SK Shahensha. An autonomous path finding robot using q-learning. In *2016 10th International Conference on Intelligent Systems and Control (ISCO)*, pages 1–6. IEEE, 2016.

O Gym and Nimish Sanghi. *Deep reinforcement learning with python*. Springer, 2021.

Jun Hong, Kaiqiang Tang, and Chunlin Chen. Obstacle avoidance of hexapod robots using fuzzy q-learning. In *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–6. IEEE, 2017.

Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore. Reinforcement learning: A survey. *Journal of artificial intelligence research*, 4:237–285, 1996.

Myounghoe Kim, Seongwon Lee, Jaehyun Lim, Jongeun Choi, and Seong Gu Kang. Unexpected collision avoidance driving strategy using deep reinforcement learning. *IEEE Access*, 8:17243–17252, 2020.

Mohamed Saber Rais, Rachid Boudour, Khouloud Zouaidia, and Lamine Bougueroua. Decision making for autonomous vehicles in highway scenarios using harmonic sk deep sarsa. *Applied Intelligence*, 53(3):2488–2505, 2023.

Ekim Yurtsever, Jacob Lambert, Alexander Carballo, and Kazuya Takeda. A survey of autonomous driving: Common practices and emerging technologies. *IEEE access*, 8:58443–58469, 2020.