# Varying Racing Line Deviation Penalties to Analyze Passing Performance

AA228 Final Project

Alaisha Alexander
*Stanford University*
alaisha@stanford.edu
December 6th, 2019

***Abstract*—Autonomous racing presents a rich test bed for not only pushing the limits of the vehicle but also addressing decision making in highly dynamic situations. Every race has clearly defined rules on passing areas as well as optimal racing lines for various tracks. For a passing scenario, we explored the performance of policies created for varying penalties to deviation from the racing line. It was analyzed as a non-banked oval racing track modeled as a Markov Decision Process (MDP) with two vehicles, ego and other. The penalties were considered in the reward functions. The resulting policies were evaluated using simulation. We evaluated how many laps it takes to overtake the other vehicle, the average speed of each policy before passing, and the number of lateral maneuvers taken to gain insight into the performance of each policy. Increased penalty on deviating from the racing line caused increased lateral maneuvers and additional laps to pass but decreased penalty resulted in higher speeds and additional passing in unsafe zones.**

## I. Introduction

### A. Background

From carts, horse-pulled wagons, and, today, cars, transportation has been central to how humans have interacted with the world around them. Much like the rapid advances of communication, transportation has made the world seem smaller and has connected people across the globe in ways that would've been unimaginable. However, those rapid advances have come with increased reliance on technology and the challenges that come with that.

It is then no surprise that when thinking about the future of transportation, autonomous driving dominates the conversation. The potential for increased safety and efficiency is evident but the technical challenges to get to the future envisioned are present at every level of the autonomous stack. When pushing the limits of the vehicles, it is not a surprise that autonomous racing is a rich opportunity to understand how to leverage the abilities of vehicles for greater safety.

Autonomous racing presents a rich test bed for not only pushing the limits of the vehicle but also addressing decision making in highly dynamic situations. Human race car drivers utilize present information (i.e. their position relative to other cars on the tracks, landmarks, etc.), knowledge of enforced rules, and their previous experience to execute various maneuvers and improve their performance. This improved performance can be seen in shorter lap times, smooth passing, and consistency in following some globally optimal racing line.

### B. Problem Description

For this scenario of passing, we explore the performance of policies created for varying penalties to deviation from the racing line. The racing line is defined as the center line for the straights of the track and the inner lane for the curves. There is further elaboration of the environment in II-E. The racing line is typically the globally optimal path for completing the lap as quickly as possible and utilizing the friction available to the vehicle. However, in a passing maneuver the optimal set of actions may not follow the racing line and may require deviations in order to pass safely [1].

Every race has clearly defined rules on passing areas as well as optimal racing lines for various tracks. These passing areas are usually restricted to long straights or wide curves to ensure the safety of the drivers involved. When it comes to vehicle dynamics, there are also limits to the lateral motion of the vehicle and the utilization of friction when turning. The scenario of interest is an ego vehicle choosing some particular method (or policy) to pass a slightly slower vehicle ahead of it on the inner line of the track. There will additionally be variation in the ego vehicle's preference for the racing line.

This scenario will be analyzed on a non-banked oval racing track that is discretized into segments. It will be modeled as a Markov Decision Process (MDP). The positions of the vehicles will act as the states; the ego vehicle decision for moving forward as the actions; and the rewards (and penalties) for passing, traveling on the outer areas of curved segments, and deviating from the racing line as the rewards.

## II. Problem Formulation

To formulate this scenario, we assume a few simplifications. This racing track is discretized to allow the problem to be represented as an MDP with a discrete state space and discrete action space, similar to how Wei et al [2] pursue a point-based approach. We also assume full observability so there is no uncertainty regarding observations of the position of the ego vehicle or the vehicle we intend to pass. An MDP is defined by its state space, action space, transition function, reward function, and discount factor and can be represented by the

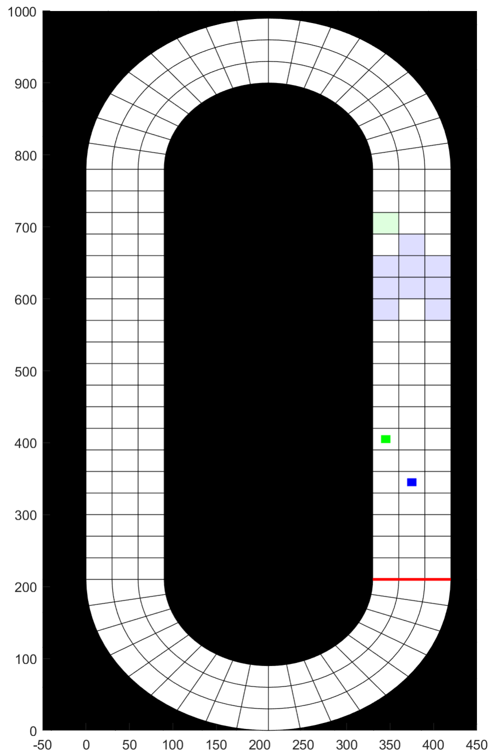


Fig. 1. Snapshot of track in state [5 2 7] where ego vehicle is in blue and other vehicle is in green. Possible actions for ego vehicle are highlighted in light blue and next position of other vehicle highlighted in light green for reference.

following tuple: $\{S, A, T, R, \gamma\}$. The discount factor is set to a value of 0.95 for all calculations to resolve the consideration of value of rewards at present and future states.

### A. Environment

The track is discretized into two straight segments separated by two curved segments. The track is also separated into three distinct lanes (inner, middle, and outer). The track begins at the bottom right of the straight segment and loops around such that once a vehicle reaches the end of the track, it immediately starts at the beginning. Fig. 1 shows an illustration of the environment with a red line marking the beginning/end of the track. The track models that of a one-mile paperclip track so the straights are much longer than the tight radius turns but the reduced speed necessary to take the turns gives them more states. The straights are discretized into 19 individual segments per lane; the turns into 18 individual segments per lane. The whole track is then 74 individual segments per lane. The vehicles would be moving counterclockwise.

### B. State Space

The state space is comprised of the position of the ego vehicle and the other vehicle along the track. For the ego vehicle this is represented as its length along the track $x_e$ and its lane $y_e$. Since we assume the other vehicle is riding only on the inner lane, its position is only represented as its length along the track $x_o$ and its lane is always considered the inner lane (or $y_o = 1$).

This results in a state space of size $|S| = 74 \times 3 \times 74 = 16,428$. Each state is of the form $[x_e \; y_e \; x_o]$. For reference, Fig. 1 shows the track with both vehicles on the track. The state shown is [5 2 7].

### C. Action Space

The action space is comprised of nine actions representing the three possible speeds of the ego vehicle and the decision to stay in the current lane or move to the right or left adjacent lanes. If the ego vehicle moves to an adjacent lane, it utilizes part of its speed to move laterally.

The ego vehicle can move $10\%$ slower than, just as fast as, or $10\%$ faster than the other vehicle. So if the other vehicle can move 10 spaces forward for each action, the ego vehicle can move 9, 10, or 11 spaces, sacrificing one space forward if it moves laterally in either direction. An example of the possible actions for the ego vehicle are shown in Fig. 1.

### D. Transition Function

The transition function $T(s'|s, a)$ provides the probability of reaching state $s'$ from state $s$ given action $a$. In this problem no matter the action, the other vehicle will always move forward at some predetermined speed so its dynamics are deterministic. The ego vehicle's dynamics are stochastic when it changes lanes. If the ego vehicle does not change lanes, it moves forward the amount given by the action. If the ego vehicle does change lanes, the lane change can fail and it moves forward, in whatever lane it was previously in, the reduced amount.

For example, if the ego vehicle planned to move 10 spaces but change lanes to the right and it fails, it will move forward in its current lane by 9 spaces. Additionally, if the ego vehicle intends to change into a nonexistent lane, it is treated as a failed lane change. The probability of changing lanes successfully was set at 0.9.

### E. Reward Function

The reward function $R(s, a)$ provides the expected reward given state $s$ and action $a$. The reward function for this problem contains several types of rewards for various maneuvers. The following are the rewards contained in the reward function determined by the state and/or the action.

- +1 – if the ego vehicle has successfully passed the other vehicle
- -10 – if the ego vehicle passes the other vehicle not in the passing zone[1]
- -1000 – if the ego vehicle crashes into the other vehicle (residing in the same space)

[1]In racing, passing zones are typically determined by the type of track and the expected skill of the drivers. In this case with an oval track, we will assume the drivers are amateurs (especially considering the slower inner lane driving of the other vehicle). As such, the allowed passing zones are only the straights of the track. This can obviously be modified.

- -1 – (per speed unit) if the ego vehicle is traveling on the outer two lanes of the curve (to penalize for style and lateral motion to maintain its increased speed on the curves)
- -1 – if the ego vehicle attempts to change into a non-existent lane
- 0.01 – (per speed unit) if the ego vehicle attempts to close the gap between it and the other vehicle
- $r$ – if the ego vehicle deviates from the racing line

We will be varying $r$ to manage the penalty of not adhering to the racing line.

### *F. Simulation Conditions*

To examine the performance of the policies, we will run 100 simulations per starting distance utilizing each policy. The simulation terminates once the ego vehicle has successfully passed the other vehicle or the other vehicle has completed 50 laps (whichever comes first). For each of the simulations, we will record how many laps it takes to overtake the other vehicle, the average speed of each policy before passing, and the number of lateral maneuvers taken. Each of these should give insight into the aggressiveness and effectiveness of the policies generated by the varying racing line penalties.

## III. Policy Creation

To determine an optimal policy given the varying reward functions, value iteration was ideal given the clearly defined model and moderately sized state space. Value iteration iteratively computes an optimal value function which allows us to find an optimal policy [3]. For each state, we compute the value function

$$U_{k+1}(s) = \max_a \left( R(s,a) + \gamma \sum_{s'} T(s'|s,a) U_k(s') \right)$$

until the value function converges. We test this by evaluating when $||U_k - U_{k-1}|| < \delta$, where $\epsilon(1-\gamma)/\gamma$. For this case, we will set $\epsilon = 0.01$ so that the value function is within $1\%$ of the optimal value function $U^*$. From that we can extract and optimal policy

$$\pi(s) = \arg\max_a \left( R(s,a) + \gamma \sum_{s'} T(s'|s,a) U^*(s') \right)$$

Each policy is then simulated in the simulation conditions described in II-F.

## IV. Results

To analyze the performance of each policy in varying scenarios, the starting distances were varied from two units to twelve and $r$ was varied from -0.1 to -1.5. The results of these is shown in how many laps it takes to overtake the other vehicle, the average speed of each policy before passing, and the number of lateral maneuvers taken displayed in Fig. 2, Fig. 3, and Fig. 4 respectively.

Regarding laps to overtake the other vehicle, it seems the more weight attributed to staying closer to the racing line, the

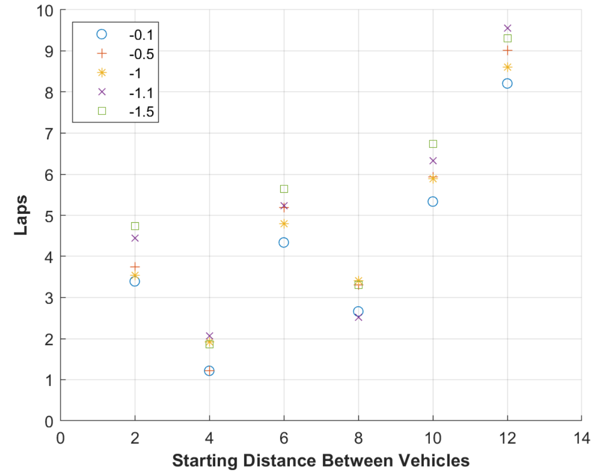


Fig. 2. Simulation Data: Average laps taken by the ego vehicle to overtake the other vehicle.

more laps needed to overtake the other vehicle. The difference between $r = -0.1$ and $r = -1.5$ being an entire lap for most of the varied starting distances between vehicles. Also, between $r = -0.5$ and $r = -1$ there is not much of a striking difference.

When also looking at speed in Fig. 3, the differences between the various policies are much narrower. With speed options in the actions varying from 8 to 11 depending on lateral movement, it's worth it to note that the average for all the policies for passing the other vehicle sticks between 10 and 11, but closet to 10 in most cases. All the policies perform very closely with the more aggressive policy being that of $r = -0.1$. In most starting distances, it is slightly faster than the others. Since it's not as beholden to the racing line, it's likely that it's moving forward more quickly without

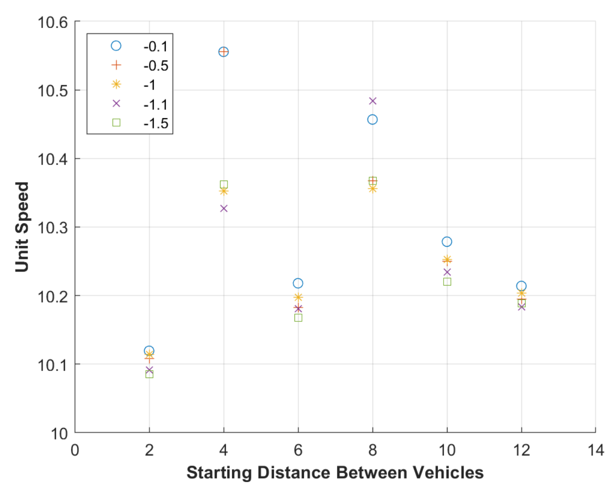


Fig. 3. Simulation Data: Average speed units per action taken by the ego vehicle in pursuit of closing the gap between itself and the other vehicle before passing.

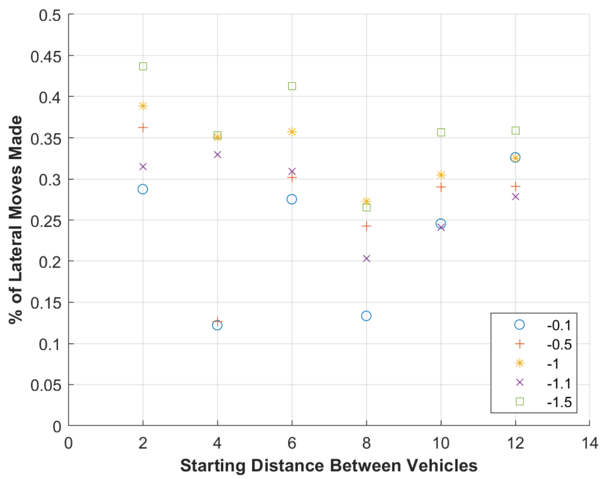


Fig. 4. Simulation Data: Lateral moves per action taken by the ego vehicle in pursuit of closing the gap between itself and the other vehicle before passing.

regard for lateral penalties on curves.

This seems especially clear when also considering Fig. 4 which shows that, especially for smaller starting distances, $r = -0.1$ makes very few lateral moves per action. As $r$ increases and there is more of a preference for the racing line, it seems to also result in more lateral moves per action, which in practice would be non-ideal. Not only would the vehicle be slowing down making it susceptible to passed by the other vehicle should it react, but it would also make the driving seem erratic and unsafe.

## V. Conclusion

Although this problem formulated is a simplified version of the complexities of racing, it gives a clear idea of the compromises faced by drivers of all skill in a fast-paced, high-stakes racing scenario. Additionally, each policy brings up additional questions about the comparative weighting and strictness of possible decision-making structures of future autonomous vehicles. We sought to compare the performance, efficiency, and safety of these different policies which balance the desire to pass with a certain "rule of the road". Though the racing line wasn't a strictly enforced rule that overtly restricted the movement of the vehicle, the expectation to follow it through, even in a situation when it is not the locally optimal response, caused increased lateral movement which would be more aggressive and frenzied. However, the lack of this racing line encouraged increased speeds and faster passing, likely in unsafe passing zones. It is really the balance of these behaviors that would be likely to prove the most efficient but safest behavior.

## VI. Further Work

Moving forward, it would beneficial to build from this point by exploring a parameter based reward function, a more finely discretized model, or additional uncertainty in the position and movement of the other vehicle. These additional layers of complexity would give additional insights and a more robust understanding of the scenario. It would also be beneficial to explore other racing scenarios, or general autonomous driving qualms, with a similar structure to explore the behaviors possible from these models.

## References

[1] J. Kunzman, J. Politi, and B. Leggett, "National Auto Sport Association: CLUB CODES AND REGULATIONS."

[2] J. Wei, J. M. Dolan, J. M. Snider, and B. Litkouhi, "A Point-based MDP for Robust Single-Lane Autonomous Driving Behavior under Uncertainties."

[3] M. J. Kochenderfer and B. Brodsky, "Sequential Problems," Chang. Anal. Nonstationary Stoch. Model., pp. 175–180, 2016.