Modeling & Simulation Project

1.9 - Reinforcement Learning for Autonomous Highway Driving

Group EIC1_09 | Bruno Machado Henrique Gardé José Felisberto

October 24, 2025

1 Brief description of the problem to be modelled

This project focuses on the simulation of highway driving scenarios to train and evaluate an autonomous vehicle (ego car) capable of performing maneuvers such as lane changing, merging, and overtaking in dynamic traffic environments. The study leverages reinforcement learning to enable the ego vehicle to make sequential decisions that balance safety, efficiency, and compliance with traffic regulations. Through interaction with a simulated environment, the RL agent learns optimal driving behaviors that adapt to varying traffic densities and vehicle behaviors.

2 Goals of the simulation project

The problem addressed in this study is the design and implementation of a simulation framework that enables an autonomous vehicle to learn safe and efficient driving strategies in realistic, dynamic traffic environments. The system must account for diverse traffic conditions, including varying vehicle densities, aggressiveness levels, and speed distributions, to ensure robustness and adaptability of the learned policy.

The main challenge lies in enabling the reinforcement learning agent (ego vehicle) to make effective driving decisions—such as when to change lanes, accelerate, or brake—while minimizing the risk of collisions and maintaining compliance with road rules. The simulation environment must therefore balance realism, controllability, and computational efficiency to allow for effective model training and evaluation.

The ultimate goal is to optimize the vehicle's driving policy for both safety (avoiding collisions and unsafe maneuvers) and efficiency (maintaining smooth traffic flow and timely goal completion) under a range of traffic scenarios.

3 What models of decision support are considered

This simulation is primarily **Prescriptive**, since the goal is to design and evaluate strategies (via reinforcement learning policies) that autonomous vehicles can use to act safely and efficiently in traffic scenarios such as lane changing, merging, and overtaking.

It also has a **Predictive** aspect, as the trained agent implicitly forecasts the outcomes of its possible actions (whether a lane change will lead to a safe maneuver or a collision) in order to maximize long-term rewards.

It is not **Descriptive**, because the model is not focused on replicating current human driver behavior. Instead, it seeks to propose better strategies. Nor is it purely **Speculative**, since the experiments are grounded in realistic traffic environments and established simulation frameworks.

4 Main entities of the system

Autonomous Vehicle (Ego Car) The reinforcement learning agent we are training. Makes driving decisions (lane changing, merging, overtaking). Learns through interaction with the environment.

Other Vehicles (Traffic Environment) Surrounding cars driven by predefined behavior models. Provide dynamic obstacles and interactions. Influence task difficulty (traffic density, aggressiveness, speed distribution).

Road/Highway Environment Multi-lane highway with traffic flow. Defines constraints: lanes, speed limits, road length, entry/exit ramps. Provides sensory input to the ego vehicle (positions, velocities).

Note: while the RL algorithm, reward function, and evaluation metrics play a central role in how the simulation operates, they are considered modeling components rather than entities of the traffic system itself.

5 Variables of the system

Entity	**Properties**	**Notes**
Ego Vehicle	Position, velocity, acceleration, lane index, action space (lane change, accelerate, brake)	Controlled by RL agent
Other Vehicles	Position, velocity, lane, driving policy	Define traffic dynamics
Road/Highway	Number of lanes, length, speed limits, entry/exit ramps	Defines environment constraints

6 Operation policies to be tested (scenarios)

The following scenarios will be used to evaluate robustness and generalization of the learned policy:

- **Traffic density levels:** low, medium, high density.
- **Driver aggressiveness:** variation in behavior models of other vehicles.
- **Speed limit configurations:** different lane-wise or global limits.
- **Algorithm variants:** compare PPO, DQN, and SAC.
- **Reward shaping variants:** safety-heavy vs. efficiency-heavy weighting.
- **Observation fidelity (optional):** ideal perception vs. noisy/partial observations.

7 Key performance indicators and decision criteria

KPIs (primary measurements)

Entity	Metric	Description	Notes
Ego Vehicle	Collision Rate	Fraction of episodes ending in collision. Measures driving safety.	Lower is better; key for safety evaluation.
Ego Vehicle	Average Speed	Mean longitudinal velocity over an episode.	Indicates efficiency; should respect speed limits.
Ego Vehicle	Lane Change Frequency	Number of lane changes per episode.	Too high → erratic; too low → passive.
Ego Vehicle	Reward per Episode	Total accumulated reward per episode.	Global RL performance indicator.
Ego Vehicle	Speed Limit Compliance	Fraction of time speed $\leq$ limit.	Adherence to road rules.
Ego Vehicle	Time-to-Collision (TTC)	Minimum time before potential collision (averaged per episode).	Safety measure.
Ego Vehicle	Episode Duration	Time until success or termination.	Lower is efficient, not at safety's expense.
Other Vehicles	Traffic Density	Vehicles per km or per lane.	Parameterizes task difficulty.
Other Vehicles	Average Relative Speed	Mean speed difference (ego vs. others).	Influences overtake/merge difficulty.
Other Vehicles	Aggressiveness Index	Tendency for abrupt accel/brake or close following.	Behavior model complexity.
Other Vehicles	Traffic Flow Stability	Standard deviation of speeds in the scene.	Lower = smoother traffic.
Road/Highway	Lane Utilization Ratio	Percentage of time ego uses each lane.	Shows strategic lane use.
Road/Highway	Global Speed Compliance	% of vehicles within legal speed range.	Realism and rule adherence.
Road/Highway	Average Throughput	Vehicles passing a point per unit time.	Overall efficiency.
Road/Highway	Road Occupancy Rate	Portion of road length occupied by vehicles.	Congestion level.
Road/Highway	Scenario Completion Rate	Fraction of successful runs (goal reached, no collision).	Stability and effectiveness.

Decision criteria (composite indicators)

Indicator	Definition / Formula	What it measures	Notes
Safety Index (SI)	$SI = w_1\,(1 - \text{CollisionRate}) + w_2\,\overline{\text{TTC}}_{\text{norm}} + w_3\,(1 - \text{ViolationRate})$	Overall safety and risk avoidance	Weights tuned per experiment.
Efficiency Index (EI)	$EI = w_1\,(\overline{v}/v_{\text{limit}}) + w_2 \cdot \text{CompletionRate}$	Traffic efficiency and task success	Balance speed and completion.
Comfort Index (CI)	$CI = 1 - \text{norm}(\text{Jerk} + \alpha \cdot \text{LaneChangeFreq})$	Smoothness and comfort	Higher = smoother.
Rule Compliance Index (RCI)	$RCI = 1 - \dfrac{\text{SpeedViolations} + \text{DistanceViolations}}{\text{TotalTime}}$	Adherence to rules	Penalizes overspeeding/unsafe following.
Learning Efficiency (LE)	$LE = \dfrac{\text{PerformanceScore}}{\text{TrainingSteps}}$	Sample efficiency of RL	Compare algorithms.
Traffic Flow Index (TFI)	$TFI = \dfrac{\text{Throughput}}{\text{Density}}$	Network-level movement quality	Higher = better coordination.
Environment Stability Index (ESI)	$ESI = 1 - \dfrac{\text{std(speeds)}}{\text{mean(speeds)}}$	Stability/consistency of flow	Low variability → stable.
Safety–Efficiency Trade-off (SET)	$SET = \beta_1 \cdot SI + \beta_2 \cdot EI$	Balance between safe/efficient	Compare policy trade-offs.
Global Performance Score (GPS)	$GPS = a \cdot SI + b \cdot EI + c \cdot CI + d \cdot RCI$	Overall system performance	Weights tuned experimentally.

The definitions and symbolic forms of the composite indicators were initially drafted with assistance from an OpenAI GPT-4–class large language model, and were subsequently verified and refined to align with the project's objectives and notation.

8 Data requirements

The simulation relies on both synthetic and benchmark traffic data to represent realistic driving conditions. The data required primarily concerns vehicle kinematics and traffic flow dynamics, which define how vehicles move and interact on a multi-lane highway.

Type of data. The simulation collects and processes information on vehicle position, velocity, acceleration, lane index, and inter-vehicle distances. Traffic flow variables such as vehicle density, average speed, and lane occupancy are also used to characterize the environment and evaluate system performance.

Sources. Data is obtained from synthetic simulations using platforms such as *Highway-Env* (Python).

Assumptions. All vehicles comply with physical kinematic constraints (e.g., maximum acceleration and braking limits). Weather and road conditions remain constant within each simulation episode to isolate learning effects. Sensor perception is assumed to be ideal, meaning no measurement noise or occlusion is introduced at this stage of development.

9 Simulation tools, environments, and languages

To implement and evaluate the autonomous driving simulation, a combination of open-source simulation environments, reinforcement learning libraries, and analysis tools are employed.

Simulation tools. The project utilizes *Highway-Env* for generating realistic highway traffic scenarios. These platforms provide controllable conditions for defining lane configurations, traffic densities, and vehicle behaviors, ensuring reproducible experiments.

Programming environment. Development and experimentation are conducted in Python (Gymnasium interface) for consistent agent–environment interactions. RL algorithms (e.g., PPO, DQN, SAC) are implemented via *Stable-Baselines3*. Performance and training progress are monitored with *TensorBoard* and custom analysis scripts.

10 Other information

Real-world testing of autonomous driving policies is costly and unsafe; simulation provides a controlled and reproducible setting for experimentation. The project emphasizes reproducibility (fixed seeds, saved configs, and logs) and ethical considerations around safety-first evaluation criteria.