Testing vehicle safety in simulated driving scenarios

Final Project Report

Author: Vakaris Paulavičius

Supervisor: Professor Mohammad Reza Mousavi

Student ID: 20062023

April 5, 2023

Abstract

The vision of a civilisation where driverless cars seamlessly navigate the lively streets and high-speed underground tunnels, transporting passengers from point A to point B, is remarkably portrayed in many futuristic movies and science fiction novels. However, this futuristic dream is not as far-off as it may seem. In fact, many vehicles operating the streets today already incorporate some level of automation, such as cruise control and automatic parking. While it is clear that we are making progress towards fully autonomous vehicles, the question remains: when will we finally witness entirely driverless cars on the roads? Although autonomous vehicles already hold the potential to surpass human drivers in terms of safety, there is currently no practical way to test this claim. Or is there? This paper proposes an innovative approach - evaluating high-level vehicle safety through simulations on virtual roads in an automated and time-efficient manner. The proposed software consists of three main components: an automated scenario generation tool that, with the help of a sophisticated machine learning algorithm, can generate diverse driving scenarios, a vehicle safety monitoring system designed for use in the simulator CARLA and data analysis tools allowing for efficient and visualised analysis of obtained data through diagrams and location marking on 2D map representations. In addition to introducing the automated safety testing software bundle, this article also dives deeper into the realm of AVs by discussing the technology, safety vulnerabilities, how AVs are classified and the comparison of human drivers to computer-controlled cars.

Originality Avowal

I verify that I am the sole author of this report, except where explicitly stated to the contrary. I grant the right to King's College London to make paper and electronic copies of the submitted work for purposes of marking, plagiarism detection and archival, and to upload a copy of the work to Turnitin or another trusted plagiarism detection service. I confirm this report does not exceed 25,000 words.

Vakaris Paulavičius

April 5, 2023

Abbreviations

Here is a list of abbreviations that are used throughout the report. It is assumed that not every person reading this article is fluent in computer scientists' terminology and thus, having a guide might come in handy.

AI artificial intelligence

AV autonomous/automated vehicle

ML machine learning

XAI explainable artificial intelligence

ISO International Organization for Standardization

NHTSA National Highway Traffic Safety Administration

Contents

Chapter 1

Introduction

This research paper aims to present a methodology for evaluating how well modern AV implementations can participate in realistic traffic scenarios and get from one location to another safely.

The fundamental obstacle that science and major car manufacturers need to overcome to set up the widespread recognition of autonomous vehicles is establishing trust in the AVs and assuring that they can be a safer and more convenient means of transportation. This is is well presented in the available literature considering autonomous vehicles [1].

However, according to the 2022 survey carried out by Policygenius in the U.S. [2], 76% of respondents said that they would feel less safe driving in cars with self-driving features, while a similar proportion (73%) said they would feel less safe knowing others are driving on the road in autonomous vehicles. This shows that although people have been restless for decades trying to establish general acceptance of technology that will massively alter how we see and use transportation today, they have not succeeded yet.

One thing that could change this perception is autonomous vehicles appearing on the streets and demonstrating that statistically, they can be better vehicle operators than human drivers. However, the law that we live by today does not allow autonomous vehicles to navigate the streets freely because there is a lack of standards and regulations for it, and it is not defined who has to be responsible if a machine makes a mistake. The recently emerged branch of AI, the XAI, promises to take this a step forward, providing the machine with the ability to explain its actions. This is particularly relevant in situations where an AV gets involved in an accident, and it is crucial to understand the circumstances that led to this and what were the vehicle operators doing to prevent this or at least minimise the impact.

An alternative that is available today is demonstrating the AI potential in a virtual, simulated world where there is no law preventing autonomous vehicles from driving on the streets and no living being able to get hurt. This article focuses on creating a framework that would allow testing the AV implementations in such a virtual setting and analysing the observations, thus determining whether the AI agent is acting safely and, if not, where the issues lie.

We start by exploring the concept of autonomous vehicles, how safe they could be and what vulnerabilities they might have if they are finally put on the streets. In Chapter 3, we continue our journey by analysing the road statistics in Europe and the U.S. from 2021 until 2022 and discussing how the scenario generation algorithm could be designed based on the observations. Subsections of the chapter mentioned above also analyse human behaviour in driving situations and discuss the relevance of statistics to this project. In Chapter 4, we discuss the chosen simulator CARLA, its advantages and shortcomings, and explain its role in the research while also covering details of how different AV implementations could be used with the simulator. In Chapter 5, we finally arrive at the scenario generation tool, one of the three main framework components. This part covers the tool's development process and inner workings while also emphasising the importance of road statistics and traffic data for its development. The following Chapter 6, dives into detail about how safety can be measured in the simulation environment, what tools are needed for that and describes in detail the safety measuring system this article is proposing that can determine how safe the vehicles are performing in the simulated scenarios. This section also provides an overview of what safety metrics are evaluated and proposes a formula to assess the performance. In Chapter 7, we introduce the data analysis tool, the last of the three main software components. We talk in detail about how it can be used to both analyse the performance scores of the participated agents and group interesting clusters of information obtained from the sensors and simulation recordings. After introducing the main software components, in Chapter 8 we will swiftly look at how the system functions as a whole and how all the modules work together. Chapter 9 focuses on evaluating how well the system functions by looking at the research involving human participants and a CARLA agent driving the generated scenarios. The chapter looks at the safety monitoring system in action and applies the data analysis tools to assess participant performance. It also compares how well the CARLA autonomous agent is executing the given scenarios compared to human participants.

The article ends in Chapter 10, having presented a working framework that generates diverse driving scenarios based on real-life data, allows human drivers and AI agents to drive in the

scenarios and measures their safety performance. This last chapter overviews what has been achieved and what could be improved. It discusses the relevance of this work and why the chosen subject is so impactful. In addition, it talks about the ethical questions concerning the project and the field of autonomous driving as a whole. The last section of the article provides a short self-reflection of the author about the challenges faced throughout the project's development.

Chapter 2

What is an autonomous vehicle

Before diving into the technical details that surround this project, it is crucial to make it clear what AV stands for and how the ideas of its development matured over the course of human history (Section 2.1). This chapter aims to give an overview of what makes an AV autonomous (Section 2.2), talks about the standardised way of classifying the autonomy levels in Section 2.3 and discusses the AV safety vulnerabilities in Section 2.4.

2.1 History of self-driving cars

The idea of a self-driving car surprisingly reaches back as far as the 1500s when Leonardo Da Vinci designed his self-propelled cart [3] that could move without being pushed or pulled. Some regard this as the grandfather of today's automated vehicles. Although it might not have looked like something revolutionary, it shows that people living more than 400 years before the emergence of the first computer were already looking for ways to reduce human interaction in such basic everyday activities as pushing a cart.

In 1925, American engineer Francis Houdini tested the first driverless car that was operated using a remote-control setup that guided the car using radio signals [4]. Despite the fact that the car riding Manhattan streets demonstrated what an autonomous vehicle could look like, it did bring little success to the constructor. The project was shut down when the operator lost control of the vehicle twice and eventually crashed it into another vehicle.

Although the two occasions mentioned above sound like attempts to produce an autonomous vehicle, both Da Vinci's and Houdini's creations needed human input in their movement, whether winding the clog to start it or using the remote control device to decide its move-

ments.

However, things changed forever when in 1958, almost twenty years after the initial concept was created, General Motors launched their prototype of a car that could ride the road on itself following a wire embedded into the asphalt. The concept was simple. The car contained sensors that could sense the current flowing through the wire, thus making it go according to the location of the wire and where it leads. It showed that a truly autonomous car has to be able to decide what to do by itself, relying on the data captured from the surrounding world.

This idea was reflected in the upcoming tries to create a self-driving vehicle. One includes a Japanese car built in 1977 that could slowly drive on the streets following the white street markers, and it could do so using cameras that were able to identify lane markings. In 1995, researchers at Carnegie Mellon University developed a car called NavLab5 [5] that could steer itself on roads and highways using computers and video sensors. It was taken on an almost 3000 miles tour through the U.S., where it manoeuvred itself 98% of the time, with human supervisors only needing to brake and throttle.

As highlighted by the organisation TWI Global - "Today, a fully autonomous vehicle, is considered one that can operate itself and perform necessary functions without any human intervention through the ability to sense its surroundings." [6]

One question that might arise is how exactly the vehicle would be able to know all its surroundings, determine what actions are legal and allowed, and navigate from one location to another without violating rules and causing harm to people or damaging properties. It might be about time to discuss the inside of an autonomous vehicle.

2.2 What makes an AV

At the core of an AV is a car with all the necessary parts for riding on the road. What distinguishes it from the cars that people are used to seeing is that it is equipped with multiple sensors providing all the necessary information about the world surrounding the vehicle. That data is then processed by the software responsible for all the logic and calculations that determine what commands need to be given to the semiconductors that make real-time driving actions.

Without sensors, the idea of an autonomous car would be impossible. They allow the car to see what is on the road, analyse what other road participants are doing, predict their movements and collect information needed to make future decisions regarding steering, braking or acceleration. Three essential sensors are needed for an autonomous vehicle[7], the first being

LiDAR.

LiDAR stands for "Light Detection and Ranging". It is by far the most crucial sensor in the realm of autonomous vehicles because it allows for object detection and range estimation. It emits laser beams that bounce off objects and return to the sensor allowing it to create a detailed 3D map of the environment around it, letting it know where every object is and how far it is from the vehicle. It might come as a surprise, but many modern phones are also equipped with LiDAR sensors allowing for features such as taking selfies with a blurred background, playing augmented reality games or simply measuring distance using the phone's camera. LiDAR sensors come in handy when determining the exact position, size, and shape of an object is needed. They are typically mounted on the vehicle's roof, providing a high-resolution 360-degree coverage of the vehicle's surroundings. They work well in dark conditions but have a limited range, typically a few hundred meters and lack accuracy in foggy and snowy conditions [8].

That is where radars come into help. Using radio waves, they can detect objects surrounding the car and determine how far they are and how quickly or in what direction they are moving. For this reason, many car manufacturers use radars for parking sensors. One advantage that radars have compared to LiDAR sensors is their range which can be up to several kilometres in some cases. In addition, radars are less affected by weather conditions because, unlike LiDAR sensors, they do not rely on lasers but radio waves. Because of the advantages and disadvantages of both LiDAR sensors and radars, they are often used in combination. Where one lacks precision, the other one contributes. Where one is affected by the rain and cannot function properly, the other provides reliable enough data.

Cameras are the third type of commonly used sensor. They capture high-resolution images and videos of the world surrounding the vehicle. Images can then be used to read road signs, determine the colour of the traffic light, and detect pedestrians and other road objects. Although cameras are much cheaper than LiDAR sensors, they have disadvantages in detecting stuff in low light or high contrast situations. Cameras also have a limited field of view and are less effective in detecting objects that are far away or not entirely in the camera's frame.

AVs also utilise GPS sensors that use satellite signals to determine a vehicle's location and track its movements, and IMU sensors that measure the vehicle's acceleration and orientation.

After the data is collected from the sensors, it is passed to the vehicle's onboard computers, which generate a model of the environment. The model is used to decide how to navigate and what obstacles to avoid.

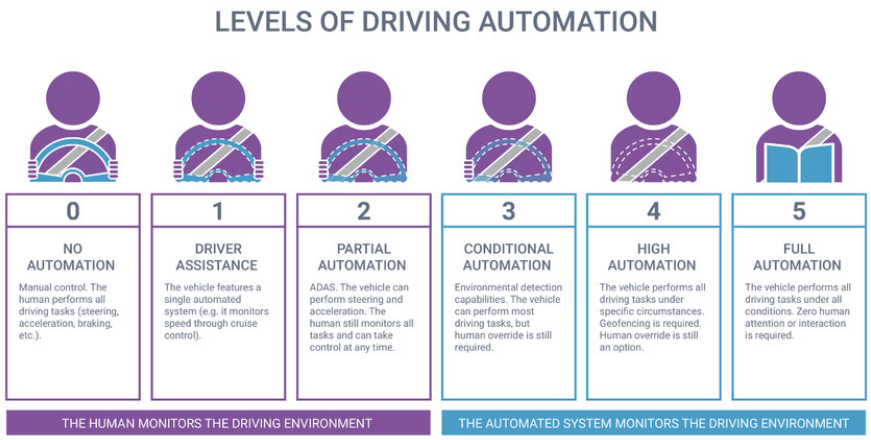

Figure 2.1: Levels of automated driving

In order to change the movement of the vehicle, computers use algorithms that analyse the data and determine the appropriate actions. For example, if sensors detect a stopped car in front of the vehicle, the algorithms may determine that the vehicle's speed needs to be reduced in order to avoid a collision. The computers then send corresponding signals to semiconductor-based actuators to engage the necessary actions, such as reducing speed or changing course.

We will not dive into more detail about how all the parts jointly work together and execute tasks because that would require writing a whole book. However, to better understand the article, it is important to have heard about the main aspects surrounding AVs because the same rules also apply to the simulated world, which is trying to be as reminiscent of the real world as possible. More about that will be covered in Chapter 4 when we talk about the virtual driving simulators and CARLA, in particular.

2.3 The six levels of autonomy

Moving on, it is relevant to mention that the classification of vehicles in terms of their autonomy is non-binary, meaning that there are not only fully autonomous vehicles and those entirely operated by humans, but various other levels of autonomy exist in between. This chapter will explain how autonomy levels are categorized and what level of autonomy is present in vehicles nowadays.

In 2014, The Society of Automotive Engineers (SAE) has proposed a standardization for categorizing the level of autonomy in vehicles, ranging from Level 0 (no automation) to level 5 (full automation) [9]. As shown in Figure 2.1, these levels are determined by the extent to which the human and the automated system are involved in monitoring the environment and

controlling the vehicle.

Currently, the majority of cars on the market fall into the lower levels of autonomy, with the driver still responsible for most aspects of driving. Some common automation features that new cars may offer to the owners include lane assisting, cruise control, automatic emergency braking, blind spot monitoring, automatic parking, autosteering and automatic lane changing.

Vehicles falling into the upper three categories are not yet available on the market, and it might be another several years until they become widely adopted. The main reasons for that are technological and regulatory challenges and public scepticism about AV safety.

Luckily, this article aims to address these questions by proposing a system for testing highly automated vehicles in a simulated environment, thus assessing whether the software agents controlling the vehicles are working as intended. More about that will be discussed in Chapter 6, when the proposed safety performance evaluation system is introduced.

2.4 Safety of autonomous vehicles

With regard to safety, there is a general belief that autonomous vehicles have the potential to improve transportation safety by addressing one of the leading causes of accidents and fatalities on the roads: human error. Drivers can be distracted by activities such as texting or talking on the phone or may become impaired or fatigued, leading to potentially dangerous situations. AVs, on the other hand, are equipped with sensors that constantly monitor their surroundings and detect potential hazards, allowing them to respond more quickly than a human driver could. This can be particularly important in emergencies, where reacting as quickly as possible might be vital.

Moreover, AVs can potentially eliminate road congestion and substantially improve traffic flow. Suppose there were no human-operated cars on the roads, and every participant would be operated by an AI with the ability to communicate among themselves and road infrastructure using radio waves. In that case, vehicles could better understand other vehicles' intentions and how the surrounding world will change. For instance, which driver might be turning soon, when the traffic light will change to green or where the road hazards are that should be avoided. If such communication were possible, travelling time would be reduced because all the vehicles would drive more efficiently.

In addition to improving overall road safety, AVs could help increase the mobility of people who cannot drive a car, for example, people with disabilities or older people.

However, it is essential to note that AVs also present potential risks, one being hardware

malfunction and the other being cyber-attacks. It is inevitable that mechanical components sometimes come with factory defects or develop a fault with time. They could contribute to inaccurate information that may lead to wrong and unsafe decision-making. For this reason, companies developing AVs must test all the critical components comprehensively to ensure their correctness. In addition, the future owners of the AVs must be provided with adequate systems that constantly check the state of the sensors and provide alerts that will announce the need for either technical service or recalibration if they are not in the optimal state.

Cyber-attacks could come into play anywhere where an AV communicates with online systems. It is inevitable that AVs, like other modern systems, will rely on external systems and services in one way or another. It could be by downloading software updates (which talking about the constantly evolving AV nature would be frequent), getting the latest traffic information for route planning, getting the weather forecast for traffic predictions or simply sending data to the manufacturer for analysis. All the information being sent can possibly be tampered with by malicious third parties and result in incorrect data, system faults and reduced safety.

Although it is crucial for AV manufacturers and regulators to prioritize safety in the design and operation of these vehicles, this research paper focuses solely on vehicle performance rather than correctness of their components and assumes that the systems are cryptographically safe and cannot be exploited by anyone. As mentioned in the book written by RAND Corporation in 2018 called "Measuring automated vehicle safety: Forging a framework": Comparing these machines at vehicle (or higher) levels fits people's intuition about safety [10].

Chapter 3

Human drivers and the road

This chapter discusses how people behave on the roads and why collecting real-life traffic data is crucial for AI model training. Section 3.1 analyses traffic incident data collected in the EU and U.S. in 2020, while the following Section 3.2 links data to human behaviour and analyses why some types of accidents happen more frequently than others. In Section 3.3, we finish the chapter by discussing how the traffic data and human psychology analysis can help us understand how the simulated scenarios should be designed and what situations they should reenact. Let us begin by looking at the statistics.

3.1 Road incident statistics

For the purpose of this research, it is important to understand the situations that lead to human driver failures on the road. These failures can occur due to a variety of factors, including distractions, participant impairment, road and weather conditions, place.

According to the European Commission's published data for road safety in the EU in 2020 [11], 52% of fatal road incidents happened in rural areas, while 40% occurred in urban areas and only 8% on motorways. The analysis also indicates that 70% of all fatalities in urban areas were made up of vulnerable road users such as cyclists, moped and motorcycle drivers and pedestrians, that account for the vast 37% of victims in city areas.

In the traffic safety facts annual report tables provided by The National Highway Traffic Safety Administration of U.S. Department of Transportation [12], it is clear that the majority of fatal road crashes in the U.S. in 2020 happened from 3PM until 11:59PM on working days and from 6PM and 3AM on weekends. The NHTSA's data set also indicates that the majority

(83%) of fatal crashes happened during normal weather conditions while rainy weather caused only 7% of fatal incidents. Out of the 83% of crashes that happened during normal weather conditions, 46% of accidents happened during daylight while 28% happened at night and 21% happened at night but in lit areas.

This shows an interesting fact that although there is a general perception that driving in snowy or foggy conditions is more prone to becoming a victim of an accident, the statistics show that it is entirely contrary. Most crashes happen when the conditions are best for all the human drivers (well-lighted areas in the middle of the day when everything is well observable).

It is also worth mentioning that 30% of all road fatalities occurred in alcohol-impaired road accidents. Before linking the data covered in this section to the goal of this project, it is important to look at why we have such a large number of incidents on the road.

3.2 Human nature and behaviour

A standard individual essentially has four main sensors when it comes to driving (two eyes and two ears) that provide information about the surrounding world and a brain which is responsible for making decisions.

Although a people have everything necessary to operate the vehicles on roads and following the established rules at all times they could do that successfully and safely, there are many factors that could make the roads unsafe regardless.

One reason is that emotions are a key aspect of human life and can heavily impact how a person makes decisions. When a person is feeling strong emotions such as anger, frustration, or stress, it can affect their ability to concentrate and make good decisions while driving. For example, a person who is feeling angry may be more likely to take risks or engage in aggressive driving behaviors, such as tailgating, speeding, or cutting off other drivers. This can increase the risk of collisions and accidents on the road.

Similarly, a person who is feeling stressed or distracted may have difficulty paying attention to the surroundings or react to situations. This is due to the humans' limited cognitive energy reserve. Neurological science has demonstrated that the human brain is incapable of focusing on two things at once as stated in the article "The Impossibility of Focusing on Two Things at Once" written by M. Hernandez [13]. This means that even listening to music or having a passenger telling a cunning story about his recent promotion reduces the driver's alertness and clouds the perception of their surroundings.

There are other factors that can have significant impact on how a person is able to operate

the vehicle one being fatigue or impairment, such as being influenced by alcohol or drugs.

Artificial intelligence has the potential to improve driving safety by addressing some of these factors. On one hand, AVs are essentially exceptionally sophisticated software systems with dozens and possibly hundreds of sensors providing reasonably accurate and reliable information about the surroundings and multiple cores for computation able to make trillions of calculations every second. Moreover, systems controlling the vehicle cannot be distracted by either delicious food advertisements on the street or a song on the radio. Even in the case of the passenger being able to communicate with the vehicle vocally, it could sacrifice some of the computational power of one of its multiple cores for that fraction of a second needed to calculate the most appropriate answer. The AV systems cannot become fatigued or get influenced by substances or emotions. If programmed well their behavior and decision making cannot be altered by external individuals or systems.

However, it is important to note that AI systems are not immune to hardware malfunctioning, software errors or cyber attacks by very sophisticated attackers. AV safety was discussed in Section 2.4.

3.3 Statistics relevance to the research

Now that we discussed the trends in road accident statistics and looked a bit into the differences between human and AI vehicle operators, it is time to understand why all this information is crucial for developing adequate systems able to generate drivable scenarios and analyse vehicle safety in situations that are the most prone to accidents.

Comprehensive road statistics analysis can help ensure that the simulated scenarios accurately reflect the real-world conditions that AVs will encounter. Scenario generation has to take into account when the most accidents occurred and what factors, such as weather, lighting, road surfaces, or places, influenced them the most. Extracting accurate data from the statistics allows testing the AVs in realistically challenging situations and shows that the research is as representative of the real world as possible.

One way road data could be used is to train AI models to predict what situations could be challenging to drive and prone to accidents. We will discuss more about scenario generation when we introduce a way of automatically generating driving scenarios using a machine learning model in Chapter 5. The proposed ML algorithm is able to predict the parameters of new scenarios based on the scenario examples it has seen previously and then put the numbers to life with the help of the CARLA modules.

From the data, analysed in Section 3.1, it is clear that most simulated scenarios should occur in rural and urban areas reflecting situations that impact road safety the most. However, highway roads also account for a substantial proportion of road incidents; thus, some scenarios should also happen on highways.

Regarding the weather statistics from Section 3.1, it is clear that to mirror the conditions prone to vehicle collisions; scenarios need to be designed in various weather conditions, with most of them based on normal weather circumstances.

The datasheet also indicates that many road incidents involved pedestrians and two-wheel vehicle drivers, e.g., cyclists and motorcycle drivers.

To sum up, it is clear that in order to develop a reliable system that can generate driving scenarios that accurately reflect real-world conditions, an enormous amount of diverse traffic data is needed that could be restructured so it can be used to train machine learning algorithms. More details about this process and the chosen way to implement the framework component responsible for scenario generation are discussed in Chapter 5.

Chapter 4

The simulated world

In this chapter, we will be focusing on the simulator used in this project - CARLA. The reasons why this simulation framework was chosen, along with its advantages and role in the research, are explained in detail in Section 4.1. Additionally, Section 4.2 covers the limitations of the simulator and explores the potential inaccuracies that may arise during the research due to action taking place in the simulated environment.

4.1 CARLA simulator

As mentioned in the introduction, this work aims to design a system for evaluating vehicle safety performance in realistic simulated driving situations and later evaluate the system's accuracy by testing it with the participants. For this, a sophisticated simulator was needed that could provide a reasonably accurate representation of the human world, from road infrastructure and people to physics and weather impacting them. In addition, it had to have a way of manipulating the world that is being simulated, launching a first-person driving mode, and connecting AV implementations to the simulation. Compared to other competing frameworks, CARLA (Car Learning to Act) [14] has proven to be an ideal choice for this project.

One of the key benefits of CARLA is its scalable client-server architecture, which allows for efficient manipulation and customization of the simulated environment. The server runs the simulation, renders objects and sensors, and maintains physics. On the other hand, the client can manipulate the simulated world by changing weather conditions, spawning actors and controlling them, and getting information about the simulated world and what is happening there at any given time step. It does so using the CARLA Python API, which provides a

comprehensive set of use cases and commands well-documented by CARLA developers. Another benefit of the client-server architecture is that it is relatively easy to connect different AV agent implementations to the simulator and have a defined behaviour of the server and all its components working separately on their own.

The third benefit is the Agent implementations provided by the CARLA development team. They allow self-driving agents to be put on the map and instructed on where to drive. Although far more sophisticated AI implementations can be bridged with CARLA, like Autoware.ai, which is briefly mentioned in section Section 9.1, the CARLA BasicAgent is used to compete against human drivers in the software testing in Chapter 9. The main reason for that is that this project aims to introduce a way of automatically testing the safety of AI implementations operating vehicles rather than connecting different implementations with the simulator. The proposed ideas work with any implementations that can run in the CARLA environment. The safest and most time-efficient way to ensure the framework performs its tasks was to use built-in and easy-to-customise AV implementations

4.2 The limitations of the simulator

Although CARLA is a powerful tool for testing how AI agents and human drivers operate the vehicles, it is important to understand that it has limitations compared to the real world.

One of them is that it cannot account for the real world's unpredictability. Many factors could impact how an AV reacts, and it is challenging to hard-code them all when designing the scenarios that could occur on the road, similar to how it is difficult to tell the AV how to react in all those situations. There are infinite possible scenarios; in some, the autonomous agent could react well; in others, not so much. This research focuses on designing scenarios with realistic traffic conditions and unpredictable events that require quick reaction and decision-making. The scenarios will test the adaptability and robustness of AVs in unexpected circumstances while also monitoring the overall vehicle's safety.

Another simulator's limitation is the fixed number of road conditions programmed into it. Although CARLA offers a wide variety of customisable stuff, it is still restrained by parameters set by the developers. To exemplify this, snowy winter conditions with iced roads cannot be simulated because there is no way to facilitate this. In addition, there is no efficient way to check if the vehicle has stopped at a stop sign, which is the opposite when talking about checking if the vehicle ran a red light.

Moreover, it is important to note that CARLA is only as accurate and good as the algorithms

that power it. If the internal calculations are inaccurate, that could affect the simulations and lead to incorrect conclusions about the performance and safety of the drivers.

Chapter 5

Designing the driving scenarios

This chapter explores various approaches for generating scenarios, with a specific focus on the method selected by this article. In Section 5.1, we start by giving an overview of some of the possible ways that scenarios could be generated, while Section 5.2 and Section 5.3 examine the chosen scenario and path generation methods and discuss their benefits and drawbacks in terms of accurately predicting new scenarios. Finally, in Section 5.4, we conclude the chapter by discussing potential improvements for the current method of scenario generation.

5.1 How the driving scenarios could be designed

Let us begin by discussing the need to design driving scenarios for autonomous vehicles. According to Ruxin Li and Runping Zhain's article "Estimation and Analysis of Minimum Traveling Distance in Self-driving Vehicle to Prove Their Safety on Road Test" [15], self-driving vehicles must drive at least 7.2 billion kilometres to prove their accuracy and confidence to an acceptable level. However, this distance is enormous, equivalent to circling the Earth approximately 180,000 times. If we imagine a car driving on the road 24/7 at an average speed of 60km/h, it would take that car 13,699 years to complete this distance. Therefore, simply putting a vehicle on the road and testing it to the fullest extent would be inadequate, given the time required. This is where the simulated environment comes into play.

A well-constructed framework that can accurately test how the software system controlling the autonomous vehicle behaves can be a huge time-saver allowing fully autonomous cars to see the daylight more quickly. However, as we will soon see, this task is very challenging, and some traditional ways of solving it might not work.

One of the traditional ways a scenario generation could be imagined is by designing the scenarios manually, putting objects into specific locations, and trying to recreate complex driving situations in the simulated world. However, time is the enemy, and doing everything by hand would not be much better than letting the autonomous vehicle drive for decades or even more. In addition, it is impossible to think of all the possible edge cases the vehicle might encounter because life is stochastic and non-fully observable. An infinite amount of possible things can happen at any given time, and hard-coding them in the software would bring poor yield. What is needed is a way of creating complex and realistic driving situations automatically, and assessing how good the generated scenarios are.

This is a perfect moment to start talking about AI in a bit more depth and discuss how it can help us in the case of scenario generation. Specifically, the rest of this section focuses on machine learning, a concept that has long been laid on the shelf and was made possible in practice only in the last few decades because of the immensely increased computational power resources and advanced hardware capabilities. Although a young member of the technology field, machine learning has already demonstrated its significance in various areas, such as medicine for predicting illnesses, economics and finance for predicting market changes, etc.

One of the fascinating aspects of machine learning applications is that they are not ordinary programs designed to perform specific tasks. One could argue that they are still meant to perform particular tasks, mainly predict values, but how they do it is markedly different. These algorithms are designed to act based on the data they have been trained on and make predictions about new data using their knowledge and gathered domain information.

To illustrate how the machine learning algorithms work and understand why we are talking about them, let's look at a specific example use case in the finance industry - fraud detection. Banks and credit card companies use machine learning algorithms to analyse customer data and transactions to identify suspicious activity. They train the algorithms on various instances of fraud, allowing algorithms to investigate patterns and detect new fraudulent behaviour, protecting customers from financial losses due to fraud and helping financial institutions avoid losses from fraudulent transactions. Companies such as Mastercard [16], PayPal [17] and others use sophisticated machine learning algorithms for fraud detection in their payment systems.

Now returning from the finance field back to autonomous driving, machine learning algorithms can also be used to predict new scenarios based on the examples they have been trained on. One thing worth mentioning before we dive into the possible ways of teaching ML algorithms to generate scenarios is that they require enormous amounts of training data to perform

well and avoid underfitting. Underfitting occurs when the ML model is too simple and fails to capture the complexity of the data it is trying to model. Another problem that may arise is overfitting. This is when the algorithm is too specific and cannot generalize the information well, meaning it has learned to classify the training data but hasn't learned the underlying patterns and thus fails to work with new inputs. That said, one cannot think that giving the algorithm a few descriptions of driving situations from the real world would make the algorithm perform well. In reality, an algorithm would have to be trained on millions and even billions of precisely described scenarios to achieve its goal - creating new scenarios.

What is meant by precisely described scenarios is not that an algorithm needs essays describing the actors and actions that took place in the real world but well-structured data acceptable by the chosen style of the algorithm. Usually, the data is translated into meaningful numerical values that the ML algorithm can synthesize. Depending on the choice of the algorithm, the training data is processed differently, usually altering the internal domain representation of the ML model. This section will not dive deeply into the machine learning world. Still, it will give an overview of what concepts could be combined with machine learning methodologies to assist in scenario generation.

One instance of a machine learning algorithm that could be used for scenario generation is multi-output regression. The way it works is it is trained on instances of scenario descriptions containing the significant metrics in numerical form. Each of the instances is labelled by one or more labels. After a considerable time of training, the algorithm is able to predict new metrics given the labels as input. This approach is implemented as part of the solution for the automatic scenario generation that this article presents and is discussed in more detail in Section 5.2.

Another way of training the model could be by using video recordings or fragments of them with precisely labelled data and objects and described dangers. A deep neural network could use this data to build an inner model of what dangerous situations look like, how the participants behave, etc. With the help of other machine learning and data analysis models, this model could be used to design new unseen situations or at least translate them into ones usable by virtual simulators. This approach is better explained in the article "Automated generation of virtual driving scenarios from test data" by Robin van der Made [18].

5.2 The selected approach for scenario generation

Now that we understand the possible approaches that could be used in scenario generation, let us dive into detail about what path this article chose and how the algorithm proposed by this project is functioning. Having already introduced the simulator used in this project in Chapter 4, it is time to investigate how CARLA's functionalities could be used to our advantage when generating the scenarios.

As a sophisticated project, CARLA provides a wide range of configurations allowing users to customise the simulated world they see fit. One of the significant components for that is the TrafficManager. Its main objective could be described in one sentence - controlling all the actors in the simulation and executing their actions according to the specified behaviour.

Before we begin analysing the scenario generation method described in this project, it is worth mentioning that for this research, there was no access to any large database whose data could be used for machine learning. Thus, the algorithm being presented is simplified and uses pseudo data that might not represent real-world conditions. In addition, the algorithm is prone to underfitting because the data used to train it was generated using a primitive approach (creating a few unique training instances and then creating copies of them with all the values increased or decreased by some percentage).

After a lengthy analysis of the simulation world and the TrafficManager with all its functionalities, it was decided to describe scenarios using attributes presented in tables in Appendix A; Table A.1, containing all the attributes that describe the weather that can be customised by the client using the CARLA simulator and Table A.2, containing all the necessary attributes describing the traffic conditions that can be created and customised using CARLA client and the TrafficManager. A scenario could be described in full by combining the attributes of both tables, assigning a value to each of them, and putting them into a JSON object, the chosen standard for scenario specification in this project. A scenario example can be seen in Figure A.1 in Appendix A.

Before moving on, it is worth noting that the algorithm training data and all the other files specifying the necessary modifications, such as what maximum value each attribute can possess, are stored in the scenario_generation_data/learning_data in the project files. The generated scenarios are stored in the directory scenario_generation_data/generated_scenarios. The machine learning algorithm can be found in the software/generating_scenarios directory. Finally, how to run the scenario generation is explained in the Section B.1 in Appendix B, where the user guide is presented.

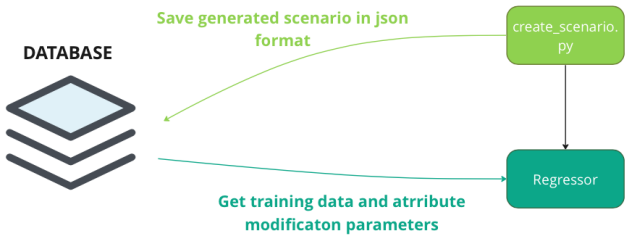

Figure 5.1: Scenario generation overview

Now that we described what the scenarios look like, gave references to where the necessary files can be found, it is time to look at how the algorithm works. The algorithm implementation can be found in the **regressor.py** file. At its core, the Regressor uses machine learning algorithms provided in the *Scikit-learn* library [19]. It is done so because *Scikit-learn* is a popular open-source library delivering a wide variety of well-implemented and accurate machine-learning tools used by many companies and developers. This ensures that, at the very core, our algorithm is behaving correctly. The algorithm also uses other library methods to split the training data into training and testing sets, work with .csv files and data frames.

Once called, the algorithm first trains itself on the training data provided in the scenarios.csv file that can be found in learning_data directory. After the training phase is complete, the algorithm tries to predict what attributes should a new scenario possess given the input labels: difficulty, wetness and sun altitude angle. The algorithm then checks if the predicted values are within their valid ranges (the ranges for each attribute are described in minimums.csv and maximums.csv files in learning_data directory), rounds them using values from rounding.csv and saves the newly generated scenario in a JSON file to the directory generated_scenarios. The high-level scenario generation infrastructure is shown in Figure 5.1. The purpose of the generated scenario file is to describe the traffic and weather conditions. It includes information used by other components that populate the world, customise the weather and set up the traffic. Both PathBuilder and Scenario classes use the generated scenario, and the ways they use it are described in Section 5.3.

One thing that was not mentioned yet is why the chosen labels are difficulty, wetness and sun altitude angle. The first one, difficulty, describes how complex the scenario is overall. For instance, the harsher the weather conditions and the more disobedient drivers there are, the more challenging the scenario is. The second label, wetness, is used to differentiate between dry

and clear weather and rainy or stormy weather. The smaller the wetness value, the smaller the attribute values related to precipitation, fog, rain deposits, etc. The third and last label, the sun altitude angle, is used to differentiate between the different times of the day. To sum up, these three labels allow the algorithm to generate the scenario of desirable difficulty, wetness level and also the preferred time of day.

5.3 Building paths and populating the simulator

The previous section mentioned that the scenarios generated by the scenario generation algorithm are used by the PathBuilder and Scenario classes. Let us begin by discussing the path generation method this article proposes.

To test how the vehicle is performing in the scenarios, it is not enough to specify the weather parameters or how many other participants will be involved. We need to give the vehicle an objective that it would try to achieve. In this case, the aim of the car is to drive some route and reach the destination safely.

Luckily, as this project aims to create a framework for automatic vehicle safety testing, manual path generation for every run is not a valid solution. That is why the PathBuilder class was developed (stored in software/carla_scripts/scenarios/path_builder.py). This class is responsible for creating a path given the scenario JSON file.

The algorithm works as follows; first, it looks at all possible maps in the data/maps/opendrive_format directory and loops through all the maps looking for one that satisfies the following criteria from the scenario file:

$$\begin{cases} number\_of\_junctions \leq total\_number\_of\_unique\_junctions\_in\_map \\ distance\_in\_metres \leq total\_map\_road\_length \end{cases} \tag{5.1}$$

If the map that is being analysed satisfies the criteria, it is chosen, and the path-building algorithm proceeds to try and create a path on that map; if the map does not meet the criteria or a path cannot be created in it, the algorithm takes another map and proceeds with the same actions until a path is generated or the algorithm is aborted meaning it is impossible to create a path in any of the maps given the scenario requirements.

It is also worth mentioning how the path itself is created when the correct map is found. If a valid map is found, the algorithm proceeds as follows: first, it selects a road element randomly, then enters a while loop and keeps looking for road successors, adding them to the

list of checkpoints until the scenario criteria are satisfied. The successor elements can be of two types: a road element or a junction. If the currently analysed road element points to another road element, the while loop proceeds with the successor road element in the next iteration. However, if the current road element points to a junction, then a random successor of the junction is selected for the next iteration. Note that the successor of the junction is always a road element. If at any point the algorithm encounters a loop (two road elements pointing to each other) or a dead-end, it starts the process all over again. The algorithm currently has a limited amount of tries that it can have on a single map $x = 10,000$. If in x tries the creation of the path is unsuccessful, the map file is either corrupt or a path there cannot be generated. This is made so that the algorithm never enters eternal while loops. The high level overview of the algorithm can be seen in the pseudocode fragment below:

Algorithm 1 High-abstraction-level path building algorithm

```
procedure PathBuilding(scenario)
    maps ← read OpenDrive maps from the database
    for map in maps do
        if map is valid then
            r ← randomly pick a road element from the map
            road_elements ← []
            total_length ← 0
            number_of_junctions ← 0
            while number_of_junctions < required_number_of_junctions
                        and total_length < required_total_length
                        and not road_elements contains dead-ends do
                total_length ← total_length + r.length
                road_elements.insert(r)
                successor ← r.successor
                if successor is junction then
                    number_of_junctions ← number_of_junctions + 1
                    r ← random successor exit that is not current r
                if successor is road element then
                    r ← successor
            path ← build path from road_elements
            if path was found then
                return path                          ▷ Else, try another map
    return None          ▷ If no path was found over all the maps, return None
```

If the maps in the opendrive_format directory are well formatted, the final generated path could look like the one presented in Figure 5.2 where the blue colour marks the beginning and the red one the end. The colourful lines, gradually moving from blue to red, indicate the lanes that a vehicle can drive on. To draw the path on the map, the draw_lanes_terminal.py script has to be run with the path's name as the argument -p. The script is located in software/-carla_scripts/helper_scripts directory.

Figure 5.2: Well-formatted path

However, the algorithm not always behaves in our best interest. In Section 4.2, when discussing the limitations of the simulator, we mentioned that CARLA is only as accurate and good working as the algorithms that power it. One of its issues revealed itself when designing this path generation algorithm. To begin with, CARLA, for its map specifications, uses OpenDRIVE format, which is a standard way of specifying map layouts and road networks[20]. The algorithm proposed by this article also uses OpenDRIVE files to analyse the road networks and thus generate paths.

As mentioned before, when a map is determined, the path generation algorithm randomly takes the first road element. If that initial pick is indeed a valid road segment, the algorithm can generate a path like the one pictured in the Figure 5.2.

Unfortunately, the bigger part of OpenDRIVE maps provided by CARLA are poorly structured in some places (in version 0.9.13), meaning that the files specify that a road element is present when it is not. An example of a generated incorrect path can be seen in figure Figure 5.3, where the points selected by the algorithm are outside the map where there is no road. After analysing the OpenDRIVE maps provided by CARLA, it was discovered that the specifications are indeed misleading. In addition, some maps possess other issues, like one road element pointing to itself or loops where two distinct road segments point to each other. The problems are illustrated in Figure C.1 and Figure C.2 in Appendix C.

An example of a generated path can be found in Figure A.2 in Appendix A. The points from the list **path_checkpoints** are used by the GlobalRoutePlanner to build a path that a vehicle has to drive. The **town** attribute indicates the map, and the **start_location** tells where

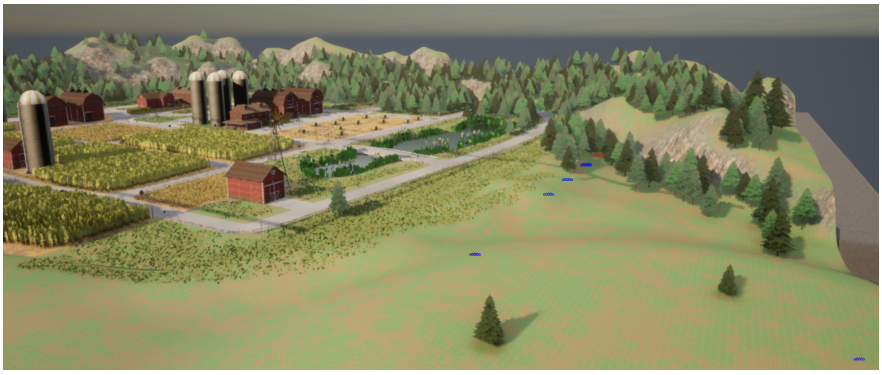

Figure 5.3: Poorly-formatted path

to spawn the car.

Now that we have talked about the PathBuilder and the issues it was facing, it is time to look at how the Scenario class uses both the generated path and the generated scenario to fulfill the simulation requirements. The Scenario class can be seen as a module responsible for spawning all the non-hero actors and specifying their behaviour according to the generated scenario attributes.

It first tries to spawn the necessary amounts of pedestrians, vehicles and two-wheel vehicles. If there are not enough spawn points for all the cars, the maximum number of them is spawned. After a successful spawn, the module customises the behaviour of all actors according to the values provided in the scenario file. For instance, if the $proportion\_of\_light\_ignoring\_vehicles = 0.2$ and the $light\_ignoring\_percent = 50$, then the algorithm randomly selects a fifth of the vehicles and sets their behaviour as such: when a red traffic light is approached, there is a 50% chance that the car will ignore the traffic light. It does so, for all the behaviour-changing scenario attributes. Once the actors are spawned and their behaviour is set, the Scenario module assigns all the actors to TrafficManager to take care of their movements in the simulation. How the simulation takes place and when the TrafficManager is called will be discussed in more detail in Chapter 8, when we talk about what the simulation session looks like and how all the components work together.

5.4 Possible future improvements

Having introduced the scenario generation methodology, it is now time to finish this chapter by discussing the possible improvements that could be implemented in the future. First, to ensure that invalid paths are never generated, well-formatted OpenDRIVE map specifications

should be used. From all the current map specifications, after some testing, only the Town01 showed the potential of being well-formed. Using other maps sometimes produced wrong paths. Another part that should be improved is the path-building algorithm itself. Right now, it creates the paths by randomly trying to create one. It would be better if it did this in some structured way, looking for interesting patterns that could be utilised. However, the OpenDRIVE files currently contain only information about the road elements and junctions, providing very little extra information. In other words, more sophisticated road network descriptions should be used, providing information about roundabouts, highways, u-turns, road signs and other things. Finally, the scenario generation algorithm could be improved in several ways. Firstly, well-described and realistic scenarios could be used for training. Currently, the algorithm uses naively generated training data, as discussed in Section 5.2. Furthermore, the machine-learning algorithm could be improved or even changed to use deep neural networks and predict scenarios from more sophisticated training data sources like video material.

Chapter 6

Assessing vehicle safety

In this chapter, the article proposes a way of assessing vehicle safety on the virtual roads. We begin by giving an overview of the suggested method and explaining why it was chosen in Section 6.1. In Section 6.2, we continue our discussion by explaining what safety metrics are monitored in the simulations and how the collected data is processed. The chapter is finished in Section 6.3 after providing a mathematical way of evaluating the vehicle's performance.

6.1 Monitoring and recording data

After implementing the scenario generation and path-building modules, the next step was to design a sub-system to monitor vehicle actions and record the behaviour, allowing for the data analysis later. As mentioned in Section 2.4, when we talked about the safety of autonomous vehicles, this article deals with evaluating the safety at the vehicle level, meaning how the car acts on the road. For the purpose of this research, it was decided to track the vehicle's actions during the simulation process and then save obtained information to XML files. The XML (Extensible Markup Language) file standard was chosen for its convenient structuring of large files, allowing for efficient data parsing in the tree structures.

The central class of the safety monitoring system is the **Manager** class, which is responsible for checking the vehicle's state at each time step and informing the other monitoring classes if data is needed to be recorded. The manager class and the monitors are built based on the observer design pattern when one object broadcasts updates to multiple other objects. Having given the overview of the system, let's continue by discussing each component participating in the safety assessment.

Let us begin by talking about the monitors of the system, the first one being the **RouteMonitor**. This class tracks the vehicle's progress along its route and ensures it stays within the allowed lanes. The route needed to drive is obtained from the GlobalRoutePlanner (provided by CARLA developers), which returns a list of waypoints given a list of coordinates. The coordinates come from path files discussed in Section 5.2. Using these waypoints, the route that is monitored by the RouteMonitor is created by recursively finding all the lanes that the vehicle could use. At each time step, the RouteMonitor checks if the car has not left the permitted lanes and tracks the number of route points that have been covered. This metric is later used to calculate route completion. At the end of the simulation, the RouteMonitor saves the route completion, whether the finish was reached, and the coordinates of all the route points and whether they have been reached to the data/recordings/participant_[A]/scenario[1]/route_data.xml file.

Another monitoring class implemented was the **CollisionMonitor**, which records data about collisions that the vehicle got involved in. Whenever a crash occurs and is detected by the CollisionSensor present in the Manager class, the CollisionMonitor records information about it, including the collision type (collision with a human, another vehicle, road objects, or buildings), the penalty points given to the driver, the time step, and the coordinates where it happened in the recording/recordings/participant_[A]/scenario[1]/collision_data.xml file.

In addition to the route and collision monitors, a **SpeedingMonitor** was developed to record data about instances when the vehicle was speeding. At each time step, the Manager class checks if the driver is exceeding the speed limit, and the SpeedingMonitor records data about it, including the allowed speed at that location, the vehicle's speed, the penalty points received for speeding, the time step, and the coordinates where it happened. This data is saved to the data/recordings/participant_[A]/scenario[1]/speeding_data.xml file.

A **LaneMonitor** was also implemented to record instances when the vehicle crossed solid and double solid road markings. In addition, instances of crossing broken lines without showing correct turn indicators do not go unnoticed. The LaneMonitor saves the lane type, the penalty points given, the time step, and the coordinates to the data/recordings/participant_[A]/scenario [1]/lane_marking_violation_data.xml file.

A **TrafficMonitor** was created to track stop sign, stop marking and traffic light violations. The Manager class observes the vehicle at each time step and records traffic violations in the data/recordings/participant_[A]/scenario[1]/road_traffic_violation_data.xml file. The saved data includes the violation type, the penalty points given, the time step, and the coordinates where it happened.

Finally, a **VehicleLightMonitor** was implemented to register data about the state of the vehicle's lights and any improper use. For example, failing to use turn indicators or not having headlights in the dark would result in light violations recorded in the data/recordings/participant_[A]/scenario[1]/vehicle_light_misuse_data.xml file. The saved data included the violation type, the penalty points given, the time step when it occurred, and the location where it happened.

In addition to the monitor classes, a **Recorder** class is responsible for creating a log file following the CARLA standards that holds all the data from the simulation. The file can later be used to replay the simulation using the **Replayer** class via the replay.sh script that can be found in software/carla_scripts directory.

When the simulation ends, the manager instructs all its monitors to write the data stored in their buffers to the corresponding files. For example, participant_A's simulation results of scenario one would be stored at data/recordings/participant_A/scenario1/ in different .xml files, e.g. collision_data.xml, speeding_data.xml. The manager also gives a command to the Recorder to stop recording and save the .log file of the simulation that can be later used to replay the simulation. This gives the ability to analyse the simulations multiple times and thus catch all the minor details that can be important or monitor some data manually if needed.

To better understand how the data is stored or how the monitoring works, the reader is encouraged to look at the implemented classes in the software/carla_scripts/recording directory in the source code. Additionally, the overview of the observer design pattern can be seen in Figure 8.1 in Chapter 8. To see what the recorded data files look like, please refer to Figure A.3 in Appendix A showing an example of collision_data.xml file or explore the data/recordings directory.

6.2 Metrics to record

In this section, we look at what quantitative metrics from the recordings we can use to evaluate a vehicle's performance in a given scenario. After long consideration, two groups of metrics were developed: time and accuracy-related metrics allowing to give positive points to the user and safety-related metrics, which give penalty points to the user for violating some safety requirements. The two groups are discussed in the following two subsections.

6.2.1 Time and accuracy-related metrics

The first one considers how neatly and optimally the vehicle is being operated and includes these metrics:

Timeliness – Each scenario has a specific optimal time to complete without violating road rules or breaking speed limits. The timeliness is calculated by dividing the optimum time value by the time the driver took to complete the route. The quicker the driver is, the larger the timeliness value will be. The value proportionally increases or decreases the driver's score depending on his swiftness.

Proportion of route completed – The route is a line of dots on the map, the first being the starting position of the vehicle and the last being the finish position of the path that needs to be driven. Each dot has coordinates associated with it and is in the middle of the lane. This metric shows what percentage of waypoints were covered. For the waypoint to be considered covered, the vehicle's centre point must pass it within a certain distance. This ensures that if a car is driving not within the boundaries of the lane that it can drive on (one wheel is on the line, for instance), the waypoint will not be covered, and thus the final score will be lower. It is worth noting that if the road has multiple parallel lanes that can be driven by the vehicle, the algorithm checks if the vehicle is within the boundaries of any of those lanes. If it is not within one of those lanes but is on some of the lane markings between two lanes and is showing a turn indicator (the vehicle is changing lanes), the waypoints within a certain distance from the vehicle's centre position will be covered. This is an exception to the rule that the vehicle must stay within the lane's bounds. However, because the action takes place in a simulated world and it is difficult to feel the vehicle and be convinced that it stays within the bounds of the lane, the gamma discount factor will be applied to reduce this metric's weight.

6.2.2 Safety related metrics

The second group is evaluating how safely the vehicle is being driven and considers these metrics:

Speeding instances – This metric indicates how many penalty points were assigned to the user due to exceeding speed limits. This is important because the vehicle's speed directly impacts the safety of the people sitting inside and the safety of other drivers, passengers, pedestrians, and animals. In an ideal situation, both AV and the human-drivers need to obey speed limits because speeding delays the reaction time, narrows the vision angle and increases the braking distance resulting in diminished safety.

Collision instances – This metric shows how many negative points the vehicle earned throughout the simulation because of collisions with any objects (pedestrians, road objects, other vehicles, buildings). Moreover, collisions while speeding will result in even more negative points added to the final score because hitting something at high speed, will typically result in more damage done and any violation done while speeding will almost always result in the driver being guilty and responsible for causing the incident.

Traffic light violations – This metric shows how many negative points the vehicle earned throughout the simulation because of running a red traffic light. In the case of a vehicle violating these rules while speeding, even more negative points will be added to the score because the risk of causing a severe road incident is increased, resulting in diminished overall road safety.

Lane marking violations – This metric shows how many negative points the vehicle earned throughout the simulation for crossing road lines other than broken ones (- - - - -). Lane markings have to be obeyed as they are there to prevent people from overtaking or turning where it is unsafe to do. Again, lane violations while speeding will give more negative points than not speeding.

Vehicle light violations – Driving at night without vehicle headlights on or in a fog without headlights and fog lights on will result in negative points given to the vehicle. In addition, changing lanes without respective turn indicators blinking will result in a worse outcome from the route metric. Showing turn indicators and having headlights on in the dark is extremely important for the vehicle's passengers and other traffic participants. The fact that all the road actors are visible and their intentions are clear directly impacts how safe the road is.

Another important metric for evaluating safety is the number of penalty points the participant receives for not stopping at stop signs and stop lane markings. This metric is of great importance as stop signs and markings are typically placed in dangerous areas. However, the current version of CARLA does not provide an efficient way to check for these violations thus this is left as a place for further improvements. More about it will be talked about in Chapter 10 when we talk about the possible future improvements of this project.

6.3 Evaluating the performance

Now that monitoring of the vehicle's actions was discussed and the metrics being captured were presented, it is time to give an overview of how the performance can be evaluated. Let us begin by introducing a table with penalty scores for violations of different safety requirements. This table can be found in Appendix A in Table A.3. It consists of three columns: the first one

highlights the category of the violation, the second one gives the number of penalty points a driver is assigned if he violates the requirement, and the third column shows the values that are assigned to the driver's score for doing the violation while speeding.

Note that the values in the table are just a proposition of what could be used, and they can and should be changed to fit the project's needs accordingly. The weights shown in the table are used to assess the negative part of the driver's performance, namely how unsafe it is. All the values introduced in this section are used to evaluate the performance of participants competing in scenarios in Chapter 9.

In order to evaluate the driver's performance, a mathematical formula was developed that consists of three main elements: the positive score, the penalty score and the discount factor γ. The discount factor is a hyper parameter and is there to deal with the inaccuracies caused by the simulated environment. It should be adjusted depending on the simulated environment and the objective of the research. The exact purpose of the γ will be discussed later when we will discuss how the penalty score is calculated because it is heavily impacted by γ.

Let us first talk about the positive evaluation element. It is made of the road completion c multiplied by the scenario difficulty d multiplied by the timeliness of the driver. The meaning of the c is what proportion of the route the vehicle completed successfully. The difficulty d is retrieved from the scenario file, indicating how challenging the scenario was. The timeliness is the optimal scenario completion time divided by the time it took for the driver to complete it. In simpler terms, the quicker the more complex scenario is completed, with the driver sticking to the path without any deviation, the higher the positive score they will receive.

The penalty score is the sum of all the penalty points from all monitors multiplied by the discount factor γ. It was mentioned before that γ is a hyper-parameter, meaning that the γ value can change the penalty score's effect on the final score. It is used to deal with the inaccuracies of the simulation and monitor malfunction. To exemplify this, consider a case where $\gamma = 1$. In that case, we are essentially saying that all the penalty scores received by the vehicle are correct and reasonable, giving the final score a maximum negative amount. However, the world is imperfect, especially the simulated world; thus, it is rational to assume that not all the sensors are 100% correctly obtaining all the data and not all the physics are impeccably implemented. As we will see in Chapter 9, there are many instances when simulator physics breaks down, and weird things start to happen. For example, upon colliding with another car, a car gets tossed into the air and lands a few streets further. Continuing about the monitors, the traffic monitor might capture that the vehicle did not stop at a stop sign, although it did, just

a few centimetres before it. Naturally, a parameter is needed to account for the unaccountable. By using the discount parameter, we are reducing the total penalty score while at the same time also reducing the miscalculation influence on the final score. Depending on the simulator and how well the physics and monitoring work, the γ value should be adjusted. Summing all up, the final formula makes an equation shown below:

$$score = c \times \frac{t_o}{t} \times d - \gamma \times \sum_{x=1}^{n_m} f(m_x) \tag{6.1}$$

where: γ = discount factor and $0 < \gamma \leq 1$

m = a set of lists of penalty values

$f(m_x)$ = sum of all the penalty points in list m_x and $m_x \in [0; +\infty)$

c = proportion of route completed and $0 < c \leq 1$

t_o = optimal amount of time (in seconds) to complete the scenario and $t_o \in (0; +\infty)$

t = amount of time to complete the scenario and $t_o \in (0; +\infty)$

d = difficulty of the scenario and $0 \leq d \leq 1000$

n_m = number of lists of penalty values in set m

However, it was not mentioned yet, how the function $f(m_x)$ is calculated. If we assume that $f(m_x) = f(x)$, since m_x is a list of penalty scores of metric x, we get the formula shown below:

$$f(x) = \sum_{i=1}^{n_x} x_i \tag{6.2}$$

where: $x_i = i_{th}$ score from list x and $x_i \in (0; +\infty)$

n_x = number of scores in list x

Now that the final formula has been introduced, it is time to return to one bit that was mentioned but not analysed yet, the optimal time value t_o.

The purpose of the t_o is to proportionally increase the final positive score if the driver is driving faster than the average and proportionally decrease it if the driver is driving more slowly than the average driver. If we look back at the score calculation formula, we can see that the optimal time value t_o is divided by the time the driver took to complete the route. The positive score is mainly determined by the scenario difficulty d, while the c is there to either retain the maximum score or decrease it if the driver did not complete the whole route. On the other

hand, the t_o divided by t increases the score if the driver is quick and decreases it if he is slow. Let us take a look at an example. Suppose the scenario's difficulty is 500 and the optimal time value to complete that scenario t_o is 200 seconds. If the driver finished the route in 100 seconds, meaning he is two times quicker than the calculated average, he gets $c \times 2 \times 500 = 1000c$ positive points for the scenario. However, if he were to complete the scenario more slowly than the optimal, he would get $score < d$.

Given the simulated environment, this article proposes a mathematical formula for t_o evaluation. It consists of three main parts. The first is the distance s divided by the average maximum allowed speed v_{avg}. This part tells us the ideal time the vehicle would take to finish the distance if it was driving in a straight line without any stops and other traffic. However, the roads are full of other drivers changing lanes, slowing down, making turns, talking on the phone and not seeing the green light, etc. For this reason, the first part of the formula is multiplied by $1 + \alpha$, where α denotes the intensity of traffic $\alpha = 1$ being the road is almost a gridlock and $\alpha = 0$ being the road is completely empty, and there are no other vehicles on it except the driver himself. Of course, intensity value should consider other factors such as road works, diversions or road surface quality. However, for the purpose of this project and the simulator's limitations, we are only considering the traffic density. The third part of the formula is a sum of all stops the vehicle has to make on the road. This includes pedestrian crossings, traffic lights, stop signs/markings and others. The complete formula for t_o can be found below:

$$t_o = \frac{s}{v_{avg}} \times (1 + \alpha) + \sum_{i=1}^{n} t_{avg\_ni} \tag{6.3}$$

where: s = length of the route in m

v_{avg} = average speed limit over the route in m/s

α = the intensity value and $0 \leq \alpha \leq 1$

n = number of of stops a vehicle has to make

t_{avg_ni} = average time in seconds the vehicle has to be stopped for at stop ni

However, a sharp-sighted reader might notice that an accurate calculation should consider the historical data, thus allowing for more confidence about the calculated value. Unfortunately, there is currently no source that could provide historical data for the simulated environment. Even if there were one, it would not necessarily be accurate one since the parameters of the simulator can be easily customised, affecting how the data is captured. Moreover, the t_o does not include variables determining how long the vehicle would take to reduce the speed when a

stop is encountered and how much time and distance it would cover to reach the speed limit again after the obstacle is passed. The formula implicitly assumes that this metric is handled by the t_{avg_ni} indicating the average time the vehicle takes to pass a probable stop. Additionally, it is worth mentioning that the last bit of the t_o formula, namely the sum of all the stops, also includes the inevitable stops generated by the scenario generation algorithm. As mentioned in Chapter 5 and Chapter 10, in the future, the scenario generation algorithm could be improved to generate individual situations and not only general scenarios. If this was the case, and the algorithm had generated a situation that involves a collision, the average time the driver would take to drive past that collision should be included in the sum of all stops. This is because every detail, affecting the speed of the vehicle is significant and should be taken into account.

In summary, although the research in this project uses the introduced formula for safety score calculation, it serves only as a proposition of what could be used for evaluation. In fact, the formula used to evaluate the software in Chapter 9 is simplified because of the simulator's limitations and poor structuring of the OpenDRIVE maps used. It considers only junctions as inevitable stops that the vehicle has to make and gives them an average waiting time of 12 seconds. This is because of the missing information in the OpenDRIVE map files the CARLA development team provided. An example of this can be seen from Figure C.3 until Figure C.6, in Appendix C that exemplify how road elements in the same map are specified differently and that on average, less than 30% of the road elements have a speed limit specified. In addition, some maps contain information about where the traffic lights are, and others do not include that information. Much more research and proving needs to happen before we can establish the correctness and quality of the safety evaluation function. A significant amount of effort was paid in search of an existing function in the literature that would have already survived the test of time and could be used here. However, no function found in the literature could suit the needs of this project. For this reason, the article aims to be inventive but risk-taking simultaneously, leaving the reader to judge the formula's correctness.

Chapter 7

Obtained data analysis

This chapter will discuss what can be done with the data obtained from the simulations. In Section 7.1, we will introduce the data analysis tool able to generalise data of many participants and extract locations where violations occurred. The Section 7.2 will present two ways the analysed data can be depicted visually, mainly via diagrams and location marking on the map. The chapter is finished in Section 7.3 after introducing a script calculating the safety performance score of a driver in a particular scenario building on the concepts discussed in Section 6.3.

7.1 The data analysis tool

Let us begin by discussing how the data obtained from the simulations could be processed. This article proposes an easily customisable way to perform data analysis that suits the user's needs. An overview of the analysis process can be seen in Figure 7.1.

Here the user specifies the parameters in the analysis_parameters.json file and runs the run_analysis.sh script, which takes the necessary data from the database, performs the analysis and saves the analysed data back to the database in a new folder whose name is specified by the user. An example of what the analysis_parameters.json file looks like can be seen in Figure A.4 in Appendix A. In the file, the user specifies which participants, scenarios and safety metrics must be analysed. The algorithm then performs the analysis and stores the data that can be divided into three categories: the overall performance of participants in the scenarios (stored in participants.xml file), the average scenario scores of the participants (stored in scenarios.xml file) and files with locations of where the violations occurred (stored in separate files in points

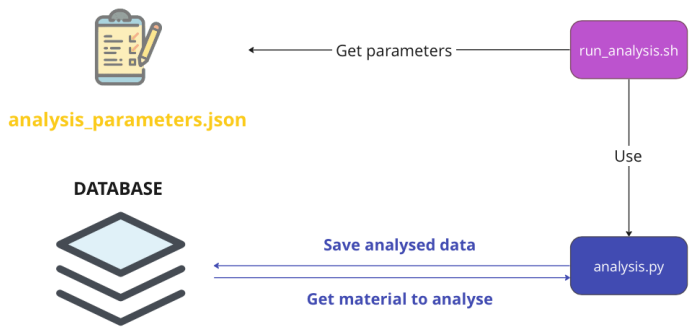

Figure 7.1: An overview of data analysis

directory). To deepen the understanding of the analysed data, the reader is highly encouraged to have a look at an example of data analysis in data/analysed_data/data/example_analysis directory.

7.2 Visualising the data

Now that we have introduced a way of analysing the data and described the basic principles, it is time to see how analysed data can be put in a visual format for further analysis. The overview of the visualisation methods can be seen in Figure 7.2.

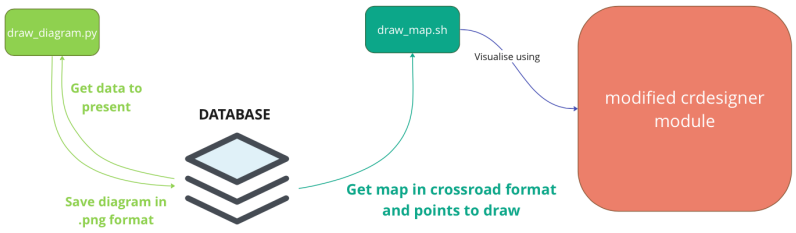

Figure 7.2: Data visualisation process

The last section mentioned that the data analysis tool creates three types of processed data: participant analysis, scenario analysis and metrics analysis. The first two categories can be used to draw diagrams indicating the general tendencies. An example of the diagram can be seen in Figure 7.3. The diagram displays the amount of penalty points each participant received for collisions in scenario 1. The information can help compare participants, including

the AI implementations, in different scenarios. Similar diagrams can be generated to compare the scenario scores concerning different metrics. For instance, to compare which scenarios were more prone to traffic rule violations or breaches of other safety metrics.

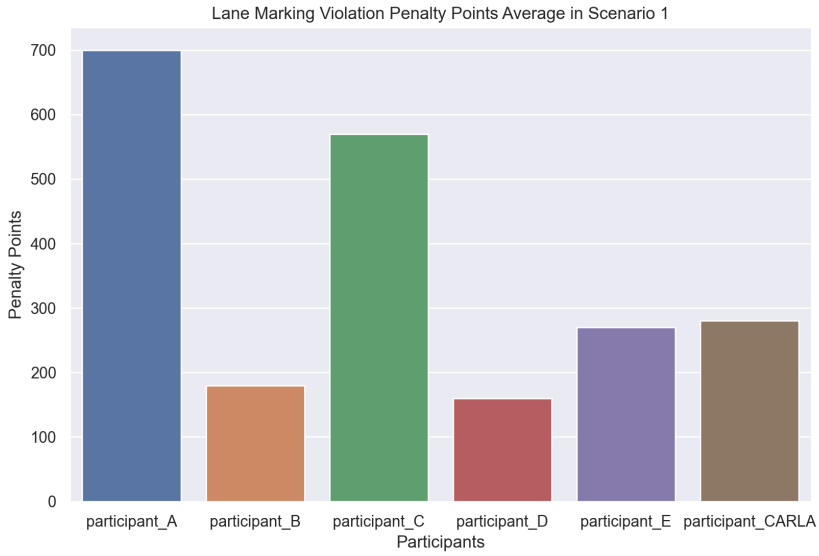

Figure 7.3: An example of a diagram

The third category of analysed data is locations where the violations occurred. Depending on the scenarios, participants and metrics specified in the analysis_parameters.json file, the files in the points directory will differ. Taking a look at an example should clarify the purpose of this type of analysis. Suppose the parameters file specifies that the analysis should be performed on all the participants (listing their names) in scenario 1 and should consider the collisions only. The analysis tool would create a file points/scenario1_collision_data.xml containing all the locations where the collision occurred during the simulation runs of that particular scenario. If we then use the **draw_map.sh** script and specify that file to be pictured, the view similar to the one in Figure 7.4 will appear on the screen.

In this figure, the blue Xs indicate the places where the analysed participants made collisions. Information like this can help identify the places where the vehicle could not react on time or did not react appropriately, and the action resulted in a collision. With the help of the **replayer.sh** script, the simulations can be replayed to analyse the cause of the collision further and thus improve the AI model or simply observe the tendency. It is worth mentioning that the **draw_map.sh** script can take two lists as parameters, depicting the points in red and blue Xs. This can be helpful in comparing the different participants or categories of participants.

As shown in the Figure 7.2, the map drawing process is facilitated by the CommonRoad

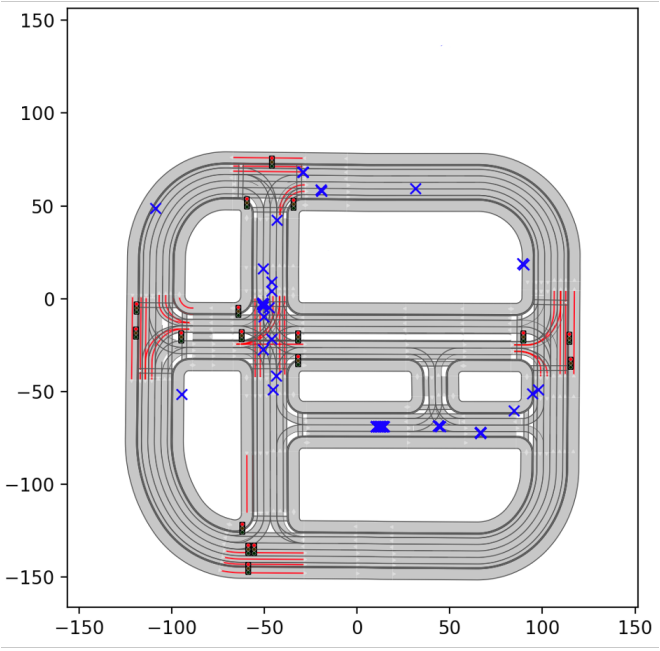

Figure 7.4: An example of a map with points marked

Scenario Designer module [21]. Note that this article uses a modified version of the CommonRoad Scenario Designer. This is due to the fact that the most recent version of it (version 0.6.0) did not have an option to provide a map and a list of points to be represented on the screen when launching the module. It only allows that after the tool was launched, and even then, it only allows for a single point to be marked on the screen. A few modifications were made to the module to launch the visualisation by just running a single command in the terminal. In addition, a bug was fixed not allowing to pass some arguments to the application. The owners of the project were informed about the error found. Because the CommonRoad Scenario Designer uses a different format for drawing the map, the OpenDRIVE files had to be converted to the CommonRoad map specification format. For this, a converter was created that can be found in the data/maps directory. The modified crdesigner module is located in the software/python_libraries directory.

7.3 Calculating the safety score

In this section, we will briefly mention how the safety evaluation formula introduced in Section 6.3 can be used to calculate the safety performance scores of participants in scenario simulations. For that, the **score_calculator.py** script is used, which is located in the software/data_analysis directory. The script takes two arguments: -p participant's name and -s

scenario that the participant drove. Both arguments have to make a path in the data/recordings directory, meaning there must be data about the participant driving that particular scenario. An example of what is given to the user once the score-calculating script is run can be seen in Figure 7.5.

The script gathers data from the database, processes it, applies the formula and prints the score to the terminal, also indicating what is considered a perfect score. More about how to perform the data analysis, visualisation and score calculation is explained in Appendix B where it is demonstrated how to use the project step by step. We will also come back to data analysis and score calculation in Chapter 9 when we analyse how human drivers and a CARLA agent performed in the scenarios.

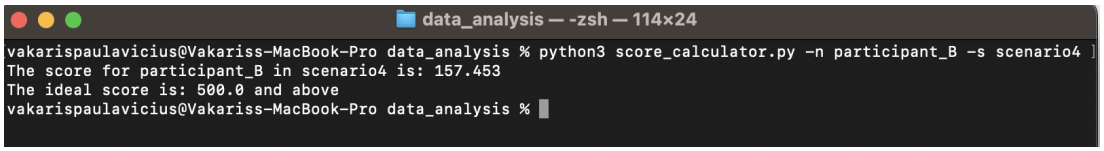

Figure 7.5: A score calculation example

Chapter 8

How the software works

This chapter aims to sum up everything discussed over the last several chapters and give an overview of how the software bundle presented in this article functions as a whole. Note that this chapter does not include instructions on how to run the software. For that, please refer to Appendix B and its sections explaining how to do that step by step.

The process of running simulations and storing the obtained data is managed by the run.py script which delegates the tasks to other scripts until the simulation session is concluded. It is called by the **run.sh** script, which the user must use to start the simulation session. It takes the participant's name as an argument and an optional keyword,"ai", indicating that an AI implementation will drive the scenarios. The run.py script creates a directory for each participant in the data/recordings directory and, within it, creates a directory for each scenario in which the participant will participate. This approach allows for easy access and organisation of the recorded data during analysis.

The scenarios that the vehicle has to drive are listed in the carla_scripts/scenario_list.json file. An example of the scenario list file can be seen in Figure A.5 in Appendix A. Here all the scenarios that the participant is supposed to drive are shown. If the scenario or path is not specified (they are null), the algorithm will automatically generate them during run-time, resulting in a new and unseen route.

For each scenario, a corresponding **scenario_manager.py** script is launched that takes care of the scenario simulation process. It first generates a path and scenario if they are not specified, then changes the map and weather, populates the map, and defines the behaviour of actors on the road. In the beginning, all the simulation actors are immobilised to ensure that each simulation begins with the same conditions for each participant.

Once everything is set up, the hero vehicle is spawned. Depending on if the "ai" keyword was specified when the run.sh script was launched, the scenario manager either launches a manually controlled vehicle or the self-driving one. For that, either driver.py (keyboard or steering wheel) or self_driver.py scripts are called.

If a human driver controls the car, then once the Python application for driving is launched, the driver, in order to start the simulation, needs to press the BACKSPACE key. This causes the Scenario to unfreeze all actors in the simulation, allowing them to move around the world according to their designated behaviour. The simulation also launches the Manager module responsible for all performance calculations and violation marking.

The simulation can end in two ways: reaching the finish line or pressing the BACKSPACE key to abort the simulation. The latter option was added to allow immediate simulation termination if something goes wrong and make it easy to rerun the simulation later (thanks to the flexible structure). Once the simulation finishes, all data and the simulation recording are saved to disk, and the subsequent simulation is launched. The process is repeated until all the scenarios have run from the scenario_list.json file. Note that manual driving can be done either using a keyboard or a steering wheel. The type used can be changed by uncommenting line 158 or line 157 in software/carla_scripts/simulation_manager.py script.

If the AV implementation is driving in the scenarios, everything is happening automatically, and no buttons need to be pressed. Only the run.sh script needs to be run with the participant's name and the keyword "ai".

The simulation process is also illustrated in Figure 8.1, providing a visual overview of how the components act together.

It is crucial to mention that not only the CARLA agent can drive in the simulations but any AV implementation that can be bridged with the CARLA simulator. After the connection to the simulator is complete, all that is needed is to give the waypoints generated by the GlobalRoutePlanner to the agent to follow, name the vehicle "hero" within the simulated environment and start the Manager module when the simulation starts and terminate it once the finish is reached. Everything else works independently, talking to the server directly and receiving the necessary information.

Once the simulations are complete, the user can run the analysis and then work with the analysed data. More about how to use all the software elements can be found in Appendix B.

Several additional Python scripts were written to understand the CARLA world better and help design the scenarios. These include print_coordinates.py, which prints the precise

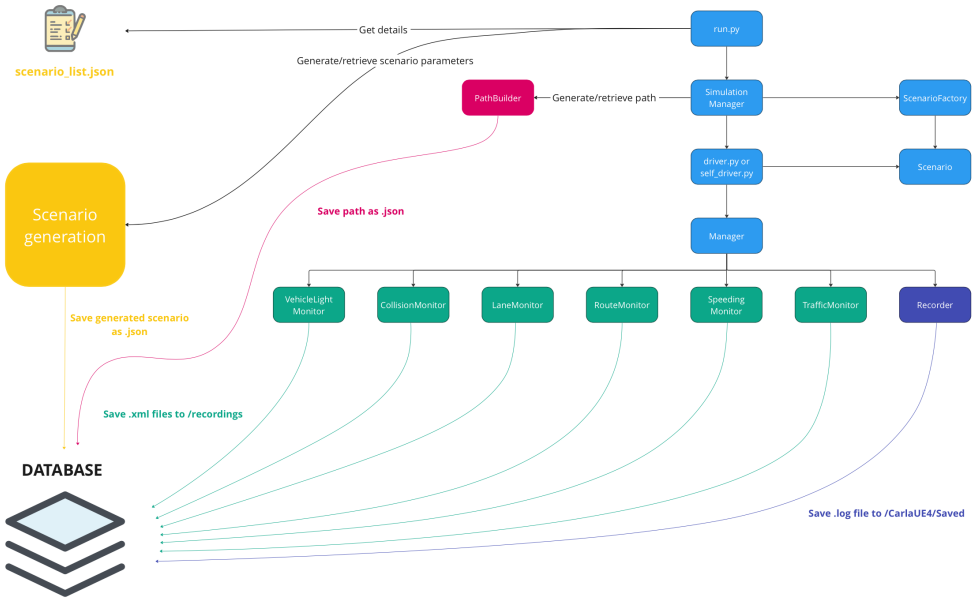

Figure 8.1: An overview of the simulation process

coordinates of the spectator, who can fly anywhere in the world, and draw_lanes_terminal.py, which generates a route when given start and finish coordinates and marks all legal lanes on the map. Other useful scripts can be found in the sub-directories of the software/carla_script directory.

One necessary script to mention is the replay.sh, which enables the replay of any simulation driven by a specific participant on the screen. For example, by executing the command "/replay.sh participant_A 1" in the terminal, the simulation of the first scenario driven by participant_A will be replayed.

Chapter 9

Software testing and evaluation

After introducing the software components proposed by this article for testing vehicles in a simulated environment, we will now look at how the tools can be used in practice. Specifically, we will discuss how the software was tested in a lab with five human participants and an AI agent. We start in Section 9.1 by discussing how the participants were chosen and what had to be done to facilitate their participation in the simulations. Next, in Section 9.2, we will look at how the simulations took place and talk about the things observed during the process. Finally, in Section 9.3, we conclude the chapter by discussing the results of the simulations, comparing the safety performance scores of the human drivers and an AI agent and evaluating how well the proposed system operates.

9.1 Selecting the participants

For this project, it was decided that the best way to test the developed software is by inviting external participants to drive in the generated scenarios, observing how monitors collected data, and running data analysis evaluating their performance.

As this project aims to create a framework for evaluating autonomous vehicle safety, some AI implementations had to be involved in the process. Initially, the leading candidate for participation was the Autoware.AI framework allowing a vehicle to be controlled by the Autoware software in the simulated environment. The reason it was such a good candidate and remains one is that Autoware.AI is a sophisticated software bundle, implementing a level four autonomous agent (discussed in Section 2.3), possessing the ability to participate in traffic. It is also relatively easy to connect it with the CARLA simulator by connecting Autoware.AI and

CARLA using ROS and Autoware bridges provided by the respective developers.

However, due to the complexity of both software systems, the high learning curve and issues not allowing the Autoware.AI agent to follow a set of waypoints, it was decided to use a more straightforward implementation. As the purpose of this project is not to connect different AI agent implementations with the simulator and run them but rather to create a software system for their safety analysis, the choice of the AI participant in the research holds very little significance. For this reason, CARLA BasicAgent was chosen to participate in the software testing.

While still in the initial steps of the project development, it was already established that for the AVs to be considered safe and finally witness the daylight, they have to fit people's safety requirements. It was then that the idea of comparing human drivers and AI agents in the simulated environment made its way onto a piece of paper. It is intriguing how both AI-controlled vehicles and humans would behave in the same situations. For this reason, from the beginning, it was clear that the project needed to be developed to facilitate user testing in the lab.

The first step was to decide how many participants should take part in the research and what prerequisites are needed to allow this process. It was decided that having five people driving in the scenarios would give a good overview of general trends and make for exciting research. As with all projects involving human research, an ethics committee approval was required. For this, an application was submitted, and a positive response came back, allowing the research to occur. The ethical review reference number is LRU-22/23-34770.

Moving on, it is evident that for a participant to reflect the behaviour of a typical driver in the simulated environment, he needs to have had some experience driving on the roads. For this reason, the only requirement for a person to participate in the research was to have a valid driver's license. The licenses were not inspected, but it was assumed that the participants willingly participating in the research were truthful. Five participants were found, and all agreed to take part. Their consent forms are stored safely together with the project files.

While the AI agent only requires the software to drive in the scenarios, the human participants require controllers to operate cars. In order to be as reflective of the real-world conditions as possible, the research had to provide the participants with the tools they are accustomed to when it comes to driving a car: the steering wheel and pedals. For this reason, it was decided to run a controlled experiment in the university lab containing all the necessary hardware.

In addition to the physical tools, software scripts allowing the vehicles to be controlled

by both people and AI were needed. For this, manual driving scripts were borrowed from the example scripts provided by the CARLA development team that can be found in the Carla/PythonAPI/examples directory. The scripts were then modified to suit the purpose of the project. The scrips are named driver_keyboard.py and driver_steeringwheel.py for the human participants and the self_driver.py for the CARLA agent. The scripts can be found in the software/carla_scripts directory.

9.2 Human drivers vs. AI

The research took place in a Bush House computer lab (30 Aldwych, London WC2B 4BG, United Kingdom). The lab contained the steering wheel and pedal-equipped powerful computers allowing the simulations to run at the highest quality graphics. This ensured that the impact of the uncontrolled variables (concerning how people perceive the environment) was minimised. The technical specifications of the computer used can be seen in Figure A.6.

All the participants were tested in five scenarios generated by the scenario generation and path-building algorithms. As discussed in Chapter 3, when we examined the road incident statistics, the generated scenarios must be diverse to reflect various road conditions. For this reason, the five scenarios cover both day and night conditions, rainy and sunny weather, urban and rural areas and highways. Some maps had their start and finish coordinates altered slightly to make the vehicle start or finish in places like a sideline or a parking lot space. This is due to the shortcomings of the path-building algorithm discussed in Chapter 5. More about it will be discussed in the following chapter, where we conclude the project by talking about the issues experienced and possible future improvements.

For the experience to be identical for all the participants, the same vehicle was used throughout all the scenarios' simulations. In addition, randomisation seeds were used with each scenario, guaranteeing identical vehicle behaviour and their arrangement on the map. The scenarios that were driven can be found in the software/carla_scripts/scenario_list.json file. The fields "name" and "details" could be changed to have values pointing to *null*, thus resulting in entirely new and random scenarios being generated. However, as this research aims to get an accurate overview of how people drive, all the scenarios had to be identical for all the participants. That is why randomisation seeds and predefined scenarios were used.

However, ensuring an equal experience for all participants proved to be a challenging endeavour, fraught with various obstacles. To begin with, technical glitches posed a significant problem, such as the steering wheel powering down during simulations or the accelerator pedal

giving impulses as if it was being pushed despite having no input. Issues like these resulted in specific scenarios being rerun. Additionally, unforeseen misbehaviour from certain vehicles within the simulation also had a significant effect, mainly on the scenarios themselves. This includes vehicles flying into the air after a collision and landing somewhere else, causing gridlock on the road and thus impacting the flow of traffic.

As reflected by the participants, although the steering wheel and pedals were there, what was lacking was the feeling of actually sitting in a car and being able to turn the head and observe the environment. This was not well conveyed by the static monitor giving the view of the world only at a specific angle at the time. In addition, the participants said they were missing the noises from the environment as the monitors and the CARLA simulator did not provide any sounds. Some also remarked that the steering wheel was too sensitive, and they did not feel the car well. Having said that, it is safe to say that the fact that the participants were driving the car in a virtual environment had an effect, presumably negative, on their performance.

What was also noted from the simulation runs was that the participants found driving at night more difficult than during bright conditions. Overall, the participants said that they enjoyed testing their driving skills in the simulator and gave positive verbal feedback about the process of the research.

All five participants successfully completed all five predefined scenarios. The data collected from their scenario runs can be found in the directory data/recordings from participant_A until participant_E.

Regarding the CARLA agent participating in the simulations, there were no technical issues with the equipment. However, there were some problems with how the vehicle executed the given path. At some points during a run, the vehicle would make a 360 degrees turn and then continue driving the path. Although the exact cause remains undetermined, it most likely was provoked by the inaccuracies caused by the GlobalRoutePlanner, which created the path. The run of the scenarios by the CARLA agent can be found in the participant_CARLA directory.

9.3 The outcome

In this section, we will evaluate the performance of the safety monitoring tools considering the data they obtained from the scenario executions. In addition, we will utilise the data analysis tools introduced in Chapter 7 to help us assess vehicle safety and provide some insights into where the most issues occurred.

Let us begin by analysing whether the monitoring tools had done their duties. As mentioned previously, the flow of each simulation session is captured in log files found in the subdirectories of the data/recordings directory. The log files are associated with each participant and record how the simulations were fulfilled, whether the monitors successfully recorded data to their intended files, etc. After examining each session_logs.log file, no errors or misbehaviours were spotted, meaning that simulation sessions were completed as planned. After looking at each scenario directory, it was confirmed that all the files were present. Having established that no errors occurred when creating the files and no issues were spotted in the simulation session records, it was time to look at the contents of the recording files to see whether they contained the right data in the correct format.

This process was performed manually because no intelligent software system was developed for checking the correctness of the files.

After observing the files, it was confirmed that the data was recorded following the XML file structure and formatting conventions described in the code. The collision files contained the instances of collisions, where they occurred and at what time; the other files contained their respective violation data. The general accuracy of data was checked by replaying some of the simulations and ensuring that the recorded violations occurred. The majority of the violation data was captured as planned.

However, there were instances that brought inaccurate data, thus affecting the safety score of the participants. While replaying the simulations, it was observed that whenever another car bumped into the participant's car, a collision was recorded, giving the participant penalty points for situations where he was the victim. In addition, there were some instances where a collision with one vehicle was recorded multiple times, giving more penalty points than intended. This can be seen in the collision data of the participant_D in scenario 2.

The first issue with the collision monitor was caused by the simplified implementation of the CollisionSensor, which only considers the instances when another object entered the vehicle's bounding box. It was not distinguishing between the driver running into another object and an object running into the driver. The second issue of the monitor recording the same data multiple times was confirmed to be temporary wrongdoing by the CollisionSensor, as no other instances like this were present. Overall, it is safe to say that the monitors were performing their tasks, although sometimes not doing them right.

Having examined the work of the safety monitors and established that they mostly recorded the right violations, it is now time to use the data analysis tools to analyse the observations

even more and evaluate the performance of each of the participants. For that, the data analysis tool was employed. After specifying all the scenarios, all the metrics and all the participants in the analysis_parameters.json file, the scenario analyser was run and completed the analysis successfully. The analysis can be found in the data/analysed_data/data directory under the name example_analysis. After observing the participants.xml and scenarios.xml file content and manually checking that the values determined by the analyser were correctly calculated, it was confirmed that the analysis tool behaved as expected, producing the correct scores. For example, the sum of all the collision penalty points gathered by the participant_A over all five scenarios is 5750. The analyser correctly tells us that the collision point average of this participant is 1150, given that he drove in five scenarios. The same calculations were manually performed on other metrics to ensure the correctness of the data analysis tool.

Similarly, the point extraction (from the points subdirectory in the analysis folder) was checked to be correct by looking at the coordinates in each file and ensuring they are present in the recording files. This check was performed on a simpler (one participant) analysis because checking for each point in the example_analysis/points would take a very long time. Nevertheless, if the algorithm behaves correctly with few participants, it is safe to assume that it will also behave correctly with many participants.

The visualisation tools were tested using various metrics from the analysed data. Although the diagrams comparing the scores were being drawn correctly, the map drawing tool marking the locations on the map was sometimes accurately reflecting the coordinates and sometimes marking the points where there was no road. The correctly represented points can be seen in Figure A.7, and the incorrectly marked points can be seen in Figure A.8 in Appendix A. Due to the complexity of the CommonRoad Scenario Designer and lack of documentation, it was challenging to determine the root cause of why the points are being represented as they are. For this reason, this issue remains unsolved.

It is now time to move on to the safety evaluation bit. For this, the score_calculator.py script was used, whose purpose was described in Chapter 7. The correctness of the algorithm was verified by running the algorithm to evaluate the performance and then manually recalculating the score and comparing the values. This method proved that the algorithm behaves as expected and accurately applies the formula introduced in Chapter 6.

However, the algorithm is not using the formula to its fullest because of the lack of information retrievable from the simulation environment and the map description files. For instance, there is no straightforward way of retrieving the allowed speed limit in a specific road element

using the CARLA simulator. Instead, they can be retrieved from the road elements specified in the OpenDRIVE map descriptions. However, as explained in Chapter 7, most road elements in the map descriptions do not provide any information about the allowed speed limits. Because of this, a lot of assumptions had to be made. For example, if the speed limit on the road element is missing, the maximum allowed speed is assumed to be 50km/h. This can affect the final score dramatically if the permitted speed is 30km/h, but the algorithm assumes that it is 50km/h, thus calculating the lower optimum time value to complete the route resulting in the driver either speeding to make it to the finish line on time or losing the points for driving too slow. Similar issues occurred with counting the mandatory stops on the map, mainly because there is no way of retrieving the number of stop signs or pedestrian crossings on the road. For this reason, an assumption was made that the vehicle would have to wait an average of 12 seconds at each junction. This value is arbitrary and does not serve as an accurate measure. However, it is impossible to accurately determine such metrics without more data.

Having said all that, the score calculator was used to calculate the scores of all the participants in each scenario. The formula used γ value 0.7, meaning that the collected data about the penalty points is 70% accurate. The highest score achieved was 435.661 by the participant_C in scenario4. The scenario's perfect score, calculated using the formula, was 500 points and above. The ideal score is achieved by completing the route in full, in the exact time given and without doing any violations. If the driver completes the course even quicker than the calculated optimum time, he gets more points. The lowest score achieved was -18979.65 by participant_D in scenario2. The score was so negative because of the issues with the collision monitor mentioned before, where one collision was recorded multiple times resulting in many penalty points given to the driver.

Overall, the scores of all the participants were mostly highly negative because of the large amounts of penalty points given for each penalty. This can be addressed by adjusting the penalty points assigned for each violation. Also, it was observed that the participants also performed lots of violations, such as accidentally crossing solid lane markings and not showing turn indicators when turning or headlights in the dark, adding to significant negative scores. Out of all six participants from the research, the CARLA agent ended up being the fourth in terms of all the penalty points obtained, which is 10237.

Chapter 10

Conclusion

This chapter concludes the discussion of AVs and the introduction of an alternative way of measuring vehicle safety in a simulated environment. In Section 10.1, an overview of the project is given, explaining what has been developed. The following section discusses the research relevance of the project and talks about the attitude with which the project was implemented. In Section 10.3, a list of possible future improvements is provided for people willing to take this project further. In Section 10.4, ethical concerns surrounding the project and the autonomous driving field are considered. The chapter is concluded in Section 10.5, after the author's self-reflection about the challenges faced over the development of this project.

10.1 What has been presented in this article

In the early chapters of this article, it was established that for autonomous vehicles to appear on the streets, their safety needs to be recognised by the general public. However, we soon discovered that traditional ways of testing autonomous vehicles, namely by exploiting them on the roads, would be inefficient given the time needed and the non-deterministic nature of the world. A new and innovative approach was needed - an approach that would look at the problem through another perspective - the simulated one. The article focused on developing a software system allowing for automated vehicle safety assessment in the virtual setting.

Built around the simulator CARLA, the introduced system consists of three main components. The first is an automatic scenario generation algorithm that, with the help of a sophisticated machine learning algorithm and a path-building script, can create diverse traffic scenarios. Generated scenarios allowed for a vehicle to be tested on the roads and its actions

recorded by the safety monitoring system, which is the second component. The system is responsible for observing the vehicle's actions and transmitting the data to its monitors that record the violations in the database. The data analysis tool then handles the recorded data by generalising it and preparing it for visualisations using diagrams and 2D map representations, allowing for further analysis of behaviour patterns. In addition, an equation for evaluating the vehicle's performance in a given scenario was developed that allows for a simulation run to be allocated a safety performance score. Moreover, besides all the technical details involving the development of the system, the article also discussed in detail how the ideas of autonomous vehicles matured throughout human history, compared the AI behaviour with that of the human drivers, tested the proposed software system with participants and evaluated how well the system works in practice.

10.2 Future improvements

As with any bigger project developed in a limited amount of time and with limited resources, there is always a place for improvement. This section aims to give some guidelines about what could be improved and in what direction.

Let us start with the scenario generation algorithm. As mentioned in the Section 5.4, the algorithm could be upgraded in several ways. Firstly, a set of real-world data could be used to have a more accurately trained algorithm whose generated scenarios would be more reflective of the real-world conditions. In addition, a more sophisticated algorithm could be used able to create even more diverse and unique situations, allowing for the automatic generation of other vehicle behaviour. For instance, generating situations where an actor in the simulation acts according to some patterns when the participant vehicle is close (suddenly changing lanes, for instance). However, the current implementation is only able to create general scenarios where the behaviour is set to all the vehicles with the hope that some interesting and unusual traffic conditions will make their way to the surface. What would also be good to have is a system able to analyse how challenging and diverse the scenarios actually are.

The path generation algorithm also has places for improvement, starting with how the paths are built. At the moment, the algorithm creates paths by trial and error, randomly choosing the starting element and then proceeding from there. It would be much better if it did that in a structured way, reasoning about why the path has to look like that and relating it to the generated scenario. This would allow for even more exciting paths and behaviour patterns to be built.

In terms of the safety monitoring system and the data analysis tools, they currently do not provide a way to use the observations for model improvement. They are there to record the violations and then look at the general behaviour patterns without linking the observations with some mistakes in the model. This, in addition to the improvements to the collision sensor (mentioned in Section 9.3, should be addressed in future updates.

Moreover, the current software does not provide automated tests to check the components' validity and correctness. As with autonomous vehicles, the systems measuring their safety must also be flawless. For that, a comprehensive test suite should be developed, ensuring that all components work as intended and that updates to one do not break anything in another.

10.3 Research relevance and project's integrity

This article aims to contribute to the field of autonomous driving by presenting a method of testing driverless cars in realistic traffic situations. As highlighted earlier in the paper, demonstrating the safety and reliability of autonomous vehicles to the general public is vital in bringing all the hard work in the field to reality. This statement is supported by other literature as well [22]. Every little step in this direction is significant, and by raising people's curiosity and awareness about the latest developments, we are getting closer to achieving this goal. By introducing this project, we hope to grab the attention of young people and inspire their engagement in this field, emphasizing that it can be both enjoyable and fruitful.

In addition to contributing to the field of autonomous driving, the project aimed to leverage existing open-source projects. By doing so, the project aimed to both benefit from existing systems and introduce innovative ways of utilizing these tools while also showing appreciation for other software developers and their creations. This article expresses gratitude towards the projects mentioned, namely the CARLA simulator and the CommonRoad Scenario Designer.

All parties involved in the project's development were aware of the BCS Code of Conduct[23] and always sought to follow the rules. That is why the project was developed with utmost care and honesty, first learning and understanding the problem and only then trying to accomplish the goals. Moreover, the research involving human participants was carried out in a professional and respectful manner resulting in mutual professional and personal growth.

10.4 Ethical concerns

The field of autonomous driving has always been concerned with ethical questions, as the introduction of such technology would have a significant impact on the lives of many people. Thus, the intentions of engineers need to be carefully considered, and the possible implications rationally weighted. There is no room for doubt regarding the well-being of living beings. Therefore, it is essential to thoroughly test the underlying technology and hardware to ensure they meet the necessary standards. As the use of artificial intelligence in the field is inevitable, proper precautions must be taken. The algorithms cannot be biased, resulting in harm to some groups of people because of the lack of training data or biases in the training data. The systems must be developed with the intention of producing rational and unbiased decisions resulting in the best possible outcome for all parties.

Although this project did not extensively explore the ethical implications of autonomous driving, it did take steps to prevent biased outcomes in the proposed scenario generation method. The article aimed to create diverse scenarios for vehicle safety testing to get comprehensive analysis as discussed in Chapter 2 and Chapter 5. Additionally, it was assumed that CARLA's set of models for vehicles and pedestrians was diverse; thus, the algorithm randomly selected models without any prejudice. While these steps may not have addressed all potential biases, they demonstrate a commitment to ethical considerations in the field. This project is in favour of ethical and unbiased project development.

10.5 Author's self-reflection

The purpose of this section is for me, as the author of this paper, to reflect on my journey and the work produced.

Doing this project allowed me to better understand the complexities surrounding autonomous driving and provided me with hands-on experience contributing to the field. The journey was challenging and full of unexpected changes and issues, constantly pushing me towards upgrading the project and looking for new ways of achieving my goals.

Having minimal programming experience in Python and no experience in machine learning, coming up with an automated way of generating the scenarios proved to be particularly challenging. The journey began with studying the machine learning methods on my own, then moved to implementations of various classification and regression algorithms and finally to applying the knowledge in solving the scenario generation in this project. Although this project

presents a simple and not the best way of generating the scenarios, attempts to have a more sophisticated algorithm guided by deep neural networks were made. However, the method was not pursued because of the lack of suitable training data and a deep understanding of the CARLA simulator and the advanced machine learning methodologies. In addition, attempts were made to access some global data sources (like the SafetyPool repository[24]) containing scenarios that could be used as training examples. Unfortunately, several tries to get access did not bring any results.

Despite having many technical issues with my computer, which was barely running the simulations at the lowest quality graphics, I managed to successfully conclude my project and develop a prototype software system for vehicle safety evaluation in a virtual environment. I wish the university had provided us with either more lab time or the right equipment to engage with the project fully. Nevertheless, I enjoyed doing this project, and the challenges faced did not discourage but motivated me to move forward and engage in problem-solving.

I believe that the prototype introduced in this article has the potential to become a tool that would be helpful to the industry. However, much more work must be done for it to be effective and trustworthy, and it should be pursued in a team of developers rather than by a single person. Looking back, I think I have taken a bite too big to chew and decided to create an extensive and comprehensive system. I should have chosen one specific aspect, for instance, scenario generation, and focused on doing that only. The end result would have been a well-constructed artefact instead of a broad but less well-implemented system. The possible ways in which the system could be continued were expressed in the previous section.

The journey also taught me about time planning and work organisation. It also showed me that developing software in groups is usually easier than working entirely independently.

In the end, I feel that I have learned many new skills, such as implementing my own machine-learning algorithms, writing shell scripts, performing multi-threading in Python, working with JSON and XML files, analysing data, automating processes and writing academic texts with LaTeX. I also realised that I need to improve my software engineering skills and deepen my knowledge of artificial intelligence to be a better expert in the future and make the field of IT even more influential.

Bibliography

[1] Jong Kyu Choi and Yong Gu Ji. Investigating the importance of trust on adopting an autonomous vehicle. *International Journal of Human-Computer Interaction*, 31(10):692–702, 2015.

[2] Rachael Brennan, Logan Sachon. Self-driving cars make 76% of Americans feel less safe on the road, Sep. 14, 2022 [Online].

[3] J Fuller. Did davinci really sketch a primitive version of the car? *HowStuffWorks. com*, page 1, 2008.

[4] C Engelking. The 'driverless' car era began more than 90 years ago. *Retrieved February*, 20:2019, 2017.

[5] T. Jochem, D. Pomerleau, B. Kumar, and J. Armstrong. Pans: a portable navigation platform. In *Proceedings of the Intelligent Vehicles '95. Symposium*, pages 107–112, 1995.

[6] Twi global, 2022. Accessed on March 15, 2023.

[7] Jelena Kocić, Nenad Jovičić, and Vujo Drndarević. Sensors and sensor fusion in autonomous vehicles. In *2018 26th Telecommunications Forum (TELFOR)*, pages 420–425, 2018.

[8] Jorge Vargas, Suleiman Alsweiss, Onur Toker, Rahul Razdan, and Joshua Santos. An overview of autonomous vehicles sensors and their vulnerability to weather conditions. *Sensors*, 21(16):5397, 2021.

[9] Sae levels of driving automation refined for clarity and international audience, 2021. Accessed on April 4, 2023.

[10] Laura Fraade-Blanar, Marjory S Blumenthal, James M Anderson, and Nidhi Kalra. *Measuring automated vehicle safety: Forging a framework*. 2018.

[11] European commission Directorate-General for Mobility and Transport. *Road safety in the EU: fatalities in 2021 remain well below pre-pandemic level*, 3 2022.

[12] National Highway Traffic Safety Administration. *Traffic Safety Facts Annual Report Tables*, 6 2022.

[13] Morela Hernandez. The impossibility of focusing on two things at once. *MIT Sloan*, 2018.

[14] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. Carla: An open urban driving simulator. In *Conference on robot learning*, pages 1–16. PMLR, 2017.

[15] Ruxin Li and Runping Zhai. Estimation and analysis of minimum traveling distance in self-driving vehicle to prove their safety on road test. In *Journal of Physics: Conference Series*, volume 1168, page 032101. IOP Publishing, 2019.

[16] Mastercard and network international launch new artificial intelligence fraud-prevention solution, 2023. Accessed on April 4, 2023.

[17] The power of data: How paypal leverages machine learning to tackle fraud, 2021. Accessed on April 4, 2023.

[18] Robin van der Made, Martijn Tideman, Ulrich Lages, Roman Katz, and Martin Spencer. Automated generation of virtual driving scenarios from test drive data. In *24th International Technical Conference on the Enhanced Safety of Vehicles (ESV) National Highway Traffic Safety Administration*, number 15-0268, 2015.

[19] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830, 2011.

[20] Marius Dupuis, Martin Strobl, and Hans Grezlikowski. Opendrive 2010 and beyond–status and future of the de facto standard for the description of road networks. In *Proc. of the Driving Simulation Conference Europe*, pages 231–242, 2010.

[21] Sebastian Maierhofer, Moritz Klischat, and Matthias Althoff. Commonroad scenario designer: An open-source toolbox for map conversion and scenario creation for autonomous vehicles. In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pages 3176–3182. IEEE, 2021.

[22] Tingru Zhang, Da Tao, Xingda Qu, Xiaoyan Zhang, Rui Lin, and Wei Zhang. The roles of initial trust and perceived risk in public's acceptance of automated vehicles. *Transportation research part C: emerging technologies*, 98:207–220, 2019.

[23] Bcs code of conduct. Accessed on April 4, 2023.

[24] Safety pool™ scenario database, 2022. Accessed on April 4, 2023.

[25] Lucas Caton. code2pdf - a simple tool for convert your source code to a pdf file, 2011. Accessed on April 4, 2023.

Appendix A

Extra Information

This appendix contains tables, various images of files, and maps to better illustrate the topics discussed throughout the report.

Attribute	Min Value	Max Value	Rounding Value
cloudiness	0	100	3
precipitation	0	100	3
precipitation deposits	0	100	3
wind intensity	0	100	3
sun azimuth angle	0	360	3
sun altitude angle	-90	90	3
fog density	0	100	3
fog distance	0	$+\infty$	3
wetness	0	100	3
fog falloff	0	$+\infty$	3
scattering intensity	0	$+\infty$	3
mie scattering scale	0	$+\infty$	3
rayleigh scattering scale	0	$+\infty$	3
dust storm	0	100	3

Table A.1: Weather related scenario attributes

Attribute	Min Value	Max Value	Rounding Value
number of pedestrians	0	$+\infty$	0
number of vehicles	0	$+\infty$	0
number of two wheel vehicles	0	$+\infty$	0
proportion of speeding vehicles	0	1	3
proportion of vehicles without lights	0	1	3
proportion of light ignoring vehicles	0	1	3
light ignoring percent	0	100	3
proportion of sign ignoring vehicles	0	1	3
sign ignoring percent	0	100	3
proportion of vehicle ignoring vehicles	0	1	3
vehicle ignoring percent	0	100	3
proportion of walker ignoring vehicles	0	1	3
walker ignoring percent	0	100	3
proportion of keeping right vehicles	0	1	3
keeping right percent	0	100	3
proportion of lane changing vehicles	0	1	3
lane change percent	0	100	3
proportion of misbehaving pedestrians	0	1	3
proportion of running pedestrians	0	1	3
proportion of road crossing pedestrians	0	1	3
number of junctions	0	$+\infty$	0
distance in metres	0	$+\infty$	3
area in square metres	0	$+\infty$	3
difficulty	0	1000	3

Table A.2: Traffic related scenario attributes

Category	Not speeding	Speeding
No beams and no fog lights	50	50
No beams	30	30
No fog lights	10	10
Hit pedestrian	600	1200
Hit vehicle	250	500
Hit two-wheel vehicle	400	800
Hit road object	150	300
Red light violation	50	100
Stop sign/marking violation	40	80
Solid lane marking	20	60
Double solid lane marking	40	100
Broken lane marking no turn indicator	10	30
Light speeding	1	1
Heavy speeding	3s	3

Table A.3: Penalty score table

```
{
    "number_of_pedestrians": 17,
    "number_of_vehicles": 47,
    "number_of_two_wheel_vehicles": 8,
    "proportion_of_speeding_vehicles": 0.22,
    "proportion_of_vehicles_without_lights": 0.24,
    "proportion_of_light_ignoring_vehicles": 0.2,
    "light_ignoring_percent": 26.569,
    "proportion_of_sign_ignoring_vehicles": 0.187,
    "sign_ignoring_percent": 28.303,
    "proportion_of_vehicle_ignoring_vehicles": 0.216,
    "vehicle_ignoring_percent": 15.7,
    "proportion_of_walker_ignoring_vehicles": 0.253,
    "walker_ignoring_percent": 16.869,
    "proportion_of_keeping_right_vehicles": 0.187,
    "keeping_right_percent": 31.601,
    "proportion_of_lane_changing_vehicles": 0.288,
    "lane_change_percent": 24.191,
    "proportion_of_misbehaving_pedestrians": 0.251,
    "proportion_of_running_pedestrians": 0.284,
    "proportion_of_road_crossing_pedestrians": 0.328,
    "number_of_junctions": 8,
    "distance_in_metres": 802.702,
    "area_in_square_metres": 3375170.561,
    "cloudiness": 20.472,
    "precipitation": 2.689,
    "precipitation_deposits": 0.0,
    "wind_intensity": 24.422,
    "sun_azimuth_angle": 313.767,
    "fog_density": 5.306,
    "fog_distance": 0.0,
    "fog_falloff": 0.101,
    "scattering_intensity": 1.047,
    "mie_scattering_scale": 0.03,
    "rayleigh_scattering_scale": 0.038,
    "dust_storm": 0.0,
    "difficulty": 300.0,
    "wetness": 0.1,
    "sun_altitude_angle": 67.0
}
```


Figure A.1: Example of a generated scenario

```
{
  "path_checkpoints": [
    {
      "x": 112.867,
      "y": -14.310,
      "z": 0.400
    },
    {
      "x": -52.180,
      "y": -6.970,
      "z": 0.400
    },
    {
      "x": 40.692,
      "y": 51.025,
      "z": 0.400
    },
    {
      "x": 43.639,
      "y": 130.790,
      "z": 0.400
    },
    {
      "x": -103.212,
      "y": -14.857,
      "z": 0.400
    }
  ],
  "start_location": {
      "x": 112.867,
      "y": -14.310,
      "z": 0.400,
      "yaw": -90
  },
  "town": "Town10HD_Opt"
}
```


Figure A.2: Example of a generated path

```xml
<?xml version="1.0" ?>
<CollisionData>
  <PenaltyPointsTotal>700</PenaltyPointsTotal>
  <NumberOfCollisionInstances>3</NumberOfCollisionInstances>
  <CollisionInstances>
    <Instance>
      <PenaltyPoints>250</PenaltyPoints>
      <Description>Hit vehicle</Description>
      <CollisionIntensity>12786.265512536726</CollisionIntensity>
      <Time>37.281s</Time>
      <Location>
        <X>-50.800228</X>
        <Y>16.336308</Y>
        <Z>-0.000915</Z>
      </Location>
    </Instance>
    <Instance>
      <PenaltyPoints>300</PenaltyPoints>
      <Description>Hit road object speeding</Description>
      <CollisionIntensity>3978.5656853911505</CollisionIntensity>
      <Time>53.346s</Time>
      <Location>
        <X>-18.755812</X>
        <Y>23.543898</Y>
        <Z>0.027082</Z>
      </Location>
    </Instance>
    <Instance>
      <PenaltyPoints>150</PenaltyPoints>
      <Description>Hit road object</Description>
      <CollisionIntensity>20890.64664324663</CollisionIntensity>
      <Time>67.504s</Time>
      <Location>
        <X>31.434811</X>
        <Y>59.229317</Y>
        <Z>0.154669</Z>
      </Location>
    </Instance>
  </CollisionInstances>
</CollisionData>

```

Figure A.3: An example of collision_data.xml file structure

```json
{
    "participants": [
        "participant_A",
        "participant_B",
        "participant_C",
        "participant_D",
        "participant_E",
        "participant_CARLA"
    ],
    "scenarios": [
        "scenario1"
    ],
    "metrics": [
        "collision_data",
        "lane_marking_violation_data",
        "vehicle_light_misuse_data",
        "route_data",
        "speeding_data",
        "road_traffic_violation_data"
    ]
}
```

Figure A.4: An example of analysis_parameters.json file

```json
{
    "scenarios": [
        {
            "name": "scenario1",
            "details": "path1",
            "vehicle": "vehicle.mercedes.coupe_2020",
            "randomization_seed": 74111222
        },
        {
            "name": "scenario2",
            "details": "path2",
            "vehicle": "vehicle.mercedes.coupe_2020",
            "randomization_seed": 23115235
        },
        {
            "name": "scenario3",
            "details": "path3",
            "vehicle": "vehicle.mercedes.coupe_2020",
            "randomization_seed": 57894561
        },
        {
            "name": "scenario4",
            "details": "path4",
            "vehicle": "vehicle.mercedes.coupe_2020",
            "randomization_seed": 34586451
        },
        {
            "name": "scenario5",
            "details": "path5",
            "vehicle": "vehicle.mercedes.coupe_2020",
            "randomization_seed": 75861238
        }
    ]
}
```


Figure A.5: An example of the scenario list

OS: Ubuntu 20.04.4 LTS
CPU: 11th Gen Intel Core i7-11700KF @ 3.60GHz × 16
RAM: 64GB
GPU: NVIDIA GeForce RTX 3090

Figure A.6: Technical specifications of the lab computer

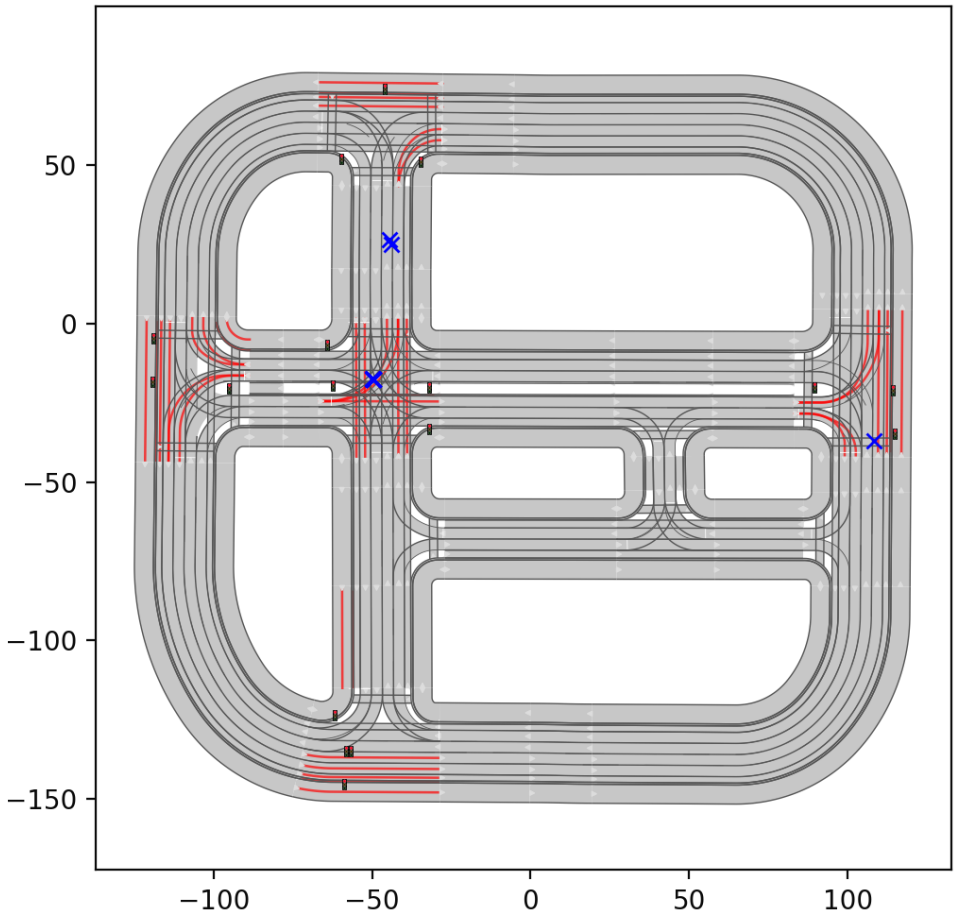

Figure A.7: Correctly marked collision locations

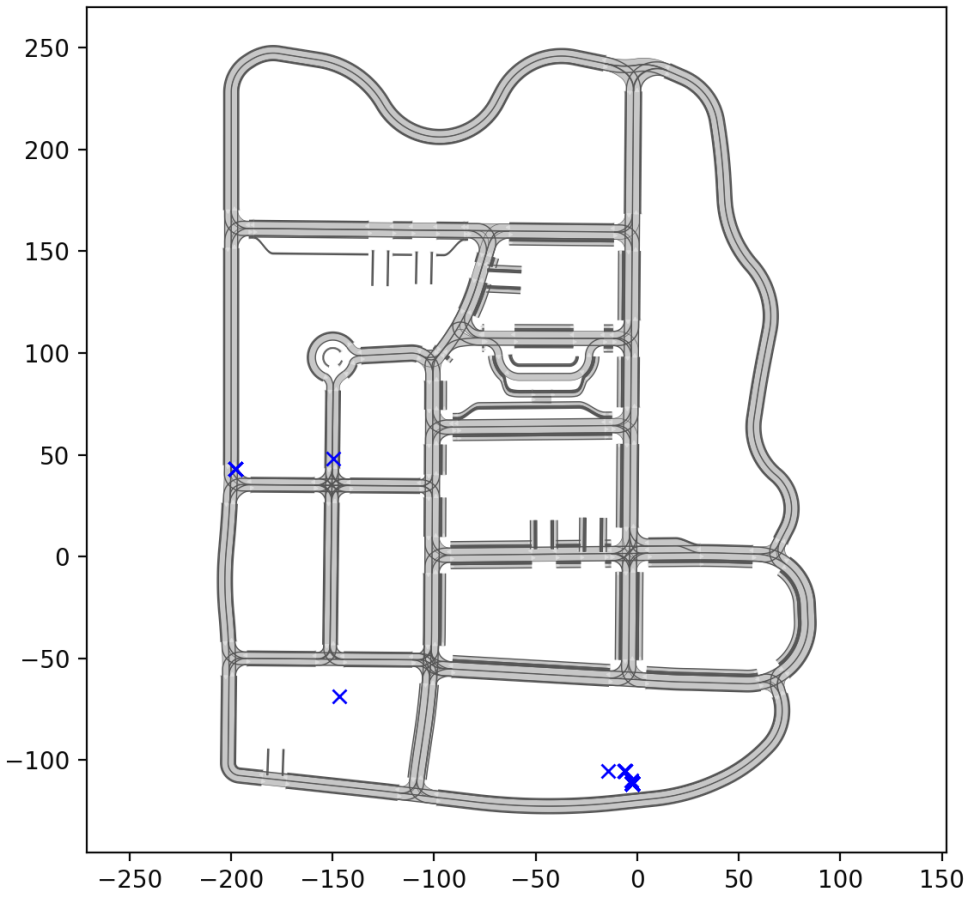

Figure A.8: Incorrectly marked collision locations

Appendix B

User Guide

The purpose of this chapter is to provide a user guide on how to use the software proposed in this article. Each section provides instructions for different subcomponents of the framework. Note that the required python modules must be installed before running any of the software. They can be found in the software/python_libraries/requirements.txt file. These dependencies can be installed by creating a virtual environment at the root of this project and running the command:

$$pip3\ install\ -r\ requirements.txt$$

Please note that most directories also contain README.md files with instructions and descriptions for using the software.

B.1 How to generate scenarios

This section gives a brief overview of how to generate the scenarios. The scenario generation script can be found in the software/generating_scenarios directory and is named **create_scenario.py**. It takes the following optional arguments:

- -d the difficulty of the scenario ($0 \leq d \leq 1000$)
- -s the sun altitude ($-90 \leq s \leq 90$), where 90 means that the sun is at its peak and -90 means that it is an absolute night
- -w the weather wetness level in the scenario ($0 \leq w \leq 100$)
- -f the name of the file to save the scenario to (without the .json ending)

- -r flag meaning that the algorithm should be trained using a randomisation seed (the same every time)

An example of the command in the terminal can be seen in the image below:

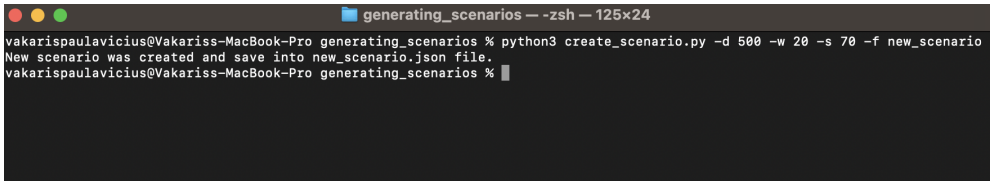

Figure B.1: How to generate a scenario

The following command would generate a new scenario in a new file **new_scenario.json** and save the file to the data/scenario_generation_data/generated_scenarios directory. The file can be later used to create a specific path, or it could be used as a predefined scenario for the simulations.

B.2 How to build paths

This section briefly overviews how a script that builds paths can be run. The script itself is located in the software/carla_scripts directory. Its name is **path_generator.py**. It takes the following mandatory arguments:

- -f the name of the file to save the built path to (without the .json ending)
- -s the name of the scenario to be referenced when designing the path (without the .json ending)

An example of the command in the terminal can be seen in the image below:

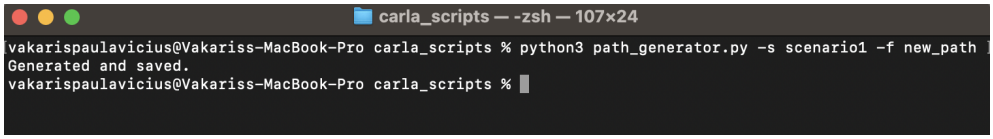

Figure B.2: How to build a path

The following command would build a new path in the file **new_path.json** and save the file to the data/paths directory. The file can be later used as a predefined path for scenarios or inspected on the map using helper scripts, for instance, draw_lanes_terminal.py.

B.3 How to run the simulations

This section presents a simple way of running the scenarios. At first, the user has to decide how many scenarios he wants the software to run and whether the scenarios should be predefined or created at run-time. If the user decides to use predefined scenarios, he must describe that in the scenario_list.json file in the software/carla_scripts directory. An example of the scenario list can be seen in Figure A.5. If the user wants the path and the scenario to be generated at run-time, the attributes "name" and "details" of each scenario element in the scenario list file should be set to *null*. The scenario simulation process requires the CARLA simulator to be running. Once the CARLA simulator is running and accessible, the user needs to navigate to the software/carla_scripts directory in the second terminal and run the following command to run the simulations with manual control:

$$/run.sh\ [name]$$

In order to change whether the steering wheel is to be used or the keyboard, the user needs to comment and uncomment lines 157 and 158 in scenario_manager.py, respectively.

The element [name] needs to be replaced by the name given to the participant. All the recorded data will be stored in the data/recordings/[name] directory, and the details of the simulations will be visible in the data/recordings/[name]/session_logs.log file. In order to launch the Carla agent, the following command has to be run:

$$/run.sh\ [name]\ ai$$

The keyword "ai" (at the end of the line) differentiates manually controlled and self-driving cars. **Important: if at any point the program crashes and the cause is unknown, the parts "> /dev/null 2 > &1" need to be commented out on lines 26 and 33 in the run.sh script in addition to line 160 in the scenario_manager.py script. This allows for message output to the terminal.** All the errors should be handled by the try-except blocks, and the appropriate messages should be written to the session_logs.log files. The steps mentioned before work as a plan b if something goes wrong.

B.4 How to perform analysis

Two types of analysis can be performed. Let us start with the analysis method meant to assist in the score calculation process called scenario analysis. The script can be found in the software/data_analysis directory and is called scenario_analyser.py. It aims to analyse both the

scenario and the path and extract information such as the average allowed speed on the path, the number of pedestrian crossings and other. Note that in order to run this script, the CARLA simulator must be running. This is needed because the script needs to use the built-in functions of CARLA and access the map attributes. The script takes two mandatory arguments:

- -p the name of the path (without the .json ending)
- -s the name of the scenario (without the .json ending)

The following command is used to perform the scenario analysis:

$$python3\ scenario\_analyser.py\ -s\ scenario1\ -p\ path1$$

As a result, it produces a file named the same as the scenario given as the argument. It saves this file into the data/simulation_details directory. This file is later used by the score calculator discussed in Section B.6

The second analysis tool is responsible for analysing the recordings of the scenario runs. It synthesises and processes information that is later used by the visualisation tools. The data analysis tool can be found in the software/data_analysis directory and is named **run_analysis.sh**. It takes a name as an argument and processes the recordings according to the specifications in the analysis_parameters.json file. How the parameters file looks can be seen in Figure A.4. It creates a new directory in data/analysed_data/data with the name provided as the argument and stores the analysed data there.

The following command can be used to run an example analysis:

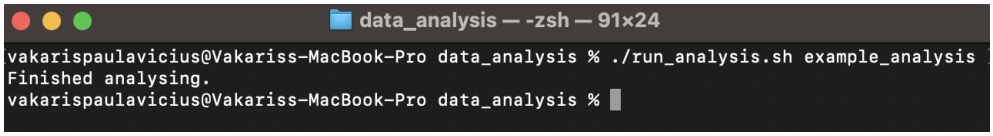

Figure B.3: How to perform data analysis

B.5 How to draw diagrams and maps

This section discusses how to visualise the analysed data. The visualisations can be done in two ways: drawing the diagrams or visualising interesting points on the map. Let us look at the diagram drawing first. The diagram drawing script can be found in the software/-data_analysis/visualisation/diagram_drawing directory. In order to make matters easier, an example file is provided in the directory, and it is called draw_diagram_example.sh. The comments in the script show which places should be edited to draw different diagrams. More about

the customisations of it was discussed in Section 7.2 when we talked about how the visualisation of data is performed. The following command would produce an image from Figure 7.3.

./draw_diagram_example.sh

For the map visualisation, the crdesgner module is required. If all the modules from the requirements.txt file were installed successfully, as mentioned at the beginning of this chapter, then it should already be present in the system. However, this project modified the module to better fit the project's needs. In order to perform the map visualisation, the modified version from software/python_libraries/crdesigner needs to be used. Simply replace the package in the system with the one provided in this project's repository.

As with the diagram drawing, an example script is provided in the software/data_analysis/visualisation/map_drawing directory. Its name is draw_map_example.sh. In order to draw a different map or a different list of points, the script's code should be edited accordingly. The following command should produce a view similar to the one in Figure 7.4.

./draw_map_example.sh

Note that if maps other than the ones provided by CARLA are used in the simulations, their OpenDRIVE description files must be added to the data/maps/opendrive_format directory. For the CommonRoad Scenario Designer to depict them, they must be converted from OpenDRIVE to CommonRoad format. It is done by the xml_converter.py script present in the data/maps directory.

B.6 How to calculate the score

This section concludes the user guide by examining how the participant's performance in a specific scenario can be evaluated. It is done by the score_calculator.py script present in the software/data_analysis directory. The script takes two mandatory arguments:

- -n the name of the participant
- -s the name of the scenario (without the .json ending)

It then looks at the files in data/recordings/[participant's name]/[scenario], extracts the necessary information and calculates the score printing it to the terminal. The following is an example of how this is done:

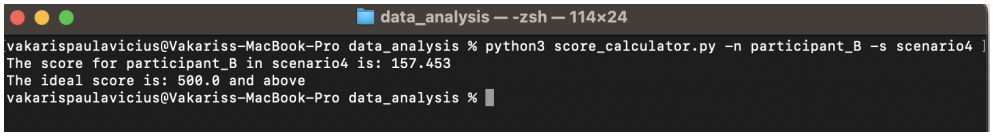

Figure B.4: How to evaluate the performance

In addition to calculating the score and printing it to the terminal, the script also provides the ideal score for that particular scenario. The scores are calculated based on the formula introduced in Section 6.3.

Appendix C

Issues

This appendix contains various images of files and terminal output to better illustrate the issues encountered. The following figures are referenced in the report.

```
<road name="Road 31" length="9.4456216548435208e+0" id="31" junction="-1">
    <link>
        <predecessor elementType="junction" elementId="1162"/>
        <successor elementType="road" elementId="41" contactPoint="end"/>
    </link>
```

Figure C.1: Road 31 pointing to road 41

```
<road name="Road 41" length="1.0000000000000000e+0" id="41" junction="-1">
    <link>
        <predecessor elementType="junction" elementId="798"/>
        <successor elementType="road" elementId="31" contactPoint="end"/>
    </link>
```

Figure C.2: Road 41 pointing back to road 31 creating a loop

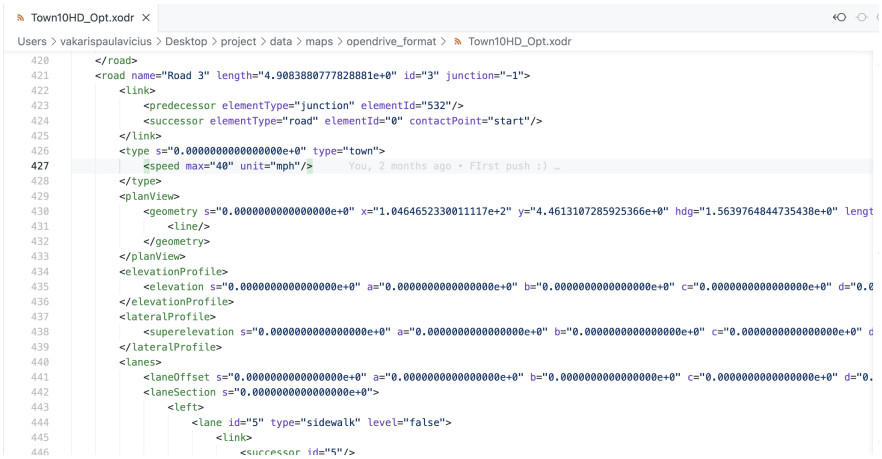

Figure C.3: The road element specifies the speed limit (Town10HD_Opt)

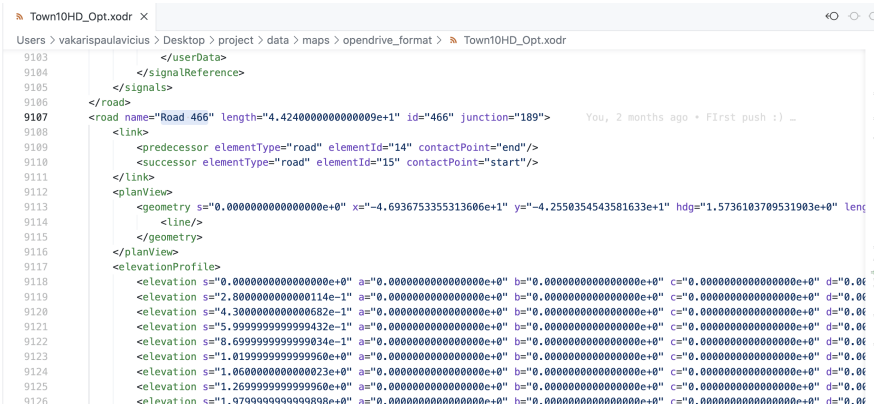

Figure C.4: The road element does not specify the speed limit (Town10HD_Opt)

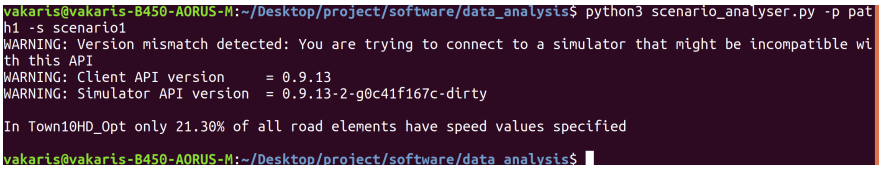

Figure C.5: The proportion of road elements indicating speed limit (Town10HD_Opt)

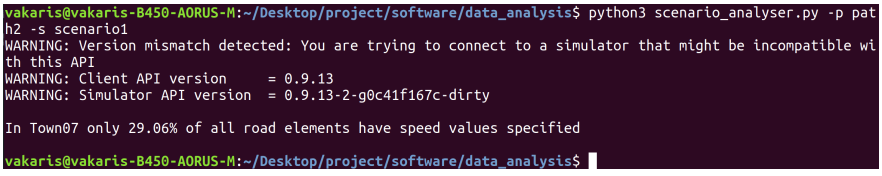

Figure C.6: The proportion of road elements indicating speed limit (Town07)