



# **Dokumentation Integrationsprojekt**

im Studiengang Information Engineering

von

**Studiengruppe HFP424**

**Prüfer: Prof. Dr. Joel Greenyer**

**Hannover, 4. September 2025**

# Erklärung der Selbstständigkeit

Hiermit versichere ich, dass ich die vorliegende selbständig verfasst und keine anderen als die angegebenen Quellen verwendet habe. Alle Textstellen und andere Inhalte wie Bilder, Quellcode oder Daten, die wörtlich oder sinngemäß aus anderen Quellen entnommen sind, habe ich eindeutig mit korrekten Quellenangaben versehen. Über Zitierrichtlinien bin ich schriftlich informiert worden. Diese Arbeit wurde bisher in gleicher oder ähnlicher Form, auch nicht in Teilen, keiner anderen Prüfungsbehörde vorgelegt.

Falls ich während der Erstellung dieser Arbeit generative KI oder andere KI-gestützte Technologien verwendet habe, die über grundlegende Werkzeuge für Übersetzungen und zur Überprüfung von Grammatik oder Rechtschreibung hinausgehen, erkläre ich, nach der Nutzung dieser Tools den Inhalt meiner überprüft und bearbeitet zu haben und übernehme die volle Verantwortung für die diesbezüglichen Ausführungen. Eine Verwendung solcher Werkzeuge habe ich in meiner Arbeit dokumentiert (bspw. im Abschnitt „Struktur der Arbeit, Methodik“ oder durch Erläuterungen zur Nutzung im Anhang).

Hannover, den 4. September 2025

---

Studiengruppe HFP424

# Inhaltsverzeichnis

<b>1</b>	<b>Motivation</b>	<b>1</b>
<b>2</b>	<b>Verwandte Arbeiten</b>	<b>3</b>
<b>3</b>	<b>Grundlagen</b>	<b>4</b>
3.1	Testumgebungen im autonomen Fahren . . . . .	4
3.2	Klassifikation von Testszenarien im autonomen Fahren . . . .	4
3.3	OpenSCENARIO als Standard zur Beschreibung von Test- szenarien . . . . .	4
3.4	Eingesetzte Simulations-Engines im Rahmen dieser Arbeit . .	4
3.4.1	HighwayEnv . . . . .	5
3.4.2	Sumo und Traci . . . . .	5
3.5	Behavioral Programming Paradigma . . . . .	6
3.5.1	Fokus auf Verhalten statt Kontrolle . . . . .	7
3.5.2	B-Threads und ereignisgesteuerte Synchronisation . .	7
3.5.3	Modularität und Erweiterbarkeit . . . . .	8
<b>4</b>	<b>Problemanalyse</b>	<b>9</b>
<b>5</b>	<b>Umsetzung</b>	<b>10</b>
5.1	Gesamt-Architektur . . . . .	10
5.2	Simulationsumgebungen . . . . .	11
5.2.1	HighwayEnv . . . . .	11
5.2.2	Sumo . . . . .	13
5.3	Modellierung von Szenarien mit BPPy . . . . .	15
5.4	Visualisierung von ausgeführten konkreten Szenarien . . . . .	16
5.5	Reinforcement Learning . . . . .	16
5.5.1	Trainieren des Modells . . . . .	18
5.5.2	Speichern des Modells . . . . .	18
5.5.3	Nutzen des Modells . . . . .	18
<b>6</b>	<b>Evaluierung</b>	<b>19</b>
<b>7</b>	<b>Fazit</b>	<b>20</b>

*INHALTSVERZEICHNIS*

iii

**8 Ausblick**

**21**

**A Anhang**

**22**

# Kapitel 1

## Motivation

Autonomes Fahren ist in der heutigen Zeit keine ferne Vision mehr, sondern bereits Realität. Unternehmen wie Waymo in San Francisco, Baidu in Wuhan oder Volkswagen in Hamburg zeigen, dass fahrerlose Fahrzeuge am Straßenverkehr teilnehmen können. Hierdurch wachsen jedoch auch die Herausforderungen, die Sicherheit solcher Systeme zu gewährleisten und dementsprechend zu validieren. Die Vielzahl potenzieller Verkehrssituationen ist praktisch unbegrenzt, von einfachen Überholmanövern bis hin zu komplexen Interaktionen mit Fußgängern, Radfahrern, Ampeln und mehreren Fahrzeugen gleichzeitig. Hinzu kommen länderspezifische Verkehrsregeln und Regularien, die eine vollständige Erprobung im Realverkehr nahezu unmöglich machen.

Simulationen stellen daher einen unverzichtbaren Baustein in der Validierung autonomer Fahrfunktionen dar, da reale Testfälle nicht in der notwendigen Breite und Tiefe skalierbar sind. Mit Standards wie ASAM OpenScenario existieren bereits Ansätze, Szenarien in strukturierter Form zu beschreiben. Allerdings sind die XML-basierten Szenarien sehr aufwändig zu modellieren und müssen für jedes Testziel individuell erstellt werden. Neuere domänenspezifische Sprachen, wie die auf Constraints basierende OpenScenario DSL, versprechen zwar leichtere Handhabung, sind aber noch nicht weit verbreitet und erfordern ebenfalls manuelle Anpassungen. Industrielle Lösungen wie Fortellix Foretify nutzen Constraint-Solver und Planungsalgorithmen, erweisen sich in der Praxis jedoch häufig als starr und schwer automatisierbar.

Vor diesem Hintergrund ist die Entwicklung neuer, flexiblerer Ansätze zu automatisierten Szenariengenerierung von zentraler Bedeutung. Behavioral Programming bietet hierfür eine attraktive Grundlage. Es erlaubt eine modulare Modellierung von Szenarien und bildet konkurrierende Verhaltensweisen auf natürliche Weise ab. Ergänzend eröffnet Reinforcement Learning die Möglichkeit, aus abstrakten Szenariobeschreibungen, etwa ein Auto überholt das VUT und bremst anschließend, konkrete, ausführbare

Szenarien zu generieren, die den abstrakten Vorgaben genügen.

Dieses Projekt verfolgt das Ziel, die Kombination von Behavioral Programming und Reinforcement Learning als neuartigen Ansatz zur automatisierten Szenariengenerierung zu untersuchen. Dabei sollen nicht nur Methoden zur Modellierung abstrakter und konkreter Szenarien entwickelt werden, sondern auch Verfahren, wie die ausgeführten Szenarien geloggt und visuell aufbereitet werden können, um Test-Engineers eine transparente Analyse zu ermöglichen. Vor diesem Hintergrund ergibt sich die zentrale Forschungsfrage:

**Wie kann BPPy zur Modellierung ausführbarer Testszenarien im autonomen Fahren eingesetzt werden ?**

Zur Beantwortung dieser Frage werden folgende Unterfragen untersucht:

- Wie können Szenarien in BPPy modelliert werden und welche grundsätzlichen Unterschiede bestehen im Vergleich zu OpenSCENARIO?
- Welche Möglichkeiten der Komposition bietet BPPy für Szenarien, etwa sequenziell oder parallel?
- Wie kann Reinforcement Learning eingesetzt werden, um einen Agenten zu trainieren, der in der Lage ist Testszenarien so zu steuern, dass diese abstrakten Szenarien genügen?
- Wie können ausgeführte Szenarien geloggt und visualisiert werden, um eine transparente Analyse zu ermöglichen?

Das Vorhaben versteht sich als Machbarkeitsstudie. Es geht nicht darum, ein industrietaugliches Framework zu entwickeln, sondern vielmehr um die grundlegende Frage, ob sich durch die Kombination von Behavioral Programming und Reinforcement Learning ein flexibler, automatisierbarer Ansatz zur Generierung konkreter Testszenarien aus abstrakten Vorgaben realisieren lässt.

## Kapitel 2

# Verwandte Arbeiten

# Kapitel 3

## Grundlagen

### 3.1 Testumgebungen im autonomen Fahren

- kurze Abgrenzung VIL, HIL, SIL
- zur Einordnung, welche unterschiedlichen Testebenen es gibt

### 3.2 Klassifikation von Testszenarien im autonomen Fahren

- Abgrenzung zwischen konkreten und abstrakten Szenarios
- am Beste mittels eines Beispiels z.B. Überhol-Manöver beschreiben

### 3.3 OpenSCENARIO als Standard zur Beschreibung von Testszenarien

- OpenScenario 1.x Standard (XML-basiert)
- OpenScenario DSL
- jeweils Beispiel zeigen, wie so etwas aussehen könnte
- Herausforderungen herausstellen

### 3.4 Eingesetzte Simulations-Engines im Rahmen dieser Arbeit

Simulationsumgebungen sind ein zentrales Werkzeug im Bereich des autonomen Fahrens. Sie ermöglichen es, Verkehrssituationen unter kontrollierten Bedingungen nachzustellen und die entwickelten Verfahren zu testen.



Abhängig vom jeweiligen Einsatzgebiet unterscheiden sich die Umgebungen vor allem im Grad der Abstraktion, in den bereitgestellten Funktionen und im benötigten Rechenaufwand. In den folgenden Abschnitten werden mit HighwayEnv sowie SUMO und dem dazugehörigen TraCI drei Umgebungen vorgestellt, die in unserem Projekt verwendet werden.

### 3.4.1 HighwayEnv

HighwayEnv ist eine Open-Source-Simulationsumgebung für autonomes Fahren, welches auf der Basis von Gymnasium beziehungsweise OpenAI Gym entwickelt wurde. Der Fokus liegt auf einer schnellen abstrahierten Simulation auf mehrspurigen Straßen und Autobahnen. Das Ziel dieser Simulationsumgebung ist die Entwicklung, Training und Evaluierung von Reinforcement-Learning-Algorithmen und Planungsverfahren im Bereich des autonomen Fahrens.

Technisch ist es mit Python implementiert und folgt einem modularen Aufbau, welche sich an der klassischen Agent-Environment-Schnittstelle aus dem Bereich des Reinforcement-Learnings bedient. Die einzelnen Szenarien werden durch Environment-Klassen definiert, etwas highway-v0 (ein klassisches Autobahn-Szenario) oder roundabout-v0 (ein Kreisverkehr). Damit eine effiziente Simulation stattfinden kann nutzt es mittels einem kinematischen Modells ein vereinfachte Fahrzeugdynamiken.

Bei der Fahrzeugmodellierung beschränkt sich HighwayEnv auf wesentliche Parameter, wie Position, Geschwindigkeit und Heading. Die Umgebungsmodellierung bietet mehrspurige Straßen, Fahrspurwechsel sowie unterschiedliche Verkehrsdichten. Ein Agent kann beschleunigen, bremsen, die Spur wechseln oder seine aktuelle Bewegung beibehalten. Die im Reinforcement-Learning benötigte Belohnungsfunktion kann je nach Zielstellung beliebig angepasst werden, z.B. auf Sicherheit oder Effizienz.

Der Fokus von HighwayEnv liegt vor allem auf abstrakten Szenarien und nicht hochrealistischen Simulationen. Es ist somit kein Ersatz für High-Fidelity-Simulatoren, wie CARLA. Zudem bietet es keine realistische Physik. Dadurch werden beispielsweise Fahrdynamiken oder Wetterbedingungen vernachlässigt. Der Vorteil von HighwayEnv liegt in den geringen Rechenkosten und ist somit geeignet für Experimente und Prototyping.

### 3.4.2 Sumo und Traci

Die Simulation of Urban MObility, kurz SUMO, ist ein Open-Source, mikroskopischer, straßenverkehrsbasierter Simulator, welcher vom Deutschen Zentrum für Luft- und Raumfahrt (DLR) entwickelt wird. Das Ziel ist es Verkehrsflüsse und -dynamiken in realistischen Straßennetzwerken zu simulieren. Es unterstützt Einzel- sowie Massensimulationen von Fahrzeugbewegungen in Städten, Autobahnen und Mischverkehr.

Die Implementierungssprache ist im Kern C++. Zudem stehen Python-APIs zur Verfügung mit denen Erweiterungen möglich sind. Mikroskopisch bedeutet in diesem Fall, dass jedes Fahrzeug einzeln simuliert wird. Die Straßennetze können aus realen Karten, wie OpenStreetMap importiert werden, sodass eine hohe Flexibilität geboten ist. Des Weiteren werden verschiedene Fahrzeugtypen von LKWs und Bussen über den PKW bis hin zum Fahrrad oder Fußgänger.

Für die Fahrdynamik werden Car-Following-Modelle wie das Krauß-Modell genutzt. Darüber hinaus finden Lane-Change-Modelle Anwendung. Innerhalb der Verkehrssteuerung können Ampeln, Zuflussregelungen oder auch Stauszenarien simuliert werden. Des weiteren gibt es eine eingebaute Routenplanung sodass Fahrzeuge entweder feste Routen folgen oder auf Basis von dynamischen Routing-Algorithmen gesteuert werden können. Im Bereich der Auswertungen werden verschiedene detaillierte Statistiken im Bezug auf Reisezeiten, Emissionen oder Staus geliefert.

Die Einsatzbereiche von SUMO sind reichen von der Verkehrsforschung, wobei Verkehrsflüsse, Engpässe oder Infrastrukturmaßnahmen untersucht werden bis hin zur Stadtplanung und Mobilitätskonzepten bei denen Szenarien wie Car-Sharing oder ÖPNV simuliert werden. Zudem kann es auch im Bereich des Autonomen Fahrens als Testumgebung für Entscheidungs- und Koordinationsstrategien dienen.

Im Gegensatz zu HighwayEnv benötigt SUMO bei sehr großen Netzen eine hohe Rechenlast. Es bietet zudem auch keine High-Fidelity, da auch hier ein vereinfachtes Fahrzeugmodell verwendet wird.

Mit dem Traffic Control Interface (TraCI) bietet das DLR eine Client-Server-Schnittstelle zur Echtzeitsteuerung und Abfrage von SUMO-Simulationen. Dies erlaubt die dynamische Interaktion mit einer laufenden SUMO-Simulation. Dabei kann ein lesender Zugriff zur Abfrage von Zuständen etwa der Position, Geschwindigkeit und Ampelphasen geschehen, aber auch eine Manipulation der Simulation indem beispielsweise Fahrzeuge hinzugefügt oder Routen verändert werden. Dies erlaubt die Einbindung von externen Steuerungsalgorithmen z.B. Reinforcement-Learning oder Verkehrsmanagementsysteme.

### 3.5 Behavioral Programming Paradigma

Behavioral Programming (BP) ist ein sprachunabhängiges Programmierparadigma zur Entwicklung reaktiver Systeme[HMW12]. Dabei steht die inkrementelle Spezifikation von Verhalten im Vordergrund, während klassische Kontrollflussentscheidungen in den Hintergrund treten[HMW12]. BP wurde ursprünglich im Kontext der szenarienbasierten Programmierung entwickelt und später auch in Umgebungen wie Java (als *BPJ*-Framework) und Python (als *BPpy*) umgesetzt[HMW10]. Die zentralen Bausteine dieses Paradigmas

sind sogenannte *Behavior Threads* (B-Threads), welche einzelne Verhaltensaspekte kapseln und nebenläufig ausgeführt werden[HMW10]. Alle B-Threads kommunizieren ausschließlich über *Ereignisse* und werden an definierten Synchronisationspunkten vom Laufzeitsystem koordiniert[HMW12].

### 3.5.1 Fokus auf Verhalten statt Kontrolle

In BP wird jeder Systemaspekt als separates Verhalten (in einem eigenen B-Thread) implementiert, anstatt einen zentralen Ablaufplan zu programmieren. Dadurch ergibt sich eine *indirekte Steuerung*: Anstatt dass ein Modul direkt ein anderes aufruft oder steuert, beeinflussen sich die Verhaltensmodule nur über das Anbieten oder Verhindern von Ereignissen[HMW12]. Dieses Prinzip erlaubt eine natürlichere Beschreibung von Anforderungen, das heißt es wird spezifiziert, was geschehen oder unterbleiben soll, ohne die Ausführung in Form von detailliertem Kontrollfluss explizit zu kodieren[HMW12]. Die Auswahl des nächsten Systemschritts (Ereignisses) erfolgt zur Laufzeit durch das BP-Laufzeitsystem, das aus den von allen B-Threads vorgeschlagenen und gegebenenfalls gesperrten Ereignissen ein zulässiges Ereignis bestimmt und ausführt[HMW12]. BP verlagert den Schwerpunkt der Programmierung somit von der Implementierung von Steuerungslogik hin zur Spezifikation von erwünschtem oder verbotenen Verhalten.

### 3.5.2 B-Threads und ereignisgesteuerte Synchronisation

Ein B-Thread ist eine in sich geschlossene Verhaltenskomponente, typischerweise implementiert als unabhängiger Prozess (Thread, Coroutine o. Ä.), der wiederholt Ereignisse abgibt oder auf Ereignisse wartet. Die Koordination aller B-Threads erfolgt an expliziten *Synchronisationspunkten*. Jeder B-Thread ruft eine Synchronisationsoperation (etwa *sync*) auf, wobei er drei Kategorien von Ereignissen deklariert[HMW10]:

- **Gewünschte Ereignisse** (*requested events*): Ereignisse, die der Thread vorschlägt und bei deren Auftreten er fortgesetzt werden möchte.
- **Erwartete Ereignisse** (*waited-for events*): Ereignisse, auf die der Thread passiv wartet.
- **Blockierte Ereignisse** (*blocked events*): Ereignisse, die der Thread momentan verhindern möchte.

Sobald alle aktiven B-Threads ihren Synchronisationspunkt erreicht haben, sammelt das Laufzeitsystem alle vorgeschlagenen Ereignisse und wählt eines davon aus, das von mindestens einem B-Thread vorgeschlagen und von keinem B-Thread blockiert wurde[HMW12]. Das ausgewählte

Ereignis wird anschließend ausgelöst; alle B-Threads, die dieses Ereignis vorgeschlagen oder darauf gewartet haben, werden aufgeweckt und setzen ihre Ausführung bis zum nächsten Synchronisationspunkt fort[HMW12]. B-Threads, die das Ereignis weder angefordert noch erwartet hatten, bleiben solange pausiert[HMW12]. Falls mehrere vorgeschlagene Ereignisse gleichzeitig ausführbar sind, kann je nach Implementierung ein Prioritätenschema oder eine feste Rangfolge zur Anwendung kommen, um ein deterministisches Verhalten zu gewährleisten[HMW12]. Anschließend erfolgt die nächste Synchronisation aller Threads, und der Zyklus wiederholt sich. Dieses ereignisgesteuerte Synchronisationsverfahren stellt sicher, dass sich unabhängige Teilverhalten konsistent zu einem Gesamtablauf verweben, ohne dass eine zentrale Steuerlogik die einzelnen Schritte vorgibt.

### 3.5.3 Modularität und Erweiterbarkeit

Ein Hauptvorteil von BP ist die *hohe Modularität* der Verhaltensbeschreibung und die einfache Erweiterbarkeit von Systemfunktionalität. Neue Anforderungen können als zusätzliche B-Threads implementiert werden, was eine *inkrementelle Entwicklung* ermöglicht, bei der bestehender Code kaum oder gar nicht geändert werden muss[HMW12].

Das Hinzufügen oder Entfernen von Verhaltensmodulen entspricht dem Ein- oder Ausschalten von Szenarien, ohne unbeabsichtigte Seiteneffekte auf den übrigen Ablauf.

Alle Module interagieren ausschließlich über das gemeinsame Ereignis-Vokabular und kennen einander nicht direkt[HMW10]. So kann beispielsweise in einem einfachen Wassermisch-System ein zusätzliches Verhalten (Interleave, abwechselndes Heiß- und Kalt-Zapfen) durch einen neuen B-Thread realisiert werden, der abwechselnd auf das eine Ereignis wartet und das andere blockiert, ohne die bestehenden B-Threads zu verändern[HMW12].

BP unterstützt damit auch Konzepte wie die merkmalsorientierte Entwicklung (Feature-Oriented Development) und Produktlinien: Unterschiedliche Ausprägungen eines Systems können durch Konfiguration der Menge der aktiven B-Threads erzeugt werden[HMW12]. Im Vergleich zu objektorientierten oder aspektorientierten Ansätzen vermeidet BP die strikte Trennung zwischen Basis- und Erweiterungslogik; stattdessen wirken alle Verhaltensmodule gleichberechtigt zusammen[HMW12]. Darüber hinaus lässt sich BP mit anderen Programmierparadigmen kombinieren. B-Threads teilen keinen gemeinsamen Zustand, sondern koordinieren sich ausschließlich indirekt über Ereignisse, was eine lose Kopplung impliziert[HMW12].

## Kapitel 4

# Problemanalyse

# Kapitel 5

## Umsetzung

### 5.1 Gesamt-Architektur

Die Architektur besteht aus drei Hauptkomponenten: den Simulations-Engines, der Logging- und Visualisierungskomponente sowie dem BProgram, in dem Szenarien modelliert und ausgeführt werden können. Das BProgram kann zur Ausführung unabhängig zwischen den beiden bereitgestellten Simulations-Engines wählen. Innerhalb des Projekts wurden zunächst konkrete Szenarios manuell implementiert. Im weiteren Verlauf wurde ein Ansatz entwickelt, um konkrete Szenarien automatisiert zu erlernen, sodass diese einem abstrakt definierten Szenario folgen. In den konkreten Szenarien kommt zudem ein Logger zum Einsatz, der Daten für die Visualisierung bereitstellt. Diese Gesamtarchitektur ist in Abbildung 5.1 dargestellt.

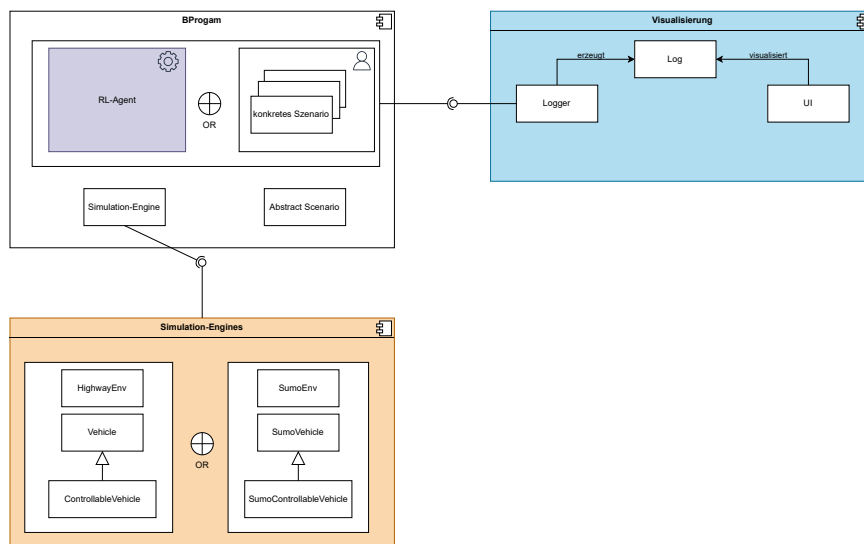


Abbildung 5.1: Gesamtarchitektur

Im Folgenden werden zunächst die beiden Simulations-Engines sowie die Hilfsklassen vorgestellt, die den Zugriff auf Fahrzeuginformationen ermöglichen. Anschließend wird die Modellierung in BPPy erläutert, die verschiedene B-Threads für die Ausführung der Simulation, abstrakten Szenarien und konkreten Szenarien nutzt. Darauf folgt eine Beschreibung der Logging- und Visualisierungskomponente. Abschließend wird der Ansatz zum automatisierten Erlernen konkreter Szenarien auf Basis eines abstrakten Szenarios vorgestellt.

## 5.2 Simulationsumgebungen

In diesem Abschnitt wird die Nutzung von HighwayEnv und SUMO sowie dessen Steuerungsschnittstelle TraCI beschrieben. Zunächst wird die Konfiguration von HighwayEnv beschrieben sowie zwei Ansätze zur Extraktion von Fahrzeuginformationen aus der Umgebung. Danach wird auf die Sumo-Verwendung innerhalb des Projektes eingegangen.

### 5.2.1 HighwayEnv

#### Konfiguration von HighwayEnv

In unserem Projekt kommt die von HighwayEnv bereitgestellte Umgebung `highway-v0` zum Einsatz. Diese simuliert eine mehrspurige Fahrbahn und kann konzeptionell als Autobahnumgebung verstanden werden.

Damit mehrere Fahrzeuge gleichzeitig steuerbar sind, muss die gewünschte Anzahl an kontrollierbaren Fahrzeugen über den Parameter `controlled_vehicles` spezifiziert werden. Zusätzlich lassen sich weitere, nicht kontrollierte Fahrzeuge über den Parameter `vehicles_count` in die Simulation integrieren. Ein spezielles VUT kann somit ebenfalls über diesen Mechanismus berücksichtigt werden.

Da unser Projekt in einem Multi-Agent-Setting ausgeführt wird, ist eine entsprechende Konfiguration notwendig. Hierfür müssen sowohl der Action-Typ (`MultiAgentAction`) als auch der Observation-Typ (`MultiAgentObservation`) gesetzt werden. Auf diese Weise wird sichergestellt, dass:

1. mehrere Agenten gleichzeitig Aktionen ausführen können und
2. die Beobachtungsdaten (`Observation`) auch für alle kontrollierten Fahrzeuge bereitgestellt werden.

#### Ansätze zum Zugriff auf Fahrzeuginformationen

Um in HighwayEnv Entscheidungen zu treffen, etwa hinsichtlich Abständen zwischen Fahrzeugen, Spurwahl oder Spurwechseln, ist eine geeignete Repräsentation von Fahrzeuginformationen erforderlich. In diesem

Kapitel werden zwei unterschiedliche Ansätze vorgestellt. Zunächst der `ObservationWrapper`, anschließend ein Vehicle-basiertes Modell. Abschließend erfolgt ein Vergleich der beiden Methoden.

Der `ObservationWrapper` kapselt die von der Environment zurückgegebene Observation und interpretiert deren Werte fachlich. Zu Beginn einer Simulation wird dieser einmalig initialisiert und nach jedem Step mit den aktuellen Daten aktualisiert. Auf Basis dieser Struktur können Methoden, wie `is_right_lane_` oder `get_distance_to_leading_vehicle` genutzt werden, um Abstände und Spurverfügbarkeit zu prüfen.

Ein zentrales Kriterium für Korrektheit ist dabei die Reihenfolge der Features innerhalb der Observation. In der aktuellen Implementierung müssen die Werte  $[\ddot{x}; \ddot{y}; \ddot{v}_x; \ddot{v}_y]$  genau in dieser Abfolge am Anfang des Arrays stehen. Nur dann können Distanz- und Geschwindigkeitsberechnungen valide durchgeführt werden. Darüber hinaus ist der Zugriff auf das Environment selbst notwendig, da bestimmte Informationen, etwa die Struktur des Straßennetzes, nicht allein aus der Observation extrahiert werden können. So lassen sich etwa Lanes über die Road-Network-Struktur bestimmen, was zusätzliche Abfragen an die Environment erfordert.

Damit wird eine relativ direkte Möglichkeit geboten mit den von HighwayEnv bereitgestellten Daten zu arbeiten. Gleichzeitig macht ihn die starke Abhängigkeit von der Feature-Reihenfolge und der Environment-Konfiguration anfälliger für Fehler.

Der zweite Ansatz modelliert Fahrzeuge explizit als Objekte. Dazu stehen die Klassen `Vehicle` und `ControllableVehicle` zur Verfügung.

1. `Vehicle` repräsentiert ein nicht kontrolliertes Fahrzeug und kapselt Methoden, die Informationen direkt aus der Simulation extrahieren. Dazu gehören unter anderem `speed()`, `lane_index()` sowie Distanzwerte wie `delta_pos` oder `delta_x_pos()`. Zusätzlich lassen sich Relationen wie `is_ahead_of()` oder `is_behind` zwischen zwei Fahrzeugen prüfen.
2. `ControllableVehicle` erweitert die Funktionalität um konkrete Aktionsmöglichkeiten. Ein solches Fahrzeug kann über Methoden wie `LANE_LEFT`, `LANE_RIGHT`, `FASTER`, `SLOWER` oder `IDLE` direkt gesteuert werden. Damit bildet es die Schnittstelle zwischen Entscheidungslogik und HighwayEnv.

Die Objekte werden beim Start der Simulation einmalig erzeugt und behalten anschließend den Zugriff auf die von ihnen zugeordneten Fahrzeuge. Eine explizite Aktualisierung interner Variablen ist nicht erforderlich, da die Methoden unmittelbar auf die jeweils aktuellen Daten der Environment zugreifen.

Im direkten Vergleich erweist sich der Vehicle-basierte Ansatz als robuster und flexibler, da die Objekte nur einmalig instanziiert werden



und ihre Methoden jederzeit aktuelle Daten liefern, entfällt die Gefahr inkonsistenter Zustände. Der `ObservationWrapper` hingegen ist stärker fehleranfällig denn er ist auf die exakte Reihenfolge der Features in der Observation angewiesen und bei Änderungen der Konfiguration können ungültige Ergebnisse geliefert werden.

### 5.2.2 Sumo

Im Verlauf des Projektes stellte sich heraus, dass die Definition eines eigenen Straßennetzes in `HighwayEnv` immer selbst programmiert werden muss und immer eine eigene Environment benötigt. Daher wurde sich nach Alternativen umgeschaut, welche uns auf die Verkehrssimulation SUMO brachte. Hierbei können individuelle Karten in SUMO-netedit leicht erstellt werden. Zudem bietet dieses Programm eine Möglichkeit der Erzeugung von Straßennetzen auf der Basis von echten Straßen.

#### SumoEnv

Im Gegensatz zu `HighwayEnv` bietet SUMO keine direkte Implementierung einer `gym.Env`-Umgebung. Um dennoch Reinforcement-Learning-Experimente in SUMO durchführen zu können wurde eine eigene Umgebung in der Klasse `SumoEnv` entwickelt. Diese basiert auf dem Gymnasium-Framework und bildet die standardisierten Methoden `reset`, `step`, `close` sowie die Definition von Beobachtungs und Aktionsräumen ab. Voraussetzung für die Nutzung ist die Installation von SUMO sowie das Python-Package TraCI, das eine bidirektionale Steuerung der SUMO-Simulation erlaubt.

Der Konstruktor übernimmt die Initialisierung der Umgebung. Neben der Referenz auf die SUMO-Konfigurationsdatei werden steuerbare Fahrzeuge als Parameter übergeben. Zudem werden die für das Reinforcement Learning relevanten Schnittstellen definiert:

- **Beobachtungsraum:** Als `spaces.Box` modelliert, repräsentiert er kontinuierliche Werte und ist in der aktuellen Version als Matrix der Dimension (20, 5) umgesetzt.
- **Aktionsraum:** Über `spaces.MultiDiscrete` werden diskrete Handlungsoptionen für jedes steuerbare Fahrzeug definiert. Dies ermöglicht eine flexible Steuerung mehrerer Agenten gleichzeitig.

Zur Verwaltung des Simulationsverlaufs werden Episoden- und Schrittzähler geführt (`episode`, `step_count`, `max_steps`).

Zum Starten der Simulation wird in `start_simulation` zunächst geprüft, ob die Umgebungsvariable `SUMO_HOME` gesetzt ist, um die ausführbare SUMO-Datei korrekt zu lokalisieren. Die Verbindung zu SUMO wird anschließend über TraCI hergestellt. Danach werden Parameter zur Zeitschrittlänge, Kollisionsbehandlung und Startoptionen gesetzt. In der

GUI werden zudem Kameraeinstellungen angepasst, um die Simulation bestmöglich visuell nachvollziehen zu können.

Das Hinzufügen von Fahrzeugen übernimmt `build_vehicle`. Neben einem vordefinierten Vehicle unter Test (`vut`) werden alle übergebenen steuerbaren Fahrzeuge erzeugt und mit individuellen Eigenschaften wie Startzeit, Fahrspur, Geschwindigkeit, Route oder Fahrverhalten können dabei individuell bestimmt werden. Besonders hervorzuheben ist, dass der Spurwechsel- und Geschwindigkeitsmodus der kontrollierbaren Fahrzeuge so gesetzt wurde, dass möglichst wenige Sicherheitsfeatures von SUMO aktiv sind. Auf diese Weise greifen SUMO-interne Automatikmechanismen, die Fahrmanöver einschränken oder korrigieren, nur minimal ein.

Ein Neustart der Umgebung erfolgt über `reset`. Falls bereits eine Simulation aktiv ist, wird diese zunächst beendet (`close`). Danach erfolgt ein Neustart der SUMO-Instanz sowie die Einbindung der Fahrzeuge. Eine kurze Verzögerung sorgt dafür, dass die Setup-Phase abgeschlossen ist, bevor auf die Fahrzeuge zugegriffen wird, um Fehler bei der Ausführung zu vermeiden.

Innerhalb der `step`-Methode wird der eigentliche Iterationszyklus einer Episode ausgeführt. Dieser besteht aus:

1. Anwendung der Agentenaktionen
2. Fortsetzung der Simulation um einen Zeitschritt
3. Erhebung der Beobachtung
4. Berechnung der Belohnung
5. Prüfung des Episodenendes

Das Rückgabeformat entspricht dem Gymnasium-Standard: (`obs`, `reward`, `terminated`, `truncated`, `info`).

Die Aktionslogik (`_apply_action`) implementiert die fünf aus HighwayEnv bekannten diskreten Handlungsoptionen: `LANE_RIGHT`, `IDLE`, `LANE_LEFT`, `FASTER` und `SLOWER`. Die Beobachtungslogik sammelt hierbei für jedes Fahrzeug relevante Zustandsgrößen wie Geschwindigkeit und Position. Fehlerfälle (etwa aus der Simulation entfernte Fahrzeuge) werden abgefangen, indem Default-Werte zurückgegeben werden.

Die Belohnungsfunktion ist aktuell als Platzhalter implementiert und gibt einen konstanten Wert zurück. Damit ist die Umgebung lauffähig, ohne dass bereits eine spezifische Optimierungslogik definiert werden muss. In unserem Anwendungsfall liegt der Fokus ohne hin nicht auf der Optimierung der Simulation, sondern auf der Steuerung der Fahrzeug, die später einem abstrakten Szenario folgen sollen. Weitere Platzhalter-Methoden wie `_get_info`, `render` oder `_render_frame` sind vorgesehen, um die Funktionalität um detaillierte Diagnoseinformationen oder zusätzliche

Visualisierungen zu erweitern. Innerhalb unserer Experimente war dies bisher nicht notwendig.

Die Methode (close) beendet die TraCI-Verbindung und sorgt für einen sauberen Abschluss der Simulation. Dies ist insbesondere für die korrekte Durchführung mehrerer Episoden im Training notwendig.

### Abbildung der Fahrzeugeigenschaften

Wie bereits in Abschnitt 5.2.1 gezeigt, hat sich der Vehicle-basierte Ansatz als robust und flexibel herausgestellt. Daher wurde er auch im Kontext von SUMO übernommen.

Zu diesem Zweck wurden die Klassen `SumoVehicle` und `SumoControllableVehicle` implementiert. `SumoVehicle` repräsentiert dabei nicht kontrollierte Fahrzeuge und stellt Methoden zur Verfügung, um relevante Fahrzeuginformationen direkt aus der SUMO-Simulation zu extrahieren. `SumoControllableVehicle` erweitert diese Funktionalität um konkrete Aktionsmöglichkeiten, die analog zur `ControllableVehicle`-Klasse im Kontext von HighwayEnv implementiert sind.

Durch diese Struktur müssen bestehende Szenario-Implementierungen lediglich einmalig auf die neue Simulations-Engine und die entsprechenden Fahrzeugobjekte angepasst werden. Die eigentliche Steuerungslogik kann unverändert übernommen werden. Gleichzeitig ermöglicht dieser Ansatz einen schnellen Wechsel der Simulations-Engine, sodass das Experiment-Setup nicht an eine spezifische Plattform gebunden ist und flexibel zwischen HighwayEnv und SUMO oder anderen kompatiblen Umgebungen adaptiert werden kann.

## 5.3 Modellierung von Szenarien mit BPpy

Für die Modellierung von Szenarien wurde BPpy genutzt. Hierbei wurden verschiedene BThreads erstellt, die unterschiedliche Aspekte der Szenarien abbilden. Einzelne BThreads bilden bestimmte Funktionalitäten und Teilszenarien ab, die dann in Kombination ein vollständiges, konkretes Szenario ergeben. So können die Funktionalitäten insgesamt besser strukturiert und wiederverwendet werden. Die BThreads werden aber außer für einzelne Funktionalitäten auch für die Simulation und Abstraktion von Szenarien genutzt, sowie gebündelt um ein konkretes Szenario wie ein Überholmanöver abzubilden. So gibt es BThreads, die ein abstraktes Szenario modellieren, also die grundlegende Anforderungen an das Szenario definieren, ohne sich auf eine konkrete Implementierung festzulegen. Diese abstrakten Szenarien werden dann im BProgram mit ausgeführt um zu überprüfen, dass die Anforderungen an das Szenario erfüllt werden. Dabei werden konkrete Angaben überprüft, wie z.B. das sich bestimmte Fahrzeugpositionen verändert haben, oder in welcher Fahrbahn sich die Fahrzeuge im Vergleich zueinander

befinden. Dazu kann dann geloggt werden, ob die Anforderungen erfüllt wurden, und wenn nicht, welche Anforderungen nicht erfüllt wurden.

Der Simulations-Thread hat die Aufgabe, die Simulation zu steuern und sicherzustellen, dass die Szenarien in der Simulationsumgebung korrekt ausgeführt werden. Er sorgt dafür, dass die Fahrzeuge entsprechend den definierten Szenarien agieren und interagieren. Dabei erhält der Simulations-Thread Informationen von den anderen BThreads, um die Simulation entsprechend anzupassen und zu steuern. Z.B. werden in diesem Thread Aktionen der Fahrzeuge ausgeführt, die von anderen BThreads definiert wurden. Dieser Thread dient insbesondere der korrekten Übersetzung der Aktionen für die jeweilige Simulationsumgebung, die in diesem Integrationsprojekt entweder HighwayEnv oder Sumo sein kann. In diesem Thread wird auch auf Kollisionen und andere terminierende Ereignisse geprüft, um die Simulation entsprechend zu beenden oder anzupassen.

Das konkrete Szenario wird durch die Kombination von verschiedenen BThreads modelliert, die zusammen die vollständige Logik und Anforderungen des Szenarios abbilden. Dabei wird spezifische Logik implementiert, die nur für dieses Szenario relevant ist und meist in weitere Methoden ausgelagert. Die entsprechende Logik wird dann im relevanten BThread aufgerufen, um den Teil des Szenarios zu steuern. Dabei können mehrere BThreads auch zu einem konkreten Szenario gehören, um verschiedene Aspekte abzudecken. Diese Threads können über ein Parallelitätskonstrukt parallel ausgeführt werden, um die gleichzeitige Ausführung verschiedener Aspekte des Szenarios zu ermöglichen. Spezifische Logik wird in den einzelnen BThreads per `yield from` aufgerufen, um diese anzuwenden.

- Erklärung der einzelnen BThreads -> Simulation, abstraktes Szenario, konkretes Szenario
- am Beispiel von follow-behind-Szenario

Die Aufteilung und Aufgaben der einzelnen BThreads wird jetzt noch einmal an einem konkreten Beispiel, dem Follow-Behind-Szenario, erläutert.

## 5.4 Visualisierung von ausgeführten konkreten Szenarien

## 5.5 Reinforcement Learning

Hier wird die Struktur und Funktionalität des Projekts, das ein Reinforcement-Learning-Modell (RL) für Überholszenarien trainiert, speichert und nutzt beschrieben. Die Hauptkomponenten des Projekts sind in mehreren Python-Dateien organisiert.

`src/envs/overtake_env.py` Diese Datei definiert die Umgebung für das Überholszenario, bei dem das VUT zum Überholen des vom Modell gesteuerten Fahrzeugs angeregt werden soll. Sie basiert auf einer Kernumgebung und erweitert diese um spezifische Logik für das Training eines RL-Agenten. Analog existiert die Datei `src/envs/intersection_env.py`, bei dem eine Umgebung für ein Kreuzungsszenario definiert wird.

- `__init__(render_mode, **config_overrides)`: Initialisiert die Umgebung mit dem übergebenen Rendermodus. Zusätzlich kann die Konfiguration der Umgebung über den `config_overrides` Parameter überschrieben werden.
- `reset(**kwargs)`: Setzt die Umgebung zurück, initialisiert die Position und Geschwindigkeit des Agenten sowie des zu überholenden Fahrzeugs (VUT).
- `step(action)`: Führt einen Schritt in der Umgebung aus, basierend auf der Aktion des Agenten. Berechnet Belohnungen anhand des Zustandes der Simulation, unter anderem der Position der Fahrzeuge im Vergleich zueinander. Abschließen wird überprüft, ob die Episode beendet ist.
- `render()`: Rendert die Umgebung.

`src/main.py` Diese Datei enthält den Einstiegspunkt für die Ausführung des Projekts und die Konfiguration der Umgebung.

- `create_env(config: Dict[str, Any])`: Erstellt eine neue Umgebung basierend auf der übergebenen Konfiguration.
- `set_config()`: Definiert die Konfiguration der Umgebung, einschließlich Fahrspuren, Fahrzeugpositionen und Beobachtungs-/Aktionsräume.
- `main()`: Führt die Hauptlogik aus, einschließlich der Initialisierung der Umgebung und der Simulation von Aktionen.

`src/training/train_overtake_agent.py` Diese Datei ist für das Training des RL-Modells verantwortlich.

- `make_env(rank: int)`: Erstellt eine Instanz der Überholumgebung für das Training.
- `RenderCallback`: Eine Callback-Klasse, der die Umgebung während des Trainings in regelmäßigen Abständen rendert. Ihre `_on_step()` Methode wird vom Modell während des Trainings aufgerufen um zu bestimmen ob die aktuelle Episode gerendert werden soll.

- `main()`: Führt das Training des RL-Modells durch. Es initialisiert das Modell und definiert die Trainingsumgebung, trainiert es mit der DQN-Methode und speichert das trainierte Modell.

### 5.5.1 Trainieren des Modells

Das Training erfolgt in der Datei `train_overtake_agent.py`. Die Hauptschritte sind:

1. Initialisierung der Umgebung mit `DummyVecEnv`, welchem das zu lernende Szenario übergeben wird.
2. Definition des RL-Modells mit DQN. Dabei werden diverse Parameter wie der Diskontierungsfaktor für das Lernen des Modells festgelegt.
3. Start des Trainings mit `model.learn()`.

### 5.5.2 Speichern des Modells

Das trainierte Modell wird mit einem Zeitstempel versehen und im Verzeichnis `models/` gespeichert:

```
model.save("models/overtake_dqn.zip" + current_time)
```

### 5.5.3 Nutzen des Modells

Das gespeicherte Modell kann später geladen und für Inferenz oder weitere Trainingsschritte verwendet werden:

```
from stable_baselines3 import DQN
model = DQN.load("models/overtake_dqn.zip")
```

## Kapitel 6

# Evaluierung

## Kapitel 7

## Fazit



## Kapitel 8

## Ausblick

## Anhang A

# Anhang

Die Zitate habe ich nur hinzugefügt, weil ich sonst einen Error bekomme, dass ich keine citations nutze... [FHD18, FHD21]

Der Anhang stellt eine Ergänzung zur eigentlichen Arbeit dar. Denkbare Einsatzmöglichkeiten sind:

- Dokumentation von Materialien, die dem Leser üblicherweise nicht zugänglich sind (Herstellung der Zitierfähigkeit): Dieser Punkt kommt bspw. zum Tragen, wenn in der Arbeit unternehmensinterne Unterlagen eingearbeitet werden (z.B. Organigramme, Auszüge aus Organisationshandbüchern, Texte von Betriebsvereinbarungen, Verkaufsunterlagen, Produktbeschreibungen).
- Detaillierte Beschreibungen von hergestellten Artefakten, z.B. ausführliche Anwendungsfallbeschreibungen, wenn im Hauptteil Anwendungsfälle schon grob beschrieben sind und dies für das primäre Verständnis ausreicht, oder ausführliche Klassendiagramme, die übersichtsartige Diagramme im Hauptteil ergänzen. Hier können auch Code-Listings, von Algorithmen oder Skripten, die Sie entwickelt haben.
- Ergebnisse von Gesprächen mit Experten in Form von Gesprächsnotizen oder Interviewprotokolle (z. B. in Frage-Antwort-Form).

Ein Anhang kann eine sinnvolle Ergänzung der eigentlichen Arbeit sein. Material sollte jedoch nicht beliebig angehängt werden, sondern durch geeignete Beschreibungen mit dem Inhalt der Arbeit in Verbindung gebracht werden. Es gilt auch hier: Beschränken Sie sich auf Wichtiges. Im Regelfall sollte der Anhang nicht 50% des Seitenumfangs des Inhaltsteils überschreiten. Im Zweifel kann eine Arbeiten auch ohne Anhang auskommen.

**Der Anhang ist insbesondere kein Ort für Abbildungen, die aus dem Hauptteil ausgelagert werden** – insbesondere nicht zu dem Zweck, dadurch den vorgegebenen Seitenumfang des hauptinhaltlichen Teils einzuhalten. Es erschwert unnötig die Lesbarkeit der Arbeit, wenn zwischen

Hauptteil und Anhang hin- und hergeblättert werden muss. **Die Arbeit muss auch ohne Anhang vollständig und verständlich sein!**

# Literaturverzeichnis

- [FHD18] FHDW Hannover. Modulhandbuch Bachelor (mit Studienbeginn ab Oktober 2018). Interne Quelle der FHDW Hannover, 2018. Stand 17.2.2020.
- [FHD21] FHDW Hannover. Modulhandbuch Master Information Engineering (Modulhandbuch HFP). Interne Quelle der FHDW Hannover, 2021. Stand Mai 2021.
- [HMW10] D. Harel, A. Marron, and G. Weiss. Programming coordinated behavior in Java. In *ECOOP 2010 – Object-Oriented Programming*, volume 6183 of *Lecture Notes in Computer Science*, pages 250–274. Springer, 2010. doi:10.1007/978-3-642-14107-2\_12.
- [HMW12] D. Harel, A. Marron, and G. Weiss. Behavioral programming. *Communications of the ACM*, 55(7):90–100, 2012. doi:10.1145/2209249.2209270.