

Schlussbericht

Matthias Althoff, Technische Universität München, Garching

Vertrauenswürdige KI mittels Propagation von Mengen (TRAITS)



Das diesem Bericht zugrunde liegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01IS21087 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei der Autorin/beim Autor.

1 Darstellung der durchgeführten Arbeiten

Um maschinelles Lernen abzusichern, haben wir einer Sicherheitsschicht entwickelt, die während der Laufzeit prüft, ob vorgeschlagene Aktionen des gelernten Agenten die Systemspezifikation erfüllen. Wenn die vorgeschlagenen Aktionen nicht sicher sind, führt die Sicherheitsschicht eine ausfallsichere Steuerung aus, die das System in einen sicheren Betriebszustand bringt. Unsere vorgeschlagene Architektur hat den großen Vorteil, dass die Techniken der künstlichen Intelligenz nicht zertifiziert werden müssen – nur die Sicherheitsschicht muss zertifiziert werden. Selbst wenn das Modul der künstlichen Intelligenz geändert wird, bleiben alle seine Entscheidungen aufgrund unserer vorgeschlagenen Sicherheitsschicht sicher. Dieses Gesamtziel wurde durch die im Folgenden beschriebenen Arbeitspakete erzielt.

1.1 Just-in-Time-Verifikation (AP 1)

Problembeschreibung: Bei der just-in-Time-Verifikation muss zur Laufzeit das Verhalten eines intelligenten Agenten verifiziert werden. Durch die Verifikation während der Laufzeit können alle auftretenden Situationen berücksichtigt werden – im Gegensatz zur klassischen Verifikation, bei der relevante Szenarien übersehen werden können.

Lösung: Wir haben eine just-in-Time-Verifikation erarbeitet, die basierend auf einer Erreichbarkeitsanalyse feststellen kann, ob eine geplante Aktion von einem intelligenten Agenten sicher durchführbar ist. Erreichbarkeitsmengen sind nur für einen endlichen Zeithorizont berechenbar – wir möchten aber die Sicherheit der vorgeschlagenen Aktionen für einen unendlichen Zeithorizont garantieren. Aus diesem Grund haben wir ein neuartiges Verfahren zur Berechnung von Invarianzmengen entwickelt [1]. Sobald eine Erreichbarkeitsmenge in einer Invarianzmenge enthalten ist, kann die Berechnung abgebrochen werden, da die Invarianzmenge alle zukünftigen Erreichbarkeitsmengen enthält.

Dadurch, dass wir erstmalig Invarianzmengen mittels Zonotopen berechnen konnten, konnten wir Invarianzmengen für verschiedenste Systeme aus der Literatur mit bis zu 20-dimensionalem Zustandsraum innerhalb weniger Minuten berechnen. Die Ergebnisse sind in Tabelle 1 zusammengefasst und führen zu folgendem Fazit: Unser Ansatz ist vielseitig anwendbar und ermöglicht die Berechnung invarianter Mengen für verschiedenste Regelungssysteme von chemischen Reaktoren bis hin zu unteraktuierten Systemen, wie einem inversen Pendel. Darüber hinaus zeigen die Rechenzeiten in der sechsten Spalte (durchschnittliche Rechenzeit je gelöstem konvexen Optimierungsproblem) und der letzten Spalte (gesamte Rechenzeit) die Skalierbarkeit unseres Ansatzes.

Tabelle 1: Übersicht der Rechenzeiten zur Berechnung von Invarianzmengen für verschiedene nichtlineare Systeme (Details siehe [1])

	n_x	n_u	n_w	Iterative Konvexifizierung		Verifikation		Gesamte Rechenzeit
				# Iterationen	Ø Rechenzeit	# Iterationen	Korrekt?	
Künstliches System	2	1	1	15	0.12 s	0	✓	3.1 s
Strahltriebwerk	2	1	1	10	0.17 s	1	✓	6.9 s
Masse-Feder-Dämpfer System I	2	1	2	12	0.11 s	0	✓	4.2 s
Masse-Feder-Dämpfer System II	2	1	2	10	0.12 s	0	✓	5.3 s
Inverses Pendel auf Wagen	2	1	1	16	0.12 s	0	✓	5.4 s
Inverses Pendel auf Wagen	4	1	1	16	0.75 s	0	✓	45.2 s
Doppeltes inverses Pendel	4	1	1	18	1.1 s	0	✓	48.3 s
Chemischer Reaktor	4	2	2	14	1.1 s	0	✓	57.5 s
Manipulator	4	2	4	10	0.64 s	0	✓	54.4 s
Multicopter	6	2	1	10	0.90 s	1	✓	36.5 s
Van-der-Pol Oszillatoren	10	5	5	14	4.1 s	0	✓	103.8 s
Masse-Feder-Dämpfer System III	20	10	10	11	7.92 s	0	✓	116.4 s

Um die Skalierbarkeit mit der Zustandsraumdimension genauer zu untersuchen, haben wir unseren Ansatz auf n_{Masse} gekoppelte Masse-Feder-Dämpfer Systeme mit nichtlinearer Feder- und Dämpfungskraft angewandt (basierend auf Masse-Feder-Dämpfer System I in Tabelle 1). Dieses System

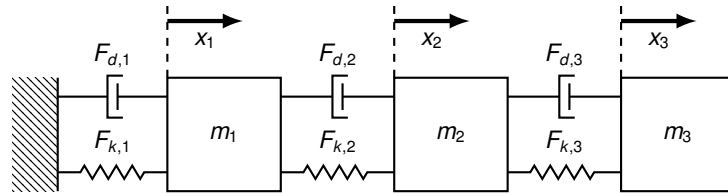
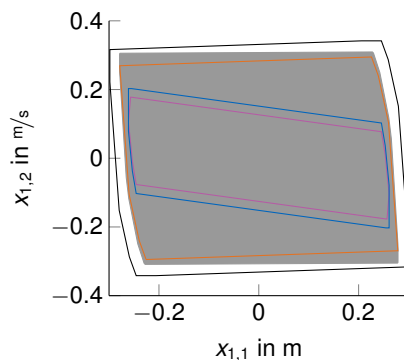


Abbildung 1: Gekoppelte Masse-Feder-Dämpfer-Systeme ($n_{Masse} = 3$).

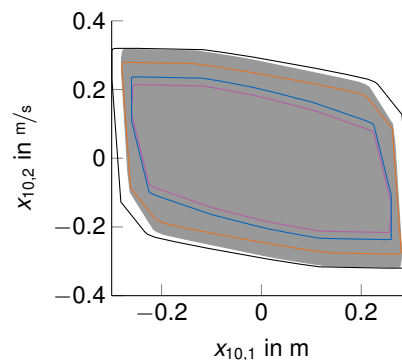
ist beispielhaft für drei gekoppelte Massen in Abb. 1 dargestellt. Die Ergebnisse für ein ($n_x = 2$) bis zehn ($n_x = 20$) gekoppelte Masse-Feder-Dämpfer Systeme sind in Tabelle 2 zusammengefasst. Dabei nimmt der Rechenaufwand nur moderat mit steigender Dimension des Zustandsraumes zu, was am besten anhand der durchschnittlichen Rechenzeit je konvexem Optimierungsproblem (siehe Spalte 3 in Tabelle 2) beobachtet werden kann. Die Projektionen der invarianten Menge für $n_{Masse} = 10$ auf die $x_{1,1}/x_{1,2}$ -Ebene (Auslenkung aus Ruhelage bzw. Geschwindigkeit der ersten Masse) und auf die $x_{10,1}/x_{10,2}$ -Ebene (Auslenkung aus Ruhelage bzw. Geschwindigkeit der zehnten Masse) sind in Abb. 2 (a) bzw. Abb. 2 (b) abgebildet.

Tabelle 2: Analyse der Skalierbarkeit mittels n_{Masse} gekoppelter Masse-Feder-Dämpfer-Systeme.

$n_{Masse} (n_x)$	Iterative Konvexifizierung		Verifikation		Gesamte Rechenzeit
	# Iterationen.	Ø Rechenzeit	# Iterationen	Zulässige Lösung?	
1 (2)	15	0.13 s	1	✓	5.7 s
2 (4)	13	0.55 s	1	✓	39.0 s
3 (6)	14	0.98 s	0	✓	41.5 s
4 (8)	12	1.79 s	0	✓	59.3 s
5 (10)	13	2.86 s	0	✓	66.4 s
6 (12)	9	2.59 s	1	✓	41.3 s
7 (14)	13	4.04 s	1	✓	79.8 s
8 (16)	16	6.4 s	0	✓	143.6 s
9 (18)	12	11.8 s	0	✓	205.4 s
10 (20)	11	7.92 s	0	✓	116.4 s



(a) Invarianzmenge von Masse m_1



(b) Invarianzmenge von Masse m_{10}

Abbildung 2: Projektion der Invarianzmengen für zehn gekoppelte Masse-Feder-Dämpfer-Systeme (Überapproximation der erreichbaren Menge zum Zeitpunkt Δt (pink) und erreichbare Menge im Zeitintervall $[0, \Delta t]$ (grau), Approximationen der erreichbaren Menge zum Zeitpunkt Δt (blau) und im Zeitintervall $[0, \Delta t]$ (orange)).

Neuheit: Unser Ansatz zur Berechnung von Invarianzgebieten nichtlinearer Regelungssysteme ermöglicht erstmals die Verifikation höher-dimensionaler Systeme mit unendlichem Zeithorizont. Durch die Verwendung von Zonotopen zur Darstellung der invarianten Mengen und der Kombination von konvexer Optimierung und mengenbasierter Erreichbarkeitsanalyse, skaliert die Rechenkomplexität unseres Ansatzes sehr gut bezüglich der Dimension des Zustandsraumes: Während Ansätze aus der

Literatur bisher nur auf niedrigdimensionale Systeme angewendet wurden, ermöglicht unser Ansatz die Berechnung invarianter Menge von Systemen mit bis zu 20 Dimensionen innerhalb weniger Minuten.

1.2 Falsifikation von Systemeigenschaften (AP 2)

Problembeschreibung: Aktionen, die unsere Just-in-Time-Verifikation nicht bestanden haben, erlauben es, ein System speziell auf kritische Szenarien zu trainieren. Die Aufgabe dieses Arbeitspaketes war Verletzungen basierend auf der Just-in-Time-Verifikation genau zu lokalisieren.

Lösung: Eine erreichbare Menge enthält zwar die Menge aller möglichen Lösungen, gibt aber keine Auskunft darüber, wie sich Systeme innerhalb dieser Mengen entwickeln können. Um in der Lage zu sein, Verhaltensweisen innerhalb der erreichbaren Mengen zu synthetisieren, die die Spezifikation verletzen, haben wir polynomiale Zonotopen als Mengenrepräsentation verwendet, die Abhängigkeiten zwischen erreichbaren Mengen zu verschiedenen Zeitpunkten enthalten können. In Abb. 3 ist beispielhaft ein generiertes Verhalten innerhalb der zuvor berechneten Erreichbarkeitsmenge gezeigt, die eine unsichere Menge erreicht. Dies ermöglicht es, Lösungen zum Erreichen von Extremwerten durch die Lösung eines einfachen statischen Optimierungsproblems zu formulieren. Ein Vergleich mit klassischen Ansätzen zur Optimierung von kontinuierlichen Systemen hat gezeigt, dass unser neuartiger Ansatz bei größeren Problemen um zwei Größenordnungen schneller ist, als der Stand der Technik. Die Anwendbarkeit unserer neuen Methode hat sich auch im Rahmen der *International Competition on Verifying Continuous and Hybrid Systems* gezeigt [2].

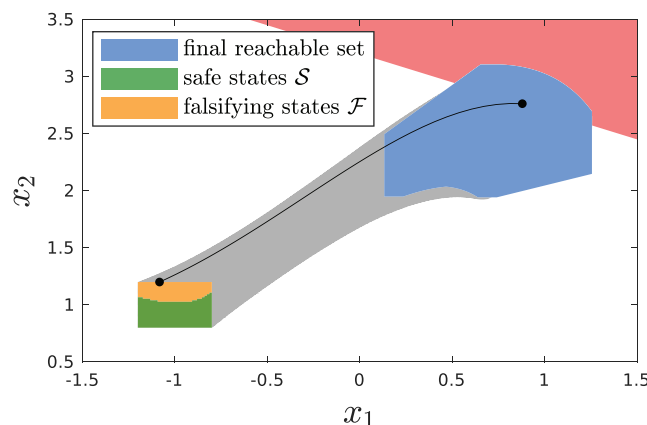


Abbildung 3: Mittels Falsifikation wird ein Verhalten innerhalb der zuvor berechneten Erreichbarkeitsmenge generiert, dass eine unsichere Menge erreicht.

Neuheit: Unser Ansatz ist der erste, der mittels Erreichbarkeitsmengen eine Falsifikation für nichtlineare Systeme berechnen kann. Da wir die Schnittmenge der erreichbaren Menge mit einer unsicheren Menge kennen, kann mittels Falsifikation gezielt ein System in dieses Gebiet gelenkt werden. Dadurch können sich Rechenzeiten um mehr als den Faktor 100 verkürzen.

1.3 Robustifizieren von Reinforcement Learning (AP 3)

Problembeschreibung: In diesem Arbeitspaket sollten neuartige Kombinationen von Just-in-Time-Verifikation mit Reinforcement-Learning dabei helfen die fehlersichere Lösung weniger häufig auszuführen.

Lösung: Wir haben drei Möglichkeiten realisiert, um beweisbare Sicherheit für Reinforcement Learning mittels Just-in-Time-Verifikation zu gewährleisten:

1. Aktionsersetzung, bei der die Sicherheitsmethode alle unsicheren Aktionen des Agenten durch sichere Aktionen ersetzt.
2. Aktionsprojektion, bei der unsichere Aktionen auf den sicheren Aktionsraum projiziert werden.
3. Aktionsmaskierung, bei der der Agent nur Aktionen aus dem sicheren Aktionsraum wählen kann.

Aktionsersetzung und Aktionsprojektion ändern die Aktion, nachdem der Agent sie zurückgegeben hat. Im Gegensatz dazu lässt die Aktionsmaskierung den Agenten ausschließlich aus dem sicheren Aktionsraum wählen. Abb. 4 zeigt das Grundkonzept dieser Methoden. Eine genaue Beschreibung dieser Konzepte, deren mathematische Formalisierung und die praktischen Implikationen der einzelnen Ansätze sind im Detail in [3] beschrieben.

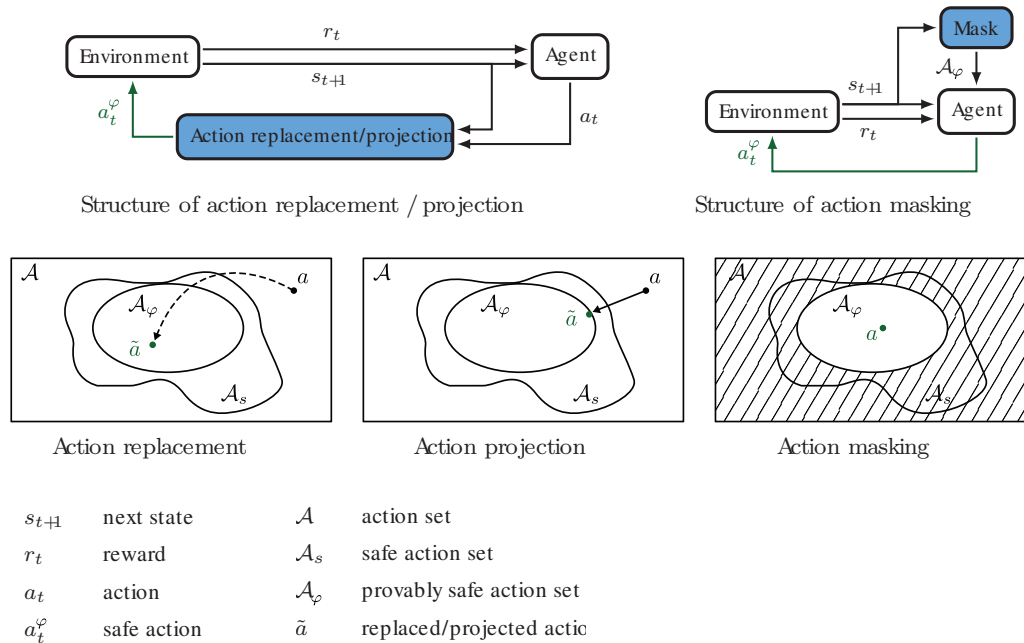


Abbildung 4: Mittels Falsifikation wird ein Verhalten innerhalb der zuvor berechneten Erreichbarkeitsmenge generiert, dass eine unsichere Menge erreicht.

Neuheit: Maskierung für kontinuierliche Aktions- und Zustandsräume, sowie ein erstmaliger Vergleich von Ansätzen für beweisbar sicheres Reinforcement Learning.

1.4 Integration von datenbasierter Modellierung und Vorhersage (AP 4)

Problembeschreibung: Ziel dieses Arbeitspaketes war zu untersuchen, inwieweit datenbasierte Ansätze in unsere Sicherheitsarchitektur integriert werden können, ohne die Sicherheit oder die Leistung zu beeinträchtigen.

Lösung: Klassisches Reinforcement Learning (RL) kann komplexe Aufgaben effizient lösen, bietet aber keine Garantien für das Systemverhalten. Um diese Lücke zu schließen, haben wir ein dreistufiges sicheres RL-Verfahren für kontinuierliche Aktionsräume entwickelt, das probabilistische Garantien bezüglich Spezifikationen in temporaler Logik bietet. Durch die Trennung von Sicherheitsüberprüfung und Leistungsverbesserung werden sowohl explizite probabilistische Sicherheitsgarantien als auch ein einfaches RL-Setup zur Verbesserung der Performanz realisiert. Die probabilistischen Sicherheitsgarantien werden realisiert, indem der Aktionsraum des RL-Agenten beschränkt wird. Wir haben unseren Ansatz für einen mobilen Roboter evaluiert, der ein Ziel erreichen muss, während er einem dynamischen Hindernis ausweicht. Unsere Ergebnisse zeigen, dass unser sicherer RL-Ansatz zu effizientem Lernen führt und gleichzeitig probabilistische Spezifikationen einhält.

Abb. 5 zeigt die Evaluierung unseres entwickelten Verfahrens. Eine genaue Beschreibung des gesamten Ansatzes findet sich in [4].

Neuheit: Die Arbeiten aus AP 3 sind zwar beweisbar korrekt, erfordern aber Modellwissen zur Absicherung von Reinforcement Learning. Der in AP 4 entwickelte Ansatz benötigt hingegen überhaupt kein Modellwissen, ist allerdings nur korrekt bezüglich einer vorgegebenen Wahrscheinlichkeit.

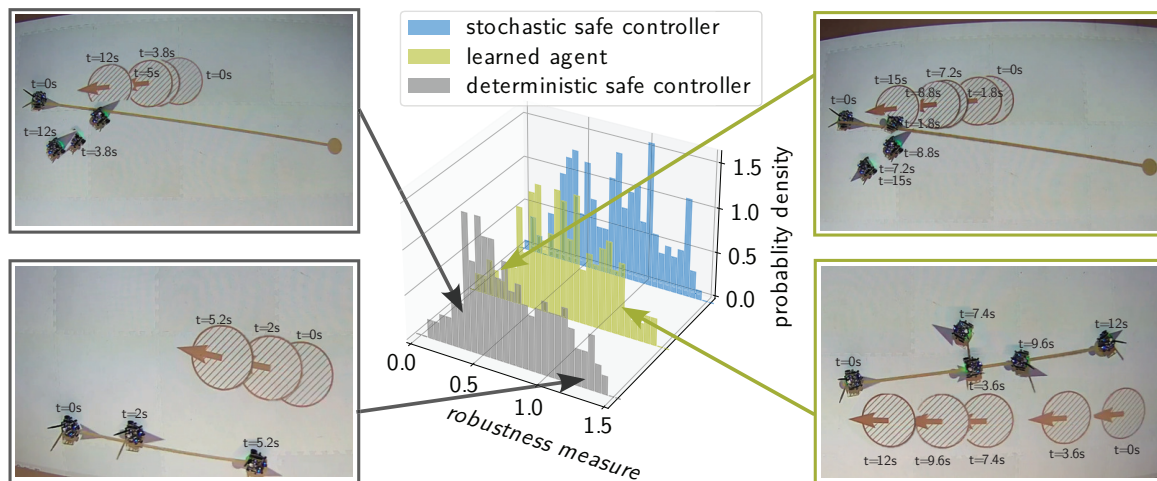


Abbildung 5: Mitte: Histogramm der Robustheitswerte zur Einhaltung der Spezifikation für 200 Stichproben. Links: Trajektorien des Roboters und des Hindernisses, wenn man einen beweisbar sicheren Regler verwendet (Robustheitswert oben links: 0.114, Robustheitswert unten links: 1.62). Rechts: Trajektorien des Roboters und des Hindernisses, wenn man den entwickelten Ansatz verwendet (Robustheitswert oben rechts: 0.089, Robustheitswert unten rechts: 1.21).

1.5 Statistische Leistungs- und Sicherheitsbewertung an realen Systemen (AP 5)

Problembeschreibung: In dieses Arbeitspaket sollten unsere sicheren Reinforcement-Learning-Ansätze auf autonomen Fahrzeugen und Robotern evaluiert werden. Die Ergebnisse sollten durch Simulationen und reale Experimente validiert werden, die von einfachen bis zu komplexen Szenarien reichen.

Lösung: Als autonome Fahrzeuge haben wir uns für ein autonomes Schiff und unser Forschungsfahrzeug EDGAR entschieden. Derzeit gibt es keine Benchmarks für autonome Schiffe, um verschiedene Ansätze zu Reinforcement Learning zu vergleichen. Daher haben wir kompositionelle Benchmarks für die Bewegungsplanung auf Ozeanen (CommonOcean) entwickelt, die unter `commonocean.cps.cit.tum.de` verfügbar sind. Wie in Abb. 6 gezeigt, besteht ein CommonOcean-Benchmark aus drei Elementen: Kostenfunktion, Schiffsmodell und Bewegungsplanungsszenario. Benchmarks können mithilfe von eindeutigen Bezeichnern für diese Elemente, die hochgradig modular sind, bequem zusammengestellt werden. CommonOcean ist einfach zu bedienen, da wir aussagekräftige Parameter für Schiffsmodelle, verschiedene Bewegungsplanungsszenarien und eine umfassende Dokumentation bereitstellen. Darüber hinaus haben wir ein Tool zur Erstellung von Szenarien entwickelt, mit dem man mühelos neue Szenarien aus Schiffsverkehrsdaten erstellen kann. Wir glauben, dass CommonOcean zu einer besseren Reproduzierbarkeit und Vergleichbarkeit der Forschung zur Bewegungsplanung von Schiffen führen wird.

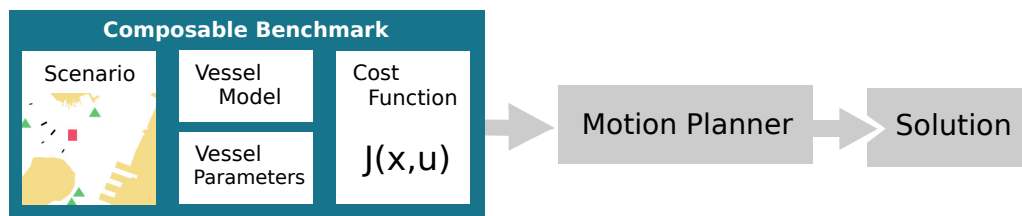
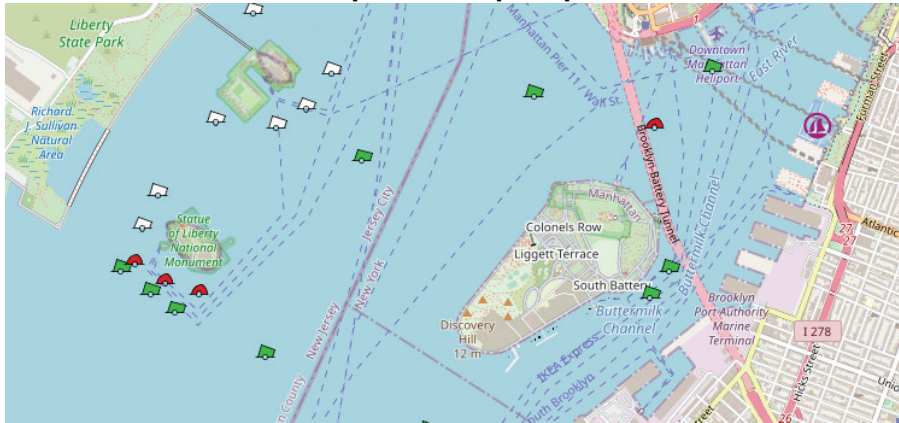


Abbildung 6: Konzept von CommonOcean: Benchmarks werden aus Kostenfunktion, Schiffsmodell und Bewegungsplanungsszenario zusammengestellt und anschließend wird die Lösung des Bewegungsplaners bewertet.

Abb. 7 zeigt die Erstellung eines Benchmarks anhand einer OpenSeaMap Karte. Eine genaue Beschreibung des gesamten Ansatzes findet sich in [5]. In Abb. 8 ist eine Testfahrt mit unserem Forschungsfahrzeug EDGAR gezeigt.

OpenSeaMap map



CommonOcean scenario

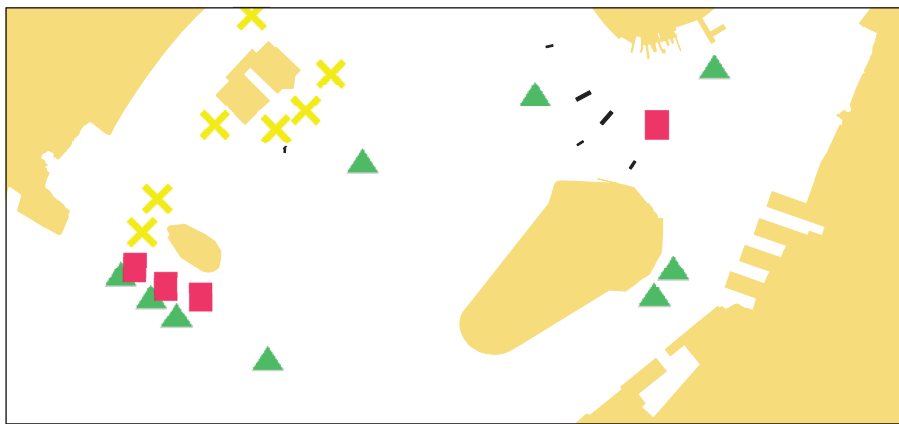


Abbildung 7: Erstellung eines Benchmarks anhand einer OpenSeaMap Karte.



Abbildung 8: Testfahrt mit unserem Forschungsfahrzeug EDGAR.

Neuheit: Wir haben die erste Benchmark-Datenbank für die Bewegungsplanung von autonomen Schiffen erstellt.

2 Wichtigsten Positionen des zahlenmäßigen Nachweis / Finanzierungsübersicht

Das Projekt wurde zu 100% aus Bundesmitteln gefördert. Die Gesamtkosten betragen 254.615,98 EUR.

3 Notwendigkeit und Angemessenheit der geleisteten Projektarbeiten

Die durchgeführten Arbeiten sowie die dafür aufgewandten Ressourcen waren notwendig und angemessen, da sie (i) der im Projektantrag detailliert dargelegten Planung entsprachen und (ii) es ermöglichten die im Arbeitsplan formulierten Aufgaben erfolgreich zu bearbeiten. Darüber hinaus mussten keine zusätzlichen Ressourcen zur Durchführung des Vorhabens aufgewandt werden.

Insbesondere realisiert das durchgeführte Vorhaben das Ziel der Künstliche-Intelligenz-Strategie der deutschen Bundesregierung¹, neue Methoden in Bezug auf Erklärbarkeit und Nachvollziehbarkeit von algorithmusbasierten Prognose- und Entscheidungssystemen zu erforschen (Seite 16). Außerdem wird der Mobilitätssektor als ein gewünschter Hauptanwendungsbereich adressiert (Seite 35).

4 Voraussichtlicher Nutzen und Verwertbarkeit der Ergebnisse

Wissenschaftlicher Nutzen Die wissenschaftlichen Ergebnisse sind in Abschnitt 1 beschrieben, die zu den in Abschnitt 6 aufgelisteten Publikationen geführt haben. Um unsere Ergebnisse vollständig reproduzieren zu können, haben wir alle Daten sowie deren Formate zugänglich gemacht und Software-Tools zum Bearbeiten der Daten bereitgestellt. TRAITS hat sich voll und ganz den Open-Access-Publikationen verschrieben. Alle Vorabdrücke von Publikationen, die in TRAITS entstanden sind, stehen auf unserer Webseite (<https://www.ce.cit.tum.de/cps/publications/>) kostenlos zur Verfügung.

Gesellschaftlicher Nutzen Die Strategie für künstliche Intelligenz der Bundesregierung aus dem Jahr 2018 zielt darauf ab, die Forschung hinsichtlich Erklärbarkeit und Nachvollziehbarkeit von Algorithmenbasierten Prognose- und Entscheidungssystemen zu fördern (Seite 16). Unser Ansatz zum sicheren Reinforcement Learning mit Sicherheitsgarantien adressiert genau das Problem, „(...) KI erklärbar, nachvollziehbar und transparent zu machen (...), um das Vertrauen der Öffentlichkeit in KI zu gewinnen“. Unsere Arbeit wird das Vertrauen wesentlich verbessern, indem wir die Sicherheit mit formalen Methoden garantieren.

Wirtschaftlicher Nutzen Die Verwertung der Projektergebnisse ist ein wichtiges Ziel von TRAITS und wird dazu beitragen, die Wettbewerbsfähigkeit Europas auf dem wachsenden Markt der sicheren künstlichen Intelligenz zu sichern. Die zukünftige Marktposition von Unternehmen, die sichere künstliche Intelligenz in den Bereichen Transport, Produktion und Energie entwickeln, wird durch ihre Fähigkeiten bestimmt werden, Lösungen anzubieten, die nachweislich hohe Sicherheitsstandards haben, die gleichzeitig flexibel sowie in einem breiten Spektrum von Arbeitsbedingungen einsetzbar sind und die zeitnah und kosteneffizient zur Marktreife gebracht werden können. Die Ergebnisse von TRAITS werden dazu beitragen, die europäische Spitzenposition in diesen Märkten gegenüber der amerikanischen und fernöstlichen Konkurrenz zu behaupten. Ganz konkret ist mithilfe der Projektergebnisse das Startup aiina hervorgegangen, das durch einen EXIST Forschungstransfer gefördert wird und den ersten Platz beim TUM IDEAward 2024 gewonnen hat.

5 Bekannt gewordener Fortschritt Anderer während des Vorhabens

Wir haben in unserem Übersichtsartikel in [3] zum Themengebiet „beweisbar sicheres Reinforcement Learning“ den Fortschritt Anderer während der Projektlaufzeit detailliert dargestellt. Der Artikel ist unter der Internetadresse <https://openreview.net/pdf?id=mcN0ezbnz0> frei zugänglich.

6 Liste der Veröffentlichungen

- [1] L. Schäfer, F. Gruber, and M. Althoff, “Scalable computation of robust control invariant sets of nonlinear systems,” *IEEE Transactions on Automatic Control*, vol. 69, no. 2, pp. 755–770, 2024.
- [2] M. Althoff, M. Forets, Y. Li, S. Mitra, C. Schilling, M. Wetzlinger, and D. Zhuang, “ARCH-COMP23 category report: Continuous and hybrid systems with linear continuous dynamics,” in *Proc. of*

¹<https://www.bundesregierung.de/resource/blob/997532/1550276/3f7d3c41c6e05695741273e78b8039f2/2018-11-15-ki-strategie-data.pdf>

10th International Workshop on Applied Verification of Continuous and Hybrid Systems, ser. EPiC Series in Computing, vol. 96, 2023, pp. 34–60.

- [3] H. Krasowski, J. Thumm, M. Müller, L. Schäfer, X. Wang, and M. Althoff, “Provably safe reinforcement learning: Conceptual analysis, survey, and benchmarking,” *Transactions on Machine Learning Research*, 2023.
- [4] H. Krasowski, P. Akella, A. D. Ames, and M. Althoff, “Safe reinforcement learning with probabilistic guarantees satisfying temporal logic specifications in continuous action spaces,” in *Proc. of the 62nd IEEE Conference on Decision and Control*, 2023, pp. 4372–4378.
- [5] H. Krasowski and M. Althoff, “CommonOcean: Composable benchmarks for motion planning on oceans,” in *Proc. of the 25th IEEE International Conference on Intelligent Transportation Systems*, 2022, pp. 1676–1682.