Fachhochschule
Dortmund
University of Applied Sciences and Arts

Betrachtung technischer Maßnahmen zur Umsetzung ethischer Sicherheitsaspekte bei der Erstellung und Anwendung von Reinforcement-Learning-Modellen

Niels Schlunder
geboren in Paderborn

2020

Praktische Informatik
Fachbereich Informatik
Fachhochschule Dortmund

1. Gutachter : Prof. Dr. Sebastian Bab

2. Gutachter : Prof. Dr. Burkhard Lenze

Abgabedatum: 18. Mai 2020

Kurzfassung

Reinforcement Learning wird als Kerntechnologie in autonomen Anwendungen der Medizin, Mobilität oder unbemannten Waffensystemen eingesetzt. Daraus ergibt sich neben dem Potenzial gesellschaftlichen und wissenschaftlichen Fortschrittes auch das Risiko zur Gefährdung von Wohlergehen und Würde des Einzelnen. Basierend auf Prinzipien der angewandten Ethik und der Betrachtung regionaler Leitlinien für ethische KI-Anwendungen werden im Rahmen dieser Arbeit ethische Werte als Grundlage für moralisches Handeln von Reinforcement Learning Agenten definiert. Um die Umsetzung dieser Werte zu unterstützen wird ein Vorgehensplan aufgezeigt, der parallel zu gängigen Vorgehensmodellen in den Entwicklungsprozess eingeführt werden kann. Innerhalb des Vorgehensplan werden technische und organisatorische Maßnahmen eingeführt, die über den tatsächlichen Implementierungsprozess hinaus die Zusicherung moralischen Handelns von Reinforcement Learning Agenten unter Miteinbeziehung aller Projektbeteiligten und möglichst verfahrensunabhängig unterstützt.

Abstract

Reinforcement Learning as a core technology for autonomous applications can be used in domains such as healthcare, transport or unmanned weapon systems. Apart from the potential for social and economic growth, the technology could possbily endanger human dignity and well-being. Based on principles of applied ethics and regional guidelines for ethical AI applications, ethical values as a foundation for moral behaviour of Reinforcement Learning agents are defined. To support the implementation of these values, process guidelines are proposed, which can be introduced simultaneous to most existing process models. The guidelines contain

technical and organisational measures as a mostly methodological independent approach, which ensures moral behaviour of Reinforcement Learning agents beyond the implementional process including all project members.

Inhaltsverzeichnis

1 Einleitung

Reinforcement Learning als maschinelles Lernverfahren ermöglicht eine vielversprechende Herangehensweise an die Entwicklung autonomer Systeme. Es unterscheidet sich zu anderen maschinellen Lernverfahren, indem keine Kenntnis über die Klassenzugehörigkeit der Daten benötigt wird. Insbesondere für die Anwendung in autonomen Systemen ist es z.B. bei überwachten Lernverfahren aufwendig, genügend große Datenmengen zu erzeugen, speziell dann, wenn alle Anwendungsfälle abgedeckt und eindeutig einer Klasse zugewiesen werden müssen. Im Vergleich zu herkömmlichen Paradigmen bietet Reinforcement Learning die Möglichkeit zur selbständigen Anpassung und Generalisierung an fremde Situationen. Durch diese Potenzial besteht wirtschaftliches und politisches Interesse am Fortschritt, weshalb es umso wichtiger ist, eine sachliche Diskussion bezüglich der Nutzung und möglicher Bedenken zu führen. Ein verantwortungsvoller Umgang bietet dann Potenzial für wirtschaftlichen und gesellschaftlichen Fortschritt. Denn auch wenn die technischen Möglichkeiten vielversprechend aussehen, müssen bei allen lernenden Systemen neben ethischen Bedenken des Nutzungskontextes auch Bedenken während des Entwicklungsprozesses betrachtet werden, denn Fehlentscheidungen können fatale Folgen [Amo+16][Haw] haben und so das Wohl und die Würde des Menschen gefährden. Es stellen sich deshalb die Fragen, welche Maßnahmen ergriffen werden können, damit Agenten zu unserem Wohl gemäß unserer Werte handeln, wie diese Werte formalisiert werden können und wie Adaption und Akzeptanz innerhalb der Gesellschaft zugesichert werden kann.

Im Rahmen dieser Arbeit soll ein Vorgehensplan zur Umsetzung technischer und organisatorischer Maßnahmen erstellt werden, um ethisches Bedenken in bestehende Prozesse der Entwicklung von Reinforcement-Learning-Anwendungen zu integrieren. Dafür werden zunächst Grundlagen des Reinforcement Learning als Kerntechnologie dieser Arbeit eingeführt. Um nicht von einer einzelnen Technologie abhängig zu sein wird dafür der Markov-Entscheidungsprozess als Grundlage zur Umgebungsmodellierung, sowie allgemeine Begriffe und Eigenschaften eingeführt, wodurch eine eigene Evaluation der Technologie gemäß des Nutzungskontexts erfolgen kann. Da im Rahmen der Arbeit Werte zum moralischen Handeln von Reinforcement-Learning-Agenten definiert und technische Maßnahmen zur Umsetzung dieser Werte aufgezeigt werden, wird anschließend ein Überblick über Grundlagen der angewandten Ethik und Moral, sowie der Maschinenethik als Bereichsethik gegeben und diskutiert, inwiefern Reinforcement-Learning-Agenten moralisch handeln können. Anschließend werden als Grundlage der Definition der ethischen Werte die Maßnahmen zur Zusicherung ethischen Umgangs mit Künstlicher Intelligenz auf internationaler, europäischer und asiatischer Ebene, und der USA betrachtet. In Kapitel 5 wird dann zunächst der Kontext der Arbeit abgegrenzt, Anwendungsgebiete sowie der Stand der Technik betrachtet und mit dieser Arbeit verglichen, sowie Herausforderungen bei der Konzeption technischer Maßnahmen ethischer Reinforcement-Learning-Anwendungen als Evaluationsgrundlage diskutiert. In der eigentlichen Konzeption in Kapitel 6 wird zunächst die Wahl ethischer Werte als Grundlage für die Erstellung des Vorgehensplan basierend auf den ethischen Grundlagen und der regionalen Maßnahmen begründet. Anschließend wird in Abschnitt 6.2 der eigentliche Vorgehensplan vorgestellt. Der Vorgehensplan ist angelehnt an einen generischen Softwareentwicklungszyklus [BK13, S. 64], zeigt technische und organisatorische Maßnahmen zur Umsetzung der ethischen Werte auf und beschreibt Möglichkeiten die in Abschnitt 5.3 genannten Herausforderungen zu lösen. Abschlie-

ßend wird der Vorgehensplan gemäß der definierten Fragestellungen, den regionalen Maßnahmen der Länder und den Herausforderungen evaluiert, sowie abschließend zusammengefasst und ein Ausblick über mögliche zukünftige Weiterentwicklungen diskutiert.

2 Grundlagen des Reinforcement Learning

Autonome Systeme bieten die Möglichkeit zur selbständigen Lösung komplexer oder für Menschen gefährlicher Probleme in potenziell unbekannten Umgebungen. Die technische Umsetzung dieser Systeme durch klassische Programmierparadigmen ist in vielerlei Hinsicht problematisch. So ist der Zustandsraum in realen Anwendungen extrem groß. Insbesondere in fremden Umgebungen mangelt es klassischen Programmen an Allgemeingültigkeit. Das System sollte das Problem auch in unbekannten Umgebungen lösen können, ohne dass eine Anpassung der Logik notwendig ist. Als zusätzliche Stufe der Komplexität ergibt sich zudem die Interaktion mit Menschen oder anderen autonomen Systemen.

Menschen lernen von früh auf, indem sie den Einfluss ihrer Aktionen auf die Umgebung [SB18, S. 1] beobachten und daraus Schlüsse ziehen. Jede Reaktion der Umwelt auf das Verhalten wird verarbeitet und beeinflusst die spätere Wahl der Aktionen. Der Mensch entwickelt sich dadurch im Laufe der Zeit [Cas, S. 634]. Formal sind viele Reinforcement-Learning-Verfahren angelehnt an das psychologische Phänomen der operanten Konditionierung [Lef86, S. 34]. Laut Skinner wird operante Konditionierung folgendermaßen beschrieben: „Wenn eine Reaktion [...] von einer Verstärkung gefolgt wird, so resultiert daraus eine Erhöhung der Wahrscheinlichkeit, dass diese Reaktion später unter ähnlichen Umständen wieder auftritt.“[Lef86, S. 34]. Analog dazu sinkt die Wahrscheinlichkeit, wenn die Aktion bestraft wird. Das Ausprobieren von Aktionen erfolgt nach dem Versuchs- und Irrtums-Prinzip [SB18,

S. 2f]. Die Schwierigkeit dabei ist, dass die Aktionen nicht nur Einfluss auf den direkten Folgezustand, sondern auch nachhaltig auf spätere Zustände nehmen. So kann die vermeintlich optimale Aktion zwar kurzfristig die Belohnung maximieren, langfristig aber nicht optimal sein. Der Agent muss dabei abwägen, ob bestehendes Wissen genutzt wird (engl. exploitation) oder neues Wissen hinzugewonnen werden soll (engl. exploration).

Anders als bei anderen maschinellen Lernverfahren, wie beim überwachten Lernen, ist das Ziel des Reinforcement Learning nicht, Wissen aus vorher manuell bewerteten Informationen, sondern aus Aktionen und dessen Auswirkung zu generalisieren. Überwachte Lernverfahren sind in vielen Domänen sinnvoll und können gute Lösungen hervorbringen. Durch die vom Menschen benötigten Information über die Klassenzugehörigkeit der Daten ist dieser Ansatz jedoch nur beschränkt für komplexe Umgebungen anwendbar.

Reinforcement Learning ist nicht ein einzelner Algorithmus, sondern ein Paradigma des maschinellen Lernens, welches aus einer Sammlung von Algorithmen und Vorangehensweisen zusammengesetzt ist. Deshalb wird im Folgenden das Konzept des Markov-Entscheidungsprozesses als Grundlage des Reinforcement Learning beschrieben. Zusätzlich werden Eigenschaften von Reinforcement-Learning-Verfahren aufgezeigt, um daraus in der späteren Konzeption Abschnitt 6.2 technische Maßnahmen ableiten zu können.

2.1 Markov-Entscheidungsprozess

Allgemein können Reinforcement-Learning-Probleme mit Hilfe eines Markov-Entscheidungsprozesses (engl. Markov decision process) [Cas, S. 636] formalisiert werden.

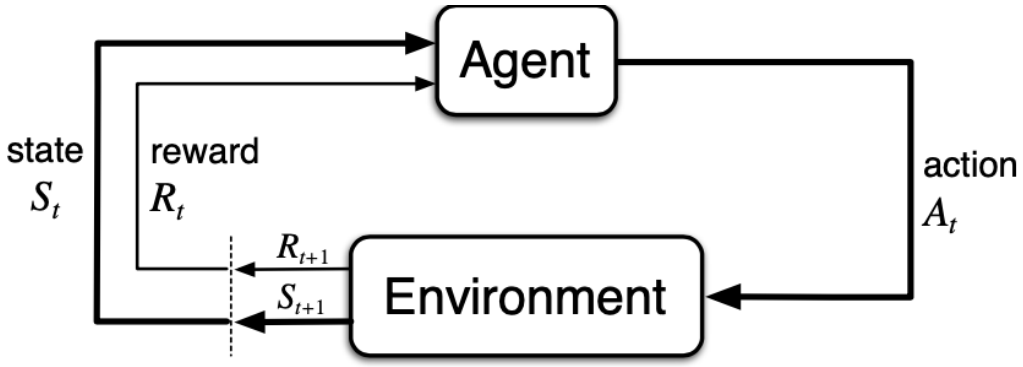

Abbildung 2.1: Ablauf des Markov-Entscheidungsprozess

Der Ablauf wird in Abbildung 2.1 dargestellt und kann laut [Cas, S. 636] als ein „[...]discrete state-time transition system [...]“ beschrieben werden. Die zwei Hauptbeteiligten Objekte sind Agent und Umwelt [SB18, S. 47ff.]. Ein Agent beschreibt eine mit seiner Umwelt interagierende Entität und kann über Software oder Hardware abgebildet werden. Alle Variablen sind abhängig von diskretisierten Zeitpunkten t. Der Agent erhält zu jedem Zeitpunkt den aktuellen Zustand (engl. state) s. Er kann dann mit der Umwelt z.B. durch Aktoren interagieren, indem er eine Aktion (engl. action) $a \in A_t$ ausführt, wobei die Menge A_t alle Aktionen, die der Agent zum Zeitpunkt t ausführen kann, beschreibt. Die Wahrscheinlichkeit für den Übergang in den Folgezustand S_{t+1} wird beschrieben durch die Verhaltensstrategie $\pi(a|s_t)$ [SB18, S. 58]. Für jedes Zustands-Aktions-Paar wird durch eine Belohnungsfunktion eine Belohnung $r_t(s_t, a_t)$ berechnet [Cas, S. 638]. Ziel ist eine möglichst optimale Abfolge von Aktionen zu ermitteln, um die Summe der Belohnungen zu maximieren.

2.2 Begriffe und Eigenschaften von Reinforcement Learning Verfahren

Dadurch, dass Reinforcement Learning nicht einen einzelnen Algorithmus beschreibt, sondern ein Paradigma bestehend aus einer Vielzahl von Algorithmen, ist es im Hinblick auf das Ziel dieser Arbeit sinnvoll, die unterscheidenden Merkmale der Verfahren zu betrachten. Die Begriffe werden als Entscheidungskriterien in Abschnitt 6.2 hinzugezogen, sodass insbesondere im Hinblick auf die Kompabilität mit den in Abschnitt 6.1 definierten Werten die Wahl eines geeigneten Verfahrens vereinfacht werden kann.

2.2.1 Belohnung und Wertfunktion

Als Reaktion der Umwelt auf eine Aktion des Agenten wird durch eine stochatische Funktion [SB18, S. 6] eine Belohnung zurückgeliefert. Die Belohnung ist die Grundlage für eine meist unverzügliche Anpassung der Verhaltensstrategie und dient als Indikator über die Güte der ausgeführten Aktion für den jeweiligen Zustand. Um nachhaltig sinnvolle Entscheidungen zu treffen existiert zusätzlich die Wertfunktion (engl. value function), welche den langfristigen Nutzen von Zuständen approximiert. So kann beispielsweise ein Zustand einzeln betrachtet stets in einer geringen Belohnung resultieren, langfristig jedoch von gut belohnten Zuständen gefolgt sein. Die Wertfunktion ordnet diesem Zustand einen hohen Wert zu, die Belohnungsfunktion hingegen einen niedrigen Belohnungswert.

2.2.2 Model-freie und Model-basierte Verfahren

Im Kontext der model-freien und der model-basierten Verfahren bezeichnet das Model die Kenntnis einer Abbildung der Umwelt und dessen Verhalten [SB18, S. 7]. Das Model der Umwelt wird abgebildet durch eine Zustandsübergangsfunktion [Li18, S. 14]. Wird ein model-basiertes Verfahren genutzt, so existiert ein solches Model über die Umwelt und kann genutzt werden, um Vorhersagen über die Auswirkungen von Aktionen zu treffen. Existiert kein Model muss ein model-freies Verfahren benutzt werden. Mit Hilfe des Versuchs- und Irrtumsprinzip [SB18, S. 7] versucht der Agent dann, je nach Verfahren ein eigenes Model der Umwelt zu erzeugen.

2.2.3 On-Policy und Off-Policy

On-Policy und Off-Policy beschreiben Lernverfahren, die sich insbesondere durch das Vorgehen bezüglich der Verwendung und Anpassung der Verhaltensstrategie bzw. im Fall des Off-Policy Lernens der Zielstrategie unterscheiden. Beim Off-Policy Lernen [Li18, S. 14] wird eine meist statische Vorgehensstrategie (engl. behavior policy) benutzt, welche das Verhalten des Agenten steuert. Zusätzlich gibt es eine Zielstrategie (engl. target policy), welche eine möglichst optimale Wertfunktion lernt. Die Zielstrategie wird basierend auf den erhaltenen Belohnungen für die ausgeführten Aktionen angepasst. On-Policy-Lernen benutzt nur eine einzelne Verhaltensstrategie, welche ähnlich wie die Zielstrategie des Off-Policy-Lernens angepasst wird, jedoch auch zur Steuerung des Verhaltens des Agenten benutzt wird.

3 Grundlagen der angewandten Ethik und Moral

Im Folgenden erfolgt eine Begriffsdefinition und Abgrenzung der Begriffe Ethik und Moral, welche als Grundlage für die in Abschnitt 6.1 definierten Werte dient. Im Rahmen dieser Arbeit wird insbesondere die angewandte Ethik betrachtet, da die spätere Konzeption des Vorgehens der Entwicklung einer Reinforcement-Learning-Anwendung stark praxisorientiert erfolgt.

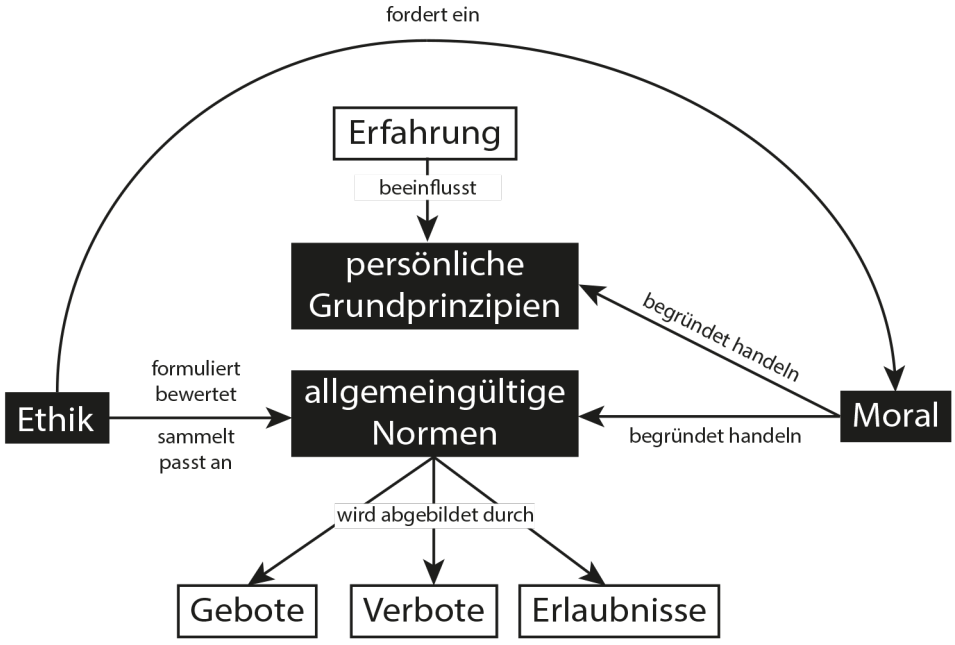

Abbildung 3.1: Überblick über die Beteiligten Komponenten des moralischen Handelns und dessen Zusammenhang.

Sowohl Ethik als auch Moral befassen sich mit Normen, welche in diesem Kontext als „[...] allgemeingültige Regeln [...]“ [SB10, S. 10 ff.] beschrieben werden. Die angewandte Ethik beschreibt dabei die Formulierung und Bewertung praxisbezogener Normen. Im Gegensatz zu praxisbezogenen Normen stehen Idealnormen als wünschenswerte aber praktisch nicht umsetzbare Werte. Als Disziplin der Philosophie fordert die Ethik, wie in Abbildung 3.1 abgebildet, nicht nur das Handeln gemäß der Normen ein, sie dient ebenfalls als Sammlung der Gesamtheit aller Normen. Nicht zuletzt stellt die Ethik auch die Grundlage für die Findung und Anpassung der Normen durch zwischenmenschlichen Diskurs in den Vordergrund. Auch wenn versucht wird, diesen Prozess rational zu begründen, kann eine Korrektheit nicht, wie etwa in naturwissenschaftlichen Disziplinen, nachgewiesen werden. Als Wissenschaft ist die Ethik unabhängig von Autoritäten, sondern basiert auf „[...] Kriterien der Rationalität, der Begründung und der Verallgemeinerungsfähigkeit [...]“ [SB10, S. 22]. Anders verhält sich die Gültigkeit von Normen im Bezug auf den aktuellen Zeitgeist [Tän08, S. 10]. Normen können zu jedem Zeitpunkt ihre Gültigkeit verlieren, was jedoch ihre gegenwärtige Gültigkeit und Relevanz nicht einschränkt. Auf persönlicher Ebene ergeben sich Werte und Normen aus persönlichen Erfahrungen und Interessen [McN88, S. 3f], was jedoch im Laufe des Lebens von äußeren Einflüssen, wie Religionen und unserer direkten Umwelt beeinflusst wird.

Moral bzw. moralisches Handeln ist als praktische Anwendung der, durch die Ethik diskutierten Normen [**werkener2017**] anzusehen. Die allgemeingültige Anwendung und Zusicherung von Normen zum Zwecke des moralischen Handelns erfolgt durch Grundprinzipien [SB10, S. 11], die es ermöglichen auch in vorher unbekannten Situationen gemäß der Normen moralisch zu handeln. Normen können durch „[...] Gebote, Verbote und Erlaubnisse [...]“ [WE17, S. 80] abgebildet und über die persönliche Moral hinaus formuliert und eingefordert werden. Sie finden speziell dann

Anwendung, wenn das Handeln nicht allein durch die Intuition entschieden werden kann.

3.1 Maschinenethik und Einordnung von Reinforcement-Learning-Agenten

Durch die Komplexität des menschlichen Lebens und Handelns ist es in der angewandten Ethik nur schwer möglich, Normen für jede Situation zu definieren, die allgemeingültig anwendbar sind und dennoch möglichst viele Fälle abbilden. Deshalb haben sich Bereitsethiken entwickelt, die explizit ein Teilgebiet des menschlichen Lebens abbilden. Bereichsethiken sind nicht unbedingt auf spezielle Verfahren oder Technologien, sondern viel mehr auf übergeordnete Anwendungen und ethische Fragen eines Systems bezogen. Ethik ist jedoch kein festgeschriebenes Regelwerk, weswegen im Rahmen dieser Arbeit für die Definition ethischer Werte eine Abstraktion und Generalisierung der verwandten Bereichsethiken und allgemeingültiger Normen erfolgt.

Inhaltlich besitzt die Maschinenethik [RKK19, S. 6] eine größe inhaltliche Nähe zur Thematik des Reinforcement Learning, indem die Anwendungen und ethischen Fragestellungen Überschneidungen aufweisen. Die Maschinenethik als solche beschreibt zum einen das Verhalten der Menschen im Bezug auf Maschinen und zum anderen, insbesondere im Kontext der lernenden Systeme, das moralische Handeln von Maschinen. Handeln Agenten z.B. durch die Begrenzung der Fähigkeiten und des Kontextes nicht implizit moralisch, so müssen Normen explizit für die Berücksichtigung in Maschinen technisch abgebildet werden [Ben19, S. 34]. Dem Agenten soll es dann möglich sein eigenständig in unbekannten Situationen moralisch zu Handeln. Als Teilgebiet der Ethik ist neben der Umsetzung und Konzeption der Normen auch

die Bewertung des Handelns zu betrachten, insbesondere mit dem Ziel das Handeln des Agenten unseren Erwartungen menschlicher Moral zu entsprechen. Grundlage der Betrachtung explizit ethischer Agenten ist die Frage, ob Maschinen überhaupt moralisch Handeln können. Laut [Ben19, S. 41 ff.] ist die Handlungsfähigkeit als Grundlage des moralischen Handelns an die Fähigkeit zur „[...] Orientierung an Gründen [...]“ [Ben19, S. 41] und an die „[...] Selbstursprünglichkeit [...]“[Ben19, S. 42] gekoppelt. Die Erfüllung dieser Bedingungen ist dabei abhängig von Art und Implementierung eines Systems. Damit eine Maschine also explizit moralisch handeln kann, muss sie zumindest grundlegend wie ein Mensch handeln. Die Fähigkeit zur Orientierung an Gründen erfüllen Maschinen, sobald ihr Verhalten an die Erfüllung eines Zieles gebunden oder zumindest dadurch motiviert ist. Zur Erfüllung der Selbstursprünglichkeit fordert [Ben19] die Eigenschaften Interaktivität, basale Autonomie und Adaptivität. Interaktivität und Adaptivität beschreiben dabei die Fähigkeit der Reaktion bzw. der Anpassung des Verhaltens auf äußere Einflüsse und basale Autonomie die Änderung des Zustands ohne äußere Einwirkung. Erfüllt ein System diese Eigenschaften, so besitzt es zumindest Handlungsfähigkeit und kann, wenn die Gründe für das Handeln einer Moral entsprechen auch moralisch handeln.

Im Bezug auf die Thematik des Reinforcement Learning besitzen die daraus entstehenden Agenten eine Orientierung an Gründen, indem die Belohnungsfunktion und das daraus resultierende Ziel die Belohnung zu maximieren das Verhalten beeinflussen. Zur Erfüllung der Selbstursprünglichkeit müssen zudem Interaktivität, basale Autonomie und Adaptivität betrachtet werden. Reinforcement-Learning-Agenten können im Sinne der Interaktivität auf äußere Einflüsse reagieren und das eigene Verhalten gemäß der Adaptivität ebenfalls anpassen. Indem die Agenten einen eigenen Antrieb haben, auch ohne äußere Einflüsse Zustände zu verändern ist auch die basale Autonomie vorhanden. Dadurch lässt sich darauf schließen, dass Reinforcement-Learning-Agenten zumindest ein gewisses Maß an Handlungsfähigkeit

besitzen. Damit die Agenten nun auch explizit moralisch handeln können, muss die Grundlage des Handelns auf Normen basieren. Insbesondere in der in Abschnitt 6.2 vorgestellten Maßnahmen zur Entwicklung von ethischen Reinforcement-Learning-Agenten wird die Umsetzung des moralischen Handelns der Agenten auf die in Abschnitt 6.1 definierten Normen bzw. Werten bezogen.

4 Regionale Zusicherung des ethischen Umgangs mit künstlicher Intelligenz

Im Folgenden werden Mittel zur Zusicherung ethischen Umgangs mit Anwendungen der künstlichen Intelligenz beispielhaft auf internationaler Ebene, sowie für die Regionen Europa, USA und Asien betrachtet. Die Zusicherung erfolgt aktuell meist durch Standards und Normen, wobei spezielle Vorgaben für Reinforcement Learning zum Zeitpunkt der Recherche nicht auffindbar sind. Deshalb wird die Betrachtung auf das Obergebiet der künstlichen Intelligenz ausgeweitet, da die meisten Vorgaben so weit gefasst sind, dass sie ihre Gültigkeit auch für den Teilbereich des Reinforcement Learning behalten. Eine Betrachtung dieser regionalen Zusicherung ist aus mehreren Gründen interessant. Die frühzeitige Beachtung der Standards ist insbesondere für die spätere technische Konzeption sinnvoll, da die Vorgaben möglicherweise in der Zukunft als Vorlage für Gesetze dienen können. Eine Auseinandersetzung mit Standards und Normen signalisiert Anwendern ein Interesse des Herausgebers an ethischen Fragestellungen. Die Werte sind zudem abhängig von den jeweiligen Rechts- und Kulturräumen. Die Vorgaben können die dahinterliegenden Wertmaßstäbe, sowie den Stellenwert der Thematik im jeweiligen Raum aufzeigen.

4.1 International

Auf internationaler Ebene werden Standards durch die International Organization for Standardization [isod] (kurz ISO) entwickelt. Die ISO besteht aus aus 164 länderspezifisischen Standardisierungseinrichtungen, welche unabhängig von den einzelnen Regierungen sind. So ist beispielsweise in Deutschland die jeweilige Standardisierungseinrichtung das Deutsche Institut für Normung [DIN] (kurz DIN), in der USA das American National Standards Institute [ANS] (kurz ANSI) und in China die Standardization Administration China [SAC] (kurz SAC). Ziel ist die Konsolidierung lokaler Standards und damit die Schaffung von Standards auf internationaler Ebene, um möglichst alle Dienste und Produkte „safe, reliable and of good quality“ [isoe] zu gestalten. Um das zu erreichen sind die Mitglieder in technische Komitees unterteilt, welche die tatsächlichen Standards entwickeln. Jedem Komitee wird ein Themenbereich zugewiesen, welcher dann Standards entwickelt, diskutiert und später aktualisiert.

Da der Prozess der Etablierung von Standards langwierig ist, werden explizit Normen betrachtet, die auf die Oberthematik der Ethik in IT-Systemen allgemein bezogen sind. Bestehende Normen, wie die ISO-Norm 13482:2014 für private Roboter und die ISO 10218:2012 [WS19, S. 50] für industriell eingesetzte Roboter geben Normen für einen Anwendungsbereich des Reinforcement Learning vor, gehen allerdings nicht auf die Technologie explizit ein. Eine speziellere Betrachtung der Thematik des maschinellen Lernens erfolgt durch das Subkomitee ISO/IEC JTC 1/SC 42. Zu den hier vorgeschlagenen Normen gehören [ISOb]:

- **ISO/IEC CD 22989** [ISOa]: Definition grundlegender Konzepte und Begriffe.
- **ISO/IEC CD 23053** [isoa]: Definition eines Frameworks zur Nutzung von maschinellen Lernens.

- **ISO/IEC AWI 23894** [isob]: Risikomanagement für Künstliche Intelligenz.
- **ISO/IEC AWI TR 24368** [isoc]: Ethische und soziale Bedenken.

Die meisten der Normen des Subkomitees sind allerdings zum aktuellen Zeitpunkt erst in der Konzeptionsphase und müssen noch diskutiert, geprüft und veröffentlicht werden, bevor sie in Kraft treten. Insgesamt liefert die ISO einen wichtigen Beitrag zur Erstellung weltweit gemeinsamer Standards. Die Normen sind nicht rechtsbindend, werden aber teilweise als Grundlage für Gesetze genutzt und bieten die Möglichkeit für Herausgeber entsprechender Systemen die Einhaltung dieser Standards nachzuweisen.

4.2 Europa

In Europa gibt es neben nationalen Vereinigungen, wie Bitkom [bit] oder der Plattform Lernende Systeme [plsb] auch übernationale Vereinigungen, wie Claire [CLA], AI4People [AI4] oder der HEG-KI [Smu]. Die Vereinigungen verstehen sich zum Teil als Dachverbände und Vertreter teilnehmender Unternehmen oder als Gruppen, welche sich aus Experten der Forschung und Wirtschaft zusammensetzen. Sie unterscheiden sich durch die Zusammensetzung der Mitglieder, sowie durch den Aufbau und die Ziele. Im Fall von AI4People bestehen die Mitglieder z.B. aus internationalen Unternehmen, wie Facebook, Intel und Microsoft. Im Gegensatz dazu besteht beispielsweise die Plattform Lernende Systeme etwa zum gleichen Anteil aus Vertretern der Wirtschaft und der Wissenschaft und ist in sieben Arbeitsgruppen zu unterschiedlichen Domänen des Oberthemas lernende Systeme aufgeteilt. Dort gibt es z.B. explizit eine Arbeitsgruppe zum Thema Ethik von lernenden Systemen [plsa].

Ethik-Leitlinien der HEG-KI

Im Folgenden sollen exemplarisch die Ethik-Leitlinien der Hochrangingen Expertengruppe für Künstliche Intelligenz (kurz HEG-KI) betrachtet werden. Die Expertengruppe wird offiziell von der europäischen Kommission eingesetzt und ist an alle Beteiligten gerichtet. Also vom Bürger über den Entwickler bis zu Unternehmen und Behörden. Es wird Wert auf den europäischen Ursprung der entwickelten Inhalte gelegt, um Unabhängigkeit zu wahren. Inhaltliche Grundlage ist die Verbesserung der Lebensqualität der Bürger, die Umsetzung von Nachhaltigkeit und die Minimierung möglicher Risiken. Das Handeln soll gemäß der europäischen Werte „Menschenrechte, Demokratie und Rechtsstaatlichkeit“ [Smu, S. 6] erfolgen. Die Pfeiler dieser Werte sind laut der HEG-KI Rechtmäßigkeit, Ethik und Robustheit. Ziel ist die Definition einer vertrauenswürdigen künstlichen Intelligenz. Dafür werden Grundsätze und Bewertungskriterien zur „Entwicklung, Einführung und Nutzung von KI-Systemen“ [Smu, S. 3] aufgezeigt. Das Vorgehen soll den gesamten Prozess der Entwicklung begleiten, möglichst unabhängig sein und sieht einen dauerhaften Abgleich mit den Anforderungen vor. Es ist eingeteilt in Fundamente, Verwirklichung und Bewertung. Die Zusicherung und Umsetzung der Maßnahmen ist unterteilt in technische- und nichttechnische Maßnahmen, wodurch möglichst alle Beteiligten miteinbezogen werden soll.

Ziel des Dokumentes ist die Standardisierung für die Entwicklung und Nutzung ethischer KI-Anwendungen. Es soll Vertrauen innerhalb der Gesellschaft schaffen, indem grundlegende Konzepte erklärt und jegliche Entscheidungen fachlich begründet werden. Allen Beteiligten soll ein Vorgehen aufgezeigt werden, mit dem ethische KI-Anwendungen strukturiert werden können. Auch wenn die Leitlinien zunächst nicht bindend sind, ist die Definition sinnvoll. So haben Herausgeber durch

die Richtlinien frühzeitig die Möglichkeit, dem Gesetzgeber, sowie Anwendern ein Bemühen bei der Umsetzung ethischer KI-Anwendungen nachzuweisen.

4.3 USA

In der USA gibt es das IEEE [IEE] als weltweit größter Fachverband mit über 420 000 Mitgliedern aus technischen Berufen in 160 Ländern [CH19, S. 287]. Ziel ist die Förderung von Innovation und Technologie zum Wohle des Menschen zu lenken. Deshalb hat das IEEE in „The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems“ diverse Aspekte des ethischen Softwaredesign betrachtet.

Neben dem allgemeinen Vorgehen werden Grundprinzipien, rechtliche Grundlagen, sowie Tipps zur Implementierung aufgelistet. Zu den Grundprinzipien zählen z.B. die Einhaltung und Beachtung der Menschenrechte, sowie die physische und emotionale Unversehrtheit des Menschen, die Nachvollziehbarkeit, Effektivität und ein hinreichendes Maß an Kompetenz der Entwickler. Inhaltlich werden explizit Entwickler angesprochen. Ziel des Dokumentes ist die Verbesserung des menschlichen Lebens durch KI-gestützte Maschinen, die explizit ethische Richtlinien befolgen und so den Menschen dienen [CH19, S. 6]. Die allgemeine Beschreibung stellt dabei den Anfang einer Reihe von tatsächlichen Standards in der IEEE-Gruppe „P7000 - Model Process for Addressing Ethical Concerns During System Design“[CH19, S. 283] [EME] dar.

4.4 Asien

Durch ein großes Engagement asiatischer Länder im Bereich der Förderung und Forschung [Uni, S. 46] Künstlicher Intelligenz werden im Folgenden Maßnahmen

asiatischer Länder betrachtet. So haben exemplarisch dafür im Jahr 2017 23 % aller KI-Unternehmen ihren Sitz in China gehabt [Din, S. 142]. China selbst hat mehrere Pläne veröffentlicht [Din, S. 8], um das Vorgehen für die Zukunft in unterschiedlichen Teilbereichen der künstlichen Intelligenz zu fördern, um bis 2030 das weltweite Zentrum der Forschung und Anwendung von KI zu sein. Deshalb wird im Folgenden betrachtet, wie ethische Kriterien asiatischer Länder zugesichert werden soll.

China, Singapur, Australien, Malaysia und Indien haben bereits Institutionen gebildet [MIT19], um ethische KI-Fragen zu diskutieren. So hat Indien beispielsweise mit #AIforAll [Aay18] die nationale Strategie veröffentlicht, in der insbesondere Fokus auf den Nutzen für die Gesellschaft und ethischen Umgang gesetzt wird. Lediglich Japan, Südkorea und Taiwan legen aktuell wenig bis keine Priorität auf ethische Bedenken von KI Anwendungen [MIT19], auch wenn einige dieser Länder in der Vergangenheit schon nicht rechtlich verbindliche Standards diskutiert haben. Viele Universitäten engagieren sich dabei für internationale Lösungen und arbeiten eng mit ihren jeweiligen Regierungen zusammen. Ziel ist der Aufbau von Vertrauen in die bestehenden Menschen- und Datenschutzrechte, ohne dabei Fortschritt und Innovation zu behindern. Nahezu alle Länder sind an der Diskussion beteiligt und sehen Potenzial in der Technologie als zukünftigen Wirtschaftszweig. Dennoch werden beispielsweise von China Anwendungen, wie die flächendeckende Nutzung von Gesichtserkennung [Moz18] genutzt, welche in der Form nicht vereinbar mit unseren europäischen Werten sind.

5 Diskussion

Im Folgenden wird zunächst der Kontext dieser Arbeit auf Grund der Komplexität der Thematik inhaltlich abgegrenzt. Anschließend werden Anwendungsgebiete des Reinforcement Learning aufgezeigt, um die Relevanz und mögliche Bedenken aufzuzeigen. Es werden zudem verwandte Arbeiten betrachtet und inhaltlich mit dieser Arbeit verglichen. Zuletzt werden Herausforderungen für die Entwicklung des Vorgehensplans betrachtet, welche als Anforderungen und spätere Bewertungsgrundlage dienen.

5.1 Abgrenzung

Wenn die Rede von einem autonomen System ist, so wird im Kontext dieser Arbeit insbesondere ein Reinforcement Learning gestütztes System betrachtet. Da Reinforcement-Learning-Systeme allerdings Überschneidungen mit anderen KI-Systemen aufweisen, sind die Aussagen teilweise auch auf diese übertragbar. Die Kernproblematik ist meist ähnlich, unterscheidet sich allerdings insbesondere durch die Lösungsansätze. Ziel der Arbeit ist die technischen Konzeption von ethischen Werten für Reinforcement-Learning-Anwendungen. Zu beachten ist dabei, dass der technische Ansatz im Rahmen dieser Arbeit nur ein Mittel ist, um die ethischen Werte umzusetzen. Nicht beachtet werden explizit technische Eigenschaften, wie beispielsweise Integrität, Interoperabilität oder Wartbarkeit, auch wenn diese

möglicherweise implizit durch die Zusicherung der ethischen Werte abgedeckt werden.

Das Oberthema der künstlichen Intelligenz ist sehr umfangreich, weswegen eine inhaltliche Abgrenzung sinnvoll ist. Zunächst wird als Technologie ausschließlich Reinforcement Learning betrachtet. Die im Rahmen der Arbeit betrachteten Agenten sind beschränkt auf die Interaktion mit der Umwelt. Eine Interaktion mit anderen Agenten in einer Multiagentenumgebung wird nicht betrachtet. Ebenso werden starke, also domänenübergreifende oder selbstverbessernde Agenten nicht betrachtet, da es zum Zeitpunkt der Arbeit im Rahmen der Recherche neben [Hal07] keine aktuellen Belege für den praktische Einsatz bzw. die Möglichkeit zur Nutzung gibt. Inhaltlich werden sowohl Reinforcement-Learning-Agenten in Form von Expertensystemen mit indirektem Einfluss, wie auch Agenten mit direktem Einfluss in Form von Robotern o.Ä. betrachtet, da beide Agententypen Anwendung finden und eine enge inhaltliche Nähe aufweisen. In Abschnitt 5.2 werden teilweise andere Arbeiten aufgezeigt, die abgegrenzte Themen behandeln.

5.2 Anwendungsgebiete und Stand der Technik

Im Folgenden wird der Stand der Technik in zweierlei Hinsicht betrachtet. So wird zunächst ein Überblick über Möglichkeiten und Anwendungsgebiete des Reinforcement Learning aufgezeigt und anschließend Ansätze der Umsetzung von Ethik in Reinforcement-Learning-Anwendungen betrachtet und mit der Konzeption des Entwicklungsprozess ethischer Agenten im Rahmen dieser Arbeit verglichen. Dadurch wird die Relevanz der Thematik, sowie Unterschiede zu bestehenden Arbeiten aufgezeigt.

5.2.1 Anwendungsgebiete

Reinforcement Learning wird speziell in Anwendungsbereichen eingesetzt, in denen der Agent autonom in fremden Situationen agieren muss. Im Folgenden werden als Anwendungsbereiche exemplarisch autonomes Fahren, autonome Waffensysteme und Einsatzgebiete im Gesundheitswesen betrachtet. Innerhalb dieser Anwendungsgebiete wird Reinforcement Learning zum einen als Aktor in Form von Robotik eingesetzt, aber auch als unterstützende Expertensysteme.

Autonomes Fahren

Fahrzeuge mit Fahrassistenzsystemen oder Teilautonomie sind schon jetzt ein relevantes Thema innerhalb der Automobilindustrie bei Unternehmen, wie beispielsweise Daimler, Google, Volkswagen oder Tesla, mit einem Gesamtmarktwert von über fünf Milliarden US-Dollar im Jahr 2018 [bus19]. Bestehende Arbeiten, wie [AS19] oder [NL09] beschreiben die Implementierung von Teilkomponenten des autonomen Fahrens in virtuellen Umgebungen, indem Aufgaben, wie die Kollisionsvermeidung gesondert betrachtet werden. [You+19] betrachtet als Teilproblem die intelligente Verkehrssteuerung, in dem das Verhalten autonomer Fahrzeuge im Bezug auf Verkehrsfluss und Verbrauchseffizienz optimiert wird. [YPB] beschreibt einen Einsatz des autonomen Fahrens, indem durch Deep-Q-Learning die Steuerung eines virtuellen Fahrzeuges erlernt und der Einfluss unterschiedlicher Belohnungsfunktionen auf den Fahrstil beobachtet wird. In [Sal+17] wird als Grundlage für die Entwicklung ein Frameworks für autonomes Fahrens vorgestellt, welches durch Kombination von Deep Reinforcement Learning und Rekurrenten Neuronalen Netzen (engl. recurrent neural network) grundlegende Funktionen, wie Steuerung, Abbildung von Umgebungen und die Interaktion mit anderen Fahrzeugen implementiert. Das Training autonomer Fahrzeuge ist durch die Natur der Fahrzeuge nur unter großem Kosten-

aufwand und mit beträchtlichem Risiko durchzuführen. So wird in [Pan+17] eine Möglichkeit vorgestellt, um simulierte Umgebungen mit Hilfe von realen Bildern abzubilden, wodurch die realitätsnähe simulierter Agenten verbessert werden soll.

Autonome Waffensysteme

Reinforcement Learning in autonomen Waffensystemen wird insbesondere bei der Entwicklung und dem Einsatz unbemannter Luftfahrzeuge (engl. unmanned aerial vehicles, kurz UAV) in Form von Quadrocoptern verwendet, da diese durch ihre Bauform und die daraus resultierende Flugstabilität diverse Vorteile gegenüber anderen Bauformen bieten [BVE10]. Ähnlich wie bei Anwendungen des autonomen Fahrens werden auch bei autonomen Waffensystemen Teilprobleme betrachtet. So beschreibt [BVE10] in [BVE10] die Nutzung verschiedener Reinforcement-Learning-Verfahren für die Steuerung von UAVs. Zusätzlich zur tatsächlichen Steuerung werden z.B. in [Zha+15] Möglichkeiten zur Identifikation einer möglichst optimalen Route für mehrere UAVs unter Betrachtung der Routenlänge und der Risikobewertung durch andere Agenten aufgezeigt. [Koc+19] beschreibt den Einsatz von Reinforcement Learning für die Höhensteuerung im Hinblick auf Stabilität und Kontrollmöglichkeiten als weitere Teilkomponente für die Steuerung von UAVs. Neben der eigentlichen Fortbewegung von autonomen Waffensystemen findet Reinforcement Learning auch Einsatz in Abwehrsystemen. So wird in [Xia+18] beschrieben, wie die Verbindungsqualität von UAVs durch Reinforcement Learning gegen Störsignale gesichert werden kann, indem die Kommunikation und mögliche Angriffe im vorhinein simuliert und das Verhalten dementsprechend angepasst wird.

Gesundheitswesen

Im Gesundheitswesen ist der Einsatz autonomer Systeme wünschenswert, um medizinisches Personal zu entlasten, die Fehlerquote zu reduzieren und die begrenzte Menge an spezialisierten Experten zu skalieren. Neben dem Einsatz von Robotik im Gesundheitswesen ist die Anwendung von Expertensystemen sinnvoll, um die Gefährdung von Patienten zu limitieren. Durch Nutzung von Expertensystemen muss die Durchführung der vorgeschlagenen Entscheidungen von Menschen getroffen werden, wodurch Entscheidungen frühzeitig hinterfragt werden können. So wird beispielsweise in [Liu+17] und in [MM18] Reinforcement Learning genutzt, um einen möglichst optimalen Behandlungsplan zu entwickeln und an mögliche Änderungen während der Behandlung anzupassen. In [ROY19] wird eine simulierte Operationsumgebung vorgestellt, welche gemäß bestehender Standards implementiert ist und mit entsprechender Schnittstelle genutzt werden kann. Ziel ist dabei, die Entwicklung weiterer Umgebungen zu motivieren und so die Vergleichbarkeit neuer Algorithmen zu gewährleisten. Zur Optimierung von Robotern im Gesundheitswesen wird in [Woo+] eine Möglichkeit zur Anpassung des Verhaltens gemäß der Vorgaben der Nutzen durch Inverse Reinforcement Learning vorgeschlagen. So können Roboter durch Beobachtung des Nutzers generalisieren und entsprechend der Erwartungen handeln.

5.2.2 Verwandte Arbeiten

Im Folgenden sollen Arbeiten aufgezeigt werden, die inhaltlich ebenfalls die Betrachtung ethischer Bedenken bei der Umsetzung von KI-Anwendungen, im Speziellen von Reinforcement-Learning-Anwendungen behandeln. Dadurch kann für die spätere Konzeption eine Grundlage der Entwicklung geschaffen werden und Alleinstellungsmerkmale dieser Arbeit herausgestellt werden.

Inhaltlich verwandt sind zunächst die in Kapitel 4 vorgestellten Beispiele zur Zusicherung ethischer Richtlinien. Darin werden beispielsweise in [Smu] Grundlagen der Technologie eingeführt und Vorstellungen bezüglich Ethik-Richtlinien von der Europäischen Union oder in [CH19] von der IEEE aufgezeigt. Dabei sind alle Maßnahmen bezogen auf künstliche Intelligenz als Oberthema. Im Vergleich zu dieser Arbeit wird der Fokus nicht speziell auf Reinforcement Learning, sondern höchstens auf potenzielle Anwendungen gelegt. Weiterhin werden nicht explizit technische Maßnahmen vorgeschlagen, sondern eher Richtlinien, die bei Umsetzung individuell erarbeitet werden müssen.

In [CBB] wird moralisches Handeln in Multiagenten-Systemen betrachtet. Dabei wird in [CBB] ein Model entwickelt, welches das eigene und das Verhalten der anderen Agenten in der Umgebung ethisch bewertet. Ähnlich zum Vorgehen im Rahmen dieser Arbeit basiert die ethische Bewertung auf Grundlage moralphilosophischer Konzepte. Unterschiedlich ist allerdings die Vorgehensweise. So werden in [CBB] die Normen und moralischen Prinzipien technisch abgebildet, um das Handeln gemäß der eigenen Einschätzung zu beschränken. Die Normen müssen zunächst definiert und bei der Anwendung regelbasiert abgefragt werden. Daraus ergeben sich zwei Probleme. Zum einen müssen die Normen formalisiert werden, was insbesondere bei komplexen Situationen nicht trivial bzw. nicht realistisch umsetzbar ist. Zum anderen wird das moralische Handeln nicht als Teil des Lernens miteinbezogen, wodurch der Agent kein eigenes Rechtsbewusstsein erhält. Anders als bei dieser Arbeit werden zudem Multiagenten-Umgebungen betrachtet und ein regelbasiertes Vorgehen vorgeschlagen. So wird im Rahmen dieser Arbeit ein allgemeines organisatorisches und technisches Vorgehen und nicht eine einzelne Maßnahme beschrieben.

In [Noo+] wird ein Agent erstellt, der mit Hilfe von Inverse Reinforcement Learning

ethischen Normen folgen soll. Im Gegensatz zu Cointe, Bonnet und Boissier werden die Normen nicht regelbasiert vor der Entscheidungsfindung abgefragt, sondern durch Beobachtung von Experten gelernt. Neben einer Verhaltensstrategie zur Belohnungsmaximierung existiert dann eine zweite Verhaltensstrategie, die gemäß des beobachten Verhaltens moralisch handelt. Je nach Situation wird dann nachvollziehbar entschieden, welche Aktion und damit welche Verhaltensstrategie sinnvoller ist. Dieser Ansatz in [Noo+] lässt, ähnlich wie in [CBB], jegliche organisatorische Maßnahmen außen vor. Ebenso kann das Verfahren durch die Wahl der jeweiligen Verfahrensstrategie lediglich entscheiden, ob effizient oder ethisch gehandelt wird. Im Rahmen dieser Arbeit ist das Ziel, den Agent möglichst zu jedem Zeitpunkt moralisch handeln zu lassen und organisatorische und technische Maßnahmen zu vereinen.

5.3 Herausforderungen

Im Folgenden werden Herausforderungen betrachtet, die bei der Konzeption technischer Maßnahmen zur Erstellung eines Vorgehensplans für die Entwicklung ethischen Reinforcement Learnings auftreten können.

5.3.1 Subjektivität moralischen Handelns

Agenten können die Werte der Entwickler widerspiegeln, indem unterschiedliche Arten der Beeinflussung (engl. bias) [Sen+18] bewusst oder unbewusst in die Datenerhebung und die spätere Entwicklung einfließen. Beispiele dafür sind z.B. Stereotypisierung oder eine einseitige Auswertung und Anpassung bezüglich der Klassenverteilung. Dadurch, dass die Entwicklung von Reinforcement-Learning-Anwendungen oft in größeren, heterogenen Teams geschieht, sollten Maßnahmen ergriffen werden,

damit die Beteiligten ein, mit den geforderten Werten an die Anwendung, kompatibles Wertesystem besitzen. Ebenso sollten Maßnahmen ergriffen werden, um Verzerrungen zu identifizieren und diese je nach Grad der Abweichung als Fehler zu betrachten und Möglichkeiten der Beseitigung dieser Verzerrung in Betracht gezogen werden.

5.3.2 Problem der Formalisierung

Moralisches Handeln ist abhängig von der jeweiligen Situation. Einfache Entscheidungen werden meist intuitiv getroffen und können dementsprechend nur schwer formalisiert werden. Im Gegensatz dazu stehen Situationen, in denen durch Komplexität und Uneindeutigkeit moralischer Prinzipien Dilemma [MdE03, S. 300] entstehen. In Diesem Fall ist die Entscheidungsfindung hochgradig dynamisch und abhängig von einer analytischen Denkweise, in der versucht wird, mögliche Folgen abzuwägen. Es stellt sich neben der Problematik der Formalisierbarkeit der intuitiv-emotionalen Entscheidungen die Problematik bezüglich der technischen Abbildung des komplexen Prozesses im Falle eines moralischen Dilemmas. Neben der Formalisierung der Entscheidungsprozesse müssen diese Werte und Normen nach denen moralisches Handeln erfolgent formalisiert werden, um in Agenten technisch abgebildet werden zu können.

5.3.3 Mangelnde Zertifizierung

Sicherheitskritische Anwendungen können Vertrauen gewinnen, indem die Korrektheit der Implementierung und Prozesse durch vertrauenswürdige Institutionen geprüft und anschließend bestätigt wird. Projektebezogene Prozesse und Grundlagenalgorithmen können durch gängige Institutionen und Normen zertifiziert werden. Durch die Aktualität und insbesondere die Komplexität von Reinforcement Learning

konnte im Rahmen der Recherche jedoch keine explizite Möglichkeit zur Zertifizierung speziell von Reinforcement Learning Systemen im Gesamten identifiziert werden. Zwar gibt es Ansätze z.B. vom Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme [Cre+] für die Zertifizierung von KI-Anwendungen, allerdings wird zum Zeitpunkt der Arbeit lediglich ein Prüfkatalog als Grundlage der späteren Zertifizierung entwickelt. Umso wichtiger ist es dann, Vertrauen aufzubauen und Korrektheit transparent und nachvollziehbar selbständig nachzuweisen.

5.3.4 Innovation, Fortschritt und Adaption

Insbesondere auf internationaler Ebene versuchen alle Beteiligten einen Vorsprung im Bereich künstlicher Intelligenz zu erlangen. Treiber dieses internationalen Wettlaufs sind dabei Innovation und Fortschritt. Dem gegenüber steht das Potenzial des Reinforcement Learning Wohl und Würde des Menschen zu gefährden. Die Definition expliziter ethischer Werte und die daraus resultierenden Maßnahmen sollen die Wahrung fördern und die Relevanz der Einhaltung eben dieser aufzeigen. Es gilt, blinden Fortschritt zu verhindern und stattdessen frühzeitig ethische Werte zu diskutieren, sowie Maßnahmen zur Zusicherung zu identifizieren. Optimalerweise unterstützt die Betrachtung ethischer Werte die Adaption in der Gesellschaft, Wirtschaft und Politik, indem Vertrauen aufgebaut wird. Um dies zu unterstützen bedarf es neben dem Einfordern der Gesellschaft der Berücksichtigung ethischer Bedenken und gesetzlichen Regelungen seitens der Politik, technische und organisatorische Maßnahmen möglichst kompatibel mit bestehenden Prozessen und Technologien zu definieren und aufzuzeigen, inwiefern ethisches Verhalten wirtschaftlich sinnvoll sein kann.

Neben dem Bestreben nach Fortschritt stehen allerdings auch Bedenken der Gesellschaft. Eine breite Adaption innerhalb der Gesellschaft ist Grundlage der Beteiligung

von Unternehmen am Fortschritt zwecks wirtschaftlichen Interesses. Die Adaption wird z.B. durch die Diffusionstheorie [Kar13, S. 513 ff.] beschrieben. Auf persönlicher Ebene entsteht die Entscheidung zur Adaption als Folge gewisser Grundvorrausetzungen, dem Wissensstand und der persönlichen Abwägung bezüglich der Nutzung. Zu den Grundvoraussetzungen gehören Erfahrungen mit ähnlichen Technologien, die Existenz von Probleme, die die Technologie löst, das Maß an Innovation und der persönliche Kontext. Durch Hinzunahme des persönlichen Verständnisses der Technologie wird abgewägt, wie groß der Nutzen im Verhältnis zur Komplexität der Adaption und wie kompatibel es mit der persönlichen Situation ist. Um eine Adaption zu fördern müssen diverse Herausforderungen betrachtet werden. Dazu gehört die digitale Kluft [Rog16] als Wissenlücke bezüglich digitaler Technologien zwischen gesellschaftlichen Gruppen. Ebenso müssen Fähigkeiten, Limitierungen der Technologie transparent gemacht und relevante Nutzungskontexte für die Gesellschaft identifiziert werden. Zusätzlich muss ein Mindestkenntnisstand etabliert, sowie die Komplexität der Anwendungen auf ein Maß reduziert werden, um die Technologie attraktiv zu gestalten.

Grundlage der persönlichen Entscheidung zur Adaption ist die Kenntnis über den tatsächlichen Einsatz der Technologie. Zum Zeitpunkt der Arbeit besteht allerdings keine Kennzeichnungspflicht bezüglich des Einsatzes von KI-Technologien. Reinforcement Learning ist dabei oft nicht eine eigenständige Anwendung, sondern Teilkomponente in bestehenden Systemen. Die Entscheidung bezüglich der Adaption ist dabei nur dann relevant, wenn der Nutzer auch einen tatsächlichen Einfluss auf die Nutzung haben kann. So haben beispielsweise Menschen in Kriegsgebieten keinen Einfluss auf den Einsatz autonomer Waffensysteme. Umso wichtiger ist dann die Schaffung gesetzlicher Rahmenbedingungen, bezüglich des Einsatzes von Reinforcement Learning und der Beachtung gewisser ethischer Werte und der Umsetzung geeigneter Maßnahmen zur Zusicherung dieser Werte.

6 Konzeption

Im Folgenden werden zunächst ethische Werte basierend auf den in Kapitel 3 eingeführten Grundlagen der Ethik und den in Kapitel 4 betrachteten Maßnahmen internationaler Institutionen begründet. Anschließend wird ein Vorgehensplan vorgestellt, der technische und organisatorische Maßnahmen enthält, mit denen entlang eines üblichen Softwareprojektablaufs gemäß [BK13, S. 64] die Zusicherung der ethischen Werte umgesetzt werden kann.

6.1 Definition ethischer Werte

Die Wahl ethischer Grundwerte für die Entwicklung und Anwendung von Reinforcement-Learning-Anwendungen wird im Folgenden auf Grundlage ethischer Prinzipien, sowie bestehender Standards und Normen begründet. Ziel ist eine Konsolidierung bestehender Ansätze für die in Abschnitt 6.2 vorgestellte Konzeptionierung eines praktischen Maßnahmekataloges. Die Werte sollen dabei klar verständlich und möglichst allgemeingültig sein. Überkulturelle Grundlage ist im Allgemeinen die Achtung der Menschenrechte [Nat48]. Oberstes Ziel sollte die Zusicherung und Wahrung eben dieser sein. Zusätzlich macht es Sinn, ethische Werte zuzusischern, um neben der Unversehrtheit von Wohl und Würde des Menschen, auch Vertrauen in die Technologie aufzubauen. Die Wahl der Werte wird in den einzelnen Unterkapiteln näher erläutert, basiert jedoch stark auf den in Kapitel 4 vorgestellten Normen und Standards der

einzelnen Regionen. Um nicht von einzelnen Verfahren innerhalb der Technologie und Domäne abhängig zu sein, sind die Werte möglichst allgemein gehalten. Bei der Definition auftretende Fragen sollen im Rahmen dieser Arbeit nicht unbedingt behandelt werden, sondern dienen viel mehr zu Veranschaulichung der Relevanz und der Problematik dieser Werte.

6.1.1 Nachvollziehbarkeit und Erklärbarkeit

Die ethische Entscheidungsfindung besteht laut [**tannsjo2018**] aus zwei Komponenten. Zum einen aus der Anwendung gelernter moralischer Prinzipien und zum anderen aus der Hinzunahme möglichst aller relevanter Informationen. Dieser Prozess der ethischen Entscheidungsfindung ist insbesondere bei der Bewertung eines Agenten relevant. Ein Agent soll gemäß unserer menschlichen Werte handeln und muss deshalb auch gemäß dieser bewertet werden. Nachvollziehbarkeit bezieht sich dabei zum einen auf die Kenntnis über vergangene und möglicherweise die Absicht der Ausführung zukünftiger Aktionen und zum anderen auf die Fähigkeiten und Limitierungen des Systems, sowie auf den Entstehungsprozess. Insbesondere für Entscheidungsprozesse von domänenfremden Verantwortungsträgern müssen Güte und Entstehungsprozess nachvollziehbar sein, dadurch dass Weiterentwicklung und Optimierung darauf basieren. Erklärbarkeit fordert zusätzlich, dass Entscheidungen nicht nur nachvollzogen, sondern begründet werden können. Die Werte Nachvollziehbarkeit und Erklärbarkeit bieten die Grundlage für die Zusicherung der folgenden zwei Werte. So kann das in Unterabschnitt 6.1.2 geforderte Vertrauen nur dann aufgebaut werden, wenn durch das System ein Verständnis seitens der verantwortlichen Personen, insbesondere des Entwicklers über das Handeln entsteht und Personen im Anwendungskontext auf Grund der Erklärbarkeit ihr Handeln ggfs. anpassen und das des Agenten einschätzen können. Die Frage der in Unterabschnitt 6.1.3 beschriebenen Verantwortung basiert zudem auf der Grundlage, dass alle verantwortwortlichen

Personen das Verhalten des Agenten kennen, nachvollziehen können und der Prozess der Entscheidungsfindung mit den persönlichen Werten übereinstimmt.

6.1.2 Vertrauen durch Kalkulierbarkeit und Zuverlässigkeit

Vertrauen ist insbesondere bei der Adaption ein wichtiger Treiber und damit Grundlage des Fortschrittes der hier betrachteten Technologien. Vertrauen entsteht dann, wenn alle beteiligten Parteien eine gemeinsame „[...] Basis geteilter Normen, Werte und positiver Zukunftserwartungen [...]“[Gil, S. 61] besitzen. Es ermöglicht Kalkulierbarkeit unabhängig von rechtlichen Verpflichtungen [Gil, S. 61 ff.]. Vertrauen kann in diesem Anwendungsbereich in der Regel nicht in einzelne Personen aufgebaut werden, da bei komplexen Softwaresystemen oft große Teams zum Einsatz kommen und das Produkt somit Ergebnis der Zusammenarbeit vieler ist. Deshalb muss das Unternehmen als ganzes ein Systemvertrauen aufbauen, indem Anwender auf die Prinzipien des Unternehmens vertrauen können. Zuverlässigkeit in Form von Korrektheit und Robustheit in unbekannten Situationen können helfen Vertrauen langfristig aufzubauen [Avi+04, S. 13]. Die Anforderungen des Systems, sowie das tatsächliche Handeln müssen zum Aufbau von Vertrauen mit den Erwartungen übereinstimmen und können nur erfüllt werden, wenn das System kalkulierbar handelt.

6.1.3 Verantwortung und Schuld

Sobald Agenten Einfluss auf Menschen nehmen können, muss geklärt werden, wer Verantwortung für die Folgen des Handelns eines Agenten und damit die Schuld im Fehlerfall trägt [HW98, S. 88]. Verantwortung ist dabei die notwendige Bedingungen, um tatsächlich Schuld zuzuweisen [DIK09, S. 75]. Die Verantwortung wird zum einen von der Umwelt gefordert [Hof19, S. 1] und zum anderen von Innen durch das

persönliche Gewissen der Verantwortlichen erbracht. Durch korrekte Dokumentation kann für einzelne Teile eines Systems zwar die Frage der verantwortlichen Person ermittelt werden, gerade bei selbstlernenden Systemen sind die verursachenden Verhaltensweisen jedoch nicht explizit implementiert, sondern entstehen möglicherweise als unerwünschte Randeffekte. Deshalb ist die Frage der Verantwortung und Schuld insbesondere bei selbstlernenden Systemen nicht trivial und sollte im Laufe des Produktlebenszyklus betrachtet werden. Eine Bestrafung im Sinne der Schuld trägt bei Maschinen keine Wirkung. Die Frage der Schuld betriftt also die Verantwortlichen. Laut [DIK09, S. 76] „[...] kann man niemanden für die Verletzung einer Norm zur Rechenschaft ziehen, die er gar nicht einhalten konnte [...]“. Im Bezug zu selbstlernenden und insbesondere autonomen Systemen stellt sich also die Frage, ob ein Verantwortlicher die Verletzung der Norm beeinflussen kann. So ist beispielsweise ein Kriegseinsatz nur dann rechtmäßig, wenn die Verantwortung für mögliche Tode eindeutig geklärt ist. Zu klären bleibt dann, inwiefern ein Einsatz gerechtfertigt ist, wenn z.B. ein autonomes Waffensystem zwar weniger Fehler macht als Menschen, die Schuldfrage jedoch nicht eindeutig geklärt werden kann.

6.2 Vorgehensplan zur Zusicherung ethischer Werte

Die Umsetzung der in Abschnitt 6.1 definierten ethischen Werte soll entlang eines üblichen Softwarelebenszyklus [BK13, S. 64] erfolgen, da die ethischen Probleme in allen Phasen Beachtung finden müssen.

Die in Abbildung 6.1 skizzierten Maßnahmen werden im Folgenden hinsichtlich der ethischen Werte begründet. Das Vorgehen soll kompatibel mit den meisten Vorgehensmodellen sein, indem die Phasen so allgemein gefasst sind, dass sie nahezu in allen Vorgehensmodellen in irgendeiner Form vorkommen. Die Phasen können parallel zu den jeweiligen Phasen im Vorgehensmodell durchgeführt werden. Viele

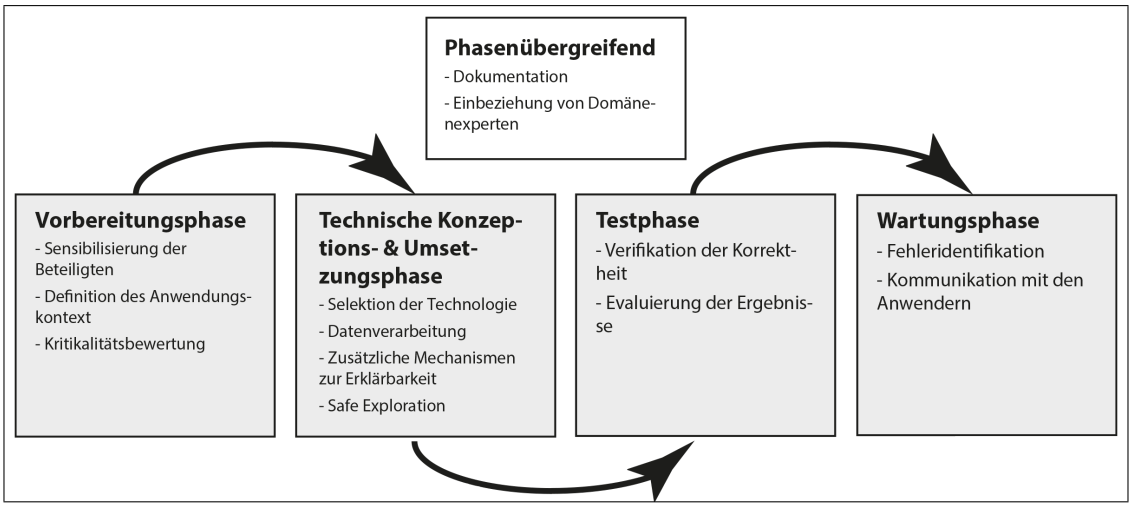

Abbildung 6.1: Phasen und jeweilige Maßnahmen der Zusicherung ethischer Werte.

Maßnahmen behalten ihre Gültigkeit über die jeweilige Phase hinweg, sollten jedoch spätestens ab der vorkommenden Phase beachtet werden. Es wird kein spezieller Fokus auf einzelne Domänen gesetzt. Stattdessen werden allgemeine Hinweise und Möglichkeiten aufgelistet, wodurch eigene Maßnahmen abgeleitet werden können. Viele der Maßnahmen beziehen sich im Speziellen auf sicherheitskritische Anwendungen. Innerhalb der Phasen wird zunächst das Ziel der jeweiligen Phase erläutert und anschließend technische und organisatorische Maßnahmen vorgestellt. Ziel dieses Vorgehens ist die Definition von Leitlinien, die mit den in Abschnitt 6.1 vorgestellten Werten kompatibel sind und in bestehende Prozesse integriert werden können.

Vorweg werden Maßnahmen eingeführt, die phasenübergreifend gültig sind und ggfs. innerhalb der einzelnen Phasen spezieller betrachtet werden. So ist Dokumentation in jeder Phase ein wichtiges Mittel, um das Verhalten nachvollziehen zu können, Vertrauen durch Transparenz aufzubauen und Verantwortung zuordnen zu können. Welche Informationen dabei in den einzelnen Phasen relevant sind, wird an passender Stelle jeweils beschrieben. Alle dokumentierten Informationen sollten

langfristig und in einem menschlich lesbaren Format gespeichert werden. Optimalerweise wird die Dokumentation ähnlich wie eine Lizenzangabe mit der Software bzw. dem Agenten ausgeliefert. In dem Fall muss die Dokumentation möglichst klar und dem Wissensniveau des Anwenders entsprechend formuliert sein. Neben einer ausführlichen Dokumentation ist phasenübergreifend eine Miteinbeziehung von Domänenexperten sinnvoll [Got+18, S. 12 ff.]. Domänenexperten können schon während des Entstehungsprozesses Einfluss auf die inhaltliche Korrektheit ausüben. Um ethisches Handeln zuzusichern, kann zudem z.B. in allen Entwicklungsschritten eine explizite Aufforderung der Übereinstimmung der formulierten Werte und der moralischen Prinzipien der Mitarbeiter gefordert und die Zustimmung oder mögliche Mängel dokumentiert werden. Ebenso fordert die Miteinbeziehung von Domänenexperten implizit Erklärbarkeit und Nachvollziehbarkeit ein, indem zum Verständnis der Experten Aktionen des Agenten nichttechnisch aufbereitet werden müssen. Domänenexperten können zudem genutzt werden, um die Realitätsnähe simulierter Umgebungen und des Aktionsraumes zu validieren. Reinforcement Learning tendiert durch Fehler, wie Reward Hacking dazu, nicht gängige Methoden zur Belohnungsmaximierung zu finden. Um dies zu vermeiden können Domänenexperten das Verhalten der Agenten einordnen und so unübliche frühzeitig Methoden identifizieren.

6.2.1 Vorbereitungsphase

In der Vorbereitsphase sollen zunächst grundlegende Informationen gesammelt und der Anwendungskontext klar abgegrenzt werden. Dafür werden zusätzlich organisatorische Anforderungen aufgezeigt, um die ethischen Werte zu berücksichtigen. Es soll eine klare Definition des Einsatzzweckes, der Absichten und der Anforderungen definiert werden und so eine solide Grundlage für die späteren Phasen entstehen.

Sensibilisierung der Beteiligten

Zur Umsetzung der Maßnahmen müssen alle beteiligten Personen für das Thema Ethik sensibilisiert sein. Dadurch können bereits frühzeitig personelle Probleme erkannt und Lösungen gefunden werden. Die Motivation für die Durchsetzung ethischer Maßnahmen kann sowohl von außen, also extrinsisch, als auch von den Personen selber, also intrinsisch, erfolgen [Bau17]. Intrinsische Motivation kann nur begrenzt beeinflusst werden und ist hauptsächlich vom Wertesystem des Einzelnen, aber auch von seiner Umgebung abhängig. Einfluss kann darauf durch extrinsische Maßnahmen ausgeübt werden, indem beispielsweise die sozialen Normen innerhalb des Unternehmens Ethik als Kernthema enthalten und durch unabhängige Instanzen überprüft und eingefordert werden. Um Teil der Gemeinschaft zu sein kann sich so beim Einzelnen eine intrinsische Motivation entwickeln, die mit den Unternehmensnormen übereinstimmt. Auf der anderen Seiten kann eine extrinsische Motivation durch das Unternehmen oder durch die Politik und Gesellschaft eingefordert werden. Maßnahmen dazu sind beispielsweise Gesetze, welche den Einsatz bestimmter Technologien verbieten. Ebenso können Vorgaben des Unternehmens und daraus potenziell resultierende Strafen oder umgekehrt Belohnungen bei Einhaltung eine Möglichkeit sein, um ethisch korrektes Handeln zu motivieren.

Definition des Anwendungskontext

Im Folgenden werden Maßnahmen und Fragestellungen betrachtet, die zur klaren Definition des Einsatzzweckes, der Absichten und Anforderungen relevant sind. Die Maßnahmen sollten bereits bei der Erhebung der Anforderungen beachtet werden.

Welche Aufgaben soll das System haben?

Eine realistische und präzise Definition der Aufgaben und Ziele des Systeme bietet die Grundlage der Entwicklung. Gemäß „AI is not magic“ [Got+18, S. 13] sind hier bereits Schwächen und Grenzen der zu nutzenden Technologie zu beachten.

In welchem Kontext soll das System eingesetzt werden?

Zu beachten sind Rechts- und Kulturräume, öffentliche und private Einsatzzwecke, sowie Einzel- und Multiagentenumgebungen und die Einbettung in andere IT-Systeme.

Welche Limitierungen soll das System haben und welchen Werten soll es folgen?

KI-gestütze Systeme sollten einen klaren Anwendungskontext besitzen und dementsprechend auch Limitierungen ihrer Funktionalität. Die klare Definition und öffentliche Kommunikation dieser Limitierungen kann helfen, Vertrauen aufzubauen. Ebenso sollte frühzeitig entschieden werden, welchen Werten das System folgen soll, da die Abbildung der Werte in gesamten Produktlebenszyklus beachtet werden muss.

Soll der Agent explizit oder implizit ethisch handeln?

Implizit ethische Agenten sind durch ihren Anwendungskontext nicht in der Lage unethisch zu handeln. In dem Fall muss ein Fokus darauf gelegt werden, den Agenten auf genau diesen Kontext zu begrenzen. Ist dies nicht der Fall, müssen für den Agenten explizit Maßnahmen ergriffen werden, um die ethischen Werte umzusetzen. Diese Maßnahmen werden in den Folgenden Phasen vorgestellt.

Wie erfolgt der Einfluss auf die Umgebung, insbesondere auf Menschen?

Agenten können z.B. als Expertensystemen einen indirekte Einfluss auf ihre Umwelt besitzen, indem die Entscheidungen von anderen Systemen ausgeführt werden müssen. Im Gegensatz dazu besitzen Agenten mit direktem Einfluss auf die Umwelt Aktoren, um selbständig mit der Umwelt zu interagieren. Die Entscheidung, um welche Art von Agent es sich im Bezug auf den Einfluss auf die Umwelt handelt, ist essentiell für die spätere Beachtung der Maßnahmen und hat einen großen Einfluss auf die Kritikalitätsbewertung.

Als Ergebnis entsteht neben einer notwendigen Grundlage für den späteren Entwicklungsprozess eine Kritikalitätsbewertung. Die Kritikalitätsbewertung gibt Aussage darüber, welche Mitarbeiter am Projekt beteiligt sein dürfen, wie der Umgang mit den dazugehörigen Daten aussehen muss, welche Verfahren zu wählen sind, was maximale Eingriffs- und Anpassungszeiten im Fehlerfall sind und in welchem Maße die Prozesse geprüft und zertifiziert werden müssen. Die in den späteren Phasen beschriebenen Maßnahmen können eine Hilfestellung geben diese Fragen zu beantworten.

6.2.2 Technische Konzeptions- und Umsetzungsphase

Im Folgenden werden Maßnahmen zum Softwaredesign basierend auf den in Unterabschnitt 6.2.1 erhobenen Anforderungen vorgestellt. Zusätzlich werden Technologien zur praktischen Umsetzung der in Abschnitt 6.1 definierten ethischen Werte im Kontext einer Reinforcement-Learning-Anwendung betrachtet. Dafür werden mögliche technische Probleme betrachtet, die mit der Umsetzung der ethischen Werte kollidieren und Maßnahmen aufgezeigt, die diesen Problemen entgegenwirken.

Selektion der Technologie

Die Wahl des Reinforcement-Learning-Verfahrens sollte durch die Beachtung ethischer Werte nicht auf wenige Einzelne beschränkt werden, sondern bestehende so angepasst und genutzt werden, dass sie den Anforderungen entsprechen können. Um Vertrauen zu gewinnen ist die Nutzung standardisierter Implementierungen von Grundlagenalgorithmen sinnvoll. Zum aktuellen Zeitpunkt konnten im Rahmen der Recherche keine Normen identifiziert werden, nach denen Reinforcement-Learning-Systeme zertifiziert werden können. Als Alternative zur Zertifizierung des Gesamtsystems ist es sinnvoll, Standardimplementationen einzelner Algorithmen zu benutzen. So bietet beispielsweise das Forschungsinstitut OpenAI [Ope] eine Sammlung an Standardimplementierungen [Dha+17]. Diese sind zwar nicht offiziell zertifiziert, werden aber durch OpenAI und durch die Open-Source-Gemeinschaft gepflegt und getestet. Auch wenn die Nutzung dieser Algorithmen Potenzial hat, eine bessere Vergleichbarkeit zu gewährleisten und die Korrektheit zu erhöhen, sollte ein intensiver Fokus auf die Testphase gelegt werden und so ein eigener Nachweis der Korrektheit erbracht werden.

Es existiert eine Vielzahl an Reinforcement-Learning-Verfahren mit unterschiedlichen Eigenschaften, die in Abschnitt 2.2 beschrieben wurden. Allgemeingültige Empfehlungen bezüglich eines einzelnen Verfahren sind nicht sinnvoll, da die Technologie sich stetig verändert und die Wahl des Verfahrens stark vom Anwendungskontext abhängig ist. Die Entscheidung sollte auf Grundlage der Abwägung der Stärken und Schwächen der einzelnen Verfahren und Eigenschaften getroffen werden. Bei der Wahl des Verfahren ist insbesondere die Komplexität der Umwelt zu beachten. Agiert der Agent ausschließlich in einer begrenzten Domäne, ist die Wahl eines model-basierten Verfahrens sinnvoll. So kann eine Repräsentation der Umwelt erzeugt werden, womit die Folgen von Aktionen im Vorhinein approximiert werden

können. Neben der Unterscheidung zwischen model-basierten und model-freien Verfahren ist zu entscheiden, ob gemäß On-Policy oder Off-Policy gelernt wird. Off-Policy-Learning tendiert eher zur Entdeckung (engl. exploration) [Her, S. 24] neuer Aktionsfolgen, wohingegen On-Policy-Learning eher in Richtung der bestehenden Verhaltensstrategie tendiert. Insbesondere bei sicherheitskritischen Anwendungen ist ein starker Fokus auf Entdeckung, zumindest in einer realen Umgebung, nicht wünschenswert.

Datenverarbeitung

Reinforcement Learning ist abhängig von den gesammelten Daten der Umwelt. Sind die Daten inkorrekt oder manipuliert, kann die Korrektheit und damit die Gesamtgüte des Systems kompromittiert werden. Auch wenn Reinforcement Learning ein Online-Learning-Verfahren ist, werden Daten möglicherweise transformiert, indem beispielsweise der Zustandsraum zeitlich diskretisiert wird. Ebenso kann von der Sammlung dieser Daten, insbesondere der Entscheidungen des Agenten profitiert werden. Um die Korrektheit der Daten sicherzustellen empfiehlt sich die Sicherstellung der Datenherkunft (engl. data provenance) [Olu+17, S. 3]. Data Provenance beschreibt die Dokumentation der Geschichte des Ursprungs und der Transformation der Daten. Durch ein entsprechendes Datenmodell wird ermöglicht nachzuvollziehen, wie die Daten entstanden sind und wie sie transformiert werden. Zusätzlich kann durch Data Provenance Poisoning als Angriff abgemildert werden, indem Veränderungen der Daten nachvollzogen werden können. Poisoning beschreibt einen Angriff, bei dem eine gezielte Veränderung der Umgebung vorgenommen wird, wodurch das Verhalten beeinflusst wird [Bar+17, S. 103]. Durch diese Maßnahmen kann konstant die Korrektheit und Neutralität der Daten-Pipeline sichergestellt werden, indem nachvollzogen werden kann, inwiefern die Transformation der Ausgangsdaten den Anforderungen entspricht.

Neben der Transformation der Daten sollte auch die Entstehung der Daten betrachtet werden. Insbesondere in praktischen Anwendungen dienen Sensoren oft als Datenquelle und können Abnutzungserscheinungen erleiden, indem sie verändert oder beschädigt werden [Kra09, S. 113]. Wünschenswert ist deshalb eine regelmäßige automatische und manuelle Überprüfung der Sensorik, wobei zu beachten ist, dass diese Veränderung ggfs. nicht auffällig ist. In diesem Fall sollte ein Fail-Safe-Mode vorhanden sein, da ein reguläres Abschalten des Systems ebenfalls eine Gefährdung darstellen kann.

Zusätzliche Mechanismen zur Erklärbarkeit

Je nach Verfahren weißt Reinforcement Learning ein unterschiedliches Maß an Erklärbarkeit auf. Wird beispielsweise ein tabulares Verfahren benutzt, so können die eigentlichen Entscheidungen transparent nachvollzogen werfen, dadurch dass die Daten und die Entscheidungsstrategie zu jedem Zeitpunkt bekannt sind. Bei Verfahren, die künstliche neuronale Netze [SB18, S. 187] nutzen gibt es das Problem der Erklärbarkeit, was bei vielen überwachten Lernverfahren üblich ist, in dem der eigentliche Entscheidungsprozess als Blackbox anzusehen ist. Durch Algorithmen wie LIME (local interpretable model-agnostic explanations) [RSG16] kann Erklärbarkeit auch bei künstlichen neuronalen Netzen gewährleistet werden. LIME erstellt ein zusätzliches interpretierbares Model, welches inhaltlich gleich zum eigentlichen Model ist. Interpretierbare Verfahren, wie Entscheidungsbäume, bieten die Möglichkeit durch nicht abstrakte Attribute eine Visualisierung anzubieten, die es dem Betrachter ermöglicht die Entscheidungsgrundlage nachzuvollziehen.

Ein anderer Ansatz für die Zusicherung von Erklärbarkeit von Reinforcement-Learning-Verfahren wird in [SG19] beschrieben. Ziel des Verfahren ist die Iden-

tifikation von „interestingness elements“ [SG19, S. 2], also von Elementen, die die Entscheidung durch einen hohen Informationsgehalt beeinflussen. Um das zu erreichen werden zusätzlich diverse Informationen gesammelt bzw. generiert. Die Informationen bestehen aus:

- Häufigkeit einzelner Beobachtungen z und wie oft daraufhin eine Aktion a ausgeführt wurde.
- Häufigkeit und Wahrscheinlichkeit, wie oft eine Beobachtung z' als Folge eines vorhergegangen Entscheidungstupels (z, a) erfolgt ist.
- Geschätzter Nutzen des Entscheidungstupels (z, a).
- Geschätzter Nutzen z zu beobachten.

Die Informationen werden analysiert, um Informationen über relevante Elemente zu erhalten und zu bewerten inwiefern diese zur Entscheidungsfindung beitragen.

Safe Exploration

Safe Exploration beschreibt eine Sammlung von Maßnahmen, um einen Agenten auf sichere Art und Weise eine Umgebung erkunden zu lassen. Eine sicherer Erkundungsprozess ist dann wichtig, wenn der Agent in unbekannten Umgebungen agiert. Auch dann sollte der Lernprozess keine Menschen in der Umgebung gefährden. In [Amo+16, S. 14 ff.] werden diverse Möglichkeiten vorgeschlagen, um Safe Exploration umzusetzen.

Veränderte Optimierung der Verhaltensstrategie

Die Verhaltensstrategie soll nicht anhand der Maximierung der Gesamtbelohnung optimiert wird, sondern auch anhand des Verhaltens in selten eintretenden Fällen.

Simulation der Umwelt

Projekte, wie OpenAI Gym [Bro+16] bieten eine standardisierte Schnittstelle zur Erstellung simulierter Umgebungen, wobei viele bekannte Szenarien, wie die Bewegung gewisser Roboter bereits implementiert sind. Viele Reinforcement-Learning-Anwendungen folgen einem iterativen Prozess des „[...] continual cycle of learning and deployment [...]"[Amo+16, S. 15]. Neben der Optimierung des Verfahrens kann es sinnvoll sein auch die Umgebung nach jedem dieser Schritte abhängig von den Resultaten und den daraus resultierenden Probleme zu optimieren. Dadurch können Fehler, wie eine zu starke Anpassung oder Reward Hacking, also das Erlernen nicht generalisierbaren Verhaltens für ein spezielles Problem verhindert werden.

Begrenzung der Umgebung

Findet das System zwangszweise Anwendung in einer realen Umwelt, kann es sinnvoll sein diese zu begrenzen. Dabei kann die Definition von sicheren Zuständen, insbesondere eines sicheren Startzustandes helfen, Risiken zu minimieren. Zusätzlich kann für jede Aktion geprüft werden, ob das Ausführen zur Folge hätte, dass der sichere Zustand verlassen wird.

Überwachung durch Menschen

Die Überwachung des Agenten kann auf unterschiedlichen Arten erfolgen. So kann beispielsweise gefordert werden, dass jede Aktion durch einen Menschen freigegeben werden muss, was allerdings bei Echtzeitanwendungen einen starken Einfluss auf die Performanz hat. Eine abgeschwächte Variante dieses Verfahrens ist die Begrenzung des autonomen Handelns auf klar definierte sichere Zustände, wobei beim Verlassen dieser Zustände eine Zusicherung eines Menschen erfordert wird. Beide Aktionen sind ggfs. sehr langsam, da je nach Modellierung Aktionen sehr schnell erfolgen können. Trotzdem sichert die menschliche Überwachung die Forderung von [Hel96] zu, dass Entscheidungen

über die Ziele der Agenten beim Anwender belassen sein sollten. Zusätzlich bietet der menschliche Einfluss ein erhöhtes Maß an Sicherheit gegenüber unerwünschter Handlungen, sowie eine klarere Definition der Verantwortung.

Sonstige Maßnahmen

Folgende Maßnahmen haben eine zu hohe inhaltliche Distanz zu den anderen Maßnahmen oder einen zu geringen Einfluss auf die Umsetzung der ethischen Werte und sollen deshalb der Vollständigkeit halber lediglich aufgezählt werden.

- Klare Kennzeichnung der Absicht des Agenten durch Statusanzeigen o.Ä.
- Apprenticeship Learning [AN04] als Möglichkeit unter menschlicher Überwachung zu lernen.
- Reward Shaping bzw. Reward Engineering [Kar+19] kann helfen das System zu optimieren. Überlicherweise wird als Optimierungsmetrik der Verhaltensstrategie ein direktes Feedback der Umgebung bzw. des Nutzers verwendet. Die Identifikation von indirektem Feedback sollte analysiert werden und möglicherweise zusätzlich zum direkten Feedback genutzt werden.
- Verzerrungen durch Daten und Entwickler beachten [TS01]. Die Daten sind stark von der Modellierung der Umwelt und der Einschätzung und Wissensstand der Entwickler abhängig und damit auch anfällig für die Verzerrung des Verhaltens durch die persönlichen Werte der Entwickler.

6.2.3 Testphase

Ziel der Testphase ist die Überprüfung der Korrektheit der vorhergangenen Implementierung gemäß Funktionalität, Robustheit und der Erfüllung der Anforderungen.

Um tatsächlich Vertrauen in die Korrektheit des Systems zu gewinnen müssen Evaluierungsprozess und Dokumentation transparent und nachvollziehbar sein. Dafür werden im Folgenden zum einen die Prozesse der Verifikation, sowie der Evaluation von Reinforcement-Learning-Systemen betrachtet.

Verifikation der Korrektheit

Gängige Verifikationsverfahren für Softwaresysteme, wie Modelprüfung (engl. model-checking) sind beim Reinforcement Learning nicht ohne Anpassung nutzbar [van17, S. 12 ff.], da das System durch den stetigen Lernprozess verändert wird. Model-Checking kann z.B. dann benutzt werden, wenn nur ein einzelner Zeitpunkt des Systems betrachtet wird. Grundlage der Verifikation ist die klare Erhebung der Anforderungen und Limitierungen in Unterabschnitt 6.2.1, welche als Spezifikation für die Verifikation benutzt werden. Zu unterscheiden ist zwischen Offline-Verifikation, welches Verifikationsverfahren beschreibt, die zu festen Zeitpunkten genutzt werden können und Online-Verifikation, welches Verifikationsverfahren beschreibt, die in dynamischen Systemen benutzt werden können.

Obwohl Reinforcement Learning ein Online-Lernverfahren beschreibt, können einige Spezifikation auch offline verifiziert werden. So können bei model-basierten Verfahren Markov-Entscheidungsprozesse über Werkzeuge, wie PRISM (Probalistic Symbolic Model Checker) [PRI] automatisch verifiziert werden. Bei model-freien Verfahren muss das Model erst gelernt werden und kann dann, analog zu den model-basierten Verfahren getestet werden. Neben des Markov-Entscheidungsprozess können bei Reinforcement-Learning-Anwendungen die eingesetzten Grundlagenalgorithmen mit regulären Offline-Verfahren getestet werden. Zusätzlich lassen sich Aussagekraft und Güte der Trainingsdaten in Form von Eigenschaften, wie der Verteilung oder Unabhängigkeit durch statistische Verfahren testen.

Durch die sich ständig verändernde Natur eines Online-Lernverfahrens ist Online-Verifizierung bzw. „runtime verification“ [van17, S. 16] rechen- und speicherintensiv. Das grundlegende Vorgehen bei Online Verification basiert auf der Annahme, dass, wenn die Spezifikation vor Beginn des Lernens und bei jeder Änderung erfüllt ist, die Spezifikation als Ganze erfüllt ist. Geprüft werden dann die Zustands-Aktionspaare hinsichtlich ihrer Übereinstimmung mit der Spezifikation. In der Spezifikation müssen neben den Anforderungen auch die Limitierungen abgebildet sein und durch das Verifikationsverfahren geprüft und zugesichert werden.

Auch wenn das Gesamtsystem als solches nicht mit nur einem Verfahren verifiziert werden kann, so kann durch die Kombination mehrerer Verfahren und die Verifikation einzelner Teilkomponenten eine Verbesserung von Korrektheit und Robustheit und damit der Zuverlässigkeit erzielt werden. Die Definition der Spezifikation muss allgemein genug definiert sein, sodass möglichst viele Sachverhalte abgedeckt sind, aber auch so speziell, dass die Eigenheiten des Systems und des Anwendungskontexts abgebildet werden.

Evaluierung der Ergebnisse

Ziel der Evaluierung ist die Bewertung der Güte des Modells anhand nachvollziehbarer Metriken. Die Metriken erbringen den Nachweis über die Zuverlässigkeit und ermöglichen einen sachlichen Vergleich mit anderen Systemen. Aussagekräftige Metriken erlauben zudem die Schaffung realistischer Erwartungen [Got+18, S. 13]. Eine Einordnung der Güte eines Systems sollte in Relation zu anderen Systemen und insbesondere zur menschlichen Leistung im Anwendungskontext gesetzt werden [dBE19]. Um die Nachvollziehbarkeit der Ergebnisse und des Evaluierungsprozesses zu gewährleisten sollte eine kontrollierte Umgebung genutzt werden. Der Aufbau

einer simulierten oder realen Umgebung sollte durch Seeds zur Zufallszahlengenerierung und andere Parameter nachvollzogen werden können. Ebenso sollten weitere Evaluationskriterien, wie die Wahl der Hyperparameter, sowie Implementationen der Algorithmen und je nach Verfahren die Netzwerkarchitektur dokumentiert werden.

Zusätzlich zur tatsächlichen Messverfahren zur Evaluation des Systems sollte die Verhaltensstrategie des Agenten [Got+18, S. 13 f] analysiert werden. Insbesondere durch Hinzunahme von Domänenexperten können Abweichungen des gewünschten Verhaltens identifiziert werden. Die technische Umsetzung der Nachvollziehbarkeit hilft dem Domänenexperten einen tieferen Einblick in das Verhalten des Agenten zu erlangen und so möglicherweise Randeffekte zu bewerten und Ursachen dafür zu identifizieren. Eine solche Analyse ist nicht nur bei der endgültigen Verhaltensstrategie, sondern auch nach jedem Iterationsschritt der Entwicklung sinnvoll, um frühzeitig Optimierungspotenzial oder Fehlverhalten zu identifizieren.

6.2.4 Wartungsphase

Grundlage dieser Phase ist das fertige System, welches im realen Anwendungsfall eingesetzt wird. Zur Wartung im Speziellen von Reinforcement-Learning-Systemen sind in Folge der intensiven Literaturrecherche im Rahmen dieser Arbeit nahezu keine Informationen entsprungen. Deshalb werden im Folgenden allgemeine Maßnahmen zur Wartung von Software-Systemen beschrieben und auf Grundlage meiner persönlichen Meinung auf die Thematik des maschinellen Lernens bzw. des Reinforcement Learning bezogen. Die Wartungsphase bietet das Potenzial, langfristig Vertrauen aufzubauen und nachhaltig zu stärken. In dieser Phase behalten insbesondere die Umsetzungsphase in Unterabschnitt 6.2.2 und die Testphase in Unterabschnitt 6.2.3 ihre Relevanz. Insbesondere die Testphase sollte bei der Umsetzung von Fehlerverbesserungen beachtet werden, um neue Fehler zu vermeiden. Die hier beschriebenen

Maßnahmen sollten stets mit dem Ziel ausgeführt werden, dass in den Phasen bis zum Endprodukt aufgebaute Vertrauen zu stärken und im Fall der Weiterentwicklung und Einführung neuer Funktionen das Vertrauen in die Neuerungen herzustellen. Als Dokumentation ist in der Wartungsphase insbesondere ein Änderungsprotokoll (engl. changelog) sinnvoll, um dem Anwender nachvollziehbar Nachweise über Änderungen und deren Inhalt zu geben. Ziel ist die Wahrung bzw. die Zusicherung eines möglichst nachhaltig zuverlässigen Zustand des Systems.

Fehler können auch durch ausgiebiges Testen nicht gänzlich ausgeschlossen werden [BL11, S. 533]. Insbesondere bei Systemen mit hoher Kritikalitätsbewertung ist eine Wartung im Sinne der Fehlerausbesserung bis zur endgültigen Außerbetriebnahme sinnvoll. Ist dies nicht der Fall, so sollte der Anwender klar über die Dauer des Wartungszeitraumes und über ein Ablaufen dieses Zeitraumes frühzeitig informiert werden. Die Wartung des Systems sollte dessen Funktionalität nicht verändern, sondern Zuverlässigkeit und Korrektheit zusichern. Grundlage der Fehlerausbesserung ist die Fehleridentifikation, welche durch die Logdatenanalyse, eine aktive Fehlersuche durch Entwickler aus Softwaresicht, von Domänenexperten durch Verhaltensanalyse oder durch Rückmeldung der Anwender erfolgen kann. Die Kommunikation mit den Anwendern ist insbesondere deshalb sinnvoll, weil die Vielzahl der Anwendungskontexte und Aufgabengebiete Situationen hervorruft, die in der Entwicklung- und Testphase nicht abgedeckt werden können. Neben den typischen Softwarefehlern bietet Reinforcement Learning Potenzial zusätzlicher Fehlerquellen, die in [Amo+16, S. 3] beschrieben werden. Dazu gehört das unerwünschte Verhalten, bei dem sich durch die stetige Veränderung des System das Verhalten so ändert, dass es nicht mehr den vorher definierten Anforderungen entspricht. Eine weitere Fehlerform ist die unnatürliche Belohnungsmaximierung (engl. reward hacking). Reinforcement-Learning-Agenten können unnatürliches Verhalten entwickeln, welches zwar Aktionen wählt, die die Belohnung maximieren,

jedoch nicht gängigen Methoden entspricht und damit potenziell gefährlich sein kann.

Werden Fehler identifiziert gilt es, diese gemäß der Kritikalität einzuordnen und dementsprechend zu handeln. Insbesondere bei Fehlern in Anwendungen mit hoher Kritikalität sollten Anwender und Betroffene informiert und das System ggfs. bis zur Nachbesserung stillgelegt werden.

7 Evaluation

Im Folgenden erfolgt eine Evaluation der Ergebnisse dieser Arbeit hinsichtlich der in Kapitel 1 gestellten Fragestellungen und der in Abschnitt 5.3 definierten Herausforderungen für die Konzeption. Zur Gewährleistung der Übersichtlichkeit werden die jeweiligen Bewertungsgrundlagen angegeben, inhaltlich passend zusammengefasst und anschließend diskutiert, inwiefern diese erfüllt werden.

7.1 Formalisierung ethischer Werte

Die Formulierung ethischer Werte ist die Grundlage des moralischen Handelns von Reinforcement-Learning-Agenten und der eigentlichen Konzeption organisatorischer und technischer Maßnahmen zur Zusicherung eben dieser. Die Forderung nach der Formalisierung ethischer Werte besteht dabei aus zwei Problemen. So sind ethische Werte abhängig vom sozialen Umfeld und persönlichen Erfahrungen und sind damit subjektiv. Zum anderen werden moralische Entscheidungen in der Regel intuitiv bzw. im Fall moralischer Dilemma hoch komplex und damit schwer formalisierbar getroffen.

Um diese Probleme zu lösen wurde deshalb zur Schaffung des inhaltlichen Bezuges zunächst Grundlagen der angewandten Ethik und Moral vermittelt und die Frage danach, ob moralisches Handeln von Agenten überhaupt möglich ist betrachtet. Zusätzlich wurden Vorschläge und Begründungen bezüglich der Zusicherung

ethischer Werte in KI-Anwendungen unterschiedlicher politischer Institutionen und Regionen vorgestellt. Auf Grundlage des Verständnisses der angewandten Ethik, sowie der Betrachtung bereits bestehender Werte wurden die Eigenschaften Nachvollziehbarkeit und Erklärbarkeit, Vertrauen durch funktionale Korrektheit und Robustheit, sowie Verantwortung und Schuld definiert und jeweils einzeln begründet und inhaltlich eingeordnet. Diese ethischen Werte können natürlich nicht die Gesamtheit aller ethischen Probleme abdecken. Ziel ist viel mehr die Schaffung von Transparenz, der Zusicherung von Vertrauen und die Sicherung und Wahrung der Menschenwürde als höchstes Gut. Ebenso decken die Werte nicht nur das Handeln des Agenten während der Nutzung ab, sondern bieten eine Grundlage für die vorhergehende Zusicherung von Adaption und für das Vorgehen im Fehlerfall, insbesondere im Bezug auf mögliche Folgen. So sollen die Werte möglichst unabhängig vom Nutzungskontext der Anwendung und den Werten einzelner Beteiligter während des Softwareentwicklungsablaufs sein.

7.2 Maßnahmen zur Umsetzung der ethischen Werte von Reinforcement-Learning-Agenten

Grundlage der Erstellung von Maßnahmen zur Umsetzung der ethischen Werte im Sinne eines moralisch handelnden Agenten wurde zunächst auf Basis der Ausarbeitung von [Ben19] begründet, dass Reinforcement-Learning-Agenten potenziell die Fähigkeit des moralischen Handelns besitzen können. In Abschnitt 6.2 wurde dann ein strukturierter Vorgehensplan entwickelt, in dem technische und organisatorische Maßnahmen entlang eines allgemein üblichen Softwareentwicklungsablaufs nach [BK13] vorgestellt wurden. Daraus haben sich die vier Phasen Vorbereitung, technische Konzeption und Umsetzung, Testen und Wartung ergeben. Unabhängig von den jeweiligen Phasen wurden allgemeingültige Maßnahmen, wie eine nachvollzieh-

bare Dokumentation und die Relevanz der Miteinbeziehung von Domänenexperten herausgestellt. Die Relevanz der Maßnahmen wird an jeweiliger Stelle explizit auf Reinforcement-Learning-Anwendungen, sowie die in Abschnitt 6.1 definierten ethischen Werte bezogen. In der Vorbereitungsphase werden so Möglichkeiten zur Sensibilisierung der Beteiligten, einer klaren Definition des Einsatzzweckes, der Absichten und Anforderungen, wie Einsatz und Relevanz einer anwendungsabhängigen Kritikalitätsbewertung eingeführt. Durch eine präzise Dokumentation und Definition des Nutzungs- und Entwicklungskontextes können so Verantwortlichkeiten identifziert werden. Anschließend wurden in der technischen Konzeptions- und Umsetzungsphase insbesondere technische Maßnahmen betrachtet. Dazu gehört neben der Wahl der geeigneten Technologie hinsichtlich ethischer Anforderungen, eine nachvollziehbare Datenverarbeitung, explizite Maßnahmen zur Zusicherung von Nachvollziehbarkeit, sowie Möglichkeiten zur sicherenden Erkundung von Agenten in fremden Umgebungen. In der Testphase wurden dann Maßnahmen aufgezeigt, mit denen die Korrektheit des Agenten nachvollziehbar geprüft werden kann. Ebenso wurde der Evaluationsprozess betrachtet, um aussagekräftige und nachvollziehbare Metriken zu produzieren, die den sachlichen Vergleich mit anderen Systemen erlauben. Abschließend wurden in der Wartungsphase Maßnahmen betrachtet, mit denen das System nachhaltig gepflegt werden kann, um Vertrauen langfristig zuzusichern, Fehler zu identifizieren und transparent zu dokumentieren, sowie Verantwortung durch Kommunikation innerhalb des Wartungszeitraumes abzugrenzen.

Die Maßnahmen und das Vorgehen sind dabei explizit so gewählt, dass sie möglichst kompatibel mit den jeweiligen Technologien und Verfahrensmodellen sind. In Folge dessen ist der Erfolg der Umsetzung der ethischen Werte abhängig von der Strenge der Durchführung. So müssen im Projektablauf die Maßnahmen aktiv integriert und die Umsetzung regelmäßig kontrolliert werden. Dadurch ist jeder Beteiligte gefragt das persönliche und kollektiv Handeln bezüglich der Anforderungen zu hinterfra-

gen. Der daraus resultierende Agent handelt, insbesondere wegen der Problematik der Formalisierbarkeit moralischen Handelns und daraus resultierender technischer Limitierungen der Abbildung eben dieser, nicht menschenähnlich, sondern gemäß eines allgemeinen ethischen Rahmens, mit dem Ziel der Wahrung von Unversehrtheit von Würde und Wohl des Menschen und dem Aufbau von Vertrauen durch Nachvollziehbarkeit, Korrektheit und der Identifikation von Verantwortung.

7.3 Adaption und Akzeptanz

Damit Reinforcement Learning als Technologie und die daraus resultierenden Anwendungen Adaption innerhalb der breiten Masse finden, müssen gewissen Hürden überwunden werden. Im Rahmen der Arbeit wurde deshalb zunächst in Abschnitt 5.3 der Prozess der Adaption auf persönlicher Ebene und die daraus entstehenden Hürden identifiziert. In der Definition der ethischen Werte in Abschnitt 6.1 wird die Adaption implizit beachtet. So dient die Zusicherung der ethischen Werte dazu, die Technologien transparent zu gestalten und so Vertrauen aufzubauen. Ebenso hilft die klare Definition innerhalb des Vorgehensplan in Abschnitt 6.2 dabei, den Nutzungskontext klar abzugrenzen und realistische Erwartungen des Anwenders herzustellen. Nicht betrachtet werden im Rahmen dieser Arbeit Möglichkeiten zur Reduzierung der Komplexität der eigentlichen Anwendungen oder der Etablierung eines Mindestkenntnisstandes, um Folgen der digitalen Kluft entgegenzuwirken. Die Etablierung der Technologie im Sinne der Adaption durch Akzeptanz erfolgt eher, indem Innovatoren [Kar13, S. 519] und frühe Übernehmer Vertrauen gewinnen und dadurch die frühe und späte Mehrheit beeinflusst wird.

7.4 Schaffung von Vertrauen trotz mangelnder Zertifizierung

Im Rahmen der Arbeit ist die Problematik der mangelnden Zertifizierung deutlich geworden. Auch wenn es die Möglichkeit zur Zertifizierung von Softwareentwicklungsprozessen gibt, so gibt es zum Zeitpunkt der Recherche keine Möglichkeit die Korrektheit von Reinforcement-Learning-Anwendungen im Speziellen von vertrauenswürdigen Institutionen nachzuweisen. Dadurch, dass Vertrauen eine der geforderten ethischen Werte im Zusammenhang mit Korrektheit ist, werden im Vorgehensplan explizit Maßnahmen zur Schaffung dessen betrachtet. Um trotz mangelnder Zertifizierungen Vertrauen zu gewinnen wird deshalb eine offene Dokumentation, insbesondere im Wartungszeitraum, gefordert. Neben einer transparenten Prozess- und Anwendungsdokumentation sind die Kernmaßnahmen in Unterabschnitt 6.2.3 in Form der Testphase definiert. So kann durch Verifikation die Korrektheit des Agenten gemäß der Anforderungen nachgewiesen werden und mit Hilfe nachvollziehbarer Metriken die Güte der Anwendung nachgewiesen werden. Ebenso resultiert die Miteinbeziehung von Domänenexperten in den gesamten Entwicklungszeitraum in einem Entgegenwirken gegen eine mögliche Wissenslücke zwischen Entwicklern und der Domäne.

8 Zusammenfassung

Im Rahmen dieser Arbeit wurden technische und organisatorische Maßnahmen zur Zusicherung ethischer Werte in Reinforcement-Learning-Anwendungen aufgezeigt. Zur Umsetzung dieser Maßnahmen wurde dann ein Vorgehensplan entwickelt, der mit üblichen Softwareentwicklungsabläufen kompatibel ist. Als Grundlage für die Definition der ethischen Werte und der Konzeption des Vorgehensplans wurden zunächst Grundlagen der angewandten Ethik und Moral vorgestellt. Dies beeinhaltet eine Abgrenzung der Begriffe Ethik und Moral, die Entstehung von Normen und Werten, sowie die Einordnung der Begriffe in einen praktischen Kontext. Innerhalb der Ethik lassen sich Situationen und Anwendungen je nach Kontext in Bereichsethiken einteilen. So wurden Grundlagen der Maschinenethik als Bereichsethik des Reinforcement Learning eingeführt und diskutiert, inwiefern Software-Agenten die Möglichkeiten besitzen, moralisch zu handeln. Neben den ethischen Grundlagen wurden technische Grundlagen des Reinforcement Learning betrachtet. Hierbei wurden explizit Markov-Entscheidungsprozess als Grundlage der Abbildung von Reinforcement-Learning-Umgebungen eingeführt, sowie allgemeine Eigenschaften und Begriffe der Technologie aufgezeigt. Als Basis für die Definition der ethischen Werte, die ein Reinforcement-Learning-Agent erfüllen soll, um moralisch handeln zu können, wurden anschließend die Maßnahmen und Forderungen von Europa, USA und Asien, sowie verschiedenen Institutionen, wie der Hochrangingen Expertengruppe für künstliche Intelligenz in Europa, IEEE in der USA und der ISO auf internationaler Ebene analysiert. Anschließend wurde der Kontext, sowie Technolo-

gien und Vorgehen innerhalb der Arbeit abgegrenzt und Anwendungsgebiete des Reinforcement Learning aufgezeigt, sowie inhaltlich verwandte Arbeiten verglichen. Als spätere Bewertungsgrundlagen wurden dann Herausforderungen aufgeführt, die bei der Definition der ethischen Werte, sowie der Konzeption der Maßnahmen zur Zusicherung eben dieser relevant sind. In der eigentlichen Konzeption wurden dann die ethischen Werte Nachvollziehbarkeit und Erklärbarkeit, Vertrauen durch Kalkulierbarkeit und Zuverlässigkeit, sowie Verantwortung und Schuld im einzelnen definiert und die Relevanz im Kontext der Arbeit begründet. Zur Zusicherung der ethischen Werte wurde anschließend ein Vorgehensplan konzipiert, welcher möglichst verfahrensunabhängig entlang eines generischen Softwareentwicklungsablaufs technische organisatorische Maßnahmen zu den verschiedenen Entwicklungsphasen Vorbereitung, Konzeption und Umsetzung, Test und Wartung aufzeigt. Die Maßnahmen umfassen Möglichkeiten zur Selektion geeigneter Verfahren, Sensibilisierung der Beteiligten, Nachvollziehbarkeit von Datenverarbeitung, Prozessen und der Entscheidungen des Agenten, sowie der sicheren Erkundung in fremden Umgebungen. Ebenso werden Maßnahmen zur Zusicherungen der Korrektheit und der nachvollziehbaren Evaluierung von Agenten, sowie zur transparenten Wartung der nachhaltigen Zusicherung von Vertrauen aufgezeigt. Abschließend wurde der im Rahmen der Arbeit erarbeitete Vorgehensplan anhand der anfänglichen Fragestellungen und Herausforderungen bewertet und Möglichkeiten zur Weiterentwicklungen betrachtet.

9 Ausblick

Insbesondere im Hinblick auf die Adaption von Reinforcement-Learning-Anwendungen der breiten Masse bietet diese Arbeit Potenzial für zukünftige Weiterentwicklungen. Dazu gehören Maßnahmen zur Schaffung von praktischen Nutzungskontexten und die Kommunikation eben dieser, der Förderung von Bildungs- und Informationsmöglichkeiten und der Reduzierung der Nutzungskomplexität.

Ebenso sollte die praktische Umsetzung der Maßnahmen, insbesondere der Dokumentation stärker vorgegeben werden, um eine nachvollziehbare Erfüllung der Maßnahmen nachweisen zu können. Maßnahmen dazu wären beispielsweise der Nachweis der Erfüllung der Maßnahmen und der Dokumentation in einer standardisierten aber erweiterbaren Form mit der Anwendung auszuliefern. Insbesondere um die Akzeptanz aus technischer Sicht zuzusichern wäre eine exemplarische Beispielanwendung gemäß des vorgestellten Vorgehens sinnvoll.

Auch wenn das Vorgehen möglichst verfahrensunabhängig ist, so ist Reinforcement Learning als Technologie schnelllebig. In Zukunft ist deshalb eine regelmäßige Evaluation der technischen Maßnahmen im Hinblick auf die Aktualität der Maßnahmen und der Nutzung neuer Verfahren, sowie der Erweiterung der organisatorischen Maßnahmen notwendig.

Abbildungsverzeichnis

Literatur

[Aay18] NITI Aayog. *National Strategy for Artificial Intelligence #AIFORALL*. Techn. Ber. 2018.

[AI4] AI4People. *AI4People | Atomium-EISMD*. en-US. In: ().

[Amo+16] Dario Amodei et al. „Concrete Problems in AI Safety“. en. In: *arXiv:1606.06565 [cs]* (Juli 2016). arXiv: 1606.06565 [cs].

[AN04] Pieter Abbeel und Andrew Y. Ng. „Apprenticeship Learning via Inverse Reinforcement Learning“. en. In: *Twenty-First International Conference on Machine Learning - ICML '04*. Banff, Alberta, Canada: ACM Press, 2004, S. 1. DOI: 10.1145/1015330.1015430.

[ANS] ANSI. *ANSI-American National Standards Institute*. In: ().

[AS19] C. S. Arvind und J. Senthilnath. „Autonomous RL: Autonomous Vehicle Obstacle Avoidance in a Dynamic Environment Using MLP-SARSA Reinforcement Learning“. en. In: *2019 IEEE 5th International Conference on Mechatronics System and Robots (ICMSR)*. Singapore: IEEE, Mai 2019, S. 120–124. ISBN: 978-1-72812-223-6. DOI: 10.1109/ICMSR.2019.8835462.

[Avi+04] A. Avizienis et al. „Basic Concepts and Taxonomy of Dependable and Secure Computing“. en. In: *IEEE Transactions on Dependable and Secure Computing* 1.1 (Jan. 2004), S. 11–33. ISSN: 1545-5971. DOI: 10.1109/TDSC.2004.2.

[Bar+17] Nathalie Baracaldo et al. „Mitigating Poisoning Attacks on Machine Learning Models: A Data Provenance Based Approach“. en. In: *Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security - AISec '17*. Dallas, Texas, USA: ACM Press, 2017, S. 103–110. ISBN: 978-1-4503-5202-4. DOI: 10.1145/3128572.3140450.

[Bau17] Seth D. Baum. „On the Promotion of Safe and Socially Beneficial Artificial Intelligence“. en. In: *AI & SOCIETY* 32.4 (Nov. 2017), S. 543–551. ISSN: 0951-5666, 1435-5655. DOI: 10.1007/s00146-016-0677-0.

[Ben19] Oliver Bendel, Hrsg. *Handbuch Maschinenethik*. de. Wiesbaden: Springer Fachmedien Wiesbaden, 2019. ISBN: 978-3-658-17482-8 978-3-658-17483-5. DOI: 10.1007/978-3-658-17483-5.

[bit] bitkom. *Bitkom e.V.* de. In: ().

[BK13] Manfred Broy und Marco Kuhrmann. „Projekt- und Produktlebenszyklus von Software“. de. In: *Projektorganisation und Management im Software Engineering*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, S. 61–83. ISBN: 978-3-642-29289-7 978-3-642-29290-3. DOI: 10.1007/978-3-642-29290-3_3.

[BL11] Helmut Balzert und Peter Liggesmeyer. *Lehrbuch der Softwaretechnik Entwurf, Implementierung, Installation und Betrieb*. German. OCLC: 1078359165. Heidelberg: Spektrum Akademischer Verl, 2011. ISBN: 978-3-8274-2246-0.

[Bro+16] Greg Brockman et al. „Openai Gym“. In: *arXiv preprint arXiv:1606.01540* (2016).

[bus19] businesswire. *Global Autonomous/Driverless Car Market Forecasts to 2024: Semi-Autonomous Vehicles Dominating the Market - ResearchAndMarkets.Com*. en. In: (Mai 2019).

[BVE10] Haitham Bou-Ammar, Holger Voos und Wolfgang Ertel. „Controller Design for Quadrotor UAVs Using Reinforcement Learning“. en. In: *2010 IEEE International Conference on Control Applications*. Yokohama, Japan: IEEE, Sep. 2010, S. 2130–2135. ISBN: 978-1-4244-5362-7. DOI: 10.1109/CCA.2010.5611206.

[Cas] Arnaldo Pérez Castaño. *Practical Artificial Intelligence*. en.

[CBB] Nicolas Cointe, Grégory Bonnet und Olivier Boissier. „Ethical Judgment of Agents' Behaviors in Multi-Agent Systems“. en. In: (), S. 9.

[CH19] Raja Chatila und John C. Havens. „The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems“. en. In: *Robotics and Well-Being*. Hrsg. von Maria Isabel Aldinhas Ferreira et al. Bd. 95. Cham: Springer International Publishing, 2019, S. 11–16. ISBN: 978-3-030-12523-3 978-3-030-12524-0. DOI: 10.1007/978-3-030-12524-0_2.

[CLA] CLAIRE. *CLAIRE*. en-US. In: *CLAIRE* ().

[Cre+] Dr Armin B Cremers et al. „HANDLUNGSFELDER AUS PHILOSOPHISCHER, ETHISCHER, RECHTLICHER UND TECHNOLOGISCHER SICHT ALS GRUNDLAGE FÜR EINE ZERTIFIZIERUNG VON KÜNSTLICHER INTELLIGENZ“. de. In: (), S. 21.

[dBE19] Thibault de Swarte, Omar Boufous und Paul Escalle. „Artificial Intelligence, Ethics and Human Values: The Cases of Military Drones and Companion Robots“. en. In: *Artificial Life and Robotics* 24.3 (Sep. 2019), S. 291–296. ISSN: 1433-5298, 1614-7456. DOI: 10.1007/s10015-019-00525-1.

[Dha+17] Prafulla Dhariwal et al. *Openai Baselines*. 2017.

[DIK09] Gunnar Duttge, Institut für Kriminalwissenschaften und Kriminalwissenschaftliches Kolloquium, Hrsg. *Das Ich und sein Gehirn: die Herausforderung der neurobiologischen Forschung für das (Straf-)Recht*. de. Göttinger Studien zu den Kriminalwissenschaften 7. OCLC: 436303860. Göttingen: Univ.-Verl. Göttingen, 2009. ISBN: 978-3-941875-01-2.

[DIN] DIN. *DIN - Deutsches Institut Für Normung*. In: ().

[Din] Jeffrey Ding. „Deciphering China's AI Dream“. en. In: (), S. 44.

[EME] IEEE EMELC-WG. *P7000 - Model Process for Addressing Ethical Concerns During System Design*. In: ().

[Gil] Dirk Ulrich Gilbert. „Vertrauen als Gegenstand der ökonomischen Theorie“. de. In: (), S. 48.

[Got+18] Omer Gottesman et al. „Evaluating Reinforcement Learning Algorithms in Observational Health Settings“. en. In: *arXiv:1805.12298 [cs, stat]* (Mai 2018). arXiv: 1805.12298 [cs, stat].

[Hal07] John Storrs Hall. „Self-Improving AI: An Analysis“. en. In: *Minds and Machines* 17.3 (Okt. 2007), S. 249–259. ISSN: 0924-6495, 1572-8641. DOI: 10.1007/s11023-007-9065-3.

[Haw] Andrew J. Hawkins. „Serious Safety Lapses Led to Uber's Fatal Self-Driving Crash, New Documents Suggest“. en. In: ().

[Hel96] Günter Hellbardt. „Die Ethik von Agenten“. de. In: *Informatik-Spektrum* 19.2 (Apr. 1996), S. 87–90. ISSN: 0170-6012, 1432-122X. DOI: 10.1007/s002870050021.

[Her] Michael Herrmann. „RL 5: On-Policy and off-Policy Algorithms“. en. In: (), S. 24.

[Hof19] Ulrich Hoffrage. „Persönliche Verantwortung und Verantwortungsübernahme in Systemen“. de. In: *Pädiatrie & Pädologie* 54.S1 (Okt. 2019), S. 4–9. ISSN: 0030-9338, 1613-7558. DOI: 10.1007/s00608-019-0679-5.

[HW98] Carey Heckman und Jacob O. Wobbrock. „Liability for Autonomous Agent Design“. en. In: *Proceedings of the Second International Conference on Autonomous Agents - AGENTS '98*. Minneapolis, Minnesota, United States: ACM Press, 1998, S. 392–399. ISBN: 978-0-89791-983-8. DOI: 10.1145/280765.280869.

[IEE] IEEE. *IEEE - The World's Largest Technical Professional Organization Dedicated to Advancing Technology for the Benefit of Humanity*. In: ().

[ISOa] ISO 22989. *ISO/IEC CD 22989*. en. In: *ISO* ().

[ISOb] ISO_SC42. *ISO - ISO/IEC JTC 1/SC 42 - Artificial Intelligence*. In: ().

[isoa] iso 23053. *ISO/IEC CD 23053*. en. In: *ISO* ().

[isob] iso 23894. *ISO/IEC AWI 23894*. en. In: *ISO* ().

[isoc] iso 24368. *ISO/IEC AWI TR 24368*. en. In: *ISO* ().

[isod] iso about. *ISO - About Us*. en. In: *ISO* ().

[isoe] iso_benefits. *ISO - Benefits of Standards*. en. In: *ISO* ().

[Kar+19] Nikos Karampatziakis et al. „Lessons from Contextual Bandit Learning in a Customer Support Bot". en. In: *arXiv:1905.02219 [cs, stat]* (Juni 2019). arXiv: 1905.02219 [cs, stat].

[Kar13] Veronika Karnowski. „Diffusionstheorie". de. In: *Handbuch Medienwirkungsforschung*. Hrsg. von Wolfgang Schweiger und Andreas Fahr. Wiesbaden: Springer Fachmedien Wiesbaden, 2013, S. 513–528. ISBN: 978-3-531-18158-5 978-3-531-18967-3. DOI: 10.1007/978-3-531-18967-3_27.

[Koc+19] William Koch et al. „Reinforcement Learning for UAV Attitude Control". en. In: *ACM Transactions on Cyber-Physical Systems* 3.2 (März 2019), S. 1–21. ISSN: 2378-962X, 2378-9638. DOI: 10.1145/3301273.

[Kra09] Oliver Kramer. *Computational intelligence: eine einführung*. de. Informatik im fokus. Dordrecht ; New York: Springer, 2009. ISBN: 978-3-540-79738-8 978-3-540-79739-5.

[Lef86] Guy R. Lefrancois. *Psychologie des Lernens*. de. Hrsg. von Peter K. Leppmann, Wilhelm F. Angermeier und Thomas J. Thiekötter. Springer-Lehrbuch. Berlin, Heidelberg: Springer Berlin Heidelberg, 1986. ISBN: 978-3-540-16192-9 978-3-662-09577-5. DOI: 10.1007/978-3-662-09577-5.

[Li18] Yuxi Li. „Deep Reinforcement Learning: An Overview". en. In: *arXiv:1701.07274 [cs]* (Nov. 2018). arXiv: 1701.07274 [cs].

[Liu+17] Ying Liu et al. „Deep Reinforcement Learning for Dynamic Treatment Regimes on Medical Registry Data". en. In: *2017 IEEE International Conference on Healthcare Informatics (ICHI)*. Park City, UT, USA: IEEE, Aug. 2017, S. 380–385. ISBN: 978-1-5090-4881-6. DOI: 10.1109/ICHI.2017.45.

[McN88] David McNaughton. *Moral Vision: An Introduction to Ethics*. en. Oxford, UK ; New York, NY: B. Blackwell, 1988. ISBN: 978-0-631-15408-2 978-0-631-15945-2.

[MdE03] Jorge Moll, Ricardo de Oliveira-Souza und Paul J. Eslinger. „Morals and the Human Brain: A Working Model:“ en. In: *NeuroReport* (März 2003), S. 299–305. ISSN: 0959-4965. DOI: 10.1097/00001756-200303030-00001.

[MIT19] Technology Review Insights MIT. *Asia's AI Agenda - The Ethics of AI*. Techn. Ber. MIT, 2019.

[MM18] Jean M. Mulcahy Levy und Martin McMahon. „Linking Brain Tumors and Epileptic Seizures“. en. In: *Nature Medicine* 24.11 (Nov. 2018), S. 1638–1639. ISSN: 1078-8956, 1546-170X. DOI: 10.1038/s41591-018-0249-6.

[Moz18] Paul Mozur. *Inside China's Dystopian Dreams: A.I., Shame and Lots of Cameras*. en. In: *NY Times* (2018).

[Nat48] Vereinte Nationen. *Resolution der Generalversammlung*. de. In: (Dez. 1948).

[NL09] Lianqiang Niu und Ling Li. „Application of Reinforcement Learning in Autonomous Navigation for Virtual Vehicles“. en. In: *2009 Ninth International Conference on Hybrid Intelligent Systems*. Shenyang, China: IEEE, 2009, S. 30–32. ISBN: 978-0-7695-3745-0. DOI: 10.1109/HIS.2009.118.

[Noo+] Ritesh Noothigattu et al. „Teaching AI Agents Ethical Values Using Reinforcement Learning and Policy Orchestration“. en. In: (), S. 5.

[Olu+17] Habeeb Olufowobi et al. „Data Provenance Model for Internet of Things (IoT) Systems“. en. In: *Service-Oriented Computing – ICSOC 2016 Workshops*. Hrsg. von Khalil Drira et al. Bd. 10380. Cham: Springer International Publishing, 2017, S. 85–91. ISBN: 978-3-319-68135-1 978-3-319-68136-8. DOI: 10.1007/978-3-319-68136-8_8.

[Ope] OpenAI. *OpenAI*. In: ().

[Pan+17] Xinlei Pan et al. „Virtual to Real Reinforcement Learning for Autonomous Driving“. en. In: *arXiv:1704.03952 [cs]* (Sep. 2017). arXiv: 1704.03952 [cs].

[plsa] pls. *AG 3 - PLS*. In: ().

[plsb] pls. *Plattform Lernende Systeme - PLS*. In: ().

[PRI] PRISM. *PRISM - Probabilistic Symbolic Model Checker*. In: ().

[RKK19] Matthias Rath, Friedrich Krotz und Matthias Karmasin, Hrsg. *Maschinenethik: Normative Grenzen autonomer Systeme*. de. Ethik in mediatisierten Welten. Wiesbaden: Springer Fachmedien Wiesbaden, 2019. ISBN: 978-3-658-21082-3 978-3-658-21083-0. DOI: 10.1007/978-3-658-21083-0.

[Rog16] Sylvia E. Rogers. „Bridging the 21st Century Digital Divide“. en. In: *TechTrends* 60.3 (Mai 2016), S. 197–199. ISSN: 8756-3894, 1559-7075. DOI: 10.1007/s11528-016-0057-0.

[ROY19] Florian Richter, Ryan K. Orosco und Michael C. Yip. „Open-Sourced Reinforcement Learning Environments for Surgical Robotics“. en. In: *arXiv:1903.02090 [cs]* (März 2019). arXiv: 1903.02090 [cs].

[RSG16] Marco Tulio Ribeiro, Sameer Singh und Carlos Guestrin. „"Why Should I Trust You?": Explaining the Predictions of Any Classifier“. en. In: *arXiv:1602.04938 [cs, stat]* (Aug. 2016). arXiv: 1602.04938 [cs, stat].

[SAC] SAC. . In: ().

[Sal+17] AhmadEL Sallab et al. „Deep Reinforcement Learning Framework for Autonomous Driving“. en. In: *Electronic Imaging* 2017.19 (Jan. 2017), S. 70–76. ISSN: 2470-1173. DOI: 10.2352/ISSN.2470-1173.2017.19.AVM-023.

[SB10] Christian Schicha und Carsten Brosda, Hrsg. *Handbuch Medienethik*. 1. Aufl. OCLC: ocn606927422. Wiesbaden: VS Verlag für Sozialwissenschaften, 2010. ISBN: 978-3-531-15822-8.

[SB18] Richard S. Sutton und Andrew G. Barto. *Reinforcement Learning: An Introduction*. en. Second edition. Adaptive Computation and Machine Learning Series. Cambridge, Massachusetts: The MIT Press, 2018. ISBN: 978-0-262-03924-6.

[Sen+18] Eishvak Sengupta et al. „Techniques to Elimenate Human Bias in Machine Learning“. en. In: *2018 International Conference on System Modeling & Advancement in Research Trends (SMART)*. Moradabad, India: IEEE, Nov. 2018, S. 226–230. ISBN: 978-1-5386-6369-1. DOI: 10.1109/SYSMART.2018.8746946.

[SG19] Pedro Sequeira und Melinda Gervasio. „Interestingness Elements for Explainable Reinforcement Learning: Understanding Agents' Capabilities and Limitations“. en. In: *arXiv:1912.09007 [cs, stat]* (Dez. 2019). arXiv: 1912.09007 [cs, stat].

[Smu] Nathalie Smuha. *Ethik-Leitlinien für eine vertrauenswürdige KI*. de. Techn. Ber. Unabhängige Hochrangige Expertengruppe für Künstliche Intelligenz, S. 51.

[Tän08] Torbjörn Tännsjö. *Understanding Ethics: An Introduction to Moral Theory*. en. 2nd ed. OCLC: 254748111. Edinburgh: Edinburgh Univ. Press, 2008. ISBN: 978-0-7486-3690-7 978-0-7486-3689-1.

[TS01] Alan J Thomson und Daniel L Schmoldt. „Ethics in Computer Software Design and Development". en. In: *Computers and Electronics in Agriculture* 30.1-3 (Feb. 2001), S. 85–102. ISSN: 01681699. DOI: 10.1016/S0168-1699(00)00158-7.

[Uni] *China AI Development Report 2018*. Techn. Ber. China Institute for Science and Technology Policy at Tsinghua University.

[van17] Perry van Wesel. „Challenges in the Verification of Reinforcement Learning Algorithms". en. In: (2017), S. 30.

[WE17] Ines-Jacqueline Werkner und Klaus Ebeling, Hrsg. *Handbuch Friedensethik*. de. Wiesbaden: Springer Fachmedien Wiesbaden, 2017. ISBN: 978-3-658-14685-6 978-3-658-14686-3. DOI: 10.1007/978-3-658-14686-3.

[Woo+] Bryce Woodworth et al. „Preference Learning in Assistive Robotics: Observational Repeated Inverse Reinforcement Learning". en. In: *Preference Learning* (), S. 19.

[WS19] Volker Wittpahl und Springer-Verlag GmbH. *Künstliche Intelligenz: Technologie, Anwendung, Gesellschaft*. de. OCLC: 1079417702. 2019. ISBN: 978-3-662-58041-7.

[Xia+18] Liang Xiao et al. „UAV Relay in VANETs Against Smart Jamming With Reinforcement Learning". en. In: *IEEE Transactions on Vehicular Technology* 67.5 (Mai 2018), S. 4087–4097. ISSN: 0018-9545, 1939-9359. DOI: 10.1109/TVT.2018.2789466.

[You+19] Changxi You et al. „Advanced Planning for Autonomous Vehicles Using Reinforcement Learning and Deep Inverse Reinforcement Learning". en. In: *Robotics and Autonomous Systems* 114 (Apr. 2019), S. 1–18. ISSN: 09218890. DOI: 10.1016/j.robot.2019.01.003.

[YPB] April Yu, Raphael Palefsky-Smith und Rishi Bedi. „Deep Reinforcement Learning for Simulated Autonomous Vehicle Control". en. In: (), S. 7.

[Zha+15] Baochang Zhang et al. „Geometric Reinforcement Learning for Path Planning of UAVs". en. In: *Journal of Intelligent & Robotic Systems* 77.2 (Feb. 2015), S. 391–409. ISSN: 0921-0296, 1573-0409. DOI: 10.1007/s10846-013-9901-z.

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Schriften entnommen wurden, sind als solche kenntlich gemacht. Die Arbeit ist in gleicher oder ähnlicher Form oder auszugsweise im Rahmen einer anderen Prüfung noch nicht vorgelegt worden.

Ort, Datum	Unterschrift