

## Prozessorientiertes Reinforcement Learning: Grafische Modellierung zur Unterstützung der Erklärbarkeit

Clemens Schreiber<sup>1</sup>, Gunther Schiefer<sup>1</sup>, Sascha Alpers<sup>2</sup>, Marius Take<sup>2</sup> und  
Andreas Oberweis<sup>1,2</sup>

**Zusammenfassung.** Das Verstärkende Lernen (Reinforcement Learning) stellt einen wichtigen Ansatz für Systeme der Künstlichen Intelligenz (KI-Systeme) dar. Dabei steigt der Anspruch an die Erklärbarkeit der KI-Systeme mit zunehmender Risikobehaftung der zu lösenden Problemstellungen. Um den Lernprozess beim Verstärkenden Lernen nachvollziehbar zu machen, verfolgen wir einen prozessorientierten Lernansatz. Zunächst soll der Lernprozess mit Hilfe eines grafischen Prozessmodells abgebildet werden, um eine Visualisierung der einzelnen Lernschritte zu ermöglichen. Diese Prozessmodellierung soll durch die Verwendung von Process Mining Methoden erfolgen. In einem weiteren Schritt soll den Anwendern die Möglichkeit gegeben werden, anhand der Prozessmodelle die Entscheidungsfindung der Algorithmen zu beeinflussen. Eine mögliche Art der Einflussnahme ist zum Beispiel die Beschränkung des Hypothesenraumes, der mit Hilfe des Verstärkenden Lernens erkundet werden soll. Auf diese Weise agiert das Prozessmodell als grafische Schnittstelle zwischen maschinellem Lernprozess und Anwender. Das wesentliche Ziel dieses neuen Ansatzes ist es, die Erklärbarkeit von KI-Systemen und die Kooperationsfähigkeit zwischen Anwendern und KI-Systemen zu verbessern. Dieser Artikel beschreibt die Grundlagen, um dieses Ziel mit Hilfe von prozessorientiertem Reinforcement Learning zu erreichen.

**Keywords:** Reinforcement Learning, Prozessmodellierung, Interaktives Maschinelles Lernen, Erklärbare Künstliche Intelligenz.

### 1 Einleitung

Systeme der Künstlichen Intelligenz (KI-Systeme) werden bereits in verschiedensten Bereichen zur Entscheidungsunterstützung genutzt und bieten das Potential, auch in kritischeren Fällen (d.h. Fällen, bei denen durch eine Fehlentscheidung ein bedeutender Schaden entsteht) Anwendung zu finden. Solche Fälle finden sich beispielsweise in der Medizin oder in der Rechtsfindung, in denen ein KI-System Entscheidungshinweise zur Patientenbehandlung oder zu Gesetzesgrundlagen geben kann. Damit die Akzeptanz für diese Entscheidungsempfehlungen, gerade in kritischen Anwendungsdomänen, bei den Anwendern und Betroffenen der KI-Systeme möglichst groß und ein verantwortungsbewusster Einsatz möglich ist, sollten die Empfehlungen sowohl nachvollziehbar als auch kontrollierbar sein. Im Forschungsbereich Explainable Artificial Intelligence

---

<sup>1</sup> Karlsruher Institut für Technologie, AIFB, Kaiserstraße 12, 76131 Karlsruhe, Germany, {andreas.oberweis, gunther.schiefer, clemens.schreiber}@kit.edu

<sup>2</sup> FZI Forschungszentrum Informatik, Haid- und Neu-Straße 10-14, 76131 Karlsruhe, Germany, {alpers, oberweis, take}@fzi.de

(XAI) wurde in den letzten Jahren eine Vielzahl an Ansätzen entwickelt, um subsymbolische Systeme nachvollziehbar zu machen [1]. Dabei geht es nicht nur um das Ergebnis selbst, sondern auch um den Lösungs- und den Lernweg [2]. Im Fokus von XAI steht daher sowohl die Transparenz, als auch die Erklärbarkeit von künstlichen Neuronalen Netzen (kNN) [3]. Bei der Transparenz geht es darum, den Lernprozess, der die Gewichtung der Verknüpfungen zwischen den einzelnen Neuronen ermittelt, vollständig nachvollziehbar zu machen. Die Erklärbarkeit verfolgt hingegen das Ziel, entscheidende Einflussfaktoren für die Entstehung der Gewichtungen zwischen den Neuronen aufzuzeigen. Der grafischen Nachvollziehbarkeit, als weitere Möglichkeit für eine bessere Erklärbarkeit von maschinellen Lernprozessen, wurde in der Forschung bisher wenig Aufmerksamkeit geschenkt. Die Modellierung des Lernvorgangs bietet potentiell die Möglichkeit, die Entwicklung des Lernvorgangs zu analysieren und Einflussfaktoren auf die Entscheidungsfindung zu identifizieren. Zusätzlich zur Analysefähigkeit können Prozessmodelle verwendet werden, um das Lernverhalten der KI-Systeme zu beeinflussen. Im Folgenden werden die Machbarkeit und die Potentiale der grafischen Prozessmodellierung für das Verstärkende Lernen (Reinforcement Learning; RL) diskutiert.

### 1.1 Erklärbarkeit von Reinforcement Learning

Neben dem Unüberwachten Lernen und dem Überwachten Lernen hat sich im Bereich des Maschinellen Lernens in den letzten Jahren vor allem das RL als erfolgreicher Ansatz zur Wissensgenerierung hervorgetan [4]. Beim RL agiert ein lernender, autonomer Agent mit seiner Umwelt und erhält dabei Rückmeldungen bezüglich der Auswirkungen seiner ausgeführten Aktionen. Durch die Maximierung der akkumulierten Rückmeldungen unterschiedlicher erprobter Aktionsfolgen nähert sich der Agent einem gestellten Ziel. RL ist zum Beispiel die Grundlage für AlphaGo Zero, einem KI-System, welches selbstständig komplexe und erfolgreiche Strategien für verschiedene Brettspiele entwickeln kann [4]. Zum Erlernen einer Erfolgsstrategie verwenden RL-Algorithmen häufig kNN (dies wird als Deep Reinforcement Learning bezeichnet; DRL). Mit Hilfe von kNN ist es möglich, eine Approximation für die beste Erfolgsstrategie eines Agenten zu errechnen. Bei dieser Methode wird nicht nur jede einzelne Aktion unabhängig von den anderen Aktionen betrachtet, sondern die gesamte Aktionsabfolge verbessert. Dabei kann die Komplexität des Entscheidungsraums die Nachvollziehbarkeit der Ergebnisse wesentlich erschweren.

Ein existierender Ansatz, um die Transparenz von DRL zu erhöhen, besteht darin, die Bedeutung der Umgebungsfaktoren für die Entscheidungsfindung genauer zu analysieren [5]. Dadurch werden Objekte in der Umgebung des Agenten veranschaulicht, die zu einer bestimmten Entscheidung des Agenten führen. Mit diesem Ansatz des Explainable RL wird zwar das Ergebnis des Lernprozesses erklärbar, der Lernvorgang allerdings nicht. Das Ziel des Prozessorientierten RL ist es hingegen, den zusammenhängenden Lernprozess des Agenten transparenter zu machen.

## 1.2 Modellierung des Reinforcement Learning Prozesses

Der Ansatz, Lernprozesse von RL-Algorithmen mit Hilfe einer grafischen Prozessmodellierungssprache zu visualisieren, ermöglicht es, Entscheidungspunkte innerhalb des Lernvorgangs zu visualisieren. Beim RL können auf diese Weise auch Zwischenergebnisse, die sich durch das Ausprobieren von unterschiedlichen Aktionsfolgen ergeben, analysiert werden. Die Zwischenergebnisse zeigen auf, welche möglichen Lösungen durch den Algorithmus bereits getestet wurden und welche Entscheidungen zu welchen Resultaten geführt haben.

Die Modellierung des Lernvorgangs kann mit Hilfe von verschiedenen Prozessmodellierungssprachen wie Ereignisgesteuerte Prozessketten (EPK), (höhere) Petri-Netze, UML-Aktivitätsdiagramme oder der Business Process Modeling Notation (BPMN) erfolgen. Als ein Auswahlkriterium für die Prozessmodellierungssprache kann, neben allgemeinen Kriterien wie Formalisierungsgrad und Ausdrucksmächtigkeit, die Kompatibilität mit Markov Decision Processes (MDP) herangezogen werden. MDP bilden in der Regel die Grundlage für die Beschreibung der Agenten-Umgebung beim RL. Sie bilden den Zustandsraum, die Menge der möglichen Aktionen, die Wahrscheinlichkeiten für die Zustandsveränderungen und das Feedback für die Aktionsausführung ab [6]. Die Kompatibilität zu MDP ist insbesondere dann hilfreich, wenn das Prozessmodell nicht nur zur Erklärbarkeit des Lernvorgangs, sondern auch zur Einflussnahme auf die Ergebnisfindung verwendet werden soll. Hierfür eignen sich bspw. Petri Netze [7].

## 1.3 Interaktives Reinforcement Learning

Weil die RL-Algorithmen auf die Veränderung der Umgebung der Agenten reagieren können, ist eine Einflussnahme auf den Lernprozess durch den Anwender des KI-Systems möglich. Die Einflussnahme geschieht hierbei indirekt durch die Veränderung der Umgebungsparameter und nicht durch die direkte Anpassung des RL-Algorithmus. Generell wird der Ansatz, bei dem der Anwender aktiv Einfluss auf das Lernverhalten der KI nehmen kann, als Interaktives Lernen bezeichnet [8]. Bezogen auf das RL, kann dieser Ansatz genutzt werden, um den Lernvorgang interaktiv zu möglichen Lösungen zu führen [9, 10]. Der Anwender übernimmt beim Interaktiven RL die Rolle des Lehrers und signalisiert dem lernenden Agenten welche Aktionsfolgen positiv bzw. negativ sind. Diese Form der Kooperation wird auch als hybride Intelligenz bezeichnet [11]. Dabei kann der Mensch, der die Rolle des Lehrers übernimmt, an diskreten Entscheidungspunkten den Lernprozess des RL-Algorithmus steuern. Ein wesentliches Ziel des Prozessorientierte RL ist es, den Lernprozess auch strukturell, mit Hilfe von Prozessmodellen, beeinflussbar zu machen.

## 2 Grundlagen des Prozessorientierten Reinforcement Learning

Beim Prozessorientierten RL wird ein Prozessmodell verwendet, um den Lernprozess bei der Ausführung eines RL-Algorithmus grafisch abzubilden. Dafür werden zunächst die Entscheidungspfade des autonom agierenden Agenten beim RL in Form eines Ent-

scheidungsbaums festgehalten (siehe Abb. 1). In einem nächsten Schritt wird der Entscheidungsbaum verwendet, um ein Prozessmodell zu erzeugen und um die Zusammenhänge der Entscheidungen zu analysieren. Das Prozessmodell dient sowohl der Nachvollziehbarkeit einer KI-Entscheidung als auch zur Einflussnahme auf das Lernverhalten eines KI-Systems durch den Anwender. Der Anwender benötigt dafür keine Programmierkenntnisse, er verfügt aber über Fachkenntnisse in der Domäne, in dem das KI-System eine Entscheidungsunterstützung liefern soll.

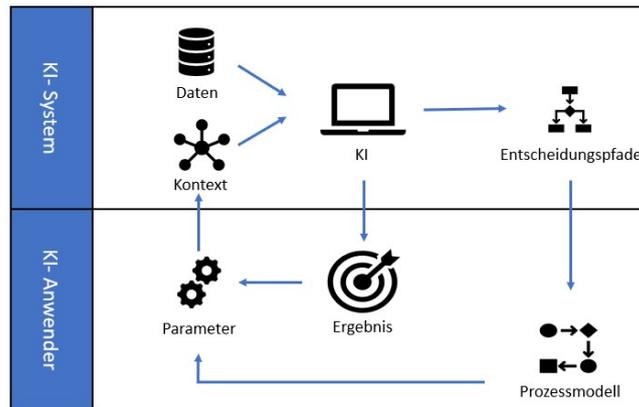


Abb. 1. Ablauf der Entscheidungsfindung durch Prozessorientiertes RL

## 2.1 Vom Lernprozess zum Prozessmodell

Durch die Modellierung des Lernprozesses wird ersichtlich, welche Faktoren zu bestimmten Entscheidungen geführt haben und welche Abhängigkeiten zwischen den Entscheidungen innerhalb des Lernverlaufes existieren. Mit Hilfe von datenbasierter Prozessanalyse (Process Mining) und dem Entscheidungsbaum kann ein Prozessmodell des Lernvorgangs erstellt werden. Als Grundlage für das Process Mining dienen die dokumentierten Aktivitäten, die der autonome Agent während des Lernprozesses sequentiell ausgeführt hat. Für dieses Vorgehen stehen eine Vielzahl an Algorithmen aus dem Bereich des Process Discovery zur Verfügung [12].

Neben der grafischen Visualisierung des maschinellen Lernprozesses bietet die Modellierung auch die Möglichkeit, Zusammenhänge zwischen den inkrementellen Entscheidungen, innerhalb des Lernprozesses, zu analysieren. Dazu wird zunächst ein Modell des Lernprozesses in Form eines Entscheidungsbaums generiert. Ein möglicher Ansatz dafür wird von Blake und Ntoutsis vorgestellt [13]. Dadurch könnten signifikante Einflussfaktoren auf die Entscheidungsfindung erkannt werden, die später durch den Anwender reguliert werden können. Für die Analyse werden Methoden aus dem Bereich des Decision Minings verwendet [14]. Mit Hilfe des Decision Minings ist es möglich, die Abhängigkeiten der Entscheidungen innerhalb eines Prozesses zu bewerten. Beim RL können auf diese Weise Ursachen für die Entscheidungspfade ermittelt werden.

## 2.2 Verwendung des Prozessmodells zur Steuerung

Auf Basis des Prozessmodells ist eine Einflussnahme auf den Lernprozess möglich. Durch eine Begrenzung des Hypothesenraums, bzw. durch die Festlegung von Nebenbedingungen, kann die Entscheidungsfindung beeinflusst werden. Nebenbedingungen können bspw. in Form eines Kontextmodells formal definiert werden [15]. Ein Beispiel für eine solche Regulierung wäre, dass bestimmte Aktionen der Agenten nur in einer bestimmten Reihenfolge ausgeführt werden dürfen.

Das Decision Mining unterstützt die Festlegung von Prozesskriterien für den Lernvorgang. Wenn bspw. die Analyse des Lernvorgangs ergibt, dass bestimmte Faktoren einen besonders großen Einfluss auf das Endergebnis haben, können die Anwender entscheiden, ob diese Faktoren aus ihrer fachlichen Sicht wirklich relevant sind. Wenn der Einfluss eines Faktors verändert werden soll, kann dies als Nebenbedingung für den RL-Algorithmus definiert werden und der Lernprozess wird dementsprechend angepasst. Auf diesem Weg kann das gelernte Modell iterativ an die Präferenzen und an das vorhandene Wissen der Anwender angepasst werden.

Die Modellierung von Lern- und Entscheidungsprozessen kann darüber hinaus die Untersuchung von ethischen, rechtlichen und sozialen Aspekten unterstützen. Hierzu gehört auch die allgemeine Technikfolgenabschätzung als Grundlage für die spätere Technikfolgenbewertung.

## 3 Ausblick

Der Fokus des Prozessorientierten RL liegt auf der Kooperationsfähigkeit zwischen KI-System und Anwender. Die Grundlage für die Kooperation ist zunächst das Verständnis des Anwenders für den KI-Algorithmus. In einem weiteren Schritt kann daraus ein Kooperationsverhältnis entstehen, durch das der Anwender gemeinsam mit dem KI-Algorithmus eine gestellte Aufgabe lösen kann. Damit eine Kooperation basierend auf Prozessorientiertem RL entstehen kann muss ermittelt werden, welche Anforderungen erfüllt sein müssen, damit der KI-Anwender die benötigten und gewünschten Informationen erhält, um aktiv in den Lernprozess eingreifen zu können. Dazu muss zunächst eine geeignete Modellierungssprache ermittelt und die entsprechende Granularität der Modellierung angepasst werden.

## Referenzen

- [1] L. H. Gilpin, D. Bau, B. Z. Yuan, A. Bajwa, M. Specter und L. Kagal, „Explaining explanations: An overview of interpretability of machine learning,“ *IEEE 5th International Conference on Data Science and Advanced Analytics*, 2018.
- [2] D. Doran, S. Schulz und T. Besold, „What Does Explainable AI Really Mean? A New Conceptualization of Perspectives,“ *in Proceedings of the First International Workshop*

- on Comprehensibility and Explanation in AI and ML 2017. Co-located with 16th International Conference of the Italian Association for Artificial Intelligence, 2017.*
- [3] I. Döbel, M. Leis, M. M. Vogelsang, D. Neustroev, H. Petzka, S. Rüping, A. Voss, M. Wegele und J. Welz, „Maschinelles Lernen: Kompetenzen, Anwendungen und Forschungsbedarf,“ Frauenhofer-Gesellschaft, 2018.
  - [4] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan und D. Hassabis, „A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play,“ *Science*, 362(6419), 1140-1144, 2018.
  - [5] R. Iyer, Y. Li, H. Li, M. Lewis, R. Sundar und K. Sycara, „Transparency and explanation in deep reinforcement learning neural networks,“ *In Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society (pp. 144-150)*, 2018.
  - [6] Q. Zhang, G. Sun und Y. Xu, „Parallel Algorithms for Solving Markov Decision Process,“ *In International Conference on Algorithms and Architectures for Parallel Processing (pp. 466-477)*. Springer, Berlin, Heidelberg, 2009.
  - [7] M. Beccuti, G. Franceschinis und S. Haddad, „Markov decision Petri net and Markov decision well-formed net formalisms,“ *In International Conference on Application and Theory of Petri Nets (pp. 43-62)*. Springer, Berlin, Heidelberg, 2007.
  - [8] J. A. Fails und D. R. Olsen Jr, „Interactive machine learning,“ *In Proceedings of the 8th International Conference on Intelligent User Interfaces (pp. 39-45)*. ACM, 2003.
  - [9] H. B. Suay und S. Chernova, „Effect of human guidance and state space size on interactive reinforcement learning,“ *In 2011 Ro-Man (pp. 1-6)*. IEEE, 2011.
  - [10] A. L. Thomaz und C. Breazeal, „Transparency and socially guided machine learning,“ *In 5th Intl. Conf. on Development and Learning (ICDL)*, 2006.
  - [11] E. Kamar, „Directions in Hybrid Intelligence: Complementing AI Systems with Human Intelligence,“ *In IJCAI (pp. 4070-4073)*, 2016.
  - [12] W. van der Aalst, „Process Discovery: An Introduction,“ in *Process Mining*, Springer, Berlin, Heidelberg, 2016, pp. 163-192.
  - [13] C. Blake und E. Ntoutsis, „Reinforcement Learning Based Decision Tree Induction over Data Streams with Concept Drifts,“ *In 2018 IEEE International Conference on Big Knowledge (ICBK) (pp. 328-335)*. IEEE, 2018.
  - [14] A. Rozinat und W. van der Aalst, „Decision mining in ProM. In International Conference on Business Process Management,“ (pp. 420-425). Springer, Berlin, Heidelberg, 2006.
  - [15] J. M. E. van der Werf, H. M. W. Verbeek und W. van der Aalst, „Context-aware compliance checking,“ *In International Conference on Business Process Management (pp. 98-113)*, Springer, Berlin, Heidelberg, 2012.