

Technische Hochschule Ostwestfalen-Lippe
University of Applied Sciences and Arts

Wintersemester 2022/2023

Vergleich verschiedener Reinforcement Learning Algorithmen

Modularbeit

Vorgelegt im Kontext des Moduls „Anwendung des maschinellen Lernens“

am Fachbereich Technische Informatik und Elektrotechnik
im Studiengang Data Science

Veranstalter:	Prof. Dr. Philipp Bruland
Vorgelegt von:	Bjarne Seen Liebigstraße 130 32657 Lemgo bjarne.seen@stud.th-owl.de
Matr. Nr.:	15467085
Vorgelegt von:	Joshua Henjes Hanseweg 11 32657 Lemgo joshua.henjes@stud.th-owl.de
Matr. Nr.:	15467024
Abgabetermin:	02.03.2023

24. Februar 2023

Inhaltsverzeichnis

Abkürzungen	I
1 Einleitung	1
2 Grundlagen	1
2.1 Definitionen	1
2.2 Algorithmen	2
2.2.1 Q-Learning	2
2.2.2 SARSA	2
3 Implementierung	2
3.1 Probleme	2
3.2 Vergleichen von Algorithmen und Paramtern	2
4 Ergebnisse	2
4.1 Taxi	2
4.2 Cliff	2
4.3 Frozen Lake	2
5 Auswertung und Diskussion	2
5.1 Taxi	2
5.2 Cliff	2
5.3 Frozen Lake	2
6 Zusammenfassung und Ausblick	2

Abkürzungen

API Application Programming Interface

1 Einleitung

Reinforcement Learning ist neben Supervised und Unsupervised Learning eins der elementaren Felder des maschinellen Lernens. Im Gegensatz zu den anderen Feldern benötigt Reinforcement Learning keine Trainingsdaten, denn der Algorithmus lernt durch wiederholtes Interagieren mit einer dynamischen Umgebung eine Strategie, um eine Belohnungsmetrik zu maximieren. Es wird daher auch als bestärktes lernen oder verstärktes lernen bezeichnet.

Bekannt wurde das Reinforcement Learning vor allem durch das Meistern von bekannten Brett- und Computerspielen, so ist Googles „AlphaGo“ in der Lage, die besten Go Spieler der Welt zu schlagen. Trotz dieser beeindruckenden Erfolge findet RL in der Industrie bisher nur geringe Anwendung.

Immer kürzer werdende Produktzyklen und steigende Produktvielfalt stellen für die heutigen Produktionsprozesse eine große Herausforderung dar. Zukünftige Produktionen müssen immer anpassungsfähiger werden. Zeitgleich soll der Personalaufwand aufgrund des anhaltenden Fachkräftemangels möglichst gering ausfallen. Maschinelles Lernen, insbesondere das Reinforcement Learning, kann bei der Bewältigung dieser Herausforderungen eine relevante Rolle übernehmen.

Auch bei der Bekämpfung des Klimawandels kann Reinforcement Learning unterstützen. Um unsere Klimaziele zu erreichen, ohne unseren Lebensstandard signifikant zu senken, ist eine Optimierung des Ressourcenbedarfs nötig. Mit ausreichender Trainingszeit sind RL-Algorithmen sehr gut in der Optimierung von Prozessen und somit auch in dessen Ressourcenverbrauches. Google, als einer der Vorreiter im Gebiet des maschinellen Lernens, konnte durch ML-Algorithmen den Energieverbrauch der Kühlung ihrer Rechenzentren um bis zu 40 Prozent reduzieren.

Mittlerweile existiert eine Vielzahl an unterschiedlichen Reinforcement Learning Algorithmen. Während die mathematischen und strukturellen Unterschiede meist gut dokumentiert und einsehbar sind, ist ein direkter Vergleich der Leistungsfähigkeit der Algorithmen in verschiedenen Umgebungen nur schwer zu finden. Aus diesem Grund beschäftigt sich diese Ausarbeitung mit dem Vergleich von beliebten RL-Algorithmen anhand von Umgebungen mit geringer Komplexität.

Während die Zeit, welche ein RL-Algorithmus zum Lernen benötigt, bei der Anwendung auf Brett- und Computerspielen eher eine untergeordnete Rolle spielt, ist sie in der Anwendung in industriellen Umgebungen deutlich relevanter. Zum einen verlangsamen hohe Trainingszeiten den Entwicklungsprozess deutlich, was wiederum zu höheren Lohn- und Entwicklungskosten führt. Zum anderen ist es in vielen Anwendungsfällen nötig, dass die Umgebung während des Trainingsprozesses dem Algorithmus zur Verfügung steht. Im Fall von Produktionsanlagen ist Trainingszeit somit sehr kostspielig. Aus diesem Grund wird neben der Leistungsfähigkeit auch die Lerngeschwindigkeit der Algorithmen im Folgenden untersucht.

2 Grundlagen

2.1 Definitionen

1. Environment
2. Action
3. State
4. Reward
5. Bellmann Equation?

- 6. Markov Decision Process
- 7. Exploration vs Exploitation
- 8. Temporal difference vs Monte Carlo
- 9. [Source](#)

2.2 Algorithmen

2.2.1 Q-Learning

2.2.2 SARSA

3 Implementierung

3.1 Probleme

- 1. Taxi
- 2. Cliff
- 3. Frozen Lake

3.2 Vergleichen von Algorithmen und Paramtern

4 Ergebnisse

4.1 Taxi

4.2 Cliff

4.3 Frozen Lake

5 Auswertung und Diskussion

5.1 Taxi

5.2 Cliff

5.3 Frozen Lake

6 Zusammenfassung und Ausblick