

Computer spielen Computerspiele mit Hilfe von reinforcement learning am Beispiel Vier Gewinnt

Pablo Lubitz



BACHELORARBEIT

eingereicht am
Universitäts-Bachelorstudiengang

Informatik

in Bremen

im September 2019

Betreuung:

Thomas Barkowsky, Holger Schultheis

Erklärung

Ich erkläre eidesstattlich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen nicht benutzt und die den benutzten Quellen entnommenen Stellen als solche gekennzeichnet habe. Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt.

Bremen, am 10. September 2019

Pablo Lubitz

Inhaltsverzeichnis

| | |
|---|------------|
| Erklärung | iii |
| Vorwort | v |
| Kurzfassung | vi |
| Abstract | vii |
| 1 Einleitung | 1 |
| 1.1 Motivation | 1 |
| 1.2 Problemstellung | 1 |
| 1.3 Umsetzung | 2 |
| 1.4 Evaluation | 2 |
| 1.5 Related work | 2 |
| 2 Grundlagen | 3 |
| 2.1 Reinforcement Learning | 3 |
| 2.1.1 psychologischer Hintergrund | 3 |
| 2.1.2 Agent | 3 |
| 2.1.3 Environment | 3 |
| 2.1.4 Policy | 3 |
| 2.1.5 Reward | 3 |
| 2.1.6 State | 3 |
| 2.1.7 Actions | 3 |
| 2.1.8 Q-Funktion | 3 |
| 2.2 Vier Gewinnt | 3 |
| 2.3 Python | 3 |
| 2.4 Tensorflow | 3 |
| 2.5 andere Software? | 3 |
| 3 Konzept | 4 |
| 3.1 Plan | 4 |
| 4 Implementierung | 5 |
| 4.1 Agent | 5 |
| 4.2 Environment | 5 |

| | |
|------------------------------|----------|
| Inhaltsverzeichnis | v |
| 4.3 Policy | 5 |
| 4.4 Reward | 5 |
| 4.5 State | 5 |
| 4.6 Actions | 5 |
| 5 Evaluierung | 6 |
| 5.1 Auswertung | 6 |
| 6 Fazit | 7 |
| 6.1 Besonderheiten | 7 |
| Quellenverzeichnis | 8 |

Vorwort

TODO

Kurzfassung

TODO

Abstract

TODO engl.

Kapitel 1

Einleitung

In der Bachelorarbeit soll ein Programm entwickelt werden das mit Hilfe von maschine learning trainiert um das Spiel Vier Gewinnt zu erlernen. Es wird ein maschine learning Ansatz gewählt da hierdurch im Vergleich zu normalen Algorithmen für jeden Spieler individuellere Gegner geschaffen werden können die das Können des Spielers besser widerspiegeln. Dies soll den Spielspaß und den Lerneffekt optimieren.

1.1 Motivation

Warum spielen wir? Wir spielen um Fähigkeiten zu erlernen, ein Kind lernt Objekte nach Formen zu sortieren oder eine Katze lernt spielerisch das Jagen von Beute. Doch auch erwachsene Menschen spielen noch und können hierdurch ihre Fähigkeiten verbessern. Um den bestmöglichen Effekt zu haben muss das Spiel eine gewisse Schwierigkeit haben die aber noch zu bewältigen sein muss. Da aber jeder Mensch, egal wie erfahren, das Spiel spielen soll ist es oft Hilfreich einen Gegner zu haben der ein ähnliches Level an Erfahrung mitbringt. Wenn nun diese Rolle des Gegners nicht von einem Menschen besetzt werden kann ist es naheliegend einen Computergegner zu erschaffen. Die Frage die ich in dieser Bachelorarbeit versuche zu beantworten ist: Ist es möglich einen Computergegner zu erschaffen der für jeden Spieler individuell genau so gut ist dass er nicht frustrierend gut oder zu einfach ist.

1.2 Problemstellung

Vier Gewinnt ist ein Spiel für 2 Spieler in dem abwechselnd Spielsteine in ein vertikales Spielfeld eingeworfen werden. Jeder Spieler kann sich in seinem Zug zwischen einem von sieben Einwurföchern entscheiden. Wird ein Spielstein eingeworfen so fällt dieser auf die niedrigste der sechs Positionen die noch nicht durch andere Spielsteine gefüllt ist. Sind schon sechs Spielsteine in ein bestimmtest Einwurfloch gesteckt wurden so darf hier kein weiterer Spielstein eingeworfen werden. Ein Spieler gewinnt das Spiel wenn er es geschafft hat, dass sich vier seiner Spielsteine in einer Reihe befinden. Eine Reihe kann horizontal, vertikal oder diagonal gebildet werden. Werden alle 42 Positionen mit Spielsteinen befüllt ohne eine Reihe von vier Steinen des selben Spielers zu bilden geht die Partie unentschieden aus. Ein Computergegner der für dieses Spiel erschaffen

wird kann auf verschiedenen Algorithmen basieren er könnte zum Beispiel in jedem Zug einfach zufällig eines der sieben möglichen Einwurflöcher bedienen oder er könnte durch einen Minimax Algorithmus zu jedem Zeitpunkt den bestmöglichen Zug errechnen. Das Problem an solchen Algorithmen ist, dass sie immer gleich agieren egal wie gut oder schlecht der Spieler ist gegen den sie spielen. Es ist natürlich möglich auch einen Computergegner zu entwerfen der je nach Anzahl der gewonnenen und verlorenen Spiele unterschiedliche Algorithmen auswählt und somit ein besseres Spielerlebnis erzeugt. Das Problem hieran ist, dass jeder einzelne Algorithmus programmiert werden muss und somit ein enormer Arbeitsaufwand entstehen kann. Ein Lösungsansatz für diese Problem sind selbstlernende Algorithmen die mit Hilfe von reinforcement learning und neuronalen Netzen selbstständig erlernen, ähnlich wie ein Mensch, wie das Spiel zu gewinnen ist.

1.3 Umsetzung

Es wird eine digitale Version des Spiels Vier Gewinnt erstellt und für eine Schnittstelle erschafft die es ermöglicht auszuwählen ob ein Mensch oder der Computer die Rolle der Spieler übernimmt. Spielt der Computer soll zwischen verschiedenen Algorithmen ausgewählt werden können. Dies ist wichtig da ein selbstlernender Algorithmus viele Spiele durchlaufen muss um einen Lernfortschritt aufzuzeigen. Somit kann in der Simulation dann trainiert werden ohne jedes einzelne Spiel gegen einen Menschen spielen zu müssen. Als Trainingsgegner wird ein Minimax Algorithmus dienen. Nachdem die Simulation dann durchlaufen ist können verschieden gute Varianten des selbstlernenden Algorithmus extrahiert werden indem die aktuelle Version zu verschiedenen Trainingsstadien abgespeichert wird.

1.4 Evaluation

Ich habe noch keine Ahnung wie die Evaluierung aussehen kann.

1.5 Related work

Kapitel 2

Grundlagen

Eine Liste der genutzten Technologien.

2.1 Reinforcement Learning

2.1.1 psychologischer Hintergrund

2.1.2 Agent

2.1.3 Environment

2.1.4 Policy

2.1.5 Reward

2.1.6 State

2.1.7 Actions

2.1.8 Q-Funktion

2.2 Vier Gewinnt

Alles was es über das Spiel zu wissen gibt

2.3 Python

2.4 Tensorflow

2.5 andere Software?

Kapitel 3

Konzept

Wie soll alles gemacht werden

3.1 Plan

Was ist der Plan?

Kapitel 4

Implementierung

Vorgehensweise bei der Implementierung

4.1 Agent

4.2 Environment

4.3 Policy

4.4 Reward

4.5 State

4.6 Actions

Kapitel 5

Evaluierung

Was kann evaluiert werden?

5.1 Auswertung

Kapitel 6

Fazit

Was wurde festgestellt auswertung ausblick

6.1 Besonderheiten

Quellenverzeichnis

Messbox zur Druckkontrolle

— Druckgröße kontrollieren! —



— Diese Seite nach dem Druck entfernen! —