

## Reinforcement Learning for Pathfinding in a partially unknown Environment

Roboter werden im Haushalt, in der Industrie oder auch in der Raumfahrt eingesetzt. Durch die jüngsten Entwicklungen im Bereich der künstlichen Intelligenz agieren sie weitgehend autonom. Ein Problem dabei stellt die Wegfindung in einer sich ständig ändernden Umgebung dar. In herkömmlichen Verfahren versucht der Roboter, sich eine Karte seiner Umgebung basierend auf seinem Vorwissen und seinen Sensordaten zu erstellen und mit Algorithmen den schnellsten Weg zu seinem Ziel zu bestimmen.

In dieser Arbeit wird folgende Leitfrage untersucht: Lässt sich ein Agent durch Reinforcement-Learning darauf trainieren, lediglich anhand der Daten seines Lidar-Sensors und der Richtung zum Ziel einen effizienten Weg um Hindernisse herum zum Ziel zu finden?

Dafür wird die Leitfrage in verschiedene Forschungsfragen aufgeteilt: Zuerst wird die triviale Aufgabe, ein Ziel in einer Umgebung ohne jegliche Hindernisse zu finden, untersucht. Weiter werden ein oder mehrere immobile und bewegliche Hindernisse zwischen Agent und Ziel platziert. Abschliessend wird die Erkennung und Meidung von einfachen Sackgassen untersucht.

Für das Training des Agenten wurde das Double-Deep-Q-Learning-Verfahren verwendet und für jede der Forschungsfragen zehn Agenten mit je einer, drei und fünf versteckten Schichten trainiert. Somit war es möglich zu ermitteln, welches Szenario welche Komplexität des neuronalen Netzwerks erfordert. Die Versuche wurden in einer selbst entwickelten Gridworld-Umgebung für OpenAI durchgeführt, um die Konfiguration des Agenten und des Belohnungssystems zu testen. Anschliessend wurde ein einfaches Szenario in der Roboter-Simulationsumgebung Gazebo trainiert, um die Übertragbarkeit der Konzepte zu testen.

Es wird gezeigt, dass in der Gridworld für jedes Szenario mindestens ein Agent gefunden werden kann, der die geforderten Kriterien erfüllt. Dabei wurde festgestellt, dass nicht immer der Agent mit der grössten Netztiefe der effizienteste ist, sondern in manchen Szenarien eine einfachere Struktur bessere Ergebnisse lieferte. Des Weiteren wurde durch einen kleinen Testlauf gezeigt, dass die Konzepte der Gridworld erfolgreich in die komplexere Roboter-Simulationsumgebung Gazebo übertragen werden können.

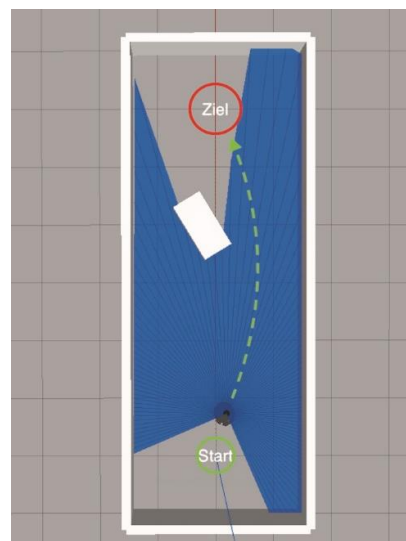


### Diplomierende

Raffael Affolter  
Stefan Morgenthaler  
Jean-Pierre Péclard

### Dozierende

Christian Jaeger  
Rudolf Marcel Fuchsli



Ein Roboter (schwarz), der mittels Reinforcement-Learning in der Gazebo-Simulationsumgebung trainiert, wird ein Hindernis umfahren, um sein Ziel zu erreichen. Dabei nutzt er lediglich Distanzmessungen seines Lidar-Sensors (blau).