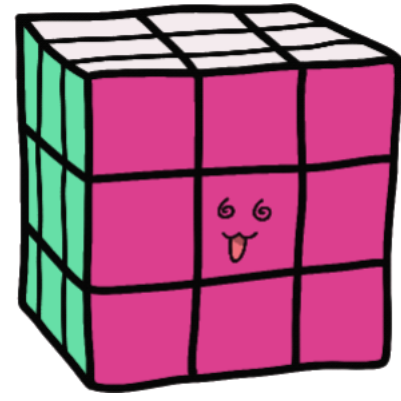


Deep Reinforcement Learning zum maschinellen Erlernen von Strategien zur Lösung von Zauberwürfeln

Masterarbeit Finn Lanz
Prof. Dr.-Ing. Marc Hensel und Prof. Dr. Sönke Appel

Hochschule für Angewandte Wissenschaften Hamburg · Fachhochschule Westküste

Der Rubik's Cube bzw. Zauberwürfel begeistert seit Jahrzehnten Menschen. Dies liegt vor allem an seinem einfachen Aufbau und seiner dennoch enormen Komplexität. Da der Würfel deutlich mehr unterschiedliche Zustände besitzt als es Menschen auf der Erde gibt, könnte es sehr viel Zeit in Anspruch nehmen, bis ein Mensch alleine Algorithmen entwickelt, um den Würfel aus jedem beliebigen Zustand zu lösen. Wie sieht es jedoch aus, wenn eine künstliche Intelligenz versucht, sich das kombinatorische Puzzlespiel selbst beizubringen? Anhand des kleineren Pocket Cubes und aufbauend auf wissenschaftlichen Veröffentlichungen sollen Methoden erarbeitet werden, welche später auf den Rubik's Cube übertragen werden könnten.



Deep Reinforcement Learning

Mit Hilfe von Deep Reinforcement Learning wird ein neuronales Netzwerk mittels Belohnungen (*Rewards*) so trainiert, dass es komplexe Entscheidungsprozesse wie das Lösen eines Zauberwürfels durch eine Ketten von Handlungen (*Actions*) erlernt.

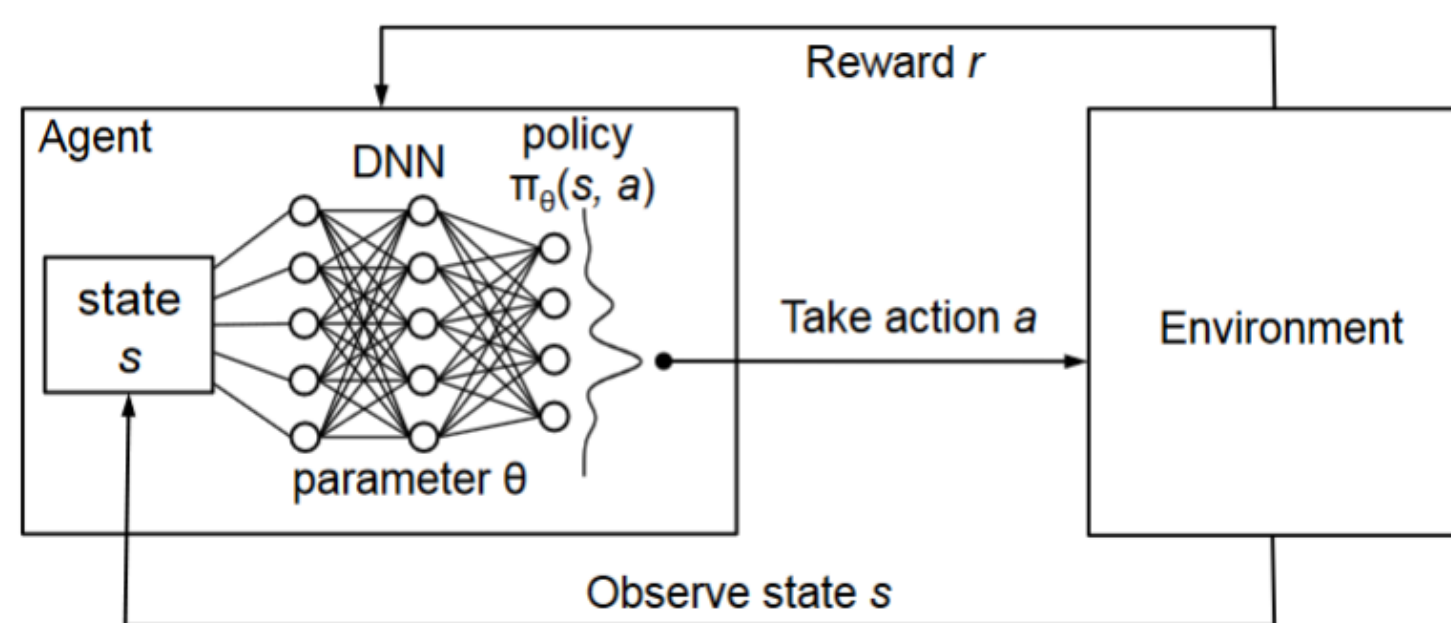


Image taken from: <https://medium.com/@vishnuvijayanpv/deep-reinforcement-learning-artificial-intelligence-machine-learning-and-deep-learning-e52cb5974420> (Access: 02.10.2023)

Autodidaktisches Iterationsverfahren

Grundlage des Trainings bildet das Autodidaktische Iterationsverfahren, bei welchem der Würfel ausgehend vom gelösten Zustand schrittweise mit einer zufälligen Aktion weiter verdreht wird. Anschließend werden die möglichen Folgezustände jedes Zustandes mit einem neuronalen Netzwerk untersucht. Somit soll das Netz den Zusammenhang zwischen aufeinander aufbauenden Zuständen lernen, anstatt zusammenhangslos zerwürfelte Zustände als Eingabe zu erhalten.

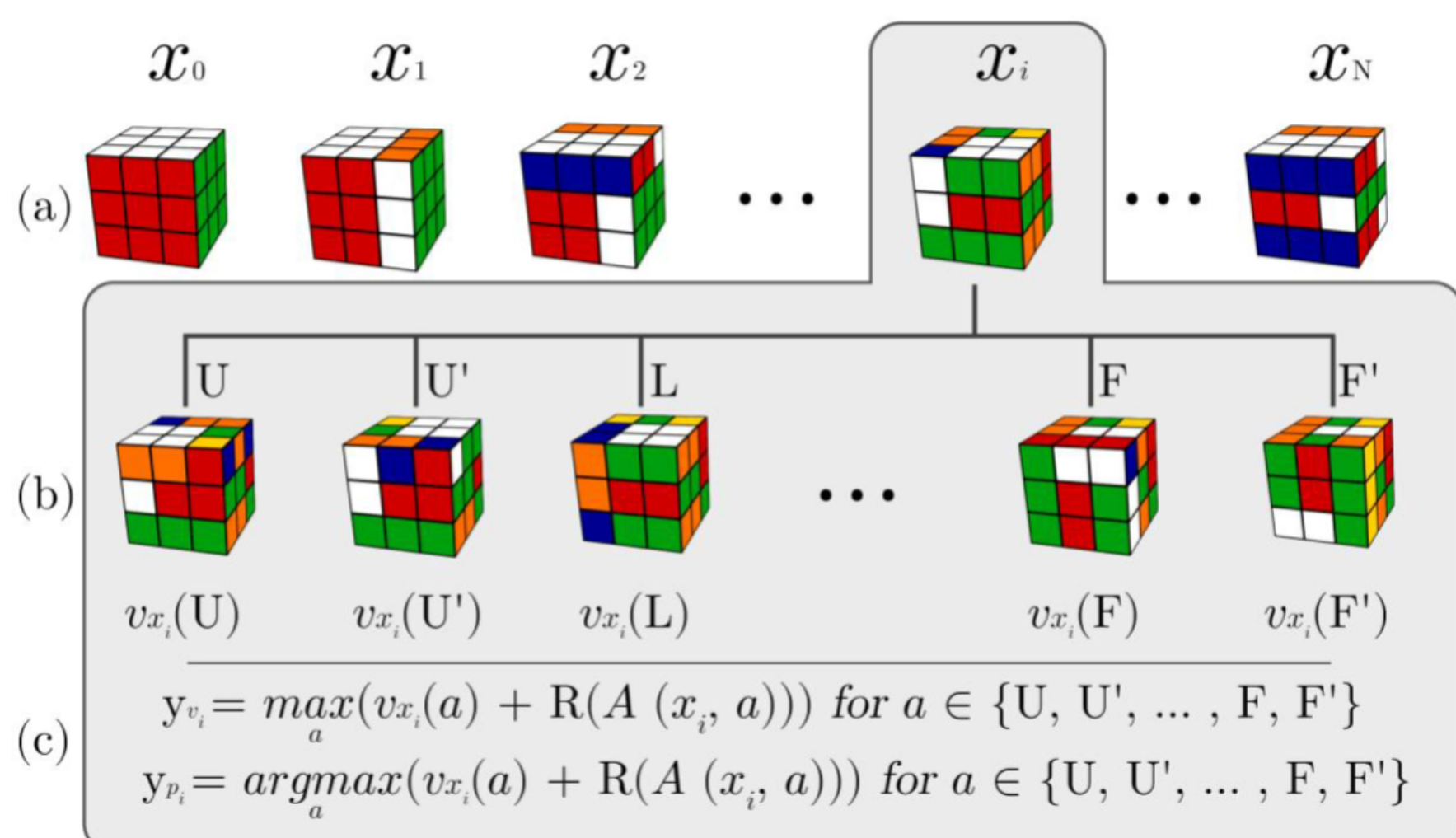


Image taken from: S. McAleer et al.: *Solving the Rubik's Cube with Approximate Policy Iteration*. ICLR 2019.

Suchalgorithmen

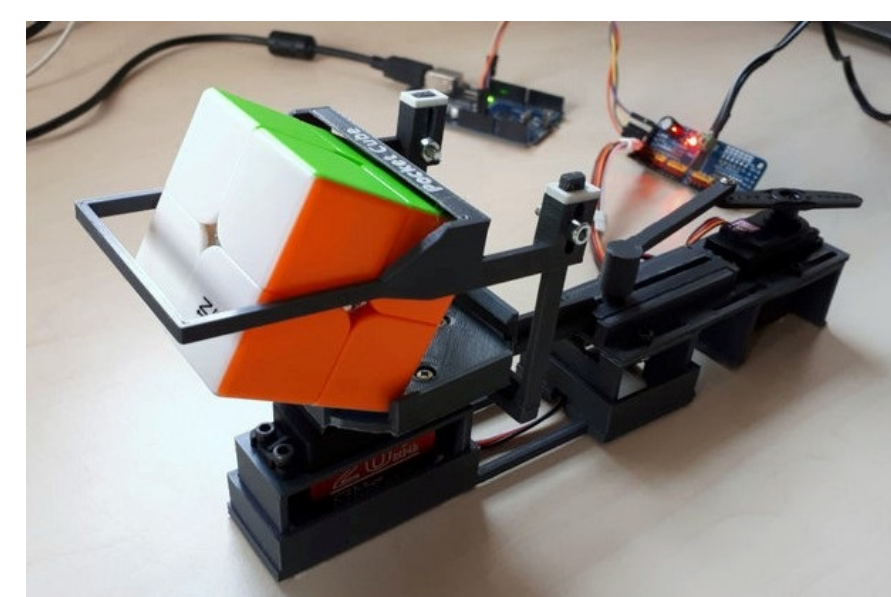
Aufgrund der Komplexität und des enormen Zustandsraums des Zauberwürfels ist eine alleinige Lösungswegfindung über ein neuronales Netzwerk nicht praktikabel. Daher werden neuronale Netze mit den Suchalgorithmen Monte-Carlo-Baumsuche bzw. dem A*-Algorithmus kombiniert. Hierbei wird das Netz als richtungsweisende Instanz verwendet, indem dieses Zustände des Würfels bewertet und somit vielversprechende Wege zum Zielzustand ermittelt.

Ergebnisse

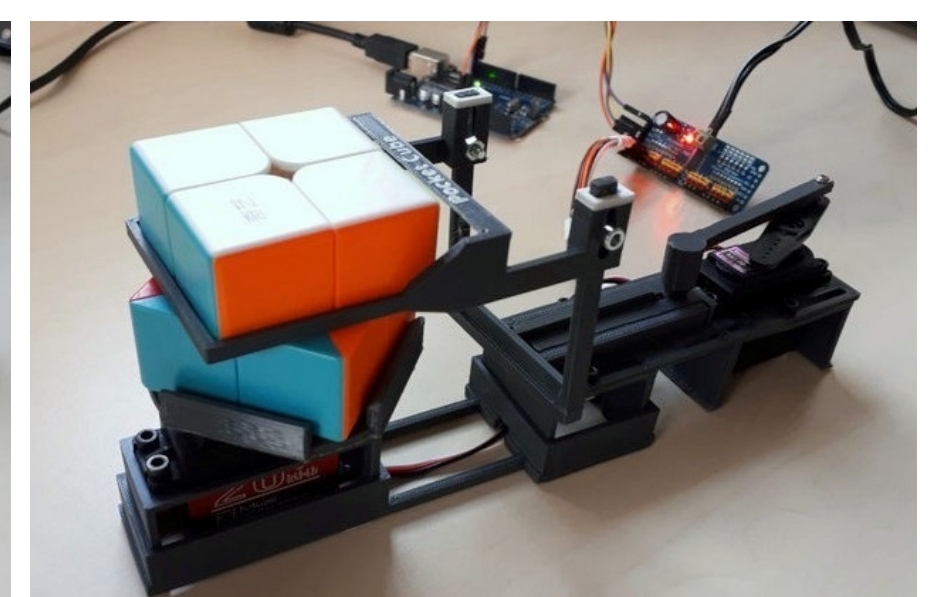
Die auf Monte-Carlo-Baumsuche bzw. dem A*-Algorithmus basierenden Modelle wurden auf Würfel, die 50 mal zufällig verdreht wurden, angewendet. Beide konnten für alle Zustände einen Lösungsweg ermitteln (100% Erfolgsquote), wobei die Monte-Carlo-Baumsuche etwas weniger Rechenzeit benötigt.

Demonstrator

Begleitend zu dieser Arbeit hat Prof. Marc Hensel einen Demonstrator mitsamt Python-API entwickelt, in welchen ein Pocket Cube eingelegt und durch die Ansteuerung zweier Servomotoren beliebig verdreht werden kann.



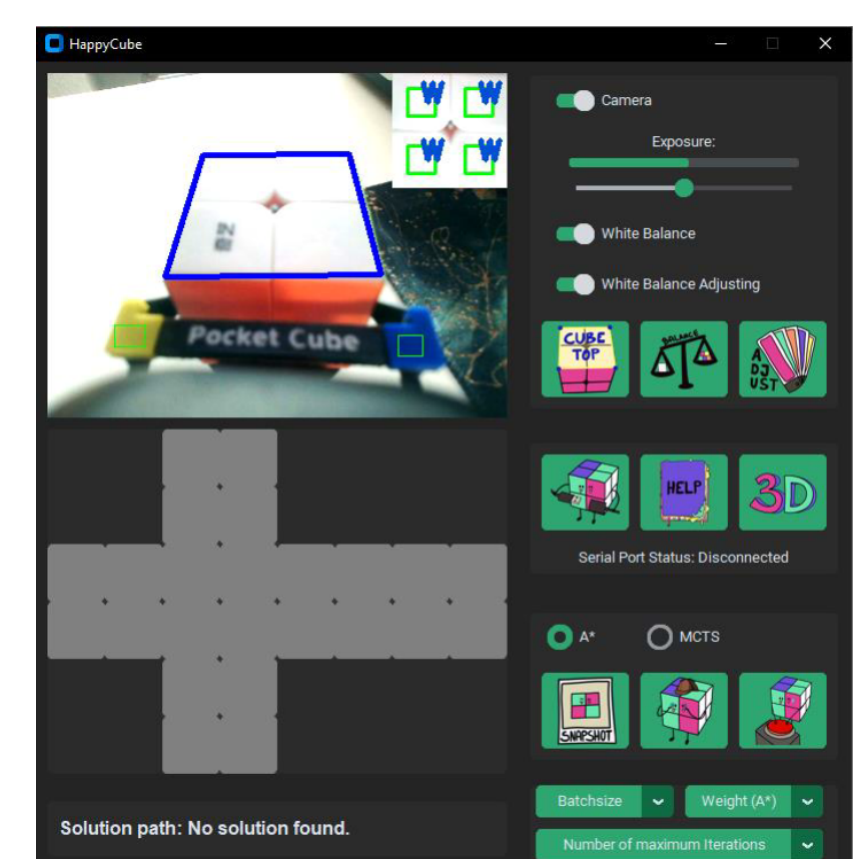
(a) Kippen des Würfels



(b) Rotation der unteren Seite

Bedienoberfläche

Der Demonstrator und die Methoden wurden in eine eigens entworfene Bedienoberfläche integriert. Durch Verwendung einer Kamera, welche auf der Hardware montiert wird, lässt sich der Würfelzustand automatisch erfassen. Nach Findung des Lösungsweges über einen der Suchalgorithmen und das zugehörige neuronale Netzwerk wird der Würfel mittels entsprechender Bewegungsabfolgen der Servomotoren zurück in den Zielzustand versetzt.



Projektseite im Internet

Masterarbeit, 3D-Druckdateien und Dokumentation des Demonstrators, Quelltexte, Modelle und mehr werden nach und nach auf der über den QR-Code verlinkten Webseite zur Verfügung gestellt.

