**Methode**

Beim Q-Learning handelt es sich um eine einfache Form des Reinforcement Learning. Das Prinzip basiert darauf, dass der KI in jedem Zustand eine Reihe von Aktionen zur Auswahl stehen. Je nachdem welche Aktion sie wählt erhält sie eine Belohnung oder Bestrafung, die sich zum einen Teil aus dem unmittelbar reichten Zustand und zum andern aus den später erreichbaren Zuständen ableiten lassen. Indem die KI versucht ihre Belohnungen zu maximieren lernt sie über mehrere Durchläufe was die bestmögliche Spielstrategie ist. [vgl.1]

Im Falle des Spiels Vier-Gewinnt Definieren sich die Zustände durch die aktuelle Position aller Steine auf dem Spielbrett. Die Aktion der KI ist das Werfen eines Spielsteins in eine der Spalten.

**Verwandte Arbeiten**

Olszowka beschäftig sich in seiner Arbeit neben einer allgemeinen Implementierung von Q-Learning für vier Gewinnt mit den damit verbundenen Schwierigkeiten, die ihm in einer früheren Arbeit begegnet sind. Unter anderem geht er auf die große Anzahl von Spielzuständen beim klassischen Vier-Gewinnt und beschäftigt sich mit Methoden um die Anzahl der für den Algorithmus relevanten Zustände zu reduzieren. [vgl. 2]

Șimșek und Barto stellen neben vielen anderen mit Q-Learning gelösten Problemen eine KI für Tic-Tac-Toe vor. [3] Tic-Tac-Toe ist zwar auf Grund der kleineren Spielfeldgröße weniger komplex als Vier-Gewinnt. Die Spielweise ist jedoch ähnlich.

[1] C. J. C. H. Watkins and P. Dayan: Q-learning. Machine Learning, 8:279–292, 1992

[2] D. Olszowka: Reinforcement-Learning in der Domane von PL-1