

Reinforcement learning Rubiks Cube

Voor deze reinforcement learning opdracht gebruik ik de meegegeven Rubiks cube code van de dlo.

Voor dit spel probeert een ai Rubiks cube op te lossen. Het kan elke keer 6 keuzen maken; naar voor, achter, links, rechts, boven of onder. Voor elke stap dat er wordt genomen komen er punten bij of gaan er punten af. Bij Q-learning wordt er geleerd om het beste keuze te nemen in een bepaalde staat van een omgeving, met behulp van een Q-tabel. De Q-tabel wordt gebruikt om Q-waardes te kunnen opslaan. Q-values zijn de geschatte waardes die bij elke stap wordt genomen. In dit geval van de Rubiks cube, krijgt de cube bij elke stap min punten en probeert het steeds stappen te maken om zo de Q-values te kunnen bereiken, als het geen slechte stap maakt, krijgt het geen min punten. Zo gaat het algoritme door totdat hij het doel bereikt heeft en een beloning krijgt bij het kunnen afronden van de cube.