

Reinforcement Learning en Computerspellen



Matthijs Gorter
Thom Brinkhorst
Pepijn van Iperen

Profielwerkstuk
onder begeleiding van
S. Rook
Christelijk Lyceum Zeist
Natuur en Techniek
Februari 2025

Voorwoord

Voorwoord inhoud hier. Bedank mensen die geholpen hebben, beschrijf het doel van het profielwerkstuk en eventuele persoonlijke motieven of ervaringen.

Matthijs Gorter, Thom Brinkhorst, Pepijn van Iperen
Christelijk Lyceum Zeist
Februari 2025

Variabelen en Notatie

| Variabele/ functie | Definitie |
|-----------------------|---|
| t | Tijdstap |
| T | Laatste tijdstap van een episode (horizon) |
| x | Toestand (state) |
| x_t | Toestand op tijdstip t |
| x' | Toestand een tijdstap na x |
| \mathcal{X} | Set van alle toestanden |
| a | Actie |
| \mathcal{A} | Alle mogelijke acties |
| a_t | Actie op tijdstip t |
| r | Beloning (reward) |
| \mathcal{R} | Set van mogelijke beloningen |
| r_t | Beloning op tijdstip t |
| $r(x, a)$ | Beloningsfunctie |
| μ | Deterministisch beleid |
| π | Stochastisch beleid |
| π^* | Optimale stochastisch beleid |
| γ | Kortingsfactor tussen 0 en 1 |
| $p(x' x, a)$ | Overgangswaarschijnlijkheidsfunctie |
| \mathcal{P} | Overgangswaarschijnlijkheidsmatrix |
| $V(x)$ | Waardefunctie |
| $Q(x, a)$ | Q-functie |
| $Q^*(x, a)$ | Q-functie met het optimale beleid |
| $\mathbb{E}[X]$ | Verwachtingswaarde van variabele X |
| $\mathbb{E}[a b]$ | Geconditioneerde verwachtingswaarde |
| $\mathbb{E}_\pi[X]$ | Verwachtingswaarde als beleid π wordt gevolgd |

Tabel 1: Variabelen en Notatie

Inhoudsopgave

| | |
|---|----|
| Voorwoord | 1 |
| Variabelen en Notatie | 2 |
| Inhoudsopgave | 3 |
| A. Inleiding | 4 |
| B. Theoretisch Kader | 5 |
| B.1. Definitie | 5 |
| B.2. Belangrijke Reinforcement Learning-Algoritmes | 6 |
| B.3. Vergelijking met andere Artificiële Leermethodes | 6 |
| B.4. Reinforcement Learning in Computerspellen | 6 |
| C. Onderzoeksmethoden | 7 |
| D. Analyse en Resultaten | 8 |
| E. Conclusie | 9 |
| F. Discussie | 10 |
| G. Bronvermelding | 11 |
| H. Bijlagen | 12 |

A Inleiding

Introductie Onderwerp

Onderwerpkeuze verantwoorden

Onderzoeksvraag/Hoofdvraag met eventuele hypothese

Deelvragen

Wat zijn de specifieke kenmerken van verschillende soorten computerspellen?

Welke reinforcement learning-algoritmes zijn beschikbaar en wat zijn hun kenmerken?

Hoe beïnvloeden de spelkenmerken de prestatie van reinforcement learning-algoritmes?

klein stukje theorie als inleiding op het theorie-onderdeel

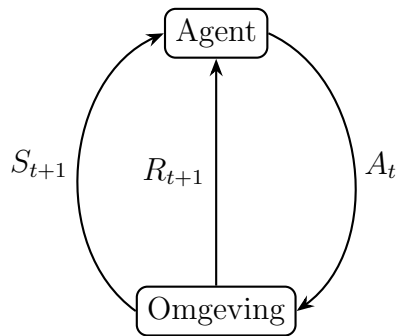
werkplan in grote lijnen, opbouw van verslag

B Theoretisch Kader

B.1 Definitie

Reinforcement Learning (RL) is een tak binnen kunstmatige intelligentie waarin een agent leert door interactie met zijn omgeving. Een agent is een entiteit die leert en acties onderneemt. Bij een zelfrijdende auto is het besturingssysteem de agent, en bij een schaakspel is de schaker de agent. De omgeving is alles waarmee de agent interageert en die reageert op de acties van de agent. Bij een zelfrijdende auto is dit de weg waar de auto op rijdt en de voertuigen om de auto heen. Bij een schaakspel is dit het schaakbord.

De agent leert door interactie met zijn omgeving. De agent ontvangt beloningen of straffen (negatieve beloningen) als gevolg van zijn acties. Het doel van de agent is om een strategie te ontwikkelen die de cumulatieve beloning maximaliseert over tijd. Bij Super Mario Bros (1985) is Mario de agent en de agent krijgt een beloning als Mario richting de eindslag beweegt (meestal naar rechts) en als Mario een coin of een power up oppakt en Mario krijgt een straf als Mario sterft en hij krijgt elke seconde straf zodat hij zo snel mogelijk het level wilt voltooien.



Figuur 1: Reinforcement learning interactiemodel tussen agent en omgeving via acties, toestanden en beloningen.

Een actie naar de beslissing die een agent neemt bij elke stap in een besluitvormingsproces. Acties worden aangeduid met a en worden gekozen uit een reeks mogelijke acties \mathcal{A} . Elke door de agent genomen actie beïnvloedt de interactie met de omgeving, wat leidt tot een verandering in de toestand en een daaruit voortvloeiende beloning.

B.2 Belangrijke Reinforcement Learning-Algoritmes

B.3 Vergelijking met andere Artificiële Leermethodes

B.4 Reinforcement Learning in Computerspellen

C Onderzoeksmethoden

D Analyse en Resultaten

E Conclusie

F Discussie

G Bronvermelding

H Bijlagen