

# Portfolio cursus Adaptive Systems

Stan Meyberg

## Leerdoelen

Begrip heeft van (verschillende vormen van) leren in artificiële agenten en de relatie met de werking van een natuurlijk brein

Bij reinforcement learning leert de agent door beloningen uit de omgeving gebaseerd op acties die hij neemt. Wanneer hij zich goed gedraagt, en acties neemt die bevorderlijk zijn, wordt hij beloond met meer beloningen. Ook mensen werken voor een groot deel op deze manier. Met kleine kinderen kunnen bijvoorbeeld afspraken gemaakt worden dat ze stickers krijgen als ze zich goed gedragen. Verder werkt de mens ook via lange en korte termijn beloningen. Deze beloningen moeten net als bij een agent goed afgesteld worden om uiteindelijk tot gewenst gedrag te komen. Wat betreft begrip van de relatie tussen RL-agents en de werking van het natuurlijk brein heb ik dit leerdoel behaald.

Bedrijfsprocessen kan analyseren en motiveren welke lerende aspecten toegevoegd kunnen worden om systemen zelflerend te maken

Bij casus over de mogelijkheden voor het gebruik van RL om de doorstroom te optimaliseren bij clusters van verkeerslichten moest eerst gekeken worden wat de huidige situatie was. Hierna is gekeken en gemotiveerd (op basis van literatuuronderzoek en experimenten tijdens het gehele vak) welke mogelijkheden er waren voor het implementeren van een zelflerend systeem. Hierbij zijn ook de voor- en nadelen toegelicht. Dit leidt ertoe dat ik kan concluderen dat dit leerdoel behaald is.

Weet hoe lerende algoritmen geïmplementeerd kunnen worden, gebruik makend van machine learning of game theory technieken

Tijdens de eerste weken zijn meerdere lerende algoritmen geïmplementeerd. Hierbij werd op basis van pseudocode en wiskundige notatie python-code geschreven. In de laatste weken is meer begrip gekweekt voor DQN in de praktijk.

Kan onderbouwen welke impact de toevoeging van lerende aspecten heeft op de bedrijfscontext en gebruikte systemen binnen een bedrijf

Binnen het bedrijfsleven zijn redelijk wat mogelijkheden voor oplossingen die gebruik maken van lerende algoritmes. Toch zijn RL-methodes niet altijd aan te raden aangezien hiervoor veel tijd en data benodigd is. Ook is niet altijd bij voorbaat zeker dat er een gewenst resultaat uitrolt.

De gewenste oplossing kan realiseren m.b.v. een bestaand framework voor de realisatie van een autonoom systeem en gebruik makend van bestaande tooling voor het integreren van AI cloud services zoals Azure of AWS

Voor het implementeren van DQN is gebruik gemaakt van Pytorch. Aangezien DQN veel training nodig heeft is er gebruik gemaakt van Google Colab voor het trainen van het model. Het voordeel van Colab is dat Google GPU's en TPU's beschikbaar stelt om het trainen te versnellen.

In staat is om ethische afwegingen te maken bij de keuzes die hij/zij maakt in de opzet van lerende agent-based oplossing

In de adviescasus voor het aansturen van verkeerslichten met RL-methodes moest ook de ethiek hiervan besproken worden. In een toegewijd hoofdstuk hieraan zijn meerdere ethische kwesties besproken die voor kunnen komen wanneer RL-methodes gebruikt zouden worden in de praktijk.

## Introducerende opdrachten

Waar heb je de meeste aandacht aan gegeven en waarom?

Tijdens de opdrachten heb ik veruit de meeste aandacht gegeven aan de theorie. Het boek maar ook de filmpjes van DeepMind gaven veel uitleg over de theorie achter de praktijk. Ik vond het belangrijk om juist deze kennis te hebben, aangezien ik af en toe het gevoel heb dat ik de achterliggende theorie van onderdelen mis. Verder heb ik de aan de Model-Free Prediction and Control opdracht ook veel tijd besteed omdat ik hier geprobeerd heb om alle functies en de omgeving zo functioneel mogelijk op te zetten zodat het makkelijker hergebruikt zou kunnen worden.

Welke leeruitkomsten leverden het maken van deze opdrachten je op (wat heb je ervan geleerd)?

Ik heb tijdens deze opdracht mijn kennis van Markov Chains en van Markov Decision processes kunnen uitdiepen. Ook heb ik geleerd hoe de Q-table in elkaar zit en heb ik intuïtie opgedaan over het gebruik van de discount en de invloed ervan op het leerproces.

## Infosupport challenge

Wat was jouw inbreng in het project?

Voor het project heb ik kleine aanpassingen gemaakt aan de tweede notebook, zodat er wat makkelijker naar verwezen kon worden in het verslag. Verder heb ik voor de opmaak en indeling van het verslag gezorgd. In het verslag heb ik aan ieder hoofdstuk wel meegeschreven, maar ik heb ook veel overleggen gehad met Esmeralda over de literatuur en de experimenten en wat wel en niet meegenomen moet worden in het verslag.

Welke leeruitkomsten leverden het maken van deze opdrachten je op (wat heb je ervan geleerd)?

Voor de eerste notebook heb ik voornamelijk het nut en het belang van discretisatie geleerd. Hierbij kon ik ook direct kennis op doen over het feit dat je bij het discretiseren van staten rekening kan houden van de waardes binnen de state-space, aangezien sommige waardes veel meer zeggen dan andere waardes. Ook vond ik de manier waarop epsilon-decay toegepast wordt in de praktijk zeer interessant om te leren.

Bij de tweede notebook had ik meer moeite om te leren, aangezien ik de opzet van het NN netwerk niet begreep in combinatie met het leren van de agent. Ik heb hier wel wat lopen spelen met verschillende parameters en het verkleinen van de state-space. Ook heb ik een grafiek gemaakt waarin de gemiddelde reward en loss per episode geplot wordt.

### Wat heb je van anderen of aan anderen geleerd?

Bij de eerste notebook heb ik veel met Esmeralda gezeten om uit te leggen hoe Q-learning toegepast moest worden op de omgeving en hoe de states gediscrètiseerd moesten worden. Bij de tweede notebook heb ik veel naar de bugfixes gekeken van Thijs en Bas om erachter te komen waarom er foutmeldingen kwamen.

### Wat heb je meegekregen en toegepast vanuit het gastcollege?

Bas heeft veel ervaring van RL in de praktijk en kon dan ook interessante voorbeelden geven uit de praktijk. Ook kon ik bijvoorbeeld de vraag stellen over wat zijn stappenplan was als hij merkte dat een agent niet leerde, wat ik zelf interessant vond om het antwoord op te horen.