

# Portfolio Adaptive Systems

Berry Hijwegen

5-6-2021

# Leerdoelen

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

Veel manieren van leren die artificiële agenten gebruiken lijken op hoe wij als mensen leren. Een voorbeeld hiervan is Q-Learning. Door allerlei neuronen aan elkaar te koppelen waar informatie verwerkt wordt, wordt er een bepaald begrip gecreëerd. Mensen observeren een bepaalde situatie (input) en op basis van wat je als persoon hebt meegemaakt en eerder hebt geobserveerd (memory) maak je een bepaalde beslissing waarvan jij denkt dat dit de uitkomst biedt. Dit is uiteindelijk ook hoe Q-Learning in elkaar zit. Uiteindelijk komen de algoritmes die we in het vak tegen zijn gekomen neer op trial-and-error: Het systeem is in een bepaalde staat, voert een actie uit, we observeren of dit een positieve of negatieve uitkomst heeft, we noteren de uitkomst, en nemen deze mee in volgende beslissingen. Dit kwam voornamelijk terug in opdracht 2 en opdracht 3. In opdracht 3 werd Q-learning gebruikt om de OpenAI CartPole en MountainCar te leren functioneren. Dit gaf een erg duidelijk beeld van wat voor stappen er nou genomen worden in het proces om dit voor elkaar te krijgen, vooral met de visuele interface erbij.

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

Veel bedrijven willen graag met zelflerende systemen werken, als AI'er zijnde is het belangrijk om te weten wanneer dit mogelijk is, en voornamelijk wanneer dit nodig is. Daarbij is het belangrijk om de risico's te kunnen aangeven die er bij komen kijken. Dit kwam erg goed naar voren in opdracht 3, het was al gauw duidelijk welke lerende aspecten er mogelijk zijn, maar een rode lijn door het verslag was voornamelijk dat er een stuk voorspelbaarheid verdwijnt bij de implementatie van een zelflerend systeem. Traditionele logische softwaresystemen zijn gebaseerd op vaste outputwaardes bij dezelfde inputwaardes. Bij een zelflerend systeem zijn deze outputwaardes dynamisch, door het lerende aspect. Dit leidt mogelijk tot meer capaciteiten en betere prestaties, maar brengt met zich mee dat dit ook mogelijke negatieve gevolgen kan hebben. Daarom is het belangrijk om bij de implementatie van zo'n zelflerend systeem ook te kijken naar een controlesysteem wat in kan grijpen. Dit was een aspect aan het zelflerende systeem waar ik voor we dit vak kregen nog niet zo goed over na had gedacht, maar wat voor ik persoonlijk een interessant onderwerp vindt binnen het implementeren van zelflerende systemen in de praktijk.

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

In de eerste twee opdrachten is er vooral gefocust op de theorie, hierin hebben we in de eerste opdracht gekeken naar de wat meer 'bewezen' theorie. Hierbij kent de agent de wereld om zich heen. In de tweede opdracht keken we naar meer experimentele theorieën. Deze gingen over een situatie waarbij de wereld niet bekend was, hierbij is zijn de algoritmen trial-and-error en leren ze door te proberen. Er moeten eerst fouten gemaakt worden om te zien dat iets fout is, in plaats van dat dat al van te voren bekend is. Deze theorieën hebben ik geïmplementeerd in de eerste en tweede opdracht. In de derde opdracht ging het voornamelijk over bestaande oplossingen met openAI Gym, voornamelijk in de praktijk lijkt me dit erg nuttig, zodat je niet zelf een basis hoeft te schrijven. Het skelet van het systeem staat dan al, en vervolgens met de theorieën die we gehad hebben kan het systeem dan ontwikkeld worden.

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

Zoals al eerder genoemd kwam dit sterk voor in het adviesrapport. Voor een bedrijf zelf brengen zelflerende systemen voordelen en risico's met zich mee. In vergelijking met een traditioneel logisch systeem kunnen de prestaties beter zijn, maar er valt bij een zelflerend systeem een stuk voorspelbaarheid weg. Daarom is het voor een bedrijf belangrijk dat een controlesysteem wordt

ingesteld bij een zelflerend systeem. Een zelflerend systeem biedt namelijk niet altijd de 100% zekerheid die nodig is voor de continuïteit van een bedrijf.

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

Dit punt sluit goed aan bij het vorige punt, de risico's die komen kijken bij zelflerende systemen. Bijvoorbeeld bij de casus van de Utrechtse verkeerslichten. Hier zitten zowel ethische als juridische aspecten aan. Zo kan het dat bij een incorrecte implementatie van het zelflerende verkeerssysteem meer doden vallen dan in de oude situatie. Is daar iemand verantwoordelijk voor? Zo ja, wie? Om dit goed op te vangen is het eerder genoemde controlesysteem een oplossing, en tevens het uitvoeren van pilots om aan te tonen dat het systeem werkt. Ook moet de AI'er grenzen kunnen aangeven, wanneer het risico op een negatief gevolg te groot is, daar zouden grenzen getrokken moeten worden en aangegeven moeten worden dat een AI-oplossing niet gewenst is, dit kan bijvoorbeeld om leed of privacy gaan.

## **Inleveropgave 1**

### **Waar heb je het meeste aandacht aan gegeven en waarom?**

In de eerste opdracht heb ik voornamelijk gefocust op het opzetten van een goede basis. Ik heb een objectgeoriënteerde oplossing bedacht die bij de eerste opdracht netjes werkte. Daarnaast heb ik goed gekeken hoe de Value Iteration in elkaar stak. Dit werd voornamelijk goed duidelijk tijdens het met de hand uitschrijven van de stappen, dit maakte de implementatie overzichtelijker voor mij.

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

- Ik heb geleerd hoe het mogelijk is om een zelflerend systeem te implementeren in een bekende omgeving.
- Ik heb geleerd hoe Markov Chains schematisch gelezen kunnen worden.
- Ik heb geleerd hoe verschillende algoritmen werken om een optimale policy te vinden in een bekende omgeving.

## **Inleveropgave 2**

### **Waar heb je het meeste aandacht aan gegeven en waarom?**

Bij de tweede opdracht heb ik de meeste aandacht gegeven aan het SARSA algoritme, ik heb hier al eerder van gehoord had en ik erg benieuwd was hoe het in elkaar stak. Echter heb ik de structuur van de vorige opdracht gekopieerd naar de tweede opdracht, maar dit bleek naderhand niet echt goed te passen bij wat de opdracht vroeg, mede hierdoor moet ik deze opdracht nog verbeteren.

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

- Ik heb geleerd hoe het mogelijk is om een zelflerend systeem te implementeren in een onbekende omgeving.
- Ik heb geleerd hoe verschillende algoritmen gebruikt kunnen worden en wat de nadelen zijn van sommige algoritmen.
- Ik heb geleerd hoe verschillende algoritmen zoeken naar een optimale policy in een onbekende omgeving.

# Infosupport challenge

## **Wat was jouw inbreng in het project?**

Het Info support challenge project heb ik samen met Selene gedaan. De Notebooks hebben wij gedaan door middel van pair-programming met screen share, omdat het lastig was om tegelijk te werken in de Google Colab notebook. Hierbij hebben we gekeken naar wat we konden aanpassen aan de notebooks om de resultaten te verbeteren en gekeken hoe dit uitpakte.

In het adviesrapport heb ik uitgezocht wat de huidige situatie is bij verkeerslichten en hoe de gewenste situatie er uit zou moeten zien. Ook heb ik hierbij de voor- en nadelen opgenoemd. Hierbij kwamen Selene en ik in gesprek over het ethische en praktische aspect, dit bracht het controlestelsel in beeld.

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

Van de notebooks heb ik vooral geleerd hoe een bestaande omgeving met OpenAI Gym gebruikt kan worden om algoritmes op te trainen. Dit lijkt me erg nuttig voor in de praktijk omdat je hiermee al snel een opzet kan bouwen in plaats van dat je zelf de omgeving bouwt.

In het adviesrapport heb ik veel geleerd over hoe een zelflerend systeem praktisch kan werken en wat er allemaal bij komt kijken, een zelflerend systeem opzichzelf kan erg risicovol zijn, en daarom is het belangrijk om goed na te denken over hoe je risico's kan indekken.

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

Selene wist al erg veel van de theorie van Reinforcement Learning door eerdere ervaringen. Ik heb hier veel van opgepikt en waarom welke parameters nou bepaalde waarden hebben en wat deze waarden doen. Hier heb ik veel van geleerd en gaf veel inzicht in de algoritmen.

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

Het gastcollege was erg interessant. Bas gaf nieuwe inzichten in hoe Reinforcement Learning gebruikt kan worden. Voor het adviesrapport vond ik het niet erg aansluiten, maar het gaf wel een inzicht hoe RL in de praktijk werkt.