



Universiteit
Leiden
The Netherlands

Affect and Learning: a computational analysis

Broekens, D.J.

Citation

Broekens, D. J. (2007, December 18). *Affect and Learning: a computational analysis*. Leiden Institute of Advanced Computer Science (LIACS), Faculty of Science, Leiden University. Retrieved from <https://hdl.handle.net/1887/12537>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/12537>

Note: To cite this publication please use the final published version (if applicable).

Samenvatting (Dutch)

Hier zal een korte samenvatting gegeven worden van de resultaten van het onderzoek gepresenteerd in dit proefschrift.

Affect, Gemoedstoestand en Informatieverwerking

Om te overleven moet elk wezen acties selecteren (“wat ga ik nu doen...”). *Actieselectie* is gedefinieerd als het probleem om te beslissen wat de volgende actie zal zijn zodanig dat de kans op overleving wordt geoptimaliseerd. *Reinforcement Learning* (RL) is een manier om leerprocessen op basis van positieve en negatieve terugkoppeling te modelleren door middel van computationele modellen. Een agent leert welke actie in welke situatie welke verwachte waarde heeft. Deze waarde wordt bepaald door de terugkoppelingen die er in de omgeving bestaan en door de ervaringen die de agent met deze omgeving heeft. In RL moet een mobiele agent dus steeds kiezen welke actie gedaan moet worden op basis van de waardes van de verschillende acties. Uiteindelijk wil die agent positieve terugkoppeling maximaliseren en negatieve terugkoppeling minimaliseren.

Een probleem dat de lerende agent hierbij moet oplossen is *exploratie* versus *exploitatie*. Exploratie (trial and error) is het proberen van nieuwe acties en het leren van wat er goed en niet goed aan is. Exploitatie is het kiezen van acties die volgens wat je geleerd hebt de beste zijn. Deze twee processen moeten bij een lerende agent afgewisseld worden. Neem als voorbeeld boodschappen doen in een nieuwe stad. Eerst zoek je naar de kortste weg naar een supermarkt (exploratie), en nadat je dit een aantal keer hebt gedaan denk je te weten wat de kortste weg is. Vervolgens neem je altijd de route waarvan je denkt dat deze het kortst is (exploitatie).

In Hoofdstuk 3 en 4 van dit proefschrift is er onderzocht hoe gemoedstoestand de keuze voor exploratie versus exploitatie kan beïnvloeden. Er is gebruik gemaakt van gesimuleerde robotjes. De robots moeten leren hoe ze door een doolhof kunnen navigeren. Ze moeten zo goed mogelijk de weg naar het doel leren. Je zou kunnen zeggen dat een gesimuleerde muis in een doolhof op zoek is naar kaas.

In Hoofdstuk 2 is een model van gemoedstoestand (affect) voor dit soort lerende robotjes ontwikkeld. Dit model gaat ervan uit dat de robotjes een stemming kunnen hebben die varieert van goed tot slecht afhankelijk van hoe

goed het met ze gaat. Hoe goed het gaat hangt af van de gemiddelde hoeveelheid straf en beloning die ze krijgen, ten opzichte van wat ze gewend zijn. Dus, als ze steeds beter weten waar de kaas is (of, als wij steeds beter weten waar de supermarkt is), gaat hun stemming vooruit. Als ze echter steeds meer tijd nodig hebben om de kaas te vinden, of steeds vaker tegen de muur aan lopen (straf) terwijl ze door de doolhof bewegen, dan gaat hun stemming achteruit.

In Hoofdstuk 3 hebben we dit model van gemoedstoestand gekoppeld aan het exploratiegedrag van de gesimuleerde robotjes. Uit dit onderzoek blijkt dat robotjes die gaan exploreren als ze zich slecht voelen en gaan exploiteren als ze zich goed voelen sneller het beste pad naar hun doel leren. Dit is niet altijd zo, maar vooral als het doel (de kaas) plotseling naar een andere plek in de doolhof wordt verplaatst. Robotjes gaan zich dan minder goed voelen (ze vinden het doel immers niet meer), en gaan daardoor exploreren (nieuwe dingen proberen). Hierdoor vinden ze het nieuwe doel sneller dan robotjes die gewoon door blijven lopen op het oude pad naar het doel. Als de affectieve robotjes de nieuwe plek hebben gevonden, gaan ze zich langzaam beter voelen. Hierdoor gaan ze meer exploiteren (goede acties kiezen op basis van wat ze geleerd hebben). Daardoor doen ze minder probeeracties (exploratie), waardoor ze sneller het beste pad leren dan robotjes die geen affectgestuurd exploratiegedrag hebben.

In Hoofdstuk 4 hebben we verder gekeken naar een andere manier om gemoedstoestand te koppelen aan leergedrag. In dit onderzoek is de stemming van de lerende robot gekoppeld aan hoe de robot vooruit denkt. Vooruit denken is in dit geval het anticiperen op de mogelijke gevolgen van een actie, voordat de actie is uitgevoerd. Het is dus een soort van intern simuleren van gedrag om te voorspellen wat er zou kunnen gebeuren. De stemming (weer variërend van goed tot slecht) wordt nu gekoppeld aan de hoeveelheid positieve gedachten die de robot heeft. Als de stemming goed is, denkt deze alleen aan positieve gedachten; als de stemming slecht is, denkt de robot aan zoveel mogelijk. Dus, goed voelen betekent goed denken, en slecht voelen betekent breed denken. Uit dit onderzoek blijkt dat er geen positief effect is van affectgestuurd denken op de leersnelheid van de robot. Deze leert dus niet sneller waar het voedsel is. Wel hoeft de robot *minder* te denken om hetzelfde resultaat te behalen. Hieruit zou dus geconcludeerd kunnen worden dat het gunstig is voor de totale hoeveelheid benodigde denkinzet tijdens een leerproces om vooral breed over mogelijke consequenties na te denken als het minder goed gaat, maar vooral over positieve mogelijke consequenties na te denken als het goed gaat.

Er is een grote kanttekening bij deze resultaten: het zijn computationele modellen! Ook zijn het modellen die getest zijn in simpele omgevingen: kleine

doolhofjes met maar een paar verschillende objecten. Het is dus niet mogelijk om deze conclusies definitief te veralgemeniseren naar bijvoorbeeld menselijk gedrag. Wat wel gezegd kan worden is het volgende. Ten eerste, gemoedstoestand lijkt een nuttige toevoeging voor lerende robots: ze kunnen in sommige situaties op gunstige wijze gebruik maken van hun stemming. Ten tweede, de relaties die onderzocht zijn tussen gemoedstoestand en leren laten aan de cognitieve psychologie zien *hoe* (het mechanisme) gemoedstoestand en leren zouden kunnen samenhangen. Hier kan vervolgens weer verder onderzoek naar gedaan worden.

Affect en Beloning

In Hoofdstuk 6 is een iets andere aanpak gekozen voor de koppeling tussen affect en leren. In de eerdere hoofdstukken was affect een signaal dat door de robot zelf werd gemaakt, en dat samenhang met gemoedstoestand: “hoe gaat het nu met me ten opzichte van wat ik gewend ben”. Dit is een lange termijn interpretatie van affect. In Hoofdstuk 6 is onderzocht hoe affect als signaal door een ander gecommuniceerd wordt aan een lerende robot. Het is voor mensen heel belangrijk dat ze emoties kunnen herkennen van anderen. Deze kunnen je bijvoorbeeld vertellen dat je iets niet meer moet doen, of juist wel. Dit principe van leren door middel van emotionele uitdrukking is onderzocht, maar dan tussen mensen en robots.

Er is onderzocht of een lerende robot beter leert als hij ook gebruik kan maken van een menselijke “ouder”. De ouder kijkt naar de gesimuleerde robot (weer in een doolhof, waar kaas in is verstopt) en de robot kijkt naar de ouder door middel van een *webcam*. De webcam vertaalt de gelaatsuitdrukkingen van de ouder (bijvoorbeeld een onderzoeker) naar een positief of negatief signaal. Dit signaal kan door de robot gebruikt worden als terugkoppeling. Het blijkt dat de robot beter leert als er een observerende ouder bij zit. De robot leert sneller wat de beste weg naar het doel is. Dit signaal werkt het best als de robot ook leert *wanneer* de ouder lacht of boos kijkt (de belangrijkste signalen die in ons model vertaald worden naar positieve of negatieve terugkoppeling). Dus, als de robot het signaal van de ouder alleen gebruikt om zijn gedrag aan te passen maar niet om ook een model op te bouwen van wat die ouder waarover vindt, dan helpt het signaal niet goed. Ook hier geldt weer: dit is een computationeel model, dus oppassen met de conclusies.

Formele Modellen.

In het laatste hoofdstuk (Hoofdstuk 7) is er een compleet andere benadering gekozen om inzicht te krijgen in emotionele processen. Er is voor een bepaald

type theorie van emotie, de *appraisal theorie*, een formele taal ontwikkeld die gebruikt kan worden om verschillende appraisal theorieën in op te schrijven. Appraisal theorieën gaan ervan uit dat emoties ontstaan door het vergelijken van een huidige situatie met toekomstige doelen en de actoren en hun rollen daarin. Als iets goed is voor mijn doelen, word ik blij, en andersom. Als iemand anders dat voor elkaar heeft gekregen ben ik die persoon dankbaar. Als ik het zelf heb gedaan ben ik trots. Als iemand anders iets doet dat slecht voor mij is word ik boos. Als er niemand verantwoordelijk voor is word ik verdrietig, etc. Het gaat hier te ver om precies uit te leggen hoe dit in zijn werk gaat; in Hoofdstuk 7 staan vele referenties naar de verschillende theorieën die er bestaan.

Het idee achter de ontwikkelde formele taal is dat verschillende theorieën allemaal beschreven kunnen worden in dezelfde taal. Hierdoor wordt het veel makkelijker om ze met elkaar te vergelijken. Ook kan de formele beschrijving van een dergelijke theorie beter gebruikt worden als basis voor het maken van computationele modellen van emotie. Waarom? Omdat een computationeel model heldere, duidelijke definities nodig heeft, en een formele beschrijving van een theorie duidelijkere definities heeft dan een in taal opgeschreven theorie. In Hoofdstuk 7 wordt laten zien hoe twee verschillende theorieën samengevoegd kunnen worden nadat ze beschreven zijn in de ontwikkelde formele taal. Vervolgens wordt er een computationeel model gemaakt op basis van deze samengestelde theorie.