

Research Paper

1. Introductie

1.1 Inleiding

Wij hebben besloten om te werken aan het container project van cofano. Hiervoor is ons groep zich gaan verdelen in verschillende deel problemen van het probleem zoals het plaatsen van de containers op de kade, visualisatie en het valideren/optimaal pakken van containers die op de kade staan voor de zee schepen die binnen komen. Hiervoor hebben we voornamelijk voor een reinforcement learning aproach gekozen om te proberen of daarmee het oplossen van het probleem mogelijk is.

Met dit onderzoek willen we een antwoord geven aan de vraag: hoe kunne we optimaal de containers op de kade plaatsen om zo min mogelijk tijd verlies te maken.

1.2 Leeswijzer

In dit research paper gaan we de probleemstelling benoemen en een doelstelling formuleren. Na de introductie gaan we beginnen aan het onderzoeksopzet dus met welke technieken en methodes we gewerkt hebben en de mogelijk meetinstrumenten om te evalueren in hoe verre ons onderzoek is gelukt bespreken. De resultaten die hieruit komen zullen worden geanalyseerd en geïnterpreteerd aan de hand van de meetinstrumenten. Uit het geïnterpreteerde data wordt er een conclusie getrokken waar er een antwoord wordt gegeven aan de onderzoeksvraag.

1.3 Probleemstelling

Het probleem is te verdelen in 4 deelproblemen: hoe kunnen we containers dynamisch op de kade plaatsen zonder dat de opstelling die al op de kade is inboxen, hoe kunnen we de opstelling zo zetten dat er rekening met onvoorziene omstandigheden gehouden wordt(zeeschepen hebben vertraging of komen in een andere volgorde dan verwacht binnen), hoe kunnen we de opstelling op de kade valideren om te voorkomen dat bijvoorbeeld bepaalde containers “geboxed” worden door andere containers , hoe kunnen we zo optimaal mogelijk de containers van de kade pakken in inladen op de zeeschepen zonder de opstelling te verpesten. Het valideren van de opstellingen is belangrijk om te kunnen evalueren hoe optimaal de kade is en of containers op de juiste plekken worden geplaatst.

1.4 Doelstelling

Ons doel is om te achterhalen in hoe verre we het proces van cofano zou optimaal mogelijk kunnen maken met voornamelijk behulp van een RL algoritme. We hopen om het proces van onze stakeholders(cofano) zo efficiënter mogelijk te maken met machine learning.

2. Onderzoeksopzet

2.1 Dataverzamelmethode

Door middel van een rl agent gaan we proberen om zien in hoeverre de agent kan leren en evalueren hoe goed een opstelling is. Hiervoor moeten we een 3x3 grid maken met containers en target containers. De agent moet de juiste weg vinden naar de target, het aantal iteraties/acties onthouden, evalueren of de target is “boxed” tussen andere containers en mogelijk containers kunnen verplaatsen om erbij te kunnen. Hiervoor zullen we synthetische data creëren door met elke game een nieuwe opstelling van containers te maken. Om dit vraagstuk te kunnen beantwoorden wordt er inspiratie genomen van de “Taxi-v3” opstelling van py gym.

2.2 Meetinstrumenten

Deelvraag3

Hoe kunnen we de opstelling op de kade valideren?

We vinden een rl agent het beste geschikt, omdat dit probleem veel te complex is voor traditionele manieren zoals lineair programmeren en algoritmes. In plaats van zelf dit enorm complex probleem oplossen laten we een neuraal netwerk het voor ons oplossen. Voor deze redenen vinden we een rl agent het meest geschikt. Het is ons plan om eerst onderzoek te doen naar vergelijkbare problemen en software. Een mogelijke kandidaat is de grid world environment wat als doel heeft om een agent zelf te leren om een target vakje te bereiken. Zo’n environment zou een mooi begin kunnen zijn voor het validatie probleem.

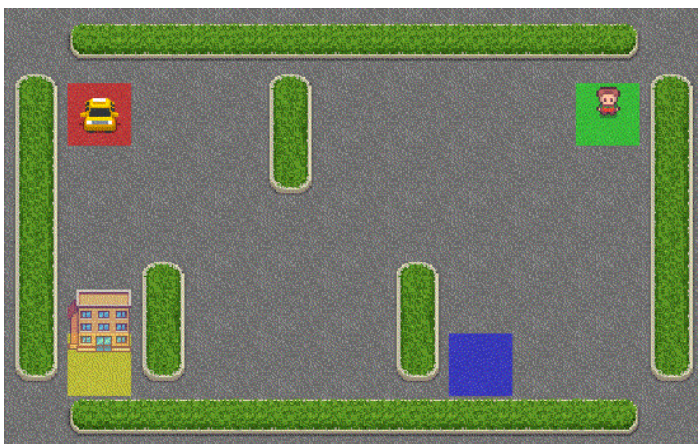
3. Resultaten

Om antwoord te kunnen op de vraag van hoe kunnen we de opstelling op de kade valideren moet er in kaart gebracht worden welke uitdagingen hierbij horen. Door het feit dat andere oplossingen voor dit probleem onrealistisch is voor ons om te uitvoeren, kunt u een reinforcement learning approach aannemen.

Om inzicht te krijgen in hoe de environment eruit zou zien is het noodzakelijk om inspiratie op te doen naar vergelijkbare environments

De Taxi-v3 is een environment met 4 locaties in de grid world, een agent en obstakels. De taxi heeft als doel om de target(passagier) naar het eindbestemming te brengen.

(*Taxi - Gym Documentation*, n.d.)



Aan de hand van deze environment zal ik nagaan of bepaalde elementen van de environment geschraapt of geïmplementeerd worden voor onze environment.

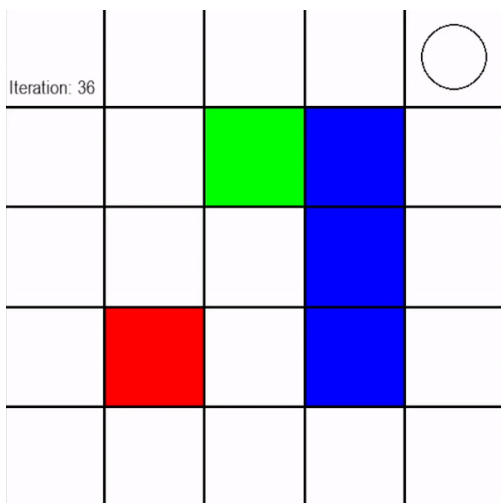
De taxi environment ziet er op een oogopslag best simpel uit, de taxi hoeft alleen van plek A naar B maar in het echt is het niet zo simpel. De taxi moet langs obstakels rijden, weten hoe ver hij van de passagier vandaan is, weten waar zijn eindbestemming en in het kortst mogelijke tijd zich dit process uitvoeren zonder vast te raken tegen een obstakel of naar de verkeerde bestemming toe rijden. Dit is al best complex. Ons container environment zou in principe het zelfde moeten verrichten met als verschil dat er niet 1 passagier is maar meerdere. Het environment zou dan als volgt uit moeten zien: Een groep containers bij elkaar in een bepaalde opstelling met een agent die zijn weg moet maken naar 1 target container.

Om de acties te definiëren is er een agent die in de grid wereld beweegt. De agent kan omhoog, omlaag, links en naar rechts. Op de kade van cofano zijn er zogenoemde "stackers" die containers kunnen pakken, neerzetten en hunzelf verplaatsen. Dit is vergelijkbaar met de taxi omdat beide moeten kunnen bewegen in de environment. Verder moet de taxi een "target" variabele(passagier) kunnen oppakken en naar zijn eind locatie brengen. Dit is vergelijkbaar met de stacker sinds de stacker ook een target variabele(de container) moet kunnen oppakken en inladen op de zeeschepen. Aangezien we rekening moeten houden met andere containers, kan de stacker niet meteen naar de target variabele rijden. In het geval dat de target container is boxed in dan moet de stacker zelf een weg er naar toe maken.

Sinds er geen data is van container opstelling zullen we die synthetische moeten maken. Zo'n opstelling kan in miljoenen mogelijk combinaties op de kade staan net als een rubiks kubus. Aangezien opstellingen maken enorm veel werk en tijd kost laten we de environment zelf random opstelling creëren om vervolgen te laten oplossen door de agent. De agent moeten worden beloond als hij op tijd bij de target variabele kan zijn en gestraft worden als hij de opstelling verpest, de verkeerde container pakt, niet van de lange zijde pakt en als het een container probeert te pakken terwijl er een container bovenop ligt.

4. Conclusie

Nu de dataverzameling en analyse is afgerond kan de definitieve algoritme worden bepaald. In onderstaande gif wordt duidelijk gemaakt wat het eindproduct is, welke features geschrapt zijn en welke niet.



Hoe kunnen we de kade nou valideren?

Om de opstellingen te kunnen valideren hebben we eerst een environment gemaakt. Deze environment neemt verschillende aspecten van de taxi-v3 en gridworldenv environments over. De environment bestaat uit een 5x5 grid met agent die noord, oost, zuid en west kan bewegen. De agent is de witte cirkel en heeft als doel om de target container te pakken (rood). De opstelling bestaat uit een 3x3 grid waar er random dummy containers(groen en blauw) geplaatst wordt met een rode container ergens tussenin. De agent heeft 200 iteraties de tijd om de target container te bereiken en in het geval dat hij faalt krijgt de agent een straf en reset de environment. Als de agent op een container vak staat gaan we ervan uit dat hij die container probeerde te pakken en checken we of de het target container was of een van de dummy containers. Als het de target was krijgt de agent een reward en reset de environment en zo niet krijgt de agent een punishment maar de environment wordt in dit geval niet reset. Verder krijgt de agent ook een punishment als het containers niet van de lange zijde pakt (noord en zuid kant).

De agent maakt gebruik van de PPO model en "MultiInputPolicy" van stable_baselines3 en de environment is gemaakt met py gym. Dit environment zou een goed begin kunnen om zijn om opstellingen op de kade te kunnen valideren en mogelijk vervolg onderzoeken.

5. Discussie

Dit onderzoeksproject was leuke maar ook heel complex. Helaas hebben we niet alle functionaliteit erin kunnen steken zoals containers kunnen oppakken en verplaatsen. De opstelling werkt alleen met een 3x3 grid en gaat dus niet de hoogte in. We begonnen deze project in opdracht van De Haagse Hogeschool en hebben het container project gekozen halverwege de minor. Omdat dit zo'n groot probleem is hopen we met dit onderzoek een impact te kunnen maken op dit het efficiënter opstellen en inladen van de kade.

Zelf hebben we ook veel geleerd over machine learning en de verschillende vormen ervan zoals supervised learning en reinforcement learning. Wij denken zelf dat dit onderzoek goed is geslaagd aangezien we elk probleem hebben kunnen oplossen.

6. Bibliografie

7. *Taxi - Gym Documentation.* (n.d.).

https://www.gymlibrary.dev/environments/toy_text/taxi/