

Vergelijken van verschillende onderzoeken naar het ontwerpen van een kunstmatige intelligentie die in staat is om het bordspel Risk te spelen

Kimberly Stessen 5609909
Bachelor Scriptie 01-07-2019
Bachelor Kunstmatige Intelligentie
Universiteit Utrecht

7,5 ECTS

Begeleider: Henry Prakken
2^e beoordelaar: Brandt van der Gaast

Abstract

In deze scriptie bespreek ik vier onderzoeken die een kunstmatige intelligentie (AI) hebben ontworpen die zelfstandig in staat is om het bordspel Risk te spelen. Risk is een oorlogsspel met als hoofddoel: alle landen op het speelbord veroveren. De onderzoeksvraag die ik in deze scriptie ga beantwoorden is: *Wat is de beste aanpak voor het ontwikkelen van een kunstmatige intelligentie voor het spel Risk?* De onderzoeken die ik in deze scriptie bespreek hebben allemaal gebruik gemaakt van een andere methode om een AI te ontwikkelen. Deze manieren zijn: heuristieken, netwerken en aanvalsheuristieken, multi-agent systems en temporal difference learning. Om de onderzoeken met elkaar te vergelijken heb ik gekeken naar de resultaten die elke AI behaald heeft en de omgang met bepaalde tactieken die een speler kunnen helpen om Risk te winnen. Door de verschillende manieren waarop Risk geïmplementeerd is en de AI's geëvalueerd zijn in de onderzoeken die ik bespreek is het niet makkelijk om een eenduidig antwoord te geven op de onderzoeksvraag. Verder onderzoek zal nodig zijn om aan te kunnen geven wat de beste aanpak is voor het ontwerpen van een AI die Risk kan spelen.

Inhoudsopgave

1. Inleiding	blz. 4
2. Spelregels Risk	blz. 5
3. Beste strategie voor het spelen van Risk	blz. 7
3.1 Kansberekening	blz. 8
4. Verschillende aanpakken voor het ontwerpen van een AI die zelfstandig Risk kan spelen	blz. 9
4.1 Heuristieken	blz. 9
4.2 Netwerken en aanvalsheuristieken	blz. 10
4.3 Multi-agent systems	blz. 11
4.4 Temporal difference learning	blz. 13
5. Onderzoeken naar het ontwerpen van een AI die Risk kan spelen vergelijken	blz. 15
5.1. Resultaten van de onderzoeken naar het ontwerpen van een AI die Risk kan spelen	blz. 15
5.2. Het gebruik van de tactieken door de onderzoeken naar het ontwerpen van een AI die Risk kan spelen	blz. 15
6. Discussie	blz. 17
7. Conclusie	blz. 18
8. Aanbevelingen	blz. 20
9. Bronnenlijst	blz. 21
10. Bijlage A: Overzicht verschillende onderzoeken	blz. 22

1. Inleiding

Het ontwikkelen van goede kunstmatige intelligenties (AI's) voor games is belangrijk binnen het vakgebied van de kunstmatige intelligentie. Spelers van games verwachten dat nieuwe spellen van betere kwaliteit zijn. Dit geldt niet alleen voor de graphics, maar ook voor de AI's die in de spellen voorkomen. Om deze reden wordt er door de game-industrie veel geld gestoken in het ontwikkelen van steeds intelligentere AI's. Deze AI's moeten steeds realistischer en menselijker gedrag vertonen, zoals een pad kunnen vinden, leren van fouten en rationeel denken. Vooral het rationeel en strategisch denken van een AI kan goed getest worden als we dit soort AI's een oorlogsspelletje laten spelen. Het spelen van een spelletje wordt vaak als puur amusement gezien, maar er wordt hierbij gebruik gemaakt van serieuze wetenschap zoals waarschijnlijkheidstheorie en speltheorie. Aan de hand van analyses van de resultaten die de AI's behalen bij het spelen van deze spellen, kunnen de AI's aangepast en verbeterd worden. De inzichten die door de game-industrie op deze manier verworven worden, kunnen daarna gebruikt worden in andere takken van de kunstmatige intelligentie, zoals bijvoorbeeld de robotica.

Risk is een typisch oorlogs-bordspel waarbij diplomatie, conflicten en veroveren centraal staan. Een goede strategie en tactiek zijn in Risk de belangrijkste elementen, naast geluk om je tegenstander te kunnen verslaan. De prestaties van een AI bij het spelen van een spelletje Risk kunnen dus inzicht geven in het niveau van de AI op het gebied van strategisch en tactisch 'denken'. Er zijn in de afgelopen jaren verschillende technieken uitgetoetst, zoals heuristieken, netwerken en aanvalsheuristieken, multi-agent systems en temporal difference learning om een AI te ontwerpen die Risk kan spelen. Meestal is het uitgangspunt van die onderzoeken het maken van een systeem dat zo goed mogelijk presteert en dus altijd het spel wil winnen. Ik ga deze verschillende onderzoeken vergelijken om er achter te komen welke manier het beste werkt. Mijn onderzoeksvraag luidt dan ook:

Wat is de beste aanpak voor het ontwikkelen van een kunstmatige intelligentie voor het spel Risk?

Om de onderzoeken die ik ga bekijken verder te kunnen analyseren en om te bepalen wat de beste aanpak is voor het ontwikkelen van een AI voor het spel Risk, heb ik ook nog twee deelvragen geformuleerd. Aan de hand van deze deelvragen wil ik bepalen welke aanpak het beste resultaat geeft door onder andere te kijken naar hoe vaak Risk gewonnen wordt door de AI en ook wil ik bepalen op welke manier de AI gebruik maakt van tactieken, die menselijke spelers kunnen helpen om Risk te winnen.

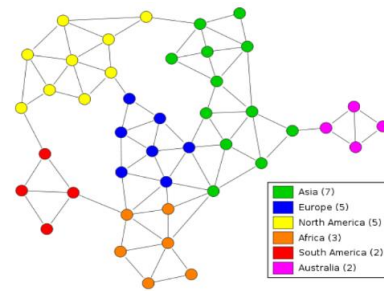
- a. *Welke aanpak voor het ontwikkelen van een kunstmatige intelligentie voor het spel Risk geeft de beste resultaten?*
- b. *Van welke tactieken maken de verschillende aanpakken voor het ontwikkelen van een kunstmatige intelligentie voor het spel Risk gebruik?*

In deze scriptie zal ik eerst uitleggen wat de spelregels van Risk zijn en vervolgens zal ik bepalen wat de beste strategie is voor het spelen van het spel. Daarna zal ik kansberekeningen bespreken die van belang zijn voor Risk. Deze zijn de basis van veel van de onderzoeken die ik vervolgens zal behandelen. Elk onderzoek gebruikt een andere aanpak om een AI te bouwen die zelfstandig Risk kan spelen. Ik zal beginnen met een onderzoek naar heuristieken en daarna zal ik de onderzoeken naar netwerken en aanvalsheuristieken, multi-agent systems en temporal difference learning bespreken. Aan het einde vergelijk ik deze onderzoeken met elkaar, om op deze manier een antwoord te kunnen formuleren op mijn onderzoeksvraag en de daarbij behorende deelvragen.

2. Spelregels Risk

Het doel van het spel is om de wereld te veroveren door het totaal van 42 landen op het speelbord in te nemen en dus al je tegenstanders te verdrijven. Het speelbord van Risk bestaat uit zes verschillende continenten. Elk continent bestaat vervolgens weer uit een verschillend aantal landen. Zie figuur 1 voor een schematische voorstelling van het speelveld en de connecties tussen de landen (Hahn, 2010).

In de 'setup'-fase, voordat het spel begint, plaatsen de spelers om de beurt één leger op een leeg land en worden op deze manier de eigenaar van dit land. Dit gebeurt totdat er geen lege landen meer over zijn. Daarna mogen de spelers hun overige troepen verdelen over de landen die ze bezitten. De hoeveelheid troepen die een speler mag plaatsen aan het begin van het spel is afhankelijk van de hoeveelheid spelers. Hierna begint het spel en komt elke speler in een vaste volgorde aan de beurt. Een beurt van een speler bestaat uit de volgende vier onderdelen, die altijd in onderstaande volgorde moeten worden uitgevoerd (UltraBoardGames, n.d.):



Figuur 1: planaire grafiekweergave van de verbindingen tussen de landen van het Risk-speelbord. Elke node representeert één land. De kleuren geven aan tot welk continent elk land behoort. Continent-bonussen worden aangegeven in de legenda tussen haakjes.

1. Het inruilen van verworven kaarten

Aan het einde van elke ronde waarin een speler minstens één land heeft veroverd, verdient hij één Risk-kaart. Als een speler drie matchende kaarten heeft verworven mag hij kiezen of hij deze wil inruilen voor extra troepen. De eerste speler die kaarten inruilt ontvangt 4 extra troepen, de volgende 6 en het aantal troepen stijgt verder elke keer als er kaarten worden ingeruild. Als een speler 5 kaarten heeft moet hij deze inruilen.

2. Plaatsen van troepen

De speler moet beslissen in welke van zijn eigen landen hij troepen wilt plaatsen. Een speler krijgt minimaal drie troepen per beurt erbij, onafhankelijk van de extra troepen uit fase 1 van zijn beurt. Verder krijgt een speler voor elke drie landen die hij bezit nog één leger. Het bezitten van een geheel continent levert ook een bonus op. Niet elk continent levert overigens hetzelfde aantal troepen op. In de legenda van figuur 1 is te zien welke bonus er per continent kan worden verdiend en dus hoeveel troepen een speler krijgt per continent.

3. Aanvallen

Zodra alle troepen geplaatst zijn, mag de speler proberen om zijn buurlanden te veroveren. Dit gebeurt aan de hand van het gooien met dobbelstenen. Als de aanvaller op het land waar hij mee aanvalt twee troepen heeft staan, mag hij één dobbelsteen gebruiken, twee als hij er drie heeft en drie dobbelstenen als hij nog meer troepen op het aanvallende land heeft staan. De verdediger mag één dobbelsteen gebruiken als er maar één leger op het land staat dat wordt aangevallen en anders twee. De hoogste worp van de aanvaller wordt vergeleken met de hoogste worp van de verdediger en (als ze beide met minstens twee dobbelstenen gooien) wordt ook de volgende hoogste worp met elkaar vergeleken. Als het resultaat van de verdediger gelijk of hoger is, dan verliest de aanvaller een leger, anders de verdediger. Dit betekent dat, als beide spelers ten minste twee dobbelstenen hebben, de aanvaller twee troepen verliest, dat de verdediger twee troepen verliest of dat ze elk een leger verliezen. Op het moment dat de verdediger geen troepen meer over heeft, heeft de aanvaller het land veroverd en verplaatst hij een gedeelte van de troepen naar het veroverde land. Als er in plaats daarvan bij de

aanvaller maar één leger overblijft, kan hij niet langer aanvallen. Een speler mag zo vaak aanvallen als hij wil, zolang hij maar genoeg troepen heeft.

4. Het versterken van de positie

Aan het einde van elke beurt kan de speler er nog voor kiezen om troepen te verplaatsen van één land dat hij bezit naar een ander aangrenzend land dat hij bezit.

Er bestaan verschillende variaties van Risk, waarbij de spelregels licht afwijken van de standaard regels. Risk wordt bijvoorbeeld ook gespeeld met een andere vormgeving van het speelbord, in de 'setup'-fase worden landen willekeurig toegewezen aan spelers in plaats van gekozen door de spelers zelf of het aantal extra troepen die spelers krijgen bij het inleveren van Risk-kaarten verschilt (Lütolf, 2013). Als er bij een onderzoek dat ik behandel wordt afgeweken van de standaardregels zoals deze hierboven beschreven staan, wordt dit expliciet vermeld.

3. De beste strategie voor het spelen van Risk

Alle geraadpleegde literatuur is het er over eens dat ‘de beste strategie’ om elk spelletje Risk te kunnen winnen simpelweg niet bestaat. Elke tegenstander is anders en ook je tegenstanders zelf hebben vaak geen vaste strategie om te winnen. Daarom is het van cruciaal belang om je tegenstander zo snel en zo goed mogelijk in te kunnen schatten, om op deze manier je eigen strategie daar op aan te passen. Ook al bestaat er dus niet zoiets als ‘de beste strategie’ voor Risk, toch zijn er een aantal tactieken die je kunt volgen, die de kans op een overwinning een stuk groter maken. Deze tips staan uitgebreid uitgelegd op de site ultraboardgames.com. Deze site verstrekt uitgebreide informatie over verschillende bordspellen. Zoals ze zelf zeggen: “This site is dedicated to promoting exciting board games. Through extensive research, we bring everything you need to know about board games. Our mission is to produce engaging articles like reviews, tips and tricks, game rules, strategies, etc.” Ook een AI kan gebruik maken van deze tactieken, die menselijke spelers kunnen helpen om te winnen, om zijn eigen kans op een overwinning te vergroten. Om te bepalen welke aanpak voor het ontwikkelen van een AI voor Risk het beste is, zal ik dus ook kijken naar hoe de verschillende onderzoeken omgaan met deze tactieken.

Verras je tegenstander

Maak een onverwachte zet om je tegenstander te verrassen en in verwarring te brengen. Hij zal niet begrijpen wat je uiteindelijke plan is, waardoor hij altijd een stap achter ligt en zal moeten reageren op jou zetten.

Beheers continenten

Het bezitten van een geheel continent versterkt je positie. Probeer in het begin van het spel kleine continenten te veroveren die weinig connecties naar andere continenten hebben. Het nadeel is wel dat bijna elke speler deze landen zal willen veroveren. Als je een ander continent kiest om te veroveren kun je zo je tegenstanders verrassen, maar ga in het begin nooit voor Azië. Dat continent is te groot en heeft te veel banden met andere continenten. Hierdoor zul je onnodig veel troepen verliezen. Het is wel altijd een goed idee om in het begin van het spel te gaan voor het continent waarvan je de meeste landen bezit.

Speel defensief

Het is belangrijk om genoeg troepen aan de rand van je territorium te plaatsen zodat andere spelers je landen niet kunnen overnemen. Houd er echter rekening mee dat je je troepen niet te veel wilt verdelen en uitdunnen. Om deze reden kan het verstandig zijn om niet te veel landen in een keer aan te vallen. In de landen, die omringd zijn door landen die je ook zelf in je bezit hebt, heeft het weinig zin om veel troepen te plaatsen. Daarnaast is het beter om met een grote groep troepen aan te vallen dan een leger van een gemiddelde grootte, omdat de kans op een overwinning daardoor groter is (Robinson, 2009). Het is hierbij wel belangrijk om het verschil in het aantal troepen van jezelf en je tegenstander in de gaten te houden. Zolang je zeker bent van een overwinning kun je de rest van je troepen sparen om je land te verdedigen.

De balans van de macht

Diegene die de meeste landen en/of troepen in zijn bezit heeft moet er voor oppassen dat de zwakkere spelers zich tegen hem zullen keren. Als een speler meer dan de helft van alle landen in zijn bezit heeft hebben de andere spelers weinig keus en zullen ze moeten samenwerken om de sterkste speler te verzwakken, om het spel zelf nog te kunnen winnen (UltraBoardGames, n.d.).

3.1 Kansberekening

Bij het spelen van Risk hebben het gooien met dobbelstenen en de kansberekeningen die daar bij horen een zeer belangrijke rol. Het is makkelijk in te zien dat bij het gebruiken van meer troepen de kans op veroveren van een land groter wordt en de kans op het verliezen van troepen kleiner. Om van één aanval deze kans uit te rekenen is niet zo moeilijk, maar om de kans van een serie aanvallen om een land te kunnen veroveren, uit te rekenen is al een stuk lastiger. In 1997 hield Tan zich hier als eerste mee bezig. Hij probeerde antwoord te geven op de vragen: Als je een land aanvalt, wat is dan de kans dat je dit land zult veroveren? En als je aanvalt, hoeveel troepen verwacht je dan te verliezen afhankelijk van het aantal troepen die op het land van je tegenstander staan (Tan, 1997)? Hij behandelde de dobbelstenen echter onafhankelijk, maar Osborne zag in dat de resultaten van de verschillende worpen gerangschikt worden en daarom niet meer onafhankelijk zijn. Zijn resultaten zijn te zien in figuur 2. Hierin staat een tabel van overwinningskansen voor een gevecht tussen een aanvaller met A-troepen en een verdediger met D-troepen. Als een speler aanvalt met 3 troepen en de verdediger 2 troepen heeft, is zijn kans om het land te veroveren 65,6%. We gaan er vanuit dat de aanvaller zijn aanval uitvoert, ongeacht of hij veel van zijn eenheden verliest bij de eerste worp. De lijn in figuur 2 geeft het keerpunt aan, waarbij de aanvaller een hogere kans heeft om de strijd te winnen dan de verdediger. Uit de tabel kunnen we aflezen dat als de aanvaller één leger meer heeft dan de verdediger zijn kans op winnen meer dan 50% is (Osborne, 2003).

A \ D	1	2	3	4	5	6
1	0.417	0.106	0.027	0.007	0.002	0.000
2	0.754	0.363	0.206	0.091	0.049	0.021
3	0.916	0.656	0.470	0.315	0.206	0.134
4	0.972	0.785	0.642	0.477	0.359	0.253
5	0.990	0.890	0.769	0.638	0.506	0.397
6	0.997	0.934	0.857	0.745	0.638	0.521
7	0.999	0.967	0.910	0.834	0.736	0.640
8	1.000	0.980	0.947	0.888	0.818	0.730
9	1.000	0.990	0.967	0.930	0.873	0.808
10	1.000	0.994	0.981	0.954	0.916	0.860

Figuur 2: kans dat een aanvaller met A-troepen de strijd wint tegen een verdediger met D-troepen, waarbij de lijn in de tabel het keerpunt aangeeft waarop de aanvaller een grotere kans heeft om de strijd te winnen dan de verdediger.

4. Verschillende aanpakken voor het ontwerpen van een AI die zelfstandig Risk kan spelen

4.1 Heuristieken

Een zoekalgoritme is een reeks instructies dat een zoekprobleem oplost door informatie op te halen uit een bepaalde datastructuur. Om een zoekalgoritme sneller en doeltreffender te maken, kan men heuristieken gebruiken. Heuristieken zijn functies die een schatting geven over de kosten van de verschillende paden die een zoekalgoritme kan verkennen. Een heuristiek kan door de ontwerper van een algoritme zelf gekozen worden en het is dan de vraag welke heuristiek het beste is (Russel & Norvig, 2018).

Franz Hahn heeft onderzoek gedaan naar het gebruik van verschillende heuristieken voor het spelen van Risk (Hahn, 2010). Bij de implementatie van Risk in dit onderzoek door Hahn is er voor gekozen om een paar versimpelde aannames door te voeren. De eerste is dat er geen Risk-kaarten in het spel zijn waardoor je extra troepen kunt bemachtigen. De tweede is dat de AI alleen de keuze krijgt met hoeveel troepen hij wil aanvallen. De AI kan er niet tussentijds voor kiezen om met de aanval te stoppen. Het resultaat wordt daarna bepaald door de tabel van Osborne. De derde versimpelde aanname is dat de AI fase 4 van zijn beurt zo vaak mag uitvoeren als hij wil, terwijl in de standaard regels staat dat dat maar één keer mag. Andere variaties op de standaardregels die zijn doorgevoerd, zijn dat de landen bij de start van het spel willekeurig aan de spelers worden toegewezen en het spel maar met twee spelers gespeeld kan worden.

Zoals we hebben gezien in 'Spelregels Risk' is de beurt van een speler verdeeld in vier fases. Elk fase heeft zijn eigen kenmerken en in elke fase wordt een andere soort actie uitgevoerd. Voor deze verschillende fases zijn dan ook verschillende strategieën van belang. Om deze reden zijn er voor elk fase andere heuristieken gekozen. De meest basale heuristiek is *Random*. Deze heuristiek wordt in elke fase gebruikt en geldt als nulmeting ten opzichte van de andere heuristieken. De heuristieken worden *Random Supply*, *Random Attack* en *Random Reinforce* genoemd. Voor het plaatsen van troepen in fase 2 van een beurt zijn de volgende twee heuristieken ontworpen, genaamd *Border Supply* en *Border Security Ratio Supply*. *Border Supply* verdeelt de troepen willekeurig over alle landen aan de rand van een territorium. *Border Security Ratio (BSR)* bepaalt de ratio tussen het aantal troepen in je eigen 'grens'-land en vijandelijke buurlanden. Een hoge BSR geeft aan dat een land meer kans heeft om veroverd te kunnen worden. Met *BSR Supply* plaats je juist op deze landen troepen en versterk je je verdediging. De heuristieken die ontworpen zijn voor fase 3 *aanvallen* heten *Full Force Attack* en *High Chance Attack*. Bij het gebruik van de heuristiek *Full Force Attack* wordt het maximum aantal troepen gebruikt om aan te vallen en bij *High Chance Attack* wordt er aangevallen met één leger meer dan de verdediger heeft. Als bij *High Chance Attack* niet wordt voldaan aan de voorwaarde dat de aanvaller één leger meer heeft dan de verdediger, dan wordt er niet aangevallen. Voor het versterken van de positie in de laatste fase van een beurt is er één heuristiek ontworpen, genaamd *Border Reinforce*. Deze heuristiek werkt ook met BSR en zorgt ervoor dat er meer troepen geplaatst worden op de landen met een hoge BSR.

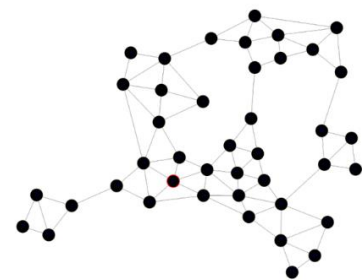
Om de heuristieken die Hahn heeft onderzocht met elkaar te vergelijken heeft hij elke mogelijk combinatie van heuristieken tegen elkaar laten spelen. Uit de resultaten blijkt dat de heuristiek *High Chance Attack* altijd het beste presteert en deze ook de algemene prestatie van de AI het sterkste beïnvloedt. Dit is ook logisch, omdat het doel van het spel immers is om landen aan te vallen en te

veroveren. Verder werkt *BSR Supply* het beste voor fase 2 van een beurt en maakt het niet uit of je *Random* of *Border Reinforce* gebruikt in fase 4. Met deze set van heuristieken wordt er 70% van de spellen gewonnen, als de AI speelt tegen een AI met dezelfde set van heuristieken (Hahn, 2010).

De sterkste combinatie van heuristieken in bovengenoemd onderzoek naar heuristieken is dus *BSR Supply*, *High Change Attack* en *Random of Border Reinforce*. Als we kijken naar de vier tips die beschreven staan in de paragraaf 'De beste strategie voor het spelen van Risk' kunnen we concluderen dat deze combinatie van heuristieken weinig rekening houdt met de tip *verras je tegenstander*. Er wordt geen gebruik gemaakt van willekeurige elementen. Behalve misschien bij het gebruik van *Random Border*, maar daar zal een tegenstander weinig van in de war raken. Ook de tactiek *beheers continenten* komt niet terug in dit onderzoek naar heuristieken. Er wordt nergens een heuristiek gebruikt die er voor zorgt dat de AI het doel nastreeft om een heel continent in zijn bezit te krijgen. *De balans van de macht* zien we in deze spelletjes Risk ook niet terug. De tactiek *speel defensief* komt wel sterk naar voren in deze aanpak. Deze kunnen we terugvinden in *BSR Supply*, waar er geprobeerd wordt zo effectief mogelijk om te gaan met het plaatsen en verdelen van troepen. Ook komt de tactiek *speel defensief* terug in de heuristiek *High Change Attack*. Deze zorgt er namelijk voor dat niet alle troepen worden ingezet bij een aanval en er daarom troepen worden overgehouden om te verdedigen. Het gedrag van de AI uit dit onderzoek naar het gebruik van heuristieken door een AI om Risk te spelen is niet weergegeven bij de resultaten. Het is daarom onmogelijk om te kijken of er uit het gedrag van deze AI wel patronen naar voren komen die lijken op de tactieken uit paragraaf 3.

4.2 Netwerken en aanvalsheuristieken

Een andere aanpak voor het implementeren van een AI voor het spelen van Risk is om gebruik te maken van netwerken. Het speelbord van Risk wordt dan gerepresenteerd door een graaf met 42 knopen (de landen) en 81 kanten (de connecties tussen de landen), zoals te zien is in figuur 3. Alle landen krijgen een kleur, afhankelijk van de eigenaar van dat land. Als twee landen naast elkaar dezelfde kleur krijgen en dus tot dezelfde speler behoren, is de kant tussen die landen niet van belang en wordt hij verwijderd. Zodra landen veroverd worden en twee verschillende kleuren krijgen, worden deze kanten weer toegevoegd. Het aantal troepen dat een land bezit, wordt opgeslagen in de knopen.



Figuur 3: representatie van het Risk speelbord in een graaf. De graaf bevat 42 Knopen en 81 kanten.

Ik heb voor deze aanpak gekeken naar het onderzoek van Jacob Munson (Munson, 2017). Bij de implementatie van het netwerk zijn een aantal versimpelde aannames gemaakt. Er is voor gekozen om in het begin de landen willekeurig toe te wijzen aan de spelers, er wordt geen gebruik gemaakt van Risk-kaarten en het spel kan maar met maximaal twee spelers gespeeld worden. Verder wordt fase 4 in dit onderzoek achterwege gelaten, dus spelers mogen na het aanvallen geen troepen meer verplaatsen.

In fase 2 van een beurt mogen de spelers hun troepen verdelen en dat gebeurt in dit onderzoek door de troepen uniform te verdelen tussen de landen die kanten hebben naar andere landen. Dit zijn dus automatisch vijandelijke landen. Landen die omringd worden door eigen landen ontvangen geen nieuwe troepen. Er worden in dit onderzoek verschillende heuristieken onderzocht om te bepalen wanneer een aanval wordt uitgevoerd in fase 3 van een beurt. Hier kom ik later op terug. Het resultaat van een aanval wordt vervolgens bepaald met de tabel van Osborne.

Om de beste strategie te bepalen voor de aanvalfase zijn er drie heuristieken onderzocht, de *naïeve strategie*, de *ratio strategie* en de *pricewise ratio strategie*. De naïeve strategie valt altijd zo veel mogelijk landen aan met zo veel mogelijk troepen, terwijl de ratio strategie meer bescheiden te werk gaat. Een land wordt aangevallen als er aan het criterium $A \geq X * D$ wordt voldaan, anders niet. Met A het aantal aanvallende troepen, D het aantal verdedigende troepen en X een getal tussen 0 en 20. Het land wordt dan aangevallen met één leger meer dan het verdedigende land bezit. De pricewise ratio strategie is een combinatie van de naïeve en ratio strategie. Als het aantal troepen laag is van het aanvallende land wordt een land pas aangevallen als aan de vergelijking van de ratio strategie wordt voldaan, maar als een aanvallend land veel troepen bezit worden zo veel mogelijk landen aangevallen met zo veel mogelijk troepen.

De naïeve strategie is gebruikt als referentiekader om de heuristieken mee te kunnen vergelijken. Als er twee spelers die beide de naïeve strategie gebruiken tegen elkaar spelen hebben beide spelers evenveel kans om te winnen. Als de waarde voor X bij de ratio strategie 0 is, worden dezelfde resultaten gevonden als bij de naïeve strategie. Wanneer deze waarde groter is, een waarde tussen de 1 en 5, stijgt de overwinningskans echter drastisch. Bij een waarde tussen 1,5 en 3 is de overwinningskans van de ratio strategie ten opzichte van de naïeve strategie 80%. Als de waarde voor X nog groter is, daalt de overwinningskans weer. Een zelfde soort curve is ook te zien in de resultaten van de pricewise ratio strategie. Waarbij deze strategie ook bij een waarde tussen 1,5 en 3 een overwinningskans heeft van 80%, maar voor de hogere waardes van X beter presteert dan de ratio strategie (Munson, 2017).

Er wordt in dit onderzoek alleen gebruik gemaakt van de tip *speel defensief* uit paragraaf 3. En dan alleen bij de heuristieken *ratio strategie* en *pricewise ratio strategie*. Bij deze heuristieken wordt er verstandiger met de troepen omgegaan en worden niet alle troepen in één keer ingezet om zo veel mogelijk landen te veroveren, wat kan leiden tot het onnodig verliezen van troepen. Uit het onderzoek blijkt dan ook, dat deze heuristieken het beste resultaat geven. Van de tips *verras je tegenstander*, *beheers continenten* en *de balans van de macht* wordt in deze aanpak geen enkel gebruik gemaakt. Ook in dit onderzoek van Munson is er in de resultaten niks te vinden over het gedrag dat de AI vertoont tijdens het spelen van Risk. Er kan dus niet bekeken worden of de tactieken die ik besproken heb in paragraaf 3 wel emergeeren uit het gedrag van de AI.

4.3 Multi-agent systems

We kunnen een agent zien als iets dat zijn omgeving kan waarnemen door middel van sensoren en op die omgeving kan reageren (Russell & Norvig, 2018). Om een multi-agent system te kunnen ontwerpen zijn er drie factoren van belang, de verdeling van de agents, de communicatie tussen de agents en het onderhandelen tussen de agents. Stefan Johansson en Fredrik Olsson hebben een AI ontworpen die Risk kan spelen door gebruik te maken van een multi-agent system (Johansson & Olsson 2006). In dit systeem is er voor gekozen om elk land zijn eigen agent toe te kennen. De agents behoren toe aan de AI zelf en zijn ook alleen toegankelijk voor de AI zelf. Toch krijgen ook de landen van de tegenspelers een agent toegewezen, zodat kan worden bepaald door de AI hoeveel troepen zijn tegenstander op de vijandelijke landen geplaatst heeft. Daarnaast is er nog een agent die verantwoordelijk is voor de Risk-kaarten en een agent die communiceert met de game server en de onderhandelingen faciliteert. Deze communicatie-agent vraagt bijvoorbeeld 'Welk land heeft nieuwe troepen het hardst nodig?' Elke agent van de aangesproken landen reageert op deze vraag en via een veilingssysteem worden de troepen gegeven aan 'de hoogste bidder'.

Johansson en Olsson hebben ervoor gekozen om elk land een waarde toe te kennen. Sommige landen zijn waardevoller om te bezitten dan anderen. Dit gegeven kan tijdens het spelen van het spel veranderen. De waarde van een land wordt dan ook bepaald door twee factoren. Een vaststaande waarde die afhankelijk is van het continent waar het land toebehoort en een dynamische waarde die wordt bepaald door onder andere de hoeveelheid troepen die er op dat land staan, hoeveel van de andere landen op hetzelfde continent je in je bezit hebt en hoeveel vijandelijke burens het land heeft. De enige versimpelde aanname die gebruikt is bij de implementatie van Risk in het onderzoek van Johansson en Olsson is om de landen in het begin van het spel willekeurig toe te wijzen aan de spelers.

In fase 2 van een beurt worden de troepen verdeeld aan de hand van een veilingssysteem. Alle agents kijken hoeveel troepen zijn land bezit en hoeveel troepen er op vijandige buurlanden staan. Hierna worden alle landen die de AI in zijn bezit heeft met elkaar vergeleken. Dit gebeurt door middel van een heuristiek die nagaat welk land de grootste win-kans heeft bij een aanval; het land met de grootste waarde krijgt de extra troepen. Vervolgens wordt in fase 3 gevraagd aan alle agents om nogmaals te kijken hoeveel troepen hij bezit en hoeveel troepen vijandige buurlanden bezitten. Alle landen die in eigen bezit zijn, een vijandig buurland hebben en meer dan één leger bezitten mogen meebieden. Het land dat de hoogste kans van veroveren heeft mag dan de aanval uitvoeren. Bij deze heuristiek speelt ook de waarde van het te veroveren land een grote rol. Het resultaat van een aanval wordt bepaald door de tabel van Osborne. Het is mogelijk dat geen enkel land aan de voorwaarden voldoet. Dan wordt er door de AI niet aangevallen. De verdeling van de troepen in de laatste fase van een beurt gebeurt weer via het veilingssysteem en hierbij is van belang hoeveel troepen een land kan missen om te verplaatsen, of het land omringd is door eigen of vijandelijke landen en hoeveel troepen er op de omringende vijandelijke landen staan.

Het multi-agent systeem, genaamd Mars, dat hierboven beschreven staat is getest op het platform Lux dat beschikt over twaalf verschillende artificiële spelers met allemaal hun eigen kwaliteiten. Een paar voorbeelden van deze AI's zijn de spelers: Communist, Pixie, Cluster en Yakool. De speler Communist heeft als eigenschap dat hij alle troepen eerlijk verdeelt over al zijn landen. Eerst spelers aanvalt die niet van het type 'Communist' zijn en daarna pas andere 'Communist' spelers aanvalt. Pixie onderzoekt eerst alle continenten en focust zich op continenten die er veel belovend uitzien. Pixie kiest verder landen om aan te vallen die de minste buurlanden hebben. De AI Cluster probeert zijn grootste aaneengesloten verzameling van landen uit te breiden en Yakool's enige doel is om tegenstanders die te sterk lijken, compleet uit te schakelen (Lütolf, 2013). Tussen de in totaal dertien spelers is een toernooi gehouden waarbij elke AI tegen elke andere AI speelt. Per spel spelen er zes AI's. Mars was de uiteindelijke winnaar en heeft de meeste wedstrijden gewonnen ten opzichte van de andere AI's, 338 van de in totaal 1716 gespeelde games (Johansson & Olsson 2006).

In dit onderzoek van Johansson en Olsson naar multi-agent systems wordt er geen gebruik gemaakt van een tactiek die er voor zorgt dat Mars zijn tegenstanders kan verrassen. Ook het samenwerken met andere spelers om de sterkste speler gezamenlijk te kunnen overwinnen is niet terug te vinden in dit onderzoek. De tips om continenten te beheersen en defensief te spelen komen wel duidelijk in het ontwerp van Mars naar voren. Het verlangen om hele continenten te beheersen in plaats van losstaande landen komt terug in de waardes die aan elk land afzonderlijk worden toegekend en ook in de keuze om een land aan te vallen. Dit zit namelijk verwerkt in de beoordeling van de 'hoogte bieder' bij het veilingssysteem. In dit systeem is ook het slim verdelen van de troepen aan het begin en het einde van een beurt vertegenwoordigd. In de resultaten van het onderzoek Johansson en Olsson is niets opgenomen over het precieze gedrag dat de AI vertoont tijdens het spelen van Risk. Het is daarom

niet duidelijk of uit het gedrag van de AI wel de tactieken emergeeren zonder dat deze expliciet zijn meegegeven bij de implementatie.

4.4 Temporal difference learning

Om een AI te bekijken die gebruik maakt van temporal difference learning heb ik gekeken naar een onderzoek van Manuela Lütolf (Lütolf, 2013). De AI in dit onderzoek is ontworpen om 'blind' te kijken naar alle mogelijke acties die hij kan ondernemen. Alle mogelijke uitkomsten worden vergeleken per keuzemoment en vervolgens wordt de actie met het beste resultaat gekozen. Er wordt bepaald wat het beste resultaat is aan de hand van een evaluatiefunctie. De evaluatiefunctie bestaat uit een aantal kenmerken. Elk kenmerk is gekoppeld aan een gewicht en die gewichten worden steeds aangepast door het temporal difference algoritme. Het onderzoek van Lütolf maakt bij de implementatie van Risk geen gebruik van versimpelde aannames.

Het temporal difference algoritme werkt, bij elk beslissingsmoment, de gewichten die gekoppeld zijn aan de kenmerken bij. Hierdoor is de AI in staat om te leren welke kenmerken het belangrijkste zijn en het zwaarst moeten meetellen. In het onderzoek van Lütolf slaat de AI bij elk keuze moment op welke beslissing er gemaakt is en achteraf wordt gekeken of de keuze een voordeel opleverde of niet. De keuzemomenten zijn alle fasen van een beurt, maar pas aan het einde van een hele game worden de gewichten bijgewerkt. De kenmerken die gebruikt worden in de evaluatie functie zijn:

- Totaal aantal landen in bezit
- Aantal startende legers
- Aantal unieke vijandelijke burens
- Paren van vriendelijke burens
- Aantal landen van een specifiek continent (aan het begin van het spel, of later)
- Aantal landen in het grootste cluster
- Aantal troepen in het grootste cluster
- Aantal troepen in je eigen grenslanden
- Totale troepen, het totale aantal legers dat je bezit, gedeeld door het totale aantal legers op het speelbord
- Het aantal kaarten

Het nadeel van het werken met kenmerken is dat de AI geen idee heeft dat het aanvallen van een land kan resulteren in het bezitten van het land in de volgende ronde. Voor de aanvalfase wordt er in dit onderzoek naar een AI die gebruik maakt van temporal difference learning daarom gebruik gemaakt van 'vooruit kijken'. De AI wordt toegestaan om zijn aanvallen ver genoeg te plannen om een eindresultaat te bereiken en aan de hand daarvan beslissingen te maken over het aanvallen. Er wordt dan bepaald of er wordt aangevallen en welk land of welke landen worden aangevallen.

De AI die is ontworpen in dit onderzoek is getest op het platform Domination. Dit is een open-source implementatie van Risk. Op dit platform is ook een artificiële speler gedefinieerd. Deze speler heeft de eigenschap om er alles aan te doen om te voorkomen dat een tegenstander een heel continent veroverd. Ook zijn de uitslagen van de dobbelstenen, bij het aanvallen en verdedigen, vaker in het voordeel van de AI van het Domination platform dan van zijn tegenspelers. Helaas is nergens duidelijk op het platform gespecificeerd hoe dit precies in zijn werk gaat. Met de artificiële speler van Domination is de AI van het onderzoek van Lütolf getraind. De waardes voor de gewichten in het begin van het leerproces zijn naar eigen inzicht door de onderzoekster gekozen, zonder feitelijke

onderbouwning. Voordat de AI getest is en de gewichten zijn aangepast wint de AI 3% van de games. Na het trainen van de AI is dit resultaat niet verbeterd. Wel duren de spellen langer. Dit wil zeggen dat de AI voor een langere periode landen in zijn bezit heeft. Hierdoor kunnen we zeggen dat de AI toch beter is na het leerproces dan daarvoor.

In het onderzoek van Lütolf is naast het aantal overwinningen van de AI ook gekeken naar het gedrag dat er ruwweg wordt vertoont na de training door middel van temporal difference learning. Bij de start van het spel worden er landen gekozen die verspreid liggen. Vervolgens zal de AI proberen een geheel continent te veroveren. Troepen worden zo veel mogelijk geplaatst in landen die in de buurt liggen van een zo groot mogelijk vijandig gebied. Dit hindert de tegenspeler wel om een continent te bemachtigen, maar vaak worden er te weinig troepen gebruikt om zelf landen te veroveren en zal de tegenspeler alsnog het continent veroveren en op dezelfde manier het spel winnen (Lütolf, 2013).

Doordat we het gedrag van de AI tijdens het spelen van Riks kunnen analyseren, kunnen we opmaken dat de AI gebruik maakt van de tip *de balans van de macht* en deze zelfs te sterk doorvoert. Er wordt door de AI dan wel niet samengewerkt met andere spelers om de sterkste speler te kunnen verslaan, maar wel elke zet die de AI doet berust er op om het de tegenstander zo moeilijk mogelijk te maken om grote clusters te kunnen vormen en hele continenten te veroveren. Dit zien we bijvoorbeeld terug in het kiezen van landen aan het begin van een game. Door landen verspreid te kiezen kan een tegenstander geen heel continent bemachtigen, maar de AI zelf ook niet. Door de AI zelf wordt dus niet genoeg geprobeerd om een continenten te beheersen, al heeft de onderzoekster dit wel proberen mee te geven door er een apart kenmerk aan te wijden. Ook komt er weinig terecht van defensief spelen. De AI valt dan wel niet alle landen aan met zo veel mogelijk troepen, maar verdedigt zijn grote groepen van landen ook niet, door daar bijvoorbeeld veel troepen aan de grenzen te plaatsen. De vierde tip *verras je tegenstander* wordt verder nergens gebruikt in het onderzoek.

5. Onderzoeken naar het ontwerpen van een AI die Risk kan spelen vergelijken

5.1 Resultaten van de onderzoeken naar het ontwerpen van een AI die Risk kan spelen

De vier aanpakken die ik besproken heb in paragraaf 4, maken allemaal gebruik van een andere manier om de door hun ontworpen AI te evalueren. De AI die getest is op verschillende heuristieken is geëvalueerd door alle verschillende combinaties van heuristieken tegen elkaar te laten spelen en de AI die gebruik maakt van een netwerk om Risk te implementeren gebruikt zijn zelfontworpen naïeve strategie als referentiekader. De onderzoeken die ik besproken heb in paragraaf 4.3 en 4.4 maken daar in tegen gebruik van externe AI's. Het onderzoek naar multi-agent systems laat zijn AI het tegen twaalf tegenspelers van het platform Lux opnemen om hem te evalueren en het onderzoek dat gebruik maakt van temporal difference learning evalueert de ontworpen AI door hem tegen een tegenstander te laten spelen van het platform Domination.

Net zoals de manieren waarop de AI's geëvalueerd zijn, zijn ook de resultaten die de verschillende aanpakken naar het ontwerpen van AI's die een spel Risk zelfstandig kunnen spelen erg verschillend. Zo wint de AI uit het onderzoek van Lütolf slechts drie van de honderd spelletjes Risk die hij speelt, zelfs na training. De AI slaagt er wel in om gemiddeld meer speelrondes te overleven na zijn training waardoor we kunnen zien dat hij dus toch vooruitgang heeft geboekt. De AI 'Mars' wint net geen 20% van de spellen die hij heeft gespeeld, maar is daarmee wel de beste speler. Mars wint verreweg meer spellen dan zijn twaalf tegenstanders. De AI uit het onderzoek van Hahn wint 70% van de games als hij gebruik maakt van de heuristieken *BSR Supply*, *High Change Attack* en *Random of Border Reinforce* en tegen een AI speelt met dezelfde heuristieken. Het vaakst wint de AI die gebruik maakt van netwerken om Risk te representeren in combinatie met de aanvalsheuristiek *pricewise ratio strategie*. Door deze AI worden 80% van de wedstrijden gewonnen als de optimale parameters worden meegegeven aan het begin van het spel. Een overzicht van de manieren waarop de verschillende onderzoeken geëvalueerd zijn en de resultaten die ze behaald hebben is te vinden in bijlage A.

5.2 Het gebruik van de tactieken door de onderzoeken naar het ontwerpen van een AI die Risk kan spelen

In deze sub-paragraaf zal ik de onderzoeken naar het ontwerpen van een AI die Risk kan spelen vergelijken op basis van hun omgang met de tactieken die ik uitgelegd heb in paragraaf 3 'De beste strategie voor het spelen van Risk'. Menselijke spelers zijn er bij gebaat als ze gebruik maken van deze tactieken. De kans om Risk te winnen wordt voor hen hierdoor vergroot. Als een AI gebruik maakt van dezelfde tactieken, kunnen we op deze manier ook conclusies trekken over hoe goed de AI is in het spelen van Risk. In het geval dat er meerdere heuristieken zijn onderzocht bij dezelfde aanpak, ben ik in deze paragraaf alleen geïnteresseerd in de heuristiek met de beste resultaten. In bijlage A is een overzicht te vinden waarin vermeld staat per onderzoek welke tactieken ze gebruiken.

De tip *verras je tegenstander* komt nauwelijks voor in alle onderzoeken die ik bekeken heb. Geen enkele aanpak kiest er expliciet voor om een element in te bouwen om de tegenstander te verrassen en daar voordeel uit te halen. Ook uit het gedrag van de AI's is deze tactiek niet op te merken. Er moet wel bij vermeld worden, dat de onderzoeken die ik besproken heb in paragraaf 4 geen gebruik maken

van menselijke tegenstanders en de computerprogramma's waar tegen gespeeld wordt ook niet bezig zijn met de tactieken van de andere spelers.

Van de tactiek *de balans van de macht* wordt ook weinig gebruik gemaakt bij de implementaties van Risk en de AI's die het spel moeten spelen. Alleen bij de aanpak die gebruik maakt van het temporal difference learning algoritme is door middel van de gewichten meegegeven dat de AI moet proberen om de sterkste speler te dwarsbomen. Uit het gedrag van deze AI is zelfs op te maken dat hij deze tactiek te sterk doorvoert en daardoor te veel bezig is met er voor zorgen dat de tegenstander niet kan winnen in plaats van te proberen om zelf te winnen. Wat leidt tot het verliezen van het spel.

Ook de tip *beheers continenten* wordt door de aanpak die gebruik maakt van het temporal difference learning algoritme meegegeven aan de AI door middel van gewichten, maar door de boven genoemde tactiek *de balans van de macht*, die het spel van de AI te sterk domineert komt dit niet tot uiting in zijn gedrag. Bij de AI Mars die gebruik maakt van multi-agent systems kom de tip *beheers continenten* wel duidelijk naar voren. Alle landen krijgen bij deze implementatie een waarde en als de AI meerdere landen van hetzelfde continent bezit worden de waardes van de overige landen op hetzelfde continent alleen maar hoger. Hierdoor worden deze landen aantrekkelijker om aan te vallen.

Alle verschillende aanpakken voor het ontwerpen van een AI die ik bekeken heb maken gebruik van de tactiek *speel defensief*. Bij de AI die gebruik maakt van het temporal difference learning algoritme wordt ook deze tip meegegeven aan de AI door middel van de gewichten en ook hier komt de tip niet tot uiting in het gedrag van de AI door de tactiek *de balans van de macht*, die het spel van de AI te sterk domineert. De tip *speel defensief* bestaat eigenlijk uit twee losse delen. Het eerste deel gaat over het verdedigen van de landen die je bezit door troepen strategisch aan grenzen te plaatsen. Bij de aanpak die gebruik maakt van netwerken en aanvalsheuristieken wordt dit minimaal gedaan door wel alle troepen in landen te plaatsen die aan vijandelijke landen grenzen, maar de troepen tussen die landen wel uniform te verdelen. De AI's die gebruik maken van alleen heuristieken en multi-agent systems gaan verstandiger om met hun troepen en gebruiken beide een vergelijking om de hoeveelheid troepen die op elk land geplaatst worden te bepalen. Het tweede deel van de tactiek *speel defensief* gaat over aanvallen en geeft als tip om niet zo veel mogelijk landen met zoveel mogelijk troepen in één keer aan te vallen. Het onderzoek dat alleen heeft gekeken naar heuristieken maakt gebruik van deze tip door de tegenstander aan te vallen met één leger meer dan zijn tegenstander. Daardoor heeft de AI wel de grootste kans om het land te veroveren, maar is hij nog lang niet zeker van een overwinning. De AI van het onderzoek dat heeft gekeken naar netwerken en aanvalsheuristieken valt zijn tegenstander ook aan met één leger meer dan zijn tegenstander, als de AI weinig troepen bezit, nadat er door middel van een vergelijking is bepaald welk land aangevallen wordt. Maar zodra de AI genoeg troepen bezit wordt er geprobeerd om zo veel mogelijk landen met zoveel mogelijk troepen aan te vallen. De AI Mars maakt ook bij het kiezen van landen om aan te vallen gebruik van zijn veilingsysteem. Er wordt dan gekeken welke landen de grootste kans hebben om veroverd te kunnen worden. Dit wordt dan vervolgens gedaan met het maximale aantal beschikbare troepen.

6. Discussie

De resultaten van de verschillende onderzoeken die ik in mijn scriptie besproken heb zijn niet eenvoudig met elkaar te vergelijken. Dit heeft verschillende redenen. Een reden is dat niet alle onderzoeken gebruik maken van precies dezelfde implementatie van Risk. Veel onderzoeken hebben versimpelde aannames doorgevoerd. Het onderzoek van Lütolf heeft bijvoorbeeld in vergelijking met de rest de meest uitgebreide versie van Risk gebruikt, maar behaalde ook minder goede resultaten. Een andere reden is dat de AI's niet op een zelfde manier geëvalueerd zijn. De AI's in de verschillende onderzoeken die ik heb besproken in paragraaf 4 zijn niet tegen hetzelfde aantal spelers getest. Het onderzoek dat gebruik maakte van multi-agent systems heeft zijn AI getest in wedstrijden met vijf tegenstanders, terwijl de andere onderzoeken gebruik hebben gemaakt van één tegenstander. Het type tegenstander bij de evaluatie van elke onderzoek verschilde ook erg van elkaar. De onderzoeken uit paragraaf 4.1 en 4.2 maken gebruik van de AI zelf om te bepalen hoe goed hun aanpak presteert en het is lastig om te bepalen wat het niveau van een AI is als er alleen maar resultaten beschikbaar zijn van spellen waarbij de AI tegen zichzelf speelt. Ook bij de onderzoeken uit paragraaf 4.3 en 4.4 is het lastig te bepalen met welk niveau de AI's Risk spelen. De AI's spelen in deze onderzoeken tegen een of meerdere externe AI's, maar van deze spelers is ook niet duidelijk hoe sterk zij precies zijn in het spelen van het spel. Een overzicht van de versimpelde aannames en de manier waarop de verschillende onderzoeken geëvalueerd zijn, is te vinden in bijlage A. De laatste reden waarom het niet makkelijk is om de verschillende onderzoeken met elkaar te vergelijken is dat we alleen bij het onderzoek dat gebruik maakt van temporal difference learning het gedrag van de AI tijdens het spelen van Risk in kunnen zien. De andere onderzoekers verstrekken hiervoor niet voldoende informatie in hun onderzoeken. Hierdoor is het niet mogelijk om bij deze onderzoeken te bepalen of de tactieken die ik besproken heb in paragraaf 3 naar voren komen tijdens het spelen van Risk. Het zou namelijk kunnen voorkomen dat een onderzoeker een tactiek in de implementatie wilde meegeven aan de AI, maar dat deze tactiek niet tot uiting komt in het gedrag of juist dat de onderzoeker een tactiek niet expliciet heeft meegegeven, maar dat deze tactiek wel tot uiting komt in het gedrag van de AI.

7. Conclusie

De aanpak voor het ontwikkelen van een kunstmatige intelligentie voor het spel Risk met de beste win kans is de aanpak die gebruik maakt van een netwerk in combinatie met aanvalsheuristieken. Als de AI met de heuristiek *pricewise ratio strategie* tegen de referentie AI speelt, wint hij maar liefst 80% van de spellen. Toch kunnen we niet met zekerheid zeggen of het onderzoek dat gebruik maakt van een netwerk om Risk te implementeren ook het antwoord is op deelvraag a *‘Welke aanpak voor het ontwikkelen van een kunstmatige intelligentie voor het spel Risk geeft de beste resultaten?’* Dit komt door de verschillende manieren waarop de onderzoeken Risk geïmplementeerd en geëvalueerd hebben, zoals ik heb uitgelegd in paragraaf 6 *‘Discussie’*. Het onderzoek dat gebruik maakt van multi-agent systems heeft bijvoorbeeld, maar een win percentage van net geen 20%, maar deze AI moest wel zien te overleven tegen vijf spelers tegelijk. Terwijl de andere AI’s geëvalueerd zijn tegen één tegenstander per game. Dit gegeven kan de uitslag van een spel Risk beïnvloeden hebben. Ook is het dus niet geheel duidelijk welk speelniveau de AI’s hebben waarmee de ontworpen AI’s in de onderzoeken zijn geëvalueerd. Het kan dus zo zijn dat de AI die gebruik maakt van een netwerk in combinatie met aanvalsheuristieken tegen een makkelijker te overwinnen tegenstander heeft gespeeld dan de AI die getraind is door middel van temporal difference learning, die maar 3% van de spellen wint. Ook hierdoor kan het zijn dat percentages van de resultaten een vertekend beeld geven.

Deel vraag b *‘Van welke tactieken maken de verschillenden aanpakken voor het ontwikkelen van een kunstmatige intelligentie voor het spel Risk gebruik?’* geeft meer inzicht in het speelniveau van de verschillende AI’s die ontworpen zijn om Risk te kunnen spelen. Als menselijke spelers gebruik maken van de tactieken die ik besproken heb in paragraaf 3 worden hun kansen om Risk te winnen vergroot. Op het moment dat een AI gebruik maakt van dezelfde tactieken, kunnen we op deze manier iets zeggen over hoe goed de AI is in het spelen van Risk. Toch is ook op deelvraag b geen eenduidig antwoord te geven. Op het eerste gezicht lijkt het erop dat het onderzoek van Lütolf, dat gebruik maakt van het temporal difference learning algoritme, bij de implementatie van de AI de meeste tactieken heeft doorgevoerd en daarom dus een sterke speler ontworpen heeft. Uit het gedrag van de AI blijkt echter dat deze tactieken helemaal niet sterk genoeg naar voren komen, onder andere hierdoor weet de AI dan ook niet zo vaak te winnen. Bij de andere drie onderzoeken is het niet mogelijk om het gedrag van de AI’s te volgen tijdens het spelen van Risk. We kunnen bij deze onderzoeken dus alleen uitgaan van de implementatie van de AI om de door hun gebruikte tactieken te evalueren. Dit kan zorgen voor een vertekening van het beeld, maar de meeste onderzoeken implementeren de tactieken op zo een manier dat er geen ruimte voor twijfel overblijft. Zoals het onderzoek dat gebruik heeft gemaakt van multi-agent systems. Bij deze aanpak is er duidelijk te zien dat de AI Mars gebruik heeft gemaakt van de tips. Door het statisch en dynamisch toekennen van waardes aan alle landen en het gebruik maken van een veilingsysteem wordt er door Mars gefocust op het veroveren van hele continenten en wordt er een wel overwogen beslissing gemaakt over hoe de troepen verdeeld worden en hoeveel landen er per beurt aangevallen worden. Mars doet dit beter dan de AI’s uit de andere onderzoeken die ik besproken heb.

Door de verschillen tussen de aanpakken die ik besproken heb in de discussie is er dus geen eenduidig antwoord te geven op de onderzoeksvraag *‘Wat is de beste aanpak voor het ontwikkelen van een kunstmatige intelligentie voor het spel Risk?’* Daarvoor zou er verder onderzoek gedaan moeten worden met een onafhankelijke tegenspeler, die zou kunnen helpen met het evalueren van de verschillende onderzoeken. In paragraaf 8 *‘Aanbevelingen’* ga ik hier verder op in. Mijn eigen voorkeur voor een aanpak van het ontwikkelen van een AI die Risk kan spelen is het onderzoek dat gebruik maakt van multi-agent systems. Een spel Risk win je als je alle landen van je tegenstander hebt veroverd. Om dit

voor elkaar te krijgen is het het belangrijkste, zo blijkt uit paragraaf 3, om niet voortijdig al je troepen kwijt te raken, waardoor je zelf makkelijk landen kwijt kunt raken. Mars maakt hierin betere keuzes ten opzichte van de andere AI's die ik besproken heb, door verstandig troepen te plaatsen en ook wat betreft het aanvallen van landen. Ook de resultaten die Mars laat zien, zijn niet slecht. In vergelijking met de AI's waartegen hij getest is, won hij ver uit de meeste wedstrijden.

8. Aanbevelingen

Om in een vervolgonderzoek makkelijker tot een antwoord te kunnen komen op de vraag ‘*Wat is de beste aanpak voor het ontwikkelen van een kunstmatige intelligentie voor het spel Risk?*’ is het mogelijk om alle AI’s van de verschillende onderzoeken die ik heb besproken te laten spelen tegen één zelfde onafhankelijke tegenstander of tegen een set van onafhankelijke tegenstanders die tijdens het onderzoek constant blijft. Van de tegenstander zou dan het speelniveau bepaald kunnen worden en aan de hand daarvan kunnen er uitspraken gedaan worden over het speelniveau van de AI die onderzocht wordt. Voor deze scriptie was dit helaas niet mogelijk, omdat er geen toegang was tot alle code van de vier verschillende onderzoeken die besproken zijn.

Het is mogelijk om de rol van de onafhankelijke tegenstander, die het makkelijker moet maken om de verschillende onderzoeken naar het ontwerpen van een AI die Risk kan spelen, in te laten vullen door een of meerdere menselijke spelers. Dat brengt echter wel een paar problemen met zich mee. Het speelniveau van menselijke spelers kan namelijk erg verschillen. Hiermee bedoel ik niet het speelniveau tussen verschillende mensen, maar het speelniveau van dezelfde speler. Dit kan afhankelijk zijn van externe factoren die moeilijk te controleren zijn. Ook zou het niet mogelijk zijn om een menselijke speler net zo vaak Risk te laten spelen als de simulaties die gebruikt zijn in de onderzoeken die ik besproken heb in paragraaf 4, omdat dat te veel tijd in beslag zou nemen. Daarnaast komt bij het gebruiken van menselijke spelers als tegenstanders voor de AI’s van de onderzoeken die ik besproken heb nog een ander fenomeen om de hoek kijken. Mensen hebben namelijk een natuurlijke neiging om eerst computers uit te schakelen en dan pas andere menselijke spelers (Johansson & Olsson, 2006). Bovengenoemde problemen kunnen opgelost worden door het speelniveau van een menselijke speler onder bepaalde omstandigheden vast te stellen en deze omstandigheden gelijk te houden tijdens de volgende spellen die de menselijke speler speelt tegen de AI. Door de menselijke spelers vervolgens tegen enkel de AI te laten spelen, kan er voorkomen worden dat er vriendelijker gespeeld wordt tegen menselijke tegenstanders dan tegen de computergestuurde AI. Dat er minder games gespeeld kunnen worden met een menselijke tegenspeler blijft een feit, maar er zullen alsnog genoeg games gespeeld kunnen worden om het niveau van de AI vast te stellen.

9. Bronnenlijst

- Hahn, F. (2010). *Evaluating Heuristics in the Game Risk An Artificial Intelligence Perspective*. Van https://project.dke.maastrichtuniversity.nl/games/files/bsc/Hahn_Bsc-paper.pdf
- Johansson, S. J., & Olsson, F. (2006). Using multi-agent system technologies in RISK bots. In *Proceedings of the Second Artificial Intelligence and Interactive Digital Entertainment Conference, (AIIDE), Marina del Rey*.
- Lütolf, M. (2013). *A Learning AI for the game Risk using the TD(λ)-Algorithm*. Van <https://ai.dmi.unibas.ch/papers/theses/luetolf-bachelor-13.pdf>
- Munson, J. M. (2017). *Neural Network Predictions of a Simulation-based Statistical and Graph Theoretic Study of the Board Game RISK*. Van <https://digitalcommons.murraystate.edu/cgi/viewcontent.cgi?article=1077&context=etd>
- Osborne, J. A. (2003). Markov chains for the RISK board game revisited. *Mathematics magazine*, 76(2), 129-135.
- Robinson, G. (2009). The Strategy of Risk. Geraadpleegd op 16 mei 2019, van <http://web.mit.edu/sp.268/www/index.html>
- Russell, S. J., & Norvig, P. (2018). *Artificial intelligence a modern approach*. Boston, Verenigde Staten: Pearson.
- Tan, B. (1997). Markov chains and the RISK board game. *Mathematics Magazine*, 70(5), 349-357.
- UltraBoardGames. (n.d.). Geraadpleegd op 28 mei 2019, van <https://www.ultraboardgames.com/risk/game-rules.php>
- Winning Strategies for Risk. (n.d.). Geraadpleegd op 16 mei 2019, van <https://www.ultraboardgames.com/risk/strategies.php>

10. Bijlage A: Overzicht verschillende onderzoeken

Onderzoek	Heuristieken	Netwerken en aanvalsheuristieken	Multi-agent systems	Temporal difference learning
Versimpelde aannames	De landen bij de start van het spel willekeurig toegewezen Geen Risk-kaarten Maximaal 2 spelers Geen limiet op uitvoeren fase 4 Kan er tussentijds niet voor kiezen om met een aanval te stoppen	De landen bij de start van het spel willekeurig toegewezen Geen Risk-kaarten Maximaal 2 spelers Geen fase 4	De landen bij de start van het spel willekeurig toegewezen	N.V.T.
Geëvalueerd met	Elke mogelijk combinatie van heuristieken speelt tegen elkaar	naïeve strategie gebruikt als referentiekader	Platform Lux, gespeeld tegen 12 artificiële spelers	Platform Domination, gespeeld tegen één artificiële spelers
Optimale heuristiek(en)	BSR Supply, High Change Attack, Random of Border Reinforce	pricewise ratio strategie	N.V.T.	N.V.T.
Resultaat	Maximaal 70% van de spellen gewonnen	Maximaal 80% van de spellen gewonnen	Maximaal 20% van de spellen gewonnen	Maximaal 3% van de spellen gewonnen
Tactieken	speel defensief	speel defensief	speel defensief beheers continenten	speel defensief beheers continenten de balans van de macht

Tabel 1: Overzicht van de versimpelde aannames die gedaan zijn, de manier waarop de AI geëvalueerd is, eventueel de optimale heuristieken, het resultaat dat de AI behaald heeft en de gebruikte tactieken per onderzoek naar het ontwerpen van een AI die zelfstandig Risk kan spelen.