

Monte Carlo Stratego

Jeroen Mets

27 juni 2008

1 Inleiding

Stratego staat vooral bekend als een spannend bordspel dat door MB [1] voor het eerst werd uitgegeven in 1961. Met een bord en twee partijen met ieder 40 stukken, geïnspireerd door 18^e-eeuwse legereenheden, wordt er een veldslag nagebootst. Het bordspel is sinds de eerste verschijning internationaal uitgegeven in verschillende edities. Sinds 2000 worden er wereldkampioenschappen gehouden, waarin vooral Nederlanders het tegen elkaar opnemen.

2 Spelregels Stratego

Stratego is van origine een spel voor 2 spelers. Het hoofddoel van het spel is de vlag van de tegenstander te veroveren. Het spel bestaat uit een speelbord en 80 speelstukken, waarvan iedere speler er 40 tot zijn beschikking heeft. Het speelbord, weergegeven in Figuur 1, is een 100-veld met 10 rijen (1 tot en met 10) en 10 kolommen (A tot en met J). De velden C6, D6, C7, D7, G6, H6, G7 en H7 zijn meren (weergegeven met \approx in Figuur 2) en kunnen niet worden bezet door de speelstukken. Ook kan er niet over de meren worden heen gesprongen.

Onder de stukken bestaat er een rangorde en er kunnen meerdere stukken met dezelfde rangorde bestaan. De verdeling en naamgeving zijn te vinden in Tabel 1, waarbij 0 de hoogste rang is en 11 de laagste rang.

Het spel bestaat uit twee fasen, te weten de opstelfase en de speelfase.

- Opstelfase: Iedere speler stelt alle 40 stukken op op het speelveld aan de boven of benedenzijde. De velden beschikbaar voor de opstelling zijn voor de beginnende speler de onderste 4 rijen en voor de andere speler de bovenste 4 rijen. De stukken worden zo opgesteld dat de rangen niet zichtbaar zijn voor de andere speler. De spelers zijn verder vrij om de stukken naar wens op te stellen, mits er op een veld niet meer dan één stuk wordt geplaatst. In Figuur 2 is een voorbeeldopstelling te vinden waarbij alle rangen zichtbaar zijn. De representaties van de rangen zijn te vinden in Tabel 1.

10											
9											
8											
7											
6			≈	≈			≈	≈			
5			≈	≈			≈	≈			
4											
3											
2											
1											
	A	B	C	D	E	F	G	H	I	J	

Figuur 1: Leeg Strategobord

10	8	9	8	ó	9	8	ó	7	ó	7	
9	9	9	5	9	5	6	8	ó	6	ó	
8	9	3	4	ó	4	S	4	5	6	9	
7	7	6	1	▷	ó	2	5	9	7	3	
6			≈	≈			≈	≈			
5			≈	≈			≈	≈			
4	2	1	9	5	6	8	3	9	5	6	
3	8	4	6	6	3	9	S	9	4	ó	
2	ó	7	8	5	9	9	6	8	ó	7	
1	7	ó	ó	5	4	9	9	ó	7	▷	
	A	B	C	D	E	F	G	H	I	J	

Figuur 2: Voorbeeld van een startopstelling waarbij alle rangen zichtbaar zijn

- Speelfase: Om beurten dienen de spelers een stuk te verzetten, waarbij de onderste speler begint. Dit wordt herhaald tot er een speler als winnaar over blijft.

Voor ieder stuk geldt dat het per beurt één veld mag verplaatsen (naar links, naar rechts, naar beneden of naar boven) naar een leeg veld of een veld waar een stuk van de tegenstander op staat. In het laatste geval vindt er een gevecht plaats waarbij er meestal geldt dat het stuk met de hoogste rang wint. Bij gelijke rang verliezen beide stukken. Verliezende stukken worden van het speelveld verwijderd. Er zijn een aantal uitzonderingsregels:

- Bommen (0) en de Vlag (11) kunnen zich niet verplaatsen.
- De Mineurs (8) maken Bommen (0) onschadelijk in een gevecht. Een Mineur neemt na het onschadelijk maken van een Bom de plaats van de Bom in.
- Verkenners (9) kunnen zoveel velden verplaatsen als gewenst is, mits er niet over meren of stukken heen wordt gesprongen en er geen diagonale stappen worden gezet.
- De Spion (10) wint bij een gevecht van de Maarschalk (1), mits de Spion aanvalt.

Rang	Representatie	Naam	Aantal
0	ó	Bom	6
1	1	Maarschalk	1
2	2	Generaal	1
3	3	Kolonel	2
4	4	Majoor	3
5	5	Kapitein	4
6	6	Luitenant	4
7	7	Sergeant	4
8	8	Mineur	5
9	9	Verkenner	8
10	S	Spion	1
11	▷	Vlag	1

Tabel 1: Verdeling strategostukken

De vlag (11) heeft hier de laagste rang omdat hij door ieder stuk kan worden geslagen. Toch is dit het belangrijkste stuk op het bord. Een speler verliest wanneer hij zijn vlag heeft verloren. In tegenstelling tot de oorspronkelijke spelregels verliest een speler ook als hij aan de beurt is en geen zetten meer kan doen. Hierdoor kan er hier geen remise worden gespeeld. Voor nu is het interessant genoeg om alleen de mogelijke uitkomsten winst en verlies te bekijken.

Aanvankelijk zijn de rangen van de stukken niet zichtbaar voor de tegenstander. Deze rangen worden bekend gemaakt wanneer er twee stukken een gevecht aangaan.

Om te voorkomen dat stukken niet oneindige achtervolgingen lopen over twee velden, is de Two-Square Rule (zie [5]) geïmplementeerd.

Er wordt tijdens Stratego bordspel kampioenschappen met aanvullende spelregels gespeeld, te lezen in [3]. Deze aanvullende regels bevatten spelregels om remise te spelen, met een klok te spelen en aangepaste regels om achtervolgingen te vermijden. Deze regels worden hier niet toegepast.

De oorspronkelijke regels zijn na te lezen op [2]. De regels die tijdens dit onderzoek worden gehandhaafd zijn op de genoemde uitzonderingen na hetzelfde, echter sommige oorspronkelijke regels zijn slechts van praktische aard voor het bordspel.

2.1 Tactiek

Om zo sterk mogelijk voor de dag te komen in een Strategospel is het nodig op de hoogte te zijn van verschillende tactieken en deze eventueel toe te kunnen passen. Met slechts een aantal basistactieken kan een speler het al een stuk langer uithouden tegen ervaren spelers. Het onderzoeken van deze tactieken kan van waarde zijn om de gebruikte kunstmatige intelligentie te ondersteunen. Tijdens de opstelfase kan er met verschillende tactieken rekening worden gehouden:

- Maak de vlag moeilijk te bereiken door er bommen omheen te zetten. Hierdoor wordt het voor de tegenstander moeilijker, zo niet onmogelijk, om zonder mineur de vlag te veroveren.
- Maak vallen voor mineurs door bommen te bewaken met sergeants.
- Zet een aantal verkenner vooraan, zodat je snel te weten kunt komen welke stukken van de tegenpartij zich als eerste in het gevecht gaan mengen. Hou ook een paar verkenner achter voor het eindspel, waar ze erg waardevol kunnen zijn.
- Zet de mineurs en de spion niet aan het front. Deze stukken komen in het spel erg goed van pas, maar vallen snel ten prooi aan stukken uit de middenklasse.
- Omdat veel spelers vaak met verkennen beginnen en informatie over de hoogste stukken veel waard is, kan het raadzaam zijn de hoogste stukken achter meren of de tweede linie te zetten. Graaf ze in ieder geval niet in.

De rest van de opstelling kan sterk afhangen van de bedachte tactiek of speelstijl. In dit onderzoek zal er meer aandacht worden besteedt aan de speelfase. Tijdens de speelfase zijn er onder andere de volgende tactieken die men tegenkomt in een Strategospel:

- Geef zo weinig mogelijk informatie weg door stukken zo lang mogelijk stil te laten staan. Een stuk dat heeft bewogen kan snel ten prooi vallen aan sterkere stukken. Daarnaast wordt de kans dat de locatie van de vlag of bommen wordt bepaald groter.
- Een grote groep beweeglijke stukken is minder mobiel dan een kleine groep. Snelheid kan belangrijk zijn in bijvoorbeeld een aanval.
- Bescherm onbekende stukken voor verkenner door er bekende stukken voor te zetten. Probeer met verkenner verderop in het spel achter de linies te verkennen.
- Agressieve onbekende stukken jagen de tegenstander soms de stuipen op het lijf, gebruik hierom ook voor verkenningdoeleinden minder sterke stukken.
- Bescherm de spion goed maar onopvallend. Zoek de maarschalk van de tegenstander en probeer deze in een val te lokken.
- Sla stukken bij voorkeur met stukken met een zo laag mogelijke rang, om de positie van de hogere stukken zo lang mogelijk geheim te houden.
- Jaag op mineurs van de tegenstander als je vlag achter bommen staat. Let goed op de eigen mineurs.

Er zijn nog vele andere tactieken en strategieën die aan de lezer zijn om te ontdekken.

3 Kunstmatige Intelligentie

Het eerste Strategospel op de computer werd in 1990 uitgebracht door Accolade [4], waarmee het voor het eerst mogelijk werd tegen de computer te spelen. De kunstmatige intelligentie (KI) component hiervan was niet heel sterk en om te verliezen was er veel geduld nodig. De laatste jaren wordt er meer aandacht besteed aan de KI-component voor computer Strategospelers. Omdat Stratego een spel is met erg veel mogelijke spelverlopen en de speler meestal geen complete informatie heeft over een situatie, is het lastig om de sterkste zetten te achterhalen.

Naarmate de processoren krachtiger worden kunnen er sterkere KI spelers worden gemaakt. Dit geldt in ieder geval voor de KI die hier wordt besproken. De intelligentie wordt met een Monte Carlo methodiek (MC) gerealiseerd. Een beschrijving van deze methodiek is verderop te vinden. De methodiek wordt steeds vaker gebruikt om intelligentie te bieden aan computerspelers en heeft in spellen als Go [6] en Scrabble [7] al veelbelovende resultaten opgeleverd.

Een van de benodigdheden voor de MC implementatie is de mogelijkheid om een partij blind of random te spelen. In het volgende hoofdstuk worden de blinde spelers besproken en kort onderzocht.

Het doel van dit Stratego onderzoek zal zijn het verkennen van de kracht van de intelligentie gerealiseerd met MC en de verschillende varianten hiervan ten opzichte van elkaar. Verwacht wordt dat deze methode een degelijke intelligentie in het spel van de speler zal brengen, echter zal de AI niet krachtig genoeg zijn om (beginnende) menselijke spelers te verslaan.

4 Blinde spelers

Blinde spelers spelen met hun ogen dicht. Ze kiezen een willekeurig (random) stuk en verplaatsen dit in een willekeurige richting, met één stap of een willekeurig aantal stappen (verkenner), mits de spelregels dit toelaten. De implementatie van blinde spelers is nodig om MC te kunnen realiseren.

Opofferingszetten, waarbij er wordt verloren terwijl de rang van het stuk van de tegenstander dat wordt aangevallen reeds bekend is, worden ook uitgesloten. Zetten van deze orde zijn de enige zetten die in alle situaties als volledig nutteloos kunnen worden beschouwd.

Om daadwerkelijk intelligentie te kunnen vaststellen worden er een aantal testresultaten van de blinde spelers bekeken. Van het (gemiddelde) aantal zetten per spel kan een bepaalde mate van intelligentie worden afgeleid. Ter referentie: tijdens toernooien worden de spelers geacht het spel binnen 450 zetten te beslissen. Mochten er meer zetten nodig zijn, kan de scheidsrechter remise beslissen. Ook kan de waarde van een bepaalde (deel-)startopstelling worden onderzocht. Zo zal blijken dat wanneer een vlag achter bommen staat deze weinig geslagen zal worden, omdat deze bommen eerst blind met een mineur dienen te worden gevonden voordat de vlag kan worden gevonden. In de eerste test wordt de vlag achter bommen geplaatst; in de tweede test wordt de vlag in open veld gezet en

10	▷	ó	ó						ó	
9	ó	ó								ó
8										
7										
6		≈	≈			≈	≈			
5		≈	≈			≈	≈			
4										
3										
2	ó							ó		ó
1		ó				ó		ó		▷
	A	B	C	D	E	F	G	H	I	J

Figuur 3: De vlaggen staan achter bommen

10		ó	ó		▷				ó	
9	ó	ó								ó
8										
7										
6		≈	≈			≈	≈			
5		≈	≈			≈	≈			
4										
3										
2	ó							ó		ó
1		ó			▷			ó		ó
	A	B	C	D	E	F	G	H	I	J

Figuur 4: De vlaggen worden niet door bommen beschermd

is deze dus makkelijker te pakken.

4.1 Test 1: blindelings spel met vlag achter bommen

Hoe de vlaggen en bommen op het bord worden geplaatst is te zien in Figuur 3. De overige stukken worden willekeurig op de helft van de speler geplaatst, maar niet ter hoogte de van meren. De testresultaten zijn te vinden in Tabel 2.

Zoals verwacht hebben de spelers het moeilijk vlaggen te veroveren en winnen ze vooral door alle stukken van de tegenstander te slaan. De gemiddelde spelduur in zetten is lang te noemen, 2.5 maal de gewenste spelduur op toernooien. De spelers zijn min of meer in evenwicht; het maakt dus niet uit wie er begint. Ter verduidelijking: een speler wint òf door de vlag te veroveren òf door alle stukken te veroveren, dus vormt de optelling van de getallen onder de kolommen “alle stukken veroverd” en “vlag veroverd” het totaal gewonnen partijen.

4.2 Test: blindelings spel met vlag in open veld

De vlaggen worden nu buiten de bommen geplaatst. De opstelling van vlag en bommen is te vinden in Figuur 4. De overige stukken worden weer willekeurig op de helft van de speler neergezet. Ook deze testresultaten staan in Tabel 2.

Test	Aantal spellen	Gemiddeld aantal zetten	Speler	Aantal keer winst	Vlag veroverd	Alle stukken veroverd
1	100000	1181	1	49901	2804	47097
			2	50099	2684	47415
2	100000	988	1	50022	19012	31010
			2	49978	18698	31280

Tabel 2: Tests met blinde spelers

5 Monte Carlo

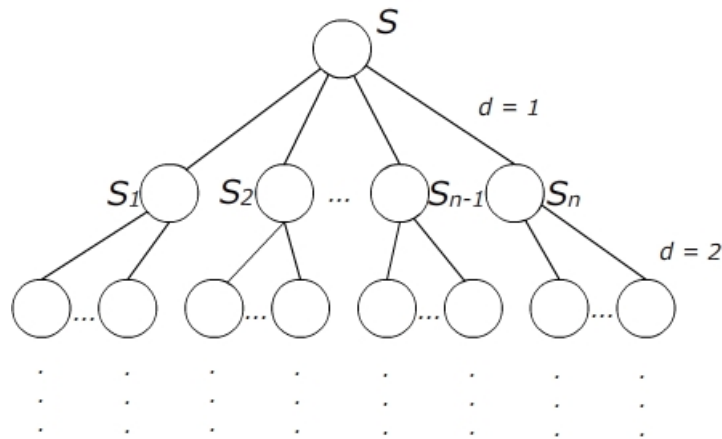
Intelligentie door middel van Monte Carlo methodiek baseert de keuzes die gemaakt dienen te worden op resultaten van vele herhalingen van situaties. Tijdens deze herhalingen worden de keuzes nog willekeurig gemaakt. Het idee hierachter is dat tussen de vele verschillende keuzes die in de herhalingen gemaakt worden er altijd wel goede keuzes zitten. Op basis van het gemiddelde resultaat van het vervolg van de keuzemogelijkheden, wordt de beste keuze geselecteerd.

De resultaten worden berekend middels een waarderingsfunctie. Een waarderingsfunctie is vaak een optelsom van positieve en negatieve factoren van een gemaakte keuze of van de situatie die uit een keuze voortvloeit. Soms is het makkelijk om een waarderingsfunctie te maken, bijvoorbeeld als er alleen wordt gekeken naar een kort spel met als uitkomst winst of verlies. Bij uitgebreider werk kunnen waarderingsfuncties moeilijker samen te stellen zijn. Het is dan bijvoorbeeld mogelijk om door middel van “leren” de factoren in een waarderingsfunctie goed af te wegen.

In de volgende paragrafen wordt beschreven hoe de Monte Carlo methodiek wordt toegepast en wat er nodig is om op deze manier een intelligente Strategospeler te realiseren.

5.1 Monte Carlo in Stratego

Om Monte Carlo methodiek toe te passen op Stratego is het volgende uitgewerkt in een programma. De speler aan zet kan n mogelijke zetten doen en overweegt iedere mogelijke zet door een aantal blinde spellen te spelen vanaf situatie (of state) S_i , met $1 \leq i \leq n$. S_i wordt gevormd door het uitvoeren van een zet in de huidige situatie S van het spel. Nu kan ieder blind spel worden uitgespeeld tot het einde of er kan d zetten vooruit worden gespeeld, d wordt hier de *lookahead* genoemd. In Figuur 5 is een schematische weergave te vinden. Ieder blind spel (van d vervolgzetten) kan een mogelijk volgend spelverloop zijn vanuit S_i . Na voor een mogelijke zet één blind spel te hebben gespeeld, wordt er voor de volgende mogelijke zet één blind spel gespeeld. Als er voor iedere mogelijke zet een vervolgspeel is gespeeld, is er een *roulatie* gedaan en wordt er begonnen met een nieuwe roulatie. Dit gaat door tot de “bedenktijd” voor de Monte Carlo



Figuur 5: Monte Carlo spelboom

speler is verstreken. Tijdens ieder blind spel worden er waarderingen gegeven aan het spelverloop. Naderhand worden de uiteindelijke waarderingen van de spelverlopen gemiddeld en opgeslagen voor iedere mogelijke zet in S . De keuzemogelijkheid met de hoogste gemiddelde waardering zal worden gekozen tot daadwerkelijke zet.

5.2 Hypothetische Opstelling

Een situatie heeft bij Stratego een aantal mogelijke onbekende factoren en dat zijn de rangen van de stukken van de tegenstander. Meestal zijn er wel een aantal stukken die hun rang al hebben bekendgemaakt, maar het overgrote deel zal nog niet bekend zijn. De informatie is dus niet perfect. Om een mogelijk spelverloop te kunnen voorspellen moeten alle factoren bekend zijn, of in ieder geval benaderd. Daarom wordt er voordat alle zetten worden gewaardeerd een hypothetische opstelling samengesteld. Aan de onbekende stukken van de tegenstander worden rangen toegekend die zij mogelijk nog kunnen hebben, met andere woorden: er worden rangen gestrooid. Sommige rangen kunnen door informatie uit het spelverloop tot dusver niet aan stukken worden toegekend: stukken die al zijn gezet kunnen geen vlag of bom zijn.

Eventueel kan er tijdens de waardering van de zetten extra worden gestrooid. Hoe vaker er tijdens de zetwaardering wordt gestrooid, hoe beter de kansen worden benaderd dat een gestrooide opstelling de juiste is.

Het is lastig om voorspellingen te doen over de opstelling van de tegenstander in een situatie. Stukken die agressief bewegen kunnen hoge rangen hebben, echter zullen slimme spelers ook lage rangen agressief laten zetten. Stukken die lang stilstaan kunnen een lagere rang hebben of vlag of bom zijn, maar kunnen ook bewust informatie achterhouden of een hoge rang hebben voor gebruik

in het midden- en eindspel. In Hoofdstuk 8 wordt er een suggestie gedaan om werkelijke opstellingen te benaderen.

5.3 Waarderingsfunctie

Om uiteindelijk een keuze te kunnen maken uit de mogelijke zetten zijn er waarderingsfuncties nodig voor de situaties die zetten opleveren. In deze waarderingsfuncties worden eigenschappen van een situatie (en dus ook zetten uit het verleden) meegenomen. Hieronder worden er een aantal besproken.

- Verkennen van een stuk. Om informatie over de opstelling van de tegenstander te weten te komen is verkennen belangrijk. Ieder stuk dat verkend is levert een aantal punten op afhankelijk van de rang van het stuk. Hoe hoger de rang, hoe waardevoller de informatie. Bij het verkennen wordt wel een stuk verloren, dat op zich weer punten in mindering brengt. Het is daarom lonender om met lage rangen te verkennen (verkenner en sergeant).
- Winnen van een stuk. Ieder stuk dat veroverd is levert punten op, afhankelijk van de rang. Verkenner leveren niet veel op, maar bommen onschadelijk maken geeft vaak een weg naar de vlag en is daarom veel waard. De vlag zelf is het meeste waard, het draait immers om de winst.
- Een stuk zetten dat nog niet heeft bewogen is informatie weggeven en levert derhalve strafpunten.

We gaan er in dit onderzoek van uit dat ieder positief punt voor de speler een negatief punt is voor zijn tegenstander. Het aantal punten van een speler is dus de negatie van het aantal punten van de tegenstander.

Een mogelijke puntenverdeling voor het veroveren, verkennen en het eerste keer zetten van een stuk per rang is in Tabel 3 te vinden.

5.4 Tests

Om aan te tonen dat MC daadwerkelijk intelligentie biedt aan een computerspeler worden er tests gedaan. Een MC-speler zal het opnemen tegen een blinde speler. De MC-speler zal in de blinde spellen 10 zetten vooruit kijken. Er wordt een tijdslimiet per zet aangegeven van 50 milliseconden. In deze tijd heeft de MC-speler de tijd om hypothetische opstellingen te strooien en blinde spellen te spelen. De resultaten van deze tests zijn te vinden in Tabel 4.

Zoals verwacht wint de MC-speler veruit de meeste spellen. Toch zijn er spellen die door de blinde speler worden gewonnen. Bij de vlag in open veld is de gemiddelde spelduur acceptabel te noemen als we deze vergelijken met de toegestane spelduur op toernooien (400 zetten). De spelduur is aanzienlijk langer als de vlag achter bommen staat. Toch heeft de MC-speler minder moeite om de vlag achter de bommen te vinden dan blinde spelers. Ongeveer een kwart

Rang	Naam	Verkend	Veroverd	Bewogen
0	Bom	200	750	-
1	Maarschalk	100	500	100
2	Generaal	50	250	100
3	Kolonel	25	100	100
4	Majoor	20	50	100
5	Kapitein	15	20	100
6	Luitenant	10	10	100
7	Sergeant	5	5	100
8	Mineur	20	50	100
9	Verkenner	-	2	100
10	Spion	-	100	100
11	Vlag	-	1000	-

Tabel 3: Waardering van de stukken

Aantal spellen	Gemiddeld aantal zetten	Speler	Aantal		Alle stukken veroverd	Gemiddeld aantal MC-spellen
			keer winst	Vlag veroverd		
1000	353	1	971	921	50	322
		2	29	25	4	-
1000	810	1	963	245	718	334
		2	37	5	32	-

Tabel 4: Stratego spellen met vlag in open veld (boven) of beschermd door bommen (onder)

van de spellen wordt gewonnen door de vlag te veroveren. Blinde spelers doen dit in ongeveer 5% van de gevallen (zie Tabel 2).

Het gemiddelde aantal MC-spellen is het gemiddelde aantal spellen k dat de MC-speler speelt om tot een zet te komen. Als er z mogelijke zetten zijn, worden er dus per mogelijke zet k/z blinde spellen gespeeld. De blinde spellen zijn in deze test maximaal 10 zetten lang.

5.5 Conclusie

Er is een mate van intelligentie bereikt door Monte Carlo methodiek toe te passen op de computerspeler. Door een aantal zetten blind vooruit te spelen kunnen zetten worden gevonden die vaak winst opleveren tegen blinde spelers. Vooral het uitroeien van de stukken van de tegenstander en het veroveren van de vlag in open veld gaat de speler makkelijker af.

Als de vlaggen echter achter bommen worden geplaatst worden de spellen veel vaker gewonnen door alle stukken te veroveren. Dit maakt de spellen aan-

zienlijk langer dan spellen met de vlaggen in open veld. Om bommen te verwijderen is er een bepaalde mate van gericht spelen nodig. Om gericht te spelen is het ten eerste belangrijk een aantal zetten vooruit te kijken. Daarnaast kan het wenselijk zijn bedachte tactieken te kunnen vasthouden. Op dit moment is dit laatste niet het geval. Per beurt worden alle mogelijke zetten steeds opnieuw bestudeerd, zonder informatie uit het verleden te gebruiken.

Het kan nog steeds voorkomen dat een blinde speler het wint van een MC-speler. Hieruit kunnen we opmaken dat deze methodiek nog zwak is te noemen zonder verdere uitbreidingen of aanpassingen. In komende hoofdstukken worden aanpassingen onderzocht en getest en worden er suggesties gedaan voor mogelijke uitbreidingen.

Als we gaan zoeken naar mogelijke oorzaken van deze zwakte, komen we al snel bij de vele mogelijke spelverlopen uit. Als de spelverlopen zouden worden gerepresenteerd door paden in een zoekboom (of *spelboom*), waarin knopen mogelijke spelsituaties en takken mogelijke zetten zouden voorstellen, zouden deze bomen al snel zeer breed kunnen worden als alle mogelijke zetten volledig zouden worden uitgetekend. In deze test, bijvoorbeeld, wordt er 10 zetten vooruit gekeken. Voor de onderste speler zijn er tijdens de startopstelling (zie Figuur 2) minimaal 6 mogelijke zetten: ieder stuk op de 4e rij kan minimaal één plaats naar boven worden verplaatst. Dit geldt analoog voor de andere speler. Stel nu dat er bij iedere beurt 6 mogelijke zetten zijn om te doen (in werkelijkheid zullen dit er meestal meer zijn). Dat betekent dat er 6^{10} (≈ 60 miljoen) mogelijke spelverlopen zijn als er 10 zetten vooruit wordt gekeken. Om een compleet beeld te krijgen van wat de beste zet is door vervolgspele te spelen, is het nodig zo veel mogelijk van deze 6^{10} spellen te spelen. Een groot aantal situaties zullen na verschillende spelverlopen hetzelfde zijn, maar dit maakt voor de te kiezen zet niet uit. Hoe beter een situatie is en hoe meer mogelijke manieren er zijn om deze te bereiken, hoe groter de kans is dat de zetten die voorafgaan aan dit spelverloop uiteindelijk worden gekozen. In de Monte Carlo methodiek worden de uitkomsten immers gemiddeld en zodoende wordt er een zet gekozen die gemiddeld het beste resultaat oplevert. Dit betekent ook dat niet per definitie de zet die het best mogelijke resultaat biedt wordt gekozen.

Gezien het feit dat er in deze test per zet gemiddeld rond de 330 spellen worden gespeeld uit miljoenen mogelijke spelverlopen, kan er niet worden gezegd dat de gekozen zet daadwerkelijk altijd (gemiddeld) de beste is. Vaak zal het tegendeel waar zijn, gezien de testresultaten. Er zou verwacht kunnen worden dat een MC-speler die meer tijd per zet krijgt beter Stratego speelt. Dit wordt in het volgende hoofdstuk onderzocht. Ook worden er een aantal andere variabele factoren binnen de methodiek die invloed hebben op de intelligentie onderzocht.

Tegenwoordig wordt er in onderzoek naar Monte Carlo methodieken veel aandacht besteed aan het versmallen van de spelboom (minder zetten meenemen in de waardering) om tot goede zetten te komen. Zie hiervoor onder andere [6]. Hier zal in dit onderzoek in Hoofdstuk 7 aandacht aan worden geschonken. In dit onderzoek zal er verder niet meer worden gekeken naar verschillende soorten startopstellingen; we plaatsen de vlag in de volgende tests in open veld. De spellen zullen korter zijn en zo kan er beter op het kiezen van de juiste zet

worden geconcentreerd.

6 Variabelen in de Monte Carlo methodiek

Er zijn een aantal variabele waarden binnen de Monte Carlo methodiek die gewijzigd kunnen worden en de mate van intelligentie kunnen beïnvloeden. De variabelen die worden onderzocht luiden als volgt:

- De tijd die een MC-speler per beurt krijgt om een zet te kiezen.
- Het maximale aantal zetten dat een MC-speler vooruit kijkt, ofwel de lookahead. Met andere woorden: de maximale spelduur in zetten die blinde spelers spelen tijdens het bepalen van de te kiezen zet.
- Het aantal keer strooien van de opstelling van de tegenstander per zet.

In de komende paragrafen wordt er eerst per variabele onderzocht welke waarden de voordeligste resultaten opleveren. Daarna wordt er gekeken naar het effect van de tijd en meer MC-spellen in combinatie met de andere variabelen. In de tests is steeds de startopstelling van Figuur 4 gebruikt met de vlag in open veld, verder willekeurig met de overige stukken aangevuld.

6.1 Tijd

Om te beginnen wordt de tijd die een MC-speler per beurt krijgt om een zet te kiezen bekeken. Er wordt verwacht dat MC-spelers die meer tijd krijgen betere zetten doen. Om direct de proef op de som te nemen wordt deze aanname getest met twee MC-spelers. Er worden 100 spellen gespeeld tussen MC-spelers die respectievelijk 10 en 100 milliseconden de tijd krijgen om een zet te doen. De resultaten van de test zijn te vinden in Tabel 5. Beide spelers kijken 10 zetten vooruit. De resultaten laten zien dat de MC-speler met langere zoektijd de sterkere is.

Dit kan vooral worden verklaart doordat het aantal MC-spellen (blinde spellen die worden gespeeld om tot een zetkeuze te komen) groter is. Om de invloed van de overige variabelen wat beter te kunnen beoordelen laten we de spelers evenveel MC-spellen spelen in een aantal komende tests. Om de invloed van de variabelen op de bedenktijd te bekijken limiteren we in de andere tests de tijd.

6.2 Lookahead

Om tot een kleinere spelboom te komen kan er ook eenvoudigweg minder zetten vooruit worden gekeken. Zo is de kans dat alle mogelijke spelverlopen worden gespeeld bij 6 mogelijke zetten als er 2 zetten vooruit wordt gekeken groot, ook als er maar korte tijd is om een zet te bepalen. Niet alle spelverlopen worden per definitie gespeeld, vanwege het willekeurige karakter dat de blinde spelers, waar de spelboom door wordt opgebouwd, hebben.

Aantal spellen	Gemiddeld aantal zetten	Speler	Aantal keer winst	Bedenktijd in ms.	Look-ahead	Gemiddeld aantal MC-spellen
100	245	1	54	100	10	463
		2	46	10	10	64
100	248	1	74	1000	10	4492
		2	26	10	10	67

Tabel 5: Stratego spellen tussen MC-spelers met verschillende toegestane bedenktijd

Het nadeel van een kleine lookahead is duidelijk. Het is lastig, zo niet onmogelijk, om bijvoorbeeld te voorzien dat de tegenstander een aanval zal doen op waardevolle stukken als de gevechten nog een aantal zetten in het verschiep liggen. Het voordeel is dat de beste zet beter benaderd kan worden door de relatief smalle spelboom, waardoor er meer spellen gespeeld kunnen worden tijdens het kiezen van de beste zet. Dit leidt uiteindelijk tot betere korte termijn keuzes.

In het bovenste gedeelte van Tabel 6 staan de resultaten van de tests waarbij spelers dezelfde bedenktijd hebben gekregen, maar verschillende lookahead. Ze laten zien dat de speler met kleine lookahead het wint van de speler met grotere lookahead. De MC-speler die slechts 2 zetten vooruit kijkt is te sterk voor de MC-spelers die 10 of 18 zetten vooruit kijken. Door de gelijke bedenktijd van 10 milliseconden waardeert speler 1 steeds duidelijk meer spellen om tot een uiteindelijke zet te komen.

Geven we de spelers echter hetzelfde aantal MC-spellen per zet, zien we dat het niet meer zo veel uitmaakt of de lookahead groot is of klein. Zie hiervoor het onderste gedeelte van Tabel 6. Toch winnen de spelers met grote lookahead het gemiddeld genomen nog niet vaker dan de spelers met zeer kleine lookahead. Het moge duidelijk zijn dat de spelers die minder MC-spellen in een vaste bedenktijd, meer tijd nodig hebben om tot het aantal vaste MC-spellen te komen.

6.3 Strooien

Omdat de informatie over de spelsituatie die beschikbaar is voor de spelers verre van perfect is, is er een methode in het leven geroepen om deze informatie te kunnen benaderen door te strooien. Zoals gezegd is het zo dat hoe vaker er gestrooid wordt, hoe beter de kans wordt benaderd dat een zekere opstelling de werkelijke opstelling is. In theorie zal deze benadering van dusdanige invloed op de waarderingszinnen zijn, dat er zetten worden gekozen die, ondanks het feit dat de informatie niet volledig is, relatief sterk zijn. Bijvoorbeeld, stukken uit het middensegment zullen minder snel ten prooi vallen aan de sterkere stukken. Dit komt doordat de kans dat tijdens een gevecht met een onbekend stuk er wordt verloren wordt meegenomen in het waarderingsgemiddelde.

We maken onderscheid tussen vier verschillende strooimodi, te weten:

Aantal spellen	Gemiddeld aantal zetten	Speler	Aantal keer winst	Bedenktijd in ms.	Look-ahead	Gemiddeld aantal MC-spellen
100	198	1	73	10	2	163
		2	27	10	10	66
100	228	1	63	10	10	74
		2	37	10	18	47
100	198	1	69	10	2	179
		2	31	10	18	47
100	159	1	48	-	2	500
		2	52	-	10	500
100	208	1	43	-	10	500
		2	57	-	18	500
100	171	1	50	-	2	500
		2	50	-	18	500

Tabel 6: Stratego spellen tussen MC-spelers met verschillende lookahead

1. Er wordt éénmaal gestrooid, en wel aan het begin van iedere beurt van de MC-speler.
2. Er wordt gestrooid aan het begin van iedere roulatie van blinde spellen.
3. Er wordt gestrooid aan het begin van ieder blind spel.
4. De daadwerkelijke opstelling van de tegenstander wordt gebruikt. De informatie wordt perfect gemaakt in deze situatie en er wordt vals gespeeld door de MC-speler.

De laatste modus, waarbij er vals wordt gespeeld, wordt gebuikt om te onderzoeken in welke mate het belangrijk is de informatie zo volledig mogelijk te maken en hoe dit het beste kan worden gedaan.

In Tabel 7 zijn de uitkomsten van verschillende tests te vinden. Alle strooi-modi zijn met elkaar vergeleken. Zoals verwacht zijn er met het valse spel van modus 4 de meeste spellen te winnen. De overige drie modi zijn redelijk aan elkaar gewaagd. Toch is het opvallend dat de speler die gebruikt maakt van strooi-modus 3 veel vaker van de valsspeler verliest dan de modus 1 en 2. De onderlinge resultaten plaatsen strooi-modus 3 ook onderaan. Door het vele strooien dat met deze modus plaatsvindt, is er minder tijd over voor MC-spellen. Dit kan een verklaring zijn voor het matige spel. Tevens is het zo dat de mogelijke zetten niet goed met elkaar worden vergeleken. Voor iedere mogelijke zet waar een MC-spel voor wordt gespeeld, wordt er een nieuwe situatie gecreëerd door te strooien.

Bij modus 2 worden de zetten beter met elkaar vergeleken; voor iedere situatie worden alle mogelijke zetten beoordeelt. Toch wint dit het niet van het gebruik van modus 1. Er wordt nog geen winst behaald met het idee vaker dan eenmaal te strooien. We zullen aan de hand van tests verderop in dit hoofdstuk onderzoeken of deze winst wel kan worden behaald door de spelers langer bedenktijd te geven, en zo de tijd te geven om het strooien effect te laten hebben. Hiervoor is het immers nodig veel spellen te spelen.

In het onderste deel van de tabel staan weer de gegevens van de spellen waarbij er per zet tot een aantal MC-spellen wordt geteld. Het is duidelijk dat modus 3 niet alleen traag is; het levert ook de minste intelligentie op. Modus 1 biedt de beste resultaten tegen het valse spel van modus 3. Hieruit kan er worden opgemaakt dat het regelmatig willekeurig strooien tijdens het kiezen van een zet geen oplossing biedt voor de onvolledige informatie in een spelsituatie.

In feite wordt er per keer strooien een nieuwe spelboom gemaakt. Iedere nieuwe gestrooide opstelling biedt een nieuwe spelsituatie, wat de wortel is van een nieuwe spelboom. Dit maakt het aantal mogelijke spellen om te spelen nog groter dan dat het al is. Hiermee is de kans kleiner dat er goede zetten worden gedaan als er veel onbekende stukken zijn. De kans op goede zetten wordt groter als er meer stukken bekend worden. Dit is een mogelijke verklaring voor het feit dat het meerdere malen strooien weinig oplevert en het net zo aantrekkelijk is om willekeurige aannames te doen.

6.4 Effect van meer MC-spellen

Er is al aangetoond dat hoe meer MC-spellen er per beurt worden gespeeld, hoe beter de zetten zijn. In theorie zouden spelers die ook verder vooruit kijken in het spel voordeel halen ten opzichte van spelers die minder ver kijken. Echter, door de grootte van de spelboom in combinatie met de korte tijd om tot een zet te komen, loont het om minder ver vooruit te kijken. De gemiddelde beste zet kan met kleinere lookahead beter worden benaderd. Om te onderzoeken of meer MC-spellen leiden tot betere resultaten van de spelers met grotere lookahead zijn er weer een aantal tests gedaan. In deze tests maken de spelers gebruik van strooimodus 1. Tabel 6 laat zien dat het aantal MC-spellen van invloed is op de resultaten van de spellen die worden gespeeld door spelers met verschillende lookahead. Een gelijk aantal MC-spellen zorgt ervoor dat spelers haast even sterk zijn. Een korte bedenktijd is dus nadelig voor spelers met grote lookahead.

Als we het aantal MC-spellen vergroten zien we dat spelers met grote lookahead elkaar in evenwicht houden. De speler met lookahead 2 kijkt niet ver genoeg vooruit en verliest het nu net in de meeste spellen. Zie hiervoor Tabel 8. Een lookahead van 2 laat de gemiddeld beste zet in de hypothetische situatie gemakkelijk benaderen door het grote aantal MC-spellen dat wordt gespeeld. De grotere lookahead van de andere spelers heeft dus een positieve invloed op het aantal spellen dat ze winnen.

Ook is het effect van meer MC-spellen op de intelligentie van spelers met verschillende strooimodi bekeken. We vergeleken strooimodi 1 en 2 met elkaar. Het vaker strooien van opstellingen om de werkelijke opstelling te benaderen wierp

Aantal spellen	Gemiddeld aantal zetten	Speler	Aantal keer winst	Bedenk-tijd in ms.	Look-ahead	Gemiddeld aantal MC-spellen	Strooi-modus
1000	211	1	532	10	10	74	1
		2	468	10	10	73	2
1000	229	1	459	10	10	55	3
		2	541	10	10	70	2
1000	175	1	618	10	10	71	4
		2	382	10	10	69	2
1000	224	1	576	10	10	70	1
		2	424	10	10	53	3
1000	167	1	713	10	10	71	4
		2	287	10	10	55	3
1000	160	1	345	10	10	72	1
		2	655	10	10	71	4
100	213	1	56	-	10	500	1
		2	44	-	10	500	2
100	284	1	41	-	10	500	3
		2	59	-	10	500	2
100	130	1	73	-	10	500	4
		2	27	-	10	500	2
100	221	1	56	-	10	500	1
		2	44	-	10	500	3
100	167	1	79	-	10	500	4
		2	21	-	10	500	3
100	114	1	34	-	10	500	1
		2	66	-	10	500	4

Tabel 7: Stratego spellen tussen MC-spelers met verschillende strooimodi.

Aantal spellen	Gemiddeld aantal zetten	Speler	Aantal keer winst	Bedenktijd in ms.	Look-ahead	Gemiddeld aantal MC-spellen
100	195	1	45	100	2	1138
		2	55	100	10	469
100	346	1	42	1000	2	7584
		2	58	1000	10	4536
100	168	1	56		10	500
		2	44	100	18	500
100	187	1	50		10	5000
		2	50		18	5000
100	250	1	47		2	5000
		2	53		10	5000
100	219	1	43		2	5000
		2	57		18	5000

Tabel 8: Meer bedenktijd, verschillende lookahead

ook met 5000 MC-spellen nog geen vruchten af. De winstaantallen verschilden dusdanig weinig dat er geen conclusies uit kunnen worden gehaald, behalve dat het effect van meer geduld bij het zetten minimaal is voor deze twee strooimodi.

6.5 Conclusie

Zolang de bedenktijd of het aantal MC-spellen veel effect heeft op de intelligentie, is het aantal MC-spellen dat kan worden gespeeld bepalend voor de mate van intelligentie. Het effect van een groot aantal MC-spellen neemt af naarmate de lookahead kleiner wordt. De gemiddeld beste zet wordt gemakkelijk benaderd met kleine lookahead. Echter, spelers die niet ver vooruit kijken worden verslagen door de spelers die dat wel doen, mits er voldoende bedenktijd is en er genoeg MC-spellen worden gespeeld. Zo heeft ook de lookahead invloed op de intelligentie.

Gezien het feit dat van de onderzochte variabelen het aantal MC-spellen de meeste invloed op de intelligentie, is het verstandig hier extra aandacht aan te geven. In Hoofdstuk 7 wordt er een methode onderzocht die de beste zet met minder MC-spellen kan benaderen. Zodoende wordt er geprobeerd meer uit de Monte Carlo methodiek te halen.

7 Monte Carlo en het snoeien van de MC spelboom

Analoog aan het snoeien (*progressive pruning*) in de MC spelboom van het spel Go beschreven in [6] kan ook de MC spelboom voor Stratego worden gesnoeid. Dit gebeurt tijdens het waarderen van de mogelijke vervolgzetten. Als er bij iedere mogelijke zet een aantal maal blinde spellen zijn gespeeld, kan de zwakste zet worden weggestreept en kunnen de overige zetten meer spellen spelen om te bepalen welke zet gemiddeld het beste resultaat oplevert.

De uitwerking van deze zeselectie strategie is als volgt. Voor iedere zet wordt er aan de hand van de waardering (W) van die zet een interval om deze waardering vastgesteld. Dit interval wordt aangegeven door een ondergrens (W_l) en een bovengrens (W_r). Van de hoogst gewaardeerde zet wordt er naar de ondergrens gekeken. Iedere andere zet waarvan de bovengrens onder deze ondergrens ligt, wordt gesnoeid. Deze zet wordt vervolgens niet meer geëvalueerd in de roulaties. Zo kunnen er meer MC-spellen worden gespeeld voor zetten die er beter uit zien, om de waardering van de beste zetten uiteindelijk beter te kunnen benaderen en een betere zetkeuze te doen.

De ondergrens respectievelijk bovengrens worden als volgt berekend:

$$W_l = W - \frac{\sigma * A}{\sqrt{n}}$$

$$W_r = W + \frac{\sigma * A}{\sqrt{n}}$$

Hierbij is σ de standaardafwijking die is vastgesteld door middel van experimenten. A is een variabele waarde die veel invloed heeft op het interval, en dus op het wel of niet snoeien van een zet. n is hier het aantal MC-spellen dat is gespeeld voor de betreffende zet. Naarmate er meer MC-spellen zijn gespeeld wordt de kans op snoeien dus steeds groter.

Een MC-speler die van de uitbreiding gebruik maakt, speelt eerst een aantal roulaties om tot een aanvankelijk gemiddelde voor iedere zet te komen. Hoeveel dit er zijn kan variëren. Vroeg snoeien houdt in dat de beste zet mogelijk gesnoeid wordt, laat snoeien leidt tot minder MC-spellen voor de beste zetten. Vervolgens wordt er per roulatie (of meerdere roulaties) een mogelijke zet gesnoeid. Samen met de parameter A geeft dit al een aantal variabelen waarvoor er optimale waarden gezocht kunnen worden.

Tabel 9 laat zien dat het snoeien van de MC-spelboom een positief effect heeft op het spel van de MC-speler. Hoewel het verschil klein is, speelt de MC-speler met lookahead 10 die snoeit beter spel dan zijn tegenstander die niet snoeit. Het snoeien van zetten heeft niet altijd een positieve uitwerking. Er is met dezelfde snoeiparameters ook getest met minder MC-spellen, waarbij de spelers even sterk bleken. Meer MC-spellen zorgen er zelfs voor dat de snoeiende

Aantal spellen	Gemiddeld aantal zetten	Speler	Aantal keer winst	Snoeien	Look-ahead	Gemiddeld aantal MC-spellen
1000	182	1	523	Ja	10	250
		2	477	Nee	10	250
1000	188	1	501	Ja	2	250
		2	499	Nee	2	250

Tabel 9: Meer bedenktijd, verschillende lookahead

speler verliest. De beste zet wordt weggesnoeit, terwijl de speler die niet snoeit genoeg MC-spellen speelt om de beste zet beter te benaderen.

Ook de lookahead heeft invloed, maar in wezen kan dit worden teruggekoppeld naar het aantal MC-spellen. Er zijn minder MC-spellen nodig om een goede zet te benaderen, en de winst die het snoeien kan opleveren is met de huidige parameters niet merkbaar. Zie Tabel 9 en de spellen met lookahead 2.

Progressive pruning kan een positieve inbreng hebben op het spel van de MC speler. Echter zijn de juiste snoeiparameters afhankelijk van het aantal spellen dat wordt gespeeld. Deze kunnen door middel van experimenten of learning worden gevonden. Uiteindelijk kan de snoefunctie dusdanig worden aangepast dat iedere MC-speler profijt kan hebben van deze uitbreiding.

8 Uitbreidingen voor de toegepaste Monte Carlo methodiek

Er zijn verschillende manieren om de gebruikte methodiek uit te bouwen om tot een intelligentere speler te komen. Deze uitbreidingen zijn niet in dit onderzoek gerealiseerd, maar zijn suggesties voor toekomstig werk.

8.1 Andere zetsselectie strategieën

In plaats van *progressive pruning* kan er ook gebruik worden gemaakt van *simulated annealing* [8]. Het idee achter *simulated annealing* is overgenomen uit de natuurkunde. Hierbij kan er worden gesnoeid tijdens het spelen van een blind spel. De zetten worden tijdens het blinde spel geëvalueerd alvorens er wordt gezet. Daarna wordt er een willekeurige zet gekozen. Indien dit de beste zet is wordt de zet geaccepteerd. Zo niet, dan is er een kans afhankelijk van de waarde-ning en de temperatuur dat de zet wordt geaccepteerd. Indien de zet niet wordt geaccepteerd, wordt er een andere zet gekozen. Vaak neemt de temperatuur af naarmate het spel verstrijkt, zodat zetten minder snel worden geaccepteerd, maar dit is optioneel. De optimale functie om een zet al dan niet te accepteren kan door middel van experimenten worden gevonden.

8.2 Kansberekening integreren in waarderingsfunctie

Om de beschreven nadelen van strooien voor een groot gedeelte te omzeilen kan er tijdens het waarden van de mogelijke zetten rekening worden gehouden met de kans dat een onbekend stuk een bepaalde rang heeft. Er wordt voor de waardering van mogelijke zetten gestrooid. Door gebruik te maken van de wetenschap welke stukken er nog op het bord aanwezig zijn kan deze kans eenvoudig worden berekend. Deze kans wordt tijdens gevechten in blinde spellen vermenigvuldigd met de waarde van de gestrooide stukken. Op deze manier worden de waarderingsfuncties van gevechten met onbekende stukken enigszins gerelativeerd, in tegenstelling tot de wijze waarop dat in de praktijk met de huidige methodiek gebeurt.

8.3 Leren

Om de variabele factoren binnen de Monte Carlo methodiek te kunnen optimaliseren, kan MC Stratego worden uitgebreid met lerende methodieken. In de volgende paragrafen wordt beschreven van welke factoren de optimale waarden dan automatisch kunnen worden gevonden.

8.3.1 Waarderingsfunctie

De waarden van de stukken die gebruikt zijn in de waarderingsfunctie zijn slechts richtlijnen en bedachte waarderingsfuncties van de stukken. Deze waarden kunnen door middel van een leerfunctie naar realistischere waarden worden geëvolueerd. Wel zullen de waarden van de stukken afhankelijk zijn van de speelstijl die de speler hanteert. Agressieve en aanvallende spelers zullen meer waarde hechten aan het veroveren van (lagere) stukken. Spelers die het spel snel willen beslissen hechten veel waarde aan het verkennen van bommen en sterke stukken. Verdedigende spelers hechten veel waarde aan bommen en de eigen spion en het veroveren van vijandelijke mineurs. De geëvolueerde waarderingsfunctie zal dus afhankelijk zijn van het karakter van de MC-speler en andersom.

Om zo weinig mogelijk informatie weg te geven kan het daarnaast aantrekkelijk zijn om bij het veroveren van stukken een zo laag mogelijke rang te gebruiken. Zo worden hoge rangen het minst snel vrijgegeven. Het wordt dus beloond als, in plaats van de generaal, een luitenant een sergeant pakt.

Ook kan er rekening worden gehouden met stukken waarvan de waarde afhangt van de spelsituatie, zoals de spion en de mineurs. Als de maarschalk van de tegenpartij is veroverd, heeft de spion enkel nog nut als verkennend stuk. Mineurs zijn vaak onmisbaar om de vlag van de tegenstander te bereiken. Als een speler nog slechts één mineur over heeft kan deze van zeer hoge waarde zijn voor de speler en dient extra beschermd te worden.

8.3.2 Optimale lookahead en strooimodus bij gegeven bedenktijd

De zoektocht naar de optimale variabelen beschreven in Hoofdstuk 6 kan worden geautomatiseerd. Om te onderzoeken wat er precies gebeurt als de variabele

factoren binnen MC wijzigen is dit voor toekomstig werk gelaten. Het zal duidelijk zijn dat hoe meer bedenktijd een MC speler tot de beschikking heeft, hoe intelligenter deze zal spelen. Binnen de bedenktijd kan er automatisch naar een optimale lookahead worden gezocht en eventueel welke stroommodus hier het best bij past.

8.3.3 Hypothetische opstellingen

Zoals aangetoond in Hoofdstuk 6 is het wenselijk zo veel mogelijk over de opstelling van de tegenstander te weten te komen. De eerder in dit hoofdstuk beschreven integratie van kansberekening is misschien niet voldoende om veel vooruitgang in de kracht van de MC speler te boeken. Er liggen mogelijkheden om aan de hand van de gegevens over stukken van de tegenstander een opstelling te benaderen middels bijvoorbeeld een neurale netwerk. Waar bommen staan zou de vlag mogelijk in de buurt kunnen staan en agressieve bewegingen van stukken kunnen leiden tot een benadering van de rang ervan.

8.4 Planning

Vanwege het aantal onbekende factoren in Stratego kan het aantrekkelijk zijn delen van het vooruit te kijken spel abstract te maken. Dit kan als er van te voren bepaalde acties worden gepland. Planning is één van de nieuwe ideeën om computerspelers in Real Time Strategie spellen krachtiger en inventiever te maken. Meer hierover is te vinden in [9].

9 Conclusie

In dit onderzoek is laten zien dat Monte Carlo methodiek voor een degelijke intelligentie kan zorgen voor Stratego computerspelers. De bedenktijd en de lookahead zijn hierbij van invloed op de kracht van de speler. Een optimale invulling van variabele waarden binnen de methodiek en uitbreidingen als *progressive pruning* verhogen het rendement van de MC-spellen die tijdens de zetbepaling worden gespeeld.

De huidige situatie biedt veel mogelijkheden tot uitbreidingen en optimalisaties om tot een nog krachtigere MC-speler te komen.

Referenties

- [1] MB Milton Bradley, Wikipedia, 16 juni 2008
http://nl.wikipedia.org/wiki/Milton_Bradley_Company
- [2] Jumbo Stratego Spelregels, 16 juni 2008
<http://www.spelregels.eu/rules/stratego/00498.pdf>

- [3] Internationale Stratego Federatie Spelregels , 16 juni 2008
<http://www.strategobond.nl/inhoud/ISF%20Stratego%20spelregels.html>
- [4] Informatie over Accolade, 16 juni 2008
<http://www.mobygames.com/company/accoladeinc->
- [5] Computer Stratego World Championship, 16 juni 2008
<http://www.strategousa.org/wiki/index.php>
- [6] G. Chaslot, J. Saito, B. Bouzy, J. W. H. M. Uiterwijk, and H. J. van den Herik. Monte-Carlo Strategies for Computer Go. In P. Schobbens, W. Vanhoof, and G. Schwanen, editors, *Proceedings of the 18th BeNeLux Conference on Artificial Intelligence, Namur, Belgium*, pages 83–91, 2006.
<http://www.cs.unimaas.nl/g.chaslot/papers/mcscg.pdf>
- [7] B. Sheppard. World-championship-caliber Scrabble. *Artificial Intelligence* 134, pages 241–275, 2002.
- [8] Simulated Annealing, Wikipedia, 16 juni 2008
http://nl.wikipedia.org/wiki/Simulated_annealing
- [9] M. Buro, M. Chung and J. Schaeffer. Monte Carlo Planning in RTS games. In Simon Lucas and Graham Kendall, editors, *Proceeding of the IEEE Symposium on Computational Intelligence and Games*, 2005.
<http://www.cs.ualberta.ca/~mburo/ps/mcplan.pdf>