



Universiteit
Leiden
The Netherlands

Calculated Moves: Generating Air Combat Behaviour

Toubman, A.

Citation

Toubman, A. (2020, February 5). *Calculated Moves: Generating Air Combat Behaviour*. *SIKS Dissertation Series*. Retrieved from <https://hdl.handle.net/1887/84692>

Version: Publisher's Version

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

Downloaded from: <https://hdl.handle.net/1887/84692>

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

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/84692> holds various files of this Leiden University dissertation.

Author: Toubman, A.

Title: Calculated Moves: Generating Air Combat Behaviour

Issue Date: 2020-02-05

Samenvatting

Het trainen met virtuele tegenstanders (ook wel *computer generated forces* of CGFs genoemd) geeft gevechtspiloten de gewenste ervaring voor vliegoperaties, zonder de hoge kosten die het trainen met echte vliegtuigen met zich mee brengt. In de praktijk zijn deze CGFs echter niet zo veelzijdig als ze zouden kunnen zijn. Dit komt vooral door een gebrek aan realistische gedragsmodellen. Het ontbreken van gedragsmodellen motiveerde mij om *air combat training* simulaties te ontwerpen en te verbeteren. In het proefschrift onderzoeken we in hoeverre het mogelijk is om de gedragsmodellen voor CGFs in een luchtgevecht automatisch te genereren, met behulp van *machine learning* (machinaal leren).

Het domein van het luchtgevecht is complex. Daarom is het belangrijk dat de *machine learning* methodes die binnen dit veld ingezet worden, opgewassen zijn tegen de uitdagingen van dit domein. In hoofdstuk 1 identificeren we vijf uitdagingen die nieuw gegenereerde gedragsmodellen het hoofd moeten bieden, om bruikbaar te zijn binnen trainingssimulaties. Dit zijn: (A) het produceren van teamcoördinatie, (B) de computationele evaluatie van CGF-gedrag, (C) het efficiënt hergebruiken van opgebouwde kennis, (D) het valideren van gegenereerde gedragsmodellen en (E) het genereren van gedragsmodellen die toegankelijk en leesbaar zijn voor mensen.

Aan de hand van de bovenstaande motivatie en de vijf uitdagingen, komen wij tot de volgende probleemstelling: *In hoeverre is het mogelijk om dynamic scripting te gebruiken om gedragsmodellen te genereren voor CGFs in een luchtgevecht, waarbij ingespeeld wordt op de vijf uitdagingen van gegenereerde gedragsmodellen?* In de probleemstelling wordt gesproken over het *dynamic scripting* (dynamisch schrijf-) algoritme. Dit algoritme produceert leesbare gedragsmodellen en komt zo tegemoet aan uitdaging E. Op basis van de overige vier uitdagingen formuleren we vijf onderzoeksvragen die we in het proefschrift verder gaan onderzoeken.

In hoofdstuk 2 presenteren we achtergrondinformatie over het proces waarmee gedragsmodellen tegenwoordig geproduceerd worden. Bovendien introduceren we (a) *machine learning* en (b) het *dynamic scripting* algoritme in het bijzonder. Daarbij kijken we terug op eerder werk over het genereren van gedragsmodellen voor luchtgevechten met behulp van *machine learning*.

In hoofdstuk 3 onderzoeken we onderzoeksvraag 1: *In hoeverre kunnen we gedragsmodellen voor CGFs in een luchtgevecht genereren waarbij sprake is van teamcoördinatie?* Vandaag de dag

is het *two-ship* (een combinatie van twee vliegtuigen, ook bekend als het *tweetje*) de kleinste eenheid die gevechtsmissies in de lucht onderneemt, bestaande uit een *lead* (leider) en een *wingman* (volgvlieger). Om in hun missies te slagen, moeten de *lead* en *wingman* in een *two-ship* nauwgezet samenwerken. Daarom is het belangrijk dit soort samenwerking terug te zien in de gedragsmodellen die worden gegenereerd voor een *two-ship* bestaande uit CGFS. Binnen het raamwerk van het *dynamic scripting* algoritme, definiëren we drie methodes voor dergelijke coördinatie: (1) een gedecentraliseerde methode zonder communicatie tussen de CGFS genaamd TACIT, (2) een gecentraliseerde methode met communicatie genaamd CENT en (3) een gedecentraliseerde methode met communicatie genaamd DECENT. Hierna voeren we drie series van geautomatiseerde simulaties uit. In elke serie gebruiken we *dynamic scripting* om gedrag voor een *two-ship* te genereren dat tegen een voorgeprogrammeerde tegenstander moet vechten, met gebruik van één van de genoemde coördinatiemethodes. Het blijkt dat elk van de drie methodes tot een flexibele rolverdeling tussen de CGFS leidt. Van de drie methodes levert de coördinatie die CENT produceert het meest effectieve gedrag met het hoogste aantal gewonnen ontmoetingen. Gebaseerd op ons onderzoek mogen we concluderen dat we met behulp van *dynamic scripting* in staat zijn om (a) verschillende vormen van team gedrag te genereren en (b) de rolverdeling die binnen een team is aangenomen goed te herkennen.

In hoofdstuk 4 bekijken we onderzoeksvraag 2: *In hoeverre kunnen we het beloningsmechanisme voor CGFS in een luchtgevecht verbeteren?* Het beloningsmechanisme (*reward function* in het Engels) is een essentieel deel van *dynamic scripting*. Het evalueert de wenselijkheid van het gedrag dat voortvloeit uit de gegenereerde gedragsmodellen en produceert vervolgens een belonend signaal dat het *dynamic scripting* algoritme stimuleert om de gedragsmodellen bij de volgende poging te verbeteren. Een veel gebruikt beloningsmechanisme is het binaire beloningsmechanisme (in het proefschrift BIN-REWARD genoemd): daarbij krijgt het algoritme een beloning van 1 als de CGFS een gesimuleerd gevecht hebben gewonnen (gewenst gedrag) aan de hand van de gegenereerde gedragsmodellen en 0 als er is verloren. In het gebruik blijkt echter dat een beloning bij dit mechanisme zowel schaars (een CGFS moet precies het juiste gedrag vertonen om een beloning te kunnen verkrijgen) als onvoorspelbaar (onverwachte gebeurtenissen in de omgeving van de CGFS kunnen ervoor zorgen dat hetzelfde gedrag een andere uitkomst heeft) is. Daarom is het goed mogelijk dat een ander beloningsmechanisme meer gewenst gedrag oplevert. Wij hebben daarop twee nieuwe beloningsmechanismen ontwikkeld voor het gebruik in het luchtgevechtsdomein: DOMAIN-REWARD, dat zich richt op het minder schaars maken van beloningen en AA-REWARD, dat gericht is op het voorspelbaarder maken van de beloningen. Beide mechanismen hebben wij getest in geautomatiseerde simulaties. Uit de resultaten maken wij op dat het de DOMAIN-REWARD niet gelukt om de resultaten van het binaire beloningsmechanisme te verbeteren. Het gebruik van AA-REWARD leidt er echter toe dat er 12,6% vaker wordt gewonnen door de CGFS.

In hoofdstuk 5 bespreken we onderzoeksvraag 3: *In hoeverre kan kennis die verkregen is door dynamic scripting succesvol uitgewisseld worden tussen CGFS in verschillende situaties?* De gedragsmodellen die gegenereerd worden door het *dynamic scripting* algoritme bevatten kennis

over verschillende situaties in luchtgevechten. In hoofdstuk 3 en 4 gebruikten we *dynamic scripting* bijvoorbeeld om gedragsmodellen te genereren voor *two-ships* in twee-tegen-één-scenario's. Wij verwachten dat de kennis in deze gedragsmodellen tot op zekere hoogte in andere scenario's hergebruikt kan worden. We plaats een *two-ship* van CGFs in scenario's waarin ze moeten leren om twee vijanden te verslaan. Daarbij gebruiken we *dynamic scripting* om gedragsmodellen voor het *two-ship* te genereren. Dit doen we twee keer. De eerste keer moet het *two-ship* zonder voorkennis leren om de vijanden te verslaan. De tweede keer krijgt het algoritme voorkennis mee, in de vorm van de gedragsmodellen die gegenereerd zijn in een eerder twee-tegen-één-scenario. In alle twee-tegen-twee-scenario's blijkt het *two-ship* met voorkennis effectiever gedrag te vertonen dan het *two-ship* zonder voorkennis. Bovendien wordt dit effectievere gedrag sneller bereikt.

In hoofdstuk 6 bekijken we onderzoeksvraag 4: *Hoe moeten we machine-gegenereerde gedragsmodellen voor luchtgevechtstrainingen valideren?* Validatie is een belangrijke stap in de ontwikkeling van gedragsmodellen, aangezien het een gestructureerde manier biedt om de bruikbaarheid van de modellen te beoordelen (in ons geval voor trainingsdoeleinden). Er is echter geen standaardmanier om alle verschillende soorten gedragsmodellen te valideren. Daarom ontwikkelen we in hoofdstuk 6 een nieuwe validatie methode, die specifiek bedoeld is voor computer-gegenereerde gedragsmodellen voor luchtgevechten. De procedure bestaat uit drie stappen. De eerste stap is het opnemen van *human-in-the-loop* simulaties (vrij vertaald: simulaties met mensen in het terugkoppelingsproces), waarbij menselijke deelnemers het opnemen tegen CGFs die aangestuurd worden op basis van (a) gedragmodellen die gemaakt zijn door menselijke professionals of (b) nieuw gegenereerde gedragsmodellen. De tweede stap bestaat uit een gestructureerde beoordeling van het gedrag van de CGFs. Deze beoordeling wordt uitgevoerd door experts in het veld, met behulp van een nieuw ontwikkeld scoreformulier voor het gedrag van CGFs in een luchtgevecht. Dit scoreformulier noemen wij de ATACC. De derde stap is het gebruik van de TOST (*two one-sided t-tests*, oftewel twee eenzijdige *t-tests*) methode om te bepalen of de beoordelingen van het gedrag dat voortkomt uit de eerder gemaakte gedragsmodellen statistisch gelijkwaardig zijn aan de beoordelingen van het nieuwe gedrag. Als dit zo is, beschouwen we de computer-gegenereerde gedragsmodellen als valide voor het gebruik in trainingssimulaties.

In hoofdstuk 7 bespreken we onderzoeksvraag 5: *In hoeverre zijn gedragsmodellen die gegenereerd zijn met dynamic scripting valide voor het gebruik in trainingssimulaties?* We passen de valideringsmethode uit hoofdstuk 6 toe op een groep nieuw gegenereerde gedragsmodellen. Als ijkpunt gebruiken we een selectie van mens-gegenereerde gedragsmodellen die gebruikt zijn voor daadwerkelijke trainingssimulaties. We draaien *human-in-the-loop* simulaties waarin vliegers van de Koninklijke Luchtmacht het opnemen tegen CGFs die aangestuurd worden door gegenereerde gedragsmodellen. Het vertoonde gedrag wordt beoordeeld door vlieginstrueteurs, met behulp van de ATACC. Tijdens de beoordeling wordt er gelet op het voorkomen van negen soorten gedragingen. Vervolgens worden de beoordelingen van de mens-gegenereerde gedragingen en de computer-gegenereerde gedragingen vergeleken. Het blijkt dat zes van de negen

soorten gedragingen gelijkwaardig scoren. Op basis van dit resultaat mogen we de computer-gegenereerde modellen niet als volledig valide beschouwen. De literatuur adviseert ons echter om verschillende gradaties van validiteit te erkennen. We beschouwen de nieuw gegenereerde gedragsmodellen daarom als eniger mate valide.

In hoofdstuk 8 ronden we dit proefschrift af met het geven van een samenvatting van de antwoorden op de vijf onderzoeksvragen en het antwoord op de probleemstelling. Ons onderzoek toont aan dat *dynamic scripting* het geautomatiseerd genereren van gedragsmodellen voor luchtgevechten faciliteert, terwijl het flexibel genoeg blijkt om de uitdagingen van het luchtgevechtdomein het hoofd te bieden. De validiteit van de nieuw gegenereerde gedragsmodellen blijft een aandachtspunt in komende onderzoeken.