Calculated Moves: Generating Air Combat Behaviour

Toubman, A.

Author: Toubman, A.
Title: Calculated Moves: Generating Air Combat Behaviour
Issue Date: 2020-02-05

8 Conclusions

In this chapter we summarise our answers to the five research questions (Section 8.1) and formulate our answer to the problem statement posed in Chapter 1 (Section 8.2). Finally, we provide two recommendations for future research (Section 8.3).

8.1 Answers to the research questions

In Section 1.3, we posed five research questions. Below, we provide a summary of our answers to the five research questions based on the research performed in the previous chapters.

Research question 1: *To what extent can we generate air combat behaviour models that produce team coordination?*

The answer to the first research question is derived from Chapter 3. We were able to implement three methods within the rule-based framework of dynamic scripting (TACIT, CENT, DECENT) that each produced a form of team coordination. We demonstrated how the three methods lead to a flexible division of roles within a two-ship of air combat CGFS. Out of the three methods, the team coordination produced by the CENT method (viz. centralised coordination by means of communication) resulted in the most effective behaviour that reached the highest win rates. Our answer to the research question is that by means of dynamic scripting, we are able to (a) generate multiple forms of team behaviour, (b) compare the effectiveness of the produced behaviour, and (c) easily inspect the roles assumed by the team members.

Research question 2: *To what extent can we improve the reward function for air combat CGFS?*

The answer to the second research question is derived from Chapter 4. The common but simple binary reward function such as BIN-REWARD (i.e., 0 for losing an encounter, 1 for winning an encounter) offers rewards that are (a) sparse and (b) unstable. We developed two new reward functions: (1) DOMAIN-REWARD, which offers less sparse but still somewhat unstable rewards, and (2) AA-REWARD, which offers somewhat sparse rewards that are entirely stable. We found that DOMAIN-REWARD did not improve the performance of air combat CGFS over the use of BIN-REWARD, but the use of AA-REWARD lead to a performance increase of 12.6% while maintaining the same learning speed as BIN-REWARD.

Our answer to the research question is that we are able to improve the reward function by making the rewards offered by the reward function (a) less sparse and (b) stable. The improvements result in a 12.6% increase in final performance.

Research question 3: *To what extent can knowledge built with dynamic scripting be transferred successfully between CGFs in different scenarios?*

The answer to the third research question is derived from Chapter 5. To answer this research question, we designed a use case for transfer learning in air combat simulations. In the use case, two distinct two-ships learn to defeat two blue opponents in a set of two-versus-two scenarios. One of the two-ships uses transferred knowledge that has been previously built up in a two-versus-one scenario. In practical terms, dynamic scripting allows for very straightforward transfers of knowledge. Knowledge is stored in the form of the weights associated with each rule in a rulebase. Therefore, a transfer simply entails copying the rules and their weights to a new rulebase. We determined the success of the transfer by comparing the performance of the two two-ships (one with, and one without transferred knowledge) by three measures. Each of the three measures indicated that the use of the transferred knowledge resulted in a significant increase in performance. Thus, our answer to the research question is that knowledge built with dynamic scripting can be successfully transferred to a large extent.

Research question 4: *How should we validate machine-generated air combat behaviour models for use in training simulations?*

The answer to the fourth research question is derived from Chapter 6. There is no one-size-fits-all solution to the validation of machine-generated behaviour models. Therefore, our answer to the research question is a newly developed validation procedure consisting of five steps.

Step 1. Selecting a sample of professionally written behaviour models (the 4P-models) which exemplify desirable behaviour.

Step 2. Generating a sample of behaviour models by means of machine learning (the 4M-models).

Step 3. Applying the 4P-models and the 4M-models in human-in-the-loop simulations.

Step 4. Assessment of the behaviour produced by the behaviour models by means of the ATACC questionnaire.

Step 5. Equivalence testing to determine whether the assessments of the behaviour produced by the 4M-models are statistically equivalent to the assessments of the behaviour produced by the 4P-models.

In Step 5, assuming the assessments are statistically equivalent, we consider the 4M-models to be valid for use in human-in-the-loop simulations. To the best of our knowledge, this is the first time a validation procedure for machine-generated air combat behaviour models has been formulated and documented.

Research question 5: *To what extent are air combat behaviour models generated by means of dynamic scripting valid for use in training simulations?*

The answer to the fifth research question is derived from Chapter 7. We applied the validation procedure from Chapter 6. As the baseline, we selected 4P-models that were designed by subject matter experts. We generated new 4M-models, and then applied both the 4P-models and 4M-models in realistic human-in-the-loop F-16 fighter jet simulations. The assessment of the behaviour of the CGFs was carried out by active duty F-16 instructor pilots. On six out of the nine rating items on the ATACC questionnaire, the assessments were statistically equivalent between the 4P-models and the 4M-models. On two of the remaining rating items, the behaviour produced by the 4M-models was perceived as more challenging, which we consider to be a positive indicator for the capabilities of dynamic scripting in the air combat domain. Although we have clearly not completely validated the 4M-models in the context of the validation procedure, we must consider that (a) the majority of the assessments of the behaviour were statistically equivalent, and that (b) "degrees of success must be recognized and accepted" (Birta and Arbez, 2013). Therefore, our answer to the research question is that the 4M-models are valid to a moderate extent.

8.2 Answer to the problem statement

In this section we answer the problem statement that was posed in Section 1.3. Our answer is based on the answers to the five research questions discussed in the previous section.

Problem statement: *To what extent can we use dynamic scripting to generate air combat behaviour models for use in training simulations, in such a way that the five challenges of generating air combat behaviour models are met?*

In Chapter 1, we stated five challenges (A-E) that must be met by machine learning techniques in order for the techniques to be considered suitable for use in the air combat domain. Below, we briefly restate the five challenges and declare how dynamic scripting, in combination with the research presented in this thesis, meets each challenge.

Challenge A: Producing team coordination. The use of the CENT coordination method enables CGFs that learn by means of dynamic scripting to coordinate their behaviour and specialise into roles.

Challenge B: Computationally evaluating CGF behaviour. The AA-REWARD reward function evaluates the behaviour of air combat CGFs and provides stable rewards. These rewards lead to a performance increase over the conventional binary reward function.

Challenge C: Efficient reuse of acquired knowledge. We have shown that CGFs that learn by means of dynamic scripting are able to improve their performance in complex scenarios by reusing the knowledge that was built in simpler scenarios.

Challenge D: Validating generated behaviour models. We have developed a validation procedure, and applied this procedure to behaviour models generated by means of dynamic scripting. We have concluded that the models are valid for use in human-in-the-loop simulations to some extent, although there is room for improvement.

Challenge E: Generating accessible behaviour models. This challenge is met by the use of dynamic scripting, as the behaviour models produced by dynamic scripting are in the form of human-readable rules.

Based on our research, our answer to the problem statement is that dynamic scripting greatly facilitates the automatic generation of air combat behaviour models, while being flexible enough to be moulded into answers to the challenges. Challenges (A-C) are met by adapting (a) the rules, (b) the reward function, and (c) the manner in which we apply dynamic scripting algorithm to the air combat domain. Meanwhile, challenge E is met by the design of the dynamic scripting algorithm itself. Still, challenge D remains somewhat open. This challenge is perhaps the most critical one, as the validation of the generated models would mean that we, as machine learning researchers, are confident that the use of our models in training simulations will teach important skills and abilities to the air force pilots of the future. We look forward to the day that challenge D is completely met, but until then this challenge is a good reminder that machine learning can make a difference not only in simulations, but also in the real world.

8.3 Recommendations for future research

Based on the research performed in the thesis, we recommend two areas for future research. They are (1) refinement of the validation procedure, and (2) quantification of the training value of CGF behaviour.

Refinement of the validation procedure: In the thesis, we have developed and applied a validation procedure for air combat behaviour models (see Chapters 6 and 7). However, as we have mentioned, there is no one-size-fits-all solution to the validation of the models, and a large part of the design of the procedure consists of the experiences and opinions of subject matter experts. Therefore, collecting more of these experiences and opinions, and putting them to the test in an empirical manner, may lead to a procedure that has more power to establish the validity of behaviour models. In the future, the validation procedure may also investigate the accessibility of the behaviour models as perceived by the subject matter experts (see, e.g., Fürnkranz, Kliegr and Paulheim, 2018). Such an investigation will allow the comparison of the accessibility of the models that are produced by different machine learning techniques.

Quantification of the training value of CGFS behaviour: The ultimate goal of generating behaviour models for CGFS by means of machine learning is to increase the training value of human-in-the-loop simulations. To this end, it is important to identify how the behaviour of the CGFS influences the learning of air combat concepts by the trainees in the simulations. Once this knowledge can be captured in a reward function (see Chapter 4), it may be possible for the machine learning techniques to optimise the generated behaviour models for the training value for the trainees, rather than towards a concept (e.g., the P_k value of missiles) that act as a proxy for the training value.