Rapid Adaptation of Air Combat Behavior[1]

Armon Toubman [a] Jan Joris Roessingh [a] Pieter Spronck [b]
Aske Plaat [c] Jaap van den Herik [c]

[a] *Netherlands Aerospace Centre NLR, P.O. Box 90502, 1006 BM Amsterdam*
[b] *Tilburg University, P.O. Box 90153, 5000 LE Tilburg*
[c] *Leiden University, P.O. Box 9500, 2300 RA Leiden*

Abstract

Fighter pilots train in simulations to practice real-life missions. However, writing the behavior for the virtual opponents is a costly task. In this paper, we use machine learning to provide the virtual opponents with the ability to adapt their behavior to that of the human trainees. This has two advantages: (1) we support the experts that create the behavior models, and (2) the trainees can be continuously challenged. To this end, we adapted the dynamic scripting technique to generate finite-state machines. The effectiveness of the method is shown through automated air-to-air combat simulations.

1 Introduction

Fighter pilots require continuous training to maintain their level of readiness. Training is mainly done in live exercises where the pilots engage in mock fights. However, due to the limited availability of jets and supporting personnel, a substantial part of training has moved to simulations.

Air combat training simulations require virtual agents that aid or oppose the trainees. While other pilots could be employed to control these agents, they might not be experts on emulating the necessary enemy tactics. Therefore, an AI solution is preferred. Currently, however, training developers still need to write new behavior for each new opponent they need to present, which is a costly undertaking.

We aim to alleviate the problem of writing behavior for the virtual agents in training simulations using machine learning. By automating the generation of behavior, we accomplish two main goals: (1) we support the subject matter experts (SMEs) that develop behavior for the virtual agents, and (2) we are able to challenge trainee fighter pilots with more interesting opponents. However, these goals require a machine learning method that is able to produce behavior models that are readable by the experts (so they can easily inspect and edit the behavior). Furthermore, the method has to be able to provide adaptation rapidly and online (i.e., during training sessions).

In this paper, we present a method for generating air combat behavior for virtual entities in training simulations. This method uses a version of dynamic scripting (DS) (a rule-based reinforcement learning technique) [2] that has been modified to work with finite-state machines (FSMs) (Section 2). We show using automated air-to-air combat simulations that our method is capable of rapidly adapting the behavior of virtual agents to that of their opponents (Section 3).

2 Method

Traditional DS selects pre-written rules from a rule base to form a script. This script is used to control an agent in some environment. Iterative feedback from the environment leads to changes in the probabilities that rules will be selected the next time a script is formed.

Scripts are relatively simple and readable behavior models. However, scripts only provide reactive behavior, and the behavior resulting from combinations of rules is hard to verify. Another behavior

[1]The full paper has been published in *ECAI 2016 - 22nd European Conference on Artificial Intelligence* (IOS Press).

model, the FSM, allows more structured behavior while remaining readable by experts. This is also the model used in Smart Bandits [1], one of the software packages that provides behavior for the virtual opponents in Royal Netherlands Air Force (RNLAF) fighter training.

We modified the DS technique to generate FSMs. States and transitions in FSMs can be easily translated to rules: e.g., for a *Patrol* state, *[if my state is Patrol, fly between points A and B]*. For each state and transition, variations can be made that are interchangeable with the original states and transitions: e.g., a variation on the previous example, *[if my state is Patrol, fly between points C and D]*. This way, a rule base can be filled with rules that are the constituent parts of an FSM, including variations. We altered the DS algorithm such that for each state and transition, only one of its variations can be selected. This way, DS will always generate a completely valid FSM, while still being able to vary its contents.

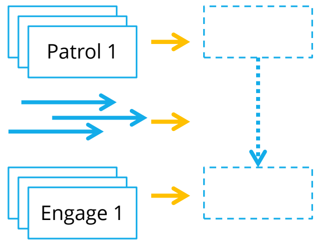

Figure 1: We altered DS to generate FSMs by selecting states and transitions.

3 Experiments and results

We used AI-versus-AI experiments to validate our method. We conducted four experiments in which two adaptive agents (AAs) had to defeat two non-adaptive agents (NAAs). The results of each experiment were averaged over ten runs. In our experiments, we used two tactics (T1 and T2). T1 was based on a tactic originally designed for Smart Bandits. T2 was newly designed to defeat T1.

In experiments 1 and 2, the NAAs used T1 and T2 respectively. The AAs were each given a rule base that contained T1 in rule form, along with several expert-made variations on the states and transitions in that tactic. As a baseline, another pair of NAAs using T1 fought the NAAs.

The results of the first two experiments are shown in Figure 2. In both experiments, the AAs reached (near-)peak performance after around 15 episodes.

Experiments 3 and 4 further demonstrate the AAs' adaptive capabilities. In these experiments, the AAs fought the NAAs using T2 after having adapted to the NAAs using T1, and vice versa.

At the point of the transfer, the win rates of the AAs quickly fall below 0.2, yet recover to >0.5 in 8 episodes. Peak performance is reached again after 35 and 28 episodes in experiments 3 and 4 respectively.

In conclusion, we have generated air combat behavior using machine learning in a way that is both transparent for the SMEs developing the virtual opponents for training simulations (i.e., with behavior in FSM form), and fast enough for use during training sessions. Human-in-the-loop experiments with RNLAF fighter pilots are currently underway.

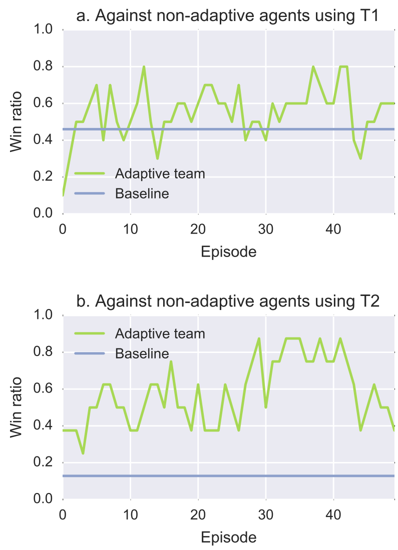

Figure 2: Win rates of the adaptive agents.

References

[1] Jan Joris Roessingh, Robbert-Jan Merk, Pieter Huibers, Remco Meiland, and Roel Rijken. Smart Bandits in air-to-air combat training: Combining different behavioural models in a common architecture. In *21st Annual Conference on Behavior Representation in Modeling and Simulation*, Amelia Island, Florida, USA, 2012.

[2] Pieter Spronck, Marc Ponsen, Ida Sprinkhuizen-Kuyper, and Eric Postma. Adaptive game AI with dynamic scripting. *Machine Learning*, 63(3):217–248, 2006.