Virtual self-learning enemies keep pilots on top of their game

Nieke Roos (Techwatch)

The Royal Netherlands Air Force is increasingly relying on computer simulators to keep its pilots mission ready. With machine learning, their virtual enemies can be made even smarter and training becomes a lot more effective, as Armon Toubman, who completed a PhD project on the topic at NLR, explains in an interview. The goal is to prevent the simulations from becoming a boring and repetitive pastime and keep the virtual dogfight challenging.

Armon Toubman is a graduate in artificial intelligence from the Vrije Universiteit Amsterdam. He obtained his PhD in February 2020, after which he joined Royal NLR.

The Royal Netherlands Air Force has seen its fleet thin out over the years. Of the 213 F-16s available to the flying army in 1992, only 68 remained in 2019. The plan is to replace them in 2023 with the first batch of 37 F-35s. Fewer aircraft also means fewer opportunities to train pilots in the real world. At the same time, the planes have become a lot smarter – the F-35 is a flying sensor – but also more complex, which only increases the importance of training. The Air Force is therefore increasingly relying on computer simulators to keep its pilots mission ready and on the Royal Netherlands Aerospace Centre (NLR) to further develop these simulators. Such a simulator consists of a seat for the pilot, a joystick for the controls, a small display with a dashboard and a large screen with a view to the outside. In the simulated airspace, there are also opponents flying around. These may be human opponents, but due to the limited availability of pilots, it's usually so-called computer-generated forces that have to provide real-life challenges. Modelling and programming the behaviour of these virtual entities is still largely manual work that takes a lot of time and requires considerable specialised knowledge. As a result, there are only a few models and the simulators can't be fully exploited. In his PhD project at NLR, Armon Toubman investigated how machine learning can offer a solution here.

More variable behaviour

'It comes down to writing rules on how such a virtual plane should behave', explains Toubman, a graduate in artificial intelligence from the Vrije Universiteit Amsterdam. 'Creating and testing these scripts each time requires quite a development effort. What's more, such a script always makes the plane move in the same way. If you, as a pilot, have to do the same trick every time, training soon becomes boring and ineffective. With machine learning, we want to give the opponents more erratic behaviour within the framework of the mission.'

Machine learning is an application of artificial intelligence in which systems automatically solve complex problems by learning new behaviour based on examples, without using explicitly programmed rules.
From the machine learning toolbox, Toubman has chosen the dynamic scripting technique. 'To put it simply, you have a box filled with small pieces of behaviour, which are combined with the technique to find the optimal configuration. The result is a behavioural model in which you can still see the constituent pieces. This is one of the big differences with a technique like deep

In a simulation, so-called computer-generated forces have to provide real-life challenges PHOTO NLR

learning, where a complete but opaque model is learned from scratch. This transparency was one of the requirements for my project. We want to keep control over the generated models so that we can still play with them. Training opponents, for example, shouldn't be perfect but should make mistakes every now and again as real people do – we want to be able to put in those mistakes ourselves.'

'The behavioural model is used to determine the state of the simulator at any given moment', details Toubman. 'Based on that, the computer calculates what action the enemy aircraft should perform. This action goes back into the simulator, which changes state as a result. The new state is fed back to the model, thereby closing the loop. You can choose when this loop ends – for example, after expiration of a set time limit or when the entire enemy team has been eliminated.'

Fixed pattern

At the end of each simulation, the virtual opponent is automatically awarded points based on the course of the battle. 'These points do not go to the model as a whole, but to its constituent pieces of behaviour', clarifies Toubman. 'The more a building block has contributed something useful, like bringing down an aircraft, the more points it receives. The more points, the greater the chance that it will be chosen for the next model. In this way, dynamic scripting gradually arrives at the optimum combination of building blocks.'

Toubman has executed a large number of runs of around 150 battles at a time. 'Not with real pilots', he says, laughingly. 'On an NLR server, I've run simulations faster than real-time with a 'red' learning team and a 'blue' team following a fixed pattern – like the scripts now used in the simulator. Before taking on real pilots, I thought it wise first to see if I could do better than the existing models. And that turned out to be the case: red emerged as the winner in 80 per cent of the battles.'

Toubman then tested his models in real life. 'In an F-16 simulator [the Fighter 4-Ship] I had sixteen pilots in teams of four fly against

PHOTO NLR

In Smart Bandits, it will be possible to generate behavioural models not only for whole battles but also for individual manoeuvres such as choosing the target or the best escape route

generated models and I recorded that. I did the same with existing, handwritten models. I presented these videos to five instructors, together with a questionnaire I designed. Their assessment confirms that the generated models are indeed at least as good as the handwritten ones.'

Smart Bandits

Toubman obtained his PhD last February, after which he joined NLR. 'The software I developed during my research has a limited scope. Together with a small team, I'm now adapting it for general use and integrating it with NLR's Smart Bandits environment. In Smart Bandits, you can define the behaviour of the enemies – 'bandits' in jargon. Currently, this is done manually, by drawing boxes and arrows, but we're expanding the environment so that we can also integrate various machine learning algorithms, including my implementation of dynamic scripting.'

In Smart Bandits, it will be possible to generate behavioural models not only for whole battles but also for individual manoeuvres such as choosing the target or the best escape route. 'If you want to find the optimum angle to fire a rocket, you only want to be able to do that part with machine learning and not everything around it', Toubman says. 'And you want to be able to supplement generated behaviour with manual input. Having that flexibility is very useful, also for research purposes.'

In the next step, Toubman wants to put an intuitive user interface on top of Smart Bandits, to increase the ease of use. 'A layer hiding the actual modelling and allowing you, as an operator, to quickly choose and adjust behaviour, for example when the computer needs to do nothing – currently, this is quite a hassle. We're looking into such an interface.' ■

'Training opponents shouldn't be perfect but should make mistakes every now and again as real people do'