NEWS & TECHNOLOGY 21 June 2017

DeepMind now learns from human preferences – just like a toddler

By **Chris Baraniuk**

ARTIFICIAL intelligence systems still need far too much hand-holding by their human masters. Now research from DeepMind and OpenAI suggests a mere nudge here and there can be enough to help AIs accomplish tricky tasks.

The team set up a series of trials in which human participants saw two clips of an AI's approach to a task. The AI did not know the desired outcome of the task. The participants were then asked to make a snap judgement about which clip showed more promising progress.

One scenario involved the AI learning to play *Space Invaders*, another involved a virtual robot learning to do backflips.

Importantly, the humans were non-experts who were simply asked to judge the clips at face value.

The human responses were used to train a part of the AI system called a reward predictor, which in turn trained the AI character that was performing the task. Over time, the agent learned how to maximise the reward and improve its behaviour in line with the humans' preferences (arxiv.org/abs/1706.03741).

Usually, humans would have evaluated progress every step of the way – which would be time-consuming. With the predictor, the AI learned to perform a perfect backflip in under an hour of the evaluators' time. According to Dario Amodei at OpenAI, this would have taken more than 100 times longer before.

Up until now, reinforcement learning systems required a hard-coded reward function to work out what the problem was they had to solve, but this technique removes that necessity.

"The AI learned to perform a perfect backflip in under an hour of the evaluators' time"

The research shows how AI is "growing up", says Pedro Domingos at the University of Washington. He says that DeepMind's previous systems were like infants, trying random things until they got a reward. "This system is more like a toddler, still trying random things, but occasionally getting some feedback from its parents and learning from it."

This article appeared in print under the headline "AI that learns like a toddler makes progress"

Magazine issue 3131, published 24 June 2017