

# AI learning technique may illustrate function of reward pathways in the brain

January 17 2020, by Bob Yirka



When the future is uncertain, future reward can be represented as a probability distribution. some possible futures are good (teal), others are bad (red). Distributional reinforcement learning can learn about this distribution over predicted rewards through a variant of the TD algorithm. Credit: *Nature* (2020). DOI: 10.1038/s41586-019-1924-6

A team of researchers from DeepMind, University College and Harvard University has found that lessons learned in applying learning techniques

to AI systems may help explain how reward pathways work in the brain. In their paper published in the journal *Nature*, the group describes comparing distributional reinforcement learning in a computer with dopamine processing in the mouse brain, and what they learned from it.

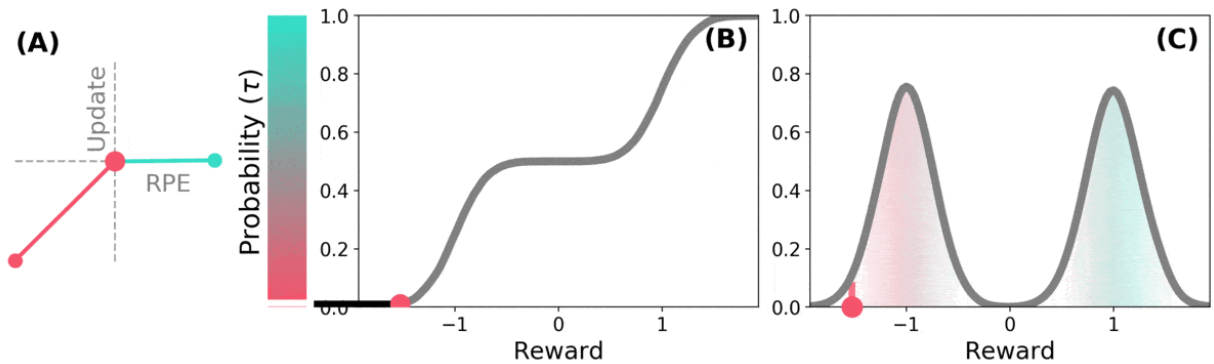
Prior research has shown that dopamine produced in the brain is involved in reward processing—it is produced when something good happens, and its expression results in feelings of pleasure. Some studies have also suggested that the neurons in the brain that respond to the presence of dopamine all respond in the same ways—an event causes a person or a mouse to feel either good or bad. Other studies have suggested that neuronal response is more of a gradient. In this new effort, the researchers have found evidence supporting the latter theory.

Distributional reinforcement learning is a type of machine learning based on reinforcement. It is often used when designing games such as Starcraft II or Go. It keeps track of good moves versus bad moves and learns to reduce the number of bad moves, improving its performance the more it plays. But such systems do not treat all good and bad moves the same—each move is weighted as it is recorded and the weights are part of the calculations used when making future move choices.

Researchers have noted that humans appear to use a similar strategy to improve their level of play, as well. The researchers in London suspected that the similarities between the AI systems and the way the [brain](#) carries out reward processing were likely similar, as well. To find out if they were correct, they carried out experiments with mice. They inserted devices into their brains that were capable of recording responses from individual [dopamine](#) neurons. The mice were then trained to carry out a task in which they received rewards for responding in a desired way.

The mouse neuron responses revealed that they did not all respond the same way, as prior theory had predicted. Instead, they responded in

reliably different ways—an indication that the levels of pleasure the mice were experiencing were more of a gradient, as the team had predicted.



Distributional TD learns value estimates for many different parts of the distribution of rewards. which part a particular estimate covers is determined by the type of asymmetric update applied to that estimate. (a) A ‘pessimistic’ cell would amplify negative updates and ignore positive updates, an ‘optimistic’ cell would amplify positive updates and ignore negative updates. (b) This results in a diversity of pessimistic or optimistic value estimates, shown here as points along the cumulative distribution of rewards, that capture (c) The full distribution of rewards. Credit: *Nature* (2020). DOI: 10.1038/s41586-019-1924-6

**More information:** Will Dabney et al. A distributional code for value in dopamine-based reinforcement learning, *Nature* (2020). DOI: [10.1038/s41586-019-1924-6](https://doi.org/10.1038/s41586-019-1924-6)

© 2020 Science X Network

Citation: AI learning technique may illustrate function of reward pathways in the brain (2020, January 17) retrieved 3 February 2023 from <https://techxplore.com/news/2020-01-ai-technique-function-reward-pathways.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.