**Pi2Go Simulator Programming: Learning Rate**

**AIM:** After completing this worksheet you should be able to explain how to use a learning rate to gradually move a machine learning program from exploring behaviours to selecting behaviours that are likely to yield a reward.

**You Need:** To complete this worksheet you need to have a virtual Pi2Go simulator (see WS1), understand how to control the robot’s motors and sensors (WS3&WS4), be able to use files to store Programs (WS5), use control structures (WS7-10), Data Types (WS12-14), functions (WS16), the time module (WS6) and have completed WS23.

**If the simulator isn’t already running: Start the Simulator, Select the Pi2Go Simulation and oval.xml, then start IDLE (open a *new IDLE window* if you have used IDLE to start the simulator).**

Your machine learning program currently calculates the average reward of an action over all attempts. However, more commonly, machine learning program use a number, called the *learning rate* to change the stored reward after each attempt. In this case we calculate the *difference* between the currently stored reward and the new reward (by subtracting the current reward from the new reward) and the modify the reward by this difference multiplied by the learning rate.

So, for instance, if the currently stored reward is 1.5, the new reward is 1.4 and the learning rate is 0.2 then we first calculate the difference between the two rewards

Difference = 1.4 – 1.5 = -0.1

Multiply this by the learning rate

Modified difference = -0.1 x 0.2 = -0.02

And then add this to the current reward

New reward = 1.5 + ( - 0.02) = 1.48

We can represent this as an equation where *R* is the current reward, *r* is the new reward and ** is the learning rate then:

R = R + (r – R) x 

**Exercise:** Replace the use of average reward in your program from WS23 with a learning rate of 0.5. Change the default value of all rewards from 0 to 1, since when working with a learning rate it is good start with a reward value somewhere in the middle of the range



University of Liverpool, 2019

This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/).