REINFORCEMENT LEARNING
CONNECT FOUR - HYPER-PARAMETERS TUNING

Hanan Benzion   026606608
Idan Azuri   302867833

January 29, 2018

Note

We assigned our project in one day delay (in the first time we assigned it on time). The reason was we found a bug in the sampling method by mistake we sampled every round the game. So we fixed this bug and re-assigned it. We will appreciate it if you can to ignore the one day delay because it was an engineering bug and not algorithmic.

Changes in the code

(1) The main thing we changed was the **exploration** rate: we changed the epsilon start from 100% and every 500 round we reduce it by 2 until 10%. We did it because we saw that the model tented to predict the same action - probably because it stacked on a local minima or on a saddle point.

(2) Batch size changed from 8 to 512.

(3) We reduced DB size to 500k from 1M in order to reduce timeouts in the late steps of the training.

(4) Additional changes:
   (a) Cleaning comments, log messages ,unused code, and fixing typos.
   (b) Adding a log message on model loading and 'load_from' to parse args.