Reinforcement Learning
Name: Eli Andrew

Exercise 1.1: Self-Play

(a) **Question:** Suppose, intead of playing against a random opponent, the reinforcement learning algorithm described above played against itself, with both sides learning. What do you think would happen in this case? Would it learn a different policy for selecting moves?

- The random opponent described in the example is assummed to play with a constant policy. So, the first change in this new scenario is that the algorithm is playing against an opponent that changes its policy over time. Now that the opponent will learn from its experience and update its policy, both players will be simulataneously adjusting their policy based on what the other is doing.