Task description

Run few replications of the experiment by using different seeds (integer numbers). You can use the pre-prepared acrobot.ini file included in the ./xacrobot directory. While the program is running check the source code of the environment available from the https://gym.openai.com/envs/ website to figure out the content of the observation vectors, the content of the action vector, and the way in which the reward is calculated. Plot performance across generations and then observe the behavior of evolved robots.

Solution

Issue

An error occurred in program file "policy.py", especially in the definition of Class Object GymPolicyDiscr at the 348 row which is showed in figure 1. Instead of $rew+ = rew$, the following equation should be written

$$rew+ = r \tag{1}$$

```
while t < self.maxsteps:
    # Copy the input in the network
    self.nn.copyInput(np.float32(self.ob))
    # Activate network
    self.nn.updateNet()
    # Convert the action array into an integer
    action = np.argmax(self.ac)
    # Perform a step
    self.ob, r, done, _ = self.env.step(action)
    # Append the reward
    rew += rew
    t += 1
    if render:
        if (self.test == 1):
            self.env.render()
            time.sleep(0.05)
```

Figure 1: Issue in file \opt\evorobotpy\bin\policy.py

Results

I change initial parameter in file "xacrobot.ini". I set maxmsteps = 10 and take the following results on different seeds:

```
Seed 1 gen 239 eval 10038000 bestfit -200.00 bestgfit -200.00 centroid -200.00
bestsam -200.00 avg -200.00 weightsize 0.26 runtime 3313.01
```

`Seed 3 gen 342 eval 10008989 bestfit -91.80 bestgfit -94.80 centroid -117.00 bestsam -95.20 avg -106.87 weightsize 0.30 runtime 3265.43`

`Seed 5 gen 245 eval 10014418 bestfit -126.00 bestgfit -130.40 centroid -148.00 bestsam -133.40 avg -161.17 weightsize 0.30 runtime 3301.66`

`Results of variation of perfomance you can see in figure 2.`

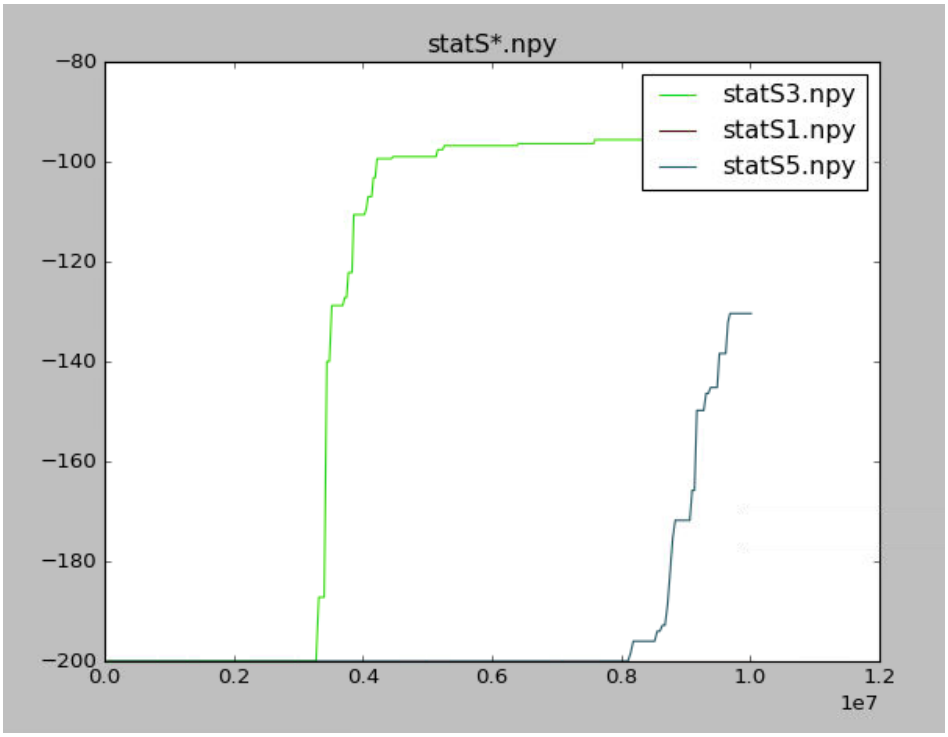

Figure 2: The graph of the variation of performance across generations