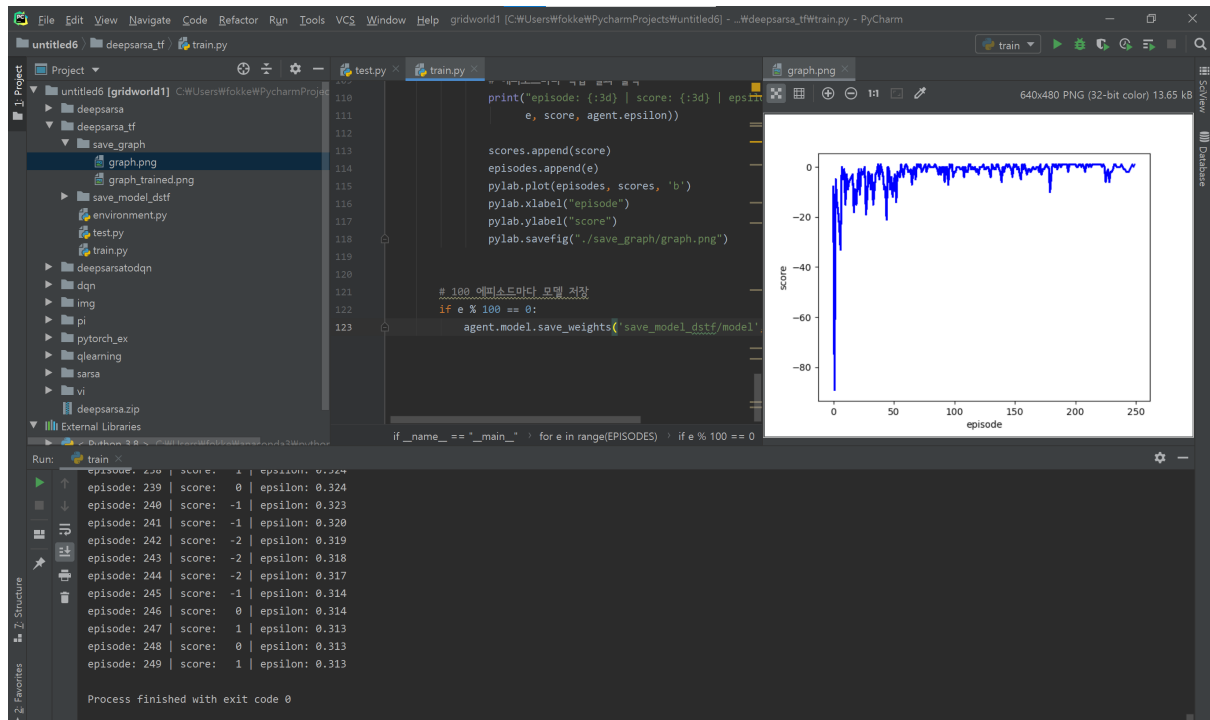
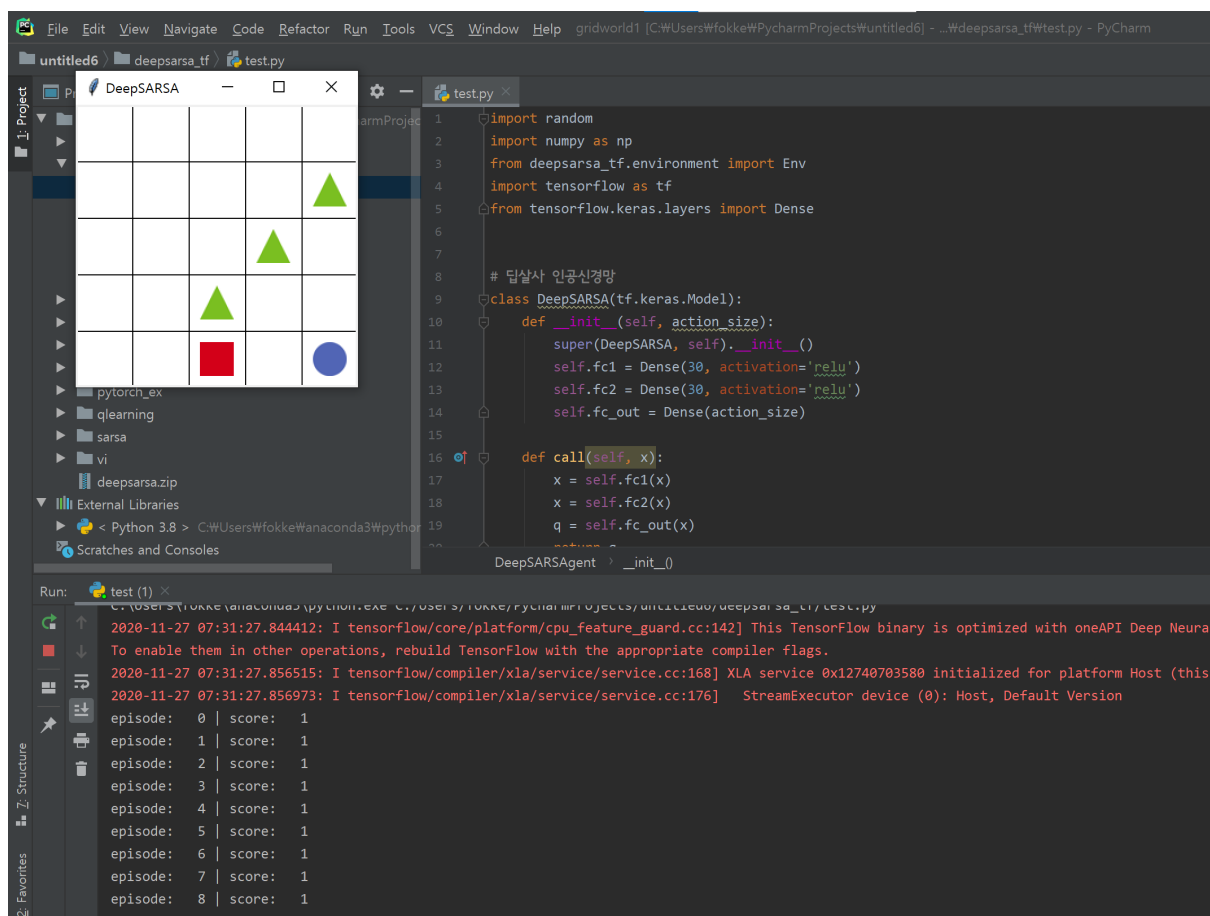


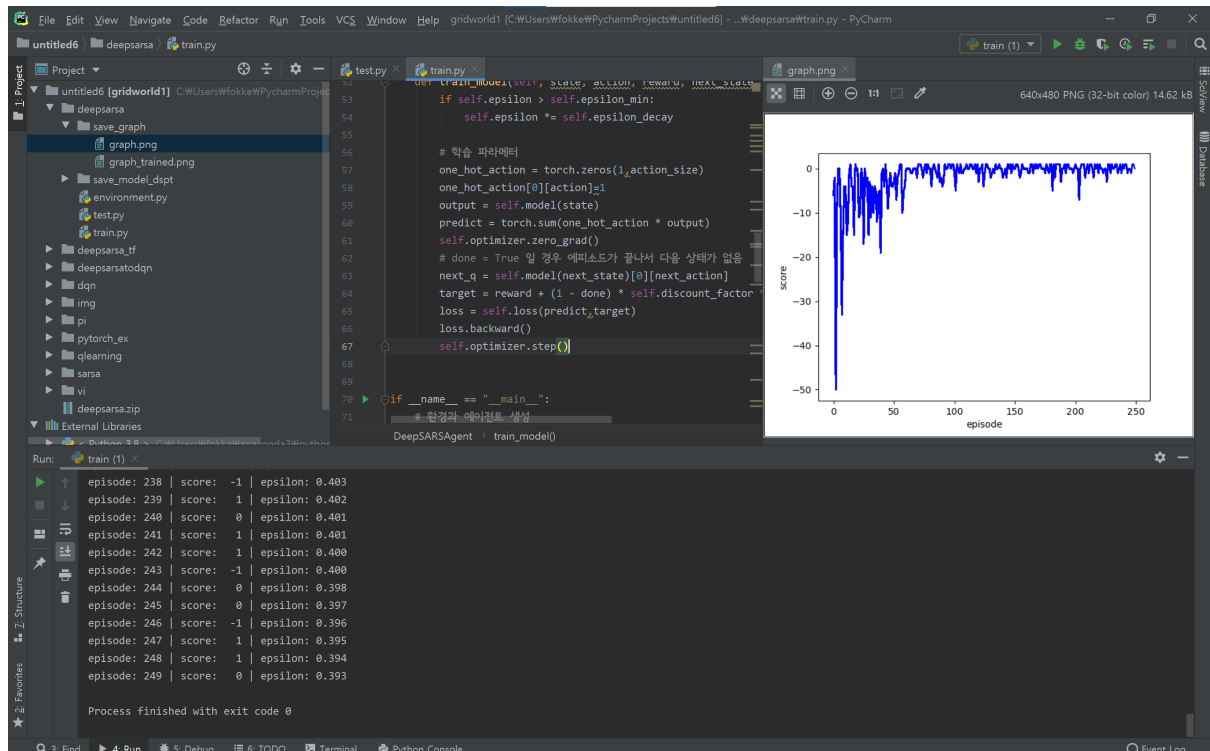
Gridworld예제를 TensorFlow를 이용한 DeepSARSA로 학습시켜보았다



성공적으로 잘 학습된 것을 볼수있다.

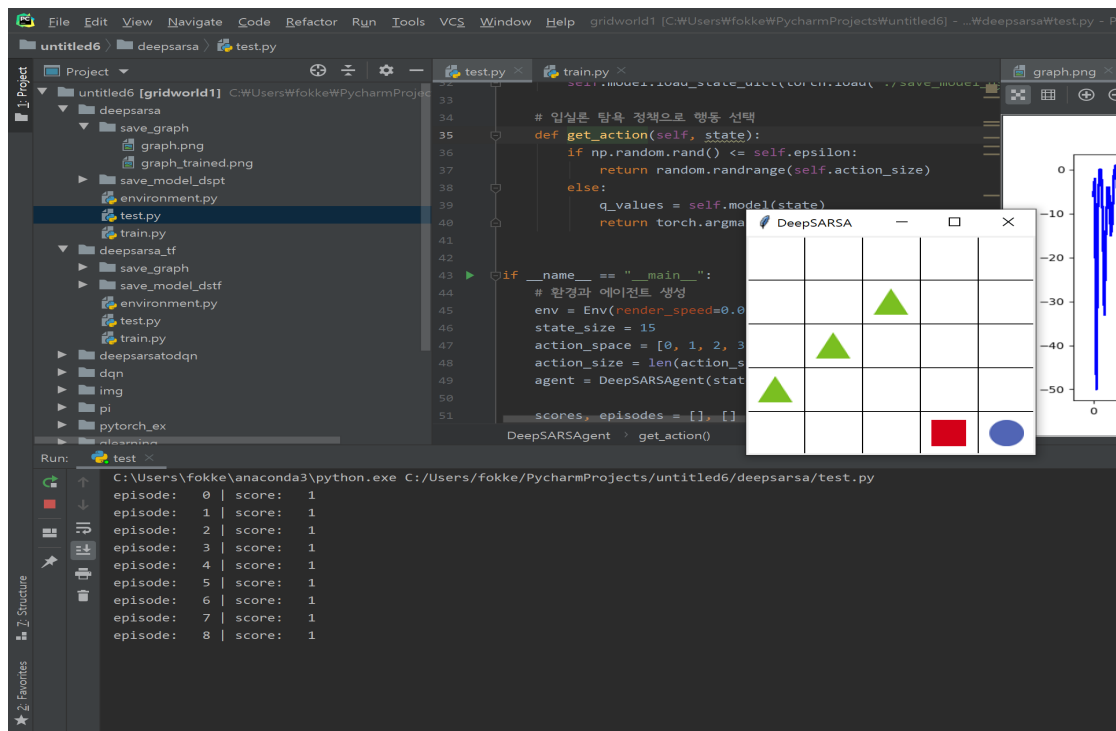


이를 pytorch로 바꾸어서 train 해보았다.

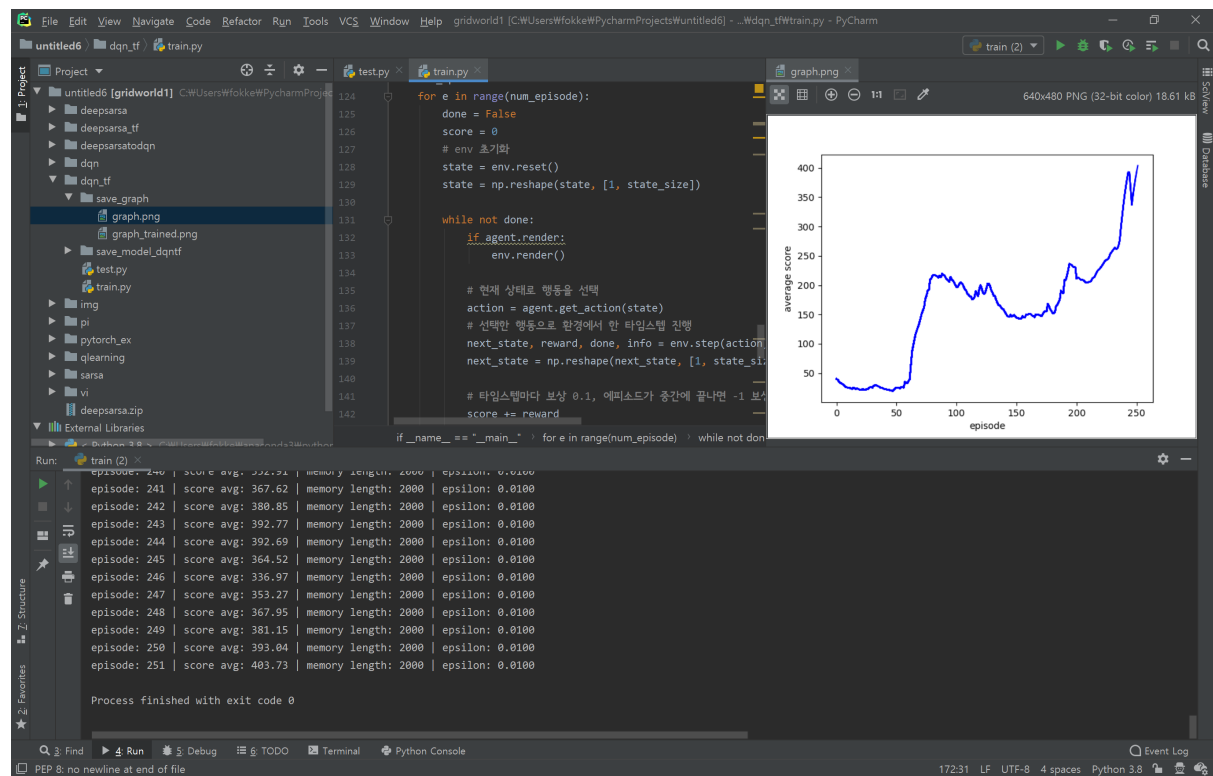


파이토치로 구현한 결과 그래프가 tensorflow로 구현했을 때보다 조금 더 튀었다. 하지만 성공적으로 잘 학습됨을 확인하였다.

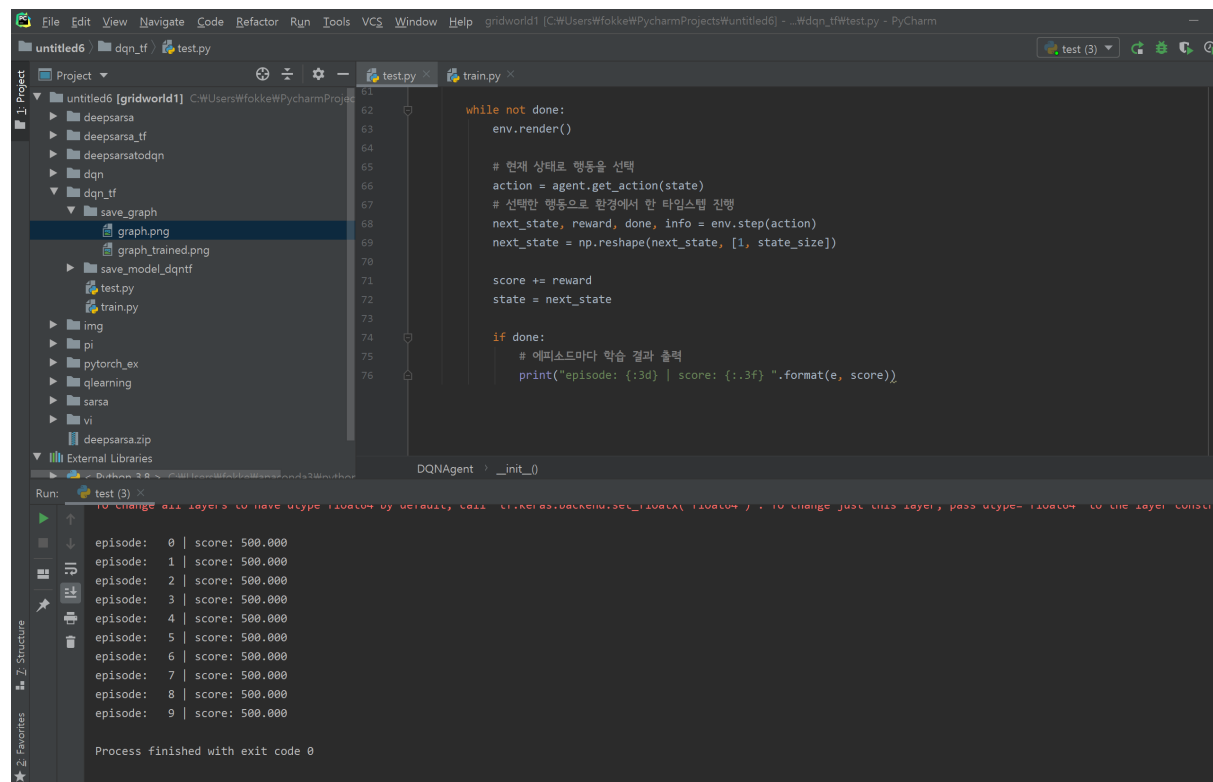
test결과도 잘나오는 것을 확인할 수 있다.



다음은 cartpole예제를 tensorflow를 이용한 DQN으로 학습시켜보았다.

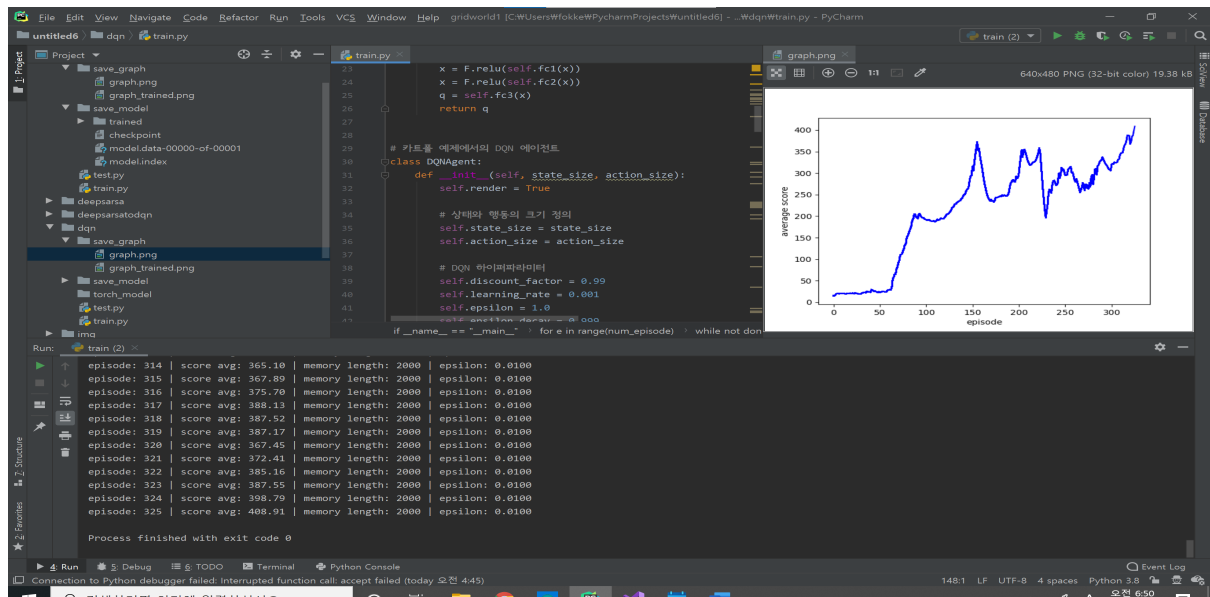


보면 완벽하게 avg score가 episode 250 만에 400을 넘어서 학습이 끝난 걸 확인하였다.



Test 실행결과도 잘나왔다.

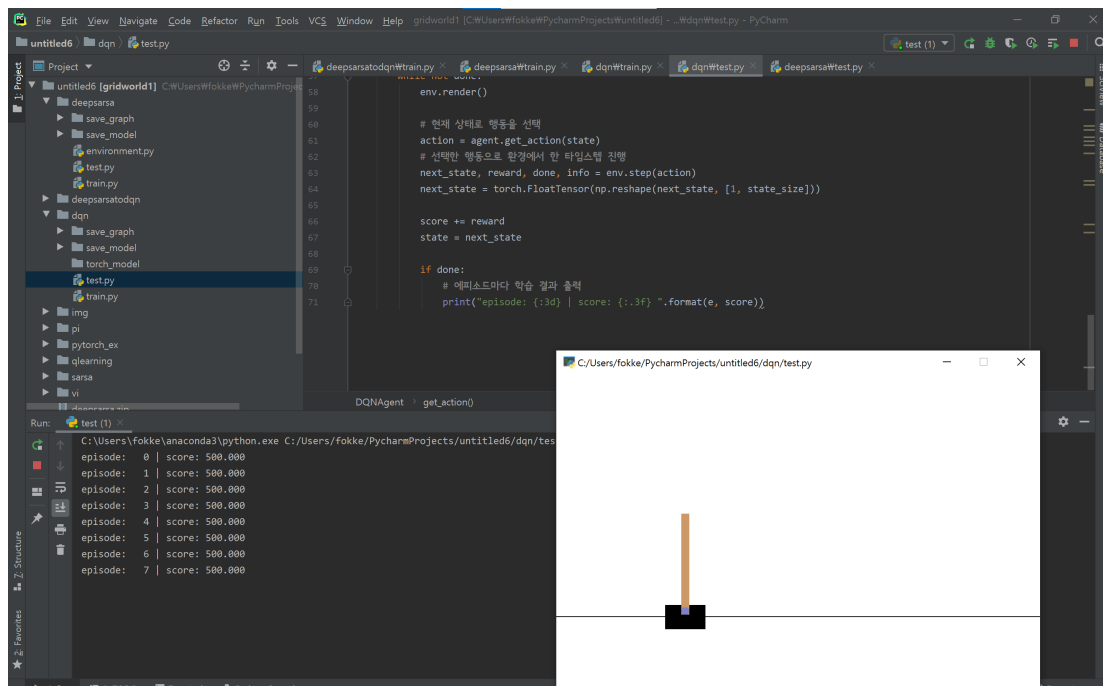
Pytorch로 구현한 DQN의 cartpole 예제 model의 score graph와 실행결과이다



325번의 episode만에 400점을 넘겨 프로그램이 종료되고 모델이 학습이 끝이 난 걸 볼 수있다.

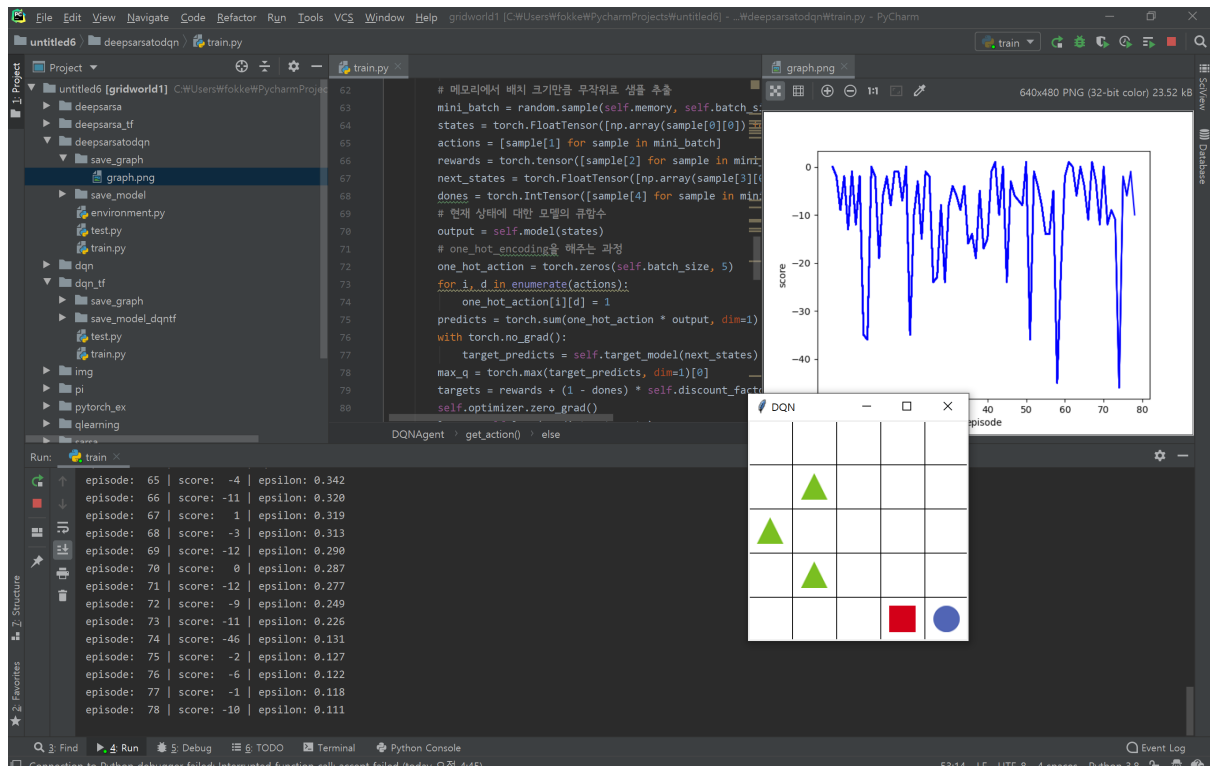
이후에 여러 번 더 학습해본 결과 운이 좋다면 230번만에도 400을 넘기는 것을 확인하였다.

스스로 구현한 pytorch가 더욱 성능이 좋아보였다.



위에서 학습한 모델로 test프로그램을 돌렸을 때도 잘 학습된 모델임을 확인할 수 있다.

다음은 gridworld의 파이토치를 활용하여 deepsarsa를 dqn방식으로 바꾼 모델의 train결과이다.



Dqn을 잘 구현했지만 어째서인지 agent가 target을 잘 찾아가지 못해 episode가 끝나지 않고 지속됨을 확인하였다. 이를 해결하기 위해 batchsize, memorysize, learning rate등을 바꾸어보았지만 해결이 되지 않았다. 이 문제는 아직까지 해결하지 못하였다. 이유를 생각해보면 안좋은 상황으로 빠질 확률이 너무 많기 때문인 것 같다. 따라서 memory size를 늘려주고 start training값도 늘려주었지만 해결이 되지 않았다.