**Exploring OpenAI Gym Environments**

**CartPole-v1**

In CartPole environment a pole is attached to a cart via an un-actuated joint and the cart moves on a frictionless surface. The goal of the agent is to maintain the pole attached to it from falling. The agent can apply a force of +1 or -1 to the cart to keep the pole upright. For each timestep the pole is upright then the agent gets the reward of +1. The episode ends when the pole is more than 15 degrees from the vertical or the cart on which the pole is mounted moves more than 2.4 units from the starting position.

**State / Observation Space:**

The state space is of type box which has 4 observations in it:

1. Cart Position
2. Cart Velocity
3. Pole Angle
4. Pole Angular Velocity at the tip

The cart position can take values from -4.8 to +4.8. The velocity of the cart can take values from -infinity to +infinity. The pole angle can take value from -24 degree to +24 degree. The angular velocity of the pole can range from -infinity to +infinity. As we can see the state space of the cart pole environment is continuous.

**Action Space:**

There are two possible actions that the agent can do in the given state:

1. 0 (push the cart to the left)
2. 1 (push the cart to the right)

**Rewards:**

The reward space for this environment consists of only one reward:

1. +1

The agent gets a reward of +1 for every timestep. This also includes the terminal step taken by the agent.

**Starting State:**

For the starting position of the environment all observations of the environment are assigned a uniform random value in -0.05 to +0.05

**Termination:**

The episode is terminated when one of the following conditions occurs:

1. Pole angle is more than 12 degrees.
2. Cart position is more than 2.4 units from the center.
3. Episode length is greater than 200.

If the pole angle is 12 degrees from the starting vertical position, then the episode terminates. The episode also terminates if the cart moves 2.4 units from the center. At this timestep the cart moves out of the screen. Therefore, the episode terminates because the cart moves out of the display. If the total number of timesteps in the episodes exceeds 200 then episode terminates.

**Solved Requirements:**

The agent is said to be learnt if the average returns of the rewards during 100 consecutive episodes is more than or equal to 195.

**Exploring OpenAI Gym Environments**

**BreakoutDeterministic-v4**

Atari breakout is a classic arcade game in which we have to clear all the bricks in a given screen. There are 8 rows of bricks on top of the screen/image, and a paddle at the bottom that deflects a ball to remove the bricks without losing its track/missing the rebound. The agent must toggle the paddle such that all the bricks break by directly deflecting the balls to the bricks or via the side wall. If the agent misses to deflect the ball beyond a specific number of tries (3 generally), the episode ends.

**State / Observation Space:**

We are given an image in with dimensions {210,160,3}, which is an RGB image. We sample k frames (4 in general), stack them sequentially to get the exact information about the state in which the agent is in. For example, the ball is going up and gets deflected after breaking a brick, it goes down. We cannot represent this in a single image. It must be represented as a series of images so that the agent understands the exact space.

**Action Space:**

There are four possible actions that the agent can do in the given state:

1. Left (shift paddle to the left)
2. Right (shift paddle to the right)
3. Noop (don’t shift paddle)
4. Fire (for starting and ending episodes)

**Rewards:**

The reward space for this environment consists of one reward:

1. +1 for each step in .

**Starting State:**

The agent starts with all bricks intact.

**Termination:**

The episode is terminated when one of the following conditions occurs:

1. All bricks are broken.
2. The agent misses to deflect the paddle beyond a specified limit.

**Solved Requirements:**

The agent is said to be learnt if manages to break all the bricks once. There is no specific reward threshold for actual definition of the “solved Atari”.