AI Presentation: Training agent for first-person shooter (FPS) game with Actor-critic Curriculum learning.

Introduction:

* What is FPS game?

First person shooter game is a sub-genre of shooter game in which a player fights against other computer control agents or human players by using 3D environment.

We are introducing a new framework for a training agent to act more advanced than human being and using reinforcement learning approach actor critic model.

* How is it related to reinforcement?

Reinforcement learning is purposed to use for a training agent in order to maximizing or minimizing future rewards/penalties on the basis of given environment. Actor critic learning is a reinforcement learning technique.

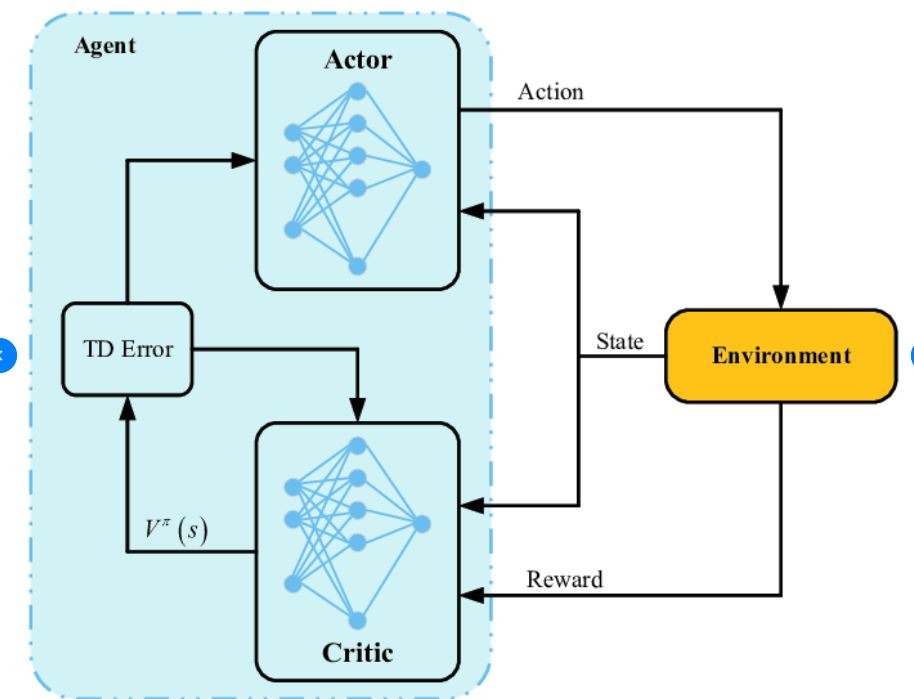
**Assalamualaikum, This is Sazzad here …thanks monjurul robin for your intro. I would like to present about.. Training agent for first-person shooter (FPS) game with Actor-critic Curriculum learning.**

* **Actor critic model:**

In Actor critic learning technique Two type of functions playing important roles: a value function that gives the expected reward of the current state, and a policy function that gives a probability distribution on the candidate actions.

TD error (**temporal difference)** improves stability by avoiding large steps in the actor update whenever the critic is highly inaccurate.

This shows the general flow of state observations and rewards signals between the algorithm and the environment, the critic update and its value estimate , which is used by the policy gradient update.



**Flatmap : GIGtrack:**

Two maps we used in the paper. FlatMap is a simple square containing four pillars . CIGTrack1 is the map used in Track1. Black dots are items (weapons, ammo, medkits, armors, etc).

the whole picture. To reduce the correlation of game experience, Asynchronous Advantage Actor Critic Model runs independent multiple threads of the game environment in parallel. These game instances are likely uncorrelated, therefore their experience in combination would be less biased.

While keeps a separate model for each asynchronous agent and perform model synchronization once in a while, we use an alternative approach called Batch A3C, in which all agents act on the same model and send batches to the main process for gradient descent optimization. The agents’ models are updated after each gradient update. Note that the contemporary work GA3C also proposes a similar architecture. In their architecture, there is a prediction queue that collects agents’ experience and sends them to multiple predictors, and a training queue that collects experience to feed the optimization.

This is my rest of presentation and next part will present rawnak Jahan taifa. Thank you.

**DOOM AS A REINFORCEMENT LEARNING PLATFORM:**

Doom is a first person shooter game released in 1993

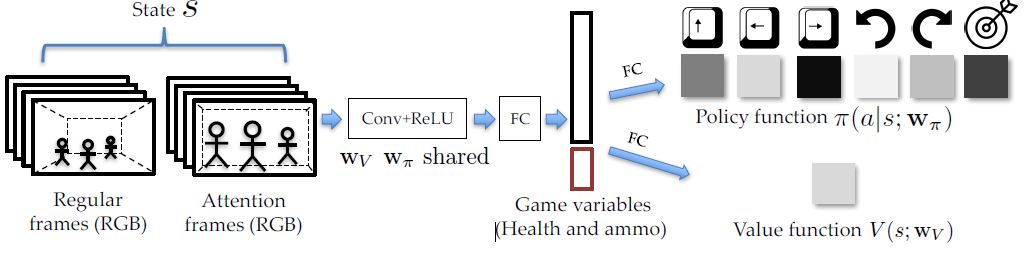
* Agent make decisions considering the environment.
* The doom game is modified by VizDoom.

In doom games the player control the agent to fight against enimies in 3D environment. The agent makes decisions on received partial information and view point of environment. The doom game is modified to viz doom which offer more fexivility including reach senarios,game variables,build in boots,evalution critaria.

* What are the steps are following for completing FPS game?

There are a number of methods which are used in order to complete in FPs game and they are-

**Network architecture :**



Network architecture is used for extract features from game frames and represents the output combined with game variables. In the following features the network architecture take the frame as the input and the output comes with two branches; policy function, value function. For input frames the center part are used of the same size frames, thus the center frames has higher regulation the policy network comes with six action move forward, move left, move right, turn left ,turn right and attack These includes the agents health (0-100) and ammo (how many bullets left). value function V (s) that gives the expected

reward of the current state s,Network architecture is related to Ai also game environment of training ,tasting and vizdoom AI competition.

**Training pipline:**

In traning pipline there are implementation of training procedure by using parameters for different maps.The main process collects different frames from different game instances to create batches

w\_ ← w\_ + α(Rt − V (st))∇w\_ log π(at|st) + β∇w\_H(π(・|st))

wV ← wV − α∇wV (Rt − V (st))2

By using these equations optimization on batches are done asynchronously. This parameter are used to accommodate agents training procedure. for example learning rate policy learning rate , discount factor etc. are set at an initial value to train agent from scratch. This training procedure requires millions of batches for multiple different scenarios.

**Curriculum learning:**

Curriculam learning includes three segments the first one is Reward Shaping. Where the agents reward are recorded according to the games activity for example if an agent kills an enemy he will receive (+1) as reward and receive (-1) while death.here als includes health loss, amo loss, health pick up amo pick up,dist\_panalty, dist\_reward.

Curriculum Design:

Here bot is trained on flat map so that it can quickly adapt more complicated maps .for each maps curriculam design are made considering the bot variation.for example moving speed,initial weapons,initial helth.

Adaptive curriculum:

Here is a probability distribution is used on different levels .probality distribution depends on the agents performance and shifts to either towards more difficult or mre easier level.here the agent perform well if the frag count is grater than 10 points.

Post-training Rules:

To examine the viz doom model here some ablation analysis are uses which are

Visualization and Internal tournament. Post training rules improve performance of a boot and train the boot for future work. These includes internal tournament Behaviors and Tactics.

Conclusion:

This paper introduces design methodology of a new frame work which helps to train a strong AI in first person shooter game like Doom game. The training agent will be able to act properly to any complicated challenges or task in 3D environment. There are some limitations in this framework. The AI bot is a reactive agent which only remember last four frames and act according to them. Bot should be able to build a map itself in an unknown environment also can localize itself. There is a scope to improve the framework in future.

* Design methodology of a new framework
* Challenges for training an agent
* Limitations and planning.

Referances:

Thank You

[robinmonjurul6@gmail.com](mailto:robinmonjurul6@gmail.com)

sazzad01794@gmail.com