

basic.cfg:

This is a super easy map that can be trained in 100k steps, although I keep the 500k and 1 million step models in the repo in: ...\\Initial Testing With Reinforcement Learning\\Training\\checkpoints\\best\_model\_PPO\_test\_basic\_2

Interestingly the 1 million step model is a lot worse than the 500k model. I do not know why.

Generated while using 2048 n steps, and the fitness function given in the [ViZDoom documentation](#):

## BASIC

The purpose of the scenario is just to check if using this framework to train some AI in a 3D environment is feasible.

The map is a rectangle with gray walls, ceiling, and floor. The player is spawned along the longer wall in the center. A red, circular monster is spawned randomly somewhere along the opposite wall. A player can only (config) go left/right and shoot. 1 hit is enough to kill the monster. The episode finishes when the monster is killed or on timeout.

### REWARDS:

- +106 for killing the monster
- -5 for every shot
- +1 for every tic the agent is alive

The episode ends after killing the monster or on timeout.

### CONFIGURATION:

- 3 available buttons: move left/right, shoot (attack)
- 1 available game variable: player's ammo
- timeout = 300 tics

**Gymnasium/Gym id:** "VizdoomBasic-v0"

**Configuration file:** [basic.cfg](#)

defend\_the\_center.cfg:

This is by far the one I tested the most,

...\\Initial Testing With Reinforcement Learning\\Training\\checkpoints\\best\_model\_PPO\_test\_defend\_the\_center\_16\\best\_model\_925000.zip

is the best model of the ones tested. It was done by first training it on the fitness function found in the [ViZDoom documentation](#) for 1 million steps with 8192 n steps:

## DEFEND THE CENTER

The purpose of this scenario is to teach the agent that killing the monsters is GOOD and when monsters kill you is BAD. In addition, wasting ammunition is not very good either. Agent is rewarded only for killing monsters so he has to figure out the rest for himself.

The map is a large circle. A player is spawned in the exact center. 5 melee-only, monsters are spawned along the wall. Monsters are killed after a single shot. After dying, each monster is respawned after some time. The episode ends when the player dies (it's inevitable because of limited ammo).

### REWARDS:

- +1 for killing a monster
- -1 for death

### CONFIGURATION:

- 3 available buttons: turn left, turn right, shoot (attack)
- 2 available game variables: player's health and ammo
- timeout = 2100
- difficulty level ( `doom_skill` ) = 3

**Gymnasium/Gym id:** `"VizdoomDefendCenter-v0"`

**Configuration file:** [defend\\_the\\_center.cfg](#)

The model was then tested for another 1 million steps with reward shaping that punished the model by 0.01 pts whenever it fired a bullet to encourage saving ammo.

Also the the map was modified to allow the Doomguy to move in any direction, the best model with the unmodified map was:

...\\Initial Testing With Reinforcement Learning\\Training\\checkpoints\\best\_model\_PPO\_test\_defend\_the\_center\_14\\best\_model\_1000000.zip

Which was trained for 1 million steps with 8192 n steps on the fitness function given on the ViZDoom website.

deadly\_corridor.cfg:

This model was trained with 8192 n steps, a clip range of 0.1, a gamma of 0.95, and a gae\_lambda of 0.9. It started by training for 400k steps on a version of the map set to difficulty 1, and then 50k steps for every difficulty of the map up to difficulty 5.

It used the fitness function on the website:

# DEADLY CORRIDOR

The purpose of this scenario is to teach the agent to navigate towards his fundamental goal (the vest) and make sure he survives at the same time.

The map is a corridor with shooting monsters on both sides (6 monsters in total). A green vest is placed at the opposite end of the corridor. The reward is proportional (negative or positive) to the change in the distance between the player and the vest. If the player ignores monsters on the sides and runs straight for the vest, he will be killed somewhere along the way. To ensure this behavior difficulty level (`doom_skill`) = 5 (config) is needed.

## REWARDS:

- +dX for getting closer to the vest.
- -dX for getting further from the vest.
- -100 for death

## CONFIGURATION:

- 7 available buttons: move forward/backward/left/right, turn left/right, shoot (attack)
- 1 available game variable: player's health
- timeout = 2100
- difficulty level (`doom_skill`) = 5

**Gymnasium/Gym id:** "VizdoomCorridor-v0"

**Configuration file:** [deadly\\_corridor.cfg](#)

Plus the following reward shaping:

Default reward - (damage taken \* 10) - (ammo used \* 5) + enemies hit \* 200

Eventually it developed the strategy of booking it to the armor, this is a worse strategy than killing every enemy because it does not work on higher difficulties, only lower difficulties.

I could keep perfecting this scenario, do other scenarios and perhaps even maps but I would prefer to move on from reinforcement learning and use evolutionary learning instead, so this is the end of the initial testing with reinforcement learning.