

Doom Bot Name: Marvin
Contact Info: islandman93@gmail.com

Marvin is a neural network based approach with two separate action outputs: moving and aiming. Marvin was trained with supervised replay of human play plus reinforcement learning against built in bots. Additionally, an aimbot was created to overwrite the network's aiming actions during training.

Important bits:

- The softmax heads (moving & aiming) are pre-trained with categorical cross entropy from human play and the critic head is trained on a one-step discounted reward from the same human dataset.
- Trained with asynchronously batched n-step actor critic over 16 threads. Each thread submits a batch of experiences asynchronously to a central controller and network which runs a training step across the agents that have submitted a batch recently. Similar to PAAC but not blocking, IE a single training step may have multiple batches from one agent and none from another. Training occurs when at least 16 batches have been submitted.
- If the network is unsure of the correct aiming action (the maximum softmax probability is less than 0.5) and the current weapon is either shotgun or minigun the aimbot will overwrite the network's aiming action with the correct action (moving the crosshairs to the closest enemy).
- Network structure is in network.py. N-step returns are calculated across 25 steps with frame skip 2, so a batch encompasses a 50 frame segment.

Moving Actions:

- Right, Left, Forward, Backward, Speed & Forward

Aiming Actions:

- No-op, Turn Left (4 degrees) & Fire, Turn Right (4 degrees) & Fire, Turn Left (16 degrees), Turn Right (16 degrees), Turn Up (2 degrees) & Fire, Turn Down (2 degrees) & Fire

Marvin Track 1 was created using the weights from Marvin track 2 and training on Track 1 for 80,000,000 steps.