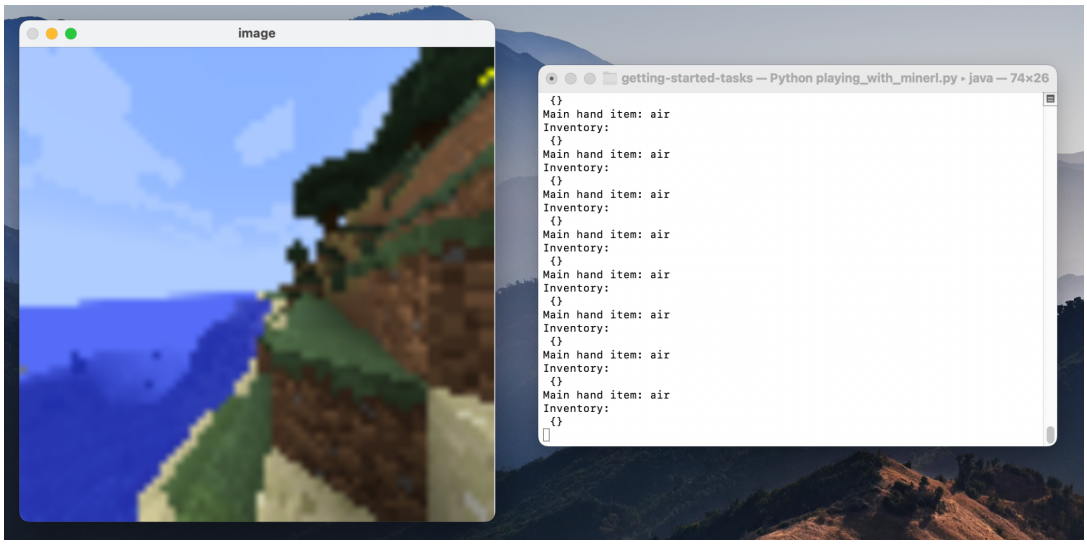
Requirement 4 Report - Getting Started With MineRL

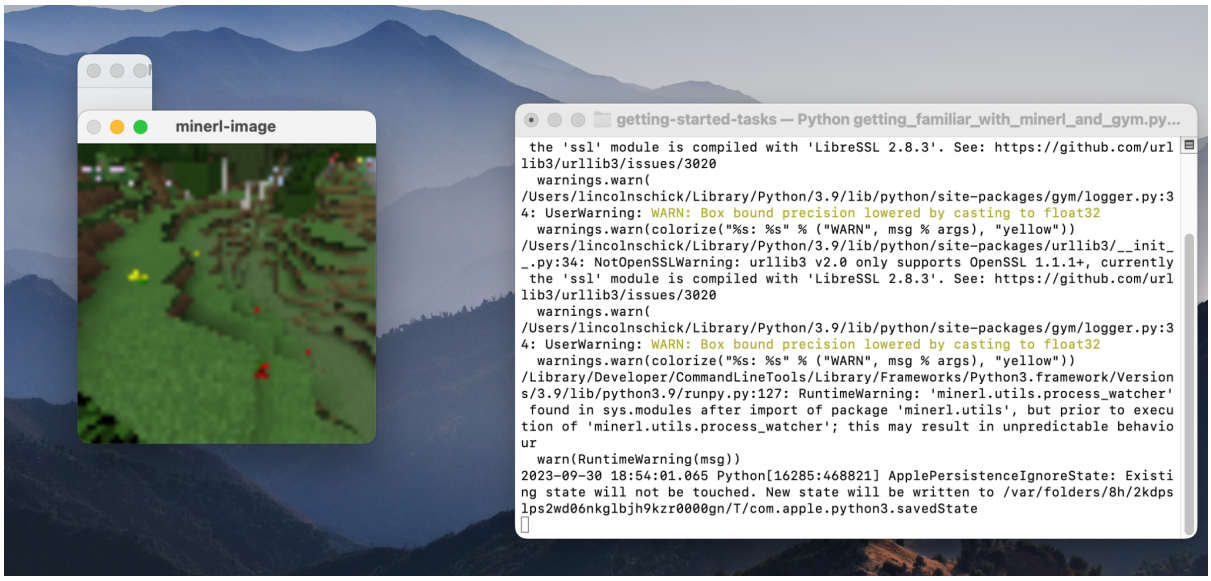
Lincoln

- Screenshot from running playing_with_minerl.py



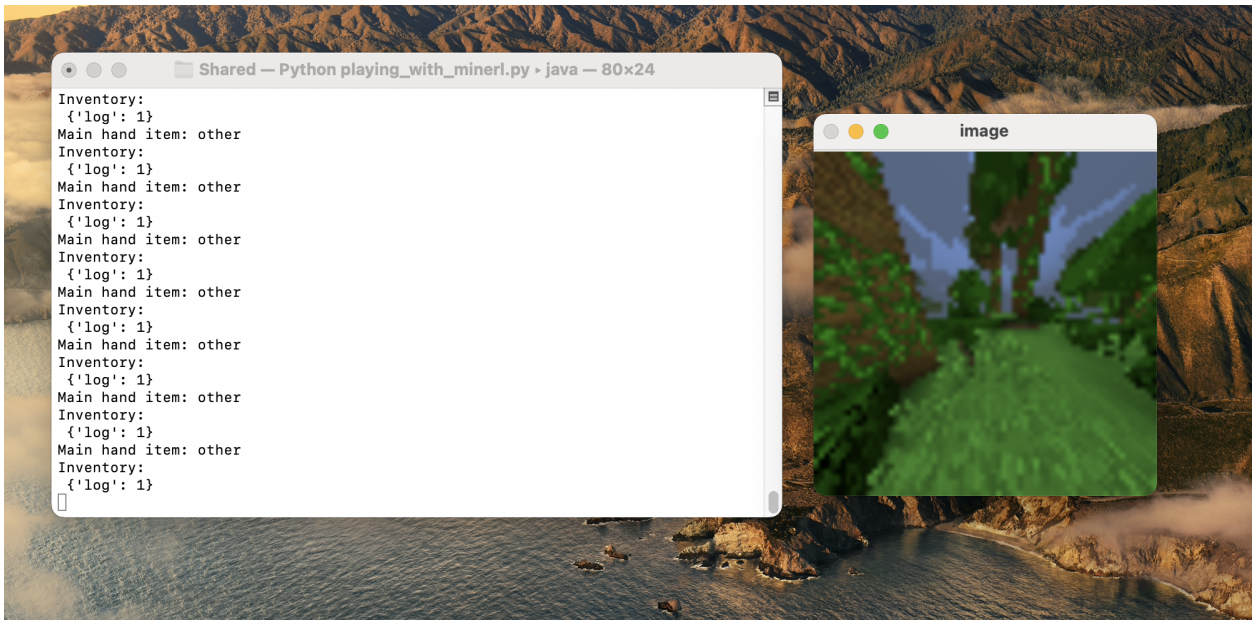
- CartPole-v0 environment
    - What is the goal?
        - The goal is to prevent the pole from falling down
    - What is the state information (what does the state from env.step(action) mean)? You may want to inspect environment.observation_space.
        - The state information tells you about the cat and pole at a particular time. Specifically, it is an array containing the cart position, the cart velocity, the pole angle, and the pole velocity.
    - What are actions (what does the action in env.step(action) do)? What are the different actions? You may want to inspect environment.action_space.
        - Actions are moves that can be taken in the game. env.step(action) will update the game with one's move. This is a discrete variable either a 0 or a 1 where a 0 moves the cart left and a 1 moves the cart right.
    - What is the reward?
        - The reward is just the number 1. This is because the goal is to stay alive and prevent the pole from falling down, so an AI would try to maximize the total reward by making moves that prevent the game from ending.
    - What is the starting state? Does it change between different episodes (games)?
        - The start state is simply an array containing initial values for the cart position, cart velocity, pole angle, and pole velocity, and it does change between different games.
    - Is the environment deterministic or stochastic? Environment is deterministic if you can perfectly predict what happens with every action in every state.
        - The environment is deterministic. Since the environment is purely based on laws of physics which are invariable, one can predict exactly what changes after each action.

- Screenshot from running getting_familiar_with_minerl_and_gym.py with the MineRLTreechop-v0 environment



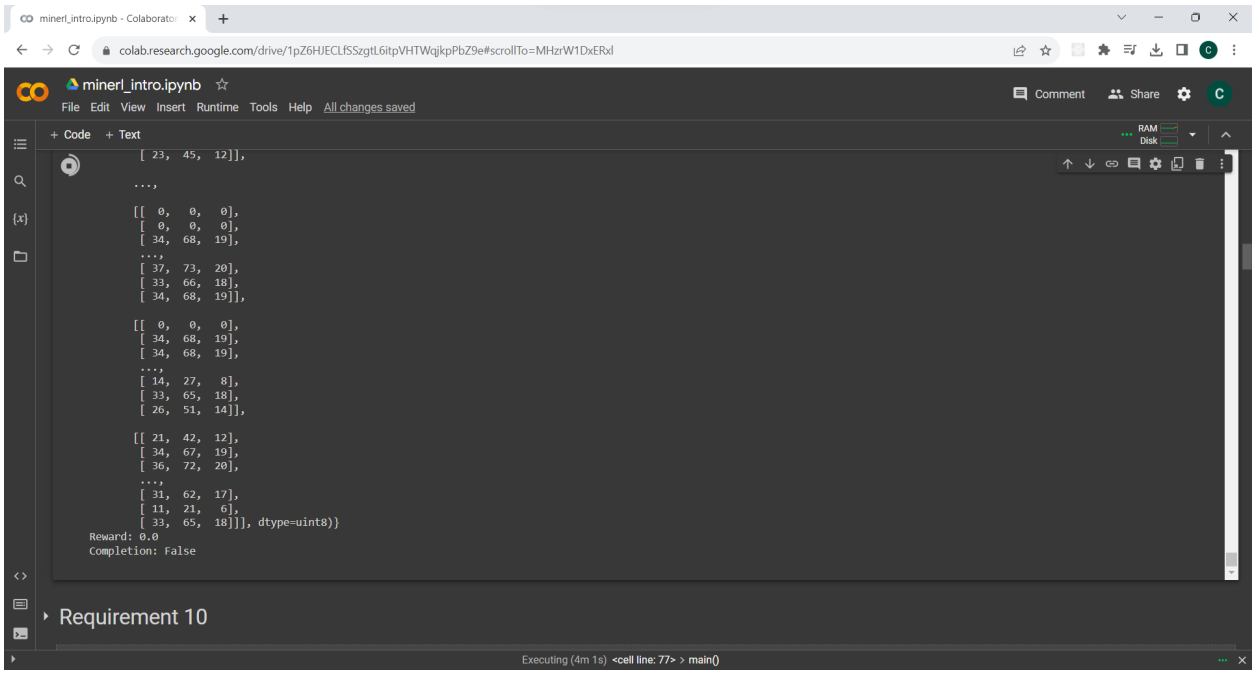
Cody

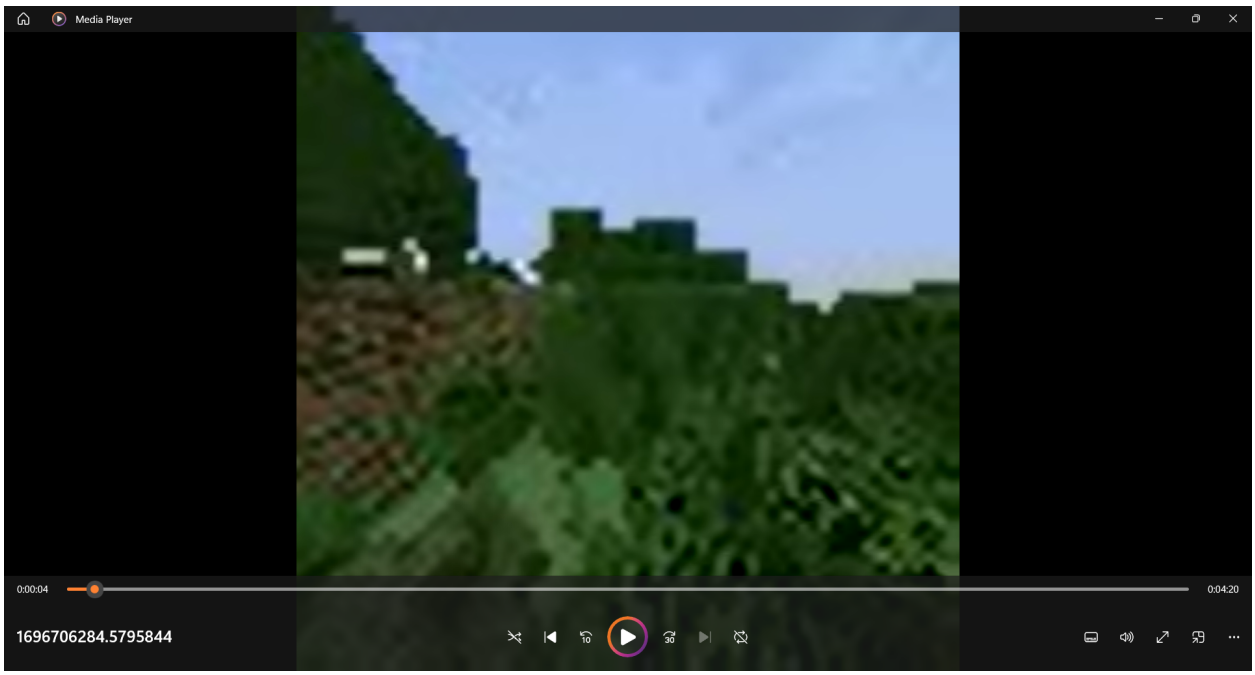
- Screenshot from running playing_with_minerl.py



- CartPole-v0 environment
    - What is the goal?
        - The goal is to keep the stick balanced upon the cart, preventing the angle of the pole from reaching a value greater than 15 degrees.
    - What is the state information (what does the state from env.step(action) mean)? You may want to inspect environment.observation_space.
        - The state information is an array consisting of four values that denote the position and velocity of the cart and the pole at any given instance of the episode.
    - What are actions (what does the action in env.step(action) do)? What are the different actions? You may want to inspect environment.action_space.
        - Actions are the moves available to the AI to interact with the environment, modifying the state for the next iteration/step of the environment.
        - In this example, the AI has the ability to move the cart right or left depending on the value passed into the step() function. A value of zero denotes left and a value of one denotes right.
    - What is the reward?
        - The reward is the value 1.0. This is the reward for the AI performing an action that does not lead to the pole falling down. This leads to an AI agent trained via reinforcement learning in this environment to learn to maximize the value by keeping the pole balanced.
    - What is the starting state? Does it change between different episodes (games)?
        - The starting state is an array containing four float values denoting the position of the cart and the pole in the environment.
        - No, the starting state does not change between episodes.

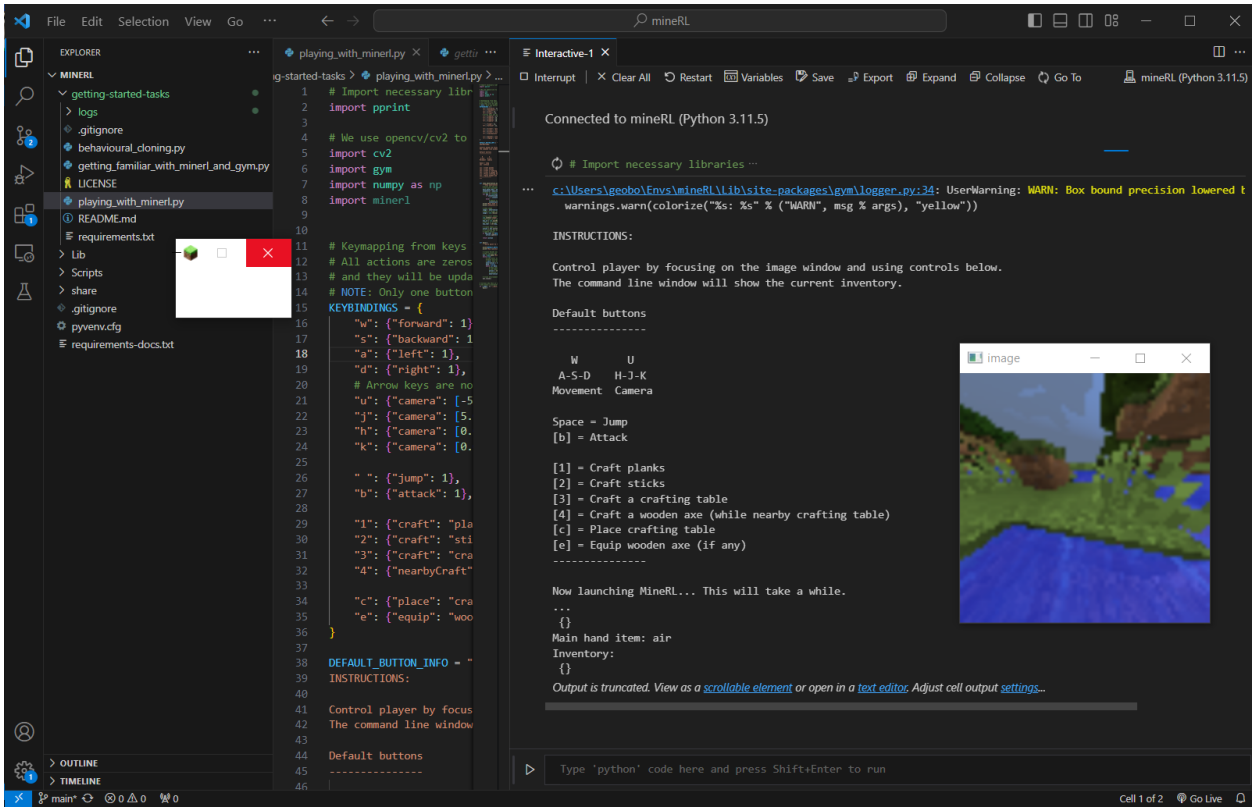
- Is the environment deterministic or stochastic? Environment is deterministic if you can perfectly predict what happens with every action in every state.
    - This environment is deterministic. The only influences on the state are the actions of the machine learning agent and the values from the previous iteration of the state (i.e. velocity causing continued motion of either the cart or the pole). Thus, it is always possible to completely predict the next state from knowledge of the agent's actions and the previous state.
- Screenshot from running getting_familiar_with_minerl_and_gym.py with the MineRLTreechop-v0 environment





Jack

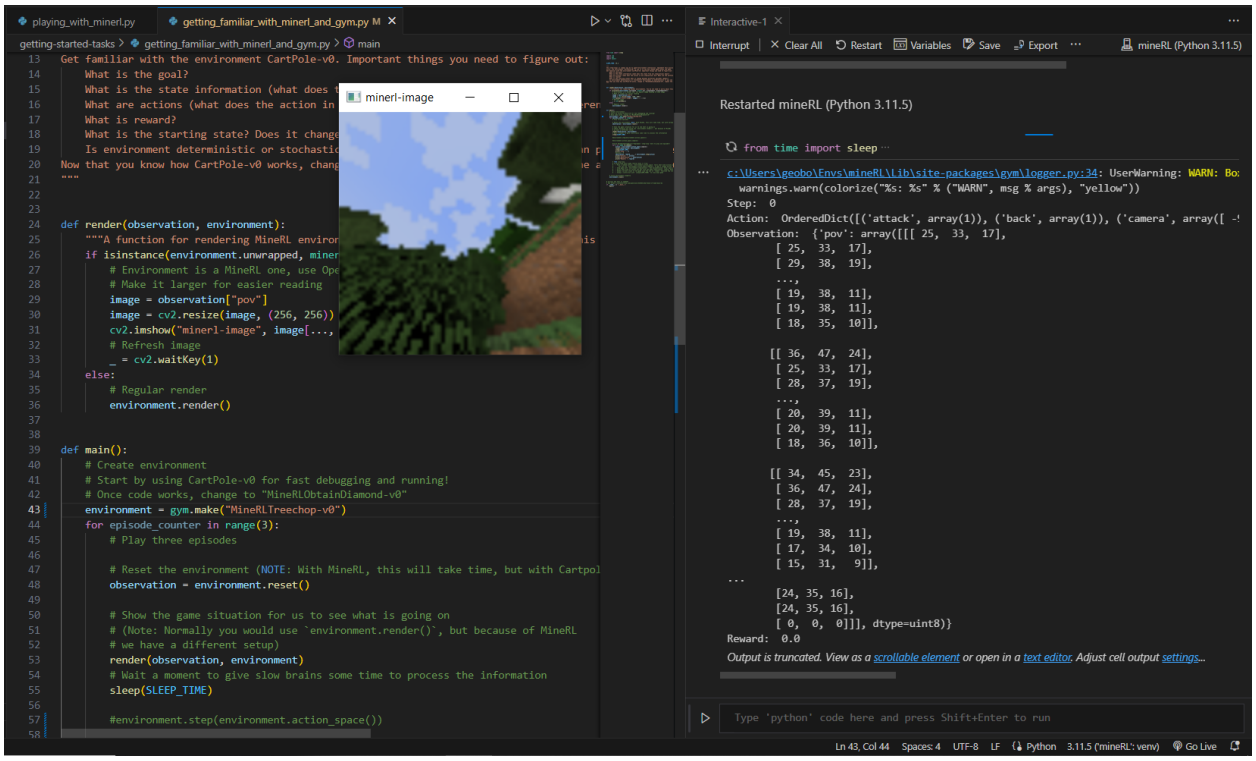
- Screenshot from running playing_with_minerl.py



- CartPole-v0 environment
    - What is the goal?
        - The goal of CartPole-v0 is to balance a wooden pole on a cart. The model has decisions it can make about how the cart is moved which may help balance or destabilize the pole depending on the model's previous actions.
    - What is the state information (what does the state from env.step(action) mean)? You may want to inspect environment.observation_space.
        - The state information the model receives the certain metrics on the position and motion of the cart, including its current position along the surface, the angle of pole, and the velocity of both of those objects.
    - What are actions (what does the action in env.step(action) do)? What are the different actions? You may want to inspect environment.action_space.
        - There are two actions the cart can take - to move left and to move right, which can be passed to env.step() to apply the action at each step. These actions may be passed randomly by sampling the action space or in a controlled manner by passing the action explicitly using some logic.
    - What is the reward?
        - The reward is a constant value that is passed each time an action is made, rewarding the model for keeping the pole balanced.

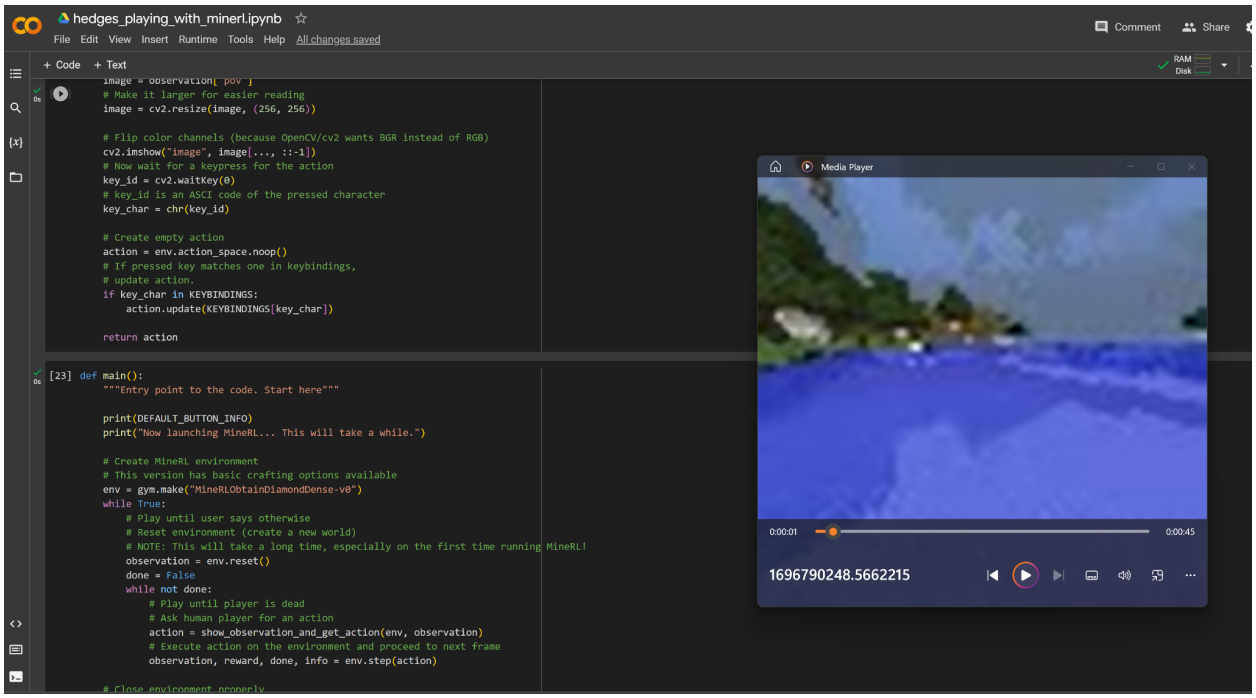
- What is the starting state? Does it change between different episodes (games)?
    - The starting state is to have the pole balanced straight up in the air. This does not change between different episodes.
- Is the environment deterministic or stochastic? Environment is deterministic if you can perfectly predict what happens with every action in every state.
    - The environment is chaotic but deterministic. Since there is a consistent starting condition, a model performing the exact same inputs at every step should perform the same across different tests.

- Screenshot from running getting_familiar_with_minerl_and_gym.py with the MineRLTreechop-v0 environment



Faith

- Screenshot from running playing_with_minerl.py:



- Answers to the following questions about the CartPole-v0 environment
    - What is the goal?
        - To maximize the amount of time that the simulated pole is balanced on the cart, via motions of the cart that the agent controls.
    - What is the state information (what does the state from env.step(action) mean)? You may want to inspect environment.observation_space.
        - It is a matrix containing the cart linear position, cart linear velocity, pole angular position, and pole angular velocity
    - What are actions (what does the action in env.step(action) do)? What are the different actions? You may want to inspect environment.action_space.
        - Actions are the manner in which an agent is able to interact with an environment; they are the rules that the entity must abide by while striving for a maximal reward. The action space in cartpole is limited, since it is a simple testing environment: move the cart a linear increment to the left or to the right.
    - What is the reward?
        - In general, rewards are points that a model aims to maximize. In this game, the reward is given based on the amount of time spent with the pole in the optimal range; the game ends once it falls out of this range (or the cart moves to a lateral out-of-range position), preventing a higher score.
    - What is the starting state? Does it change between different episodes (games)?
        - For cartpole, the starting state is always the same, since consistency provides simplicity, the goal of this test.

    - Is the environment deterministic or stochastic? Environment is deterministic if you can perfectly predict what happens with every action in every state.
        - It is deterministic, since the laws of physics are simulated within this environment; given an input cart velocity, the radial pole velocity can be precisely calculated.
- Screenshot from running getting_familiar_with_minerl_and_gym.py with the MineRLTreechop-v0 environment: