Lab 9: Training a Room Explorer using RL

Brandon Trinkle

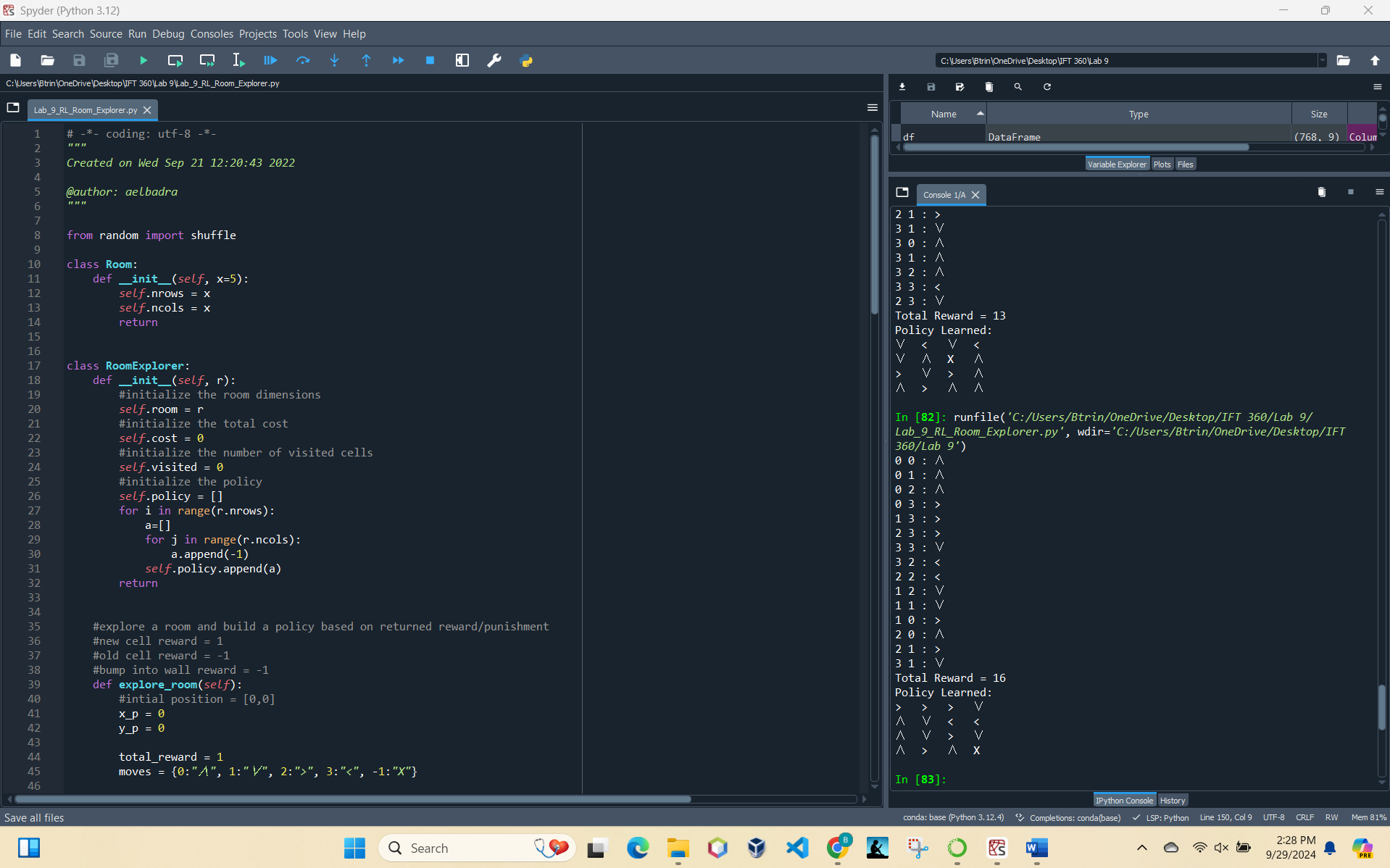
Arizona State University

Course Number: IFT 360

Professor Durgesh Sharma

9/29/24

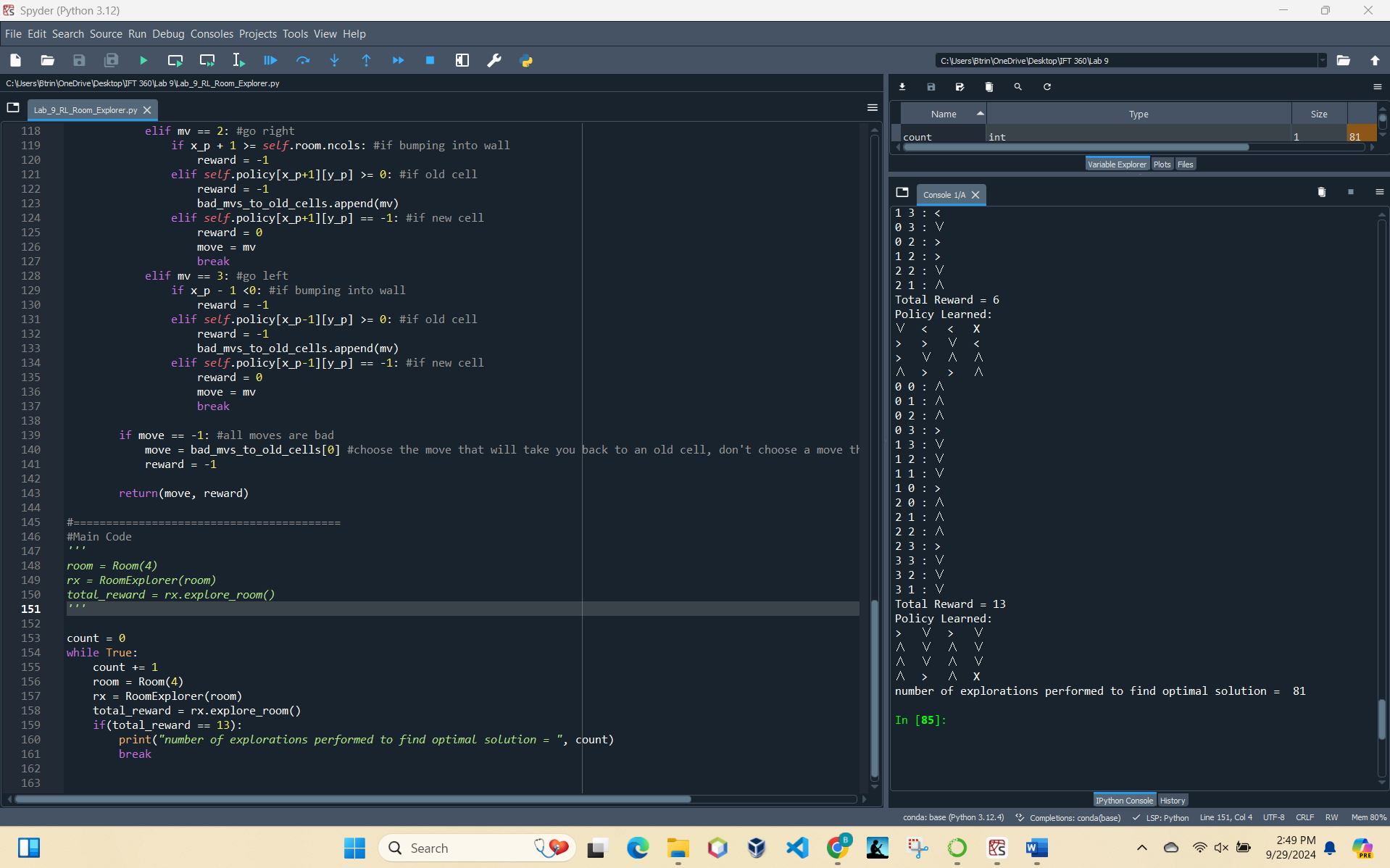
**Question 1:** *How does the optimal policy path look like? Draw the path on a 4 X 4 grid.*



Optimal path:

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Question 2:** *Does it contain as much right or left turns as the policies generated by the old reward function?*

****

No this does not contain as many left or right turns as the old reward function gave us. We had 7 left and right turns to obtain a reward of 16 with the old polices. The updated policy only required 3 left and right turns.

**Question 3**: *What is the number of times the exploration task was carried out until the optimal policy was found?*

As seen in the screenshot, it performed this 81 times until it found the most optimal path to achieve the reward.