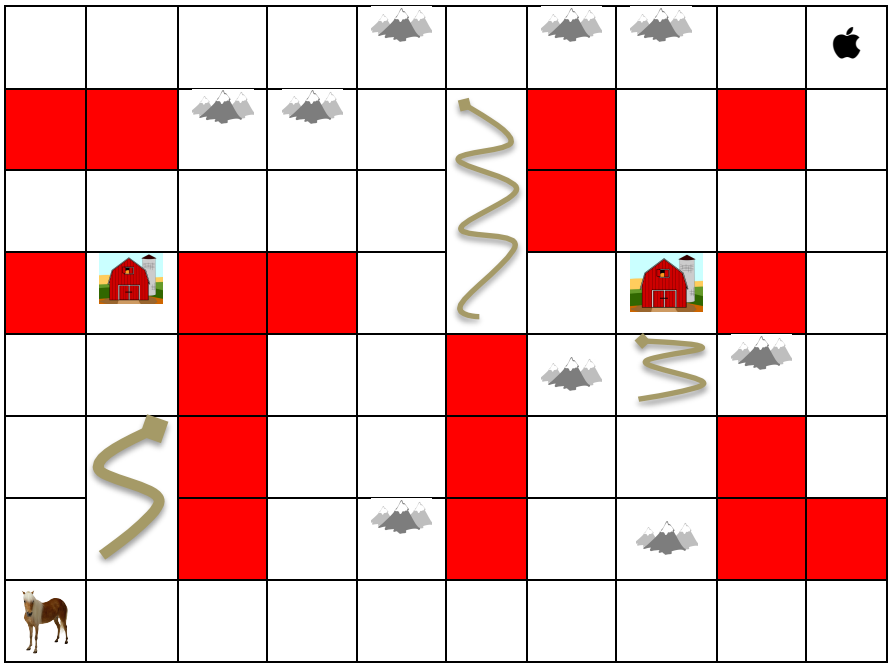
CSCI 3202 - Introduction to Artificial Intelligence
Instructor: Hoenigman
Assignment 5
Due Wednesday, October 7 by 4pm

Problems:

Github
In your *ai3202* repository, create a folder called *Assignment5*, and place all code for this assignment in that repository. We will be grading your assignment by pulling your code from that repository.

Markov Decision Processes
The following pictures show fictional worlds where the horse needs to navigate the maze to get to the bushel of apples. Along the way, there are walls that he needs to go around (shown in red), treacherous mountains to navigate, and snakes to avoid. There are also warm and cozy barns where he can recharge. Use the Value Iteration algorithm to find the path through the maze that has the highest utility.



World 1. Navigate the maze to get the horse to the apple. Watch out for mountains and snakes.

Rules for the MDP:

- There is a discount factor, γ, of 0.9.
- There is an additional negative reward in the mountains of -1.0. The mountains take extra time to navigate and they can be treacherous.
- There is an additional negative reward going past the snakes of -2.0. Snakes bite. Ouch!

- There is a positive reward of 1.0 if the horse gets to a barn.
- The horse can't move through a wall.
- The horse can move in four directions only: left, right, up, down.
- The reward for getting the apples is 50.0.
- The value for ϵ in $\delta < \epsilon(1-\gamma)/\gamma$ is 0.5.
- Use the transition model discussed in class, where the horse is successful 80% of the time, 10% of the time he will go left of his intended move and 10% of the time he will go right of his intended move.

Implementation details

- The world is represented as a matrix, where 0 is an open square, 1 is a mountain square, 2 is a wall, 3 is a snake, and 4 is a barn.
- The goal state is represented in the matrix with the value of 50.
- The world is on Moodle, called World1MDP.txt. The values in the text file show the category of each cell in the matrix and the reward for the goal state. Your program needs to read in the file, using the filename as a command line argument for your program.
- The value for ϵ should also be a command-line argument.
- The output of your program needs to include:
    - The utility scores along the optimal path.
    - The locations along the path.
- It should be possible to modify you're A* search implementation from Assignment 3 to create your MDP. In the Value Iteration implementation, you will want to modify the algorithm we used in class to store both the utility for the state and either the action or the parent for the state.

Question to Answer:

- Run your program multiple times with different values for ϵ. Can you find a value that changes the solution path of your program? Include your explanation of the different values for ϵ that you tried and whether any of these values changed the solution.