Consider a 5x10 gridworld. There is a door at the lower right hand side of this grid (D1, red box). The agent starts at a random location and has four actions (move in each of the four directions). There is a reward of 100 to reach D1, and a reward of -2 for every time step the agent is in this gridworld.

1- Assuming “magic” sensors of your choice (you know exactly where you are, etc) devise a navigation algorithm (see note) that can reach D1. Clearly state all system parameters (inputs, states, outputs, training method, etc.)

2- Now, the agent is placed in a similar location with a new door at the upper left corner (D2, yellow box). Use the EXACT navigation algorithm you designed in step 1. What is the outcome?
   a) if the agent can find the door and get out, great! Explain why and how this worked.
   b) If the agent cannot find the door, explain why your algorithm failed to generalize to this case. What was the problem? How should you modify your algorithm to avoid this in the future?

Note: The navigation algorithm can be a neural network or a reinforcement learning algorithm. Clearly state the reward structure you select in either case.