

The first page allows you to select a character for training and select the map size. Note that larger map size means longer training time.

Cat and Mouse 2: Reinforcement Learning Training Configuration

A Project by Joseph Ma

Choose Cat Character:

Regular Cat

Choose Mouse Character:

Regular Mouse

Maze Dimensions:

10 x 10

Design Your Own Map

The next page allows you to design your own map for the training environment. Place walls down as you would like (in black).

Cat and Mouse 2: Reinforcement Learning Training Map Configuration

A Project by Joseph Ma

Selected Cat: **Regular Cat**
Selected Mouse: **Regular Mouse**
Selected Dimensions: **10**
Map Validity (Ensure all paths are connectable from one another): ●

Exit

Entrance

Start Game

Clear Map

Note that if a portion of the map is disconnected from the others this will invalidate it. Ensure all paths are connectable from one another. There should also be at least one path from entrance to exit.

Cat and Mouse 2: Reinforcement Learning Training Map Configuration

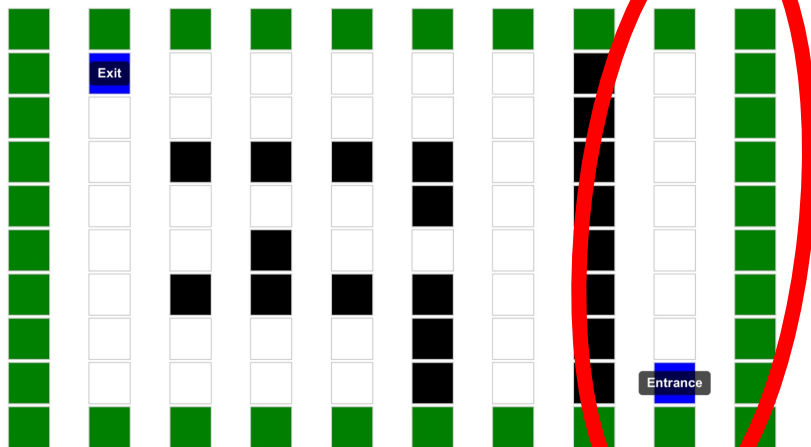
A Project by Joseph Ma

Selected Cat: **Regular Cat**

Selected Mouse: **Regular Mouse**

Selected Dimensions: **10**

Map Validity (Ensure all paths are connectable from one another): ●



Start Game

Clear Map

Let's begin training

Cat and Mouse 2: Reinforcement Learning - a Project by Joseph Ma

Movement Speed: 85

Current Epsilon: 0.884084

Current Training Episode: 5

Current Reward: 26

Cumulative Episode Rewards: 26


Exploration / Exploitation: ☒ Exploration ☐ Exploitation

Prey Status: ☐ Dead ☒ Alive ☐ Flee

Dead End Status: ☐ Dead End ☒ Not Dead End

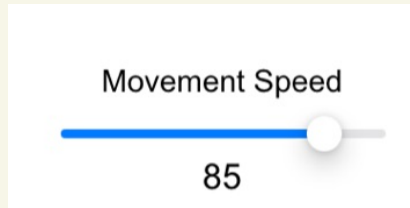
Phase:

Show A* Shortest Path:



The image shows a maze environment with a green border and black walls. A cat icon is at the top, and a mouse icon is in the center. The cat is currently in a dead end. The mouse is in a path that leads to the cat. The interface includes various controls and status indicators.

Adjust the speed at which
the episodes play out



We use the greedy epsilon strategy, this is the chance the mouse chooses the best action to take. And $1 - \epsilon$ the chance the mouse chooses a random action.

Current Epsilon
0.884084

The current reward is a reflection of how “good” the mouse’s action is. The more positive the better, and the more negative the worse.

Current Reward
26

The total rewards accumulated through the current episode.

Cumulative Episode Reward:
-318

Indicates whether the most recent action by the mouse was exploration or exploitation

Exploration / Exploitation

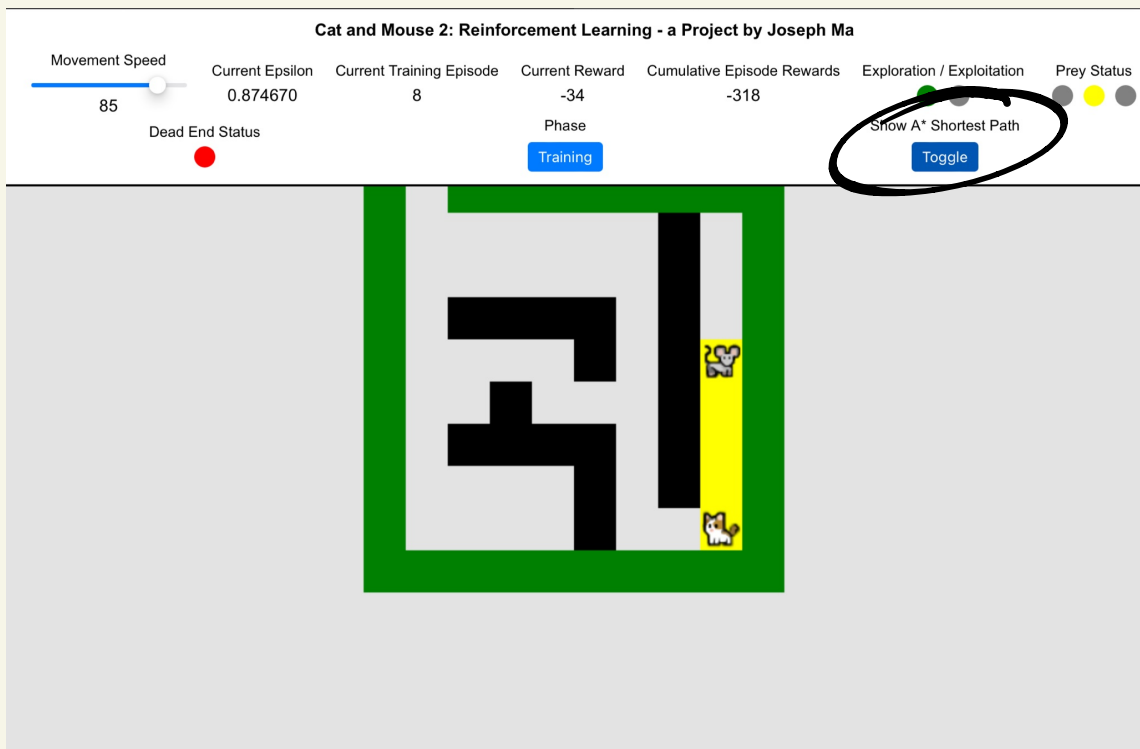
A green circle followed by a grey circle, indicating that the most recent action was exploration.

Prey status: Green will light up at the most left if mouse escapes, Yellow will light up in the middle if mouse is still alive, Red will light up at the most right if mouse gets caught by the cat

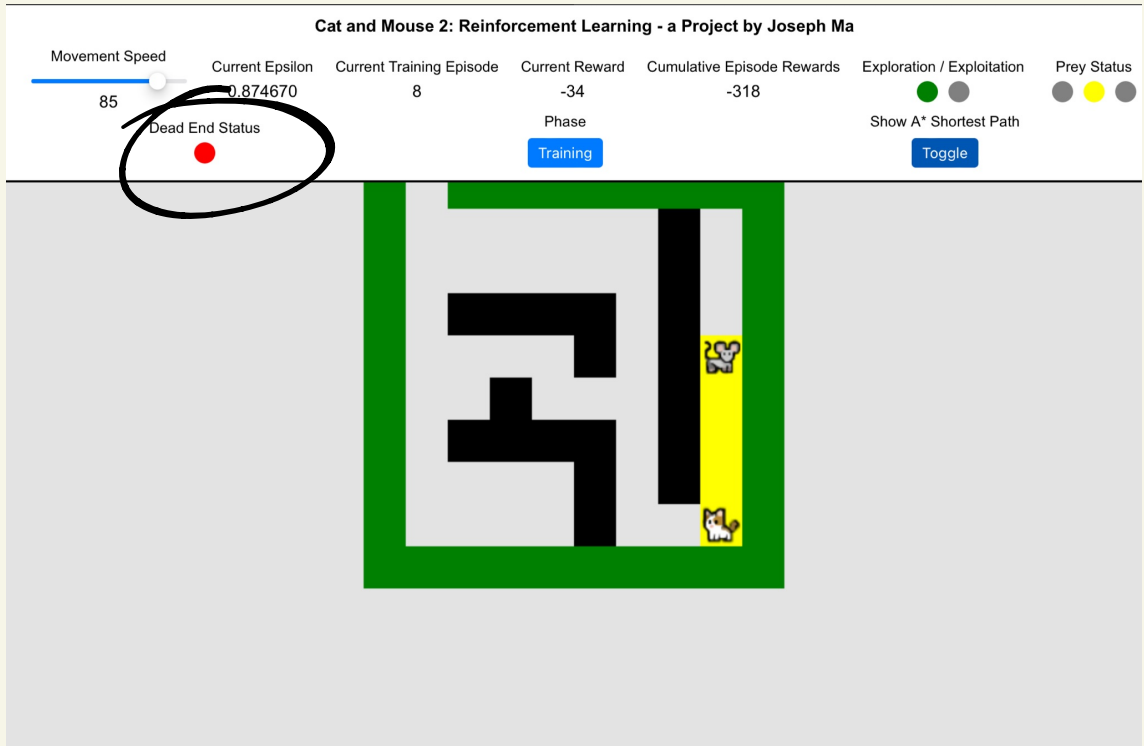
Prey Status



Click toggle to show the A* shortest path calculated by the cat to reach the mouse. The cat always takes the shortest path to the mouse.



If the mouse finds itself cornered by the cat, the dead end status will light up. A double DFS algorithm was written to check for this.



Once epsilon has decayed, toggle the training button. This will cause the mouse to take the best action in each move. In other words it will exploit every action. In this mode, the Qtable and epsilon will remain frozen until we return to regular training mode. This mode is to visualize how intelligent the mouse has gotten.

