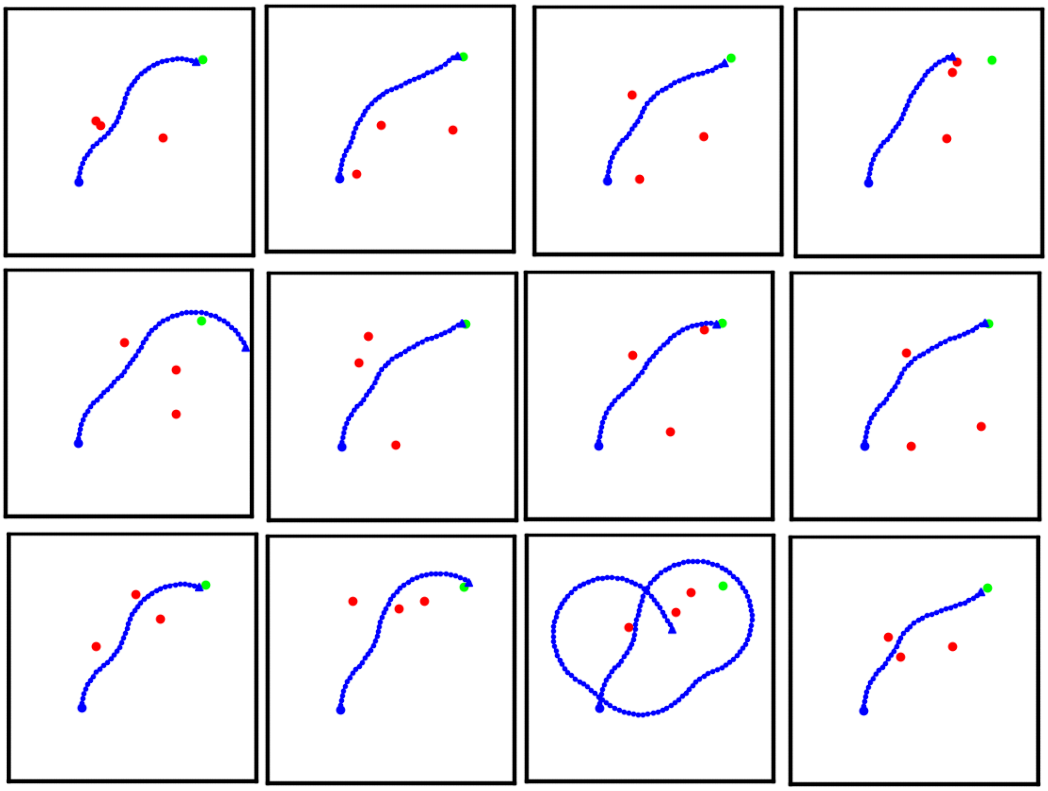
**Obstacle Avoidance**

- This environment has the simplest task of finding an optimal path from start to goal point while avoiding static obstacles scattered around the map.
- There are 3 possible states: free space, obstacle, or goal.
- At each timestep where the goal is not visible to the agent, a virtual goal will be generated to guide the agent towards the goal. The distance between agent and virtual goal will be half of the observation radius.
- Simulation parameters:
    - Map dimension: W100 x H100
    - Start – goal points: (30, 30) – (80, 80)
    - Static obstacles: randomly generate 3 around the map (especially the area between start and goal)
    - Observation radius: 40
    - Grid size: 5
    - Distance to waypoint/virtual goal: 20 or 25
    - Turn rate: 5
    - Speed: 2
    - Maximum steps: 150
- Reward function:
    - Collide with obstacle/boundary: -5
    - Reach goal: +10
    - On free space: 0
- Training hyperparameters:
    - Learning rate: 0.0001
    - Batch size: 30000
    - Number of minibatch (n_epochs): 10
    - Discount factor (gamma): 0.99
    - Clipping parameter (clip_range): 0.1
    - Value function coefficient (vf_coef): 0.5
    - Entropy coefficient (ent_coef): 0.01
- Results:

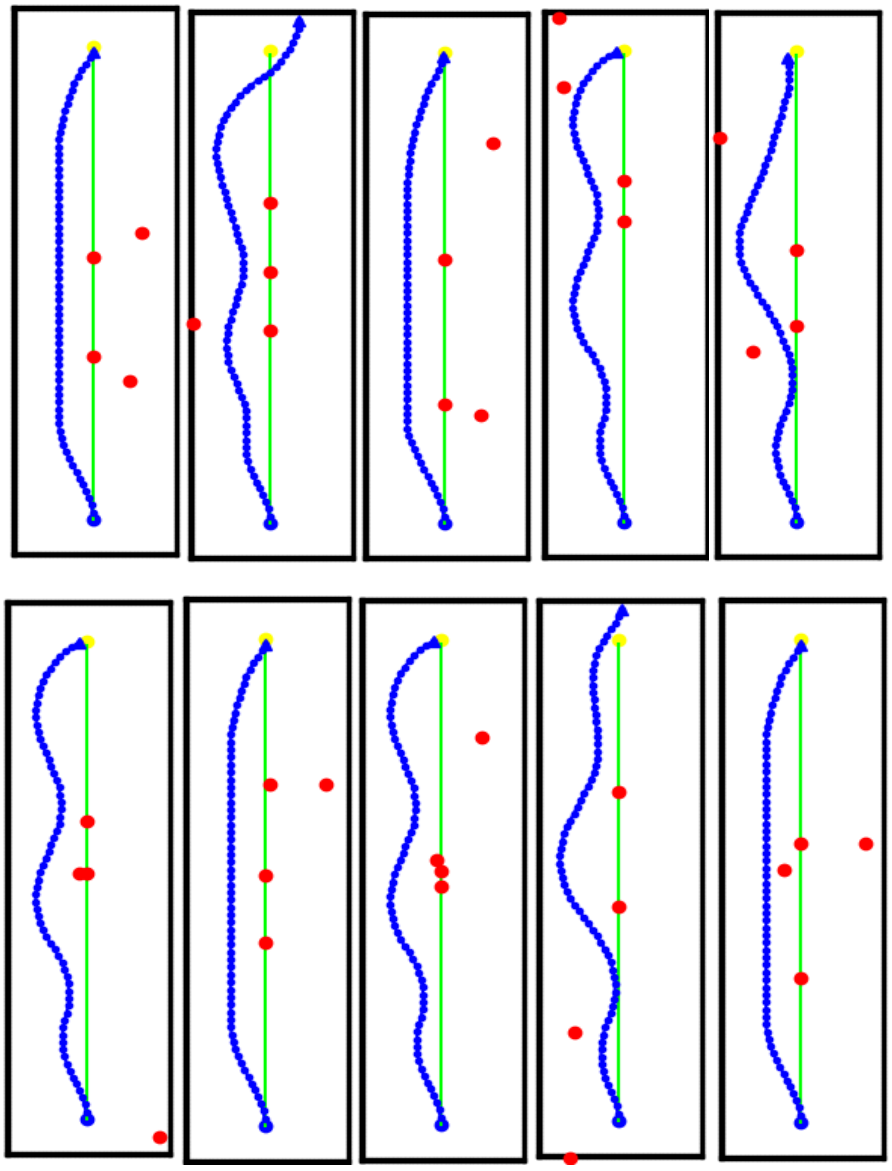


- Comments:
    - Generally, it performs good for about 90% of the time.
    - For the unsuccessful cases, the agent was able to avoid all obstacles but cannot turn in time and miss the goal point, and after that the agent was not able to turn back, either hitting the boundary or wandering around until reaching the maximum number of steps.
    - Need to improve on mobility of the ASV.

**Static Obstacles**

- Addition to the previous case, this environment has 4 possible states: free space, path, obstacle, or goal.
- The agent needs to navigate from start to goal point, avoid static obstacles, and stay on the path as much as possible, where the path is a straight line connecting start and goal points.
- Similar to Obstacle Avoidance, a virtual goal will be generated and updated at each timestep to guide the agent towards goal point. The distance from agent to virtual goal is half of observation radius.
- Simulation parameters:
    - Map dimension: W40 x H150
    - Start – goal points: (20, 10) – (20, 140)
    - Static obstacles: randomly generate 2 along the path, and 2 around the map.
    - Observation radius: 40
    - Grid size: 5
    - Distance to waypoint/virtual goal: 20
    - Turn rate: 5
    - Speed: 2

    - Maximum steps: 200
- Training hyperparameters:
    - Learning rate: 0.0001
    - Batch size: 30000
    - Number of minibatch (n_epochs): 10
    - Discount factor (gamma): 0.99
    - Clipping parameter (clip_range): 0.1
    - Value function coefficient (vf_coef): 0.5
    - Entropy coefficient (ent_coef): 0.01
- Reward function version 1 (does not consider path)
    - Collide with obstacle/boundary: -10
    - Reach goal: +5
    - On path or free space: 0
- Results v1:



- Reward function version 2 (consider path following)
    - Collide with obstacle/boundary: -50
    - Reach goal: +10
    - On path: +1
    - On free space: 0
- Results v2: