1: K-Means

(a)

Initial conditions:

Cluster centers: 1,6
Data assignment:
- Cluster 1: 1,2,3
- Cluster 2: 4,9,12,6,10,9

Iteration 1:

Cluster centers: 2,8.3333333
Data assignment:
- Cluster 1: 1,2,3,4
- Cluster 2: 9,12,6,10,9

Iteration 2:

Cluster centers: 2.5,9.2
Data assignment:
- Cluster 1: 1,2,3,4
- Cluster 2: 9,12,6,10,9

(b)

Yes because no data switched cluster assignment between the last two iterations, thus the cluster centers will remain unchanged indicating the the cluster assignments will also remain unchanged.

2: K-Means and Variance

(a)

Variance decreases because the average distance to a cluster center decreases as there are more cluster centers.

(b)

K=n because each cluster will have exactly one data point and the variance of one data point is zero.

3: Reinforcement Learning

The reward function is not effectively communicating the goal to the agent. Instead of simply just having one large reward which is given for escaping the maze, the reward function needs to include a type of discounted living reward for steps leading toward the exit of the maze. Currently, the reward is always 0 until time step T when the agent leaves the maze. In order to motivate the agent to leave the maze more quickly, we can penalize it for each time step it remains in the maze and potentially even increase this penalty as time goes on. For example, we may choose to have a penalty of -0.01*t where t is the time elapsed since the beginning of the episode.

4: Reinforcement Learning II

(a) The signs of the rewards are important because if we were to implement the same reward system, but using the absolute values of the rewards outlined in the question, our agent would learn to both score goals and run into the edge of the world. Depending on the space of the environment, this could lead to a situation in which the rewards were easier to reap by simply running into a near by edge and would cause the agent to learn unproductive behaviors.

(b)

$$R_t = \sum_{k=0}^{\infty} \gamma^k * r_t + k + 1 + c$$
$$V^\pi(s) = E[R_t | s_t = s]$$
$$V^\pi(s) = E\left[\sum_{k=0}^{\infty} \gamma^k * r_t + k + 1 + c | s_t = s\right]$$
$$V^\pi(s) = E\left[c * \infty + \sum_{k=0}^{\infty} \gamma^k * r_t + k + 1 | s_t = s\right]$$
$$V^\pi(s) = E\left[K + \sum_{k=0}^{\infty} \gamma^k * r_t + k + 1 | s_t = s\right]$$
$$V^\pi(s) = K + E\left[\sum_{k=0}^{\infty} \gamma^k * r_t + k + 1 | s_t = s\right]$$
$$\therefore$$
$$V^\pi(s) = E[R_t | s_t = s] + K$$

(c)

$$K = n * c$$

Where n = number of time-steps t in R

II 1: Image Segmentation Using K-Means

New Mexico: K=7

 

Polar Bear: K=7

 

Skis: K=7