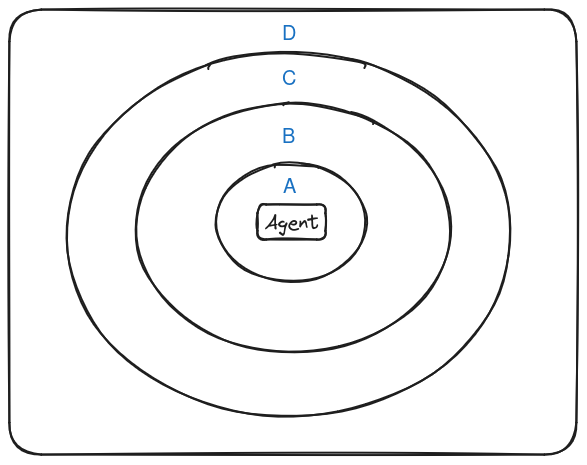
Quiz - 2

1. What are the parameters to a state-value function?
2. What are the parameters to an action-value function?
3. What is meant by a policy(π)? Is it deterministic or stochastic?
4. Under an optimal policy π\*, it is not necessary that **every** state has the highest state-value compared to any other policy. (TRUE/FALSE)
5. You are trying to train an agent that has **1000 chances** to play a game. Player/Agent tries to throw a ball standing at center and gets **different** **points** when the ball reaches **different regions** as follows:

A: 10  
B: 20  
C: -30  
D: 40

There is always a 90% chance that the ball reaches the region that was aimed for and a 10% chance that it reaches an adjacent next region(either above or below).

* How would you choose the rewards to formulate this as an RL problem? Select the reward you would give the agent for throwing the ball to each of these regions such that it learns to get maximum total points by playing the game a 1000 times.
* Which region would be the best to aim for each time?

Note: This is a stateless RL problem as there are no states involved here. Therefore no state-transition probabilities.

1. Can there be more than one optimal policy for a given value function?