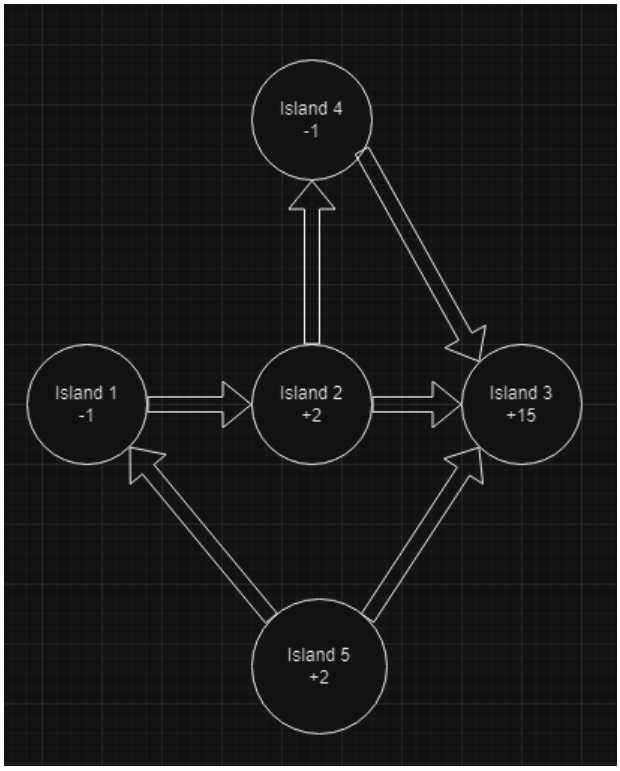
CSCI 397: Building the Value Function

Due Date: Friday October 6th @ 11:59 PM

Information

In this assignment, you have already built a Markov Reward Process, now I want you to calculate the value function for your MRP. This can involve the iterate process of applying the Bellman equation to each state on a repetitive process or using the close form solution. Once you do that, I want you to visually show me what the result of your value function is (aka tell me about the optimal policy through your MRP).

To make this a bit more manageable, I would like you to remove the resemblance of treasures from your MDP and make the reward -1 for traversing to 6/10 islands and +2 for 3/10 islands. You can keep the same structure as you had in assignment 2 but you're also welcome to change it. This should make the application of the closed form solution to the value function simple. The terminal island should have a reward of +15 which will incentivize the agent to terminate the simulation as quickly as possible. Below is a brief example of what a smaller MDP might look like. Your MDP should have at least 10 states, but this is a small example.



Report the values of each state, this could be a separate function within your assignment that purely calculates the value at each state.

Task 1: Calculate the Value Function

1. Understand what the value function's purpose is within the context of your MRP
2. Apply the Bellman equation iterative OR use the close form solution
3. Report the calculations of the value of each state
4. Report the optimal policy through the network (this part doesn't need to be in code, but it should follow your value calculations.
5. Tell me if your MDP is solved given your optimal policy and report the max reward possible.