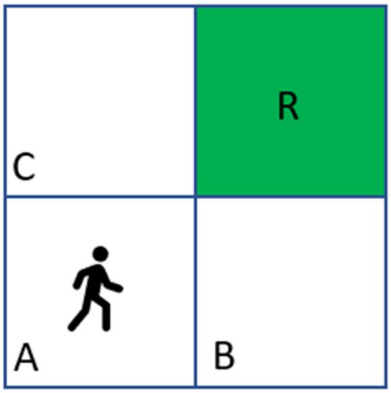
MDL Assignment 2

PART 1

Problem World:



The top right state is the terminal state with a **reward** given as per the instructions below:

Actions, Transition Probabilities and Rewards:

If at square A: The person has 2 actions, move Right and move Up.

- **Move Right:**
    - Moves to Square B with Probability 0.8 and Step Cost -1
    - Remains in Square A with Probability 0.2 and Step Cost -1
- **Move Up**:
    - Moves to Square C with Probability 0.8 and Step Cost -1
    - Remains in Square A with Probability 0.2 and Step Cost -1

If at Square B: The person has 2 actions: move Left and move Up.

- **Move Left:**
    - Moves to Square A with Probability 0.8 and Step Cost -1
    - Remains in Square B with Probability 0.2 and Step Cost -1
- **Move Up**:
    - Moves to terminal state with Probability 0.8 and Step Cost -4
    - Remains in Square B with Probability 0.2 and Step Cost -1

If at Square C: The person has 2 actions: move Right and move Down.

- **Move Right:**
    - Moves to Terminal State with Probability 0.25 and Step Cost -3
    - Remains in Square C with Probability 0.75 and Step Cost -1
- **Move Down**:
    - Moves to Square A with Probability 0.8 and Step Cost -1
    - Remains in Square C with Probability 0.2 and Step Cost -1

Parameters:

- **Gamma** (Discount Factor) = 0.20
- **Delta** (Bellman Error) = 0.01

Reward R is given as **R = Arr[(roll_number) % 15],** where

Arr = [8.8, 9, 10, 11, 12, 13, 13.9, 15, 16, 16.5, 16.6, 17, 18, 19, 20]

Questions:

Q1. Write the Transition Table (it has to be in the form of a table).

Q2. What do you think would be the best path for the person standing at Square A to reach the Terminal State? Do not calculate anything here, just try to guess a solution through appropriate reasoning.

Q3. Perform Value Iteration by Hand until Convergence. Clearly draw the new values at each state after each iteration.

Q4. Find the optimal path for the person at Square A to the Terminal State using the result from Value Iteration. Was your initial guess correct?

Q5. Try to make a guess of what could be the importance of specific values of reward and what could be the possible trend with the different values of the reward.

Marks and Deadline:

- This part of the assignment is for 20 marks with each of the above 5 questions comprising 4 marks.
- For Q2 and Q5, you will not be graded on the correctness of your guess, but the validity of your logic.
- The deadline for this part of the assignment is **20th March 2021**.