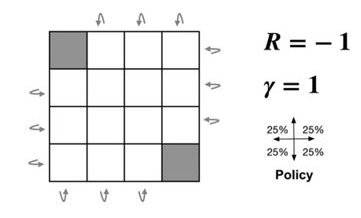
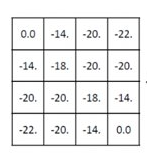
TT2 RL Question Bank (Modules 4-6)

**MODULE 4**

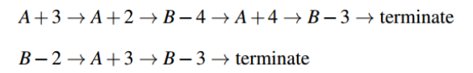
1. Explain the concept Iterative Policy Evaluation with the help of appropriate diagram.
2. Write the pseudocode for Iterative Policy Evaluation.
3. Write the Bellman update rule for Iterative Policy Evaluation and explain each term in it.
4. Explain and solve a given Grid World problem using Iterative Policy Evaluation. (3 iterations)



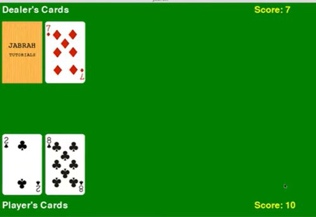
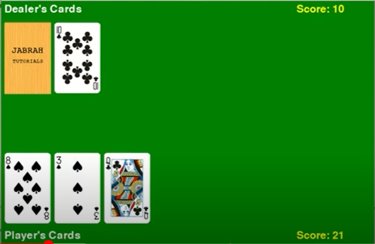
1. Explain the Policy Improvement Theorem stepwise for given V\* and give the optimal policy as a result.



1. Explain the concept of Policy Iteration with an example.
2. Write the pseudocode for Policy Iteration.
3. What are the drawbacks of Policy iteration.
4. Explain the concept of Value Iteration with the help of its backup rule.
5. Explain the relation between policy and value iteration using backup diagrams.
6. What is Asynchronous DP?
7. Which method can be used to solve the Gambler’s Problem – Policy Iteration or Value Iteration? Justify.
8. What is GPI? Explain with proper diagram.
9. What is the concept of bootstrapping? Is it used in DP or in MC method?
10. What are the advantages of Monte Carlo methods over DP?
11. Write the pseudocode for First-visit Monte Carlo Prediction.
12. Calculate value of states ‘A’ and ‘B’ for given scenario (A and B are states and numbers are rewards) using Every visit MC prediction.



1. Identify the RL elements for Blackjack Problem along with the rules of the game.
2. What are the factors to be considered by Blackjack player while taking decision of Hit or Stick? For given instances of game, identify the action chosen by the player.

A) B) 

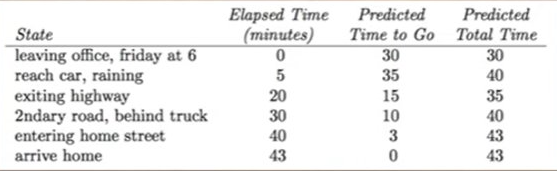
1. Explain the term usable ace with example
2. What are the assumptions made to obtain this guarantee of convergence for the Monte Carlo method?
3. Explain on-policy and off-policy control in Monte Carlo
4. How to calculate Importance sampling ratio?
5. What are the equations for state value estimate using Ordinary IS and Weighted IS? For given data of return at time steps and probability distribution for target and behavior policies (consider only three time steps), calculate state value i.e. V(s) using Ordinary IS and Weighted IS .

|  |  |  |
| --- | --- | --- |
| G1= 2 | π 1= 1/3 | µ1=4/3 |
| G2= 3 | π 2= 1/3 | µ2=1/3 |
| G3= 4 | π 3= 2/3 | µ3=2/3 |

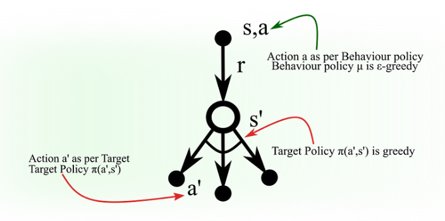
1. Derive the Incremental update rule of Monte Carlo

**MODULE 5**

1. What is temporal difference learning?
2. What are the advantages of Temporal Difference Learning over Monte Carlo methods?
3. Consider the real life example of examination patterns like annual examination and weekly examination to get the final score at the end of year. Identify the apprpriate RL approach for both patterns and justify the mapping.
4. Give the update rule for TD(1) and TD(0) algorithm and explain each term in it.
5. Define TD error.
6. Write pseudocode for TD(0) algorithm.
7. For given data of Driving Home example, explain the corrections (changes in predictions) suggested by both MC and TD approach with the help iof graphs.



1. Identify the TD algorithm for given backup diagram and also state its update rule.



1. Is SARSA an on-policy or off-policy method? Explain.
2. Explain the difference between SARSA and Q learning with the help of cliff walking problem
3. Write the pseudocode for Q learning.
4. Numericals based on Q learning and SARSA algorithm to update Q table.

**MODULE 6**

1. What are the drawbacks of tabular implementation?
2. Function approximation attempts to solve the issues of tabular implementation. Justify
3. What are the different function approximators available and which one is widely used in RL?
4. What are the advantages of function approximation?
5. How is deep learning introduced or used for reinforcement learning tasks?
6. Compare Deep Learning and Reinfoecement Learning.
7. Explain the training workflow of deep learning.
8. Name the different categories of Deep Learning.
9. What is Deep Reinforcement Learning?
10. What are the different applications of Deep Reinforcement Learning?Explain any one in brief.
11. Explain the cost function in Deep Q-Network.
12. State the challenges in Deep Q- learning and explain any one technique to address them.
13. Explain the difference between Q learning and Deep Q-learning