Roll No:.....

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -2 EXAMINATION- May 2023

M.Tech. CSE/IT 2nd Semester

COURSE CODE: 22M1WCI235 MAX. MARKS: 25

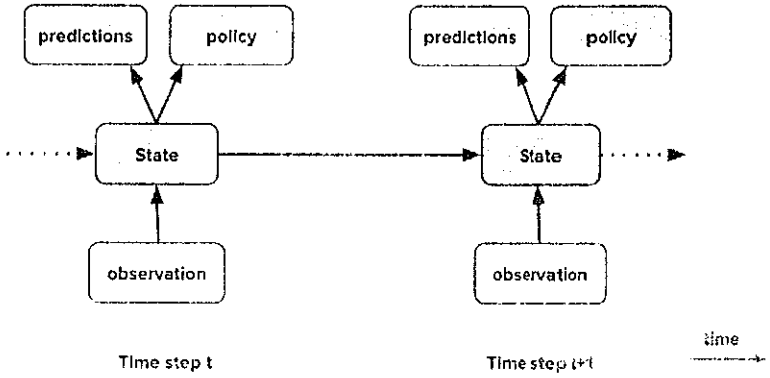
COURSE NAME: REINFORCEMENT LEARNING

COURSE CREDITS: 03 MAX. TIME: 1.5Hr

COURSE COORRDINATOR: Prof (Dr.) Vivek Kumar Sehgal

Note: All questions are compulsory. Carrying of mobile phone during examinations will be treated as case of unfair means.

1. Among the following agent components what is the role of?



a. Agent State
b. Policy
c. Value functions
d. Model

How is the history used to construct the agent state S_t? CO- 1 [5]

2. (a) Explain the following mathematical framework for Markov decision processes.

$$p(r, s \mid S_t, A_t) = p(r, s \mid \mathcal{H}_t, A_t)$$ CO- 2 [2.5]

(b) What is the difference between Fully Observable Environments and Partially Observable Environments? CO- 2 [2.5]

Roll No:.....

3. What is the role of value function in Reinforcement learning? Write the value function for:
   - Actual value function in the form of $v_\pi(s)$including discount factor $\gamma[0,1]$
   - Value Function using Bellman equation.

CO- 3 [5]

4. (a) Explain the following model which predicts what the environment will do next:

$$\mathcal{P}(s, a, s') \approx p(S_{t+1} = s' \mid S_t = s, A_t = a)$$ CO- 3 [2.5]

(b) Differentiate the following agent categories:

i. Value Based
ii. Policy Based
iii. Actor Critic CO- 3 [2.5]

5. (a) What is the role of Prediction and Control for a given policy? How these can be strongly related? CO- 3 [2.5]

(b) Explain the following Learning Agent Components:

- Policies: $\pi: \mathcal{S} \to \mathcal{A}$ (or to probabilities over $\mathcal{A}$)
- Value functions: $v: \mathcal{S} \to \mathbb{R}$
- Models: $m: \mathcal{S} \to \mathcal{S}$ and/or $r: \mathcal{S} \to \mathbb{R}$
- State update: $u: \mathcal{S} \times \mathcal{O} \to \mathcal{S}$

CO- 3 [2.5]