JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -3 EXAMINATION- December-2018

B.Tech VII Semester

COURSE CODE: 10B1WCI731   MAX. MARKS:35

COURSE NAME: Artificial Intelligence

COURSE CREDITS: 03   MAX. TIME: Two Hours

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

1. What is learning from interactions with environment? Describe reinforcement learning in this regards incorporating state transition diagram, relevant example and key components of RL in your example.   [Marks 2+2+3]

2. Define multi-arm bandit problem. Describe a) exploration dilemma, b) action value methods and c) epsilon greedy action selection.   [Marks 1+2+2+2]

3. In what condition a RL task become Markov decision process. Derive the Bellman equations.   [Marks 3+4]

4. Show how dynamic programming can be used to compute value functions and optimal policies of a Markov decision process. Explain efficiency and utility of dynamic programming for this.   [Marks 3+2+2]

5. Describe Monte Carlo policy evaluation method. Explain first visit Monte Carlo algorithm.

[Marks 3+4]