Indian Institute of Technology Kharagpur
Quiz 01 2021-22

Date of Examination: 27 Aug. 2021 **Duration:** 30 minutes
Subject No.: CS60077 **Subject:** Reinforcement Learning
Department/Center/School: Computer Science **Credits:** 3 **Full marks:** 20

Instructions

i. This question paper contains 3 pages and 4 questions. All questions are compulsory. Marks are indicated in parentheses. This question paper has been cross checked.
ii. Please write your name, roll number and date on top of the answer script.
iii. **Organize your work**, in a reasonably neat and coherent way. Work scattered all across the answer script without a clear ordering will receive very little marks.
iv. **Mysterious or unsupported answers will not receive full marks**. A correct answer, unsupported by calculations, explanation, will receive no marks; an incorrect answer supported by substantially correct calculations and explanations may receive partial marks.
v. In the online setting, you need to upload your answer scripts as **pdf file**. We will prefer a single pdf file. If you happen to have multiple files, please zip them and then uplaod as a single file. You can scan your worked out example or you can use latex to produce the pdf.

---

1. (2 points) In a behavioral experiment a group of subjects suffering from depression is provided with sweetened softdrinks that has anti-depression drug in it. For the first few weeks, it is noticed that symptoms of depression improve significantly. Then the drug is removed from the soft-drink but the sweetening agent remains. On continuation of the experiment, it is noticed that the symptoms are still improved when the patients consume the soft-drink. Which of the following is the conditioned stimulus in this experiment?

   A. The anti-depression drug B. The sweetened soda C. Depression D. The removal of depression

   'B' as the conditioned stimulus is a stimulus that gets paired with an unconditioned stimulus.

2. (2 points) Suppose, in the first lecture of this course you logged in to the meeting using Zoom in browser. But you got disconnected frequently and audio-visual quality was poor. So you made it a point to install the Zoom app to avoid such disturbance. This behavior can be best described in terms of

   A. Positive Reinforcement B. Negative Reinforcement C. Stimulus omission D. Negative punishment

   'B' as it is the removal of aversive consequence

3. (8 points) An RL agent has to choose among $K+1$ slot machines. The probability to win on the i^{th} machine is i/K. The agent randomly selects a machine and repeatedly plays on it for $n+1$ times. If it wins for all the first n times, what is the conditional probability that it will win the $(n+1)^{th}$ time?

   Let us define some of the events as follows:

M_i : be the event that the agent selects the i^{th} machine, $i = 0, 1, 2, ..., K$.
F_n : be the event that the agent wins for the first n times.
W : be the event that the agent wins the $(n+1)^{th}$ time.

So, the desired probability becomes,

$$P(W|F_n) = \sum_{i=0}^{K} P(W|F_n M_i) P(M_i|F_n) \tag{1}$$

Now, given that the i^{th} machine is choosen, it is reasonable to assume that the outcomes (win/lose) will be conditionally independent, with each one having a win probability of i/K. Hence,

$$P(W|F_n M_i) = P(W|M_i) = \frac{i}{K} \tag{2}$$

Now,

$$P(M_i|F_n) = \frac{P(M_i F_n)}{P(F_n)} = \frac{P(F_n|M_i)P(M_i)}{\sum_{j=0}^{K} P(F_n|M_j)P(M_j)} = \frac{(i/K)^n[1/(K+1)]}{\sum_{j=0}^{K}(j/K)^n[1/(K+1)]} = \frac{i^n}{\sum_{j=0}^{K} j^n} \tag{3}$$

Hence, from the equations (1), (2), and (3),

$$P(W|F_n) = \frac{1}{K} \sum_{i=0}^{K} \frac{i^{(n+1)}}{\sum_{j=0}^{K} j^n} \tag{4}$$

4. (8 points) Suppose that the joint density of X and Y is given by:

$$f(x,y) = \frac{e^{-\frac{x}{y}} e^{-y}}{y}; \qquad 0 < x < \infty, \quad 0 < y < \infty.$$

Compute $\mathbb{E}[X|Y = y]$, i.e. the conditional expectation of X given $Y = y$.

Since,

$$\mathbb{E}[X|Y] = \int x f(x|y)\, dx$$

Computing $f(x|y)$,

$$\begin{aligned} f(x|y) &= \frac{f(x,y)}{f_Y(y)} \\ &= \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)\, dx} \\ &= \frac{(1/y)e^{-x/y}e^{-y}}{\int_0^{\infty} (1/y)e^{-x/y}e^{-y}\, dx} \\ &= \frac{1}{y} e^{-x/y} \end{aligned}$$

Finally,

$$\begin{aligned}
\mathbb{E}[X|Y] &= \int x f(x|y)\, dx \\
&= \int_0^\infty \frac{x}{y} e^{-x/y}\, dx \\
&= -xe^{-x/y}\Big|_0^\infty - \int_0^\infty -ye^{-x/y}\, dx \text{ [Using Integration by parts} \\
&= y
\end{aligned}$$