

A?

HomeDashboardMy own coursesSchoolsCourse feedbackService LinksIntelliboard

?

AssignmentsForumsQuizzesResources

Course feedback

Syllabus

ELEC-E8125 - Reinforcement learning D, Lecture, 4.9.2023-29.11.2023

departm... / Sections / assignm... / Quiz 7

Started on	Tuesday, 31 October 2023, 5:31 PM
State	Finished
Completed on	Tuesday, 31 October 2023, 5:55 PM
Time taken	23 mins 45 secs
Grade	10.00 out of 10.00 (100%)

Question 1

Flag questionMark 1.00 out of 1.00Correct

The multi-armed bandit arm with the highest value is always known.

Select one:

☐ True

☒ False

Yes, in a multi-armed bandit arm values are not always known but must be learned.
The correct answer is 'False'.

Question 2

Flag questionMark 1.00 out of 1.00Correct

Multi-armed UCB bandit chooses the arm

☐ a. which has the highest value when summing an exploitation and and exploration term

☐ b. which has the highest probability to be the best arm

☐ c. by sampling an arm from a probability distribution that specifies a probability for each arm to be the best one

Your answer is correct.
The correct answer is:
which has the highest value when summing an exploitation and and exploration term

Question 3

Flag questionMark 1.00 out of 1.00Correct

Thompson Sampling chooses the multi-armed bandit arm

☐ a. which has the highest probability to be the best arm

☐ b. by sampling an arm from a probability distribution that specifies a probability for each arm to be the best one

☐ c. which has the highest value when summing an exploitation and an exploration term

Your answer is correct.
The correct answer is:
by sampling an arm from a probability distribution that specifies a probability for each arm to be the best one

Question 4

Flag questionMark 1.00 out of 1.00Correct

In a Bernoulli bandit, the reward is binary.

Select one:

☐ True

☐ False

Yes, this is correct.
The correct answer is 'True'.

Question 5

Flag questionMark 1.00 out of 1.00Correct

There is a multi-armed bandit with three arms. You have collected information and inferred that the arms yield rewards according to Gaussian distributions each arm following a Gaussian probability distribution as follows:
Arm 1: $N(0, 20)$
Arm 2: $N(2, 10)$
Arm 3: $N(1, 1)$
where in $N(\mu, \sigma^2)$, μ denotes the mean and σ^2 the variance of a Gaussian.
Which arm would a greedy policy choose?

☐ a. Arm 2: $N(2, 10)$

☐ b. Arm 3: $N(1, 1)$

☐ c. Arm 1: $N(0, 20)$

Your answer is correct.
The correct answer is:
Arm 2: $N(2, 10)$

Question 6

Flag questionMark 1.00 out of 1.00Correct

For a three-armed bandit, which arm would the UCB rule in Sutton & Barto Ch. 2.7 choose with $c = 2$ at time step $t = 15$, when the following statistics have been collected for each arm:

☐ a. $Q_t = 4, N_t = 12$

☐ b. $Q_t = 2, N_t = 1$

☐ c. $Q_t = 3, N_t = 2$

Yes, plugging in the values into the UCB rule $Q_t + 2 * \sqrt{\ln(15) / N_t}$ yields roughly 5.33 for $Q_t = 3, N_t = 2$

Your answer is correct.
The correct answer is:
 $Q_t = 3, N_t = 2$

Question 7

Flag questionMark 1.00 out of 1.00Correct

Monte Carlo tree search (MCTS) usually uses Monte Carlo value estimates.

Select one:

☐ True

☐ False

Yes, Monte Carlo value estimation is typical for MCTS
The correct answer is 'True'.

Question 8

Flag questionMark 1.00 out of 1.00Correct

Monte Carlo tree search (MCTS) uses always UCB at the root tree node and Thompson Sampling at the leaf tree nodes.

Select one:

☐ True

☒ False

Yes, there is no general requirement like this for MCTS although there can be different varieties of MCTS methods
The correct answer is 'False'.

Question 9

Flag questionMark 1.00 out of 1.00Correct

Which of the four (select 0 to 4 choices) operations are usually executed in an MCTS iteration according to Sutton & Barto?

☒ a. Backup

☒ b. Selection

☒ c. Expansion

☒ d. Simulation

Yes, an MCTS iteration consists of Selection, Expansion, Simulation, and Backup
Yes, an MCTS iteration consists of Selection, Expansion, Simulation, and Backup
Yes, an MCTS iteration consists of Selection, Expansion, Simulation, and Backup
Yes, an MCTS iteration consists of Selection, Expansion, Simulation, and Backup

Your answer is correct.
The correct answers are:
Selection,
Expansion,
Simulation,
Backup

Question 10

Flag questionMark 1.00 out of 1.00Correct

MCTS generates a search tree where all tree branches have equal depth / the same horizon / same number of time steps

Select one:

☐ True

☒ False

Yes, MCTS explores "promising" actions more resulting in a search tree where some branches are deeper
The correct answer is 'False'.

Finish review

Previous activity

Next activity

Quiz 6

Quiz 8