
Chapter 8: Rewards for Robots

Please ...

1. Read “Chapter 8: Rewards for Robots” of Melanie Mitchell’s “Artificial Intelligence: A Guide for Thinking Humans” book.
2. With respect to the questions presented for the “Chapter 8: Rewards for Robots” reading, construct a document containing question/answer pairs, one pair of each question, where the answers, with perhaps just a small number of exceptions, are more or less lifted from Melanie Mitchell’s text. Save your document as a **pdf** file.
3. Post your question/answer document to your web worksite.
4. Do your best to internalize your question/answer pairs in some sort of semantic sense, so that the answers are likely to come back to you when prompted by the questions.
5. Come to class for the discussion of “Chapter 8: Rewards for Robots,” when the time rolls around, prepared to participate in the discussion.
6. Please do all of this within one week of the “distribution” of this assignment.

Chapter 8: Rewards for Robots - Questions

1. What is the primary method used by animal trainers?
2. What is meant by the term “operant conditioning?”
3. TRUE/FALSE - Operant conditioning inspired an important machine-learning approach called reinforcement learning.
4. TRUE/FALSE - Reinforcement learning requires labeled training examples.
5. TRUE/FALSE - In reinforcement learning, an *agent* – the learning program – performs *actions* in an *environment* (usually a computer simulation) and occasionally receives *rewards* from the environment. These intermittent rewards are the only feedback the agent uses for learning.
6. TRUE/FALSE - The technique of reinforcement learning is a relatively new addition to the AI toolbox.
7. TRUE/FALSE - Reinforcement learning played a central role in the program that learned to beat the best humans at the complex game of Go in 2016.

8. In just a few sentences, describe the “illustrative example” that MM used to communicate the basic concepts associated with reinforcement learning, in general, and the variant of reinforcement learning known as Q Learning, in particular.
9. TRUE/FALSE - The promise of reinforcement learning is that the agent can learn flexible strategies on its own simply by performing actions in the world and occasionally receiving rewards (that is, *reinforcement*) without humans having to manually write rules or directly teach the agent every possible circumstance.
10. TRUE/FALSE - In general, the **state** of an agent in a reinforcement learning situation is the agent’s perception of its current situation.
11. TRUE/FALSE - A crucial notion in reinforcement learning is that of the *value of performing a particular action in a given state*.
12. In reinforcement learning, what is the *value* of action A in state S?
13. What is the “Q-table” in Q-learning?
14. Why the name “Q-learning?”
15. The Q-learning manifestation of reinforcement learning is a process that iterates over “episodes” until the learning is accomplished. What is an **episode** in this learning technique?
16. List a couple of issues, other than the “exploration versus exploitation balance” issue, that reinforcement-learning researchers face for complex tasks.
17. Deciding how much to *explore* new actions and how much to *exploit* (that is, stick with) tried-and-true actions is called the exploration versus exploitation balance. Achieving the right balance is a core issue for making reinforcement learning successful. **What real world example does MM use to illustrate the exploration versus exploitation balance?**
18. MM identifies two “stumbling blocks” to using reinforcement learning in the real world. Please briefly describe each of these stumbling blocks.