# Reading Assignment from Melanie Mitchell's "Artificial Intelligence: A Guide for Thinking Humans"
## Assignment by Carrie Corcoran

## Chapter 8: Rewards for Robots

1. What is the primary method used by animal trainers?

The primary method used by animal trainers is operant conditioning.

2. What is meant by the term "operant conditioning?"

Operant conditioning is a style of training where the desired behavior is rewarded, and all other behaviors are ignored.

3. TRUE/FALSE - Operant conditioning inspired an important machine-learning approach called reinforcement learning.

True

4. TRUE/FALSE - Reinforcement learning requires labeled training examples.

True

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.

True

6. TRUE/FALSE - The technique of reinforcement learning is a relatively new addition to the AI toolbox.

True

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.

True

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.

MM uses the example of a robot dog learning to kick a ball. The robot, named Rosie, is placed at varying distances from a ball, and has actions such as moving forward, moving backward, and kicking the ball. Rosie is rewarded for kicking the ball.

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.

True

10. TRUE/FALSE - In general, the state of an agent in a reinforcement learning situation is the agent's perception of its current situation.

True

11. TRUE/FALSE - A crucial notion in reinforcement learning is that of the value of performing a particular action in a given state.

True

12. In reinforcement learning, what is the value of action A in state S?

The value of Action A in state S is a prediction that the given action will eventually result in a reward, given future correct actions.

13. What is the "Q-table" in Q-learning?

The Q table is a record of all possible, states, actions, and their values.

14. Why the name "Q-learning?"

This is because the letter V was used for something else in the original paper, so the term Q-Table was used.

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?

An episode is a series of steps that the robot takes until it performs the desired behavior and is rewarded.

16. List a couple of issues, other than the "exploration versus exploitation balance" issue, that reinforcement-learning researchers face for complex tasks.

One issue is that the state of the world is also significantly more difficult to quantify than it is in the given example.  Another issue is that the effects of a given action can also be uncertain.

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?

MM uses the example of going to a restaurant and deciding between ordering something you know you'll like and ordering something new that you may or may not like.

18. MM identifies two "stumbling blocks" to using reinforcement learning in the real world. Please briefly describe each of these stumbling blocks.

The first is that real-world tasks are exponentially more complicated than the example, so creating a Q table for most problems isn't feasible.  The second is that having a robot carry out the given tasks repeatedly in order to learn can be time-consuming and even dangerous.