Chapter 3 Exercises

Kyler Krenzke

1. **Exercise 3.1: Devise three example tasks of your own that fit into the MDP framework, identifying for each its states, actions, and rewards. Make the three examples as different from each other as possible. The framework is abstract and flexible and can be applied in many different ways. Stretch its limits in some way in at least one of your examples.**

   (a) *Resistance training routine*: The state would be defined over the hidden variables of the human musculature. The actions would be sets/exercises to try to increase your muscular strength, endurance, size, etc. Rewards would be given proportional to the individual's increase in exercise benchmarks (i.e. bench press 1rm going up by 5lb could reward +5)

   (b) *Rocket League player*: The state space for Rocket League is defined as the position and velocity vectors for all the players and the ball, as well as game information such as the score and time remaining. The action space is the possible control inputs that can be used. The rewards would be given for winning the game.

   (c) *TV programming scheduler*: The state space is a set of possible TV show episodes, advertisers/commercials, and time slots to fit it all in. The action space is filled with selections of certain shows or advertisements at different times. Rewards could be given proportional to the number of total view time (viewers * time watched) to maximize advertising revenue.

2. **Exercise 3.2: Is the MDP framework adequate to usefully represent all goal-directed learning tasks? Can you think of any clear exceptions?**

   I cannot, lol

3. **Exercise 3.3: Consider the problem of driving. You could define the actions in terms of the accelerator, steering wheel, and brake, that is, where your body meets the machine. Or you could define them farther out—say, where the rubber meets the road, considering your actions to be tire torques. Or you could define them farther in—say, where your brain meets your body, the actions being muscle twitches to control your limbs. Or you could go to a really high level and say that your actions are your choices of where to drive. What is the right level, the right place to draw the line between agent and environment? On what basis is one location of the line to be preferred over another? Is there any fundamental reason for preferring one location over another, or is it a free choice?**

   The decision of where to define the action space depends on what the agent has absolute control over.