

APS360 Reinforcement Learning Problem Setup Worksheet

For each of the RL Problems, what information should be in the **state** provided to the agent? What possible **actions** can the agent choose at each time step? What should the **reward** be?

Problem:	State	Actions	Reward
Go			
Mario			
Breakout			
Biped			

APS360 Policy- vs Value-Based Model Worksheet

Suppose we want to train an agent to play the game of tic-tac-toe.

How would you represent the **state** of the game?

How would you represent the **actions** of the program?

Describe the architecture of a **policy-based**, neural-network agent for this problem. What are the shapes of its input layer and output layer? How would you interpret the output layer activations?

Describe the architecture of a **value-based**, neural-network agent for this problem. What are the shapes of its input layer and output layer? How would you interpret the output layer activations?