**Graded Discussion 9**

Reinforcement learning models are trained to maximize a reward function. The model uses experience to learn policies which have increased reward function values, known as improved policies. However, for many real-world problems, creating a good reward function can be difficult. These difficulties can include defining a single task the RL agent can perform, rewarding only the desired behavior and not encouraging spurious behavior, and maintaining safety when encountering unexpected state or incomplete state information.

Think of an RL application you might like to create. What reward function could you use to ensure only desired behavior? How can prevent unexpected behavior, including unsafe behavior, when the RL agent is presented with unusual or incomplete state information?