Project Description

Using your favorite programming environment, write reinforcement learning algorithms that simulate the three-state, two action Markov decision problem from Figure 10.5.

Implement this two ways: using model-free reinforcement learning (either Q-learning or the actor-critic, from section 10.4), and also using model-based reinforcement learning (section 10.4.2).

For each algorithm, simulate learning the task shown in Figure 10.5A – while the modeled rat is hungry but not thirsty. Then make him thirsty but not hungry, substituting the reward function shown in Figure 10.5 (and updating the model-based learner's reward function accordingly), and simulate additional trials.

How do the two algorithms compare in their behavior on the different phases?

Finally, simulate a rat that runs both model-based and model-free learning algorithms in parallel. For this, you need some mechanism for *arbitration* – deciding which algorithm's output controls the simulated animal's actions. Design and implement some simple arbitration mechanism, guided by the classic empirical finding that animals tend to rely on model-based learning initially, and transition to model-free control after overtraining. Simulate again the experiment from Figure 10.5, using more or less initial training to explore this change in control.