1) PREDICTION: $Q(s_t)$

In this step the agent knows the action values from the current state s_t, and we predict the Q-values $Q(s_t)$.

2) PREDICTION: $Q'(s_{t+1})$

In this step the we predict the Q-values $Q'(s_{t+1})$ taking into consideration that the agent know the action values starting from the state s_{t+1}.

3) TRAINING

The input of this step is the current state (s_t) and the updated $Q(s_t,a_t)$ contains the maximum expected future reward and this is the desired output.

4) UPDATE: $Q(s_t,a_t)$

Now we will update $Q(s_t,a_t)$, which represents the value of a specific action (a_t) during simulation.