

Deep Q Learning: From Paper to Code

Coding The Dueling Deep Q Agent Class

What We Need to Accomplish

- Memory, target network updating, model saving, and decrementing epsilon are unchanged
- Action selection just needs advantage stream
- For learn, have to combine V, A streams

$$Q(s, a; \theta, \alpha, \beta) \stackrel{\text{def}}{=} V(s, \theta, \beta) + (A(s, a; \theta, \alpha) - \frac{1}{|A|} \sum_{a'} A(s, a'; \theta, \alpha))$$