

Standard TD

$$\delta_t = r_t + V_{TD}(s_{t+1}) - V_{TD}(s_t)$$

$$V_{TD}(s) \xleftarrow{\text{Update target}} V_{TD}(s) + \alpha_{TD} \overset{\text{Prediction}}{\delta_t}$$

Dual Process

$$V_{NET}(s_{t+1}) = k V_{TD}(s_t) + (1 - k) V_{MB}(s_{t+1})$$

$$\delta_t = r_{t+1} + V_{NET}(s_{t+1}) - V_{TD}(s_t)$$

$$V_{TD}(s) \xleftarrow{\text{Update target}} V_{TD}(s) + \alpha_{TD} \overset{\text{Prediction}}{\delta_t}$$