

更新后的状态值 \leftarrow 当前状态值 + 学习率(奖励 + 折扣因子 \times 下一状态值 - 当前状态值)

实际值

预测值

时序差分误差