Equation sheet:

$$V^*(s) = \max_a Q^*(s, a) \qquad \text{(Bellman equations)}$$

$$Q^*(s, a) = \sum_{s'} T(s, a, s')[R(s, a, s') + \gamma V^*(s')]$$

$$V^*(s) = \max_a \sum_{s'} T(s, a, s')[R(s, a, s') + \gamma V^*(s')]$$

$$\pi^*(s) = \operatorname{argmax}_a \sum_{s'} T(s, a, s')[R(s, a, s') + \gamma V^*(s')] \qquad \text{(Optimal policy)}$$

$$Q(s, a) = (1 - \alpha)Q(s, a) + \alpha(r + \gamma \max_{a'} Q(s', a')) \qquad \text{(Q-Learning)}$$

$$Q(s, a) = w_1 f_1(s, a) + w_2 f_2(s, a) + ... + w_n f_n(s, a) \qquad \text{(Approximate Q-Learning)}$$

$$\text{diff} = [r + \gamma \max_{a'} Q(s', a')] - Q(s, a)$$

$$w_i = w_i + \alpha[\text{diff}] f_i(s, a)$$