

$$a^* = \arg \max_{a_i} \{q(s_0, a_i)\} \leftarrow \text{-----} s_0$$

$$q(s_0, a_i) = \mathcal{R}(s_0, a_i) + \sum_{s_j} \mathcal{T}(s_0, a_i, s_j) h(s_j) \leftarrow \text{-----} a_1 \quad a_2$$

$$h(s_j) = \max_{a_i} \{\mathcal{R}(s_j, a_i)\} \leftarrow \text{-----} s_1 \quad s_2 \quad s_3 \quad s_4$$

