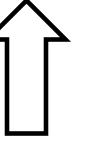
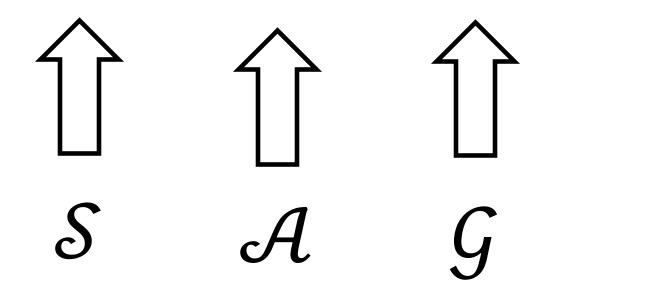
$$\mathbb{T}_Q(\mathcal{G})$$

Stage 2 – Infer



Tabular Q-learning





$$O = \langle s_0, a_0, s_1, a_1, ..., s_k, a_k \rangle$$