

$$(r + \gamma \max_{a'} Q(s', a'; \theta) - Q(s, a; \theta))$$

Objetivo

Predicción

Q'

Q

Sincronización cada
N iteraciones

Red objetivo

Red de predicción

