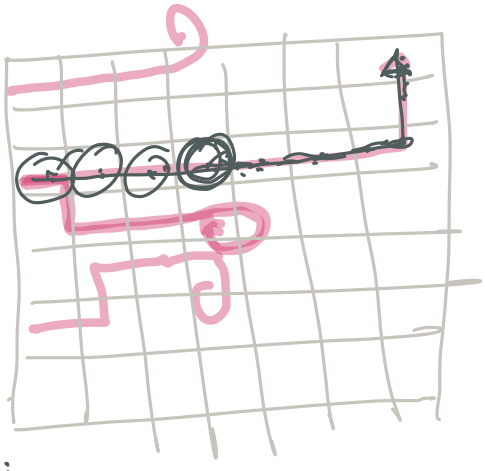

ESTIMATE $V^\pi(s)$

(1) MONTE CARLO

$$V^\pi(s) = \sum_{t=1}^{T} c(s_t, a_t)$$

Maintain a running average of value

ROLLOUTS:

$s_1, a_1, c_1, s_2, a_2, c_2, \ldots$

(2) BELLMAN !!!!

TD for $Q^\pi(s_t, a_t)$

for $(s_t, a_t, c_t, s_{t+1}, a_{t+1}) \sim \pi$

$$Q^\pi(s_t, a_t) \leftarrow Q^\pi(s_t, a_t) + \alpha\left(c(s_t, a_t) + \gamma Q^\pi(s_{t+1}, a_{t+1}) - Q^\pi(s_t, a_t)\right)$$

for any (s_t, a_t, c_t, s_{t+1})

$$Q^*(s_t, a_t) \leftarrow Q^*(s_t, a_t) + \alpha\left(c_t + \gamma \min_{a_{t+1}} Q^*(s_t, a_{t+1}) - Q^*(s_t, a_t)\right)$$