

Brainstorm-1

① Eligibility Traces - TD

→ Apply to robotics

Credit assign

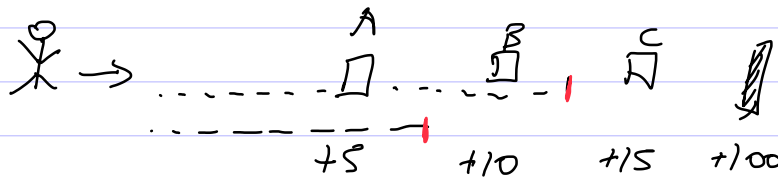
$$Z \leftarrow \gamma Z + \nabla Q$$

$$[r + \gamma \max Q - Q]$$

$$Z [r + \gamma \max Q - Q]$$

Q-Learning

3 papers!



② CQL

→ Apply to robotics

Diets. Shift!

"Learning by cheating" - offline

$$Q_{CQL} \leq Q$$

③ $r + I(x, z)$

→ x

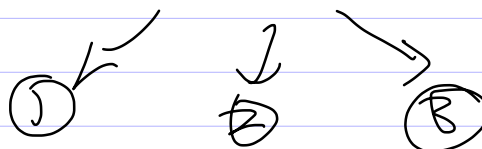
→ z

Problem?

Problem



Unif. Sol.



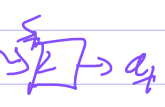
Code

① Elig. Traces

② CQL

③ Learning by Cheat.

Bin



① Run

for my intentions!

→ ① Run LC online (ξ_t, a_t, ξ_{t+1}) $\sim \underline{B_t}$.

② Train Co2+El based B_t

3 ideas → 1 big code based!

① { MuJoCo
App. DM Control

2 papers!

② Restrict Pair Inf. \rightarrow Pair. $\rightarrow \boxed{N_t} \rightarrow a_t$
 $\rightarrow \boxed{N_t} \rightarrow a_t$

SAC

Dagger!
+ LSTM

③ LQ2 $-\frac{F}{s}(\hat{v})$ $\rightarrow e^Q$

LC on policy
replay new!
 \rightarrow SAC
off policy