

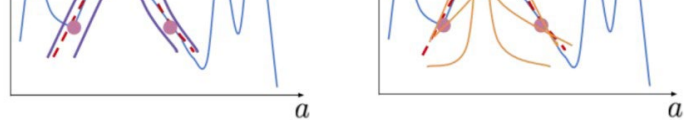
- Would you implement epsilon-greedy in a self-driving car?



Universidad
Zaragoza



(e.g.: Conservative Q-Learning)



Universidad
Zaragoza



Universidad
Zaragoza



\mathbf{a}

$\sum_{i=0}$

Bellemare*, Dabney*, Munos (ICML 2017).



Universidad
Zaragoza



Bellemare*, Dabney*, Munos (ICML 2017).



(Ian Goodfellow et al., 2014)

(Jonathan Ho & Stefano Ermon, 2016)
(Yuke Zhu et al, 2018)



improve the policy

optimize $\pi_{\theta}(\mathbf{a}|\mathbf{s})$ (model-based)

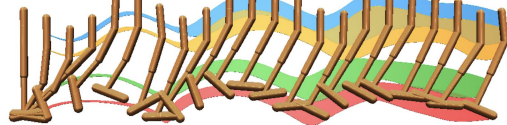


Universidad
Zaragoza

trial #1 (random actions)



Universidad
Zaragoza



(c) 300 steps

(d) 600 steps

(e) 3600 steps

(f) 12000 steps



Universidad
Zaragoza



Universidad
Zaragoza

<https://ai.googleblog.com/2023/04/robotic-deep-rl-at-scale-sorting-waste.html>



- Csaba Szepesvári, Algorithms of Reinforcement Learning

- <https://sites.ualberta.ca/~szepesva/rlbook.html>



Universidad
Zaragoza

