Author : REBOUL Sebastian

Title: « Learning Upper–Lower Value Envelopes to Shape Online RL»

Short Abstract:

We accelerate online reinforcement learning through a two-stage framework that derives flexible, decoupled upper and lower value envelopes from offline data.
Our theoretical analysis bridges pre-training and fine-tuning by establishing high-probability regret bounds based on data-driven, rather than fixed, shaping functions.
Empirical tests on tabular MDPs demonstrate that these tighter approximations yield substantial regret reductions compared to UCBVI and prior methods.

Télécom SudParis • Site d'Évry : 9 rue Charles Fourier - 91011 Évry-Courcouronnes Cedex - France • Site de Palaiseau : 19 place Marguerite Perey - 91120 Palaiseau France • Tél. +33 (0)1 60 76 40 40 • Siret : 180 092 025 00055 • NDA 11910803791 • APE : 8542Z
www.telecom-sudparis.eu