Cornell University Library

We gratefully acknowledge support from
the Simons Foundation
and member institutions

arXiv.org > stat > arXiv:1305.1027

Search or Article-id

(Help | Advanced search)

All papers   Go!

Statistics > Machine Learning

Regret Bounds for Reinforcement Learning with Policy Advice

Mohammad Gheshlaghi Azar, Alessandro Lazaric, Emma Brunskill

(Submitted on 5 May 2013 (v1), last revised 17 Jul 2013 (this version, v2))

In some reinforcement learning problems an agent may be provided with a set of input policies, perhaps learned from prior experience or provided by advisors. We present a reinforcement learning with policy advice (RLPA) algorithm which leverages this input set and learns to use the best policy in the set for the reinforcement learning task at hand. We prove that RLPA has a sub-linear regret of \tilde O(\sqrt{T}) relative to the best input policy, and that both this regret and its computational complexity are independent of the size of the state and action space. Our empirical simulations support our theoretical analysis. This suggests RLPA may offer significant advantages in large domains where some prior good policies are provided.

Download:

- PDF
- PostScript
- Other formats

Current browse context:
stat.ML
< prev | **next >**
new | recent | 1305

Change to browse by:

cs
    cs.LG
stat

References & Citations

- NASA ADS

Bookmark(what is this?)

ScienceWISE

Subjects:	**Machine Learning (stat.ML)**; Learning (cs.LG)
Cite as:	**arXiv:1305.1027** [stat.ML]
	(or **arXiv:1305.1027v2** [stat.ML] for this version)

Submission history

Which authors of this paper are endorsers?

Link back to: arXiv, form interface, contact.