Metadata

differences DRA maxEnt

Different trade-off parameter per state

One idea is that DRA (or a generalization thereof) may be more general than max entropy RL because in max entropy RL, the trade-off parameter between the reward and entropy is the same in each state, whereas DRA allows finer grained

Trajectory vs time-step

Behavioral difference in @ziebartMaximumEntropyInverse2008's task (last figure)

> ⊘ **Question:**
>
> Could this difference be resolved by using this Alternative to softmax instead of softmax for SAC?