Inverse Q-function

Topics: Reinforcement Learning, Inversion, Q-Function

Modulation is best avoided in ML. There should not be a separation between the Q_θ function (the judge) and the action function U_ϕ (the executioner).

A simple way of avoiding this separation is to approximate Q by an easily invertible parameterized function Q_θ. The ease of inversion should remain after conditioning on any state s.

If Q_θ generalizes well the mapping Q of the subspace S to the subspace T, there should be a guaranteed generalization of Q_θ^{-1} from T to S.

Then $Q_\theta^{-1}(\mathrm{High\ Return}; s)$ should yield a good action if $s \in S$.