Improving upon Iterative Approaches for Solving Fully-Specified POMDPs

Benjamin Kern

October 2019

Abstract

Markov Decision Processes (MDPs) are discrete, mathematical formulations used to simulate generalized sequential decision making. An agent acting within an MDP must optimize expected future reward by making a decision on which action to take given the current state. Partially Observable MDPs (POMDPs) are MDPs in which the state is not known by the agent, and instead the agent must act upon observations given by the environment at each time step. This work presents a novel solution that is more sample-efficient than traditional methods for fully-specified POMDPs, viz. when transition probabilities between states as well as observations are given to the agent. Currently, traditional approaches for solving POMDPs require iterative learning processes that converge slowly when not exploiting opportunities for linear parallelism. Consequently, we present a closed-form solution that is derived algebraically from traditional iterative update rules. Solving this closed-form solution yields an accelerated learning rate that enables a jump start unemployed by traditional iterative methods. Simulation results support the efficacy of this method on traditional POMDPs. Additionally, applied and theoretical impliciations of this method are discussed.