1139-68-318   **Michael N Katehakis*** (`mnk@rutgers.edu`), 100 Rockafeller Road, Piscataway, NJ 08854,
Wesley Cowan (`c.wes.cowan@gmail.com`), 100 Rockafeller Road, Piscataway, and **Daniel Pirutinsky**, 100 Rockafeller Road, Piscataway, NJ. *Reinforcement learning: connections between MDPs and MAB problems.* Preliminary report.

This talk considers a basic reinforcement learning model dealing with adaptively controlling an unknown Markov Decision Process (MDP), in order to maximize the long term expected average value. In this work, we consider a factored representation of the MDP problem that allows it to be decoupled into a set of individual MAB-problems on a state by state basis. In this way, we show sufficient conditions for efficiently extending classical MAB-type policies to corresponding MDP policies. These constructed MDP policies largely inherit the properties of their MAB-generators, allowing the simple construction of asymptotically optimal MDP policies. We additionally show the construction of a simple UCB-type MDP policy, dramatically simplifying an earlier proof of its optimality. Additional extensions to other MAB policies (e.g., Thompson Sampling) are discussed. (Received February 15, 2018)