Combinations of Optimal Policies in Unichain Markov Decision Processes are Optimal

Ronald Ortner[1] <ronald.ortner@unileoben.ac.at>

We show that combinations of optimal (stationary) policies in unichain Markov decision processes are optimal. That is, let $\mathcal{M}$ be a unichain Markov decision process with state space S, action space A and policies π_1^*, $\pi_2^* : S \to A$ with optimal average infinite horizon reward. Then all *combinations* of these policies π where for each state $s \in S$ either $\pi(s) = \pi_1^*(s)$ or $\pi(s) = \pi_2^*(s)$ are optimal as well.

[1]Montanuniversität Leoben