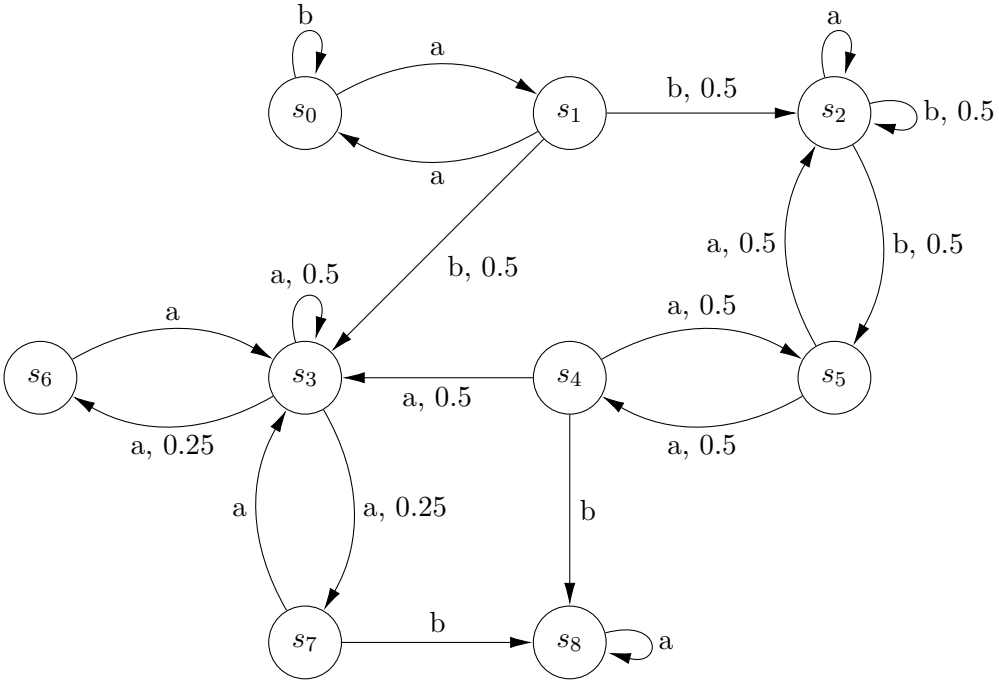
Technische Universität München
Institut für Informatik I07
Theoretical Computer Science

Quantitative Verification 12

Ex 1: MEC Decomposition

Compute the MECs (both states and actions) of the following MDP.



Ex 2: Reachability as special case

In the lecture, we saw expected step-bounded reward and expected long-run average reward. How can you rephrase (bounded) reachability as an instance of these problems?

Ex 3: Discounted Reward

As a "trade-off" between step-bounded reward and long-run average, one can define discounted reward. Given a "discount factor" γ with $0 < \gamma < 1$, the reward obtained by a single run $s_0 s_1 \cdots$ is

$$\sum_{i=0}^{\infty} \gamma^i r(s_i).$$

One can easily define a value iteration algorithm by iterating

$$x_{n+1}(s) = r(s) + \gamma \max_{a \in \mathsf{Act}(s)} \sum_{s'} P(s, a, s') x_n(s')$$

Prove that this iteration converges in the limit by using the Banach fixed-point theorem (also known as contraction mapping theorem).

Hint: Use the $\mathcal{L}^\infty$-norm.