Quantitative Verification 12 - Solutions

Ex 1: MEC Decomposition

MECs:

- $(\{s_0, s_1\}, \{(s_0, a), (s_0, b), (s_1, a)\})$
- $(\{s_2\}, \{(s_2, a)\})$
- $(\{s_3, s_6, s_7\}, \{(s_3, a), (s_6, a), (s_7, a)\})$
- $(\{s_8\}, \{(s_8, a)\})$

Ex 2: Reachability as special case

- Bounded reachability: Make target states absorbing, i.e. remove all outgoing transitions and add a self loop, set their reward to 1 and all other rewards to 0.
- Unbounded reachability: Can't be phrased easily as bounded reward problem. For unbounded, apply the above idea.

Ex 3: Discounted Reward

Define the operator $T : \mathbb{R}^S \to \mathbb{R}^S$ as the value iteration, i.e.

$$T(x)_s = r(s) + \gamma \max_{a \in \mathsf{Act}(s)} \sum_{s'} P(s, a, s') x_{s'}.$$

Clearly, the correct state values for the discounted reward are a fixed point of this operator, i.e. $T(x^\star) = x^\star$.

But this operator also is a contraction. Fix some state s and value vectors x, y. Assume w.l.o.g. that $T(x)_s \geq T(y)_s$ and let $a^\star$ be the action maximizing $T(x)_s$.

$$\begin{aligned}
T(x)_s - T(y)_s &= \gamma\left(\max_{a \in \mathsf{Act}(s)} \sum_{s'} P(s, a, s') x_{s'} - \max_{a \in \mathsf{Act}(s)} \sum_{s'} P(s, a, s') y_{s'}\right) \leq \\
&\leq \gamma\left(\sum_{s'} P(s, a^\star, s') x_{s'} - \sum_{s'} P(s, a^\star, s') y_{s'}\right) = \\
&= \gamma \sum_{s'} P(s, a^\star, s')(x_{s'} - y_{s'}) \leq \gamma \sum_{s'} P(s, a^\star, s') \|x - y\|_\infty = \gamma \|x - y\|_\infty.
\end{aligned}$$

Together, we get that $|T(x)_s - T(y)_s| \leq \gamma \|x - y\|_\infty$ and hence $\|T(x) - T(y)\|_\infty \leq \gamma \|x - y\|_\infty$. Consequently, T is a contraction w.r.t. the $\mathcal{L}^\infty$ norm and by the Banach fixed-point theorem, it (i) has a unique fixed point $x^\star$ and (ii) $T^n x_0$ converges to $x^\star$ for any x_0.