Reinforcement Learning
Exercise 2 - Solution

Jonathan Schnitzler - st166934
ErickVillanuevaVillasenor - st190300
Eric Choquet - st160996

April 26, 2024

Proofs

a) Bellman optimality operator is a gamma-contraction We want to show

$$(\mathcal{T}v)(s) = \max_a \sum_{s',r} p(s',r|s,a)[r + \gamma v(s')] \tag{1}$$

fullfills the γ-contraction property, namely

$$\|\mathcal{T}v - \mathcal{T}w\|_\infty \leq \gamma \|v - w\|_\infty \tag{2}$$

Inspired by the lecture for the Bellman expectation backup operator, we will similarly use the definition of the infinity norm to show the contraction property

$$\|\mathcal{T}v - \mathcal{T}w\|_\infty = \| \max_a \sum_{s',r} p(s',r|s,a)[r + \gamma v(s')] - \max_a \sum_{s',r} p(s',r|s,a)[r + \gamma w(s')]\| \tag{3}$$

$$\leq \| \max_a \sum_{s',r} p(s',r|s,a)[r + \gamma v(s') - (r + \gamma w(s'))]\| \tag{4}$$

$$= \gamma \| \max_a \sum_{s',r} p(s',r|s,a)[v(s') - w(s')]\| \tag{5}$$

$$\leq \gamma \| \max_a \sum_{s',r} p(s',r|s,a)\|v(s') - w(s')\|_\infty\| \tag{6}$$

$$\leq \gamma \|v - w\|_\infty \tag{7}$$