Reinforcement Learning Exercise 1

, ,

1 Proofs

a)

$$
\begin{aligned}
||\mathcal{T}v - \mathcal{T}v'||_\infty &= \max_s |\mathcal{T}v(s) - \mathcal{T}v'(s)| \\
&= \max_s |\max_a \sum_{s',r} p(s',r|s,a)[r + \gamma v(s')] - \max_a \sum_{s',r} p(s',r|s,a)[r + \gamma v'(s')]| \\
&= \gamma \max_s \max_a |\sum_{s',r} p(s',r|s,a)[(v(s') - v'(s')]| \\
&\leq \gamma \max_s |\max_s |(v(s') - v'(s')|| \\
&= \gamma ||v - v'||_\infty
\end{aligned}
$$

b)

$$
\begin{aligned}
\because v(s) &= \mathbb{E}_\pi[G_t|S_t = s] \\
&= \mathbb{E}_\pi[\sum_{i=0}^{\infty} \gamma^i R_{t+i+1}|S_t = s] \\
&= \mathbb{E}_\pi[\frac{R_{t+i+1}}{1-\gamma}|S_t = s], \\
R_{t+i+1} &\in [r_{min,r_{max}}] \\
\therefore \frac{r_{min}}{1-\gamma} &\leq v(s) \leq \frac{r_{max}}{1-\gamma} \\
\therefore ||v - v'||_\infty &= \max_s |v(s) - v(s')| \\
&= \frac{r_{max} - r_{min}}{1-\gamma} \\
\therefore |v(s) - v(s')| &\leq \frac{r_{max} - r_{min}}{1-\gamma}
\end{aligned}
$$

2 Value Iteration