Name	Student Number
Oluwatomilayo Adegbite	500569283
Nikolas Maier	500461990

Show that if $\left|V^*(s) - V_{apr}(s)\right| \leq \epsilon$ for every state s, then

$$\max_{s \in S} \{L^{\pi'}(s)\} \leq \frac{2\gamma\epsilon}{1-\gamma}$$

Let s be a state of maximum difference between $V^*(s) - V_{apr}(s)$, by some optimal action being different from the action proscribed by π.

Optimal Action y vs policy Action x

$$V^*(s) = R_s^y + \gamma \sum_{s' \in S} P_{ss'}^y V^*(s') \text{ vs}$$

$$V_{apr}(s) = R_s^x + \gamma \sum_{s' \in S} P_{ss'}^x V^{\pi'}(s')$$

If we followed the non-optimal policy π but behave according to the optimal policy following that we would have a difference of $\gamma(V^*(s) - V_{apr}(s))$.

For ease of use, we shall say Δ = $V^*(s) - V_{apr}(s)$, so the above will be (ɣ * Δ), which is worse than our V^* by (Δ + ɣ*Δ) ie the difference between V* and $V_{apr}(s)$ is the normal discounted value (ɣ) vs the normal discounted value times the less optimal option (ɣ*Δ).

Given that between the two actions we know the optimal action is y, the following is still true.

$$(\Delta + \gamma^*\Delta) \leq (R_s^y + \gamma \sum_{s' \in S} P_{ss'}^y V^*(s')) - (R_s^x + \gamma \sum_{s' \in S} P_{ss'}^x V^*(s'))$$

We also need to remember that since action x looked better according to $V_{apr}(s)$ that we also have the following:

$$0 \geq (R_s^y + \gamma \sum_{s' \in S} P_{ss'}^y V_{apr}(s')) - (R_s^x + \gamma \sum_{s' \in S} P_{ss'}^x V_{apr}(s'))$$

When we consider these two equations and subtract them we cancel out a bunch of them

$$(\Delta + \gamma^*\Delta) \leq (R_s^y + \gamma \sum_{s' \in S} P_{ss'}^y V^*(s')) - (R_s^x + \gamma \sum_{s' \in S} P_{ss'}^x V^*(s')) -$$

$$(R_s^y + \gamma \sum_{s' \in S} P_{ss'}^y V_{apr}(s')) - (R_s^x + \gamma \sum_{s' \in S} P_{ss'}^x V_{apr}(s'))$$

$$(\Delta + \gamma^*\Delta) \leq (\cancel{R_s^y} + \gamma \sum_{s' \in S} P_{ss'}^y V^*(s')) - (\cancel{R_s^x} + \gamma \sum_{s' \in S} P_{ss'}^x V^*(s')) -$$

$$(\cancel{R_s^y} + \gamma \sum_{s' \in S} P_{ss'}^y V_{apr}(s')) - (\cancel{R_s^x} + \gamma \sum_{s' \in S} P_{ss'}^x V_{apr}(s'))$$

Keeping in mind that we know $V^*(s) - V_{apr}(s) <= \epsilon$

$$(\Delta + \gamma^*\Delta) \leq \gamma \sum_{s' \in S} P^{y}_{ss'} \epsilon + \gamma \sum_{s' \in S} P^{x}_{ss'} \epsilon$$

The RHS can't be bigger than 2 *ɣ*ε

$\Delta(1 + \gamma) \leq 2 {}^*\gamma^*\varepsilon$

$\Delta \leq (2 {}^*\gamma^*\varepsilon)/(1 + \gamma)$