Notes on the Bellman Expectation Equation

In the previous video, we derived one equation for each environment state. For instance, for state s_1, we saw that:

$$v_\pi(s_1) = \frac{1}{2}(-1 + v_\pi(s_2)) + \frac{1}{2}(-3 + v_\pi(s_3)).$$

We mentioned that this equation follows directly from the Bellman expectation equation for v_π.

> $v_\pi(s) = \mathbb{E}_\pi[R_{t+1} + \gamma v_\pi(S_{t+1})|S_t = s] = \sum_{a \in \mathcal{A}(s)} \pi(a|s) \sum_{s' \in \mathcal{S}, r \in \mathcal{R}} p(s', r|s, a)(r$
> (**The Bellman expectation equation for v_π**)

In order to see this, we can begin by looking at what the Bellman expectation equation tells us about the value of state s_1 (where we just need to plug in s_1 for state s).

$$v_\pi(s_1) = \sum_{a \in \mathcal{A}(s_1)} \pi(a|s_1) \sum_{s' \in \mathcal{S}, r \in \mathcal{R}} p(s', r|s_1, a)(r + \gamma v_\pi(s'))$$

Then, it's possible to derive the equation for state s_1 by using the following:

- $\mathcal{A}(s_1) = \{\text{down}, \text{right}\}$ (*When in state s_1, the agent only has two potential actions: down or right.)*
- $\pi(down|s_1) = \pi(\text{right}|s_1) = \frac{1}{2}$ (*We are currently examining the policy where the agent goes down with 50% probability and right with 50% probability when in state s_1.*)
- $p(s_3, -3|s_1, \text{down}) = 1$ (and $p(s', r|s_1, \text{down}) = 0$ if $s' \neq s_3$ or $r \neq -3$) (*If the agent chooses to go down in state s_1, then with 100% probability, the next state is s_3, and the agent receives a reward of -3.*)
- $p(s_2, -1|s_1, \text{right}) = 1$ (and $p(s', r|s_1, \text{right}) = 0$ if $s' \neq s_2$ or $r \neq -1$) (*If the agent chooses to go right in state s_1, then with 100% probability, the next state is s_2, and the agent receives a reward of -1.*)
- $\gamma = 1$ (*We chose to set the discount rate to 1 in this gridworld example.*)

If this is not entirely clear to you, please take the time now to plug in the values to derive the equation from the video. Then, you are encouraged to repeat the same process for the other states.