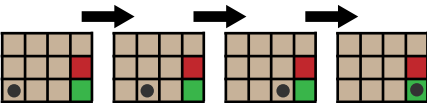
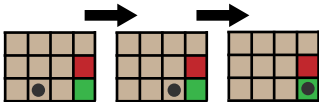


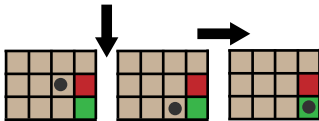
Policy (π): Considering: $\gamma=1$



$$V^{\pi}(s_1): -2$$



$$V^{\pi}(s_2): -1$$



$$V^{\pi}(s_6): -1$$

- $V^{\pi}(s)$ depends of policy (π) and initial state (s_t).
- Agent needs to start in $(s_{3,4,5,7,8,9,10})$ to know its value $V^{\pi}(s)$.