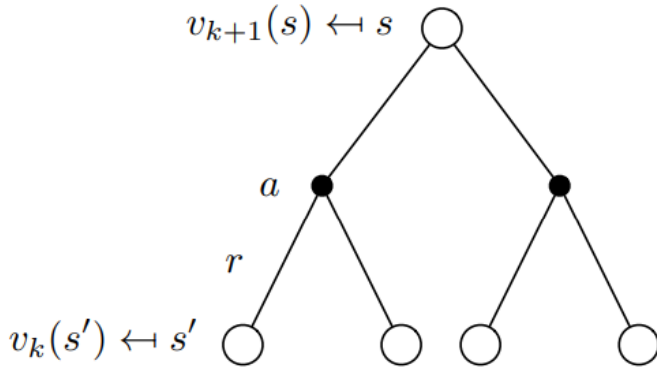


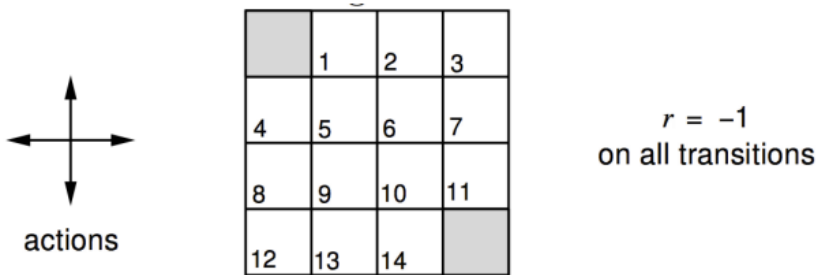
CSSE490 DRL - Quiz: Dynamic Programming

Notation	Description
\mathcal{R}_s	reward at state s
$\mathcal{P}_{ss'}^a$	environment transition probability from state s to the next state s' when taking action a
$\pi(a s)$	policy: the probability of taking action a at the state s



1. **Policy Evaluation:** Based on the diagram above, write the update equation of $v(s)$ in terms of $v(s')$. Note: we know $\pi(a|s)$ in this case.

2. The policy is uniform random in all possible direction, we want to evaluate this policy by calculating $v(s)$ for all states.

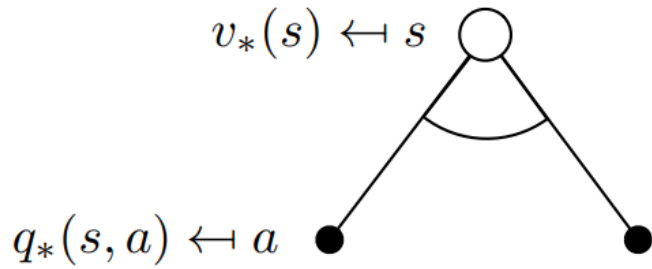


Given the result of the first iteration as shown below, calculate the result of the next iteration, i.e., $k = 2$.

$k = 1$

0.0	-1.0	-1.0	-1.0
-1.0	-1.0	-1.0	-1.0
-1.0	-1.0	-1.0	-1.0
-1.0	-1.0	-1.0	0.0

3. **Value Iteration:** Given that $v_*(s) = \max_a q_*(s, a)$, write the update equation of $v_*(s)$ in terms of $q_*(s, a)$.



4. **Value Iteration:** Based on the top, write the update equation of $v_*(s)$ in terms of $v_*(s')$. Note: we don't have $\pi(a|s)$ in this case.

