

K=0

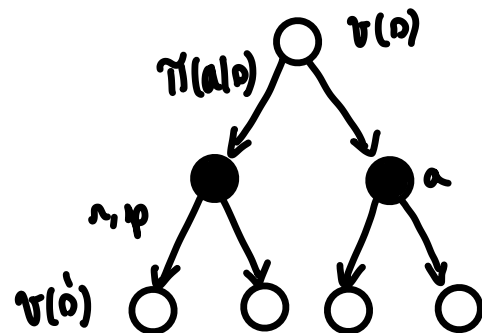
0,0 1	-1 2	-2 3	-2 4
-2 5	-3 6	-3 7	-2 8
-2 9	-3 10	-3 11	-1 12
-2 13	-2 14	-1 15	0,0 16

K=1

0,0 1			
	X 6		
			0,0 16

Policy 0

0,0 1	↕↕ 2	↕↕ 3	↕↕ 4
↕↕ 5	↕↕ 6	↕↕ 7	↕↕ 8
↕↕ 9	↕↕ 10	↕↕ 11	↕↕ 12
↕↕ 13	↕↕ 14	↕↕ 15	0,0 16



$$v_1(6) = \max_a \sum_{s'} p(s'|a,6) [r + \gamma v_0(s')]$$

$$v_1(6) = \max_a \begin{cases} p(2|a,6) [-1 - 1] \leftarrow \text{Max} \\ p(10|a,6) [-1 - 3] \\ p(7|a,6) [-1 - 3] \\ p(5|a,6) [-1 - 2] \end{cases}$$

$$v_1(6) = -2$$