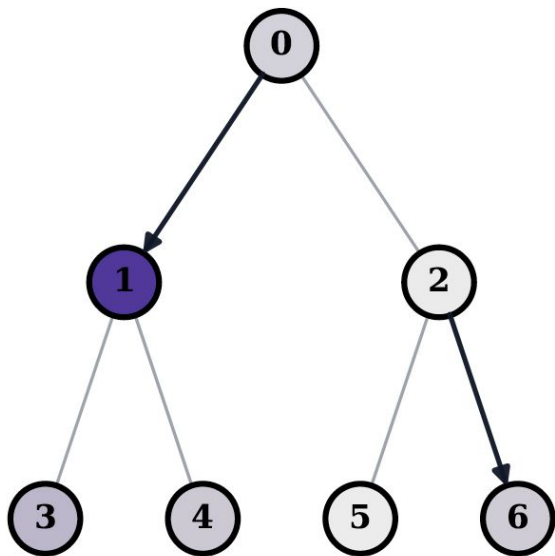


Result A: Example optimal policy and MDP

actions at all nodes = {left child, right child, stay}



near deterministically (>90%) transition from 0 to 1 if take left child, 0 to 2 if take right child.

taking action left child or right child only succeeds <10% of the time; otherwise, stays stationary with ~90% probability

Result B: Example optimal policy and MDP

