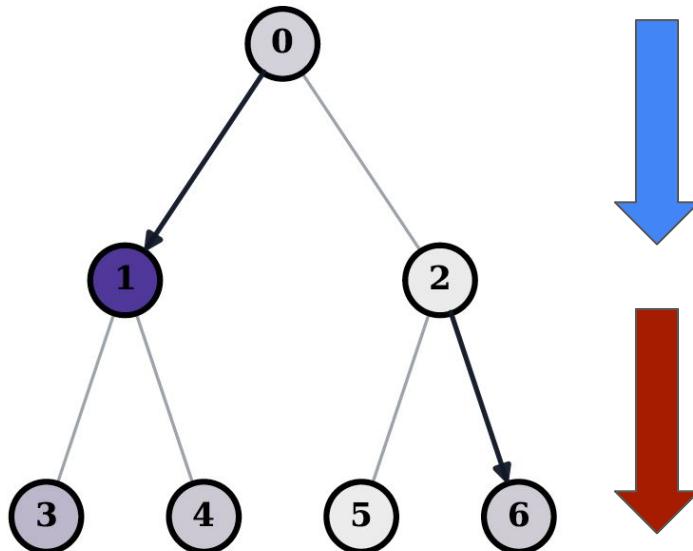


Result A: Example optimal policy and MDP

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



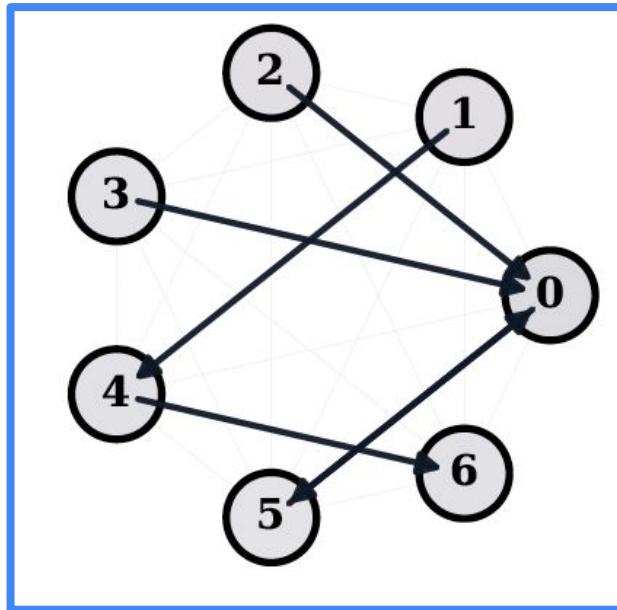
near deterministically (>90%) transition from 0 to 1 if take left child, 0 to 2 if take right child.

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

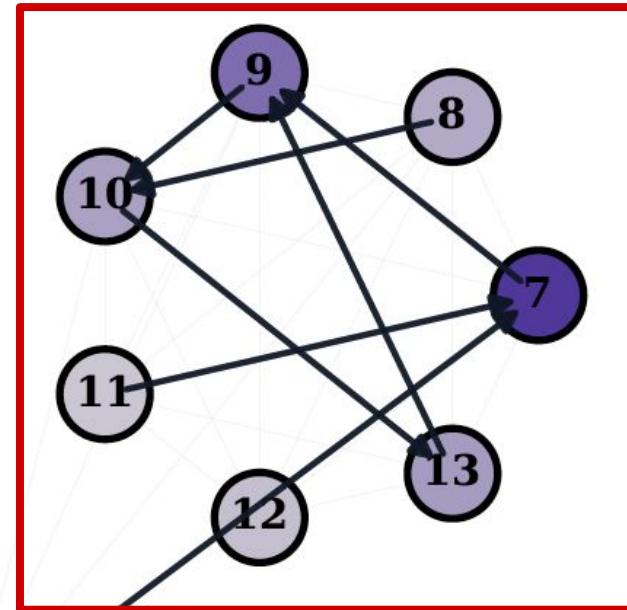
dynamic left side: cannot stay at current state,
transitions almost deterministically to specified node

actions at $i = \{0, 1, 2, 3, 4, 5, 6\} \setminus i \cup \{\text{go to random other node}\}$



"sticky" right side: 90% probability of staying at
current state regardless of action

actions at $i = \{7, 8, 9, 10, 11, 12, 13\}$



actions at 14 = {go to left side, go to right side}

