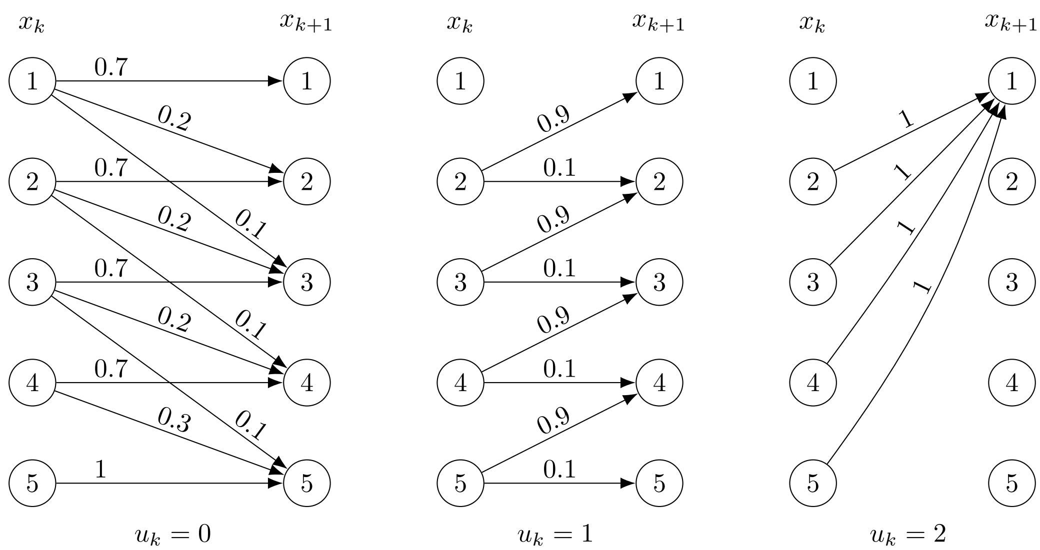
Example – Machine repair scheduling (revisited)

Same as in the previous case, we have a machine that we use to print money. The condition of this machine can be in one of five states: {new, good, normal, bad, broken}, which naturally deteriorate over time. The amount of money the machine prints depends on its condition. If the machine is not in perfect condition (new), we can schedule a repair, that should put the machine into a better state, or we can buy a new machine. Our task is to come up with a plan of making repairs and buying new machines to maximize our profit, this time over an infinite time period.

- x_0: Initial machine state, $x_0 \in \{\text{new (1), good (2), normal (3), bad (4), broken (5)}\}$,
- α: Discount factor,
- x_k: State of the machine at time $k = 1, \ldots, N$,
- u_k: Chosen action at time $k = 0, \ldots, N-1$, $u_k \in \{\text{do nothing (0), repair (1), buy new (2)}\}$
- w_k: Machine deterioration at time k, depends on x_k and u_k.



Amount of money printed (g^p) and cost of actions (g^a) at k:

$$g^p(x_k) = \begin{cases} 1, & \text{for } x_k = 1, \\ 0.9, & \text{for } x_k = 2, \\ 0.7, & \text{for } x_k = 3, \\ 0.5, & \text{for } x_k = 4, \\ 0, & \text{for } x_k = 5, \end{cases} \qquad g^a(u_k) = \begin{cases} 0, & \text{for } u_k = 0, \\ -0.4, & \text{for } u_k = 1, \\ -1, & \text{for } u_k = 2. \end{cases}$$

Profit at time k: $g_k(x_k, u_k) = g^p(x_k) + g^a(u_k)$

Assignment: Compute the optimal control for $x_0 = 1, \alpha = 0.9$, i.e., the one that minimizes

$$J_\pi(x_0) = \lim_{N\to\infty} E_{w_0,w_1,\ldots} \left\{ \sum_{k=0}^{N-1} \alpha^k g(x_k, \mu_k(x_k), w_k) \right\}$$

by

a) using the value iteration algorithm

b) using the policy iteration algorithm