**Decision Models – Reinforcement Learning assignment**
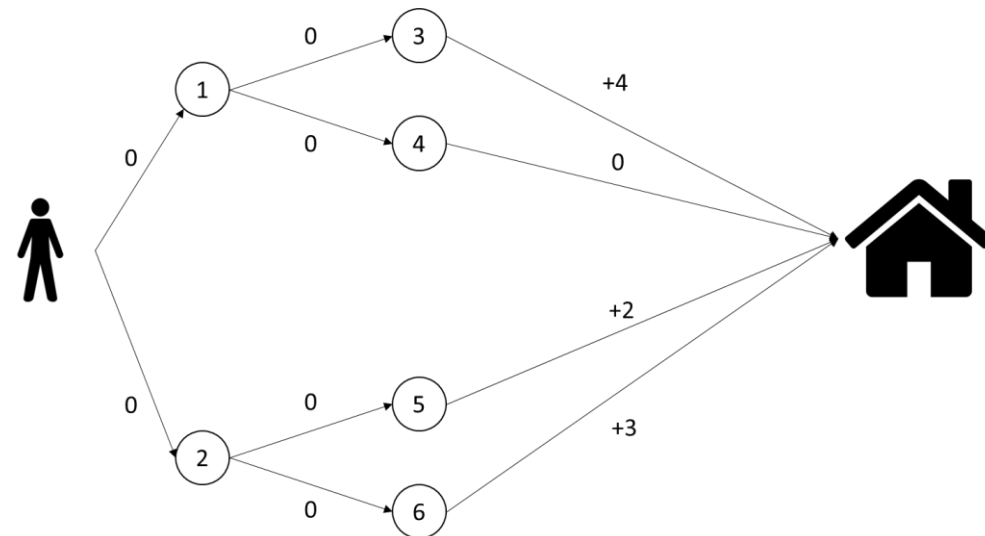
Consider the deterministic setting reported in the following figure. There are 8 different states:

- One initial state (where the man is)
- Six "intermediate" states (from 1 to 6)
- One final state (where the home is)

At all states, but the final one, a decision must be taken, according to two different options: going right or going left.



The following initial policy $\pi_0$ is given:

- $0 \rightarrow 2$
- $2 \rightarrow 5$
- $1 \rightarrow 4$
- (from 3, 4, 5, and 6 only one action is possible, leading to home, that is the final state)

You have to:

1. Compute the state function $V^{\pi_0}$ for each state
2. Improve the policy to obtain the new $\pi_1$
3. Compute the new state function $V^{\pi_1}$ for each state
4. Improve again the policy to obtain the new $\pi_2$

Did you converge to the optimal policy?