States:
Hot (reward 10)
Mild (reward 0)
Cold (reward -10)

Actions: East, West

Transitions:
From Hot go West (.9 > Hot, .1 > Mild)
From Hot go East (.5 > Hot, .5 > Mild)
From Mild go West (1.0 > Hot)
From Mild go East (1.0 > Cold)
From Cold go West (.5 > Mild, .5 > Cold)
From Cold go East (.1 > Mild, .9 > Cold)

VALUE ITERATION

\[ V_0(H) = 10 \]
\[ V_0(M) = 0 \]
\[ V_0(C) = -10 \]
\[ \gamma = 0.5 \]

\[ V_1(H) = 10 + 0.5 \max(W, E) = 10 + 0.5 \max(0.9 \times 10 + 0.1 \times 0, 0.5 \times 10 + 0.5 \times 0) = 14.5 \]

POLICY ITERATION

\[ \Pi_0(H) = E, \Pi_0(M) = E, \Pi_0(C) = W \]

Check these numbers (They are right)

\[ V(H) = 10 + 0.5(0.5 \times V(H) + 0.5 \times V(M)) \]
\[ V(M) = 0 + 0.5(1.0 \times V(C)) \]
\[ V(C) = -10 + 0.5(0.5 \times V(M) + 0.5 \times V(C)) \]

\[ \Pi_1(H) = W \]

\[ \Pi_1(H) = \arg \max_a \{W, E\} = \{(9.5 \times 10 + 0.5 \times 10 + 0.5 \times 8), 10 + 0.5(5 \times 10.67 + 0.5 \times 8)\} \Rightarrow \Pi_1(H) = W \]