Problem 7a

Some people got philosophical on this one, asking what I meant by "eventually." Formally, my intent was "with probability 1 as time goes to infinity," in which case the answer was yes, the robot will eventually make it to the goal state. Regardless of the policy followed, the randomness in its execution (0.05 probability of moving in one of the three directions other than intended) ensures that, with probability 1, it will eventually take enough moves in the right direction to get out of any dead ends.

In most cases, I gave partial credit to reasonable, but differing interpretations of "always" and "eventually".

Problem 7b

Begin with open loop policy of "always go right". Given the noise model, this means that the robot will go right 85% of the time, up 5% of the time, down 5% of the time, and left 5% of the time. If there's a wall in the direction it attempts to go, it will find itself back in the same state.

Initially, assume "cost to go" (which we'll write as $v()$) from each state (except G) is 1.

Conditions before first backup of S:

$$\begin{aligned}
C_S &= \text{immediate cost of step} = 1 \\
p_{S4} &= \text{prob of going right} = 0.85 \\
p_{S1} &= \text{prob of going up} = 0.05 \\
p_{SS} &= \text{prob of staying put (by trying to go left)} = 0.05 \\
v(S) &= 1, v(1) = 1, v(4) = 1, v(8) = 1
\end{aligned}$$

First backup of S:

$$\begin{aligned}
v(S) &= C_s + \sum_{S'} p_{SS'} \cdot v(S') \\
&= 1 + 0.85 \cdot 1 + 0.05 \cdot 1 + 0.05 \cdot 1 + 0.05 \cdot 1 = 2
\end{aligned}$$

Conditions before first backup of 4:

$$\begin{aligned}
C_4 &= \text{immediate cost of step} = 1 \\
p_{45} &= \text{prob of going right} = 0.85 \\
p_{44} &= \text{prob of staying put (trying up or down)} = 0.1 \\
p_{4S} &= \text{prob of going left} = 0.05 \\
v() &= 2, v(4) = 1, v(5) = 1
\end{aligned}$$

First backup of 4:

$$\begin{aligned} v(4) &= C_4 + \sum_{S'} p_{4S'} \cdot v(S') \\ &= 1 + 0.85 \cdot 1 + 0.1 \cdot 1 + 0.05 \cdot 2 = 2.05 \end{aligned}$$

Conditions before first backup of 5:

$$\begin{aligned} C_5 &= \text{immediate cost of step} = 1 \\ p_{56} &= \text{prob of going right} = 0.85 \\ p_{52} &= \text{prob of going up} = 0.05 \\ p_{59} &= \text{prob of going down} = 0.05 \\ p_{54} &= \text{prob of going left} = 0.05 \\ v(2) &= 1, v(4) = 2.05, v(5) = 1, v(9) = 1 \end{aligned}$$

First backup of 5:

$$\begin{aligned} v(5) &= C_5 + \sum_{S'} p_{5S'} \cdot v(S') \\ &= 1 + 0.85 \cdot 1 + 0.05 \cdot 1 + 0.05 \cdot 1 + 0.05 \cdot 2.05 = 2.0525 \end{aligned}$$

Conditions before first backup of 6:

$$\begin{aligned} C_6 &= \text{immediate cost of step} = 1 \\ p_{67} &= \text{prob of going right} = 0.85 \\ p_{66} &= \text{prob of staying put (trying up or down)} = 0.1 \\ p_{65} &= \text{prob of going left} = 0.05 \\ v(5) &= 2.0525, v(6) = 1, v(7) = 1 \end{aligned}$$

First backup of 6:

$$\begin{aligned} v(6) &= C_6 + \sum_{S'} p_{6S'} \cdot v(S') \\ &= 1 + 0.85 \cdot 1 + 0.1 \cdot 1 + 0.05 \cdot 2.0525 = 2.05265 \end{aligned}$$

Conditions before first backup of 7:

$$\begin{aligned} C_7 &= \text{immediate cost of step} = 1 \\ p_{7G} &= \text{prob of going right} = 0.85 \\ p_{73} &= \text{prob of going up} = 0.05 \\ p_{710} &= \text{prob of going down} = 0.05 \\ p_{76} &= \text{prob of going left} = 0.05 \\ v(3) &= 1, v(6) = 2.05265, v(10) = 1, v(G) = 0 \end{aligned}$$

First backup of 7:

$$\begin{aligned} v(7) &= C_7 + \sum_{S'} p_{7S'} \cdot v(S') \\ &= 1 + 0.85 \cdot 0 + 0.05 \cdot 1 + 0.05 \cdot 1 + 0.05 \cdot 2.05265 = 1.2026 \end{aligned}$$

Conditions before second backup of 6:

$$\begin{aligned} C_6 &= \text{immediate cost of step} = 1 \\ p_{67} &= \text{prob of going right} = 0.85 \\ p_{66} &= \text{prob of staying put (trying up or down)} = 0.1 \\ p_{65} &= \text{prob of going left} = 0.05 \\ v(5) &= 2.0525, v(6) = 2.05265, v(7) = 1.2026 \end{aligned}$$

Second backup of 6:

$$\begin{aligned} v(6) &= C_6 + \sum_{S'} p_{6S'} \cdot v(S') \\ &= 1 + 0.85 \cdot 1.2026 + 0.1 \cdot 1 + 0.05 \cdot 2.05265 = 2.329 \end{aligned}$$

Conditions before second backup of 5:

$$\begin{aligned} C_5 &= \text{immediate cost of step} = 1 \\ p_{56} &= \text{prob of going right} = 0.85 \\ p_{52} &= \text{prob of going up} = 0.05 \\ p_{59} &= \text{prob of going down} = 0.05 \\ p_{54} &= \text{prob of going left} = 0.05 \\ v(2) &= 1, v(4) = 2.05, v(6) = 2.329, v(9) = 1 \end{aligned}$$

Second backup of 5:

$$\begin{aligned} v(5) &= C_5 + \sum_{S'} p_{5S'} \cdot v(S') \\ &= 1 + 0.85 \cdot 2.329 + 0.05 \cdot 1 + 0.05 \cdot 1 + 0.05 \cdot 2.05 = 3.18215 \end{aligned}$$

Conditions before second backup of 4:

$$\begin{aligned} C_4 &= \text{immediate cost of step} = 1 \\ p_{45} &= \text{prob of going right} = 0.85 \\ p_{44} &= \text{prob of staying put (trying up or down)} = 0.1 \\ p_{4S} &= \text{prob of going left} = 0.05 \\ v(S) &= 2, v(4) = 2.05, v(5) = 3.18215 \end{aligned}$$

Second backup of 4:

$$\begin{aligned} v(4) &= C_4 + \sum_{S'} p_{4S'} \cdot v(S') \\ &= 1 + 0.85 \cdot 3.18215 + 0.1 \cdot 1 + 0.05 \cdot 2 = 4.00 \end{aligned}$$

Conditions before second backup of S:

$$\begin{aligned} C_S &= \text{immediate cost of step} = 1 \\ p_{S4} &= \text{prob of going right} = 0.85 \\ p_{S1} &= \text{prob of going up} = 0.05 \\ p_{SS} &= \text{prob of staying put (by trying to go left)} = 0.05 \\ v(S) &= 2, v(1) = 1, v(4) = 4.00, v(8) = 1 \end{aligned}$$

Second backup of S:

$$\begin{aligned} v(S) &= C_s + \sum_{S'} p_{SS'} \cdot v(S') \\ &= 1 + 0.85 \cdot 4.00 + 0.05 \cdot 1 + 0.05 \cdot 1 + 0.05 \cdot 2 = 4.6 \end{aligned}$$

Conditions before second backup of 7:

$$\begin{aligned} C_7 &= \text{immediate cost of step} = 1 \\ p_{7G} &= \text{prob of going right} = 0.85 \\ p_{73} &= \text{prob of going up} = 0.05 \\ p_{710} &= \text{prob of going down} = 0.05 \\ p_{76} &= \text{prob of going left} = 0.05 \\ v(3) &= 1, v(6) = 2.329, v(10) = 1, v(G) = 0 \end{aligned}$$

Second backup of 7:

$$\begin{aligned} v(7) &= C_7 + \sum_{S'} p_{7S'} \cdot v(S') \\ &= 1 + 0.85 \cdot 0 + 0.05 \cdot 1 + 0.05 \cdot 1 + 0.05 \cdot 2.329 = 1.216 \end{aligned}$$

Conditions before first backup of 3:

$$\begin{aligned} C_3 &= \text{immediate cost of step} = 1 \\ p_{33} &= \text{prob of staying put (by trying left, right or up)} = 0.95 \\ p_{37} &= \text{prob of going down} = 0.05 \\ v(3) &= 1, v(7) = 1.216 \end{aligned}$$

First backup of 3:

$$\begin{aligned} v(3) &= C_3 + \sum_{S'} p_{3S'} \cdot v(S') \\ &= 1 + 0.95 \cdot 1 + 0.05 \cdot 1.216 = 1.9108 \end{aligned}$$

Conditions before second backup of 3:

$$\begin{aligned}
C_3 &= \text{immediate cost of step} = 1 \\
p_{33} &= \text{prob of staying put (by trying left, right or up)} = 0.95 \\
p_{37} &= \text{prob of going down} = 0.05 \\
v(3) &= 1.9108, v(7) = 1.216
\end{aligned}$$

Second backup of 3:

$$\begin{aligned}
v(3) &= C_3 + \sum_{S'} p_{3S'} \cdot v(S') \\
&= 1 + 0.95 \cdot 1.9108 + 0.05 \cdot 1.216 = 2.68
\end{aligned}$$

Problem 7c

Initially, all neighboring states look equally good to the robot, so the max of the expected next value doesn't depend on which next (valid) state you try to move to. The initial backups S-4-5-6-7-6 are the same as before.

However, when we next try to backup state 5, unexplored states 2 and 9 still have (incorrect) value estimates of 1, so our policy of going to the lowest neighbor leads us astray. Let's arbitrarily say we decide to go up (rather than down)

Conditions before second backup of 5:

$$\begin{aligned}
C_5 &= \text{immediate cost of step} = 1 \\
p_{56} &= \text{prob of going right} = 0.05 \\
p_{52} &= \text{prob of going up} = 0.85 \\
p_{59} &= \text{prob of going down} = 0.05 \\
p_{54} &= \text{prob of going left} = 0.05 \\
v(2) &= 1, v(4) = 2.05, v(6) = 2.329, v(9) = 1
\end{aligned}$$

Second backup of 5:

$$\begin{aligned}
v(5) &= C_5 + \sum_{S'} p_{5S'} \cdot v(S') \\
&= 1 + 0.05 \cdot 2.329 + 0.85 \cdot 1 + 0.05 \cdot 1 + 0.05 \cdot 2.05 = 2.122
\end{aligned}$$

We won't correct the erroneous estimates of states 2 and 9 until we actually back *them* up.

Now, due to our underestimates of state S, our policy again points us in the wrong direction. Conditions before second backup of 4:

$$C_4 = \text{immediate cost of step} = 1$$

$$
\begin{aligned}
p_{45} &= \text{prob of going right} = 0.05 \\
p_{44} &= \text{prob of staying put (trying up or down)} = 0.1 \\
p_{4S} &= \text{prob of going left} = 0.85 \\
v(S) &= 2, v(4) = 2.05, v(5) = 2.122
\end{aligned}
$$

Second backup of 4:

$$
\begin{aligned}
v(4) &= C_4 + \sum_{S'} p_{4S'} \cdot v(S') \\
&= 1 + 0.85 \cdot 2 + 0.1 \cdot 2.05 + 0.05 \cdot 2.122 = 3.011
\end{aligned}
$$

Conditions before second backup of S:

$$
\begin{aligned}
C_S &= \text{immediate cost of step} = 1 \\
p_{S4} &= \text{prob of going right} = 0.05 \\
p_{S1} &= \text{prob of going up} = 0.85 \\
p_{SS} &= \text{prob of staying put (by trying to go left)} = 0.05 \\
v(S) &= 2, v(1) = 1, v(4) = 3.011, v(8) = 1
\end{aligned}
$$

Second backup of S:

$$
\begin{aligned}
v(S) &= C_s + \sum_{S'} p_{SS'} \cdot v(S') \\
&= 1 + 0.85 \cdot 1 + 0.05 \cdot 2 + 0.05 \cdot 1 + 0.05 \cdot 3.011 = 2.15
\end{aligned}
$$

Conditions before second backup of 7:

$$
\begin{aligned}
C_7 &= \text{immediate cost of step} = 1 \\
p_{7G} &= \text{prob of going right} = 0.85 \\
p_{73} &= \text{prob of going up} = 0.05 \\
p_{710} &= \text{prob of going down} = 0.05 \\
p_{76} &= \text{prob of going left} = 0.05 \\
v(3) &= 1, v(6) = 2.329, v(10) = 1, v(G) = 0
\end{aligned}
$$

Second backup of 7:

$$
\begin{aligned}
v(7) &= C_7 + \sum_{S'} p_{7S'} \cdot v(S') \\
&= 1 + 0.85 \cdot 0 + 0.05 \cdot 1 + 0.05 \cdot 1 + 0.05 \cdot 2.329 = 1.216
\end{aligned}
$$

Conditions before first backup of 3:

$$
\begin{aligned}
C_3 &= \text{immediate cost of step} = 1 \\
p_{33} &= \text{prob of staying put (by trying left, right or up)} = 0.15 \\
p_{37} &= \text{prob of going down} = 0.85 \\
v(3) &= 1, v(7) = 1.216
\end{aligned}
$$

First backup of 3:

$$\begin{aligned} v(3) &= C_3 + \sum_{S'} p_{3S'} \cdot v(S') \\ &= 1 + 0.15 \cdot 1 + 0.85 \cdot 1.216 = 2.183 \end{aligned}$$

Conditions before second backup of 3:

$$\begin{aligned} C_3 &= \text{immediate cost of step} = 1 \\ p_{33} &= \text{prob of staying put (by trying left, right or up)} = 0.15 \\ p_{37} &= \text{prob of going down} = 0.85 \\ v(3) &= 2.183, v(7) = 1.216 \end{aligned}$$

Second backup of 3:

$$\begin{aligned} v(3) &= C_3 + \sum_{S'} p_{3S'} \cdot v(S') \\ &= 1 + 0.15 \cdot 2.183 + 0.85 \cdot 1.216 = 2.36 \end{aligned}$$

The take home message on this is that you need to back up everything (like states 2 and 9), or your values may never point you in the direction you need, and, for a fixed policy, the value function converges more quickly than if one's policy is dictated by the current value estimates (though both are guaranteed to converge if all states are backed up infinitely often.