Part a: Absorbing states are stopping states by definition (since the Markov chain does not move from there). So states 0 and 10 are stopping states. Also, state 4 gets the maximal payoff and so there is no gain from continuing and it is a stopping state.

Consider now the strategy that says we stop at states 2 and 4 but continue at 3. The expected gain under this strategy, if we start at 3, is given by 

\[
\frac{f(2) + f(4)}{2} = \frac{4 + 10}{2} = 7.
\]

Since the optimal strategy would give the best expected gain \(v(3)\) we see that \(v(3) \geq 7\). In particular, \(v(3) > f(3)\) and so we need to keep going at 3.

To continue, recall that \(v(x) \geq f(x)\) and \(v(x) \geq Pv(x)\) for all \(x\). From this, we get that

\[
v(2) \geq Pv(2) = (v(1) + v(3))/2 \geq (f(1) + f(3))/2 = (2 + 7)/2.
\]

So \(v(2) \geq 9/2\) and in particular it is strictly bigger than \(f(2)\) and we should go at 2. Similarly,

\[
v(1) \geq Pv(1) = (v(0) + v(2))/2 = (f(0) + v(2))/2 \geq (f(0) + 9/2)/2 = 9/4 > f(1)
\]

and so we go at 1.

Also,

\[
v(8) \geq Pv(8) = (v(7) + v(9))/2 \geq (f(7) + f(9))/2 = 7/2 > f(8)
\]

and we should go at 8. Similarly, \(v(7) \geq (f(6) + f(8))/2 = 9/2 > f(7)\) and we should go at 7. We could have also argued that

\[
v(7) \geq Pv(7) = (v(6) + v(8))/2 \geq (f(6) + v(8))/2 \geq (f(6) + 7/2))/2 = 19/4 > f(7).
\]

State 5 is argued similarly: \(v(5) \geq (f(4) + f(6))/2 = 8 > f(5)\). So 5 is a go state. Now for state 6 we see that

\[
v(6) \geq Pv(6) = (v(5) + v(7))/2 \geq (8 + 9/2)/2 = 25/4 > f(6).
\]

So 6 is a go state.

Alternatively: Running the algorithm we get the hint that in addition to the trivial stopping states, 9 is also a stopping state. So then consider this strategy (which we cannot yet assume to be the optimal one, but we are free to consider it and see what happens!). If you solve the linear equations for this strategy (these equations are the exact same equations as for the \(v\) below), then you get that the values for the states \(\{1, 2, 3, 5, 6, 7, 8\}\) are all strictly larger than the corresponding payoffs. Since \(v\) can only give larger (or equal) values, we know that the \(v\) at these sites is strictly larger than the payoffs and so these are indeed go states.

Either way, it remains to sort out state 9. The strategy that says stop at 8 and 10 and go at 8 gives an expected gain of \((f(8) + f(10))/2 = 3/2 < f(9)\). This only says this particular strategy is worse than stopping at 9. But it does not tell us whether or not there is a better strategy that actually gains us more than \(f(9)\). However, since we have actually determined all the other states we only have two possible strategies to consider: the stopping states are \(\{0, 4, 10\}\) or they are \(\{0, 4, 9, 10\}\). Let us consider the former, i.e. suppose 9 is a go state. Then starting at 9 we keep going until we reach either 4 or 10. By gambler's ruin, the
probability we reach 4 is \((10 - 9)/(10 - 4) = 1/6\) and the probability we reach 10 first is \((9 - 4)/(10 - 4) = 5/6\). So the expected gain would then be

\[
\frac{1}{6} \cdot f(4) + \frac{5}{6} \cdot f(10) = \frac{10}{6} < f(9).
\]

So this is not a good strategy and that leaves us with the strategy corresponding to \(\{0, 4, 9, 10\}\) being the stopping states.

If you used the alternative method, then now you have confirmed that that indeed is the optimal strategy and you already computed the expected gain for it, which is exactly what \(v\) is.

If not, then now that we know the stopping states we can compute the expected gain \(v\) by solving the system of equations:

\[
\begin{align*}
v(1) &= (f(0) + v(2))/2 = v(2)/2 \\
v(2) &= (v(1) + v(3))/2 \\
v(3) &= (v(2) + f(4))/2 = (v(2) + 10)/2,
\end{align*}
\]

and

\[
\begin{align*}
v(5) &= (f(4) + v(6))/2 = (10 + v(6))/2 \\
v(6) &= (v(5) + v(7))/2 \\
v(7) &= (v(6) + f(8))/2 \\
v(8) &= (v(7) + f(9))/2 = (v(7) + 3)/2.
\end{align*}
\]

The first system is solved quickly because the first and third equations can be substituted in the second, giving \(v(2)\), and thus \(v(1)\) and \(v(3)\). For the second system substitute the first and fourth equations to eliminate \(v(5)\) and \(v(8)\) from the second and third equations. Then this is just a system of two equations in two unknowns.

For this particular problem, if we do not want to solve the above equations, then we can use gambler’s ruin which says that starting at \(x \in [a, b]\) the probability we reach \(a\) before \(b\) is \((b - x)/(b - a)\). So for a go state \(x\) that is between two stopping states \(a\) and \(b\) we have

\[
v(x) = \frac{b - x}{b - a} \cdot f(a) + \frac{x - a}{b - a} \cdot f(b).
\]

For example,

\[
v(7) = \frac{9 - 7}{9 - 4} \cdot f(4) + \frac{7 - 4}{9 - 4} \cdot f(9) = \frac{2}{5} \cdot 10 + \frac{3}{5} \cdot 3 = \frac{29}{5}.
\]

**Part b:** First, we know that any states that were stopping states without a discount continue to be stopping states with the discount. So \(\{0, 4, 9, 10\}\) are all stopping states. Also, since \(f(5) = 0\) a discount does not change the fact that we get nothing if we stop at 5 and we have nothing to loose if we do not. So 5 is again a go state.

Before continuing notice the following: if we fix a set of stopping states \(S_0\) and consider the strategy that corresponds to this set and if we let \(u(x)\) denote the expected gain under
this strategy, starting at $x$, and $T$ the stopping time for this strategy, then decomposing into the possible values of $X_T$ (the stopping point) we get

$$u(x) = \sum_{y \in S_0} f(y) E[\alpha^T 1\{X_T = y\} \mid X_0 = x].$$

Suppose now that to get from $x$ to $y$ we need at least $m(y)$ steps. Then we have $T \geq m(y)$ in the above expectation and so $\alpha^T \leq \alpha^{m(y)}$. We thus have

$$u(x) \leq \sum_{y \in S_0} \alpha^{m(y)} f(y) P(X_T = y \mid X_0 = x).$$

So for example, if we consider the strategy where we stop at $\{0, 4\}$ and we go at $\{1, 2, 3\}$, then starting at 1 we need at least 3 steps to get to 4 and at least one step to get to 0. So the expected gain for this strategy satisfies

$$u(1) \leq 0.9 f(0) P(X_T = 0 \mid X_0 = 1) + 0.9^3 f(4) P(X_T = 4 \mid X_0 = 1).$$

By gambler’s ruin we know the first probability is $3/4$ and the second is $1/4$. So

$$u(1) \leq 0.9^3 \times 10 \times 1/4 < 2 = f(1).$$

So this is not a good strategy because just stopping at 1 does better. This does not yet tell us that 1 is a stopping state. Similarly, if we say we go at 1 and 2 but stop at 3 we have

$$u(1) \leq 0.9^2 \times 3 \times 1/3 < 2 = f(1).$$

So this is not a good strategy either (actually even without a discount!). And if we consider going at 1 and stopping at 2 (and 0), then

$$u(1) \leq 0.9 \times 2 \times 1/2 < 2 = f(1).$$

So this is again not a good strategy (even without the discount). So the conclusion is that 1 must be a stopping state.

To decide on 2 and 3 we solve the linear equations for the strategy that says go at both 2 and 3 and stop at 1 and 4:

$$u(2) = 0.9(f(1) + u(3))/2 \quad \text{and} \quad u(3) = 0.9(u(2) + f(4))/2.$$ 

Solving we get $u(2) = 1179/319$ and $u(3) = 1962/318$. Since $u(3) > f(3)$ and we know that $v(3) \geq u(3)$ we see that $v(3) > f(3)$ and 3 must be a go state. But also the fact that $u(2) < f(2)$ says that the strategy that says go at 2 and 3 and stop at 1 and 4 is not good. Now the only remaining strategy is the one that says go at 3 but stop at 1 and 2 and 4.

The situation with states 6, 7, and 8 is worked out similarly and I leave it to the students to work out, without having to solve more than two equations in two unknowns. (So first, e.g. determine that 6 must be a stop state. Then solve for 7 and 8 being go states and deduce that they are indeed both go states.)

Once we know all the stopping states we can solve the system of linear equations that gives the optimal expected gain $v$.

**Alternative method:** Here too you could use the algorithm and it hints to you that states $\{0, 1, 2, 4, 6, 9, 10\}$ are stopping states. So you can solve the linear equations for the expected gain under this strategy and determine that the solutions at the other states are strictly larger than the payoffs. This confirms that the other states are indeed all go states.
But then you need to still argue why $\{1, 2, 6\}$ are stopping states. For this, you can use the arguments mentioned above to determine that 1 is a stopping state and then that 2 cannot be a go state and that 6 must be a stopping state. Ans once we have confirmed that what the algorithm suggested is indeed what the optimal strategy says we should do, we get as a result that the expected gains we already computed are actually the optimal ones.

**Part c:** Again, states $\{0, 4, 9, 10\}$ are trivial stopping states. Now suppose that $T$ is the stopping time corresponding to the optimal (yet unknown to us) strategy. Suppose $x$ is not a stopping state under this strategy. Then if we denote by $v_c(x)$ the optimal expected gain, starting at $x$, and playing with a cost $c$, we have

$$v_c(x) = E[f(X_T) - cT | X_0 = x] \leq E[f(X_T) | X_0 = x] - c \leq v(x) - c,$$

where $c$ is the constant cost we pay each time we continue and $v(x)$ is the optimal expected gain without a cost (so what we found in part a). The above inequality is because we said that $x$ is not a stopping state and so $T \geq 1$.

So any time $v(x) - c < f(x)$ we would get $v_c(x) < f(x)$, which is a contradiction and so our assumption that $x$ is a stopping state is not true for such an $x$. This immediately tells us that states $\{1, 2, 6, 8\}$ are now stopping states. (For example, $v(1) - c = 2.5 - 2 < 2 = f(1).$)

So we now only have states $\{3, 5, 7\}$ to consider. These states are all surrounded by stopping states and so the problem is now very easy. For example, assuming 3 is a go state we get

$$u(3) = (f(2) + f(4))/2 - 2 = 7 - 2 > 3 = f(3).$$

So 3 is indeed a go state. And so on.

Once we know the stopping states we can solve the system of linear equations that gives the optimal expected gain (what we called here $v_c$).

**Alternative method:** Once again, we can use the algorithm and get a hint of what the stopping states are. Then we solve the linear equations that give us the expected gain for the strategy with these stopping states. This will confirm the go states because they will have expected gains that are strictly larger than the payoffs. Then we need to still confirm that the stopping states suggested by the algorithm are truly optimal stopping states and for this we can use the above reasoning. Once we have confirmed that what the algorithm suggested is indeed what the optimal strategy says we should do, we get as a result that the expected gains we already computed are actually the optimal ones.

Note how the alternative method is the most efficient way to do things (and not only with a random walk like in this problem). Namely, you use the algorithm to get a guess of the stopping states. Next, you say let us try this strategy. You set up the linear equations for the expected gain for this strategy and you discover that the go states in this strategy all have expected gain strictly more than the payoffs. So they are go states for the optimal strategy as well. Next, you eliminate the trivial stop states (e.g. absorbing and maximal payoff states) and hopefully you are now left with a small number of suggested stopping states that you need to confirm. For these, you have to use the tricks that are mentioned in this write-up or look at all the cases (or do some tricks for some of the states and then look at all the remaining cases). At the end, you will have confirmed that your initial guess, coming from the algorithm, is indeed correct and so you will have as a bonus that the expected gain you calculated at the very beginning is actually the optimal expected gain.