Question 1

We have the following SDE for W_t^*:

$$dW^* = ((\pi(\mu - r) + r)\, W^* - c)\; dt + \pi\sigma W^*\, dz$$

where $W^* = W_t^*, c = c_t, \pi = \pi_t(t, W_t^*)$ and z is the Brownian motion. We aim to maximize:

$$\mathbb{E}\left(\int_t^T e^{-\rho(T-s)} \log c_s\, ds + e^{-\rho(T-t)} \log W_t \,\middle|\, W_t \right)$$

The corresponding HJB equation is:

$$\rho V^* = \max_{\pi,c} \left[\frac{\partial V^*}{\partial t} + \pi(\mu - r)W\frac{\partial V^*}{\partial W} + rW\frac{\partial V^*}{\partial W} - c \cdot \frac{\partial V^*}{\partial W} + \frac{1}{2}\pi^2\sigma^2 W^2 \frac{\partial^2 V^*}{\partial W^2} + \log(c) \right]$$

with the boundary condition, $V^*(T, W_T) = \log W_T$. Let us denote the term within the square braces above as Φ. By taking the derivative of the argument in the `max` function above wrt π and c, we get:

$$\begin{aligned}
\frac{\partial \Phi}{\partial \pi} &= 0 \Rightarrow (\mu - r)W\frac{\partial V^*}{\partial W} + {\pi^*}^2\sigma^2 W^2 \frac{\partial^2 V^*}{\partial W^2} = 0 \Rightarrow \pi^* = -\frac{(\mu - r)\cdot \frac{\partial V^*}{\partial W}}{\sigma W \cdot \frac{\partial^2 V^*}{\partial W^2}} \\
\frac{\partial \Phi}{\partial c} &= 0 \Rightarrow -\frac{\partial V^*}{\partial W} + \frac{1}{c^{*2}} = 0 \Rightarrow c^* = \frac{1}{\frac{\partial V^*}{\partial W}}
\end{aligned}$$

Therefore, we get:

$$\frac{\partial V^*}{\partial t} - \frac{(\mu - r)\cdot \frac{\partial V^*}{\partial W}}{\sigma W \cdot \frac{\partial^2 V^*}{\partial W^2}} \cdot (\mu - r)W\frac{\partial V^*}{\partial W} + rW\frac{\partial V^*}{\partial W} - \frac{1}{\frac{\partial V^*}{\partial W}} \cdot \frac{\partial V^*}{\partial W} + \frac{1}{2}\left(-\frac{(\mu - r)\cdot \frac{\partial V^*}{\partial W}}{\sigma W \cdot \frac{\partial^2 V^*}{\partial W^2}} \right)^2 \sigma^2 W^2 \frac{\partial^2 V^*}{\partial W^2} - \log\left(\frac{\partial V^*}{\partial W}\right) = \rho V^*$$

$$\therefore \frac{\partial V^*}{\partial t} - \frac{(\mu - r)^2 \cdot \left(\frac{\partial V^*}{\partial W}\right)^2}{2\sigma^2 \frac{\partial^2 V^*}{\partial W^2}} + rW\frac{\partial V^*}{\partial W} - \log\left(\frac{\partial V^*}{\partial W}\right) - 1 = \rho V^*$$

Let us guesstimate the solution for V^* to be $V^* = A_t + \log W_t$. Therefore, we get the following ODE:

$$\begin{aligned}
&\frac{dA_t}{dt} = \rho A_t + k \; \left(\text{where } k = (\rho - 1)\log W_t - \frac{(\mu - r)^2}{2\sigma^2} - r - 1 \right) \\
&\therefore A_t = ce^{\rho t} - \frac{k}{\rho} \;\; (\text{with } A_T = 0) \\
&\therefore A_t = \frac{k}{\rho}\left(e^{-\rho(T-t)} - 1 \right)
\end{aligned}$$

Therefore, we get:

$$V^* = \frac{k}{\rho}\left(e^{-\rho(T-t)} - 1 \right) + \log W_t$$

is the solution to the HJB equation.

Question 3

Let the state space be $\mathcal{S} = \{(S_t, d_t) \mid d_t \in \mathbb{Z}_{\geq 0}\}$ where S_t is the skill learned at day t, while d_t is the number of days unemployed. The action space is $\mathcal{A} = \{\alpha_t \mid \alpha_t \in [0,1]\}$ where α_t denotes the fraction of the day spent on working on day t.

The transition function is:

$$\mathcal{P}\left((S_t, d_t), \alpha_t, (S_{t+1}, d_{t+1})\right) = \begin{cases} 1-p, & \text{if } S_{t+1} = S_t \cdot (1 + g(S_t)) \text{ and } d_{t+1} = 0 \\ p, & \text{if } S_{t+1} = S_t \cdot (1 + g(S_t)) \text{ and } d_{t+1} = 1 \\ h(S_t), & \text{if } S_{t+1} = S_t \cdot e^{-\lambda \cdot d_t} \text{ and } d_{t+1} = 0 \\ 1 - h(S_t), & \text{if } S_{t+1} = S_t \cdot e^{-\lambda \cdot d_t} \text{ and } d_{t+1} = d_t + 1 \end{cases}$$

The first case corresponds to when I've been working today and I continue working tomorrow.
The second case corresponds to when I've been working today but I lose the job the next day. Hence, we set $d_{t+1} = 1$ to account for the fact that we've been unemployed for 1 day only.
The third and fourth cases are similar to the first two, just that the starting state is that I've been unemployed, so I must change my next state and adjust d_t as well.

With the above formulation of the problem, we also have the following reward function:

$$\mathcal{R}\left((S_t, d_t), \alpha_t\right) = \begin{cases} U\left(\alpha_t \cdot f(S_t) \cdot (60 \times 24)\right), & \text{if } d_t = 0 \text{ (aka, still employed)} \\ 0, & \text{if } d_t > 0 \text{ (aka, unemployed)} \end{cases}$$

where $U(\cdot)$ is my utility function and the 60×24 comes from the fact that α_t is the fraction of the day, while $f(\cdot)$ is the rate per minute. For the purpose of finding the optimal policy α_t^*, we can drop this constant and use the MDP as is.

The optimal policy in the above case would be to learn as much skills as possible in the beginning and then switch to the case where $\alpha = 1$ (you work more than you learn). In this way, my lifetime utilities of earnings will be optimal.

Let us now discuss other variants on this MDP:

- Finite versus infinite horizon:
  In the finite horizon case, we can use $\gamma = 1$ since the rewards wonn't blow up to infinity.
  To ensure that this doesn't occur in the infinite-horizon case, we must use $\gamma < 1$.

- Multiple skills or jobs:
  In the case of multiple skills or job options, we can maintain a list of the different skills and different jobs. In either case, we will need to rely on linear algebra to represent the different probabilities and calculate the transition states accordingly.

- Consumption:
  We'd have to add the daily consumption at the end of the day c_t to the action space.
  We will also have to account for this consumption when calculating the reward for transiting to the next state (the 60×24 cannot be ignored in this case).