2) a) If we consider that $d = \begin{bmatrix} d(s_1) \\ d(s_2) \end{bmatrix} = \begin{bmatrix} a_{1,1} \\ a_2 \end{bmatrix}$, where it is given that $r_d = \begin{bmatrix} 5 \\ -1 \end{bmatrix}$ indicating reward of $a_{1,1}$ for 5 and reward of a_2 for -1

$P_d = \begin{bmatrix} 0.5 & 0.5 \\ 0 & 1 \end{bmatrix}$ in which $P(s_1 \to s_1) = 0.5$ (wrt $a_{1,1}$) $P(s_1 \to s_2) = 0.5$
$P(s_2 \to s_1) = 0$ (wrt a_2) $P(s_2 \to s_2) = 1$

Taking into account that value function during the stationary policy will be provided as follows: $d^\infty = (d, d, \dots)$ and $V_\lambda^{d^\infty} = \lambda P_d V_\lambda^{d^\infty} + r_d$

This will follow that $(I - \lambda P_d) V_\lambda^{d^\infty} = r_d$ will be satisfied

$\boxed{V_\lambda^{d^\infty} = (I - \lambda P_d)^{-1} r_d}$ Plugging back the results, $I - \lambda P_d =$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} \frac{\lambda}{2} & \frac{\lambda}{2} \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 1 - \frac{\lambda}{2} & -\frac{\lambda}{2} \\ 0 & 1-\lambda \end{bmatrix}$$

$$\begin{bmatrix} 1 - \frac{\lambda}{2} & -\frac{\lambda}{2} \\ 0 & 1-\lambda \end{bmatrix}^{-1} = \frac{1}{\left(1-\frac{\lambda}{2}\right)(1-\lambda)} \begin{bmatrix} 1-\lambda & \frac{\lambda}{2} \\ 0 & 1 - \frac{\lambda}{2} \end{bmatrix} = \frac{1}{\left(\frac{2-\lambda}{2}\right)(1-\lambda)} \begin{bmatrix} 1-\lambda & \frac{\lambda}{2} \\ 0 & 1-\frac{\lambda}{2} \end{bmatrix}$$

$$V_\lambda^{d^\infty} = \frac{1}{\left(1-\frac{\lambda}{2}\right)(1-\lambda)} \begin{bmatrix} 1-\lambda & \frac{\lambda}{2} \\ 0 & 1-\frac{\lambda}{2} \end{bmatrix} r_d = \frac{1}{\left(1-\frac{\lambda}{2}\right)(1-\lambda)} \begin{bmatrix} 1-\lambda & \frac{\lambda}{2} \\ 0 & 1-\frac{\lambda}{2} \end{bmatrix} \begin{bmatrix} 5 \\ -1 \end{bmatrix} =$$

$$= \frac{1}{\left(1-\frac{\lambda}{2}\right)(1-\lambda)} \begin{bmatrix} 5 - 5\lambda - \frac{\lambda}{2} \\ \frac{\lambda}{2} - 1 \end{bmatrix} = \frac{1}{\left(1-\frac{\lambda}{2}\right)(1-\lambda)} \begin{bmatrix} 5 - \frac{11\lambda}{2} \\ \frac{\lambda}{2} - 1 \end{bmatrix}$$

, In the end, we can conclude that

$$V_\lambda^{d^\infty} = \frac{1}{\frac{2-\lambda}{2}(1-\lambda)} \begin{bmatrix} \frac{10-11\lambda}{2} \\ \frac{\lambda-2}{2} \end{bmatrix} = \frac{1}{(2-\lambda)(1-\lambda)} \begin{bmatrix} 10 - 11\lambda \\ \lambda - 2 \end{bmatrix} = \begin{bmatrix} (10-11\lambda)/(2-\lambda)(1-\lambda) \\ 1/\lambda - 1 \end{bmatrix}$$

So, we found that $\boxed{V_\lambda^{d^\infty} = \begin{bmatrix} \dfrac{10-11\lambda}{(\lambda-2)(\lambda-1)} \\ \dfrac{1}{\lambda-1} \end{bmatrix}}$ $\underline{V}$

b) Assume $d = \begin{bmatrix} a_{1,1} \\ a_2 \end{bmatrix}$ where 2^{nd} decision rule $\rightsquigarrow \begin{bmatrix} a_{1,2} \\ a_2 \end{bmatrix}$

It's clear that there exist two possible alternatives:

Decision rule d^* – conserving whenever $L_{d^*} V_\lambda^* = L V_\lambda^* = V_\lambda^*$

Since $d^* = \arg\max_{d \in D^{MD}} \{ r_d + \lambda P_d V_\lambda^* \}$

$\underline{L_d V = r_d + \lambda P_d V}$ $\rightsquigarrow$ indicating L_d as linear transformation

At the end, this implies $L_{d^*} V = V$ in the conserving decision problem

c) The major goal in this part will be to generalize the results occurring in $\underline{\text{part (a)}}$ Considering we found $V_\lambda^{d^\infty} = \begin{bmatrix} \dfrac{10-11\lambda}{(\lambda-2)(\lambda-1)} \\ \dfrac{1}{\lambda-1} \end{bmatrix}$

it's crucial to compute $V_\lambda \rightsquigarrow (I - \lambda P_d)^{-1} r_d = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \lambda P_d \right)^{-1} r_d =$

$$= \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 10 \\ -1 \end{bmatrix} = \left(\begin{bmatrix} 1 & -\lambda \\ 0 & 1-\lambda \end{bmatrix} \right)^{-1} \begin{bmatrix} 10 \\ -1 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & -\lambda \\ 0 & 1-\lambda \end{bmatrix}^{-1} \begin{bmatrix} 10 \\ -1 \end{bmatrix} = \frac{1}{1-\lambda} \begin{bmatrix} 1-\lambda & \lambda \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 10 \\ -1 \end{bmatrix} =$$

$$= \begin{bmatrix} \dfrac{10-11\lambda}{1-\lambda} \\ \dfrac{1}{\lambda-1} \end{bmatrix}$$ where $(1-\lambda)10 - \lambda = 10 - 11\lambda$ and $\dfrac{-1}{1-\lambda} = \dfrac{1}{\lambda-1}$

$$\boxed{V_\lambda^{d^\infty} = \begin{bmatrix} \dfrac{10-11\lambda}{1-\lambda} \\ 1/\lambda-1 \end{bmatrix}}$$

$$V_\lambda^{f^\infty} = \begin{bmatrix} \frac{10-11\lambda}{1-\lambda} \\ \frac{1}{\lambda-1} \end{bmatrix} \text{ and } V_\lambda^{d^\infty} = \begin{bmatrix} \frac{10-11\lambda}{(\lambda-1)(\lambda-2)} \\ \frac{1}{\lambda-1} \end{bmatrix}$$ found in previous problem

We compare first components of these 2 column vectors

$\lambda = \frac{10}{11} \Rightarrow$ both are equal

$\lambda < \frac{10}{11}$, therefore we choose f^∞

$\lambda > \frac{10}{11}$, we get that d^∞ should be chosen