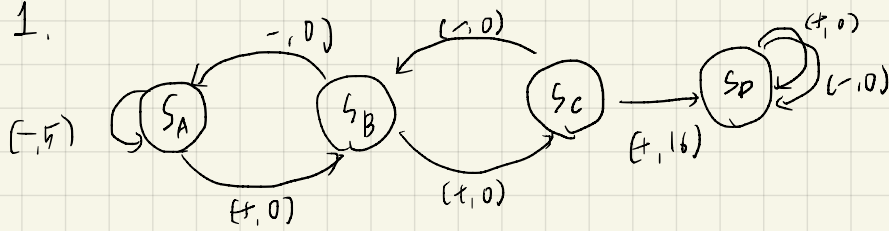


1.



all episodes have length 3

1a) \boxed{SA} if policy \oplus

$$\begin{aligned} SA \xrightarrow{\oplus} SB \xrightarrow{\oplus} SA \xrightarrow{\oplus} SA & 0 + 0 \times r + 5 \times r^2 \\ SA \xrightarrow{\oplus} SB \xrightarrow{\oplus} SC \xrightarrow{\oplus} SD & 0 + 0 \times r + 16 \times r^2 \\ SA \xrightarrow{\oplus} SB \xrightarrow{\oplus} SC \xrightarrow{\oplus} SB & 0 + 0 \times r + 0 \times r^2 \\ SA \xrightarrow{\oplus} SB \xrightarrow{\oplus} SA \xrightarrow{\oplus} SB & 0 + 0 \times r + 0 \times r^2 \end{aligned}$$

Value = expected utility of episodes

$$\boxed{21r^2}$$

$$\frac{21r^2}{4}$$

$\ominus \ominus$

$$\begin{aligned} SA \xrightarrow{\ominus} SA \xrightarrow{\ominus} SB \xrightarrow{\ominus} SC & 5 + 0 \times r + 0 \times r^2 \\ SA \xrightarrow{\ominus} SA \xrightarrow{\ominus} SB \xrightarrow{\ominus} SA & 5 + 0 \times r + 0 \times r^2 \\ SA \xrightarrow{\ominus} SA \xrightarrow{\ominus} SA \xrightarrow{\ominus} SA & 5 + 5 \times r + 5 \times r^2 \\ SA \xrightarrow{\ominus} SA \xrightarrow{\ominus} SA \xrightarrow{\ominus} SB & 5 + 5 \times r + 0 \end{aligned}$$

$$20 + 10r + 5r^2$$

$$\frac{20 + 10r + 5r^2}{4}$$

if $r = 0.001$ \ominus policy $\left(\frac{20 + 10r + 5r^2}{4} \right)$ is bigger than $\frac{21r^2}{4}$

so answer is " \ominus " $\pi(SA) = \ominus$

1b) if $r = 0.999$ \ominus policy is also bigger than \oplus

so answer is " \ominus " $\pi(SA) = \ominus$

1c) if policy $\oplus \oplus$

$$\begin{aligned} \oplus \oplus \oplus & 0 \\ \oplus \oplus \ominus & 16r \\ \oplus \oplus \oplus & 16r \\ \oplus \ominus \oplus & 0 \end{aligned}$$

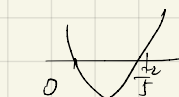
$\ominus \ominus$

$$\begin{aligned} \ominus \oplus \oplus & 0 \\ \ominus \ominus \ominus & 5r + 5r^2 \\ \ominus \ominus \oplus & 6r \\ \ominus \oplus \ominus & 0 \end{aligned}$$

$$\ominus \oplus \quad 10r + 5r^2 < 32r$$

$$5r^2 - 22r < 0$$

$$5r(r - \frac{22}{5}) < 0$$



$$\rightarrow \text{True} \quad [0 < r < 1]$$

$$\text{so } \pi(SB) = \oplus$$