4 (a)

Return: $G_t = \sum_{k=0}^{T-t-1} \gamma^k R_{t+k+1}$, $R$ is reward $\begin{cases} -1 & \text{failure} \\ 0 & \text{success} \end{cases}$

The end of the episode is when failure occures. That means

$$G_t = 0 + 0 + \cdots 0 + \gamma^{T-t-1}(-1) = -\gamma^{T-t-1}$$

So the return is $-\gamma^{T-t-1}$ if there is discouting factor

Others is same where we have return as $-\gamma^k$, $k$ is the time step before failure.

(b) Let $G_5 = 0$ terminal

$G_4 = R_5 = 2$

$G_3 = \gamma G_4 + R_4 = 4$

$G_2 = \gamma G_3 + R_3 = 8$

$G_1 = \gamma G_2 + R_2 = 6$

$G_0 = \gamma G_1 + R_1 = 2$

(c) $G_t = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}$, $G_0 = R_1 + \underbrace{\gamma \sum_{k=0}^{\infty} \gamma^k R_{k+2}}_{\gamma G_1}$

$R_{2\sim\infty} = 7$, $\gamma = 0.9 \rightarrow G_0 = 2 + \underbrace{\frac{0.9 \times 7}{1-0.9}}_{\text{등비급수의 합}} = 65$