

[2 points] Soit un MPD avec $S = \{s_0, s_1, s_2, s_3\}$ où s_2 est terminal, l'ensemble d'actions $\{a_1, a_2, a_3\}$ et le facteur d'escompte $\gamma = 0.5$. Supposez qu'un plan donné génère les essais suivants :

$$\begin{aligned} (s_0)_1 &\xrightarrow{a_1} (s_1)_2 \xrightarrow{a_2} (s_1)_2 \xrightarrow{a_2} (s_2)_{10} \\ (s_0)_1 &\xrightarrow{a_1} (s_0)_1 \xrightarrow{a_1} (s_1)_2 \xrightarrow{a_2} (s_1)_2 \xrightarrow{a_2} (s_0)_1 \xrightarrow{a_1} (s_2)_{10} \end{aligned}$$

Quelle sont les valeurs de ce plan en s_0 et s_1 , telles que calculées par l'estimation directe (vous pouvez simplement écrire l'expression numérique correspondant à ce calcul sans le faire vous-même) ?