

【片思いNAIST生の問題】

もうすぐクリスマスである..意中の相手にどうアピールすべきか...?

朝($t = 1$), 昼($t = 2$), 夜($t = 3$)からなる, 3ステップのエピソードを考えた場合に, 学生Aが朝($t = 1$)にとるべき適切な行動を求めよ.

- 意中の相手は, 研究室(S_1), 自宅(S_2), それ以外(S_3)のいずれかにいる

$$S = \{S_1, S_2, S_3\}$$

- ✓ 初期状態はランダム
- ✓ 以降は遷移行列に基づいて動く
- 学生Aは何らかの手段で意中の相手の居場所が観測可能
- 学生AはLINEをする(a_1)か, 何もしない(a_2)を選択しなければならない

$$S = \{S_1, S_2, S_3\}$$

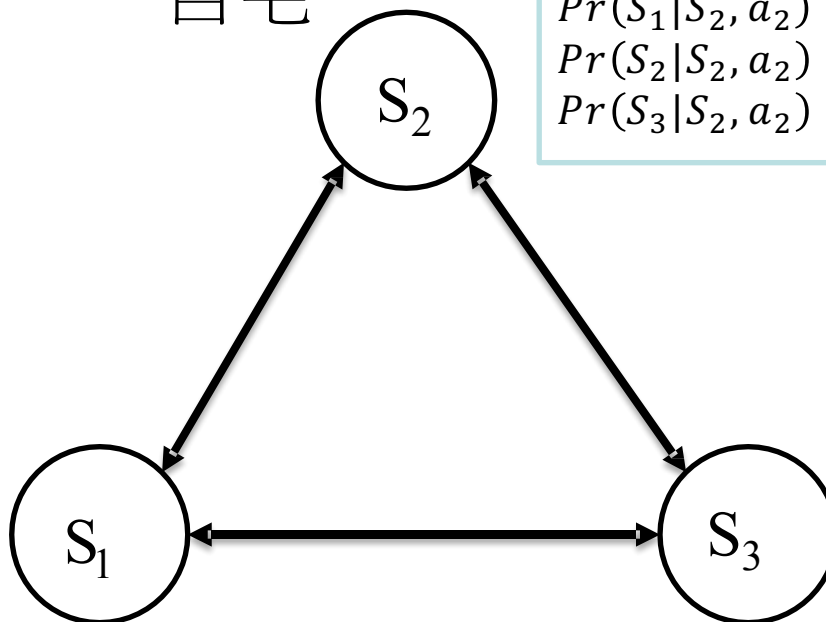
- 報酬 R は遷移先のみ依存
- $t = 3$ から $t = 4$ については, 必ず自宅に戻り, 報酬0を受け取る
- 学生Aは $t = 2$ において, 適切な行動を取る
- **割引率 = 0.7**

意中の相手の状態遷移関数

自宅

$$\begin{aligned}Pr(S_1|S_2, a_1) &= 0.2 \\Pr(S_2|S_2, a_1) &= 0.7 \\Pr(S_3|S_2, a_1) &= 0.1 \\Pr(S_1|S_2, a_2) &= 0.5 \\Pr(S_2|S_2, a_2) &= 0.1 \\Pr(S_3|S_2, a_2) &= 0.4\end{aligned}$$

研究室



スタバ

$$\begin{aligned}Pr(S_1|S_1, a_1) &= 0.5 \\Pr(S_2|S_1, a_1) &= 0.3 \\Pr(S_3|S_1, a_1) &= 0.2 \\Pr(S_1|S_1, a_2) &= 0.3 \\Pr(S_2|S_1, a_2) &= 0.2 \\Pr(S_3|S_1, a_2) &= 0.5\end{aligned}$$

$$\begin{aligned}Pr(S_1|S_3, a_1) &= 0.6 \\Pr(S_2|S_3, a_1) &= 0.2 \\Pr(S_3|S_3, a_1) &= 0.2 \\Pr(S_1|S_3, a_2) &= 0.2 \\Pr(S_2|S_3, a_2) &= 0.5 \\Pr(S_3|S_3, a_2) &= 0.3\end{aligned}$$

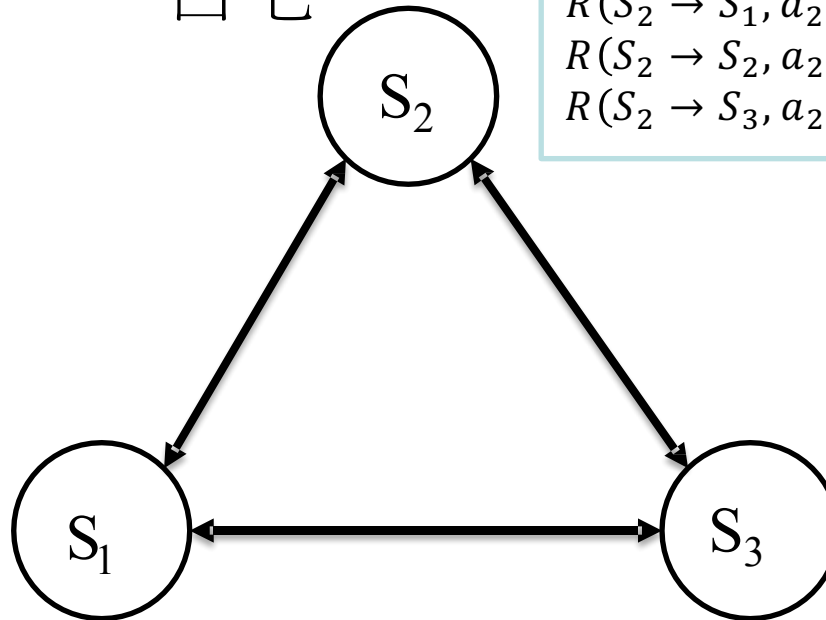
意中の相手の報酬関数

- 自宅に移動 or とどまる 0
- 研究室に移動 or とどまる 2
- その他に移動 or とどまる 3

自宅

$$\begin{aligned}R(S_2 \rightarrow S_1, a_1) &= 2 \\R(S_2 \rightarrow S_2, a_1) &= 0 \\R(S_2 \rightarrow S_3, a_1) &= 3 \\R(S_2 \rightarrow S_1, a_2) &= 2 \\R(S_2 \rightarrow S_2, a_2) &= 0 \\R(S_2 \rightarrow S_3, a_2) &= 3\end{aligned}$$

研究室



$$\begin{aligned}R(S_1 \rightarrow S_1, a_1) &= 2 \\R(S_1 \rightarrow S_2, a_1) &= 0 \\R(S_1 \rightarrow S_3, a_1) &= 3 \\R(S_1 \rightarrow S_1, a_2) &= 2 \\R(S_1 \rightarrow S_2, a_2) &= 0 \\R(S_1 \rightarrow S_3, a_2) &= 3\end{aligned}$$

スタバ

$$\begin{aligned}R(S_3 \rightarrow S_1, a_1) &= 2 \\R(S_3 \rightarrow S_2, a_1) &= 0 \\R(S_3 \rightarrow S_3, a_1) &= 3 \\R(S_3 \rightarrow S_1, a_2) &= 2 \\R(S_3 \rightarrow S_2, a_2) &= 0 \\R(S_3 \rightarrow S_3, a_2) &= 3\end{aligned}$$

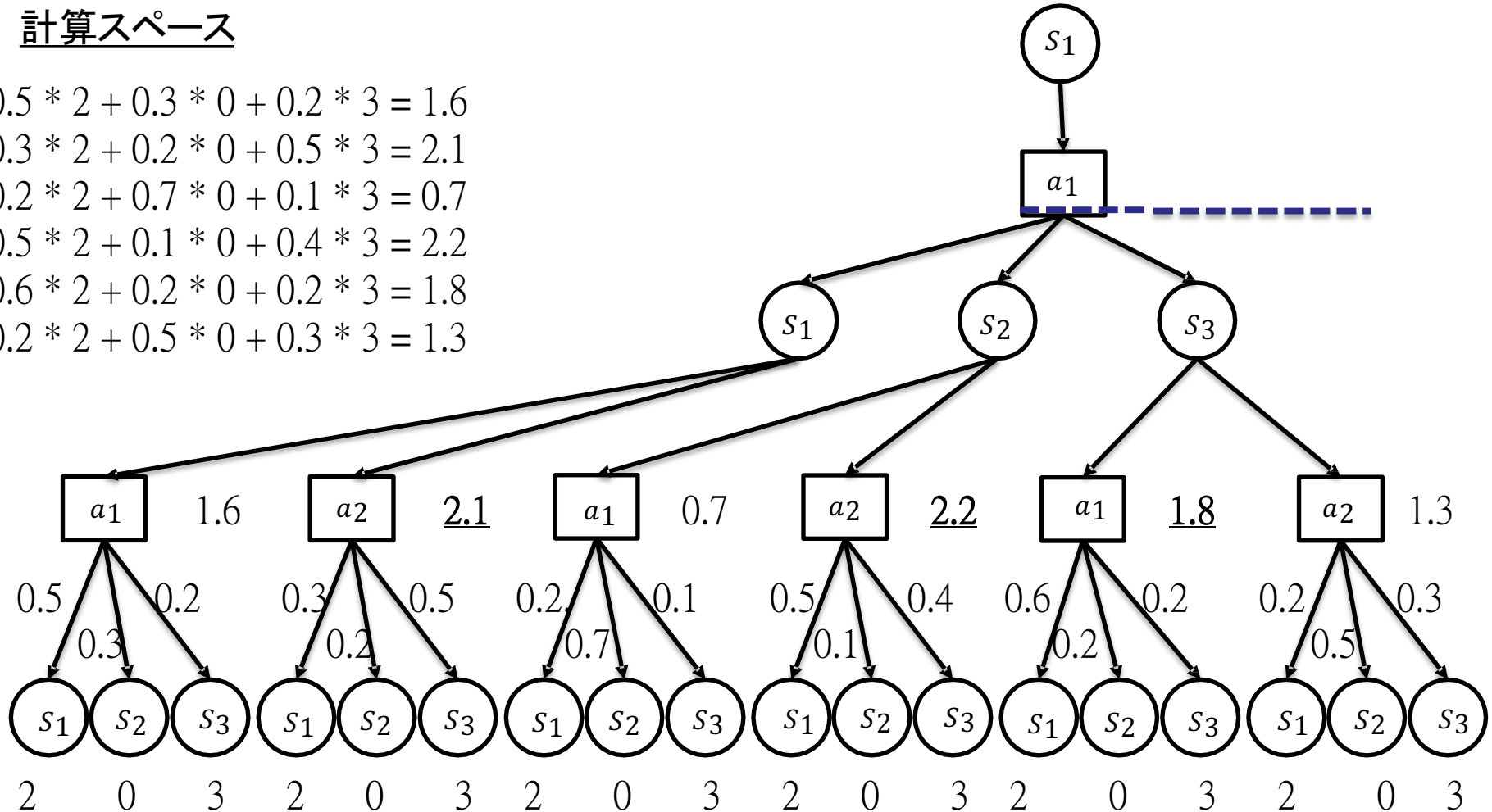
ヒント & 解答用紙

破線部 ----- に各行動後の報酬期待値を記入せよ・

報酬期待値 = (行動直後に得られる報酬期待値) + 割引率 * (次の行動で得られる報酬期待値)

計算スペース

- $0.5 * 2 + 0.3 * 0 + 0.2 * 3 = 1.6$
- $0.3 * 2 + 0.2 * 0 + 0.5 * 3 = 2.1$
- $0.2 * 2 + 0.7 * 0 + 0.1 * 3 = 0.7$
- $0.5 * 2 + 0.1 * 0 + 0.4 * 3 = 2.2$
- $0.6 * 2 + 0.2 * 0 + 0.2 * 3 = 1.8$
- $0.2 * 2 + 0.5 * 0 + 0.3 * 3 = 1.3$

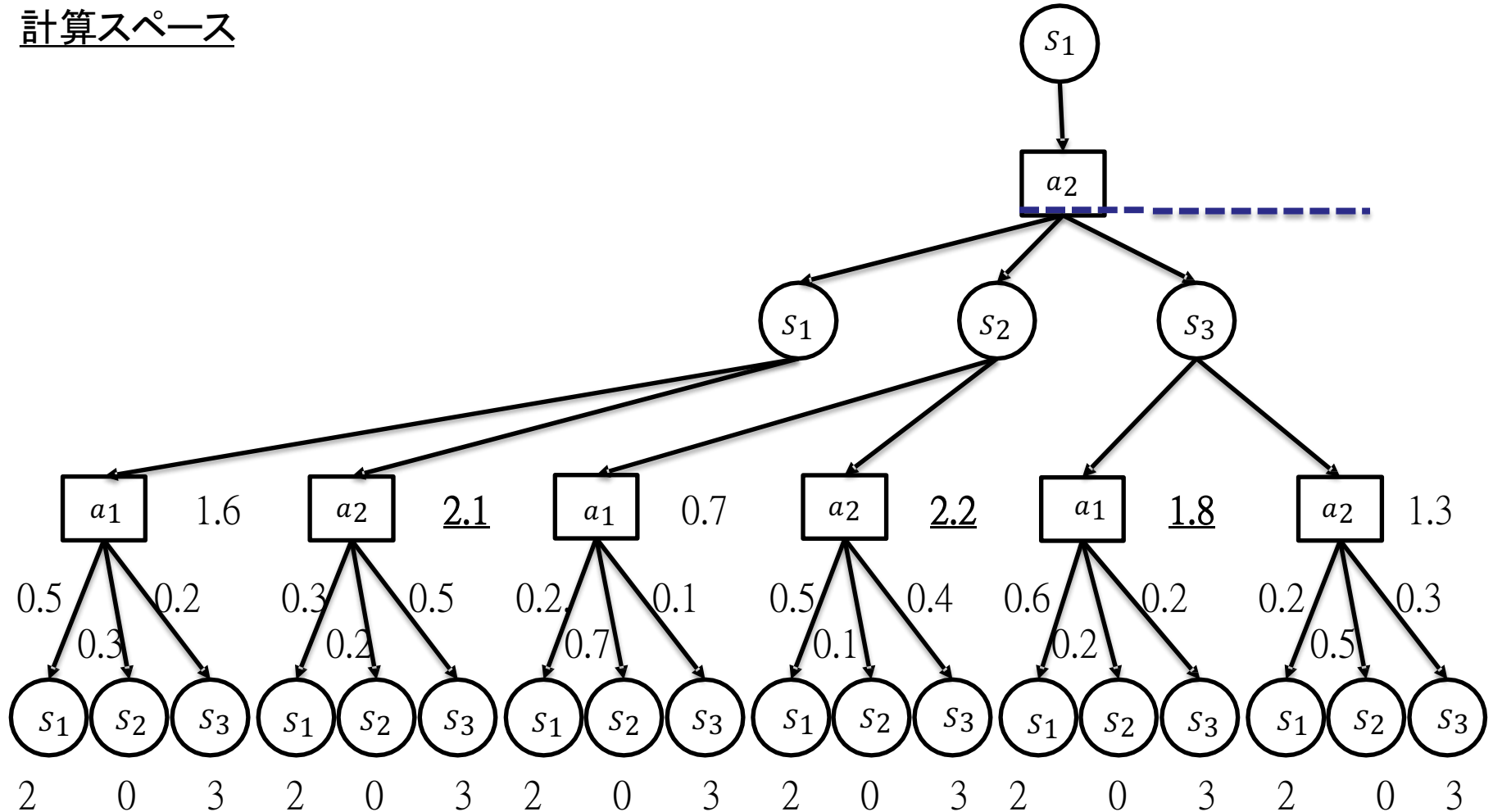


ヒント & 解答用紙

破線部 ----- に各行動後の報酬期待値を記入せよ・

報酬期待値 = (行動直後に得られる報酬期待値) + 割引率 * (次の行動で得られる報酬期待値)

計算スペース

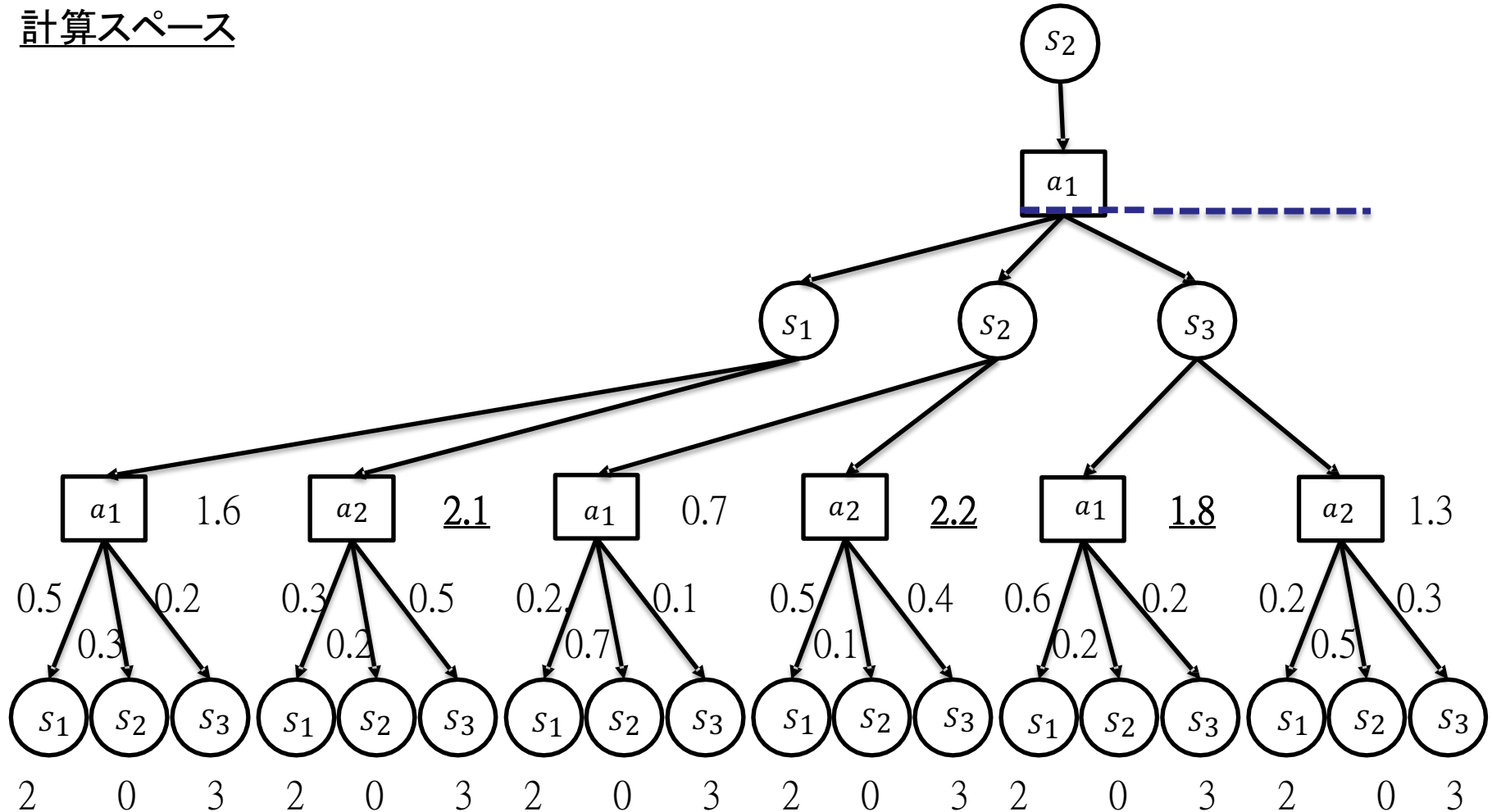


ヒント & 解答用紙

破線部 ----- に各行動後の報酬期待値を記入せよ・

報酬期待値 = (行動直後に得られる報酬期待値) + 割引率 * (次の行動で得られる報酬期待値)

計算スペース

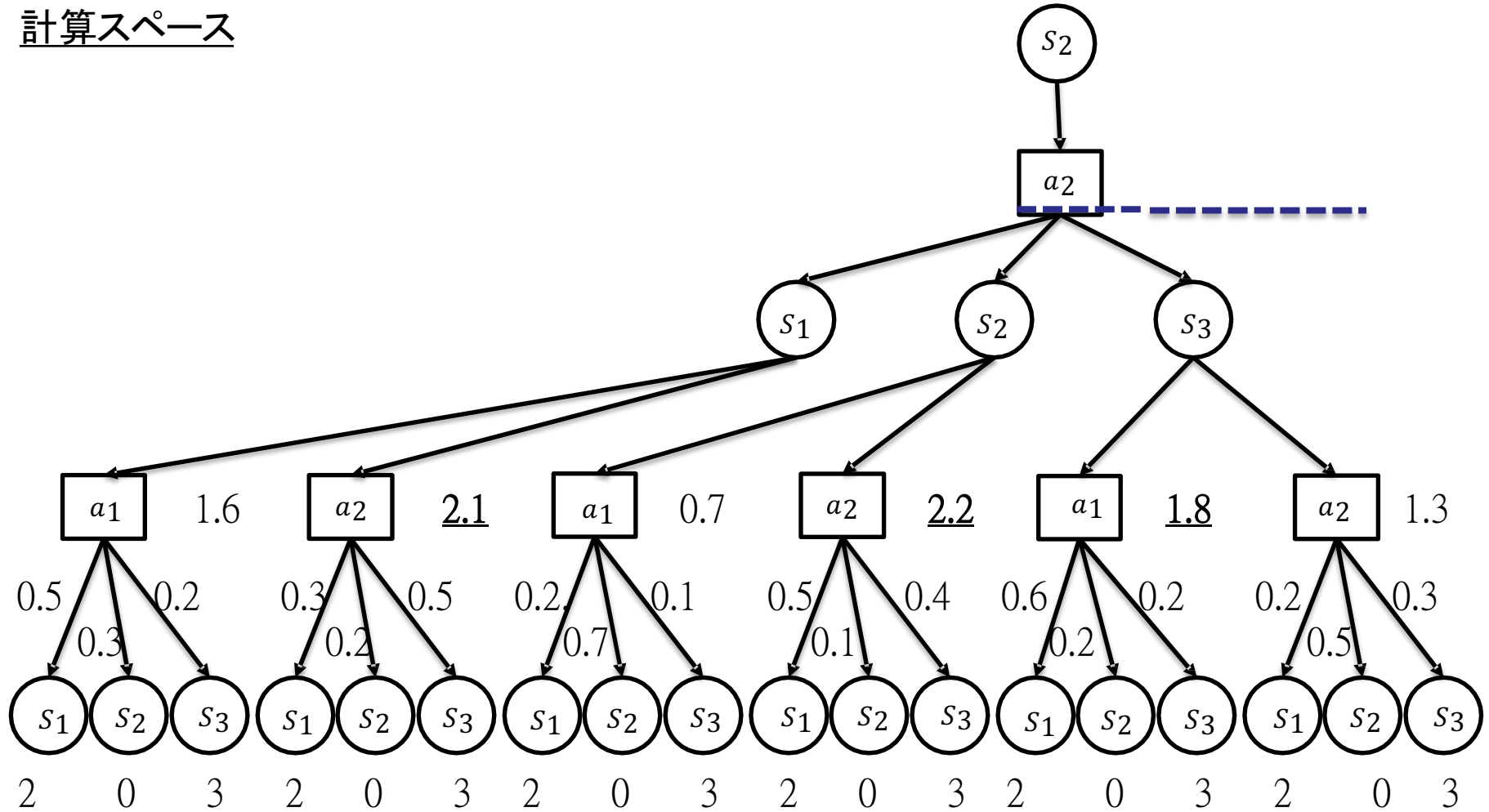


ヒント & 解答用紙

破線部 ----- に各行動後の報酬期待値を記入せよ・

報酬期待値 = (行動直後に得られる報酬期待値) + 割引率 * (次の行動で得られる報酬期待値)

計算スペース

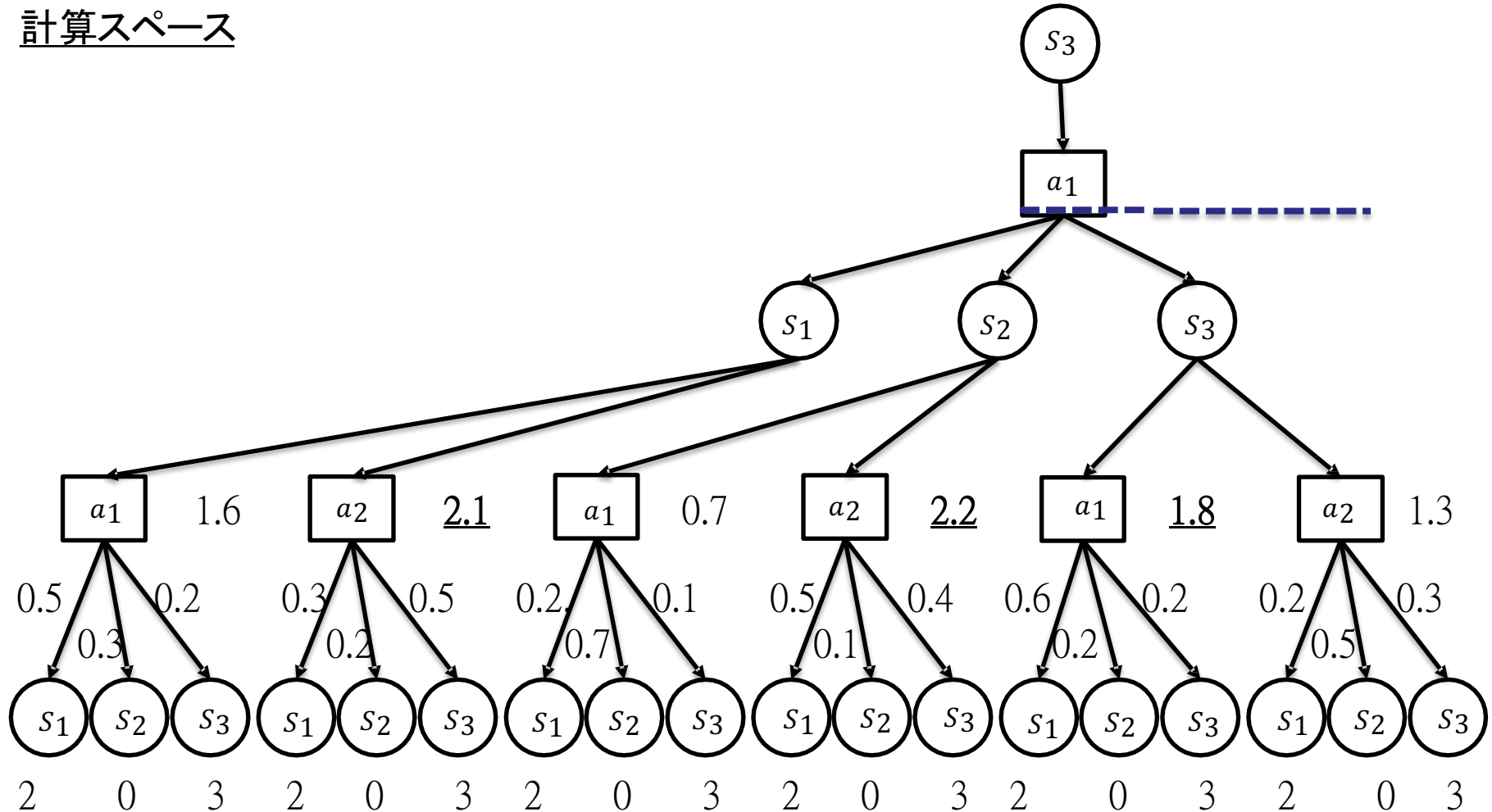


ヒント & 解答用紙

破線部 ----- に各行動後の報酬期待値を記入せよ・

報酬期待値 = (行動直後に得られる報酬期待値) + 割引率 * (次の行動で得られる報酬期待値)

計算スペース

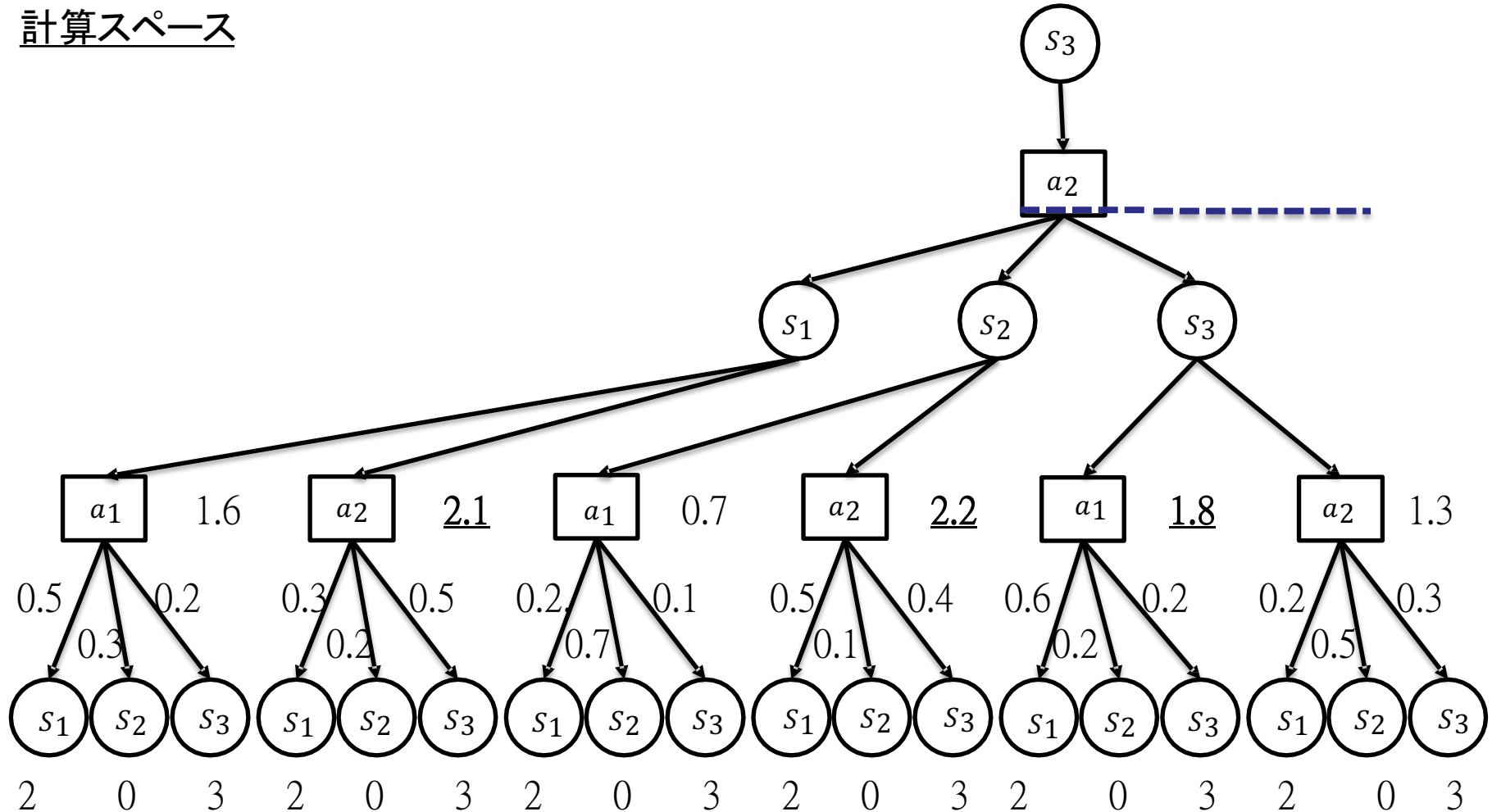


ヒント & 解答用紙

破線部 ----- に各行動後の報酬期待値を記入せよ・

報酬期待値 = (行動直後に得られる報酬期待値) + 割引率 * (次の行動で得られる報酬期待値)

計算スペース



解答用紙

自宅(s_1)における適切な行動は _____

期待値は _____

研究室(s_2)における適切な行動は _____

期待値は _____

その他(s_3)における適切な行動は _____

期待値は _____