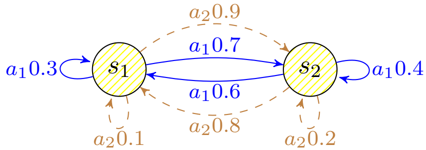
Discrete-time Partially Observalbe Markov Decision Process: 6-tuple $(\mathbb{S}, \mathbb{O}, \mathbb{A}, \mathbf{T}, \mathbf{r}, \mathbf{O})$

e.g.



$\mathbb{S} = \{s_1, s_2\}$,$\mathbb{O} = \{o_1, o_2\}$,$\mathbb{A} = \{a_1, a_2\}$, Transition model $\mathbf{T}$, Reward model $\mathbf{r}$, Sensor Model $\mathbf{O}$

Transition Model under Markovian Property: $T(s'|s,a)$, for discrete states we use transition matrix

$$\mathbf{T} = [T(s'|s,a)] = \begin{array}{cc} & \\ s_1 & \begin{array}{c} a_1 \\ a_2 \end{array} \\ s_2 & \begin{array}{c} a_1 \\ a_2 \end{array} \end{array} \begin{array}{c} \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{bmatrix} 0.3 & 0.7 \\ 0.1 & 0.9 \\ 0.6 & 0.4 \\ 0.8 & 0.2 \end{bmatrix} \end{array} = \begin{array}{cc} & \\ a_1 & \begin{array}{c} s_1 \\ s_2 \end{array} \\ a_2 & \begin{array}{c} s_1 \\ s_2 \end{array} \end{array} \begin{array}{c} \begin{array}{cc} o_1 & o_2 \end{array} \\ \begin{bmatrix} T_{111} & T_{112} \\ T_{121} & T_{122} \\ T_{211} & T_{212} \\ T_{221} & T_{222} \end{bmatrix} \end{array}$$

Reward Model: $r(s,a)$ or $r(s,a,s')$ below is an example of deterministic reward table:

$$\mathbf{r} = [r(s,a)] = \begin{array}{cc} & \\ s_1 & \begin{array}{c} a_1 \\ a_2 \end{array} \\ s_2 & \begin{array}{c} a_1 \\ a_2 \end{array} \end{array} \begin{array}{c} r \\ \begin{bmatrix} r_{11} \\ r_{12} \\ r_{21} \\ r_{22} \end{bmatrix} \end{array} \text{ or } \mathbf{r} = [r(s,a,s')] = \begin{array}{cc} & \\ s_1 & \begin{array}{c} a_1 \\ a_2 \end{array} \\ s_2 & \begin{array}{c} a_1 \\ a_2 \end{array} \end{array} \begin{array}{c} \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{bmatrix} r_{111} & r_{112} \\ r_{121} & r_{122} \\ r_{211} & r_{212} \\ r_{221} & r_{222} \end{bmatrix} \end{array}$$

Sensor model: Conditional Probability of Observation $O(o'|s')$ or $O(o'|s',a)$

$$\mathbf{O} = [O(o'|s')] = \begin{array}{c} \\ s_1 \\ s_2 \end{array} \begin{array}{c} \begin{array}{cc} o_1 & o_2 \end{array} \\ \begin{bmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{bmatrix} \end{array} \text{ or } \mathbf{O} = [O(o'|s',a)] = \begin{array}{cc} & \\ a_1 & \begin{array}{c} s_1 \\ s_2 \end{array} \\ a_2 & \begin{array}{c} s_1 \\ s_2 \end{array} \end{array} \begin{array}{c} \begin{array}{cc} o_1 & o_2 \end{array} \\ \begin{bmatrix} O_{111} & O_{112} \\ O_{121} & O_{122} \\ O_{211} & O_{212} \\ O_{221} & O_{222} \end{bmatrix} \end{array}$$