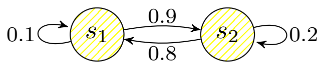
Markov Chain: 2−tuple $(\mathbb{S}, \mathbf{P})$

$\mathbb{S} = \{s_1, s_2\}, \mathbf{P} = \mathbf{P}(s)$



Transition Probability Matrix

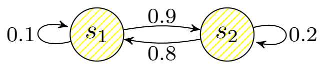
$$\mathbf{P} = [\pi(s'|s)] = \begin{array}{c} \\ s_1 \\ s_2 \end{array} \begin{array}{c} \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{bmatrix} 0.1 & 0.9 \\ 0.8 & 0.2 \end{bmatrix} \end{array}$$

$\pi(s'|s)$: Next state s' is determined only by the current state s

States are completely observable i.e. $p(o_1|s_1) = 1$, $p(o_2|s_2) = 1$

Hidden Markov Model: 3−tuple $(\mathbb{S}, \mathbb{O}, \mathbf{P})$

$\mathbb{S} = \{s_1, s_2\}, \mathbb{O} = \{o_1, o_2\}, \mathbf{P} = \mathbf{P}(s)$



Transition Probability Matrix

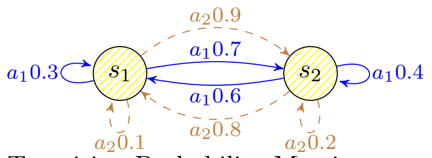
$$\mathbf{P} = [\pi(s'|s)] = \begin{array}{c} \\ s_1 \\ s_2 \end{array} \begin{array}{c} \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{bmatrix} 0.1 & 0.9 \\ 0.8 & 0.2 \end{bmatrix} \end{array}$$

$\pi(s'|s)$: Next state s' is determined only by the current state s

We are unsure which state we are in e.g. $p(o_1|s_1) = 0.75$, $p(o_2|s_2) = 0.75$

Markov Decision Process: 4−tuple $(\mathbb{S}, \mathbb{A}, \mathbf{P}, \mathbf{R})$

$\mathbb{S} = \{s_1, s_2\}$, $\mathbb{A} = \{a_1, a_2\}$, $\mathbf{P} = \mathbf{P}(s, a)$, $\mathbf{R} = \mathbf{R}(s, a)$



Transition Probability Matrix

$$\mathbf{P} = [\pi(s'|s, a)] = \begin{array}{cc} & a_1 \\ s_1 & a_2 \\ s_2 & a_1 \\ & a_2 \end{array} \begin{array}{c} \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{bmatrix} a_1 & 0.7 \\ 0.1 & 0.9 \\ 0.6 & 0.4 \\ 0.8 & 0.2 \end{bmatrix} \end{array}$$

$\pi(s'|s, a)$: Next state s' is determined by current state s and current action a

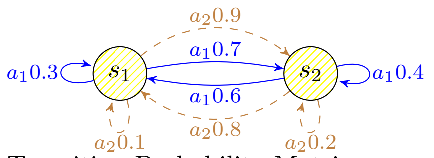
Reward Matrix

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

States are completely observable i.e. $p(o_1|s_1) = 1$, $p(o_2|s_2) = 1$

Partially Observalbe Markov Decision Process: 5-tuple $(\mathbb{S}, \mathbb{O}, \mathbb{A}, \mathbf{P}, \mathbf{R})$

$\mathbb{S} = \{s_1, s_2\}, \mathbb{O} = \{o_1, o_2\}, \mathbb{A} = \{a_1, a_2\}$, $\mathbf{P} = \mathbf{P}(s, a)$, $\mathbf{R} = \mathbf{R}(s, a)$



Transition Probability Matrix

$$\mathbf{P} = [\pi(s'|s, a)] = \begin{array}{cc} & a_1 \\ s_1 & a_2 \\ s_2 & a_1 \\ & a_2 \end{array} \begin{array}{c} \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{bmatrix} 0.3 & 0.7 \\ 0.1 & 0.9 \\ 0.6 & 0.4 \\ 0.8 & 0.2 \end{bmatrix} \end{array}$$

$\pi(s'|s, a)$: Next state s' is determined by current state s and current action a

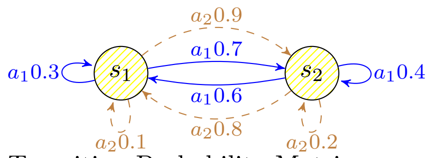
Reward Matrix

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

We are unsure which state we are in e.g. $p(o_1|s_1) = 0.75$, $p(o_2|s_2) = 0.75$

Semi-Markov Decision Process: 5−tuple $(\mathbb{S}, \mathbb{A}, \mathbf{P}, \mathbf{R}, \mathbf{F})$

$\mathbb{S} = \{s_1, s_2\}$, $\mathbb{A} = \{a_1, a_2\}$, $\mathbf{P} = \mathbf{P}(s, a)$, $\mathbf{R} = \mathbf{R}(s, a)$, $\mathbf{F} = \mathbf{F}(s, a)$



Transition Probability Matrix

$$\mathbf{P} = [\pi(s'|s, a)] = \begin{array}{cc} & a_1 \\ s_1 & a_2 \\ s_2 & a_1 \\ & a_2 \end{array} \begin{array}{c} \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{bmatrix} 0.3 & 0.7 \\ 0.1 & 0.9 \\ 0.6 & 0.4 \\ 0.8 & 0.2 \end{bmatrix} \end{array}$$

$\pi(s'|s, a)$: Next state s' is determined by current state s and current action a

Reward Matrix

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

States are completely observable i.e. $p(o_1|s_1) = 1$, $p(o_2|s_2) = 1$

Duration distribution $F(s, a, t) = Pr(t_s \leq t|a, s)$, which represents the probability of staying in state (s) for a duration less equal than t after taking action a