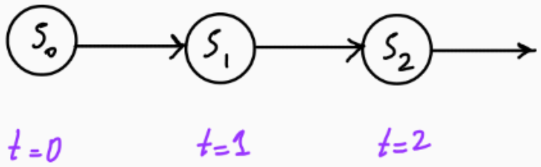
Graphical Representation of MC and MDP

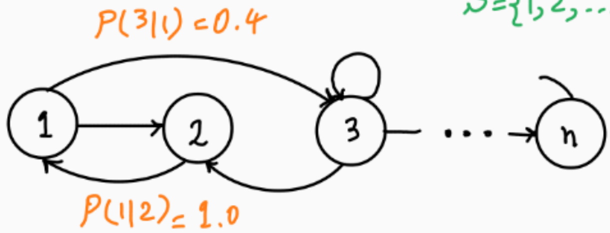
Temporal Unrolling

Transition - Based

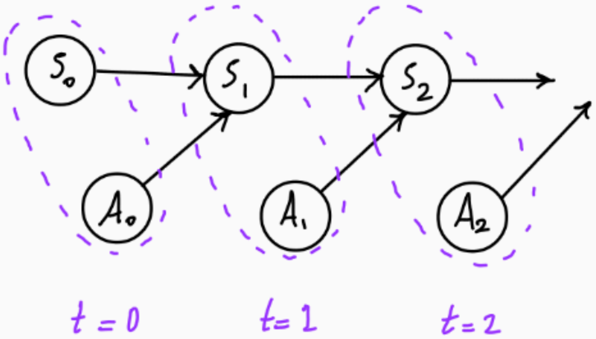
$S = \{1, 2, \dots, n\}$

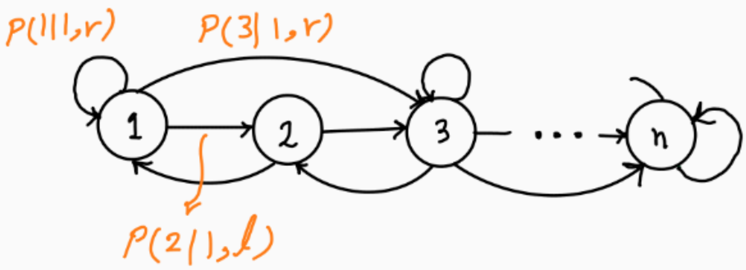
MC





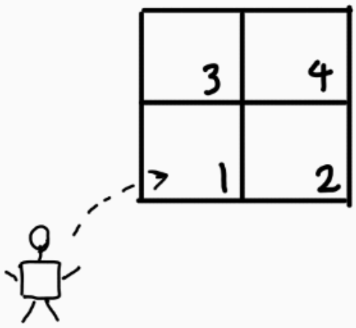
MDP





MDP + policy $\longrightarrow$ (induced) MC

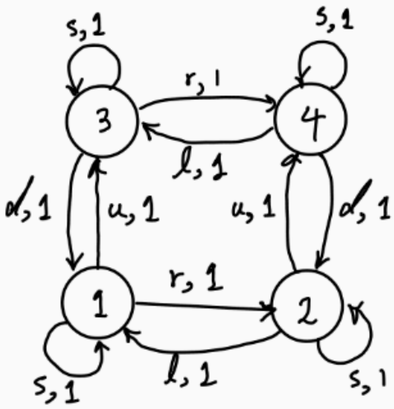
edge label: (action, transition probability)



$S = \{1, 2, 3, 4\}$

$A = \{r, l, u, d, s\}$

MDP



Policy

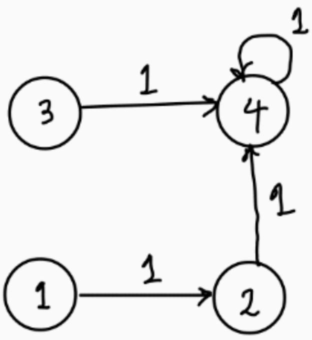
if $s = 1 \longrightarrow a = r$

if $s = 2 \longrightarrow a = u$

if $s = 3 \longrightarrow a = r$

if $s = 4 \longrightarrow a = s$

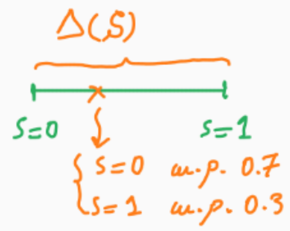
(induced) MC



State Distribution over MC

$$\mu_t \in \Delta(S)$$

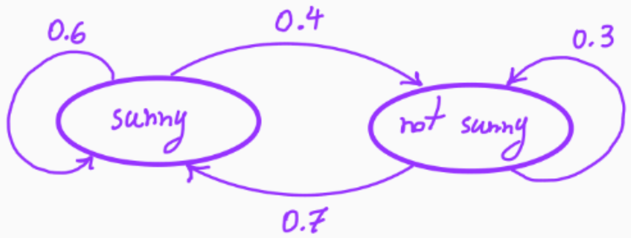
$$\mu_t(s) = \mathbb{P}[S_t = s]$$



$$\mu_{t+1}(s') = \mathbb{P}[S_{t+1} = s']$$

$$= \sum_{s\in S} \mathbb{P}[S_{t+1} = s', S_t = s]$$

$$= \sum_{s\in S} \underbrace{\mathbb{P}[S_{t+1} = s' \mid S_t = s]}_{P(s'|s)} \underbrace{\mathbb{P}[S_t = s]}_{\mu_t(s)}$$



$$\begin{cases} \mu_{10}(\text{sunny}) = 1 \\ \mu_{10}(\text{not sunny}) = 0 \end{cases}$$

$$\begin{cases} \mu_{11}(\text{sunny}) = 0.6 \\ \mu_{11}(\text{not sunny}) = 0.4 \end{cases}$$

$$\begin{cases} \mu_{12}(\text{sunny}) = 0.6\times 0.6 + 0.4\times 0.7 \\ \mu_{12}(\text{not sunny}) = 1 - \mu_{12}(\text{sunny}) \end{cases}$$

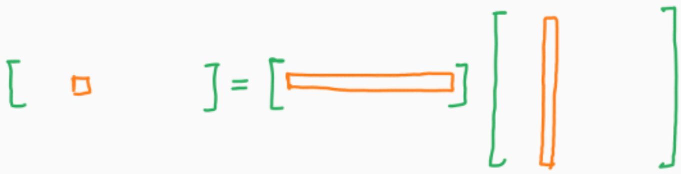
row vector

Let $\mu_t = \big[\, \mathbb{P}[S_t=1] \;\; \mathbb{P}[S_t=2] \;\; \cdots \;\; \mathbb{P}[S_t=n] \,\big]$

$$P = \begin{array}{cc} & j \\ i & \left[\begin{array}{ccc} & & \\ & & \circ \\ & & \end{array}\right]_{n\times n} \end{array} \quad \circ \to \mathbb{P}[j|i]$$

$$\mu_{t_0+1} = \mu_{t_0} P$$

$$\mu_{t_0+2} = \mu_{t_0+1} P = \mu_{t_0} P^2$$



$$\vdots$$

$$\mu_{t_0+t} = \mu_{t_0+t-1} P = \mu_{t_0+t-2} P^2 = \cdots = \mu_{t_0} P^t$$

t-step transition matrix

$$\boxed{\text{Chapman - Kolmogorov Equation}}$$

$$\mu_t = \begin{bmatrix} \mathbb{P}[S_t=1] \\ \mathbb{P}[S_t=2] \\ \vdots \\ \mathbb{P}[S_t=n] \end{bmatrix}$$

column vector

$$\mu_{t_0+1} = P^T \mu_{t_0} \longrightarrow \mu_{t_0+t} = (P^t)^T \mu_{t_0}$$

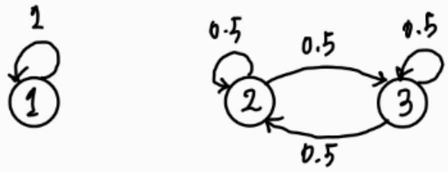
Stationary (steady state) distribution: A state distribution $\bar{\mu}$

such that $\begin{cases} \bar{\mu}(s) \geq 0 \quad \forall s \in S, \quad \sum_{s \in S} \bar{\mu}(s) = 1 \\ \bar{\mu} = \bar{\mu} P \end{cases}$

$$\bar{\mu} P^2 = \bar{\mu} P P = \bar{\mu} P = \bar{\mu}$$

Example:

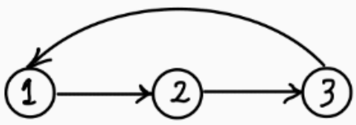
not irreducible 1)



$\mu_0 = [1 \quad 0 \quad 0]$ stationary

$\mu_0' = [0 \quad 0.5 \quad 0.5]$ stationary

periodic 2)



$\mu_0 = [1 \quad 0 \quad 0] \to \mu_1 = [0 \quad 1 \quad 0] \to \mu_2 = [0 \quad 0 \quad 1]$

$\to \mu_3 = [1 \quad 0 \quad 0]$ not stationary

$\mu_0' = [0.7 \quad 0.3 \quad 0] \to \mu_1' = [0 \quad 0.7 \quad 0.3]$ not stationary

$\bar{\mu} = [\frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{3}]$ stationary

Theorem: An ergodic (irreducible and aperiodic) has a unique stationary distribution $\bar{\mu}$.

Perron-Frobenius Theorem

$$\lim_{t \to \infty} P^t = \mathbf{1} \bar{\mu}$$

$$\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} [\bar{\mu}(1)\ \bar{\mu}(2) \cdots \bar{\mu}(n)] = \begin{bmatrix} \bar{\mu}(1)\ \bar{\mu}(2) \cdots \bar{\mu}(n) \\ \text{---------} \\ \text{---------} \\ \vdots \\ \text{---------} \end{bmatrix}$$

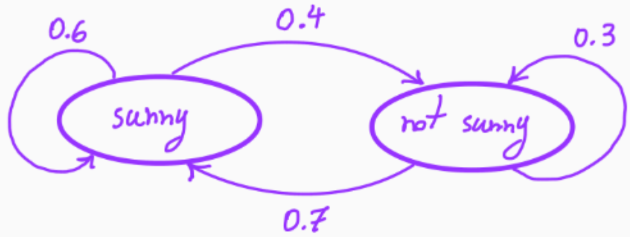
Note: The speed of convergence (starting from any $\mu_0 \in \Delta(S)$) to the stationary distribution $\bar{\mu}$ is characterized by $\frac{\lambda_2}{\lambda_1}$.

$$1 = \lambda_1 > |\lambda_2| \geq |\lambda_3| \cdots$$

$\frac{\lambda_2}{\lambda_1} \downarrow \longrightarrow$ speed of convergence $\uparrow$

Example :

$$S = \begin{array}{cc} & \\ [\ \text{sunny} & \text{not sunny}\ ] \\ S & NS \end{array}$$



$$P = \begin{array}{cc} S & N \\ \left[\begin{matrix} 0.6 & 0.4 \\ 0.7 & 0.3 \end{matrix}\right] \end{array} \begin{array}{c} S \\ N \end{array}$$

$S \longrightarrow S$ 1

$S \longrightarrow NS \longrightarrow S$ 2

$S \longrightarrow NS \longrightarrow NS \longrightarrow S$ 3

$$\mu_0 = [\mu_0(S) \quad \mu_0(NS)]$$

$$\bar{\mu} = \bar{\mu} P \longrightarrow \bar{\mu} I = \bar{\mu} P \longrightarrow \bar{\mu}(P - I) = \mathbf{0}_{1\times n}$$

$$\bar{\mu} = \left[\tfrac{7}{11} \quad \tfrac{4}{11}\right] \qquad \left[\tfrac{7}{11} \quad \tfrac{4}{11}\right] \begin{bmatrix} 0.6 & 0.4 \\ 0.7 & 0.3 \end{bmatrix} = \left[\tfrac{7}{11} \quad \tfrac{4}{11}\right]$$

$$\mu_0 \longrightarrow \mu_1 \longrightarrow \mu_2 \longrightarrow \cdots \quad \mu_t$$

for large t

$$\mu_t \approx \bar{\mu}$$