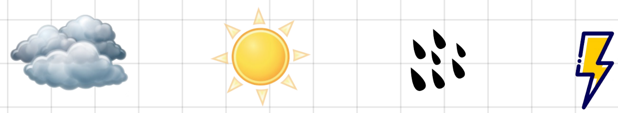
CS70 – Spring 2024

Lecture 24 – April 16

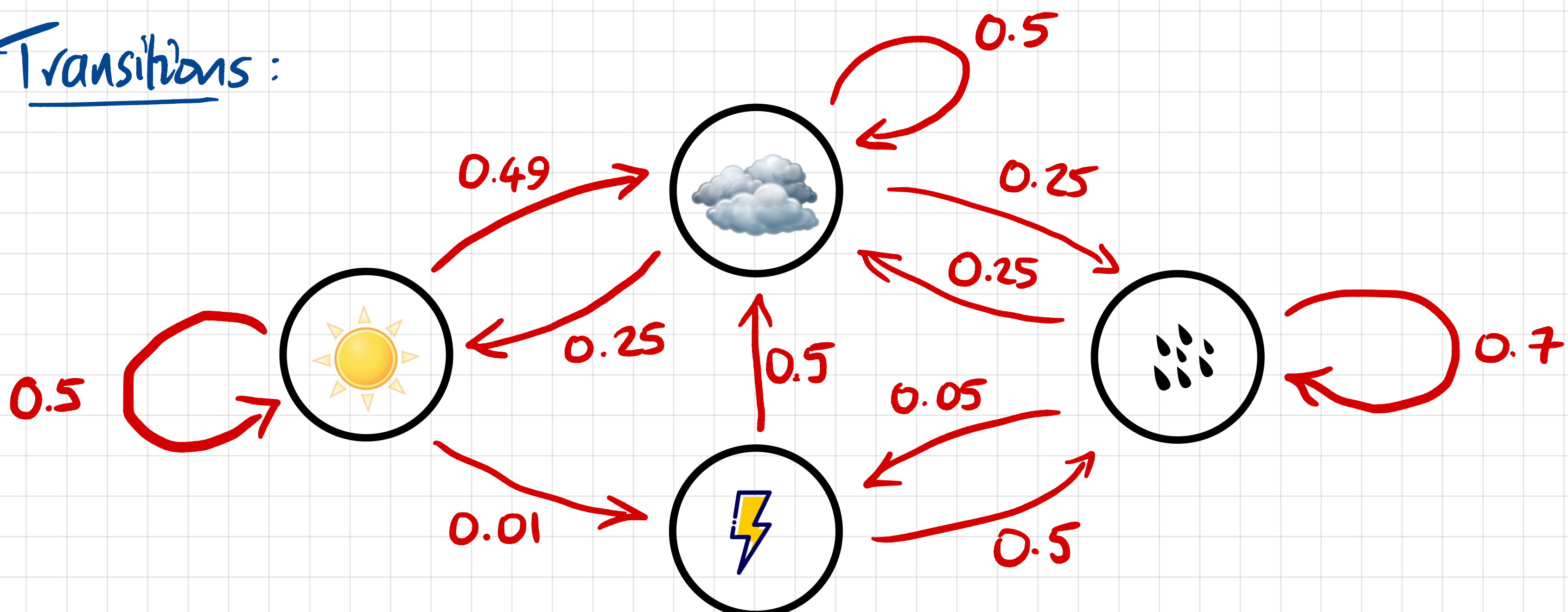
Markov chains

Model for describing systems that move from state to state via random transitions

Example: simple weather system

4 states: 

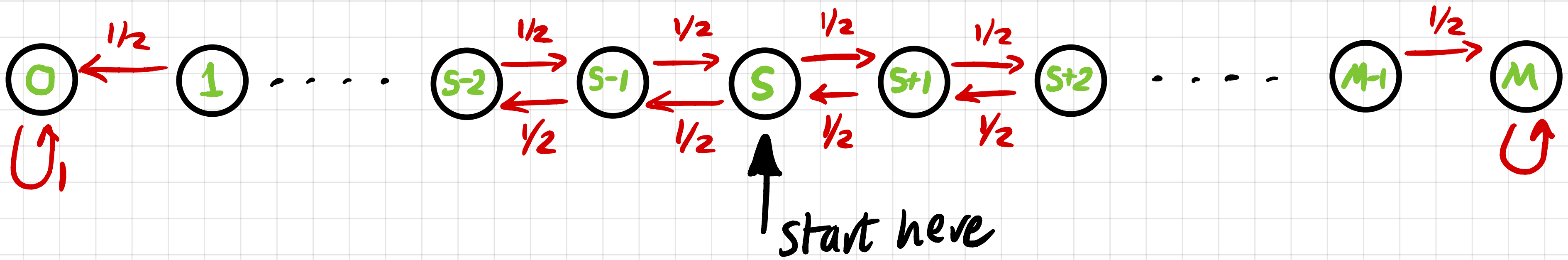
Transitions:



Key property: Distribution of next state depends only on current state

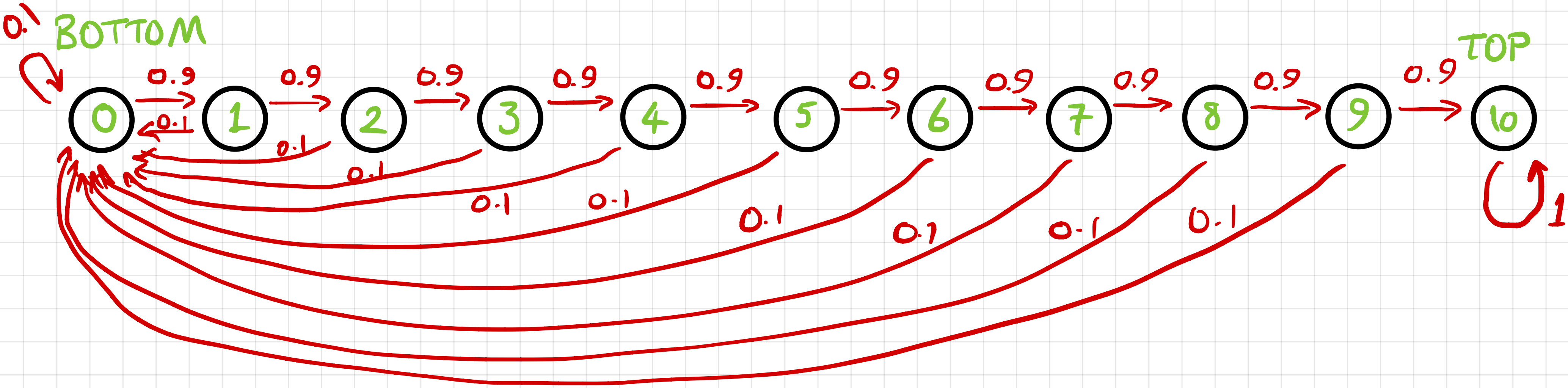
Example: Fair game: win/lose \$1 each with prob. $1/2$

Start with \$S, end when reach \$0 or \$M



Example: Climbing a (very slippery) 10-rung ladder

On each step, slip down to bottom w. prob. 0.1



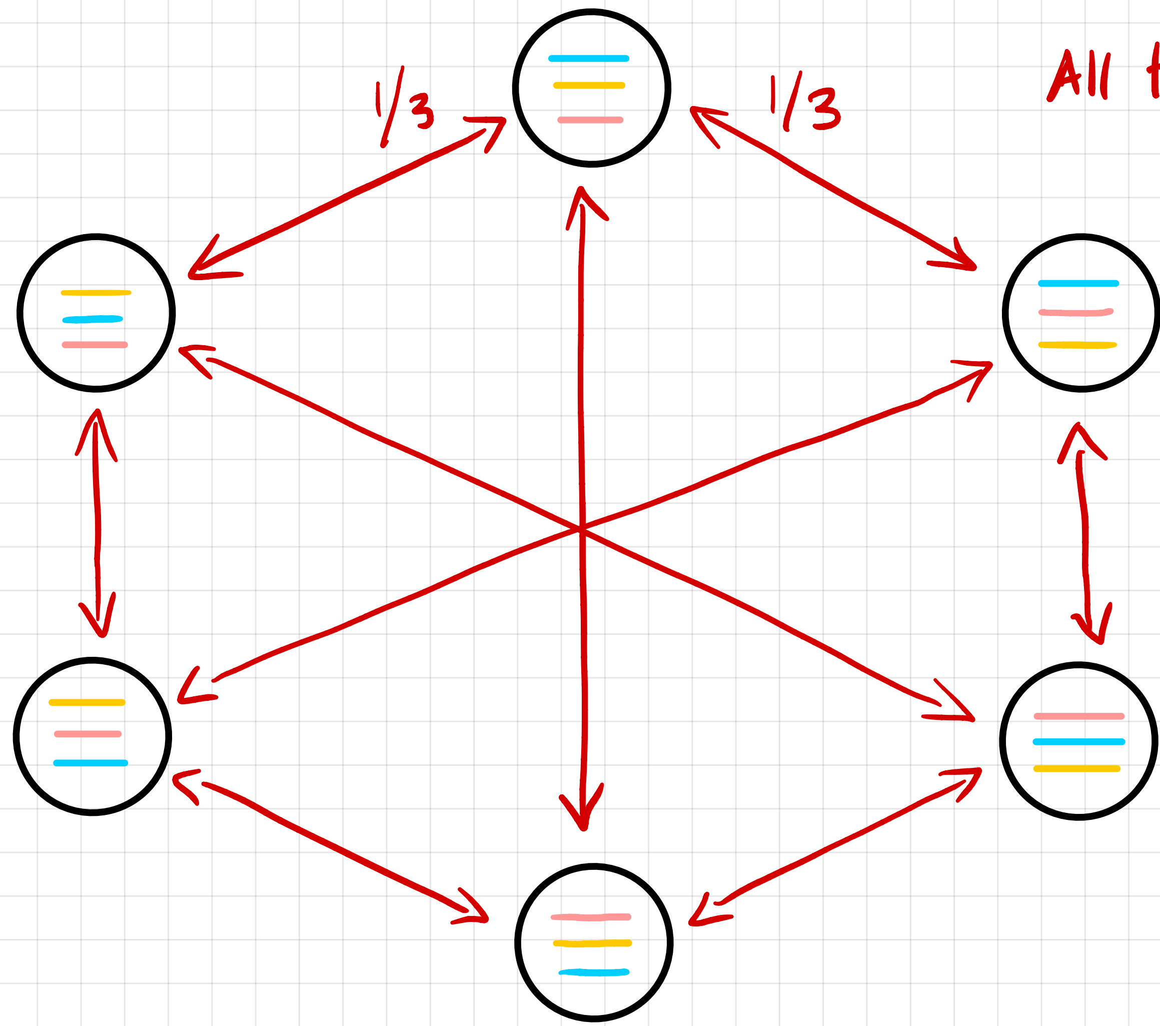
Example: Shuffling cards (slowly!)

States: all $n!$ permutations of the deck (n cards)

Transitions: pick 2 random cards & switch them

$\boxed{n=3}$

All trans. probs. $= 1/3$



$\longleftrightarrow$ denotes 2 edges $\rightleftarrows$

$$i \xrightarrow{P(i,j)} j$$

Formal Set-Up

<u>State space</u>: $\mathcal{K} = \{1, 2, \dots, K\}$ for finite K

<u>Transition matrix</u>: P, a $K \times K$ real matrix satsifying:

P is called "stochastic"

$$P(i,j) \geq 0 \quad \forall i,j \in \mathcal{K} \qquad \text{[non-negative]}$$

$$\sum_j P(i,j) = 1 \quad \forall i \in \mathcal{K} \qquad \text{[row sums = 1]}$$

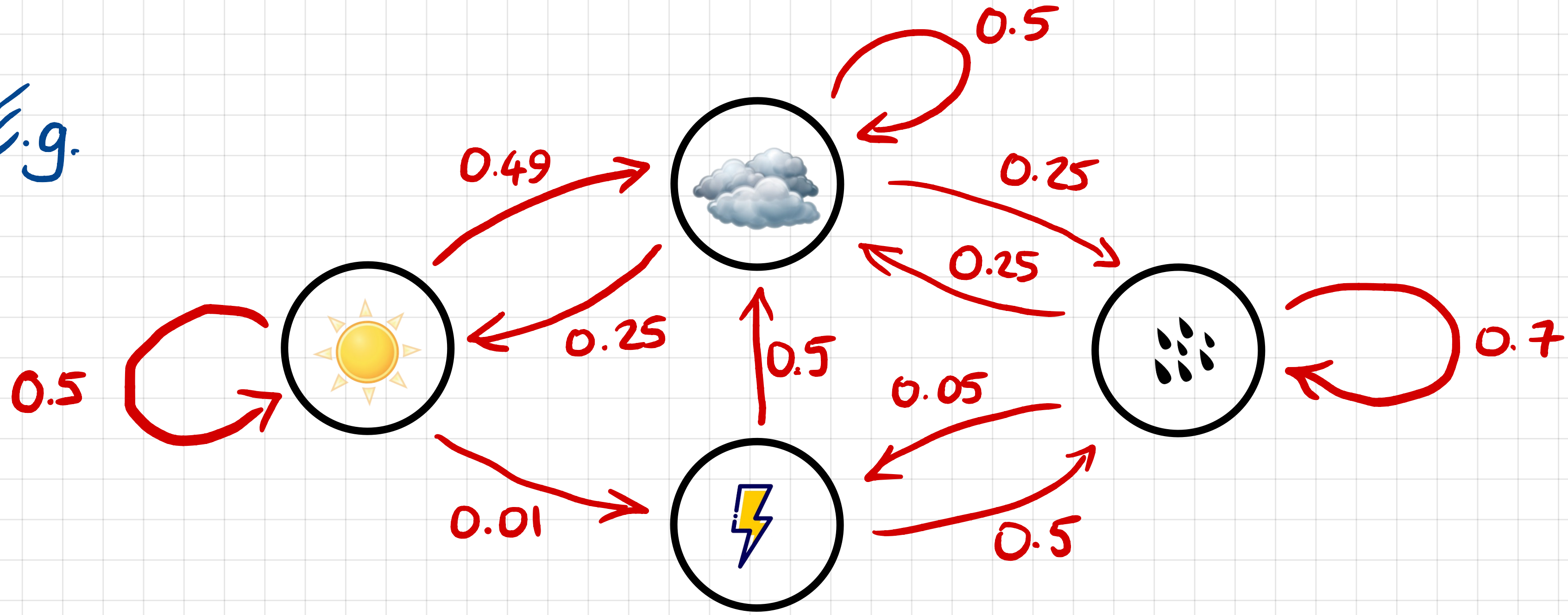
Given any $X_0 \in \mathcal{K}$, define random seq. $X_0, X_1, X_2, \dots$ by

$$\Pr[X_{n+1} = j \mid X_n = i, X_{n-1}, \dots, X_0] = P(i,j)$$

<u>Note</u>: This transition probability depends <u>only on</u> $X_n = i$!

More generally: X_0 has any <u>probability distribution</u> on $\mathcal{K}$

E.g.



Transition matrix $P =$

	☁️	☀️	💧	⚡
☁️	0.5	0.25	0.25	0
☀️	0.49	0.5	0	0.01
💧	0.25	0	0.7	0.05
⚡	0.5	0	0.5	0

Matrix-vector formulation

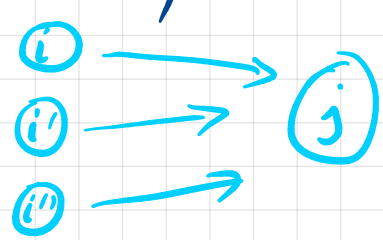
Let π_n be a row vector describing the probability distribution over states after n transitions, i.e.,

$$\pi_n(i) := \Pr[X_n = i]$$

Given π_n, what does π_{n+1} look like?

$$\pi_{n+1}(j) = \sum_{i \in \mathcal{X}} \pi_n(i) \Pr[X_{n+1} = j \mid X_n = i] = \boxed{\sum_{i \in \mathcal{X}} \pi_n(i) P(i,j)}$$

$$[\,-\ \pi_{n+1}\ -\,] = [\,-\ \pi_n\ -\,] \begin{pmatrix} & & \\ & P & \\ & & \end{pmatrix}$$



So: $\boxed{\pi_{n+1} = \pi_n P}$

$$\pi_{n+1} = \pi_n P$$

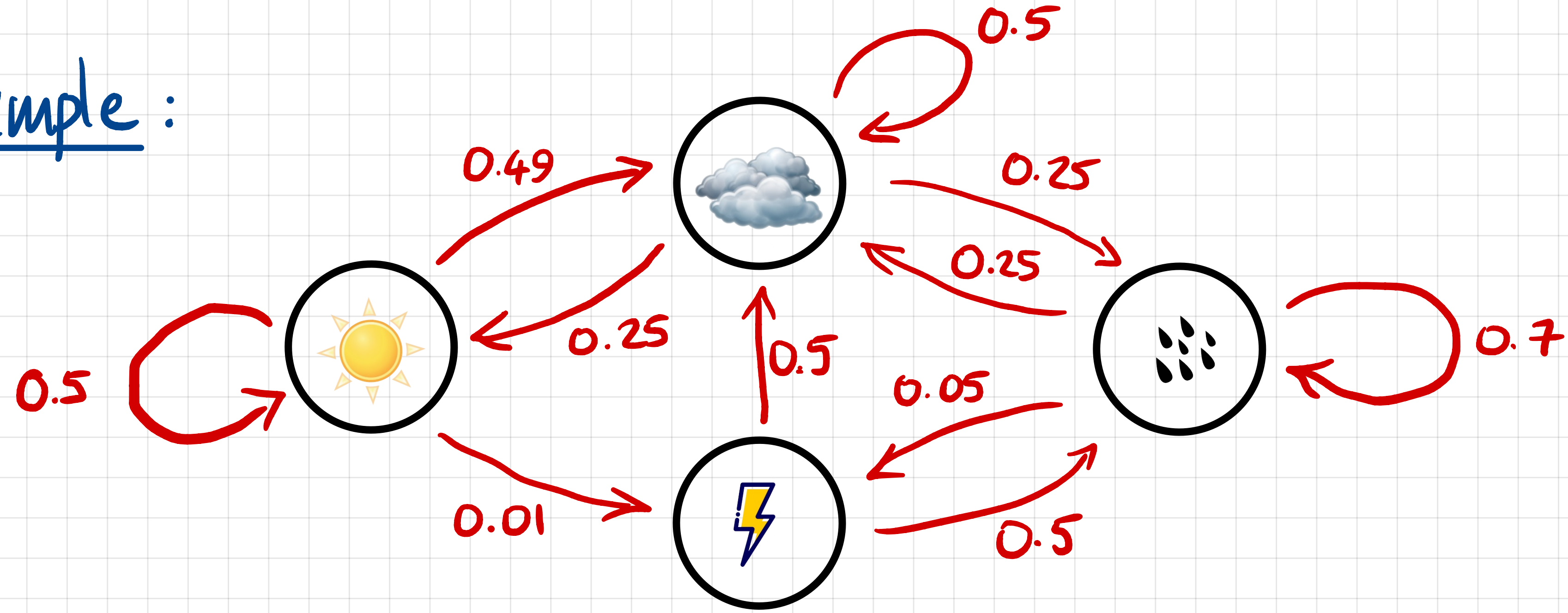
$\Rightarrow$ By induction on n:

$\boxed{\pi_n = \pi_0 P^n}$ where π_0 is the <u>initial distribution</u> (distribution of X_0)

<u>Proof</u>: Base case: $\pi_0 = \pi_0 P^0 = \pi_0$ ✓

Inductive step: $\pi_{n+1} = \pi_n P \underset{\substack{\uparrow \\ \text{induction hypothesis}}}{=} (\pi_0 P^n) P = \pi_0 P^{n+1}$ ✓

Example :



Transition matrix $P =$

$$
\begin{array}{c}
\\ \text{☁️} \\ \text{☀️} \\ \text{🌧️} \\ \text{⚡}
\end{array}
\begin{array}{c}
\begin{array}{cccc} \text{☁️} & \text{☀️} & \text{🌧️} & \text{⚡} \end{array} \\
\begin{pmatrix}
0.5 & 0.25 & 0.25 & 0 \\
0.49 & 0.5 & 0 & 0.01 \\
0.25 & 0 & 0.7 & 0.05 \\
0.5 & 0 & 0.5 & 0
\end{pmatrix}
\end{array}
$$

$$P = \begin{array}{c} \\ \text{☁️} \\ \text{☀️} \\ \text{🌧️} \\ \text{⚡} \end{array} \begin{array}{c} \begin{array}{cccc} \text{☁️} & \text{☀️} & \text{🌧️} & \text{⚡} \end{array} \\ \begin{pmatrix} 0.5 & 0.25 & 0.25 & 0 \\ 0.49 & 0.5 & 0 & 0.01 \\ 0.25 & 0 & 0.7 & 0.05 \\ 0.5 & 0 & 0.5 & 0 \end{pmatrix} \end{array}$$

Take $\pi_0 = \begin{array}{c} [\,0,\ 1,\ 0,\ 0\,] \\ \text{☁️ ☀️ 🌧️ ⚡} \end{array}$

(i.e., start on a <u>sunny</u> day)

$$\overbrace{[\,0,\ 1,\ 0,\ 0\,]}^{\pi_0} \begin{pmatrix} 0.5 & 0.25 & 0.25 & 0 \\ 0.49 & 0.5 & 0 & 0.01 \\ 0.25 & 0 & 0.7 & 0.05 \\ 0.5 & 0 & 0.5 & 0 \end{pmatrix} = \begin{array}{c} \overbrace{[\,0.49,\ 0.5,\ 0,\ 0.01\,]}^{\pi_1} \\ \text{☁️ ☀️ 🌧️ ⚡} \end{array}$$

$$\overbrace{[\,0.49,\ 0.5,\ 0,\ 0.01\,]}^{\pi_1} \begin{pmatrix} 0.5 & 0.25 & 0.25 & 0 \\ 0.49 & 0.5 & 0 & 0.01 \\ 0.25 & 0 & 0.7 & 0.05 \\ 0.5 & 0 & 0.5 & 0 \end{pmatrix} = \begin{array}{c} \overbrace{[\,0.495,\ 0.3725,\ 0.1275,\ 0.005\,]}^{\pi_2} \\ \text{☁️ ☀️ 🌧️ ⚡} \end{array}$$

... and so on!

Invariant Distribution (a.k.a. Stationary Distribution)

Defn: A distribution π over $\mathcal{K}$ is <u>invariant</u> for P if

$$\boxed{\pi P = \pi}$$

I.e., π does not change under the action of P

Note: If π_0 is invariant then

$$\boxed{\pi_n = \pi_0 P^n = \pi_0 \quad \forall n}$$

Defn: A distribution π over $\mathcal{K}$ is <u>invariant</u> for P if

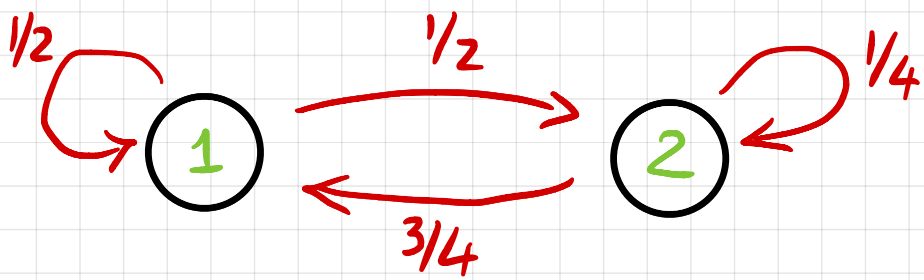
$$\boxed{\pi P = \pi}$$

Finding an invariant distribution: the condition $\pi P = \pi$ corresponds to K linear equations:

$$\pi(j) = \sum_{i \in \mathcal{K}} \pi(i) P(i,j)$$

"balance equations"

<u>Simple example</u>:

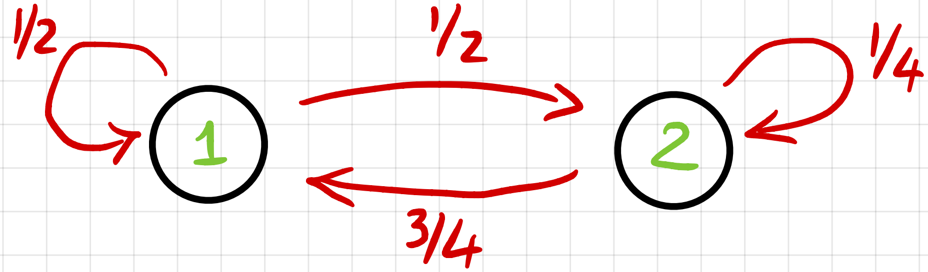


$$P = \begin{pmatrix} 1/2 & 1/2 \\ 3/4 & 1/4 \end{pmatrix}$$

$$\pi(1) = \pi(1) P(1,1) + \pi_2 P(2,1) = \tfrac{1}{2}\pi(1) + \tfrac{3}{4}\pi(2)$$

$$\pi(2) = \pi(1) P(1,2) + \pi_2 P(2,2) = \tfrac{1}{2}\pi(1) + \tfrac{1}{4}\pi(2)$$

Simple example :



$$P = \begin{pmatrix} 1/2 & 1/2 \\ 3/4 & 1/4 \end{pmatrix}$$

$$\pi(1) = \pi(1)\,P(1,1) + \pi_2\,P(2,1) = \frac{1}{2}\pi(1) + \frac{3}{4}\pi(2)$$

$$\pi(2) = \pi(1)\,P(1,2) + \pi_2\,P(2,2) = \frac{1}{2}\pi(1) + \frac{1}{4}\pi(2)$$

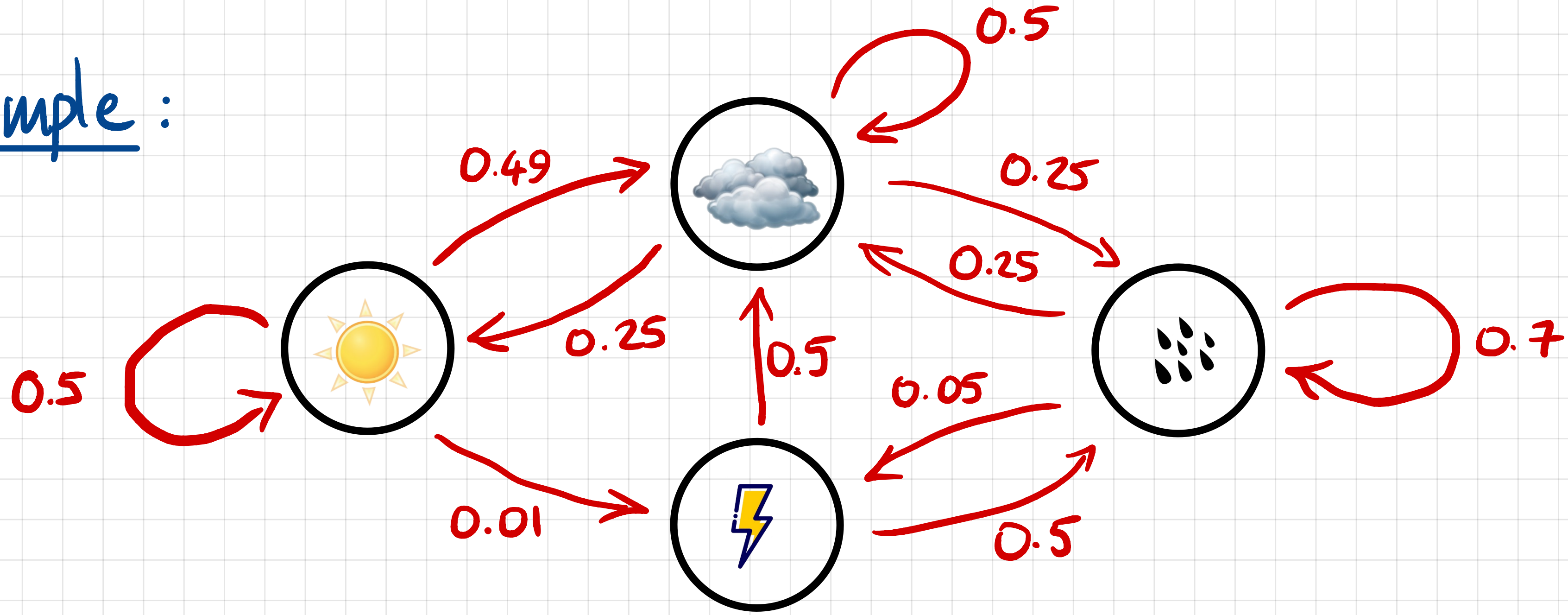
$$\left.\begin{aligned} \frac{1}{2}\pi(1) - \frac{3}{4}\pi(2) &= 0 \\ \frac{1}{2}\pi(1) - \frac{3}{4}\pi(2) &= 0 \end{aligned}\right\}\ \text{redundant}$$

Extra equation : $\pi(1) + \pi(2) = 1$

So: $\pi(1) = \frac{3}{2}\pi(2) \implies \pi = \frac{2}{5}\left(\frac{3}{2}, 1\right) = \boxed{\left(\frac{3}{5}, \frac{2}{5}\right)}$

normalizing factor

Example :



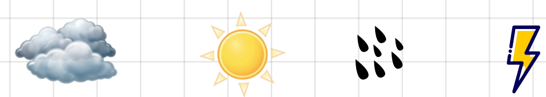
Transition matrix $P = \begin{pmatrix} 0.5 & 0.25 & 0.25 & 0 \\ 0.49 & 0.5 & 0 & 0.01 \\ 0.25 & 0 & 0.7 & 0.05 \\ 0.5 & 0 & 0.5 & 0 \end{pmatrix}$

Balance equations $\pi P = \pi$:

$$[\pi(1), \pi(2), \pi(3), \pi(4)] \begin{pmatrix} 0.5 & 0.25 & 0.25 & 0 \\ 0.49 & 0.5 & 0 & 0.01 \\ 0.25 & 0 & 0.7 & 0.05 \\ 0.5 & 0 & 0.5 & 0 \end{pmatrix} = [\pi(1), \pi(2), \pi(3), \pi(4)]$$

$$\Rightarrow \begin{aligned} 0.5\,\pi(1) + 0.49\,\pi(2) + 0.25\,\pi(3) + 0.5\,\pi(4) &= \pi(1) \\ 0.25\,\pi(1) + 0.5\,\pi(2) &= \pi(2) \\ 0.25\,\pi(1) + 0.7\,\pi(3) + 0.5\,\pi(4) &= \pi(3) \\ 0.01\,\pi(2) + 0.05\,\pi(3) &= \pi(4) \end{aligned}$$

solve $\longrightarrow$ $\boxed{\pi = \frac{1}{1358}\,[550,\ 275,\ 505,\ 28]}$ $\approx [0.405, 0.202, 0.372, 0.021]$

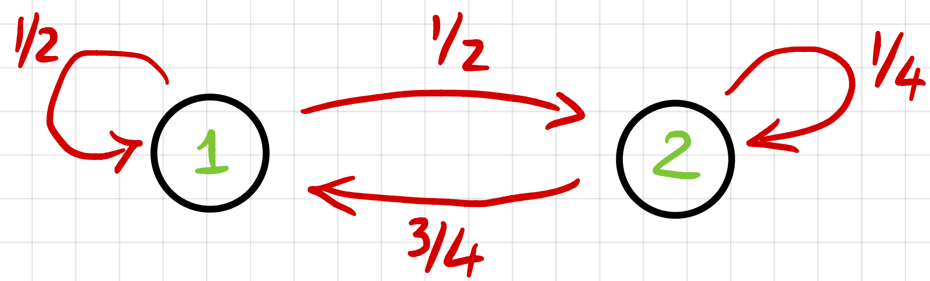


normalizing factor

Convergence to Invariant Distribution

(Informal) Theorem: Under mild conditions, a Markov chain converges to a unique invariant distribution, for any initial distribution π_0

Simple example:



$$P = \begin{pmatrix} 1/2 & 1/2 \\ 3/4 & 1/4 \end{pmatrix}$$

$$P^n = \begin{pmatrix} \frac{3}{5} + \frac{2}{5}\cdot\left(-\frac{1}{4}\right)^n & \frac{2}{5} - \frac{2}{5}\cdot\left(-\frac{1}{4}\right)^n \\ \frac{3}{5} - \frac{3}{5}\cdot\left(-\frac{1}{4}\right)^n & \frac{2}{5} + \frac{3}{5}\left(-\frac{1}{4}\right)^n \end{pmatrix} \xrightarrow[n\to\infty]{} \begin{pmatrix} \frac{3}{5} & \frac{2}{5} \\ \frac{3}{5} & \frac{2}{5} \end{pmatrix}$$

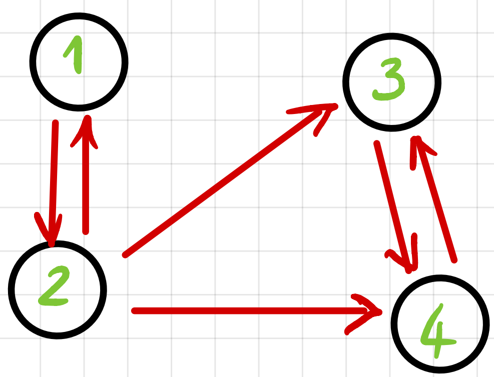
Hence $\pi_n = \pi_0 P^n \xrightarrow[n\to\infty]{} \pi_0 \begin{pmatrix} \frac{3}{5} & \frac{2}{5} \\ \frac{3}{5} & \frac{2}{5} \end{pmatrix} = \boxed{\left(\frac{3}{5}, \frac{2}{5}\right)}$ (any π_0)

Condition 1: Irreducibility

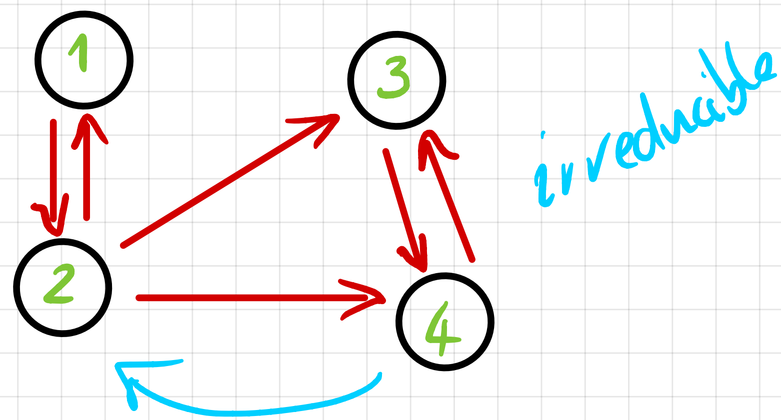
Defn: A Markov chain with trans. matrix P is irreducible if

$$\boxed{\forall i,j \in \mathcal{K} \ \exists n \ \text{s.t.} \ [P^n](i,j) > 0}$$

I.e., $\forall i,j$ $\exists$ a path of transitions leading from i to j



Not irreducible



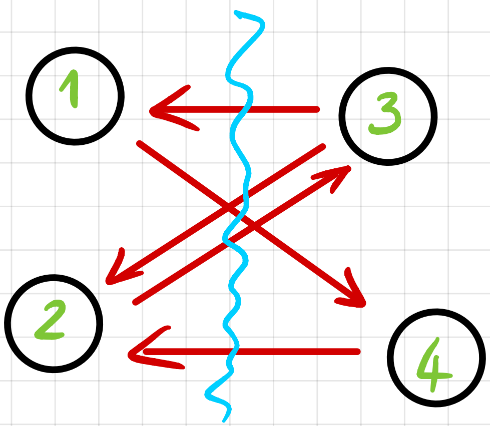
[Equivalent to graph of transitions being strongly connected]

Condition 2: Aperiodicity

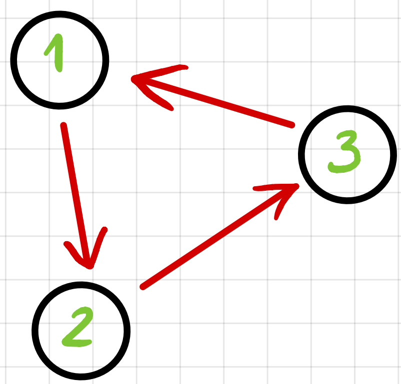
Defn: A Markov chain with trans. matrix P is *aperiodic* if

$$\boxed{\forall i,j \in \mathcal{K} \quad \gcd\{n : [P^n](i,j) > 0\} = 1}$$

I.e., the lengths of paths $i \rightsquigarrow j$ do not have a non-trivial period



Not aperiodic

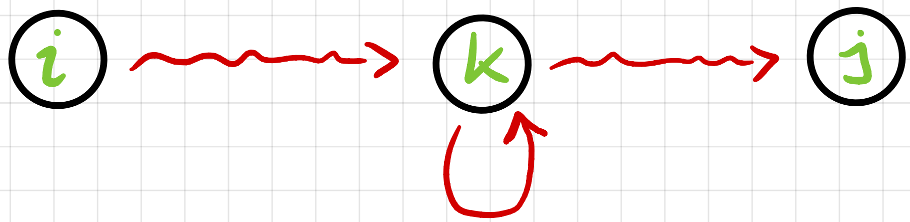


Not aperiodic

Claim: If P is irreducible and $P(k,k)>0$ for some k then P is aperiodic

Proof: Let $i,j \in \mathcal{K}$ be arbitrary

By irreducibility $\exists$ paths $i \rightsquigarrow k$ & $k \rightsquigarrow j$



Sp. total length of path $i \rightsquigarrow k \rightsquigarrow j$ is ℓ

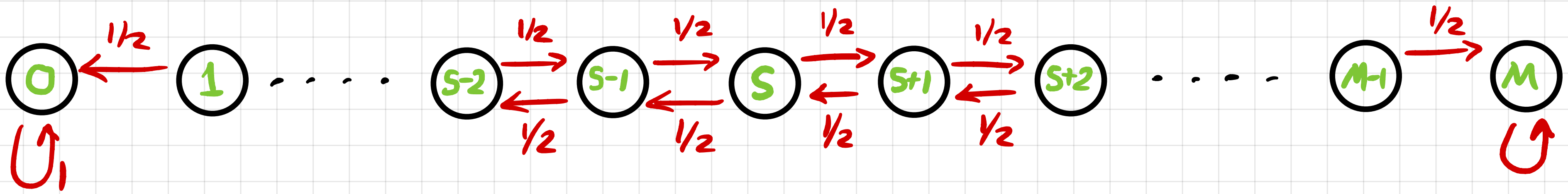
Inserting the loop at k gives paths of lengths ℓ, $\ell+1$

$$\Rightarrow \gcd\{n : [P^n](i,j) > 0\} = 1$$

✓

Note: Actually sufficient to have $\gcd\{n : [P^n](k,k) > 0\} = 1$

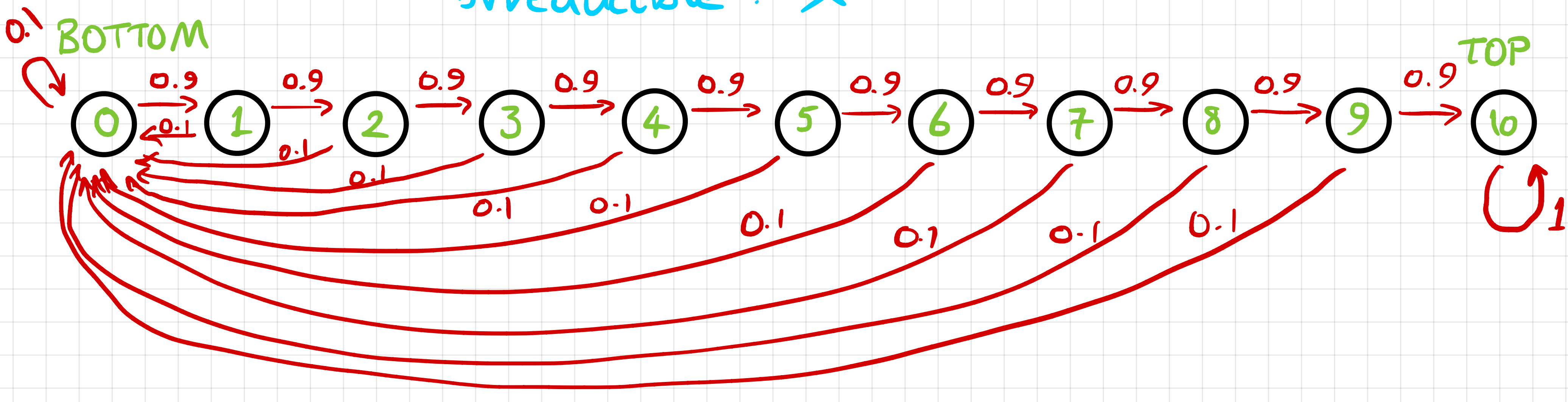
Example: Fair game: win/lose \$1 each with prob. 1/2
Start with \$S, end when reach \$0 or \$M



Irreducible? ✗

Example: Climbing a (very slippery) 10-rung ladder
On each step, slip down to bottom w. prob. 0.1

Irreducible? ✗

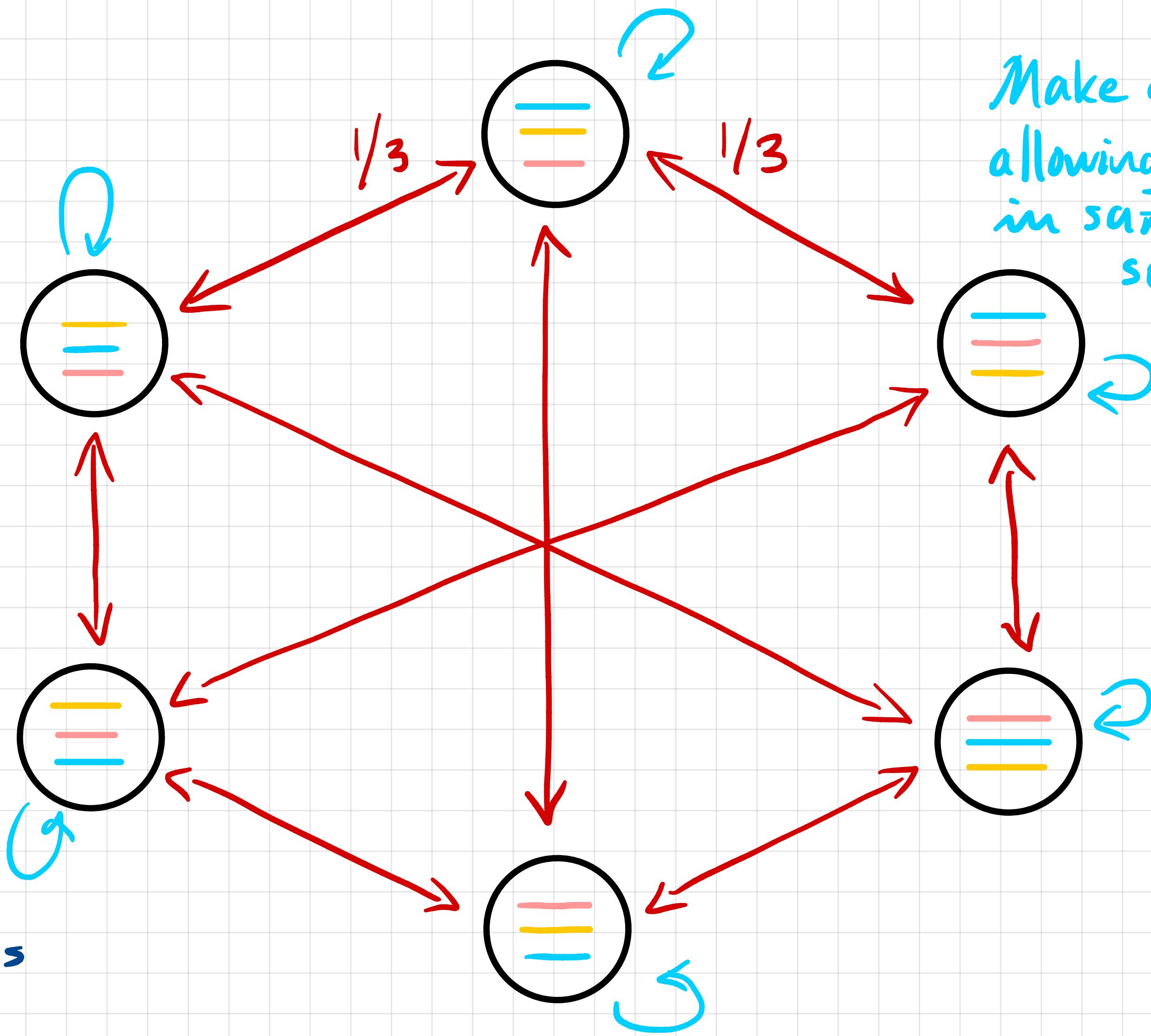


Example: Shuffling cards (slowly!)

States: all $n!$ permutations of the deck (n cards)

Transitions: pick 2 random cards & switch them

$\boxed{n=3}$



Make aperiodic by allowing chain to stay in same state with some prob (same for all states)

Irreducible? ✓

Aperiodic? ✗

$\longleftrightarrow$ denotes 2 edges $\rightleftarrows$

<u>Note</u>: Irreducibility & aperiodicity depend only on the <u>non-zero pattern</u> of P (i.e., the transitions with non-zero probability) — not on the actual values of the transition probabilities

Fundamental Theorem of Markov Chains

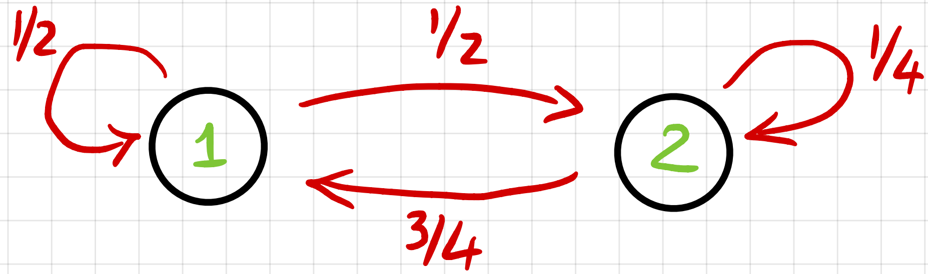
If P is irreducible & aperiodic, then it has a unique invariant distribution π with $\pi(i) > 0 \ \forall i$.

Also, the distribution after n steps converges to π as $n \to \infty$, for any initial distribution π_0

I.e., $\boxed{\forall i \quad Pr[X_n = i] \longrightarrow \pi(i) \quad \text{as } n \to \infty}$

<u>Proof</u>: Out of scope

Simple example :

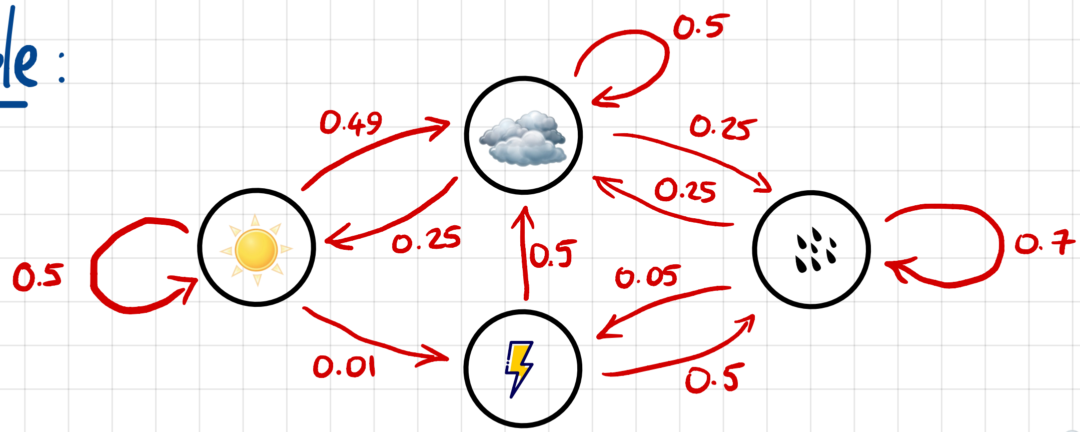


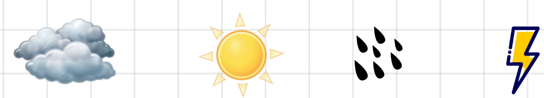
$$P = \begin{pmatrix} 1/2 & 1/2 \\ 3/4 & 1/4 \end{pmatrix}$$

$$\pi = \left[\tfrac{3}{5}, \tfrac{2}{5}\right]$$

$$\Pr[X_n = 1] \xrightarrow[n\to\infty]{} \frac{3}{5} \quad \text{for any } X_0$$

Example :





$$\pi = \frac{1}{1358}[550, \ 275, \ 505, \ 28] \approx [0.405, 0.202, 0.372, 0.021]$$

$$\Pr[X_n = \quad$$  $$] \xrightarrow[n\to\infty]{} \frac{550}{1358} \approx 0.405 \quad \text{for any } X_0$$