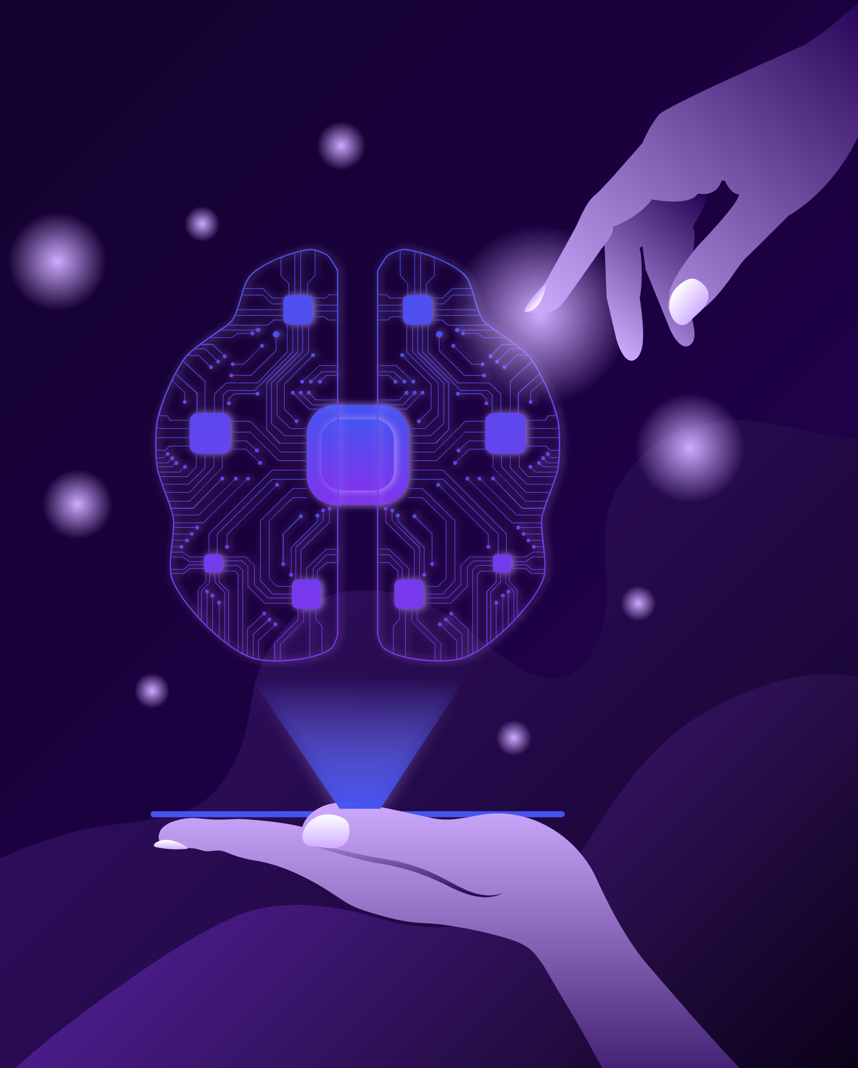
ARC_101

Markov Chains

Start from the basics



Ch_01

INTRODUCTION

Table of contents_

01

Definitions

MC History

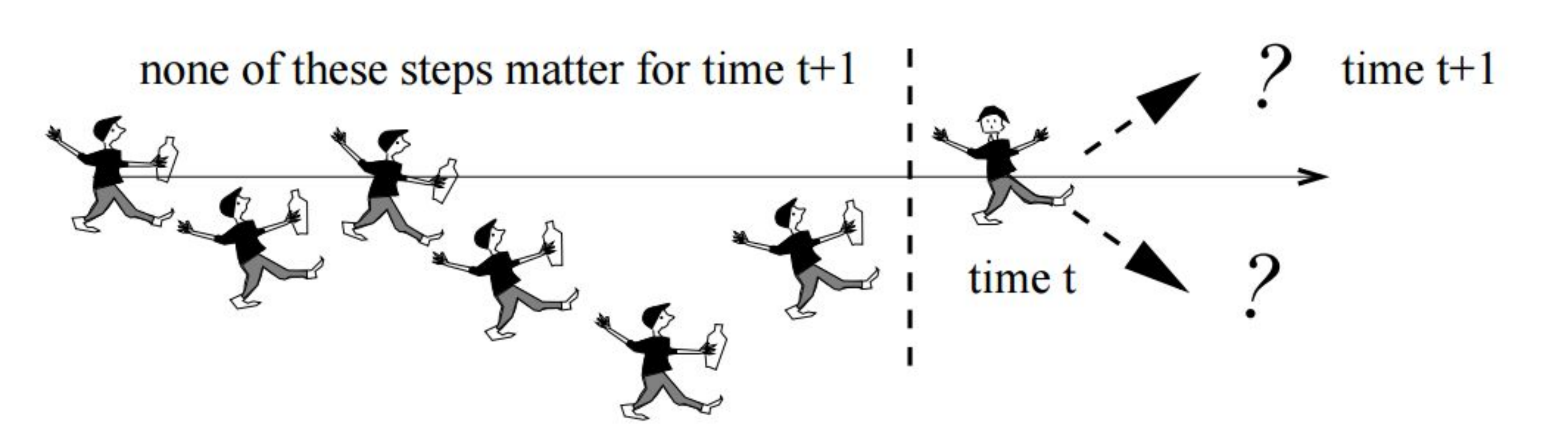
⇒ Markov chains were introduced in 1906 by a russian mathematian Andrei Andreyevich Markov (1856–1922)

⇒ His work focused on probability theory and stochastic processes.

⇒ The famous PageRank algorithm (1998) by Google relies on Markov chains.

MC @ core

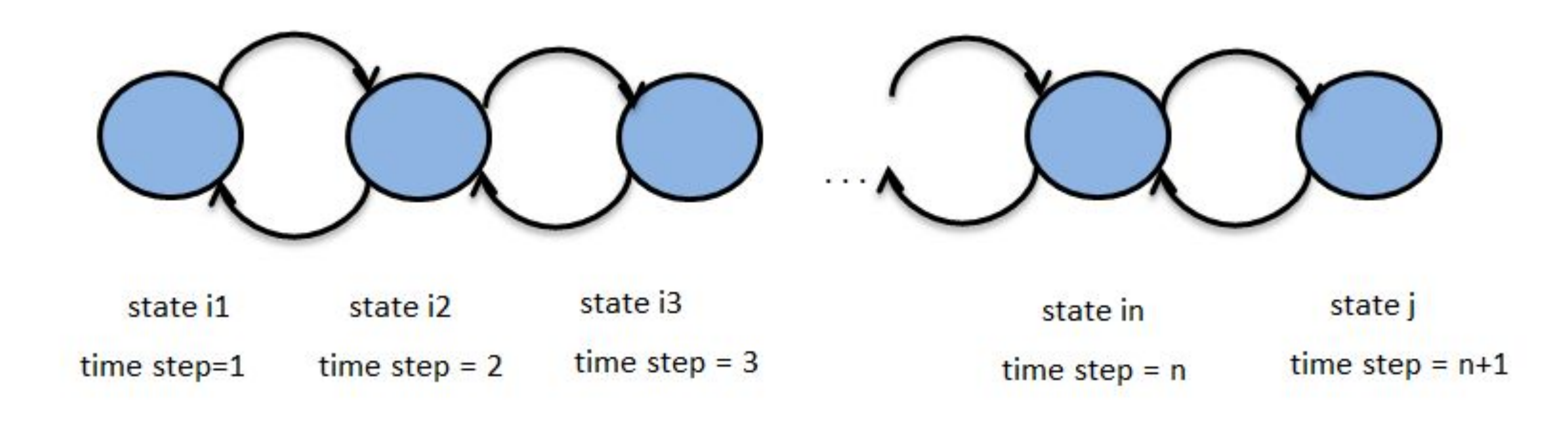
⇒ Mathematical models that describe a sequence of events where the outcome of each event depends only on the previous state



MC Definitions ⇒ state

⇒ The **state** of a Markov chain at time t is the value of X_t

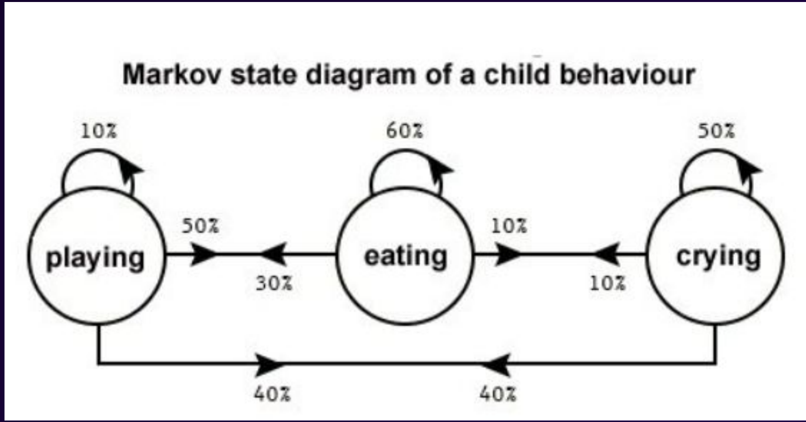
↪ if $X_t = 6$, we say the process is in state 6 at time t.



MC Definitions ⇒ state space

⇒ The state space of a Markov chain refers to the set of all possible states that the system can occupy.

⇒ Each state represents a specific condition or configuration of the system.



MC Definitions ⇒ trajectory

⇒ A trajectory of a Markov chain is a particular set of values for X_0, X_1, X_2 ...

↪ if $X_0 = 1$, $X_1 = 5$, and $X_2 = 6$, then the trajectory up to time $t = 2$ is 1, 5, 6.

‘TRAJECTORY’ IS JUST A WORD MEANING ‘PATH’

02

Markov Property

Markov Property ⛉ Memorylessness property

⇒ It means that Xt+1 depends upon Xt, but it does not depend upon $X_{t-1},\ldots,X_1, X_0$

$$\underset{\substack{\uparrow\\ \text{distribution}\\ \text{of } X_{t+1}}}{\mathbb{P}(X_{t+1}=s} \;\Big|\; \underset{\substack{\uparrow\\ \text{depends}\\ \text{on } X_t}}{X_t = s_t,} \;\; \underbrace{\cancel{X_{t-1}=s_{t-1}, X_{t-2}=s_{t-2},\ldots,X_1=s_1, X_0=s_0}}_{\substack{\uparrow\\ \text{but whatever happened before time } t\\ \text{doesn't matter.}}})$$

↪ If $\{X_0, X_1, X_2, \ldots\}$ is a sequence of discrete random variables. They are a ***Markov chain*** if they satisfies the **Markov property**

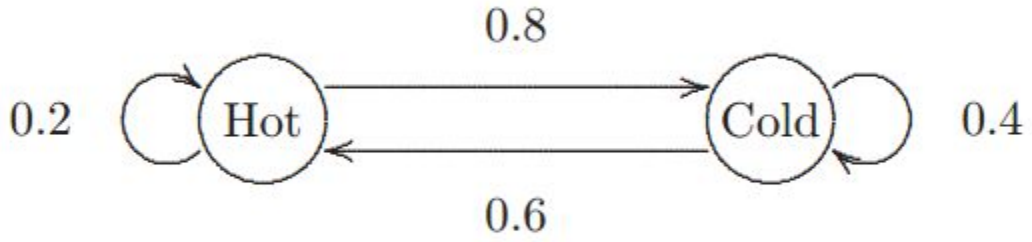
02

Transition Matrix

MC Transition Matrix

⇒ The matrix describing the Markov chain is called the transition matrix.

↪ most important tool for analysing Markov chains.



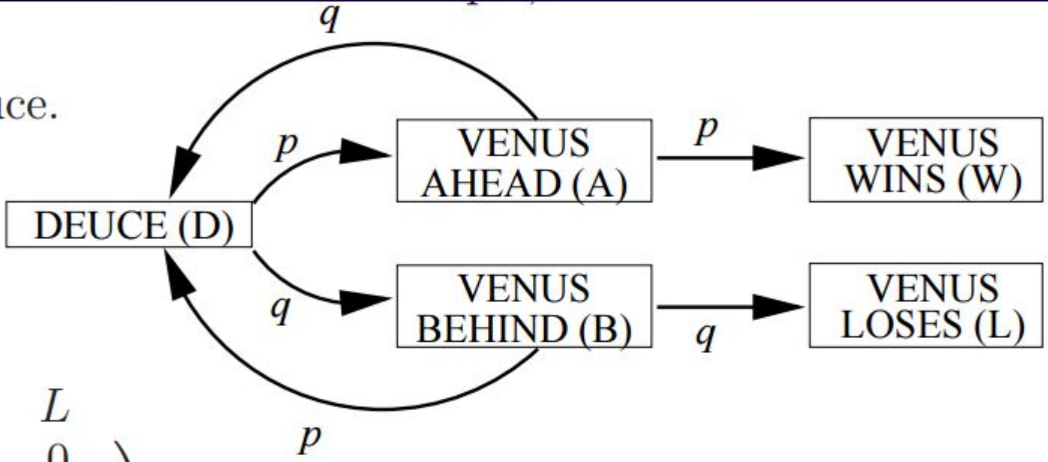
$$\begin{array}{ccc} & & \overbrace{\begin{array}{cc}\text{Hot} & \text{Cold}\end{array}}^{X_{t+1}} \\ X_t \left\{ \begin{array}{c} \\ \\ \end{array} \right. & \begin{array}{c}\text{Hot} \\ \text{Cold}\end{array} & \begin{pmatrix} 0.2 & 0.8 \\ 0.6 & 0.4 \end{pmatrix} \end{array}$$

Transition Matrix ⇒ Notes

1. The transition matrix P must list all possible states in the state space S.
2. P is a square matrix (N × N), because X_{t+1} and X_t both take values in the same state space S (of size N)
3. The **rows** of P *should* each sum to 1
4. The **columns** of P *do not in general* sum to 1

Transition Matrix ⇒ Example

Example: Tennis game at Deuce.



$$
\begin{array}{c} \\ D \\ A \\ B \\ W \\ L \end{array}
\begin{array}{c} \begin{array}{ccccc} D & A & B & W & L \end{array} \\
\begin{pmatrix}
0 & p & q & 0 & 0 \\
q & 0 & 0 & p & 0 \\
p & 0 & 0 & 0 & q \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix} \end{array}
$$

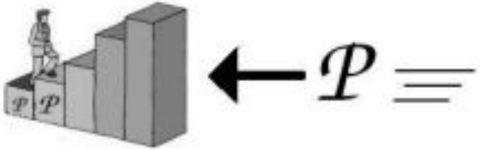
Transition Matrix ⇒ t-step transition

★ Let $\{X_0, X_1, X_2,\ldots\}$ be a Markov chain with $N \times N$ transition matrix P ⇒ the t-step transition probabilities are given by the matrix P^t

$$
\begin{aligned}
X_0 &\sim \pi^T \\
X_1 &\sim \pi^T P \\
X_2 &\sim \pi^T P^2 \\
&\ \vdots \\
X_t &\sim \pi^T P^t.
\end{aligned}
$$

$$\mathbb{P}(X_t = j \mid X_0 = i) = \mathbb{P}(X_{n+t} = j \mid X_n = i) = \left(P^t\right)_{ij} \quad \text{for any } n.$$

take 1 step...



...multiply by $\mathcal{P}$ on the right

n-step Transition Probabilities

Given the chain is in state i at a given time, what is the probability it will be in state j after n transitions? Find it by conditioning on the initial transition(s).

$$P_{ij}^{n} = P\left\{X_{m+n} = j \middle| X_m = i\right\}$$

$$= \sum_{k=0}^{\infty} P\left\{X_{m+n} = j \middle| X_m = i, X_{m+1} = k\right\} P\left\{X_{m+1} = k \middle| X_m = i\right\}$$

$$= \sum_{k=0}^{\infty} P\left\{X_{m+n} = j \middle| X_{m+1} = k\right\} P\left\{X_{m+1} = k \middle| X_m = i\right\}$$

$$= \sum_{k=0}^{\infty} P_{kj}^{n-1} P_{ik}$$

MC State Vector

- A state vector is a mathematical representation of the probabilities associated with different states.
- Initially, the state vector represents the probabilities at time (t = 0).
- As the Markov chain evolves, we update the state vector based on transition probabilities (using a transition matrix).
- State vectors are used in predicting future states, analyzing steady-state behavior, and understanding the long-term behavior of Markov chains.

Steady-State Probabilities (Stationary Distribution)

- The steady-state probabilities represent the long-term behavior of a Markov chain

- When the system reaches a stable state, the probabilities of being in each state no longer change with time

- The steady-state probabilities are denoted by (π), where (π) represents the probability of being in state in the long run.

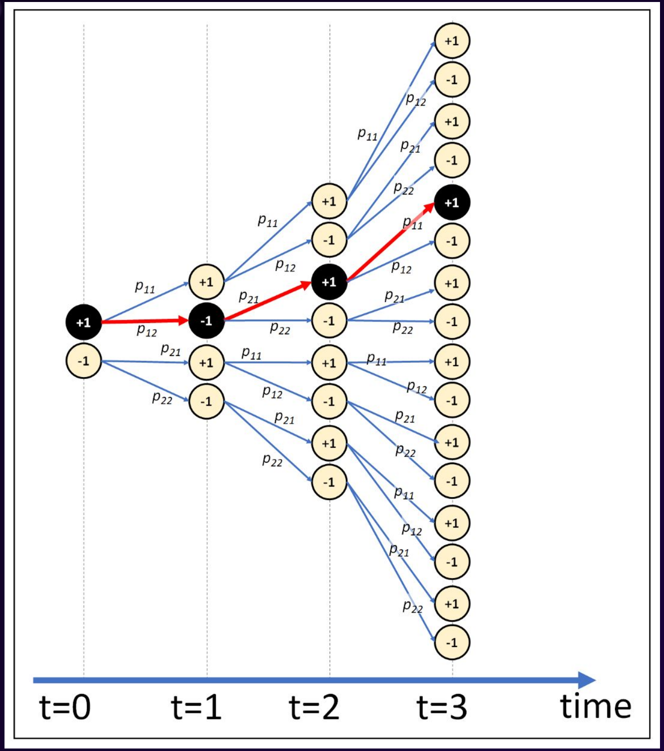
Computing Stationary Distribution

👌There are 3 ways to compute it:

1. The Vanilla way
2. The Repeated Matrix Multiplication way
3. The Left Eigenvector way

➔ To compute the probability of a specific trajectory (a sequence of states)

- Let's say it consists of states (S_1,S_2,S_3...)
- Multiply the initial state probability with the with the consequantive t-step transitions.



👉 Taking one step in the Markov chain corresponds to multiplying by transition p on the right.

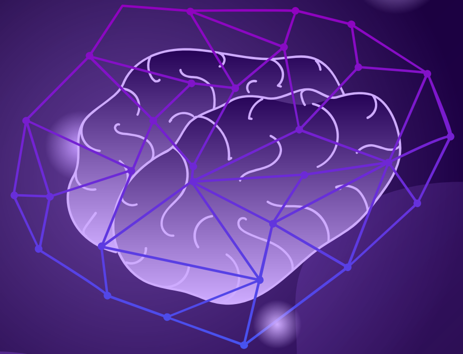
$$\pi_{s_1} \cdot P(s_1 \to s_2) \cdot P(s_2 \to s_3) \cdot \ldots \cdot P(s_{k-1} \to s_k)$$

Thanks!_

Do you have any questions?

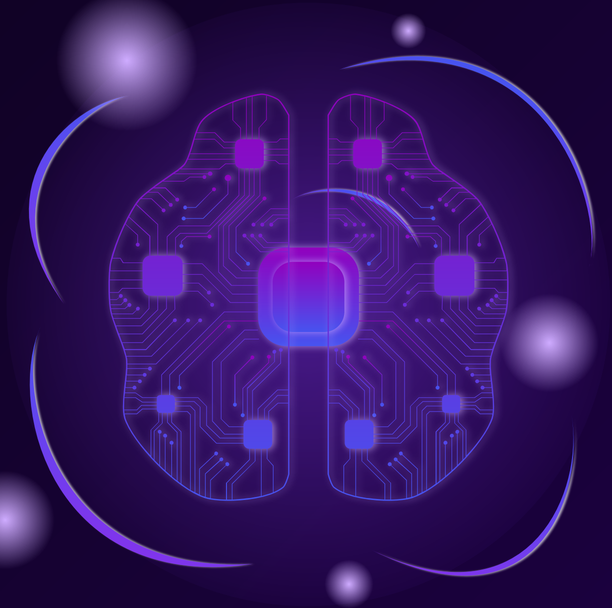
kavehkarimodini@gmail.com

https://www.linkedin.com/in/kaveh-karimadini-a049b3195



CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon** and infographics & images by **Freepik**

Please keep this slide for attribution



Coding Time

Jupyter Notebook:
ARC_101_Markov_Chain_Simulation.ipynb