Introducción a Cadenas de Markov

Rodrigo Maranzana

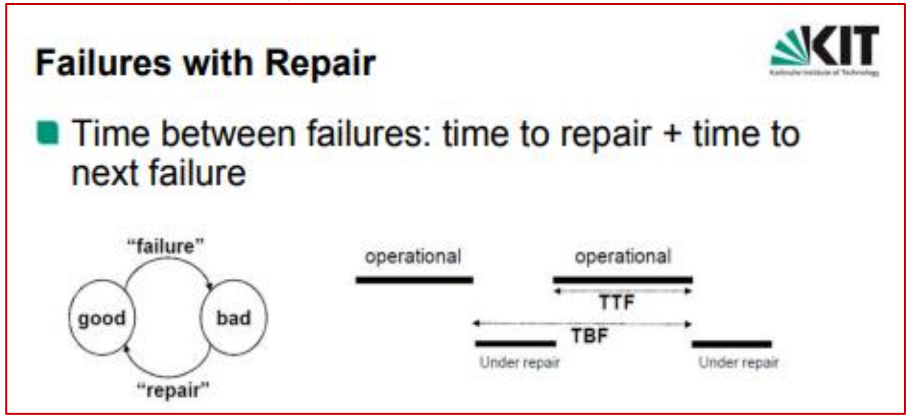
Repaso clasificación procesos estocásticos

Variable: Discreta / Continua

Evolución anual del rating crediticio de una institución

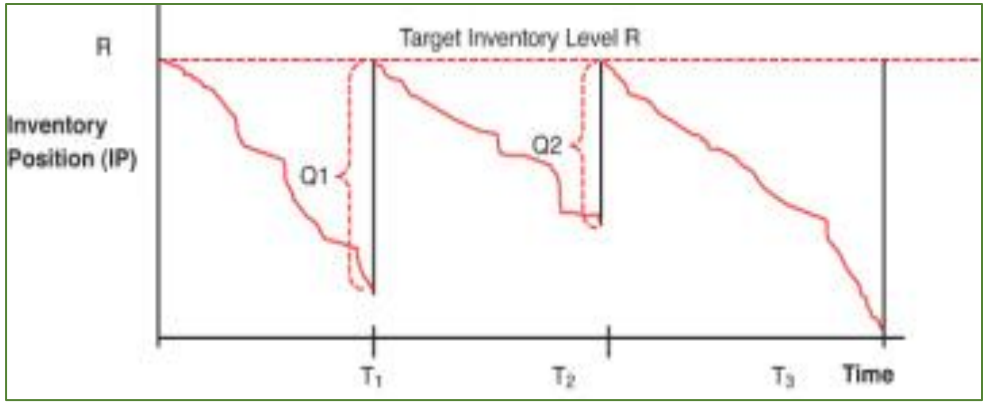
AAA	AAA	AAA	AAA
AA	AA	AA	AA
A	A	A	A
BBB	BBB	BBB	BBB
BB	BB	BB	BB
B	B	B	B
CCC	CCC	CCC	CCC
Default	Default	Default	Default

Estado de falla y reparación de una máquina



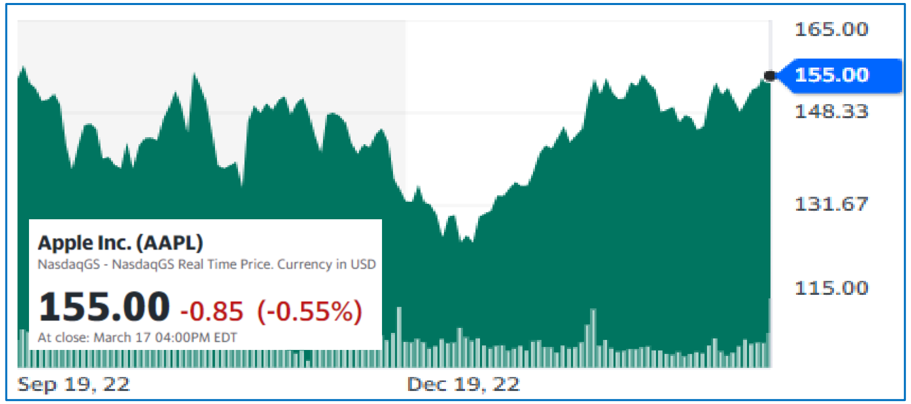
Fuente: https://cdnc.itec.kit.edu/downloads/RC1_WS_2011_lecture5.pdf

Sistema de control de inventarios de tiempo fijo



Fuente: https://www.informit.com/articles/article.aspx?p=2167438&seqNum=7

Dinámica de stocks en la bolsa

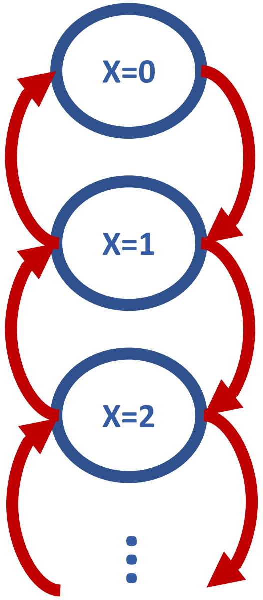


Fuente: https://finance.yahoo.com/quote/AAPL/

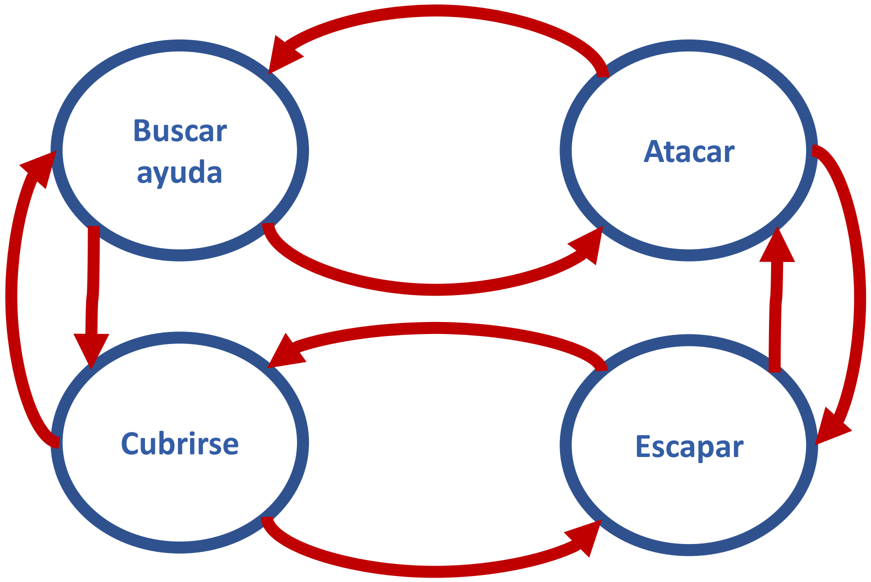
Discreto — *Parámetro* — Continuo

Repaso de estados en procesos estocásticos

Cantidad de gente esperando en una fila:



Comportamiento de una inteligencia artificial en un juego:



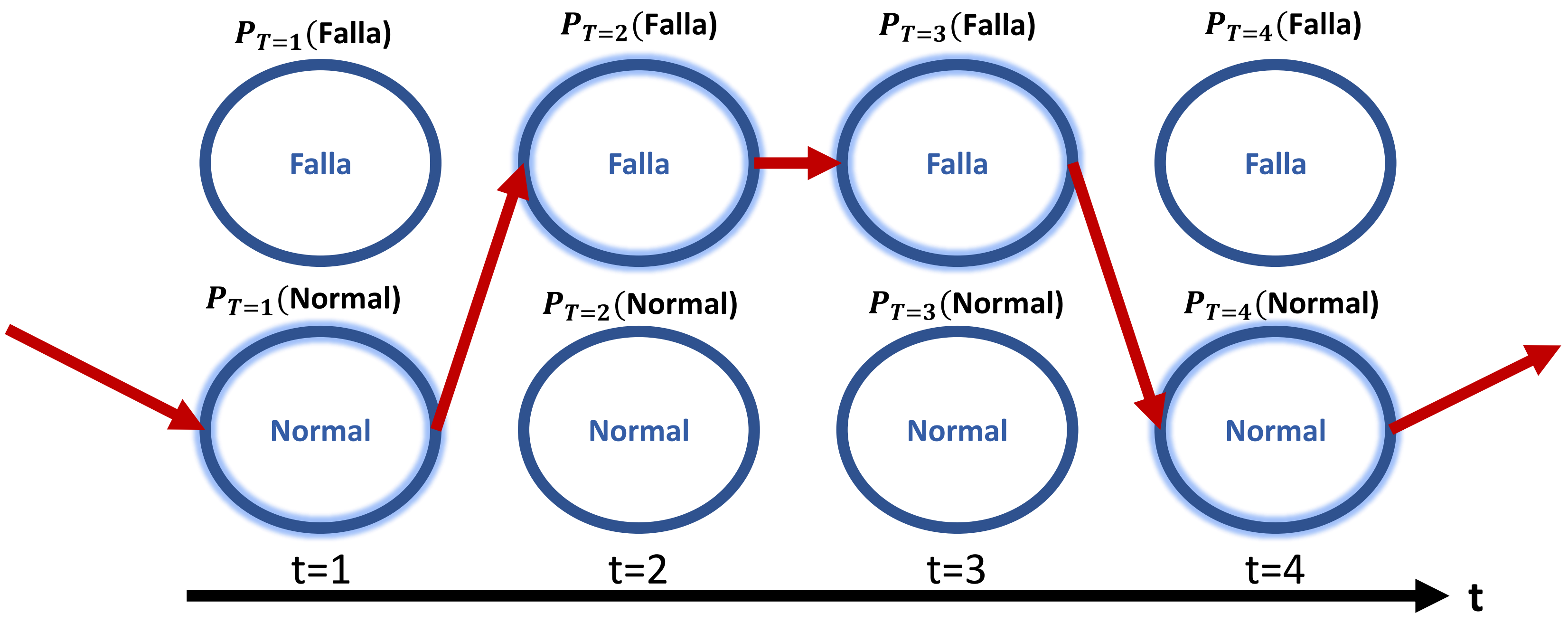
**Caso de autómata finito: set de estados finitos.*

Repaso de transición de estados

(Ej: Modelo de falla con parámetro discreto)

P(Estado|<historia de estados>): Probabilidad de transición, de cambiar a un estado, dada la historia de estados recorrida.

$P_{T=t}$(Estado): probabilidad de estar en un estado en un tiempo t



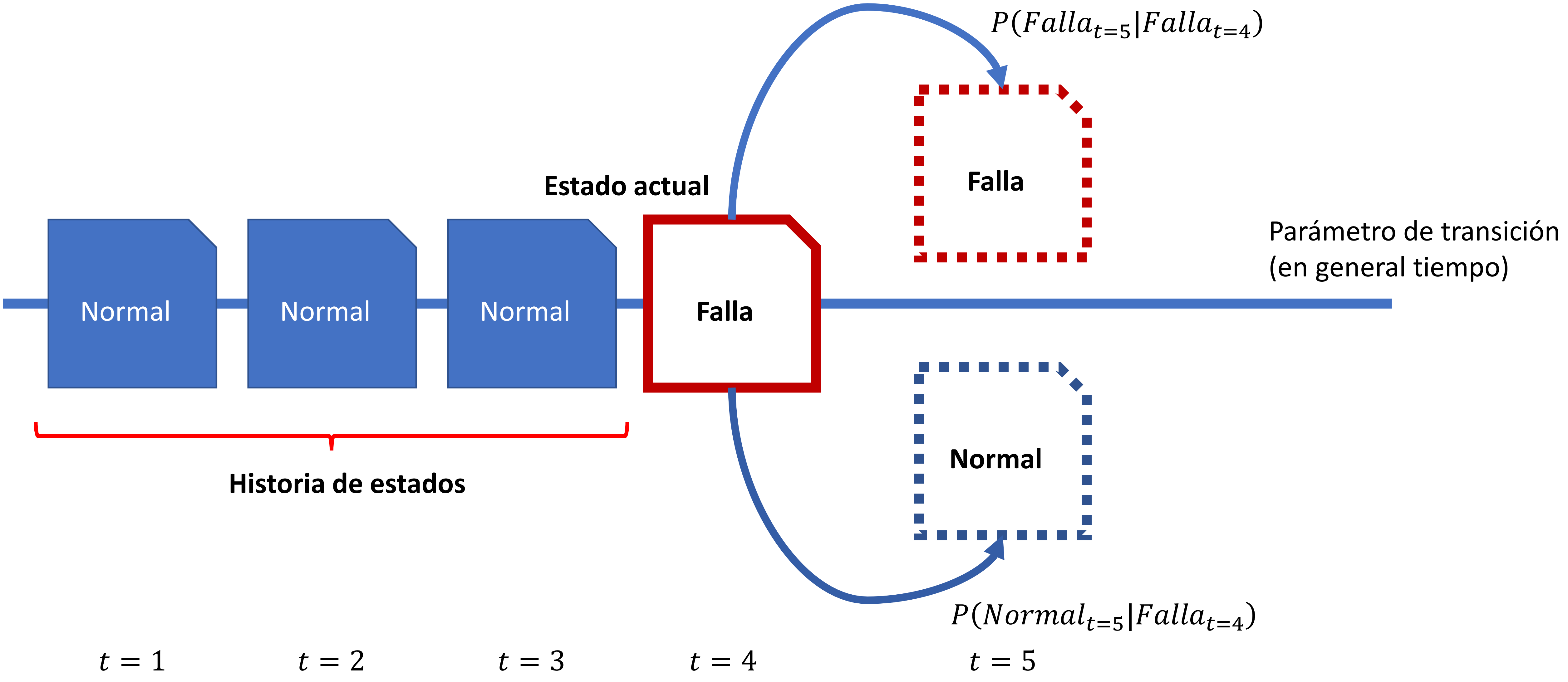
Cadenas de Markov

Es un modelo estocástico que describe una secuencia de estados de un sistema.

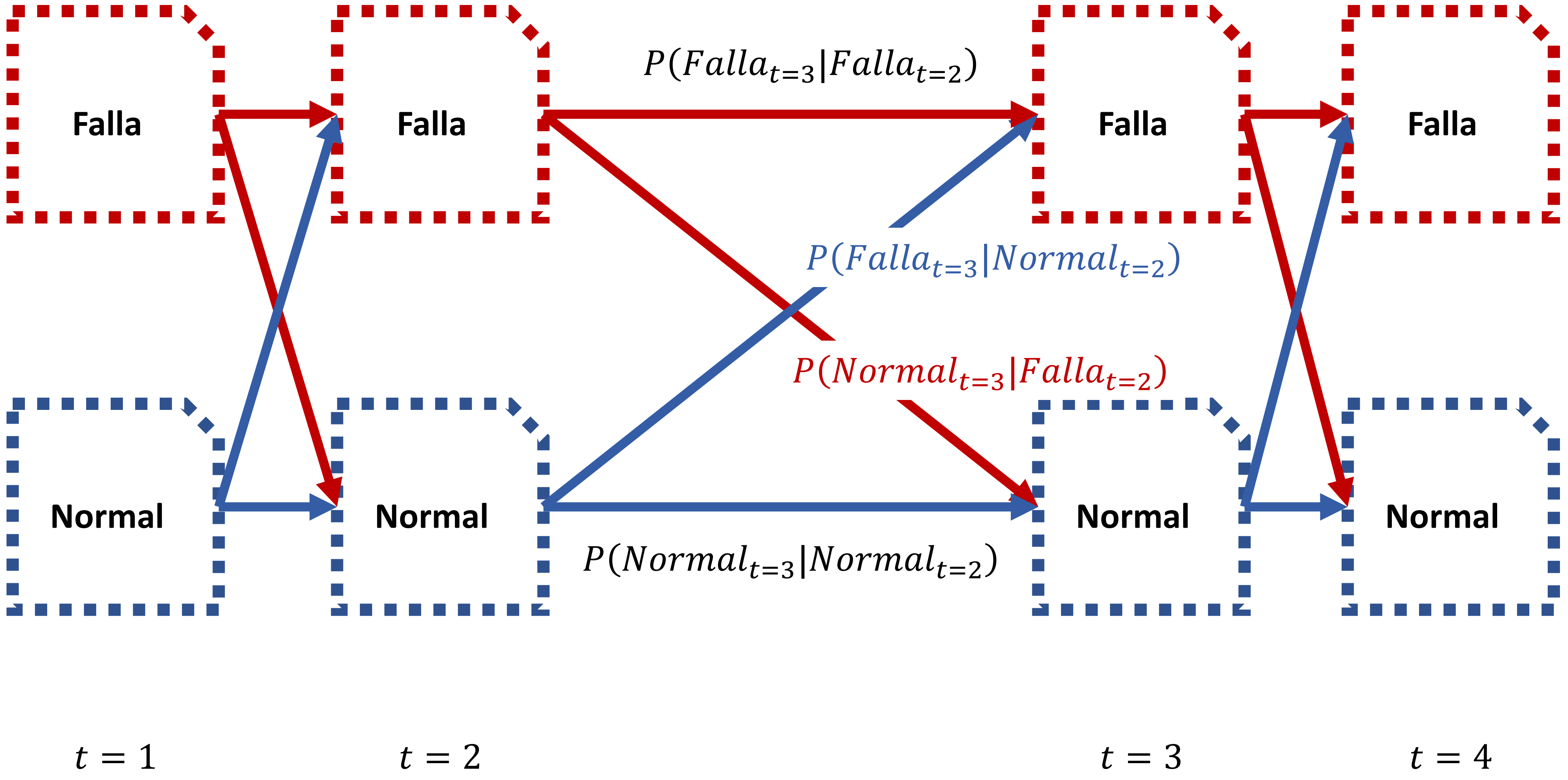
- La probabilidad de transición entre cada estado, solo depende del estado actual.
- Es un proceso estocástico sin memoria.
- Las cadenas de Markov, son un caso particular de procesos Markovianos con parámetro discreto.

$$P(X_{i+1} \mid \underbrace{X_i}_{\text{Dependencia del estado actual}}) = P(X_{i+1} \mid \overbrace{X_i, X_{i-1}, X_{i-2}, \dots, X_{i-n}}^{\text{Dependencia de toda la historia}})$$

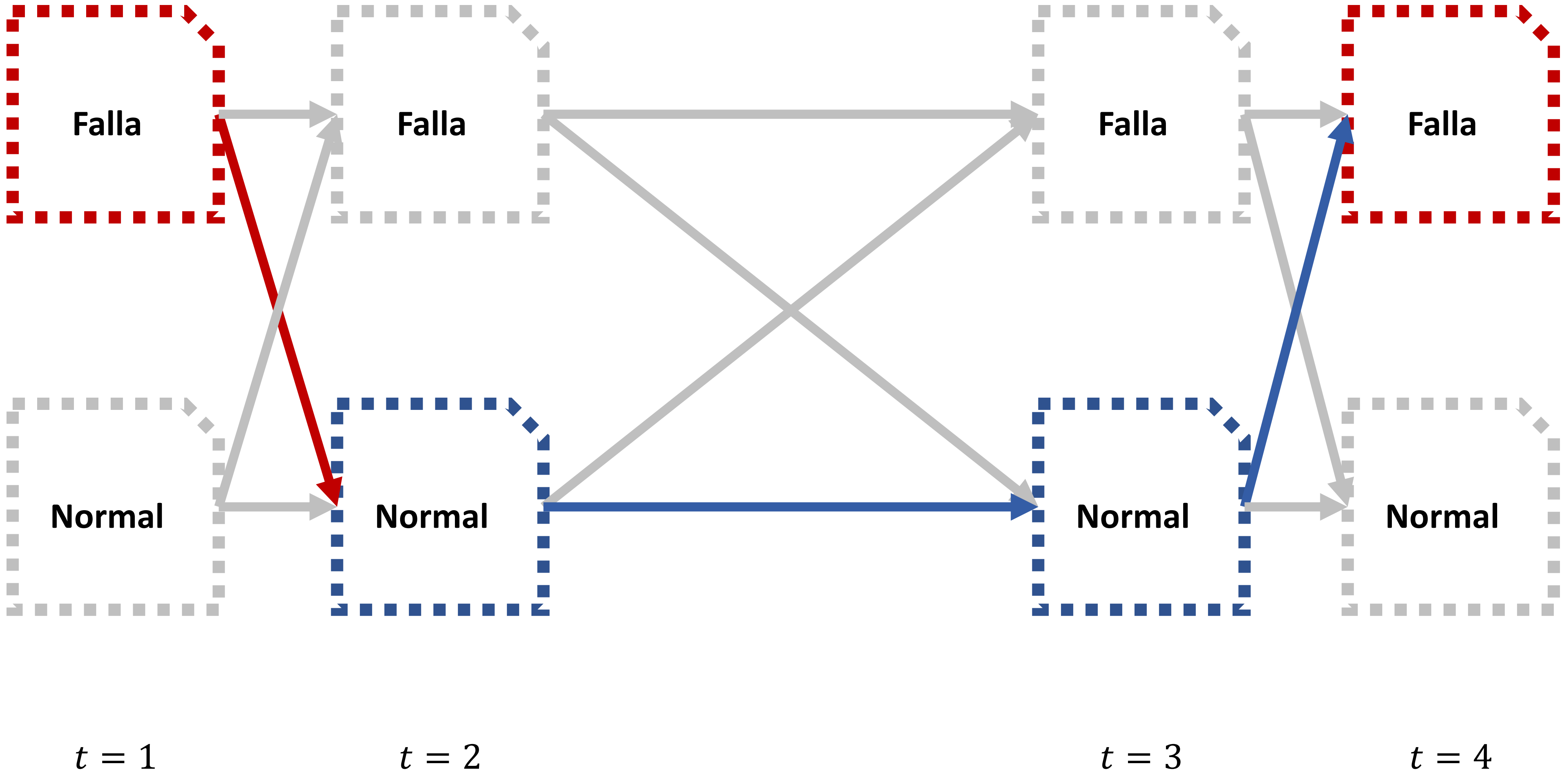
Cadena de eventos realizados y probables



Cadena de eventos posibles



Un camino probable



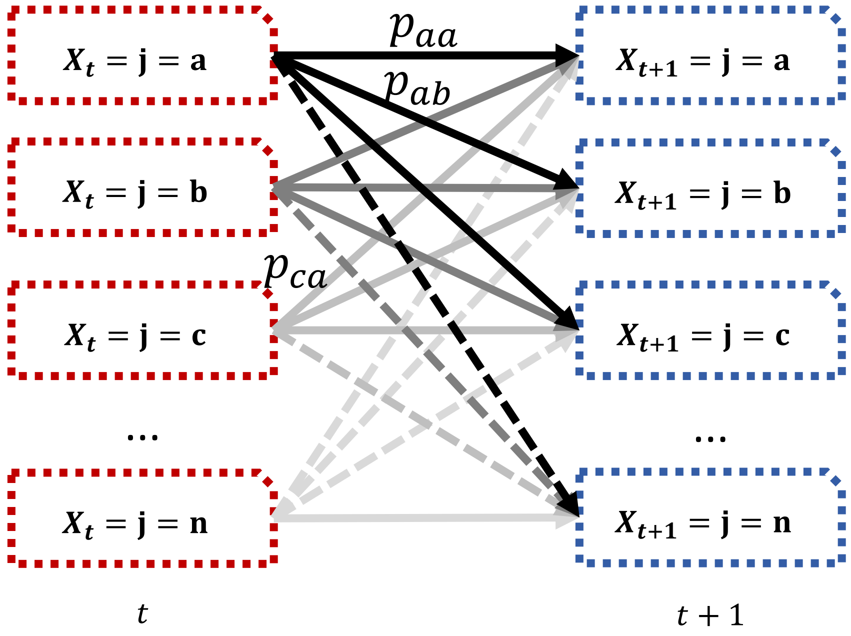
Generalización Markov parámetro discreto

Siendo:

- el tiempo el parámetro del proceso, discreto.
- "t" un instante de tiempo.
- "Δt" el salto discreto de tiempo del parámetro del proceso.
- "i" el estado actual (t) y "j" el estado siguiente (t+1).

La probabilidad condicional de transición de estados se define como:

$$p_{ij}(\Delta t) = P(X_{t+1} = \mathrm{j} \mid X_t = i)$$



Matriz de transición discreta de un paso

- Una forma más práctica de modelizar el paso entre estados es mediante una Matriz de Transición de un paso.

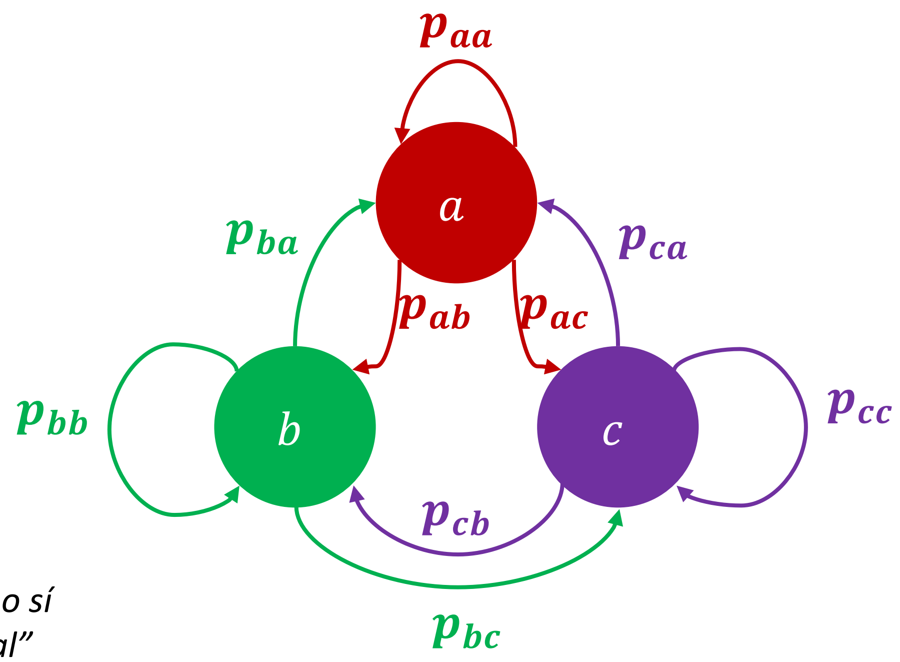
$$T(\Delta t) = \begin{array}{cc} \longleftarrow\; j \;\longrightarrow & \\ \begin{bmatrix} p_{aa} & p_{ab} & p_{ac} & \cdots & p_{an} \\ p_{ba} & p_{bb} & p_{bc} & \cdots & p_{bn} \\ p_{ca} & p_{cb} & p_{cc} & \cdots & p_{cn} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{na} & p_{nb} & p_{nc} & \cdots & p_{nn} \end{bmatrix} & \updownarrow i \end{array}$$

Condiciones:

$$\boxed{\begin{array}{c} 0 \leq p_{ij} \leq 1 \\ \sum_{j} p_{ij} = 1 \;\; \forall i \end{array}}$$

Grafo de transición discreta

$$T(\Delta t) = \begin{bmatrix} p_{aa} & p_{ab} & p_{ac} \\ p_{ba} & p_{bb} & p_{bc} \\ p_{ca} & p_{cb} & p_{cc} \end{bmatrix}$$



Condiciones:

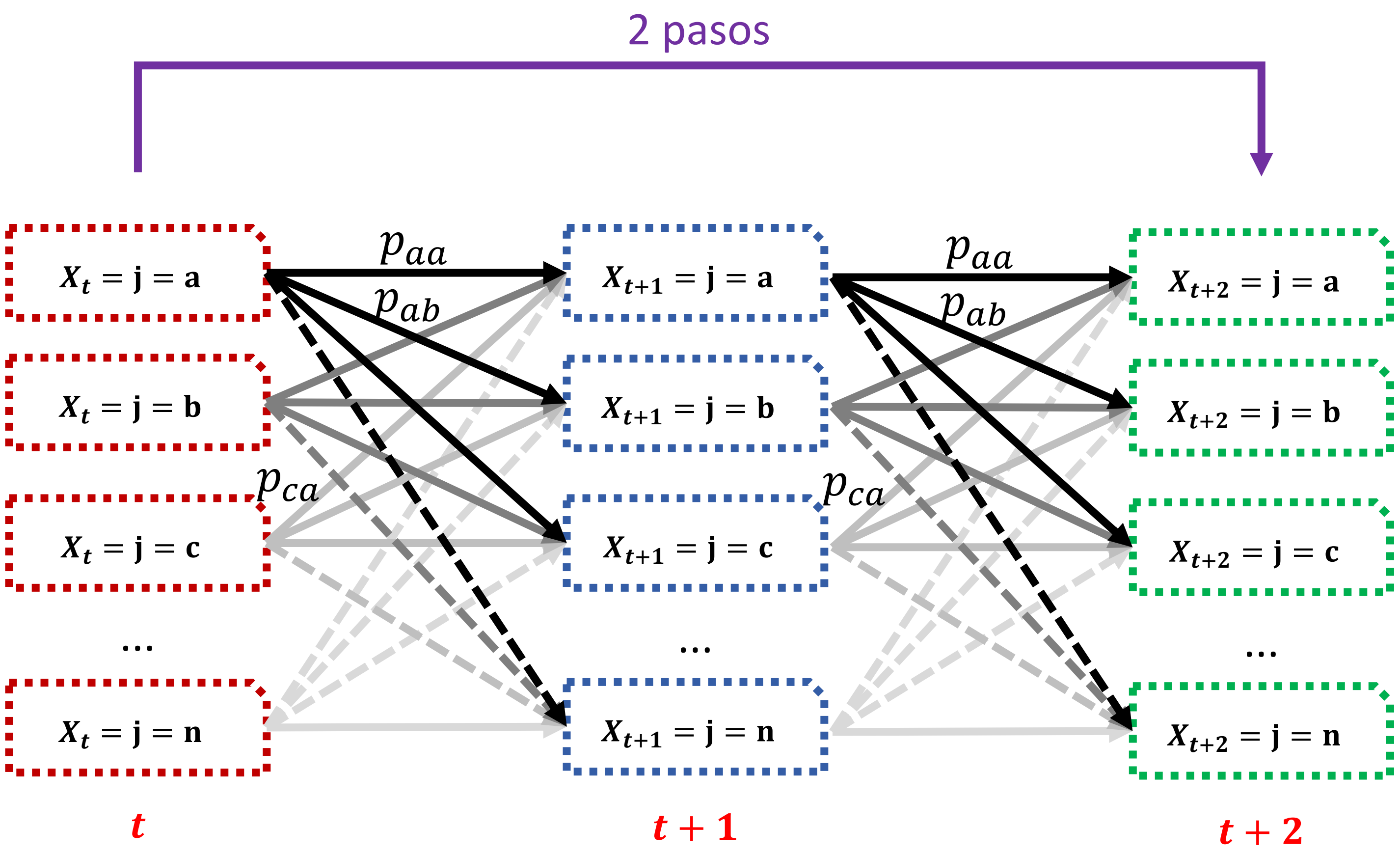
$$\boxed{\begin{gathered} \mathbf{0 \le p_{ij} \le 1} \\ \sum_j \mathbf{p_{ij} = 1} \ \ \forall \boldsymbol{i} \end{gathered}}$$

"Es una probabilidad"

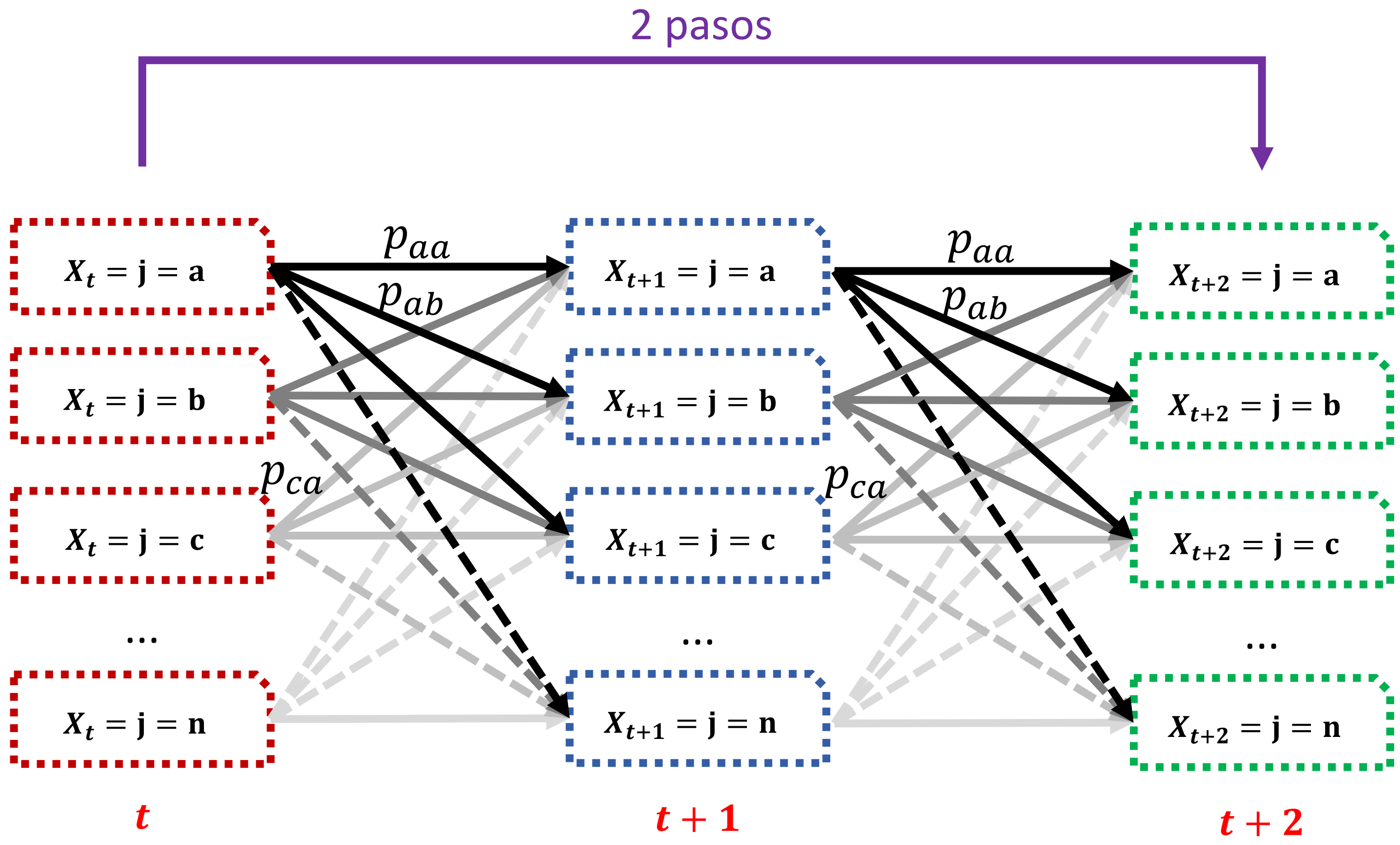
"Dado que estoy en un estado, sí o sí debo ir a otro o seguir en el actual"

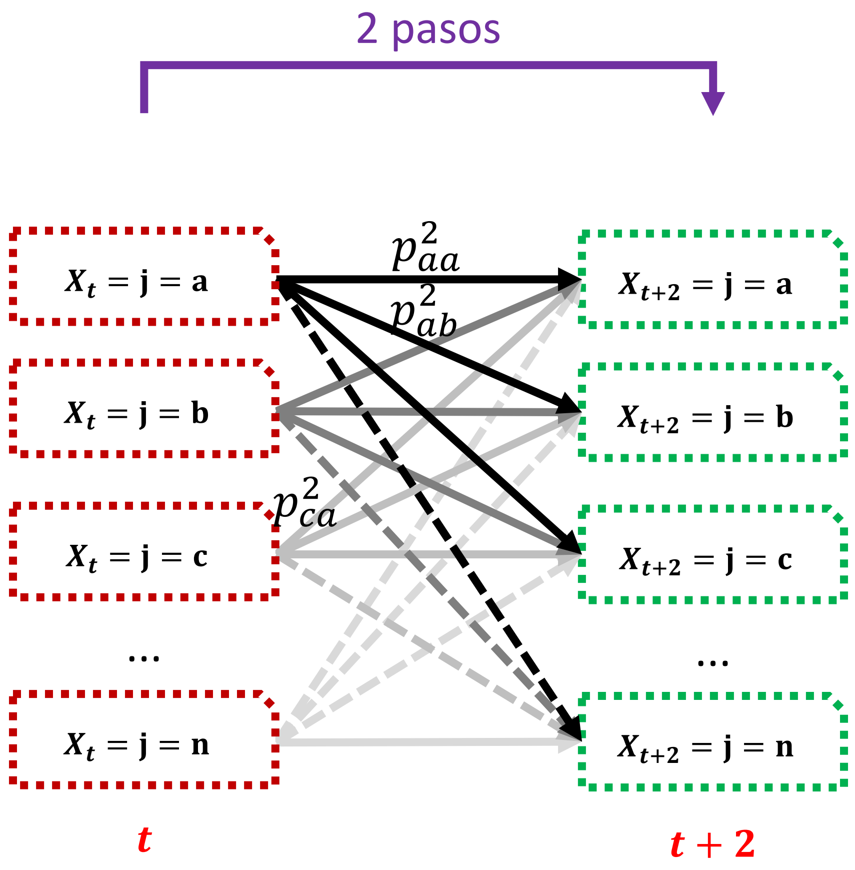
Jugando con pasos en cadenas de markov: *https://setosa.io/ev/markov-chains/*

Transición de “m” pasos, caso m=2



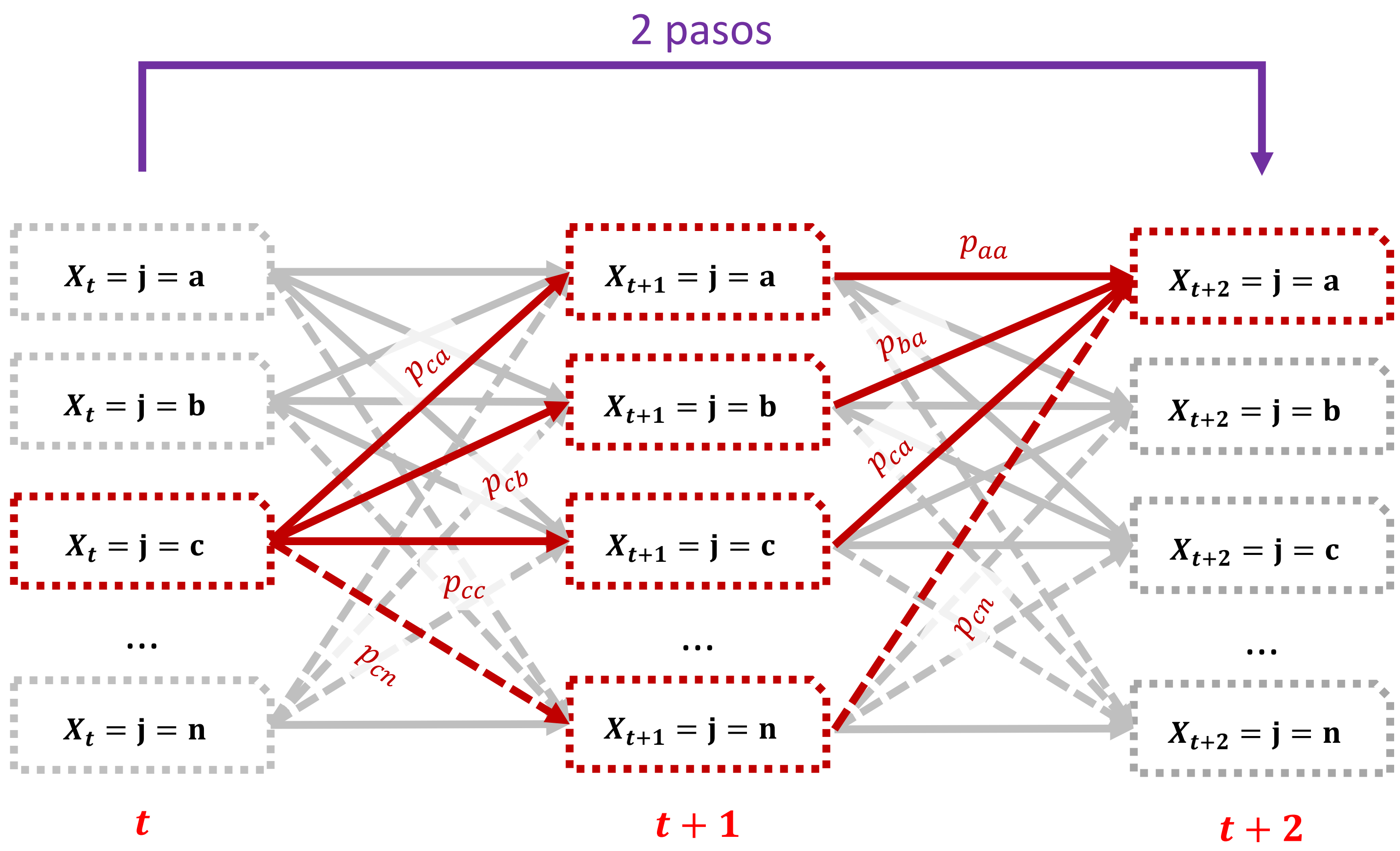
Transición de “m” pasos, caso m=2

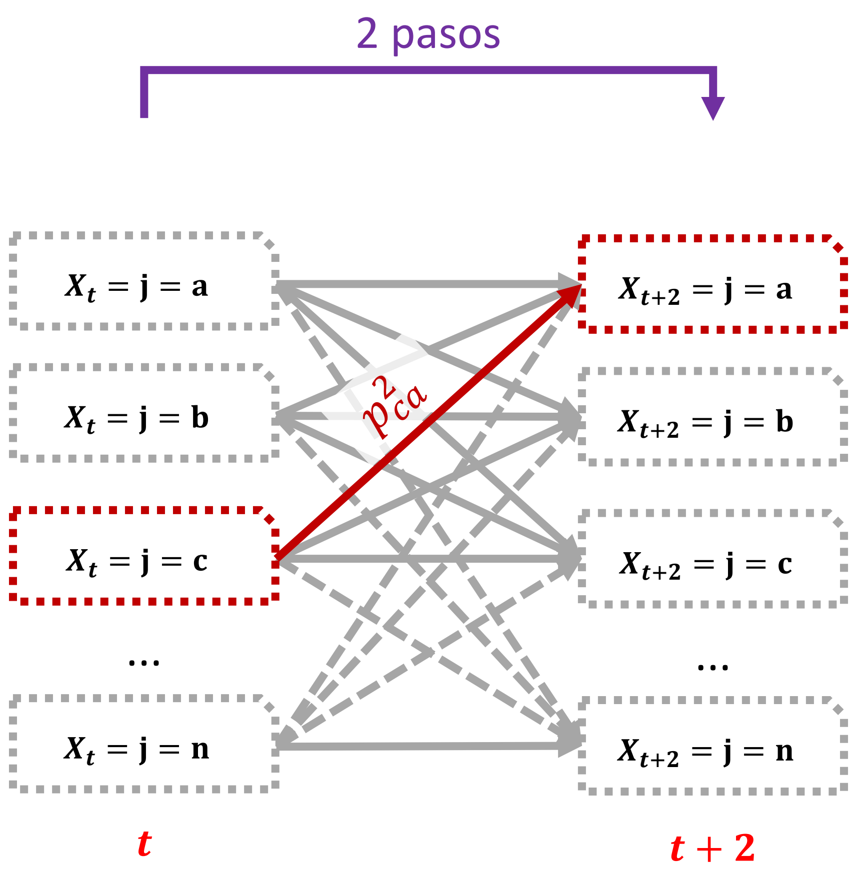




$$p^2_{ij} = P(X_{t+2} = j \mid X_t = i)$$

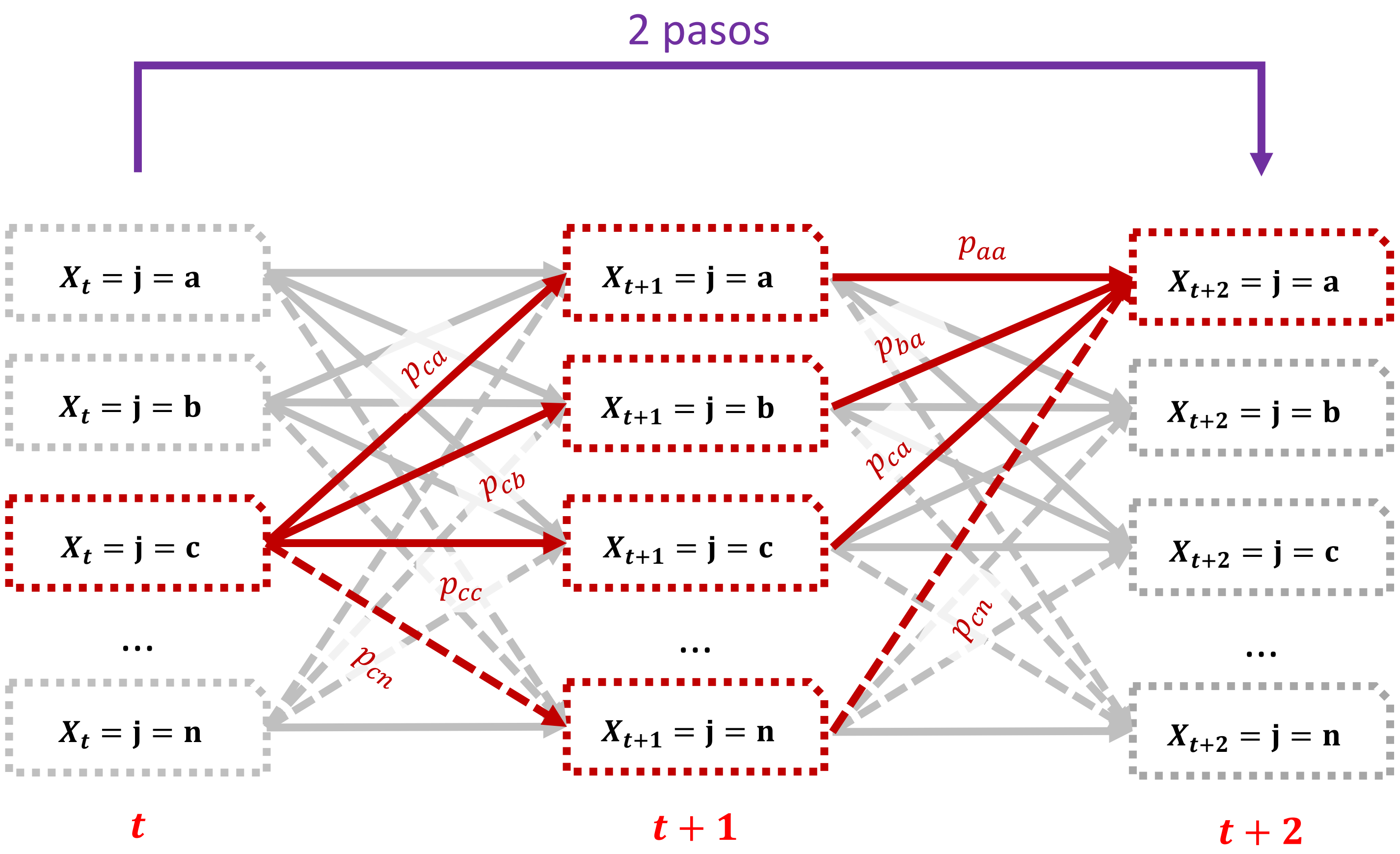
Transición de "m" pasos, caso m=2





$$p_{ca}^2 = P(X_{t+2} = \mathrm{a} \mid X_t = c)$$

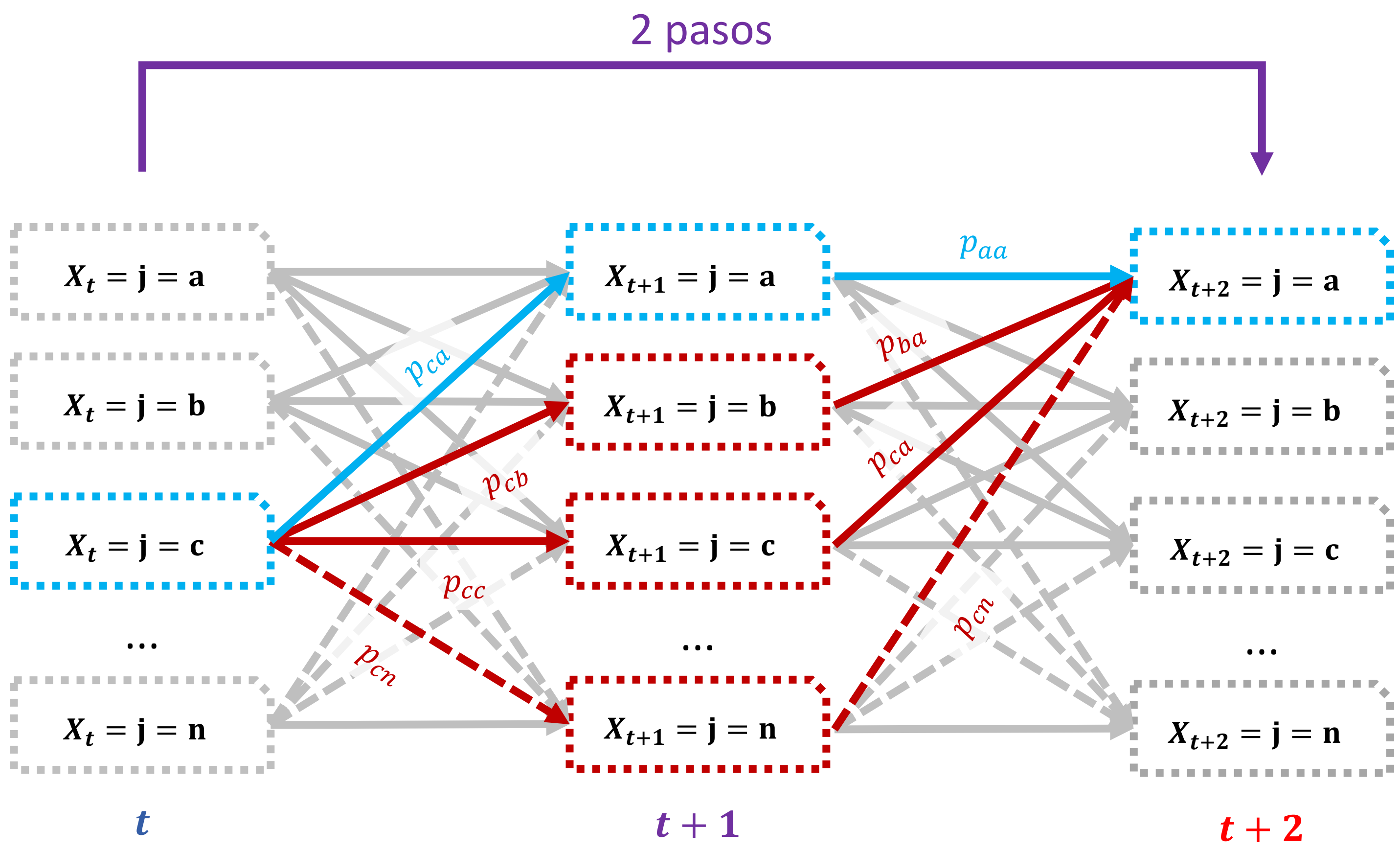
Transición de "m" pasos, caso m=2



"Ir de i a j en m pasos, implica considerar la probabilidad de haber estado en el resto de los estados intermedios"

$$p_{ca}^2 = P(X_{t+2} = \mathrm{a} \mid X_t = c)$$

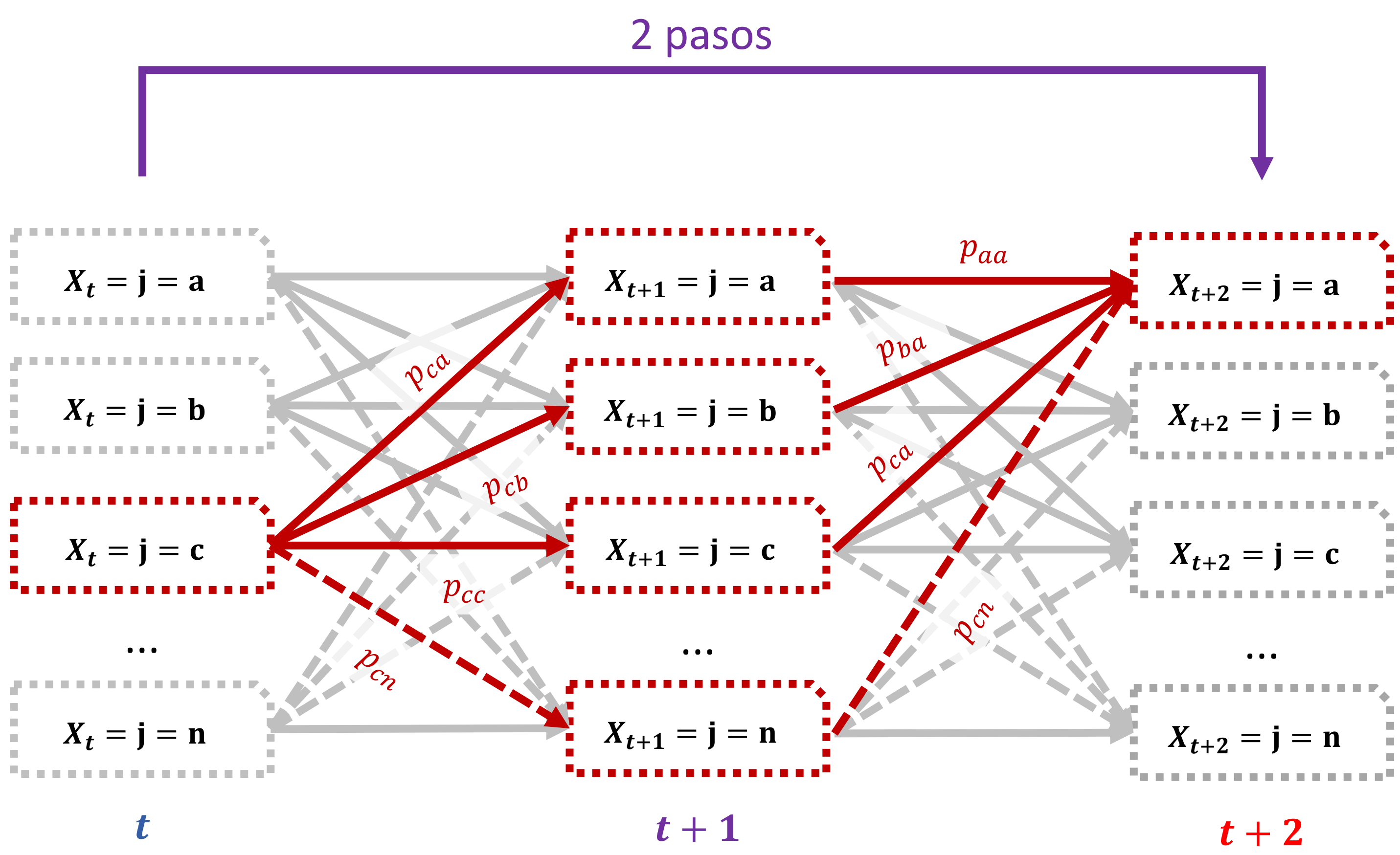
Transición de "m" pasos, caso m=2



Uno de los caminos probables:

$$P(X_{t+1} = a \mid X_t = c) * P(X_{t+2} = a \mid X_{t+1} = a) = p_{ca} * p_{aa}$$

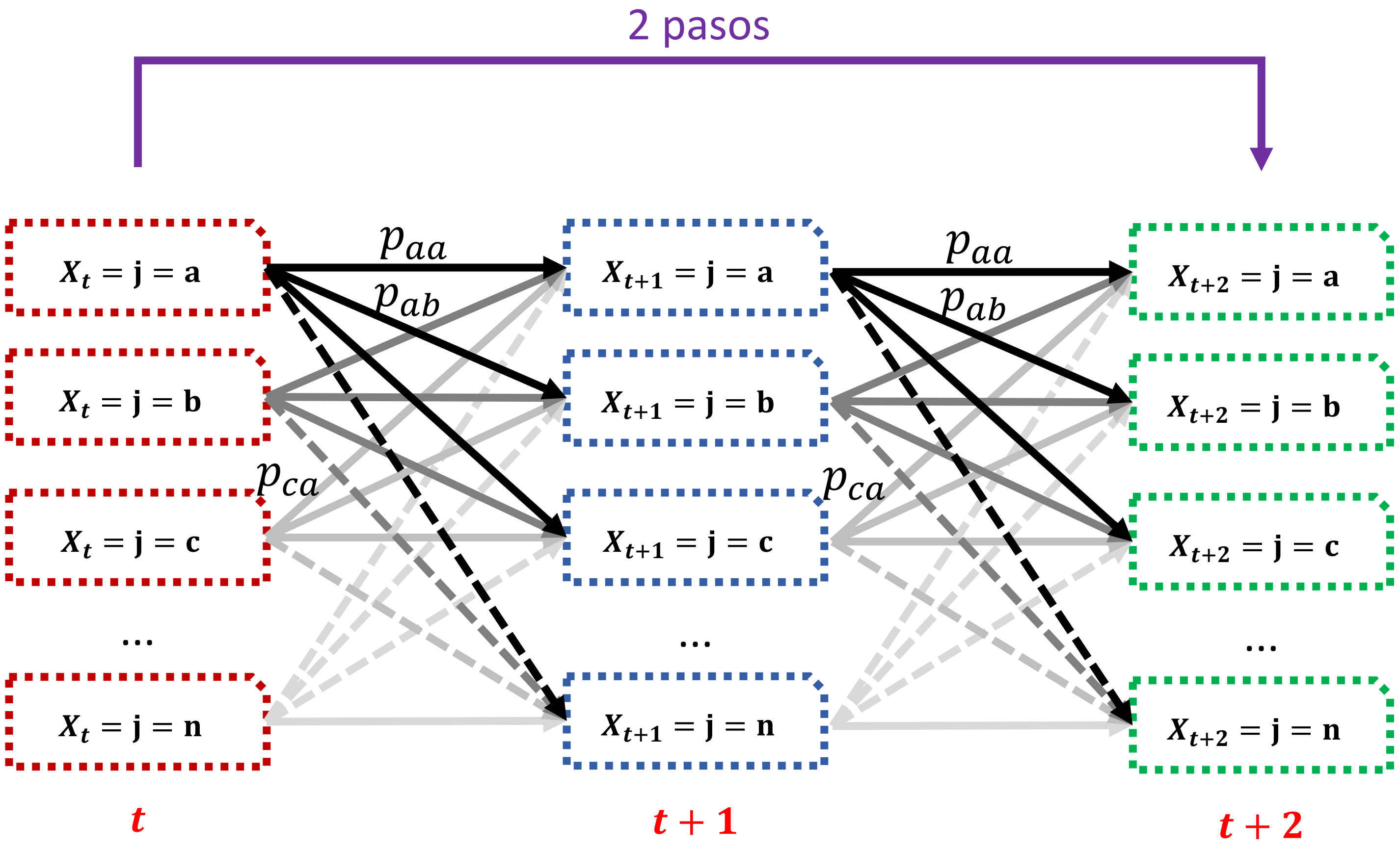
Transición de "m" pasos, caso m=2

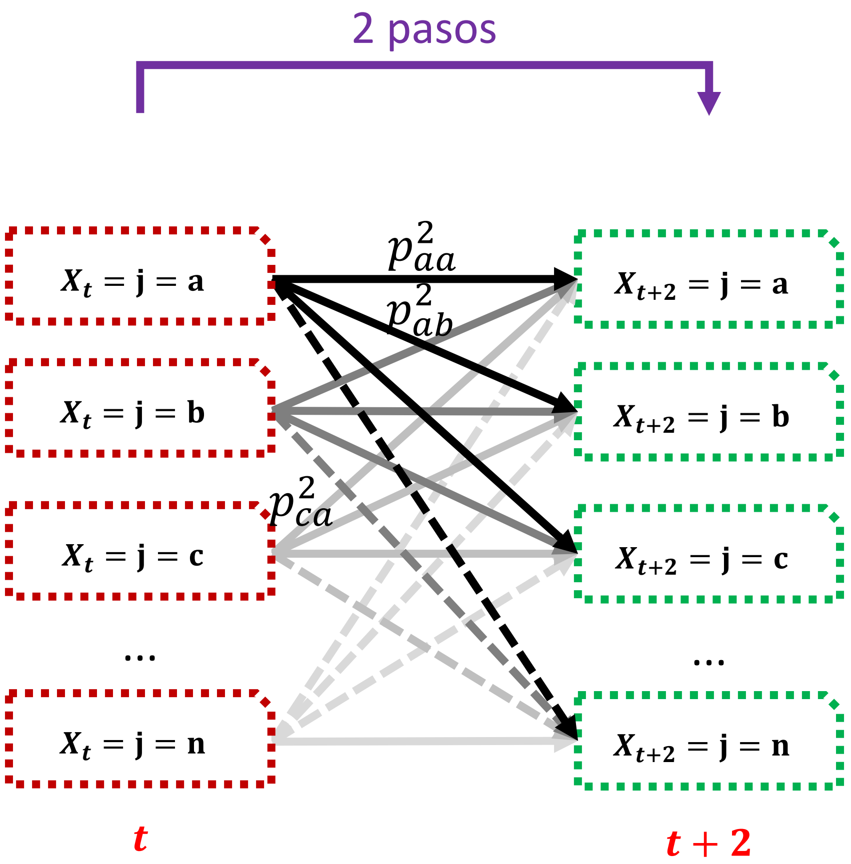


Suma de caminos probables:

$$p_{ca}^2 = p_{ca} * p_{aa} + p_{cb} * p_{ba} + p_{cc} * p_{ca} + \cdots + p_{cn} * p_{na}$$

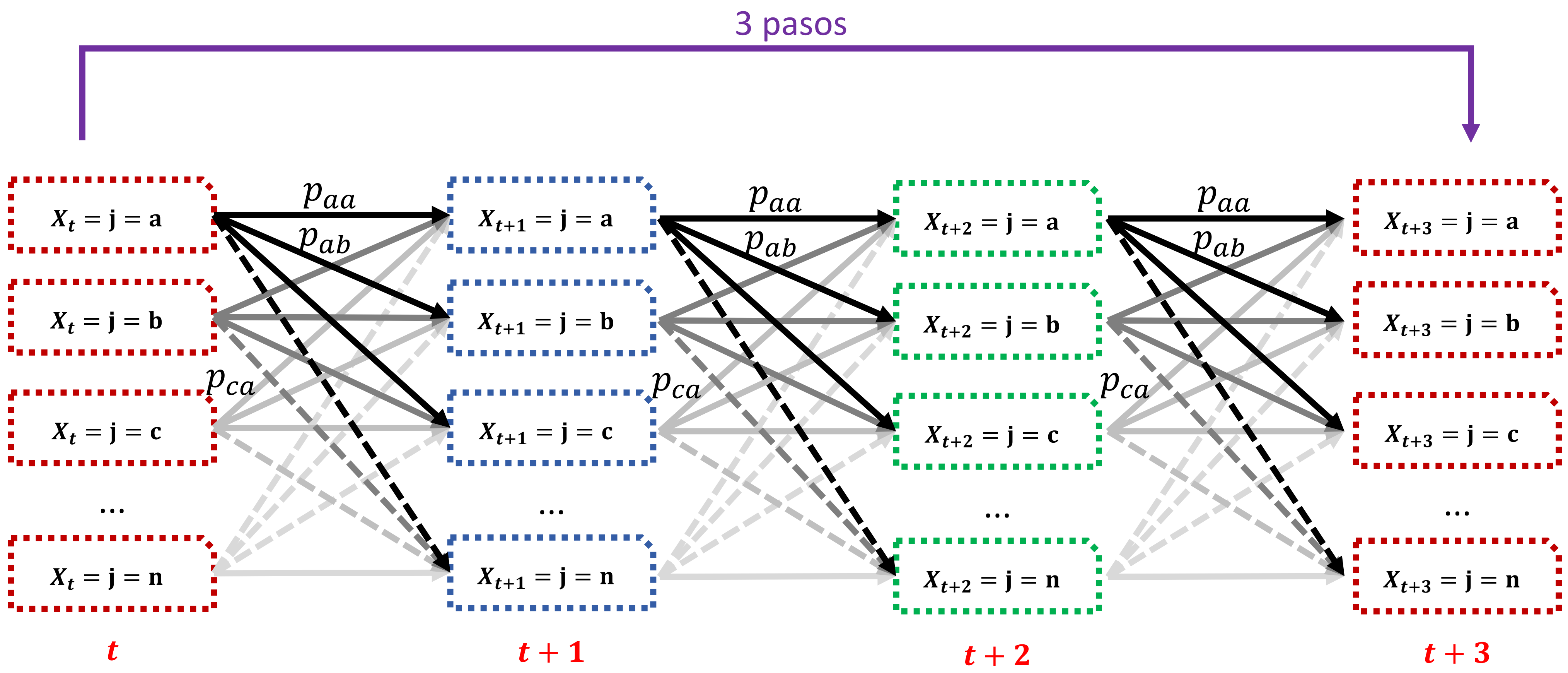
Transición de "m" pasos, caso m=2



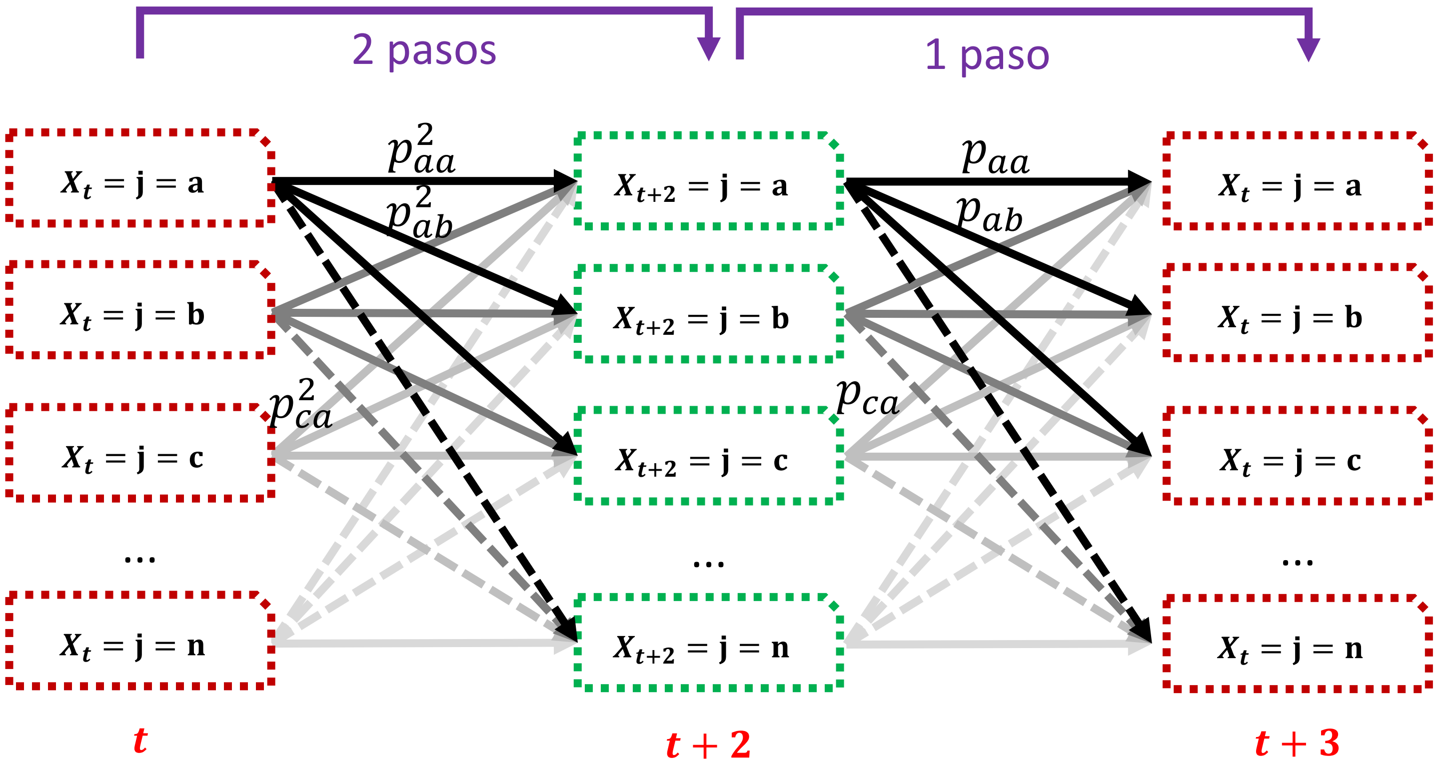


Generalización para trayectoria de i a j en 2 pasos: $$p_{ij}^2 = \sum_k p_{ik} * p_{kj} \quad \forall i,j$$

Transición de “m” pasos, caso $m \geq 2$



Transición de "m" pasos, caso $m \geq 2$



Mismo razonamiento, usamos resultado de 2 pasos, para calcular 3 pasos:

$$p_{ij}^3 = \sum_k p_{ik}^2 * p_{kj} \quad \forall i, j$$

Cadenas de Markov Homogéneas

La probabilidad de transición no depende del parámetro, por ejemplo, no depende del tiempo.

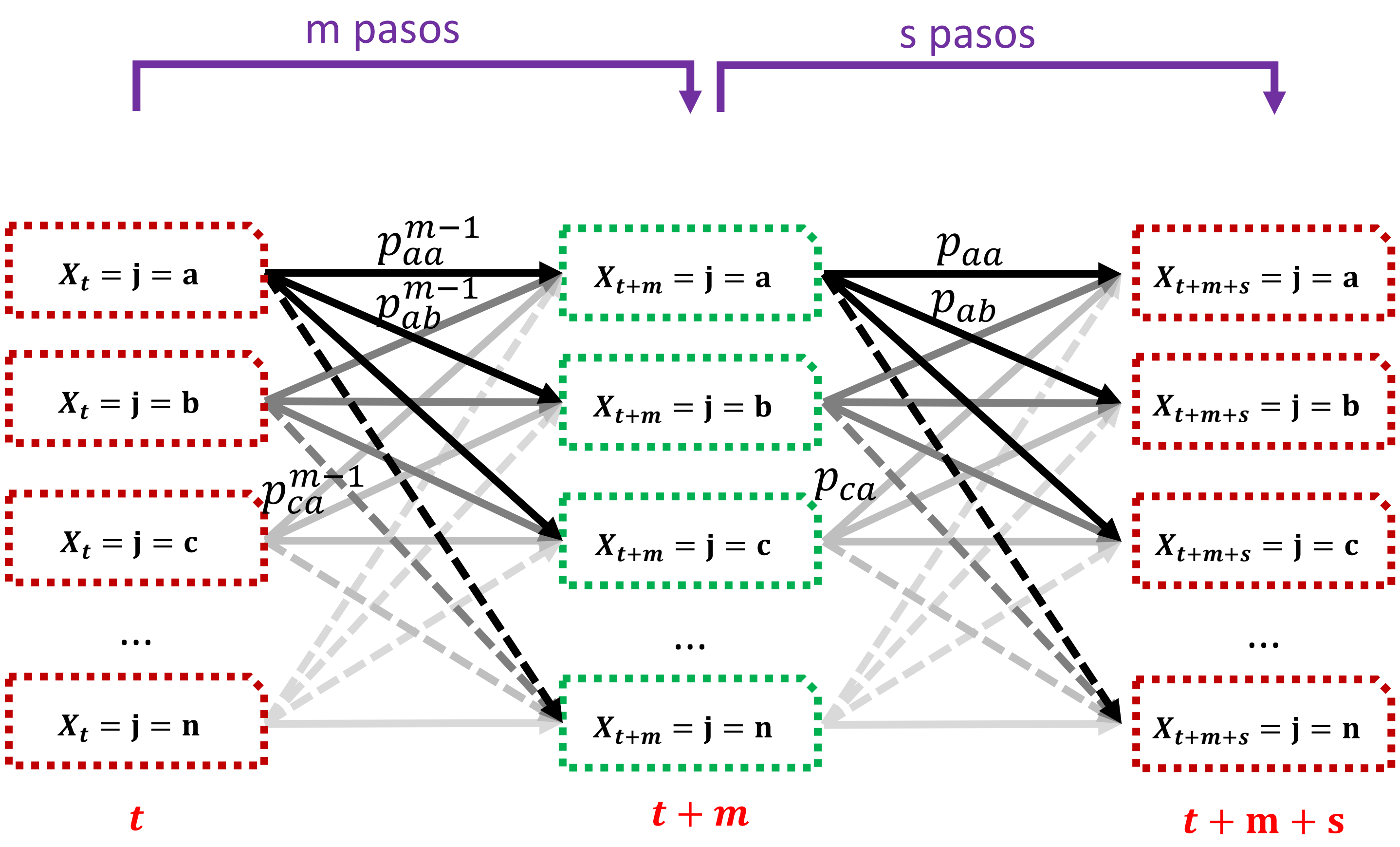
$$P(X_{t+1} \mid X_t) = P(X_{t+m} \mid X_{t+m-1}) \quad \forall m$$

Ej: considero que la probabilidad de transición de rotura y funcionamiento de una máquina, gobernada por una cadena de Markov es la misma a las 18.00 hrs que a las 9.00 hrs.

No homogéneas: dependen del parámetro.

Por ejemplo, filas en el subte a las 15.00 un Domingo o a las 8.30 hrs un Lunes

Regla de Chapman-Kolmogórov



- Para Cadenas de Markov, **procesos discretos**.
- En cadenas **Homogéneas**.

$$\boxed{p_{ij}^{m+s} = \sum_{k} p_{ik}^{m} * p_{ik}^{s} \quad \forall i, j}$$

Forma matricial

Los mismos cálculos pueden expresarse de forma matricial, por ejemplo, un camino de m=2:

$$\boldsymbol{T^1 \times T^1} = \begin{bmatrix} \boxed{\boldsymbol{p_{aa}} & \boldsymbol{p_{ab}} & \boldsymbol{p_{ac}} & \cdots & \boldsymbol{p_{an}}} \\ \boldsymbol{p_{ba}} & \boldsymbol{p_{bb}} & \boldsymbol{p_{bc}} & \cdots & \boldsymbol{p_{bn}} \\ \boldsymbol{p_{ca}} & \boldsymbol{p_{cb}} & \boldsymbol{p_{cc}} & \cdots & \boldsymbol{p_{cn}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{p_{na}} & \boldsymbol{p_{nb}} & \boldsymbol{p_{nc}} & \cdots & \boldsymbol{p_{nn}} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{p_{aa}} & \boldsymbol{p_{ab}} & \boldsymbol{p_{ac}} & \cdots & \boldsymbol{p_{an}} \\ \boldsymbol{p_{ba}} & \boldsymbol{p_{bb}} & \boldsymbol{p_{bc}} & \cdots & \boldsymbol{p_{bn}} \\ \boldsymbol{p_{ca}} & \boldsymbol{p_{cb}} & \boldsymbol{p_{cc}} & \cdots & \boldsymbol{p_{cn}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{p_{na}} & \boldsymbol{p_{nb}} & \boldsymbol{p_{nc}} & \cdots & \boldsymbol{p_{nn}} \end{bmatrix}$$

Suma de caminos probables:

$$p_{ca}^2 = p_{ca} * p_{aa} + p_{cb} * p_{ba} + p_{cc} * p_{ca} + \cdots + p_{cn} * p_{na}$$

Forma matricial

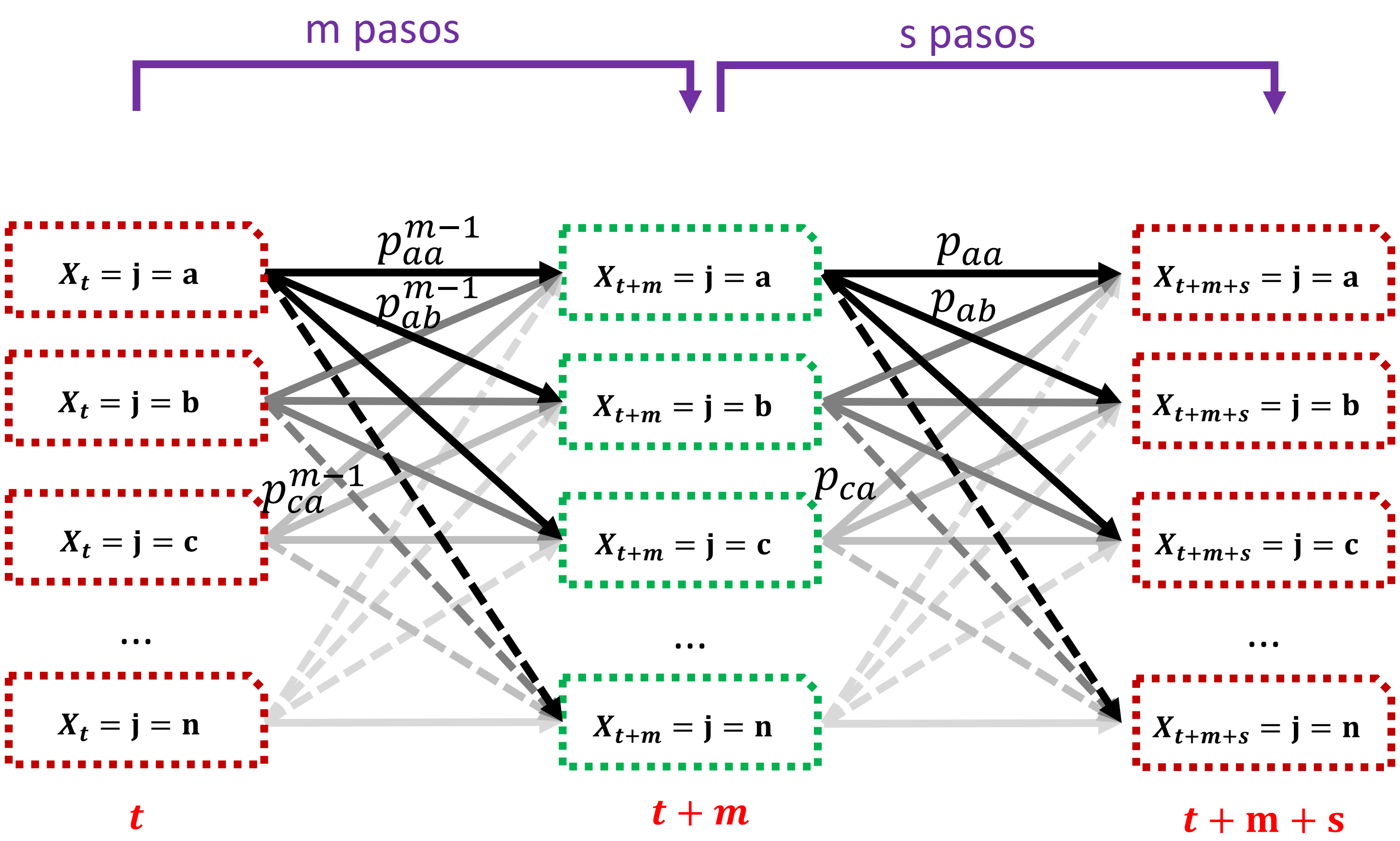
Matriz de transición de 2 pasos:

$$\boldsymbol{T^2} = \boldsymbol{T^1} \times \boldsymbol{T^1}$$

Cada elemento de esa matriz coincide con los cálculos anteriores:

$$p_{ij}^2 = \sum_k p_{ik} * p_{kj} \quad \forall i, j$$

Regla de Chapman-Kolmogórov, matricial



- Para Cadenas de Markov, **procesos discretos**.
- En cadenas **Homogéneas**.

$$\boxed{T^{m+s} = T^m \times T^s}$$

Cada elemento de esa matriz, resulta:

$$p_{ij}^{m+s} = \sum_k p_{ik}^m * p_{ik}^s \quad \forall i, j$$

Vector de distribución de probabilidad de estados

Es un vector fila, que muestra la probabilidad de estar en un estado "i" en tiempo "t". No expresa transición, es un concepto estático.

$$p(t) = [p_1(t) \quad p_2(t) \quad \boxed{p_3(t)} \quad \dots \quad p_i(t)]$$

Probabilidad de estar en el estado 3, en t.

Debe cumplir:

$$\sum_i p(t) = 1$$

Nota: estamos usando la misma letra p para el vector de estados, y para las probabilidades de la matriz de transición. Pero no son lo mismo.

Vector de distribución de probabilidad de estados

Existe un vector de estados para cada tiempo.

$$p(t=1) = [p_1(t=1) \quad p_2(t=1) \quad p_3(t=1) \quad \dots \quad p_i(t=1)]$$
$$p(t=2) = [p_1(t=2) \quad p_2(t=2) \quad p_3(t=2) \quad \dots \quad p_i(t=2)]$$
$$\dots$$
$$p(t=m) = [p_1(t=m) \quad p_2(t=m) \quad p_3(t=m) \quad \dots \quad p_i(t=m)]$$

Es útil para expresar la distribución inicial de mi cadena o encontrar una distribución final.

Vector de distribución de probabilidad de estados

Si el vector de estados tiene una probabilidad con "1", quiere decir que estoy seguro que el agente está en ese estado en un tiempo determinado.

Ej:

$$p(t = 1) = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \end{bmatrix}$$

En t=1, el agente está en el primer estado.

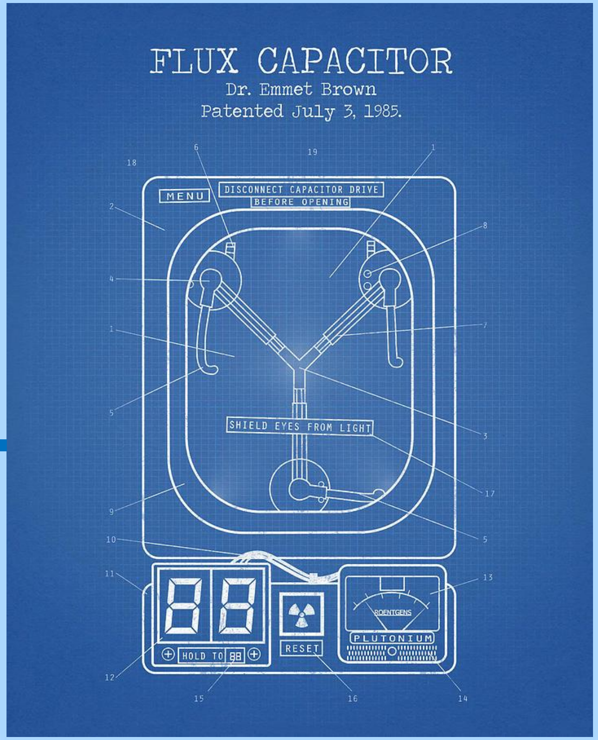
Vector de estado inicial y final

Dados:

- una Matriz de transición de "m" pasos, T^m.
- Un vector de estados inicial en tiempo t, $p(t)$.

Podemos calcular un vector de probabilidad de estados final en t+m $p(t+m)$.

$$p(t+m) = p(t).T^m$$



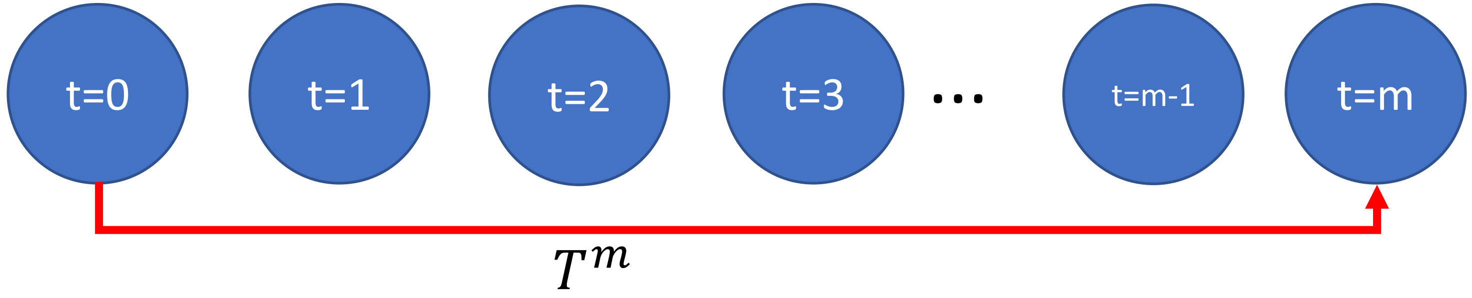
Dennson Creative (2020)
Flux Capacitor Art Work

Vector de estado inicial y final

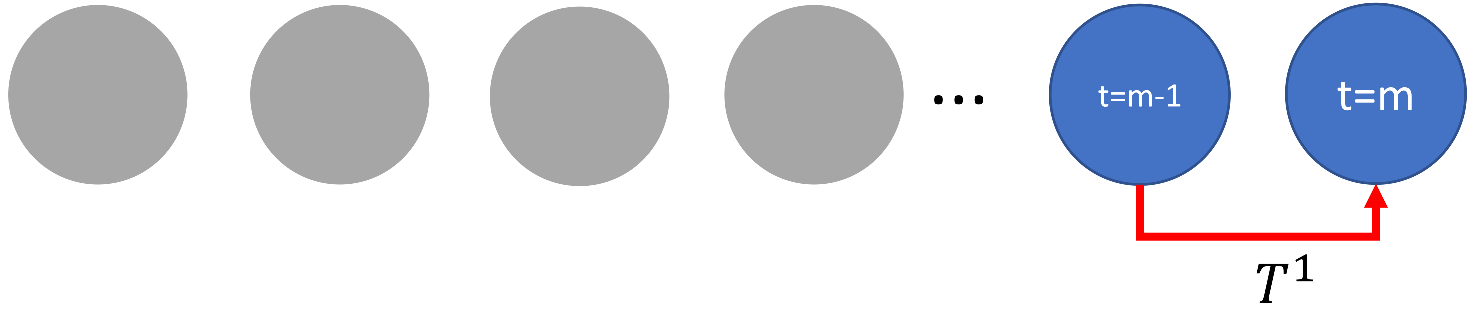
Sabemos que la cadena de Markov es homogénea.

Por lo tanto, estas dos expresiones llevan al mismo estado:

$$p(m) = p(0).T^m$$



$$p(m) = p(m-1).T^1$$



Ejemplo: AGVs en un centro de distribución

En un centro de distribución de una e-commerce se quiere modelizar con una cadena de Markov el trabajo de un AGV.

Se recogen datos de la máquina cada un instante de tiempo de 40 segundos.

- Si en un instante está cargando o descargando mercadería, hay una probabilidad del 90% que en el próximo instante esté en viaje.
- Si está en viaje, hay una probabilidad del 75% que descargue o cargue en el próximo instante.

SRSI https://www.srsi.com/6-reasons-to-implement-a-warehouse-agv/

Ejemplo: AGVs en un centro de distribución

Parámetro:

Tipo de parámetro:

Paso del parámetro:

Estados:

Matriz de transición:

Grafo:

Ejemplo: AGVs en un centro de distribución

Parámetro: *tiempo (t)*

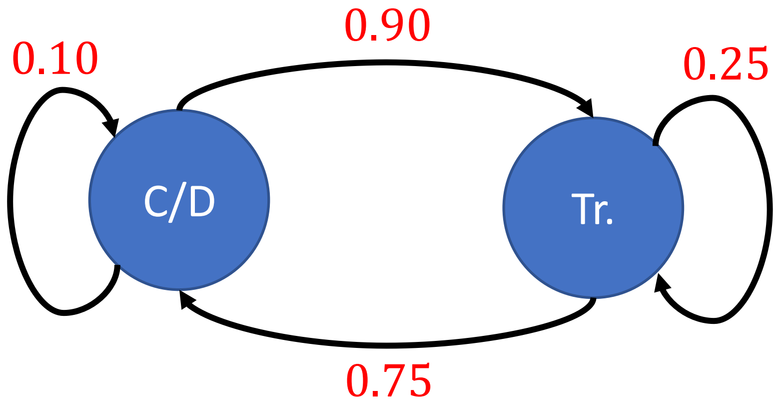
Tipo de parámetro: *discreto*

Paso del parámetro: *40 segundos (Δt)*

Estados: *1) Cargando / Descargando 2) En tránsito*

Matriz de transición:

$$\begin{array}{r c} & \begin{array}{cc} \text{Cargando / Descargando} & \text{En tránsito} \end{array} \\ \begin{array}{r} \text{Cargando / Descargando} \\ \text{En tránsito} \end{array} & \begin{bmatrix} 0.10 & 0.90 \\ 0.75 & 0.25 \end{bmatrix} \end{array} = T^1$$



Ejemplo: AGVs en un centro de distribución

Si en este momento está en tránsito,

¿Cuál es la probabilidad que en 3 instantes de tiempo esté cargando o descargando?

Ejemplo: AGVs en un centro de distribución

Dada la matriz de transición de 1 paso:

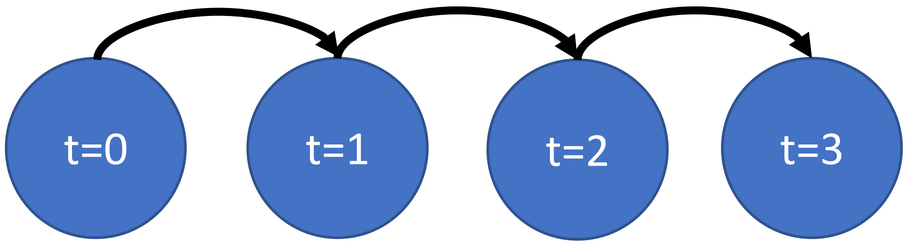
$$T = \begin{bmatrix} 0.10 & 0.90 \\ 0.75 & 0.25 \end{bmatrix}$$

Dada la probabilidad de estado inicial:

$$p(0) = \begin{bmatrix} 0 & 1 \end{bmatrix}$$

Podemos calcular el vector de estado futuro, en t=3

$$p(3) = p(0)T^3$$



Ejemplo: AGVs en un centro de distribución

Matriz de transición de 3 pasos:

$$T^3 = \begin{bmatrix} 0.3047 & 0.6952 \\ 0.5793 & 0.4206 \end{bmatrix}$$

Cálculo de estado futuro:

$$p(3) = p(0)T^3$$

$$p(3) = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0.3047 & 0.6952 \\ 0.5793 & 0.4206 \end{bmatrix}$$

$$p(3) = \begin{bmatrix} 0.5793 & 0.4206 \end{bmatrix}$$

La probabilidad de estar cargando o descargando es de 57.93%

Ejemplo: AGVs en un centro de distribución

Cálculo en Python:

```python
## Matriz de transición como numpy array:
T = np.array([[0.10, 0.90],
              [0.75, 0.25]])

# Generación del vector inicial p_0:
p_0 = np.array([0, 1])
```

```python
# Cálculo del estado a tiempo 3, p_3:
p_3 = np.dot(p_0, T3)

# Estado Carga/Descarga:
print(p_3[0])

>> 0.579375
```

Ejemplo: AGVs en un centro de distribución

Simulación de 10 estados con Python.

```python
# Estados posibles. 0: carga, 1: transito.
estados = [0, 1]

# Cantidad de iteraciones.
h = 10

# Estado inicial.
estado_actual = 1

# Historia de estados.
historia_estados = [estado_actual]
```

```python
# Simulación.
for i in range(0, h):

    # Probabilidades de salto al próximo estado.
    probas = T[estado_actual, :]

    # Cálculo de próximo estado.
    estado_actual = np.random.choice(estados, 1, p=probas)[0]

    # Guardado en historia.
    historia_estados.append(estado_actual)

print(historia_estados)
```

```
>> [1, 1, 0, 1, 0, 1, 1, 1, 0, 1, 0]
```

Ejemplo: AGVs en un centro de distribución

¿Se estabiliza en algún momento la distribución?

Por fuerza bruta:

Podemos calcular distintos vectores de probabilidad de estado $p(m)$ con el mismo estado inicial.

Ejemplo: AGVs en un centro de distribución

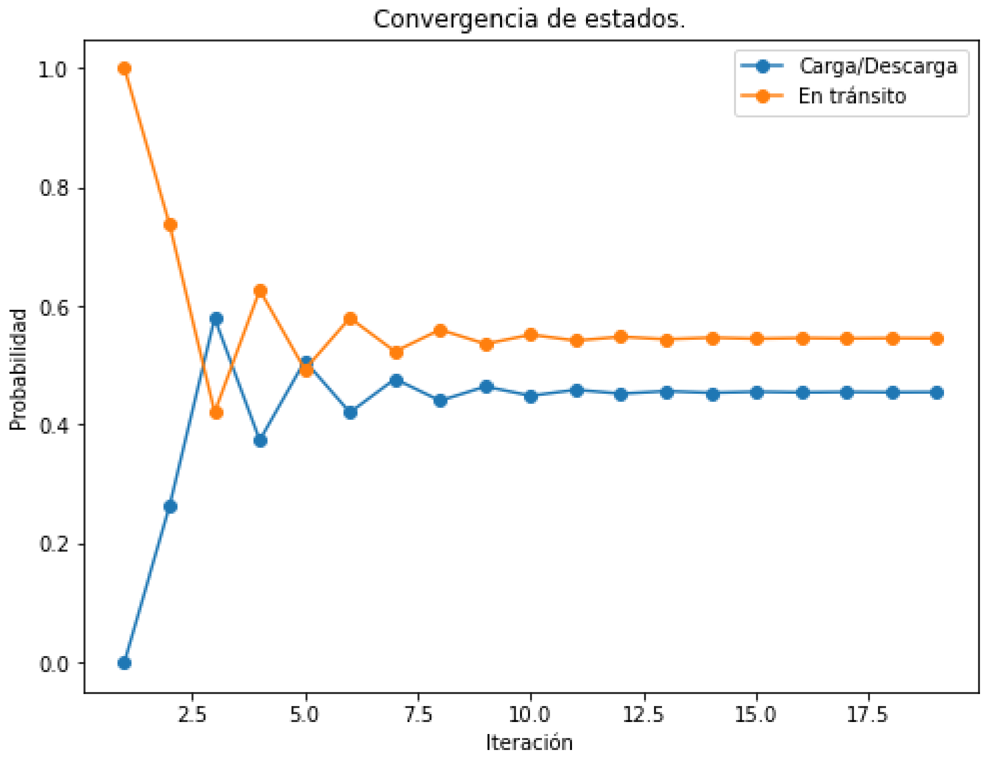
¿Se estabiliza en algún momento la distribución?

```python
def calcular_estado_futuro(m, p_0):

    # Cálculo de la matriz de transición a tiempo m:
    Tm = np.linalg.matrix_power(T, m)

    # Cálculo de vector de proba de estado a m:
    p_m = np.dot(p_0, Tm)

    return p_m
```

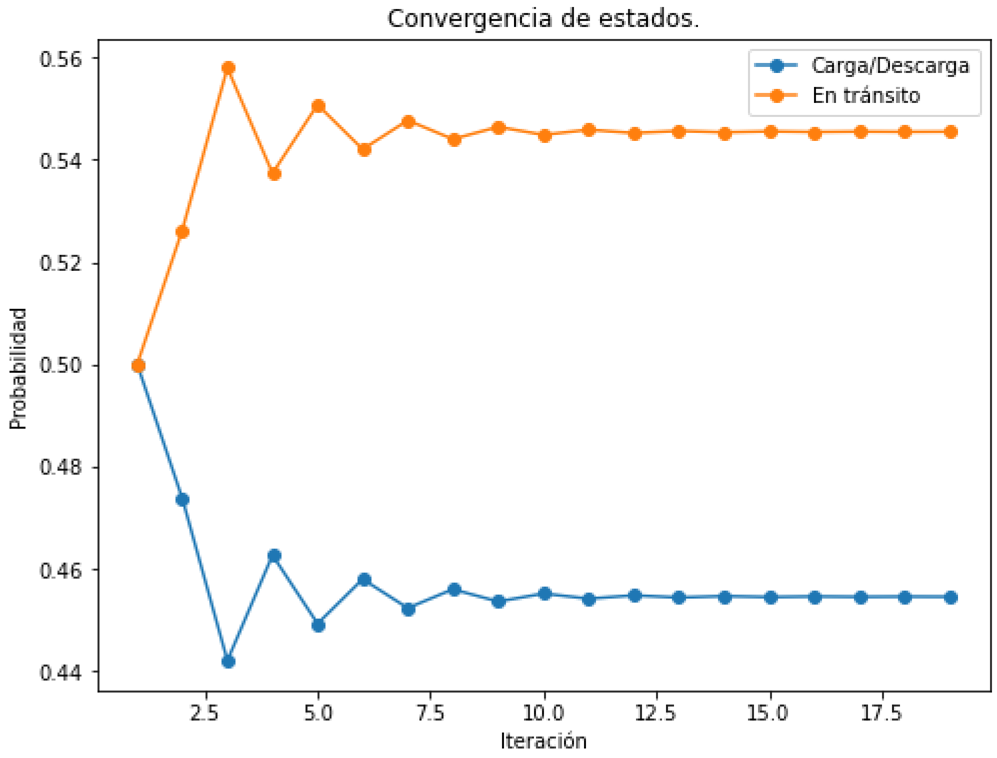
```python
# Cantidad de pasos:
m_tot = 20

# Generación del vector inicial p_0:
p_0 = np.array([0.5, 0.5])

# vectores para guardar estados:
estado_carga = [p_0[0]]
estado_transito = [p_0[1]]

# For loop:
for m in range(2, m_tot):

    p_m = calcular_estado_futuro(m, p_0)

    # guardamos estados:
    estado_carga.append(p_m[0])
    estado_transito.append(p_m[1])

# Plotteamos
plt.plot(range(1, m_tot), estado_carga)
plt.plot(range(1, m_tot), estado_transito)
```

Ejemplo: AGVs en un centro de distribución

¿Se estabiliza en algún momento la distribución?



Estado inicial $p(0) = [0 \quad 1]$



Estado inicial $p(0) = [0.5 \quad 0.5]$