Stochastic Modelling and Simulation
Discrete Time Markov Chains

Luca Bortolussi[1]

[1] Dipartimento di Matematica e Geoscienze
Università degli studi di Trieste

Office 328, third floor, H2bis
lbortolussi@units.it

Data Science and Scientific Computing

DISCRETE-TIME MARKOV CHAINS (DTMC)

$(X_n)_{n\geq 0}$, $n \in \mathbb{N}$, TIME (TIME STEP

$X_n \in S$, STATE SPACE $S = \{s_1, \dots, s_n, \dots\}$

$X_0, X_1, X_2, \dots, X_n, \dots$

STOCHASTIC PROCESS IN DISCRETE TIME and SPACE.

Example : X_n = NUMBER of INFECTED INDIVIDUALS ON AN EPIDEMIC AT DAY n.

$P(X_0, \dots, X_n)$, $\forall n \geq 0$ CHARACTERIZES THE STOCHASTIC PROCESS

$$P(X_0, \dots, X_n) = P(X_0, \dots, X_{n-1})\, P(X_n \mid X_{n-1}, \dots, X_0)$$

MEMORYLESS OR MARKOV PROPERTY

$$\boxed{P(X_n \mid X_{n-1}, \dots, X_0) = P(X_n \mid X_{n-1})}$$ in DTMC

$$\boxed{P(X_n \mid X_{n-1})}$$

IF DEPENDS ON n: TIME INHOMOGENEOUS DTMC

IF IT DOES NOT DEPEND ON n:

TIME HOMOGENEOUS DTMC.

$$P(X_n = s_j \mid X_{n-1} = s_i) = P(X_1 = s_j \mid X_0 = s_i) = \pi_{ij}$$

$\Pi = (\pi_{ij})_{i,j \in S}$ • TRANSITION MATRIX has rows that sum up to 1

$$\pi_{ij} = P(X_1 = s_j \mid X_0 = s_i) \qquad \sum_j \pi_{ij} = 1$$

$\Downarrow$

STOCHASTIC MATRIX

$$A = (a_{ij}) \quad a_{ij} \geq 0 \quad \sum_j a_{ij} = 1$$

$P(X_0 = s_i)$: • INITIAL DISTRIBUTION

$$P(X_0 = s_0, \dots, X_n = s_n) = \overbrace{P(X_n = s_n \mid X_{n-1} = s_{n-1})}^{\pi_{s_{n-1}, s_n}} \cdot \overbrace{P(X_{n-1} = s_{n-1} \mid X_{n-2} = s_{n-2})}^{\pi_{s_{n-1}, s_{n-2}}} \cdot$$

$$\cdot \dots \cdot \underbrace{P(X_1 = s_1 \mid X_0 = s_0)}_{p_{s_0, s_1}} \, \underline{P(X_0 = s_0)}$$

DTMC - DEFINITION

A DTMC $(X_n)_{n\geq 0}$ (TIMEHOMOGENEOUS) IS (S, p_0, Π):

- $S = \{s_1, \ldots, s_n, \ldots\}$ STATE SPACE
- $p_0 = p_0[j] = P(X_0 = s_j)$ INITIAL DISTRIBUTION
- $\Pi = (\pi_{ij}) = P(X_1 = s_j \mid X_0 = s_i)$ TRANSITION MATRIX

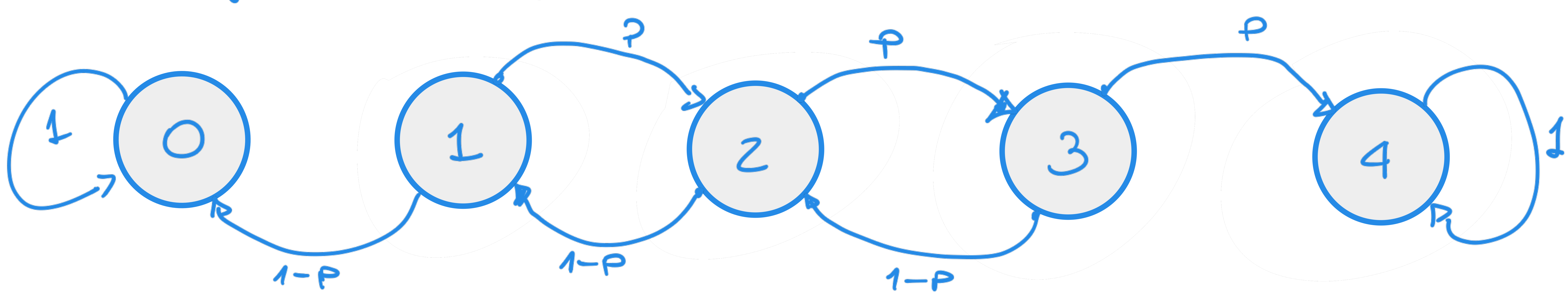
SUCH THAT:

- $P(X_0 = s_j) = p_0[j]$
- $P(X_n = s_j \mid X_{n-1} = s_i, X_{n-2}, \ldots X_0) = P(X_n = s_j \mid X_{n-1} = s_i) = \pi_{ij}$

EXAMPLE: GAMBLER'S RUIN

Consider a gambling game. On any turn you win \$1 with probability $p = 0.4$ or lose \$1 with probability $1 - p = 0.6$. You quit playing if your fortune reaches \$$N$ or \$0.

$S = \{0, 1, \_, N\}$ (\$) $N=4$



EXAMPLE: GAMBLER'S RUIN

Consider a gambling game. On any turn you win \$1 with probability $p = 0.4$ or lose \$1 with probability $1 - p = 0.6$. You quit playing if your fortune reaches $\$N$ or \$0.

The state space is $S = \{0, 1, \ldots, N\}$.

$$
\Pi = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array}
\begin{array}{c}
\begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0.6 & 0 & 0.4 & 0 & 0 \\
0 & 0.6 & 0 & 0.4 & 0 \\
0 & 0 & 0.6 & 0 & 0.4 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}
\end{array}
$$

What is the probability of winning?

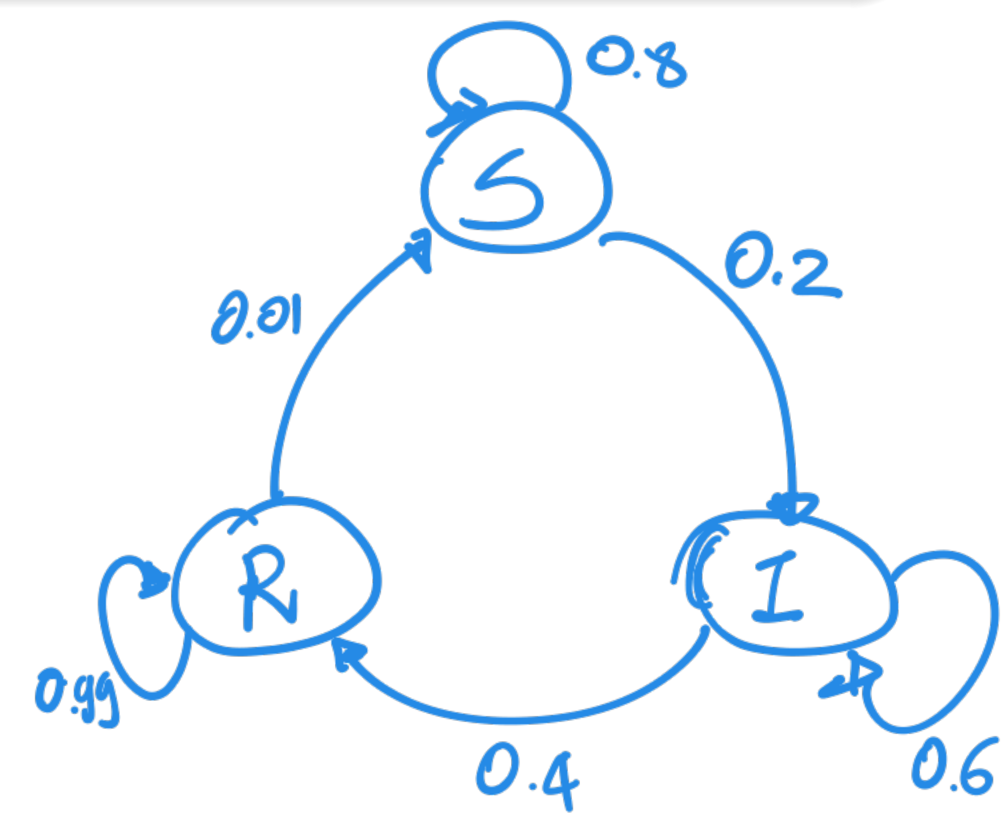
EXAMPLE: FLU

SIRS

A person can be susceptible to flu, infected, or immune (usually after recovery). Susceptibles can be infected with probability 0.2, while infected individuals can recover and become immune with probability 0.4. Immunity is lost with probability 0.01.

State space $S = \{S, I, R\}$

$$\Pi = \begin{array}{c} \\ S \\ I \\ R \end{array} \begin{array}{c} \begin{array}{ccc} S & I & R \end{array} \\ \begin{pmatrix} 0.8 & 0.2 & 0.0 \\ 0.0 & 0.6 & 0.4 \\ 0.01 & 0.0 & 0.99 \end{pmatrix} \end{array}$$



EXAMPLE: FLU

A person can be susceptible to flu, infected, or immune (usually after recovery). Susceptibles can be infected with probability 0.2, while infected individuals can recover and become immune with probability 0.4. Immunity is lost with probability 0.01.

State space $S = \{S, I, R\}$

$$\Pi = \begin{pmatrix} 0.8 & 0.2 & 0.0 \\ 0.0 & 0.6 & 0.4 \\ 0.01 & 0.0 & 0.99 \end{pmatrix}$$

What is the fraction of time one spends ill?

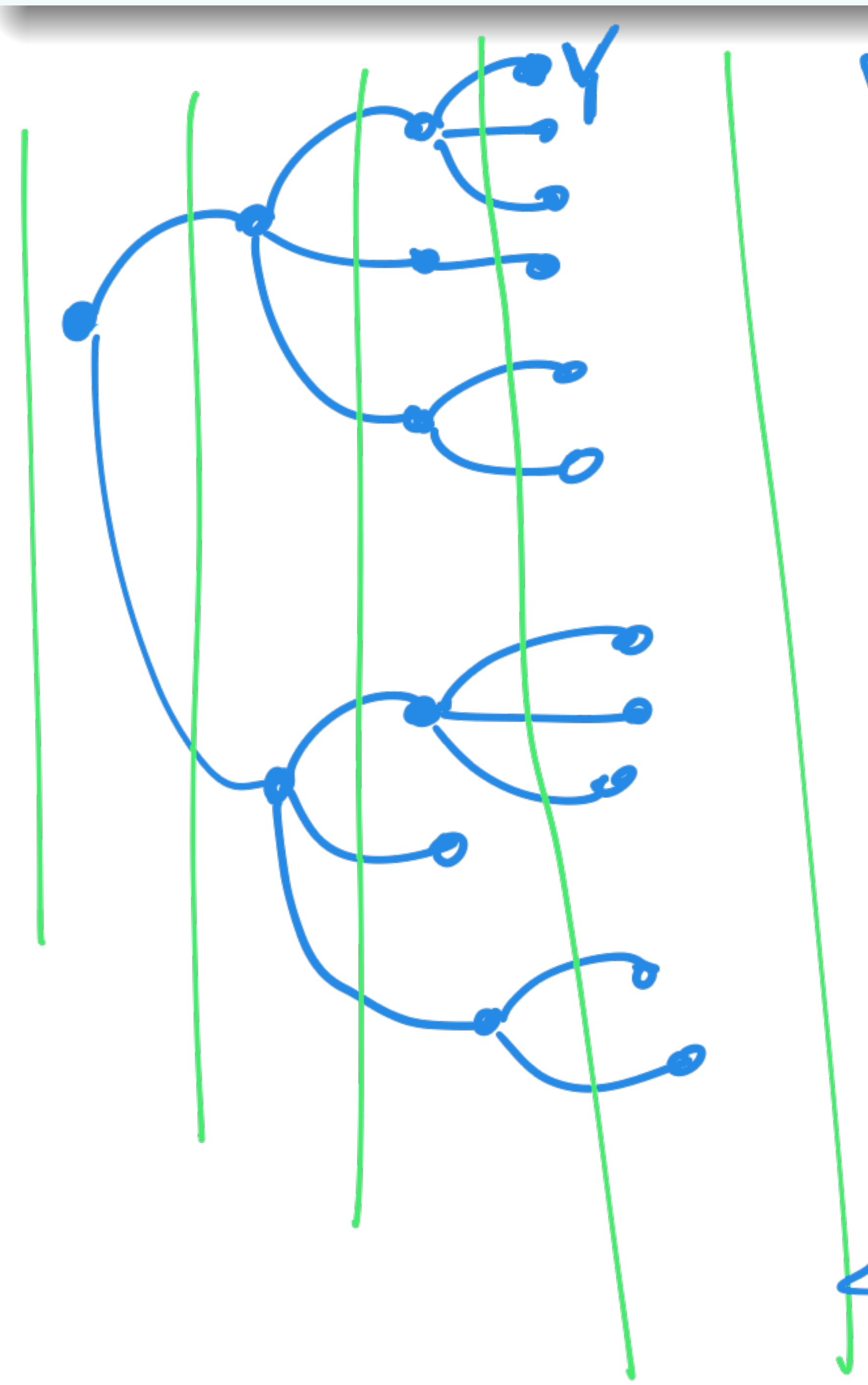
QUESTION

Do the fraction of infected individuals stabilise? To which value?

EXAMPLE: BRANCHING PROCESS

$S = \mathbb{N}_{\geq 0}$

Consider a population, in which each individual at each generation independently gives birth to k individuals with probability p_k. These will be the members of the next generation.



$P(Y=k)=p_k$

$$P(X_{n+1}=j \mid X_n=i) = P(Y_1+\dots+Y_i=j)$$

$$P(X_1=j \mid X_0=1) = P(Y_1=j) = p_j$$

$$P(X_2=3 \mid X_1=2) = P(Y_1+Y_2=3) = P(Y_1=3, Y_2=0) + P(Y_1=2, Y_2=1) + P(Y_1=1, Y_2=2) + P(Y_1=0, Y_2=3) = \boxed{2p_3p_0 + 2p_1p_2}$$

EXAMPLE: BRANCHING PROCESS

Consider a population, in which each individual at each generation independently gives birth to k individuals with probability p_k. These will be the members of the next generation.

The state space is $S = \mathbb{N}$, hence infinite.

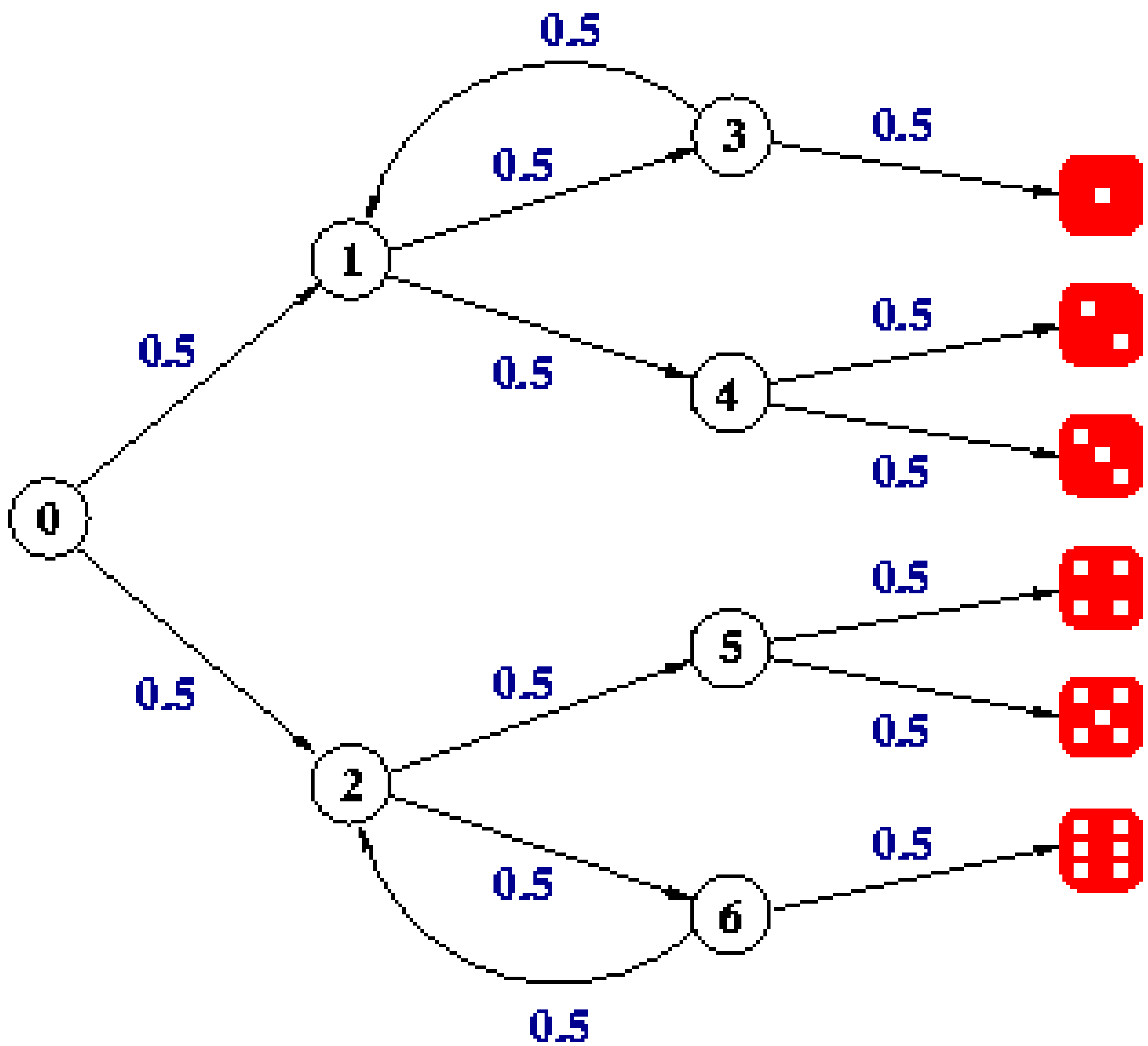
$$\pi(i,j) = P(Y_1 + \ldots + Y_i = j) \text{ for } i > 0 \text{ and } j \geq 0$$

where Y_j is an independent random variable on $\mathbb{N}$ with $\mathbb{P}\{Y_j = k\} = p_k$.

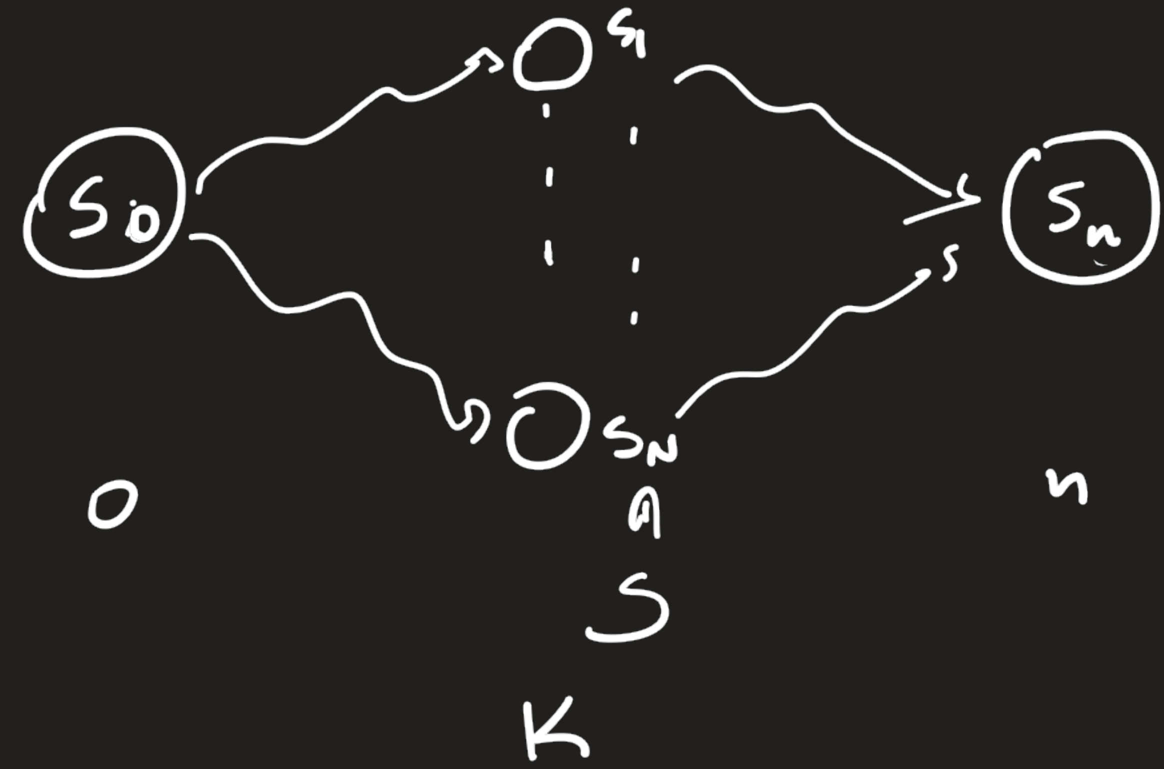
QUESTION

What is the probability of extinction of the population?

EXAMPLE: SIMULATING A DICE WITH A COIN (KNUTH)



Outline



$$\rightsquigarrow \boxed{\underbrace{P(X_n=s_n \mid X_0=s_0)} = \left(\Pi^n\right)_{s_0,s_n}}$$

$$= \sum_{\underbrace{s_1,\dots,s_{n-1}\in S}_{|S|^{n-1}}} P(X_n=s_n, X_{n-1}=s_{n-1},\dots,X_1=s_1 \mid X_0=s_0)$$

$$s_0 \rightsquigarrow s_n \iff s_0 \rightsquigarrow s_k \rightsquigarrow s_n$$

any state

$$s_0 \to s_1 \to s_2 \to \dots \to s_n$$

$$P(X_0=s_0,\dots,X_n=s_n) = P(X_0=s_0)\prod_{j=1}^{n} P(X_j=s_j \mid X_{j-1}=s_{j-1})$$

$$\boxed{\begin{aligned} P(X_n=s_n \mid X_0=s_0) &= \sum_{s_k} P(X_n=s_n, X_k=s_k \mid X_0=s_0) \\ &= \sum_{s_k} P(X_n=s_n \mid X_k=s_k)\, P(X_k=s_k \mid X_0=s_0) \end{aligned}}$$

CHAPMAN KOLMOGOROV EQNS

$$P(X_2=s_2 \mid X_0=s_0) = \sum_{s_1} \underbrace{P(X_2=s_2 \mid X_1=s_1)}_{\Pi_{s_1,s_2}} \underbrace{P(X_1=s_1 \mid X_0=s_0)}_{\Pi_{s_0,s_1}} = \left(\Pi^2\right)_{s_0,s_2}$$

EXAMPLE: FLU

QUESTION

What is the probability that an individual is initially infected, remains infected for one time unit, then recovers just before loosing immunity?

EXAMPLE: FLU

QUESTION

What is the probability that an individual is initially infected, remains infected for one time unit, then recovers just before loosing immunity?

$$\underbrace{I \longrightarrow I \longrightarrow R \longrightarrow S}$$

$P(X_n = I \mid X_0 = S) =$

$= [\Pi^n]_{1,2}$

$$\Pi = \begin{pmatrix} 0.8 & 0.2 & 0.0 \\ 0.0 & 0.6 & 0.4 \\ 0.01 & 0.0 & 0.99 \end{pmatrix}$$

$$\mathbb{P}\{2 \to 2 \to 1 \to 3\} = p_2 \cdot \underbrace{\pi_{2,2} \cdot \pi_{2,3} \cdot \pi_{3,1}} = 0.33 \cdot 0.6 \cdot 0.4 \cdot 0.01 = 0.000792$$

What if $\quad P(X_n = s_n)$?

$$P(X_n = s_n) = \sum_{s_0} \underbrace{P(X_n = s_n \mid X_0 = s_0)}_{\Pi^n} \underbrace{P(X_0 = s_0)}_{p_0} = \Pi^n p_0$$

$$P(X_{n-1}) = \Pi^{n-1} p_0$$

$$P(X_n) = \Pi \cdot P(X_{n-1})$$

matrix x vector.
multiplication

$$P(X_0) = p_0 \qquad P(X_1) = \Pi \cdot P(X_0) \quad - \quad -$$

$$\ldots \quad P(X_n) = \Pi \cdot P(X_{n-1})$$

Outline

1. Discrete Time Markov Chains

2. Chapman-Kolmogorov Equations

3. Absorption probabilities

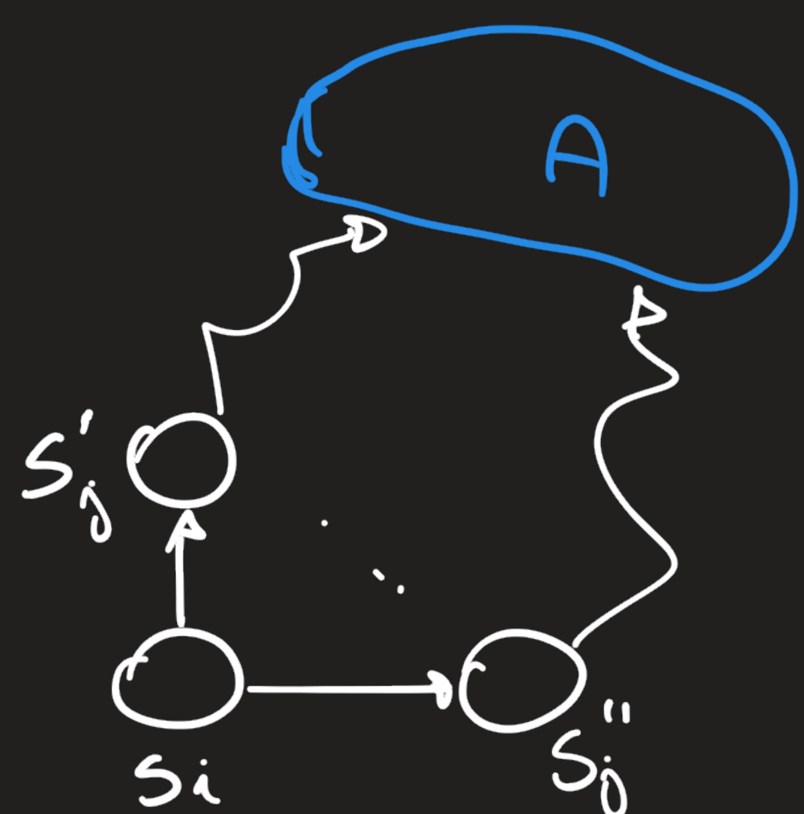
4. Steady state behaviour

ABSORPTION PROBABILITY (REACHABILITY)

$A \subseteq S$, absorption probability of A from state s_i

$h_A[i]$

$$h_A[i] = P(\underbrace{\text{eventually } A}_{\exists n \geq 0 \,:\, X_n \in A} \mid X_0 = s_i)$$



h_A is the LEAST NON-NEGATIVE solution of

$$h_A[i] = \begin{cases} \sum_j \underbrace{\pi_{ij}}_{s_i \to s_j} h_A[j] & \text{if } s_i \notin A \\ 1 & \text{if } s_i \in A \end{cases}$$

EXPECTED HITTING TIME

$A \subseteq S$, THE HITTING TIME

$$\underbrace{\xi_A[i]}_{\text{RANDOM VARIABLE}} = \min\{n \mid X_n \in A \wedge X_0 = s_i\}$$

$\mathbb{E}[\xi_A[i]]$ expectation

$$\begin{cases} \mathbb{E}[\xi_A[i]] = 1 + \sum_j \pi_{ij} \mathbb{E}[\xi_A[j]] & \text{if } s_i \notin A \\ \mathbb{E}[\xi_A[i]] = 0 & \text{if } s_i \in A \end{cases}$$

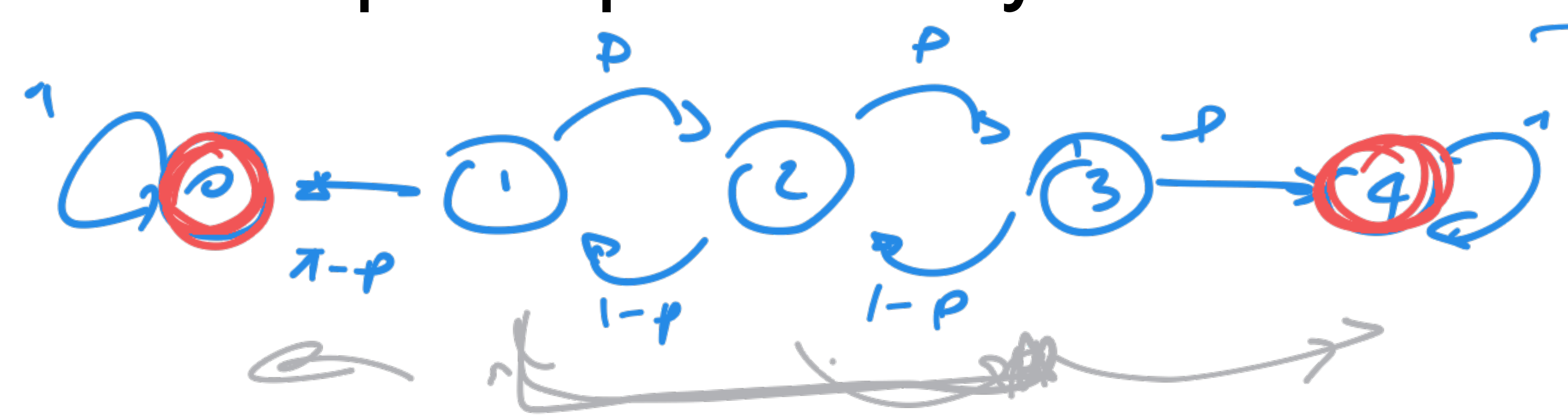
least non-negative solution

EXAMPLE: GAMBLER'S RUIN

QUESTION

What is the probability of the game eventually terminating?

It is the absorption probability of the set $A = \{0, N\}$. For $N = 4$:



$$\begin{cases} h_0 = 1 \\ h_1 = p \cdot h_2 + (1-p) \cdot h_0 \\ h_2 = p \cdot h_3 + (1-p) \cdot h_1 \\ h_3 = p \cdot h_4 + (1-p) \cdot h_2 \\ h_4 = 1 \end{cases}$$

EXAMPLE: GAMBLER'S RUIN

QUESTION

What is the probability of the game eventually terminating?

It is the absorption probability of the set $A = \{0, N\}$. For $N = 4$:

$p = 0.4$

$$\begin{cases} h_0^A = 1 \\ h_1^A = 0.6h_0^A + 0.4h_2^A \\ h_2^A = 0.6h_1^A + 0.4h_3^A \\ h_3^A = 0.6h_2^A + 0.4h_4^A \\ h_4^A = 1 \end{cases}$$

with solution $h^A = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$

Example: gambler's ruin

Question

What is the probability of being ruined?

It is the absorption probability of the set $A = \{0\}$. For $N = 4$:

$$\begin{cases} h_0^A = 1 \\ h_1^A = 0.6h_0^A + 0.4h_2^A \\ h_2^A = 0.6h_1^A + 0.4h_3^A \\ h_3^A = 0.6h_2^A + 0.4h_4^A \\ \boxed{h_4^A = h_4^A} \end{cases}$$

$h_4 = 0$

EXAMPLE: GAMBLER'S RUIN

QUESTION

What is the probability of being ruined?

It is the absorption probability of the set $A = \{0\}$. For $N = 4$:

$$\begin{cases} h_0^A = 1 \\ h_1^A = 0.6h_0^A + 0.4h_2^A \\ h_2^A = 0.6h_1^A + 0.4h_3^A \\ h_3^A = 0.6h_2^A + 0.4h_4^A \\ h_4^A = h_4^A \end{cases} \quad \text{with solution } h^A = \begin{pmatrix} 1.0000 \\ 0.8769 \\ 0.6923 \\ 0.4154 \\ 0 \end{pmatrix}$$

EXAMPLE: GAMBLER'S RUIN

QUESTION

What is the probability of being a happy winner?

It is the absorption probability of the set $A = \{N\}$. For $N = 4$:

$$\begin{cases} h_0^A = h_0^A \\ h_1^A = 0.6h_0^A + 0.4h_2^A \\ h_2^A = 0.6h_1^A + 0.4h_3^A \\ h_3^A = 0.6h_2^A + 0.4h_4^A \\ h_4^A = 1 \end{cases}$$

$h_0 = 0$

EXAMPLE: GAMBLER'S RUIN

QUESTION

What is the probability of being a happy winner?

It is the absorption probability of the set $A = \{N\}$. For $N = 4$:

$$\begin{cases} h_0^A = h_0^A \\ h_1^A = 0.6h_0^A + 0.4h_2^A \\ h_2^A = 0.6h_1^A + 0.4h_3^A \\ h_3^A = 0.6h_2^A + 0.4h_4^A \\ h_4^A = 1 \end{cases} \quad \text{with solution } h^A = \begin{pmatrix} 0 \\ 0.1231 \\ 0.3077 \\ 0.5846 \\ 1.0000 \end{pmatrix}$$

HOW TO COMPUTE THE ABSORPTION PROBABILITY

DTMC (S, P_0, Π) $\qquad A \subseteq S$

S_A ABSORBING

eventually s_A

$$P(\exists n \le \bar{n} : X_n = s_A) = P(X_{\bar{n}} = s_A)$$

① MODIFY the DTMC, to make all states $s_i \in A$ ABSORBING

↓

TAKE Π_{ij} of Π and we replace by e_i ($e_{ij} = 1$ iff $i = j$, other. $e_{ij} = 0$) Π_A

② h_0: $h_0[i] = \begin{cases} 1 & \text{if } s_i \in A \\ 0 & \text{if } s_i \notin A \end{cases}$

③ $h_n = \Pi_A \cdot h_{n-1}$

UNTIL $\|h_n - h_{n-1}\|_\infty \le \varepsilon$ TOLERANCE

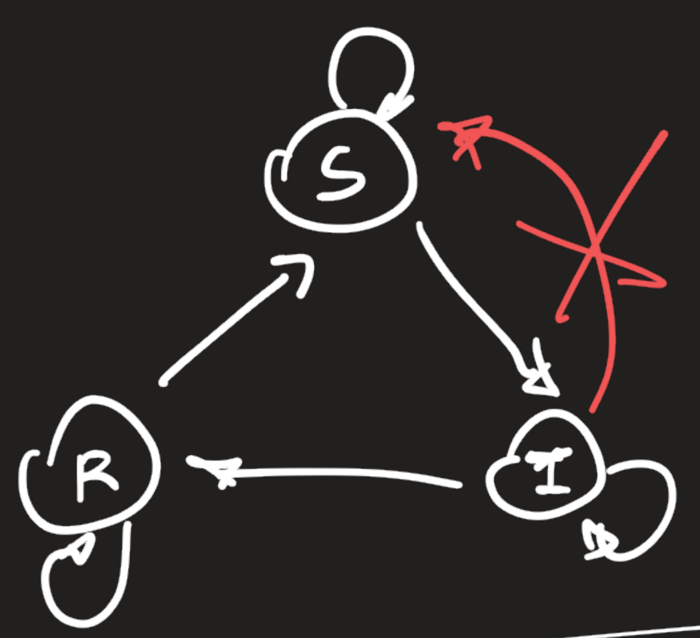
works because $\lim_{n \to \infty} h_n = h_A$

Outline

1. Discrete Time Markov Chains
2. Chapman-Kolmogorov Equations
3. Absorption probabilities
4. Steady state behaviour

DTMC $(S, p_0, \Pi) \rightsquigarrow$ SUPPORT GRAPH $G = (S, \underline{E})$ directed

$E \ni (s_i, s_j)$ iff $\pi_{ij} > 0$.



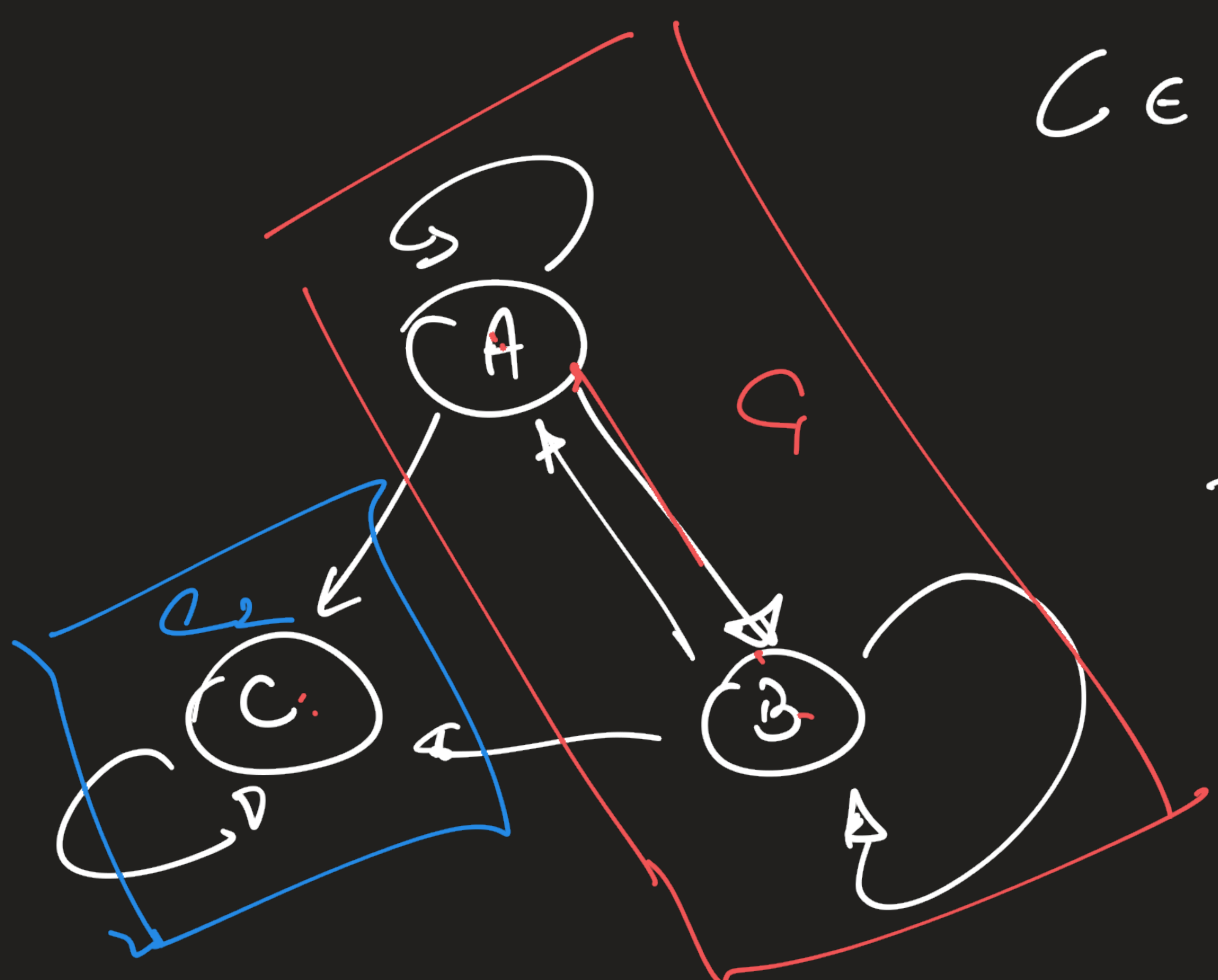
COMMUNICATING CLASSES.

$\underline{s_i \leftrightsquigarrow s_j}$ iff $\exists$ path $s_i \to s_j$ and $s_j \to s_i$ in G

mutually communicating states

$\leftrightsquigarrow$ is an equivalence relation

$C \in S/\leftrightsquigarrow \quad \forall s_i, s_j \in C, \; s_i \leftrightsquigarrow s_j$



$s_1 \in C_1 \rightarrow s_2 \in C_2$

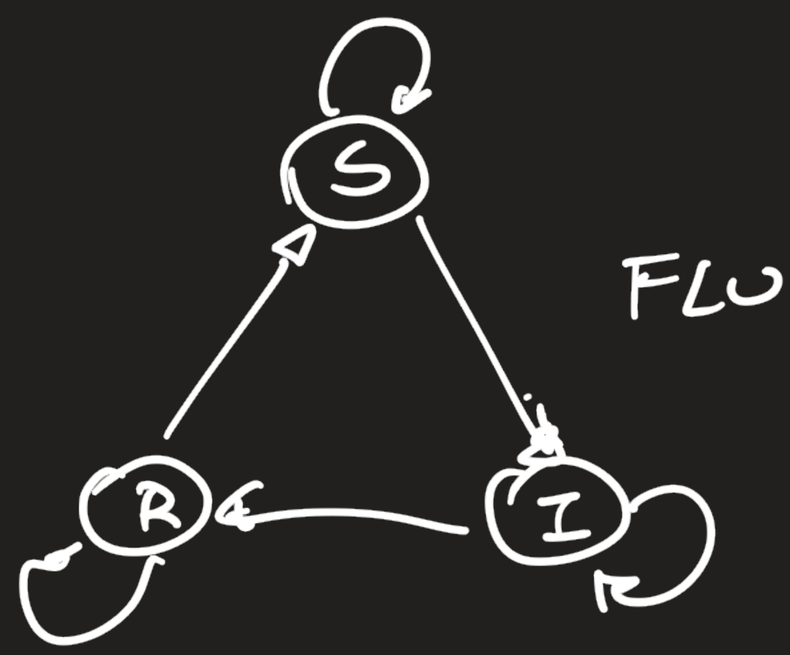
$C_1 \longrightarrow C_2$

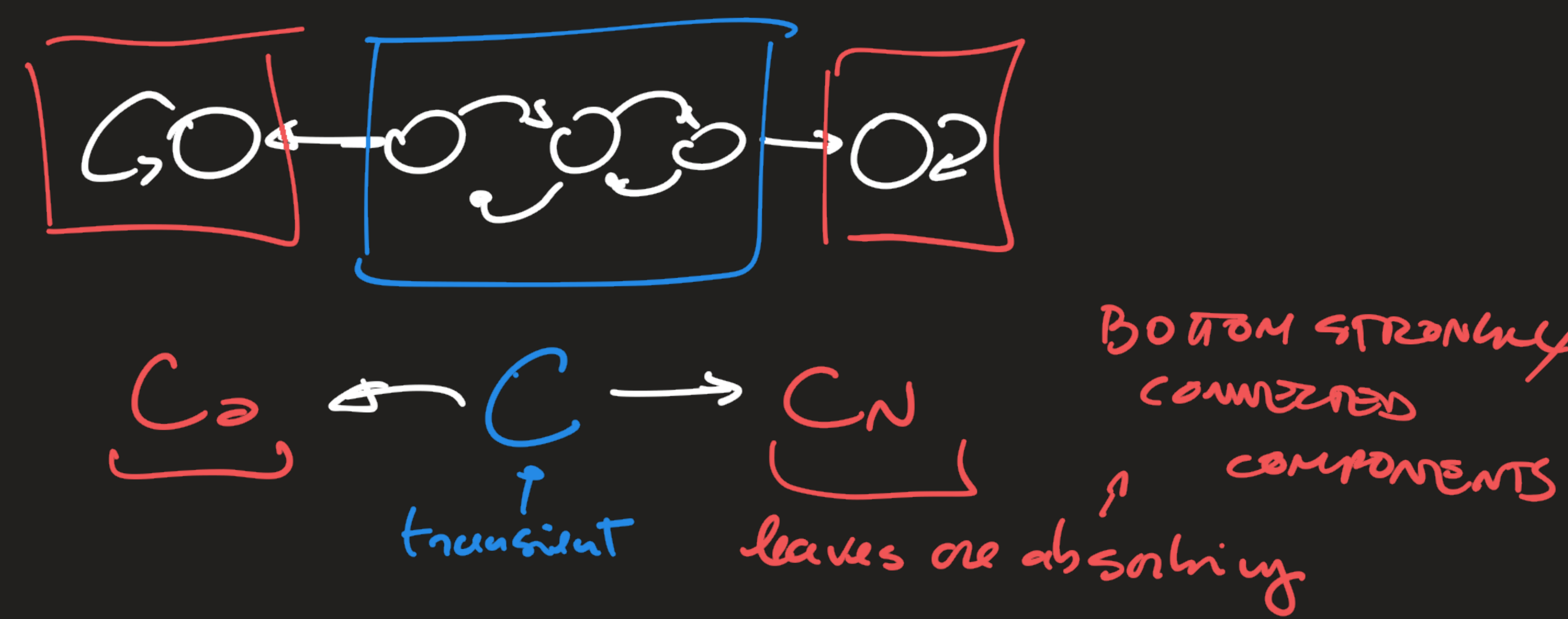
The graph of communicating classes is acyclic (DAG)

STRONGLY CONNECTED COMPONENTS OF G

A DTMC is $\boxed{\text{IRREDUCIBLE}}$ (ergodic) iff

G is STRONGLY CONNECTED (1 SINGLE COMM. CLASS = G)





DTMC : (S, p_0, Π)

POSITIVE RECURRENT

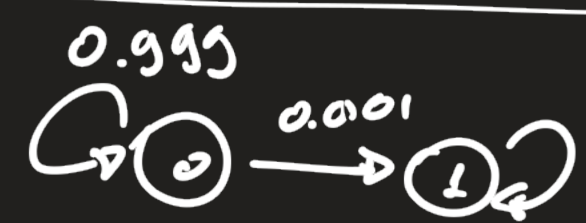
<u>RETURN TIME</u> TO STATE $s_i \in S$

$$T_i = \min\{n > 0 \mid X_n = s_i \wedge X_0 = s_i\} \quad \text{(could be } +\infty\text{)}$$

A STATE IS POSITIVE RECURRENT iff. $\mathbb{E}[T_i] < \infty$

$T_0 = +\infty$

$E[T_0 = +\infty]$



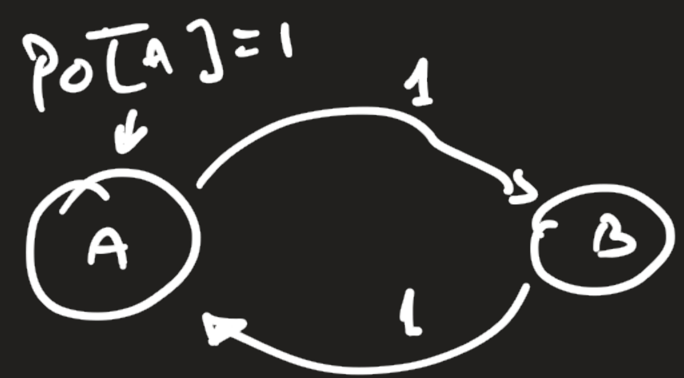
$\mathbb{E}[T_i] = +\infty$

state 0 is not positive recurrent.

APERIODICITY

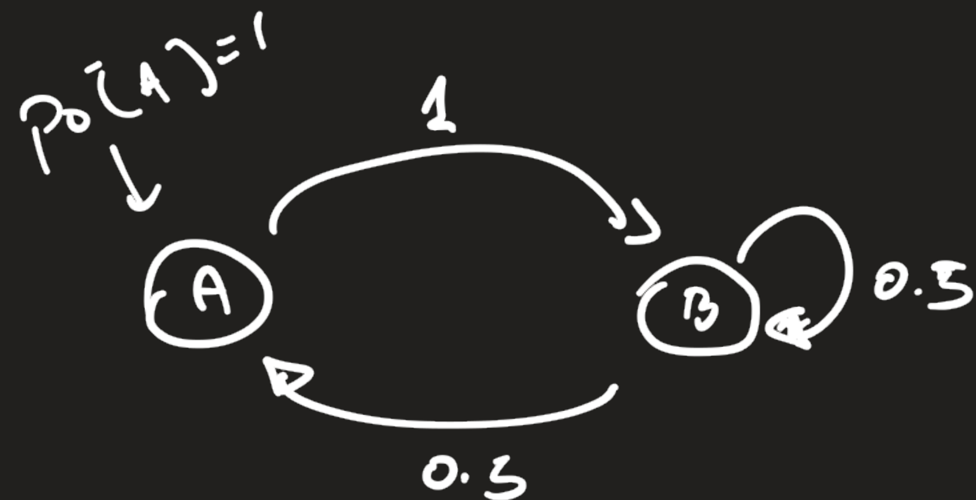
DTMC (S, p_0, Π)

$s_i \in S$ is APERIODIC if $\exists n_0 \geq 0 : (\Pi^n)_{ii} > 0$ for all $n \geq n_0$



$s_A\, s_B\, s_A\, s_B\, s_A\, s_B \dots$ periodic trajectory

$$(\Pi^n)_{AA} = \begin{cases} 1, & \text{if } n \text{ is even} \\ 0, & \text{if } n \text{ is odd} \end{cases}$$



$(\Pi^n)_{AA} > 0, \quad n \geq 2 = n_0$. A is aperiodic

INVARIANT MEASURE

measure: prob distribution over S of a DTMC (S, p, Π)
μ

A MEASURE μ IS INVARIANT iff $\boxed{\mu \Pi = \mu}$

Let DTMC (S, p_0, Π) be IRREDUCIBLE.

These statements are equivalent

1) $\forall s \in S$, s is POSITIVE RECURRENT
2) $\exists s \in S$, s is POSITIVE RECURRENT
3) Π has an invariant MEASURE μ and $\mathbb{E}[T_i] = \frac{1}{\mu_i}$

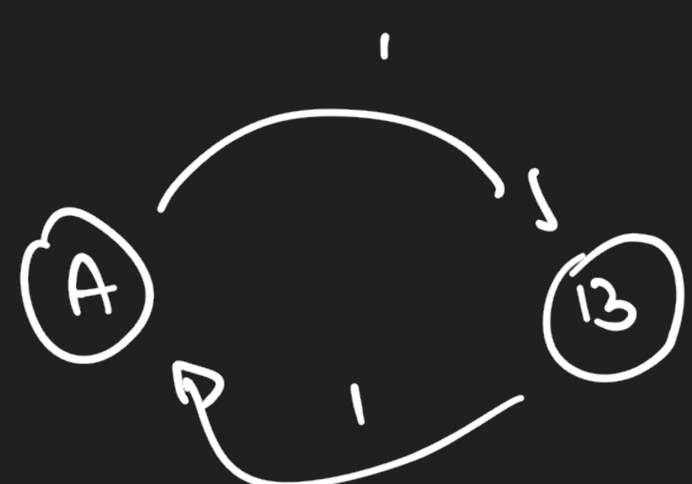
STEADY STATE BEHAVIOUR OF (S, p_0, Π)

IF the DTMC is IRREDUCIBLE and APERIODIC, let μ be an INVARIANT MEASURE for the DTMC

Then

$$\forall s_j \in S \quad P(X_n = s_j) \to \mu_j \text{ as } n \to +\infty$$

and μ is UNIQUE, and does not depend on p_0.



$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \Pi$ has an invariant measure

$$\left(\frac{1}{2}, \frac{1}{2}\right)$$

EXAMPLE: FLU SPREADING

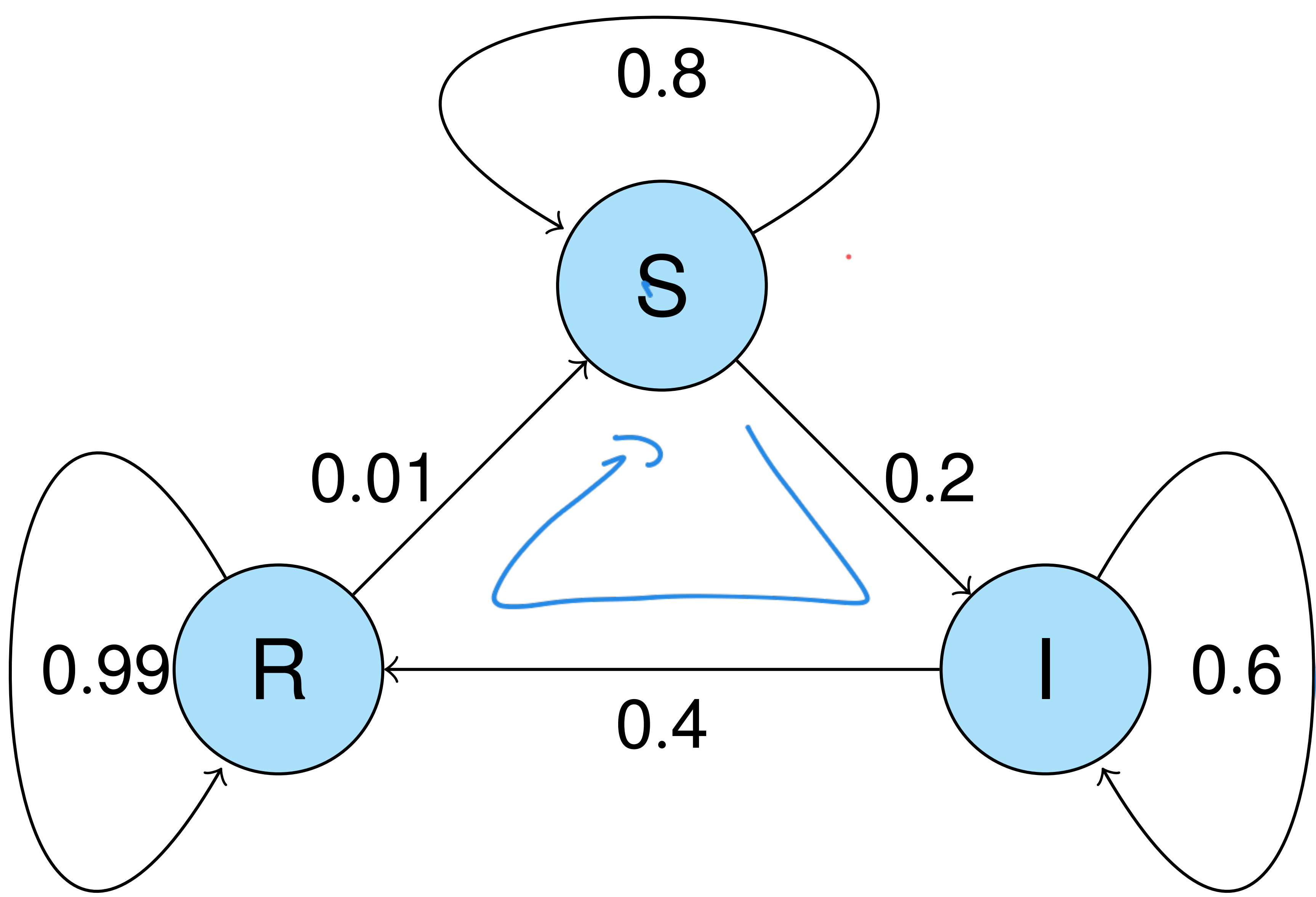
QUESTION

Is there a steady state probability/ invariant measure for the flu example? What is it?

EXAMPLE: FLU SPREADING

QUESTION

Is there a steady state probability/ invariant measure for the flu example? What is it?



$\mu \Pi = \mu$

$$\Pi = \begin{pmatrix} 0.8 & 0.2 & 0.0 \\ 0.0 & 0.6 & 0.4 \\ 0.01 & 0.0 & 0.99 \end{pmatrix}$$

EXAMPLE: FLU SPREADING

QUESTION

Is there a steady state probability/ invariant measure for the flu infection example? What is it?

The invariant measure is given by the unique solution of

$$\mu\begin{pmatrix} 0.8 & 0.2 & 0.0 \\ 0.0 & 0.6 & 0.4 \\ 0.01 & 0.0 & 0.99 \end{pmatrix} = \mu, \text{ which is } \mu = \begin{pmatrix} 0.0465 \\ 0.0233 \\ 0.9302 \end{pmatrix}$$

COMPUTATION OF INVARIANT MEASURE

SOLUTION $\left[\mu \Pi = \mu\right]$ $\quad \mu(\Pi - I_n) = 0$, with condition $\sum \mu_n = 1$

$$\mu \begin{pmatrix} \Pi - I_n \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

---

p_0 -

$$p_n = p_{n-1} \cdot \Pi \quad , \quad \lim_{n \to \infty} p_n = \mu$$

$\|p_n - p_{n-1}\|_\infty < \varepsilon$ TOLERANCE

n MIXING TIME OF the DTMC (depends on second largest eigenvalue of Π)

STEADY STATE FOR NON-IRREDUCIBLE CHAINS

DTMC (S, p_0, Π)

$G = (S, E)$ has more than one s.c.c. $C_1 \dots C_n$

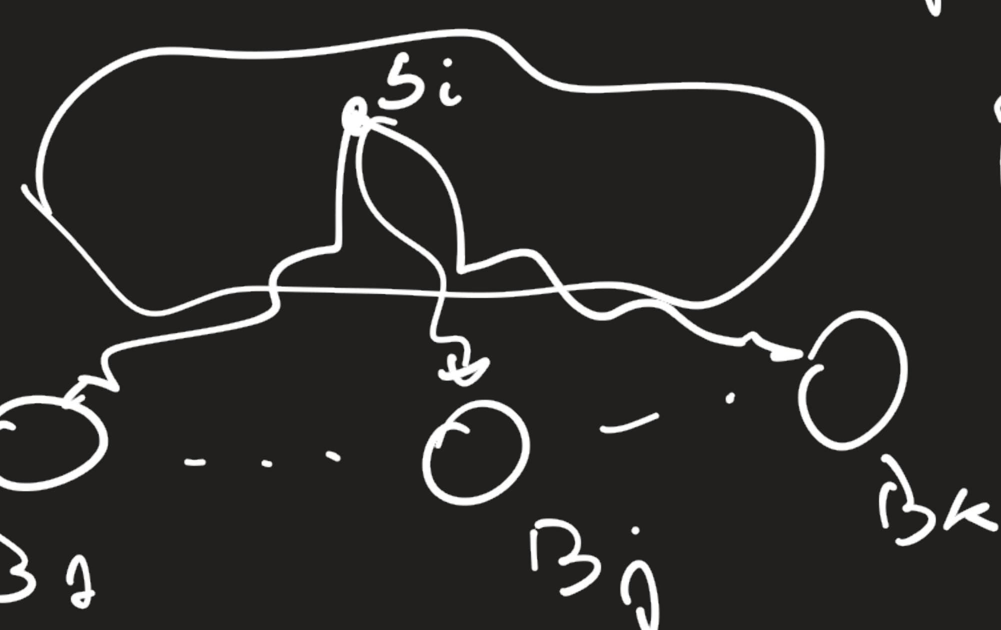
$C_1, \dots, C_n$ are arranged in an acyclic graph.

$\{C_1, \dots, C_n\} \supseteq \{B_1, \dots, B_k\}$ bottom s.c.c.

Assume B_j are APERIODIC.

$\Downarrow$

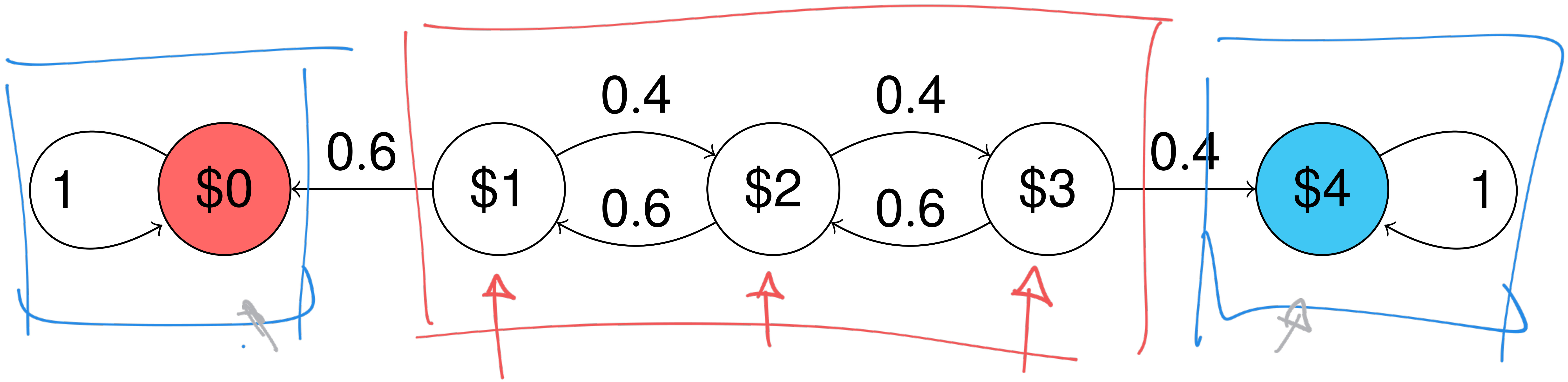
μ_j INVARIANT MEASURE for B_j



$P(X_n \mid X_0 = s_i) \xrightarrow[n \to \infty]{} \boxed{\sum_j h_{B_j}[i] \cdot \mu_j}$

$h_{B_1}[i]$

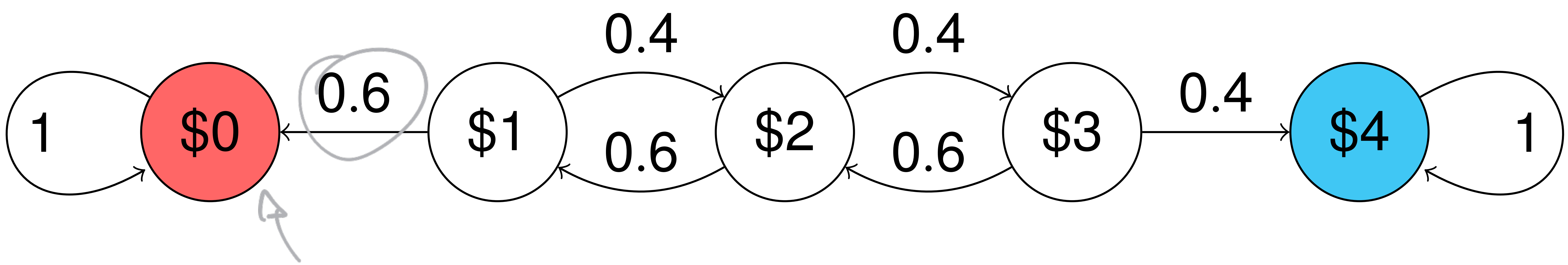
EXAMPLE: GAMBLER'S RUIN



$\mu_0 = 1$ if $s = 0$

$$P(X_n \mid X_0 = \$2) = \begin{pmatrix} h_0[2] \\ 0 \\ 0 \\ 0 \\ h_4[2] \end{pmatrix}$$

EXAMPLE: GAMBLER'S RUIN



In the gambler's ruin model, we have two single-state bottom s.c.c.: 0 and N.

Hence we have the following steady state (conditional on the initial state):

$$
\begin{aligned}
\mu(\cdot \mid 0) &= (1, 0, 0, 0, 0) \\
\mu(\cdot \mid 1) &= (0.8769, 0, 0, 0, 0.1231) \\
\mu(\cdot \mid 2) &= (0.6923, 0, 0, 0, 0.3077) \\
\mu(\cdot \mid 3) &= (0.4154, 0, 0, 0, 0.5846) \\
\mu(\cdot \mid 4) &= (0, 0, 0, 0, 1)
\end{aligned}
$$

REFERENCES

- J.R. Norris. Markov Chains, Cambridge University Press, 1998.
- R. Durrett, Essentials of Stochastic Processes, Springer-Verlag, 1998.

SIMULATION OF DTMC

DTMC $(S, \underline{p_0}, \Pi)$

$\cancel{\left[P_n[x] \right]}$

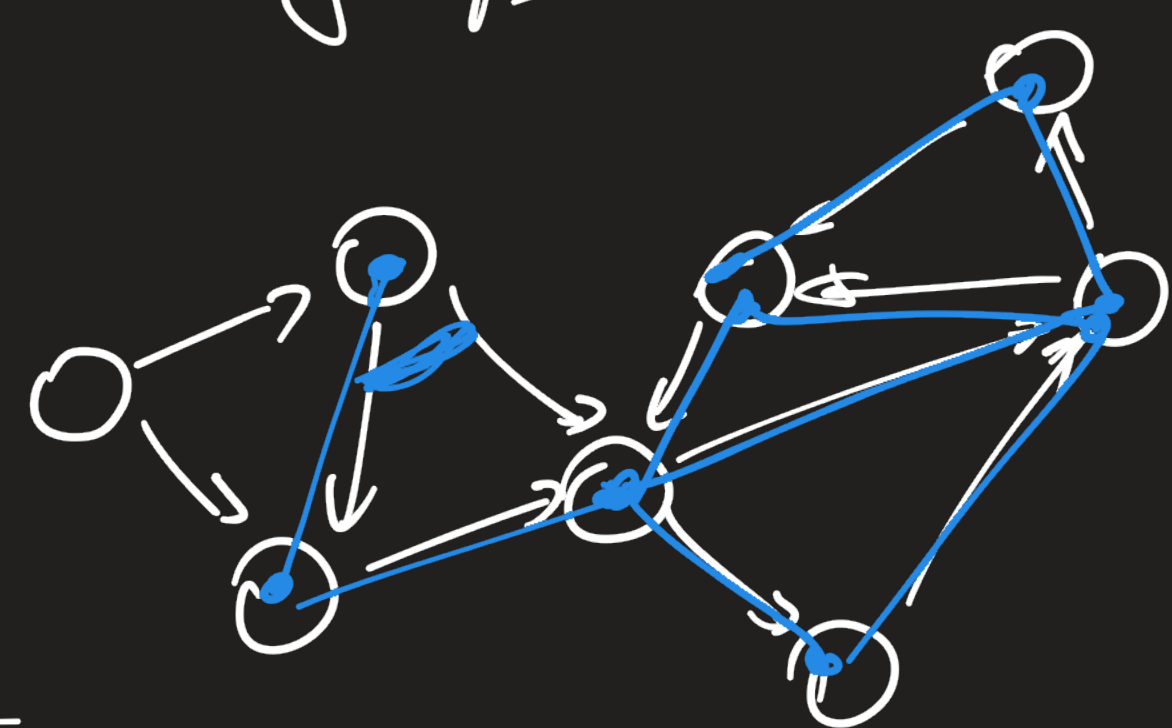
We want to generate PATHS according to the p.d. of the DTMC

sample $s_0 \sim p_0$ ← FINITE SUPPORT: sampling from a discrete distribution.

FOR i in $1:N$ $\quad$ $\#\underline{N}$ = number of steps

sample $s_i \sim \Pi_{s_{i-1},\cdot}$ ←

RETURN $s_0 \cdots s_N$



$A \subseteq S$

$$P_N(A) \approx \frac{|\{s_N \in A\}|}{K}$$