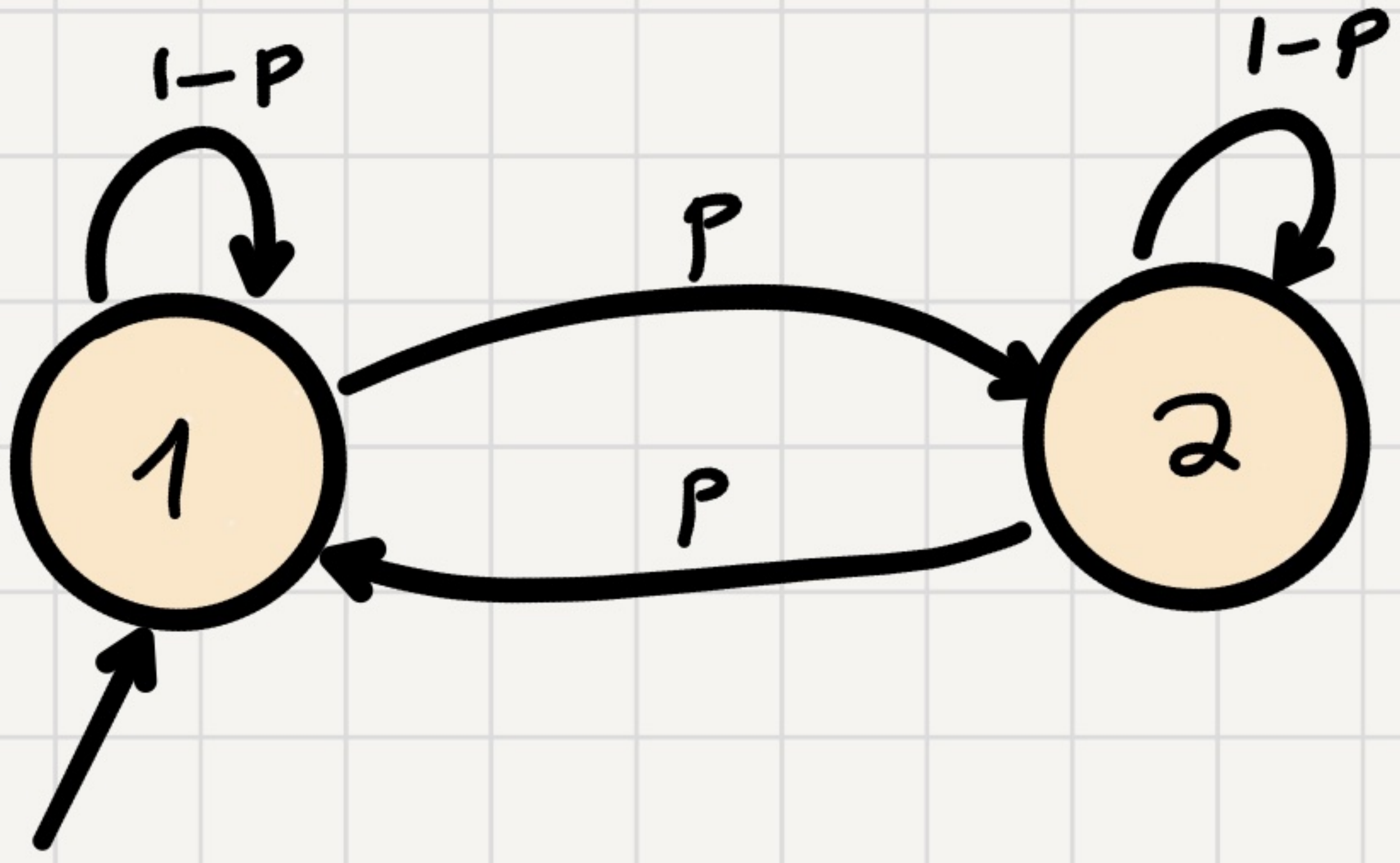
Lecture 26

Markov Chains

Example



- Start at 1.
- Every step, you stay in place w.p. $1-p$

  or move to the other state w.p. p.

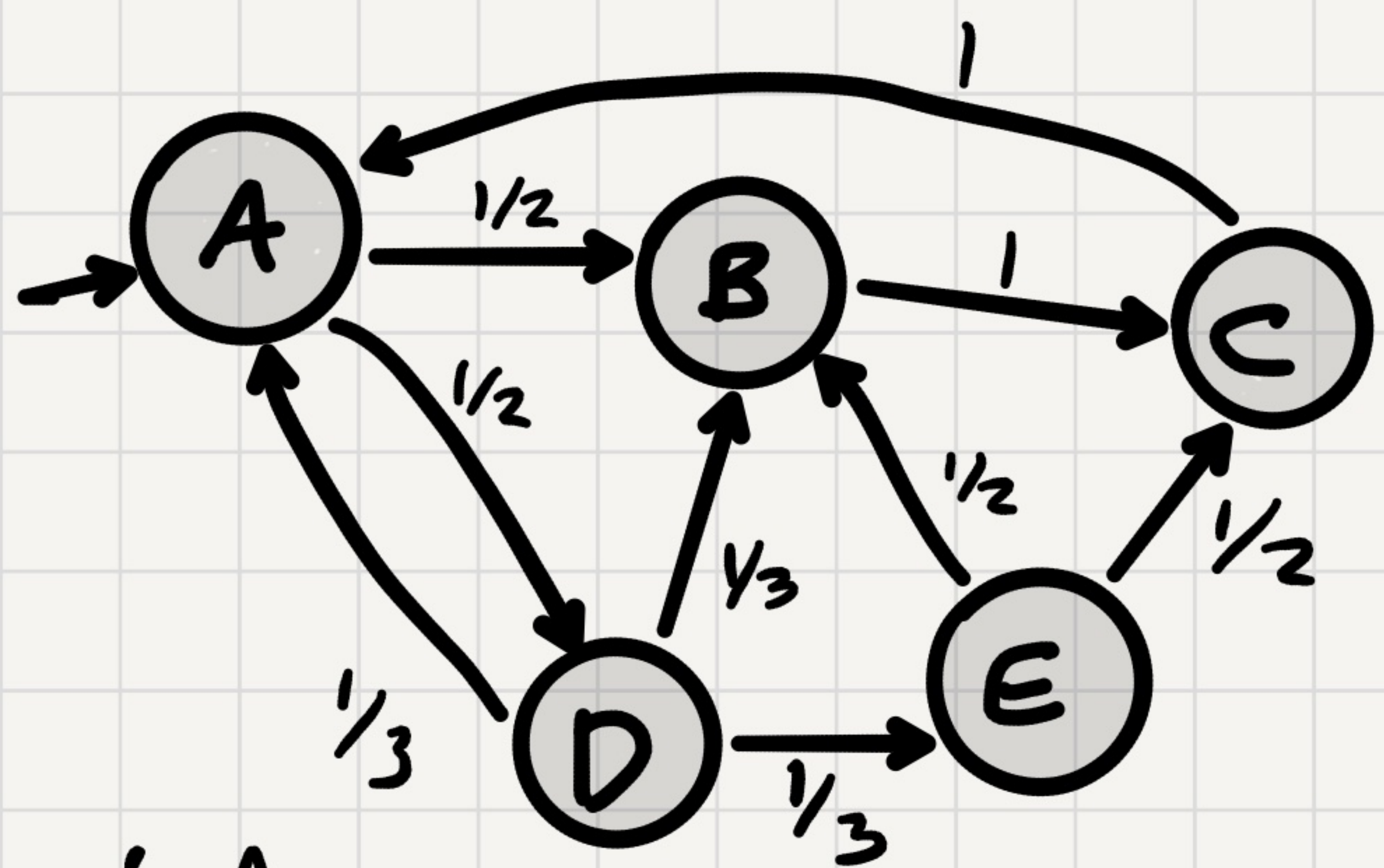
$p=0.2$, a sequence might be:

1 1 1 1 2 2 2 2 2 1 1 1 1 2 2 2 2 2 2 1 1 1

$p=0.5$,

1 2 2 2 1 2 1 1 2 1 2 ...

Another Example



Start at A.

At each step, follow one of the edges of the current state, with equal prob.

Simulation:

A B C A B C ABCADADBC ABCADADEB ...

Markov Chain – Definition

Ingredients:

- A finite set of states $\mathcal{X} = \{1, 2, 3, \dots, K\}$
- An initial prob. dist. π_0 on $\mathcal{X}$.
- Transition probabilities $P(i,j)$ for $i, j \in \mathcal{X}$

(the probability to move from i to j in one step).

We define a sequence of random variables

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

- X_0 is sampled from π_0, i.e., $\Pr[X_0 = i] = \pi_0(i) \ \forall i \in \mathcal{X}$
- $\forall n \geq 0 \ \forall i_0, \dots, i_{n-1}, i, j \quad \Pr[X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0]$

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

the only thing we remember about the past is the current state!

Markov Chain – Definition

Ingredients:

- A finite set of states $\mathcal{X} = \{1, 2, 3, \ldots, K\}$
- An initial prob. dist. π_0 on $\mathcal{X}$.
- Transition probabilities $P(i,j)$ for $i, j \in \mathcal{X}$

We define a sequence of random variables

$$X_0, X_1, X_2, \ldots, X_n, \ldots$$

- X_0 is sampled from π_0, i.e., $\Pr[X_0 = i] = \pi_0(i)$ $\forall i \in \mathcal{X}$
- $\forall n \geq 0$ $\forall i, j$ $\quad \Pr[X_{n+1} = j \mid X_n = i, X_{n-1}, \ldots, X_0] = P(i,j)$

What's the distribution of X_1?

What's the distribution of X_n?

What's the Distribution of X_1?

$$Pr[X_0 = i] = \pi_0(i) \quad \forall i \in \{1,2,\dots,k\}$$

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

Let's denote the dist. of X_1 by π_1

$$\forall j \in \{1,\dots,k\} \quad \pi_1(j) = Pr[X_1 = j] = \sum_{i=1}^{k} Pr[X_0 = i, X_1 = j]$$

$$= \sum_{i=1}^{k} Pr[X_0 = i]\, Pr[X_1 = j \mid X_0 = i]$$

$$= \sum_{i=1}^{k} \pi_0(i) \cdot P(i,j)$$

If we think of π_0, π_1 as row vectors, P as a matrix

then $\pi_1 = \pi_0 \cdot P$.

What's the Distribution of X_n?

$$\Pr[X_0 = i] = \pi_0(i) \quad \forall i \in \{1,2,\dots,k\}$$

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

Let's denote the dist. of X_n by π_n

$$\forall j \in \{1,\dots,k\} \quad \pi_n(j) = \Pr[X_n = j] = \sum_{i=1}^{k} \Pr[X_{n-1} = i, X_n = j]$$

$$= \sum_{i=1}^{k} \Pr[X_{n-1} = i] \cdot \Pr[X_n = j \mid X_{n-1} = i]$$

$$= \sum_{i=1}^{k} \pi_{n-1}(i) \cdot P(i,j)$$

In vector-matrix form: $\pi_n = \pi_{n-1} \cdot P$.

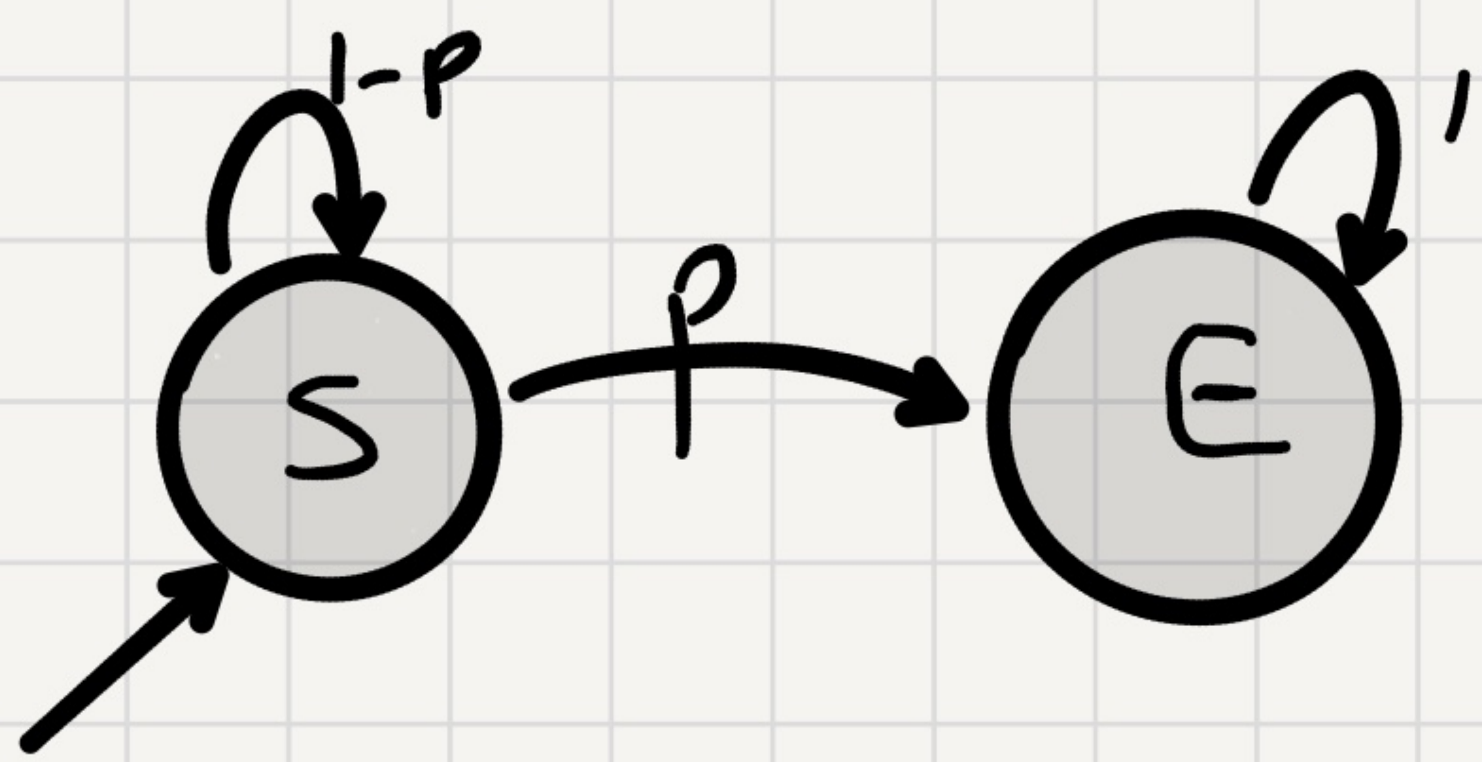
Thus, $\pi_n = \pi_{n-1} \cdot P = \pi_{n-2} \cdot P \cdot P = \dots = \pi_0 \cdot \underbrace{P \cdot P \cdots P}_{n \text{ times}} = \pi_0 \cdot P^n$.

Hitting Time – Example

Let's flip a coin w. heads prob. p until we get H.

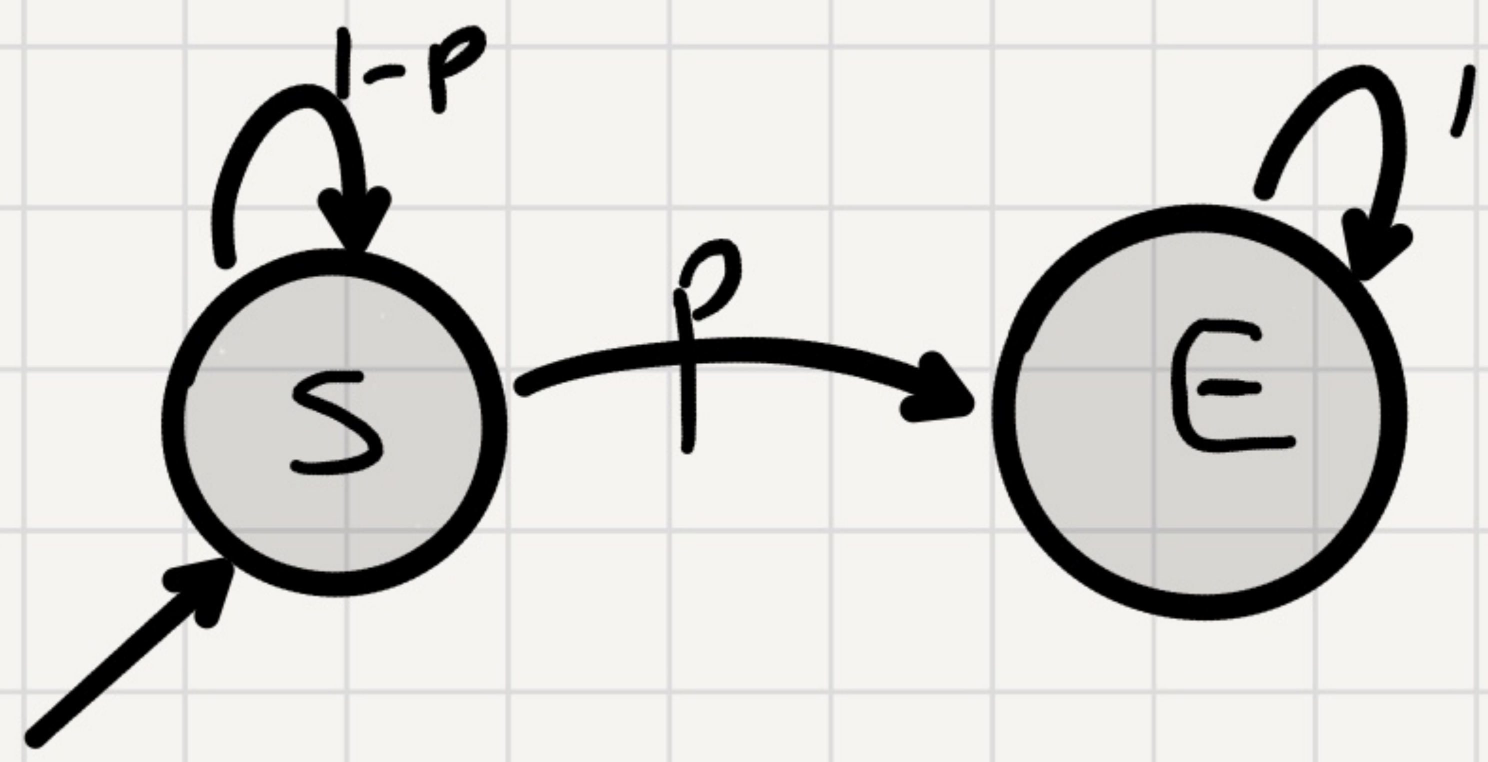
How many flips on average?

Let's model this as a Markov Chain.



Let $\beta(S)$ = average time until a MC starting from S would reach E.

Hitting Time - Example



Let $\beta(S)$ = average time until a MC starting from S would reach E.

What's true:

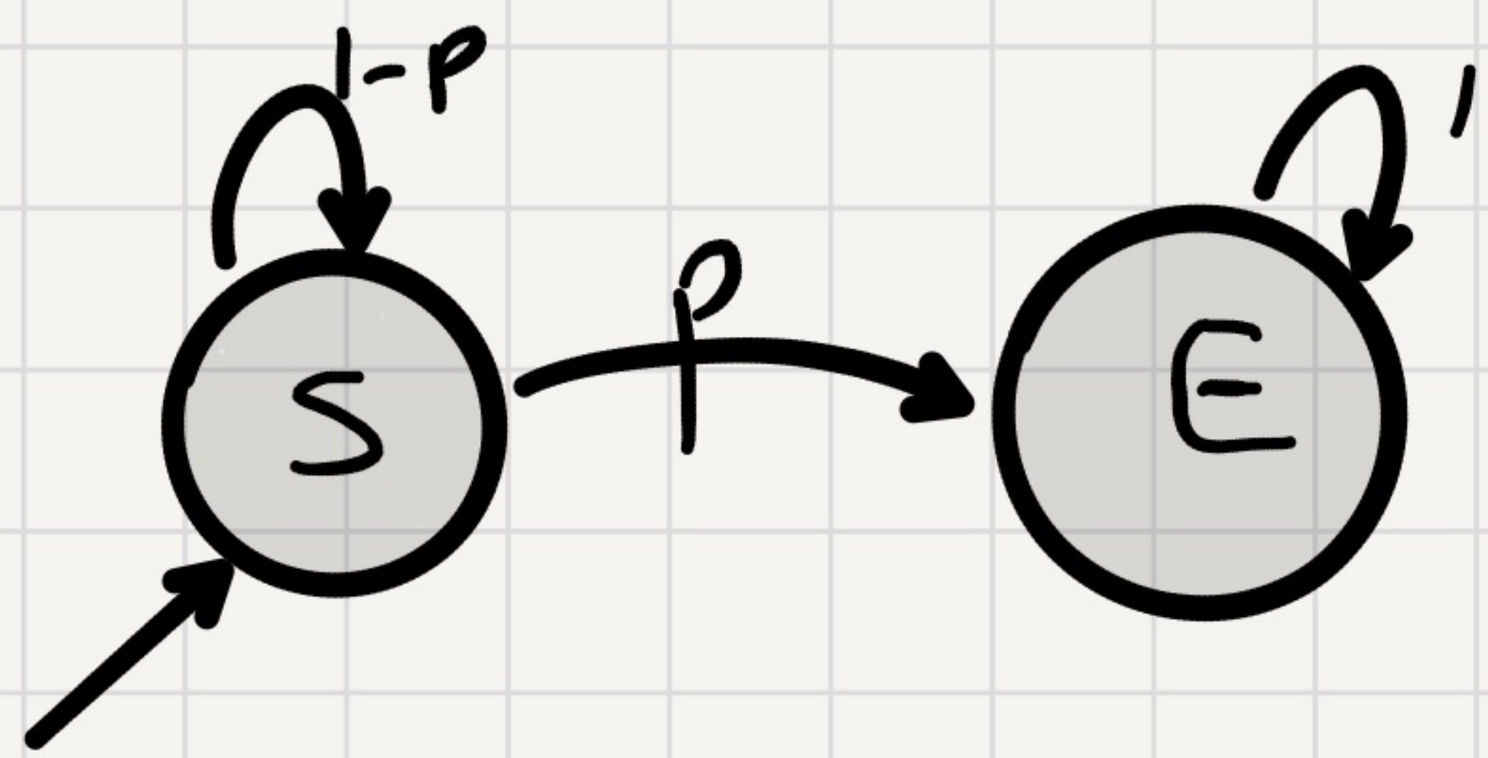
A. $\beta(S) \geq 1$.

B. From S we stay in S w.p. $1-p$.

C. From S we go to E w.p. p.

D. $\beta(S) = 1 + (1-p)\cdot\beta(S) + p\cdot 0$.

Hitting Time – Example



Let $\beta(S)$ = average time until a MC starting from S would reach E.

$$\boxed{\beta(S) = 1 + (1-p)\cdot\beta(S) + p\cdot 0.}$$

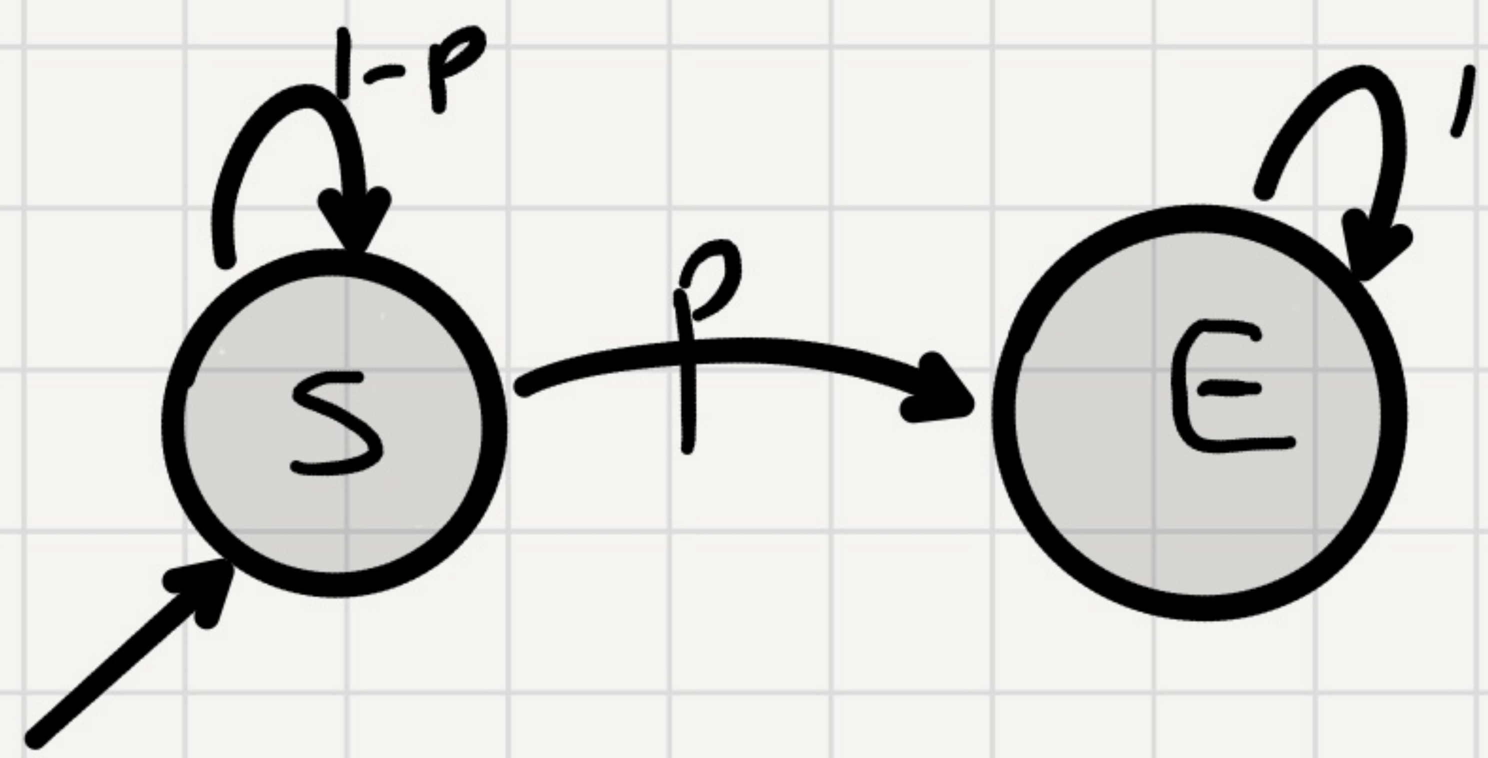
Justification:

We start at S and spend one step there.

Then, w.p. p we move to E and stop.

w.p. $1-p$ we stay in S, then the expected remaining time to hit E is again $\beta(S)$.

Hitting Time – Example



Let $\beta(S)$ = average time until a MC starting from S would reach E.

$$\boxed{\beta(S) = 1 + (1-p)\cdot\beta(S) + p\cdot 0.}$$

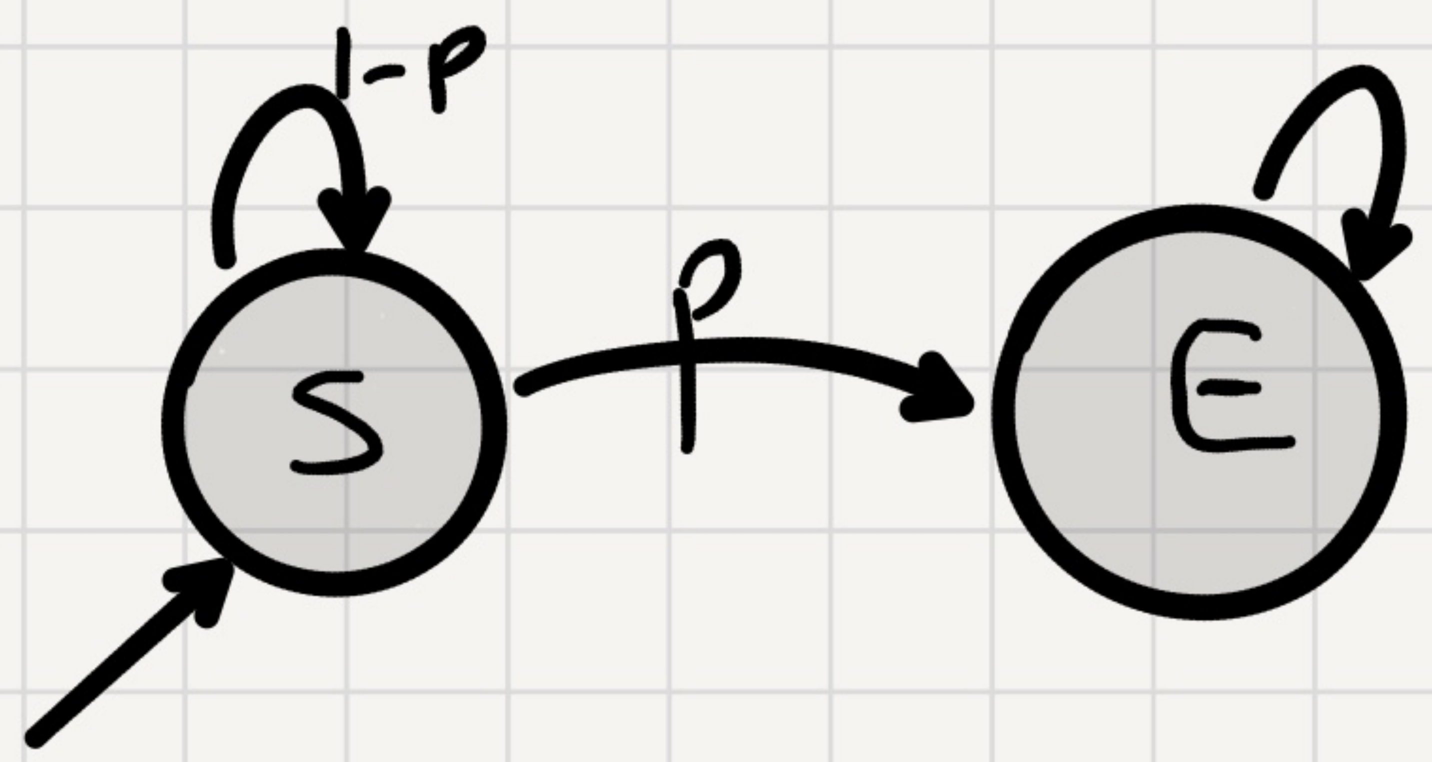
A bit more formal:

Let N be the r.v. capturing the number of steps until we hit E.

$$\beta(S) = \mathbb{E}[N] = \Pr[H]\cdot\mathbb{E}[N|H] + \Pr[T]\cdot\mathbb{E}[N|T]$$

$$= p\cdot 1 + (1-p)\cdot(1+\beta(S)) = 1 + (1-p)\beta(S).$$

Hitting Time – Example



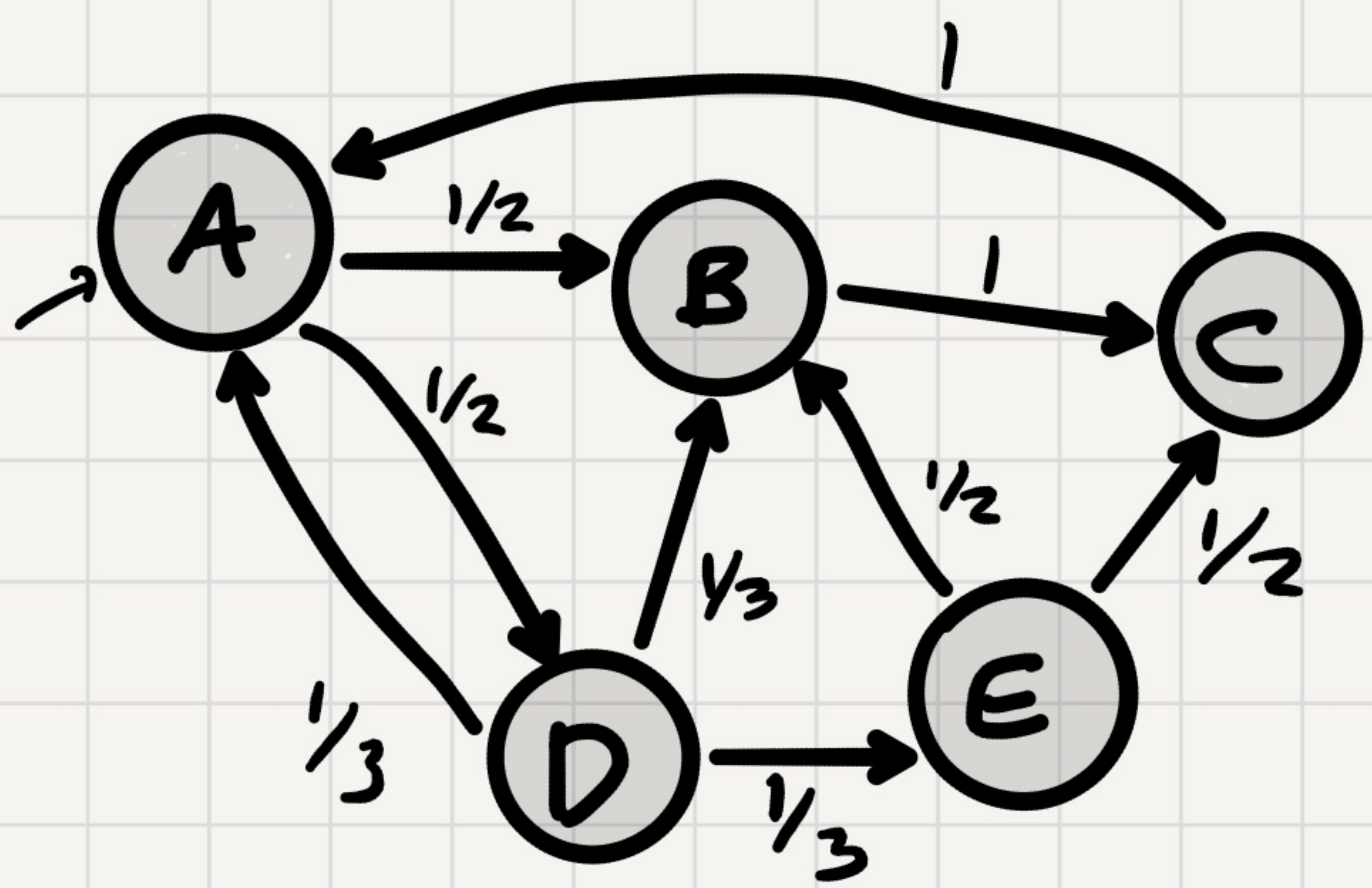
Let $\beta(S)$ = average time until a MC starting from S would reach E.

$$\boxed{\beta(S) = 1 + (1-p)\cdot\beta(S) + p\cdot 0.}$$

Solving: $p\cdot\beta(S) = 1 \implies \beta(S) = 1/p.$

We calculated the mean of a geometric R.V. without any infinite sums...

Hitting Time - Example 2



for $i \in \{A, B, C, D, E\}$ $\quad \beta(i) =$ average time until a MC starting at i would reach E.

$$\left.\begin{aligned}
\beta(E) &= 0 \\
\beta(A) &= 1 + \tfrac{1}{2}\beta(B) + \tfrac{1}{2}\beta(D) \\
\beta(B) &= 1 + \beta(C) \\
\beta(C) &= 1 + \beta(A) \\
\beta(D) &= 1 + \tfrac{1}{3}\beta(A) + \tfrac{1}{3}\beta(B)
\end{aligned}\right\}$$

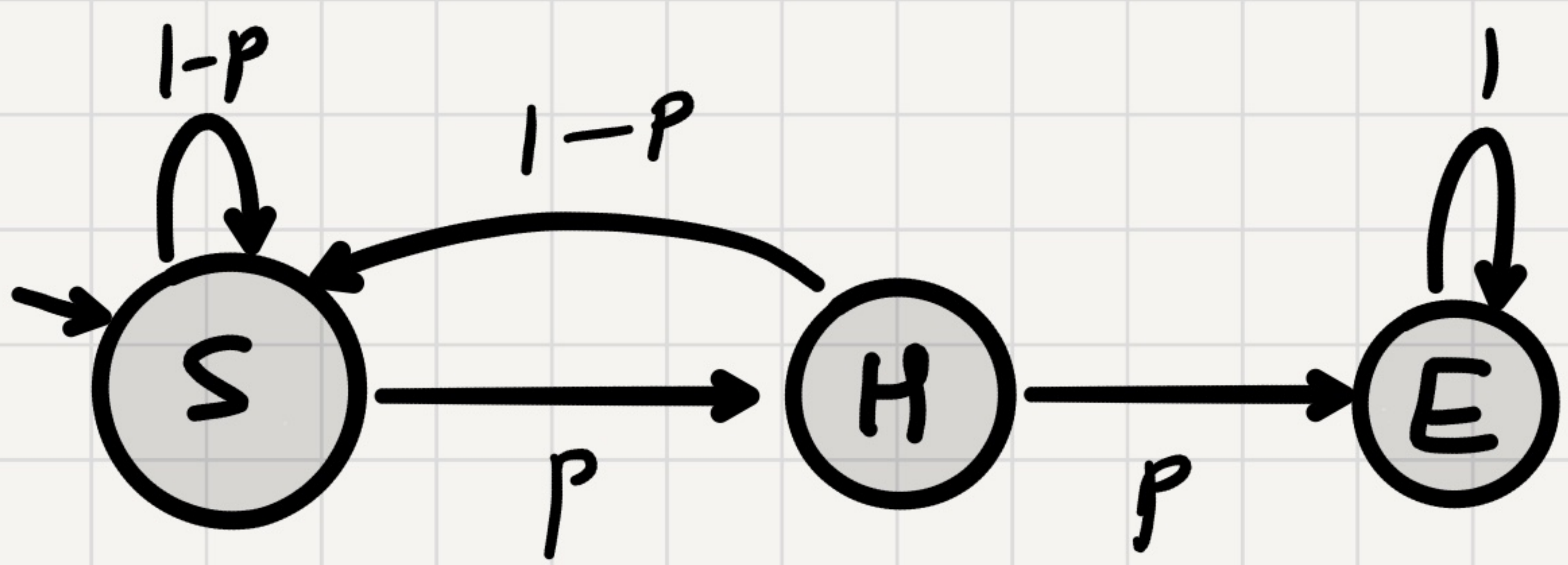
Solve the system of linear equations:

$\beta(A) = 17$
$\beta(B) = 19$
$\beta(C) = 18$
$\beta(D) = 13$

Hitting Time - Example 3

You flip a coin w. heads prob. p until you get two consecutive H.

How many flips on average?



$\beta(E) = 0$

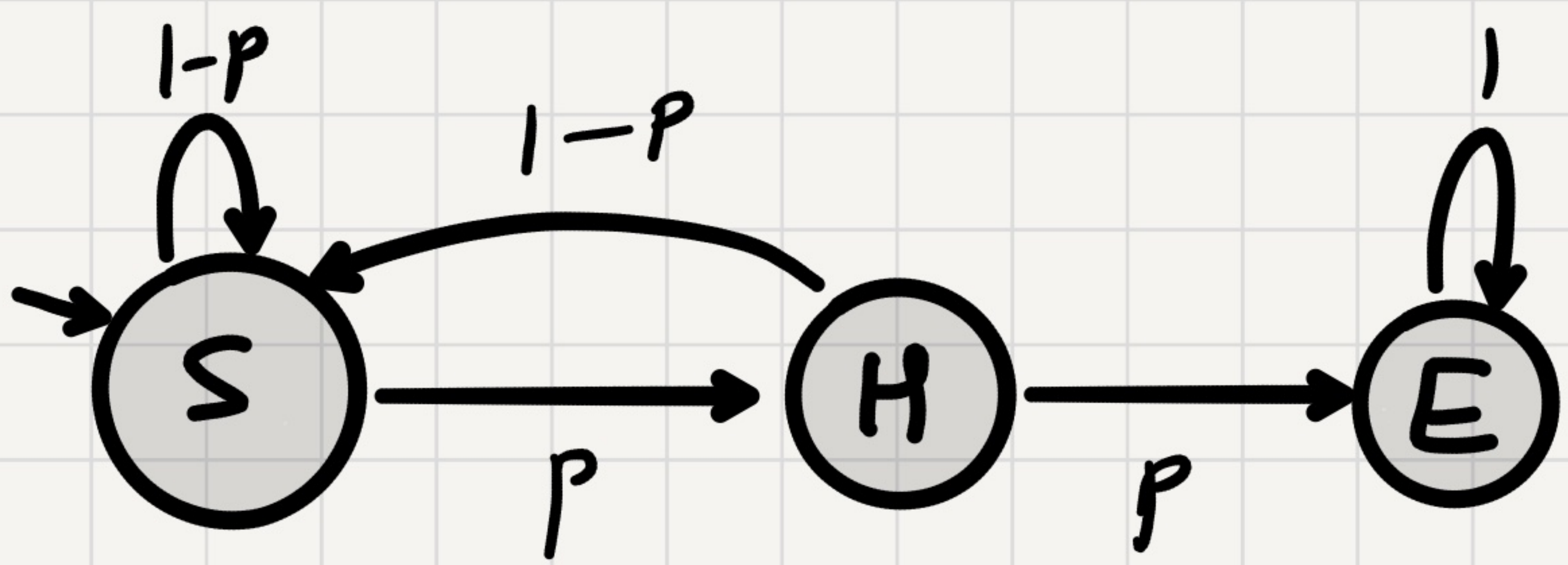
$\beta(H) = 1 + p \cdot \beta(E) + (1-p) \cdot \beta(S)$

$\beta(S) = 1 + p \cdot \beta(H) + (1-p) \cdot \beta(S).$

Hitting Time - Example 3

You flip a coin w. heads prob. p until you get two consecutive H.

How many flips on average? $\boxed{\frac{1+p}{p^2}}$



$\beta(E) = 0$

$\beta(H) = 1 + (1-p)\cdot\beta(s) + \cancel{p\cdot\beta(E)}$

$\beta(S) = 1 + (1-p)\beta(s) + p\cdot\beta(H)$

$= 1 + (1-p)\beta(s) + p\cdot(1 + (1-p)\beta(s))$

$\therefore\ \beta(s) = \frac{1+p}{p^2}$.

If $p = 1/2$ $\beta(s) = \frac{3/2}{1/4} = 6$.

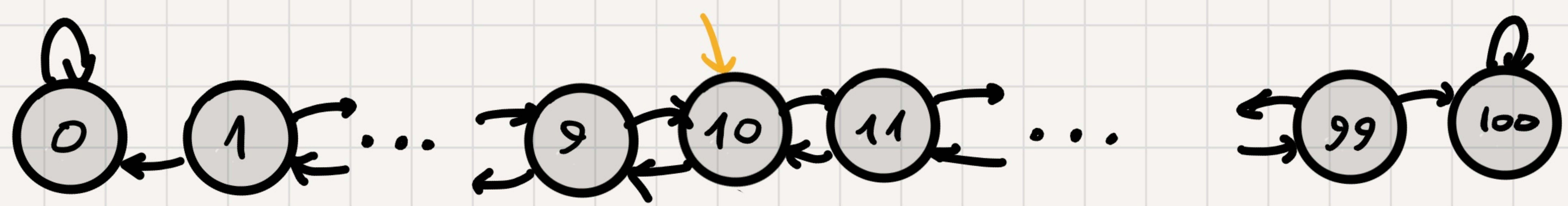
Here Before There

You go to a nice casino. You start with 10 \$

Every round, you win 1 \$ w.p. 50%

lose 1 \$ w.p. 50%

What's the probability you'll reach 100 \$ before 0 \$?

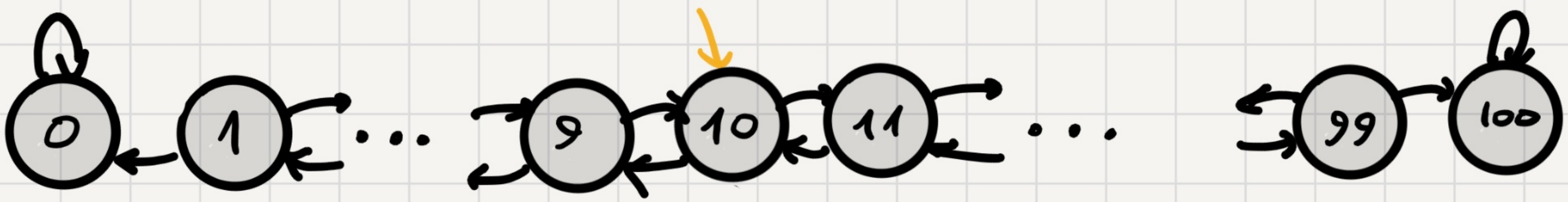


For $i \in \{0, 1, \dots, 100\}$

let $\alpha(i) =$ Prob. of reaching 100 before 0 starting from i.

Here Before There

What's the probability you'll reach 100\$ before 0\$?



For $i = 0, 1, \ldots, 100$,

$\alpha(i)$ = Prob. of reaching 100 before 0 starting from i.

What's true:

A. $\alpha(0) = 0$

B. $\alpha(0) = 1$

C. $\alpha(100) = 1$

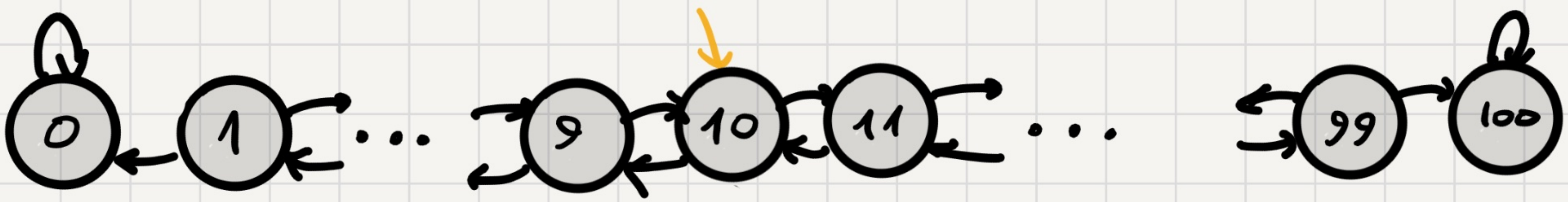
D. $\alpha(i) = 1 + \frac{1}{2}\alpha(i-1) + \frac{1}{2}\alpha(i+1)$ for $1 \le i \le 99$.

E. $\alpha(i) = \frac{1}{2}\alpha(i-1) + \frac{1}{2}\alpha(i+1)$ for $1 \le i \le 99$.

Here Before There

What's the probability you'll reach 100\$ before 0\$?

0.1.



For $i=0,1,\dots,100$,

$\alpha(i) =$ Prob. of reaching 100 before 0 starting from i.

$\alpha(0)=0 \qquad \alpha(100)=1$

For $i=1,\dots,99 \qquad \alpha(i)=\frac{1}{2}\alpha(i-1)+\frac{1}{2}\alpha(i+1)$

$\alpha(i) =$ avg of $\alpha(i-1), \alpha(i+1)$.

What's the solution? $\qquad \alpha(i)=\frac{i}{100}$.

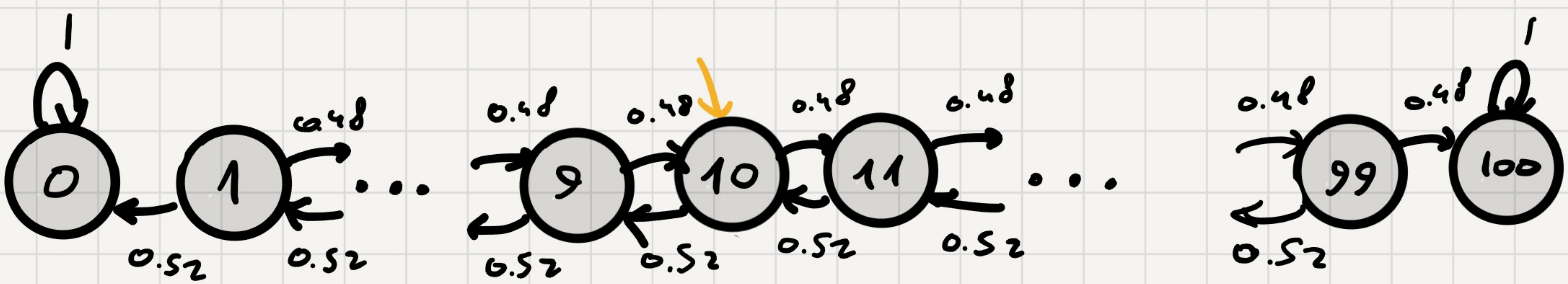
Here Before There

You go to a real casino. You start with 10 \$

Every round, you win 1 \$ w.p. 48 %

lose 1 \$ w.p. 52 %

What's the probability you'll reach 100 \$ before 0 \$?

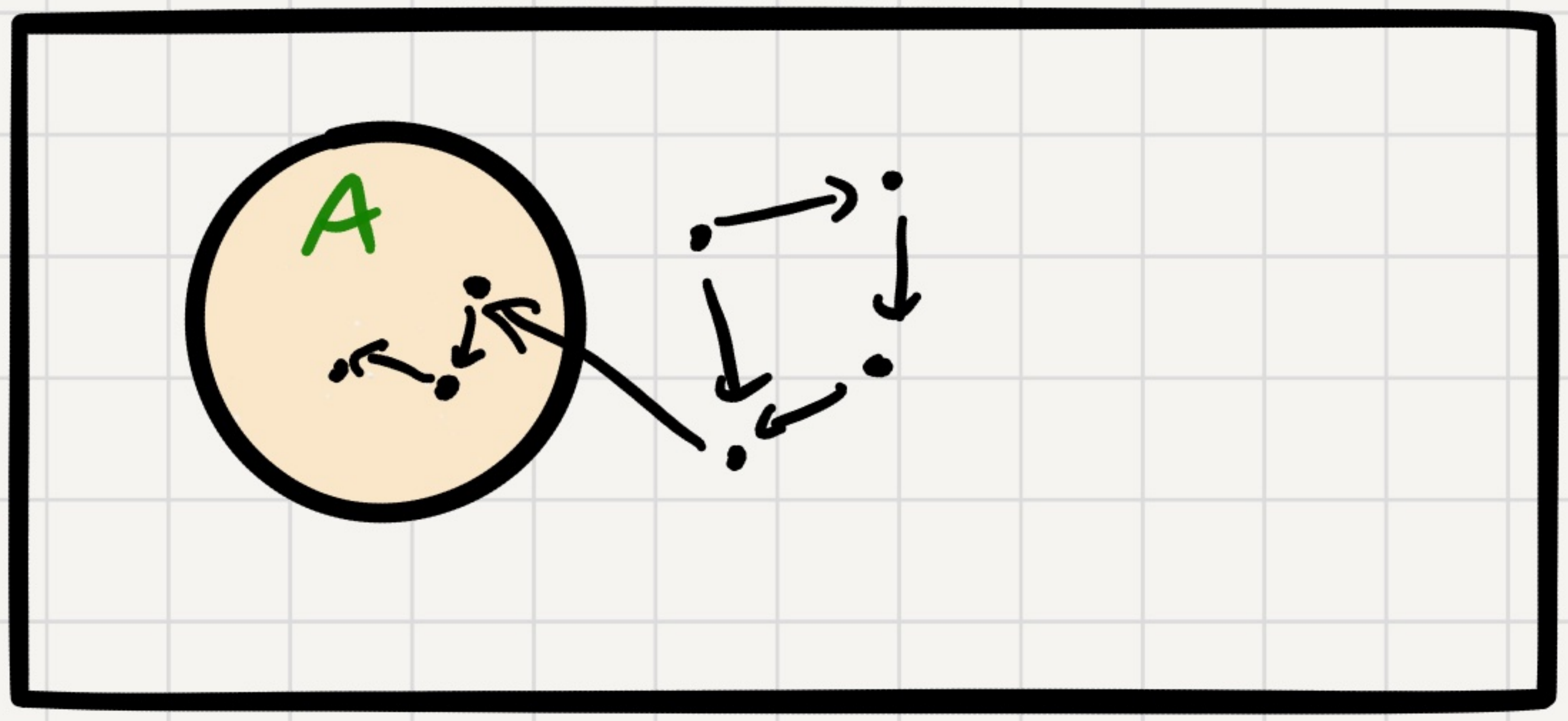


$\alpha(0) = 0 \qquad \alpha(100) = 1$

$$\alpha(i) = 0.48 \cdot \alpha(i+1) + 0.52 \cdot \alpha(i-1)$$

Solution: $\alpha(i) = \dfrac{\rho^i - 1}{\rho^{100} - 1}$ where $\rho = \dfrac{0.52}{0.48}$. $\alpha(10) \approx \dfrac{1}{2{,}440}$

First Step Equations



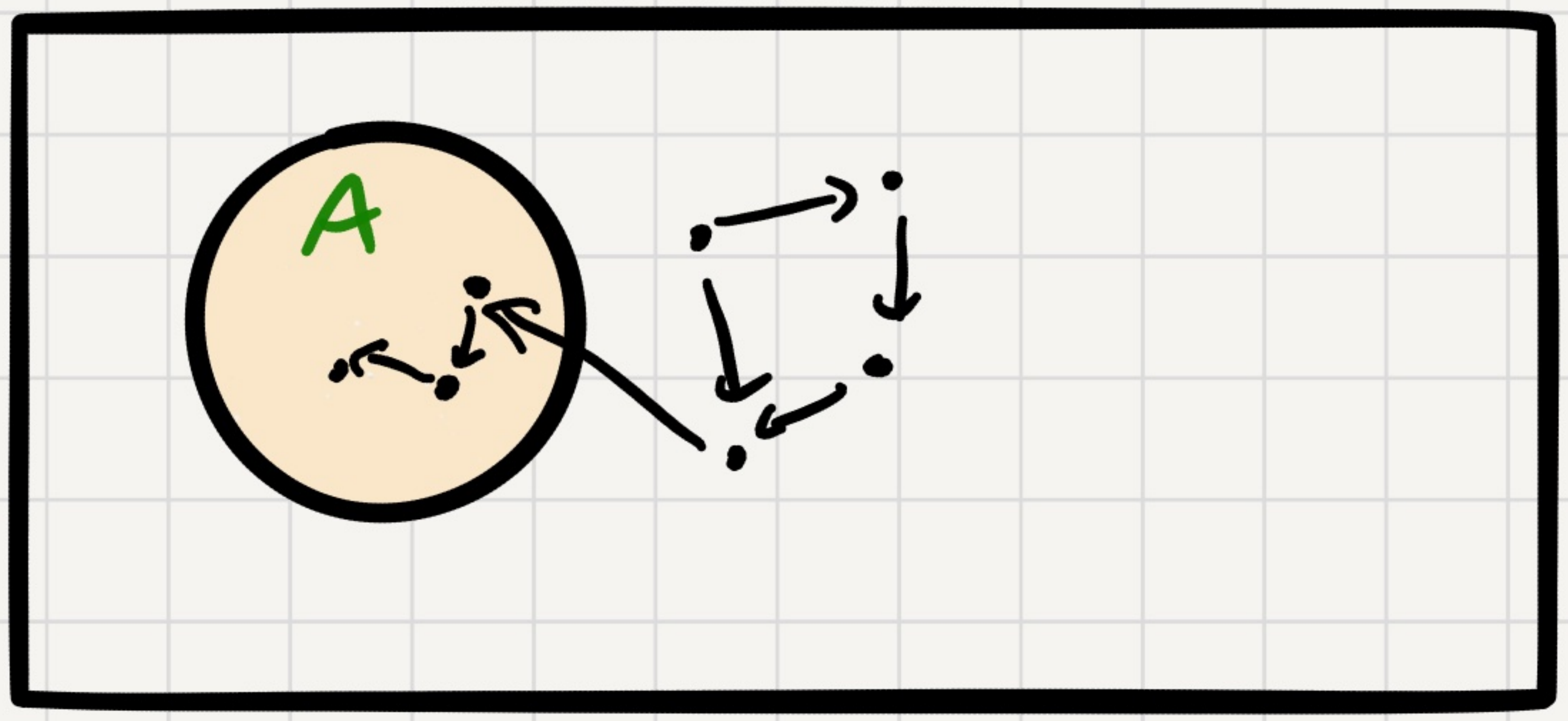
Let $\{X_n\}_{n=0}^{\infty}$ be a MC on $\mathcal{X}$ $\quad A \subseteq \mathcal{X}$

$\beta(i)$ = expected time to reach A starting from i.

Formally: Define the r.v. $T_A = \min\{n \geq 0 : X_n \in A\}$

$$\beta(i) = \mathbb{E}[T_A \mid X_0 = i]$$

First Step Equations



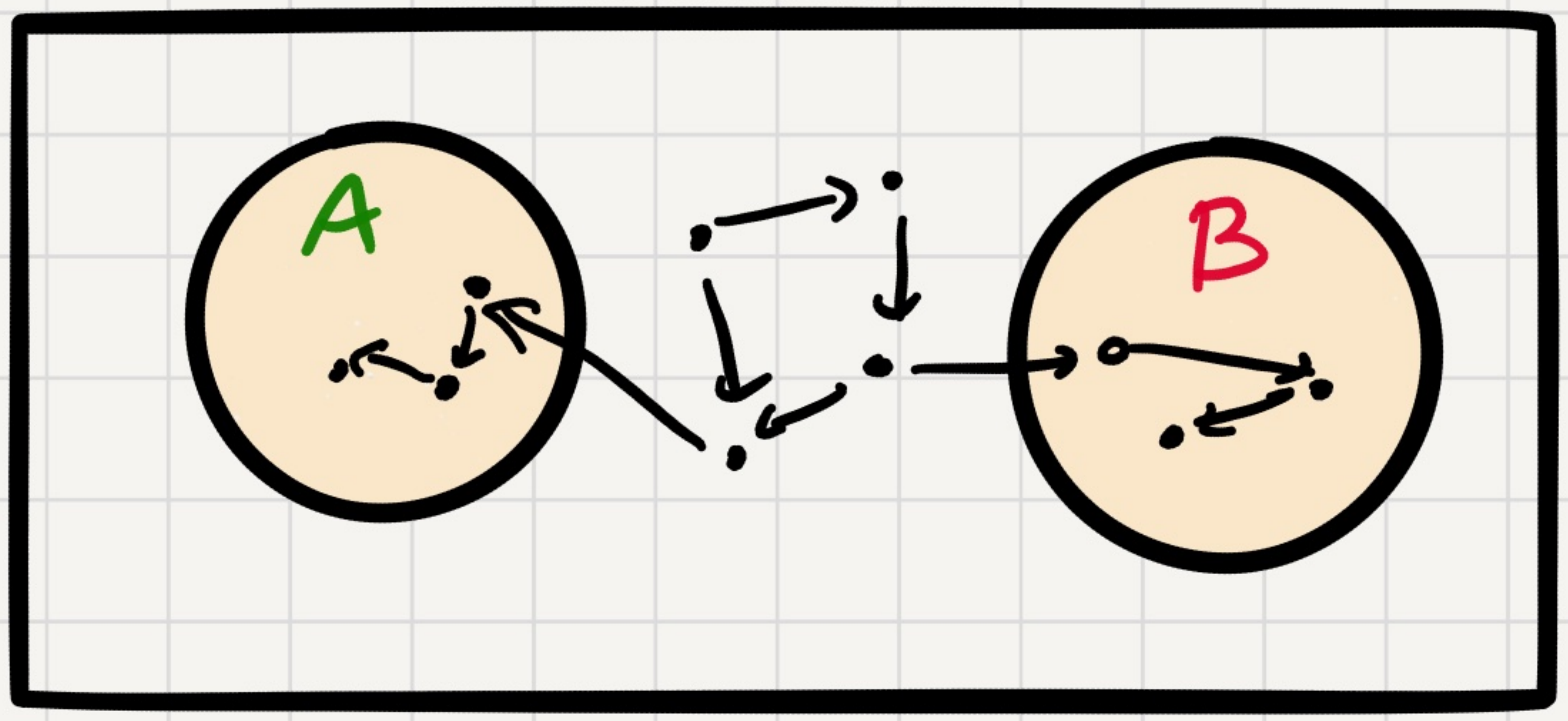
Let $\{X_n\}_{n=0}^{\infty}$ be a MC on $\mathcal{X}$ $A \subseteq \mathcal{X}$

$\beta(i)$ = expected time to reach A starting from i.

$\beta(i) = 0$ for $i \in A$

$\beta(i) = 1 + \sum_j P(i,j)\,\beta(j)$ for $i \notin A$.

First Step Equations



Let $\{X_n\}_{n=0}^{\infty}$ be a MC on $\mathcal{X}$ $A, B \subseteq \mathcal{X}$

A & B disjoint.

$\alpha(i) = P_r$[reaching A before B, starting from i]

$$\begin{cases} \alpha(i) = 0 & \text{for } i \in B \\ \alpha(i) = 1 & \text{for } i \in A \\ \alpha(i) = \sum_j P(i,j) \cdot \alpha(j). & \text{for } i \notin A \cup B. \end{cases}$$

Distribution of X_n

Recall if a MC starts with distribution π_0

and has transition matrix P

then the dist of X_1 is $\pi_0 \cdot P$.

and $\forall n$ the dist of X_n is $\pi_0 \cdot P^n$.

Definition:

A distribution π is stationary if $\pi = \pi \cdot P$.

Suppose π_0 is a stationary dist.

What's π_1 ?

What's π_n ?

Distribution of X_n

Recall if a MC starts with distribution π_0

and has transition matrix P

then the dist of X_1 is $\pi_0 \cdot P$.

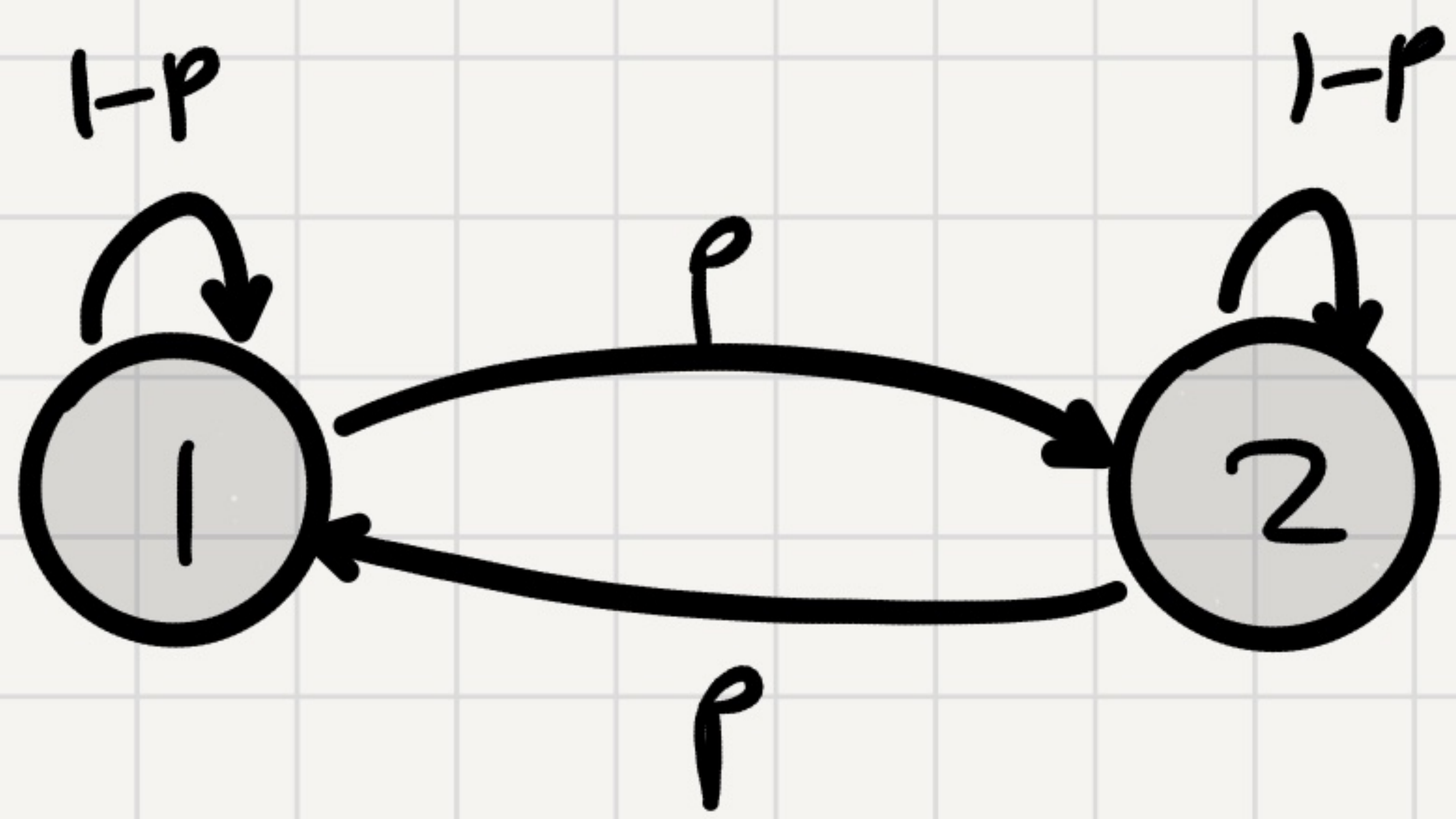
and $\forall n$ the dist of X_n is $\pi_0 \cdot P^n$.

Definition:

A distribution π is stationary if $\pi = \pi \cdot P$.

Theorem: π_0 is stationary iff $\forall n \geq 0$: $\pi_n = \pi_0$.

Stationary Distribution - Example



π is stationary iff $\begin{pmatrix}\pi(1) & \pi(2)\end{pmatrix} = \begin{pmatrix}\pi(1) & \pi(2)\end{pmatrix} \cdot \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}$

$$\pi(1) = \pi(1)\cdot(1-p) + \pi(2)\cdot p \iff \pi(1) = \pi(2)$$

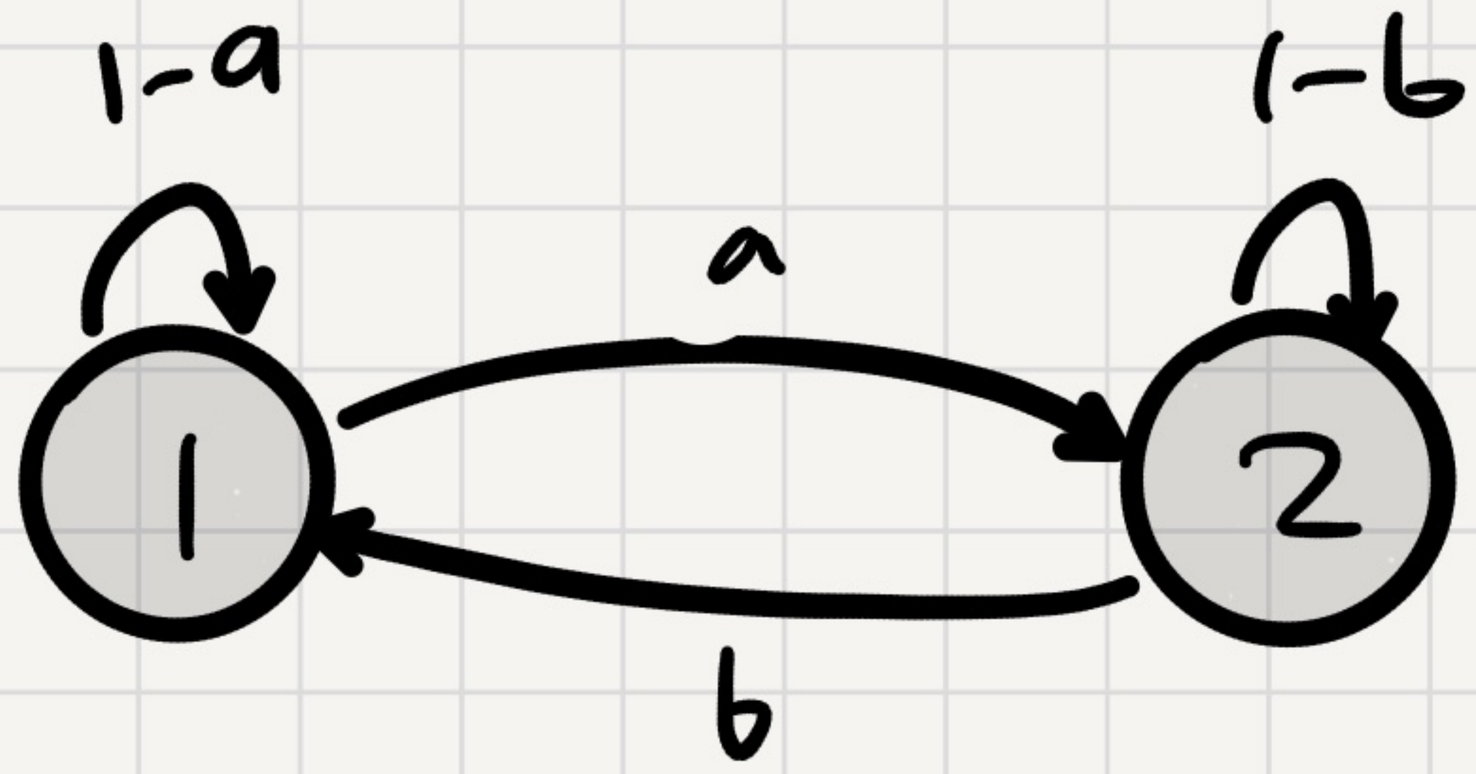
$$\pi(2) = \pi(1)\cdot p + \pi(2)\cdot(1-p) \iff \pi(1) = \pi(2)$$

The two equations are redundant.

Is there a unique solution?

Yes, if we also use $\pi(1) + \pi(2) = 1$.

Stationary Distribution - Example 2



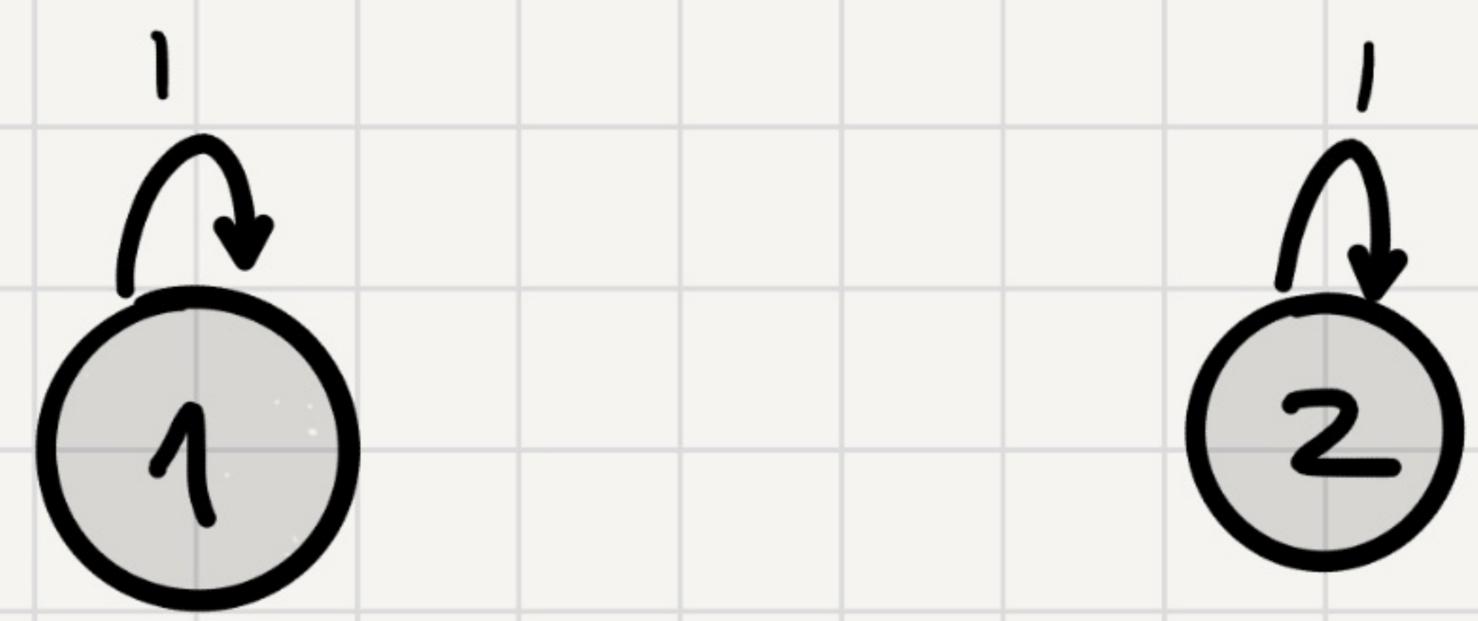
π is stationary iff $\begin{pmatrix}\pi(1) & \pi(2)\end{pmatrix} = \begin{pmatrix}\pi(1) & \pi(2)\end{pmatrix} \cdot \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$

and $\pi(1) + \pi(2) = 1$

$$\pi(1) = \pi(1)\cdot(1-a) + \pi(2)\cdot b \iff a\cdot\pi(1) = \pi(2)\cdot b$$

Unique Solution: $\pi = \left[\frac{b}{a+b} , \frac{a}{a+b} \right]$

Stationary Distributions - Example 3



Which distributions are stationary?

all of them.

$$\forall \pi \qquad \pi = \pi \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

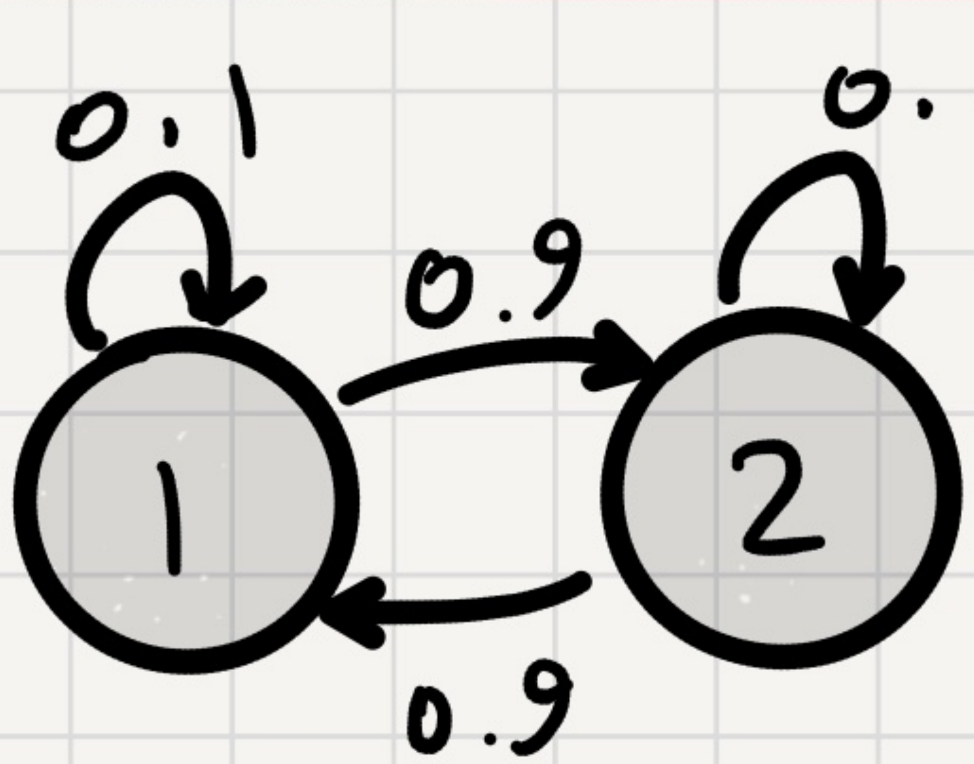
Q: Which Markov Chains have a unique stationary distribution?

Irreducible Markov Chains

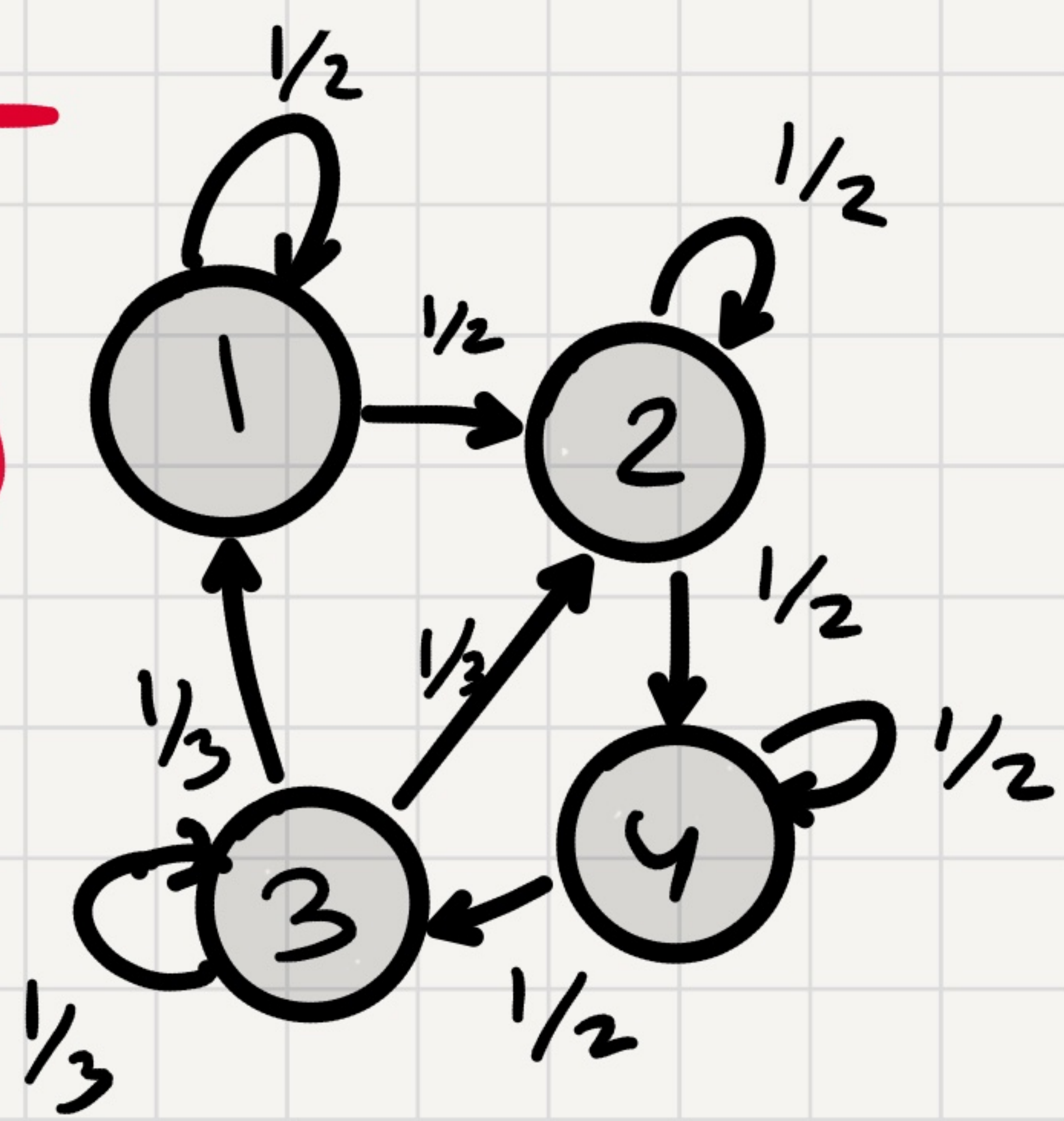
A MC is irreducible if you can go from every state i to every state j (possibly in multiple steps).

Which MC are irreducible?

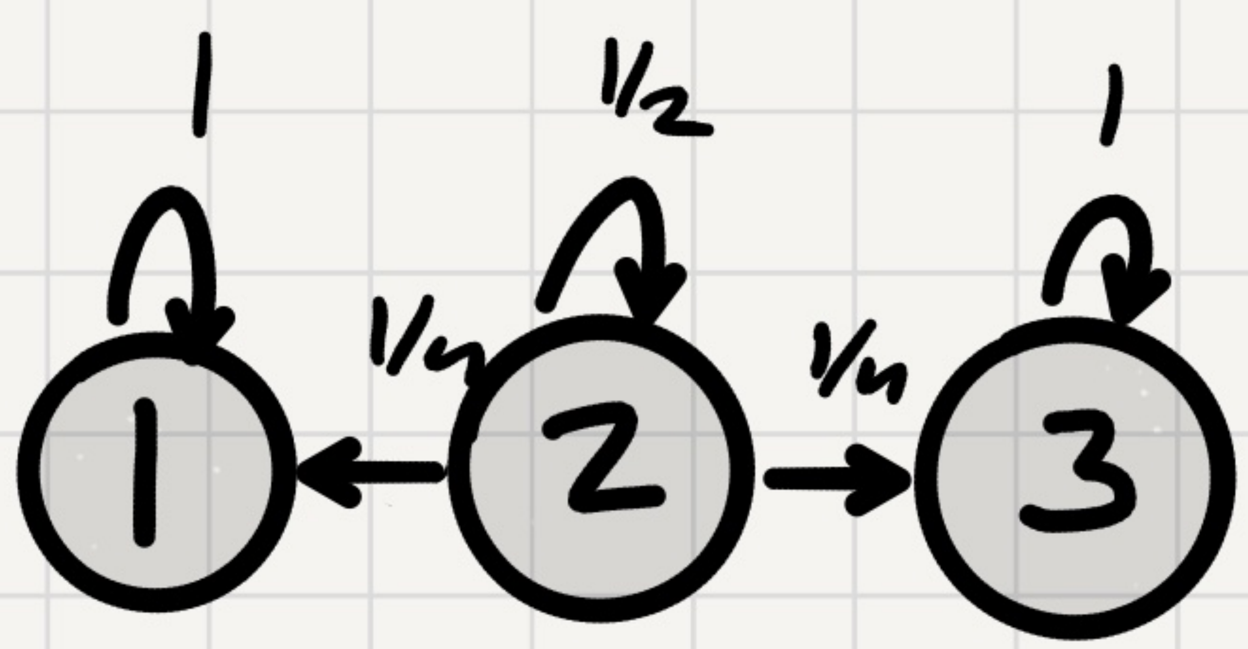
(A)



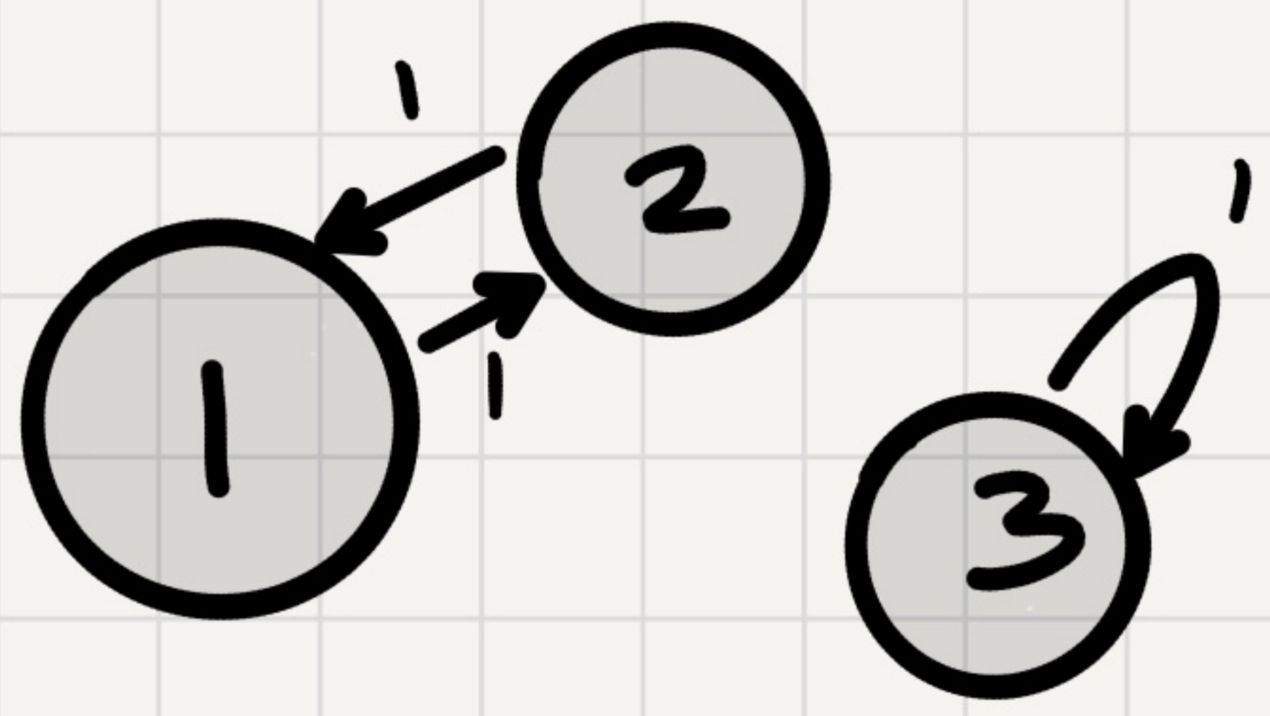
(B)



(C)



(D)



Theorem:

Any finite irreducible MC has one and only one stationary distribution.

Theorem 2: (Long Term Fraction of Time in States)

If $(X_n)_{n=0}^{\infty}$ is an irreducible MC on $\{1, \dots, k\}$ with stationary distribution π.

Then, for any start dist. π_0, for all $i \in \{1, \dots k\}$

$$\frac{1}{n} \sum_{m=0}^{n-1} 1\{X_m = i\} \xrightarrow[n \to \infty]{} \pi(i).$$