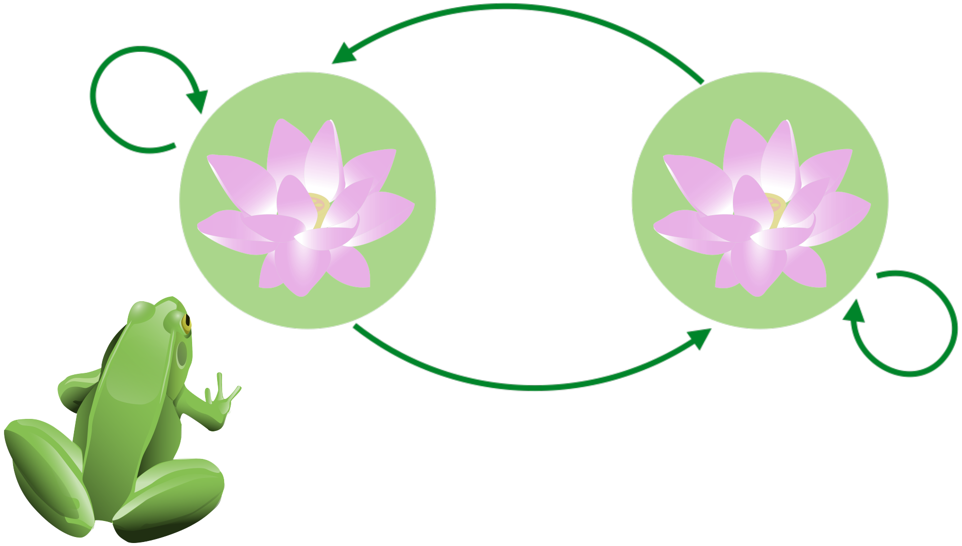
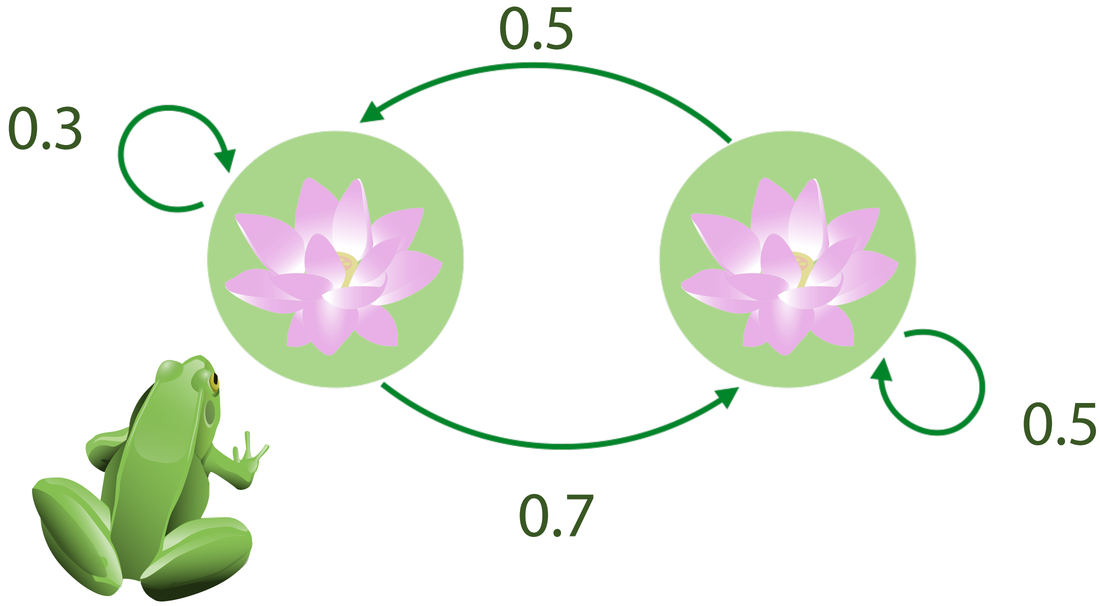
Markov Chain Monte Carlo (MCMC)

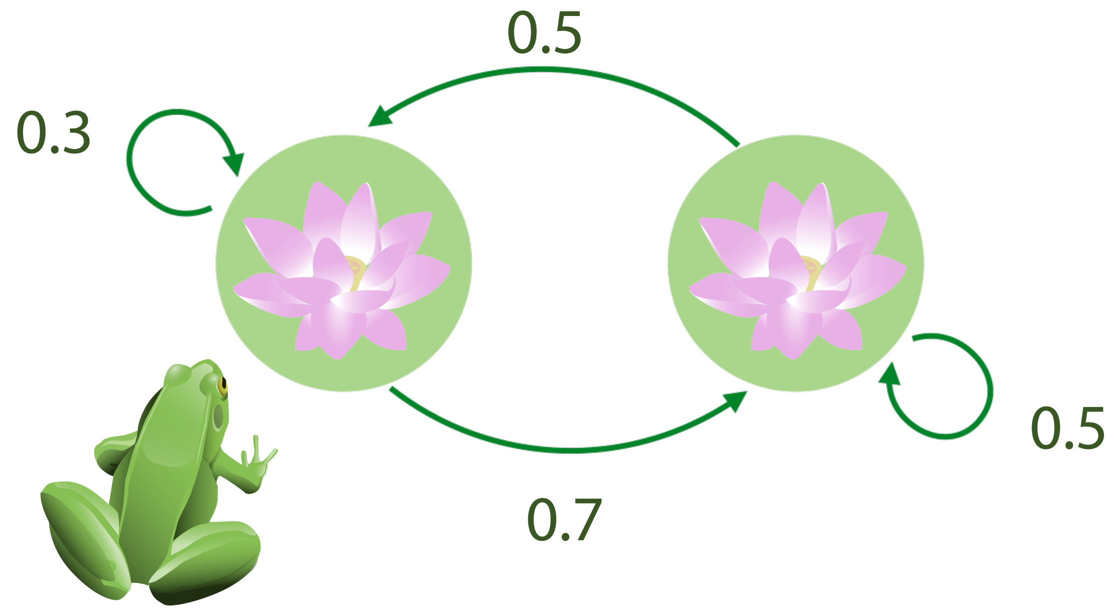
Markov Chains



Markov Chains



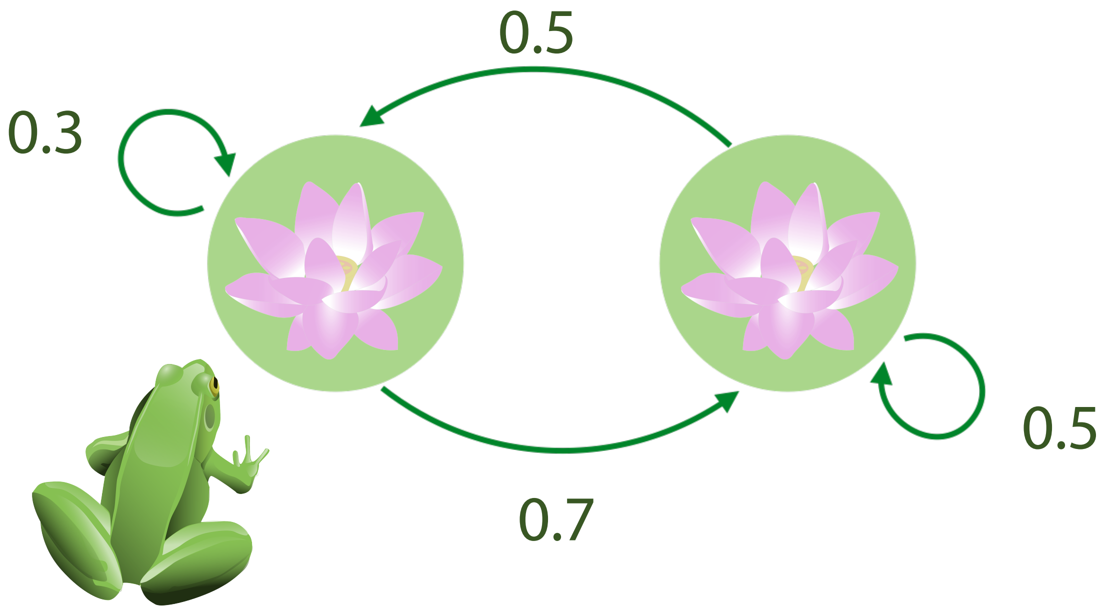
Markov Chains



$$T(\mathrm{L} \to \mathrm{L}) = 0.3 \qquad T(\mathrm{R} \to \mathrm{L}) = 0.5$$

$$T(\mathrm{L} \to \mathrm{R}) = 0.7 \qquad T(\mathrm{R} \to \mathrm{R}) = 0.5$$

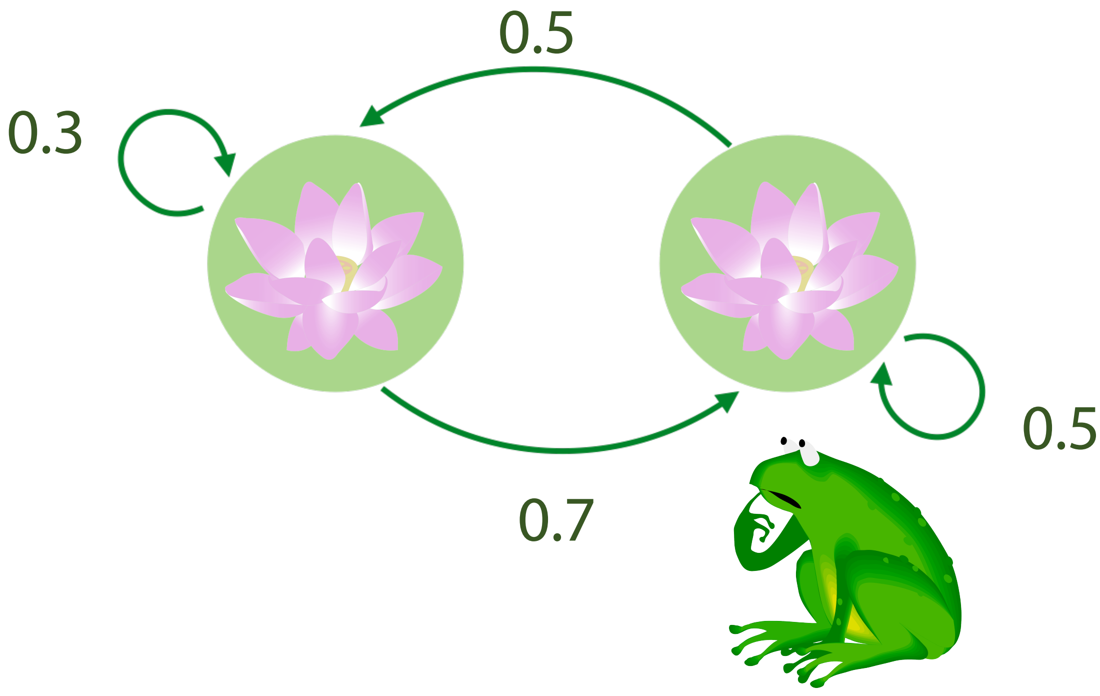
Markov Chains



Timestamp: **1**

Log: **L**

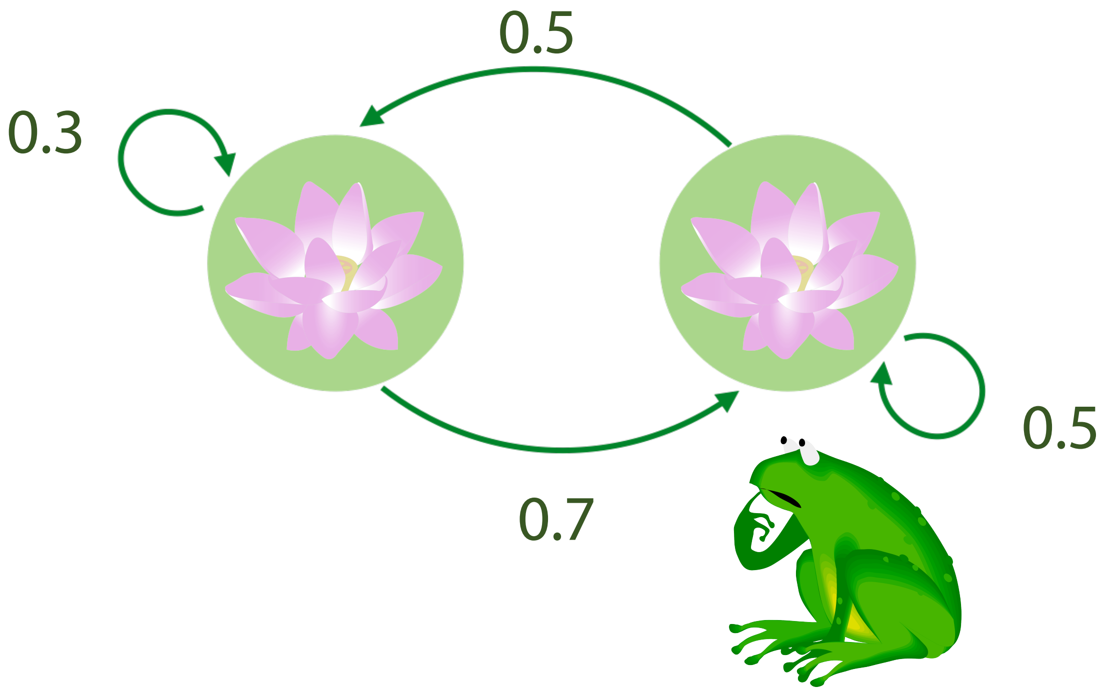
Markov Chains



Timestamp: **2**

Log: **L R**

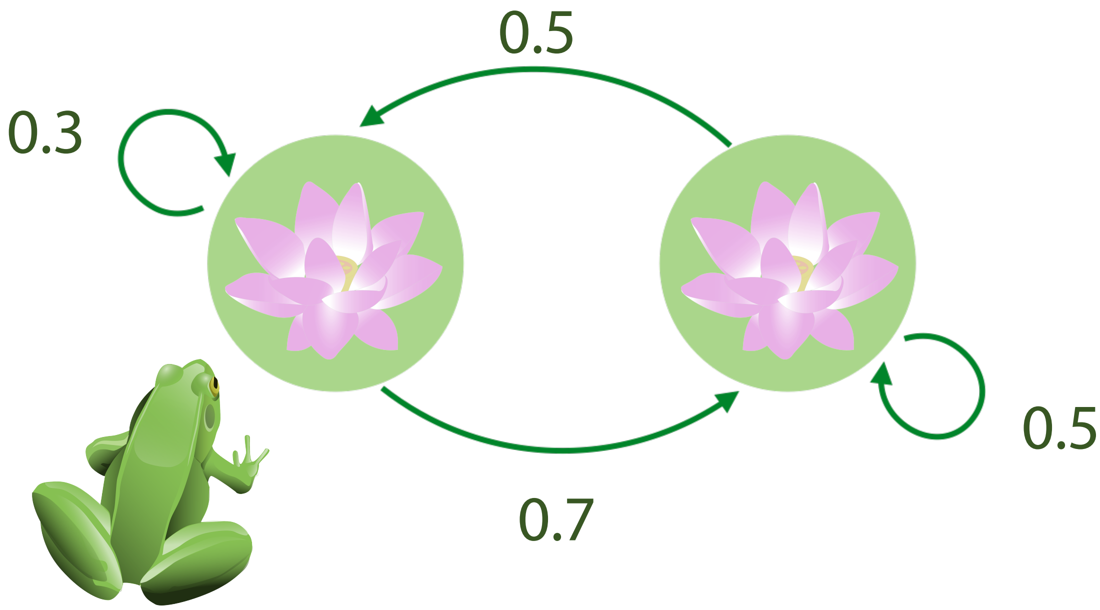
Markov Chains



Timestamp: **3**

Log: **L R R**

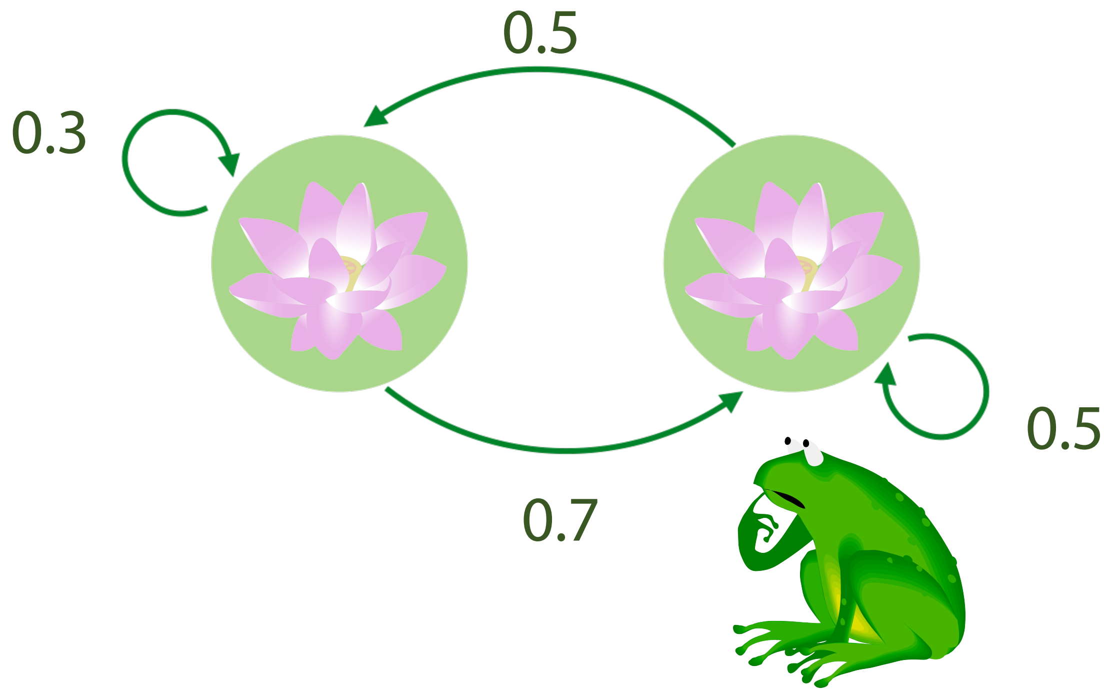
Markov Chains



Timestamp: **4**

Log: **L R R L**

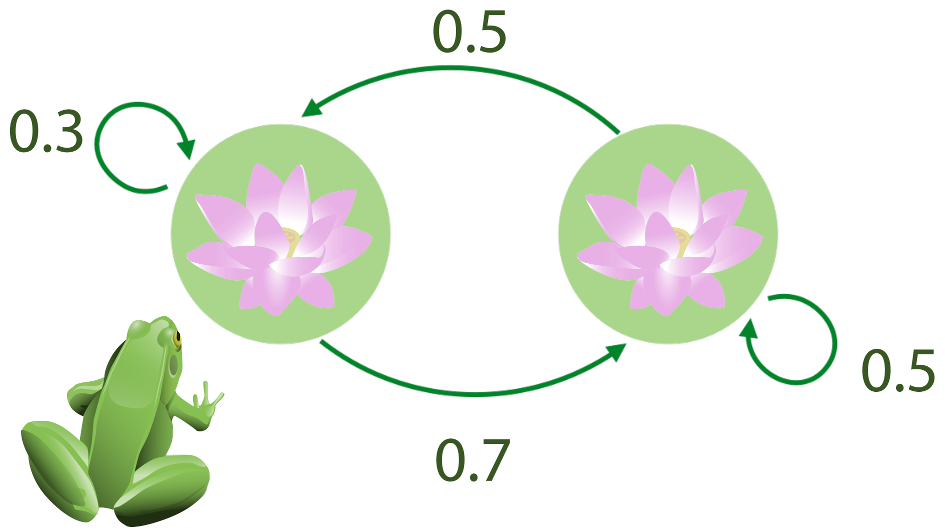
Markov Chains



Timestamp: **5**

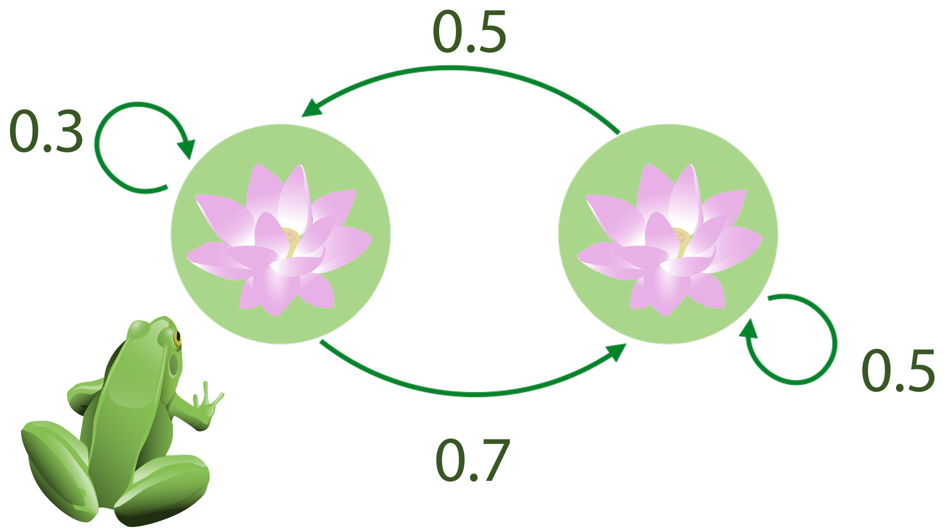
Log: **L R R L R**

Markov Chains



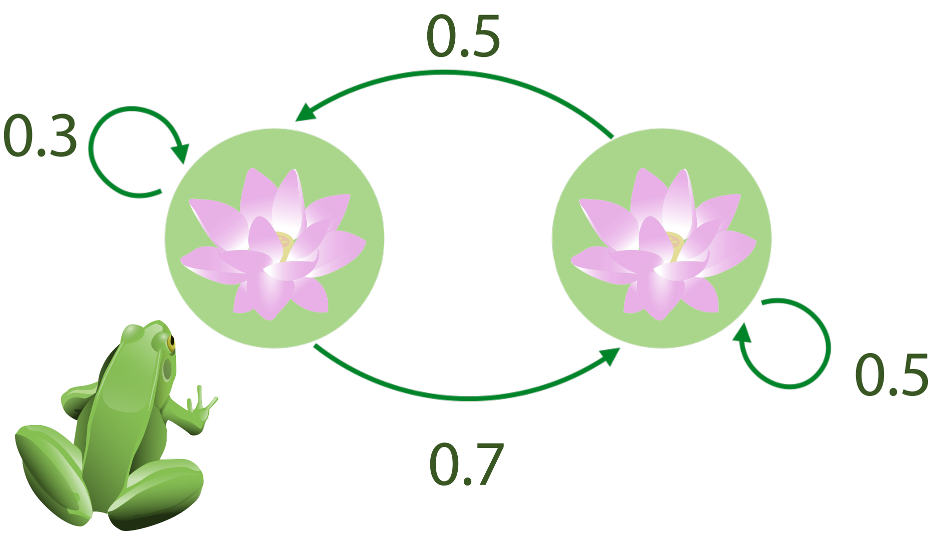
	p(Left)	p(Right)
x^1	1	0

Markov Chains



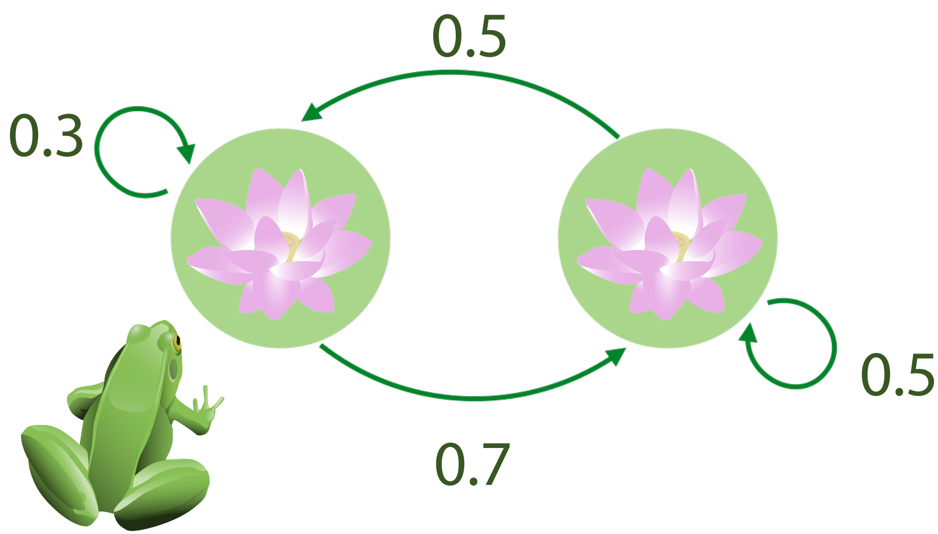
	p(Left)	p(Right)
x^1	1	0
x^2	0.3	0.7

Markov Chains



	p(Left)	p(Right)
x^1	1	0
x^2	0.3	0.7
x^3		

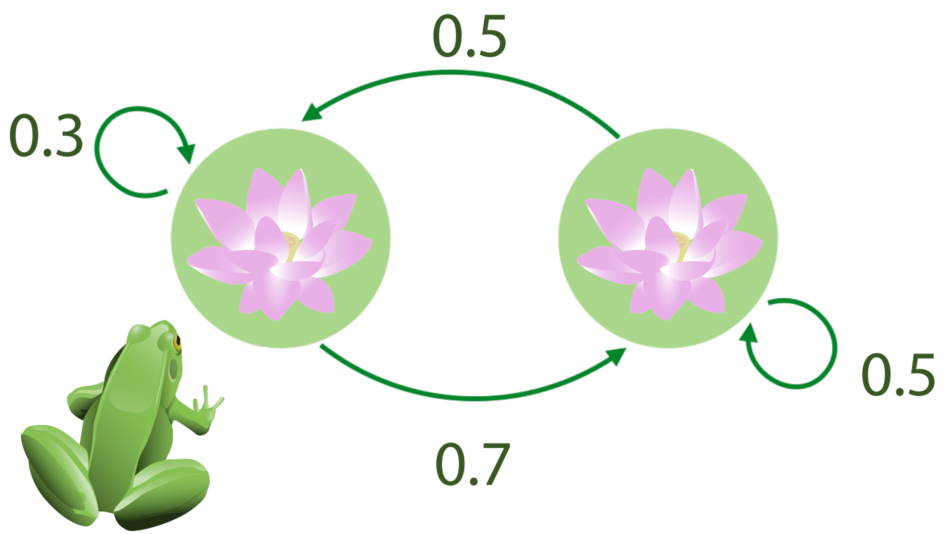
Markov Chains



	p(Left)	p(Right)
x^1	1	0
x^2	0.3	0.7
x^3		

$$p(x^3) = p(x^3 \mid x^2 = \mathrm{L})p(x^2 = \mathrm{L}) + p(x^3 \mid x^2 = \mathrm{R})p(x^2 = \mathrm{R})$$

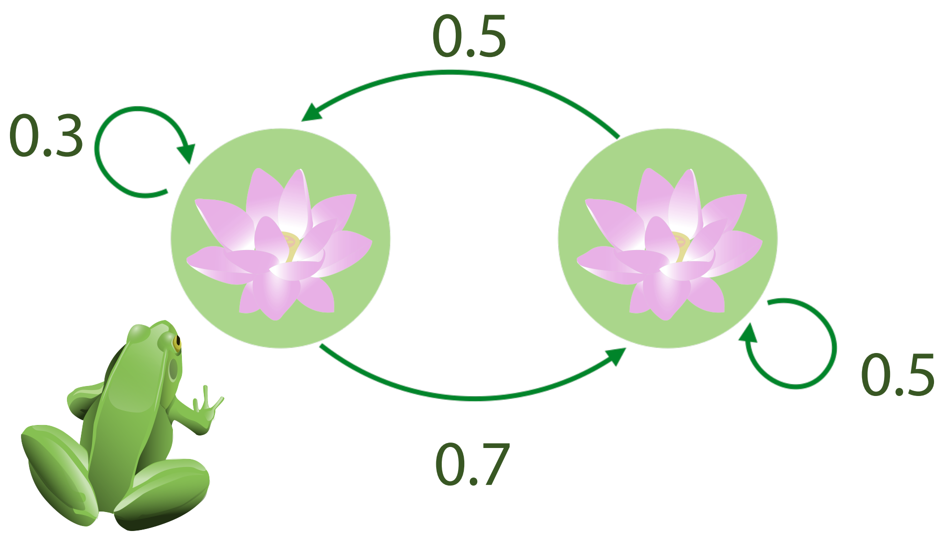
Markov Chains



	p(Left)	p(Right)
x^1	1	0
x^2	0.3	0.7
x^3	$0.3^2 + 0.7 \cdot 0.5$	

$$p(x^3) = p(x^3 \mid x^2 = \mathrm{L})p(x^2 = \mathrm{L}) \\ + p(x^3 \mid x^2 = \mathrm{R})p(x^2 = \mathrm{R})$$

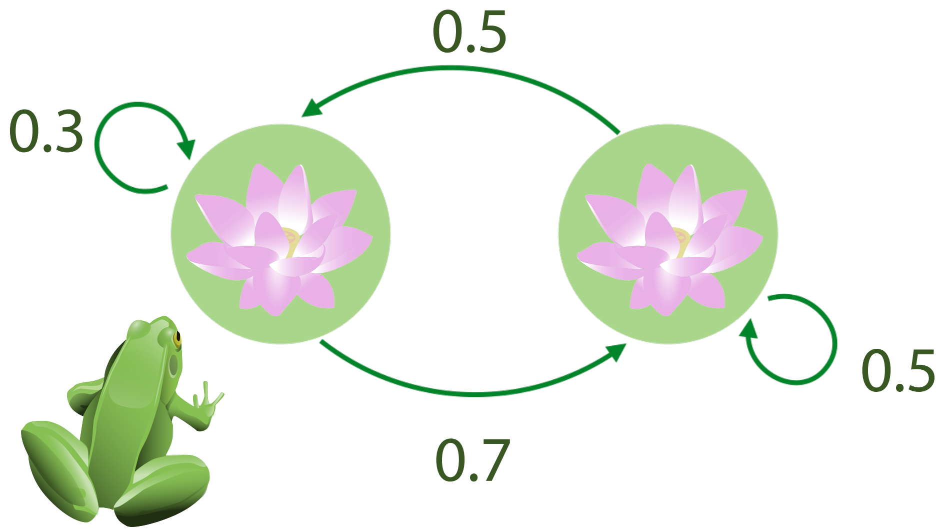
Markov Chains



	p(Left)	p(Right)
x^1	1	0
x^2	0.3	0.7
x^3	$0.3^2 + 0.7 \cdot 0.5$	$0.3 \cdot 0.7 + 0.7 \cdot 0.5$

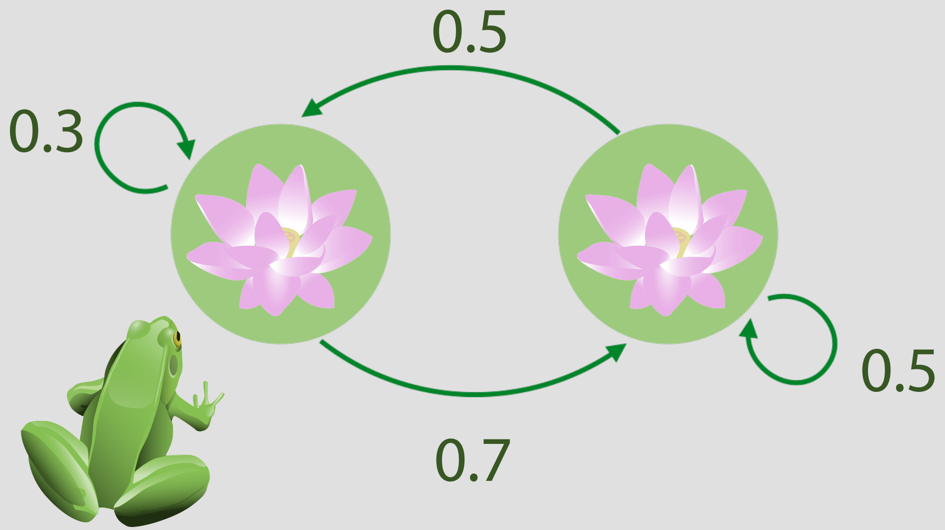
$$\begin{aligned} p(x^3) &= p(x^3 \mid x^2 = \mathrm{L})p(x^2 = \mathrm{L}) \\ &+ p(x^3 \mid x^2 = \mathrm{R})p(x^2 = \mathrm{R}) \end{aligned}$$

Markov Chains



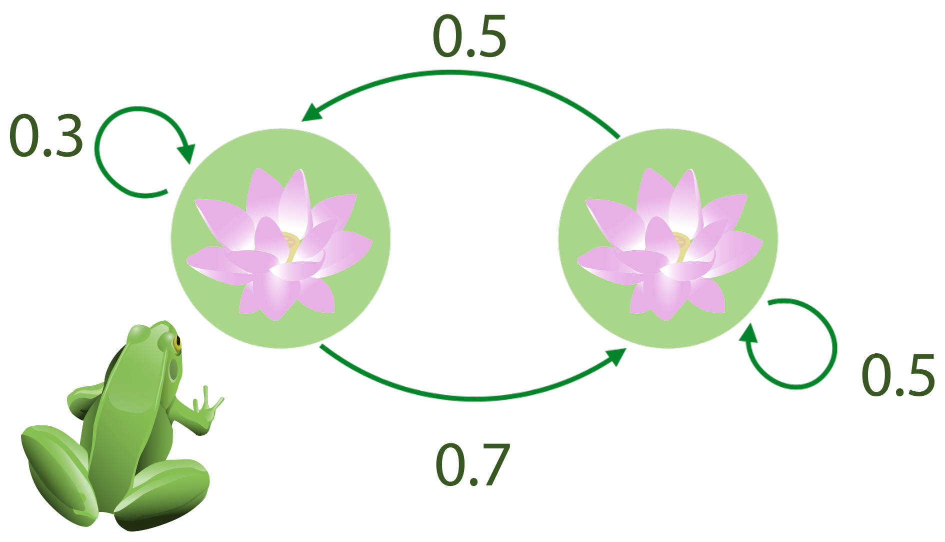
	p(Left)	p(Right)
x^1	1	0
x^2	0.3	0.7
x^3	0.44	0.56

Markov Chains



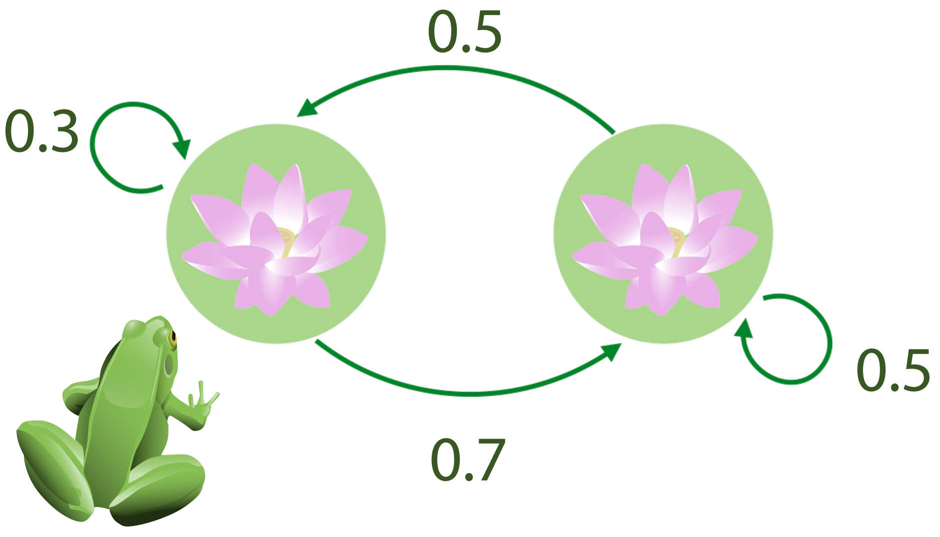
	p(Left)	p(Right)
x^1	1	0
x^2	0.3	0.7
x^3	0.44	0.56
$\dots$	$\dots$	$\dots$
	≈ 0.42	≈ 0.58

Markov Chains



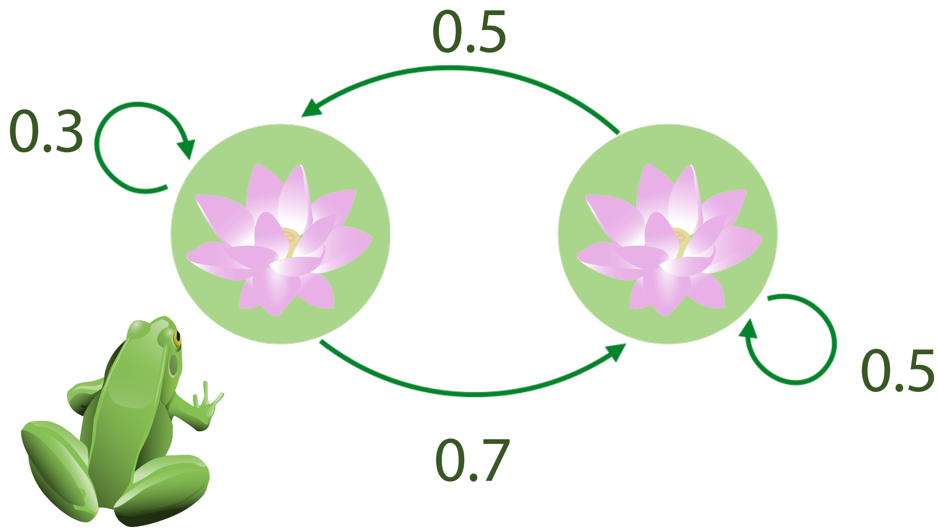
L R R L R … L L

Markov Chains



L R R L R … L **L**

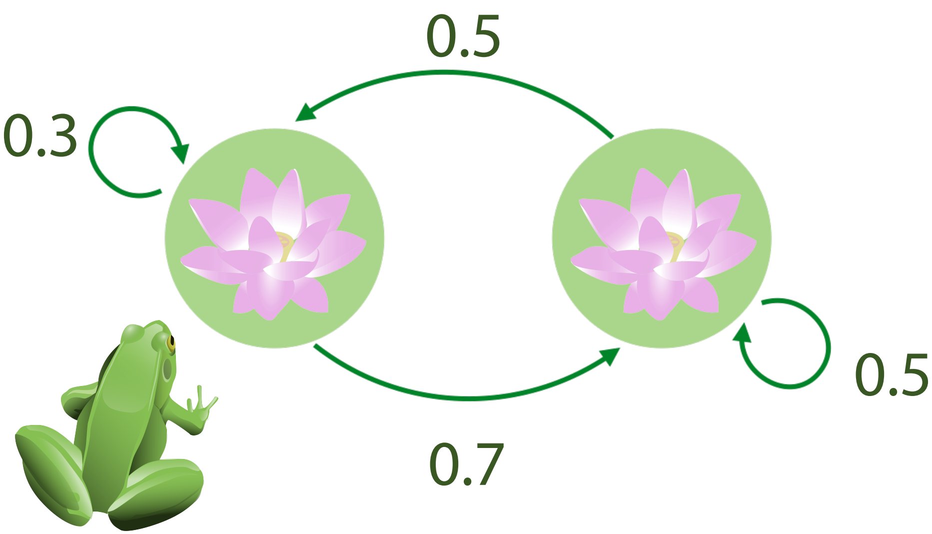
Markov Chains



L R R L R … L **L**

L R R L R … L R

Markov Chains

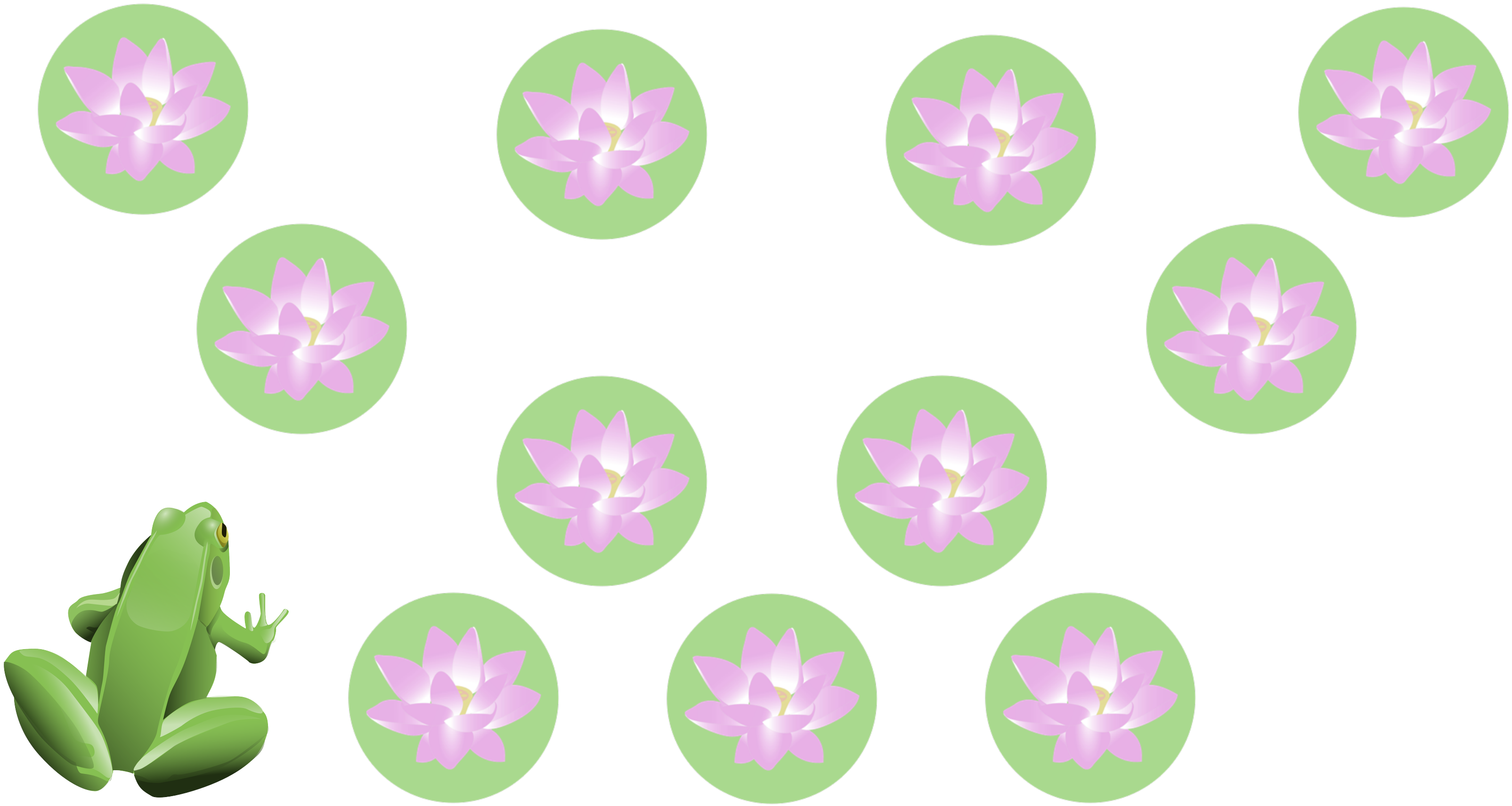


L R R L R … L **L**
L R R L R … L **R**
L R L R R … R **R**
L R R L L … L **R**
L L R L R … R **L**

$$p(L) \approx 0.42$$

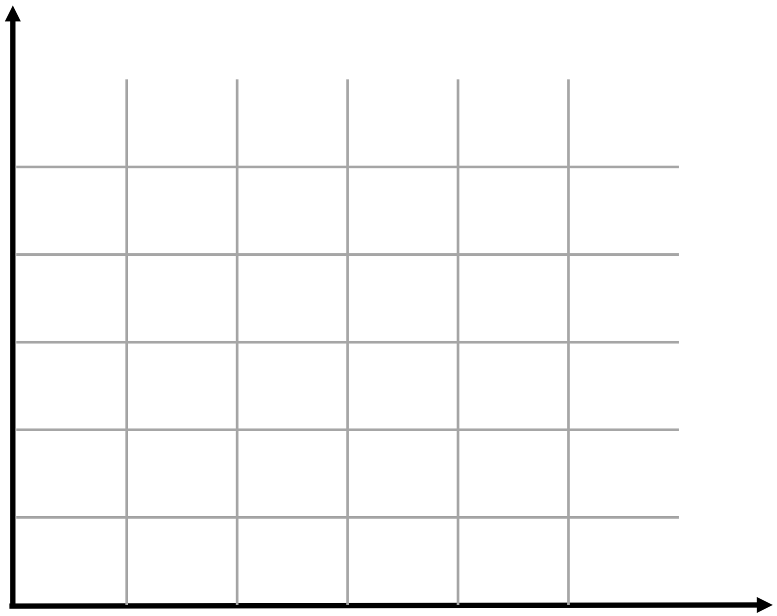
$$p(R) \approx 0.58$$

Markov Chains



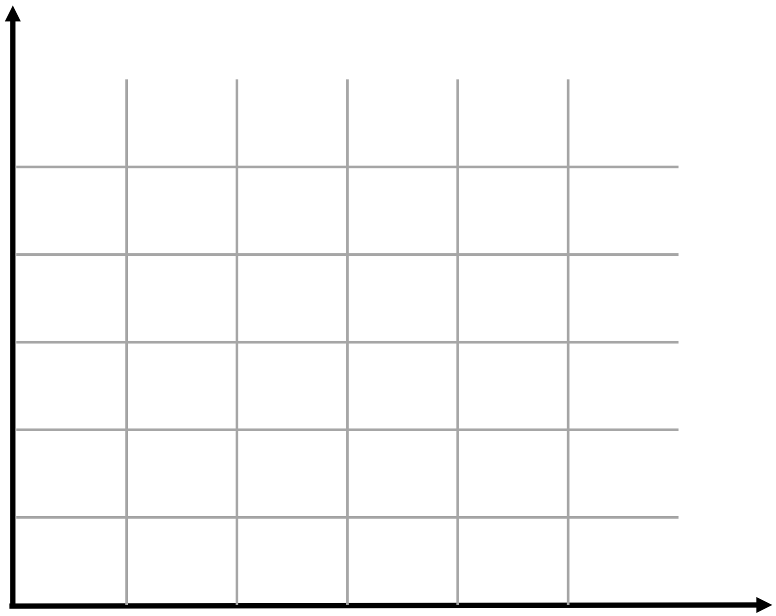
But what if there are 10 lilies? Or a billion?

Markov Chains



But what if there are 10 lilies? Or a billion?
Or maybe frog position is continuous?

Markov Chains



But what if there are 10 lilies? Or a billion?
Or maybe frog position is continuous?
You can still sample!

Using Markov Chain

- We want to sample from $p(x)$

Using Markov Chain

- We want to sample from $p(x)$
- Build a Markov chain that converge to $p(x)$

Using Markov Chain

- We want to sample from $p(x)$
- Build a Markov chain that converge to $p(x)$
- Start from any x^0

Using Markov Chain

- We want to sample from $p(x)$
- Build a Markov chain that converge to $p(x)$
- Start from any x^0
- For k = 0, 1, …

$$x^{k+1} \sim T(x^k \rightarrow x^{k+1})$$

Using Markov Chain

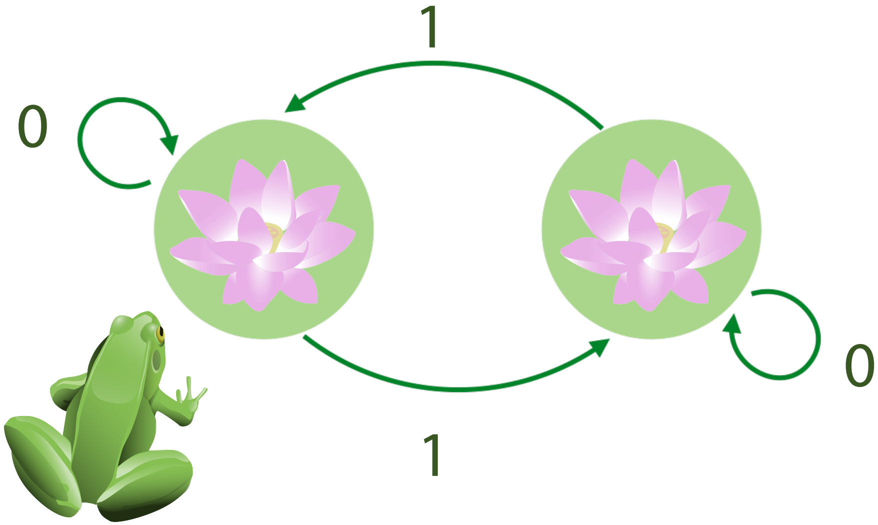
- We want to sample from $p(x)$
- Build a Markov chain that converge to $p(x)$
- Start from any x^0
- For k = 0, 1, …

$$x^{k+1} \sim T(x^k \to x^{k+1})$$

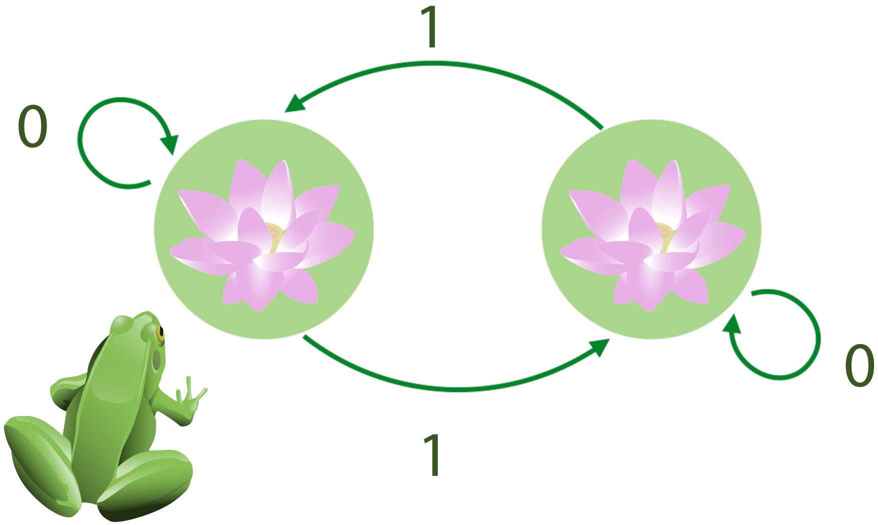
- Eventually x^k will look like samples from $p(x)$

Do Markov chains always converge?

Do Markov chains always converge?



Do Markov chains always converge?



	p(Left)	p(Right)
x^1	1	0
x^2	0	1
x^3	1	0
$\cdots$	$\cdots$	$\cdots$

Does not converge

Markov Chains

Definition:

A distribution π is called stationary if

$$\pi(x') = \sum_x T(x \to x')\pi(x)$$

Markov Chains

Theorem:

If $T(x \to x') > 0$ for all x, x' then exists unique π:

$$\pi(x') = \sum_{x} T(x \to x')\pi(x)$$

Markov Chains

Theorem:

If $T(x \rightarrow x') > 0$ for all x, x' then exists unique π:

$$\pi(x') = \sum_x T(x \rightarrow x')\pi(x)$$

And Markov chain converges to π from any starting point