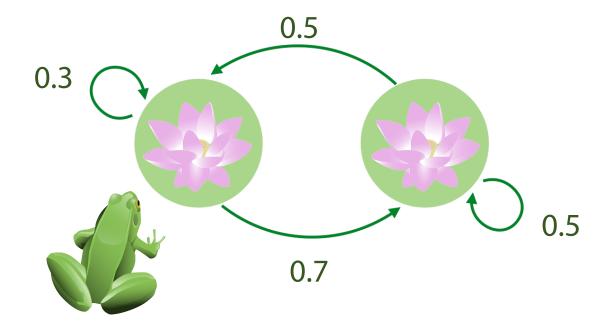


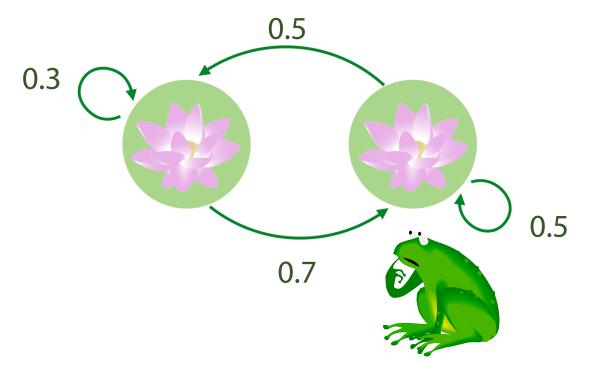
$$T(L \rightarrow L) = 0.3$$
 $T(R \rightarrow L) = 0.5$

$$T(L \rightarrow R) = 0.7$$
 $T(R \rightarrow R) = 0.5$



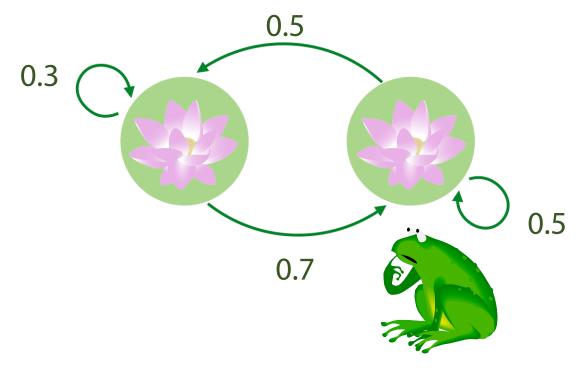
Timestamp: 1

Log: L



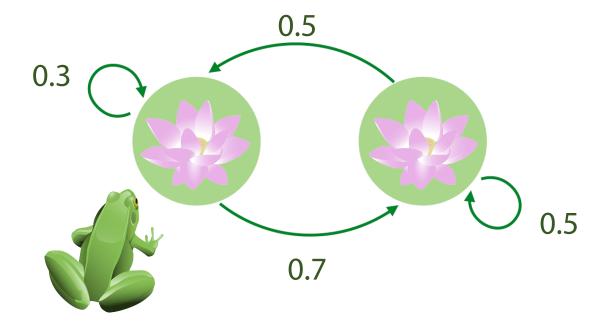
Timestamp: 2

Log: LR



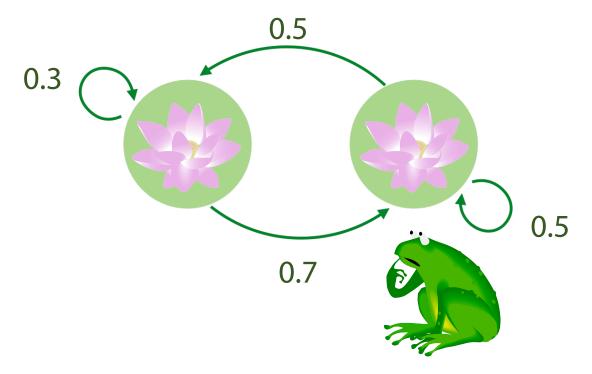
Timestamp: 3

Log: LRR



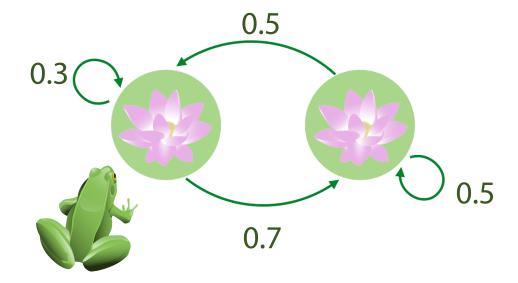
Timestamp: 4

Log: LRRL

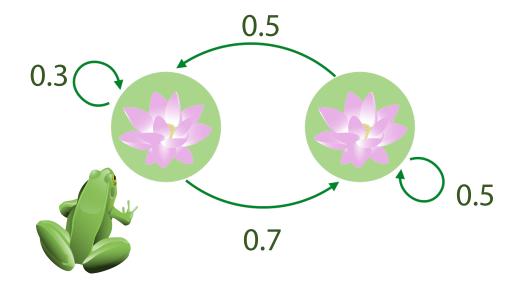


Timestamp: 5

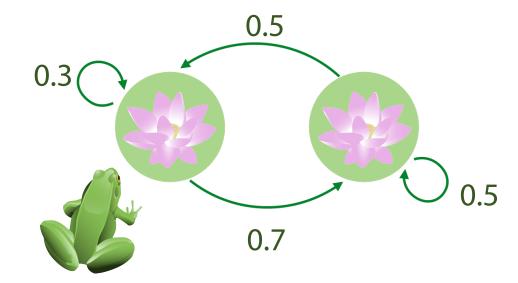
Log: LRRLR



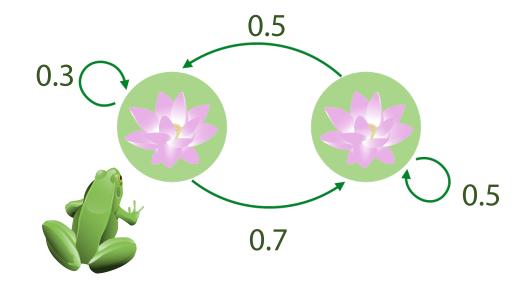
	p(Left)	p(Right)
x^1	1	0



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

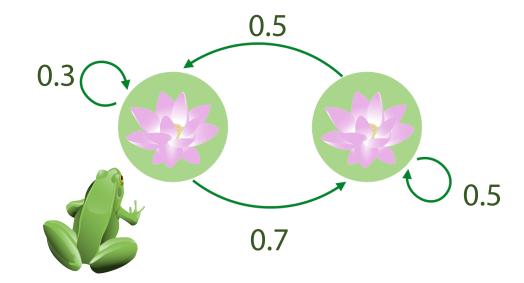


	p(Left)	p(Right)
x^1	1	0
x^2	0.3	0.7
x^3		



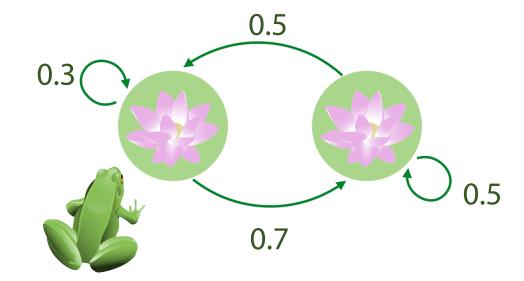
	p(Left)	p(Right)
x^1	1	0
x^2	0.3	0.7
x^3		

$$p(x^3) = p(x^3 | x^2 = L)p(x^2 = L) + p(x^3 | x^2 = R)p(x^2 = R)$$



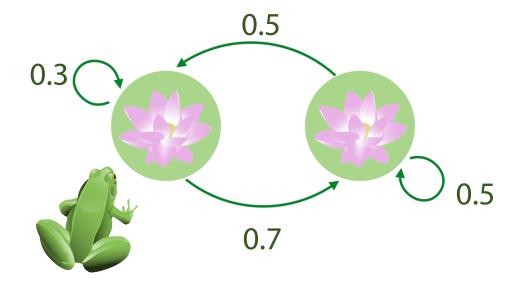
	p(Left)	p(Right)
x^1	1	0
x^2	0.3	0.7
x^3	$0.3^2 + 0.7 \cdot 0.5$	

$$p(x^3) = p(x^3 | x^2 = L)p(x^2 = L) + p(x^3 | x^2 = R)p(x^2 = R)$$

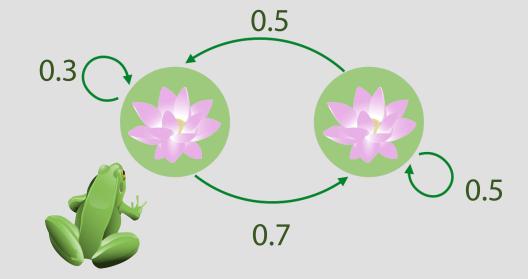


	p(Left)	p(Right)
x^1	1	0
x^2	0.3	0.7
x^3	$0.3^2 + 0.7 \cdot 0.5$	$0.3 \cdot 0.7 + 0.7 \cdot 0.5$

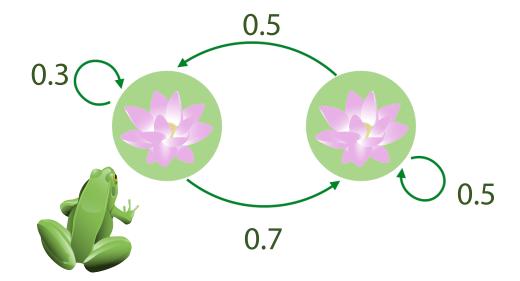
$$p(x^3) = p(x^3 | x^2 = L)p(x^2 = L) + p(x^3 | x^2 = R)p(x^2 = R)$$



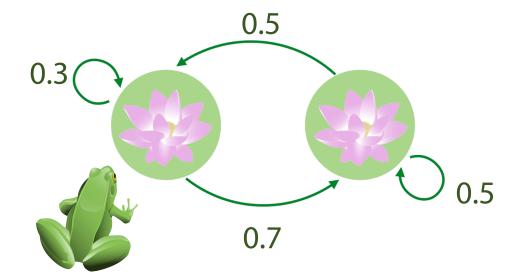
	p(Left)	p(Right)
x^1	1	0
x^2	0.3	0.7
x^3	0.44	0.56



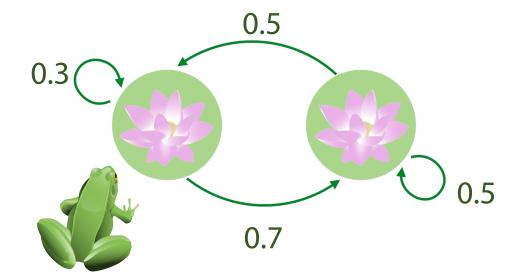
	p(Left)	p(Right)
x^1	1	0
x^2	0.3	0.7
x^3	0.44	0.56
• • •	• • •	
	≈ 0.42	≈ 0.58



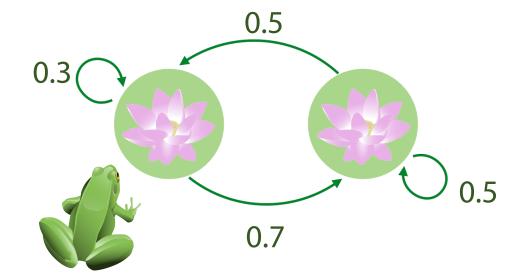
LRRLR...LL



LRRLR...LL



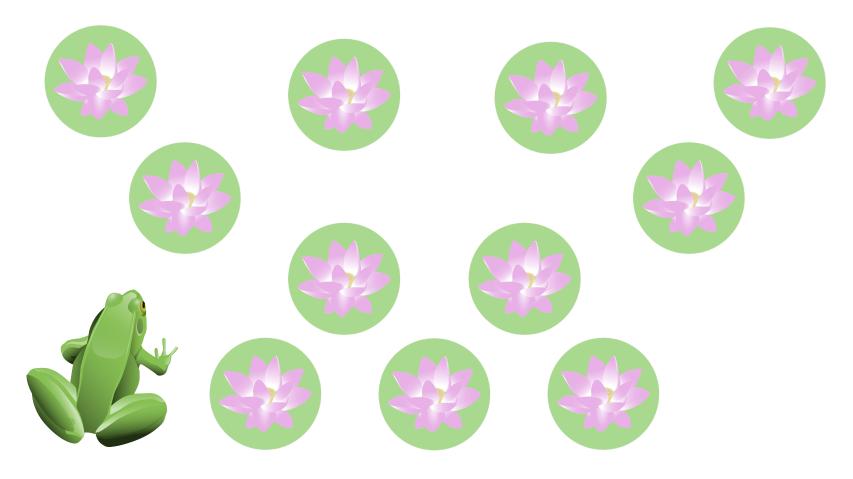
LRRLR...LL LRRLR...LR



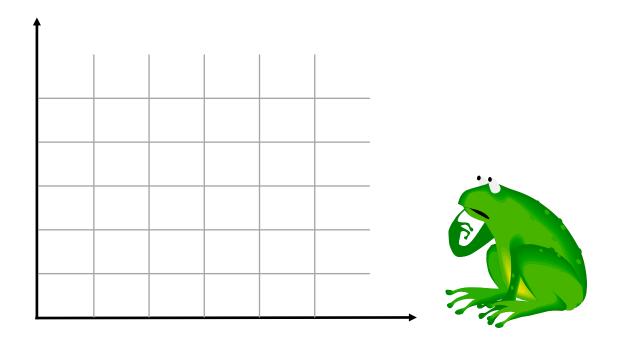
LRRLR...LL LRRLR...LR LRLRR...RR LRRLL...LR LLRLR...RL

$$p(L) \approx 0.42$$

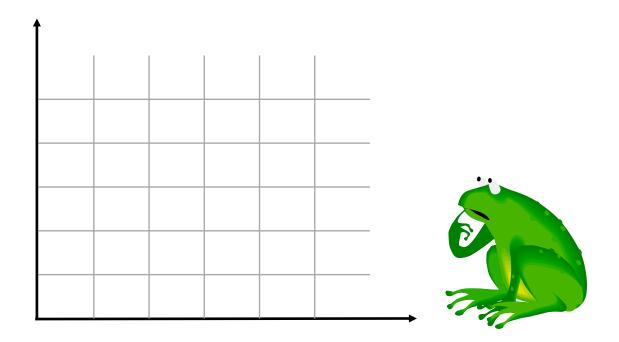
$$p(R) \approx 0.58$$



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



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



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

• We want to sample from p(x)

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

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

- 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, ...

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

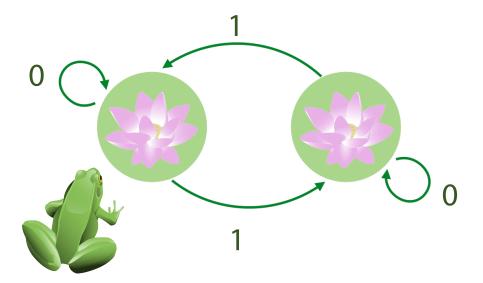
- We want to sample from p(x)
- Build a Markov chain that converge to p(x)
- Start from any x^0
- For k = 0, 1, ...

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

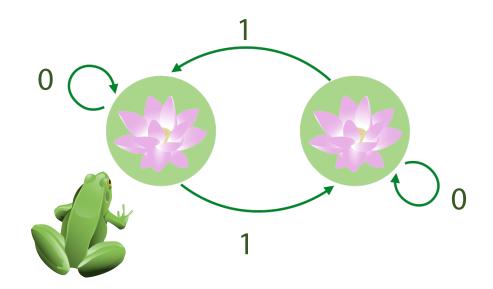
• Eventually x^k will look like samples from p(x)



Do Markov chains always converge?



Do Markov chains always converge?



	p(Left)	p(Right)
x^1	1	0
x^2	0	1
x^3	1	0
• • •	• • •	• • •

Does not converge

Definition:



A distribution π is called stationary if

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

Theorem:

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

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

Theorem:

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

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

And Markov chain converges to π from any starting point