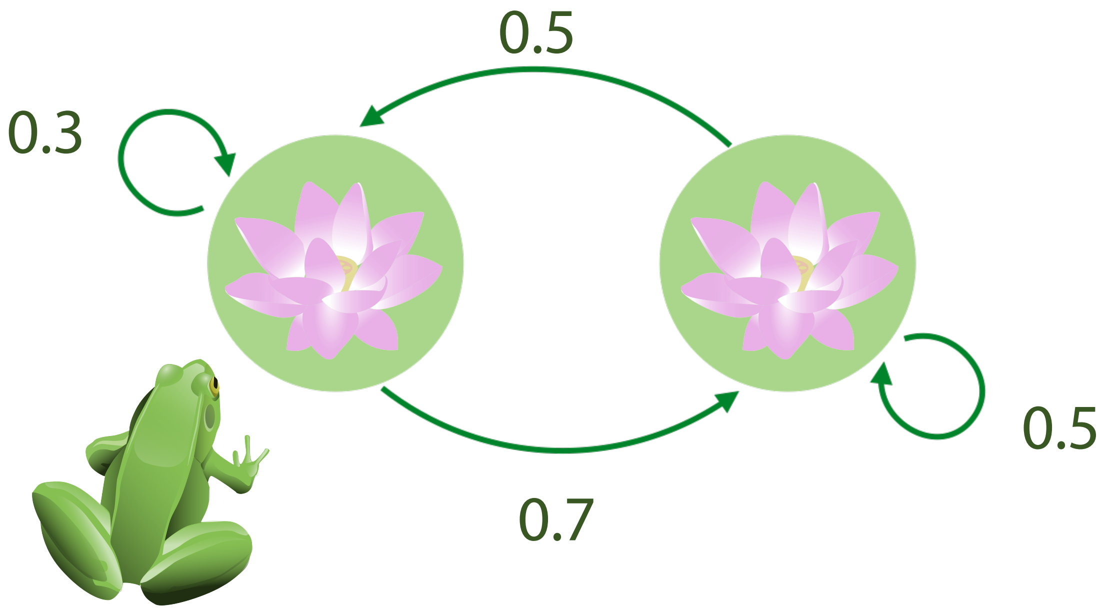
# Markov Chains

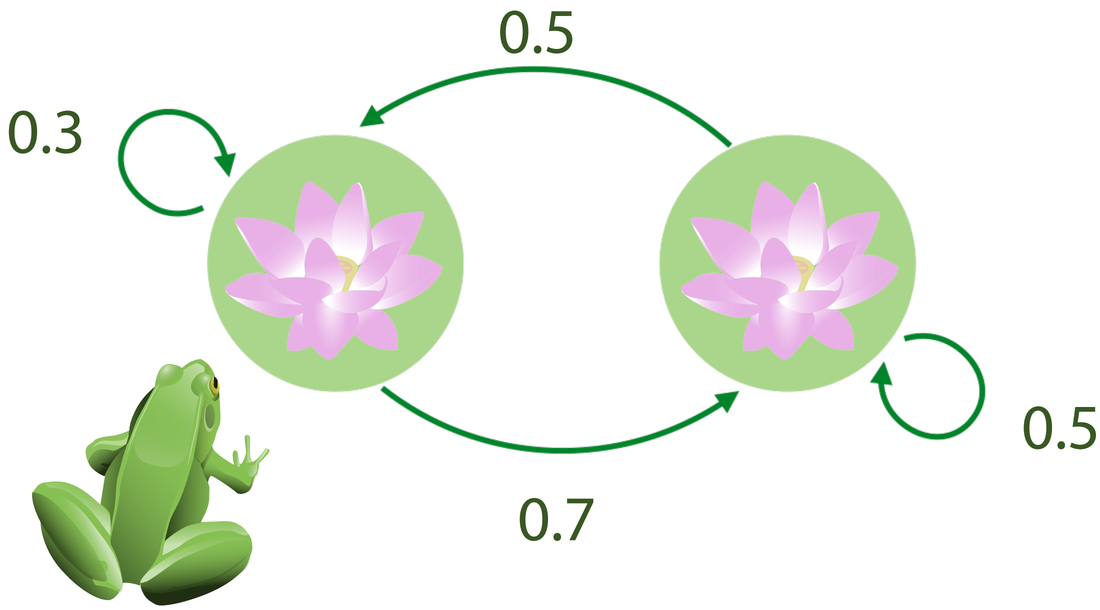



$$T(\mathrm{L} \to \mathrm{L}) = 0.3 \qquad T(\mathrm{R} \to \mathrm{L}) = 0.5$$

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

Markov chains depend only from the previous state.

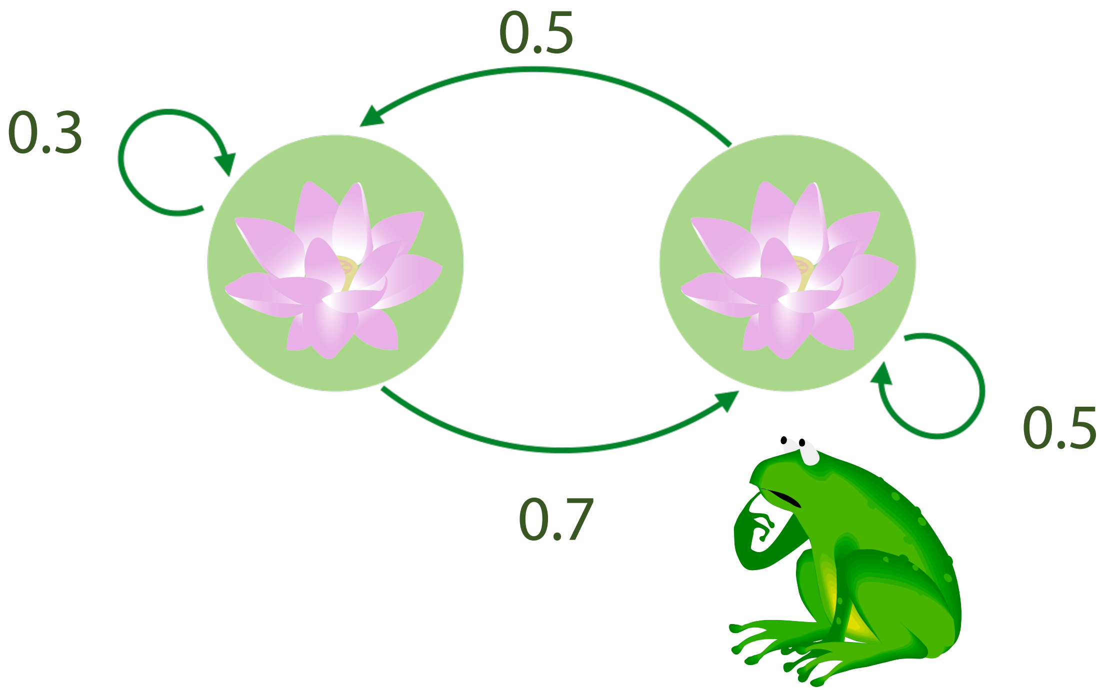
# Markov Chains




Timestamp: **1**

Log: **L**

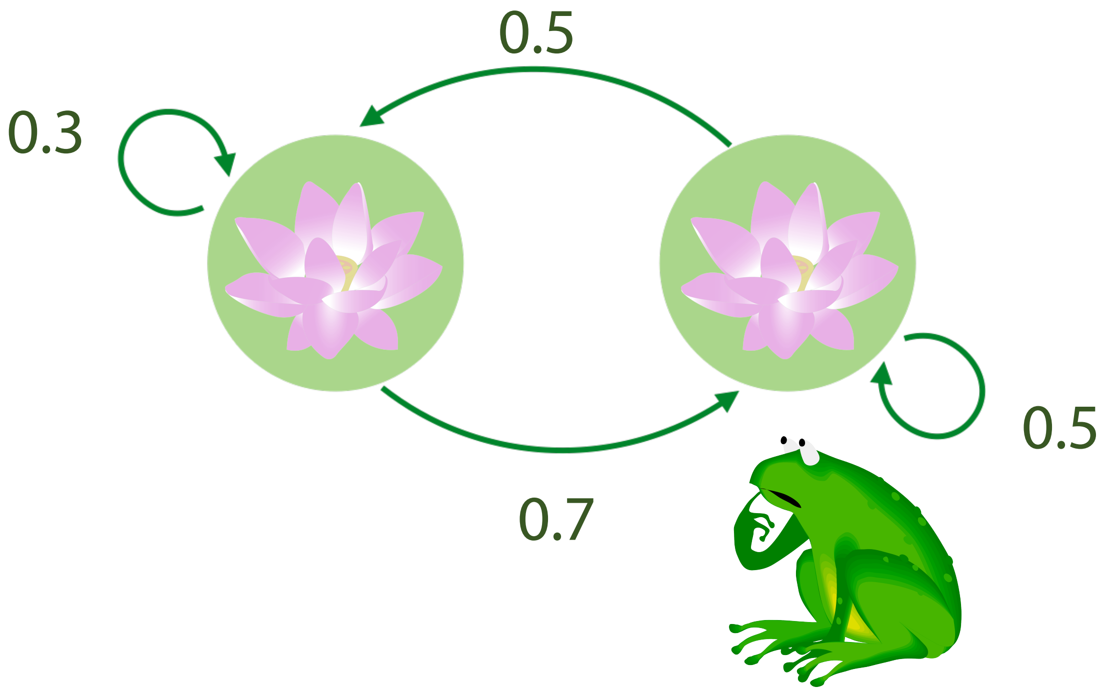
# Markov Chains




Timestamp: **2**

Log: **L R**

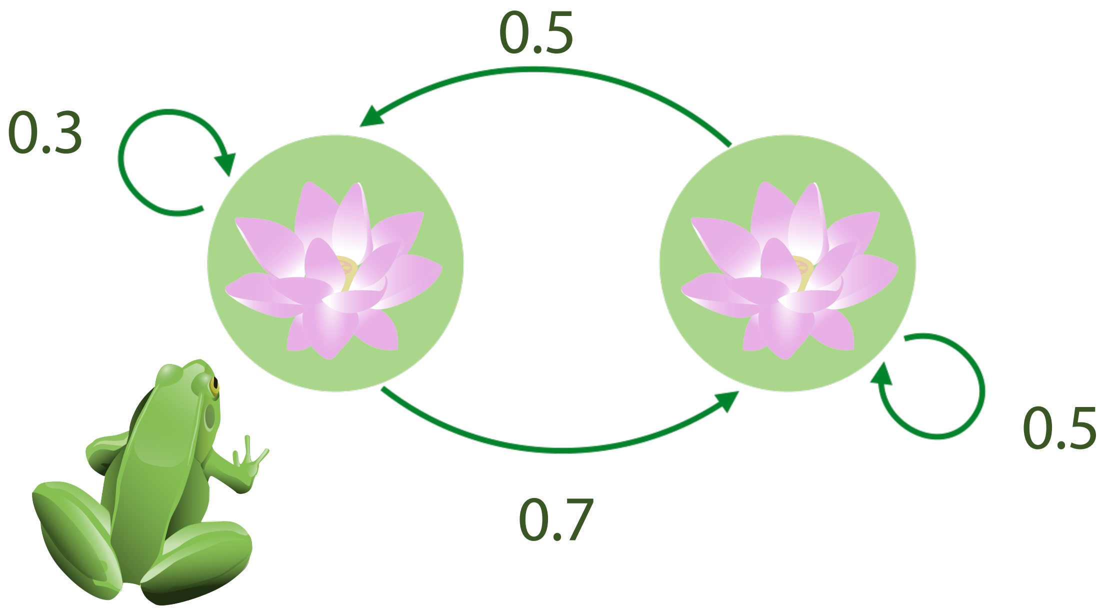
# Markov Chains




Timestamp: **3**

Log: **L R R**

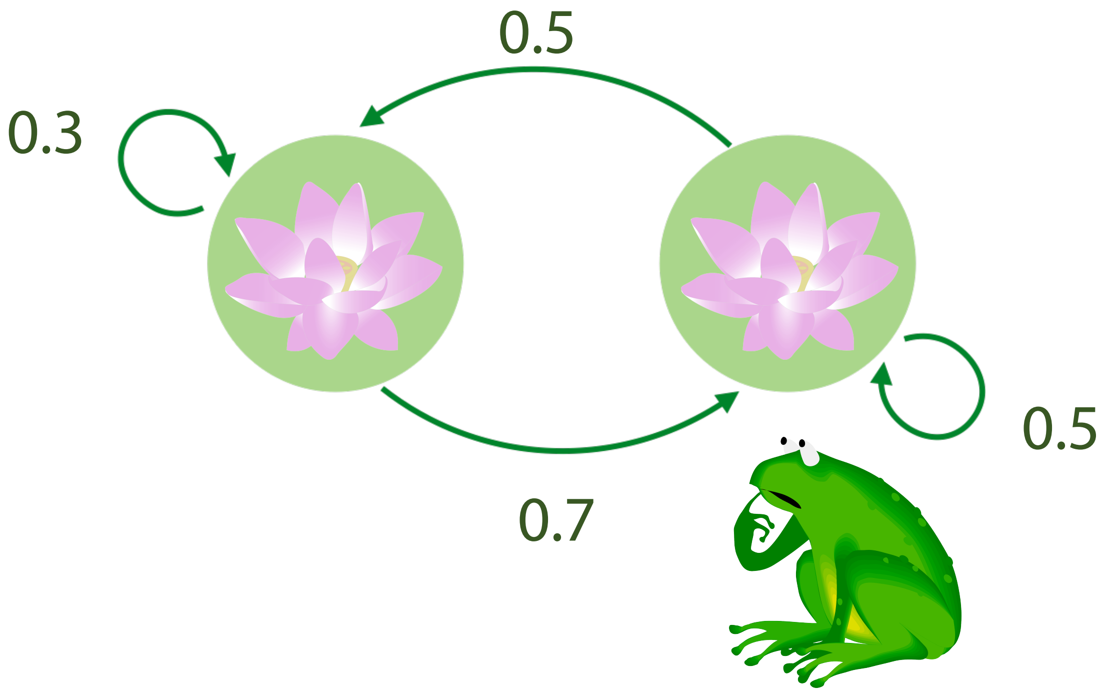
# Markov Chains




Timestamp: **4**

Log: **L R R L**

# Markov Chains



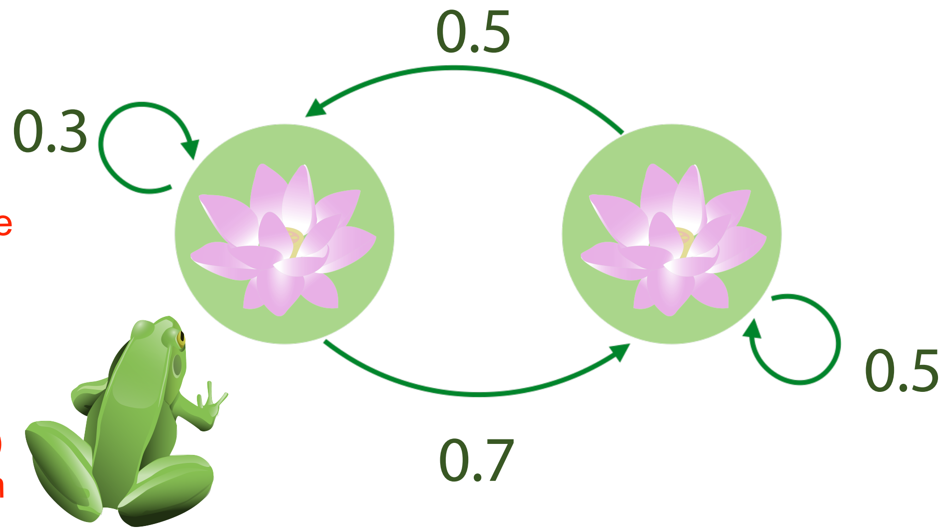

Timestamp: **5**

Log: **L R R L R**

# Markov Chains

For position at point x^1, it starts on the left.
For x^2, from x^1, it can move to the left (stay) with proba 0.3, OR go to the right with proba 0.7.

To compute p(left) or p(right) for x^n, we always refer from the old x^{n-1}.

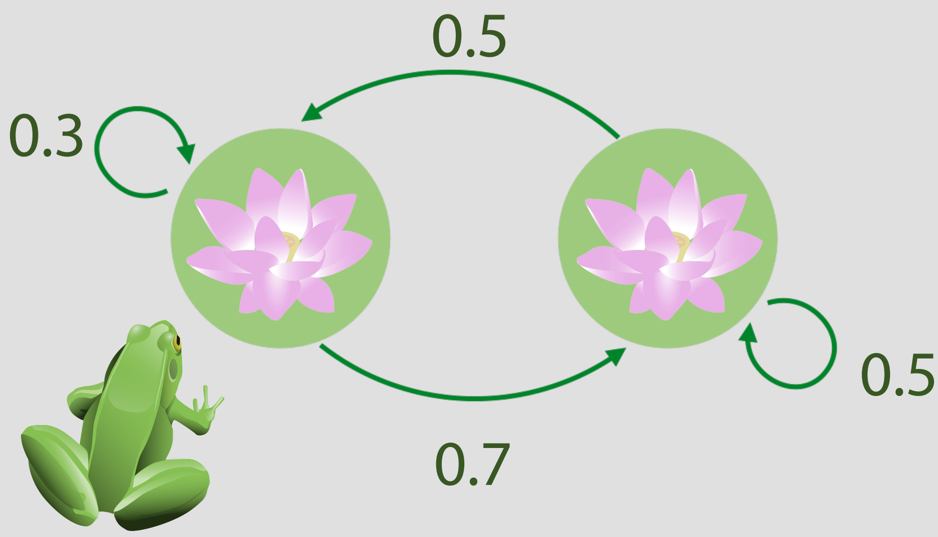



| | p(Left) | p(Right) |
|---|---|---|
| $x^1$ | $1$ | $0$ |
| $x^2$ | $0.3$ | $0.7$ |
| $x^3$ | $0.3^2 + 0.7 \cdot 0.5$ | $0.3 \cdot 0.7 + 0.7 \cdot 0.5$ |

This can be generalized for x^n >>

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

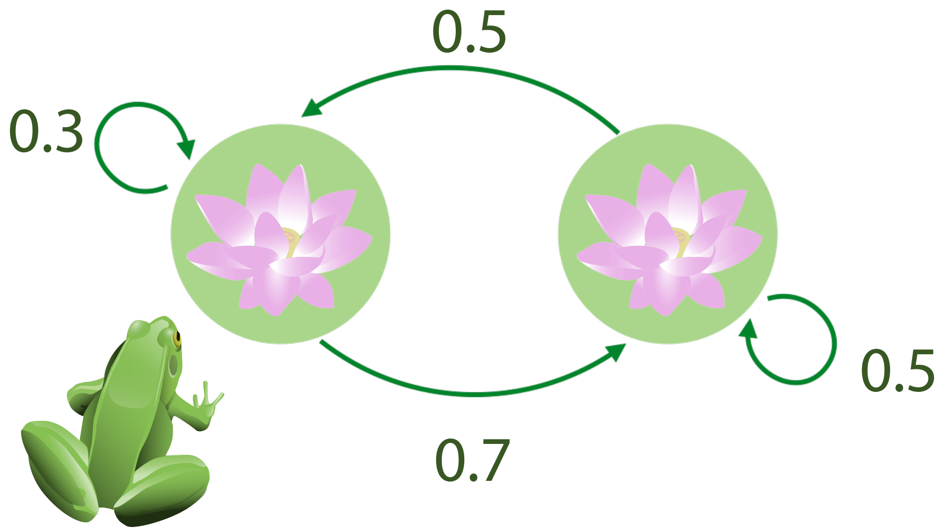
# Markov Chains




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

Markov chain can eventually converge to some value after many iterations.

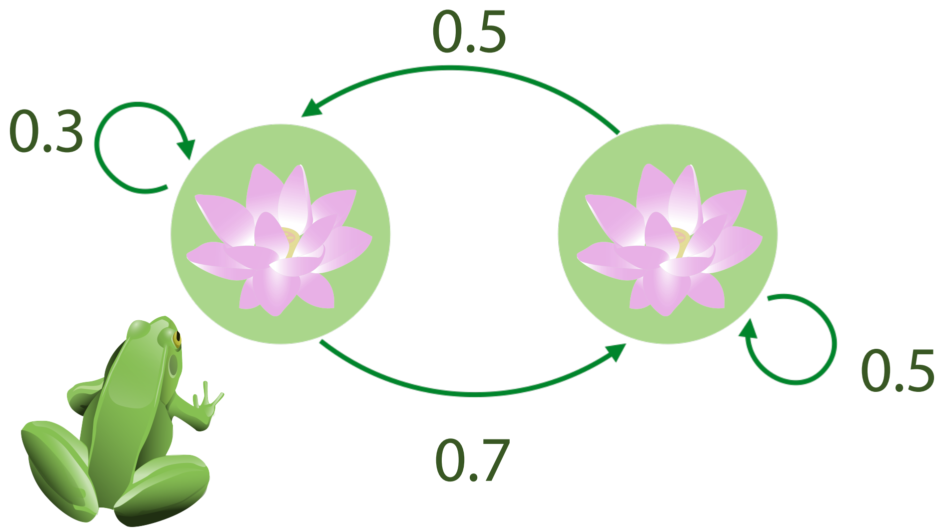
# Markov Chains




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**

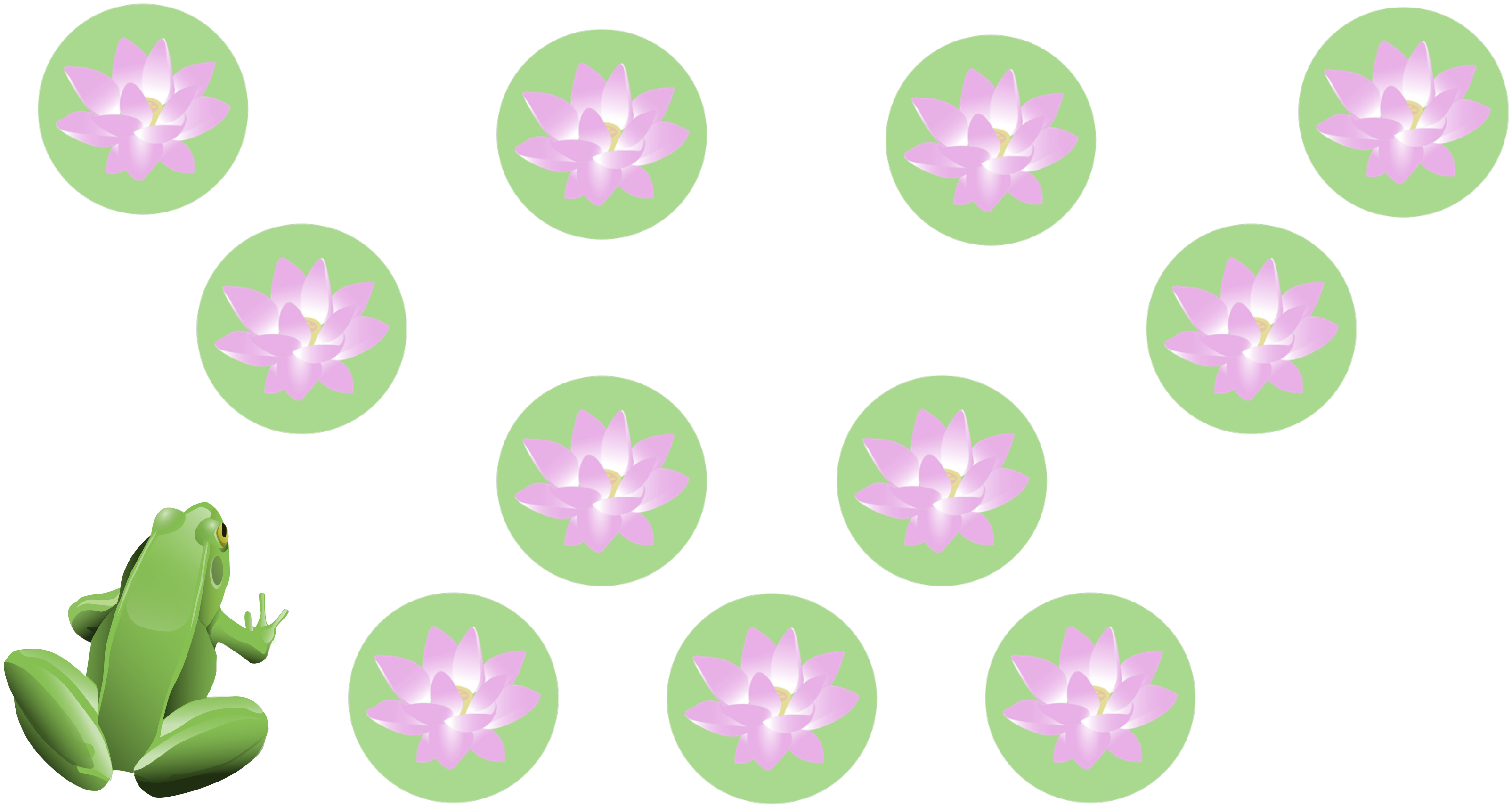
L L R L R ... R **L**

$$p(L) \approx 0.42$$

$$p(R) \approx 0.58$$

Converge to some value.

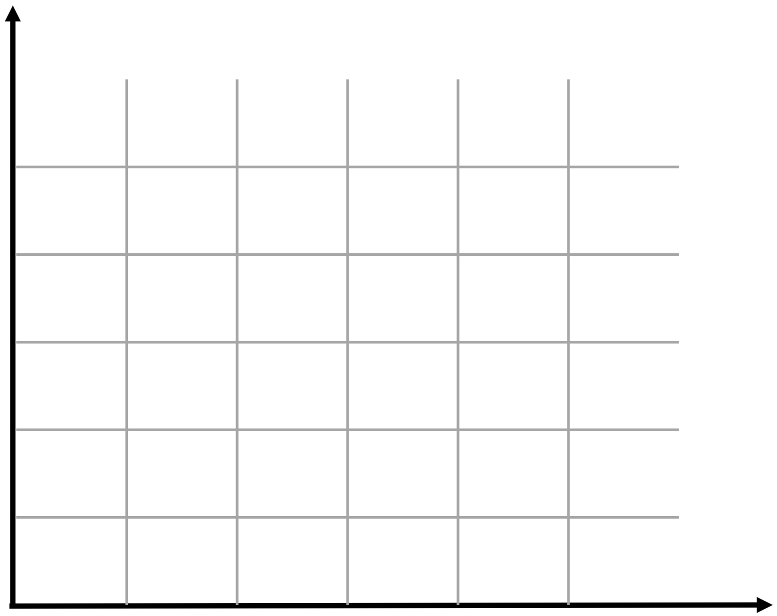
# Markov Chains



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

We can still use the markov chain.
We can simulate complicated distributions by using a markov chain.

# 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

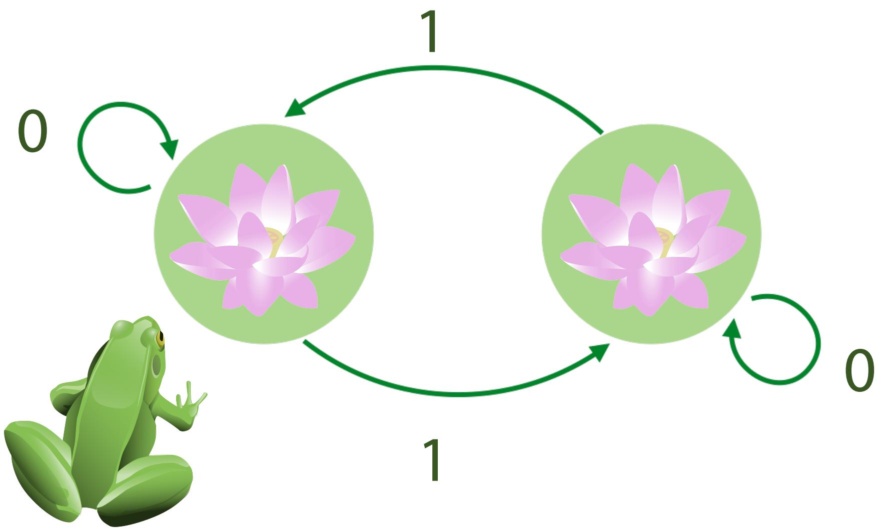
- We want to sample from $p(x)$
- Build a Markov chain that converge to $p(x)$
- Start from any $x^0$
- For k = 0, 1, …

generate x^{k+1} using the transition probabilities.

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

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

# Do Markov chains always converge?




| | p(Left) | p(Right) |
|---|---|---|
| $x^1$ | 1 | 0 |
| $x^2$ | 0 | 1 |
| $x^3$ | 1 | 0 |
| $\ldots$ | $\ldots$ | $\ldots$ |

This one obviously does not converge.

**Does not converge**

# Markov Chains

## Definition:

A distribution $\pi$ is called stationary if

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

For a markov chain T.

If we start from Pi, and marginalize out the current position x, then we'll get the distribution of the position of the next step.

If we encounter this pi, that means we will stay at the distribution.
It converges to THIS distribution.

# Markov Chains

oooo i see.
If all transition probabilities are no nzero, then we will converge.

## Theorem:

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

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

And Markov chain converges to $\pi$ from any starting point

Is there a proof for this?