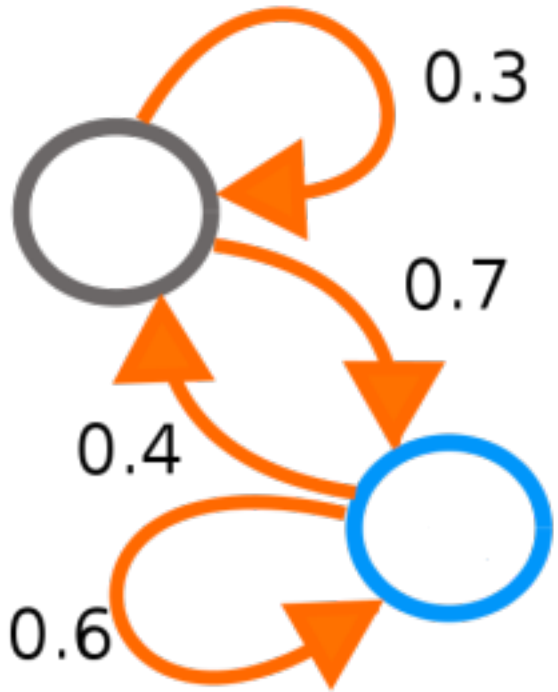
Markov Models

Anil Ada



Oct 22nd, 2014

A day in the life of me

9:00am

Work

$$f(x) = \sum_{S \subseteq [n]} \widehat{f}(S)\chi_S(x)$$

A day in the life of me

9:01am

Work

$$f(x) = \sum_{S \subseteq [n]} \widehat{f}(S) \chi_S(x)$$

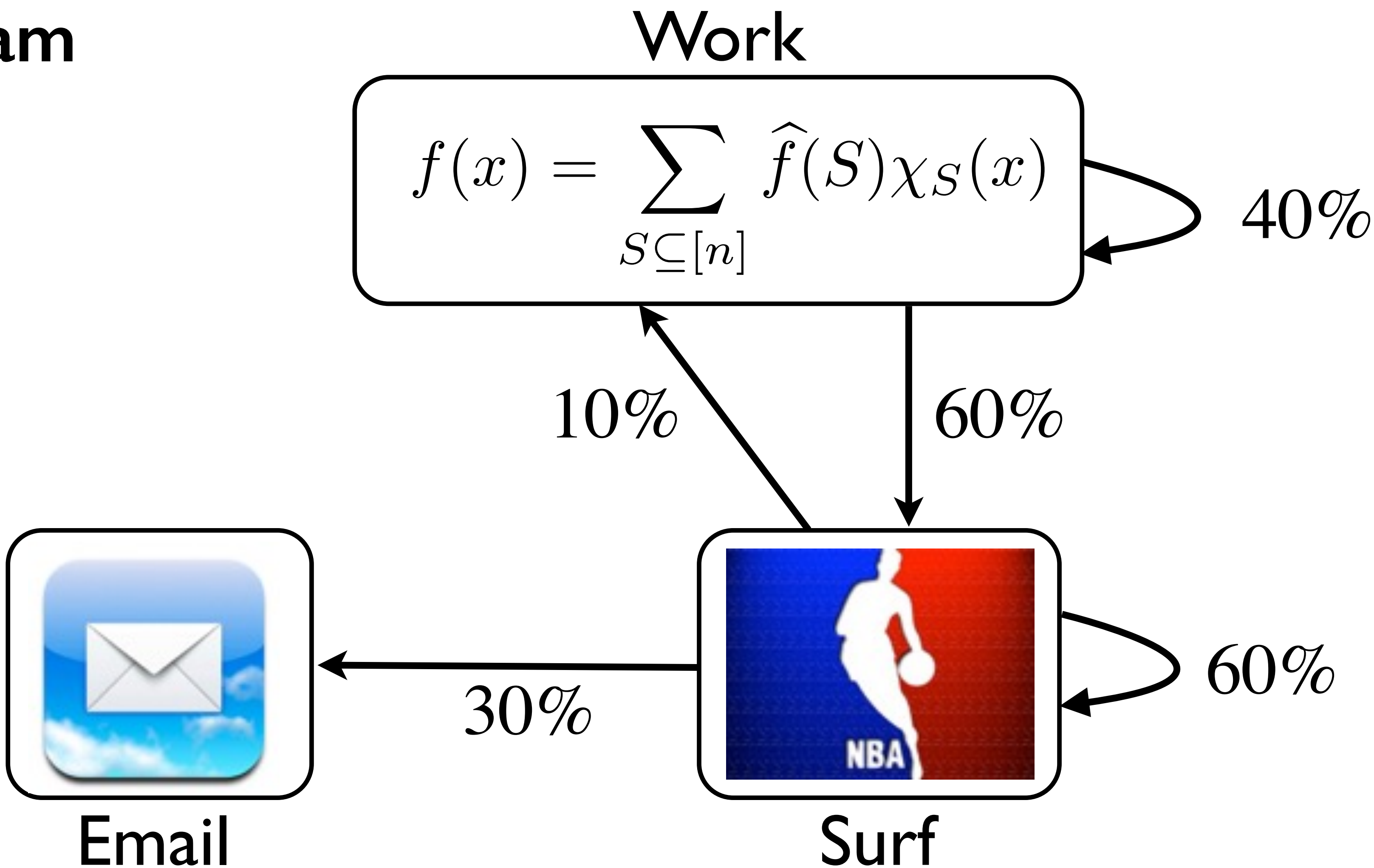
40%

60%

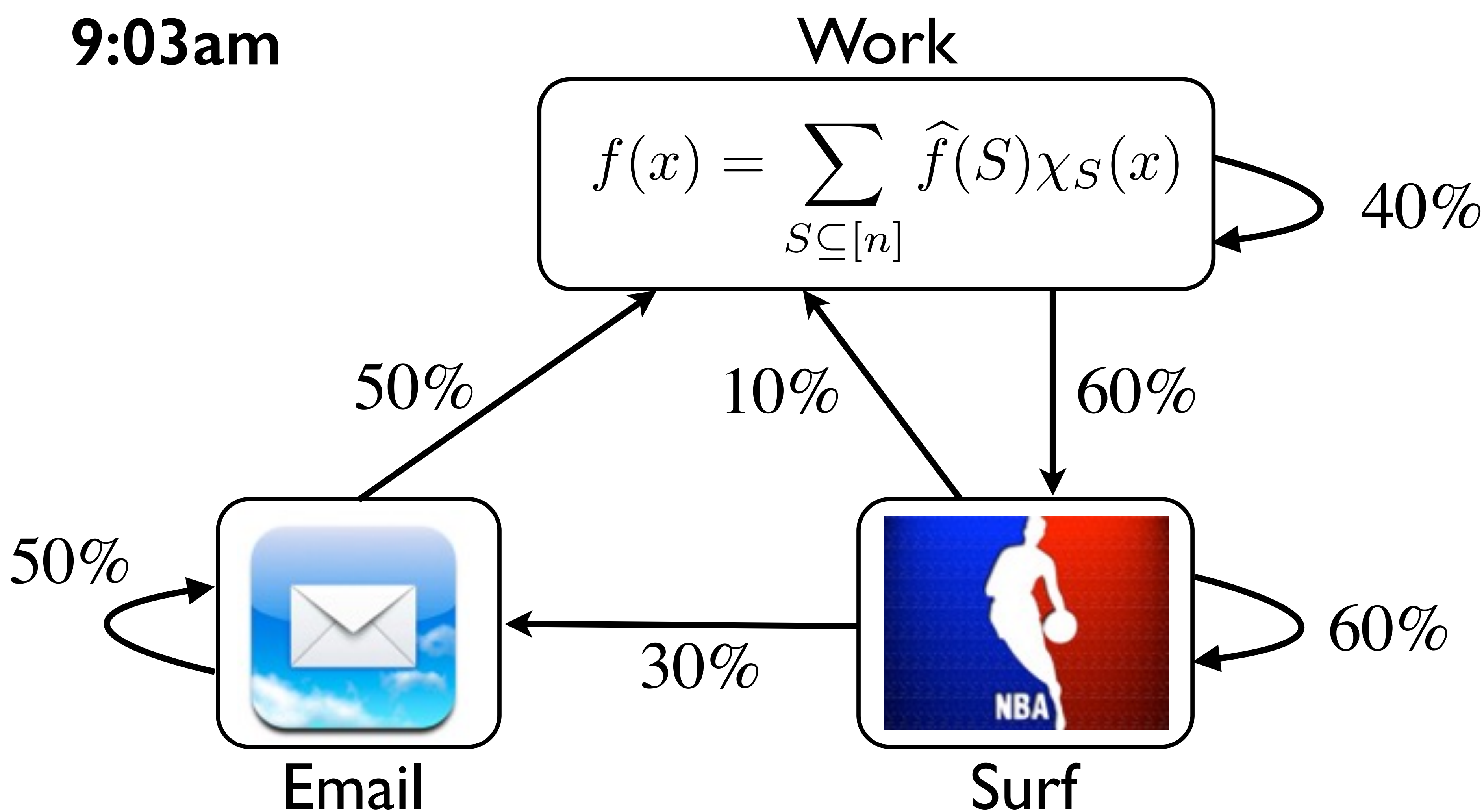
Surf

A day in the life of me

9:02am

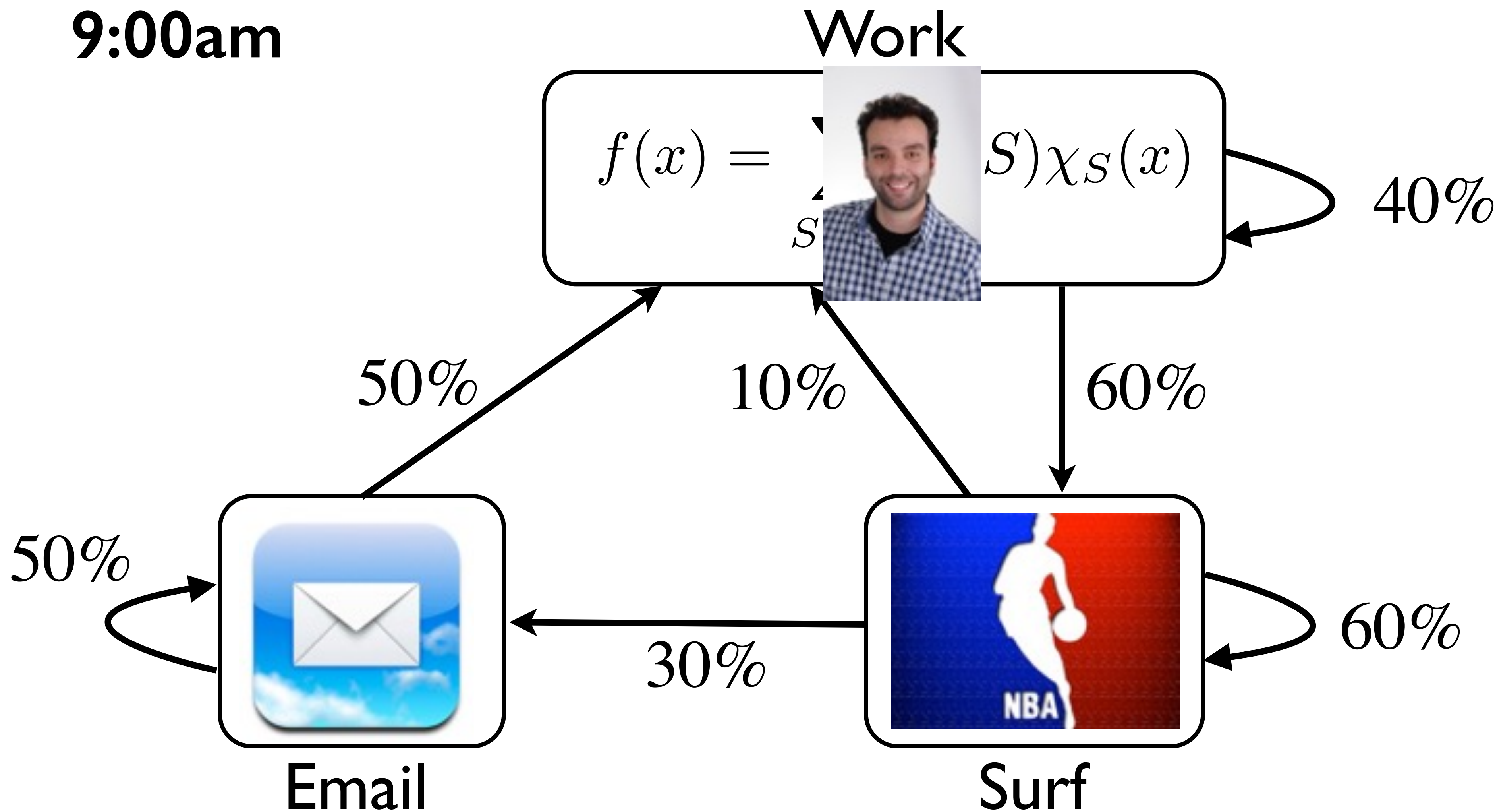


A day in the life of me



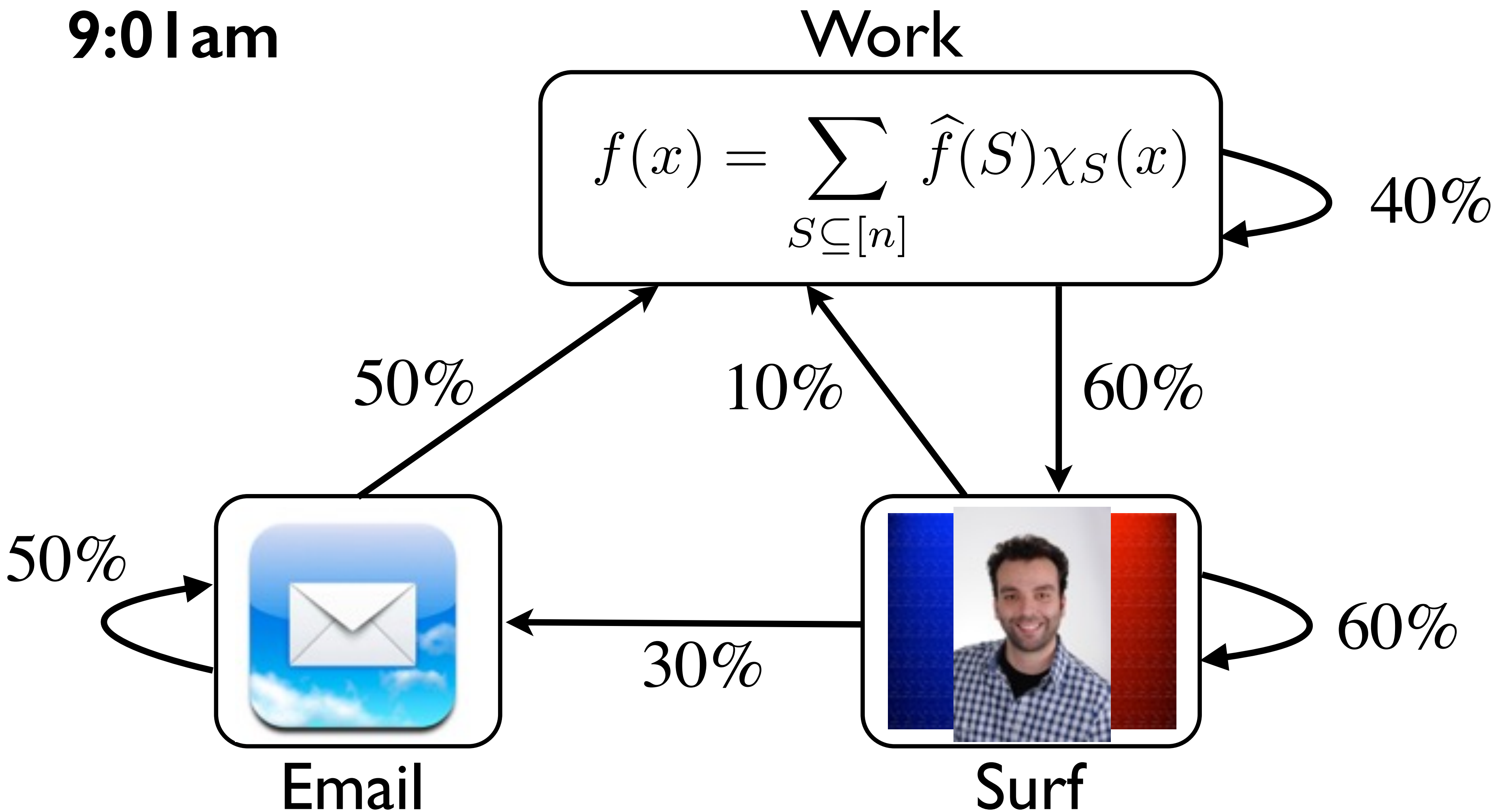
A day in the life of me

9:00am



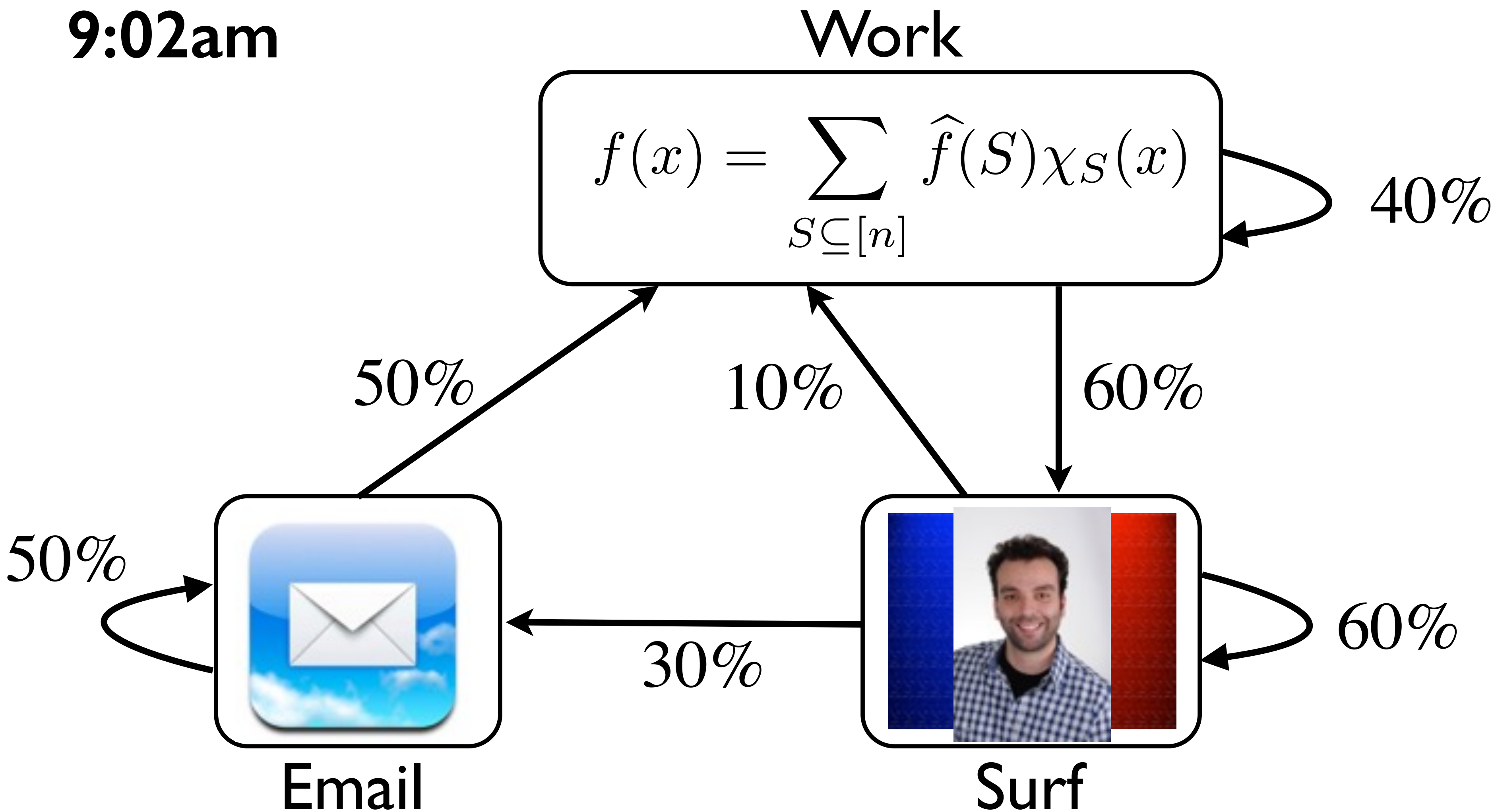
A day in the life of me

9:01am



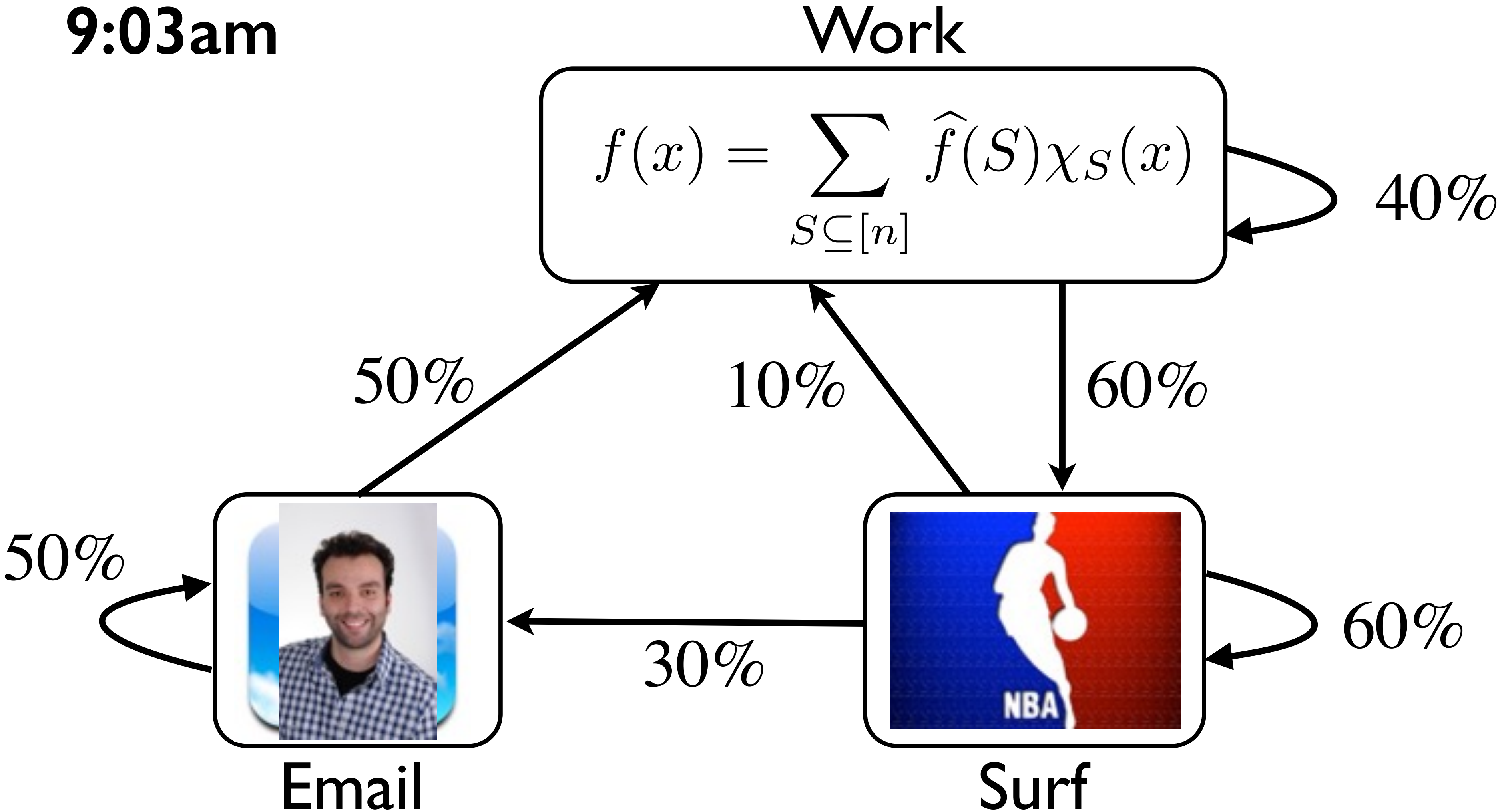
A day in the life of me

9:02am



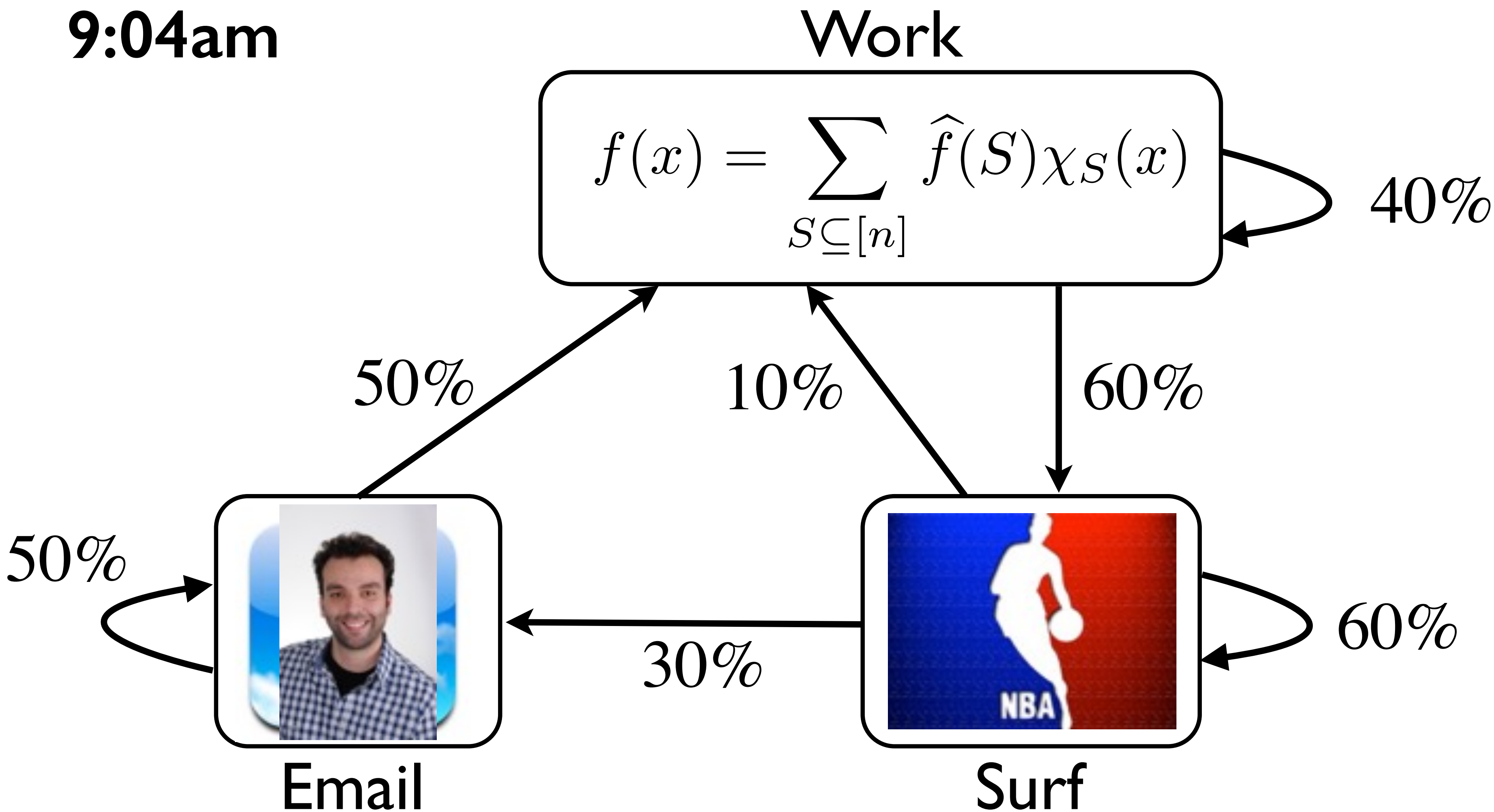
A day in the life of me

9:03am



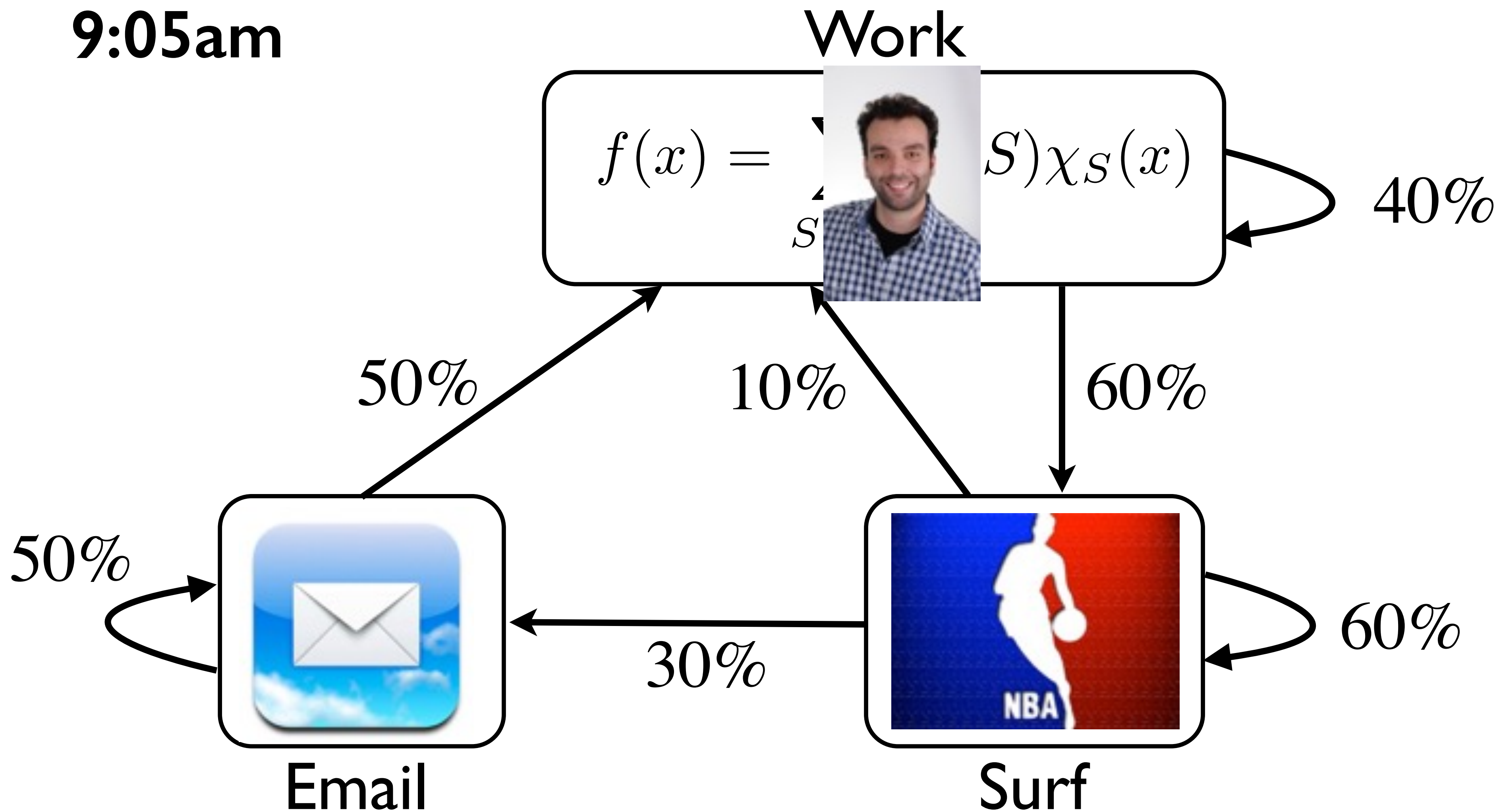
A day in the life of me

9:04am



A day in the life of me

9:05am



Markov Model

Markov Model

Andrey Markov (1856 - 1922)

- Russian mathematician.
- Famous for his work on <u>random processes</u>.

Markov Model

Andrey Markov (1856 - 1922)

Russian mathematician.

Famous for his work on random processes.

The canonical probabilistic model for sequential data.

A model for the evolution of a random system.

The future is independent of the past, given the present.

Cool Things About Markov Model

It is a very general and natural model.

- Extraordinary number of applications in many different disciplines:
  - computer science, mathematics, biology, physics, chemistry, economics, psychology, music, baseball,...

The model is simple and neat.

A beautiful mathematical theory behind it.

- Starts simple, goes very deep.

Outline

Motivating examples and applications

Basic mathematical representation

Applications

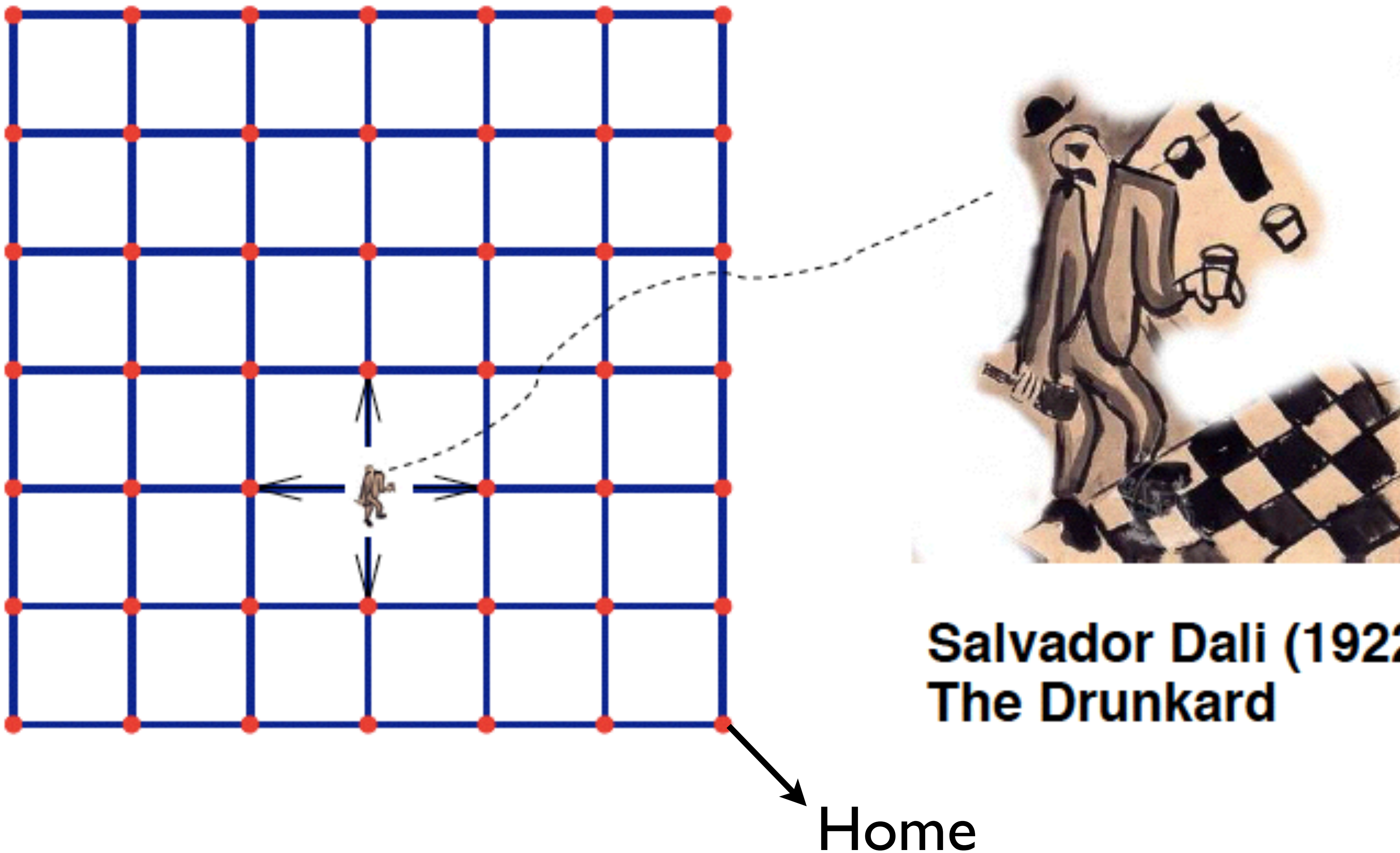
Outline

Motivating examples and applications

The future is independent of the past, given the present.

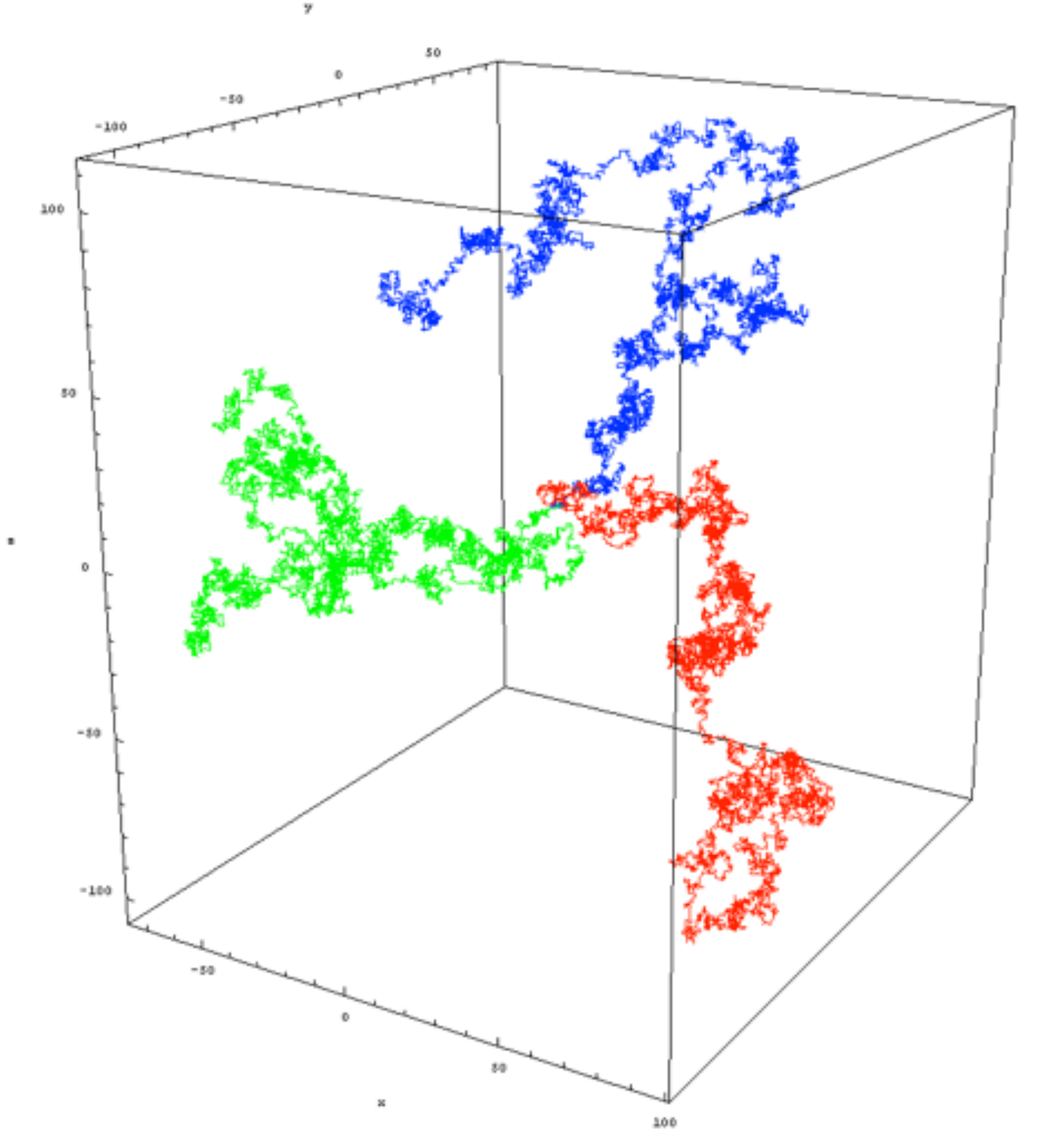
Some Examples of Markov Models

Example: Drunkard Walk



Salvador Dali (1922)
The Drunkard

Example: Diffusion Process



Example: Weather

A very(!) simplified model for the weather.

Probabilities on a daily basis:

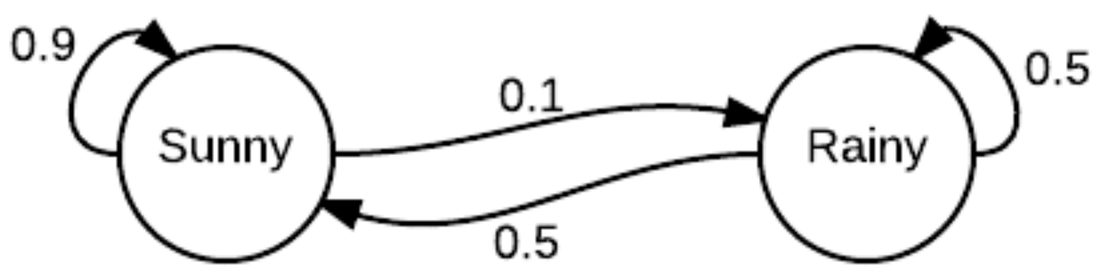
Pr(sunny to rainy) = 0.1

Pr(sunny to sunny) = 0.9

Pr(rainy to rainy) = 0.5

Pr(rainy to sunny) = 0.5

$$\begin{array}{c c} & \begin{array}{cc} \textbf{S} & \textbf{R} \end{array} \\ \begin{array}{c} \textbf{S} \\ \textbf{R} \end{array} & \begin{bmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{bmatrix} \end{array}$$



Encode more information about current state for a more accurate model.

Example: Life Insurance

Goal of insurance company:

figure out how much to charge the clients.

Find a model for how long a client will live.

Probabilistic model of health on a monthly basis:

Pr(healthy to sick) = 0.3
Pr(sick to healthy) = 0.8
Pr(sick to death) = 0.1
Pr(healthy to death) = 0.01
Pr(healthy to healthy) = 0.69
Pr(sick to sick) = 0.1
Pr(death to death) = 1

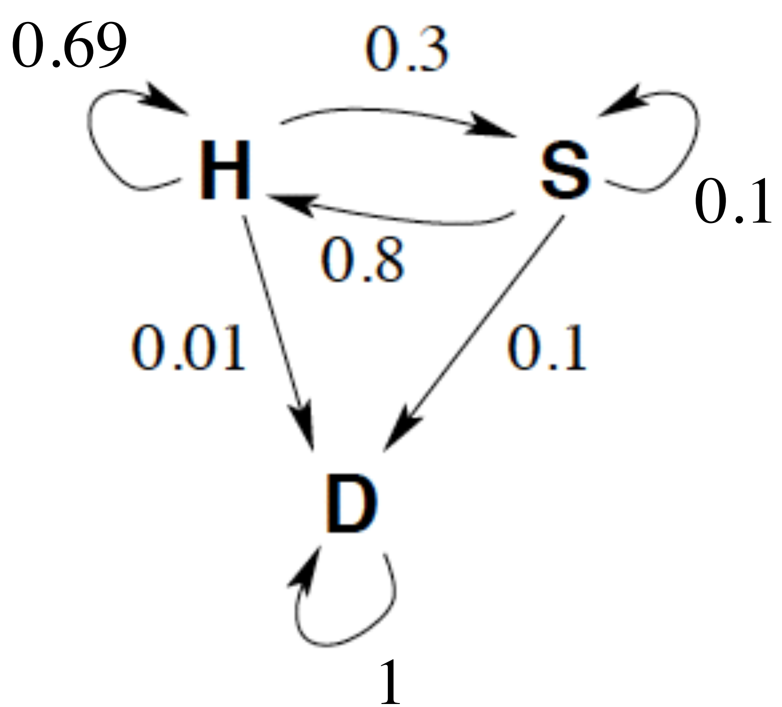
Example: Life Insurance

Goal of insurance company:

figure out how much to charge the clients.

Find a model for how long a client will live.

Probabilistic model of health on a monthly basis:

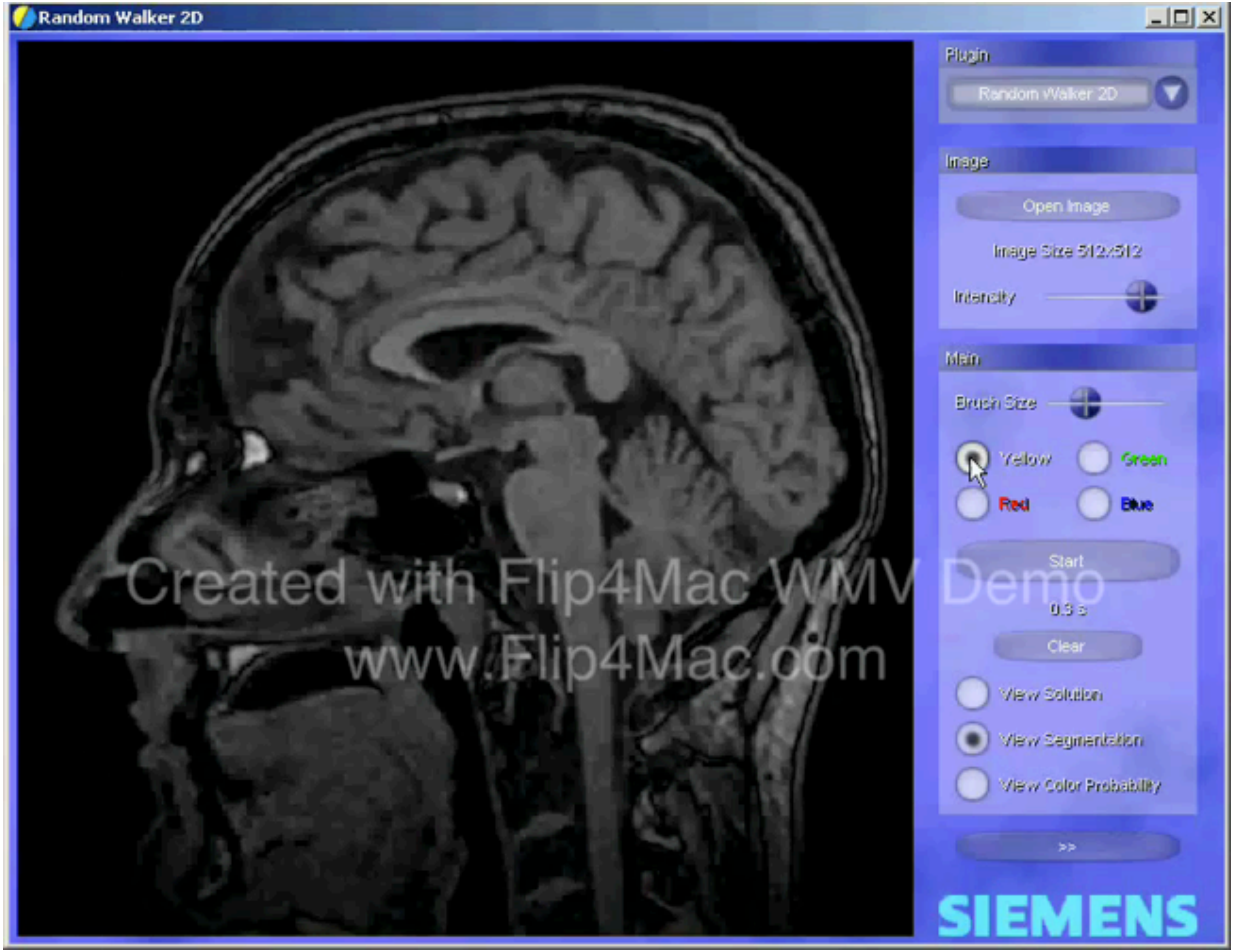


$$\begin{array}{cc} & \begin{array}{ccc} \mathbf{H} & \mathbf{S} & \mathbf{D} \end{array} \\ \begin{array}{c} \mathbf{H} \\ \mathbf{S} \\ \mathbf{D} \end{array} & \begin{bmatrix} 0.69 & 0.3 & 0.01 \\ 0.8 & 0.1 & 0.1 \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

Some Applications of Markov Models

Application: Algorithmic Music Composition

Application: Image Segmentation



Application: Automatic Text Generation

Random text generated by a computer
(putting random words together):

"You can't settle the issue. It seems I've forgotten what it is, but I don't. I know about violence against women, and I really doubt they will ever join together into a large number of jokes."

"While at a conference a few weeks back, I spent an interesting evening with a grain of salt."

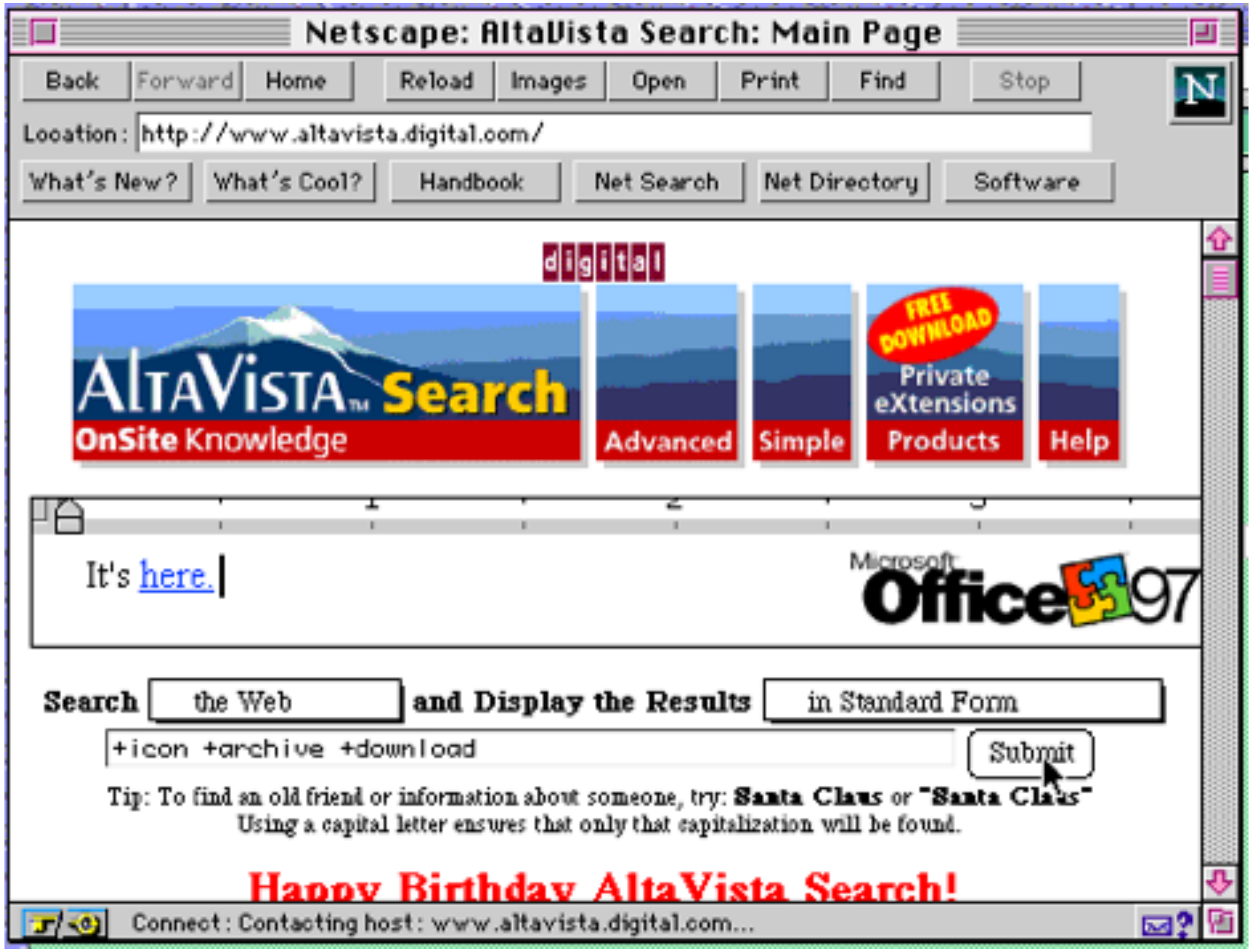
Google: Mark V Shaney

Application: Speech Recognition

Speech recognition software programs use Markov models to listen to the sound of your voice and convert it into text.

Application: Google PageRank

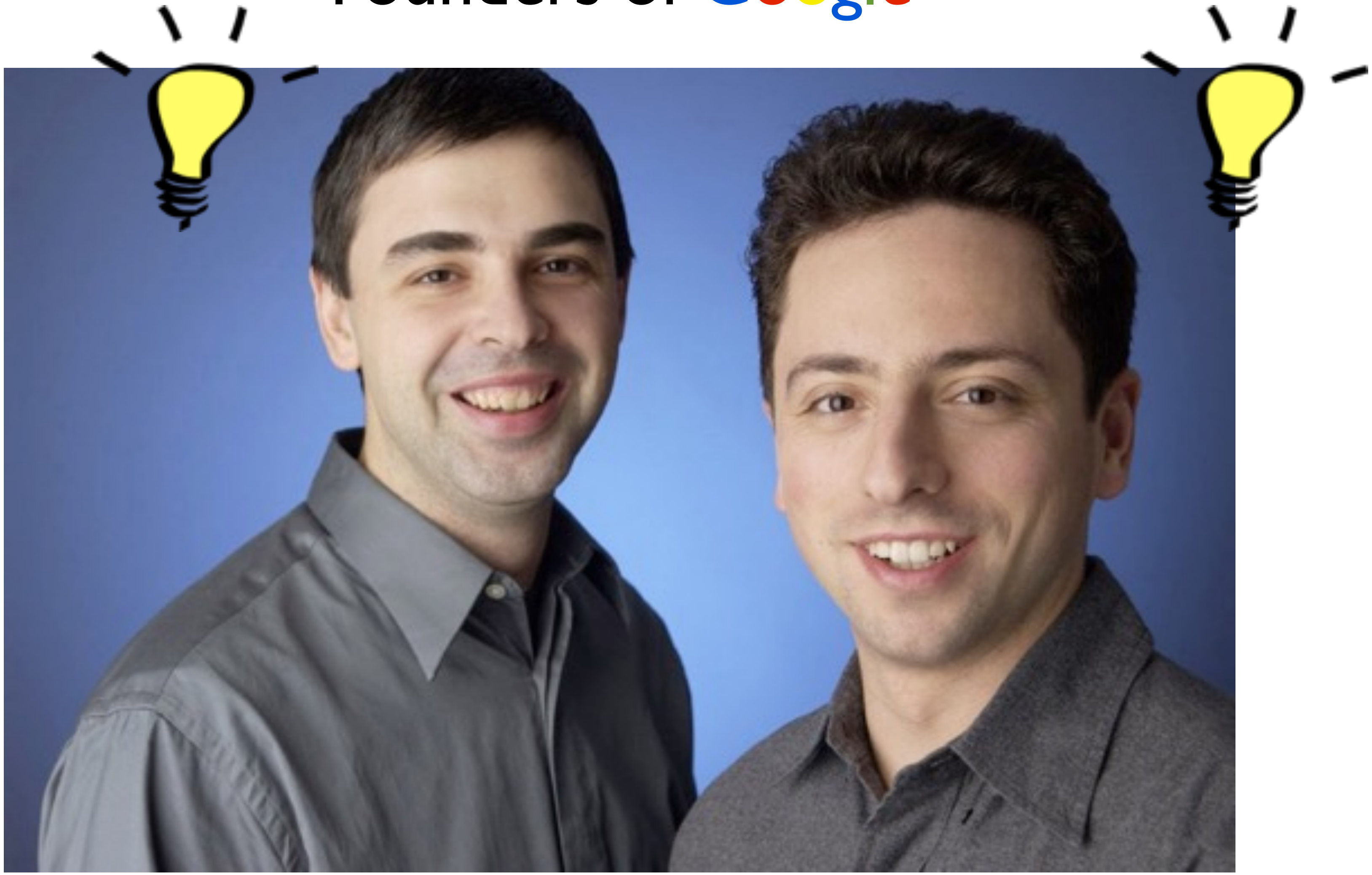
1997: Web search was horrible



You search "Garrett popcorn", finds all pages containing "Garrett popcorn" & sorts by number of occurrences.

Application: Google PageRank

Founders of Google



$20Billionaires

Application: Google PageRank

How does Google order the webpages displayed after a search?

2 important factors:

Relevance of the page.

Reputation of the page.

The number and reputation of links pointing to you.

Reputation is measured using PageRank.

PageRank is calculated using a Markov model.

Outline

Motivating examples and applications

Basic mathematical representation

Applications

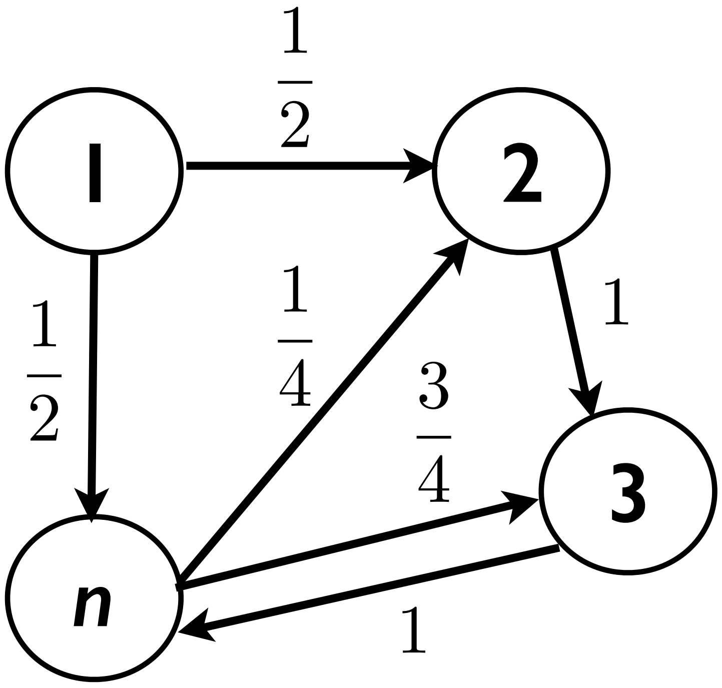
Outline

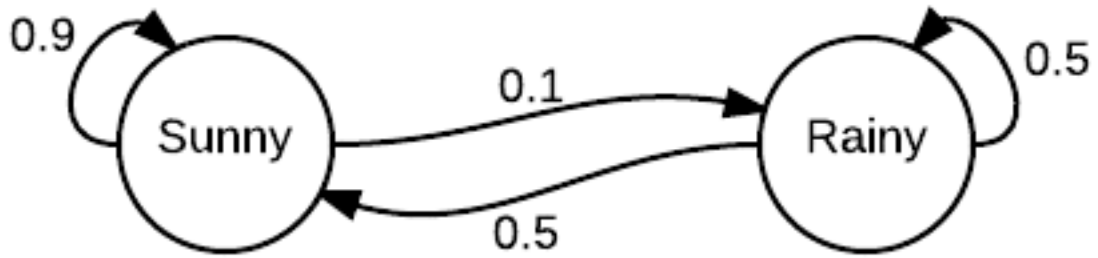
Basic mathematical representation

The Setting

There is a system with *n* possible states/values.

At each time step, the state changes probabilistically.

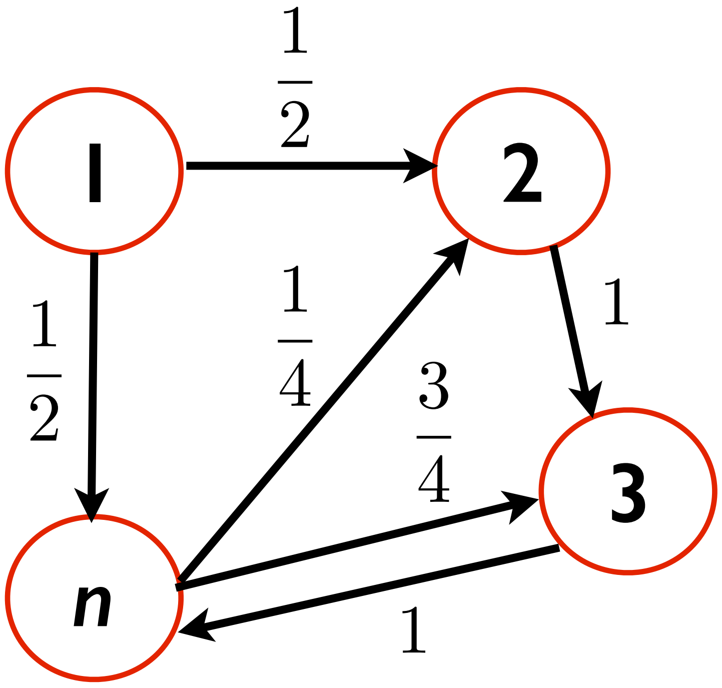


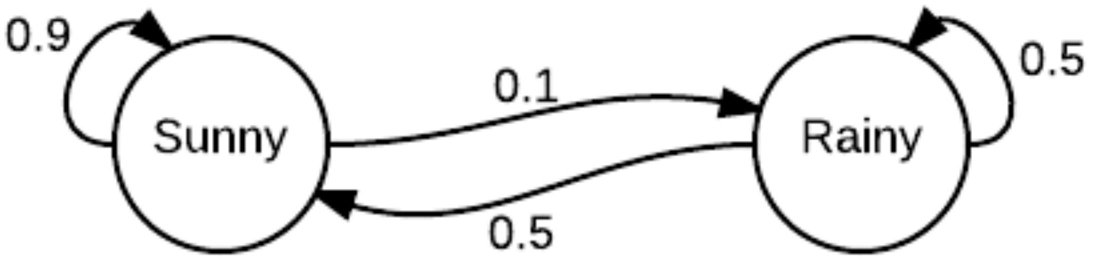


The Setting

There is a system with *n* possible states/values.

At each time step, the state changes probabilistically.



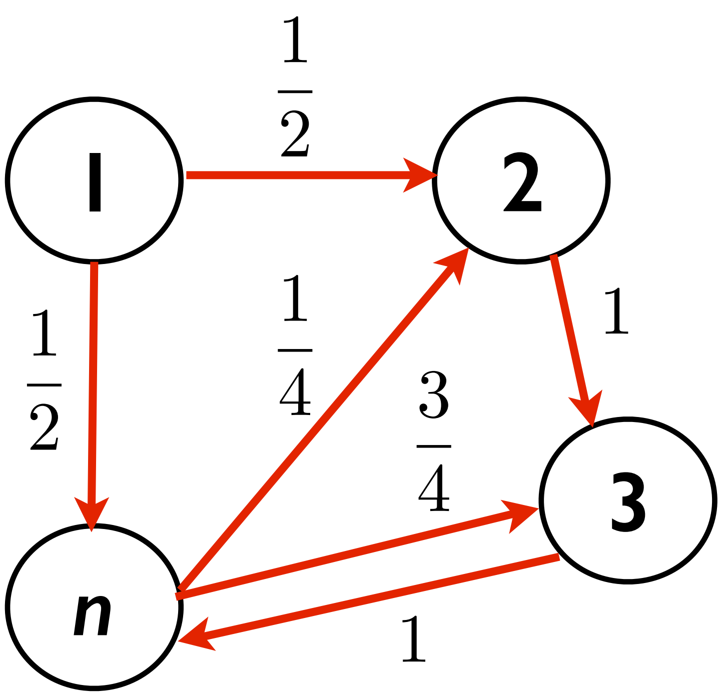


nodes

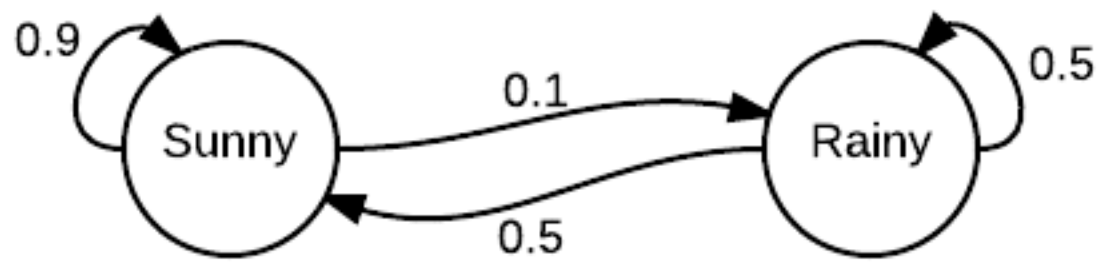
The Setting

There is a system with *n* possible states/values.

At each time step, the state changes probabilistically.



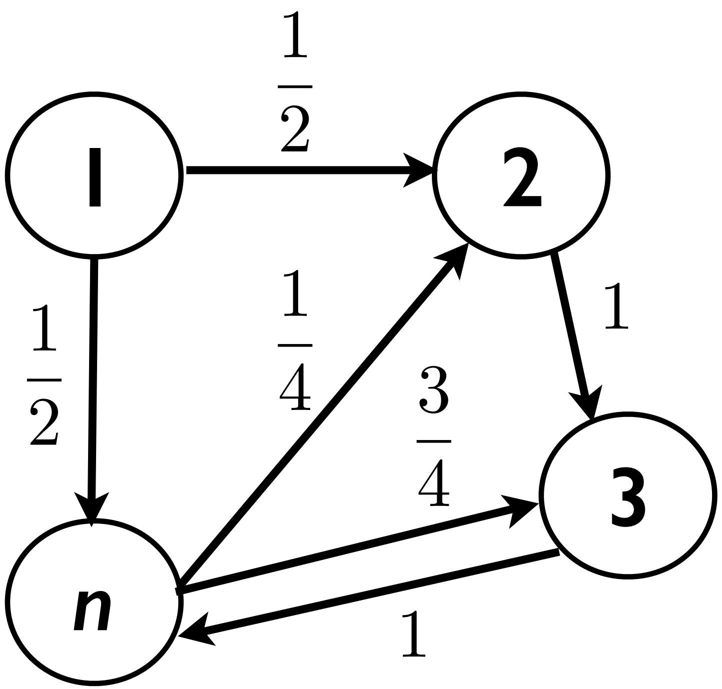
edges

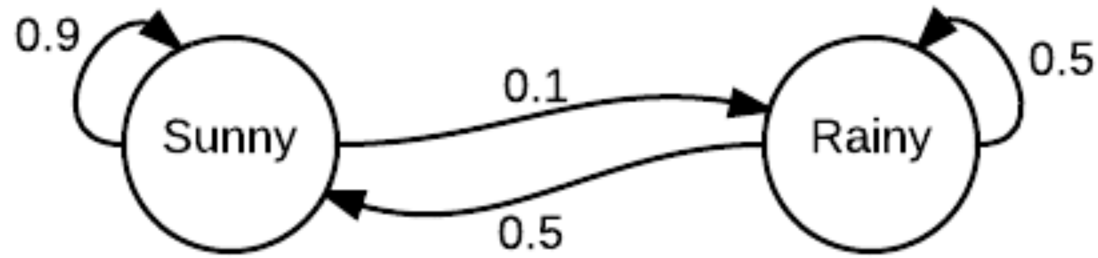


The Setting

There is a system with *n* possible states/values.

At each time step, the state changes probabilistically.





Memoryless

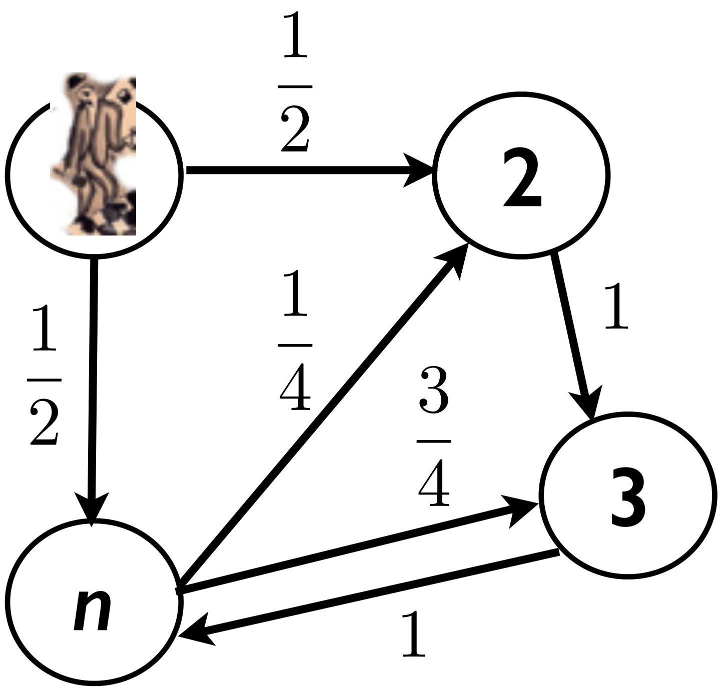
The next state only depends on the current state.

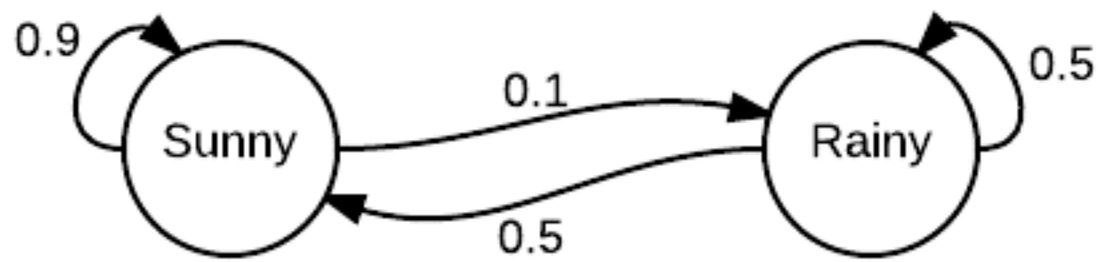
Evolution of the system: random walk on the graph.

The Setting

There is a system with *n* possible states/values.

At each time step, the state changes probabilistically.



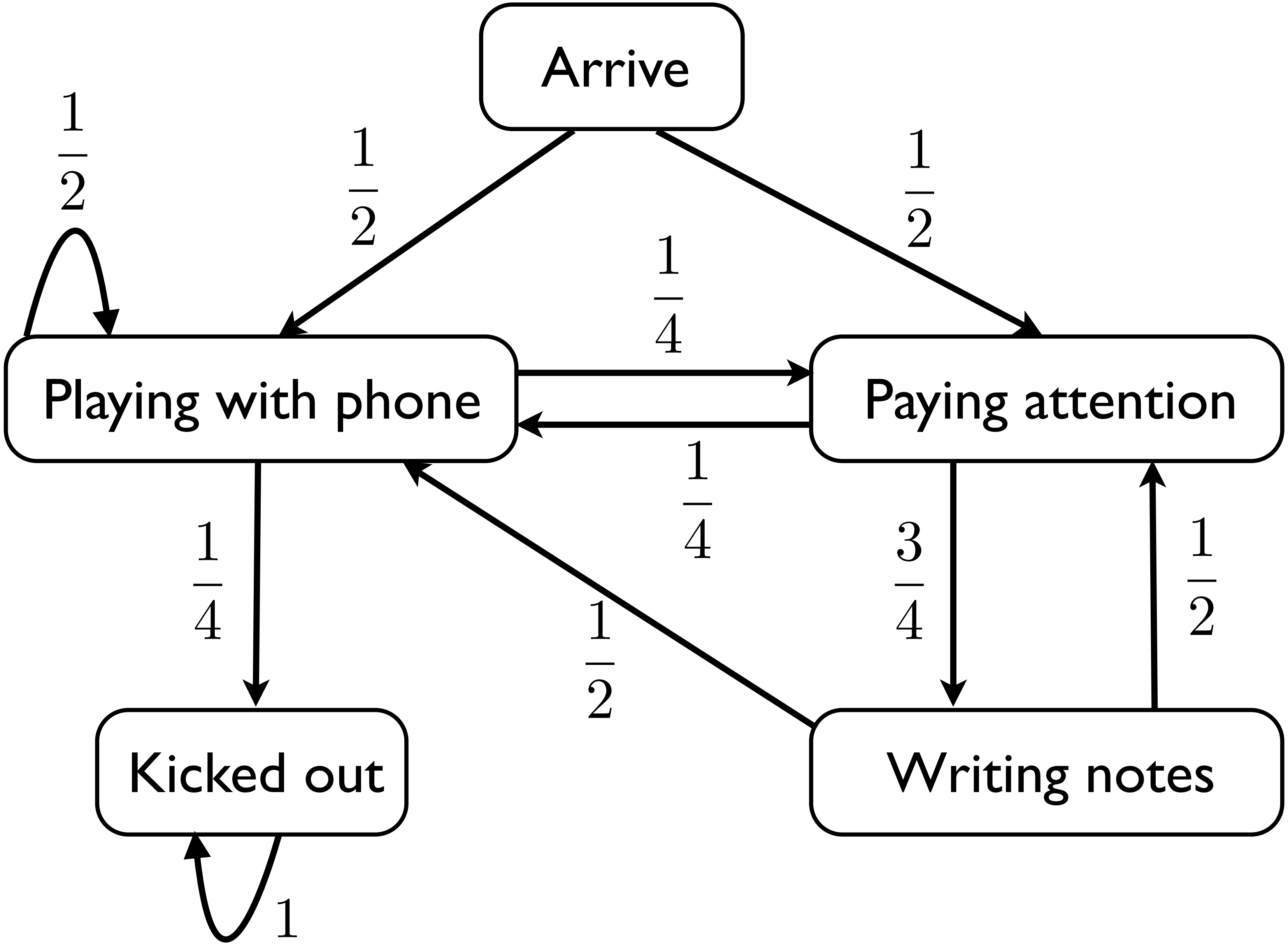


Memoryless

The next state only depends on the current state.

Evolution of the system: random walk on the graph.

Example: Markov Model for a Lecture



Some Natural Questions

What is the probability of being in state *i* after *t* steps?

What is the probability of visiting state *i* in at most *t* steps?

What is the expected time of having visited every state?

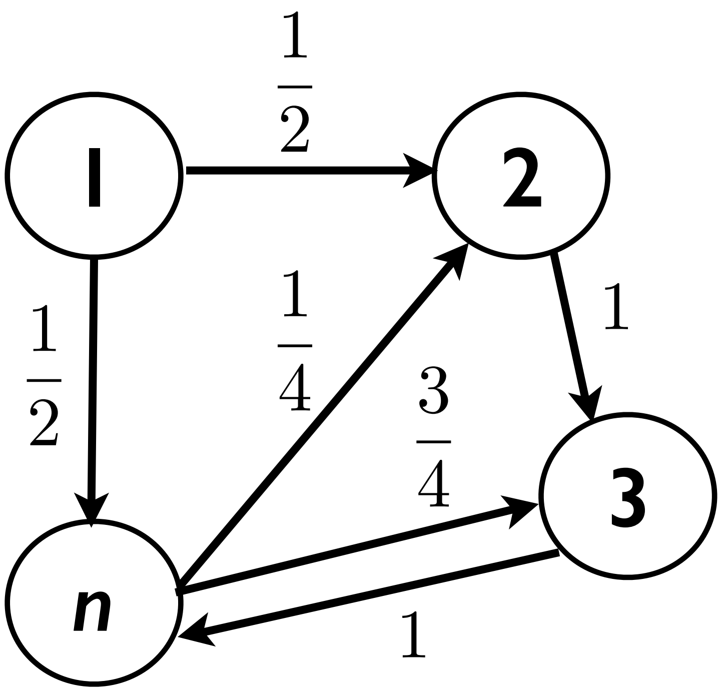
What is the expected time of the first return to state *i* when starting at state *i*?

…

Mathematical Formulation

Suppose we start at state **1** and let the system evolve.

How can we mathematically represent the evolution?



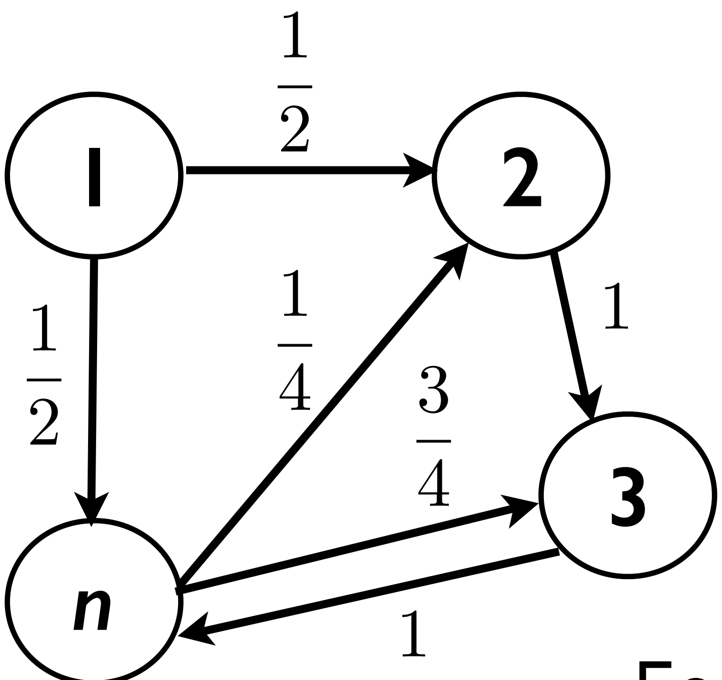
$$
\begin{array}{c c}
 & \begin{array}{cccc} \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} \end{array} \\
\begin{array}{c} \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \end{array} &
\begin{bmatrix}
0 & \frac{1}{2} & 0 & \frac{1}{2} \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & \frac{1}{4} & \frac{3}{4} & 0
\end{bmatrix}
\end{array}
$$

Transition Matrix

Mathematical Formulation

Suppose we start at state 1 and let the system evolve.

How can we mathematically represent the evolution?



$$\begin{array}{c|cccc} & \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} \\ \hline \mathbf{1} & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \mathbf{2} & 0 & 0 & 1 & 0 \\ \mathbf{3} & 0 & 0 & 0 & 1 \\ \mathbf{4} & 0 & \frac{1}{4} & \frac{3}{4} & 0 \end{array}$$

$$\begin{array}{cccc} \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} \\ 1 & 0 & 0 & 0 \end{array} \begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & \frac{1}{4} & \frac{3}{4} & 0 \end{bmatrix} = \begin{array}{cccc} \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{array}$$

Mathematical Formulation

The probability of states after 1 step:

$$\begin{array}{cccc} \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} \\ \left[1\right. & 0 & 0 & \left.0\right] \end{array} \begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & \frac{1}{4} & \frac{3}{4} & 0 \end{bmatrix} = \begin{array}{cccc} \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} \\ \left[0\right. & \frac{1}{2} & 0 & \left.\frac{1}{2}\right] \end{array}$$

the new state (probabilistic)

The probability of states after 2 steps:

$$\begin{array}{cccc} \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} \\ \left[0\right. & \frac{1}{2} & 0 & \left.\frac{1}{2}\right] \end{array} \begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & \frac{1}{4} & \frac{3}{4} & 0 \end{bmatrix} = \begin{array}{cccc} \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} \\ \left[0\right. & \frac{1}{8} & \frac{7}{8} & \left.0\right] \end{array}$$

the new state (probabilistic)

Mathematical Formulation

In general:

If the current probabilistic state is $\begin{bmatrix} p_1 & p_2 & \cdots & p_n \end{bmatrix}$

$p_i =$ probability of being in state *i* ,

$p_1 + p_2 + \cdots + p_n = 1$,

after *t* steps, the new probabilistic state is:

$$\begin{bmatrix} p_1 & p_2 & \cdots & p_n \end{bmatrix} \begin{bmatrix} \text{Transition} \\ \text{Matrix} \end{bmatrix}^{t}$$

Assignment:

Prove this.

Remarkable property of most Markov models

Suppose every state is reachable from any other state.

Suppose the Markov model is aperiodic.

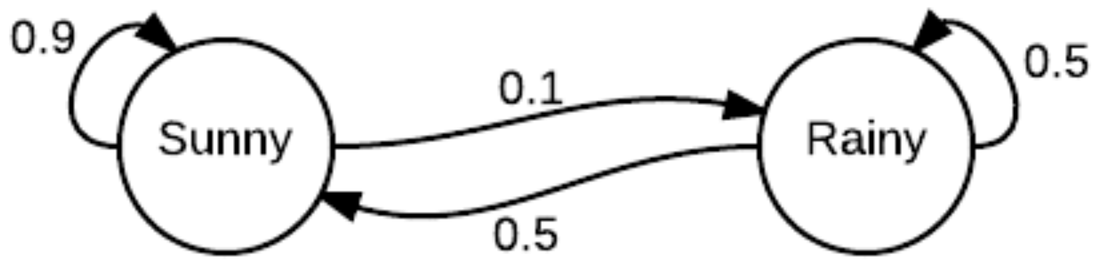
Then, as the system evolves, the probabilistic state **converges** to a limiting state.

As $t \to \infty$

$$\begin{bmatrix} p_1 & p_2 & \cdots & p_n \end{bmatrix} \begin{bmatrix} \text{Transition} \\ \text{Matrix} \end{bmatrix}^t \longrightarrow \begin{bmatrix} \pi_1 & \cdots & \pi_n \end{bmatrix}$$

stationary distribution

The stationary distribution example



Stationary distribution is

$$\begin{bmatrix} \frac{5}{6} & \frac{1}{6} \end{bmatrix}$$

In the long run, it is sunny 5/6 of the time,
it is rainy 1/6 of the time.

Summary

Markov model can be characterized by the transition matrix.

What is the probability of being in state *i* after *t* steps?
Can calculate using the transition matrix.

As *t* increases, the probability of being in state *i* converges to a fixed value.

(doesn't matter where you start)

Outline

Motivating examples and applications

Basic mathematical representation

Applications

Outline

Applications

How are Markov models applied ?

2 common types of applications

1. Build a Markov model as a statistical model of a real-world process.

   Use the Markov model to simulate the process.

   e.g. text generation, music composition.

2. Use a measure associated with a Markov model to approximate a quantity of interest.

   e.g. Google PageRank, image segmentation

How are Markov models applied ?

2 common types of applications

1. Build a Markov model as a statistical model of a real-world process.

   Use the Markov model to simulate the process.

   e.g. text generation, music composition.

2. Use a measure associated with a Markov model to approximate a quantity of interest.

   e.g. Google PageRank, image segmentation

Automatic Text Generation

Generate a superficially real-looking text given a sample document.

Idea:

From the sample document, create a Markov model.

Use a random walk on the Markov model to generate text.

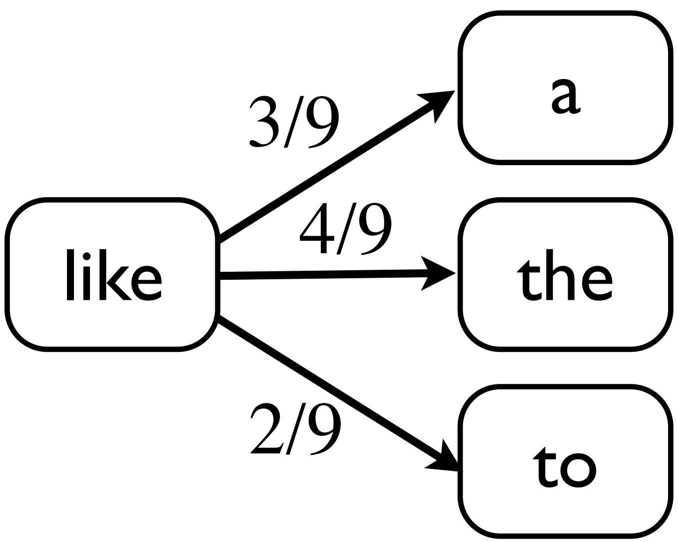
Automatic Text Generation

From the sample document, create a Markov model.

For each word in the document, create a node/state.

Put an edge word1 ---> word2
if there is a sentence in which word2 comes after word1.

Edge probabilities reflect frequency of the pair of words.



like a	3 times
like the	4 times
like to	2 times

Automatic Text Generation

Another use:

Build a Markov model based on speeches of Obama.

Build a Markov model based on speeches of Bush.

Given a new quote, can predict if it is by Obama or Bush.

(by testing which Markov model the quote fits best)

Assignment: Can you find a way to modify the Markov model to improve its performance?

Image Segmentation

Simple version

Given an image of an object, figure out:
which pixels correspond to the object,
which pixels correspond to the background

i.e., label each pixel "object" or "background"

User labels a small number of pixels with known labels

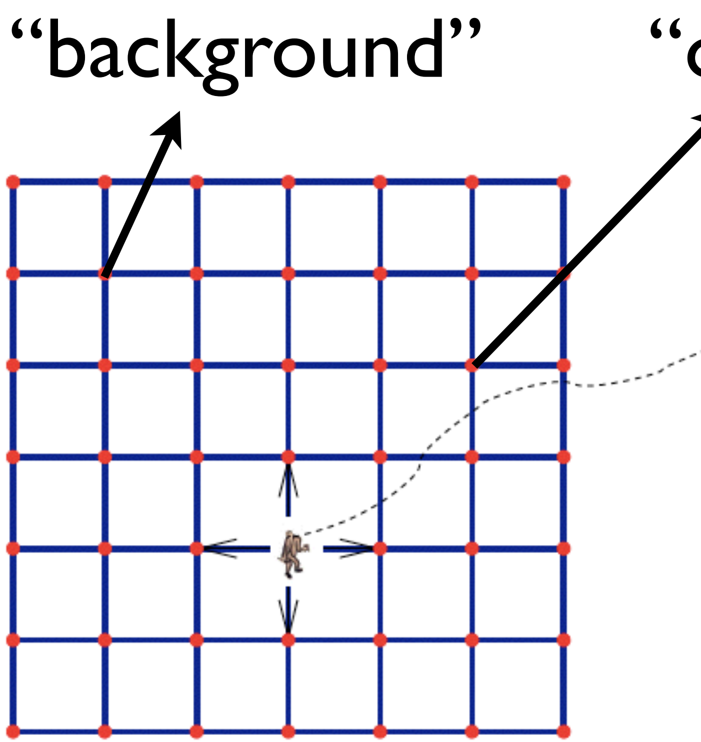
Image Segmentation

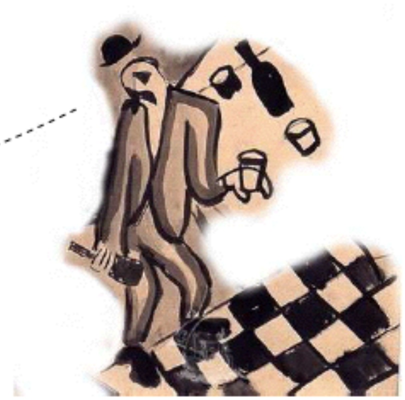
<u>Underlying Markov Model</u>

Each pixel is a node/state.

There is an edge between adjacent pixels.

Edge probabilities reflect similarity between pixels.





Salvador Dali (1922)
The Drunkard

Which one is more likely:

random walker first visits
“background”
or
“object” ?

Google PageRank

PageRank is a measure of reputation:

The number and importance of links pointing to you.

<u>The Markov Model</u>

Every webpage is a node/state. (In total n webpages)

Each hyperlink is an edge.

if webpage A has a link to webpage B, A ---> B

If A has m outgoing edges, each gets label $1/m$

If A has no outgoing edges, put an edge A ---> B for all B

Google PageRank

We want to make sure that the Markov model has a stationary distribution.

Tweak the model slightly to ensure this.

Stationary distribution:

probability of being in state i in the long run

PageRank of a webpage

=

The stationary probability corresponding to the webpage

Google:

"*PageRank continues to be the heart of our software*"

How are Markov models applied ?

2 common types of applications

1. Build a Markov model as a statistical model of a real-world process.

   Use the Markov model to simulate the process.

   e.g. text generation, music composition.

2. Use a measure associated with a Markov model to approximate a quantity of interest.

   e.g. Google PageRank, image segmentation

Outline

Motivating examples and applications

Basic mathematical representation

Applications

Take-Home Message

Markov Models

A random system that undergoes transitions from one state to another.

The next state only depends on the current state. (and not on the sequence of events preceding it)

A simple and neat model with beautiful math underlying it.

Extraordinary number of applications in various areas.