01 December 2022

UNIVERSITY OF GLOUCESTERSHIRE

CT4031
Maths for Data Science

Week 8 – Probability Part 4

Markov chains

- A Markov model is a stochastic model to represent randomly changing systems.
- Assumes that future states depend only on the present state and not on the sequence of events.

Stochastic model

- A way to calculate probability distributions by allowing for random variation over time.
- Based on fluctuations in historical data using time-series techniques.

Assumptions for a Markov model

- Fixed set of states;
- Fixed transition probabilities;
- Possibility of getting from any state to another through transitions;
- What happens in the long run doesn't depend on where the process started (or the way taken);

Markov chain

- States and transitions.
- The probability of a next state depends ONLY on the current state and NOT on a sequence of states.

Markov chain – Example 1 (boring)

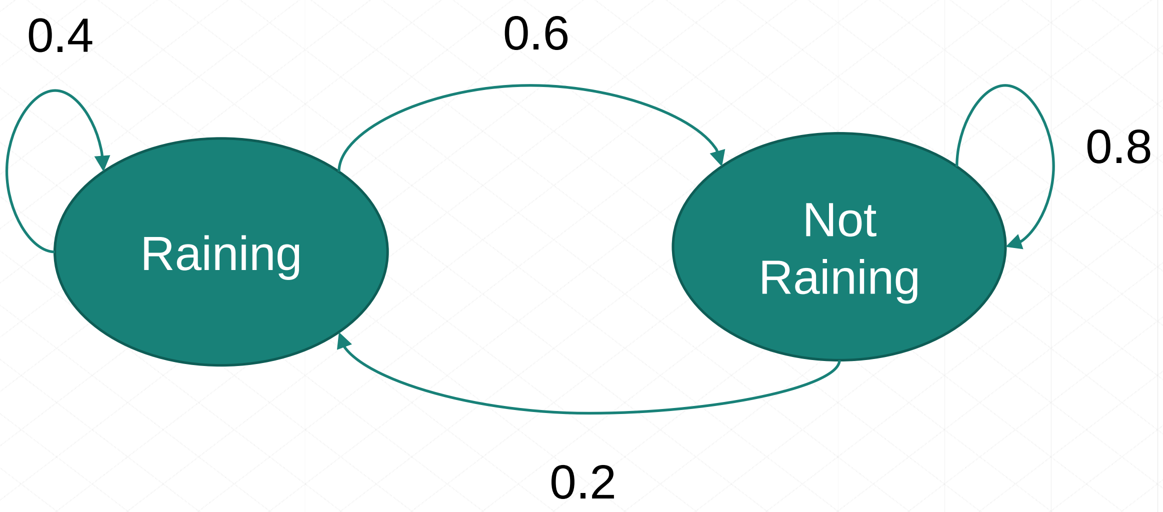
Weather

Raining today => Raining tomorrow

Raining today	60% Raining tomorrow 40% Not raining tomorrow
Not raining today	20% Raining tomorrow 80% Not raining tomorrow

Markov chain – Example 1 (boring)

Raining today	40% Raining tomorrow 60% Not raining tomorrow
Not raining today	20% Raining tomorrow 80% Not raining tomorrow



Markov chain – Example 1 (boring)

A Markov Process is usually represented in a transition matrix

$$\begin{array}{cc} \text{Raining} & \text{Not Raining} \end{array}$$

$$P = \begin{pmatrix} 0.4 & 0.6 \\ 0.2 & 0.8 \end{pmatrix}$$

Markov chain – Example 1 (boring)

A Markov Process is usually represented in a transition matrix

Rows sum up to 1, covering the 100%

$$P = \begin{array}{c} \\ \\ \end{array} \begin{array}{cc} \text{Raining} & \text{Not Raining} \\ \end{array}$$

$$P = \begin{pmatrix} 0.4 & 0.6 \\ 0.2 & 0.8 \end{pmatrix} \begin{array}{l} = 1 \\ = 1 \end{array}$$

Markov chain – Example 1 (boring)

- Let X_i be the weather of the day *i.*
- The probability of X_{i+1} will depend only on the probability of $X_{i.}$
- Markov property
- $P(X_{i+1} \mid X_{1,} X_{2,} \dots X_n) = P(X_{i+1} \mid X_i)$
- Stationary assumption
  - The transition probabilities are independent of time.

Markov chain – Example 2 (much more interesting one)

Food!

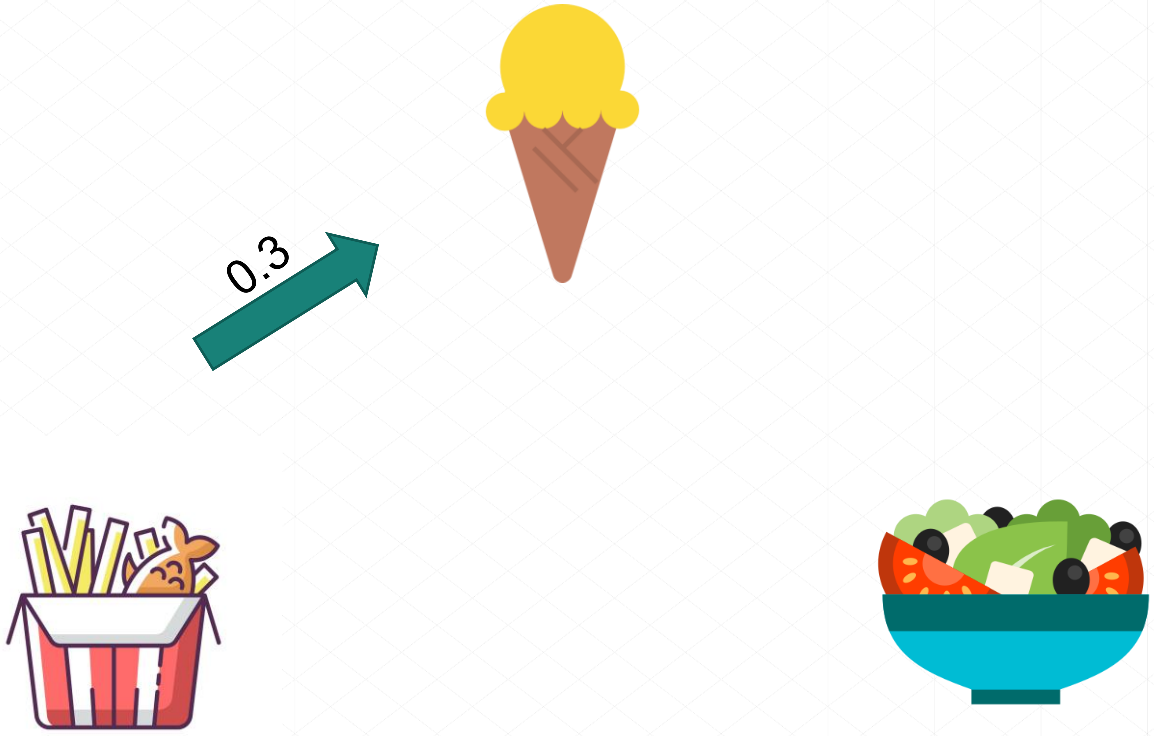
Markov chain – Example 2 (much more interesting one)

Imagine that a place serves 3 types of food, and only one type a day:

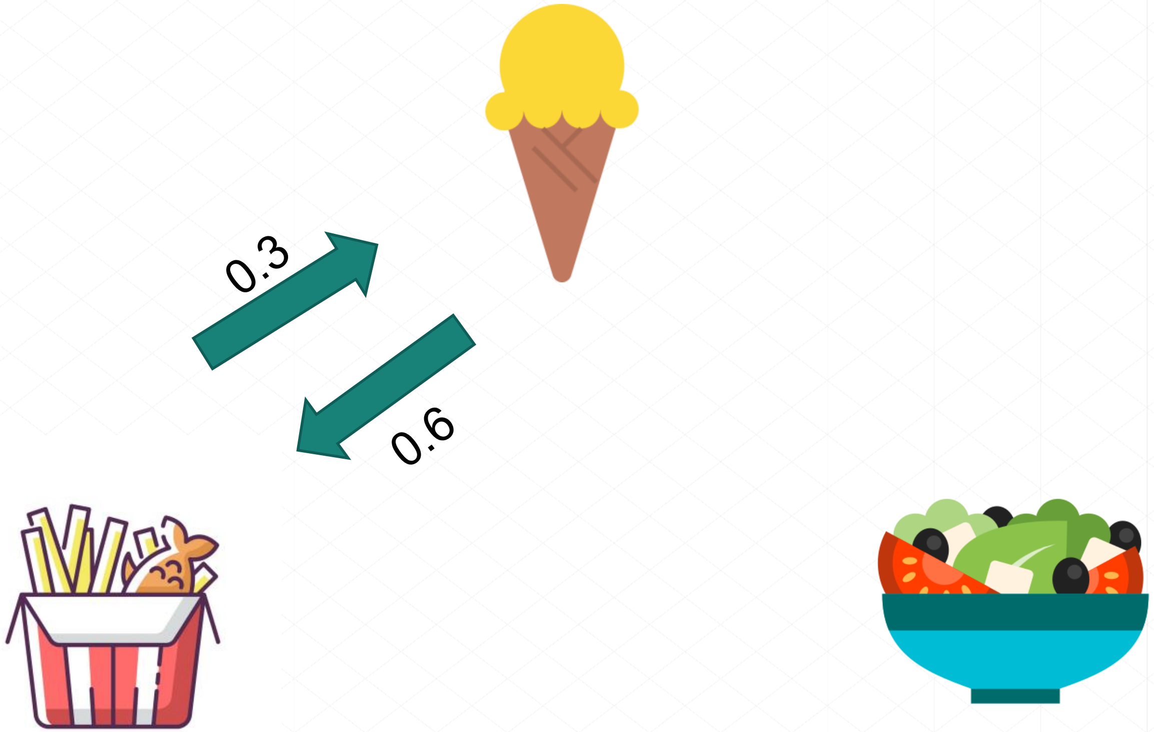
Markov chain – Example 2 (much more interesting one)

Food served one day impacts the probability of the food served on the following day

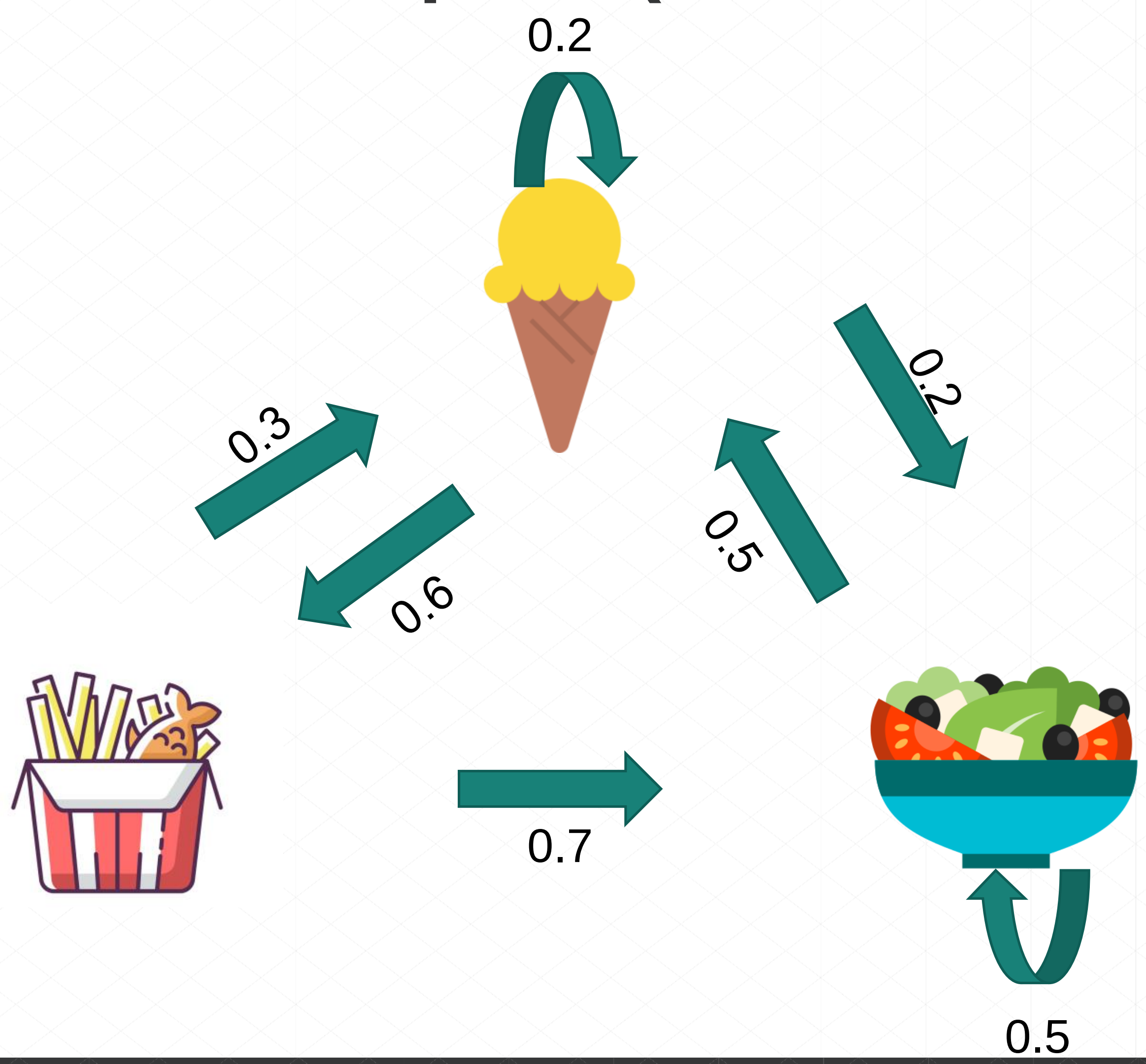
Markov chain – Example 2 (much more interesting one)



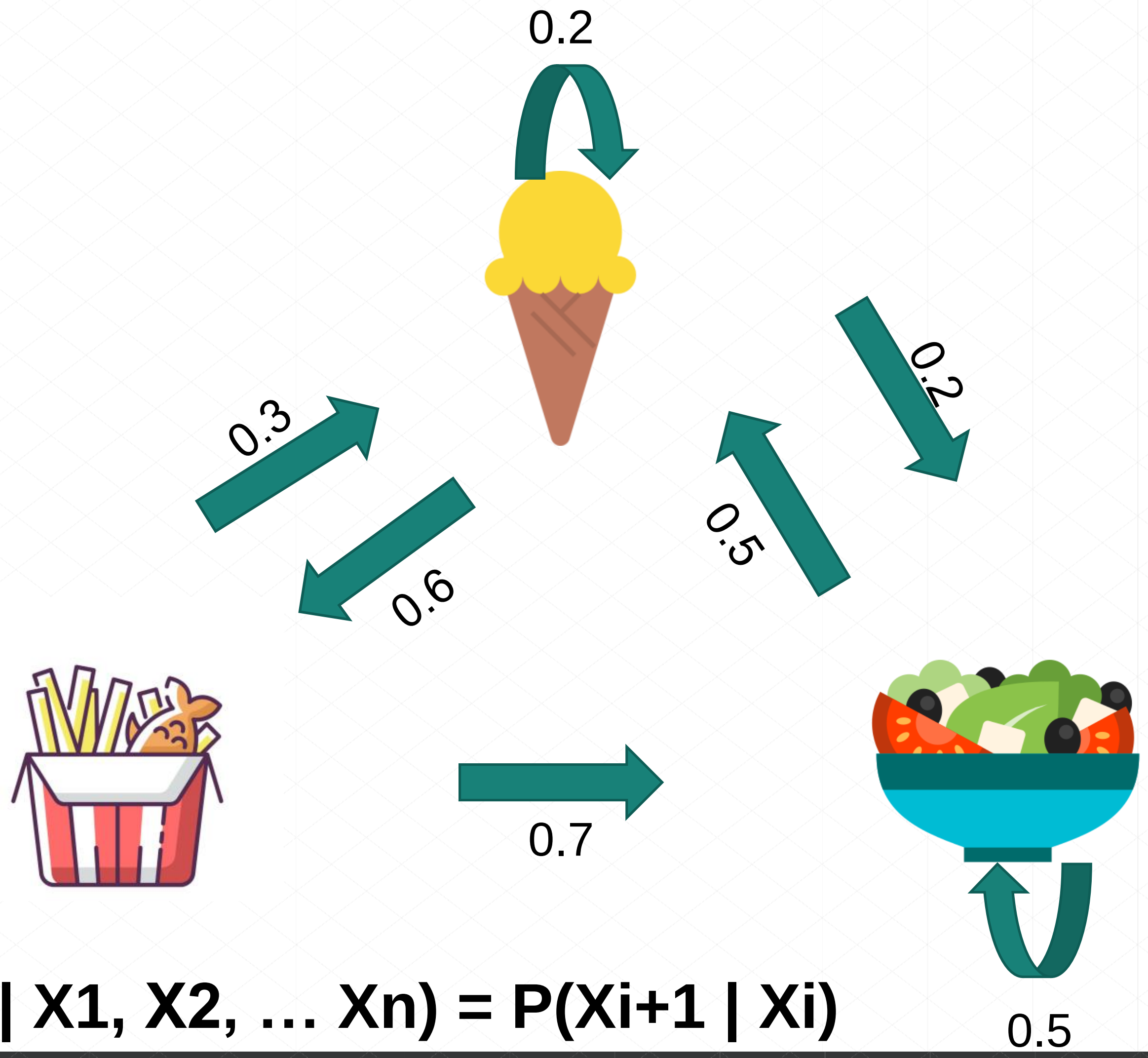
Markov chain – Example 2 (much more interesting one)



Markov chain – Example 2 (much more interesting one)

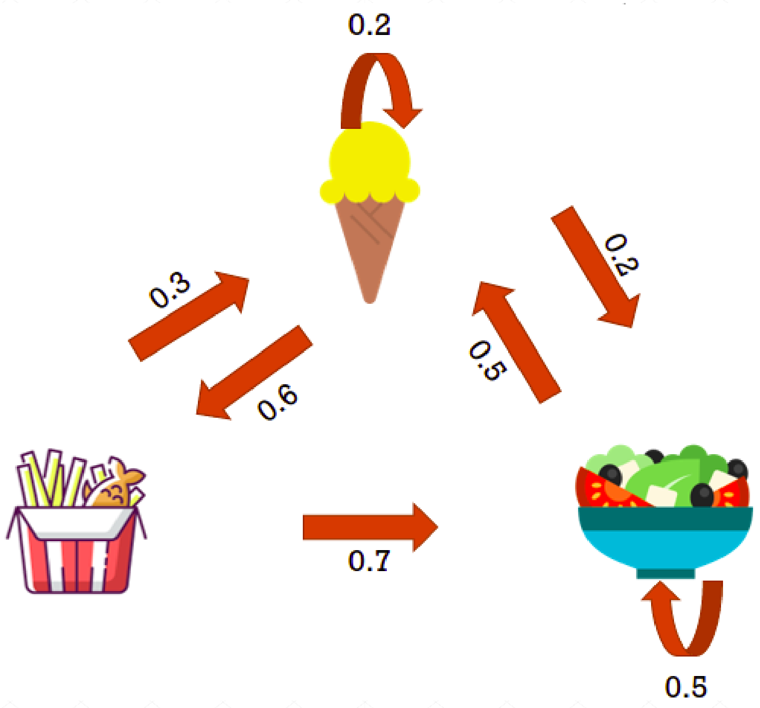


Markov chain – Example 2 (much more interesting one)



P(Xi+1 | X1, X2, … Xn) = P(Xi+1 | Xi)

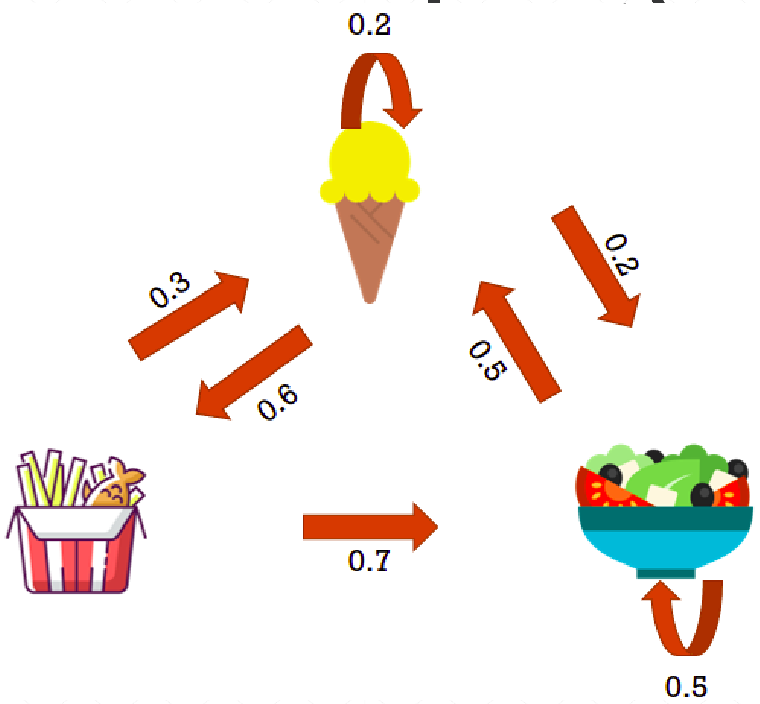
Markov chain – Example 2 (much more interesting one)



Knowing this, how can we predict the probability of the next one being salad?

Markov chain – Example 2 (much more interesting one)

UNIVERSITY OF GLOUCESTERSHIRE

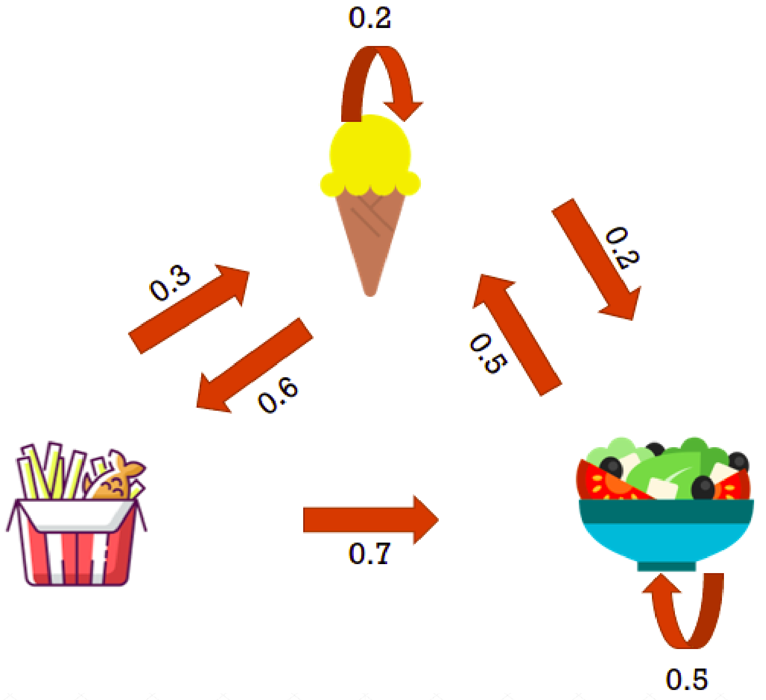


Knowing this, how can we predict the probability of the next one being salad?

P(X4=  |X1=  , X2=  , X3= )

Markov chain – Example 2 (much more interesting one)

UNIVERSITY OF GLOUCESTERSHIRE

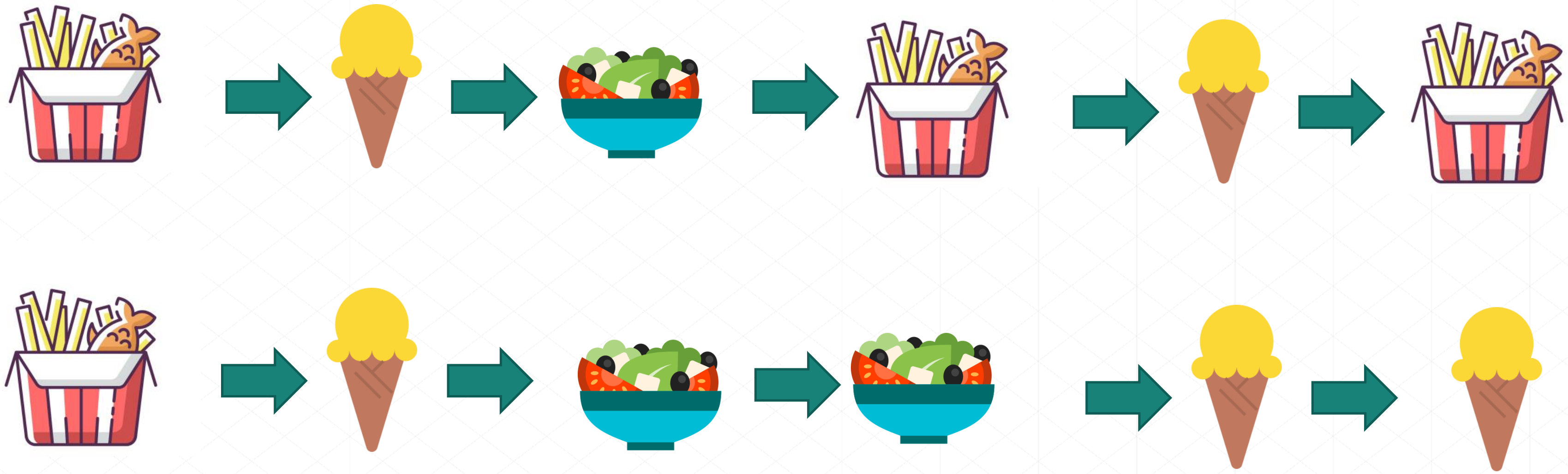


The Markov property:

P(X4=  |X3= )

Markov chain – Example 2 (much more interesting one)

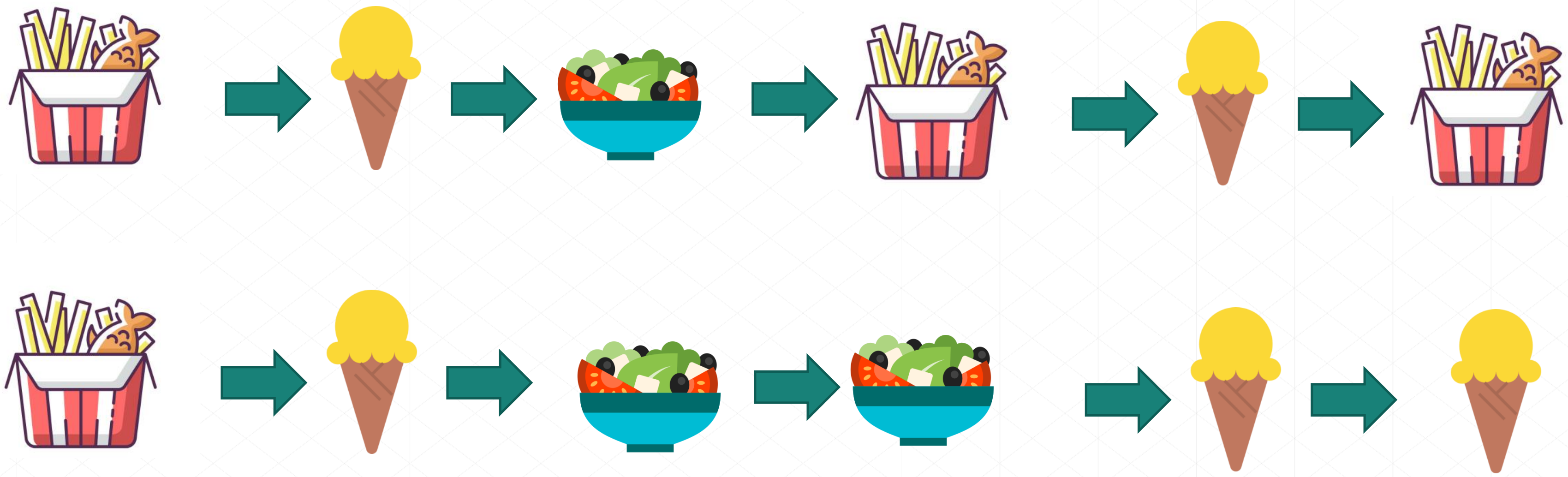
Imagine a Random walk on the model:



Markov chain – Example 2 (much more interesting one)

UNIVERSITY OF GLOUCESTERSHIRE

Imagine a Random walk on the model:



P() =

= 4/12

P() =

= 5/12

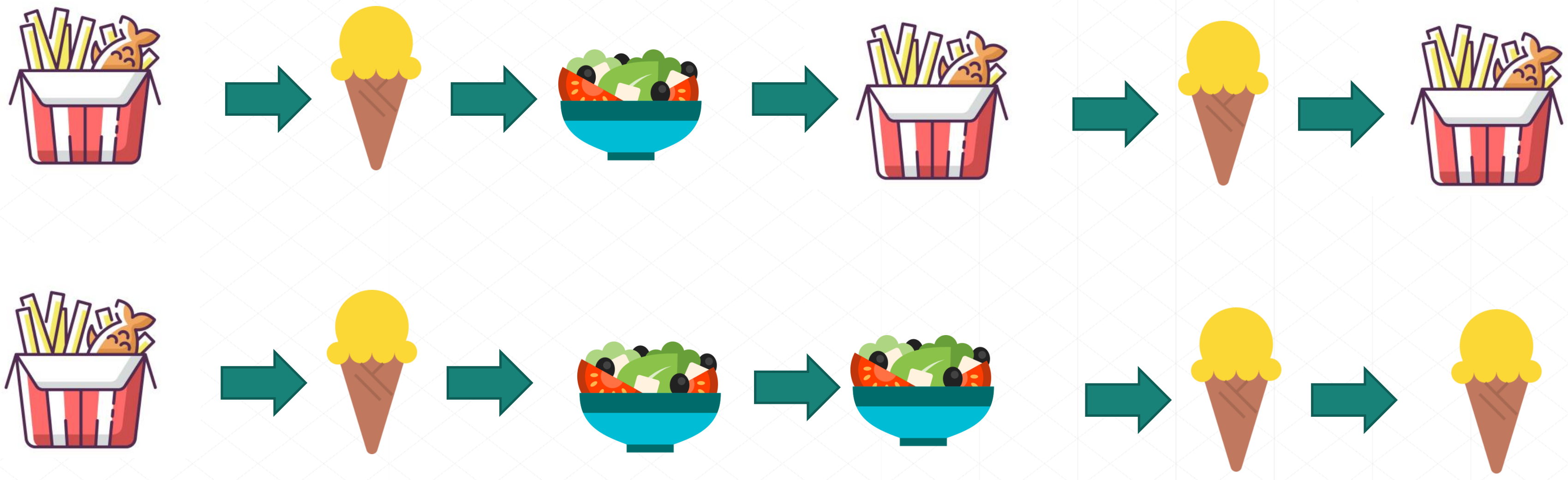
) =

=3/12

Markov chain – Example 2 (much more interesting one)

UNIVERSITY OF GLOUCESTERSHIRE

Imagine a Random walk on the model:



P() =

= 33%

P() =

= 42%

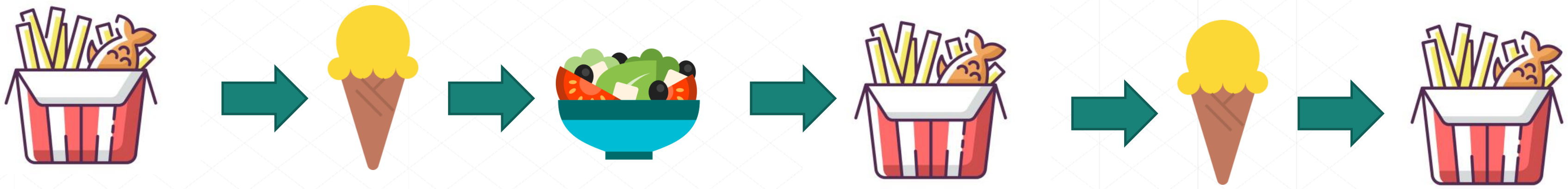
) =

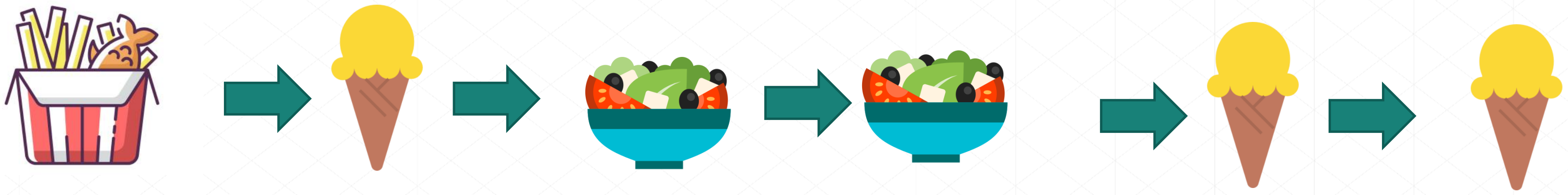
=25%

Markov chain – Example 2 (much more interesting one)

UNIVERSITY OF GLOUCESTERSHIRE

Imagine a Random walk on the model:





P() = 33%

WRONG!!!

P( = 42%

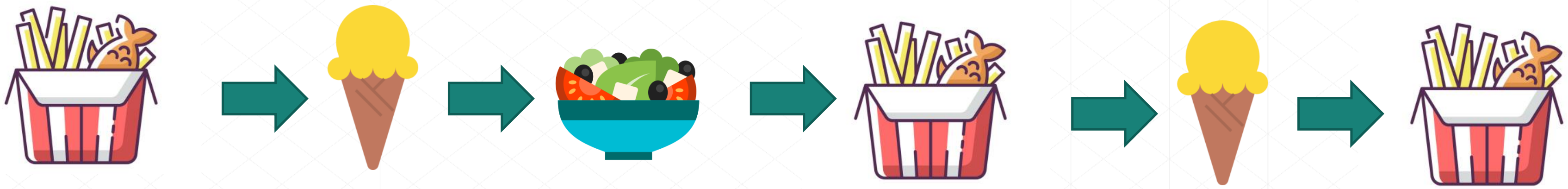
WRONG!!!

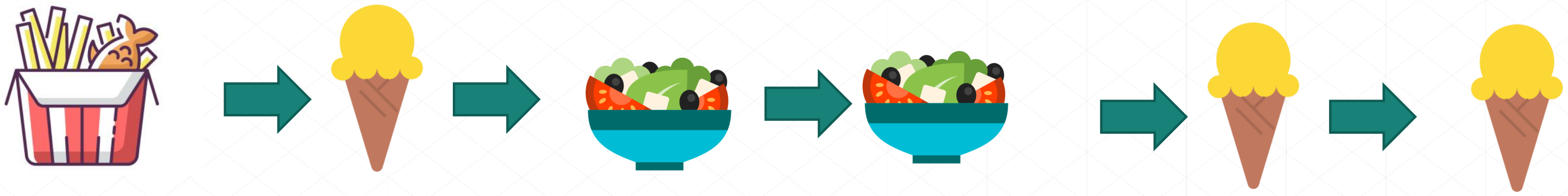
 =25%

WRONG!!!

Markov chain – Example 2 (much more interesting one)

Imagine infinite steps (*back to limits and calculus*)





P() =

= 0.351

P() =

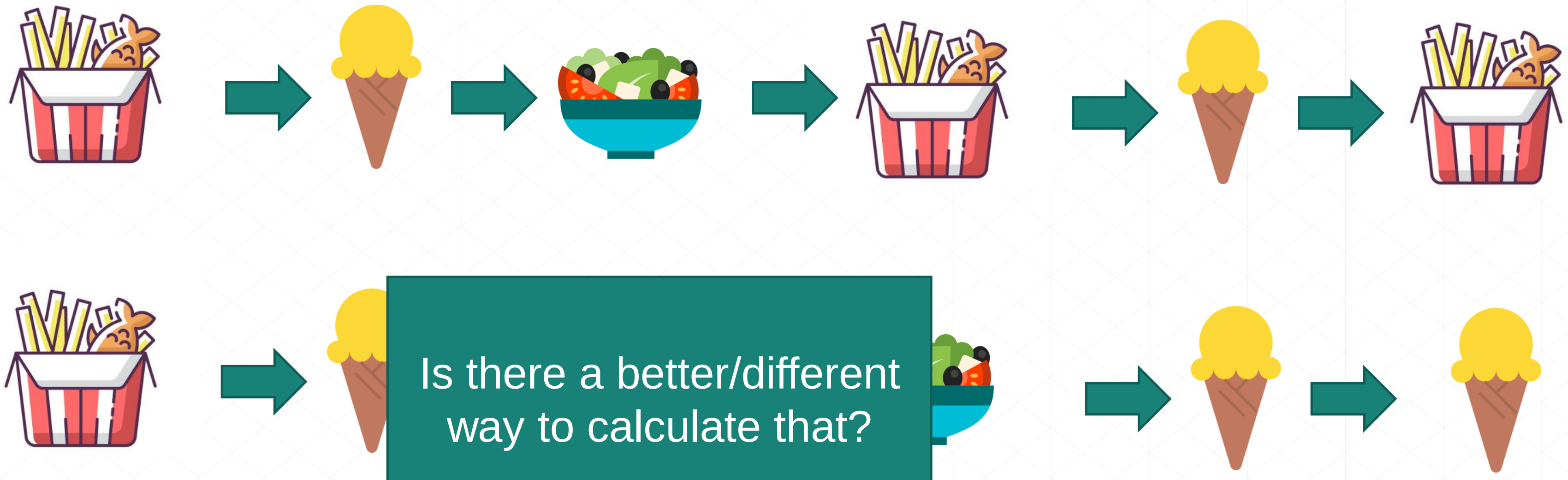
= 0.212

) =

=0.435

Markov chain – Example 2 (much more interesting one)

Imagine infinite steps (*back to limits and calculus*)



Is there a better/different way to calculate that?

P() =

= 0.351

P() =

= 0.212

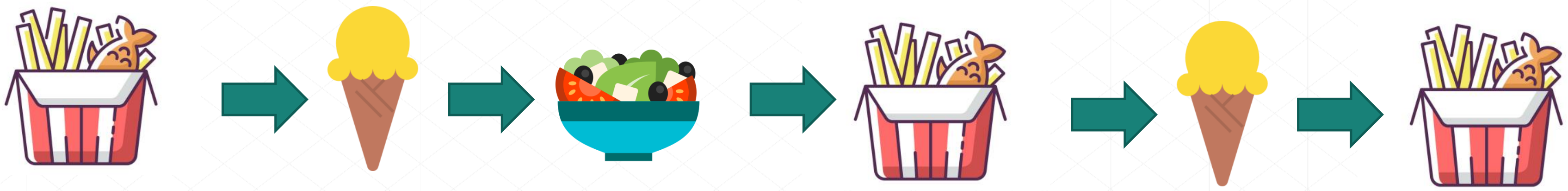
) =

=0.435

Markov chain – Example 2 (much more interesting one)

UNIVERSITY OF GLOUCESTERSHIRE

Imagine infinite steps (*back to limits and calculus*)

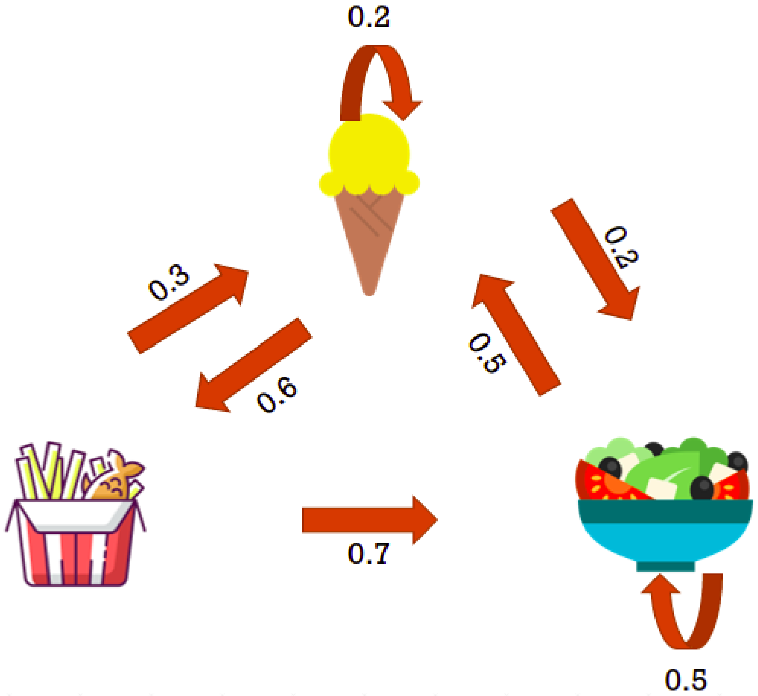


P() = = 0.351

P() = = 0.212

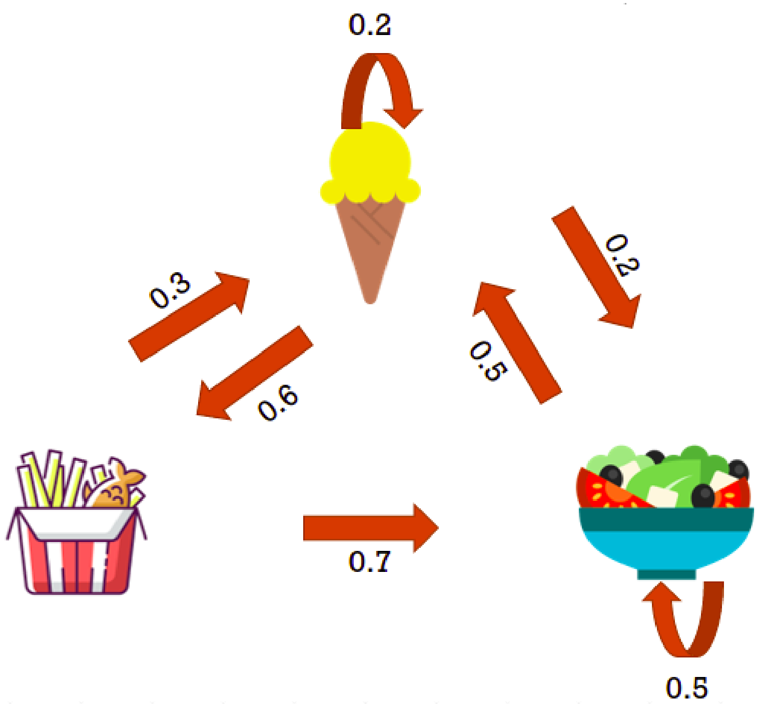
) = =0.435

Markov chain – Example 2 (much more interesting one)



$$\begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix}$$

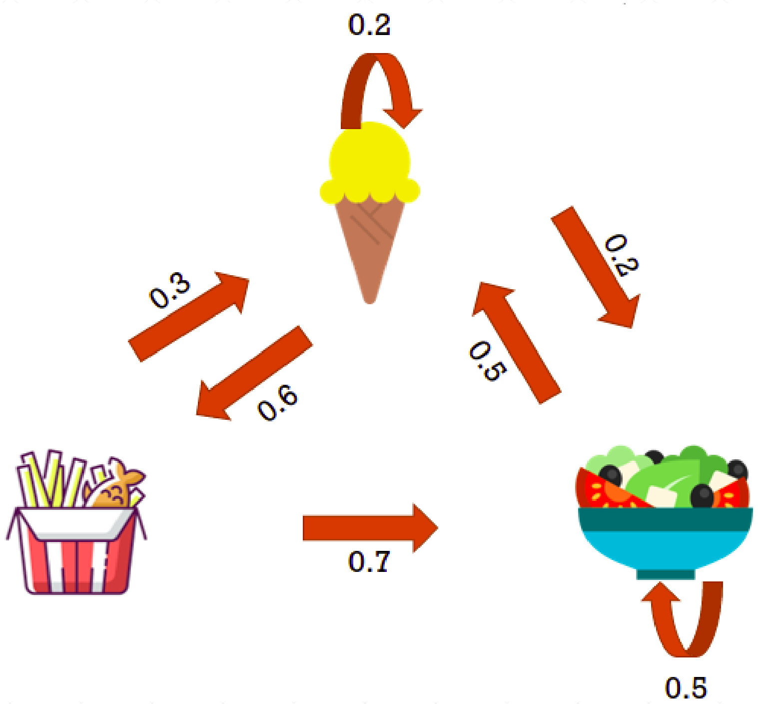
Markov chain – Example 2 (much more interesting one)



$$\begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix}$$

The answer is 

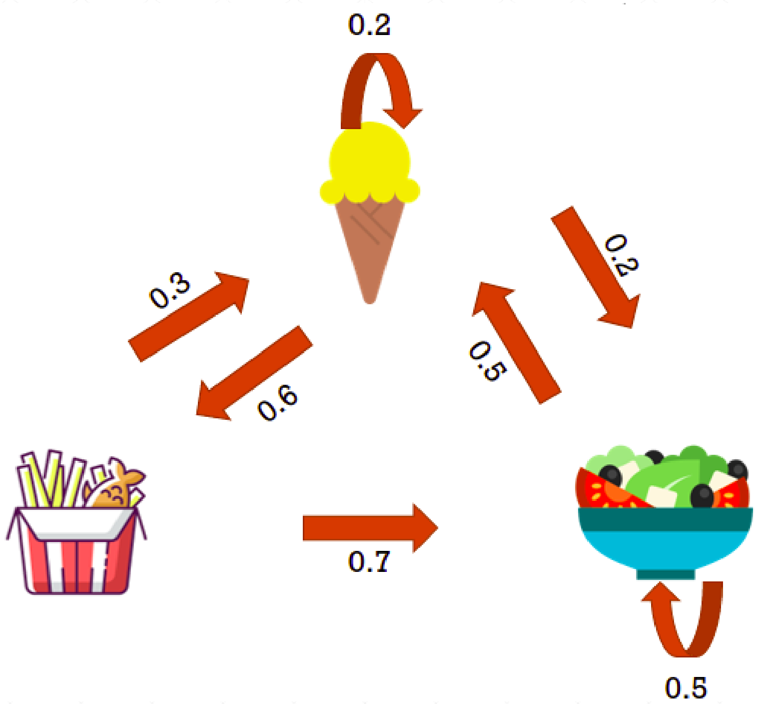
Markov chain – Example 2 (much more interesting one)



0.2	0.6	0.2
0.3	0.0	0.7
0.5	0.0	0.5

Not this one...

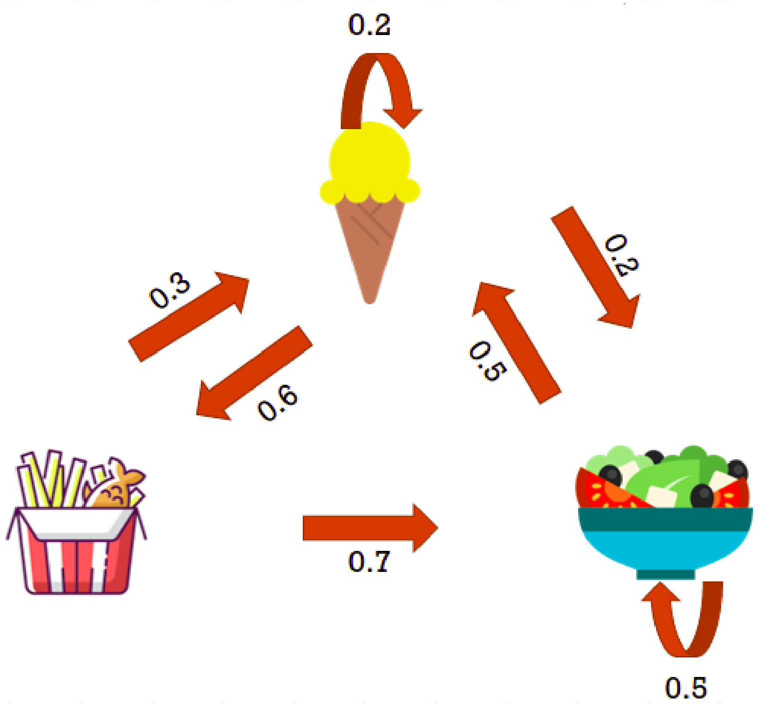
Markov chain – Example 2 (much more interesting one)



$$\begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix}$$

This one π

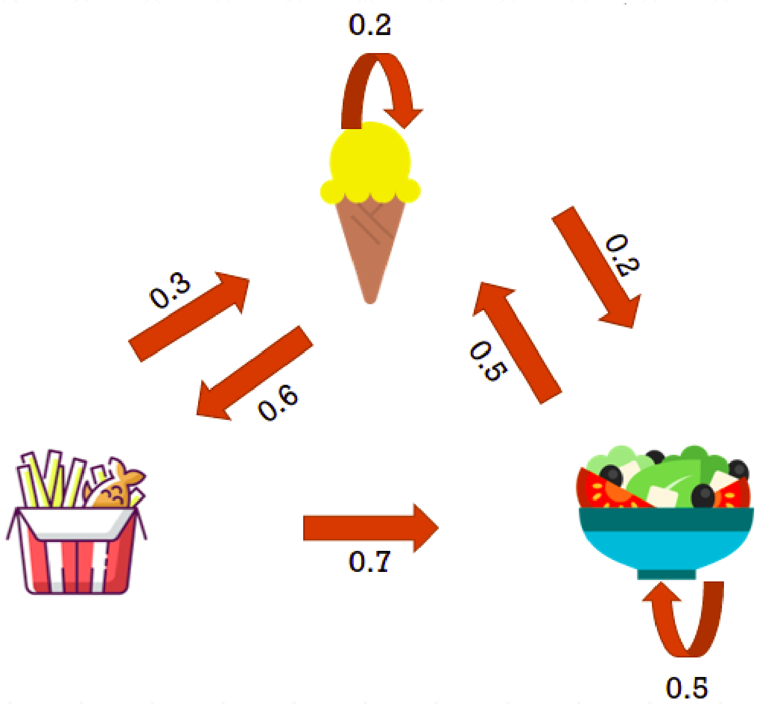
Markov chain – Example 2 (much more interesting one)



$$\begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix}$$

$$\pi \begin{bmatrix} 0.0 & 1.0 & 0.0 \end{bmatrix}$$

Markov chain – Example 2 (much more interesting one)



$$\begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix}$$

$$\pi \begin{bmatrix} 0.0 & 1.0 & 0.0 \end{bmatrix}$$

Fish and chips days

$$\begin{array}{cc} & \begin{array}{ccc} \text{🍦} & \text{🍟} & \text{🥗} \end{array} \\ \begin{array}{c} \text{🍦} \\ \text{🍟} \\ \text{🥗} \end{array} & \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix} \end{array}$$

$$\pi \begin{bmatrix} 0.0 & 1.0 & 0.0 \end{bmatrix}$$

$$\begin{bmatrix} 0.0 & 1.0 & 0.0 \end{bmatrix} \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.3 & 0.0 & 0.7 \end{bmatrix}$$

Markov chain – Example 2 (much more interesting one)

$$\begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix}$$

$$\pi \begin{bmatrix} 0.0 & 1.0 & 0.0 \end{bmatrix}$$

Fish and chips days

$$\begin{bmatrix} 0.0 & 1.0 & 0.0 \end{bmatrix} \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.3 & 0.0 & 0.7 \end{bmatrix}$$

Markov chain – Example 2 (much more interesting one)

$$\begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix}$$

$$\pi \begin{bmatrix} 0.0 & 1.0 & 0.0 \end{bmatrix}$$

$$\begin{bmatrix} 0.0 & 1.0 & 0.0 \end{bmatrix} \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.3 & 0.0 & 0.7 \end{bmatrix}$$

Markov chain – Example 2 (much more interesting one)

$$\begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix}$$

$$\pi \begin{bmatrix} 0.0 & 1.0 & 0.0 \end{bmatrix}$$

$$\begin{bmatrix} 0.3 & 0.0 & 0.7 \end{bmatrix} \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.41 & 0.18 & 0.41 \end{bmatrix}$$

Markov chain – Example 2 (much more interesting one)

$$\begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix}$$

$$\pi \begin{bmatrix} 0.0 & 1.0 & 0.0 \end{bmatrix}$$

Keep doing it.....

$$\begin{bmatrix} 0.3 & 0.0 & 0.7 \end{bmatrix} \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.41 & 0.18 & 0.41 \end{bmatrix}$$

Markov chain – Example 2 (much more interesting one)

$$\begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix}$$

$$\pi \begin{bmatrix} 0.0 & 1.0 & 0.0 \end{bmatrix}$$

Keep doing it.....

$$\begin{bmatrix} 0.41 & 0.18 & 0.41 \end{bmatrix} \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.34 & 0.25 & 0.41 \end{bmatrix}$$

Markov chain – Example 2 (much more interesting one)

$$\begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix}$$

$$\pi \begin{bmatrix} 0.0 & 1.0 & 0.0 \end{bmatrix}$$

Until input = output

In that case, we have a stationary state and that is the "common" probability of each item

$$\begin{bmatrix} 0.41 & 0.18 & 0.41 \end{bmatrix} \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.34 & 0.25 & 0.41 \end{bmatrix}$$

UNIVERSITY OF GLOUCESTERSHIRE

Markov chain – Example 2 (much more interesting one)

$$\begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.3 & 0.0 & 0.7 \\ 0.5 & 0.0 & 0.5 \end{bmatrix}$$

$$\pi \begin{bmatrix} 0.0 & 1.0 & 0.0 \end{bmatrix}$$

For this example, it will be

$$\pi = \begin{bmatrix} 0.352 & 0.211 & 0.436 \end{bmatrix}$$

Markov chain – Example 2 (much more interesting one)

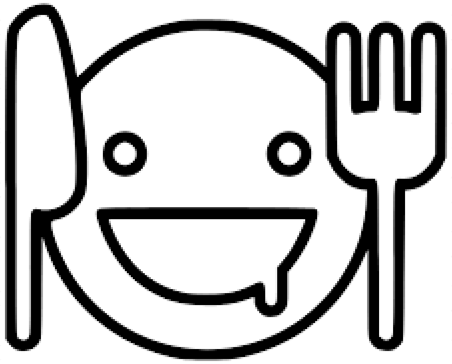
That is very close to the previous one

$$\pi = \begin{bmatrix} 0.352 & 0.211 & 0.436 \end{bmatrix}$$

$$\begin{bmatrix} 0.351 & 0.212 & 0.435 \end{bmatrix}$$

Markov chain – Example 2 (much more interesting one)

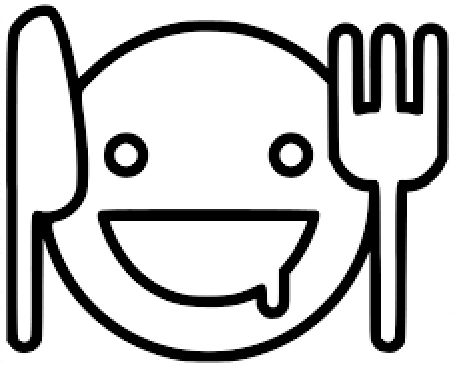
Now... I guess that you are hungry



UNIVERSITY OF GLOUCESTERSHIRE

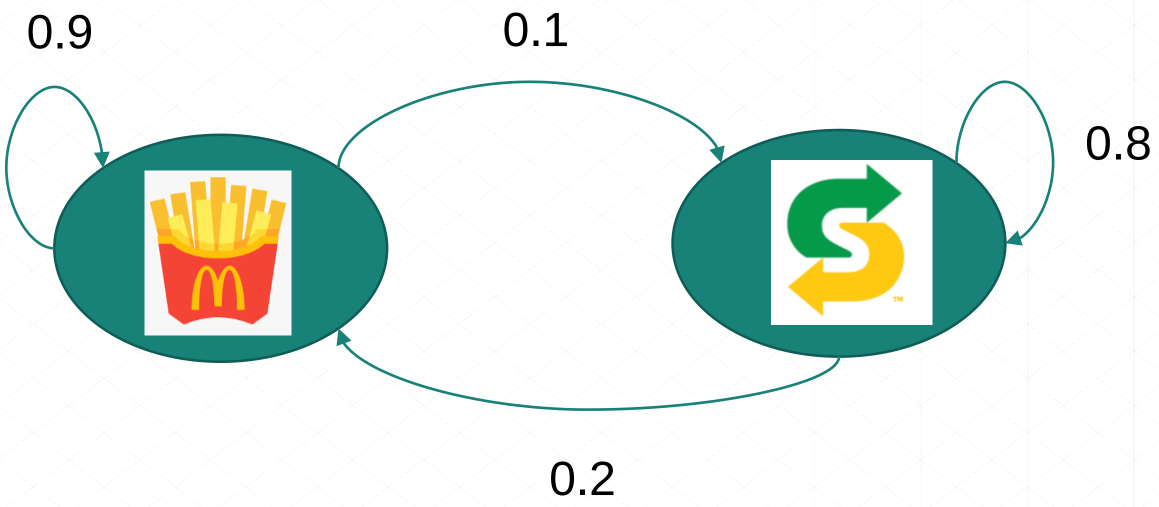
Markov chain – Example 2 (much more interesting one)

Let's use this in our favor and work on an exercise



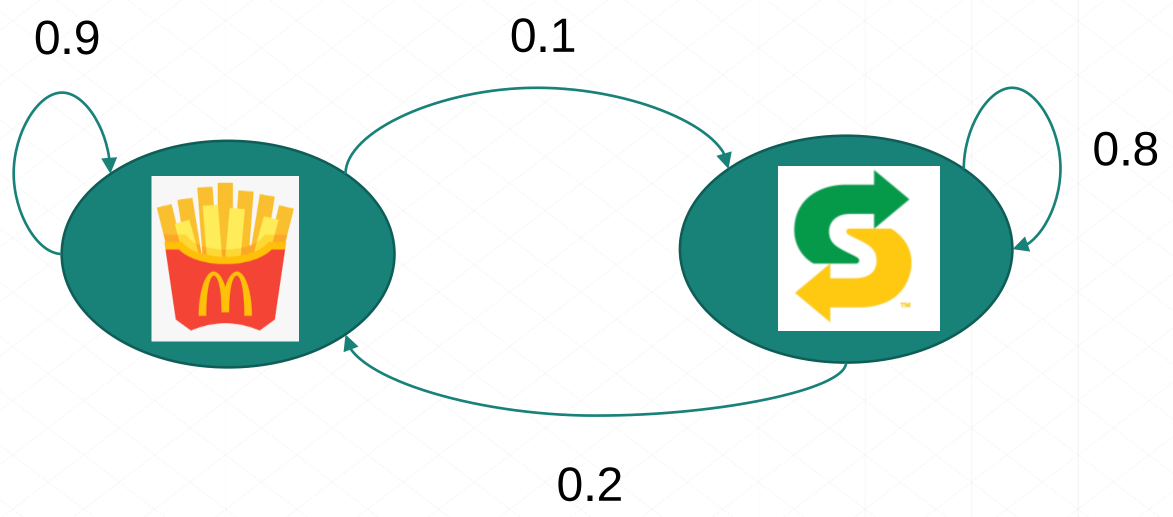
Markov chain – Exercise

- Given that a person buys McDonalds and Subway at a random choice.
- Given that the purchase can be represented by the Markov model below.



Markov chain – Exercise

Given that a person just bough a Subway, what is the probability that they will purchase McDonalds two purchases from now?



Markov chain – Exercise

P(Sub => ??? => McDon) =

P(Sub => McDon => McDon) + P(Sub => Sub => McDon) =

$$\begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$$

Markov chain – Exercise

P(Sub => ??? => McDon) =

P(Sub => McDon => McDon) + P(Sub => Sub => McDon) =

0.2 * 0.9 + 0.8 * 0.2 = 0.3
4

$$\begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$$

Markov chain – Exercise

P(Sub => ??? => McDon) =

P(Sub => McDon => McDon) + P(Sub => Sub => McDon) =

0.2 * 0.9 + 0.8 * 0.2 = 0.3
4

$$\pi^2 \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix}$$

Markov chain – Exercise

P(Sub => ??? => McDon) =

P(Sub => McDon => McDon) + P(Sub => Sub => McDon) =

0.2 * 0.9 + 0.8 * 0.2 = 0.3
4

$$\pi^2 \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix}$$

Markov chain – Exercise

P(Sub => ??? => McDon) =

P(Sub => McDon => McDon) + P(Sub => Sub => McDon) =

0.2 * 0.9 + 0.8 * 0.2 = 0.3
4

$$\pi^2 \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix}$$

Markov chain – Exercise

P(Sub => ??? => McDon) =

$$\pi^2 \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix}$$

Differences

Task

- What are the main differences between a Markov Chain and a Bayesian Network?

Differences

- Bayesian network is a directed graphical model, while a Markov network is an undirected graphical model.
- Cause/Effect (aka causality)
- Markov property
  - List and level of dependencies
- Acyclic vs cyclic
- Induced dependencies vs Cyclic dependencies

Post sessional work

- What is a hypothesis?
- Provide an example of two hypotheses in the area of Computer Science / Cybersecurity / Digital Forensics.

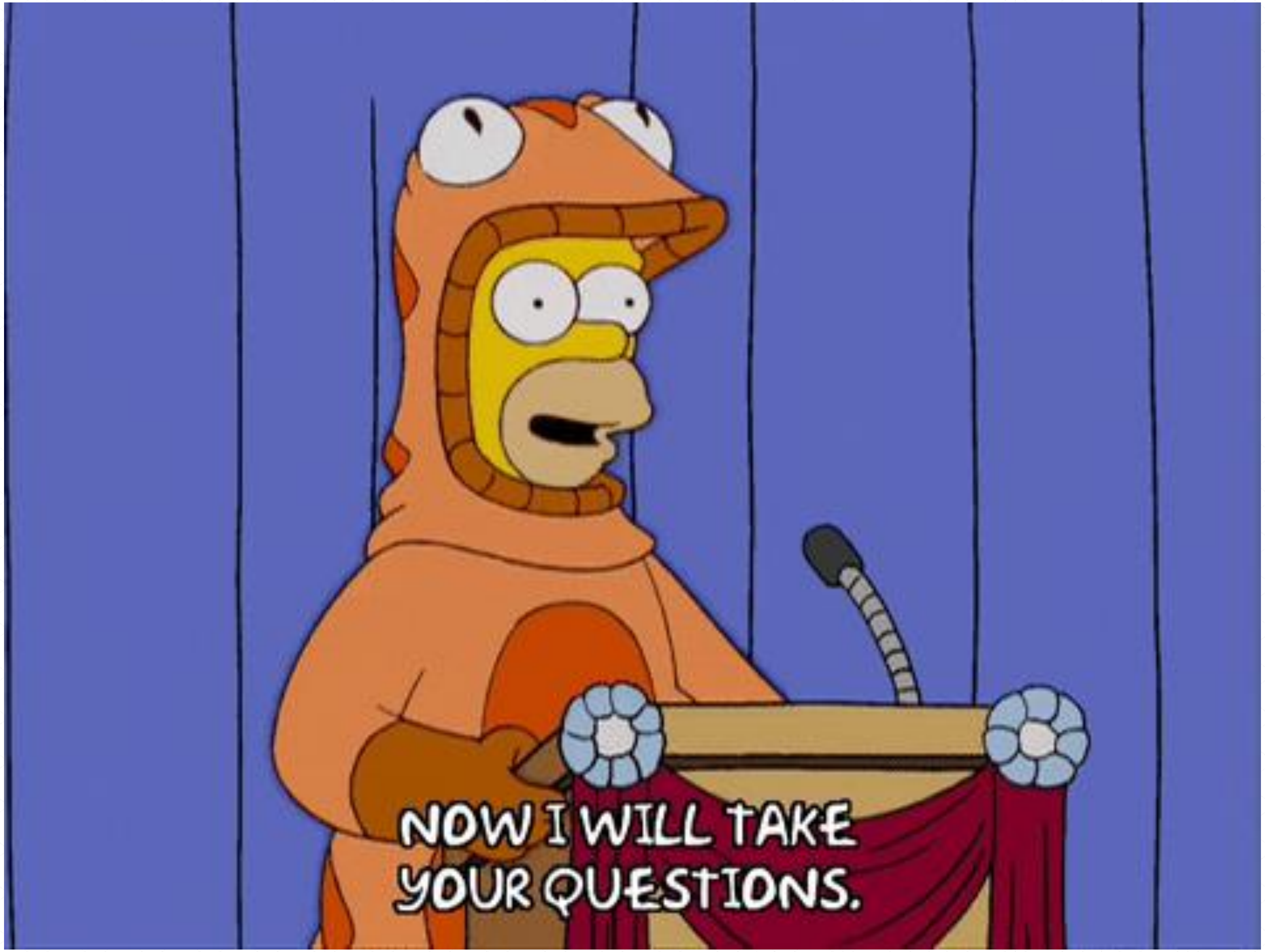
References

Yogesh Khandelwal, ‘Markov chain’, 13:03:05 UTC. Accessed: Jul. 15, 2021. [Online]. Available: https://www.slideshare.net/yogesh_khandelwal/markov-chain-46504704

Normalized Nerd, *Markov Chains Clearly Explained! Part - 1*. Accessed: Jul. 15, 2021. [Online Video]. Available: https://www.youtube.com/watch?v=i3AkTO9HLXo

Next Session !

- Hypothesis testing


NOW I WILL TAKE
YOUR QUESTIONS.