Markov Chains

Or how to create fake Wikipedia articles about Dinosaurs

By: Anthony Hevia

9/17/2020

Introduction 💬

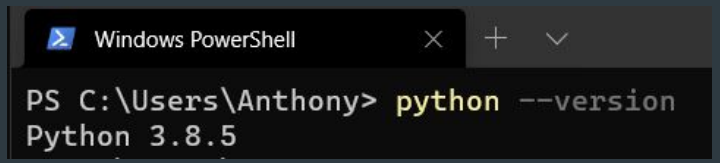
- Senior CS Major
- Officer with AI@UCF
- Undergrad researcher in evolutionary computation and bioinformatics
- Check out the AI Club discord: https://ucfai.org/discord
- Shameless Twitter plug: @starch_wars

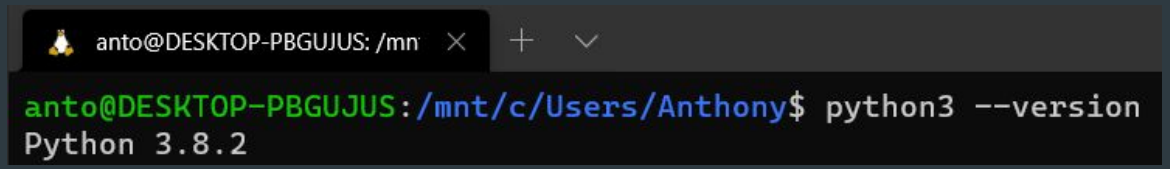
Agenda 

1. Prereqs, where to find the code.
2. What are sequences?
3. What problems do they solve and why do we care?
4. Markov Chains
    a. Examples
    b. The algorithm behind them
    c. Code walkthrough
    d. Where to go next?

Prereqs, Installing Python....

- Install latest version of Python 3 at: https://www.python.org/downloads/
- Any version of Python 3.6 or higher is good
- Code can be found at: https://github.com/Hevia/workshops
- You can check which version of Python you have by typing:
  - “python3 --version” or “python --version” in your command line





Prereqs, Downloading the repo.... 

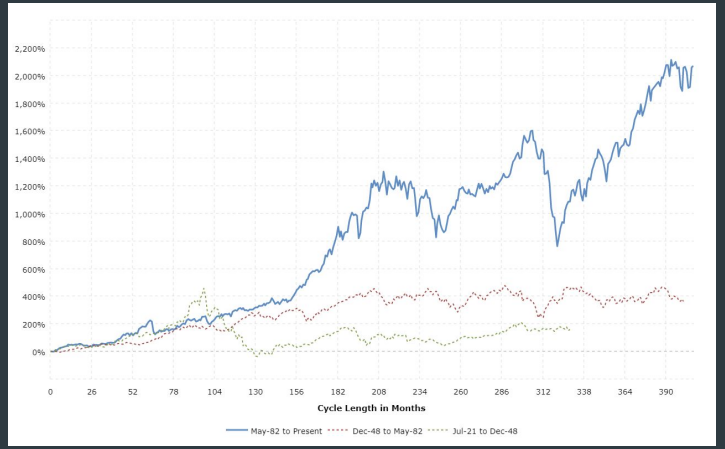
- Code can be found at: https://github.com/Hevia/workshops
- Install git here: https://git-scm.com/
- Getting the code is as simple as:

  “git clone https://github.com/Hevia/workshops”

```
PS C:\Users\Anthony> git clone https://github.com/Hevia/workshops
```

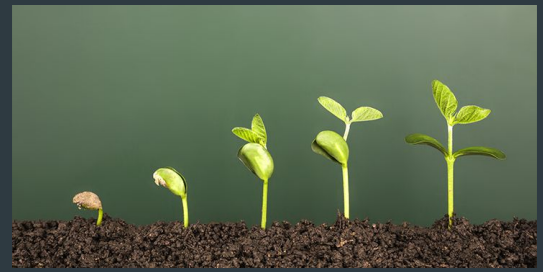
Sequences... What are they? Sequences... What are t....

- Sequences are a set of events that follow some sort of order
- Also called “Time-Series” data
- Examples: Weather, Stonks, Language, Cooking, Lifecycle of an organism



Sequences... the numbers Mason, what do they mean?

- Intuition: If we know the present, we can predict the future.
- Examples:
  - If you see dark clouds, you know it might rain soon
  - If a company announced a hit new product, stocks might go up
  - If an invasive species is introduced, we might see populations of native species decline.
  - Language itself is a sequence, the rules of grammar create restrictions on which words come after what.

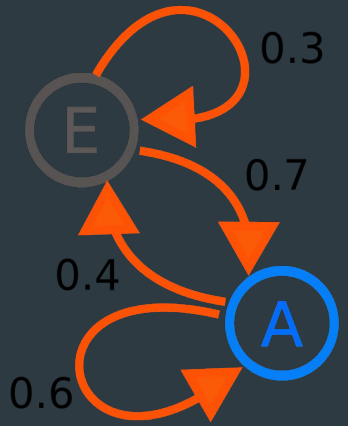


What problems do they solve? 

- Predicting the future or, the next event in a “sequence” is a core application of Computer Science.
- Examples:
  - Weather forecasting : The weather tomorrow?
  - Climate Modeling : The climate 10 years from now?
  - Advertising : Which Ad will get the most clicks?
  - Social Media : What happens after a user does action X?

Markov Chains: Introduction

- Markov Chains are one of the mains way of mathematically modeling a sequence of events
- Each event is a “state” and there is a “probability” of going from state A or state B



Markov Chains: Applications 💻

- Google’s original page ranking algorithm was a Markov Chain.
- The predictive keyboard on your phone is a Markov Chain (or at least used to be).
- Early versions of speech recognition used Hidden Markov Models, a version of a Markov Chain.
- Have been used to model population dynamics of animal communities.
- Generative Art

Markov Chains: How does it work?

- Consider the sentences: [“I love to eat pizza”, “I like to eat porridge”, and “I like to eat pancakes”].
- We need to figure out what is the probability of what word comes after the current word we’re on.
- We can use a for loop to loop over our words in the sentences.
- If we are at position “i”, we want to know the word at “i + 1”

Markov Chains: How does it work?

- Consider the sentences: [“I love to eat pizza”, “I like to eat porridge”, and “I like to eat pancakes”].
- i = 0, then i + 1 = 1
  - Sentence1[0] = ‘I’
  - Sentence1[0 + 1] = ‘love’
- In code, we will use a data structure, to keep track of what word comes after our current word
- We repeat this process over all our sentences and all the words in our sentences

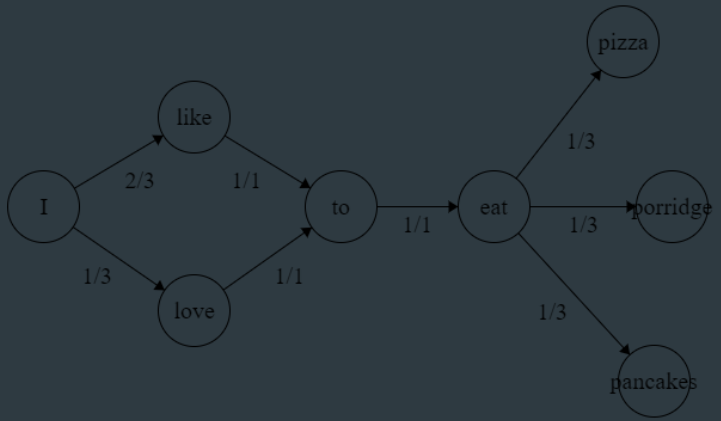
Markov Chains: How does it work?

- Python Pseudocode:
  - For every sentence in sentences:
    - For every word in the sentence
      - mysterious _data_structure[current word].append(next word in the sentence)

- Here are some questions (answers shown in the code):
  - What is this mysterious data structure?
  - What is the runtime?
  - Why do we have to go through each word in the sentence?

Markov Chains: How does it work?

- Markov Chain from the example sentences: [“I love to eat pizza”, “I like to eat porridge”, and “I like to eat pancakes”].



Let's look at the code!

Markov Chains: Where to go next? 🚀

- Check out the exercises included in the README.md
- Implement and deploy your own application that uses Markov Chains!
- Read up on Hidden Markov Models for a more advanced version of Markov Chains
- Learn about ngram Language Models which are Markov Chains with a slight increase in complexity
- Quantum Markov Chains are a thing as if the world wasn't crazy enough.
- You can even find Markov Chains in cutting edge fields like Reinforcement Learning! Check out the "Markov Decision Process"

Questions? 

Appendix

Python 101: Dictionaries

- Dictionaries are a data structure that stores data in (key, value) pairs
- Think of them as quite literally, real world dictionaries!