Random Story Generation - Worksheets 1-3

Random Story Generation 1

Key points:
- Punctuation is considered part of the word.
- Words are separated by whitespace.
- Capitalization matters: what is considered a different word than What.

(1) If the current 1-word context is what then the words that follow it are: I I
(2) If the current 2-word context is I gave, then the word(s) that follow it are: I

(a.1) play. do! go play Sit!
(a.2) to day. to
(b.1) shine. like
(b.2) If we pick the word shine, then the new current context is not shine.

Here are 2 examples of a story (11 words long, using one word of context):

We did not shine. It was too cold to go out
We did not shine. It was too wet to Sit! Sit!

Random Story Generation 2

How to represent:
(a) the new story: str
(b) the number of words in the new story: int
(c) the number of words of context: int
(d) the context and its corresponding next word(s): dict of {tuple of str : list of str}
The dictionary comes in handy here because it provides us a way to map any context to its possible next words.

Important design decision: When you filled in the tables in this 2nd worksheet you realized there’s no word following the word net, as we’ve reached the end of our training text. This is the point where you need to decide what to do. You could decide that you’ll wrap around and pick the first word of your text or you can say you’ll randomly pick one word from your entire text and start again or you may choose to do something else. All are valid options. We’ll follow the random word approach.
Random Story Generation 3

story = training_file.read().split()

{
('And', 'the') : ['fan', 'ship'],
('the', 'fan') : ['and'],
('fan', 'and') : ['the'],
('and', 'the') : ['cup', 'fish'],

# Complete the rest.
# Remember that a key should exist in the dictionary only once.
# And don't forget the quotes!!
}