Eilidh Dunsmore

Computer Science

Healey

February 10, 2019

Homework- Project Proposal

Goal 1: Create a program capable of producing randomly-generated poetry based on a system of text-predictability that can draw from a choice of texts.

Goal 2: Create a pretty user interface for the poem to be displayed in that has different font size for things like titles & lines

Day 1: Research & learn about the basics of text-predictability (i.e.: bigrams, conditional frequency distribution, & defaultdict) and then build a program that uses/creates bigrams & then mapping the conditional frequency distribution of those bigrams.

Day 2: Be able to input a text from corpus’ inside NLTK into the nascent poetry generator (i.e.: research different texts/corpus’ contained within NLTK & how to install & then install them) & tokenize, bigram, CFD, and then generate some words based on CFD of the text.

Day 3: Clean up nascent poetry generator by shake-filtering the text, potentially removing punctuation (Jane Austen’s dashes come to mind here…) & research phonetic alphabets and pronunciation dictionaries as well as syllable counting & NLTK syntax trees/parsing & part-of-speech tagging

Day 4: Research neural network based generation & using it to generate text (paper: <http://www.cs.utoronto.ca/~ilya/pubs/2011/LANG-RNN.pdf>).

Day 5: Poetry generator should be up & running, capable of generating funky poems by either RNN or just plain drawing on corpus material. (Draw from all of my own poems?? Self-indulgent, but ultimately revealing??) Begin work on researching pretty ways to display the poems. (Text Widgets using Tkinter? Ask for suggestions)

Day 6: Overflow day/finish any remaining work on creating a nicer & more readable presentation for newly-generated poems.