

PS6 Random Writer

For this assignment, we had to create a program to construct a Markov chain that can analyze k-grams (a fixed number of characters) and keep track of the probability of each character's occurrence. The program would then produce text using any input text and a Markov chain of given order k.

Key concepts

The RandWriter class utilizes a map with a kgram key and a map of characters and their frequency. The Mersenne Twister random number generator is also used by RandWriter in the kRand() function. RandWriter constructor creates a map that may be shown as a table-like output, containing each kgram and its frequency, as well as probability for each subsequent letter. When generating a new string from of previously produced characters, the function generate() employs the helper function kRand(), which picks a random next character from a kgram string.

What I Accomplished

Using probabilistic analysis on text to determine the next character/s in a sequence of length k words called kgrams is the Markov Model's name of the game. With the input of a string and order k the RandWriter class maps each of the kgrams in the string to it's following character and frequency. With the given kgram it can then generate a new string based on the probability of each of the following characters. Using this function, the TextWriter class is able to analyze words in text file and generate a pseudorandom string of L length.

What I Learned

Prior to beginning this project, I took it upon myself to absorb more knowledge about how Markov chains are commonly used in systems like online search engines, information retrieval, speech recognition, and gene prediction. I was able to observe firsthand how the Markov model can generate decent text using a trajectory through a table of probabilities of k-grams throughout the assignment.