**Problem 1:**

* Markov models used to generate artificial text with sufficient spatial dependence to capture the style of text on which it was trained
* GOAL: Design and implement a Markov model that is capable of generating stylized text, where the style is determined by the input text used to train the model
  + The result is a program that writes an English short story for you in the style of Joseph Conrad, or a poem in the style of Milton
* Characters in English text are not spatially independent
  + If the previous 5 characters in a string of English text are arith, then the next letter is almost surely an m
* A kth order Markov model predicts that each character in a text occurs with a fixed probability no matter where it occurs in the text, but that the probability can depend on the previous k characters.
  + The probability values are parameters of the model and we can perform parameter estimation to learn the probability values from observed data
* This is known as parameter estimation and uses approaches like maximum likelihood (empirically observed frequencies, which are based on occurrence counts) or maximum a posteriori (based on counts plus psuedocounts)
* In particular, you can estimate the parameters of a Markov model of order k by counting up how often each character occurs following each sequence of k characters.
  + Ex: if a text has 100 occurences of “th”, 50 of those are “the”, 25 are “thi”, 20 are “tha”, and 5 are “tho”, a second order Markov model would predict the next character following a th is:
    - ½ probability – e| ¼ probability – i| 1/5 probability – a|1/20 probability o