CS 1.2: Intro to Data Structures & Algorithms

Higher Order Markov Chains Worksheet

Name: _____

Text: *"I like dogs and you like dogs. I like cats but you hate cats."* (ignore all punctuation)

Q1: Outline a window of *three words* centered on each occurence of the word "like" in the text.

Q2: Draw a conceptual diagram of a *second-order* Markov chain generated from analyzing the text above. Each *state* should hold a *pair of words* and each *transition arc* leaving a *state* will represent the *next word* observed after the *pair of previous words* represented by the state.

Q3: Write the dictionary data structure you would build to store this *second-order Markov chain* (as it would look if you printed it out in Python). Put each *key-value entry* that represents a state on a separate line. A *key* is a *pair of previous words* and a *value* is a *histogram of next words*.

Q4: Write a *new sentence* that can be generated by doing a *random walk* on this Markov chain.