I knew immediately as I jogged into class, a minute late, that this period was not going to be fun. I hastily sat down, frantically got out a piece of paper and furiously started jotting down bullet points. I vaguely heard my teacher rambling on about last night’s homework and how he had made the instructions very clear on what to do, and that NO ONE should’ve forgotten. At this point I looked up from my pitiful notes only to see him glaring down at me. “Ryan,” he snarled. “Why don’t you present what you have.” I swallowed my nervousness, looked him in the eye and said, “I’m sorry, but I didn’t know how to do the homework.” His response was immediate. “I thought I made it CLEAR on how to do it!” I replied, “I’m sorry Mister-” “I don’t CARE what your excuse is, you are presenting what you have prepared.” he roared with a smirk on his smug face. Taken back by this sudden burst of outrage, I quickly stood up and launched into the meager presentation I had worked on for the best of two minutes. As I concluded my brief performance I hurriedly sat down and stared at my desk as other people commenced their presentations. Never had I felt so embarrassed in my life.

I knew immediately as I jogged into class, a minute late, that this period was not going to be fun. I hastily sat down, frantically got out a piece of paper and furiously started jotting down bullet points. If we’re going to detect plagiarism, we need some samples to work with. Generate many files, around 10 or so, that have varying degrees of content stolen from each other. Here are some ideas: Take paragraphs from Wikipedia, your old papers you’ve written, or from any other source. Make sure there are different amounts of overlap between files. One file should be an entire copy of some of the text of another file, two should share a large paragraph, etc. Make a pair of files where one is a direct copy of the other with a bunch of synonymous words substituted throughout. Plenty of pairs of files that are totally unrelated to each other.