Team9 HW4 – Maria Chrysanthou (ID: 1123114) & Maria Hadjichristoforou (ID: 1104894)

First of all, I’d like to clarify that we mostly used the extension codetogether on eclipse in order to do this project.

During this project when we first did each one function asked by the homework in one big class. However, we both agreed it would be a better idea if we would separate some of the functions in separate classes for a few reasons:

1. First reason, is we remembered Mr.Pyrros mentioned quite a few times about the Single Responsibility Principle (SRP). The SPR states that a class should have only one responsibility/job to do.
2. Second reason is that with that the code was more maintainable, and easier to spot any kind of errors/bugs.

Our final classes consist of:

* SentenceParser: responsible for parsing a text into lists of sentences and words
* FileHandler: responsible for handling files and extracting lists of sentences and words from them
* SemanticDescriptorBuilder: responsible for building semantic descriptors from a list of sentences.
* SimilarityTester: tests the semantic similarity between words using the cosine similarity of their respective semantic descriptor vectors.
* ResultsTOEFL: is used to run the program which uses the functions of the previous classes

To run the program, you go to the ResultsTOEFL class. (keep in mind that you will also need the brown-train-sentences , pg2600 and pg7178). We created a similarity-test.txt file (the one we used is attached) which included inside lines with words

The first word is the question word. The word right next to the first one is the correct answer.

Next are the choices the program has.