Google Search Engine Simulator Report

By: Saad Ahmed Siddiqui

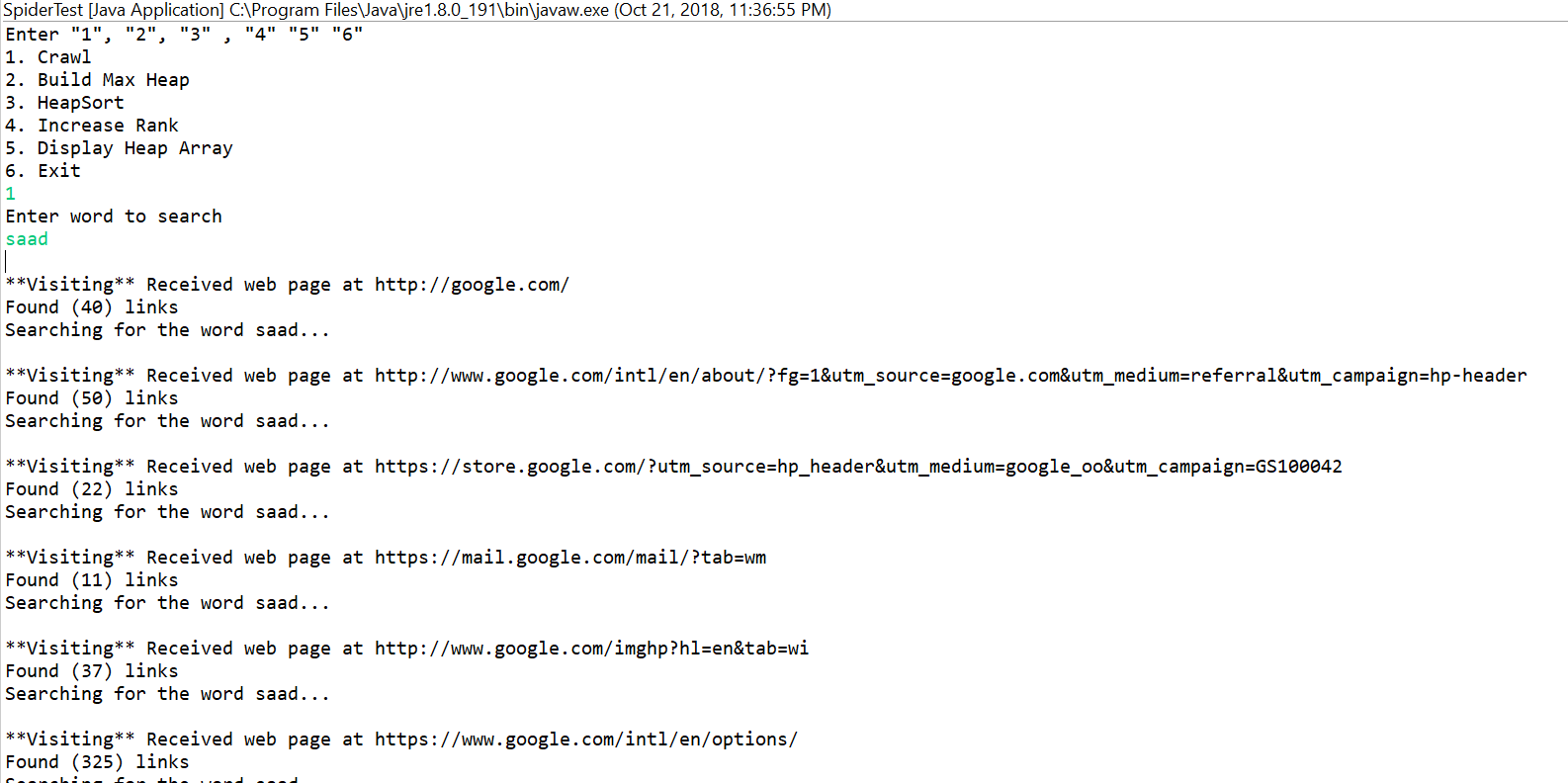
Section 7 1:30 pm – 3pm

**Explanation**

This program is a simulation of Googles Searching algorithm called PageRank. This program crawls the internet and extracts the links it visits and stores it in a heap. It is given a rank and using a priority queue it is displayed. If a person wants to increase the rank, simply select the option 4. And it would ask to enter the index and the rank you would like to give. 0 being the highest and 30 being the lowest in terms of ranking. It then sorts the links according to the ranks.

**Screenshot:**

1. Crawl



1. Build Max Heap And display



1. Heapsort and display



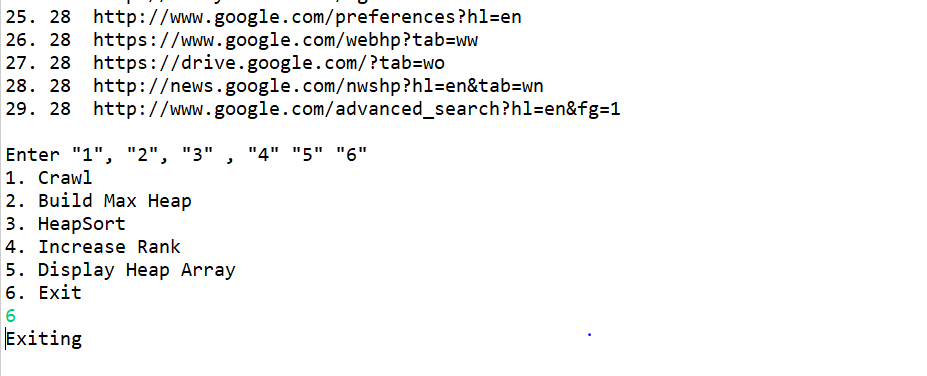
1. Increase Rank and Display







1. Exit



**Procedures to run:**

1. Unzip the PA1-Saad-Siddiquizip file
2. Direct to Java File Directory
3. Direct to src directory
4. Run:  java SpiderTest.java
5. Run: java SpidetTest

**Problems:**

My first [problem was storing the rank and links together I overcame that my creating a links Objects and storing the links and the rank together. Other problems were to implement the heapsort algorithm I implemented it using arrays.

**Lesson Learned**

I learnt how to implement the Heap sort [pseudocode using the algorithms books. Towards the end I learnt how to implement using array list to due to time limit wasn’t able to. I learnt how to crawls the web and how to manipulate the pages into ranking them in the order. I learned how to increase the rank of a page. The most important thing I learnt is now I have an idea of how google handles its data.