

课程名称：Java应用技术 指导老师：翁恺 学生姓名：尹嘉权

实验名称：Log Analyzer 实验类型：编程实现 学生学号：3120000419

1. 实验目的和要求

目的：

进一步了解 Java 基本使用和语法，熟悉Java容器和String类的使用。

要求：

Web servers typically maintain log files of client accesses to the web pages that they store. Given suitable tools, these logs enable web service managers to extract and analyze useful information such as:

* which are the most popular pages they provide;
* which ip took the most pages from the site;
* whether other sites appear to have broken links to this site’s pages;
* how much data is being delivered to clients;
* the busiest periods over the course of a day, or week, or month.

Such information might help managers to determine, for instance, whether they need to upgrade to more powerful server machines, or when the quietest periods are in order to schedule maintenance activities.

The log analyzer project contains an application that performs an analysis of data from such a web server. The server writes a log line to a file each time an access is made. A sample log file called access.log is provided in the Resource division.

1. 实验内容和原理

内容：

编写一个日志文件分析器，给出log文件导入，通过分析器解析之后，把结果输出到result.txt文件里面，主要分析结果应该要有：

* which are the most popular pages they provide;
* which ip took the most pages from the site;
* whether other sites appear to have broken links to this site’s pages;
* how much data is being delivered to clients;
* the busiest periods over the course of a day, or week, or month.

原理：

了解log文件的基本结构[[1]](#footnote-1)，对于log文件进行一行一行地解析，主要需要运用到的技术是对Java String类的字符串处理，截取相关内容；然后使用HashMap来统计log相关信息，最后遍历整个HashMap来找到最大的信息输出。

1. 主要仪器设备

MacBook Pro, MacOSX 10.10, 2.6GHz Intel Core i5, 8G RAM

JSE: JavaSE-1.8

JDK: oracle-java8-jdk

1. 操作方法和实验步骤

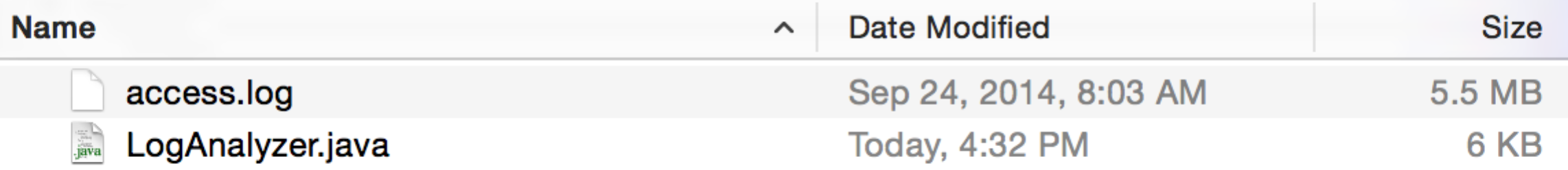
分析题目，然后打开Eclipse建立相对应的工程文件，输入代码，进行初步的调试运行之后把所有的.java文件整合到一个文件夹中，打开terminal，，找到相对应的程序目录，用javac通过JDK编译得到相对应的class文件，然后用java命令运行调试，注意log文件要和.java文件存放到一个目录下面。

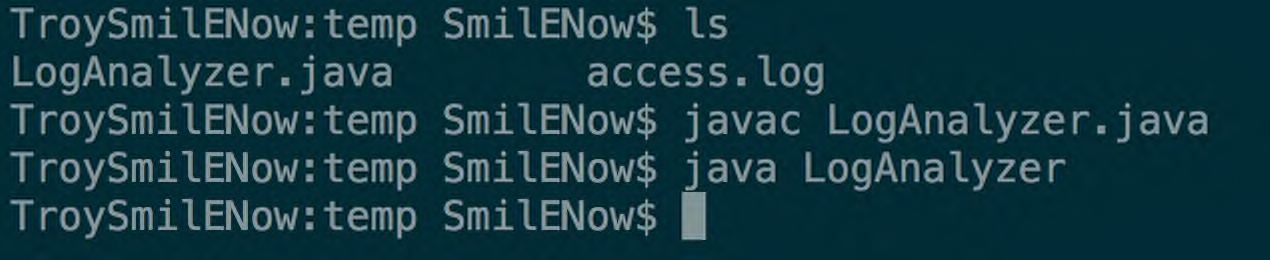
程序运行结束之后，在代码目录（或者说log文件目录）下找到相对应的result.txt，里面存放的就是分析结果信息。

具体的代码可以参见/src文件夹下的各个.java文件。

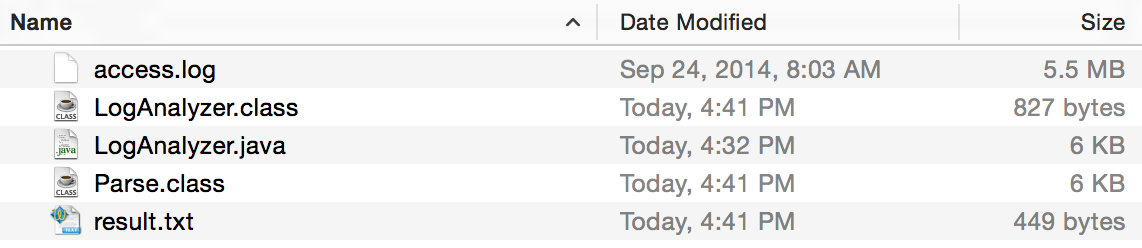
1. 实验数据记录和处理

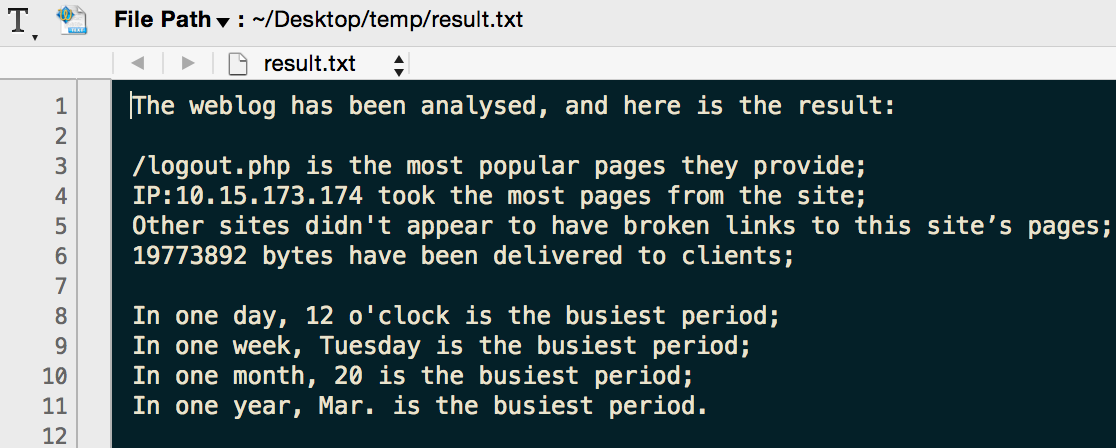
先把代码文件和log文件存放在一个目录下，然后打开terminal通过cd命令转移到对应的目录，然后编译LogAnalyzer.java，之后执行：





随后打开result.txt查看结果：





1. 实验结果与分析

实验输出正确。

1. 讨论心得

在这次实验中，让我进一步了解了Java中容器的使用方法，同时对于String类的substring有了初步的了解，实验并不难，主要是一开始要弄懂log文件内容的每一项具体指的是什么，当在做log的parse的时候，实质就是对字符串进行截取操作，把对应的内容存入/更新相对应的容器；另外，在计算星期几的时候，使用了基姆拉尔森计算公式来高效地计算，在统计总数据传输量的时候使用了大数类；也通过网络上学习到了对于Map的两种遍历方法[[2]](#footnote-2)（代码中有注释，被注释掉的就是相对低效的一种），对比它们的优劣，并采取了相对高效的一种来实现最后的输出。

1. <http://www.ittribalwo.com/article/1994.html> [↑](#footnote-ref-1)
2. <http://www.cnblogs.com/meieiem/archive/2011/11/02/2233041.html> [↑](#footnote-ref-2)