CSci2111: Assignment and Project Set # 3
Handed Out on: Tueday March 7, 2000
Due Date: Tuesday March 21, 2000

1 Written Exercises

Question 1: Problems 3, 7 and 8 pp. 284-285 (Chapter 7).

Question 2: Problem 5 and 8 pp. 364-365 (Chapter 8).

2 Programming Project, Part III (of IV)

Use Java or C++ to do the following (for grading purposes, this will be considered the programming project even though you will not have to implement the project from the textbook anymore):

Question 1:

1. Download the Kansas Court Case Name List from http://borg.cs.dal.ca/~nat/Courses/DFS-Course/court-cases
2. Each case is fomatted as follows: Name1 v. Name2 (d1d2d3d4d5) where Name1 and Name2 are strings of variable size and d1, d2, d3, d4, and d5 are digits (0, 1, .. 9).

Note: In certain cases, “v.” is actually replaced by “V.”

3. Think of this data as a series of variable length records of variable length fields with fields: “key”=d1d2d3d4d5, “party1”= Name1, “party2”= Name2, and “year”= d1d2.
4. Write a program that reads in this data file and creates a primary index file (where “key” corresponds to the primary key).

Hints: In this part of the assignment, you may want to use a string tokenizer. In C or C++, you can use the “strtok” function. In Java, you can use the StreamTokenizer class. Class StringBuffer will also be very useful.

5. Write a program that performs a Keysort (Section 6.4). Do not, however, sort the file directly nor re-write the file in sorted order once the keys have been sorted. Use insertion sort or any sorting method of your choice to sort the index.

6. Print all the records in sorted key order.

Question 2:

1. Write a program that creates two secondary index files. The first one should be indexed by Name2 while the second one should be indexed by year. Both secondary indexes should refer to addresses in the primary index.

2. Sort the two secondary indexes.

3. Perform a binary search of each index so as to output 1) all the records whose second party is "State" and 2) all the records of cases that took place in 76, 77, or 78 (Section 6.3).

Note 1: be careful to find *all* the records whose second party is "State". i.e., once you found one instance using a binary search, you did not necessarily find the first instance that occurs in the index.
Note 2: Do not output records named "State of Ohio" or "The Ohio State Fair". Only output those named "State".

Extra Credits:

1. Write a program that performs cosequential processing of the two secondary indexes. (Section 8.1).

2. Use your program to output all the records involving "State" as a second party and that took place in 76 and 78 (but not in 77).