University of Toronto
Department of Computer Science
CSC 228 - Summer 1995

Assignment 1: Miami Vice meets 228

Due: in tutorial, June 15

This assignment will give you some practice with separate compilation, file I/O, master file update, and file-based linked lists.

You are to write two programs, as described below, make up test data for all the necessary files, and write a report. The report is intended to convince your tutor that your programs are correct and that you have tested their performance. Read pages 1-2 of the course guide for more information.

The problem

Crocket and Tubbs are police officers working on a major drug bust. Their operation requires them to maintain a log book of suspects, with observations of each suspect's activities. To assist the police in this effort, you have been hired to write a computer program that will store the log book information in a database for easy access by other members of the force.

Your database will consist of two files, organized as follows:

A master suspects file, sorted by suspect name:

name	alias	whereabouts	pointer to first record in activity file

An activity file, arranged as a series of chronologically ordered linked lists, one per suspect:

time	activity description	pointer to next activity record for this suspect

Part I

The first program you write will update the database from a daily transactions file, using a variant of the master file update algorithm. Each day, Crocket and Tubbs will create a transactions file, sorted by suspect name, and apply it to the suspects database. Transactions file records will be of three types: new, move, and observation. The formats for each are described below**:**

transaction	fields		
new	name	alias	whereabouts
move	name	whereabouts	
observation	name	time	activity description

Crocket, unfortunately, is a chain smoker, and due to years of brain damage from nicotine, he tends to keep highly disorganized log books. As a result, the observations of each suspect will not necessarily be entered in chronological order.

Part II

In order to access the database, other members of the police force will submit a requests file with the names of various suspects. Your second program must generate a report containing the name, alias, address, and a list of observations for each of these suspects. The requests file will not be sorted, so you have two choices:

a) first sort the requests file, then search the master file sequentially for the desired suspects

b) repeatedly use binary search on the master file to locate the appropriate record for each suspect

Again, due to his smoking habit, Crocket has a very short temper. Keeping in mind the fact that he also carries a big gun, you want to make sure that his requests for information on suspects are processed as quickly as possible.

In order to decide between the two processing options, you should implement both and time their performance on a number of sample requests files. The UNIX `time()` command can be used for this purpose. Your program may also make use of the `qsort()` command to sort the requests file. Finally, your report should discuss the measurements you made, and try to explain why one method is superior to the other.

What to submit

You must submit the report, program code, and some relevant test files, demonstrating the input and output of your program. Read the course guide for more details. The actual number and content of your test files is up to you, but some guidelines will be discussed in your tutorials. Don’t forget to supply timing measurements for your performance analysis in part II.

Things to think about

- What to do with duplicate names?
- How many runs of the program are required to get reasonable performance measurements?
- Do your measurements agree with what you expected in part II of the assignment?