>>> Assignment #1 for Simulation (CIS 4930) <<<

>>> SOLUTIONS <<<

This assignment covers material from the first week of class lecture and reading.

Problem #1 (5 points)

If is very important that you know what the class website contains. So, for this first fun problem you are to go on a frog hunt. There are images of frogs hidden on one, or more, page(s) that I have created found somewhere on, or linked to, the class website. So, go find the frogs! Give the URL of each page that contains a frog. It is a good idea to be aware of the content of the pages that contain the frogs.

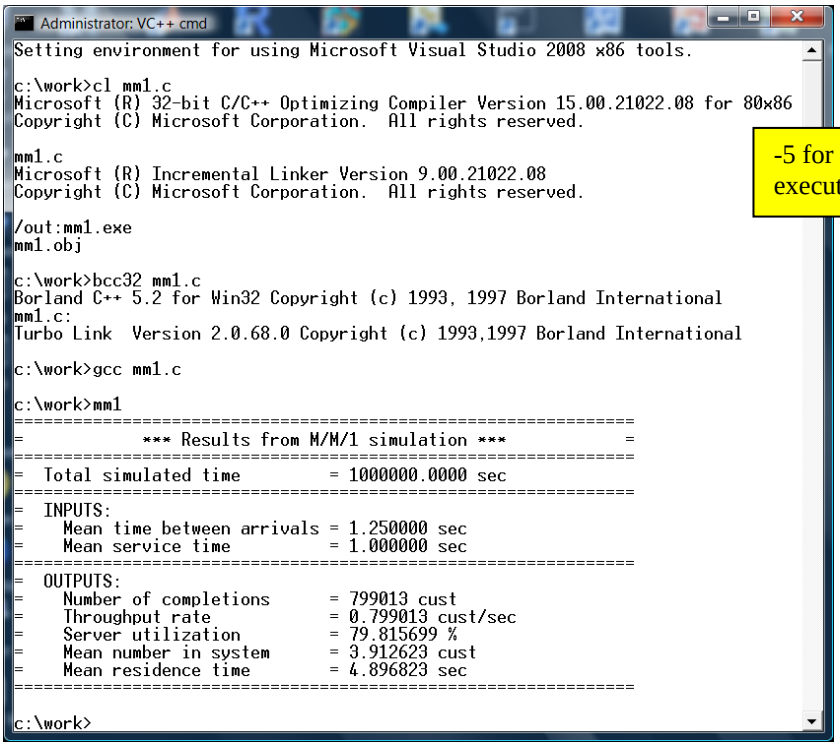
-2 for a missed frog

One frog can be found at the bottom of the page http://www.csee.usf.edu/~christen/class3/cstyle.html. This page is found on the (top level) page http://www.csee.usf.edu/~christen/class3/misc3.html. A second frog can be found on the bottom of the page http://www.csee.usf.edu/~christen/class3/hidden/reading3.html. This page is linked-to from the course outline on http://www.csee.usf.edu/~christen/class3/outline3.html. The third frog is on the miscellaneous page (http://www.csee.usf.edu/~christen/class3/misc3.html).

Problem #2 (5 points)

It is *absolutely critical* that you have a C programming environment that you are comfortable with. Download the program mm1.c from the Christensen tools page (http://www.csee.usf.edu/~christen/tools/toolpage.html). Compile it and run it. Take a screenshot of your compilation and execution. Submit the screenshot. Later in the semester you will need to use Visual C++ Express Edition 2008 as your development environment. So, if you do not yet have a C development environment on your PC I suggest this may be a good time to download and install VC++ Express 2008. You can find VC++ Express 2008 here: https://www.dreamspark.com/Products/product.aspx?productid=9. Why not VC++ Express 2010? It is not clear that the CSIM software we will use later in the semester will work with 2010. Personally, I use command line for all of my C compilation – I do not use the IDE. You should use whatever you are comfortable with (that is, command line versus IDE).

Here is a screenshot showing builds with cl, bcc32, and gcc, and then execution.



-5 for bad compile or failed execution

Problem #3 (50 points)

You are to find a paper in an ACM or IEEE conference that uses simulation methods to evaluate an ICT component or system. The paper must have been published in the last 10 years. For the paper, answer the following questions:

a) What is the problem or question being addressed?
b) Very briefly, what is the solution to the problem?
c) Describe the simulation model developed and used
d) Identify and describe the response variables
e) Identify and describe the factors and factor levels
f) Describe how the simulation results were presented

-8 points for each of (a) thru (f). Can get -8, -4, or 0 on each. Note that selected paper must be valid, else –25 just to start.

Include the paper in your submission package. Give the full and correct citation (in IEEE-CS style) for the paper. As you probably already know, the USF library gives you access to IEEE Xplore and the ACM Digital Library. I suggest using scholar.google.com (and not plain google.com) for searching for academic/research papers. Note that the paper must not be about simulation, but rather must be about something else (say, a new design or method) that uses simulation as the tool to evaluate the design or method being proposed and developed in the paper.

The paper I selected to review was: T. Xie and Y. Sun, "PEARL: Performance, Energy, and Reliability Balanced Dynamic Data Redistribution for Next Generation Disk Arrays," *Proceedings of the 16th Annual Meeting of the IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS)*, Baltimore, Maryland, USA, September 8-10, 2008.

a) The paper explores a new hybrid storage architecture that combines hard disks and flash disks. Hard disks and flash disks have different performance, energy, and reliability (and cost) properties. For example, flash disks are well suited to read data (no seek time), but less so to write data (due to erasure time). The problem addressed is data placement (on hard disk or flash disk) to trade-off performance, energy use, and reliability.
b) The solution is a new algorithm with data structures to determine data placement.
c) The simulation model itself is not described in the paper beyond saying that the authors develop an "execution-driven simulator that models a hybrid disk array which has one hard disk array an done flash disk array." The simulator is parameterized from real (commercial) equipment and is trace driven.
d) The key response variables are mean response time and energy consumption.
e) Control variables for the system include storage capacity, disk access time, disk transfer rate, disk idle power, disk active power, flash read speed, flash write speed, flash idle power, flash active power. Control variables for the workload (which are traces) include number of reads, number of writes, trace length, average size, and average interarrival time.
f) The simulation results are presented in bar graphs showing mean response time and energy consumed as a function of storage size and number of disks.

The paper is attached to these solutions.

Problem #4 (40 points)

Consider a generic PC. What are the response variables of interest to performance? What are the factors (and factor levels) of interest to performance? What are possible workloads of interest? Identify factors (and factor levels) related to both the system itself and the workload.

20 points for control variables and 20 points for response variables. -5 for confusing variable types. Need a "reasonable" list, else -10.

The two key response variables for a PC are:

1) Job throughput – how many jobs per second can the PC complete
2) Response time – how long it takes to complete a job (this includes how long it takes for the PC to start a program and how long it takes to finish executing a program)

Factors of interest for a PC are many and they include for the system (listed here with level measures):

1) Processing (including width of data path, MIPS rating, and number of cores)
2) Memory (including amount and access speed)
3) Storage (including type, size, seek time, and access speed)
4) I/O (including data rate – wired and wireless)
5) Energy use (including battery lifetime if a laptop)

For the workload the following are key:

1) Type of jobs (e.g., spreadsheet, work processing, numerical intensive, I/O intensive, GUI or batch)
2) Number of simultaneous jobs

Note:

The TA and I are here to help you! Make use of help if you need it.