Institut for Matematik og Datalogi
Syddansk Universitet

Assignment 3 — Introduction to Computer Science 2013

This is your third assignment in DM534. The assignment is due at 8:15 on Monday, December 9. You may write this either in Danish or English. It must be made in LaTeX. (though you do not need to include your LaTeX code). Write your full name, your section number, and your "instruktor"s name (Uffe Thorsen or Magnus Gausdal Find) clearly on the first page of your assignment (on the top, if it's not a cover page). Turn in both a paper copy of this assignment (to Uffe Thorsen's mailbox, which is among all the "instruktor"s mailboxes, if you are in S17; or to Magnus Gausdal Find's mailbox, which is in IMADA's "sekretariat", if you are in S7) and an electronic version as a PDF file via Blackboard through your DM534 course (choose the correct one, S7 or S17). The assignment hand-in is in the menu for the course and is called "SDU Assignment". Keep the receipt it gives you proving that you turned your assignment in on time. Blackboard will not allow you to turn in an assignment late.

Cheating on this assignment is viewed as cheating on an exam. You are allowed to talk about course material with your fellow students, but working together on this assignment is cheating. If you have questions about the assignment, come to Joan Boyar or your "instruktor" for DM534.

Please note that you must have this assignment approved in order to pass DM534. If it is not turned in on time, or if you do not get it approved, it will count as one of your two retries in the course, and you must have it approved on your only allowed retry for this assignment.

Assignment 3

Do the following problems and write your solutions in LaTeX. Write clear, complete answers, but not longer than necessary. Do not include the statements of the problems or other information not asked for in the problems.

1. Do one of the following two problems. The second is the more challenging.

   (a) Assume sets of numbers are represented by lists sorted on element value. For example, the set $\{14, 27, 13, 9, 32\}$ is represented by a list of length 5 containing $[9, 13, 14, 27, 32]$. Write a procedure in pseudocode for constructing $A \cup (B \cap C)$. Use an algorithm similar to that in Figure 9.15 (which goes through each list only once, never reading any element more than once). Assume that you can check for the end of a list, similarly to how you can check for the end of a file, say with "EOL". As in problem 6 (from the discussion section described on the note for Lecture 11), process the three lists simultaneously (do not first calculate $B \cap C$ and then union with A). Make sure your algorithm works correctly in all cases, and explain why it does.

   (b) Consider the problem of merging four files, each of length n. Call the files A, B, C, and D. There are two obvious ways to do this. One is to first merge the files A and B, then merge the files C and D, and then merge the two resulting files. The second processes the four files simultaneously (as in the above problem, never reading any element more than once). Analyze both of these algorithms, calculating the number of reads, the number of writes, and the number of comparisons, separately. Express your results in the form $i \cdot n + j$ in all cases. The value i should be as low as possible, and both i and j should always be constants, not functions of n. In order to show the the i value for the number of comparisons is the same for both algorithms, you need to specify exactly which comparisons are done by the second algorithm. Explain your answers.

2. If a hash file is partitioned into 9 buckets, what is the probability of at least two of three arbitrary records hashing to the same bucket? Assume that the hash function is such that a randomly chosen record is equally likely to hash to any of the buckets. How many records must be stored in the file until it is more likely for collisions to occur than not? Assume again that there are 9 buckets. (You may use a program to check your answer, but show intermediate steps of your calculations here.)