Chair for Algorithms
and Data Structures
Prof. Dr. Hannah Bast
Marjan Celikik

Search Engines WS 09/10

http://ad.informatik.uni-freiburg.de/teaching

UNI FREIBURG

Exercise Sheet 14

complete until Thursday, February 18th

Exercise 1

Write a program that takes two *sorted* sequences of integers, and merges them into one sorted sequence. Use the obvious linear-time algorithm.

Exercise 2

Write a program that takes the same input as in Exercise 1, and produces the same output. But this time concatenate the two input sequences, and then do a standard sort of the resulting sequence (thus ignoring the fact that the two parts were already sorted each).

Exercise 3

Run the two programs on random input sequences, both of the same length. As lengths, take $n = 10^3, 10^4, 10^5, \ldots$ (as far as you can go on your machine). For each n, run your programs 10 times. For each run, the random sequence should be the same for both programs. Display the average running times for the two programs for the various values of n in a table. Which program seems to be more efficient? Are you surprised?

Exercise 4

Use statistical hypothesis testing, as discussed in the lecture, to determine (separately, for each value of n) whether the difference in the average running times is statistically significant or not (with the standard significance level of 5%).

Assume your averages were obtained from 100 runs instead of from 10. Does that affect the answer to the question whether the difference is statistically significant or not?