Datenbanken II (SS 2017)
Exercise 6

Task 1 Statical Hashing

(a) Which characteristics should a good hash function exhibit? How can collisions be handled?

(b) The following students (MatrNr, Name) have to be inserted into a hash table of size 4, bucket size 2 with overflow buckets: 24002 Xenokrates, 25403 Jonas, 26120 Fichte, 26830 Aristoxenos, 27550 Schopenhauer, 28106 Car- nap, 29120 Theophrastos, 29555 Feuerbach

i. Make use of a simple modulo function for this task.

ii. Add a digit sum function in advance of the computation of the modulo function. How does the distribution change?

Task 2 Dynamical Hashing

Data series 1: 8, 10, 12, 14, 24, 33, 37, 41
Data series 2: 6, 59, 76, 11, 95, 66, 4, 26, 92, 60, 38, 57, 29, 99, 53

The initial size of the hash table is 2; the size of the buckets is 2; overflow pages are possible; splitting takes place after every collision.

(a) Linear hashing
Explain the approach of linear hashing. Insert the values of data series 1 in the given order into the hash table.

(b) Extendable Hashing
Explain the approach of extendable hashing. Insert the values of data series 1 in the given order into the hash table.

Task 3 Sorting

(a) Why are external sorting algorithms needed? Sketch their functionality.

(b) What is the complexity of external sort algorithms? Is a logarithmic external sort possible?

(c) How much recursion steps does the sort-merge technique need for sorting 100.000.000 tuples (10 tuples per block) with a buffer size of mem = 3, 11, 101?
Task 4 Scans

(a) What are the different kinds of a scan?

(b) Sketch the principles of the scan strategies?

(c) When do you use which scan strategy?

(d) What is a scan semantic and what is it used for?

Good Luck!