Implementation of DBMS
Exercise Sheet 2
Klingemann, WS 2024 / 2025

1) Suppose you use Two-Phase Multiway Merge Sort in the scenario described in the lecture slides. Tell how many disk I/O's are needed for the sort if the following changes are applied:
a) The number of records of the file is doubled.
b) The size of blocks is doubled, to 8192 bytes.

•10^7records of 100 bytes = 10^9bytesfile.
•Stored on a diskwith 4KByteblocks, each holding 40 records + header ..... information.
•Entire file takes 250,000 blocks.
•50MByte available main memory = 12,800 blocks 1/20th of file.
•Task: Sort records of file by primary key field. We have calculated in the lecture that 6.71 * 10^9 record

2) Suppose we have a relation whose n tuples each require R bytes, and we have a machine whose main memory M and disk-block-size are just sufficient to sort the n tuples using Two-Phase Multiway Merge Sort. How would the maximum n change if we made one of the following modifications of parameters?
a) Double B
b) Double R
c) Double M

3) You want to use the Two-Phase Multiway Merge Sort (or its extension to a different number of phases as appropriate) to sort a file. The file consists of 119,985 records. Each block can contain 20 records. We have 10 main memory blocks available.
a) How many blocks do we need to store the file if each block is as full as possible?
b) How many phases do we need?
c) What is the required number of I/O's?
d) How many sorted sublists do we have after each phase?

4) Consider the following relations:

R:

A	**B**
a	b
c	b
d	e

S:

B	**C**
b	c
f	a
b	d

Calculate the following relations:
a) $\pi_A(R)$
b) $\sigma_{A=d}(R)$
c) $\pi_A(\sigma_{B=b}(R))$
d) $R \times S$
e) $R \bowtie S$
f) $R \bowtie_{A=C} S$
g) $\pi_A(R \bowtie S)$
h) $\pi_A(R) \bowtie S$