Implementation of DBMS
Exercise Sheet 5
Klingemann, WS 2024 / 2025

1) Suppose that if we swizzle all pointers automatically, we can perform the swizzling in half the time it would take to swizzle each one separately. If the probability that a pointer in main memory will be followed at least once is `p`, for what values of `p` is it more efficient to swizzle automatically than on demand?

2) Suppose blocks hold either ten records, or 50 key-pointer pairs. We have a data file with 10^6 records. In this task we assume that each block will be as full as possible.
a) How many blocks do we need for the data file?
b) How many blocks do we need for a dense index?
c) How many blocks do we need for a sparse index?
d) We want to add higher level index structures to the index considered in b). Describe how many blocks we have in these higher levels until we end up in a level with just one block.
e) Repeat d) for the sparse index in c).

3) Use the same scenario as in task 2. We assume that nothing is in memory initially, and that the search key is the primary key for the records. What is the average number of disk I/O's needed to retrieve a particular record that is stored in our data file given its search key using the following approaches?
a) We do not use any index but sequentially inspect the data file from the beginning until we find the record.
b) We just use the dense index described in task 2b to retrieve the record. We sequentially scan the index from the beginning until we find the search key value we are looking for.
c) We use the multi-level index described in task 2d to retrieve the record.
d) We use the multi-level index described in task 2e to retrieve the record.