Introduction to Database Systems

2020/21

2nd exam, FAMNIT

The time available for the exam is 120 minutes. Good luck!

NAME AND SURNAME:	
STUDENT NUMBER:	
PROGRAM:	
SIGNATURE:	

The most important tables in the information system of a Big Library are as follows.

```
Books(bid,author,title,publisher,year);
Members(mid,name,surname,street,city,phone);
Borrow(rid,bid,mid,sid,date);
Staff(sid,name,surname,address,phone)
```

In addition, we have the following data.

1 disk page = 8KB

|Books| = 3.000.000 records, 320 bytes/record, 25 records/page, 120.000 pages
|Members| = 100.000 records, 200 bytes/record, 40 records/page, 2.500 pages
|Borrow| = 30.000.000 records, 40 bytes/record, 200 records/page, 150.000 pages
|Staff| = 400 records, 200 bytes/record, 40 records/page, 10 pages

A database management system has a buffer pool with 1000 pages. Write all the additional assumptions!

Exercise 1. (25%)

Print the data about the members that borrowed in this year (>21/1/1) more than 5 books of the author with the surname Knuth. Print the names, surnames, telephone numbers, and the number of borrowed books (of the author Knuth).

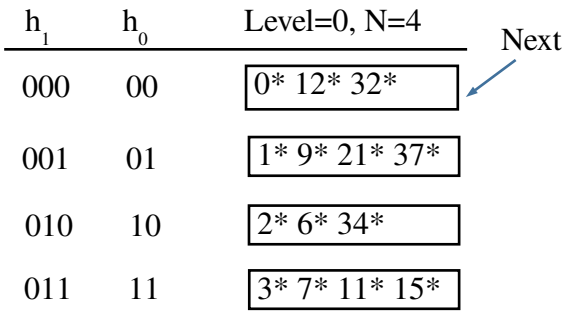
a) (10%) Use SQL.

a) (10%) Use QBE.

a) (5%) Describe briefly the data structures used for the representation of a relational table on a disk.

Exercise 2.

a) (20%) An example of the linear hash index is given below. Every time a new overflow page is added, the bucket pointed by Next is split, and Next is incremented by one. Show the state of the index after adding the keys 5, 22, 13 in 23.



b) (5%) Describe the algorithm for inserting a new key into a B+ tree in few sentences. Present the alternatives.

Exercise 3.

The following query expressed in SQL is given.

```
SELECT Borrow.bid, Borrow.date
FROM Members, Borrow
WHERE Members.mid=Borrow.mid AND Members.city='Koper'
  AND Borrow.date>2020/1/1
```

a) (6%) Translate SQL statement into relational algebra.

b) (14%) Each table has a hash index defined on its key. A B+ tree is created for the attribute Borrow`.date`. How many blocks reads the query if our system uses the index nested-loops join solely.

c) (5%) Briefly describe the alternative implementations of the GROUP BY statement.

Exercise 4.

The following schedule of transactions T1, T2 and T3 is given.

T1: R(X) W(X)
T2: W(X) W(Y) R(Z)
T3: R(Z) R(X) R(Y)

a) (6%) Identify the examples of RW, WR, and WW conflicts.

b) (14%) Draw a dependency graph for the given transactions. Is the schedule conflict serializable? To which serial schedule is the above schedule equivalent?

c) (5%) Present the main ideas of the optimistic concurrency control.