Assignment 1

Databases: Fundamentals

Recap on databases

Exercise 1 – SQL

The following relational model describes aircrafts, airlines, the airports they fly in (may be more than one), and the flights they operate. Primary keys are underlined and foreign keys are marked with [FK].

airline(cnr, cname, address)
aircraft(pnr, pname, capacity, owner_cnr[FK])
airport(anr, aname, city, nation)
flight(fnr, day, time, using_pnr[FK], from_anr[FK], to_anr[FK])

1. Write SQL DDL-statements to create all necessary tables.
2. Write sample INSERT-statements to fill records into each table.
3. Write SQL queries to determine …

    a. … all aircrafts with a capacity higher than 100 and lower than 300.

    b. … all flights of the airline *Austrian* departing from Vienna.

    c. … all airlines operating aircrafts with name *A380.*

    d. … all aircrafts going to Venice or Trieste.

    e. … all airports with incoming flights from Austria and not from Germany.

    f. … all airlines with flights to Klagenfurt or Dubai (but not both).

    g. … the average aircraft capacity for each airline.

    h. … the name of the airline(s) with the highest number of aircrafts.
        Note: The query must return multiple flights when required!

Exercise 2 – Indexes

1. Explain briefly what a primary index is.
2. Explain briefly what a secondary index is.
3. Propose indexes which (may) speed up the queries of Exercise 1.

Exercise 3 – Physical Data Organization

Table *customer* has following properties:

Relation: customer (cid, name, city, sum_sales)

Attributes: cid (4 bytes), name (20 bytes), city (20 byte), sum_sales (10 bytes)

Data-volume: 70.000 customer records

Cardinalities: about 15.000 different values for sum_sales (equally distributed)

Indexes: primary index on cid

Basic system properties:

block size is 1 KB (1024 byte);

a block pointer requires 10 bytes;

a record pointer 16 bytes.

The following two queries are frequently and equally often issued:

Q1: SELECT * FROM customer WHERE sum_sales = ?;

Q2: SELECT * FROM customer WHERE cid = ?;

1. How many block accesses are required for both queries, if they are issued once?

2. A programmer wants to create a dense secondary index on sum_sales to increase the performance of query Q1. Decide for or against this proposal, based on block accesses. Or do you find an even better solution?

Exercise 4 – Transactions

1. What is ACID?
2. What is a transaction?
3. Explain the difference between commit and rollback.
4. Non-serialized multi-user environment:

   a. Identify all problems that might arise when you execute schedules T1 and T2 concurrently in a non-serialized multi-user environment.

   b. Which statements or statement-combinations are responsibly for these problems?

   c. How can concurrency-problems be avoided?

T1		T2
BOT	1	
read(x)	2	BOT
x:=x-100	3	read(z)
write(x)	4	
read(y)	5	read(x)
y:=y+100	6	z:=z+x
write(y)	7	read(y)
	8	z=z+y
	9	write(z)
	10	read(y)
	11	y=y*2
	12	write(y)
	13	COMMIT
ABORT	14	

Exercise 5 – Serialization   - 3 -

1. Is this schedule serializable?
    Justify your answer.

	T1	T2	T3
1	update(a)		
2		read(b)	
3			read(a)
4		update(b)	
5			read(b)
6			read(c)
7			update(c)
8	read(c)		
9	read(b)		
10		read(a)	