

Database Systems, Fall 2006

Rasmus Pagh

November 13, 2006

Exercises for the lecture on November 14

1. **Machine exercise.** The standard behavior of Oracle is to represent relations in an unclustered (i.e., unordered) way. If a primary key is declared, there will be an (unclustered) index on it. If we desire the relation to be clustered according to the primary key, the relation must be declared “index organized”. The goal of the following is to compare these three possibilities. Consider the following relation, that supposedly contains data on what things have been sold together (Amazon.com stores data like this):

```
CREATE TABLE CoBuys (  
  ID int,  
  LOCATION varchar(20),  
  WHAT1 varchar(40),  
  WHAT2 varchar(40),  
  QUANTITY1 int,  
  QUANTITY2 int  
);
```

You should create this relation and import some data into it by the command: `INSERT INTO CoBuys (SELECT * FROM pagh.CoBuys)`; Secondly, create another relation with the same schema but the primary key declaration `PRIMARY KEY (location,what1,what2,id)`. Finally, create a third identical relation which is index organized: Specify `ORGANIZATION INDEX` after the schema. Now run various queries with autotrace enabled, and compare the query plans, and number of physical disk accesses. Note that these numbers can be blurred by various caches in the system. For example, try the following queries on all three relations:

```
select sum(quantity1) from CoBuys where location='London';  
select sum(quantity1) from CoBuys where location='London' AND what1='Camel saddle';
```

2. RG exercise 20.1
3. RG exercise 20.5
4. RG exercise 20.7