Question 1: Skyline Queries (1 P.)

Given the TPC-H[1] table `part`.

We want to compute the skyline over the following four dimensions:

- *p_size:* larger is better,
- *p_retailprice:* less is better,
- *p_container:* Use the function `getContainerSize` from below: less is better.
- *p_brand:* Cannot be compared.

The function `getContainerSize` can be created by using the following command:

```
CREATE OR REPLACE FUNCTION getContainerSize(d1 TEXT) RETURNS
INT AS $$
DECLARE
  d1SizeStr TEXT;
BEGIN
  d1SizeStr=split_part(d1, ' ', 1);
  if (d1SizeStr) = 'WRAP' then return 0; end if;
  if (d1SizeStr) = 'SM' then return 1; end if;
  if (d1SizeStr) = 'MED' then return 2; end if;
  if (d1SizeStr) = 'LG' then return 3; end if;
  if (d1SizeStr) = 'JUMBO' then return 4; end if;
END;
$$ LANGUAGE plpgsql;
```

a) Create a SQL query to calculate the number of elements in the skyline defined above (without using the `SKYLINE`-operator). Submit the query and the result.

   Required submission: SQL Skyline query; Result of query

b) Write the same query using the `SKYLINE`-operator. (You do not have to install the plugin and execute the query.)

   Required submission: SQL Skyline query, using skyline operator

[1] http://dbis.informatik.uni-kl.de/files/teaching/ws1819/dbs/protected/tpch.dmp.gz

Question 2: Skyline NN (1 P.)

Implement the nearest neighbor (NN) method to compute skylines in a language of your choice. A basic Java template is available in OLAT. This template provides a (fake) R-Tree implementation which can be used to query for the nearest neighbor of a point, all points within a rectangle, and the nearest neighbor within a rectangle. If you are using another language, you are allowed to also implement a fake R-Tree with the methods described above (or use a real implementation without a included skyline function).

The program should output the calculated skyline (the template already takes care of this).

Submit your code and the output for the following points (If you use the template, the points are already included):

$$(10, 20), (12, 10), (8, 11), (16, 19), (6, 4), (5, 6), (14, 12), (2, 5), (3, 10), (13, 19), (17, 5), (9, 3), (20, 8), (8, 10)$$

Required submission: Source code; Output of program (must contain skyline points)

Question 3: Transactions (1 P.)

a) A database has the consistency condition $0 \leq A \leq B$. Describe, for the following transactions, if the condition is fulfilled. Explain your answer.

Required submission: Explanation how the consistency condition holds or breaks

T_1 : `B:= A+B; A:= A+B;`

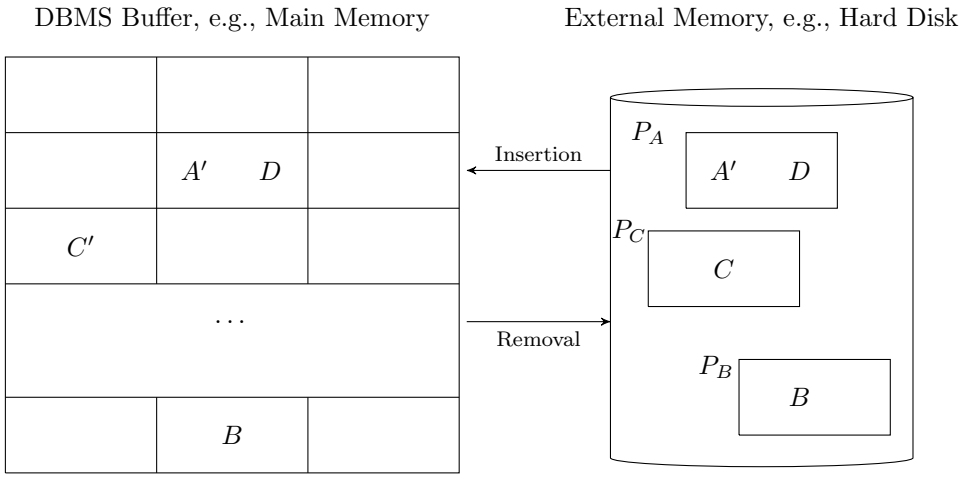
T_2 : `A:= B+1; B:= A+1;`

T_3 : `A:= A+B; B:= A+B;`

b) For each of the previous transactions, provide the read and write operations. Describe the effect of these operations on the main memory and the hard disk. The initial values are $A = 5$ and $B = 10$.

Required submission: List of INPUT/READ/WRITE operations;

c) Given two transactions T_1 and T_2 and the following situation (' marks changed entries):



The following table describes the operations of T_1 and T_2 at different times.

Time	T_1	T_2
0	READ(A,a)	READ(C,c)
10	a:=a+10	c:=c*2
20		READ(B,b)
30	WRITE(A,a)	
40	READ(D,d)	b:=b+c/4
50	d:=17*d+42	
60	OUTPUT(A)	WRITE(C,c)
70	WRITE(D,d)	OUTPUT(C)
80	COMMIT	
90		COMMIT

Discuss if the illustration matches the entries of the table. Does it only represent the table at a specific time (if yes, when and why) or does it not represent it at all?

If the system crashes during the execution, what would have to be done to guarantee ACID? Which parts of ACID are affected? How does the situation change at timestamp 91?

Required submission: Discussion

Question 4: Recovery (1 P.)

In a DBMS three transactions T_1, T_2, and T_3 are executed concurrently. The data accessed performed by the transactions are stored in the log (Table 1).

a) The system crashes after step 18. None of the changes were written to the database, but the log is complete. Using this log, perform the three stages of recovery and explain what happens.

Required submission: Detailed recovery procedure with intermediate steps

b) Is the log changed after completing the recovery process? If yes, explain what changed.

	T_1	T_2	T_3	Log
				[LSN, TA, PageID, Redo, Undo, PrevLSN]
1.	BOT			[#01, T_1, -, BOT, -, 0]
2.	$r(A, a_1)$			
3.	$a_1 = a_1 - 2.3$			
4.		BOT		[#02, T_2, -, BOT, -, 0]
5.	$w(A, a_1)$			[#03, T_1, P_A, $A = A - 2.3$, $A = A + 2.3$, #01]
6.		$r(B, b_1)$		
7.			BOT	[#04, T_3, -, BOT, -, 0]
8.		$b_1 = b_1 - 10$		
9.		$w(B, b_1)$		[#05, T_2, P_b, $B = B - 10$, $B = B + 10$, #02]
10.			$r(B, b_2)$	
11.			$b_2 = \frac{b_2}{2}$	
12.			$w(B, b_2)$	[#06, T_3, P_b, $B = \frac{B}{2}$, $B = B * 2$, #04]
13.		$r(C, c_1)$		
14.		$c_1 = c_1 + 5$		
15.		$w(C, c_1)$		[#07, T_2, P_c, $C = C + 5$, $C = C - 5$, #05]
16.	$r(E, e_1)$			
17.		abort		[#08, T_2, -, abort, -, #07]
18.			commit	[#09, T_3, -, commit, -, #06]
				Crash

Tabelle 1: Log