



Implementation of Databases (WS 17/18)

Exercise 4

Due until December 5, 2017, 10am.

Please submit your solution *in a single PDF file* before the deadline to the L²P system!

Please submit solutions in groups of three students.

Exercise 4.1 (Tableaux Containment and Minimization)

(10 pts)

Given are the following tableaux:

T_1			T_2			T_3		
a1	a2		a1	a2		a1	a2	
b3	a2	(R)	b4	a2	(R)	b3	a2	(R)
a1	b4	(R)	b1	a2	(R)	a1	b2	(R)
5	b3	(R)	a1	b3	(R)	b4	b1	(R)
b4	5	(R)	b2	b4	(R)	b1	b2	(R)
			b2	b1	(R)	b2	b3	(R)
			b3	b2	(R)	b1	b3	(R)

1. Find out if $T_i \subseteq T_j$ i.e., $T_i \equiv T_j$ for $i \neq j, i, j \in \{1, 2, 3\}$.

2. Write down the minimal tableau for $T_i, i \in \{1, 2, 3\}$.

Exercise 4.2 (Quant Graphs)

(12 pts)

Given is the Chinook database schema (See the schema given in L2P).

- Specify the following queries in TRC and draw the corresponding quant graphs, subsequently. Determine for each graph, if it contains a cycle. Explain in your own words what this means for the query.
 - EmployeeId and LastName of employees who have supported at least two different customers living in the same city.
 - Track name and composer of all tracks, which are from the genre “Rock” and where the artist is also the composer.

- Given is the following query as tableau. Translate it into Tuple Relational Calculus (TRC) and draw the quant graph for the query. Does the graph contain a cycle? What is the meaning of the query?

TAG	TrackName	Milliseconds	GenreId	GenreName	MediaTypeId	MediaType
ANSWER	tn					
Track	tn	≤ 90000	gid		mid	
Genre			gid	"Rock"		
MediaType					mid	"MPEG audio file"

Exercise 4.3 (External Sorting)

(8 pts)

Suppose that you just finished inserting several records into a heap file and now want to sort those records. Assume that the DBMS uses external sort and makes efficient use of the available buffer space when it sorts a file. Here is some potentially useful information about the newly loaded file and the DBMS software available to operate on it:

The number of records in the file is 7500. Each record is a total of 36 bytes long. The page size is 1024 bytes. Each page has 64 bytes of control information on it. Four buffer pages are available.

- How many sorted subfiles will there be after the initial pass of the sort, and how long will each subfile be?
- How many passes (including the initial pass just considered) are required to sort this file?
- What is the total I/O cost for sorting this file?
- What is the largest file, in terms of the number of records, you can sort with just four buffer pages in two passes? How would your answer change if you had 42 buffer pages?