CHALMERS
EXAMINATION / TENTAMEN

Course code/kurskod	Course name/kursnamn			
DIT 033	Data Management			
Anonymous code Anonym kod	Examination date Tentamensdatum	Number of pages Antal blad	Grade Betyg	
427	22-06-09	6	5.	

* I confirm that I've no mobile or other similar electronic equipment available during the examination.
Jag intygar att jag inte har mobiltelefon eller annan liknande elektronisk utrustning tillgänglig under eximinationen.

Solved task Behandlade uppgifter No/nr		Points per task Poäng på uppgiften	Observe: Areas with bold contour are to completed by the teacher. Anmärkning: Rutor inom bred kontur ifylles av lärare.
1	X	14.5	
2	X	20	
3	X	12.	
4	X	18.	
5	X	20.	
6	X	6.	
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
Bonus poäng			
Total examination points Summa poäng på tentamen		90.5	

CHALMERS	Anonymous code Anonym kod	DIT033 427	Points for question (to be filled in by teacher) Poäng på uppgiften (ifylles av lärare)	Consecutive page no. Löpande sid nr	1
				Question no. Uppgift nr	1

1.1 A weak entity is an entity that cannot exist without a "strong" entity and has no unique key but a partial key which together with the key of the other entity is unique.

Example: A room cannot exist without a building it can be found in. The room number is not globally unique, but together with the address of the building it is.

3 ✓

1.2 Semi-structured data has a ~~structured~~ minimum required structure but still allows multiple values ~~and~~ to be stored as an array as well as composite attributes as objects. Additionally, optional additional information can be stored.

3.5

Unstructured data can be anything since it enforces no type of structure. Structured data is very strict and only specified attributes in the schema can be stored. Also each attribute can only store one value.

1.3 A full table scan goes over all the entries in a table to check for a certain query.

1.5

Example: SELECT * FROM R; where= ? clause.

1.4 Views allow you to give a name ✓ to the result of a query, making it easier to understand/write more complicated queries.

1

Types: window view, mountain view, city view — no, I have no idea.

1.5 Atomacity - transaction either happens fully or not at all.

4

Consistency - the results of an action should be consistent

Isolated - each transaction should be isolated and independant from others.

Durability - the changes are stored in the database.

1.6 Sharding is splitting up the data in shards, this works great when the data and its relations are scoped to some level like in slack the workspaces (I believe that is what they are called), GitHub organizations/users. It allows you to store data of let's say the first 100 users in the first shard, the next 100 in the second shard and so on. And when most of the data is scoped like that there will be little to no relations between shards.

2

To find which shard to ask for the data, a mapping function is used, probably by looking at a part/hash of the ids.

Answer only one question on this page. Do not write on the back of this paper
Behandla endast en uppgift på detta blad. Skriv ej på baksidan

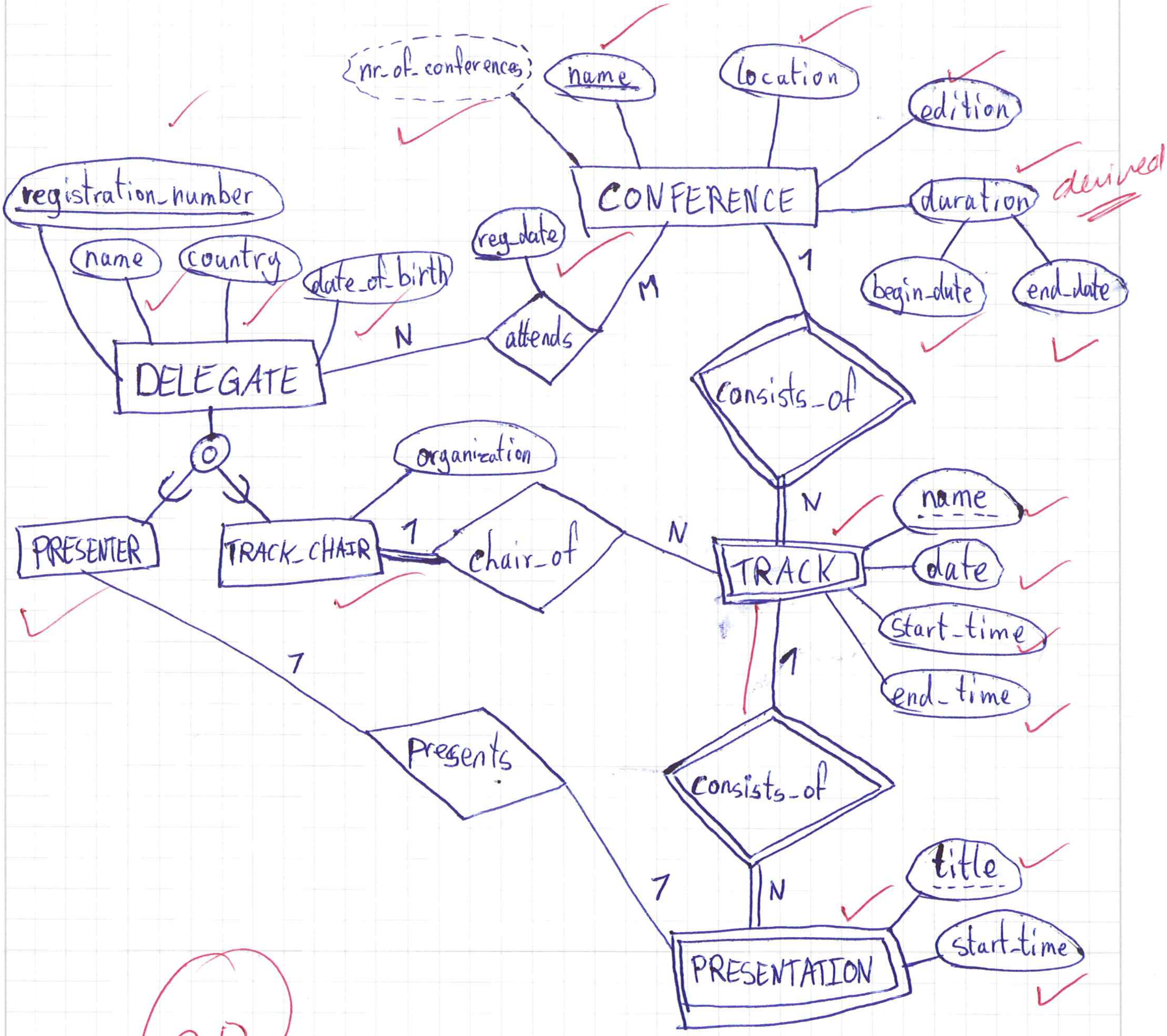
CHALMERS

Anonymous code
Anonym kod
DIT033
427

Points for question
(to be filled in by teacher)
Poäng på uppgiften
(ifylles av lärare)

Consecutive page no.
Löpande sid nr
2

Question no.
Uppgift nr
2



20

Excellent!

2

CHALMERS

Anonymous code
Anonym kod
DIT033
427

Points for question
(to be filled in by teacher)
Poäng på uppgiften
(ifylles av lärare)

Consecutive page no.
Löpande sid nr 3

Question no.
Uppgift nr 3

3

CONFERENCE (name, location, edition, start_date, enddate)

DELEGATE (registration_number, name, country, date_of_birth, organization)

ATTENDS (conference, delegate, reg_date)

conference -> CONFERENCE.name
delegate -> DELEGATE.registration_number

TRACK (conference, name, date, start_time, end_time, chair)

conference -> CONFERENCE.name
chair -> DELEGATE.registration_number

PRESENTATION (conference, track, title, start_time, presenter)

{conference, track} -> {TRACK.conference, TRACK.name}
presenter -> DELEGATE.registration_number

12

Excellent

Notes:

- Creating an own table for conference durations would make it unnecisarily complex, especially since it are only two values with no clear single key candidate.
- One table for all seems the best approach for delegate since track chair only adds a single attribute. The existance of this attribute for all track chairs can be easily checked in a hook.

* Single table is a good approach but it should have attribute is_chair and is_presenter.

Answer only one question on this page. Do not write on the back of this paper
Behandla endast en uppgift på detta blad. Skriv ej på baksidan

2

CHALMERS

Anonymous code
Anonym kod
DIT033
427

Points for question
(to be filled in by teacher)
Poäng på uppgiften
(ifylles av lärare)

Consecutive page no.
Löpande sid nr 4

Question no.
Uppgift nr 4

4.1 $\pi_{fname \rightarrow Josh\_names}(\sigma_{lname = "Josh"}(PERSON))$

4

4.2 ~~$\pi_{isbn, title, name}(BOOK \bowtie$~~

~~$\pi_{isbn, title, name}(BOOK \bowtie_{Book.isbn =} (BOOK\_LIB \bowtie$~~

5

$LIBRARY\_OF\_BOOK \leftarrow BOOK\_LIB \bowtie_{BOOKLIB.library = LIBRARY.id} LIBRARY$

$\pi_{isbn, title, name}(BOOK \bowtie_{BOOK.isbn = LIBRARY\_OF\_BOOK.book} LIBRARY\_OF\_BOOK)$

4.3 $HC\_BOOK\_LIB \leftarrow BOOK\_LIB \bowtie_{BOOK\_LIB.book = BOOK.isbn} (\sigma_{cover = "hc"} BOOK)$ ✓

$\pi_{title, page\_count, name, location}(LIBRARY \bowtie_{LIBRARY.id = HC\text{-}BOOK\text{-}LIB.library} HC\_BOOK\_LIB)$

5

4.4 $_{library}\mathcal{F}_{COUNT(book)}(BOOK\_LIB)$

4

Answer only one question on this page. Do not write on the back of this paper
Behandla endast en uppgift på detta blad. Skriv ej på baksidan

2

CHALMERS

Anonymous code
Anonym kod
DIT033
427

Points for question
(to be filled in by teacher)
Poäng på uppgiften
(ifylles av lärare)

Consecutive page no.
Löpande sid nr
5

Question no.
Uppgift nr
5

5.1

```
CREATE TABLE PERSON (
    ssn NUMBER PRIMARY KEY, ✓
    fname VARCHAR(64), ✓
    lname VARCHAR (64), ✓
    dob DATE, ✓
    location VARCHAR(64), ✓
    FOREIGN KEY location REFERENCES CITY(name), ✓
    CHECK (dob < to_date("2004-06-09"))
    ~~CHECK (dob < to_date "2004-06-09"))~~ ✓
);
```

5

5.2

```
SELECT fname, lname, location FROM PERSON
  WHERE location LIKE "%köping";
```

✓ 5

5.3

```
SELECT title, location FROM BOOK_LIB
  INNER JOIN BOOK ON BOOK_LIB.book = BOOK.isbn
  INNER JOIN LIBRARY ON BOOK_LIB.library = LIBRARY.id;
```

5 ✓

5.4

```
SELECT count(isbn) FROM LIBRARY
  INNER JOIN BOOK_LIB ON LIBRARY.id = BOOK_LIB.library
  INNER JOIN BOOK ON BOOK_LIB.book = BOOK.isbn
  WHERE location = "Gothenburg" AND cover = "SC";
```

✓ 5

Answer only one question on this page. Do not write on the back of this paper
Behandla endast en uppgift på detta blad. Skriv ej på baksidan

2

CHALMERS	Anonymous code Anonym kod	DIT033 427	Points for question (to be filled in by teacher) Poäng på uppgiften (ifylles av lärare)
		Consecutive page no. Löpande sid nr	6
		Question no. Uppgift nr	6

6

1. conference is missing its closing tag
2. the first presentation misses the > at the end of the opening tag
3. the title of the first presentation should be in quotes "
4. the first presentation misses ~~two~~ its closing tag

The double id at the second presentation is a semantic error, not a syntactical.

its a syntactical error.

6