AALBORG UNIVERSITY

Cassiopeia

Database Systems – Spring 2014
DPT Group, Aalborg University
Handed out: 12.03.2014

Teacher: Katja Hose
Self study 3: exam preparation 1
Deadline: 18.03.2014

Exam preparation guideline:
If you are well prepared for the exam, solving these exercises should not take you much longer than 2 hours. At the moment, however, you have not finished exam preparation, so it is expected that you are still slower.

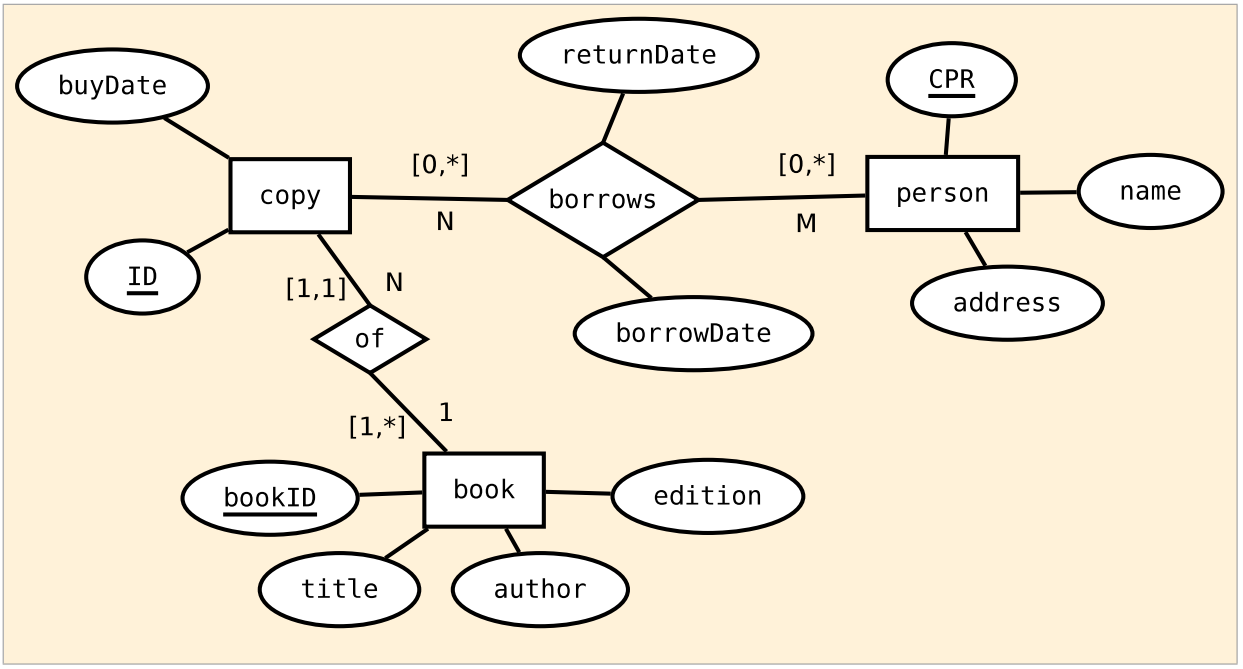
1 ER Modeling

We would like to create a database that captures information about borrowing books from the university's library. We want to model the following information:

- Persons can borrow copies of books.
- Books are uniquely identified by a bookID.
- For each book we also store title, author, and edition.
- The university's library can have multiple copies of a book. We store information about each individual copy and assign to it a unique identifier so that we can unambiguously identify the person that borrowed a particular copy of a book.
- In addition, we also store the date when a particular copy was bought.
- For each person we store name, address, and CPR number.
- We also store the date when a copy was borrowed as well as the return date.

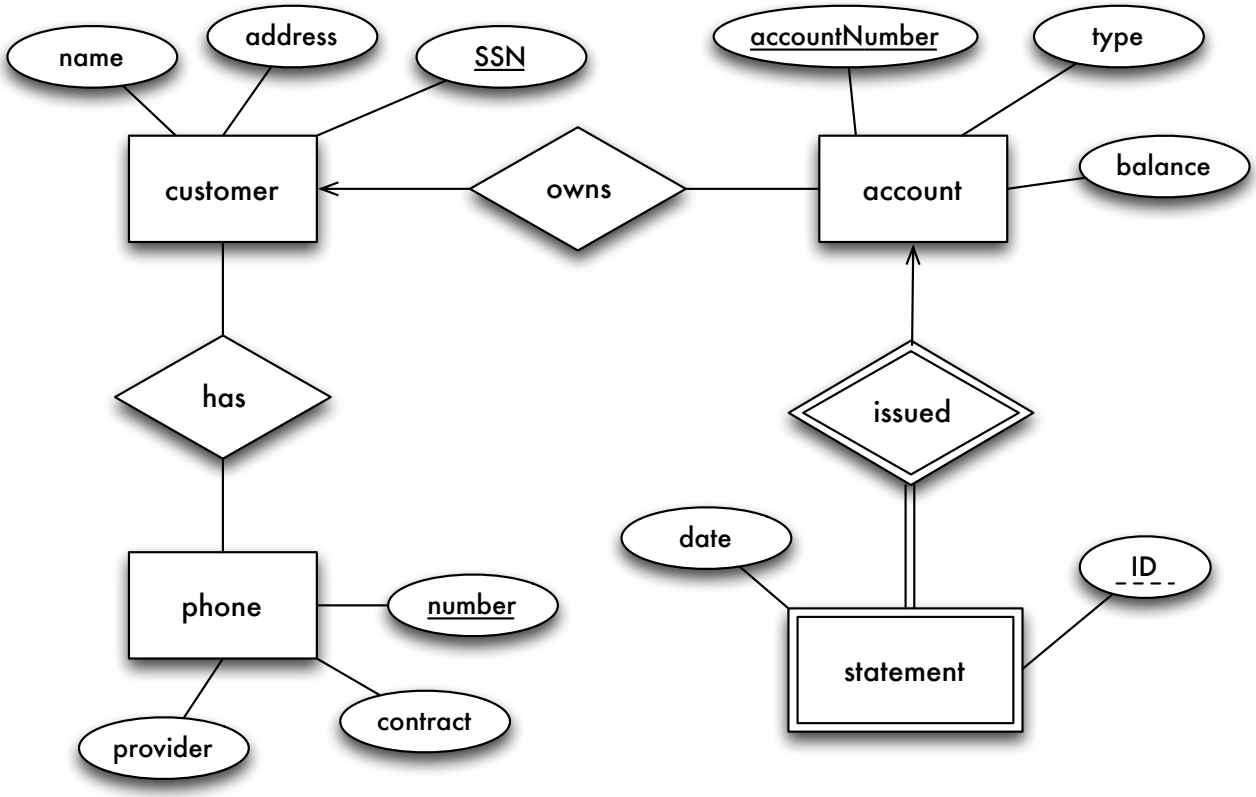
Please create an ER diagram that models this information. Please mark the keys and add information about cardinality ratios (Chen notation) and cardinality limits ([min,max] notation). The design should be minimal, do not model additional ID attributes if they are not necessary.

Solution.


buyDate
returnDate
CPR
copy
[0,*]
borrows
[0,*]
person
name
N
M
ID
[1,1]
N
of
borrowDate
address
[1,*]
1
bookID
book
edition
title
author

2 Banking System

Please map the following ER diagram to relations. Only a few null values are expected, hence do not create extra relations if an alternative solution with fewer relations is possible. The relations should be minimal, do not add additional attributes.



Solution.

- *customer: {[SSN, name, address]}*
- *account: {[accountNumber, type, balance, owner → customer]}*
- *phone: {[number, provider, contract]}*
- *has: {[SSN → customer, number → phone]}*
- *statement: {[number → account, ID, date]}*

3 Relational Algebra

Given the following relations:

animals						
animalID	nickname	species	gender	zooID	father	mother
1	Tally	giraffe	female	1	3	2
2	Kathy	giraffe	female	2	–	–
3	Billy	giraffe	male	3	–	–
4	Stan	ape	male	1	5	–
5	Pam	ape	male	2	–	–
6	Uhu	owl	male	2	9	8
7	Jahoo	owl	male	1	10	11
8	Boo	owl	female	1	10	11
9	Wohoo	owl	male	2	–	–
10	Huhuu	owl	male	1	–	–
11	Eule	owl	female	1	–	–

zoos			
zooID	name	city	country
1	Zoo Frankfurt	Frankfurt	Germany
2	Tiergarten Schönbrunn	Vienna	Austria
3	Aalborg Zoo	Aalborg	Denmark

Please determine the results for the following relational algebra expressions:

1. $(\pi_{species,zooID} \quad (animals)) \quad \div \quad (\pi_{zooID} \quad (\sigma_{country='Germany'} \quad (zoos)))$

2. $\pi_{T1.nickname} \quad (\sigma_{T1.animalID=T2.father} \quad \vee \quad T1.animalID=T2.mother($
   $(\rho_{T1}(animals)) \quad \bowtie_{T1.zooID=T2.zooID} \quad (\rho_{T2}(animals))))$

Solution.

1.

species
giraffe
ape
owl

2.

nickname
Wohoo
Huhuu
Eule

4 Relational Calculus

Consider the following relations:

`Suppliers(`*sid*`: integer,` *sname*`: string,` *address*`: string)`
`Parts(`*pid*`: integer,` *pname*`: string,` *color*`: string)`
`Catalog(`*sid*`: integer,` *pid*`: integer,` *cost*`: real)`

The `Catalog` relation lists the prices charged for parts by suppliers.

Please find expressions for the following queries in

- relational algebra
- tuple relational calculus
- domain relational calculus

1. Find the *names* of suppliers who supply some red part.

Solution.

- *RA:* $\pi_{sname}((\sigma_{color='\text{red}'}(Parts) \bowtie Catalog) \bowtie Suppliers)$
- *TRC:* $\{s.sname \mid s \in Suppliers \wedge \exists p \in Parts(p.color = 'red' \wedge \exists c \in Catalog(c.pid = p.pid \wedge c.sid = s.sid))\}$
- *DRC:* $\{< y >| \exists x, z(< x, y, z >\in Suppliers \wedge \exists p, q, r(< p, q, r >\in Parts \wedge r = 'red' \wedge \exists k(< x, p, k >\in Catalog)))\}$

2. Find the *sids* of suppliers who supply some red or green part

Solution.

- *RA:* $\pi_{sid}((\sigma_{color='\text{red}' \vee color='\text{green}'}(Parts) \bowtie Catalog)$
- *TRC:* $\{c.sid \mid c \in Catalog \wedge \exists p \in Parts((p.color = 'red' \wedge p.color = 'green') \wedge c.pid = p.pid)\}$
- *DRC:* $\{< x >| \exists y, z(< x, y, z >\in Catalog \wedge \exists b, c(< y, b, c >\in Parts \wedge (c = 'red' \vee c = 'green')))\}$

3. Find the *sids* of suppliers who supply some red part and some green part.

Solution.

- *RA:* $\rho_{R1}(\pi_{sid}(\sigma_{color='red'}(Catalog \bowtie Parts))) \cap \rho_{R1}(\pi_{sid}(\sigma_{color='green'}(Catalog \bowtie Parts)))$ – *renaming is not necessary*

- *TRC:* $\{c1.sid \mid c1 \in Catalog \wedge \exists p1 \in Parts(p1.color = \text{'red'} \wedge p1.pid = c1.pid \wedge \exists c2 \in Catalog \wedge \exists p2 \in Parts(p2.color = \text{'green'} \wedge p2.pid = c2.pid \wedge c2.sid = c1.sid))\}$
- *DRC:* $\{< x > | \exists y, z(< x, y, z > \in Catalog \wedge \exists b, c(< y, b, c > \in Parts \wedge c = \text{'red'} \wedge \exists q, r(< x, q, r > \in Catalog \wedge \exists f, g(< q, f, g > \in Parts \wedge g = \text{'green'}))))\}$

4. Find pairs of *sids* such that the supplier with the first *sid* charges more for some part than the supplier with the second sid.

Solution.

- *RA:* $\pi_{R1.sid,R2.sid}(\sigma_{R1.sid \neq R2.sid \wedge R1.cost > R2.cost}(\rho_{R1}(Catalog) \bowtie_{R1.pid=R2.pid} \rho_{R2}(Catalog)))$
- *TRC:* $\{t1.sid, t2.sid \mid t1 \in Catalog \wedge t2 \in Catalog(t2.pid = t1.pid \wedge t2.sid \neq t1.sid \wedge t2.cost < t1.cost)\}$
- *DRC:* $\{< x, p > | \exists y, z(< x, y, z > \in Catalog \wedge \exists r(< p, y, r > \in Catalog \wedge p \neq x \wedge r < z))\}$

5. Find the *pids* of parts supplied by at least two different suppliers.

Solution.

- *RA:* $\pi_{R1.pid}(\sigma_{R1.sid \neq R2.sid}(\rho_{R1}(Catalog) \bowtie_{R1.pid=R2.pid} \rho_{R2}(Catalog)))$
- *TRC:* $\{t1.pid \mid t1 \in Catalog \wedge \exists t2 \in Catalog(t2.pid = t1.pid \wedge t2.sid \neq t1.sid)\}$
- *DRC:* $\{< y > | \exists x, z(< x, y, z > \in Catalog \wedge \exists a, c(< a, y, c > \in Catalog \wedge a \neq x))\}$

5 Functional dependencies

Consider the following table with three columns (A, B, and C):

	A	B	C
1	a_1	b_1	c_1
2	a_1	b_2	c_2
3	a_2	b_3	c_1
4	a_2	b_3	c_2

For each of the functional dependencies listed below, indicate whether the above tuples violate it or not. If the functional dependency is not violated, write *OK*. If it is, indicate a combination of tuples in the above table that violates the functional dependency – if there are more, one combination is sufficient.
Refer to the tuples as 1,2,3,4; for example, for $A \rightarrow C$ you may answer with "*violated: tuples 3, 4*".

FD	OK or violated?
$A \rightarrow C$	*violated: tuples 3,4*
$B \rightarrow A$	
$C \rightarrow A$	
$A \rightarrow B$	
$B \rightarrow C$	
$BC \rightarrow A$	
$AC \rightarrow B$	

Solution.

FD	*Holds?*
$B \rightarrow A$	*OK*
$C \rightarrow A$	*violated: tuples 1,3 or 2,4*
$A \rightarrow B$	*violated: tuples 1,2*
$B \rightarrow C$	*violated: tuples 3,4*
$BC \rightarrow A$	*OK*
$AC \rightarrow B$	*OK*

Instructions to prepare the report

We will try to have a double-blinded review process, i.e., you do not know the group whose report you are reviewing and you do not know which group is reviewing your own report. Hence, state your group name as well as the members of your group on the first page only. Do not repeat your group name and the names of group members on other pages. What will be handed over to the peer group is then everything except the first page.

Course goals covered by this self study

- Understand the relational model and apply relational algebra as well as relational calculus
- Conceptually design a database (ER model, conceptual design)
- Basic understanding of functional dependencies

Instructions for peer review

After the deadline to hand in your own reports (18.03.2014), you will receive a report handed in from another group. Your task is to "grade" the handed in solutions and decide whether they are correct and how many points should be awarded.

To give you some guidance, we will provide you with a set of solutions for the exercises. Your task is then to decide how correct the solutions of your peer group are. As there is often more than one correct solution, you first need to decide on the correctness and then identify how much weight you want to give to coverage and mistakes in the solutions (regarding both the semantics of the solution and the notation).

Hence, for each exercise and handed in solution, you should give a short statement on how correct (in the range of 0% to 100%) the solution is. Furthermore, you should explain how and why you gave less then 100% – be explicit, e.g., we gave 80% because..., we reduced by 5% because.... In addition, please also give a final grade to the report that you were given – use the 7-trins-skala and add a short explanation.

As a consequence of the double-blinded review process, for the feedback please again prepare a first page listing group name and members and do not repeat this information on other pages.

Selfstudy: 12.03.2014
The report must be handed in via Moodle no later than
18.03.2014, 23:55 CET
Peer reviews must be handed in via Moodle no later than
25.03.2014, 23:55 CET