Datenbanken
Übungsblatt 6 Abgabe der Lösungen: 5. Mai 2011

Aufgabe 1

Studieren Sie das umseitig gegebene Datenbankschema. Beschreiben Sie umgangssprachlich, welche Abfragen die folgenden Ausdrücke der relationalen Algebra realisieren:

1. $\pi_{title, price}(\sigma_{type = "psychology" \wedge price > 7 \wedge price < 19.99}(titles))$
2. $\pi_{title, price}(\sigma_{lorange \leq total-sales \wedge total-sales \leq high-range}(titles \bowtie roysched))$

Korrigieren sie die untenstehenden Abfragen und erläutern Sie die Fehler.

1. $\pi_{au-lname, au-fname, pub-name}(((authors \bowtie titleauthors) \bowtie titles) \bowtie publishers)$

   Wieso ergibt die Abfrage nicht wie erwartet die Autoren zusammen mit den Verlegern, bei denen sie publiziert haben? Korrigieren Sie.

2. $\pi_{pub_id}(\sigma_{no_titles \leq 2}({}_{pub-id}\mathcal{G}_{count(title-id)\ AS\ no-titles}(\sigma_{type = "psychology"}(titles))))$

   Wieso ergibt die Abfrage nicht wie gewünscht alle Verleger, die höchstens zwei Psychologiebucher verlegt haben? Korrigieren Sie.

Aufgabe 2

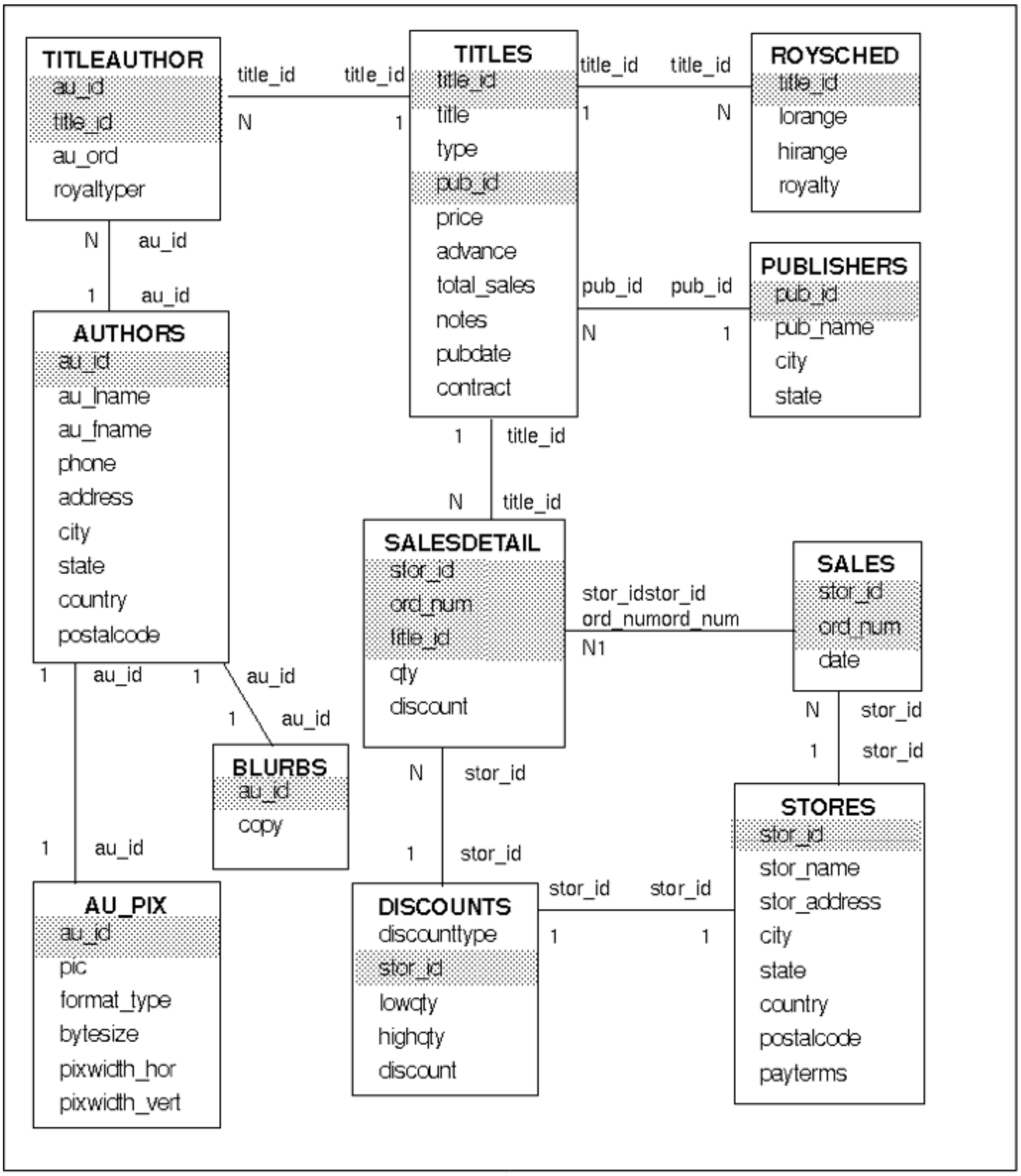
Zeigen Sie:

1. dass $\sigma_p(r \times s) = \sigma_p(r) \times s$ für alle Relationen r und s und für alle Prädikate p, die sich nur auf die Attribute von r beziehen, gilt.
2. weshalb $(r \div s) \times s$ eine Teilmenge von r liefert, falls die Attribute von s auch in r vorkommen. Unter welchen Umständen sind beide Mengen gleich?

Aufgabe 3

Berechnen Sie die für beiden untenstehende Abfragen die Grössen der bei der Auswertung entstehenden Zwischentabellen, wenn *borrower* und *loan* je 4000 Tupel enthalten und die Selektionsprädikate jeweils auf 25% der Tupel zutreffen. Welche der beiden Abfragen ist effizienter?

- $\pi_{customer-name}(\sigma_{branch-name = "Perryridge"}(\sigma_{borrower.loan-number = loan.loan-number}(borrower \times loan)))$
- $\pi_{customer-name}(\sigma_{loan.loan-number = loan.loan-number}(\sigma_{branch-name = "Perryridge"}(loan) \times borrower))$


TITLEAUTHOR
au_id
title_id
au_ord
royaltyper
TITLES
title_id
title
type
pub_id
price
advance
total_sales
notes
pubdate
contract
ROYSCHED
title_id
lorange
hirange
royalty
PUBLISHERS
pub_id
pub_name
city
state
AUTHORS
au_id
au_lname
au_fname
phone
address
city
state
country
postalcode
SALESDETAIL
stor_id
ord_num
title_id
qty
discount
SALES
stor_id
ord_num
date
BLURBS
au_id
copy
AU_PIX
au_id
pic
format_type
bytesize
pixwidth_hor
pixwidth_vert
DISCOUNTS
discounttype
stor_id
lowqty
highqty
discount
STORES
stor_id
stor_name
stor_address
city
state
country
postalcode
payterms
title_id
title_id
N
1
title_id
title_id
1
N
pub_id
pub_id
N
1
N
au_id
1
au_id
1
title_id
N
title_id
stor_idstor_id
ord_numord_num
N1
1
au_id
1
au_id
1
au_id
1
au_id
N
stor_id
1
stor_id
N
stor_id
1
stor_id
stor_id
stor_id
1
1

Aufgabe 4
Gegeben sei folgendes Relationenschema:

$dreieck(DreieckId,\ EckeNr,\ x,\ y,\ z)$.

Der Primärschlüssel sei gegeben durch die Menge $\{DreieckId, EckeNr\}$.

Dann gibt es für jedes Dreieck genau drei Einträge in dieser Relation, nämlich einen für jede Ecke mit den entsprechenden Koordinaten.

Geben Sie einen relationaen Ausdruck an, der für jedes Dreieck die Koordinaten seines Schwerpunktes ausgibt.

Hinweis: Der Schwerpunkt eines Dreiecks ist die Summe der Ortsvektoren der drei Ecken geteilt durch 3.

Beispiel: Das Dreieck mit den Ecken $(0, 0, 0)$, $(0, 1, 0)$ und $(0, 1, 1)$ hat den Schwerpunkt $(0, \frac{2}{3}, \frac{1}{3})$