Lehrstuhl für Praktische Informatik III
Prof. Dr. Guido Moerkotte
Email: moer@db.informatik.uni-mannheim.de

Datenbanksysteme 1 — 3. Übungsblatt
Herbst-/Wintersemester 2016

Aufgabe 1

Gegeben seinen die beiden folgenden Relationenausprägungen L und R.

L		
a	b	c
a_1	b_1	c_1
a_2	b_2	c_2
a_3	b_3	c_3

R		
c	d	e
c_1	d_1	e_1
c_3	d_3	e_3
c_4	d_4	e_4

Geben Sie das Ergebnis der folgenden Ausdrücke an.

Aufgabe 1 a)

$L \bowtie_{L.c=R.c} R$ Lösung

$L \bowtie R$					
a_1	b_1	c_1	c_1	d_1	e_1
a_3	b_3	c_3	c_3	d_3	e_3

Aufgabe 1 b)

$L ⟕_{L.c=R.c} R$ Lösung

$L ⟕ R$					
a_1	b_1	c_1	c_1	d_1	e_1
a_2	b_2	c_2	-	-	-
a_3	b_3	c_3	c_3	d_3	e_3

Aufgabe 1 c)

$L ⟖_{L.c=R.c} R$ Lösung

$L \bowtie R$					
a_1	b_1	c_1	c_1	d_1	e_1
a_2	b_2	c_2	-	-	-
a_3	b_3	c_3	c_3	d_3	e_3
-	-	-	c_4	d_4	e_4

Aufgabe 1 d)

$L \ltimes_{L.c=R.c} R$ Lösung

$L \ltimes R$		
a_1	b_1	c_1
a_3	b_3	c_3

Aufgabe 1 e)

$L \rhd_{L.c=R.c} R$

Lösung

$L \rhd R$		
a_2	b_2	c_2

Aufgabe 1 f)

$L \div R$

Lösung

$L \div R$

Aufgabe 1 g)

$L \div \Pi_{R.C}(\sigma_{R.D=d_1}(R))$

Lösung

$L \div \Pi_{R.C}(\sigma_{R.D=d_1}(R))$	
a_1	b_1

Aufgabe 2

Gegeben sei folgendes relationales Schema für ein Weingut.

Rebsorte	{[<u>Sorte</u>, Name, Farbe]}
Wein	{[<u>ID</u>, Name, RSorte]}
Jahrgang	{[<u>WeinID, Jahr</u>, Preis, Qualität]}

In der Datenbank werden Weine verschiedener *Rebsorten* verwaltet. Die verschiedenen Weine werden in der Tabelle *Wein* verwaltet. Jeder Wein ist über den Fremdschlüssel *RSorte* genau einer Rebsorte zugeordnet. Für jeden *Jahrgang* wird für jeden Wein ein Preis und die Qualität vermerkt. Die Qualitäten "exzellent", "sehr gut" und "gut" usw. werden auf die Ordinalzahlen 1, 2, 3 usw. abgebildet.

Im folgenden sind Beispielausprägungen der Tabellen zu finden.

Rebsorte		
Sorte	Name	Farbe
1	Merlot	Rot
2	Riesling	Weiß
3	Spätburgunder	Rot

Wein		
ID	Name	RSorte
1	Marbuzet	1
2	Dr. Bassermann-Jordan	2
3	Herrgottsacker	2

Jahrgang			
WeinID	Jahr	Preis	Qualität
2	2003	20.00	1
2	2004	14.00	2
3	2001	50.00	1

Formulieren Sie die folgenden Anfragen in relationaler Algebra.

Aufgabe 2 a)

Geben die Farbe der Rebsorte Merlot an.

Lösung

$$\Pi_{Farbe}(\sigma_{Name=\text{“Merlot”}}(\text{Rebsorte}))$$

Aufgabe 2 b)

Aus welcher Rebsorte ist der Wein mit dem Namen "Bassermann-Jordan" gemacht.

Lösung

$$\Pi_{Name}(\text{Rebsorte} \bowtie_{RSorte=Sorte} (\Pi_{RSorte}(\sigma_{Name=\text{“BJ”}}(\text{Wein}))))$$

Aufgabe 2 c)

Listen Sie den Namen und die Preise aller Riesling-Weine auf.

Lösung

$$\Pi_{Name,Preis}((\Pi_{Sorte}(\sigma_{Name=\text{“Riesling”}}(\text{Rebsorte})))\bowtie_{RSorte=Sorte}(\text{Wein}\bowtie_{ID=WeinID}\text{Jahrgang}))$$

Aufgabe 2 d)

Für welche Rebsorten sind keine Weine in der Datenbank? Geben Sie deren Namen an.

Lösung

mittels Antijoin
$\Pi_{Name}((\text{Rebsorte}) \rhd_{Sorte=RSorte} (\text{Wein}))$
oder
$\Pi_{Name}((\text{Rebsorte}) \bowtie_{Sorte=RSorte} ((\Pi_{RSorte}(\rho_{RSorte \leftarrow Sorte}(\text{Rebsorte}))) - (\Pi_{RSorte}(\text{Wein}))))$

Aufgabe 2 e)

Welche Weine haben ausschließlich exzellente Qualität (Qualität = 1)? Listen Sie deren IDs auf.

Lösung

$\Pi_{WeinID}(\text{Jahrgang}) - \Pi_{WeinID}(\sigma_{Qualität \neq 1}(\text{Jahrgang}))$
oder
$\Pi_{WeinID}(\text{Jahrgang}) - (\Pi_{WeinID}((\text{Jahrgang}) \ltimes_{WeinID=ID} (\rho_{ID \leftarrow WeinID}(\sigma_{Qualität \neq 1}(\text{Jahrgang})))))$

Aufgabe 3

Die Informationen über das Streckennetz einer Bahngesellschaft werden in einer relationalen Datenbank gespeichert. Die Datenbank besitzt folgendes Schema:

- Zug($\underline{\text{ZugNr}}$, Zugtyp)
- Verbindung($\underline{\text{ZugNr, Wochentag}}$, StartBhf, ZielBhf)
- Teilstrecke($\underline{\text{ZugNr, Wochentag, vonBhf}}$, nachBhf, AbfahrtZeit, AnkunftZeit, Preis, Entfernung)

Die Relation *Zug* enthält Informationen über den Zugtyp eines Zuges (also z.B. IC, ICE, IR, etc.). In *Verbindung* ist vermerkt, an welchen Wochentagen ein Zug verkehrt und in welchem Bahnhof er startet und endet. Da ein Zug in der Regel mehrere Teilstrecken während einer Reise absolviert, existiert die Relation *Teilstrecke*, in der gespeichert wird, zwischen welchen Bahnhöfen im Einzelnen der Zug verkehrt (z.B.: Hamburg Hbf → Hannover Hbf, Hannover Hbf → Kasssel-Wilhelmshöhe, Kassel-Wilhelmshöhe → Frankfurt Hbf). Außerdem sind hier der Preis einer Fahrkarte für diese Teilstrecke und diesen Zug sowie die jeweilige Abfahrt- und Ankunftzeit zu finden.

Entscheiden Sie für die folgenden Anfragen, ob sie in relationaler Algebra formulierbar sind. Falls ja, geben Sie den algebraischen Ausdruck an. Falls nein, begründen Sie kurz Ihre Antwort.

Aufgabe 3 a)

Geben Sie die ZugNr, den Zugtyp und den Fahrpreis aller Züge auf der Teilstrecke Mannheim Hbf → Karlsruhe Hbf an. (Betrachten Sie dabei nur Direktverbindungen.)

Lösung

$$\pi_{\text{ZugNr,Zugtyp,Preis}}(\sigma_{\text{vonBhf='MannheimHbf'} \land \text{nachBhf='KarlsruheHbf'}}(\text{Teilstrecke} \bowtie \text{Zug})$$

Aufgabe 3 b)

Geben Sie die Start- und Zielbahnhöfe aller Teilstrecken an, auf denen alle Zugtypen eingesetzt werden.

Lösung

$$\pi_{\text{vonBhf,nachBhf,Zugtyp}}(\text{Teilstrecke} \bowtie \text{Zug}) \div \pi_{\text{Zugtyp}}(\text{Zug})$$

Es ergben sich folgende temporäre Relationen:

T		
vonBhf	nachBhf	Zugtyp

Z
Zugtyp

$$\begin{array}{lrcc} t \in T \div Z \quad \rightarrow & \forall t_Z \in Z & & \exists t_T \in T : \\ & t_T.Zugtyp & = & t_Z.Zugtyp \\ & t_T.(vonBhf, nachBhf) & = & t \end{array}$$

Aufgabe 3 c)

Geben Sie die Entfernung der längsten Teilstrecke an.

Lösung

$\pi_{\text{Entfernung}}(\text{Teilstrecke}) \ \backslash$
$(\pi_{\text{Entfernung}}($
$\text{Teilstrecke} \bowtie_{\text{Entfernung<Entfernung2}} \pi_{\text{Entfernung2}}(\rho_{\text{Entfernung2} \leftarrow \text{Entfernung}}(\text{Teilstrecke}))))$

Alternativ kann hier auch die Teilstrecke umbenannt werden und auf die Attribute dann explizit über den Relationennamen zugegriffen werden.

Anmerkung: Ich habe in der Übung die Notation zur Umbenennung von Attributen vertauscht. Der neue Attributname sollte *links* vom Pfeil stehen, wie oben!

Aufgabe 3 d)

Geben Sie die Zugnummern und alle Zwischenstationen derjenigen Züge an, die von München Hbf nach Hamburg Hbf fahren.

Lösung

Folgende Lösungen gehen davon aus, dass Züge nicht um Mitternacht unterwegs sind (das vereinfacht die Anfragen etwas, da sonst Sonderfälle eingeführt werden müssen, die vom eigentliche Lösungsprinzip ablenken). Falls Züge an jedem Wochentag immer die gleiche Strecke fahren, braucht man die Vergleiche `TSx.Wochentag = TSy.Wochentag` nicht (wurde in der Übung weggelassen).

- einfache Lösung:

$\pi_{\text{ZugNr,vonBhf,Abfahrtzeit,nachBhf,Ankunftzeit}}$
$(\sigma_{\text{StartBhf='MünchenHbf'} \land \text{Zielbhf='HamburgHbf'}}$

(Teilstrecke $\bowtie$ Verbindung))

- für Züge, die nicht unbedingt in München eingesetzt werden oder in Hamburg enden:

  $\pi_{\text{ZW.ZugNr,ZW.vonBhf,ZW.Abfahrtzeit,ZW.nachBhf,ZW.Ankunftzeit}}[$
  $[\sigma_{\text{vonBhf='MünchenHbf'}}(\rho_{\text{TS1}}(\text{Teilstrecke}))$
  $\bowtie_{\text{TS1.ZugNr=TS2.ZugNr} \land \text{TS1.Wochentag=TS2.Wochentag} \land \text{TS1.Abfahrtzeit<TS2.Ankunftzeit}}$
  $\sigma_{\text{nachBhf='HamburgHbf'}}(\rho_{\text{TS2}}(\text{Teilstrecke}))]$
  $\bowtie_{\substack{\text{TS1.ZugNr = ZW.ZugNr} \land \text{TS1.Wochentag = ZW.Wochentag} \land \\ \text{ZW.Abfahrtzeit < TS2.Ankunftzeit} \land \text{ZW.Ankunftzeit > TS1.Abfahrtzeit}}}$
  $(\rho_{\text{ZW}}(\text{Teilstrecke}))]$

- für Züge, die nicht in München eingesetzt werden und in Hamburg enden:

  $\pi_{\text{ZW.ZugNr,ZW.vonBhf,ZW.Abfahrtzeit,ZW.nachBhf,ZW.Ankunftzeit}}[$
  $[\sigma_{\text{vonBhf='MünchenHbf'}}(\rho_{\text{TS1}}(\text{Teilstrecke}))$
  $\bowtie_{\text{TS1.ZugNr=TS2.ZugNr} \land \text{TS1.Wochentag=TS2.Wochentag} \land \text{TS1.Abfahrtzeit<TS2.Ankunftzeit}}$
  $(\sigma_{\text{nachBhf='HamburgHbf'}}(\rho_{\text{TS2}}(\text{Teilstrecke}))$
  $\rhd_{\text{TS2.ZugNr=NTS.ZugNr} \land \text{TS2.Ankunftzeit<NTS.Abfahrtzeit} \land \text{TS2.Wochentag=NTS.Wochentag}}$
  $\sigma_{\text{vonBhf='HamburgHbf'}}(\rho_{\text{NTS}}(\text{Teilstrecke})))]$
  $\bowtie_{\substack{\text{TS1.ZugNr = ZW.ZugNr} \land \text{TS1.Wochentag = ZW.Wochentag} \land \\ \text{ZW.Abfahrtzeit < TS2.Ankunftzeit} \land \text{ZW.Ankunftzeit > TS1.Abfahrtzeit}}}$
  $(\rho_{\text{ZW}}(\text{Teilstrecke}))]$

- 1x umsteigen:

  $[\sigma_{\text{vonBhf='MünchenHbf'}}(\rho_{\text{TS1}}(\text{Teilstrecke}))$
  $\bowtie_{\text{TS1.ZugNr=TS2.ZugNr} \land \text{TS1.Wochentag=TS2.Wochentag} \land \text{TS1.Abfahrtzeit<TS2.Ankunftzeit}}$
  $(\rho_{\text{TS2}}(\text{Teilstrecke}))]$
  $\bowtie_{\substack{\text{TS2.nachBhf = TS4.vonBhf} \land \text{TS2.Wochentag = TS4.Wochentag} \land \text{TS2.ZugNr} \neq \text{TS4.ZugNr} \land \\ \text{TS2.Ankunftzeit + 5} \leq \text{TS4.Abfahrtzeit} \land \text{TS2.Ankunftzeit + 45} \geq \text{TS4.Abfahrtzeit}}}$
  $[\sigma_{\text{nachBhf='HamburgHbf'}}(\rho_{\text{TS3}}(\text{Teilstrecke}))$
  $\bowtie_{\text{TS3.ZugNr=TS4.ZugNr} \land \text{TS3.Wochentag=TS4.Wochentag} \land \text{TS3.Ankunftzeit<TS4.Abfahrtzeit}}$
  $(\rho_{\text{TS4}}(\text{Teilstrecke}))]$
  Das Ergebnis dieses Ausdruckes entspricht nicht wirklich einer Antwort auf die ursrprüngliche Frage, sondern ist eher als Baustein zum Abbilden des Umsteigevorganges zu verstehen. Es werden hier nur die Teilstrecke direkt vor und direkt nach dem Umsteigen ausgegeben. Es könnten aber mehrere Zwischenstationen zwischen München und den Umsteigebahnhof bzw. zwischen dem Umsteigebahnhof und Hamburg liegen, die so nicht ausgegeben werden.

- 2x umsteigen:
  analog zu 1x Umsteigen, nur dass noch ein weiterer Join nötig ist

- beliebig oft umsteigen:
  nicht möglich, da transitive Hüllen in relationaler Algebra nicht berechnet werden können

Aufgabe 3 e)

Geben Sie zusätzlich zu (d) noch den Gesamtfahrpreis für die Fahrt von München Hbf nach Hamburg Hbf aus.

Lösung

Nicht möglich, da kein Summierungsoperator in der relationalen Algebra existiert. Einzige Möglichkeit, die Preise zu summieren, wäre eine Gruppierung nach der ZugNr:
$\Gamma_{ZugNr;sum(Preis)}(\pi_{\text{ZugNr,vonBhf,Abfahrtzeit,nachBhf,Ankunftzeit}}$
$(\sigma_{\text{StartBhf='MünchenHbf'} \land \text{Zielbhf='HamburgHbf'}}$
$(\text{Teilstrecke} \bowtie \text{Verbindung})))$
Damit gehen aber alle Informationen über die Zwischenstationen verloren.