

Datenbanken
Wintersemester 11/12
 Prof. Dr. W. May

2. Übungsblatt: Algebra

Besprechung voraussichtlich am ab 29.11.2011

Aufgabe 1 (Division mit Basisoperationen) Beweisen Sie, daß die in der Vorlesung angegebene Darstellung der Division durch relationale Basisoperatoren als

$$r \div s = \pi[Z](r) - \pi[Z]((\pi[Z](r) \bowtie s) - r)$$

mit $r(X)$, $s(Y)$ und $Z = X \setminus Y$ äquivalent zu der gegebenen Definition

$$r \div s = \{\mu \in \text{Tup}(Z) \mid \{\mu\} \bowtie s \subseteq r\}$$

ist.

Veranschaulichen Sie sich Ihre Überlegungen anhand des Beispiels “Geben Sie die Namen derjenigen Organisationen an, die auf jedem Kontinent mindestens ein Mitglied haben”.

Aufgabe 2 (Äquivalenzen: Join, Division, Differenz) Seien $R(X), S(Y)$ Relations-Schemata. Zeigen oder widerlegen Sie:

(a) Sei $X \cap Y = \emptyset$.

$$(R \bowtie S) \div S \equiv R.$$

(b) Sei $X = Y$ und $Z \subseteq X$.

$$\pi[Z](R - S) \equiv \pi[Z]R - \pi[Z]S.$$

Aufgabe 3 (Korrektheit der Join-Algorithmen) Die Definition des relationalen Joins-Operators, \bowtie , auf den Folien ist nicht konstruktiv, sondern beschreibt “nur” *deklarativ*, aus welchen Tupeln das Ergebnis bestehen soll.

Beweisen Sie, dass die beiden folgenden Join-Algorithmen (a) und (b) korrekt sind, d.h. diese Menge von Tupeln liefern.

Gegeben seien Relationen $R(\bar{X})$ und $S(\bar{Y})$; berechnet werden soll $R \bowtie Y$.

a) Nested-Loop-Join:

```

let  $\bar{Z} := \bar{X} \cap \bar{Y}$ ;
let  $T := \emptyset$ ;
for  $r \in R$  do
  begin
    for  $s \in S$  do
      if  $r[\bar{Z}] = s[\bar{Z}]$  then
        begin

```

```

     $\mu :=$  ein neues Tupel so dass  $\mu(x) = r(x)$  für alle  $x \in \bar{X}$  und  $\mu(y) = s(y)$  für alle  $y \in \bar{Y} \setminus \bar{X}$ ;
     $T := T \cup \{\mu\}$ ;
end;
end;
return  $T$ ;

```

- b) Loop-Index-Join: Hierbei ist ein Baumindex über den Join-Attributen $\bar{Z} := \bar{X} \cap \bar{Y}$ von S in der Datenbank vorhanden.

Nehmen Sie an, dass \bar{Z} Schlüssel von S ist (dann erstellt ein Datenbanksystem diesen Index automatisch). Nehmen Sie weiter an, dass \bar{Z} nur ein Attribut enthält, welches numerisch ist, und stellen Sie sich als Index einen Binärbaum wie (hoffentlich) in Info I besprochen vor. In den jeweiligen Baumknoten ist nicht nur der Wert, sondern auch eine Referenz auf das Tupel, das diesen Wert hat, enthalten.

```

let  $\bar{Z} := \bar{X} \cap \bar{Y}$ ;
let  $B :=$  der Baumindex über  $S.\bar{Z}$ ;
let  $T := \emptyset$ ;
for  $r \in R$  do
  begin
    if Suche nach  $r[\bar{Z}]$  in  $B$  erfolgreich
      begin
         $s =$  das vom Baumknoten referenzierte Tupel in  $S$ ;
         $\mu :=$  ein neues Tupel so dass  $\mu(x) = r(x)$  für alle  $x \in \bar{X}$  und  $\mu(y) = s(y)$  für alle  $y \in \bar{Y} \setminus \bar{X}$ ;
         $T := T \cup \{\mu\}$ ;
      end;
    end;
  return  $T$ ;

```

- c) Wie ist die Komplexität der beiden Algorithmen?
d) Wenn \bar{Z} Fremdschlüssel in R ist, existiert in der Datenbank auch darauf ein Baumindex. Skizzieren Sie, wie dieser hilfreich sein kann, um $R \bowtie S$ noch effizienter zu berechnen.

Aufgabe 4 (Algebra: Minimale- und Maximale Anzahl von Tupeln) Die Relationen $R(\bar{X})$ und $S(\bar{Y})$ enthalten n bzw. m Tupel. Wie groß ist die maximale und minimale Anzahl von Tupeln, die das Ergebnis folgender Operationen (bei geeigneten \bar{X}, \bar{Y}) enthalten kann?

- $R \cup S$
- $R \bowtie S$
- $\sigma[C](R) \times S$, für eine Bedingung C
- $\pi[Y](R) - S$
- $R \div S$

Aufgabe 5 (Äquivalenz von Ausdrücken) Gegeben seien folgende Relationen:

- $R(A,B,C)$
- $S(A,E,F)$
- $T(A,H)$

Die Wertebereiche aller nicht namensgleichen Attribute seien voneinander verschieden. Gegeben sei nun folgender relationaler Ausdruck:

$$\pi[E, H](\sigma[B = 10]((R \bowtie T) \bowtie S))$$

Sind die folgenden Ausdrücke äquivalent zu obigem Ausdruck? Begründen Sie Ihre Antwort.

- $\pi[E, H](\sigma[B = 10](R) \bowtie (\pi[A, E](S) \bowtie T))$
- $\pi[E, H](\sigma[B = 10](\pi[B](R) \bowtie (\pi[A, E](S) \bowtie (\pi[A, H](T))))$
- $\pi[E, H](\pi[A, B](\sigma[B = 10](R)) \bowtie ((\pi[A](S) \bowtie T))$

Aufgabe 6 (Relationale Anfragen an Mondial: Bedingungen) Geben Sie Ausdrücke der relationalen Algebra für die folgenden Anfragen an die Mondial-Datenbank an:

- Die Namen aller Städte, die mehr als 1.000.000 Einwohner haben.
- Die Namen aller Städte, die mehr Einwohner als Neuseeland haben.
- Die Namen aller Städte, in denen mehr als 25% der Bevölkerung des jeweiligen Landes leben.

Für spätere Übungsblätter:

- Geben Sie dieselben Anfragen in SQL an.
- Geben Sie dieselben Anfragen im relationalen Kalkül an.

Aufgabe 7 (Relationale Anfragen an Mondial: Organisationen) Geben Sie Ausdrücke der relationalen Algebra für die folgenden Anfragen an die Mondial-Datenbank an:

- Die Namen aller Organisationen, die auf jedem Kontinent mindestens ein Mitgliedsland haben.
- Die Namen aller Organisationen, in denen alle Staaten mit mehr als 50.000.000 Einwohnern Mitglied (unabhängig von der Art der Mitgliedschaft) sind.

Für spätere Übungsblätter:

- Geben Sie dieselben Anfragen in SQL an.
- Geben Sie dieselben Anfragen im relationalen Kalkül an.

Aufgabe 8 (Relationale Anfragen an Mondial: Schweizer Sprachen) Geben Sie Ausdrücke der relationalen Algebra für die folgenden Anfragen an die Mondial-Datenbank an:

- Alle Landescodes von Ländern, in denen eine Sprache gesprochen wird, die auch in der Schweiz gesprochen wird.
- Alle Landescodes von Ländern, in denen ausschliesslich Sprachen gesprochen werden, die in der Schweiz nicht gesprochen werden.
- Alle Landescodes von Ländern, in denen nur Sprachen gesprochen werden, die auch in der Schweiz gesprochen werden.
- Alle Landescodes von Ländern, in denen alle Sprachen gesprochen werden, die in der Schweiz gesprochen werden.

Für spätere Übungsblätter:

- Geben Sie dieselben Anfragen in SQL an.
- Geben Sie dieselben Anfragen im relationalen Kalkül an.

Aufgabe 9 (Transitive Hülle) Gegeben sei eine Relation $R(A,B)$. Skizzieren Sie einen Algorithmus, der, bestehend aus Operationen der relationalen Algebra und einer while-Schleife, die transitive Hülle der Relation R berechnet.

Hinweis: Die transitive Hülle einer Relation R , bezeichnet als R^* ergibt sich wie folgt: betrachte z.B. eine Relation $R(\text{von}, \text{nach})$ von Flugverbindungen. R^2 ist dann die Menge aller Verbindungen, die über eine Zwischenlandung zustandekommen, etc; R^n sind also diejenigen, Verbindungen, die

sich aus n Teilverbindungen zusammensetzen. Die unendliche Vereinigung $R \cup R^2 \cup R^3 \cup \dots$ für $R \rightarrow \infty$ wird dann als R^* bezeichnet. In einer endlichen Datenbasis benötigt man nur endlich viele Schritte um diese zu berechnen. Ein anderes beliebtes Beispiel ist die aus $Kind(x, y)$ berechnete Vorfahren-Relation.

Aufgabe 10 (Relationale Vollständigkeit) Was versteht man unter *Relationaler Vollständigkeit*? Wenn eine Sprache relational vollständig ist, ist es dann möglich jede beliebige Anfrage in dieser Sprache zu formulieren?