Kapitel 5: Relationenmodell und algebraorientierte Anfragesprachen

5.1 Grundbegriffe des Relationenmodells

5.2 Relationenalgebra

5.3 Erweiterte Relationenalgebra

5.1 Grundbegriffe des Relationenmodells

DB: Relationen (Tables)

Bücher

ISBN	Autor	Titel	...	Kategorie
1111	Grass	Blechtrommel		Roman
2222	Gates	Robin Hood		Märchen
3333	King	Windows 2000		Horror

Verkäufe

ISBN	KNr	Datum	...
3333	999	1.4.00	...

Kunden

KNr	Name	Ort	PLZ	Strasse	...
901	Becker	Saarlouis	...		
902	Caesar	Rom			

Grundbegriffe des Relationenmodells

DB: Relationen
(Tables)

Bücher

ISBN	Autor	Titel	...	Kategorie
1111	Grass	Blechtrommel		Roman
2222	Gates	Robin Hood		Märchen
3333	King	Windows 2000		Horror

Ausprägung

$\subseteq$ Dom (ISBN) $\times$ Dom (Autor) $\times$ ... Dom (Kategorie)

Verkäufe

ISBN	KNr	Datum	...
3333	999	1.4.00	...

Kunden

KNr	Name	Ort	PLZ	Strasse	...
901	Becker	Saarlouis	...		
902	Caesar	Rom			

Schema

Attribut (Column)

Tupel (Record, Row)

Relationenmodell: Definitionen

Definition:
Gegeben sei eine Menge von Wertebereichen primitiver Datentypen {D1, ..., Dm}, die als ***"Domains"*** bezeichnet werden.
Eine ***Relation*** R ist ein Paar R = (s,v) mit
- einem *Schema* s = {A1, ..., An}, das aus einer Menge von ***Attributen (Attributnamen)*** besteht und für jedes Attribut Ai einen Domain dom(Ai) $\in$ {D1, ..., Dm} festlegt,
- und einer ***Ausprägung (Wert)*** v $\subseteq$ dom(A1) $\times$ dom(A2) $\times$ ... $\times$ dom(An).

Schema und Ausprägung von R werden mit *sch(r)* und *val(R)* bezeichnet.

Häufige Schreibweise für Relationenschemata:
R (A1, ..., An)
Bücher (ISBN, Autor, Titel, ...)
Kunden (KNr, Name, Stadt, ...)

Beispieldatenbank (1)

Kunden

KNr	Name	Stadt	Saldo	Rabatt
1	Lauer	Merzig	-1080.00	0.10
2	Schneider	Homburg	-800.00	0.20
3	Kirsch	Homburg	0.00	0.10
4	Schulz	Merzig	0.00	0.10
5	Becker	Dillingen	0.00	0.05
6	Meier	Saarlouis	-3800.00	0.05

Produkte

PNr	Bez	Gewicht	Preis	Lagerort	Vorrat
1	Papier	2.000	20.00	Homburg	10000
2	Platte	1.000	2500.00	Saarbrücken	400
3	Drucker	5.000	2000.00	Merzig	200
4	Bildschirm	5.000	3000.00	Merzig	80
5	Disketten	0.500	20.00	Homburg	5000
6	Maus	0.250	100.00	Homburg	200
7	Speicher	0.100	200.00	Saarbrücken	2000

Beispieldatenbank (2)

Bestellungen

BestNr	Monat	Tag	KNr	PNr	Menge	Summe	Status
1	7	16	1	1	100	1800.00	bezahlt
2	7	21	1	1	100	1800.00	bezahlt
3	9	30	1	2	4	9000.00	bezahlt
4	9	30	1	3	1	1800.00	bezahlt
5	9	30	1	4	10	27000.00	bezahlt
6	10	15	1	5	50	900.00	bezahlt
7	10	28	1	6	2	180.00	geliefert
8	11	2	1	7	5	900.00	neu
9	10	26	2	1	100	1600.00	bezahlt
10	11	2	2	5	50	800.00	neu
11	9	28	3	5	50	900.00	bezahlt
12	10	28	3	7	10	1800.00	bezahlt
13	4	15	4	1	50	900.00	bezahlt
14	5	31	6	1	200	3800.00	bezahlt
15	6	30	6	7	10	1900.00	geliefert
16	7	31	6	1	100	1900.00	geliefert

Einschränkungen der "1. Normalform"

Studenten

Name	**Fach**	**...**	**Systemkenntnisse**
Meier	Informatik		{Oracle, mySQL, PHP}
Schmidt	Informatik		{Java, Oracle}
Kunz	Informatik		{Oracle}
Müller	Mathematik		$\emptyset$

ist im („flachen") Relationenmodell nicht erlaubt !

Repräsentation in “1. Normalform”

Studenten

Name	Fach	...	Systemkenntnis
Meier	Informatik		Oracle
Meier	Informatik		mySQL
Meier	Informatik		PHP
Schmidt	Informatik		Java
Schmidt	Informatik		Oracle
Kunz	Informatik		Oracle
Müller	Mathematik		---

oder besser:

Studenten

Name	Fachbereich	...
Meier	Informatik	
Schmidt	Informatik	
Kunz	Informatik	
Müller	Mathematik	

Kenntnisse

Name	Systemkenntnis
Meier	Oracle
Meier	mySQL
Meier	PHP
Schmidt	Java
Schmidt	Oracle
Kunz	Oracle

Integritätsbedingungen des Relationenmodells (1)

Definitionen:

Eine Attributmenge $K \subseteq sch(R)$ einer Relation R heißt ***Schlüsselkandidat***, wenn zu jedem Zeitpunkt für je zwei Tupel $t1, t2 \in val(R)$ gelten muß:

$t1.K = t2.K \Rightarrow t1 = t2$

und wenn es keine echte Teilmenge von K gibt, die diese Eigenschaft hat. (Dabei bedeutet $t1.K = t2.K$: $\forall A \in K: t1.A = t2.A$.)

Ein Attribut einer Relation R, das in mind. einem Schlüsselkandidaten vorkommt, heißt ***Schlüsselattribut***.

Der ***Primärschlüssel*** einer Relation ist ein Schlüsselkandidat, der explizit ausgewählt wird.

Attributmenge $F \subseteq sch(S)$ einer Relation S ist ein ***Fremdschlüssel*** in S, wenn es eine Relation R gibt, in der F Primärschlüssel ist.

Ein Attribut A eines Tupels t hat einen ***Nullwert***, wenn der Wert t.A undefiniert oder unbekannt ist.

Integritätsbedingungen des Relationenmodells (2)

Primärschlüsselbedingung (Entity Integrity):

Für jede Relation muß ein Primärschlüssel festgelegt sein.
Der Primärschlüssel eines Tupels darf niemals den Nullwert annehmen.

Fremdschlüsselbedingung (Referential Integrity):

Für jeden Wert eines Fremdschlüssels in einer Relation R muß in den referenzierten Relationen jeweils ein Tupel mit demselben Wert als Primärschlüssel existieren, oder der Wert des Fremdschlüssels muß der Nullwert sein.

5.2 Relationenalgebra (RA)

Eine Operation der RA hat eine oder mehrere Relationen als Operanden und liefert eine Relation als Ergebnis.
(Abgeschlossenheit der Algebra)

Mengenoperationen:

Für zwei Relationen R, S mit sch(R) = sch(S) sind die üblichen Mengenoperationen definiert:

- **Vereinigung (Union) $R \cup S$:**
  $sch(R \cup S) = sch(R)$
  $val(R \cup S) = \{t \mid t \in val(R) \vee t \in val(S)\}$
- **Durchschnitt (Intersection) $R \cap S$:**
  $sch(R \cap S) = sch(R)$
  $val(R \cap S) = \{t \mid t \in val(R) \wedge t \in val(S)\}$
- **Differenz (Difference) $R - S$:**
  $sch(R - S) = sch(R)$
  $val(R - S) = \{t \mid t \in val(R) \wedge t \notin val(S)\}$

Selektion und Projektion

Selektion σ (Filterung, Auswahl von Zeilen einer Tabelle):
Das Resultat einer Selektion σ[F](R) auf einer Relation R ist:

$$sch(\sigma[F](R)) = sch(R)$$

$$val(\sigma[F](R)) = \{t \mid t \in R \wedge F(t)\}$$

Die Menge der möglichen Filterformeln F ist:

1) Für Attr. A, B von R mit dom(A)=dom(B), Konstanten c ∈ dom(A) und Vergleichsoperationen θ ∈ {=, ≠, <, >, ≤, ≥} sind A θ B und A θ c zulässige Filterbedingungen.
2) Falls F1 und F2 zulässige Filterbedingungen sind, dann sind auch F1 ∧ F2, F1 ∨ F2, ¬ F1 und (F1) zulässig.
3) Nur die von 1) und 2) erzeugten Filterbedingungen sind zulässig.

Projektion π (Auswahl von Spalten einer Tabelle):
Sei A ⊆ sch(R). Das Resultat einer Projektion π[A](R) ist:

$$sch(\pi[A](R)) = A$$

$$val(\pi[A](R)) = \{t \mid \exists r \in val(R): t.A = r.A\}$$

Beispiele Selektion

1) Finde alle Homburger Kunden.

$\rightarrow \sigma$[Stadt='Homburg'] (Kunden)

KNr	Name	Stadt	Saldo	Rabatt
2	Schneider	Homburg	-800.00	0.20
3	Kirsch	Homburg	0.00	0.10

2) Finde alle Homburger Kunden, die einen Rabatt von mind. 15 % haben.

$\rightarrow \sigma$[Stadt='Homburg' $\wedge$ Rabatt >= 0.15] (Kunden)

KNr	Name	Stadt	Saldo	Rabatt
2	Schneider	Homburg	-800.00	0.20

Beispiele Projektion

1) Gib alle Produktbezeichnungen aus.

$\rightarrow \pi[\text{Bez}]$ (Produkte)

Bez
Papier
Platte
Drucker
Bildschirm
Disketten
Maus
Speicher

2) Gib alle Lagerorte von Produkten aus.

$\rightarrow \pi[\text{Lagerort}]$ (Produkte)

Lagerort
Homburg
Saarbrücken
Merzig

(Natural) Join $|\times|$ und Zuweisung

(Natural) Join $|\times|$ (Natürlicher Verbund von Relationen über gleiche Attributnamen und Attributwerte):

Das Resultat von R $|\times|$ S mit A=sch(R) und B=sch(S) ist:

$$sch(R \,|\times|\, S) = sch(R) \cup sch(S)$$

$$val(R \,|\times|\, S) = \{t \mid \exists r \in val(R)\ \exists s \in val(S): t.A = r.A \wedge t.B = s.B\}$$

Zuweisung:

Seien R, S Relationen mit sch(R)={A1, ..., An} und sch(S)={B1, ..., Bn}, so daß für alle i gilt: dom(Ai)=dom(Bi).

Die Zuweisung *R := S* bedeutet: val(R) = val(S).

Ausführlicher schreibt man auch R(A1, ..., An) := S(B1, ..., Bn).

Für Ausdrücke E1, ..., En, die über B1, ..., Bn und k-stelligen Operatoren ψ: dom(B_{i1}) x … x dom(B_{ik}) $\rightarrow$ D mit skalarem Resultat im Bereich W bedeutet ***R(A1, ..., An) := S(E1, ..., En):***

$$val(R) = \{t \mid \text{es gibt } s \in val(S) \text{ und } t.Ai = Ei(s) \text{ für alle } i\},$$

wobei der Wertebereich von Ei gleich dom(Ai) sein muß.

Beispiele für Join

1) Alle Bestellungen zusammen mit den dazugehörigen Produktdaten.
   → Bestellungen |×| Produkt

2) R |×| S

R

R1	**J**
a	1
b	2
c	2
d	3

S

S1	**J**
w	2
x	2
y	3
z	4

→

R1	**J**	**S1**
b	2	w
b	2	x
c	2	w
c	2	x
d	3	y

Kartesisches Produkt und Division

Kartesisches Produkt × :

Seien R, S Relationen mit Schemata A=sch(R) und B=sch(S).
Sei A' ein Schema, bei dem alle Ai, die auch in B vorkommen, unbenannt sind in R.Ai, und sei B' ein analoges Schema mit Attributnamen S.Ai.
Das Resultat von R × S ist:

$$sch(R \times S) = A' \cup B'$$

$$val(R \times S) = \{t \mid \exists r \in val(R) \; \exists s \in val(S): \; t.A' = r.A \text{ und } t.B' = s.B\}$$

Division ÷ :

Seien R, S Relationen mit A=sch(R) und B=sch(S), so daß B ⊂ A.
Das Resultat der Division R ÷ S ist:

$$sch(R \div S) = A - B = Q$$

$$val(R \div S) = \{t \mid \forall s \in val(S) \; \exists r \in val(R): r.B = s.B \wedge t = r.Q\}$$

Intuitive Bedeutung: ein Tupel t ist in R ÷ S genau dann, wenn für alle S-Tupel s ein Tupel <t,s> in R enthalten ist.

Satz: Für R(A,B,C), T(A,B), S(C) mit R=T × S gilt: R ÷ S = T.

Beispiel Division

Kundennummern derjenigen Kunden, die alle überhaupt lieferbaren Produkte irgendwann bestellt haben

$\rightarrow \pi$[KNr, PNr] (Bestellungen) $\div$ π[PNr] (Produkte)

KNr	PNr
1	1
1	2
1	3
1	4
1	5
1	6
1	7
2	1
2	5
3	5
3	7
4	1
6	1
6	7

PNr
1
2
3
4
5
6
7

Rückführung Division auf andere Operationen

Satz:

Seien R und S Relationen mit A=sch(R) und B=sch(S) mit $A \supset B$, und sei Q=sch(R) - sch(S). Es gilt (bis auf Umbenennungen von Attributen):

$$R \div S = \pi[Q](R) - \pi[Q]\left(\left(\pi[Q](R) \times S \right) - R \right)$$

Beweisskizze (kanonisches Beispiel):

$R := \pi[KNr, PNr]$ (Bestellungen)

$S := \pi[PNr]$ (Produkte)

$T1 := \pi[KNr](R) \times S$	alle überhaupt möglichen Bestellungen
$T2 := T1 - R$	potentiell mögliche, aber nicht erfolgte Bestellungen
$T3 := \pi[KNr]\ (T2)$	Kunden, die nicht alle Prod. bestellt haben
$T4 := \pi[KNr]\ (R) - T3$	Kunden, die alle Produkte bestellt haben

θ-Join

θ-Join:

Seien R, S Relationen mit $sch(R) \cap sch(S) = \emptyset$. Seien $A \subseteq sch(R)$ und $B \subseteq sch(S)$, und sei θ eine der Operationen =, ≠, <, >, ≤, ≥.

Das Resultat des θ-Joins R |×|[A θ B] S ist:

$$sch(R\ |\times|[A\ \theta\ B]\ S) = sch(R \times S) = sch(R) \cup sch(S)$$

$$val(R\ |\times|[A\ \theta\ B]\ S) = val(\sigma[A\ \theta\ B](R \times S))$$

<u>**Beispiele:**</u>

1) alle Kunden, die an einem Lagerort wohnen.

   → π[KNr, Name, Stadt, ...] (Kunden |×|[Stadt=Lagerort] Produkte)

2) alle bisherigen Bestellungen, die den momentanen Vorrat erschöpfen würden.

   → B(BestNr, KNr,B.PNr, ...) := Bestellungen(BestNr, KNr, PNr, ...)

   π[BestNr, ...] (B |×|[B.PNr=PNr ∧ Menge >= Vorrat] Produkte)

3) alle Paare von Kunden, die in derselben Stadt wohnen.

   → K1(K1.KNr, K1.Name, ...) := Kunden(KNr, Name, ...)

   K2(K2.KNr, K2.Name, ...) := Kunden(KNr, Name, ...)

   π[K1.KNr, K2.KNr]

   (K1 |×|[K1.Stadt=K2.Stadt ∧ K1.KNr < K2.KNr] K2)

Outer Joins

Seien R, S Relationen mit A=sch(R), B=sch(S) und J=sch(R) $\cap$ sch(S).
Bezeichne ferner ω den Nullwert.
Das Resultat des **Outer Joins R |*| S** ist:

$$\text{sch}\,(R\ |{*}|\ S) = \text{sch}\,(R) \cup \text{sch}\,(S)$$

$$\text{val}\,(R\ |{*}|\ S) = \text{val}\,(R\ |{\times}|\ S) \cup$$

$$\{t | \exists\ r \in \text{val}\,(R) : t.A{=}r.A \wedge \neg\,(\exists\ s \in \text{val}\,(S) : t.J{=}s.J) \wedge t.(B{-}J){=}\omega\ \} \cup$$

$$\{t | \exists\ s \in \text{val}\,(S) : t.B{=}s.B \wedge \neg\,(\exists\ r \in \text{val}\,(R) : t.J{=}r.J) \wedge t.(A{-}J){=}\omega\ \}.$$

Left Outer Join:

$$\text{sch}\,(R\ ||{*}|\ S) = \text{sch}\,(R) \cup \text{sch}\,(S)$$

$$\text{val}\,(R\ ||{*}|\ S) = \text{val}\,(R\ |{\times}|\ S) \cup$$

$$\{t\ |\ \exists\ r \in \text{val}\,(R) : t.A{=}r.A \wedge \neg\,(\exists\ s \in \text{val}\,(S) : t.J{=}s.J) \wedge t.(B{-}J){=}\omega\ \}$$

Right Outer Join:

$$\text{sch}\,(R\ |{*}||\ S) = \text{sch}\,(R) \cup \text{sch}\,(S)$$

$$\text{val}\,(R\ |{*}||\ S) = \text{val}\,(R\ |{\times}|\ S) \cup$$

$$\{t\ |\ \exists\ s \in \text{val}\,(S) : t.B{=}s.B \wedge \neg\,(\exists\ r \in \text{val}\,(R) : t.J{=}r.J) \wedge t.(A{-}J){=}\omega\ \}.$$

Beispiel Outer Join

Gib alle Kunden mit 5 % Rabatt zusammen mit ihren Bestellungen aus, und zwar auch, wenn für einen Kunden gar keine Bestellungen vorliegen.

σ [Rabatt = 0.05] (Kunden |*| Bestellungen)

KNr	**Name**	**Stadt**	**Saldo**	**Rabatt**	**BestNr**	**...**
5	Becker	Dillingen	0.00	0.05	ω	...
6	Meier	Saarlouis	-3800.00	0.05	14	...
6	Meier	Saarlouis	-3800.00	0.05	15	...
6	Meier	Saarlouis	-3800.00	0.05	16	...

Äquivalenzregeln ("Rechenregeln") der RA

Kommutativitätsregeln:

1) $\pi[R1]\ \sigma[P]\ (R) = \sigma[P]\ \pi[R1]\ (R)$ falls P nur R1-Attribute enthält

2) $R\ |x|\ S = S\ |x|\ R$

Assoziativitätsregeln:

3) $R\ |x|\ (\ S\ |x|\ T\ ) = (\ R\ |x|\ S\ )\ |x|\ T$

Idempotenzregeln:

4) $\pi[R1]\ (\ \pi[R2]\ (R)\ ) = \pi[R1]\ (R)$ falls $R1 \subseteq R2$

5) $\sigma[P1]\ (\ \sigma[P2]\ (R)\ ) = \sigma[P1 \wedge P2]\ (R)$

Distributivitätsregeln:

6) $\pi[R1]\ (\ R \cup S\ ) = \pi[R1](R) \cup \pi[R1](S)$

7) $\sigma[P]\ (\ R \cup S\ ) = \sigma[P](R) \cup \sigma[P](S)$

8) $\sigma[P]\ (\ R\ |x|\ S\ ) = \sigma[P](R)\ |x|\ S$ falls P nur R-Attribute enthält

9) $\pi[R1,S1](R\ |x|\ S) = \pi[R1](R)\ |x|\ \pi[S1](S)$ falls Joinattribute $\subseteq R1 \cup S1$

10) $R\ |x|\ (\ S \cup T\ ) = (R\ |x|\ S) \cup (R\ |x|\ T)$

Invertierungsregeln:

11) $\pi[sch(R)]\ (R\ ||{*}\,|\ S)=R$

Äquivalenzregeln ("Rechenregeln") der RA

Kommutativitätsregeln:

1) $\pi[R1]\ \sigma[P]\ (R) = \sigma[P]\ \pi[R1]\ (R)$ falls P nur R1-Attribute enthält

2) $R\ |x|\ S = S\ |x|\ R$

Assoziativitätsregeln:

3) $R\ |x|\ (\ S\ |x|\ T\ ) = (\ R\ |x|\ S\ )\ |x|\ T$

Idempotenzregeln:

4) $\pi[R1]\ (\ \pi[R2]\ (R)\ ) = \pi[R1]\ (R)$ falls $R1 \subseteq R2$

5) $\sigma[P1]\ (\ \sigma[P2]\ (R)\ ) = \sigma[P1 \wedge P2]\ (R)$

Distributivitätsregeln:

6) $\pi[R1]\ (\ R \cup S\ ) = \pi[R1](R) \cup \pi[R1](S)$

7) $\sigma[P]\ (\ R \cup S\ ) = \sigma[P](R) \cup \sigma[P](S)$

8) $\sigma[P]\ (\ R\ |x|\ S\ ) = \sigma[P](R)\ |x|\ S$ falls P nur R-Attribute enthält

9) $\pi[R1,S1](R\ |x|\ S) = \pi[R1](R)\ |x|\ \pi[S1](S)$ falls Joinattribute $\subseteq R1 \cup S1$

10) $R\ |x|\ (\ S \cup T\ ) = (R\ |x|\ S) \cup (R\ |x|\ T)$

Invertierungsregeln:

11) $\pi[sch(R)]\ (R\ ||^{*}\mid S) = R$

Ausdrucksmächtigkeit der RA

Die Menge der **relationenalgebraischen Ausdrücke** über einer Menge vo
Relationen R1, ..., Rn ist wie folgt definiert:

(i) R1, ..., Rn sind Ausdrücke.

(ii) Wenn R, S, T, Q Ausdrücke sind, F eine Filterformel über sch(R) ist,
$A \subseteq$ sch(R), sch(S)=sch(T und sch(R) $\supset$ sch(Q) gilt, dann sind
σ[F](R), π[A](R), R |x| S, R x S, R |*| S, S $\cap$ T, S $\cup$ T, S - T, R $\div$ Q
auch Ausdrücke.

(iii) Nur die von (i) und (ii) erzeugten Ausdrücke sind RA-Ausdrücke.

Satz:

$\times$, π, σ, $\cup$ und - bilden eine minimale Menge von Operationen,
mit denen sich alle Operationen der RA ausdrücken lassen.

Eine Anfragesprache heißt *relational vollständig*, wenn sich damit alle
Anfragen der (minimalen) Relationenalgebra ausdrücken lassen.

5.3 Erweiterte Relationenalgebra

Definition:

Eine ***Multimenge*** (engl. multiset, bag) M über einer Grundmenge G ist eine Abbildung $M: G \rightarrow N_0$.

M(x) wird als die Häufigkeit von x in M bezeichnet.

Definition:

Eine ***Multirelation*** R ist ein Paar R = (s,v) mit einem Schema s = {A1, ..., An} und einer Ausprägung v, die eine Multimenge über dom(A1) × ... × dom(An) ist.

Beispiel:

Name	Stadt	Rabatt
Lauer	Merzig	0.10
Schneider	Homburg	0.20
Schneider	Homburg	0.20
Schulz	Merzig	0.10
Schulz	Merzig	0.05
Meier	Saarlouis	0.05

Die Funktion χ konvertiert eine Multirelation R in eine Relation $\chi(R)$ mit: $val(\chi(R)) = \{t \mid R(t)>0\}$.

Operationen auf Multimengen und -relationen

Für Multirelationen R, S mit sch(R) = sch(S) sind:

Vereinigung R $\cup_+$ S:

$sch(R \cup_+ S) = sch(R)$, $val(R \cup_+ S) = t \mapsto R(t) + S(t)$

Durchschnitt R $\cap_+$ S:

$sch(R \cap_+ S) = sch(R)$, $val(R \cap_+ S) = t \mapsto \min(R(t), S(t))$

Differenz R $-_+$ S:

$sch(R -_+ S) = sch(R)$, $val(R -_+ S) = t \mapsto R(t)-S(t)$ falls $R(t) \geq S(t)$, 0 sonst

Das Resultat der **Selektion σ_+ [F](R)** auf einer Multirelation R ist:

$sch(\sigma_+ [F](R)) = sch(R)$,

$val(\sigma_+ [F](R)) = t \mapsto R(t)$ falls $R(t) \wedge F(t)$, 0 sonst

Das Resultat der **Projektion π_+ [A](R)** auf einer Multirelation R
mit $A \subseteq sch(R)$ ist:

$sch(\pi_+ [A](R)) = A$,

$val(\pi_+ [A](R)) = t \mapsto \Sigma \{R(r) \mid r \in val(R) \text{und } r.A = t.A\}$

Aggregation und Gruppierung

Aggregation α_+ (Zusammenfassen von Zeilen einer Tabelle):
Sei $A \in sch(R)$ und sei f eine Funktion, die Multimengen über dom(A) in einen Wertebereich W abbildet (z.B. max, min, sum, count, median).
Das Resultat der Aggregation **α_+ [A,f](R)** auf der Multirelation R ist:

$sch(\alpha_+ [A,f](R)) = A'$ mit $dom(A')=W$

$val(\alpha_+ [A,f](R)) = f(\pi_+[A](R))$

Gruppierung γ_+ (Zusammenfassen von Äquivalenzklassen von Zeilen):
Sei $X \subseteq sch(R)$, $A \in sch(R)$, und sei f eine Funktion, die Multimengen über dom(A) in einen Wertebereich W abbildet.
Das Resultat einer Gruppierung **γ_+ [X,A,f](R)** auf der Multirelation R ist:

$sch(\gamma_+ [X,A,f](R)) = X \cup \{A'\}$ mit $dom(A')=W$

$val(\gamma_+ [X,A,f](R)) = \{ t \mid$ es gibt eine Äquivalenzklasse G von $\pi_+[X](R)$ unter der Wertegleichheit und t.X ist der Wert der Tupel in G und $t.A' = f(\pi_+[A](G)) \}$

Beispielanfragen auf Multirelationen (1)

PNr	Bez	Gewicht	Preis	Lagerort	Vorrat
1	Papier	2.000	20.00	Homburg	10000
2	Platte	1.000	2500.00	Saarbrücken	400
3	Drucker	5.000	2000.00	Merzig	200
4	Bildschirm	5.000	3000.00	Merzig	80
5	Disketten	0.500	20.00	Homburg	5000
6	Maus	0.250	100.00	Homburg	200
7	Speicher	0.100	200.00	Saarbrücken	2000

1) Produkte unter 50 DM sowie deren Lagerorte und Preise:

$\sigma_+[\text{Preis}<50.00](\pi_+[\text{Lagerort, Preis}](\text{Produkte}))$

Lagerort	Preis
Homburg	20.00
Homburg	20.00

2) Gesamtstückzahl (aller Produkte) über alle Lager:

$\text{Resultat(Gesamtvorrat)} := \alpha_+[\text{Vorrat,sum}](\text{Produkte})$

Gesamtvorrat
17880

Beispielanfragen auf Multirelationen (2)

3) Bestimme für jedes Lager die Gesamt-
stückzahl aller dort gelagerten Produkte:

Resultat (Lagerort, Gesamtvorrat) :=
γ+ [{Lagerort}, Vorrat, sum](Produkte)

Lagerort	**Gesamtvorrat**
Homburg	15200
Saarbrücken	2400
Merzig	280

4) Bestimme für jedes Lager die
Gesamtkapitalbindung
(Stückzahl * Preis)
aller dort gelagerten Produkte:

Lagerort	**Kapitalbindung**
Homburg	320 000
Saarbrücken	1 400 000
Merzig	640 000

P1 := π+ [Lagerort, Preis, Vorrat] (Produkte)
P2 (Lagerort, Wert) := P1(Lagerort, Preis*Vorrat)
Resultat (Lagerort, Kapitalbindung) :=
γ+ [{Lagerort}, Wert, sum](P2)

Weitere Erweiterungen der RA

Transitive Hülle R^+ einer binären Relation R:

Sei R(A, B) eine binäre Relation mit dom(A)=dom(B).

Das Resultat der transitiven Hülle R^+ ist:

$sch(R^+) = sch(R)$

$val(R^+)$ ist die kleinste Menge, für die gilt:

(i) für jedes $r \in R$ gilt $r \in R^+$ und

(ii) für $t \in R^+$ und $r \in R$ mit $t.B = r.A$ gilt $(t.A, r.B) \in R^+$

Beispiel: Flugverbindungen := (π[Abflugort, Zielort] (Flüge))$^+$

Flüge

FlugNr	**Abflugort**	**Zielort**	**...**
LH58	Frankfurt	Chicago	
AA371	Chicago	Phoenix	
DA77	Phoenix	Yuma	
AA70	Frankfurt	Dallas	
AA351	Dallas	Phoenix	
UA111	Chicago	Dallas	

Flugverbindungen

Abflugort	**Zielort**
Frankfurt	Chicago
Frankfurt	Dallas
Frankfurt	Phoenix
Frankfurt	Yuma
...	...