3.3 Das Relationenmodell

- von Codd, 1970
- anschaulich:

Attribute

STUDENT

Relationen-name

Matrnr	Vorname	Nachname	Fach
603415	Alfred	Albers	Chemie
606237	Willi	Winzig	Physik
617543	Herta	Hurtig	Chemie

Relationenschema

Tupel

Relation

Attributwert

mehrere Relationenschemata ergeben *Datenbank-schema*
mehrere Relationen ergeben *Datenbank*

Das Relationenmodell

formal:

- Eine endliche Menge *U* heißt *Universum*, ein A $\in$ *U* heißt *Attribut*. Sei *D* = $\{D_1,..,D_m\}$ eine Menge endlicher, nicht-leerer Mengen D_i; jedes D_i wird *Wertebereich* (*Domäne*) genannt.

  Es existiert total definierte Funktion *dom*: *U* $\rightarrow$ *D*, *dom*(A) heißt *Wertebereich von A*, w $\in$ *dom*(A) *Attributwert* für A.

- R $\subseteq$ *U* heißt *Relationenschema*. Eine *Relation* r über R=$\{A_1,..,A_n\}$ (kurz: r(R)) ist eine endliche Menge von Abbildungen t: R $\rightarrow \bigcup_{i=1}^{m} D_i$

  t werden *Tupel* genannt, t(A) ist Restriktion von t auf A $\in$ R mit t(A) := t $|_A \in$ *dom*(A). Für X $\subseteq$ R heißt t(X) =: t $|_X$ *X-Wert von t*.

- Die Menge aller Relationenschemata S := $\{R_1,..,R_p\}$ heißt *Datenbankschema*. Eine *Datenbank* d über S = $\{R_1,..,R_p\}$ (kurz: d(S)) ist eine Menge von Relatonen d = $\{r_1,..,r_p\}$, wobei $r_i(R_i)$ für alle i $\in$ $\{1,..,p\}$ gilt. Ein r $\in$ d heißt *Basisrelation*.

Das Relationenmodell

Beispiel:

S = { $\mathtt{ARTIKEL}$, $\mathtt{ABTEILUNG}$, $\mathtt{FÜHRT}$, $\mathtt{AUFTRAG}$, $\mathtt{ENTHÄLT}$ }
$\mathtt{ARTIKEL}$ = { Artnr, EAN, Artname}
$\mathtt{ABTEILUNG}$ = { Abtlgsnr, Abtlgsname}
$\mathtt{FÜHRT}$ = { Abtlgsnr, Artnr}
$\mathtt{AUFTRAG}$ = { Auftragsnr, Datum}
$\mathtt{ENTHÄLT}$ = { Auftrnr, Artnr, Menge}

dom	Attribut	Domäne
	Artnr	INTEGER
	Artname	STRING(30)
	Abtlgsnr	INTEGER
	Abtlgsname	STRING(40)
	Auftragsnr	INTEGER
	Datum	DATE
	Menge	INTEGER
	EAN	STRING(15)

t_1(Artnr) = 1046,
t_1(EAN) = '4-902588-306625',
t_1(Artname) = '10 Disketten maxell'

t_1 = < 1046,'4-902588-306625','10 Disketten maxell'>
r(ARTIKEL) = {t_1,t_2,t_3,...}

Das Relationenmodell

Integritätsbedingungen:

- Eine *identifizierende Attributmenge* für eine Relation r(R) ist eine Menge $K := \{B_1, .., B_k\} \subseteq R$ mit $\forall\, t_1, t_2 \in r\ [\ t_1 \neq t_2 \Rightarrow \exists\ B \in K\!: t_1(B) \neq t_2(B)]$
  Ein *Schlüssel* ist eine bzgl. $\subseteq$ minimale, identifizierende Attributmenge. Jedes Attribut eines Schlüssels heißt *Primattribut*. Ein *Primärschlüssel* ist (genau) ein ausgezeichneter Schlüssel einer Relation r(R).

  Beispiel:

  { Artnr, Artname, EAN} (automatisch) identifizierende Attributmenge von `ARTIKEL`
  { Artnr} Schlüssel für `ARTIKEL`
  { EAN} Schlüssel für `ARTIKEL`
  etwa: { Artnr} Primärschlüssel für `ARTIKEL`

Das Relationenmodell

- Ein *Fremdschlüssel* ist eine Attributmenge einer Relation, die in einer anderen Relation Primärschlüssel ist.
  Eine Relation r *genügt der Fremdschlüsslbedingung* X, wenn X Fremdschlüssel ist und r $|_X \subseteq$ r' $|_X$, wobei X für r' Primärschlüssel ist.

  Beispiel:

  { Artnr} ist Fremdschlüssel in `FÜHRT` und in `ENTHÄLT`, da Artnr Primärschlüssel in `ARTIKEL` ist.

- Übersicht:

Begriff	Bedeutung
Attribut	Spalte einer Tabelle
Wertebereich/ Domäne	mögl. Werte eines Attributs
Attributwert	Element eines Wertebereichs
Relationenschema	Menge von Attributen (Tabellenkopf)
Relation	Menge von Zeilen einer Tabelle (Tabellenrumpf)
Tupel	Zeile einer Tabelle
DB-Schema	Menge von Relationenschemata
Datenbank	Menge von Basisrelationen
Schlüssel	minimale Menge von Attributen, deren Werte ein Tupel eindeutig identifizieren
Primärschlüssel	beim DB-Entwurf ausgezeichneter Schlüssel
Fremdschlüssel	Attributmenge, die in anderer Relation Primärschlüssel ist
Fremdschlüssel-bedingung	die Attributwerte eines Fremdschlüssels müssen in anderer Relation als Werte des Primärschlüssels auftauchen

Das Relationenmodell

Operationen:

- im Modell implizit enthaltene Operationen (= *generische* Operationen)

- Relationenalgebra umfaßt Selektion (σ), Projektion (π), natürlicher Verbund ($\bowtie$), Mengenoperationen ($\cup, \cap, -$), Umbenennung (β)

- Selektion wählt einige Tupel aus Relation aus, Tupel müssen Selektionsbedingung erfüllen

  Bsp.:
  $\sigma_{Abtlgsnr=5}(\ \mathrm{r}(\texttt{FÜHRT}))$ oder $\sigma_{Artnr=EAN}(\ \mathrm{r}(\texttt{ARTIKEL}))$

- Projektion wählt bestimmte Spalten einer Tabelle aus

  Bsp.: $\pi_{Artnr,Menge}(\mathrm{r}(\texttt{ENTHÄLT}))$

- natürlicher Verbund (Join) ist Verknüpfung zweier Relationen über allen gemeinsamen Attributen, jeweils Tupel mit gleichen Attributwerten werden zu neuen Tupeln verbunden

- Bsp.:
  a) r($\texttt{ARTIKEL}$) $\bowtie$ r($\texttt{ENTHÄLT}$) liefert

Artnr	EAN	Artname	Auftragsnr	Menge

  b) $\pi_{Artname,Menge}$(r($\texttt{ARTIKEL}$) $\bowtie$ r($\texttt{ENTHÄLT}$)) könnte dann folgendes Ergebnis liefern:

Artname	Menge
10 Disketten maxell	23
Hemdbluse	12
Heftklammern 24/6	45
Hemdbluse	4
...	...

  -> Operationen der Relationenalgebra können bel. kombiniert werden

- Mengenoperationen bilden die Vereinigung, den Durchschnitt bzw. die Differenz von zwei Relationen, die das gleiche Relationenschema besitzen müssen

- Umbenennung benennt Attribute um; ist sinnvoll, wenn z.B. Attribut beim Join nicht zur Verknüpfung genommen werden soll, obwohl es mit Attribut(name) in anderer Relation übereinstimmt

  Bsp.: $\beta_{EAN \leftarrow Artnr}$ (r($\texttt{ENTHÄLT}$))

Das Relationenmodell

Relationenalgebra formal:

Ω sei Menge an Operationen, Datenbank d ist Erzeugermenge, $[d]_\Omega$ ist Erzeugnis unter den Operationen aus Ω, $([d]_\Omega,\Omega)$ die Relationenalgebra

- Projektion $\pi_X(r) := \{t(X) \mid t \in r\}$

- Konstantenselektion mit $\theta \in \{<,\leq,=,>,\geq,\neq\}$, x X-Wert, $X \subseteq R$
  $\sigma_{X\theta x}(r) := \{t \mid t \in r \wedge t(X)\theta x\}$

- Attributselektion mit $X,Y \in R$
  $\sigma_{X\theta Y}(r) := \{t \mid t \in r \wedge t(X)\ \theta\ t(Y)\}$

- Selektion ist entweder Konstanten- oder Attributselektion

- natürlicher Verbund von $r_1(R_1),r_2(R_2)$
  $r_1 \bowtie r_2 := \{t \mid t \in (R_1 \cup R_2) \wedge [\ \forall i \in \{1,2\}\ \exists\ t_i \in r_i{:}\ t_i = t(R_i)]\}$

  wenn $R_1=R_2$: mengentheoretischer Durchschnitt
  wenn $R_1 \cap R_2 = \{\}$: kartesisches Produkt

- Mengenoperationen
  $r_1 \cup r_2 := \{t \mid t \in r_1 \vee t \in r_2\}$
  $r_1 \cap r_2 := \{t \mid t \in r_1 \wedge t \in r_2\}$
  $r_1 - r_2 := \{t \mid t \in r_1 \vee t \notin r_2\}$

- Umbenennung
  $\beta_{B\leftarrow A}(r) := \{t' \mid \exists\ t \in r{:}\ t'(R\text{-}\{A\}) = t(R\text{-}\{A\}) \wedge t'(B) = t(A)\}$ mit $B \notin R\text{-}\{A\}$, $R'=(R\text{-}\{A\})\cup\{B\}$, $dom(A)=dom(B)$