

Kapitel 7

Dr. Jérôme Kunegis

Logische Kalküle

Lernziele

- Grundideen des Domain-Relationenkalküls (DRK) und des Tupel-Relationenkalküls (TRK)
- Relationale Datenbank als Formelmenge
- Sicherheit der Anfragen
- Mächtigkeit relationaler Anfragesprachen

Domain-Relationenkalkül (DRK)

Anfragen sind prädikatenlogische Mengenspezifikationen der Form

$$\{[x_1, \dots, x_n] \mid F(x_1, \dots, x_n)\}$$

wobei F ein prädikatenlogischer Ausdruck über einer Menge von **Domain-Variablen** ist mit $x_1 \dots x_n$ als einzigen freien Variablen.

Domain-Relationenkalkül (DRK)

Zulässige prädikatenlogische Ausdrücke für F sind:

- 1) Für Variablen x_1, \dots, x_n und eine Relation R mit n Attributen ist $\langle x_1, \dots, x_n \rangle \in R$ (auch $R(x_1, \dots, x_n)$) ein zulässiger Ausdruck.
- 2) Für Variablen x, y , Konstanten c und Vergleichsoperationen $\theta \in \{=, \neq, <, >, \leq, \geq\}$ sind $x \theta y$ und $x \theta c$ zulässige Ausdrücke.
- 3) Falls F_1 und F_2 zulässige Ausdrücke sind, dann sind auch $F_1 \wedge F_2, F_1 \vee F_2, \neg F_1$ zulässig.
- 4) Falls F ein zulässiger Ausdruck mit einer freien Variablen x ist, dann sind auch $\exists x F(x)$ und $\forall x F(x)$ zulässige Ausdrücke.
- 5) Falls F ein zulässiger Ausdruck ist, dann ist auch (F) zulässig.
- 6) Nur die durch 1) bis 5) erzeugten Ausdrücke sind zulässig.

Datenbank als Formelmenge

Eine **Datenbank** über einem Schema mit Relationen

$$R_1(A_{11}, \dots, A_{1n_1}), \dots, R_m(A_{m1}, \dots, A_{mn_m})$$

ist eine Formelmenge H , die für jede Relation R_i ein Prädikatsymbol R_i verwendet und für jedes Tupel $[a_1, \dots, a_{n_i}] \in \text{val}(R_i)$ eine atomare Formel $R_i(a_1, \dots, a_{n_i})$ enthält, bzw. ein Modell von H .

Das Ergebnis einer **Anfrage** $\{[x_1, \dots, x_n] \mid F(x_1, \dots, x_n)\}$ über der Datenbank H ist:

$$\{[a_1, a_2, \dots, a_n] \mid a_1, a_2, \dots, a_n \in U \wedge H \models F[x_1/a_1, x_2/a_2, \dots, x_n/a_n]\}$$

Sonderfall:

wenn F keine freien Variablen enthält,

ist das Ergebnis 1 wenn $H \models F$ und 0 sonst.

DRK: Beispiel

Anfrage: MatrNr und Namen der Prüflinge von Curie

$$\{ [m, n] \mid$$
$$\quad \exists s ([m, n, s] \in \text{Studenten} \wedge$$
$$\quad \exists v, p, g ([m, v, p, g] \in \text{prüfen} \wedge$$
$$\quad \exists a, r, b ([p, a, r, b] \in \text{Professoren} \wedge$$
$$\quad \quad a = \text{'Curie'})))) \}$$

Professoren			
PersNr	Name	Rang	Raum
2125	Sokrates	C4	226
2126	Russel	C4	232
2127	Kopernikus	C3	310
2133	Popper	C3	52
2134	Augustinus	C3	309
2136	Curie	C4	36
2137	Kant	C4	7

prüfen			
MatrNr	VorlNr	PersNr	Note
28106	5001	2126	1
25403	5041	2125	2

Studenten		
MatrNr	Name	Semester
24002	Xenokrates	18
25403	Jonas	12
26120	Fichte	10
26830	Aristoxenos	8
27550	Schopenhauer	6
28106	Carnap	3
29120	Theophrastos	2
29555	Feuerbach	2

Tupel-Relationenkalkül (TRK)

Anfragen sind prädikatenlogische Mengenspezifikationen der Form

$$\{t \mid F(t)\} \quad \text{bzw.} \quad \{t_1.A_1, \dots, t_n.A_n \mid F(t_1, \dots, t_n)\}$$

wobei F eine prädikatenlogische Formel über einer Menge von Tupelvariablen ist mit t bzw. t_1, \dots, t_n als einzigen freien Variablen und A_1, \dots, A_n Attributnamen sind.

Tupel-Relationenkalkül (TRK)

Zulässige prädikatenlogische Ausdrücke für F :

- 1) Für eine Variable r und eine Relation R ist $r \in R$ (auch $R(r)$) ein zulässiger Ausdruck.
- 2) Für Variablen r, s , Attribute A, B mit $dom(A)=dom(B)$, Konstante $c \in dom(A)$ und Vergleichsoperationen $\theta \in \{=, \neq, <, >, \leq, \geq\}$ sind $r.A \theta s.B$ und $r.A \theta c$ zulässige Ausdrücke.
- 3) Falls $F1$ und $F2$ zulässige Ausdrücke sind, dann sind auch $F1 \wedge F2$, $F1 \vee F2$, $\neg F1$ zulässig.
- 4) Falls F ein zulässiger Ausdruck mit einer freien Tupelvariablen r ist, dann sind auch $\exists r: F(r)$ und $\forall r: F(r)$ zulässige Ausdrücke.
- 5) Falls F ein zulässiger Ausdruck ist, dann ist auch (F) zulässig.
- 6) Nur die aufgrund von 1) bis 5) erzeugten Ausdrücke sind zulässig.

Übersetzung TRK-DRK

Übersetzung in den DRK:

In einer Anfrage der Form $\{t \mid F(t)\}$ oder $\{t.B_1, \dots, t.B_n \mid F(t)\}$ wird

- jeder Term der Form $r.A_i$ mit einer Tupelvariablen r durch eine Domain-Variable ra_i ersetzt und
- jeder Term der Form $r \in R$ über einer Relation mit Schema $R(A_1, \dots, A_m)$ durch einen Term $R(ra_1, \dots, ra_m)$ mit Domain-Variablen ra_1, \dots, ra_m .

Beispiele für TRK-Anfragen

Anfrage:

C4-Professoren

$$\{p \mid p \in Professoren \wedge p.Rang = 'C4'\}$$

Anfrage:

Studenten, die mindestens eine Vorlesung von Curie besuchen

$$\begin{aligned} \{s \mid & s \in Studenten \\ & \wedge \exists h \in hoeren (s.MatrNr = h.MatrNr \\ & \wedge \exists v \in Vorlesungen (h.VorlNr = v.VorlNr \\ & \wedge \exists p \in Professoren (p.persNr = v.gelesenVon \\ & \wedge p.Name = 'Curie'))))\} \end{aligned}$$

Beispiele für TRK-Anfragen (2)

Anfrage:

Studenten, die **alle** vierstündigen Vorlesungen gehört haben

$$\{s \mid s \in \text{Studenten} \wedge \forall v \in \text{Vorlesungen} (v.SWS = 4 \Rightarrow \exists h \in \text{hoeren} (h.VorlNr = v.VorlNr \wedge h.MatrNr = s.MatrNr))\}$$

Domain-unabhängige Anfragen

Problem: Anfragen liefern u.U. unendliche Resultatmengen.

Beispiel: $\{ t \mid \neg (t \in R) \}$ für eine endliche Relation R .

Wahl des Universums U bei der Interpretation einer TRK-Formel F :

- **unbeschränkte** Interpretation: U besteht aus der Vereinigung aller Domains der Datenbank
- **beschränkte** Interpretation: U besteht aus der aktiven Domäne ($ADom$) der Datenbank und der Formel F , d.h.:
 - alle Attributwerte von Relationen, die in F referenziert werden
plus
 - alle in der Formel vorkommenden Konstanten

Domain-unabhängige Anfragen (2)

Eine Anfrage des TRK heißt **domain-unabhängig** g.d.w. für jede mögliche Ausprägung der Datenbank die Resultatmenge von der Datenbank, nicht aber von den zugrundeliegenden Domains abhängt.

Definition:

F heißt **domain-unabhängig** g.d.w. ihre unbeschränkte Interpretation für jede Domäne Dom , die die aktive Domäne $ADom$ umfasst, mit ihrer beschränkten Interpretation übereinstimmt.

Domain-unabhängigkeit: Beispiel

Sei R eine Relation mit $sch(R)=\{A\}$, $dom(A)=\{1,2\}$, $val(R)=\{\langle 1 \rangle\}$.

a) Der Ausdruck $\{t \mid \neg (t \in R)\}$ hat

als unbeschränkte Interpretation den Wert $\{\langle 2 \rangle\}$ und
als beschränkte Interpretation den Wert \emptyset .

b) Der Ausdruck $\{t \mid \exists x (\neg (x = t) \wedge x.A = 1)\}$ hat

als unbeschränkte Interpretation den Wert $\{\langle 2 \rangle\}$ und
als beschränkte Interpretation den Wert \emptyset .

c) Der Ausdruck $\{t \mid t \in R \wedge \forall x (x = t)\}$ hat

als unbeschränkte Interpretation den Wert \emptyset und
als beschränkte Interpretation den Wert $\{\langle 1 \rangle\}$.

Beschränkte und unbeschränkte Ausdrücke

$F(x)$ mit Variablen x heisst **beschränkt**,
wenn für jede Interpretation von F gilt:

- wenn $F(x)$ wahr ist,
dann liegt die Interpretation von x in $ADom$ von F .

$F(x)$ mit Variablen x heisst **unbeschränkt**,
wenn für jede Interpretation von F gilt:

- wenn die Interpretation von x nicht in $ADom$ von F liegt,
muss $F(x)$ wahr sein.

Sicherheit durch Beschränkung / Unbeschränkung

Durch syntaktische Einschränkungen von Formeln können

- **freie** Variablen **beschränkt** werden
(so dass in $\{t \mid F(t)\}$ $F(t)$ nur für t aus $ADom$ erfüllbar ist),
- durch **Existenzquantoren** gebundene Variablen **beschränkt** werden
(so dass in $\exists x G(x)$ $G(x)$ nur für x aus $ADom$ erfüllbar ist)
- durch **Allquantoren** gebundene Variablen **unbeschränkt** werden
(so dass in $\forall x G(x)$ $G(x)$ für Interpretationen von x , die nicht im $ADom$ liegen, immer wahr ist).

Rezept für sichere Anfragen

Satz:

Eine Anfrage des TRK ist sicher und damit garantiert domain-unabhängig, wenn

- die Anfrage die Form $\{t \mid t \in R \wedge F(t)\}$ hat und
- jede existenzquantifizierte Teilformel die Form $\exists x: x \in R \wedge F(x)$ hat und
- jede allquantifizierte Teilformel die Form $\forall x: x \in R \Rightarrow F(x)$ hat.

Mächtigkeit relationaler Anfragesprachen

Satz:

- Alle Anfragen, die sich mit der **RA** ausdrücken lassen, lassen sich auch mit dem **sicheren TRK** oder dem **sicheren DRK** ausdrücken.
- Alle Anfragen, die sich mit dem **sicheren TRK** ausdrücken lassen, lassen sich auch mit der **RA** oder dem **sicheren DRK** ausdrücken.
- Alle Anfragen, die sich mit dem **sicheren DRK** ausdrücken lassen, lassen sich auch mit der **RA** oder dem **sicheren TRK** ausdrücken.