Реляционная модель данных. Часть 2

Дж.Ульман

Основы систем баз данных

Глава 4

Реляционное исчисление

- Реляционное исчисление на кортежах
- Реляционное исчисление на доменах

В основе лежат понятия математической логики – исчисление предикатов

Реляционное исчисление на кортежах

Выражения реляционного исчисления на кортежах имеют вид:

$$\{\ t \mid \varphi\ (t)\ \}$$

где t – переменная кортеж, а φ – формула, построенная из атомов и совокупности операторов

<u>*Атомы*</u> формул φ могут быть 3 типов

1. $R\ (s)$, где R – имя отношения, а s – переменная-кортеж. Этот атом означает, что s есть кортеж в отношении R
2. $s[i]\ \theta\ u[j]$, где s и u являются переменными-кортежами, а θ – арифметический оператор сравнения (<, >, = и т.д.) Этот атом означает, что i *–ый* компонент s находится в отношении θ с j *–ым* компонентом u
3. $s[i]\ \theta\ \alpha$, где α – константа

Реляционное исчисление на кортежах

- В выражениях РИК переменные-кортежи могут быть свободными или связанными.
- Переменная связана, если в формуле ей предшествует квантор $\exists$ или $\forall$.

Формулы

1. Каждый атом есть формула. Все вхождения переменных-кортежей в атоме являются свободными в этой формуле.
2. Если φ_1 и φ_2 – формулы, то $\varphi_1 > \varphi_2$, $\varphi_1 \vee \varphi_2$, $\neg \varphi_1$ – также формулы. Переменные-кортежи являются свободными или связанными в построенных формулах точно так же какими он были в исходных формулах.
3. Если φ – формула, то $\exists$ s (φ) также формула
4. Если φ – формула, то $\forall$ s (φ) также формула
5. В формулах могут быть скобки
6. Ничто иное не является формулой

Реляционное исчисление на кортежах

- Выражение РИК есть выражение вида

$$\{\ t \mid \varphi\ (t)\ \},$$

где t единственная переменная-кортеж в формуле φ .

<u>Примеры выражений РИК для отношений</u>

Если $\{\ t \mid R\ (t)\ \}$ – множество кортежей отношения R,
а $\{\ t \mid S\ (t)\ \}$ – множество кортежей отношения S, то

- Объединение R U S выражается как
  $\{\ t \mid R\ (t) \vee S\ (t)\ \}$
- Разность R \ S выражается как
  $\{\ t \mid R\ (t) > \neg\ S\ (t)\ \}$

Ограничение реляционного исчисления

- Ранее определенное РИК может представлять **бесконечные** отношения.
- Пример $\{ t \mid \neg R(t)\}$.
  Означает множество кортежей длины t , но не принадлежащих R .
  Так как НЕ ВОЗМОЖНО перечислить все такие кортежи (над каким доменом?) рассматривают безопасные выражения $\{ t \mid \varphi(t) \}$.
- Для таких выражений можно продемонстрировать, что каждый компонент t, удовлетворяющий φ , принадлежит конечному множеству $DOM(\varphi)$

Ограничение реляционного исчисления

- $DOM(\varphi)$ определяется как множество всех символов, которые либо **явно упоминаются** в φ, либо служат **компонентами кортежей отношений упоминаемых** в φ.
- Заметим, что $DOM(\varphi)$ определяется не в зависимости от вида φ, а как **функция фактических отношений**, которые должны подставляться вместо переменных отношений в φ
- Так как все отношения – конечны, то и $DOM(\varphi)$ - всегда **конечно**.

Пример Пусть φ (t) имеет вид $t[1]=a \wedge R(t)$,

где $R(A,B)$ – бинарное отношение,

тогда $DOM(\varphi)$ есть унарное отношение $\{a\} \cup \pi_A(R) \cup \pi_B(R)$

Безопасные выражения РИК

Выражение РИК $\{ t / \varphi(t) \}$ называется безопасным если удовлетворяются следующие условия:

1. Всякий раз, когда t удовлетворяет φ, каждый компонент t есть элемент $DOM(\varphi)$
2. Для любого подвыражения φ вида $\exists u(\omega(u))$ каждый компонент u принадлежит $DOM(\omega)$, если u удовлетворяет ω
3. Если для любого подвыражения φ вида $\forall u(\omega(u))$ каждый компонент u не принадлежит $DOM(\omega)$, то u удовлетворяет ω

Условия 2,3 позволяют устанавливать истинность формул вида $\exists u(\omega(u))$ или $\forall u(\omega(u))$, рассматривая только u, составленные из принадлежащих $DOM(\omega)$ символов

Примеры

1. Любая формула $\exists\, u(R(u) \vee \ldots)$ удовлетворяет условию 2
2. Любая формула $\forall u\, (\neg R(u) \vee \ldots)$ удовлетворяет условию 3

Безопасные выражения РИК

- Хотя условие 3 может показаться не интуитивным, заметим, что $\forall u(\omega(u))$ логически эквивалента $\neg\exists u(\neg\omega(u))$
- Последняя не является безопасной т. и т. т. когда существует некоторое u_0, для которого истинно $\neg\omega(u_0)$ и u_0 не принадлежит домену формулы $\neg\omega$.
- Так как домены ω и $\neg\omega$ совпадают, то условие 3 устанавливает, что формула $\forall u(\omega(u))$ безопасна, когда безопасна формула $\neg\exists u(\neg\omega(u))$

Редукция РА к РИК

<u>Теорема.</u> Если E – выражение реляционной алгебры, то существует эквивалентное ему безопасное выражение в реляционном исчислении с переменными-кортежами.

<u>Доказательство.</u>
Индукция по числу вхождений операторов в E.

<u>Базис.</u> Нуль операторов.

Тогда E – либо постоянное выражение $\{\mathrm{t}_1, \mathrm{t}_2, \ldots, \mathrm{t}_\mathrm{n}\}$, либо переменная R, обозначающая отношение.

В последнем случае E эквивалентно $\{\ \mathrm{t} \mid \mathrm{R}\ (\mathrm{t})\}$, которое безопасно.

В первом случае Е эквивалентно $\{t \mid t=t_1 \vee t=t_2 \vee \ldots \vee t=t_n\}$, где $t=t_i$ – краткая запись $t[1]=t_i[1] > t[2]=t_i[2] > \ldots > t[n]=t_i[n]$

Редукция РА к РИК

<u>Индукция.</u> Предположим, что E имеет не менее одного оператора и теорема истинна для выражений с числом вхождений операторов меньшим, чем в E.

<u>Случай 1</u>. $E = E_1 \cup E_2$.

Пусть E_1 эквивалентно безопасному выражению $\{t|\psi_1(t)\}$

Пусть E_2 эквивалентно безопасному выражению $\{t|\psi_2(t)\}$

Тогда для E строим эквивалентное выражение

$$\{t|\ \psi_1(t) \vee \psi_2(t)\}.$$

Если t удовлетворяет $\psi_1(t) \vee \psi_2(t)$, то каждый компонент t принадлежит $DOM(\psi_1)$ или $DOM(\psi_2)$.

Так как $DOM\ (\psi_1(t) \vee \psi_2(t)){=}DOM(\psi_1(t)) \cup DOM(\psi_2(t))$, то E эквивалентно безопасному выражению $\{t|\ \psi_1(t) \vee$

Редукция РА к РИК

<u>Случай 2</u>. $E = E_1 - E_2$.

Пусть E_1 эквивалентно безопасному выражению $\{t|\psi_1(t)\}$

Пусть E_2 эквивалентно безопасному выражению $\{t|\psi_2(t)\}$

Тогда для E строим эквивалентное выражение

$$\{t|\ \psi_1(t) > \neg\ \psi_2(t)\}.$$

Так как $DOM\ (\psi_1(t) > \neg\ \psi_2(t)) = DOM(\psi_1(t))\ \cup\ DOM(\psi_2(t))$,

то E эквивалентно безопасному выражению

$$\{t|\ \psi_1(t) > \neg\ \psi_2(t)\}$$

Редукция РА к РИК

Случай 3. $E = E_1 \times E_2$.

Пусть E_1 – отношение арности k и эквивалентно безопасному выражению $\{t|\psi_1(t)\}$

Пусть E_2 – отношение арности s и эквивалентно безопасному выражению $\{t|\psi_2(t)\}$

Тогда для E эквивалентно выражению

$$\psi(t) = \{t^{(k+s)}| \; \exists u \; \exists v \; (\; \psi_1(u) > \psi_2(v) \; \& \; t[1]=u[1] \; \& \; \ldots \; \& \; t[k]=u[k] \; \& \; t[k+1]=v[1] \; \& \; \ldots \; \& \; t[k+s]=v[s] \;) \; \}.$$

$DOM(\psi(t))=DOM(\psi_1(t)) \cup DOM(\psi_2(t))$.

Выражение $\psi(t)$ безопасно, так как $t[i]$ ограничено значениями $u[i]$, если $i \le k$, и значениями $v[i]$, если $k < i \le k+s$

Редукция РА к РИК

Случай 4. $E = \pi_{i1,\ i2,\ \ldots\ ik}(E_1)$.

Пусть E_1 – отношение арности n и эквивалентно безопасному выражению $\{t|\psi_1(t)\}$

Тогда для E эквивалентно выражению

$$\psi(t) = \{t^{(k)}|\ \exists u\ (\ \psi_1(u)\ \&\ t[1]=u[i_1]\ \&\ \ldots\ \&\ t[k]=u[i_k]\ )\ \},$$

где $k \leq n$

DOM $(\psi(t))$=DOM$(\psi_1(t))$.

Выражение $\psi(t)$ безопасно, так как $t[i]$ ограничено значениями $u[i_j]$, если $j \leq n$

Редукция РА к РИК

Случай 5. $E = б_F (E_1)$.

Пусть E_1 – эквивалентно безопасному выражению

$\{t|\psi_1(t)\}$

Тогда для E эквивалентно выражению

$\psi(t) = \{t|\psi_1(t) \ \& \ F' \}$, где F' есть F , (1)

в которой каждый операнд обозначающий компонент i заменяется на $t[i]$.

Можно DOM $(\psi(t))$ = DOM$(\psi_1(t))$. Выражение (1) безопасно, так как каждый компонент t ограничивается символами, которые есть в DOM$(\psi_1(t))$.

Реляционное исчисление на доменах

Выражения реляционного исчисления на доменах имеют вид:

$$\{ x_1, x_2, \dots , x_k \mid \psi (x_1, x_2, \dots , x_k) \}$$

а ψ – формула, построенная из атомов и совокупности операторов, а $x_1, x_2, \dots , x_k$ переменные на доменах отношений, входящих в формулу

Атомы формул ψ могут быть

1. $R (x_1, x_2, \dots , x_k)$, где R – k- арное отношение и каждое x_i есть константа или переменная на домене.
2. $x\theta y$, где x и y – константы или переменные на доменах, а θ – арифметический оператор сравнения (<, >, = и т.д.)

$R (x_1, x_2, \dots , x_k)$ означает, что значения x_i, которые являются переменными должны быть выбраны так, чтобы $x_1, x_2, \dots , x_k$ был кортежем в отношении R

Реляционное исчисление на доменах

Выражение реляционного исчисления с переменными на доменах

$\{x_1, x_2, \ldots , x_k | \psi(x_1, x_2, \ldots , x_k)\}$ является безопасным, когда

1. из истинности $\psi\ (x_1, x_2, \ldots , x_k)$ следует, что x_i принадлежит $DOM(\psi)$
2. если $\exists u\ (\omega\ (u))$ – подформула ψ, то из истинности $\omega(u)$ следует, что u принадлежит $DOM(\omega)$
3. если $\forall u\ (\omega\ (u))$ – подформула ψ, то из истинности $\neg\omega(u)$ следует, что u принадлежит $DOM(\omega)$

Редукция РИК к РИД

Для каждого выражения $\{t \mid \varphi(t)\}$ РИК выражение РИД строится следующим образом.

Если t имеет арность k, то

- введем k новых переменных на доменах $t_1, t_2, \ldots, t_k$
- и заменим $\{t \mid \varphi(t)\}$ на $\{t_1, t_2, \ldots, t_k \mid \varphi'(t_1, t_2, \ldots, t_k)\}$, где

2) φ' есть φ, в которой

3) каждый атом $R(t)$ заменен атомом $R(t_1, t_2, \ldots, t_k)$,

4) каждое свободное вхождение $t[i]$ – переменной t_i