Реляционное исчисление

Общая характеристика

Запрос – формула некоторой формально-логической теории; описывает свойства желаемого результата.

Ответ – множество объектов из области интерпретации (базы данных), на котором истинна формула, соответствующая запросу.

Формально-логическая теория – ***теория исчисления предикатов первого порядка,*** в которой формула задается в виде ***предиката***.

Понятие предиката (1)

Даны произвольные множества $D_1, D_2, \ldots, D_n$, $D_i \cap D_j = 0$ для любых $i \neq j$, и переменные $x_1, x_2, \ldots, x_n$, $x_i \in D_i$ для любых $i = 1, 2, \ldots, n$.

Предикатом (или предикатной функцией) называется функция $P(x_1, x_2, \ldots, x_n)$, принимающая одно из двух значений – 1 или 0 (истина или ложь).

$x_1, x_2, \ldots, x_n$ – <u>предикатные переменные</u>

$D_1, D_2, \ldots, D_n$ – <u>область интерпретации предиката</u>

Понятие предиката (2)

- Логические операции – $\wedge$ (и), $\vee$ (или), $\neg$ (не)
- Кванторы – $\forall$ (всеобщности), $\exists$ (существования)

$\forall x\ (f(x))$ – для всех значений x из области интерпретации предиката формула $f(x)$ имеет значение "истина";

$\exists x\ (f(x))$ – существует, по крайней мере, одно значение x из области интерпретации предиката, для которого формула $f(x)$ имеет значение "истина"

$\forall x\ (f(x))$ эквивалентно $\neg\exists x\ (\neg f(x))$

Связь предиката с базой данных

Область интерпретации предиката – база данных

Соответствие между предикатом $P(x_1, x_2, \ldots, x_n)$ и отношением $r(R)$, $R(A_1{:}D_1, A_2{:}D_2, \ldots, A_n{:}D_n)$:

$a_1 \in D_1$, $a_2 \in D_2$, …, $a_n \in D_n$

1. Если $P(a_1, a_2, \ldots, a_n) = 1$, то $\langle a_1, a_2, \ldots, a_n \rangle$ есть *выборка* отношения $R(A_1{:}D_1, A_2{:}D_2, \ldots, A_n{:}D_n)$, т.е. $\langle a_1, a_2, \ldots, a_n \rangle \in r$

2. Если $P(a_1, a_2, \ldots, a_n) = 0$, то $\langle a_1, a_2, \ldots, a_n \rangle \notin r$

Реляционное исчисление с переменными-кортежами

1. Областью определения переменных являются отношения
2. Переменные-кортежи должны удовлетворять определенной схеме отношения R
3. Предикат – это правильно построенная формула (*wff* – *well formulated formula*) $\psi(t)$. Выбираются те кортежи *t*, для которых $\psi(t)$ дает значение 1

Атомы wff (1)

1. Пусть *r(R)* – некоторая реализация отношения, удовлетворяющая схеме *R;*
   t – некоторая переменная-кортеж, удовлетворяющая схеме *R.*

Тогда *r(t)* – атом; означает, что *t* есть кортеж в отношении *r* (т.е. формула истинна, если
$t \in r$)

Атомы wff (2)

2. Пусть *r(R)* – некоторая реализация отношения, удовлетворяющая схеме *R;*
   u и *v* – переменные-кортежи из отношения *r(R)* (т.е. $u \in r$, $v \in r$); θ – арифметическая операция сравнения ($<$, $=$, $>$, $\geq$, $\neq$, $\leq$);
   A, *B* – атрибуты схемы отношения *R*, сравнимые по операции θ.

Тогда $u[A]\ \theta\ v[B]$ – атом

t[X] – значение переменной *t* по атрибуту *X*

Атомы wff (3)

3. Пусть *u* – переменная-кортеж из отношения *r(R)* (т.е. $u \in r$);
θ – арифметическая операция сравнения
($<, =, >, \geq, \neq, \leq$);
A, *B* – атрибуты схемы отношения *R*, сравнимые по операции θ;
c – константа из домена, на котором определен атрибут *B*.

Тогда $u[A]\ \theta\ c$ (или $c\ \theta\ u[A]$) – атом

Выражение реляционного исчисления (1)

$$\{t(R) \mid \psi(t)\},$$

где t – переменная-кортеж, удовлетворяющая схеме отношения R; единственная переменная, имеющая свободное вхождение в формулу $\psi(t)$;
$\psi(t)$ – правильно построенная формула

Интерпретация: множество кортежей t, удовлетворяющих схеме отношения R, таких, для которых правильно построенная формула $\psi(t)$ истинна

Выражение реляционного исчисления (2)

Пример

Есть отношение R(*Имя*, *Стипендия*);
атрибут *Стипендия* определен на домене
D = {«да», «нет»}.

Получить из отношения имена всех студентов, получающих стипендию:

{ t(*Имя*) | $\exists$x(R) (r(x) $\wedge$ x[*Стипендия*] = «да» $\wedge$ x[*Имя*] = t[*Имя*]}

Безопасные выражения

$\{t \mid \neg r(t)\}$ – в общем случае, определяет бесконечное отношение, что недопустимо.

Б*езопасные* выражения вида $\{ t \mid \psi(t) \}$ гарантированно дают ограниченное множество кортежей.

Значения атрибутов кортежей *t* являются элементами некоторого ограниченного универсального множества – $DOM(\psi)$.

$DOM(\psi)$ – унарное отношение, элементами которого являются символы, которые либо явно появляются в ψ, либо служат компонентами какого-либо кортежа в некотором отношении *R*, упоминаемом в ψ

Реляционное исчисление с переменными на доменах (1)

Атомы:

- $r(x_1, x_2, \dots, x_n)$, где r – отношение, удовлетворяющее схеме $R(A_1, A_2, \dots A_n)$, и каждое x_i есть константа или переменная на домене;
- $u\ \theta\ v$, где u и v – константы или переменные, определенные на доменах, совместимых по операции θ, θ – арифметическая операция сравнения ($<, =, >, \geq, \neq, \leq$);

Реляционное исчисление с переменными на доменах (2)

Формула реляционного исчисления $\psi(t)$, а также свободные и связанные вхождения переменных определяются так же, как и для исчисления с переменными-кортежами.

Реляционное исчисление с переменными на доменах (3)

Пример. Пусть мы имеем базу данных служащих. Будем считать, что мы определили доменные переменные, имена которых совпадают с именами атрибутов отношения СЛУЖАЩИЕ

WFF исчисления доменов:

СЛУЖАЩИЕ (СЛУ_НОМ:2934, СЛУ_ИМЯ:'Иванов', СЛУ_ЗАРП:22400.00, ПРО_НОМ:1)

примет значение true в том и только в том случае, когда в теле отношения СЛУЖАЩИЕ содержится кортеж <2934, 'Иванов', 22400.00, 1>. Соответствующие значения доменных переменных образуют область истинности этой WFF.

Реляционное исчисление с переменными на доменах (4)

Пример. (продолжение)

Запрос: "Выдать номера и имена служащих, не получающих минимальную заработную плату":

СЛУ_НОМ, СЛУ_ИМЯ WHERE EXISTS СЛУ_ЗАРП1 (СЛУЖАЩИЕ (СЛУ_ЗАРП1) AND СЛУЖАЩИЕ (СЛУ_НОМ, СЛУ_ИМЯ, СЛУ_ЗАРП) AND СЛУ_ЗАРП > СЛУ_ЗАРП1)