

**INSTITUTO FEDERAL DE EDUCAÇÃO CIÊNCIA E TECNOLOGIA DE SÃO
PAULO**

MARCOS VINICIUS DE SOUZA PEREIRA

CALCULO RELACIONAL

**CAMPOS DO JORDÃO-SP
2024**

INTRODUÇÃO

O calculo relacional é uma forma de consulta formal, utilizando-se de uma forte base de lógica matemática de modo que se utilizando de uma forma declarativa pode-se especificar uma consulta.

Uma expressão de calculo relacional permite que sua descrição seja feita sem procedimentos para a obtenção de resultados sendo assim não-procedural, diferente de álgebra relacional onde se faz necessário o uso de procedimentos para se obter uma consulta, contudo a mesma consulta feita com calculo deve ser feita de maneira formal e precisa contendo muita precisão nas informações inseridas, pois nas expressões em calculo não se especifica o modo de se obter o resultado mas sim o tipo de informação que se pretende adquirir.

Existem dois tipos de expressões de cálculos: a expressão de calculo por tuplas (CRT) e a expressão de calculo por domínio (CRN) que são subconjuntos simples de lógica de primeira ordem. As expressões em cálculos são chamadas de fórmulas onde dentro destas fórmulas podem ter variáveis, constantes, operadores lógicos, operadores de comparação e quantificadores.

Uma estrutura de repostas é essencialmente uma mistura de constantes e variáveis e levam uma fórmula a um estado verdadeiro.

CALCULO RELACIONAL POR TUPLAS

A CRT é baseada da especificação de variáveis sobre os tuplos, onde cada variável especificada pode tomar como valor qualquer tuplo de determinada relação.

Sua forma mais simples de expressão é:

$\{ t \mid \text{COND}(t) \}$

onde t é uma variável que representa os tuplos de uma relação e $\text{COND}(t)$ é uma condição sobre t . O resultado de tal expressão é conjunto de todas as variáveis tuplas t que satisfazem a condição indicada como verdadeira $\text{COND}(t)$.

De maneira genérica CRT pode ser expressa como:

$\{ t1.A1, t2.A2, \dots, tn.An \mid \text{COND}(t1, t2, \dots, tn) \}$

de modo que $t1, t2, \dots, tn$ são as variáveis que representam os tuplos de relações, a cada Ai é um atributo da relação na qual ti recebe os valores e $\text{COND}(t1, t2, \dots, tn)$ é uma fórmula aprimorada do calculo demonstrado.

As fórmulas atômicas de cálculos de predicados podem ser as seguintes:

1 – A fórmula atômica $R(ti)$ sendo R é o nome de uma relação e ti é uma variável de tupla.

Identificando a extensão da variável de tupla ti com a relação cujo o nome seja R .

2 – A fórmula atômica $ti.A \text{ op } tj.B$, onde op é um dos operadores de comparação no conjunto $\{=, >, <, \dots\}$, ti e tj são variáveis de tuplas, A é um atributo de relação onde ti se encontra, B é um atributo da relação na qual tj se encontra.

3 – Uma fórmula atômica $ti.A \text{ op } c$ ou $c \text{ op } tj.B$, sendo op um operador de comparação no conjunto $\{=, >, <, \dots\}$, ti e tj são variáveis de tuplas, A é um atributo de relação no qual ti se encontra e B é um atributo de relação no qual tj se encontra e c é um valor contante.

Cada fórmulas atômicas demonstradas anteriormente tem seu valor verdade avaliado como TRUE ou FALSE para uma determinada combinação de tuplas. Para fórmulas tipo 1, caso a variável de tupla seja atribuída a uma tupla da relação R fornecida, assume-se então o valor TRUE; caso contrário, FALSE. Já nas fórmulas tipo 2 e tipo 3, se as variáveis de tuplas forem designadas de forque que os valores atribuídos especificados satisfaçam o predicado, esta assumirá o valor TRUE como verdadeiro. Já que nenhuma variável tupla mostrada ate que aparece quantificada as mesmas podem ser consideradas variáveis livres, entretanto com as definições acima, quantificadores (universal (\forall) ou existencial(\exists)) podem aparecer nas definições das fórmulas, neste caso as variáveis que sucedem são denominadas variáveis limite.

Uma formula é definida por uma ou mais formulas atômicas unidas por operadores lógicos AND, OR, NOT como segue os próximos exemplos:

Começando por qualquer fórmula atômica já é considerada uma fórmula. Se F1 e F2 são fórmulas atômicas, então (F1 AND F2), (F1 OR F2), NOT(F1) e NOT (F2) também são levando em consideração que F1 e F2 tem seus valores verdades derivados da seguinte forma:

- . F1 AND F2 será TRUE apenas se ambos forem verdadeiro.
- . F1 OR F2 será TRUE apenas quando qualquer uma for verdadeiro.
- . NOT F1 será TRUE quando F1 for falso e vise e verso (o mesmo ocorrera com F2).

Se F1 é uma fórmula atômica, então $(\exists t)(F1)$ também será e seu valor verdadeiro apenas será TRUE se a fórmula F for verdadeira para, pelo menos, uma tupla atribuída para ocorrências livres de t em F.

Se F1 é uma fórmula atômica, então $(\forall t)(F1)$ também será, e seu valor verdadeiro apenas será TRUE se a fórmula F for verdadeira para todas as tuplas atribuídas para ocorrências livres de t em F.

É possível escrever expressões equivalentes manipulando operadores e quantificadores em alguns casos dessa manipulação podem ser declarados da seguinte forma: $F1 \text{ AND } F2 \equiv \text{NOT} (\text{NOT } F1 \text{ OR } \text{NOT } F2)$; $(\forall t) \in r (F1(t)) \equiv \text{NOT} (\exists t) \in r (\text{NOT } F1(t))$; $F1 \Rightarrow F2 \equiv \text{NOT } F1 \text{ OR } F2$; $(\forall x) (F(x)) \equiv \text{NOT} (\exists x) (\text{NOT } (F(x)))$; $(\exists x) (F(x)) \equiv \text{NOT} (\forall x) (\text{NOT } (F(x)))$; $(\forall x) (F(x) \text{ AND } P(x)) \equiv \text{NOT} (\exists x) (\text{NOT } (F(x)) \text{ OR } \text{NOT } (P(x)))$; $(\forall x) (F(x) \text{ OR } P(x)) \equiv \text{NOT} (\exists x) (\text{NOT } (F(x)) \text{ AND } \text{NOT } (P(x)))$; $(\exists x) (F(x) \text{ OR } P(x)) \equiv \text{NOT} (\forall x) (\text{NOT } (F(x)) \text{ AND } \text{NOT } (P(x)))$; $(\exists x) (F(x) \text{ AND } P(x)) \equiv \text{NOT} (\forall x) (\text{NOT } (F(x)) \text{ OR } \text{NOT } (P(x)))$.

A seguir alguns exemplos de consultas feitas em CRT.

Ex: Encontre todos os funcionários com salários acima de R\$3.500,00 e de os nomes e sobrenomes destes mesmos funcionários, em seguida mostre os nomes dos funcionários que trabalham na área de informática da empresa.

$\{t \mid \text{EMPREGADO}(t) \text{ AND } t.\text{SALARIO} > 3500\}$

$\{t.\text{NOME}, t.\text{SOBRENOME} \mid \text{EMPREGADO}(t) \text{ AND } t.\text{SALARIO} > 3500\}$

$\{t.\text{NOME}, t.\text{SOBRENOME}, t.\text{ENDERECO} \mid \text{EMPREGADO}(t) \text{ AND } (\exists d) (\text{DEPARTAMENTO}(d) \text{ AND } d.\text{NOME} = 'informatica' \text{ AND } d.\text{NUMERODEP} = t.\text{NUD})\}$

As únicas variáveis de tupla livres em uma expressão de calculo relacional devem ser aquelas a esquerda da barra pipe (|).

EXPRESSÕES SEGURAS

Uma expressão realizada em CRT podem gerar diversas relações. Para a expressão $\{t \mid \text{NOT}(R(t))\}$ pode ser gerada uma infinidade de tuplas que não estão em R, dessa forma esta maneira não é segura. De certa forma a maior parte dessas tuplas contem valores que não estão no banco de dados, assim trazendo valores que não são relevantes.

Uma expressão segura em calculo relacional é uma expressão que possa fornecer um número finito tuplas como resultado. Para ser mais específico sobre expressão segura podemos sutilizar o conceito de domínio. O domínio de uma expressão P é o conjunto de todos os valores referenciados por P sendo assim incluindo os valores propriamente dito as por P e os valores que utilizam da menção de uma referência por P.

CALCULO RELACIONAL DE DOMÍNIO

Podemos entender o CRD pois em sua diferença do CRT é que as variáveis se estendem sobre valores únicos de domínios de atributos. Para se que se forme a relação de grau n para um resultado de consulta, se faz necessário cria n variáveis de domínio para cada atributo.

Uma expressão do calculo de relacional de tuplas tem a forma $\{x_1, x_2, \dots, x_n \mid \text{predicado}(x_1, x_2, \dots, x_n, x_{n+1}, x_{n+2}, \dots, x_{n+m})\}$ onde $x_1, x_2, \dots, x_n, x_{n+1}, x_{n+2}, \dots, x_{n+m}$ são variáveis de domínio aplicada sobre o domínio dos atributos exigidos na consulta e predicado, a fórmula atômica do CRD pode ser especificada da seguinte forma:

Uma fórmula atômica $R(x_1, x_2, \dots, x_n)$ onde R é o nome de uma relação de grau j e cada $x_i, 1 \leq i \leq j$, é uma variável de domínio. Isto resulta que uma lista de valores de $\langle x_1, x_2, \dots, x_j \rangle$ deve ser uma tupla de relação R, onde x_i é o valor do i-ésimo valor de atributo da tupla. Outra maneira é a fórmula atômica $x_i \text{ op } x_j$, onde op é um operador de comparação $\{=, <, >, \dots\}$ e x_i e x_j são variáveis de domínio, podendo também ter as formula onde acrescentamos uma constante qualquer ficando então $x_i \text{ op } c$ ou $c \text{ op } x_j$ sendo c a constante.

EXPRESSÕES SEGURAS

As expressões consideradas seguras em CRD devem seguir as seguintes orientações. Todos os valores que nas colunas da expressão deveram ser valores dentro do domínio da mesma, para todas as sub-formulas $(\exists x) (P(x))$ a sub-formula só será verdadeira se o P(x) for verdadeiro para todos os valores de x dentro do domínio P. para a fórmula de $(\forall x) (P(x))$ considera-se o mesmo raciocínio para o domínio de P, as proposições anteriores são meios de testarmos todas as sub-formulas sem que precise de infinitas possibilidades

de ocorrência. Em seguida usaremos os mesmos exemplos já feitos em CRT feitos agora em CRD para fins de comparação do modo.

Ex: Encontre todos os funcionários com salários acima de R\$3.500,00 e de os nomes e sobrenomes destes mesmos funcionários, em seguida mostre os nomes dos funcionários que trabalham na área de informática da empresa.

$\{qrstuvwxyz \mid (\exists x) \text{EMPREGADO}(qrstuvwxyz) \text{ AND } x > 3500\}$

$\{qs \mid (\exists x) \text{EMPREGADO}(qrstuvwxyz) \text{ AND } x > 3500\}$

$\{qsv \mid (\exists z) (\exists l) (\exists m) (\text{EMPREGADO}(qrstuvwxyz) \text{ AND } \text{DEPARTAMENTO}(lmno) \text{ AND } l = \text{'Informática'} \text{ AND } m = z)\}$

UTILIZAÇÃO

A linguagem SQL é um padrão de linguagem de consulta que faz uso de álgebra relacional e calculo relacional portanto todas essas consultas são encontrados em SQL. O SQL permite que uma tabela tenha duas ou mais tuplas identificadas em todos os seus valores de atributos, portanto no geral uma relação SQL não é um conjunto de tuplas pois um conjunto de tuplas não permite elementos idênticos ao invés disso SQL é um multiconjunto de tuplas também conhecido como bag.

CONCLUSÃO

Considerando a pesquisa realizada e combinado a matéria explicada em sala, podemos concluir que o calculo relacional e a álgebra relacional andam em conjunto para formar as diversas maneiras de apresentar uma consulta feita em bancos de dados, os dois assuntos demonstram como é o raciocínio por trás dos comandos utilizados na busca de resultados por dados no dia a dia de uma empresa que precise gerar relatórios com a utilização destes dados.]

REFERENCIA

O.K. Takai; I.C.Italiano; J.E. Ferreira. 2005 - Introdução a Banco de Dados