5. Álgebra relacional

definição
- linguagem formal que oferece um conjunto de operações básico para consulta e manipulação de dados
- oferece uma base teórica para o SQL, utilizado na prática

caracterização
- as suas expressões são uma sequência de operações de álgebra relacional que permitem formular pedidos básicos de recuperação de informação sobre uma ou mais relações
- a cada conjunto de operadores que operam sobre as relações é devolvida uma nova relação

operações verticais

seleção

σ <condição> (R)

seleciona um subconjunto de tuplos que satisfazem critérios de seleção

relação resultante tem o mesmo esquema relacional

operadores = ≠ ≥ ≤ > < AND OR NOT

exemplo σ Idade>60 (Participantes)

projeção

π <atributos> (R)

seleciona atributos da relação

remove linhas duplicadas

exemplo π Nome, Idade (Participantes)

com o encadeamento das operações por vezes é necessário renomear atributos, ou mesmo relações

<novonome> ← <nomeatual>
relações

R(<nattr1>, <nattr2>) ←
RELACAONOME
atributos

renomeação

ρ R2 (B1, B2, …, Bn) (R1) renomear relação e atributos

ρ R2 (R1) renomear relação

ρ (B1, B2, …, Bn) (R1) renomear atributos

união

R1 ∪ R2 comutativa e associativa

para tabelas com o mesmo número de atributos com domínios compatíveis

tuplos duplicados são eliminados

exemplo ParticipantesA ∪ ParticipantesB

interseção

R1 ∩ R2 comutativa e associativa

mesmas condições que união

diferença

R1 − R2 NÃO comutativa

devolve tuplos de R1 que não existem em R2

mesmas condições que união

operações horizontais

Produto cartesiano

$R1 \times R2$

combina cada tuplo da relação R1 com todos os de R2

tendo R1 n tuplos e R2 m tuplos, $Q = R1 \times R2$ terá $n.m$ tuplos

junção
θ join

$R1 \bowtie_{<condition>} R2 = \sigma_{<condition>} (R1 \times R2)$

- equijoin: qnd há o operador = na condição; vai ter duas colunas repetidas
- junção natural: quando há duas colunas com o mesmo nome, fica apenas uma; os atributos repetidos são removidos

divisão

$R1 \div R2 = \pi_{R1-R2}(R1) - \pi_{R1-R2}((\pi_{R1-R2}(R) \times S) - R)$

seja $R1(A1, \ldots, Ar, B1, \ldots, Bk)$ e $R2(B1, \ldots, Bk)$

$R1 \div R2$ vai ter todos os tuplos de $R1(A1, \ldots, AR)$ que tenham correspondencia com todos os tuplos de R2 em $(B1, \ldots, Bk)$

⚠ R1 tem de ter mais atributos que R2!

exemplo departamentos que existem em todas as localizações!

semi join

- left: $R1 \ltimes R2 = \pi_{R1}(R1 \bowtie R2)$; projeção dos atributos de R1 na junção com R2
- right: $R1 \rtimes R2 = \pi_{R2}(R1 \bowtie R2)$

inner join

as operações de junção vistas acima combinam dados de duas tabelas descartando os tuplos que não estão relacionados

outer join

operações de junção, mas mantendo os tuplos que não se relacionam, preenchendo os seus atributos na outra relação a null

R1 ⟗ R2   R1 ⟕ R2   R1 ⟖ R2

exemplo obter empregado que não pertence a nenhum departamento

operações verticais

agregação
VIP

<grupos agregação> $\mathfrak{I}$ <função> (R)

funções avg() min() max() sum() count()

também pode ser usada em projeção

exemplo sexo $\mathfrak{I}$ avg(salario) (Pagamentos)

π sexo, Avg-salario = avg(salario) (Pagamentos)