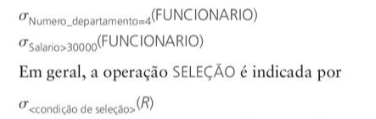
**8.1 Operações relacionais unárias: Seleção e Projeção**

8.1.1 A operação SELEÇÃO

A operação SELEÇÃO é usada para escolher um subconjunto das tuplas de uma relação que satisfaça uma condição de seleção. Pode-se considerar que a operação SELEÇÃO seja um filtro que mantém apenas as tuplas que satisfazem uma condição qualificadora. Como alternativa podemos considerar que essa operação restringe as tuplas em uma relação para apenas aquleas que satisfazem a condição. A operação SELEÇÃO também pode ser visualizada como uma partição horizontal da relação em dois conjunto de tuplas, aquelas que satisfazem a condição e são selecionadas e aquelas que não satisfazem a condição e são descartadas.

Por exemplo, para selecionar a tupla FUNCIONARIO cujo departamento é 4, ou aquelas cujo salário é maior que 30.000, podemos especificar individualmente cada uma dessas duas condições com uma operação de SELEÇÃO da seguinte forma:



Em que o símbolo sigma é usado para indicar o operador SELEÇÃO e a condição de seleção é uma expressão booleana(condição) especificada nos atributos da relação R.

A expressão booleana especificada em <condição de seleção> é composta de uma série de cláusular da forma:

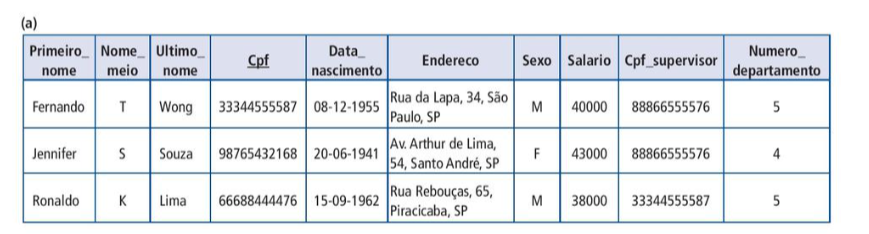
<nome atributo><operador de comparação><valor constante> ou

“ “ “ “ <nome atributo>

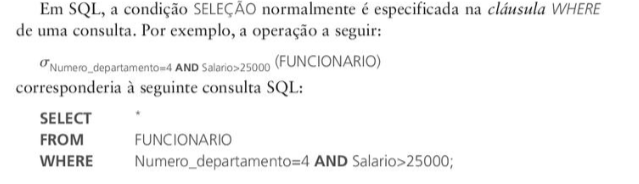
As cláusular podem ser conectadas pelos operadores booleanos padrões, and, or , not, para forma uma condição de seleção geral.

Todos os operadores de comparação no conjunto podem ser aplicados aos atributos cujos domínios são valores ordenados, como númercios ou de data. O domínio de cadei de caracteres também são considerados ordenados com base na ordem alfabética dos caracteres. Se o domíno de um atributo for um conjunto de valores desordenados, somente os operadores de comparação podem ser usados no conjunto. Um exemplo é Cor.

O resultado de seleção à cima é este:

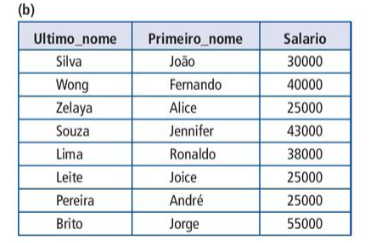


O operador SELEÇÃO é unário, ou seja, é aplicado a uma única relação. Além disso o operação de seleção é aplicada a cada tupla individualmente, logo as condições de seleção não podem envolver mais de uma tupla. O grau da relação resultante de uma operação SELEÇÃO- seu número de atributos- é o mesmo que o grau de R. O número de tuplas resultantes é sempre menor ou igual ao número de tuplas de R. A fração de tuplas selecionadas por uma condição de SELEÇÃO é conhecida como seletividade da condição.



8.1.2 A operação PROJEÇÃO

A operação PROJEÇÃO, por sua vez, seleciona certas colunas da tabela e descarta as outras. Se estivermos interessados apenas em certos atributos de uma relação, usamo a operação PROJEÇÃO para projetar a relação apenas por esse atributos. Portanto, o resultado da operação PROJEÇÃO pode ser visualizado, como uma partição vertical da relação em duas relações: uma tem as colunas(atributos) necessárias e contém o resultado da operação, e a outra contém as colunas descartadas. Por exemplo, para lista último nome, primeiro nome e salário de cada funcionário, podemos usar a operação PROJEÇÃO da seguinte forma:



A forma geral da operação PROJEÇÃO é:



Em que pi é o símbolo usado para representar a operação PROJEÇÃO, e <lista de atributos> e a sublista desejada de atributos da relação R. Mais uma vez, observe que R, em geral, é uma expressão da álgebra relacional cujo resultado e uma relação que no caso mais simples é apenas o nome de uma relação do BD. O resultado da operação PROJEÇÃO tem apenas os atributos especificados em <lista de atributos> na mesma ordem em que eles aparecem na lista. Logo seu grau é igual ao número de atributos em <lista de atributos>

Se a lista de atributos inclui apenas atributos não chave de R, provavelmente haverá tuplas duplicadas. A operação PROJEÇÃO remove quaisquer tuplas duplicadas, de modo que o resultado dessa operaçãp é um conjunto de tuplas distintas, e, portanto, uma relação válida. Isso é conhecido como eliminação de duplicatas.

Por exemplo, considere a seguinte operação PROJEÇÃO:

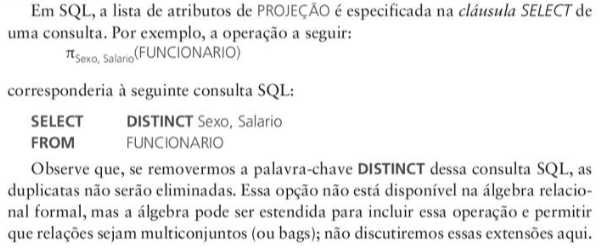


O resultado é:



Observe que a tupla <”F”, 25000> só aparece uma vez na figura acima, embora essa combinação de valores apareça duas vezes na relação FUNCIONARIO. A eliminação de duplicatas envolve a classificação, ou alguma outra técninca para detectar duplicatas, e portanto, aumenta o processamento. Se as duplicatas não fossem eliminadas, o resultado seria um multiconjunto ou bag de tuplas, em vez de um conjunto.

O número de tuplas em uma relação resultado de uma operação PROJEÇÃO é sempre menor igual ao número de tuplas em R. Se a lista de atributos da projeção é uma superchave de R, ou seja, inclui alguma chave de R, a relação resultante tem o mesmo número de tuplas que R. Diferentemente da SELEÇÃO a comutatividade não é mantida na PROJEÇÃO.

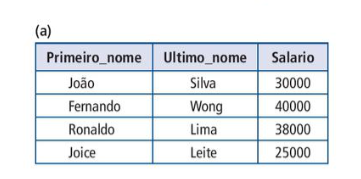


8.1.3 Sequências de operações e a operação RENOMEAR

As relações mostradas nas figura acima, que representam resultado de operação, não possuem nome. Em geral, para a maioria das consultas, precisamos aplicar várias operações da álgebra relacional uma após a outra. Ou podemos escrevê-las como uma única expressão de álgebra relacional, aninhando operações, ou aplicar uma operação de cada vez e criar relações de resultado intermediário. No último caso, temos de dar nomes às relações que mantém os resultados que mantêm os resultados intermediários.

Por exemplo, para recuperar o primeiro nome, o ultimo nome e o salário de todos os funcionários que trabalham no depto 5, devemos aplicar uma operação SELEÇÃO e uma PROJEÇÃO.

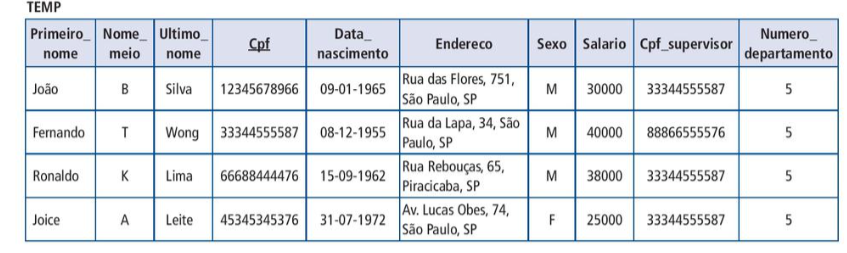


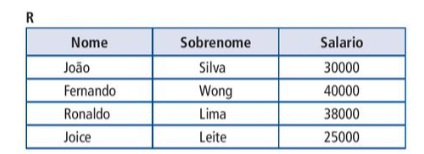


Como alternativa, podemos explicitamente mostrar a sequência de operações dando uma nome a cada relação intermediária, e usando o operador de atribuição indicador por <-, da seguinte forma:



As vezes é mais simples desmembrar uma sequência complexa de operações especificando relações de resultado intermediário, que escrever uma única expressão da álgebra relacional. Também podemos usar essa técnica para renomear os atributos nas relações intermediárias e de resultado. Isso pode ser útil em conexão com operações mais complexas, como UNIÃO ou JUNÇÃO. Para renomear os atributos em uma relação, simplesmente listamos os novos nomes de atributo entre parênteses, como a seguir:



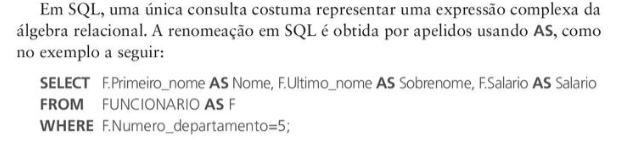


Caso não haja renomeação, os nomes dos atributos se mantém igual aos da relação original.

A operação RENOMEAR formal, que pode renomear o nome da relação, os nomes de atributos, como um operador unário. Em geral, quando aplicada à relação R de grau n, é indicaa por qualquer uma das três formas a seguir:



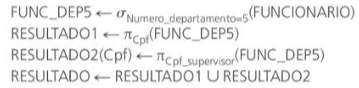
Em que o simbolo rho(p) é usado para indicar o operador RENOMEAR, S é o nome da nova relação, e B1,B2,...,Bn são os novos nomes de atributos. A primeira expressão renomeia tanto a relação quantos seus atributos, a segunda renomeia apenas a relação, e a terceira renomeia apenas os atributos.



**8.2 Operações de álgebra relacional com base na teoria dos conjuntos**

8.2.1 As operações UNIÃO, INTERSEÇÃO e DIFERENÇA(SUBTRAÇÃO)

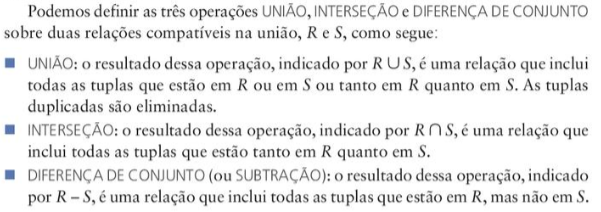
As operações matemáticas são muito importantes na álgebra relacional. Por exemplo, para recuperar os número de Cadastro de Pessoa Física, famoso CPF, de todos os funcionários que trabalham no departamento 5, ou supervisionam diretamente um funcionário que trabalha nesse departamento, podemos usar a operação UNIÃO da seguinte forma:



A realação RESULTADO1 tem o Cpf de todos os funcionários que trabalham no depto 5, enquanto RESULTADO2 tem o Cpf de todos os funcionários que supervisionam diretamente um funcionário que trabalha nesse departamento. A operação UNIÃO produz tuplas que estão ou no RESULTADO1 ou no RESULTADO2, ou em ambos, enquanto elimina quaisquer duplicatas. Assim, por exemplo, o valor de Cpf “33344555587” aparece uma vez no resultado.

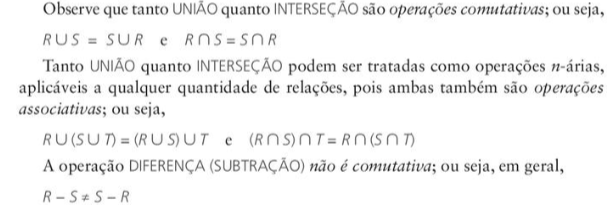


Várias operações de teoria de conjunto são usadas para mesclar os elementos de dois conjuntos de diversas maneiras, incluindo UNIÃO, INTERSEÇÃO e DIFERENÇA DE CONJUNTO. Estas são operações binárias, ou seja, cada uma é aplicada a dois conjuntos de tuplas. Quando essas operações são adaptadas aos BD relacionais, as duas relações sobre as quais qualquer uma dessa três são aplicadas precisam ter o mesmo tipo de tuplas, essa condição é chamada de compatibilidade de união, ou compatibilidade de tipo. Duas relações R(A1,A2,...,AN) e S(B1,B2,...Bn) são compatíveis na união se tiverem o mesmo grau n e se dom(Ai) = dom(Bi). Isso significa que as duas relações tem o mesmo número de atributos e cada par correspondente de atributos tem o mesmo domínio.

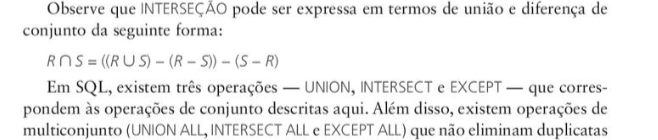


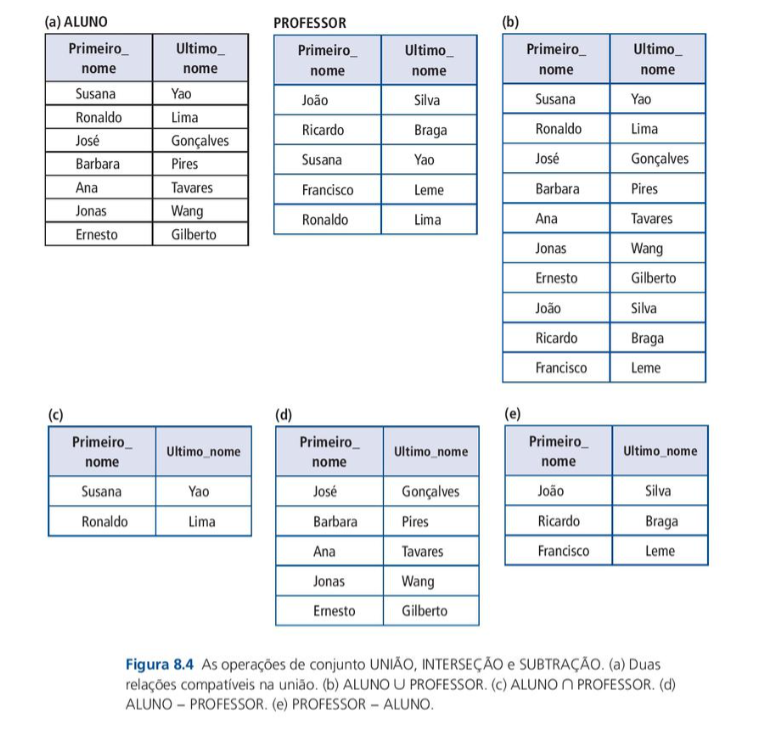
Adotaremos a convenção de que a relação resultante tem os mesmo nomes de atributo da primeira relação R. Sempre é possível renomear os atributos no resultado usando o operador de renomeação.

A figura abaixo ilistra as três operações. As relações ALUNO e PROFESSOR são compatíveis na UNIÃO, e suas tuplas representam os nomes dos alunos e professores respectivamente. O resultado da UNIÃO na figura b) mostra os nomes de todos os alunos e professor. Observe que tuplas duplicas aparecem apenas uma vez. O resultado de operação INTERSEÇÃO em c) inclui apenas aqueles que são tanto alunos quanto professores.



A figura d), mostra os nomes dos alunos que não são professores, e a figura e) mostra os nomes dos professores que não são alunos.





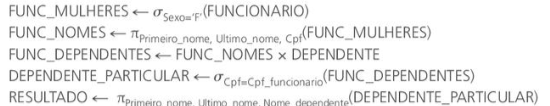
8.2.2 A operação PRODUTO CARTESIANO (PRODUTO CRUZADO)

Está é uma operação de conjunto binária, mas as relações sobre as quais ela é aplicada não precisam ser compatíveis na união. Em sua forma binária, esta operação de conjunto produz um novo elemento combinando cada membro(tupla) de uma relação(conjunto) com cada membro(tupla) da outra relação(conjunto).

Em geral, o resultado de R(A1,A2,...,An) x S(B1,B2,...,Bn) é uma relação Q, com grau n + m atributos Q(A1,A2,..An,B1,B2,...,Bn), nessa ordem. A relação resultante Q, tem uma tupla para cada combinação de tupla, uma de R e uma de S, logo, se R tem nR tuplas e S tem nS tuplas, então R x S(Produto cruzado) terá nR \* nS tuplas, em que nR é a cardinalidade de R.

A operação n-ária PRODUTO CARTESIANO é uma extensão desse conceito, que produz novas tuplas ao concatenar todas as possíveis combinações de tuplas de n relações básicas. A operação PRODUTO CARTESIANO aplicada isoladamente não tem significado. Ela é mais útil quando seguida por uma seleção que combina valores de atributos vindos das relações componentes.

Por exemplo, suponha que queiramos recuperar uma lista dos nomes dos dependentes de cada funcionária. Para isso:





O PRODUTO CARTESIANO cria tuplas com os atributos combinados de duas relações, e assim podemos continuar o processo de busca.