

Álgebra Relacional

■ Tópicos:

- ★ Definição da Álgebra Relacional
- ★ Operadores básicos da Álgebra Relacional
- ★ Operadores adicionais (1ª parte)
- ★ Operações estendidas em Álgebra Relacional
- ★ Modificações de Bases de Dados
- ★ Vistas

■ Bibliografia:

- ★ Secções 2.5 e 2.6 e 6.1 do livro recomendado (6ª edição)
- ★ Secções 2.5 e 2.6 do livro recomendado (7ª edição)
- ★ Capítulos 2 e 3 do livro *The theory of relational databases*

Esquema de Relação e Relação

- A_1, A_2, \dots, A_n são (nomes de) *atributos* com *domínios* D_1, D_2, \dots, D_n , respetivamente.
 - ★ E.g. *customer-name* com domínio {Jones, Smith, Curry, Lindsay}
customer-street com domínio {Main, North, Park}
customer-city com domínio {Harrison, Rye, Pittsfield}
- $R = (A_1, A_2, \dots, A_n)$ é um *esquema de relação*
 - ★ E.g. *Customer-schema* = (*customer-name*, *customer-street*, *customer-city*)
- $r(R)$ é uma *relação* no *esquema de relação* R
 - ★ E.g. *customer*(*Customer-schema*) ou *vip*(*Customer-schema*)
- Formalmente, dados os conjuntos D_1, D_2, \dots, D_n , uma *relação* r é um subconjunto do produto cartesiano
$$D_1 \times D_2 \times \dots \times D_n$$
 - ★ i.e., uma relação é um conjunto de tuplos (a_1, a_2, \dots, a_n) em que $a_i \in D_i$
 - ★ E.g. *customer* = {(Jones, Main, Harrison), (Smith, North, Rye), (Curry, North, Rye), (Lindsay, Park, Pittsfield)}

Instância de Relação

- Os valores (*instância*) de uma relação podem ser representados por uma tabela
- Um elemento t de r é um *tuplo*, representado por uma *linha* da tabela

id	nome	morada	cidade
13123	Luís Trindade	Rue Central	Paris
43242	Pedro Silva	Rua da Sofia	Coimbra
36645	Joana Sobral	Rua D ^a Maria	Coimbra
21313	Susana Dias	Av do Brasil	Lisboa

Linguagem de Consulta/Interrogação

- São necessárias linguagens que acessem e modifiquem os dados
 - ✦ Linguagens de interrogação e manipulação
- Categorias de linguagens:
 - ✦ Procedimentais (para dizer **como** obter os dados)
 - ✦ Declarativas (para dizer **que** dados queremos)
- Linguagens Declarativas
 - ✦ Como estamos a trabalhar sobre um modelo que não reflete necessariamente a forma como os dados estão armazenados, as procedimentais não são muito apropriadas
 - ✦ As linguagens declarativas tornam mais fácil lidar com a complexidade e relações entre os dados
- Linguagens declarativas como a Álgebra Relacional formam a base das linguagens de consulta utilizadas na prática.

Interrogação de bases de dados

- Como deve ser uma linguagem de interrogação?
- O que são perguntas (ou interrogações)?
- E como são as respostas?
- Uma resposta a uma pergunta a uma base de dados relacional é uma relação (ou tabela).
 - ★ Eg. a resposta à pergunta quais os clientes com contas com saldos inferiores a 10, e de quanto são esses saldos, é uma relação com atributos nome (ou n^o) de cliente e saldo, e cujos tuplos são aqueles que se pretendem
- Logo uma pergunta é uma “função” (ou, melhor, uma operação) que, dado um conjunto de relações, devolve uma relação
- Para “formalizar” uma pergunta precisamos de um conjunto de operadores que operem sobre relações
- **Álgebra relacional** como linguagem de interrogação

Álgebra Relacional

■ Seis operadores básicos

- ✧ seleção
- ✧ projeção
- ✧ união
- ✧ diferença de conjuntos
- ✧ produto cartesiano
- ✧ renomeação

■ Os operadores têm como argumentos relações de entrada e devolvem uma relação como resultado (i.e. são fechados sobre relações)

Operação de Seleção

- Notação: $\sigma_p(r)$
- p é designado por **predicado de seleção**
- Definida como:

$$\sigma_p(r) = \{t \mid t \in r \text{ e } p(t)\}$$

onde p é uma fórmula do cálculo proposicional constituída por **termos** ligados por: \wedge (**e**), \vee (**ou**), \neg (**não**)

★ Cada **termo** é da forma:

$\langle \text{atributo} \rangle \text{ op } \langle \text{atributo} \rangle$

ou $\langle \text{atributo} \rangle \text{ op } \langle \text{constante} \rangle$

onde op pode ser: $=$, \neq , $>$, \geq , $<$ e \leq

- Exemplo de seleção:

$\sigma_{\text{branch-name}='Perryridge' \wedge \text{balance} > 100}(\text{account})$

- Propriedades da seleção:

★ Comutatividade: $\sigma_{P_1}(\sigma_{P_2}(r)) = \sigma_{P_2}(\sigma_{P_1}(r)) = \sigma_{P_1 \wedge P_2}(r)$

Operação de Seleção – Exemplo

■ Relação r

r

A	B	C	D
α	α	1	7
α	β	5	7
β	β	12	3
β	β	23	10

■ $\sigma_{A=B \wedge D > 5}(r)$

$\sigma_{A=B \wedge D > 5}(r)$

A	B	C	D
α	α	1	7
β	β	23	10

Operação de Projeção

- Notação:

$$\Pi_{A_1, A_2, \dots, A_k}(r)$$

onde A_1, \dots, A_k são nomes de atributos e r é uma relação.

- O resultado é a relação com os k atributos (k colunas) obtidos eliminando os atributos (colunas) que não estão listadas
- Relembrar: o resultado não tem tuplos duplicados, pois as relações são conjuntos
- E.g. eliminar o atributo *branch-name* de *account*

$$\Pi_{\text{account-number, balance}}(\text{account})$$

- Propriedades da Projeção:

★ Comutatividade com seleção: $\Pi_{Cn}(\sigma_P(r)) = \sigma_P(\Pi_{Cn}(r))$

P só com atributos de Cn!

Operação de Projeção – Exemplo

■ Relação r :

r

A	B	C
α	10	1
α	20	1
β	30	1
β	40	2

■ $\Pi_{A,C}(r)$

A	C
α	1
α	1
β	1
β	2

\Rightarrow

$\Pi_{A,C}(r)$

A	C
α	1
β	1
β	2

(Integridade de referência com álgebra relacional)

- Agora que definimos o operador de projeção temos outra forma de definir chaves estrangeiras!
- Os atributos A na relação r são chave estrangeira referenciando os atributos B na relação s sse

$$\Pi_A(r) \subseteq \Pi_B(s)$$

- Exemplo:
 - ★ id em *customer* ser chave estrangeira referenciado id em *person*, significa/impõe que:

$$\Pi_{id}(customer) \subseteq \Pi_{id}(person)$$

Operação de União

■ Notação: $r \cup s$

■ Definida como:

$$r \cup s = \{t \mid t \in r \text{ ou } t \in s\}$$

■ Para $r \cup s$ ser válida:

1. r, s devem ter a **mesma aridade** (igual número de atributos)
2. Os **domínios** dos atributos devem ser **compatíveis** (e.g., os valores da 2ª coluna de r são do mesmo tipo dos valores da 2ª coluna de s)

■ E.g. determinar quais os clientes que têm uma conta ou um empréstimo

$$\Pi_{customer-name}(depositor) \cup \Pi_{customer-name}(borrower)$$

Operação de União – Exemplo

■ Relações r , s :

r

A	B
α	1
α	2
β	1

s

A	B
α	2
β	3

■ $r \cup s$:

$r \cup s$

A	B
α	1
α	2
β	1
β	3

Propriedade da União

■ Propriedades

- ✱ Comutatividade: $r \cup s = s \cup r$
- ✱ Associatividade: $r \cup (s \cup t) = (r \cup s) \cup t$
- ✱ Distributividade sobre seleção: $\sigma_P(r \cup s) = \sigma_P(r) \cup \sigma_P(s)$
- ✱ Distributividade sobre projeção: $\Pi_L(r \cup s) = \Pi_L(r) \cup \Pi_L(s)$

Operação de Diferença de Conjuntos

- Notação: $r - s$

- Definida como:

$$r - s = \{t \mid t \in r \text{ e } t \notin s\}$$

- As diferenças de conjuntos só podem ser efectuadas entre relações compatíveis.

 - ✳ r e s devem ter a **mesma aridade**

 - ✳ os **domínios** dos atributos de r e s devem ser **compatíveis**

- Propriedades

 - ✳ Distributividade sobre seleção: $\sigma_P(r - s) = \sigma_P(r) - \sigma_P(s)$

Operação de Diferença de Conjuntos-Ex.

■ Relações r, s :

r

A	B
α	1
α	2
β	1

s

A	B
α	2
β	3

■ $r - s$:

r - s

A	B
α	1
β	1

Operação de Produto Cartesiano

- Notação: $r \times s$

- Definida como:

$$r \times s = \{tq \mid t \in r \text{ e } q \in s\}$$

- Assume que os atributos de $r(R)$ e $s(S)$ são disjuntos. (Ou seja, $R \cap S = \emptyset$).

- Se os atributos de $r(R)$ e $s(S)$ não são disjuntos, então têm que se utilizar renomeações:

- ★ Se o nome das relações for diferente, distinguem-se os atributos com o mesmo nome prefixando-os com o nome da relação.
- ★ Se não for possível usar o nome da relação para desambiguar os atributos com o mesmo nome (e.g. as relações têm o mesmo nome, ou o nome é ambíguo por resultar de operações compostas), veremos mais à frente...

Operação de Produto Cartesiano-Ex.

■ Relações **r** e **s**:

r

A	B
α	1
β	2

s

C	D	E
α	10	a
α	13	a
β	20	b
γ	10	b

■ **r x s**:

r x s

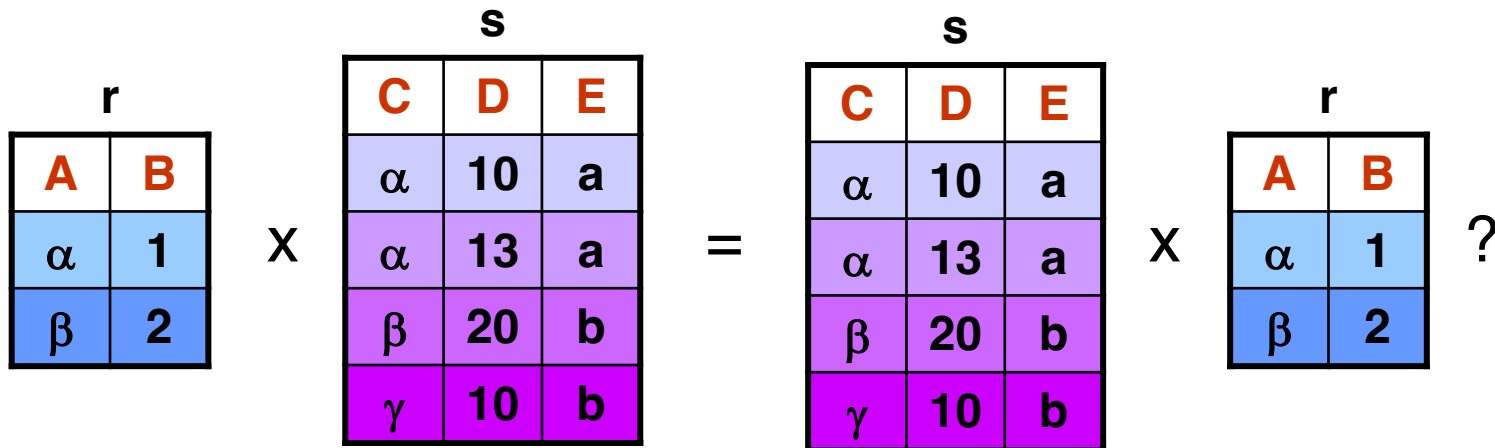
A	B	C	D	E
α	1	α	10	a
α	1	α	13	a
α	1	β	20	b
α	1	γ	10	b
β	2	α	10	a
β	2	α	13	a
β	2	β	20	b
β	2	γ	10	b

Propriedades do Produto Cartesiano

■ Propriedades

✦ Associatividade: $r \times (s \times t) = (r \times s) \times t$

✦ Comutatividade: $r \times s = s \times r$?



Em rigor não!

Composição de Operações

- Pode-se construir expressões combinando várias operações
- Exemplo: $\sigma_{A=C}(r \times s)$

r

A	B
α	1
β	2

s

C	D	E
α	10	a
α	13	a
β	20	b
γ	10	b

r x s

A	B	C	D	E
α	1	α	10	a
α	1	α	13	a
α	1	β	20	b
α	1	γ	10	b
β	2	α	10	a
β	2	α	13	a
β	2	β	20	b
β	2	γ	10	b

$\sigma_{A=C}(r \times s)$

A	B	C	D	E
α	1	α	10	a
α	1	α	13	a
β	2	β	20	b

Operação de Renomeação

- Permite dar um nome aos resultados de expressões de álgebra relacional.
- Permite que uma relação seja referida por mais de um nome.
- A expressão:

$$\rho_X(E)$$

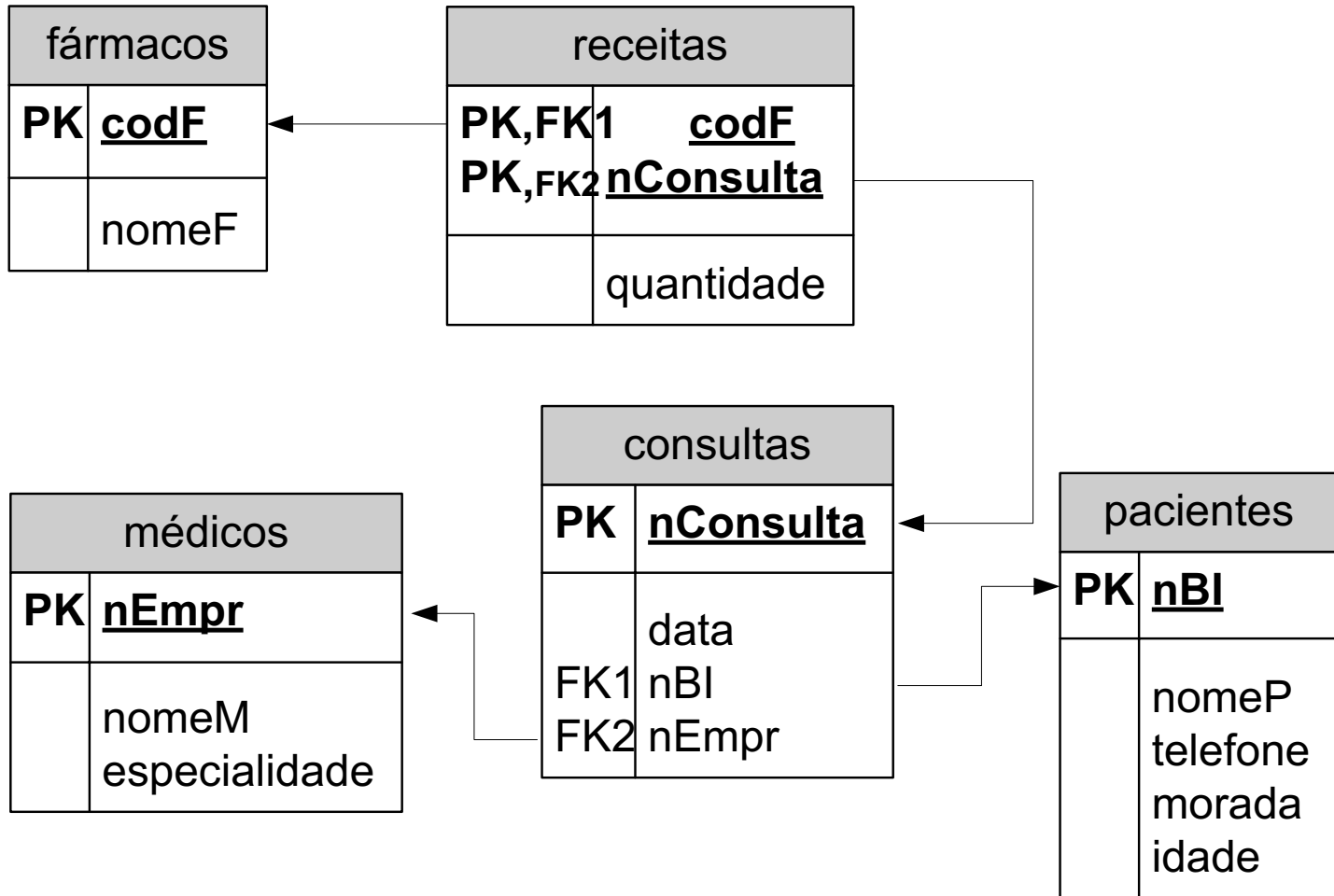
devolve a expressão E com o nome X

- Se uma expressão de álgebra relacional E tem aridade n , então

$$\rho_{X(A1, A2, \dots, An)}(E)$$

devolve a expressão E com o nome X , e com os atributos renomeados para $A1, A2, \dots, An$.

Exemplo Clínica



Exemplo Clínica

- *médicos(nEmpr,nomeM,especialidade)*
- *pacientes(nBl,nomeP,telefone,morada,idade)*
- *fármacos(codF,nomeF)*
- *consultas(nConsulta,data,nBl,nEmpr)*
- *receitas(codF,nConsulta,quantidade)*

Exemplo Clínica

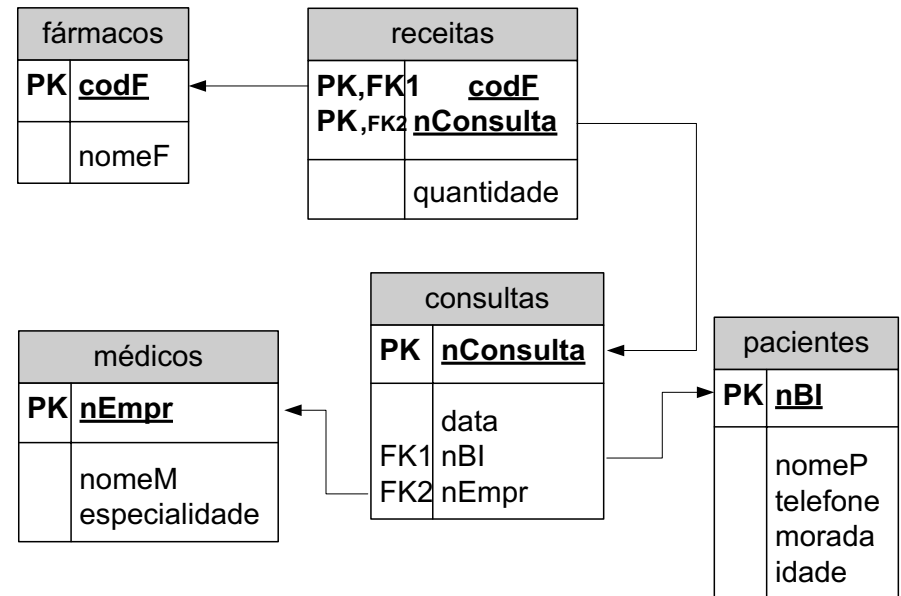
médicos(nEmpr, nomeM, especialidade)

fármacos(codF, nomeF)

pacientes(nBI, nomeP, telefone, morada, idade)

consultas(nConsulta, data, nBI, nEmpr)

receitas(codF, nConsulta, quantidade)



- Quais os pacientes com mais de 50 anos de idade?

$$\sigma_{idade > 50}(pacientes)$$

- Quais os nomes dos pacientes com mais de 50 anos de idade?

$$\Pi_{nomeP}(\sigma_{idade > 50}(pacientes))$$

Exemplo Clínica

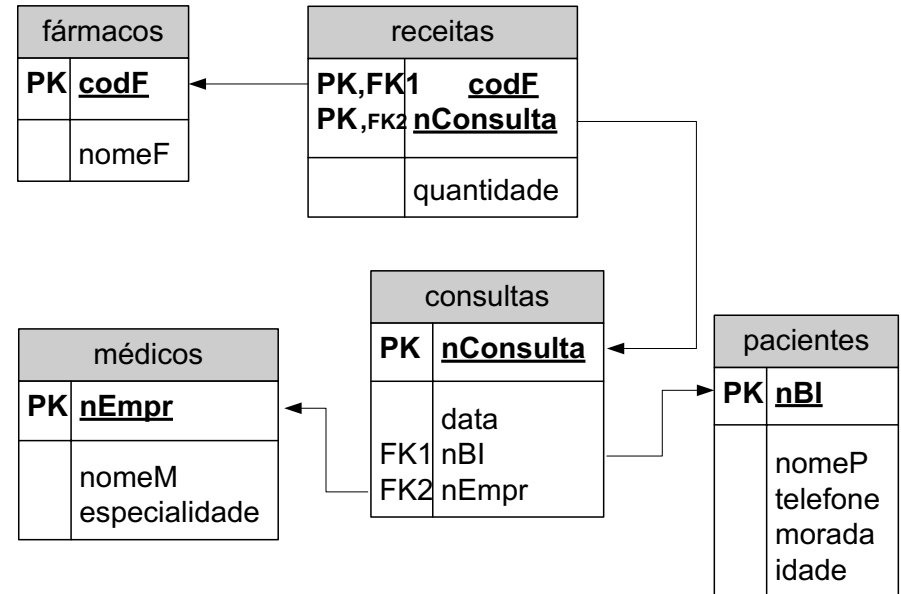
médicos(nEmpr, nomeM, especialidade)

fármacos(codF, nomeF)

pacientes(nBI, nomeP, telefone, morada, idade)

consultas(nConsulta, data, nBI, nEmpr)

receitas(codF, nConsulta, quantidade)



- Quais os nomes dos fármacos que já foram receitados em consultas da clínica?

$$\Pi_{nomeF}((\sigma_{receitas.codF = fármacos.codF}(receitas \times fármacos)))$$

Exemplo Clínica

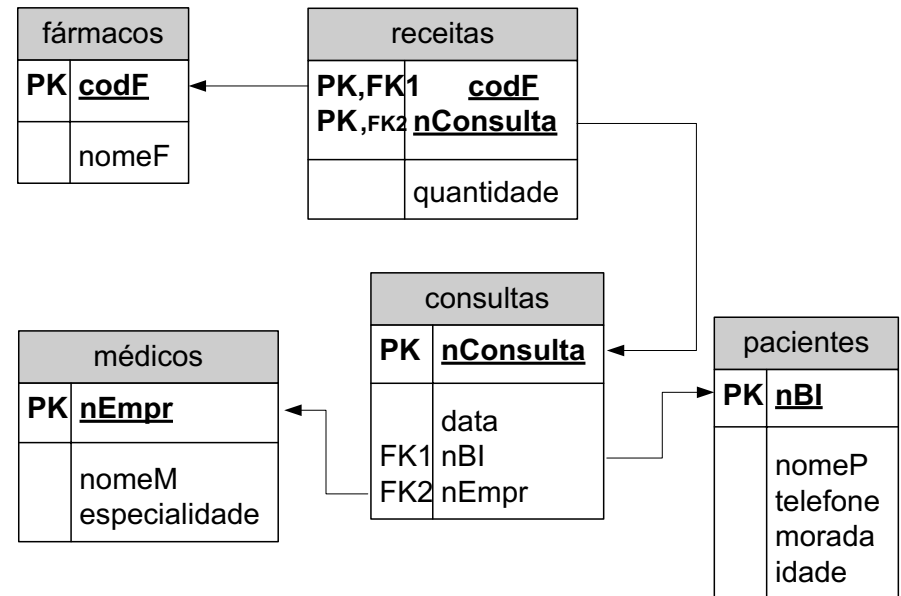
médicos(nEmpr, nomeM, especialidade)

fármacos(codF, nomeF)

pacientes(nBI, nomeP, telefone, morada, idade)

consultas(nConsulta, data, nBI, nEmpr)

receitas(codF, nConsulta, quantidade)



- Quais os nomes e códigos dos fármacos que nunca foram receitados?

fármacos –

$$\Pi_{receitas.codF, nomeF}(\sigma_{receitas.codF \neq fármacos.codF}(receitas \times fármacos)))$$

Exemplo Clínica

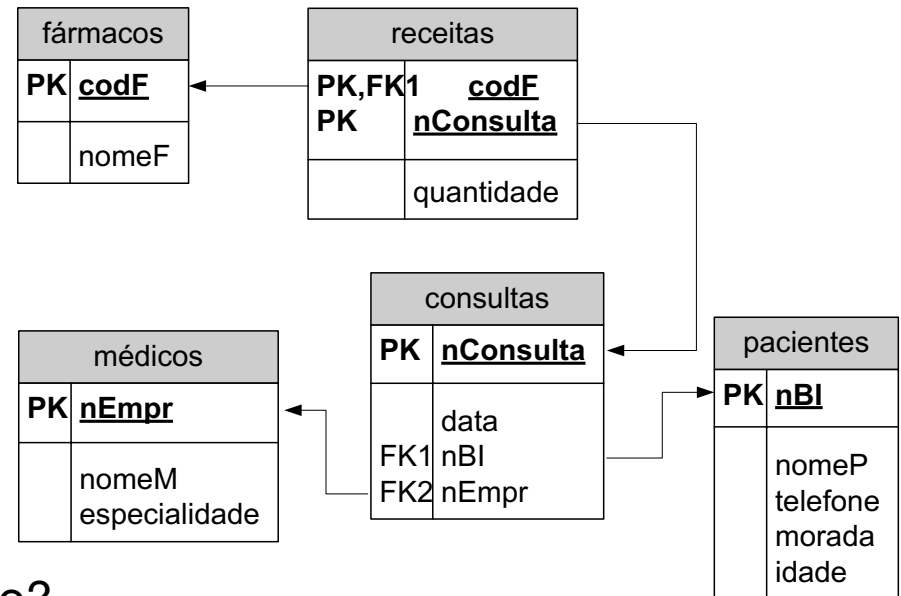
médicos(*nEmpr*, *nomeM*, *especialidade*)

fármacos(*codF*, *nomeF*)

pacientes(*nBI*, *nomeP*, *telefone*, *morada*, *idade*)

consultas(*nConsulta*, *data*, *nBI*, *nEmpr*)

receitas(*codF*, *nConsulta*, *quantidade*)



■ Qual a idade do paciente mais velho?

✳ Renomear a relação *pacientes* como *d*

✳ A consulta é:

$$\Pi_{idade}(pacientes) - \Pi_{pacientes.idade}(\sigma_{pacientes.idade < d.idade} (pacientes \times \rho_d (pacientes)))$$

Exemplo Clínica

pacientes

Nome	Idade
Ana	30
Rui	20
Carla	25

$\rho_d(\text{pacientes})$

d.Nome	d.Idade
Ana	30
Rui	20
Carla	25

pacientes x $\rho_d(\text{pacientes})$

Nome	Idade	d.Nome	d.Idade
Ana	30	Ana	30
Ana	30	Rui	20
Ana	30	Carla	25
Rui	20	Ana	30
Rui	20	Rui	20
Rui	20	Carla	25
Carla	25	Ana	30
Carla	25	Rui	20
Carla	25	Carla	25

$\sigma_{\text{pacientes.Idade} < \text{d.Idade}}$
(pacientes x $\rho_d(\text{pacientes})$)

Nome	Idade	d.Nome	d.Idade
Rui	20	Ana	30
Rui	20	Carla	25
Carla	25	Ana	30

$\Pi_{\text{pacientes.Idade}(\sigma_{\text{pacientes.Idade} < \text{d.Idade}}(\text{pacientes x } \rho_d(\text{pacientes})))}$

Idade
20
25

$\Pi_{\text{Idade}}(\text{pacientes})$

Idade
30
20
25

$\Pi_{\text{Idade}}(\text{pacientes}) -$

$\Pi_{\text{pacientes.Idade}(\sigma_{\text{pacientes.Idade} < \text{d.Idade}}(\text{pacientes x } \rho_d(\text{pacientes})))}$

\Rightarrow

Idade
30

Exemplo Clínica

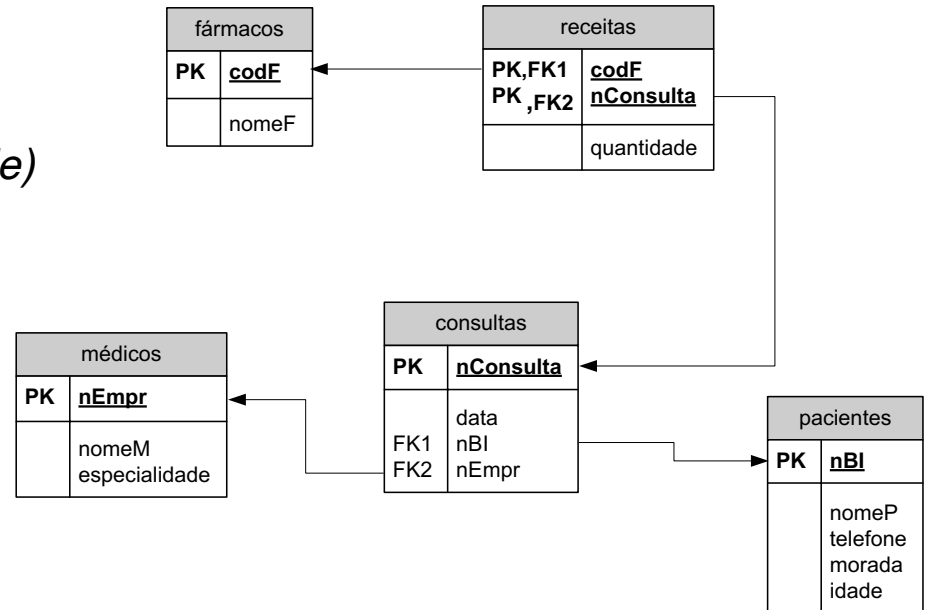
médicos(nEmpr, nomeM, especialidade)

pacientes(nBI, nomeP, telefone, morada, idade)

fármacos(codF, nomeF)

consultas(nConsulta, data, nBI, nEmpr)

receitas(codF, nConsulta, quantidade)



■ E quais os (nomes dos) pacientes com essa idade?

★ Seja r a relação da pergunta anterior:

$$\Pi_{nomeP}(\sigma_{pacientes.idade = r.idade}(pacientes \times r))$$

Definição Formal

- Uma expressão básica na álgebra relacional é:
 - ✱ Uma relação na base de dados
 - ✱ Uma relação constante
- Sejam E_1 e E_2 expressões de álgebra relacional; então todas as expressões abaixo são expressões de álgebra relacional:
 - ✱ $E_1 \cup E_2$
 - ✱ $E_1 - E_2$
 - ✱ $E_1 \times E_2$
 - ✱ $\sigma_P(E_1)$, P é um predicado nos atributos de E_1
 - ✱ $\Pi_S(E_1)$, S é uma lista com alguns dos atributos de E_1
 - ✱ $\rho_X(E_1)$, x é um novo nome para o resultado de E_1

Operações Adicionais

- Definem-se outras operações que não aumentam o poder expressivo da álgebra relacional, mas simplificam algumas consultas habituais.
 - ✧ Interseção de conjuntos
 - ✧ Junção Natural
 - ✧ Divisão
 - ✧ Atribuição

Operação de Interseção de Conjuntos

- Notação:

$$r \cap s$$

- Definida por:

$$r \cap s = \{ t \mid t \in r \text{ e } t \in s \}$$

- Assume:

- ★ r, s têm a mesma aridade
- ★ Os atributos de r e s são compatíveis

- Notar que:

$$r \cap s = r - (r - s) = s - (s - r)$$

Interseção de Conjuntos - Exemplo

■ Relações r , s :

r

A	B
α	1
α	2
β	1

s

A	B
α	2
β	3

■ $r \cap s$

$r \cap s$

A	B
α	2

Operação de Junção Natural

- Notação: $r \bowtie s$
- Sejam r e s relações nos esquemas R e S , respetivamente. O resultado é uma relação no esquema $R \cup S$ que é obtido considerando cada par de tuplos t_r de r e t_s de s .
- Se t_r e t_s têm o mesmo valor em cada um dos atributos em $R \cap S$, um tuplo t é adicionado ao resultado, em que
 - ★ t tem os mesmos valores que t_r em r
 - ★ t tem os mesmos valores que t_s em s
- Exemplo: $R = (A, B, C, D)$ $S = (E, B, D)$
 - ★ Esquema resultado: (A, B, C, D, E)
 - ★ $r \bowtie s$ define-se por:

$$\Pi_{r.B, r.D, r.A, r.C, s.E}(\sigma_{r.B = s.B \wedge r.D = s.D}(r \times s))$$

Junção Natural – Exemplo

■ Relações r, s:

r

A	B	C	D
α	1	α	a
β	2	γ	a
γ	4	β	b
α	1	γ	a
δ	2	β	b

s

B	D	E
1	a	α
3	a	β
1	a	γ
2	b	δ
3	b	ε

■ $r \bowtie s$

A	B	C	D	E
α	1	α	a	α
α	1	α	a	γ
α	1	γ	a	α
α	1	γ	a	γ
δ	2	β	b	δ



$r \bowtie s$

B	D	A	C	E
1	a	α	α	α
1	a	α	α	γ
1	a	α	γ	α
1	a	α	γ	γ
2	b	δ	β	δ

Exemplo Clínica

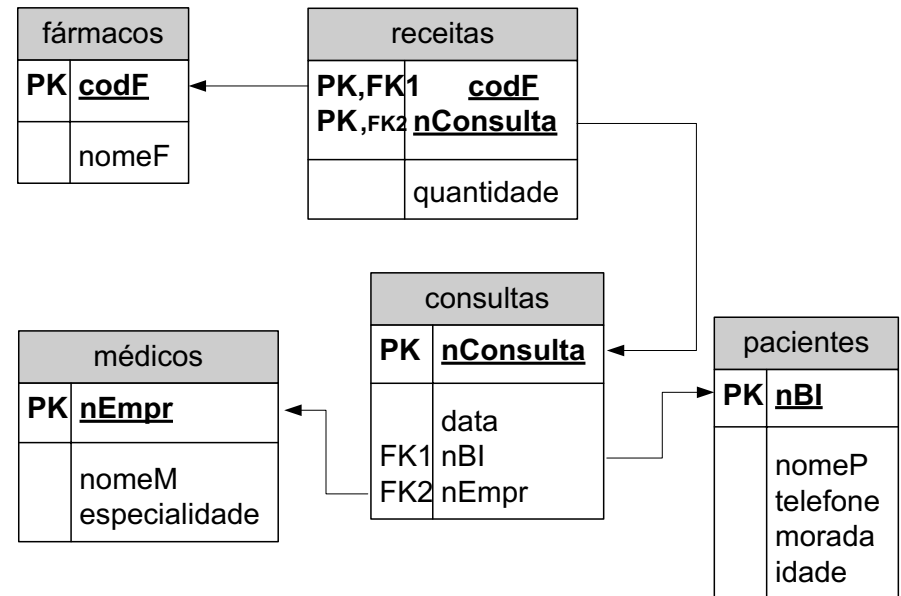
médicos(nEmpr, nomeM, especialidade)

fármacos(codF, nomeF)

pacientes(nBl, nomeP, telefone, morada, idade)

consultas(nConsulta, data, nBl, nEmpr)

receitas(codF, nConsulta, quantidade)



- Quais os (nomes dos) fármacos alguma vez prescritos por cardiologistas?

$\Pi_{nomeF} (\text{fármacos} \bowtie \text{receitas} \bowtie \text{consultas} \bowtie \sigma_{esp = \text{"car..."}} (\text{médicos}))$

ou

$\Pi_{nomeF} (\sigma_{esp = \text{"car..."}} (\text{fármacos} \bowtie \text{receitas} \bowtie \text{consultas} \bowtie \text{médicos}))$