



LABORATÓRIO DE BANCO DE DADOS

IBD100

ÁLGEBRA RELACIONAL

- A Necessidade de especificar a Álgebra Relacional
- A Solução e Conceitos fundamentais
- O SQL e a evolução 'forçada' da AR: Novos Símbolos [simplificados]
- Operações
 - Seleção / Projeção
 - Produto Cartesiano / Junção (natural)
 - União / Subtração / Intersecção
 - Requisitos e implicações no processamento
- Exercícios
- Operações de Ordenação e Agrupamento de Dados
- Usando funções de processamento [contagem, totalização, máximos, mínimos, médias]
- Procedimento da Divisão de tabelas
- Revisão da junção [Natural] e definindo
 - Junção Aberta pela direita
 - Junção Aberta pela esquerda
 - Junção Aberta pelos dois lados ou Completa
- Exercícios

- O início da necessidade de TI
 - As organizações sociais, na sua dinâmica de desenvolvimento, necessitam de dispositivos de coleta, armazenamento, processamento e distribuição de dados adequados para o suporte à tomada de decisão.
 - O desenvolvimento das ciências propiciaram a oferta de soluções computacionais que foram aceleradas durante e após a segunda metade do séc. 20.
 - Os computadores projetados desde o início do séc. 20 se baseiam na arquitetura de uma CPU processando os dados e instruções acessados em mecanismos de armazenamento de dados (discos rígidos e memórias de acesso randômico (RAM)).
 - Por conta de uma abordagem apresentada pelos pequenos desenvolvimentos, a estrutura de implementação e execução de Sistemas de Informação se dividiu em duas partes fundamentais para a mesma finalidade de um lado os **processos** e de outro os **dados** manipulados por estes processos.

- O início da necessidade de TI
 - Durante as décadas de 50 até 90 os desenvolvimentos em torno dos PROCESSOS resultaram nas linguagens de programação (compiladas ou interpretadas). A aplicação das linguagens computacionais permitiu, então, a estruturação dos Programas Aplicativos (PA).
 - Em 1970 Edgar F. Codd apresentou um artigo "A Relational Model of Data for Large Shared Data Banks", fundamentando o Modelo Relacional (MR) como forma de organização e acesso a dados.
 - O MR se baseia no conceito matemático da Teoria dos Conjuntos e apresenta um formalismo matemático para processamento das linhas das tabelas na execução das operações referentes às operações da Teoria dos Conjuntos.
 - São apresentadas em AR as operações fundamentais: Seleção(σ), Projeção(π), União($\cup$), Diferença($-$) [ou subtração], Produto Cartesiano($\times$) e Junção($\bowtie$).

Note que a Intersecção não figura como operação fundamental, pois pode ser derivada de subtrações sucessivas. Sendo A e B passíveis de se fazer União, pode-se escrever a intersecção na forma: $A - (A - B) = A \cap B$.

- O início da necessidade de TI
 - Abordam-se aqui, as definições mais detalhadas de cada operação fundamental e uma exemplificação de sua operação.
 - Ainda no artigo de E. F. Codd se apresenta a especificação do conceito matemático de conjuntos união compatíveis 'traduzidos' para o contexto de tabelas.
 - Dadas duas tabelas A e B aderentes ao MR (sem linhas duplicadas), diz-se que A e B são União Compatíveis se e somente se:
 - Possuem o mesmo número de colunas (campos).
 - As colunas correspondentes possuem o mesmo domínio (tipo de dados).
 - Tendo, então, como referência estes conceitos, segue-se a apresentação da Simbologia das operações.
 - É frequente que o resultado do processamento de uma operação seja direcionado para uma tabela para ser posteriormente processado com outra operação.

- O início da necessidade de TI
 - É frequente que o resultado do processamento de uma operação seja direcionado para uma tabela para ser posteriormente processado com outra operação.
 - Nesta situação diz-se que a tabela processada inicialmente passa a ser a ORIGEM e que a tabela para onde se direciona o resultado do processamento passa a ser o DESTINO do processamento.
 - Na simbologia da AR veremos que o símbolo que atribui o resultado de um processamento para uma outra tabela é simplesmente o sinal de igual '='.
 - Portanto , chega-se ao momento de apresentação da simbologia proposta para AR.

- Simbologia da AR

<u>Símbolo</u>	<u>Significado</u>
Destino = Origem	Significa a criação de uma Tabela, processando dados na Origem e criando os dados no Destino;
[[c1, c2,...]]	Nomes de campos.
[[condição]]	Condição de comparação entre campos de tabelas ou de campos com valores.
][condição]]	Condição de comparação entre campos de tabelas com lista completa pela ESQUERDA.
[[condição][	Condição de comparação entre campos de tabelas com lista completa pela DIREITA.
][condição][	Condição de comparação entre campos de tabelas com lista completa pelos dois lados – Junção completa.
→	Muda o nome de um campo ou uma tabela dinamicamente. Pode também direcionar o resultado e um função para um campo ou variável.
()	Contém a especificação de uma função ou pode conter comandos encadeados.
{ }	Contém a especificação de agrupamento ou ordenação.

Agora, seguem-se as especificações e exemplos das operações da AR.

• Seleção

- Cria uma lista de linhas de uma tabela que atendam a uma CONDIÇÃO. A lista criada na Origem pode ser direcionada para uma tabela de Destino dos dados processados.

- Nota: A Estrutura (Esquema) da tabela no Destino é igual á Estrutura da tabela de Origem.

- Em AR: A[[Condição]]

- Exemplo: A [[C2 contenha w]]

- Esta lista pode ser direcionada para uma tabela R na forma:

- R = A [[C2 contenha w]]

R

C1	C2
Asd	Qwe
Fgh	Erw

A

C1	C2
Asd	Qwe
Rfv	Tgb
Zxc	Rty
Fgh	Erw
Wsx	Edc

C2 tem W

C2 NÃO
tem W

- Observações:

- O Resultado da operação pode ser uma tabela SEM TUPLAS (o que é o conceito de **conjunto vazio**).

- A [[condição]] aceita qualquer tipo de operação, dependendo do SGBD aceitar ou não estas operações;

- A [[condição]] pode processar todos os operadores de comparação matematicamente aceitos, tais como:

= , ≥ ou >= , ≤ ou <= , > , < , <> ou != ou ≠ , E, OU, OU exclusivo e NÃO (operadores lógicos).

- **Projeção**

- Cria uma lista de linhas copiando os dados de um ou mais campos de uma tabela [Origem dos dados].
- Esquema da tabela Destino é igual à lista de campos da tabela de Origem da operação.

- Em AR: A[[Lista de Campos]]

- Exemplo (destinando para R):

$$R = A[[C2,C4]]$$

- Note a repetição de linhas em R

- Se for preciso mostrar a lista SEM repetição, escreve-se:

- $R = A[[C2,C4]](\text{distintos})$

- Observações:

- TODOS os Valores dos campos escolhidos

são copiados para o Destino

(Destino e Origem com a mesma quantidade de linhas).

- O Resultado da operação pode ser uma tabela SEM TUPLAS (o que seria o conceito de conjunto vazio).

- Se a tabela Destino apresentar tuplas repetidas podemos complementar o comando com (distintos)

- $R = A[[C1,C4]](\text{distintos})$

- Na projeção é possível trocar os nomes dos campos, para isso usa-se o simbolo $\rightarrow$ para indicar a troca, desse modo:

- $R = A[[C1 \rightarrow \text{casa}, C2 \rightarrow \text{tel}]]$ produz uma tabela R com campos (casa,tel).

R	C2	C4	A	C1	C2	C3	C4
	Qwe	Dfg		Asd	Qwe	Fr4	Dfg
	Tgb	Edc		Rfv	Tgb	Gt5	Edc
	Rty	Rfv		Zxc	Rty	Hy6	Rfv
	Ert	Tgb		Fgh	Ert	Ju7	Tgb
	Rty	Rfv		Wsx	Rty	Ki8	Rfv

- Produto Cartesiano

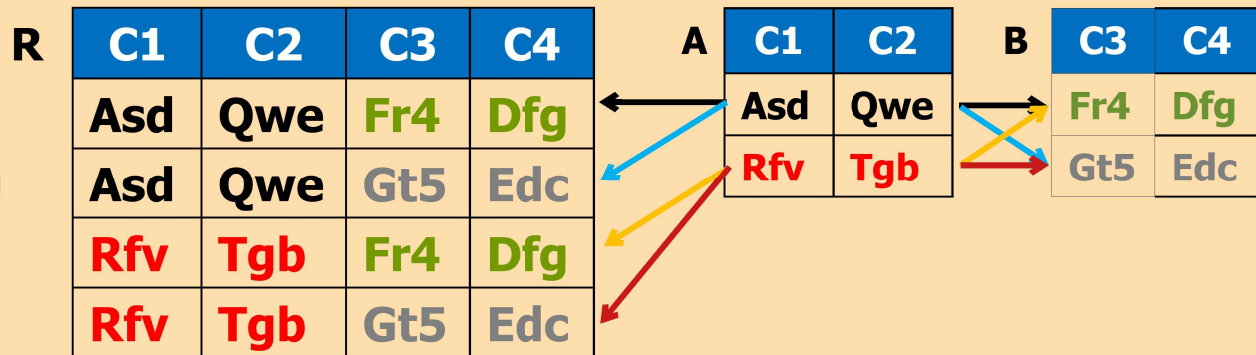
- Cria uma lista de linhas tendo duas tabelas na Origem e colocando cada linha da 1ª tabela ao lado de cada linha da 2ª tabela. Isto se diz "emparelhamento de linhas ou tuplas";
- Esquema da tabela criada é igual aos esquemas das tabelas da Origem colocados lado a lado, obedecendo, inclusive, a ordem de campos e pode ser direcionada para uma tabela Destino.

- Em AR: $A \times B$

- Por Exemplo: $R = A \times B$

- Observações

- A Quantidade de linhas da tabela Destino será igual à quantidade de linhas da 1ª tabela multiplicada pela quantidade de linhas da 2ª tabela [da origem].

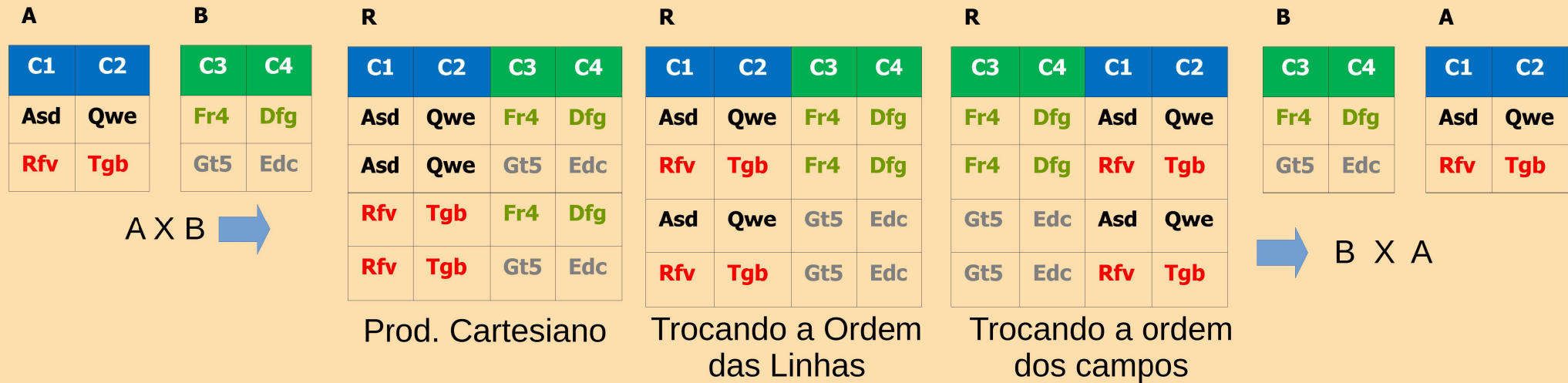


- Produto Cartesiano

- Observações

- É uma operação comutativa. Demonstra-se:

- $A \times B = B \times A$



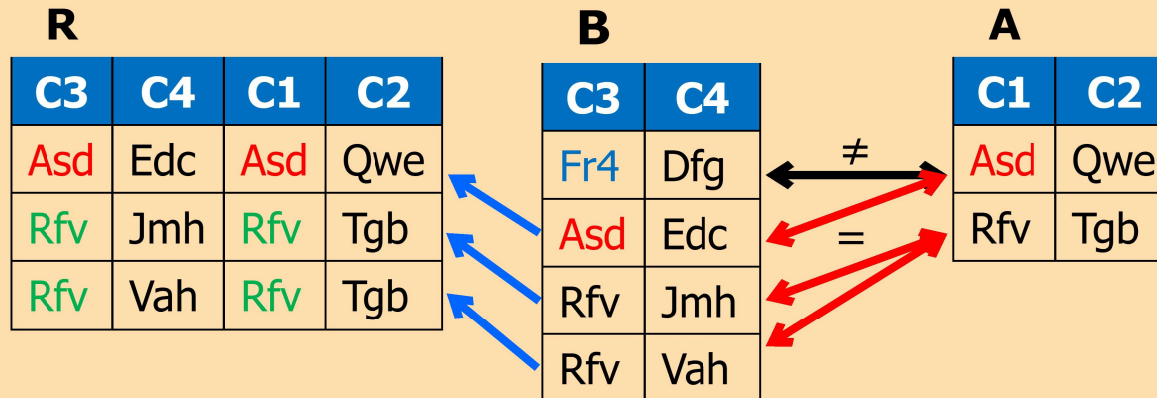
- **Junção [Natural]**

- Cria uma lista de linhas partindo de duas tabelas e emparelhando as linhas obedecendo a condição.
- As linhas geradas podem ser direcionadas para uma tabela Destino.
- Do mesmo modo como Prod. Cartesiano, o Esquema da tabela Destino é igual aos esquemas das tabelas da Origem colocados lado a lado, obedecendo, inclusive a ordem de campos.
- O esquema gerado das tabelas Origem pode apresentar campos com nomes repetidos. Como veremos a maioria dos SGBDs permitem isso mas não permitem a criação da Tabela Destino com os campos repetidos. Veremos como é possível resolver este paradoxo, caso ocorra.
- Na condição da junção, na maioria das vezes, se compara o valor de uma Chave Primária de uma tabela com a correspondente Chave estrangeira presente na outra tabela (ambas na Origem). Por enquanto vamos aceitar esse contexto, mas Peter Chen questionou justamente isso na definição do conceito de relacionamento.

- Junção [Natural]

- Em AR: B [[Condição]] A

- Por exemplo: $R = B \text{ } [[B.C3=A.C1]] \text{ } A$ ou $R = B \text{ } [[C3=C1]] \text{ } A$



- Observação: A tabela Destino pode ser um conjunto vazio. Basta a Condição da Junção não ser atendida por nenhuma das tuplas, entretanto a tabela Destino será criada do mesmo modo.

- União

- Cria uma lista de dados com as linhas comuns e não comuns de duas tabelas da Origem.
- A Operação pode ser feita se e somente se as tabelas de origem forem União Compatíveis;
- O Esquema da tabela resultado é igual ao esquema de cada uma das tabelas.

- Em AR: $R = A \cup B$

R	C1	C2
	Zxc	Rty
	Fgh	Ert
	Rfv	Tgb
	Asd	Qwe
	Rfv	Tgb

A	C1	C2
	Asd	Qwe
	Rfv	Tgb

B	C1	C2
	Zxc	Rty
	Fgh	Ert
	Rfv	Tgb

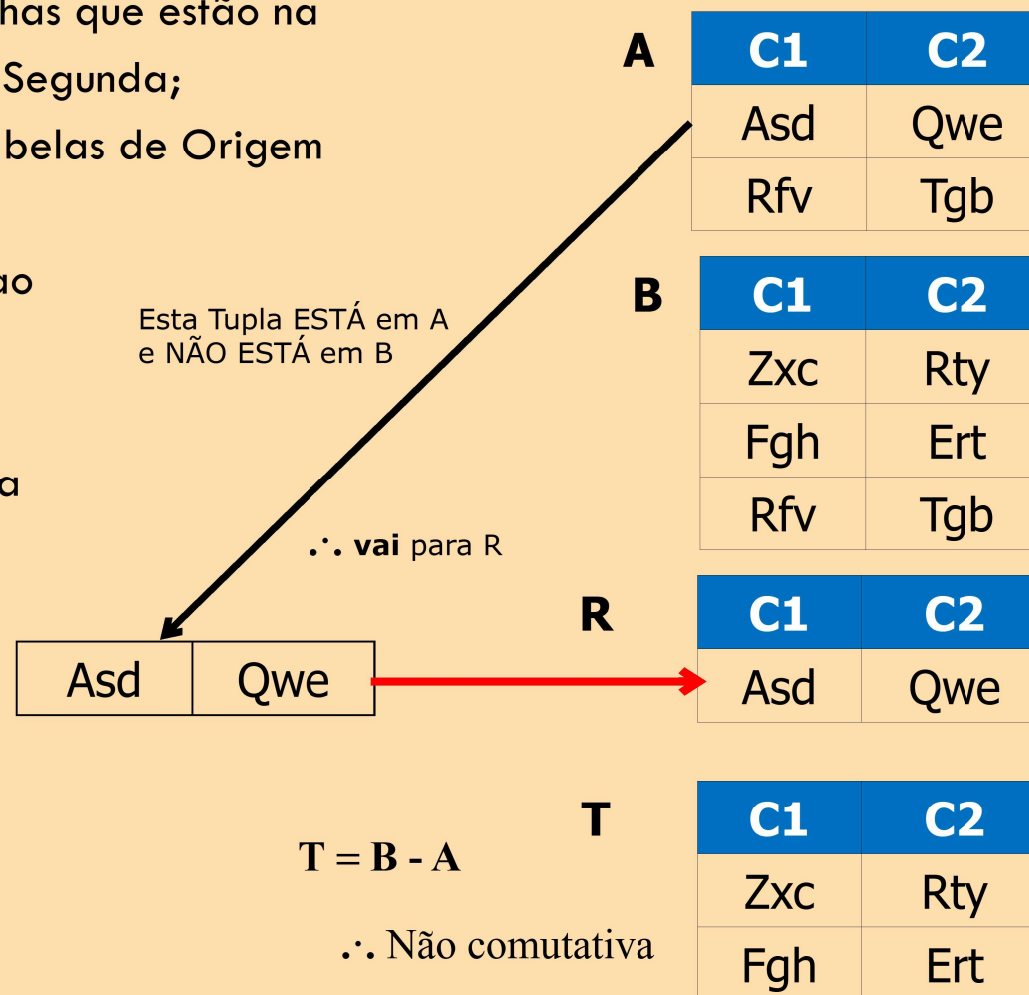
- Esta repetição é 'correta'.

Na implementação as linguagens colocam operadores que omitem a repetição.

- Observações:
- A União é comutativa (a ordem dos fatores na operação não altera o resultado da operação);
 $A \cup B = B \cup A$
- As Tabelas geradas por operações de Álgebra Relacional são Livres podem ter linhas repetidas;
- Podemos escrever complementos aos comandos da Álgebra Relacional para tratamento de prováveis desvios do Modelo. A Operação acima pode ser escrita assim: **$R = A \cup B$ (Distintos)**

- Diferença (subtração)

- Cria uma lista de dados levando as linhas que estão na Primeira Tabela e que NÃO estão na Segunda;
- A Operação só pode ser feita se as tabelas de Origem forem União Compatíveis;
- Esquema da tabela resultado é igual ao esquema de cada uma das tabelas.
- EM AR: $A - B$
- O resultado pode ser direcionado para uma tabela no Destino na forma:
 $R = A - B$
- O Resultado da operação pode ser uma tabela SEM TUPLAS (conceito de conjunto vazio)



- Intersecção

- A AR, como proposta inicialmente, não apresentava a operação de intersecção. Embora presente na teoria dos Conjuntos, na AR a intersecção pode ser uma derivação de outras operações.
- Note: $A - (A - B)$ produz a intersecção de A e B.
- Entretanto, como o MR foi desde o início desenvolvido sem a pretensão de geração de propriedade intelectual, logo ela recebeu propostas de evolução de conceitos e operações.
- Desse modo, define-se: Intersecção: Gera uma lista de linhas com as linhas comuns entre duas tabelas na origem.
- A Operação só pode ser feita se as tabelas de origem forem *União Compatíveis*;
- Esquema da tabela destino é igual ao esquema de cada uma das tabelas [Origem].
- EM AR: $A \cap B$
- Observações:

- O símbolo $\cap$ pode ser difícil de ser escrito em alguns editores de texto.

Por conta disso existe o Símbolo alternativo '^', assim:

$A \wedge B$ seria o modo 'mais fácil' de escrever $A \cap B$.

R	C1	C2
	Rfv	Tgb

A	C1	C2
	Asd	Qwe
	Rfv	Tgb

B	C1	C2
	Zxc	Rty
	Fgh	Ert
	Rfv	Tgb

- Exercício

- Tendo agora as tabelas de uma base de dados podemos resolver perguntas que exijam o processamento de linhas das tabelas. O processamento destas tabelas é referenciado como "recuperação de informação", embora seja preferencial citar "recuperação de dados".
- Dadas as tabelas abaixo responda à pergunta: Quais os dias que os clientes com nome José estacionaram um de seus carros?

Clientes

CdCli	Nome
A	José
...	...

Carros

CdCli	CdCarro	Marca
A	F1	Gol
.....		

Paradas

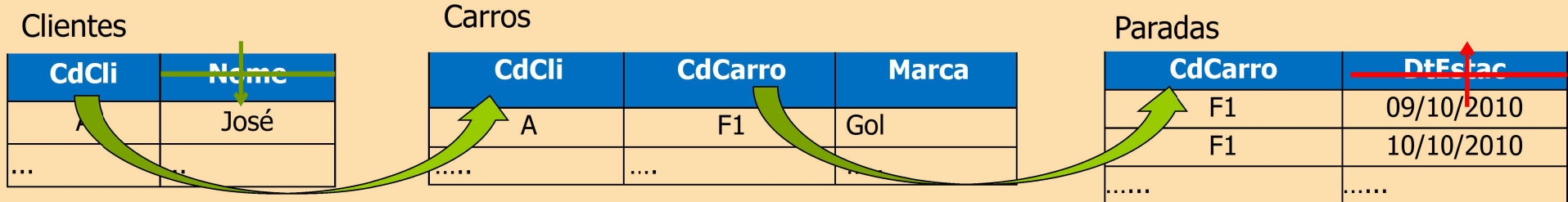
CdCarro	DtEstac
F1	09/10/2010
F1	10/10/2010
.....	

- Note: Os dados para responder esta pergunta estão espalhados pelas tabelas. Será necessário pesquisar (seleção), combinar (junção) e separar colunas (projetar). Mas... quais tabelas devem ser envolvidas? Em ordem executar as operações? Não existe uma só resposta. Quase sempre acontece **mais de uma** sequência de resposta que pode atender a uma recuperação de dados. É boa pratica então 'mapear a pergunta' indicando com uma seta para baixo argumentos de seleção e com uma seta para cima argumentos (ou campos de projeção). Na forma...

- Exercício

- Mapeando a pergunta: ~~Quais os dias~~ que os ~~clientes com nome José~~ estacionaram um de seus carros?
- Agora analisando o modelo... como conseguimos 'relacionar' os dados necessários no BD?

Começamos a colocar o mapeamento da pergunta no modelo.



- Agora... Colocamos no modelo os mapeamentos de 'junções' que podem ser feitas entre as linhas das tabelas. Assim...
- A seguir marcamos o mapeamento da pergunta no modelo. Assim...
- Isso nos ajuda a determinar uma sequência de operações que podemos fazer para responder esta pergunta...

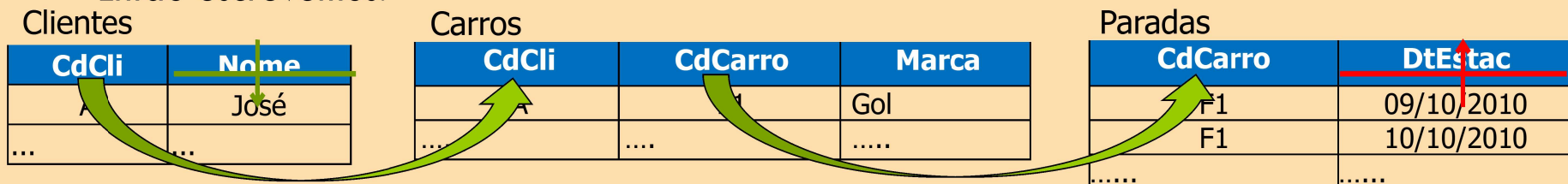
Seleção, depois uma junção, a seguir uma outra junção e finalmente uma projeção.

Seleção em clientes com nome "José", depois uma junção (com o cdcli), a seguir uma outra junção (com o cdcarro) e finalmente uma projeção (tirando o dtestac).

- Exercício

- Sugestão fortemente recomendada (pelo menos no início de pratica dos exercícios): CONSTRUA a estrutura das tabelas DESTINO abaixo de cada operação da AR. Isso facilita a alteração de operações que talvez sejam necessárias.

- Então escrevemos:



- Seleção.: $W = \text{Clientes } [[\text{Nome} = \text{"José"}]]$
W

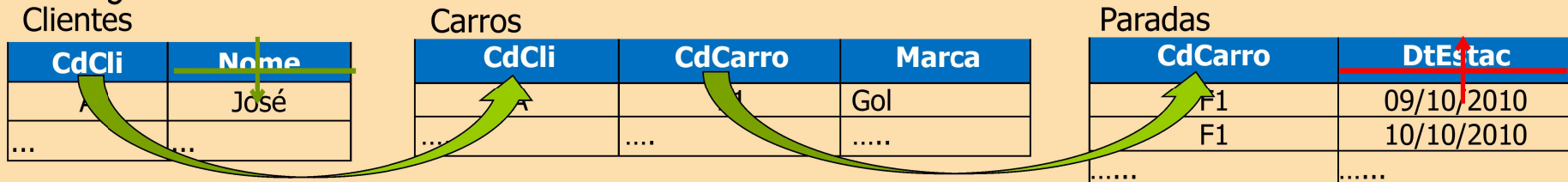
CdCli	Nome
-------	------

- Junção...: $T = W [[W.cdcli = \text{carros.cdcli}]]$ Carros
W

CdCli	Nome	cdCli	cdCarro	Marca
-------	------	-------	---------	-------

- Exercício

- Seguindo com a AR...



- Seleção.: $W = \text{Clientes} \ [[\text{Nome} = "José"]]$

CdCli	Nome
-------	------

- Junção...: $T = W \ [[W.cdcli = \text{carros}.cdcli]] \text{ Carros}$

CdCli	Nome	cdCli	cdCarro	Marca
-------	------	-------	---------	-------

- Junção...: $K = T \ [[T.cdcarro = \text{Paradas}.cdcarro]] \text{ Paradas}$

CdCli	Nome	cdCli	cdCarro	Marca	CdCarro	DtEstac
-------	------	-------	---------	-------	---------	---------

- Alcançamos o campo 'dtEstac' e podemos fazer a
Projeção: $K \ [[dtEstac]] \text{ (distintos)}$

DtEstac

• Exercício

- Agora é a sua vez. Resolva este exercício e grave sua resposta em um arquivo em formato TXT com nome: SeuNomeCompleto-AR-EX01.txt e envie-o pela AR-tar01 do Teams.
- Dadas as tabelas abaixo responda à pergunta: Quais as marcas dos carros cujos os donos tenham nome "Jose" e que não fizeram serviço de "pintura" no mês de "Maio de 2010"?

Clientes

CdCli	Nome
A	José
B	João
C	Ana
D	José
E	Fábio
...	...

Carros

CdCli	CdCarro	Marca
A	F1	Gol
B	F2	Celta
A	F3	Corsa
C	F4	Gol
D	F5	Fox
A	F6	Corsa
...	...	...

Aplicação

CdCarro	CdSer	DtServiço
F1	1	09/10/2010
F1	5	10/10/2010
F2	2	09/10/2010
F2	3	10/10/2010
F2	4	11/10/2010
F3	1	09/10/2010
F3	4	10/10/2010
F5	3	11/10/2010
F5	4	09/10/2010
F4	2	10/10/2010
F4	3	10/10/2010
F3	5	11/10/2010
F4	1	11/10/2010
...	...	...

Serviços

CdSer	Nome
1	Pintura
2	Suspensão
3	Motor
4	Elétrica
5	Painel e revestimento
...	...

Dica:

- 1 Faça o mapeamento da pergunta
- 2 Faça o mapeamento de junções no modelo e marque o mapeamento da pergunta
- 3 Escreva as operações de AR que consigam
- 4 Responder a pergunta
- 5 Para cada operação 'MONTE' o esquema da tabela.



LABORATÓRIO DE BANCO DE DADOS

IBD100