

REFATORAÇÃO EM BANCOS DE DADOS (PARTE 3)

Sérgio Mergen

Introdução

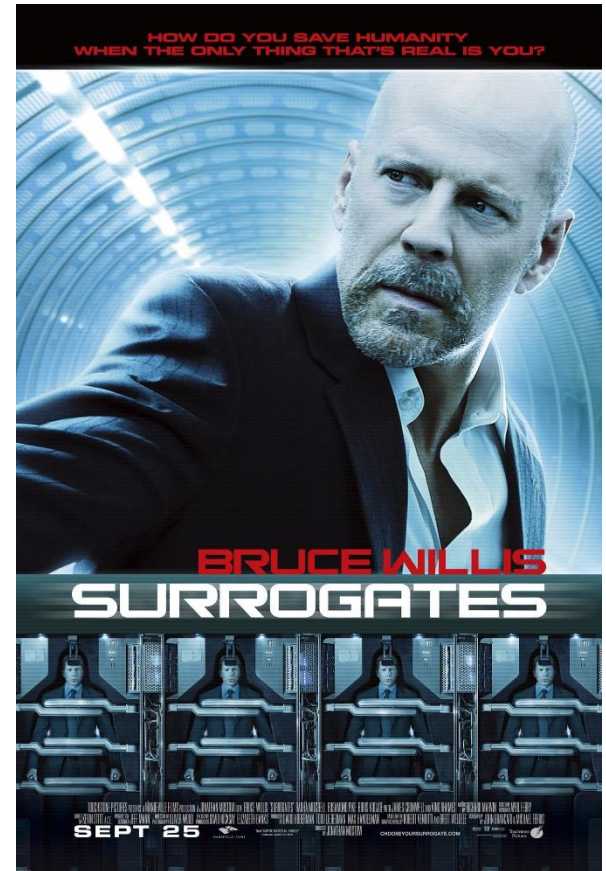
- A chave primária de uma tabela deve ser definida com cautela
- Problemas decorrentes de uma escolha equivocada
 - **Tamanho insuficiente:** chave pequena demais pode não comportar todos os registros futuros.
 - **Dados voláteis:** usar informações que podem mudar ou perder relevância (ex.: CPF, número de matrícula) compromete integridade e relacionamentos.
 - **Falsa unicidade:** campos que não deveriam ser únicos acabam bloqueando inserções válidas.

Introdução

- Como vimos em aulas anteriores, alterar a chave primária de uma tabela pode ser algo trabalhoso
 - Envolve remoção/recriação de restrições de integridade
 - Em alguns bancos (como MySQL), a tabela inteira pode precisar ser recriada
- Para evitar esse tipo de problema
 - é preferível usar como chave primária campos estáveis, imutáveis e de tamanho suficiente
- **Surrogate keys** podem ajudar.

Surrogate Keys

- Surrogate key é uma coluna artificial usada como chave primária
- **Características principais**
 - Não tem valor semântico: não representa dados reais, como CPF ou matrícula.
 - Geralmente é um número sequencial (INT, BIGINT) ou um identificador único (UUID).



Surrogate Keys

- Alguns cenários de uso
 - Substituir uma chave primária cujo valor pode ser modificado
 - Substituir uma chave primária cujo tipo de dados ocupe muitos bytes
 - Substituir uma chave primária composta

Surrogate Keys

- Substituição de uma chave primária cujo valor pode ser modificado
- Ex. sem surrogate key

Projeto	
*nome	char(30)
custo	decimal

- A chave primária de Projeto é uma coluna que possui significado dentro do sistema
- O nome é uma informação que eventualmente pode ser atualizada

Surrogate Keys

- Substituição de uma chave primária cujo valor pode ser modificado
- Ex. com surrogate key

Projeto	
*idProj	int
nome	char(30)
custo	decimal

- Agora a chave primária de Projeto deixou de ter um significado
- O idProj não tem relevância fora do banco de dados. Não existe o interesse em atualizar essa informação

Surrogate Keys

- Substituição de uma chave primária cujo tipo de dados ocupe muitos bytes
- Ex. sem surrogate key

Projeto	
*cod	char(30)
custo	decimal

- A chave primária de Projeto ocupa 30 bytes
- Junções envolvendo projeto serão relativamente caras
 - Cada entrada do índice para nome ocupará 30 bytes

Surrogate Keys

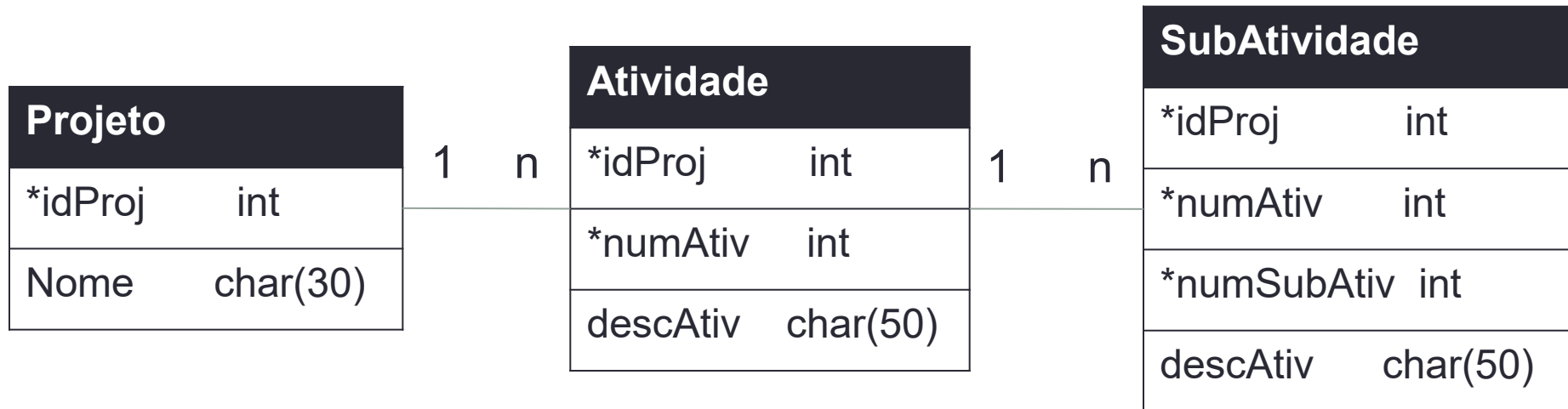
- Substituição de uma chave primária cujo tipo de dados ocupe muitos bytes
- Ex. com surrogate key

Projeto	
*idProj	int
cod	char(30)
custo	decimal

- Agora a chave primária de Projeto ocupa 4 bytes
- Junções envolvendo projeto serão mais baratas

Surrogate Keys

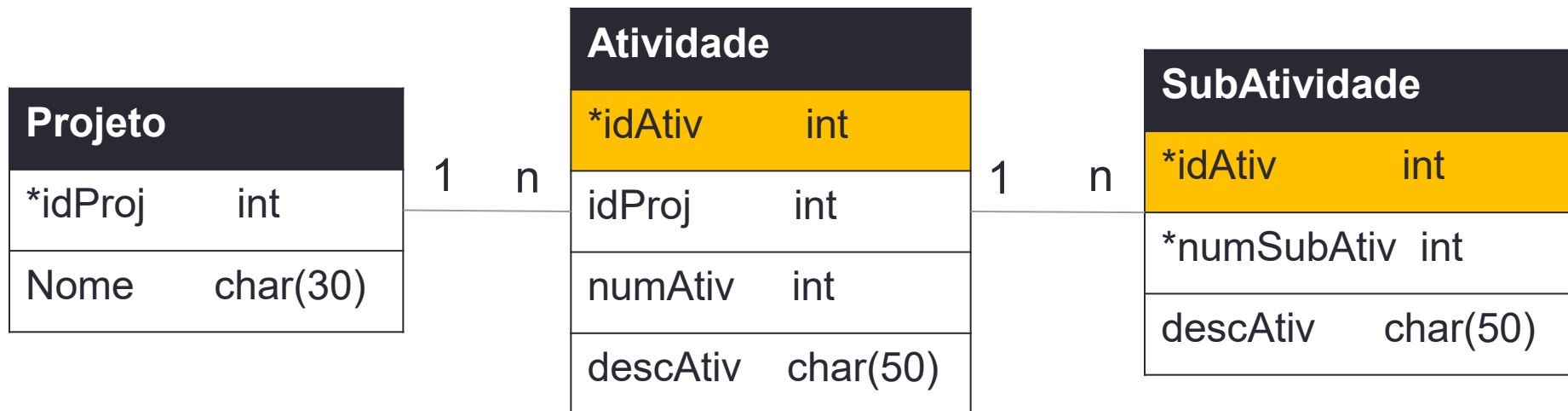
- Substituição de uma chave primária composta
- Ex. sem surrogate key



- A chave primária de Atividade possui duas colunas
- Junções envolvendo atividade serão um pouco mais caras
 - As comparações de chaves serão feitas sobre dois valores

Surrogate Keys

- Substituição de uma chave primária composta
- Ex. com surrogate key



- A chave primária de Atividade possui uma única coluna
- Junções envolvendo atividade serão um pouco mais baratas
 - As comparações de chaves serão feitas sobre um único valor

Exercício

- O modelo abaixo **não** usa surrogate keys

Empresa (**nEmp**, nomeEmp)

Depto (**nDepto**, **nEmp**, nomeDepto)

nEmp referencia empresa

Setor (**nSetor**, **nDepto**, **nEmp**, **nomeSetor**)

nDepto, nEmp referencia depto(nDepto, nEmp)

Equip (**nEquip**, **nSetor**, **nDepto**, **nEmp**, **nomeEquip**)

nSetor, nDepto, nEmp referencia setor (nSetor, nDepto, nEmp)

- Como listar os equipamentos do setor “Montagem”?

Exercício

- O modelo abaixo **não** usa surrogate keys

Empresa (**nEmp**, nomeEmp)

Depto (**nDepto**, **nEmp**, nomeDepto)

nEmp referencia empresa

Setor (**nSetor**, **nDepto**, **nEmp**, nomeSetor)

nDepto, nEmp referencia depto(nDepto, nEmp)

Equip (**nEquip**, **nSetor**, **nDepto**, **nEmp**, nomeEquip)

nSetor, nDepto, nEmp referencia setor (nSetor, nDepto, nEmp)

- Como listar os equipamentos do setor “Montagem”?
- Seria mais barato se fossem usadas surrogate keys?

Exercício

- O modelo abaixo **usa** surrogate keys (prefixo **sur**)

Empresa (**idEmp**, nEmp, nomeE)

Depto (**idDepto**, nDepto, idEmp, nomeDepto)

idEmp referencia empresa(idEmp)

Setor (**idSetor**, nSetor, idDepto, **nomeSetor**)

idDepto referencia Depto (idDepto)

Equip (**idEquip**, nEquip, idSetor, **nomeEquip**)

idSetor referencia Setor (idSetor)

- Como listar os equipamentos do setor “Montagem”?
- Com surrogate keys é mais barato.
 - Menos chaves comparadas

Exercício

- O modelo abaixo **usa** surrogate keys (prefixo **sur**)

Empresa (**idEmp**, nEmp, **nomeE**)

Depto (**idDepto**, nDepto, idEmp, nomeDepto)

idEmp referencia empresa(idEmp)

Setor (**idSetor**, nSetor, idDepto, nomeSetor)

idDepto referencia Depto (idDepto)

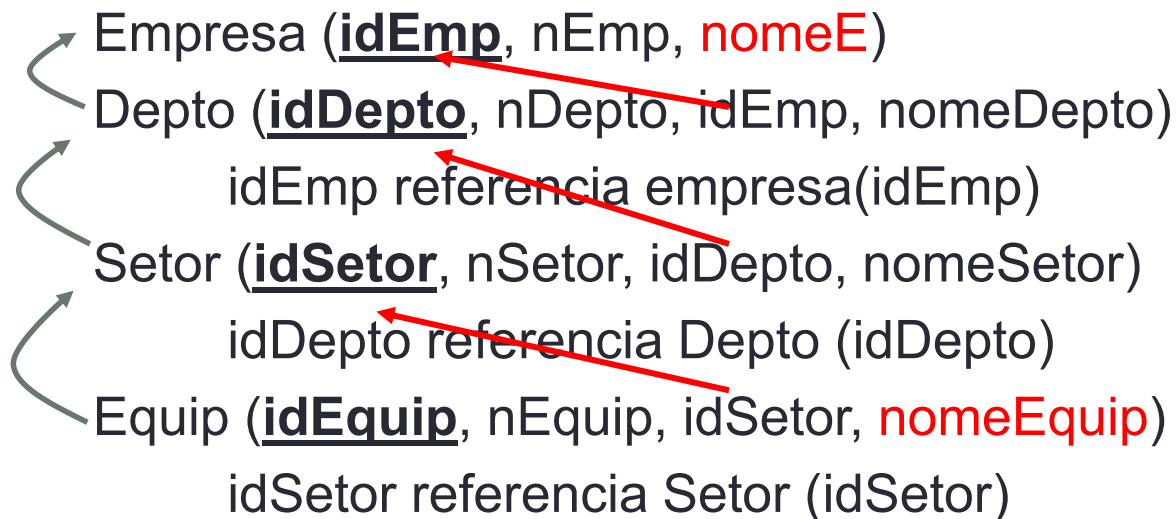
Equip (**idEquip**, nEquip, idSetor, **nomeEquip**)

idSetor referencia Setor (idSetor)

- E para listar os equipamentos da empresa “ACME”?

Exercício

- O modelo abaixo **usa** surrogate keys (prefixo **sur**)



- E para listar os equipamentos da empresa “ACME”?
- Seria mais barato com chaves compostas?

Exercício

- O modelo abaixo **não** usa surrogate keys

Empresa (**nEmp**, **nomeEmp**)

Depto (**nDepto**, **nEmp**, nomeDepto)

nEmp referencia empresa

Setor (**nSetor**, **nDepto**, **nEmp**, nomeSetor)

nDepto, nEmp referencia depto(nDepto, nEmp)

Equip (**nEquip**, **nSetor**, **nDepto**, **nEmp**, **nomeEquip**)

nSetor, nDepto, nEmp referencia setor (nSetor, nDepto, nEmp)

- E para listar os equipamentos da empresa “ACME”?
- Com chaves compostas é mais barato
 - Pode-se atalhar o caminho até a empresa

Considerações

- Vantagens das surrogate keys
 - **Chaves menores** → PKs e FKs ocupam menos espaço quando comparadas a compostas.
 - **Flexibilidade** → mudanças nos atributos naturais não afetam a PK.
 - **Junções mais simples e rápidas** (na maioria dos casos) → evita múltiplas comparações em PKs compostas.
 - **Clareza no modelo** → facilita leitura e entendimento (uma única coluna identificadora).

Considerações

- Desvantagens das surrogate keys
 - **Artificialidade** → introduzem dados que não têm significado no domínio.
 - **Perda de “atalhos” nas junções** → com PKs compostas, certos relacionamentos podem ser resolvidos diretamente pelos atributos naturais.
 - **Limitação do espaço numérico** → pode estourar dependendo do tipo escolhido (ex.: INT).
 - Mitigação: usar tipos maiores (BIGINT).

Exercício

- Modifique o modelo abaixo para que sejam usadas surrogate keys onde elas forem aplicáveis

Func (idFunc, nome, salario)

Proj (idProj, nome)

Aloc (idFunc, idProj, funcao)

idFunc referencia Func

idProj referencia Proj

Atividade (idFunc, idProj, dia, atividade)

idFunc, idProj referencia Aloc (idFunc, idProj)

Solução

- Modifique o modelo abaixo para que sejam usadas surrogate keys onde elas forem aplicáveis

Func (idFunc, nome, salario)

Proj (idProj, nome)

Aloc (idAloc, idFunc, idProj, funcao)

idFunc referencia Func

idProj referencia Proj

Atividade (idAtiv, idAloc, dia, atividade)

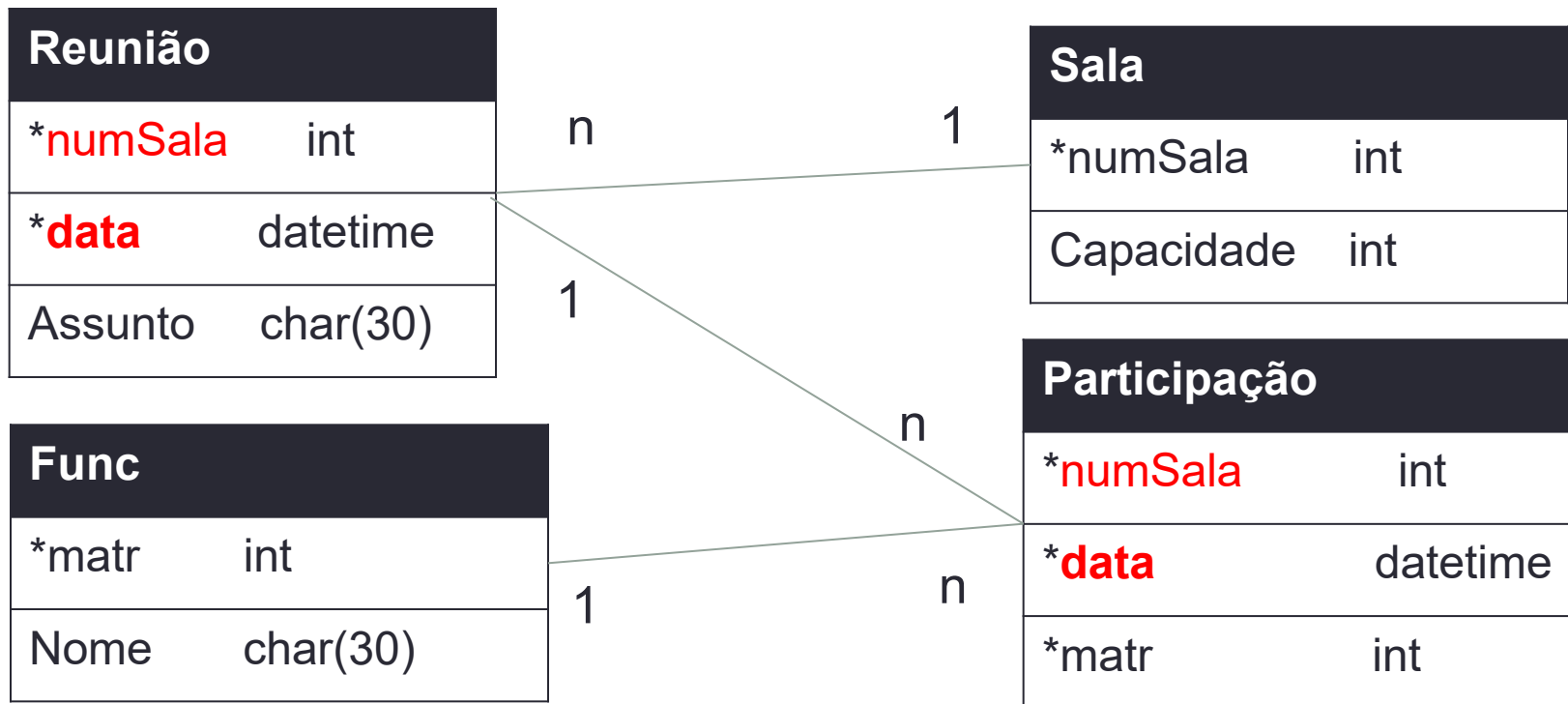
idAloc referencia Aloc (idAloc)

Processos de Refatoração

- Processos de refatoração voltados à surrogate keys
 - Introdução de surrogate key
 - Substituição de surrogate key por chave natural

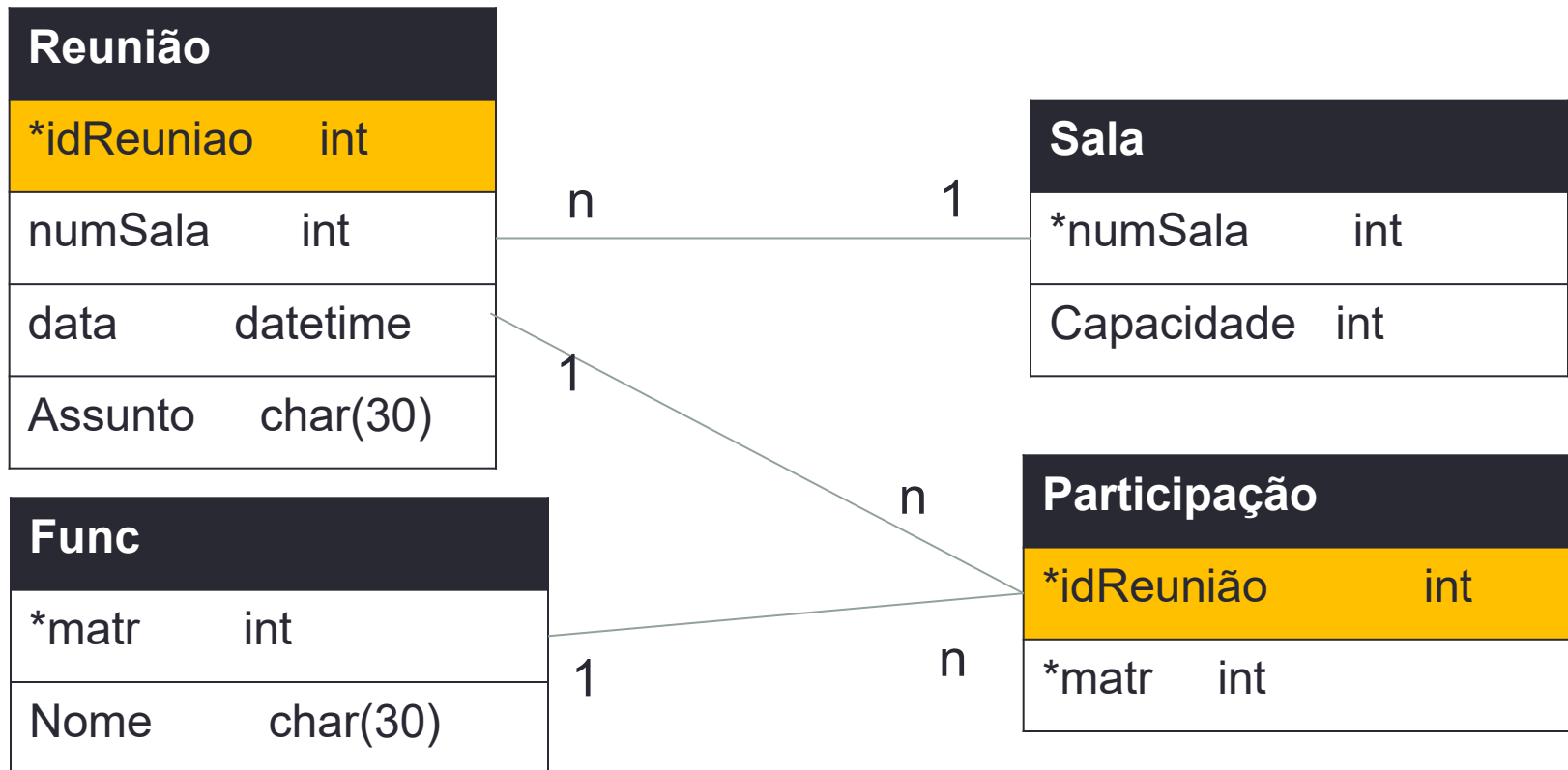
Introdução de Surrogate Key

- Por uma mudança de boas práticas da empresa, decidiu-se trocar todas chaves primárias com colunas do tipo data por surrogate keys



Introdução de Surrogate Key

- Por uma mudança de boas práticas da empresa, decidiu-se trocar todas chaves primárias com colunas do tipo data por surrogate keys
 - Desse modo, as tabelas reunião e participação precisam ser modificadas



Introdução de Surrogate Key

- Script para introdução de surrogate key
 - Passo 1: Remover restrições de chaves primárias e estrangeiras
 - Passo 2: Criar e alimentar surrogate key
 - Passo 3: Criar e alimentar surrogate keys nas tabelas relacionadas
 - Passo 4: Adequar chave primária de tabelas relacionadas

Introdução de Surrogate Key

- Script para introdução de surrogate key
 - Passo 1: Remover restrições de chaves primárias e estrangeiras

- Ex.

//remover restrição de chave estrangeira em participacao

```
ALTER TABLE participacao DROP FOREIGN KEY fk_part_reuniao;
```

//remover restrição de chave primária em reuniao

```
ALTER TABLE reuniao DROP PRIMARY KEY;
```

Introdução de Surrogate Key

- Script para introdução de surrogate key
 - Passo 2: Criar e alimentar surrogate key

- Ex.

//criar e alimentar surrogate key

```
ALTER TABLE reuniao ADD COLUMN idReuniao  
INT PRIMARY KEY AUTO_INCREMENT;
```

- Obs.

- A alimentação automática de surrogate keys pode ser feita com colunas auto incrementáveis
 - Registros já existentes também receberão um valor
- Em MySQL, a coluna é definida com “AUTO_INCREMENT “

Introdução de Surrogate Key

- Script para introdução de surrogate key
 - Passo 3: Criar e alimentar surrogate keys nas tabelas relacionadas

- Ex.

//criar surrogate key nas tabelas relacionadas

```
ALTER TABLE participacao ADD COLUMN idReuniao INT  
                                                    DEFAULT 0 NOT NULL;
```

//alimentar as chaves estrangeiras

```
UPDATE participacao p SET idReuniao =  
    (SELECT idReuniao FROM reuniao r  
    WHERE r.numSala = p.numSala AND r.data = p.data);
```

//definir a coluna como chave estrangeira

```
ALTER TABLE participacao ADD CONSTRAINT fk_part_reuniao  
    FOREIGN KEY (idReuniao) REFERENCES reuniao (idReuniao);
```

Introdução de Surrogate Key

- Script para introdução de surrogate key
 - Passo 3: Criar e alimentar surrogate keys nas tabelas relacionadas
 - O DEFAULT é usado porque a coluna NOT NULL precisa ter valor
- Ex.

//criar surrogate key nas tabelas relacionadas

```
ALTER TABLE participacao ADD COLUMN idReuniao INT  
DEFAULT 0 NOT NULL;
```

//alimentar as chaves estrangeiras

```
UPDATE participacao p SET idReuniao =  
(SELECT idReuniao FROM reuniao r  
WHERE r.numSala = p.numSala AND r.data = p.data);
```

//definir a coluna como chave estrangeira

```
ALTER TABLE participacao ADD CONSTRAINT fk_part_reuniao  
FOREIGN KEY (idReuniao) REFERENCES reuniao (idReuniao);
```

Introdução de Surrogate Key

- Script para introdução de surrogate key
 - Passo 3: Criar e alimentar surrogate keys nas tabelas relacionadas
 - A ordem dos dois últimos comandos precisa ser respeitada. Por quê?
- Ex.

//criar surrogate key nas tabelas relacionadas

```
ALTER TABLE participacao ADD COLUMN idReuniao INT  
                                                    DEFAULT 0 NOT NULL;
```

//alimentar as chaves estrangeiras

```
UPDATE participacao p SET idReuniao =  
    (SELECT idReuniao FROM reuniao r  
    WHERE r.numSala = p.numSala AND r.data = p.data);
```

//definir a coluna como chave estrangeira

```
ALTER TABLE participacao ADD CONSTRAINT fk_part_reuniao  
    FOREIGN KEY (idReuniao) REFERENCES reuniao (idReuniao);
```

Introdução de Surrogate Key

- Script para introdução de surrogate key
 - Passo 4: Adequar chave primária de tabelas relacionadas
- Ex.

//precisa remover a restrição de chave primária antes de remover as colunas

```
ALTER TABLE participacao DROP PRIMARY KEY
```

//colunas parte da antiga chave estrangeira agora podem ser removidas

```
ALTER TABLE participacao DROP COLUMN data
```

```
ALTER TABLE participacao DROP COLUMN numSala
```

//nova chave primária criada

```
ALTER TABLE participacao ADD CONSTRAINT pk_part  
PRIMARY KEY (idReuniao, matr)
```

Introdução de Surrogate Key

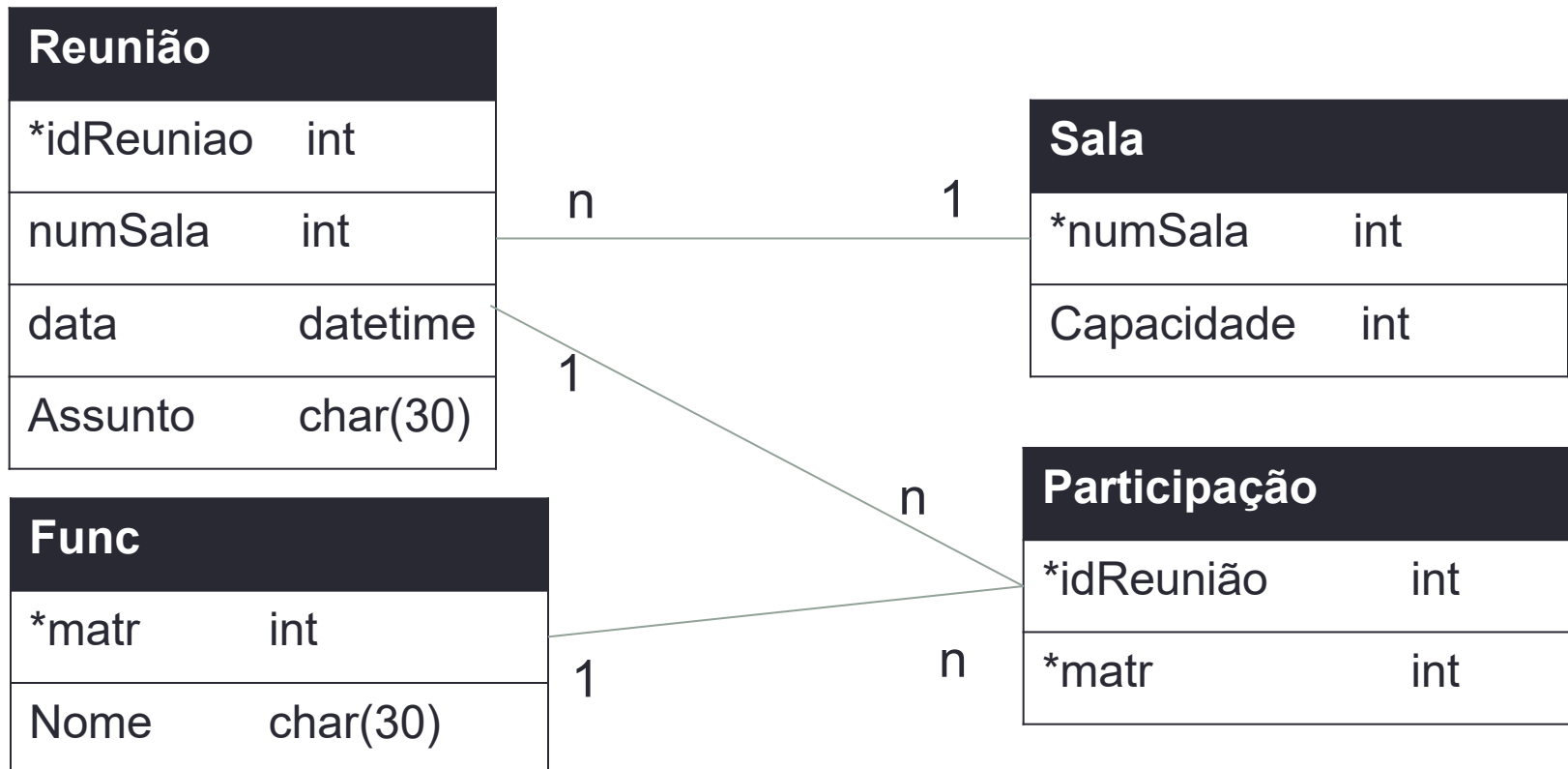
- Como visto no exemplo, a introdução de uma surrogate key gerou uma série de alterações
 - Tanto no esquema
 - Quanto nos dados
- A decisão pela introdução de surrogate keys deve ser tomada com cuidado
 - Afinal, o impacto pode ser muito grande

Processos de Refatoração

- Processos de refatoração voltados à surrogate keys
 - Introdução de surrogate key
 - Substituição de surrogate key por chave natural

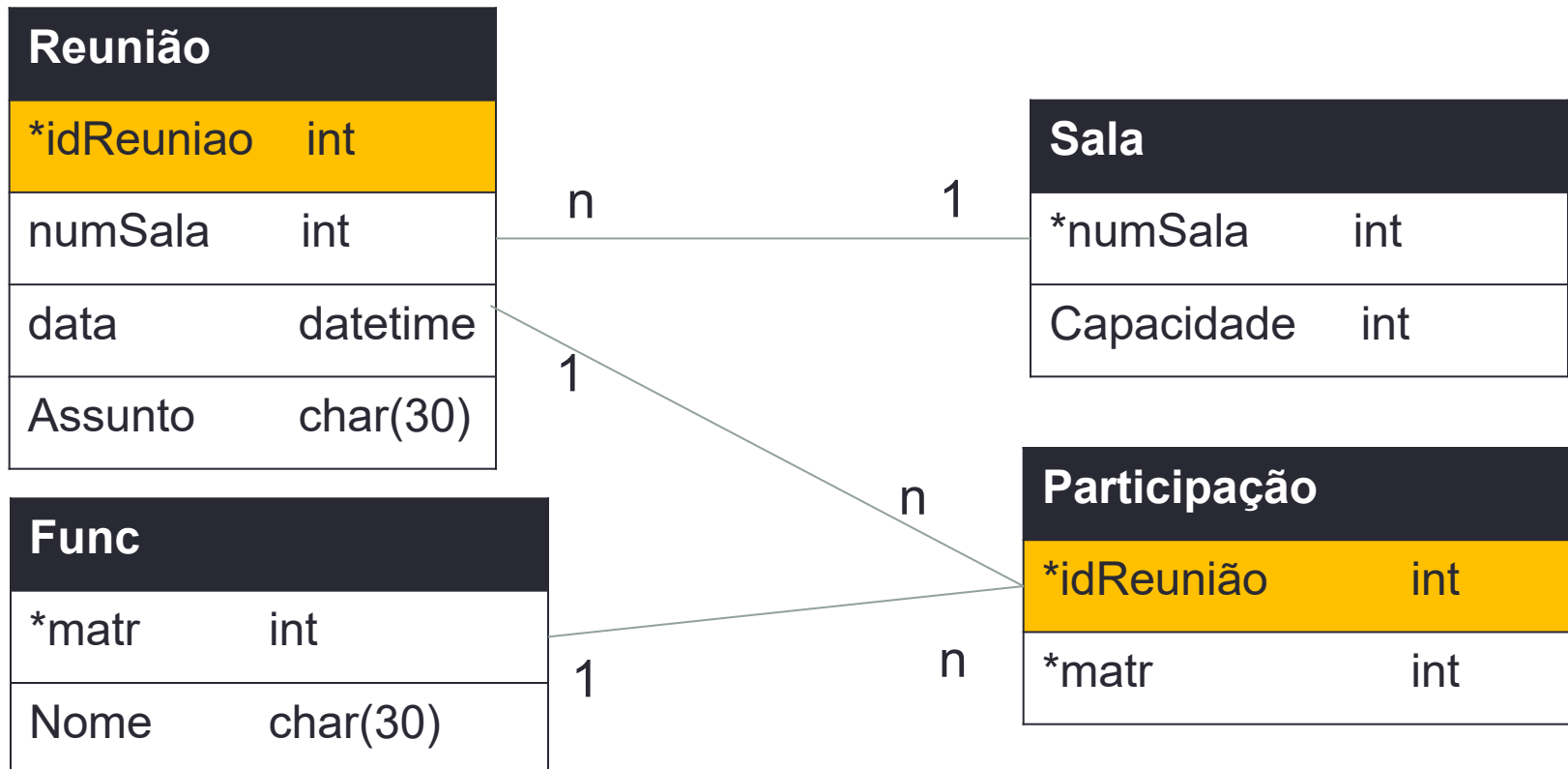
Substituição de Surrogate Key por chave natural

- Por uma mudança de boas práticas da empresa, decidiu-se passar a usar chaves naturais ao invés de surrogate keys



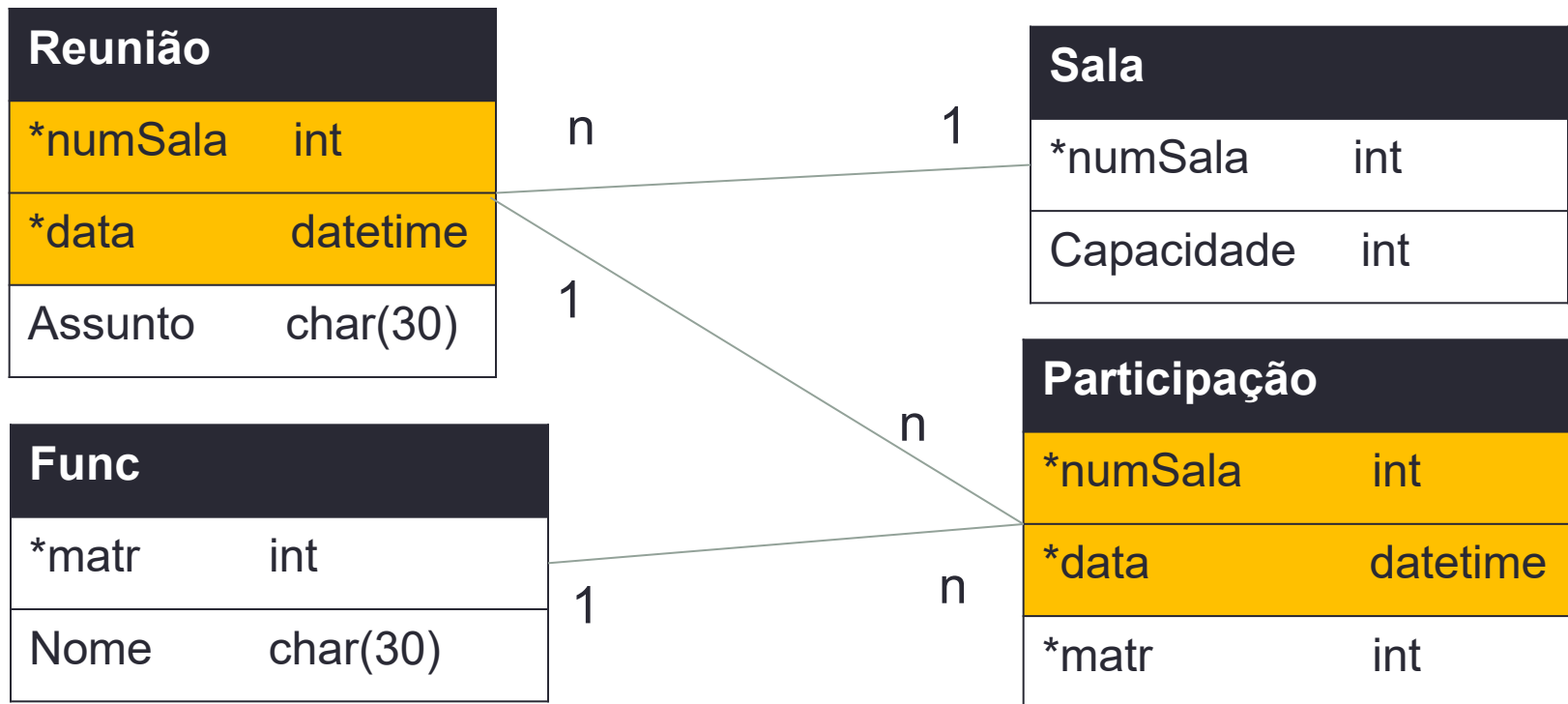
Substituição de Surrogate Key por chave natural

- Por uma mudança de boas práticas da empresa, decidiu-se passar a usar chaves naturais ao invés de surrogate keys
 - Desse modo, as tabelas reunião e participação precisam ser alteradas



Substituição de Surrogate Key por chave natural

- Assim ficaria o modelo após a alteração

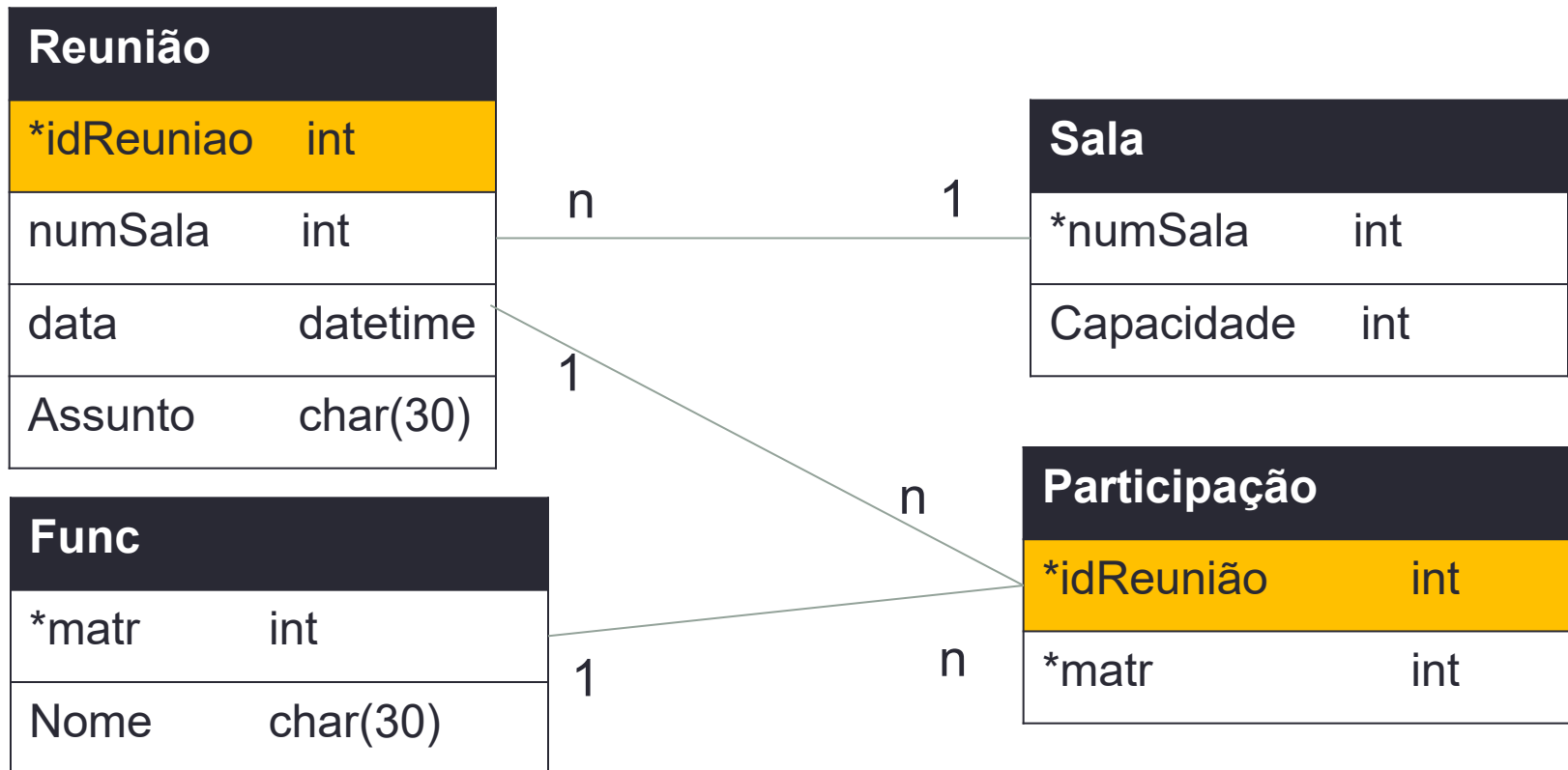


Substituição de Surrogate Key por chave natural

- Script para substituição de surrogate key por chave natural
 - Passo 1: Remover restrições de chaves primárias e estrangeiras
 - Passo 2: Criar e alimentar a nova chave estrangeira nas tabelas relacionadas
 - Passo 3: Adequar chave primária de tabelas relacionadas
- Os passos são semelhantes aos passos de criação de uma surrogate key
 - Um pouco mais simples

Exercício

- Elimine as surrogate keys no exemplo abaixo



Atividade Individual

- Use surrogate key no exemplo abaixo

