

Engenharia Reversa de Modelos Relacionais

De forma geral, um processo de **engenharia reversa** pode ser definido como sendo um processo de abstração, que parte de um modelo de implementação e resulta em um modelo conceitual que descreve abstratamente a implementação em questão. O termo **engenharia reversa** vem do fato de usar-se como ponto de partida do processo um produto implementado (o modelo de implementação) para se obter sua especificação (o modelo conceitual).

No caso de banco de dados, fala-se de **engenharia reversa** quando transforma-se modelos de banco de dados mais ricos em detalhes de implementação em modelos de dados mais abstratos.

Um caso específico de **engenharia reversa** de banco de dados é o da **engenharia reversa de modelos relacionais**. Neste tipo de **engenharia reversa** tem-se, como **ponto de partida**, um **modelo lógico de um banco de dados relacional** e, como **resultado**, um **modelo conceitual**, no nosso caso, na abordagem ER. Este é o processo inverso ao de projeto lógico.

A **engenharia reversa de modelos relacionais** pode ser útil quando não se tem um modelo conceitual para um banco de dados existente. Isso pode acontecer quando o banco de dados foi desenvolvido de forma empírica, sem o uso de uma metodologia de desenvolvimento, ou quando o esquema do banco de dados sofreu modificações ao longo do tempo, sem que as mesmas tenham sido registradas no modelo conceitual.

Conforme será visto nesta disciplina, a **engenharia reversa de modelos relacionais** é um passo dentro de um processo mais amplo de **engenharia reversa de arquivos e documentos convencionais**.

O processo de **engenharia reversa de um modelo relacional** consta dos seguintes passos:

1. Identificação da construção ER correspondente a cada tabela
2. Definição de relacionamentos **1:n** e **1:1**
3. Definição de atributos
4. Definição de identificadores de entidades e relacionamentos

Estes passos são detalhados nas seções que se seguem. O processo será exemplificado sobre um banco de dados para um sistema acadêmico, cujo esquema é apresentado abaixo.

DISCIPLINA (<u>CodDisc</u> , NomeDisc)
CURSO (<u>CodCr</u> , NomeCr)
CURRIC (<u>CodCr</u> , <u>CodDisc</u> , Obr/Opc)
CodCr referencia CURSO
CodDisc referencia DISCIPLINA
SALA (<u>CodPr</u> , <u>CodSl</u> , Capacidade)
CodPr referencia PREDIO
PREDIO (<u>CodPr</u> , Endereco)
TURMA (<u>Anosem</u> , <u>CodDisc</u> , <u>SiglaTur</u> , Capacidade, CodPr, CodSl)
CodDisc referencia DISCIPLINA
(CodPr, CodSl) referencia SALA
LABORATORIO (<u>CodPr</u> , <u>CodSl</u> , Equipam)
(CodPr, CodSl) referencia SALA

1. Identificação da construção ER correspondente a cada tabela

Na primeira etapa da engenharia reversa de um banco de dados relacional define-se, para cada tabela do modelo relacional, qual a construção que lhe corresponde a nível de modelo ER.

Uma tabela pode corresponder a:

- uma entidade
- um relacionamento **n:n**
- uma entidade especializada

O fator determinante da construção ER que corresponde a uma tabela é a composição de sua chave primária. Tabelas podem ser classificadas em três tipos, de acordo com sua chave primária:

- **Regra 1: Chave primária composta por mais de uma chave estrangeira**

A tabela que possui uma chave primária composta de múltiplas chaves estrangeiras implementa um **relacionamento n:n** entre as entidades correspondentes às tabelas referenciadas pelas chaves estrangeiras. Um exemplo de tabela deste tipo é a tabela CURRIC que tem como chave primária **CodCr** e **CodDisc**. Ambas colunas são chave estrangeira em relação às tabelas CURSO e DISCIPLINA, respectivamente. Portanto, a tabela CURRIC representa um relacionamento entre as entidades correspondentes às tabelas CURSO e DISCIPLINA. No exemplo, a única tabela deste tipo é a tabela CURRIC.

- **Regra 2: Toda a chave primária é uma chave estrangeira**

A tabela cuja chave primária é toda ela uma chave estrangeira representa uma entidade que forma uma **especialização** da entidade correspondente à tabela referenciada pela chave estrangeira. Um exemplo de tabela deste tipo é a tabela LABORATÓRIO que possui como chave primária as colunas (CodPr, CodSI), as quais são chave estrangeira da tabela de salas. A restrição de integridade referencial em questão especifica que uma linha na tabela de laboratórios somente existe quando uma linha com a mesma chave existir na tabela de salas. A nível de modelo ER, isso significa que uma ocorrência da entidade laboratório somente pode existir quando a correspondente ocorrência da entidade sala existe, ou seja, significa que a entidade laboratório é uma especialização de sala. No exemplo, a única tabela deste tipo é a tabela LABORATÓRIO.

- **Regra 3: Demais casos**

Quando a chave primária da tabela não for composta de múltiplas chaves primárias (regra 1) nem for toda uma chave estrangeira (regra 2), a tabela representa uma **entidade**. Exemplificando, a tabela CURSO, cuja chave primária, a coluna CodCr não contém chaves estrangeiras, representa uma entidade. Da mesma forma, a tabela SALA também representa uma entidade. Sua chave primária (colunas CodPr e CodSI) contém apenas uma chave estrangeira (coluna CodPr). Assim, não obedece ao requisito da multiplicidade de chaves estrangeiras (regra 1), nem ao requisito de toda a chave primária ser chave estrangeira (regra 2), enquadrando-se na presente regra. O mesmo é válido para as tabelas DISCIPLINA, PREDIO e TURMA.

Tendo feita a classificação de tabelas segundo a composição da chave primária e com isso identificando as construções ER correspondentes a cada tabela, é possível montar um diagrama ER inicial, conforme mostra a figura 1.

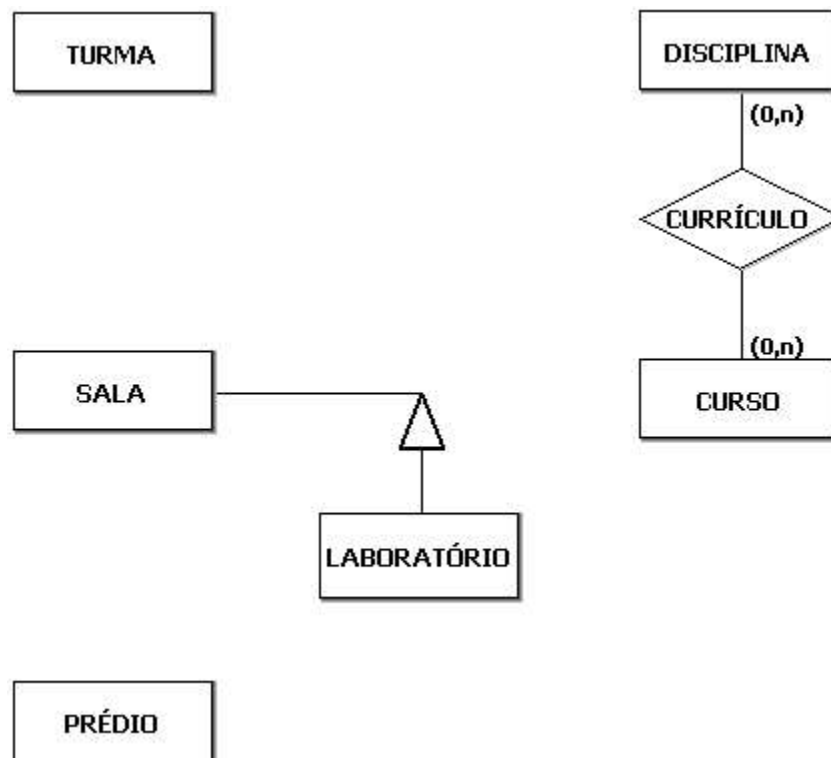


Figura 1. Diagrama ER inicial para o BD acadêmico

2. Identificação de relacionamentos 1:n ou 1:1

Toda chave estrangeira que não se enquadra nas regras 1 e 2 anteriores, ou seja, toda chave estrangeira que não faz parte de uma chave primária composta por múltiplas chaves estrangeiras nem é toda ela uma chave primária, representa um relacionamento **1:n** ou **1:1**. Em outros termos, toda chave estrangeira que não corresponde a um relacionamento **n:n** nem a uma entidade especializada representa um relacionamento **1:n** ou **1:1**. A regra não permite definir se a cardinalidade do relacionamento é **1:n** ou **1:1**. Para definir qual dos dois tipos de relacionamentos está sendo representado pela chave estrangeira é necessário verificar os possíveis conteúdos do banco de dados. No caso do exemplo, as chaves estrangeiras que representam relacionamentos **1:n** ou **1:1** são as seguintes:

Tabela SALA CodPr referencia PREDIO Tabela TURMA CodDisc referencia DISCIPLINA (CodPr, CodSl) referencia SALA

Com isto podemos completar a definição dos relacionamentos no diagrama ER, conforme mostra a figura 2. No caso do exemplo, todos os relacionamentos referentes às chaves estrangeiras acima são do tipo **1:n**.

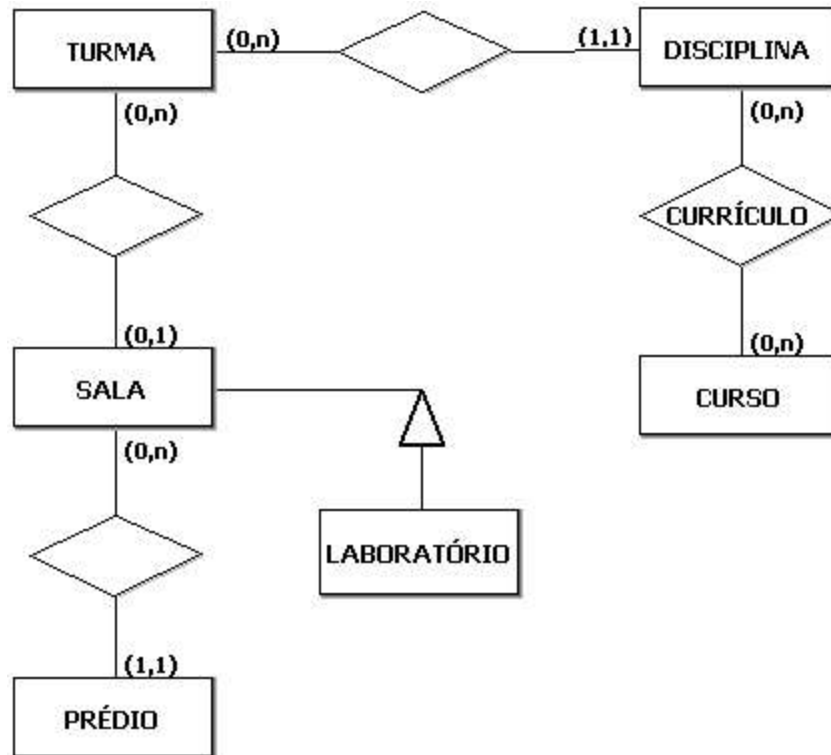


Figura 2. Definição dos relacionamentos 1:n e 1:1

3. Definição de atributos

Nesta etapa, para cada coluna de uma tabela que não seja chave estrangeira, é definido um atributo na entidade/relacionamento correspondente à tabela. Observe-se que colunas chave estrangeira não correspondem a atributos no diagrama ER, mas sim a relacionamentos e, por isso, já foram tratados nas etapas anteriores. Para o caso do exemplo, a execução deste passo da engenharia reversa resulta no diagrama ER da figura 3.

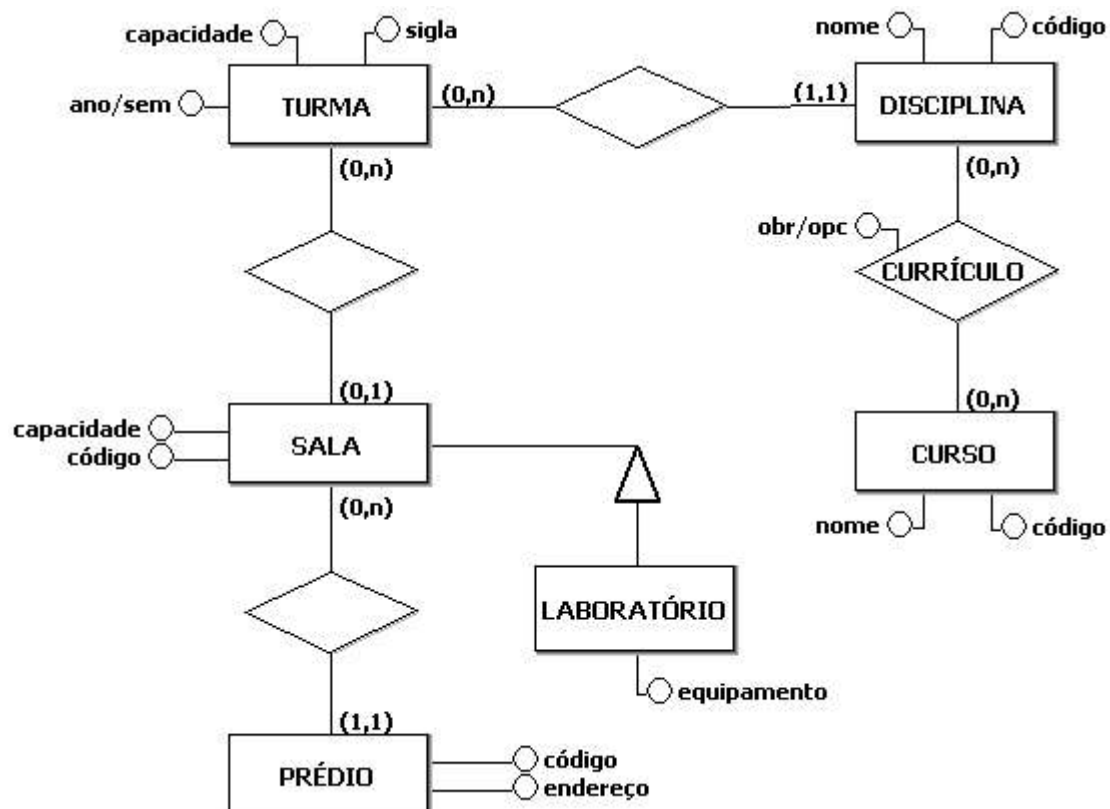


Figura 3. Diagrama dos atributos

4. Definição de identificadores de entidades

No último passo da **engenharia reversa** são definidos os identificadores das entidades e dos relacionamentos. A regra para definição dos identificadores é a seguinte:

- **Coluna da chave primária que não é chave estrangeira**

Toda coluna que faz parte da chave primária e que **não é** chave estrangeira corresponde a um **atributo identificador** da entidade ou relacionamento.

- **Coluna da chave primária que é chave estrangeira**

Toda coluna que faz parte da chave primária e que **é** chave estrangeira corresponde a um identificador externo da entidade. Exemplificando, a coluna CodDisc, que é parte da chave primária da tabela TURMA é também chave estrangeira em relação a tabela DISCIPLINA. Portanto, a entidade TURMA é identificada também pelo relacionamento com DISCIPLINA.

Executando este passo da **engenharia reversa** sobre o modelo do exemplo, chegamos à figura 4.

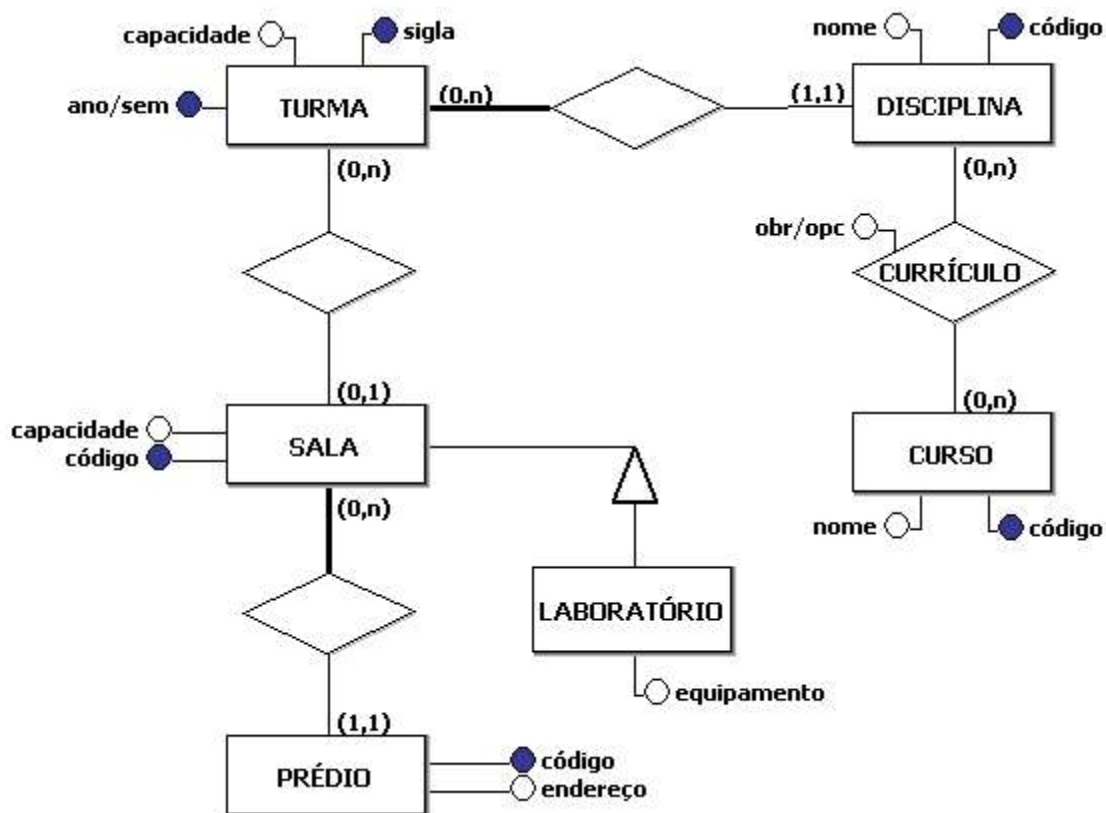


Figura 4. Definição dos identificadores de entidades