

Resumo 1

Banco de dados 1

Marcos Geraldo Braga Emiliano 19.1.4012

Fonte: Slides das aulas

Conceitos Básicos

Dados são fatos conhecidos que possuem significados implícitos e que podem ser armazenados. **Banco de Dados** é uma coleção de dados relacionados com as seguintes propriedades implícitas: representa algum aspecto do mundo real, ou seja, diz respeito a um mini-mundo (Universo de Discurso), corresponde a uma coleção coerente de dados com significado inerente e é projetado, implantado e povoado com dados para um propósito específico.

Um banco de dados é criado e mantido por um **Sistema de Gerência de Banco de Dados (SGBD)**. Um SGBD é um software de propósito geral que facilita os seguintes processos envolvendo bancos de dados: definir que envolve especificar os tipos de dados e as restrições para os dados a serem armazenados no banco de dados, construir que envolve armazenar dados válidos, que seguem as restrições estabelecidas, no banco de dados e manipular, que envolve possibilitar a realização de consultas e a alteração de dados armazenados no banco de dados.

Banco de Dados X Sistema Tradicional de Arquivos

Independência entre programas e dados, em um sistema de banco de dados, as informações estruturais dos dados são armazenadas no Catálogo e são chamadas de metadados, gerando independência entre programas e dados. Em um sistema tradicional de arquivos, a estrutura dos dados está embutida nos programas que realizam o acesso aos mesmos, não havendo independência entre programas e dados.

Definição e manutenção de um único banco de dados a ser usado por vários usuários. Um SGBD multiusuário deve permitir que vários usuários acessem o banco de dados ao mesmo tempo. Em um sistema tradicional de arquivos, para cada aplicação específica, são definidos e implementados os arquivos necessários.

Usuários de um SGBD

Projetista do Banco de Dados: são os responsáveis pela identificação dos dados a serem armazenados no banco de dados e pela escolha da estrutura apropriada para representar e armazenar esses dados.

Administradores do Banco de Dados: são os responsáveis pela administração do banco de dados e do SGBD. Autorizam o acesso dos usuários, coordenam e monitoram o uso do SGBD, verificam a necessidade de novos recursos de software e hardware para o bom funcionamento do SGBD.

Usuários finais: são os usuários que acessam o banco de dados para realizar consultas, promover modificações e gerar relatórios. Os tipos de usuário final são: ingênuos ou parametrizados, usuários cujo trabalho envolve a realização de consultas e modificações no

banco de dados constantemente, por meio de interfaces próprias implementadas e testadas para tais atividades, casuais, usuários que raramente necessitam acessar o banco de dados, mas podem precisar de diferentes informações em cada acesso, especializados, usuários completamente familiarizados com as facilidades do SGBD, realizando atividades complexas.

O que deve fornecer um SGBD?

Controle de redundância, dados redundantes podem gerar problemas como espaço em disco desperdiçado, repetição de uma alteração lógica e inconsistência de dados, **restrição de acesso**, em um banco de dados utilizado por vários usuários, nem todos podem acessar todos os dados e nem todos podem realizar operações de escrita sobre o banco de dados, **suporte a múltiplas visões dos dados**, uma visão consiste em um subconjunto do banco de dados que pode ser destinado a um determinado grupo de usuários, garantia das restrições de integridade, as restrições podem ser simples, como as relacionadas ao tipo de dado, ou complexas, como as que verificam se um item em um arquivo está relacionado a itens em outro arquivo, **backup e recuperação**, caso ocorra uma falha de hardware ou de software, um SGBD deve fornecer facilidades para recuperar os dados, se for necessário, e gerar um banco de dados confiável e consistente, **múltiplas interfaces de usuários**, os usuários do banco de dados possuem um variado nível de conhecimento técnico, **armazenamento persistente para objetos e estruturas de dados de programas**, um objeto é persistente se ele continua existindo após o término da execução de um programa e pode ser acessado diretamente por um outro programa, inferência em banco de dados usando regras de dedução, os bancos de dados que possuem essa capacidade são chamados Bancos de Dados Dedutivos.

Quando não usar um SGBD?

Para se usar um SGBD, existe um aumento de custo em relação ao sistema tradicional de arquivos que se deve ao investimento inicial em software, hardware e treinamento, generalidade para definir e processar dados e fornecimento de mecanismos de integridade, segurança, controle de concorrência e recuperação a falhas.

O uso de sistemas de arquivos é desejável quando o banco de dados e as aplicações são simples, bem definidas e a expectativa de mudança é pequena, acesso de vários usuários aos dados não é requerido e requisitos de tempo real de programas são necessários.

Modelo de Dados

Um modelo de dados é um conjunto de conceitos que podem ser usados para descrever a estrutura de um banco de dados, é categorizado de acordo com os tipos de conceito que fornece para descrever a estrutura do banco de dados, podendo ser conceitual, lógico ou físico e a estrutura de um banco de dados envolve os tipos de dados, os relacionamentos e as restrições sobre os dados.

Conceitual (ou alto-nível): fornece conceitos para descrever os dados como a maioria dos usuários os percebe. Esses conceitos são: entidade, representa objetos e conceitos do mundo real, relacionamento, representa a interação entre entidades e Atributo, representa alguma propriedade de uma entidade ou de um relacionamento.

Lógico (ou representacional): fornece conceitos que podem ser compreendidos pelos usuários, mas não estão muito longe de como os dados serão organizados. Representam os dados usando a estrutura de registro.

Físico (ou baixo-nível): fornece conceitos para descrever os detalhes de como os dados são fisicamente armazenados no computador. Alguns conceitos são, formato dos registros, ordenação dos registros, vias de acesso: são utilizadas para recuperar de forma eficiente os registros de um banco de dados.

Esquema

Esquema é a descrição textual ou gráfica de um banco de dados de acordo com um determinado modelo de dados, o esquema de banco de dados, na forma gráfica, é chamado de diagrama do esquema.

Instância

Instância é o conjunto de dados armazenado em um banco de dados em um determinado instante do tempo, correspondendo ao estado do banco de dados, os dados em um banco de dados podem mudar frequentemente, alterando assim a instância do banco. O SGBD deve garantir que toda instância do banco de dados corresponda a um estado válido, o estado inicial de um banco de dados é obtido quando os primeiros dados são armazenados.

Arquitetura de Três Níveis (Três Esquemas)

O objetivo da arquitetura de três níveis é separar as aplicações dos usuários e o banco de dados físico. Possui os seguintes níveis, interno, tem um esquema interno que descreve a estrutura física de armazenamento do banco de dados, conceitual, tem um esquema conceitual que descreve a estrutura completa do banco de dados para os usuários e externo ou de visão, possui um conjunto de esquemas externos ou visões de usuários que descrevem partes do banco de dados que um grupo particular de usuários pode estar interessado.

Os três esquemas são apenas descrições de dados, o único local onde os dados realmente existem é no nível físico. Mapeamento entre os esquemas, o SGBD deve transformar as requisições especificadas no esquema externo em requisições para o esquema conceitual e as do esquema conceitual em requisições para o esquema interno para serem processadas sobre o banco de dados armazenado.

Independência de dados é a capacidade de mudar o esquema em um nível da arquitetura de três níveis sem ter que mudar o esquema do nível acima, independência de dados lógica é a capacidade de mudar o esquema conceitual sem ter que mudar algum esquema externo ou programa de aplicação e a independência de dados física é a capacidade de mudar o esquema interno sem ter que mudar o esquema conceitual ou algum esquema externo.

Linguagens de um SGBD

Linguagem de Definição de Dados (LDD): especifica esquema conceitual, Linguagem de Definição de Armazenamento (LDA): especifica o esquema interno, Linguagem de Definição de Visões (LDV): especifica as visões dos usuários e o mapeamento para o esquema conceitual, Linguagem de Manipulação de Dados (LMD): usada para consultar, inserir, remover e modificar dados do banco de dados. A linguagem de banco de dados relacional SQL representa uma combinação de LDD, LDA, LDV e LMD.

Interfaces de um SGBD

Interfaces baseadas em menus para browsing, baseadas em formulários, gráficas, de linguagem natural, para o Administrador de Banco de Dados.

Módulos Componentes de um SGBD

Gerenciador do banco de dados, controla o acesso aos dados armazenados no disco, Compilador da LDD, compila as definições de um esquema, armazenando-as no catálogo do SGBD, Processador do banco de dados em tempo de execução, recebe as operações de recuperação e modificação e as executa sobre o banco de dados, Compilador de consultas, compila as consultas de alto nível que são fornecidas interativamente, Pré-compilador: extrai comandos LMD de um programa de aplicação escritos em uma linguagem de programação hospedeira e Compilador da LMD: recebe e compila os comandos LMD extraídos pelo pré-compilador, gerando o código objeto referente, o código objeto gerado e o resto do programa compilado pela linguagem hospedeira da aplicação são unidos formando uma só transação.

Utilitários

Carregador, é utilizado para povoar o banco de dados com arquivos de dados existentes, Backup, cria uma cópia de backup do banco de dados, Monitor de eficiência, monitora o uso do banco de dados e oferece estatísticas para o administrador do mesmo e Reorganizador de arquivos, é utilizado para reorganizar arquivos do banco de dados no intuito de melhorar a eficiência na execução de operações.

Critérios de Classificação de SGBDs

Modelo de dados lógico, relacional, hierárquico, rede, orientado a objeto e objeto relacional. Número de servidores, centralizado, o banco de dados e o SGBD estão armazenados em um único servidor e distribuído (SGBDD): o banco de dados e o SGBD estão em vários servidores conectados por uma rede, SGBDD Homogêneo: usa o mesmo SGBD nos servidores ou SGBDD Federado: os SGBDs participantes são heterogêneos; torna-se necessária uma forma eficiente e eficaz de comunicação entre eles.

Modelagem de Dados

Modelo ER

O Modelo de Entidades e Relacionamentos (MER) é um modelo conceitual usado para projeto de aplicações de banco de dados, é um modelo baseado na percepção do mundo real como conjuntos de objetos básicos chamados entidades e nos relacionamentos entre esses objetos e é independente de aspectos de implementação.

Exemplo de Aplicação de Banco de Dados

Deseja-se criar um BD "Empresa" que controla empregados, departamentos e projetos de uma empresa, após a fase de coleta e análise dos requisitos, os projetistas declararam a seguinte descrição do "mini-mundo": A empresa é organizada em departamentos. Cada departamento possui um nome único, um código único e um determinado empregado que gerencia o departamento. Acompanha-se a data inicial que o empregado começou a gerenciar o departamento. Um departamento pode possuir diversas localizações. Um departamento controla um número de projetos onde cada um deles possui um nome único, um código único e uma localização única. Armazenam-se o nome, o número da seguridade social, o endereço, o

salário, o sexo e a data de nascimento de cada empregado. Um empregado é alocado a um único departamento, mas pode trabalhar em diversos projetos, que não são necessariamente controlados pelo mesmo departamento. Acompanha-se o número de horas por semana que um empregado trabalha em cada projeto. Também se acompanha o supervisor direto de cada empregado. Acompanham-se os dependentes de cada empregado para fins de seguridade social. São mantidos, para cada dependente, o nome, o sexo, a data de nascimento e o grau de parentesco com o empregado.

Entidade

Entidade é um objeto que existe no mundo real e é distinguível dos outros objetos. Tipo de Entidade define uma coleção de entidades que possuem os mesmos atributos. Descreve o esquema para um conjunto de entidades. Conjunto de Entidades é uma coleção das instâncias de entidades de um determinado tipo de entidade.

Atributo

Atributo é uma propriedade que descreve uma característica particular de uma entidade. Os atributos podem ser dos seguintes tipos, simples (atômicos) ou compostos: os atributos simples são indivisíveis e os atributos compostos podem ser divididos em subpartes menores, Mono-valorados ou Multi-valorados: os atributos mono-valorados têm um único valor para uma entidade particular e os atributos multi-valorados podem ter um conjunto de valores, Armazenados ou Derivados: os atributos derivados podem ser obtidos a partir dos atributos armazenados (e portanto não precisam ser armazenados no banco de dados).

Atributos complexos: os atributos podem ser compostos e multi-valorados ao mesmo tempo. Valores nulos: uma entidade pode não ter valor aplicado a um atributo. Neste caso, é usado o valor especial nulo. É também usado quando não se conhece o valor de um atributo. Cada atributo de um tipo de entidade está associado com um conjunto de valores (ou domínio de valores), o qual especifica os valores que podem ser atribuídos àquele atributo para cada instância de entidade.

Chave

Atributo Chave (identificador): é um atributo cujos valores são distintos (únicos) para cada instância de entidade de um mesmo tipo de entidade, a chave pode ser formada por mais de um atributo (chave composta). Neste caso, a combinação dos valores desses atributos é que é única para cada instância de entidade. Superchave é qualquer conjunto de um ou mais atributos cujos valores são distintos para cada instância de entidade. Chave Candidata é uma superchave que não possui subconjuntos próprios de superchaves, ou seja, é uma superchave com um conjunto mínimo de atributos. Chave Primária é a chave candidata escolhida arbitrariamente pelo projetista do banco de dados como meio principal de identificação de entidades para o tipo de entidade em questão, as demais chaves candidatas são chamadas de chaves alternativas ou chaves secundárias.

Projeto Conceitual inicial do BD Empresa

De acordo com a descrição do "mini-mundo" da empresa, identifica-se os seguintes tipos de entidade, departamento com os atributos nome, código, localizações, gerente, data_início_gerência. Localizações é o único atributo multivalorado. Pode-se especificar que tanto nome quanto código são atributos chaves (separados), porque cada um foi especificado

para ser único, projeto com os atributos nome, código, localização e departamento_controlador. Tanto o nome quanto o código são atributos chaves (separados). Empregado com os atributos nome, seguridade_social (carteira profissional), sexo, endereço, salário, data_nascimento, departamento, supervisor. Tanto o nome quanto o endereço podem ser atributos compostos (não foi especificado). O fato de que um empregado poder trabalhar em vários projetos, um certo número de horas por semana para cada projeto, pode ser representado pelo atributo composto multi-valorado trabalha_em com os componentes (projeto, horas), dependente com os atributos empregado, nome_dependente, sexo, data_nascimento, parentesco (com o empregado).

Relacionamento

Relacionamento é uma associação entre entidades que representa um fato do mundo real. Tipo de Relacionamento R sobre n tipos de entidade E1, E2, ..., En define um conjunto de associações entre entidades destes tipos. Grau de um tipo de relacionamento é o número de tipos de entidades participantes. Tipos de relacionamento podem ter atributos, similares àqueles dos tipos de entidade. Tipo de relacionamento recursivo ocorre quando o mesmo tipo de entidade participa mais de uma vez em um tipo de relacionamento desempenhando diferentes papéis.

Restrição de cardinalidade especifica o número de instâncias de um relacionamento que uma instância de entidade pode participar. Para um tipo de relacionamento binário R entre os tipos de entidade A e B, tem-se as relações de cardinalidade, 1:1 (um-para-um), uma entidade de A está associada a, no máximo, uma entidade de B, e uma entidade de B está associada a, no máximo, uma entidade de A, 1:N (um-para-muitos), uma entidade de A está associada a várias (zero ou mais) entidades de B, mas uma entidade de B está associada a, no máximo, uma entidade de A, M:N (muitos-para-muitos), uma entidade de A está associada a várias entidades de B (zero ou mais), e uma entidade de B está associada a várias entidades de A (zero ou mais).

Restrição de participação especifica se a existência de uma entidade depende de sua associação a outra entidade por meio de um relacionamento. A participação pode ser, total (dependência de existência), todas as instâncias do tipo de entidade devem participar de um relacionamento e parcial, as instâncias podem participar de um relacionamento.

Uma restrição estrutural alternativa em relacionamentos no Diagrama ER é a **Cardinalidade Mínima e Máxima**, envolve associar um par de números inteiros (min, max) a cada participação de um tipo de entidade E num tipo de relacionamento, onde $0 \leq \min \leq \max$ e $\max \geq 0$, os números min e max significam que para cada instância de entidade e em E, e deve participar de, no mínimo, min e, no máximo, max instâncias do relacionamento em questão.

Entidade Fraca

Tipo de Entidade Fraca é um tipo de entidade que não possui atributos chaves (não tem identificação própria), está sempre associada a um tipo de entidade forte (ou proprietária) através de um relacionamento identificador, sua chave é formada pela combinação da chave primária do tipo de entidade forte + uma chave parcial própria, sempre tem uma restrição de participação total (dependência de existência) em relação ao relacionamento identificador, é representada no diagrama por contornos duplos e algumas vezes também pode ser modelada como atributos complexos (compostos e multi-valorados).

Refinamento do Projeto do BD Empresa

Os atributos que representam relacionamentos devem ser transformados em tipos de relacionamento. Assim, têm-se os seguintes tipos de relacionamento no BD Empresa, gerencia é um tipo de relacionamento 1:1 entre empregado e departamento. A participação de empregado é parcial. A participação de Departamento é total já que foi dito que um departamento sempre deve possuir um gerente. O atributo data_inicio_gerência é designado ao relacionamento, trabalha_para é um tipo de relacionamento N:1 entre empregado e departamento, ambos têm participação total, controla é um tipo de relacionamento 1:N entre departamento e projeto. A participação de projeto é total, enquanto que a de departamento foi determinada como sendo parcial (após consulta aos usuários), Supervisão é um tipo de relacionamento 1:N entre empregado (supervisor) e empregado (supervisionado). Ambas as participações são parciais, já que os usuários mencionaram que nem todo empregado possui um supervisor e vice-versa.

Trabalha_em é um tipo de relacionamento M:N entre empregado e projeto, com o atributo horas indicando quantas horas um empregado trabalha em um projeto. Ambas as participações foram determinadas totais. Dependentes_de é um tipo de relacionamento N:1 entre dependente e empregado, sendo um relacionamento identificador para a entidade fraca dependente. A participação de empregado é parcial, enquanto que a de dependente é total.

Modelo ERE

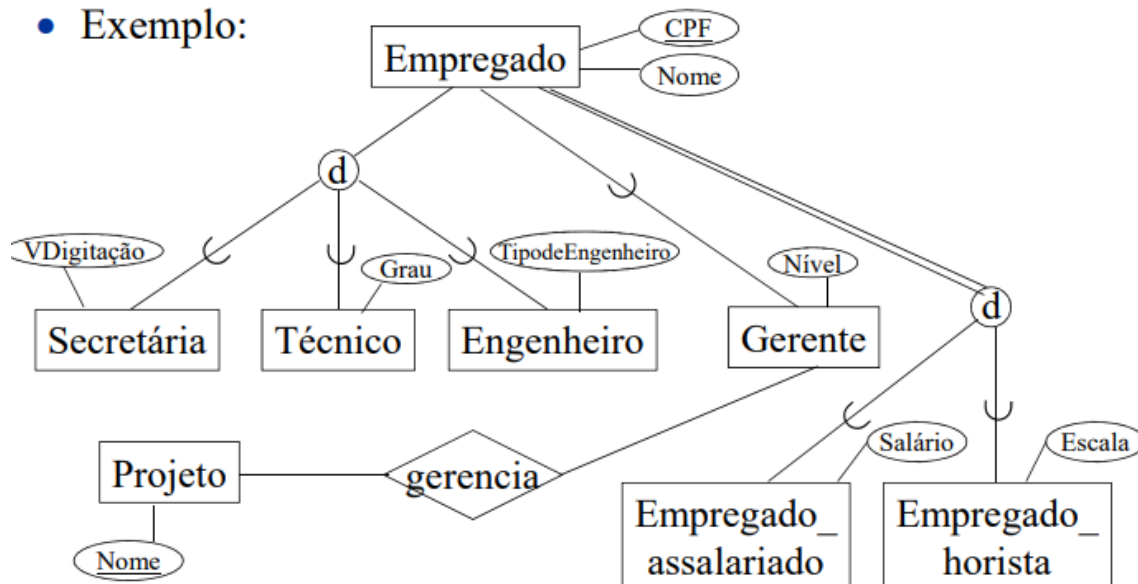
Herança

Em muitos casos, um tipo de entidade tem vários subconjuntos de entidades que são significativos para a aplicação, cada um dos subconjuntos é chamado de subclasse do tipo de entidade Empregado, e o tipo de entidade Empregado é chamado de superclasse, não é necessário que toda entidade em uma superclasse seja membro de alguma subclasse, como uma entidade na subclasse representa a mesma entidade no mundo real da superclasse, ela possui valores de seus atributos específicos bem como valores de seus atributos como um membro da superclasse, a entidade também herda todos os relacionamentos dos quais a superclasse participa. Também, pode ter seus próprios relacionamentos como subclasse.

Especialização

Especialização é o processo de definir um conjunto de subclasses a partir de um tipo de entidade (superclasse).

- Exemplo:



Algumas características do exemplo anterior são: O conjunto de subclasses {Secretária, Técnico e Engenheiro} é uma especialização de Empregado que distingue as entidades pelo tipo de trabalho, o conjunto de subclasses {Empregado_assalariado e Empregado_horista} é uma especialização de Empregado que distingue as entidades pela forma de pagamento, somente as entidades da subclasse Engenheiro possuem o atributo TipodeEngenheiro, os tipos de entidade Secretária, Técnico, Engenheiro, Gerente, Empregado_assalariado e Empregado_horista herdam os atributos da superclasse Empregado e somente as entidades da subclasse Gerente podem participar do tipo de relacionamento gerencia.

Generalização

Generalização é o processo inverso da especialização.

Restrições sobre Especialização/Generalização

Restrição de Disjunção: Disjunção: uma entidade da superclasse pode ser membro de, no máximo, uma das subclasses da especialização ou generalização. Sobreposição: uma entidade

da superclasse pode ser membro de mais de uma subclasse da especialização ou generalização. Restrição de Completude, total: toda entidade da superclasse deve ser membro de, pelo menos, uma subclasse da especialização ou generalização ou parcial: nem toda entidade da superclasse precisa ser membro de alguma subclasse da especialização ou generalização.

Herança Múltipla

Uma subclasse pode ser descendente de mais de uma superclasse. Neste caso, tem-se uma herança múltipla.

Tipo União ou Categoria

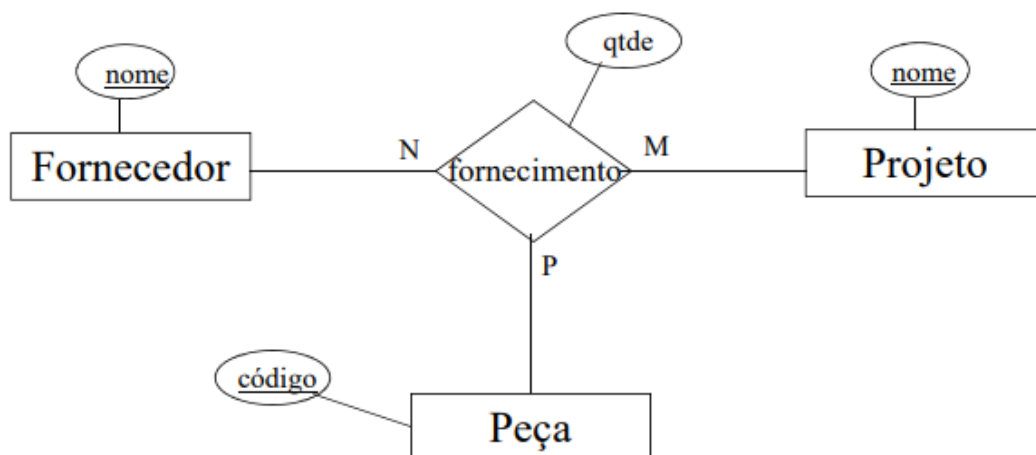
Em algumas aplicações, é necessário modelar um único relacionamento superclasse/subclasse com mais de uma superclasse, onde as superclasses representam tipos de entidades diferentes e a subclasse representa uma coleção de objetos que é a UNIÃO de instâncias das superclasses (ou um subconjunto da união). Esta subclasse é chamada de tipo união ou categoria. Uma categoria pode ser total ou parcial. No exemplo acima, a categoria Correntista é parcial (simbolizado pela linha simples), indicando que podem haver pessoas ou empresas que não são correntistas. Entretanto, todo correntista ou é uma pessoa ou é uma empresa.

Herança Múltipla versus Categoria

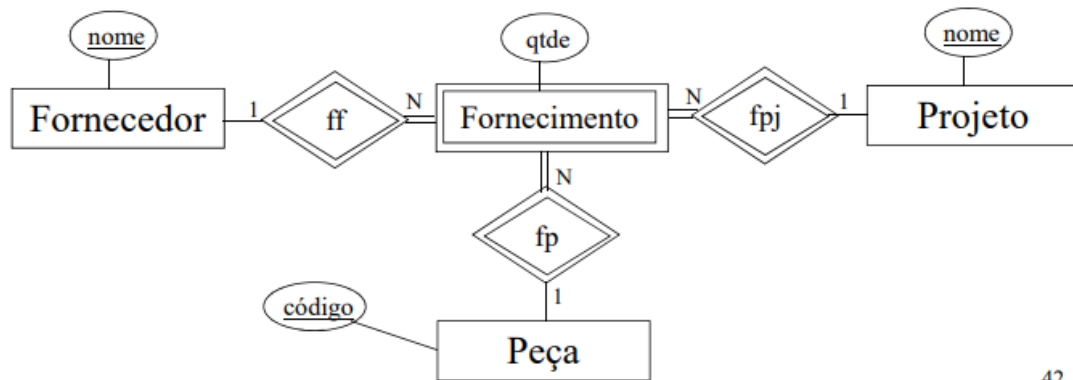
Comparação Herança múltipla versus Categoria, na herança múltipla, uma entidade na subclasse deve existir também em todas as superclasses. Ex.: uma entidade em Gerente_de_Engenharia existe também em Engenheiro, em Gerente e em Empregado_Assalariado, uma categoria é um subconjunto da união de suas superclasses. Assim, uma entidade na subclasse (categoria) deve existir somente em uma das superclasses, na herança múltipla, uma entidade na subclasse herda todos os atributos de todas as suas superclasses e na categoria, uma entidade na subclasse herda os atributos somente da superclasse da qual ela pertence.

Relacionamento Ternário

Um tipo de relacionamento ternário é um tipo de relacionamento de grau 3, ou seja, que possui 3 tipos de entidades participantes. Considere o exemplo abaixo, onde fornecedores fornecem peças a projetos.



Uma representação equivalente ao tipo de relacionamento ternário, usando apenas tipos de relacionamentos binários, deve criar um tipo de entidade fraca, sem chaves parciais e com três relacionamentos identificadores. Uma instância do tipo de entidade fraca Fornecimento é identificada pela combinação de suas três entidades proprietárias.

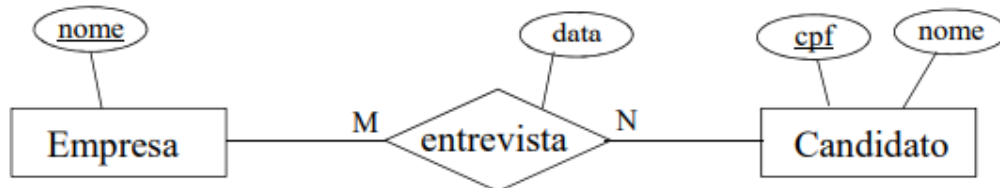


42

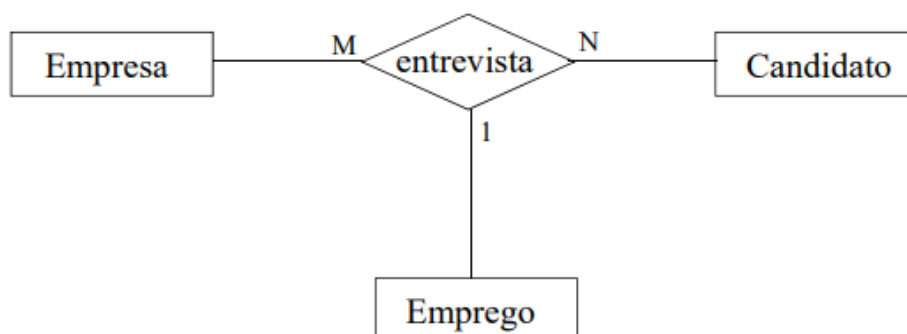
Agregação

Agregação é uma abstração usada para construir objetos compostos a partir de seus componentes, informalmente, uma agregação corresponde a um tipo de relacionamento que, devido aos requisitos de modelagem, deve ser transformado em um tipo de entidade.

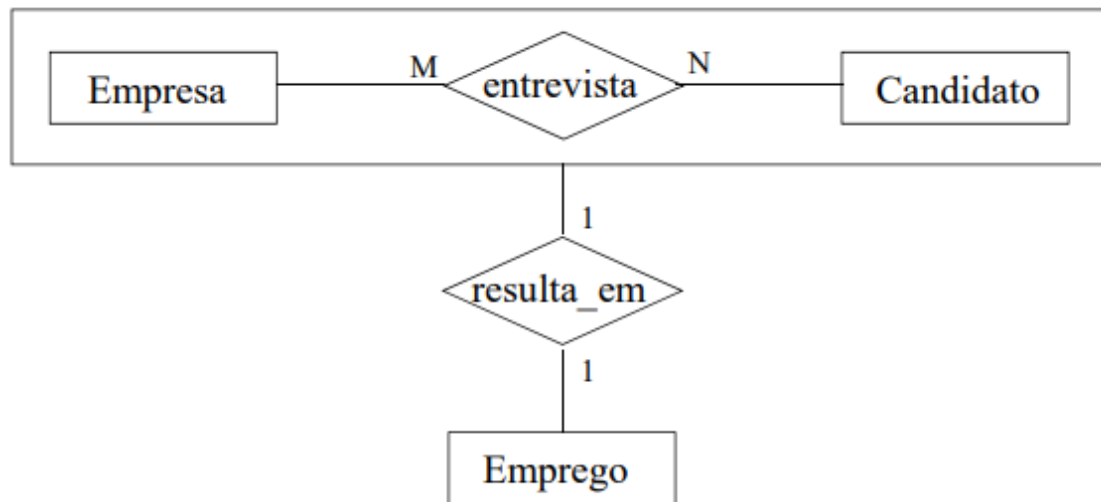
Exemplo: o diagrama abaixo representa um banco de dados de entrevistas para emprego de uma agência de RH.



Suponha que algumas entrevistas resultam em oferta de emprego e outras não. Como representar este fato? Usar um tipo de relacionamento ternário não é correto, pois requer que cada instância do relacionamento entrevista tenha uma oferta de emprego garantida.



Uma maneira de representar essa situação é criar uma entidade agregada de nível mais alto composta por Empresa, Candidato e entrevista e, então, relacioná-la a Emprego, como mostrado abaixo:



Outra maneira de representar a mesma situação é criando um tipo de entidade fraca Entrevista, como mostrado abaixo:

