

# Oportunidades de Desenvolvimento e Pesquisa sobre Projeto de Bancos de Dados usando IDEF1X

**Vinícius Medina KERN**

Professor da UNIVALI/Campus de São José/UNITEC-Grupo de Pesquisa em Tecnologia e Sistemas  
E-mail: kern@sj.univali.rct-sc.br

**Márcio Umberto BRAGAGLIA**

Mestrando do PPGEP-Programa de Pós-Graduação em Engenharia de Produção  
E-mail: marciob@floripa.com.br

## Resumo

IDEF1X é uma linguagem e método para modelagem da informação publicada como padrão norte-americano. Um novo processo de especificação da linguagem IDEF1X, usando a abordagem objeto-orientada, está em curso. Este artigo sumaria as principais diferenças entre a versão original de IDEF1X e a nova versão. Resultados de um esforço de desenvolvimento de ferramentas gráficas para projeto de bancos de dados usando IDEF1X são apresentados. Questões de pesquisa surgidas a partir deste esforço de desenvolvimento são aventadas e discutidas.

**Palavras-chaves:** IDEF1X, modelagem da informação, bancos de dados orientados a objeto

## Abstract

IDEF1X is an information modeling language and method published as American standard. This paper summarizes progresses in IDEF1X developed by a new specification team, using the object-oriented approach, in comparison to the original version of IDEF1X. Results from a development initiative on graphic tools for IDEF1X design are outlined. Research issues that arise from this development initiative are discussed.

**Keywords:** IDEF1X, information modeling, object databases

## INTRODUÇÃO

A força aérea norte-americana empreendeu, em meados da década de 1970, um esforço para aumentar a produtividade da indústria de manufatura através do uso sistemático de tecnologia computacional. Durante a realização deste trabalho, conhecido como programa ICAM (*Integrated Computer Aided Manufacturing*), reconheceu-se a necessidade da modelagem da informação para a construção de esquemas de bancos de dados.

O programa ICAM desenvolveu uma série de técnicas conhecidas como Métodos IDEF, que incluem (IEEE 1998):

- IDEF0, dedicado à modelagem dos processos ou atividades em um ambiente.
- IDEF1, dedicado à modelagem da estrutura e semântica da informação em um ambiente.
- IDEF2, dedicado à modelagem das características de comportamento dinâmico, variante no tempo, de um ambiente.

Dentre os três métodos, IDEF1 (*Integration Definition for Information Modeling*) é o que alcançou maior disseminação. IDEF2 foi descontinuado e IDEF0 ainda é usado em indústrias e instituições governamentais, embora não seja a técnica mais popular para

modelagem funcional. IDEF1 foi publicado em 1981, tendo como base o modelo relacional (Codd 1970) e o modelo Entidade-Relacionamento (Chen 1976).

O governo norte-americano publicou o padrão IDEF1X (IDEF1 *Extended*) baseado em IDEF1 e em outra especificação conhecida como LDDT (*Logical Database Design Technique*) (Brown 1982). Além de definir a sintaxe e semântica da linguagem, o padrão IDEF1X inclui uma definição da linguagem usando lógica de primeira ordem e a especificação de um método de modelagem.

As principais construções da linguagem IDEF1X são entidades, relacionamentos e atributos. Os relacionamentos são implementados segundo a abordagem do modelo relacional, através da migração de chaves. IDEF1X permite modelar as características (atributos) das entidades, mas não prevê a modelagem do comportamento (métodos) associado a entidades.

A orientação a objetos, por outro lado, é uma tecnologia que vem recebendo atenção crescente, tendo como característica a concepção de objetos, instâncias de classes (da mesma forma que ocorrências são instâncias de entidades) que encapsulam atributos e métodos. Relacionamentos, segundo esta abordagem, são estabelecidos através de ponteiros, sem migração de chaves (que funcionam, no modelo relacional, como “ponteiros lógicos”).

Por causa destes avanços, e pela demanda crescente pela aplicação da tecnologia de orientação a objetos em bancos de dados, um novo esforço de padronização foi lançado, com o objetivo de produzir uma especificação de IDEF1X adequada à modelagem de bancos de dados orientados a objeto (BDOO). Esta nova especificação, conhecida como IDEFObject (IEEE 1998), é debatida a seguir. Na seqüência, é relatada a experiência de desenvolvimento da interface gráfica IDEFClasses para projeto de BDOO, e são discutidas questões de pesquisa surgidas a partir deste esforço de desenvolvimento.

## A LINGUAGEM IDEFOBJECT

Uma nova especificação IDEF1X está em desenvolvimento pelo IEEE (*The Institute of Electrical and Electronics Engineers*). A nova especificação é conhecida como IDEFObject ou IDEF1X97, para diferenciar da especificação original, também conhecida como IDEF1X93.

A figura 1 ilustra as mudanças operadas sobre a versão original da linguagem para produzir a nova versão. Algumas das construções sintáticas de IDEF1X93 continuam sendo importantes em IDEFObject, tais como: classe (entidade), relacionamento e atributo. Outras construções são desnecessárias, como chave primária e relacionamento identificador.

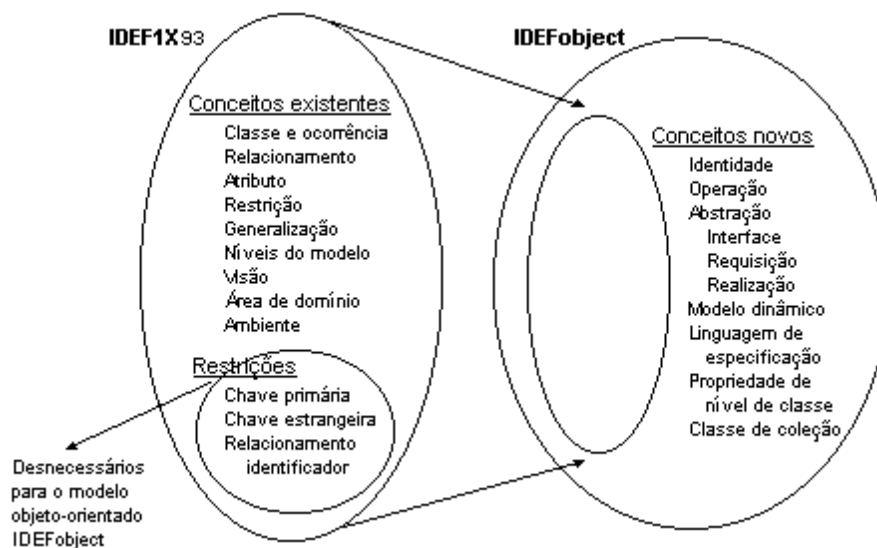


Figura 1 - Correspondência de conceitos nas duas versões de IDEF1X (IEEE 1998)

Chaves primárias são desnecessárias porque o conceito de identidade de objeto exclui a necessidade de haver atributos chaves cujo valor identifica uma ocorrência (objeto). Os

relacionamentos são estabelecidos através de ponteiros; não dependem de migração de chaves. Portanto, os relacionamentos identificadores (aqueles que incluem os atributos chaves da entidade genitora na chave da entidade filha) são desnecessários, também.

Embora desnecessárias, as construções citadas foram mantidas como parte da linguagem IDEFObject, para manter a compatibilidade retroativa. É possível construir modelos completamente objeto-orientados, bem como modelos que mantêm características da versão relacional. Esta possibilidade visa a garantir a transição gradual de uma abordagem para a outra.

As principais diferenças entre a abordagem objeto-orientada de IDEFObject e a abordagem baseada em dados e processos de IDEF1X93 são:

- As classes IDEFObject encapsulam atributos e métodos, enquanto as entidades IDEF1X93 descrevem apenas atributos.
- Não há processo independente em IDEFObject. Todo processo é disparado por um método de uma classe e manipula atributos de um objeto da mesma classe. Em IDEF1X93, os processos estão isolados e podem acessar quaisquer atributos.
- Não existe o conceito de identidade em IDEF1X93. É preciso que haja um ou mais atributos cujo valor identifica uma ocorrência. Em IDEFObject, dois objetos podem ter os mesmos valores de atributos. A identidade é intrínseca a cada objeto.

As principais construções de IDEFObject são (IEEE 1998):

- Classe: uma abstração sobre o conhecimento (atributos) e comportamento (método) de um conjunto de objetos de mesma natureza. Uma classe pode ser de estado, quando representa objetos mutáveis, ou de valor, quando representa objetos que são puramente valores, imutáveis. Classes de valor têm significado semelhante ao de domínios em IDEF1X93.
- Generalização: uma abstração que estabelece a herança entre subclasses e uma superclasse.
- Relacionamento: uma conexão entre duas classes de estado.
- Responsabilidade: o conjunto de características representadas pelos atributos, métodos e regras de cada objeto de uma classe.
- Requisição: uma mensagem enviada a partir de um objeto emissor para um objeto receptor.
- Realização: uma sentença lógica que estabelece as condições necessárias e suficientes para que uma responsabilidade seja atendida. A realização de uma responsabilidade específica como esta é atingida.
- Construções de infra-estrutura: elementos de suporte à organização e documentação: visões, níveis, ambientes, glossários, modelos de informações.

## DESENVOLVIMENTO DE FERRAMENTAS GRÁFICAS

O documento original de IDEF1X (NIST 1993) previa a implementação de ferramentas gráficas para o suporte à modelagem da informação. As ferramentas comerciais existentes suportam, em parte, a linguagem IDEF1X original. ERwin, a interface gráfica mais difundida, não suporta a documentação da equipe de projeto, da lista indexada de material-fonte, e da declaração de propósito e escopo, documentos fundamentais para a interpretação de um modelo da informação.

Na UNIVALI, o desenvolvimento de ferramentas gráficas para modelagem da informação com suporte completo a IDEF1X tem sido tema de alguns trabalhos de conclusão do Curso de Ciência da Computação. Scheidt (1998) abordou o problema de oferecer ao modelador uma interface gráfica de apoio ao projeto usando IDEF1X em sua versão original (NIST 1993). Bragaglia (1999) tratou do projeto gráfico de bancos de dados objeto-orientados, conforme a nova especificação de IDEF1X (IEEE 1998).

O primeiro trabalho teve como plataforma de desenvolvimento o ambiente Delphi, que usa a linguagem de programação Object Pascal. Delphi é bastante usado para a construção de aplicações de processamento de dados, mas revelou-se pouco apropriada nas etapas mais avançadas deste projeto, pois a construção de uma ferramenta gráfica requer funcionalidades

que estão muito além daquelas providas (ainda que se procedesse a uma busca, aquisição e experimentação de componentes dedicados à manipulação gráfica).

O segundo trabalho foi implementado usando Visual C++. O desenvolvimento em C++ demandou esforço maior na etapa inicial, principalmente porque era objetivo criar uma ferramenta portátil para outras plataformas de software e hardware. Para tanto, os programas fontes foram codificados em ANSI C++. O software IDEFClasses (Bragaglia 1999), ilustrado na figura 2, é o principal resultado deste trabalho, que implementa uma interface gráfica de apoio à modelagem da informação com IDEFObject.

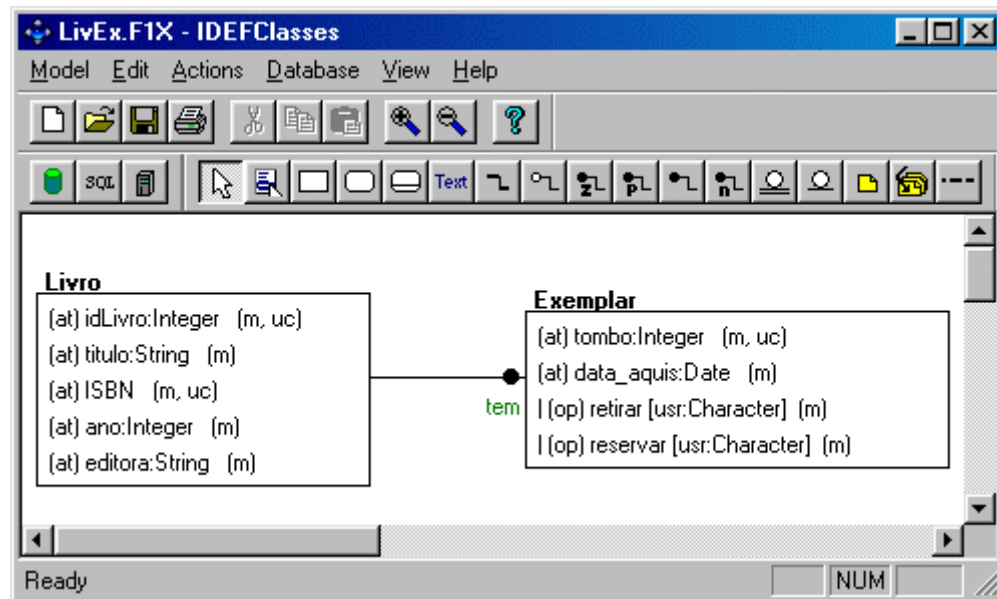


Figura 2 - Janela de IDEFClasses, interface para modelagem IDEFObject

A versão atual (Beta 6, de Junho de 1999) de IDEFClasses tem como características um tamanho de arquivo executável e um tempo de resposta aproximadamente iguais a um décimo daqueles relativos às primeiras versões escritas em Delphi. IDEFClasses permite ao modelador de banco de dados executar as seguintes tarefas:

- Desenhar classes e suas responsabilidades (atributos, métodos, relacionamentos), sem limite de tamanho do modelo, mantendo a sintaxe da linguagem IDEFObject.
- Especificar e editar detalhes acerca das classes e responsabilidades.
- Manipular graficamente as construções do modelo.
- Incluir notas e referências a notas, bem como comentários na forma de caixas de texto.
- Gravar o modelo e imprimir cópia do diagrama.
- Especificar um sistema de gerência de banco de dados e um nome de usuário para a geração de definição de dados a partir do modelo.
- Gerar um *script* SQL com a definição de dados para um modelo e criar uma conexão com o sistema de gerência, implementando o catálogo do sistema.

Além das funcionalidades já disponíveis, uma série de desenvolvimentos futuros virá a complementar a interface. Em especial, o modelador deve ser capaz de realizar as seguintes tarefas com auxílio de IDEFClasses:

1. Documentar a Fase 0 (iniciação de projeto), incluindo a declaração de propósito e escopo, a definição da equipe de projeto, e a organização da lista indexada de material-fonte.
2. Especificar as definições relativas às fases 1, 2, 3 e 4 (entidades, relacionamentos, chaves e atributos).

3. Projetar classes e relacionamentos compatíveis com a especificação relacional de IDEF1X, utilizando chaves e migração de chaves.
4. Representar visões e áreas de domínio em um modelo.
5. Reorganizar o leiaute de um modelo através de uma função de redesenho automático.
6. Produzir um modelo da informação a partir da engenharia reversa de uma definição de banco de dados existente.

Em adição às atividades de desenvolvimento para o suporte completo à linguagem e método IDEF1X, IDEFClasses enseja oportunidades para a condução de pesquisa inovadora envolvendo várias disciplinas, especialmente modelagem da informação, teoria de grafos e processamento de linguagem natural. Alguns tópicos de pesquisa são discutidos a seguir.

## TÓPICOS DE PESQUISA

Vários autores (Chen *et al.* 1998, Kusiak *et al.* 1997) discutem direções de pesquisa promissoras na área de modelagem da informação. A experiência da utilização de IDEF1X como linguagem e técnica de modelagem da informação e a construção de IDEFClasses oportunizaram o surgimento de algumas perguntas de pesquisa. Especificamente em relação a IDEFClasses, as seguintes questões devem fazer parte dos futuros esforços de investigação:

1. Como oferecer suporte à aprendizagem da linguagem e método IDEF1X a peritos em área de interesse de modelagem (isto é, aqueles que revisam e aceitam um modelo e, portanto, precisam interpretá-lo) e modeladores novatos?
2. Quais as técnicas melhor adequadas à tarefa de redesenho de um modelo, de modo que seja possível oferecer em tempo adequado à equipe de projeto um diagrama IDEF1X cujo leiaute facilita a leitura?
3. Como detectar a existência de relacionamentos de caminhos duais em um modelo e conduzir a ação do modelador para detalhar as regras do negócio associadas?

A primeira pergunta diz respeito à correspondência na forma de expressar regras de negócios em IDEF1X e em linguagem natural. Por exemplo, algumas das regras do negócio que podem ser vertidas para o português a partir do modelo na figura 3 são: “um departamento administra zero, um, ou vários cursos”; “um curso não pode existir a menos que exista o departamento que o administra”; “é possível que a administração de um curso seja transferida de um para outro departamento”; etc. Como classificar estas regras? Como conduzir peritos na aprendizagem da interpretação de modelos? Como orientar modeladores novatos de forma a facilitar sua aprendizagem sobre as regras de negócio expressas em modelos IDEF1X?

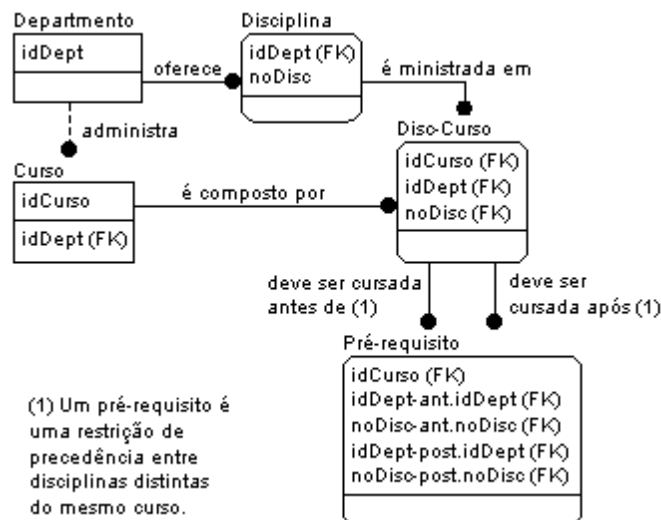


Figura 3 - Caminhos duais e asserção de caminhos em IDEF1X

A segunda pergunta de pesquisa corresponde ao item 5 listado entre os desenvolvimentos futuros na seção anterior. A abordagem tradicional para o problema do redesenho de um modelo da informação IDEF1X usa teoria de grafos, atribui níveis de precedência às entidades, e impõe a regra “entidade genitora ou genérica acima e/ou à esquerda da entidade filha ou categoria” (o modelo da figura 3, por exemplo, atende a esta regra). Uma inovação digna de investigação é a abordagem do problema através de algoritmos genéticos, como proposto por Mangano (1994). Esta nova abordagem poderia ser comparada com a tradicional quanto ao desempenho e eficácia para a obtenção de redesenhos legíveis.

O problema da detecção e projeto de relacionamentos de caminhos duais, objeto da terceira pergunta, pode ser explicado com o apoio da figura 3, que ilustra a organização de currículos de cursos de graduação em uma universidade com organização departamental, onde cada curso tem um currículo composto por disciplinas oferecidas por vários departamentos, e há restrições de precedência de disciplinas. Pode-se observar que há, para a entidade Pré-requisito, dois caminhos de relacionamentos específicos desde Curso, Disciplina e Disc-Curso, e quatro caminhos desde Departamento.

Relacionamentos de caminhos duais podem ser iguais, diferentes, ou indeterminados, conforme as ocorrências específicas da entidade ancestral associadas a uma ocorrência da entidade descendente em questão sejam necessariamente iguais, diferentes, ou não haja qualquer imposição. Por exemplo, na figura 3, uma ocorrência de Pré-requisito deve estar associada, necessariamente, a ocorrências distintas de Disc-Curso, bem como a ocorrências distintas de Disciplina. Portanto, os caminhos duais são diferentes. Já os caminhos duais entre Pré-requisito e Curso são iguais, pois só faz sentido haver restrição de pré-requisito entre disciplinas em um mesmo currículo de curso. Por outro lado, os caminhos duais entre Pré-requisito e Departamento e entre Disc-Curso e Departamento são indeterminados, pois um curso administrado por um departamento pode ter disciplinas oferecidas pelo mesmo ou outros departamentos, sem restrição.

As regras de negócios envolvendo caminhos duais em IDEF1X são expressas usando migração de chaves e notas. No último caso, diz-se que a nota enuncia uma asserção de caminho. No modelo da figura 3, os caminhos diferentes foram projetados usando migrações de chaves distintas, o que foi expresso através da criação de diferentes nomes de papéis para os atributos chaves migrados (por exemplo, noDisc-ant e noDisc-post para o atributo-base noDisc) e uma nota que registra que as chaves migradas devem, necessariamente, ter valores diferentes em uma ocorrência.

Os caminhos iguais na figura 3 foram implementados através da unificação dos dois atributos chaves idCurso em Pré-requisito, de forma que não há como especificar cursos diferentes em uma restrição de pré-requisito. Os caminhos indeterminados não necessitam de qualquer anotação, pois representam a falta de restrição quanto aos caminhos duais. Nos casos de caminhos duais nos quais a chave da entidade genitora migra até a chave da entidade descendente, nomes de papéis distintos (sem usar nota) resolvem o problema de permitir a associação com ocorrências distintas da entidade ancestral.

A solução tradicional para a detecção de caminhos duais usa técnicas da teoria de grafos, concebendo o modelo da informação como um digrafo acíclico (isto é, um grafo dirigido e sem ciclos, com a exceção de auto-relacionamentos opcionais um-para-vários). Também neste caso, no entanto, é possível abordar o problema utilizando algoritmos genéticos, segundo Mangano (1994), e comparar esta solução com a tradicional.

Quanto à forma de oferecer suporte ao projeto dos caminhos duais, IDEFClasses deverá interrogar o modelador sobre o caráter igual, diferente ou indeterminado dos caminhos. Porém, a forma de conduzir esta interação leva a outra questão que merece investigação: é válido particionar o problema, reduzindo uma malha de caminhos duais a uma configuração mais simples?

No exemplo da figura 3, é válido analisar apenas os caminhos duais entre Disc-Curso e Pré-requisito, e entre Departamento e Disc-Curso? Esta questão necessita de uma abordagem formal, usando Lógica. A alternativa é proceder a um

detalhamento exaustivo dos caminhos duais para cada par de entidades no modelo, sendo que, apenas na figura 3, há caminhos duais para cinco pares de entidades.

## Tríades

Um tipo peculiar de relacionamento de caminhos duais é a estrutura conhecida como tríade, ilustrada na figura 4, onde um dos caminhos é estabelecido por um relacionamento direto entre a entidade ancestral e a descendente. Em alguns destes casos, o relacionamento direto é redundante e deve ser eliminado do modelo (figura 4 (a)), enquanto em outros o relacionamento não pode ser eliminado sob risco de perda de informação. No exemplo da figura 4 (b), a Filial onde está lotado um Funcionário pode mudar, mas a Filial que emitiu determinada Nota Fiscal nunca muda, nem o Funcionário que imprimiu (processou) a Nota Fiscal. Portanto, os caminhos são indeterminados.

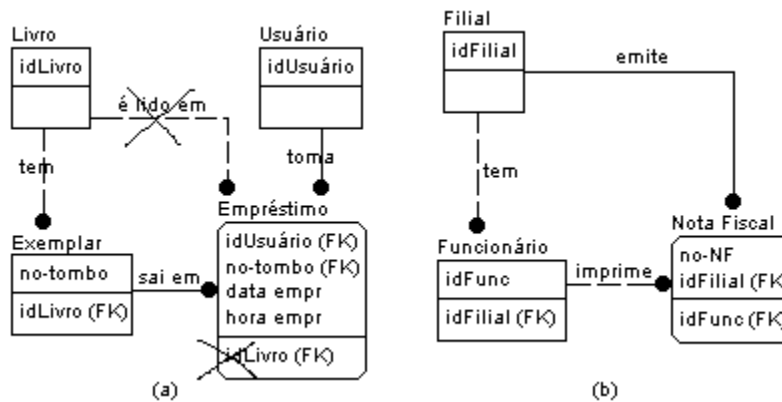


Figura 4 - Tríades com (a) e sem (b) relacionamento redundante

Modeladores novatos têm maior propensão para projetar relacionamentos redundantes (NIST 1993). Uma ferramenta gráfica de modelagem como IDEFClasses deve oferecer suporte ao projeto de tríades, de forma a facilitar o aprendizado e aumentar a produtividade do esforço de modelagem.

Uma vez detectada uma tríade, seu detalhamento deve levar em conta as dependências de identificador das entidades descendentes, e também o caráter dinâmico ou estático (isto é, mutável ou imutável) das ocorrências dos relacionamentos envolvidos na tríade, para determinar se os caminhos são iguais, diferentes ou indeterminados.

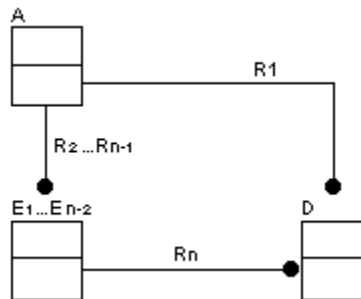


Figura 5 - Tríade hipotética

A figura 5 representa uma tríade genérica, com  $n$  entidades e relacionamentos participantes. A tabela 1 determina como detalhar o projeto de tríades para cada caso, em função dos tipos de relacionamentos – identificador ou categorização (casos em que a chave da entidade filha ou categoria depende da chave da entidade pai ou genérica), ou não-identificador (caso em que a chave da entidade filha ou categoria não inclui qualquer chave de entidade pai ou genérica).

O esquema para a análise e projeto de tríades apresentado na tabela 1 foi obtido de modo empírico. O resultado obtido, entretanto, não oferece uma solução genérica para o problema de projeto de caminhos duais, que continua sendo tema para investigação.

Tabela 1 - Análise e projeto de tríades em IDEF1X

Análise do tipo dos relacionamentos			Projeto dos caminhos duais ou eliminação do relacionamento direto		
R1	R2...Rn-1	Rn	Caminhos iguais	Caminhos diferentes	Caminhos indeterminados
Identificador ou categorização	Todos identificadores ou categ.	Identificador ou categ.	Eliminar R1	Usar nomes de papéis distintos; Usar nota para enunciar caminhos diferentes	Usar nomes de papéis distintos
Identificador ou categorização	Todos identificadores ou categ.	Não-identificador	Usar mesmo ou nenhum nome do papel; unificar chave da ancestral na chave da descendente	Usar nomes de papéis distintos; Usar nota para enunciar caminhos diferentes	Usar nomes de papéis distintos
Identificador ou categorização	Pelo menos um não-identificador	(Não importa o tipo)	Usar nota para enunciar caminhos iguais	Usar nota para enunciar caminhos diferentes	Usar apenas as construções básicas de IDEF1X
Não-Identificador	Todos identificadores ou categ.	Identificador ou categ.	Eliminar R1	Usar nomes de papéis distintos; Usar nota para enunciar caminhos diferentes	Usar nomes de papéis distintos
Não-Identificador	Todos identificadores ou categ.	Não-Identificador	Eliminar R1	Usar nomes de papéis distintos; Usar nota para enunciar caminhos diferentes	Usar nomes de papéis distintos
Não-Identificador	Pelo menos um não-identificador	(Não importa o tipo)	Eliminar R1	Usar nota para enunciar caminhos diferentes	Usar apenas as construções básicas de IDEF1X

## CONSIDERAÇÕES FINAIS

Este artigo discutiu brevemente a evolução da linguagem IDEF1X para modelagem da informação. O desenvolvimento de uma nova especificação baseada na orientação a objetos, conhecida como IDEFobject, foi introduzido. As principais diferenças entre as versões relacional e objeto-orientada de IDEF1X foram sumariadas.

O desenvolvimento de ferramentas gráficas para modelagem da informação com IDEF1X foi relatado, apresentando o estado atual e as oportunidades de desenvolvimento futuro da interface IDEFClasses. Foram apresentados tópicos de pesquisa envolvendo a analogia entre linguagem natural e IDEF1X no enunciado de regras de negócios, a comparação de técnicas alternativas para o redesenho de um modelo, e o projeto de caminhos duais. Resultados iniciais sobre o projeto de caminhos duais, nos casos de tríades, foram detalhados.

## REFERÊNCIAS

- M.U. Bragaglia. **Software para Modelagem de Bancos de Dados Objeto-Orientados usando IDEFobject**. Trabalho de Conclusão do Curso de Ciência da Computação. São José: UNIVALI, Junho de 1999.
- R.G. Brown. **Logical Database Design Techniques**. Mountain View, CA: The Database Design Group, Inc., 1982.
- P.P. Chen. "The Entity-Relationship Model - Toward a Unified View of Data". **ACM Transactions on Database Systems** 1 (1), pp. 9-36, 1976.
- P.P. Chen, B. Thalheim & L.Y. Wong. "Future Directions of Conceptual Modeling". In: **Conceptual Modeling: Current Issues and Future Directions**. P.P. Chen et al. (eds.). Berlin: Springer-Verlag, Lecturing Notes in Computer Sciences, No. 1565, 1998, pp. 294-308.
- E.F. Codd. "A Relational Model of Data for Large Shared Data Banks". **Communications of the ACM** 13 (6), 1970, pp. 377-87.
- IEEE (Institute of Electrical and Electronic Engineers) IDEF1X Standards Working Group. **Standard for Conceptual Modeling Language Syntax and Semantics for IDEF1X97 (IDEFobject)**. IEEE 1320.2 Standards Committee, document P1320.2. Release Draft 0.91 – May 1, 314 p., 1998.
- A. Kusiak, T. Letsche & A. Zakarian. **Data modelling with IDEF1x**. International Journal of Computer Integrated Manufacturing, vol.10, no.6, p. 470-86. Iowa City, USA, Nov.-Dec. 1997.
- S. Mangano. "Algorithms for Directed Graphs: A Unique Approach Using Genetic Algorithms". **Dr. Dobb's Journal**, April 1994.
- NIST (National Institute of Standards and Technology). **Federal Information Processing Standards Publication 184. Integration Definition for Information Modeling (IDEF1X)**. Gaithersburg, MD (USA), December 1993.
- L.S. Scheidt. **Desenvolvimento de uma Ferramenta Gráfica de Modelagem da Informação Baseada na Metodologia IDEF1X**. Trabalho de Conclusão do Curso de Ciência da Computação. São José: UNIVALI, Novembro de 1998.