

UNIVERSIDADE DE BRASÍLIA

Faculdade do Gama

Sistemas de Banco de Dados 2

Tecnologias de Banco de Dados (TI-BD)

Bancos de Dados Dedutivos

Guilherme Brito Vilas – Bôas - 190108011

Brasília, DF

2023

1. Introdução

Com a popularização da adoção de Sistemas de Bancos de Dados, os usuários necessitavam cada vez mais de funcionalidades adicionais tendo por finalidade a facilitação da implementação de aplicações mais avançadas e complexas, surgindo então diversas outras tecnologias de Bancos de Dados, trazendo essas implementações necessárias para os diversos casos de uso.

Neste trabalho, a tecnologia que será abordada é a: Bancos de Dados Dedutivos, uma área que se interliga com bancos de dados, lógica e inteligência artificial.

2. Definição

De acordo com Ramez Elmasri e Shamkant B. Navathe, autores do livro “Sistemas de Bancos de Dados”, um sistema de banco de dados dedutivo é “um sistema que inclui capacidade para definir regras (dedutivas), que podem inferir informações adicionais sobre os fatos que estão armazenados em um banco de dados”, ou seja, esses sistemas, tendo por base a lógica matemática, utilizam um mecanismo de inferência para obter informações que não necessariamente estão armazenadas no banco.

Essa tecnologia se baseia muito no paradigma lógico de programação, tendo como representante principal a linguagem de programação Prolog.

2.1 Datalog

Esses sistemas de bancos geralmente utilizam a linguagem de programação Datalog.

De acordo com Stefano Ceri, no artigo: “What You Always Wanted to Know About Datalog (And Never Dared to Ask)”, Datalog é uma linguagem de consulta de banco de dados baseada no paradigma lógico de programação, que se utiliza de predicados e seus respectivos argumentos, sendo definido da seguinte forma:

$$PREDICADO(a_1, a_2, a_3, \dots, a_n)$$

sendo cada a , os argumentos do predicado.

Caso os argumentos se tratem de valores constantes podemos chamar o predicado de fato, caso contrário ele pode ser uma consulta ou parte de uma regra ou restrição. A seguir, os conceitos de Fato e Regra serão introduzidos.

2.2 Fatos

Definidos como sentenças que podem ser tratadas como verdade, uma constatação do mundo real, por exemplo: “João é aluno da UnB”. Os fatos são semelhantes ao modo como relações são especificadas em bancos de dados relacionais, porém não são incluídos os nomes dos atributos, sendo a posição que eles ocupam o determinante para a semântica do fato.

- Ex: Vamos supor que queremos armazenar que fernando SUPERVISIONA joão. Em um banco de dados dedutivo, podemos guardar o fato da seguinte forma:
 - SUPERVISIONA(fernando, joão).

2.3 Regras

Especificam relações virtuais que não estão armazenadas, mas podem ser formadas com base nos fatos, baseado na aplicação dos mecanismos de inferência. Se assemelham bastante as visões relacionais, porém diferente das visões, as regras podem envolver recursão.

- Ex: Se baseando em vários fatos SUPERVISIONA(a,b), sendo que cada pessoa a supervisiona uma pessoa b, podemos definir que se uma pessoa x é superior de uma pessoa y, então isso implica que x supervisiona y. Essa regra pode ser escrita da seguinte forma:
 - SUPERIOR(X,Y) :- SUPERVISIONA(X, Y)
 - :- , leia-se 'implica'
- Um exemplo de regra recursiva:
 - SUPERIOR(X,Y):-SUPERVISIONA(X,Z) , SUPERIOR(Z,Y)
 - Essa regra pode ser entendida como: Se X é superior de Y, então X supervisiona uma certa pessoa Z e (representado pela ',') Z é superior de uma pessoa Y.

Abaixo, podemos ver como uma hierarquia de supervisão de uma empresa pode ser definida em Datalog:

FATOS

- SUPERVISIONA(fernando,joao).
- SUPERVISIONA(fernando,ronaldo).
- SUPERVISIONA(fernando,joice).
- SUPERVISIONA(jennifer,alice).
- SUPERVISIONA(jennifer,andre).
- SUPERVISIONA(jorge,fernando).
- SUPERVISIONA(jorge,jennifer).

REGRAS

- $\text{SUPERIOR}(X,Y) :- \text{SUPERVISIONA}(X,Y).$
- $\text{SUPERIOR}(X,Y) :- \text{SUPERVISIONA}(X,Z) , \text{SUPERIOR}(Z,Y).$
- $\text{SUBORDINADO}(X,Y) :- \text{SUPERIOR}(Y,X).$

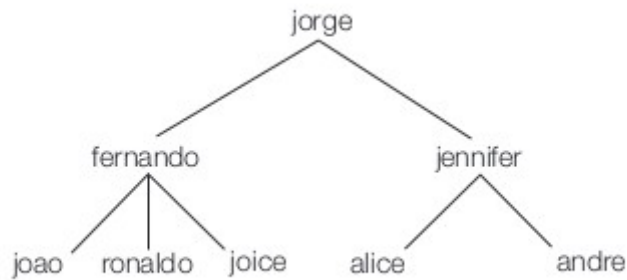


Figura 1: Árvore formada a partir dos fatos definidos acima

Fonte: Livro: Sistemas de Bancos de Dados

2.4. Consultas

A partir dos fatos e especificação das regras, as consultas podem ser feitas da seguinte forma:

- $\text{SUPERIOR}(\text{jorge}, Y)?$
 - sendo responsável pela recuperação de todas as tuplas onde jorge é o superior.
- $\text{SUPERIOR}(\text{jorge}, \text{joice})?$
 - Caso jorge seja superior de joice, ele retornará a tupla $\text{superior}(\text{jorge}, \text{joice})$, caso contrário o retorno será vazio.

3. Objetivo

Essa tecnologia de Banco de Dados tem por objetivo a derivação de novos conhecimentos e informações a partir de dados existentes por meio das regras.

4. Vantagens

Uma das maiores vantagens dessa tecnologia, em contra-partida da tecnologia relacional de SGBD, é a sua incrível expressividade. Nessa tecnologia, é possível que se ateste um conhecimento geral sobre um grupo todo de objetos sem que se saiba exatamente os indivíduos presentes nesse grupo.

Por exemplo, voltando ao exemplo mostrado no tópico 2, queremos obter todos os profissionais os quais jorge é superior.

- Na abordagem relacional, pode ser criada uma view para que se obtenha os superiores:
 - `CREATE VIEW jorge_eh_superior as SELECT nome FROM PROFISSIONAIS p where p.superior = 'jorge';` (Esquemas meramente ilustrativos)
 - Porém, como a view não aceita recursividade, só será retornado as tuplas onde jorge supervisiona diretamente, ou seja Fernando e Jennifer.
- Já na abordagem deducional, como as regras aceitam recursividade, pode-se então obter em todos os níveis quais são os funcionários os quais jorge é superior. (regras definidas no tópico 2).

Outra importante vantagem é o uso da lógica. A lógica é uma linguagem formal para expressar conhecimento, assim como um conjunto de regras que ditam como novas sentenças verdadeiras podem ser derivadas de outras conhecidas. Assim, a lógica provê um framework unificado para modelagem de dados e dedução de conhecimento bastante próxima do mundo real.

Uma outra vantagem importante a ser citada é a restrição explícita dos possíveis estados que o banco de dados pode estar. Em sistemas convencionais, esse tipo de conhecimento é implantado em trechos de programas que controlam a manipulação de dados, necessitando de um conhecimento procedural prévio, sendo assim mais difícil de entender e modificar (manutenção custosa).

5. Desvantagens

As maiores desvantagens decorrem da falta de popularidade dessa tecnologia, portanto, dentre estas desvantagens podemos citar: a falta de suporte amplo, uma comunidade pequena e , coisas que não acontecem com a tecnologia relacional.

Podemos tomar como uma simples comparação: a presença de 659 mil questões no Stackoverflow sobre 'mysql' e 500 questões sobre 'datalog'.

6. Bancos que incorporam essa tecnologia

Os sistemas mais populares desse tipo de tecnologia são as aplicações construídas em LDL e VALIDITY.

6.1 LDL

O projeto LDL foi iniciado em 1984, sendo o primeiro sistema de banco de dados dedutivo disponível em larga escala. Foi distribuído em universidades e para o consórcio de microeletrônica e computadores (MCC). Alguns dos domínios de aplicações que se beneficiam desse sistema: Modelagem empresarial, teste de hipóteses e drenagem de dados.

6.2 VALIDITY

O Projeto VALIDITY se apresenta como um Banco de dados orientado a objetos dedutivo (DOOD). Como um Sistema DOOD, o VALIDITY combina as capacidades dedutivas com a habilidade de manipular objetos complexos utilizando da linguagem DEL (Datalog Extended Language). Com esse sistema, pode ser construída varias aplicações, como por exemplo: Comércio Eletrônico, Processos baseados em Regras, Programação Concorrente e “Descobrimeto de Conhecimento”.

7. Referência Bibliográficas

ELMASRI, Ramez; NAVATHE, Shamkant B.. **Sistemas de Banco de Dados**. 7. ed. S. L: Pearson Universidades, 2019. 1152 p.

CERI, Stefano *et al.* What You Always Wanted to Know About Datalog: (and never dared to ask). **IEEE Transactions On Knowledge And Data Engineering**, S.L, v. 1, n. 1, p. 146-166, mar. 1989. Disponível em: <https://www2.cs.sfu.ca/CourseCentral/721/jim/DatalogPaper.pdf>. Acesso em: 14 abr. 2023.

BRISCHKE, Eduardo João Marcondes Marcelo. **Banco de Dados Dedutivos**. 2005. 31 f. Monografia - Curso de Bacharelado em Informática, Centro de Ciências Exatas e Tecnológicas, Universidade Estadual do Oeste do Paraná, Cascavel, 2005.

NARDON, Fabiane Bizinella. **Estudo e construção de um sistema gerenciador de banco de dados dedutivo**. 1996. 143 f. Dissertação (Mestrado) - Curso de Pós-Graduação em Ciência da Computação, Instituto de Informática, Universidade Federal do Rio Grande do Sul, S.L, 1996. Disponível em: <https://lume.ufrgs.br/bitstream/handle/10183/24105/000121783.pdf?sequence=1&isAllowed=y>. Acesso em: 13 abr. 2023.

MAHALI, Ankit. **Applications of Commercial Deductive Database Systems**. 2023. Disponível em: <https://www.geeksforgeeks.org/applications-of-commercial-deductive-database-systems/>. Acesso em: 14 abr. 2023.

RAMAKRSHINAN, Raghu; ULLMAN, Jeffrey D.. A Survey of Deductive Database Systems. **The Journal Of Logic Programming**. Nova York, p. 125-149. jan. 1995. Disponível em: [hhttps://www.sciencedirect.com/science/article/pii/0743106694000399](https://www.sciencedirect.com/science/article/pii/0743106694000399) Acesso em: 12 abr. 2023.