

NGÔN NGỮ TRUY VẤN DỮ LIỆU

Hình thức đánh giá

- 10%: Chuyên cần
- 20%: Điểm giữa kỳ
- 70%: BTL
 - *Quá trình – nộp bài*
 - *Test trực tiếp*

Mục đích

- Cung cấp khái niệm về truy vấn dữ liệu
- Giới thiệu về ngôn ngữ truy vấn SQL
 - DDL, **Data Definition Language**
 - DML, **Data Manipulation Language**
 - DCL, **Data Control Language**
- Thiết lập các lệnh truy vấn từ đơn giản đến phức tạp
- Áp dụng trên bài toán cụ thể

Nội dung

- Nhắc lại một số khái niệm cơ bản khi làm việc với cơ sở dữ liệu.
- Ngôn ngữ truy vấn
 - ngôn ngữ định nghĩa dữ liệu – DDL
 - ngôn ngữ thao tác dữ liệu – DML
 - *Đơn giản*
 - *Phức tạp*
 - ngôn ngữ điều khiển dữ liệu – DCL

Tài liệu tham khảo

- **Giáo trình Hệ QT Cơ sở dữ liệu**, Khoa CNTT, Trường CĐKT KT TW
- **Fundamentals of Database Systems**, 6rd+ edition, John Wiley & Sons, Inc,...2011.
- **Book Online SQL Server 2005**

Phần mềm sử dụng

- SQL Server

MỘT SỐ KHÁI NIỆM CƠ BẢN

NỘI DUNG CẦN NHẮC LẠI

- Các phép toán đại số quan hệ
 - *Chiếu, chọn, nối*
 - *Hợp, giao, trừ*
- Các bước thiết kế CSDL
- Mô hình cơ sở dữ liệu quan hệ, các thành phần: *bảng, trường, bản ghi*
 - Mối liên kết
 - Ràng buộc
 - Dạng chuẩn

■ Quan hệ

- Là tập hợp các thông tin lưu trữ trong CSDL được tổ chức dưới dạng **bảng 2 chiều**, gồm:
 - **Tên quan hệ** \Leftrightarrow **Table**
 - **Tập hợp các cột** \Leftrightarrow *các thuộc tính của quan hệ* \Leftrightarrow *Field*
 - * Cố định, Được đặt tên,
 - * Có kiểu dữ liệu xác định, các giá trị trong một cột phải cùng kiểu,
 - **Tập hợp các dòng** \Leftrightarrow *các bản ghi của quan hệ* \Leftrightarrow *Record*

■ Liên kết

- Một – một
- Một – nhiều

■ Ràng buộc dữ liệu

- Ràng buộc khóa (*Key Constraints*)
- Ràng buộc miền giá trị (*Domain Constraints*)
- Ràng buộc liên bộ
- Ràng buộc liên thuộc tính
- Ràng buộc toàn vẹn tham chiếu (*Referential Integrity Constraints*)

Các phép toán đại số quan hệ

■ Chọn

- Phép chọn được sử dụng để chọn một tập hợp các bộ thoả mãn điều kiện chọn từ một quan hệ. Ta có thể xem phép chọn như một bộ lọc, nó chỉ giữ lại các bộ thoả mãn điều kiện đặt ra

Các phép toán đại số quan hệ

- Chọn: $\sigma_{\langle \text{Mã số ĐV} = 4 \rangle \text{ AND } \langle \text{lương} > 3000 \rangle \text{ OR } \langle \text{Mã số ĐV} = 5 \rangle \text{ AND } \langle \text{lương} > 3500 \rangle}$ (NHÂNVIÊN)

Mã số NV	Họ tên	Tên	Ngày sinh	Địa chỉ	Giới tính	Lương	Mã số NG	Mã số ĐV
NV002	Trần Đức	Nam	14/02/66	Hà nội	Nam	4000	NV061	5
NV014	Phạm	Bằng	26/06/52	Bắc ninh	Nam	4300	NV061	4
NV016	Nguyễn	Sơn	14/08/73	Hà nam	Nam	3800	NV002	5

Các phép toán đại số quan hệ

■ Chiếu

- Phép chiếu là phép toán chọn một số cột của bảng

Các phép toán đại số quan hệ

- Chiếu: $\pi_{\langle \text{Mã số NV} \text{ Họ tên} \text{ Địa chỉ} \text{ Lương} \rangle}(\text{NHÂNVIÊN})$

Mã số NV	Họ tên	Tên	Địa chỉ	Lương
NV001	Lê	Vân	Hà nội	3000
NV002	Trần Đức	Nam	Hà nội	4000
NV010	Hoàng	Thanh	Nghệ an	2500
NV014	Phạm	Bằng	Bắc ninh	4300
NV016	Nguyễn	Sơn	Hà nam	3800
NV018	Vũ Hương	Giang	Nam định	2500
NV025	Trần Lê	Hoa	Phú thọ	2500
NV061	Hoàng	Giáp	Hà tĩnh	5500

Các phép toán đại số quan hệ

■ Hợp

- Hợp của hai quan hệ R và S , được ký hiệu là $R \cup S$, cho kết quả là một quan hệ chứa tất cả các bộ có trong R hoặc ở trong S hoặc ở trong cả hai. Các bộ trùng lặp bị loại bỏ

■ Giao

- Giao của hai quan hệ R và S , được ký hiệu là $R \cap S$, cho kết quả là một quan hệ chứa tất cả các bộ có trong cả hai quan hệ R và S .

■ Trừ

Các phép toán đại số quan hệ

■ Trừ

- Phép trừ quan hệ R và S , được ký hiệu là $R - S$, cho kết quả là một quan hệ chứa tất cả các bộ có trong R nhưng không có trong S . Phép trừ quan hệ R và S , được ký hiệu là $R - S$, cho kết quả là một quan hệ chứa tất cả các bộ có trong R nhưng không có trong S .

Các phép toán đại số quan hệ

R	Họ tên	Tuổi	Giới tính
	AA	20	Nam
	BB	18	Nữ
	CC	21	Nam
	DD	25	Nữ

S	Họ tên	Tuổi	Giới tính
	BB	18	Nữ
	EE	20	Nam
	DD	25	Nữ
	FF	21	Nam

Các phép toán đại số quan hệ

$R \cup S$	Họ tên	Tuổi	Giới tính
	AA	20	Nam
	BB	18	Nữ
	CC	21	Nam
	DD	25	Nữ
	EE	20	Nam
	FF	21	Nam

$R \cap S$	Họ tên	Tuổi	Giới tính
	BB	18	Nữ
	DD	25	Nữ

$R - S$	Họ tên	Tuổi	Giới tính
	AA	20	Nam
	CC	21	Nam

Các phép toán đại số quan hệ

■ Nối

- Phép nối được ký hiệu là \bowtie và được dùng để kết hợp các bộ có liên hệ với nhau từ hai quan hệ thành một bộ. Dạng tổng quát của phép nối trên hai quan hệ $R(A_1, A_2, \dots, A_n)$ và $S(B_1, B_2, \dots, B_m)$ là $R \bowtie S$ < Điều kiện nối >
- Kết quả của phép nối là một quan hệ $Q(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$ có $n+m$ thuộc tính

Các phép toán đại số quan hệ

R	A1	A2	A3
	Aa	Ca	Ba
	Ab	Cb	Bb
	Ac	Ca	Ba
	Ad	Cc	Null
	Ae	Cd	Bb

S	B1	B2	B3
	Ba	Aaa	Bbb
	Bb	Ccc	Ddd

Các phép toán đại số quan hệ

- Khi đó kết quả của phép nối tên-ta R và S với điều kiện $A3 = B1$ sẽ cho kết quả là:

$R \bowtie S$	A1	A2	A3	B1	B2	B3
$\langle A3 = B1 \rangle$	Aa	Ca	Ba	Ba	Aaa	Bbb
	Ab	Cb	Bb	Bb	Ccc	Ddd
	Ac	Ca	Ba	Ba	Aaa	Bbb
	Ae	Cd	Bb	Bb	Ccc	Ddd

Dạng chuẩn

■ Chuẩn 1 – 1NF

- Không có thuộc tính đa trị và thuộc tính suy dẫn

■ Chuẩn 2 - 2NF

- Không có thuộc tính không khóa phụ thuộc một phần vào khóa

■ Chuẩn 3 – 3NF

- Không có thuộc tính không khóa phụ thuộc bắc cầu vào khóa

■ Chuẩn BoyCode - BCNF

- Tất cả các thuộc tính không khóa phụ thuộc trực tiếp vào khóa

Các bước thiết kế CSDL

- Liệt kê, chính xác hóa và lựa chọn thông tin
 - Phát biểu bài toán
 - Xác định các tài liệu nguồn (nếu có)
- Xác định các kiểu đối tượng và thuộc tính của nó
 - *Phân loại thuộc tính (khóa – đa trị - suy dẫn)*
- Xác định mối quan hệ - liên kết và thuộc tính của quan hệ (nếu có) giữa các đối tượng
- Áp dụng nguyên tắc chuyển đổi mô hình ER về mô hình CSDL quan hệ
- Xác định các ràng buộc dữ liệu (nếu có)
- Chuẩn hóa và rút gọn sơ đồ (nếu có)

Ví dụ “Bài toán Quản lý đề án công ty”

- Công ty được tổ chức thành các phòng ban . Mỗi phòng ban có một tên duy nhất, một mã số duy nhất, một nhân viên cụ thể quản lý đơn vị. Việc nhân viên quản lý đơn vị được ghi lại bằng ngày nhân viên đó bắt đầu quản lý. Một đơn vị có thể có nhiều địa điểm. Nhân viên quản lý đơn vị là một người thuộc công ty.
- Mỗi phòng ban thực hiện một số dự án. Một dự án có một tên duy nhất, một mã số duy nhất, một địa điểm thực hiện và thời gian bắt đầu thực hiện dự án.
- Với mỗi nhân viên trong công ty, ta lưu giữ lại thông tin như: Họ tên, Mã số, địa chỉ, hệ số lương, lương (được tính dựa vào hệ số lương và mức lương cơ bản hiện tại), giới tính, ngày sinh, ngoại ngữ (mỗi nhân viên có thể biết nhiều ngoại ngữ). Một nhân viên chỉ làm việc cho một phòng ban nhưng có thể làm việc trên nhiều dự án do nhiều phòng ban thực hiện. Trong một phòng ban chỉ cho phép tối đa 10 nhân viên dưới quyền quản lý của phòng ban đó. Chúng ta cần lưu giữ lại số giờ làm việc của mỗi nhân viên trên một dự án nào đó.
- Mỗi nhân viên có những người thân kèm theo. Những người này được hưởng bảo hiểm theo nhân viên. Với mỗi người thân của nhân viên, chúng ta lưu giữ Họ tên, giới tính, ngày sinh, tuổi, quan hệ với nhân viên.

Ví dụ áp dụng – Tạo CSDL “QLyNS”

Khảo sát việc quản lý nhân viên của một công ty thu được:

Công ty được chia thành các phòng ban khác nhau, mỗi phòng ban được xác định: mã phòng ban, tên phòng ban, địa chỉ và số điện thoại. Mỗi phòng ban được xác định bởi mã PB và có nhiều địa điểm đặt ở các phòng khác nhau

Các nhân viên của công ty được chia về các phòng ban quản lý, mỗi nhân viên có các thông tin: mã NV, họ tên, giới tính, ngày sinh, tuổi, quê quán, điện thoại. Nhân viên được xác định thông qua mã nhân viên.

Mỗi nhân viên được giữ một chức vụ khác nhau, mỗi chức vụ được xác định thêm hệ số phụ cấp tương ứng

Mỗi nhân viên cũng có nhiều trình độ học vấn khác nhau, trình độ học vấn được xác định: mã trình độ, tên trình độ và chuyên ngành

Công ty cũng theo dõi quá trình thay đổi bậc lương của nhân viên theo từng thời điểm khác nhau. Mỗi bậc lương được xác định: mã bậc lương, hệ số lương, hệ số phụ cấp, tổng tiền