

ĐẠI HỌC QUỐC GIA HÀ NỘI  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ

LÊ TRUNG HÙNG

# CƠ SỞ DỮ LIỆU THỜI GIAN

LUẬN VĂN THẠC SĨ

Hà Nội - 2006

**ĐẠI HỌC QUỐC GIA HÀ NỘI  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ**

*Lê Trung Hưng*

***CƠ SỞ DỮ LIỆU THỜI GIAN***

Ngành: Công nghệ thông tin  
Mã số: 1.01.10

**LUẬN VĂN THẠC SĨ**

**NGƯỜI HƯỚNG DẪN KHOA HỌC:  
PGS. TS. Hồ Thuần**

Hà Nội – 2006

## LỜI CẢM ƠN

Trước hết, tôi xin được trân trọng bày tỏ lòng biết ơn sâu sắc đến thầy giáo PGS.TS. *Hồ Thuận*, thầy đã cho tôi những định hướng và ý kiến quý báu, hướng dẫn và giúp đỡ tận tình để tôi có thể hoàn thành được luận văn này.

Đồng thời, tôi cũng xin gửi lời cảm ơn chân thành tới quý thầy cô giáo Khoa Công nghệ thông tin - Trường Đại học Công nghệ, những người đã dạy dỗ, trang bị cho tôi các kiến thức khoa học vô cùng bổ ích và cần thiết trong quá trình học tập và nghiên cứu tại trường.

Xin cảm ơn các đồng nghiệp công tác tại Cục Tin học & Thống kê Tài chính - Bộ Tài chính đã tạo điều kiện và giúp đỡ tận tình tôi trong suốt thời gian qua. Xin cảm ơn gia đình và bạn bè, những người luôn gắn bó, động viên khuyến khích tôi học tập, nghiên cứu và vươn lên.

*Xin chân thành cảm ơn!*

*Hà Nội, tháng 3 năm 2006.*

*Lê Trung Hưng*

---

# MỤC LỤC

DANH MỤC CÁC BẢNG VÀ HÌNH VẼ .....	3
<b>MỞ ĐẦU</b> .....	7
<b>CHƯƠNG 1: TỔNG QUAN VỀ CSDL VÀ CSDL THỜI GIAN</b> .....	9
1.1- GIỚI THIỆU .....	9
1.2 NHẮC LẠI MỘT SỐ KHÁI NIỆM TRONG CSDL QUAN HỆ:.....	10
1.3- CƠ SỞ DỮ LIỆU CÓ YẾU TỐ THỜI GIAN . <b>Error! Bookmark not defined.</b>	
1.3.1- Dữ liệu thời gian và một số khái niệm cơ bản: <b>Error! Bookmark not defined.</b>	
1.3.2- Khái niệm về Thời gian hợp thức và Thời gian giao dịch: <b>Error! Bookmark not defined.</b>	
1.3.3- Các vấn đề về thời gian hoá: .....	<b>Error! Bookmark not defined.</b>
1.3.3.1 - Bán thời gian hoá: .....	<b>Error! Bookmark not defined.</b>
1.3.3.1.1 - Các ràng buộc: .....	<b>Error! Bookmark not defined.</b>
1.3.3.1.2 - Truy vấn .....	<b>Error! Bookmark not defined.</b>
1.3.3.2 -Thời gian hóa toàn phần: .....	<b>Error! Bookmark not defined.</b>
1.3.3.2.1 – Ràng buộc .....	<b>Error! Bookmark not defined.</b>
1.3.3.2.2 – Truy vấn: .....	<b>Error! Bookmark not defined.</b>
1.4- BIỂU DIỄN YẾU TỐ THỜI GIAN TRONG CSDL THỜI GIAN. <b>Error! Bookmark not</b>	
1.4.1- Khoảng thời gian.....	<b>Error! Bookmark not defined.</b>
1.4.2- Các kiểu khoảng:.....	<b>Error! Bookmark not defined.</b>
1.4.2- Toán tử vô hướng trên khoảng:.....	<b>Error! Bookmark not defined.</b>
1.4.3- Các toán tử gộp nhập trên khoảng .....	<b>Error! Bookmark not defined.</b>
1.4.5- Ràng buộc trên các khoảng .....	<b>Error! Bookmark not defined.</b>
1.4.6- Toán tử Update trên khoảng: .....	<b>Error! Bookmark not defined.</b>
<b>CHƯƠNG 2: THIẾT KẾ CƠ SỞ DỮ LIỆU THỜI GIAN</b> <b>Error! Bookmark not defined.</b>	
2.1- QUÁ TRÌNH THIẾT KẾ .....	<b>Error! Bookmark not defined.</b>
2.1.1- Di thừa dữ liệu và các dị thường cập nhật ....	<b>Error! Bookmark not defined.</b>
2.1.2- Chiều thời gian của thiết kế logic .....	<b>Error! Bookmark not defined.</b>
2.2. KIỂU VÀ MODUL THỜI GIAN: .....	<b>Error! Bookmark not defined.</b>

---

2.2.1- Kiểu thời gian:.....	<b>Error! Bookmark not defined.</b>
2.2.2 - Module thời gian:.....	<b>Error! Bookmark not defined.</b>
2.3. PHỤ THUỘC HÀM THỜI GIAN: .....	<b>Error! Bookmark not defined.</b>
2.3.1 - Các tiên đề suy diễn cho TFDs: .....	<b>Error! Bookmark not defined.</b>
2.3.2 - Bao đóng của một tập các thuộc tính: .....	<b>Error! Bookmark not defined.</b>
2.4. CHUẨN HOÁ THỜI GIAN: .....	<b>Error! Bookmark not defined.</b>
2.5. DẠNG CHUẨN BOYCE-CODD CÓ YẾU TỐ THỜI GIAN.	<b>Error! Bookmark not defined.</b>
2.5.1 Tách các lược đồ module thời gian về dạng TBCNF:	<b>Error! Bookmark not defined.</b>
2.6. BẢO TOÀN PHỤ THUỘC: .....	<b>Error! Bookmark not defined.</b>
2.7. DẠNG CHUẨN BA CÓ YẾU TỐ THỜI GIAN	<b>Error! Bookmark not defined.</b>
2.7.1 Tách lược đồ module thời gian về dạng T3NF:	<b>Error! Bookmark not defined.</b>
<b>CHƯƠNG 3: NGÔN NGỮ TRUY VẤN TSQL2</b>	<b>Error! Bookmark not defined.</b>
3.1- GIỚI THIỆU .....	<b>Error! Bookmark not defined.</b>
3.2- MỘT SỐ TÍNH NĂNG CẦN THIẾT CỦA NGÔN NGỮ TSQL2:	<b>Error! Bookmark not defined.</b>
3.3- MỘT SỐ ĐẶC TẢ VỀ SỰ MỞ RỘNG CỦA TSQL2:	<b>Error! Bookmark not defined.</b>
3.4-VÍ DỤ MINH HOẠ .....	<b>Error! Bookmark not defined.</b>
3.4.1 - Lược đồ cho CSDL thử nghiệm.....	<b>Error! Bookmark not defined.</b>
3.4.2 - Lược đồ trong TSQL2.....	<b>Error! Bookmark not defined.</b>
3.4.3 - Cập nhật dữ liệu thử nghiệm.....	<b>Error! Bookmark not defined.</b>
3.4.4 - Một số truy vấn: .....	<b>Error! Bookmark not defined.</b>
<b>MỘT SỐ HƯỚNG NGHIÊN CỨU TIẾP THEO</b>	<b>Error! Bookmark not defined.</b>
<b>KẾT LUẬN</b> .....	<b>Error! Bookmark not defined.</b>
<b>TÀI LIỆU THAM KHẢO</b> .....	11

---

## DANH MỤC CÁC BẢNG VÀ HÌNH VẼ

CHƯƠNG 1: TỔNG QUAN VỀ CSDL VÀ CSDL THỜI GIAN.....	9
Bảng 1.1 CSDL Nhà cung cấp & Mặt hàng (Phiên bản dạng ảnh hiện tại) .....	<b>Error! Bookmark not defined.</b>
Bảng 1.2 CSDL Nhà cung cấp & Mặt hàng (Phiên bản bán thời gian hoá với các giá trị đơn giản)	<b>Error! Bookmark not defined.</b>
Bảng 1.3 CSDL Nhà cung cấp & Mặt hàng (Phiên bản thời gian hoá toàn phần thứ nhất, Sử dụng các nhãn thời gian)	<b>Error! Bookmark not defined.</b>
Bảng 1.4 CSDL Nhà cung cấp & Mặt hàng (Phiên bản thời gian hoá toàn phần sau cùng , Sử dụng các khoảng)	<b>Error! Bookmark not defined.</b>
CHƯƠNG 2: THIẾT KẾ CƠ SỞ DỮ LIỆU THỜI GIAN	<b>Error! Bookmark not defined.</b>
Hình 2.1 Quá trình thiết kế một CSDL .....	<b>Error! Bookmark not defined.</b>
Hình 2.2: Ví dụ về quan hệ các khóa học.....	<b>Error! Bookmark not defined.</b>
Hình 2.3. Phép tách tự nhiên.....	<b>Error! Bookmark not defined.</b>
Hình 2.4: Thuật toán tìm $\overline{X}^+$ .....	<b>Error! Bookmark not defined.</b>
Hình 2.5: Thuật toán tách về dạng chuẩn TBCNF...	<b>Error! Bookmark not defined.</b>
Hình 2.6. Hàm Raise .....	<b>Error! Bookmark not defined.</b>
Hình 2.7 Thuật toán tách về dạng chuẩn T3NF.....	<b>Error! Bookmark not defined.</b>
CHƯƠNG 3: NGÔN NGỮ TRUY VẤN TSQL2	<b>Error! Bookmark not defined.</b>
Bảng 3.1 - Nhanvien_Tbl .....	<b>Error! Bookmark not defined.</b>
Bảng 3.2 - Kynang_Tbl .....	<b>Error! Bookmark not defined.</b>
Bảng 3.3 - Phong_ban_Tbl.....	<b>Error! Bookmark not defined.</b>

## MỞ ĐẦU

Hầu hết các ứng dụng kỹ thuật về cơ sở dữ liệu (CSDL) trong tự nhiên đều mang yếu tố thời gian. Ví dụ: các ứng dụng về tài chính như: quản lý đầu tư, kế toán, ngân hàng; các ứng dụng lưu trữ hồ sơ như: quản lý nhân sự, hồ sơ bệnh án, quản lý kiểm kê, kiểm toán; các ứng dụng về lập lịch trong ngành đường sắt, hàng không, đặt chỗ khách sạn; các ứng dụng khoa học như: dự báo thời tiết, thảm họa tự nhiên...v.v. Tất cả những ứng dụng đó đều dựa trên những CSDL ghi lại dữ liệu tham chiếu tới thời gian và được gọi là “CSDL có yếu tố thời gian” hay “CSDL thời gian” (*Temporal Database*).

Với những ứng dụng rộng rãi trong mọi lĩnh vực đời sống, quản trị CSDL thời gian được xem là một lĩnh vực nghiên cứu đầy sức lôi cuốn, đã có hàng trăm nhà nghiên cứu với trên 2000 bài viết về vấn đề này trong suốt 2 thập kỷ qua. Việc xây dựng và ứng dụng một CSDL có yếu tố thời gian thành công mang một ý nghĩa rất quan trọng trong thực tiễn. Một CSDL thời gian phải phản ánh hay biểu diễn được một cách chân thực nhất, chính xác nhất về dữ liệu tại mọi thời điểm. Một hệ quản trị CSDL có yếu tố thời gian sẽ quản lý những dữ liệu tham chiếu tới thời gian (*Time-referenced data*), do đó các yếu tố thời gian gắn liền với các thực thể trong CSDL.

Nói một cách không chặt chẽ, cơ sở dữ liệu (CSDL) thời gian là một CSDL bao gồm dữ liệu mang tính lịch sử thay vì dữ liệu hiện tại. Một vài trong số những nghiên cứu này đã thừa nhận một quan điểm cực đoan là dữ liệu trong một CSDL đó chỉ được chèn (*Insert*), không bao giờ xoá (*Delete*) hoặc được cập nhật (*Update*), trong trường hợp đó CSDL chỉ bao gồm dữ liệu mang tính lịch sử. Ở một thái cực khác là một CSDL dạng ảnh, chỉ bao gồm dữ liệu hiện tại và dữ liệu đã được xoá hoặc cập nhật khi các sự kiện biểu diễn bởi dữ liệu không còn đúng (hay nói cách khác, một CSDL dạng ảnh chỉ là một CSDL như cách hiểu thông thường, chứ không phải là CSDL biểu thị thời gian).

Trong quá trình lưu trữ và thao tác trên CSDL thì giá trị của một số thuộc tính ổn định theo thời gian, nhưng với một số thuộc tính khác thì giá trị của chúng lại thay đổi theo thời gian. Luận văn này nghiên cứu, hệ thống lại khái niệm về mô hình dữ liệu quan hệ có yếu tố thời gian kéo theo các phép toán đại số, mở rộng các phụ thuộc hàm và các dạng chuẩn có yếu tố thời gian.

Trong thực tế, chúng ta sẽ gặp phải rất nhiều khó khăn khi quản trị CSDL thời gian bằng những mô hình dữ liệu và các ngôn ngữ truy vấn thông thường (không mang yếu tố thời gian). Từ những khái niệm cơ sở về CSDL thời gian, luận văn tập trung vào nghiên cứu và trình bày cách giải quyết một số vấn đề liên quan như:

- Biểu diễn yếu tố thời gian trong CSDL.
- Xây dựng những mô hình CSDL thời gian, thêm yếu tố thời gian vào ngôn ngữ truy vấn và chuẩn hoá các lệnh truy vấn.
- Thiết kế CSDL thời gian ở mức Logic và Khái niệm.
- Xử lý các câu lệnh truy vấn, các toán tử đại số, các chỉ mục (*index*) trên CSDL thời gian.

Bản luận văn được chia thành bốn phần. Phần thứ nhất nhắc lại một số khái niệm cơ bản về CSDL quan hệ, tiếp theo đó tập trung vào giới thiệu những nét tổng quan về CSDL thời gian, những thuật ngữ cũng như những cơ sở để tiếp cận với CSDL thời gian. Trong phần thứ hai, luận văn đi sâu vào việc nghiên cứu cách xây dựng, thiết kế những CSDL có yếu tố thời gian, những vấn đề liên quan đến phụ thuộc hàm thời gian, các dạng chuẩn thời gian và những thuật toán tách lược đồ module thời gian về các dạng chuẩn. Để có những minh hoạ rõ nét và cụ thể hơn về CSDL thời gian, phần ba sẽ giới thiệu ngắn gọn về ngôn ngữ truy vấn thời gian (TSQL2) thông qua một ví dụ cụ thể. Phần cuối của luận văn là phần nhận xét, kết luận và một số hướng nghiên cứu tiếp theo.



## **CHƯƠNG 1: TỔNG QUAN VỀ CSDL VÀ CSDL THỜI GIAN**

### **1.1- GIỚI THIỆU**

Một CSDL là một tập hợp các dữ liệu có liên quan với nhau chứa thông tin về một tổ chức nào đó (như một trường đại học, một ngân hàng, một công ty, hay một nhà máy...), được lưu trữ trên các thiết bị nhớ thứ cấp (như đĩa từ, đĩa quang học...) để đáp ứng nhu cầu khai thác thông tin của nhiều người sử dụng với nhiều mục đích khác nhau. Do đó, trước hết CSDL phản ánh được thông tin về hoạt động của một tổ chức nhất định, nghĩa là biểu thị một “góc” của thế giới thực (tạm gọi là một thế giới nhỏ). Như vậy CSDL cũng phải phản ánh được một cách trung thực sự thay đổi của thế giới nhỏ đó, nó phải tập hợp thông tin mang tính hệ thống chứ không thể là một tập hợp dữ liệu tùy tiện chứa những thông tin rời rạc không có mối quan hệ với nhau. Thông tin lưu trữ trong CSDL được chia sẻ cho nhiều người sử dụng hay cho nhiều ứng dụng khác nhau. Từ đó cho thấy việc thiết kế, xây dựng một CSDL phản ánh thế giới nhỏ đó một cách khách quan, chân thực và chính xác đã trở thành một nhu cầu tất yếu. Từ những năm 80 của thế kỷ hai mươi, các hệ CSDL dựa trên mô hình quan hệ được sử dụng một cách rộng rãi bởi tính đơn giản và các cơ sở toán học của nó. Tuy nhiên cũng giống như những các hệ CSDL thông thường, CSDL quan hệ cũng chỉ biểu thị được những dữ liệu ở dạng tĩnh (dạng ảnh) với mỗi sự kiện của thực thể tương ứng.

Trong một CSDL dạng ảnh, yếu tố thời gian của bức ảnh thông thường được xem là thời điểm hiện tại (nghĩa là thời gian khi mà CSDL thực sự được xem xét. Ngay cả khi yếu tố thời gian của CSDL dạng ảnh đôi khi diễn ra khác với “hiện tại”, điều đó không tạo ra sự khác biệt nào đáng kể đến cách quản lý và sử dụng dữ liệu. Tuy nhiên, như chúng ta sẽ thấy dữ liệu trong một CSDL thời gian được quản lý và sử dụng rất khác trong hàng loạt phương diện quan trọng so với cách nó được quản lý và sử dụng trong một CSDL dạng ảnh; điều này được xem là phần quan trọng nhất và sẽ được trình bày cụ thể sau phần nhắc lại một số những khái niệm cơ bản về CSDL quan hệ.

## 1.2 NHẮC LẠI MỘT SỐ KHÁI NIỆM TRONG CSDL QUAN HỆ:

Mô hình quan hệ biểu thị dữ liệu trong một CSDL như là một bảng giá trị gồm các hàng và các cột. Mỗi hàng trong bảng là một tập các giá trị có liên quan đến nhau, các giá trị này biểu thị một sự kiện tương ứng với một thực thể hay một mối quan hệ trong thế giới thực [1].

- **Miền** (Domain): Một miền  $D$  là một tập hợp các giá trị nguyên tố, hiểu theo nghĩa mỗi giá trị trong miền không thể phân chia trong phạm vi mô hình quan hệ.
- **Quan hệ** (Relation): Một quan hệ trên một tập các thuộc tính là một tập con của tích Đề các của một hay nhiều miền:  
Cho một quan hệ  $r$  xác định trên tập thuộc tính  $\Omega = \{A_1, A_2, \dots, A_n\}$ . Khi đó  $r \subseteq Dom(A_1) \times Dom(A_2) \times \dots \times Dom(A_n)$ , trong đó  $Dom(A_i)$  là miền của thuộc tính  $A_i$ , là tập các giá trị có thể của nó.
- **Lược đồ quan hệ** (Relation schema): Một lược đồ quan hệ  $R$  là một cặp có thứ tự  $R = \langle \Omega, F \rangle$ , trong đó  $\Omega$  là tập hữu hạn các thuộc tính của quan hệ,  $F$  là tập các điều kiện giữa các thuộc tính ( $F$  còn gọi là tập các ràng buộc toàn vẹn).
- **Siêu khoá** của một lược đồ quan hệ  $R$  là một tập hợp gồm một hay nhiều thuộc tính của lược đồ  $R$  có tính chất xác định duy nhất một bộ trong mỗi thể hiện của  $R$ .
- **Khoá** của một lược đồ quan hệ là một siêu khoá của lược đồ này sao cho mọi tập con thực sự của nó không là một siêu khoá.
- **Khoá ngoài** của một lược đồ quan hệ là tập hợp gồm một hay nhiều thuộc tính là khoá của một lược đồ quan hệ khác.
- **Phụ thuộc hàm**: Cho  $U$  là một tập thuộc tính và  $R(U)$  là một lược đồ quan hệ trên  $U$ . Giả sử  $X, Y \subseteq U$ . Khi đó,  $Y$  được gọi là phụ thuộc hàm vào  $X$  trên lược đồ  $R(U)$ , ký hiệu là  $X \rightarrow Y$ , nếu với mọi quan hệ  $r$  trên lược đồ  $R(U)$ , với hai bộ bất kỳ  $t_1, t_2 \in r$ , mà  $t_1[X] = t_2[X]$  thì  $t_1[Y] = t_2[Y]$ .

## TÀI LIỆU THAM KHẢO

### A. TÀI LIỆU THAM KHẢO TIẾNG VIỆT

[1] Hồ Thuần & Hồ Cẩm Hà, (Tập I và II, 2004,2005) "*Các hệ Cơ sở dữ liệu- Lý thuyết và thực hành* ", Nhà xuất bản Giáo dục, Hà Nội.

### B. TÀI LIỆU THAM KHẢO TIẾNG ANH

[2] *Advanced Database Technology and Design* / Mario G.Piattini, Oscar Díaz, (Artech House computing library – Aug-2000).

[3] *Temporal Database Management* / dr.techn. thesis by Christian S. Jensen, defended April 2000.

### C. MỘT SỐ BÀI BÁO VÀ TÀI LIỆU THAM KHẢO KHÁC:

[4] *A Temporal Data Model for Multimedia Database Systems* /Ilan Eini, Vera Goebel and Bjørn Skjellaug. (June-1997)

<http://citeseer.ist.psu.edu/eini97temporal.html>

[5] *CASE Tool Support for Temporal Database Design* / Virginie Detienne, Jean-Luc Hainaut:

Institut d'Informatique, University of Namur rue Grandgagnage,

21 - B-5000 Namur - Belgium

tel: +32 81 724985 - fax: +32 81 724967

\* **Conceptual Modeling-Er 2001** ngày xuất bản 1-12-2001

[6] ***Logical Design for Temporal Databases with Multiple Granularities /***

**X. SEAN WANG**

George Mason University, Fairfax, VA

**CLAUDIO BETTINI**

University of Milan, Milan, Italy

**ALEXANDER BRODSKY and SUSHIL JAJODIA**

George Mason University, Fairfax, VA

(ACM Transactions on Database Systems, Vol. 22, No. 2, June 1997, Pages 115–170)

[7] ***Representing and reasoning about temporal granularities (2004)***

Carlo Combi (corresponding author)

Department of Computer Science, University of Verona,

Ca' Vignal 2 strada le Grazie 15

I - 37134 Verona - VR - Italy

phone: ++39 045 802 7985; fax: ++39 045 802 7068

E-mail: [combi@sci.univr.it](mailto:combi@sci.univr.it)

Massimo Franceschet

Department of Sciences, University 'G. d'Annunzio' of Pescara, Italy

E-mail: [francesc@sci.unich.it](mailto:francesc@sci.unich.it)

Adriano Peron

Department of Physical Sciences, University 'Federico II' of Napoli, Italy

E-mail: [peron@na.infn.it](mailto:peron@na.infn.it)

(<http://citeseer.ist.psu.edu/combi04representing.html>)

[8] ***Temporal Data Management***

Christian S. Jensen, Senior Member, IEEE, and Richard T. Snodgrass, Senior Member,  
IEEE

(IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 11, NO. 1,  
JANUARY/FEBRUARY 1999)

[9] ***Time in Database Systems***

Jan Chomicki

Department of Comp. Sci. and Eng.

University at Buffalo, U.S.A.

[chomicki@cse.buffalo.edu](mailto:chomicki@cse.buffalo.edu)

David Toman

School of Computer Science

University of Waterloo, Canada

[david@uwaterloo.ca](mailto:david@uwaterloo.ca)

(30th January 2004 - WorldScientific/ws-b9-75x6-50)

**[10] *TSQL2 Language Specification (September, 1994)***

Richard T. Snodgrass (chair)

Ilsoo Ahn, Gad Ariav, Don Batory, James Clifford

Curtis E. Dyreson, Ramez Elmasri, Fabio Grandi

Christian S. Jensen, Wolfgang Käfer, Nick Kline, Krishna Kulkarni

T. Y. Cliff Leung, Nikos Lorentzos, John F. Roddick

Arie Segev, Michael D. Soo, Suryanarayana M. Sripada