

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CNTT & TRUYỀN THÔNG

PHẠM THỊ HƯƠNG

**NGHIÊN CỨU CÁCH CHUYỂN ĐỔI CƠ SỞ DỮ LIỆU
QUAN HỆ VÀ XML, ỨNG DỤNG HỆ THỐNG VÀO QUẢN LÝ THƯ VIỆN**

Chuyên ngành : Khoa học máy tính
Mã số : 60.48.01

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

NGƯỜI HƯỚNG DẪN KHOA HỌC : PGS.TS. ĐOÀN VĂN BAN

Thái Nguyên - 2012

LỜI CAM ĐOAN

Tôi xin cam đoan bản luận văn “*Nghiên cứu cách chuyển đổi cơ sở dữ liệu quan hệ và XML, ứng dụng vào hệ thống quản lý thư viện*” là công trình nghiên cứu riêng của tôi. Các số liệu trong luận văn được sử dụng trung thực. Kết quả nghiên cứu được trình bày trong luận văn này chưa từng được công bố tại bất kỳ công trình nào khác. Tôi xin chân thành cảm ơn các Thầy, Cô trường Đại học Công nghệ thông tin & Truyền thông đã truyền đạt cho tôi kiến thức trong suốt những năm học ở trường. Tôi xin chân thành cảm ơn **PGS.TS Đoàn Văn Ban** đã tận tình hướng dẫn tôi hoàn thành tốt luận văn này.

Thái nguyên, ngày 13 tháng 05 năm 2012

Phạm Thị Hương

MỤC LỤC

Trang phụ bìa	
Lời cam đoan.....	i
Mục lục.....	ii
Danh mục các ký hiệu, các chữ viết tắt.....	v
Danh mục các bảng	vi
Danh mục các hình.....	vii
LỜI MỞ ĐẦU	1
1. Lý do chọn đề tài.....	1
2. Mục tiêu nghiên cứu.....	1
3. Phương pháp nghiên cứu.....	2
4. Tổng quan luận văn	2
Chương 1: Tổng quan về XML và cơ sở dữ liệu quan hệ	3
1.1. Tổng quan về XML.....	3
1.1.1. Giới thiệu về XML	3
1.1.2. Mục tiêu ra đời của tài liệu XML	3
1.1.3. Các lợi ích khi sử dụng XML	3
1.2. Cấu trúc tài liệu XML	4
1.2.1. Định nghĩa tài liệu XML	4
1.2.2. Cấu trúc tài liệu XML	4
1.2.3. Các tài liệu cơ bản của tài liệu XML.....	5
1.2.4. Cấu trúc một tài liệu XML hợp khuôn dạng	13
1.2.5. Định nghĩa kiểu DTD	15
1.2.6. Lược đồ XML.....	17
1.2.7. Bảng định nghĩa kiểu CSS.....	18
1.3. Phân tích tài liệu XML theo DOM.....	20
1.3.1. Mô hình đối tượng tài liệu.....	20
1.3.2. Phân tích tài liệu XML theo mô hình đối tượng dữ liệu DOM.....	23
1.4. Cơ sở dữ liệu quan hệ.....	24

1.4.1. Hệ quản trị cơ sở dữ liệu	24
1.4.2. Mô hình dữ liệu quan hệ.....	25
1.4.3. Kiến trúc cơ sở dữ liệu quan hệ.....	25
1.4.4. Ràng buộc toàn vẹn cơ sở dữ liệu quan hệ.....	26
1.5. Kết luận chương 1	28
Chương 2: Chuyển đổi giữa cơ sở dữ liệu quan hệ và XML	29
2.1. Phương pháp luận.....	29
2.1.1. Cấu trúc, kiểu cơ chế của XML và lược đồ CSDL quan hệ.....	29
2.1.2. Mối quan hệ giữa một CSDL quan hệ và một tài liệu XML.....	33
2.1.3. Tên kiểu phân tử.....	33
2.1.4. Giá trị rỗng và giá trị mặc định	34
2.1.5. So sánh các quan hệ trong CSDL quan hệ và XML.....	35
2.1.6. Thứ tự kiểu phân tử trong lược đồ XML.....	37
2.2. Cơ chế chuyển một CSDL quan hệ thành một tài liệu XML.....	39
2.2.1. Phi chuẩn hóa các lược đồ quan hệ	40
2.2.2. Kết nối các CSDL đã phi chuẩn hóa	41
2.2.3. Ánh xạ một CSDL quan hệ thành một tài liệu XML	42
2.3. Chuyển một tài liệu XML sang một CSDL quan hệ.....	45
2.3.1. Phương pháp ánh xạ một tài liệu XML thành CSDL quan hệ	46
2.3.2. Phương pháp băm nhỏ hay phân tách các tài liệu XML	49
2.4. Thuật toán chuyển đổi giữa cơ sở dữ liệu quan hệ và XML.....	50
2.4.1. Biểu diễn tài liệu XML theo mô hình DOM	51
2.4.2. Biểu diễn tài liệu XML theo dạng bảng của CSDL quan hệ.....	52
2.4.3. Chuyển một CSDL quan hệ sang một tài liệu XML	53
2.5. Cài đặt thuật toán	53
2.5.1. Đọc tài liệu XML dưới dạng DOM.....	54
2.5.2. Chuyển đổi một tài liệu XML thành CSDL quan hệ.....	55
2.5.3. Chuyển một CSDL quan hệ thành một tài liệu XML	56
2.6. Kết luận chương 2	60

Chương 3: Thực nghiệm	61
3.1. Mô tả bài toán và giả thuyết.....	61
3.2. Yêu cầu hệ thống.....	62
3.3. Giao diện và một số chức năng của chương trình.....	62
3.3.1. Màn hình chính của hệ thống	62
3.3.2. Chức năng đăng nhập hệ thống	63
3.3.3. Chức năng nhập dữ liệu.....	63
3.3.4. Chức năng tìm kiếm thông tin.....	64
3.3.5. Chức năng báo cáo thống kê	65
3.4. Một số kết quả cài đặt thuật toán	65
3.4.1. Biểu diễn tài liệu XML theo dạng cây	65
3.4.2. Biểu diễn tài liệu XML dạng bảng.....	66
3.4.3. Chuyển CSDL sang tài liệu XML	66
3.4.4. Chuyển tài liệu XML sang CSDL quan hệ	67
3.5. Kết luận chương 3	67
KẾT LUẬN	68
TÀI LIỆU THAM KHẢO	70

BẢNG KÝ HIỆU CÁC CHỮ VIẾT TẮT

Chữ viết tắt	Nghĩa tiếng anh	Nghĩa tiếng việt
1NF	First Normal Form	Dạng chuẩn 1
2NF	Second Normal Form	Dạng chuẩn 2
3NF	Third Normal Form	Dạng chuẩn 3
BCNF	Boyce Codd Normal Form	Dạng chuẩn Boyce Codd
CSDL	Database	Cơ sở dữ liệu
CSS	Cascading Style Sheets	Bảng kiểu xếp chồng
DHTML	Dynamic HyperText Markup Language	Ngôn ngữ định dạng siêu văn bản động
DOM	Document Object Model	Mô hình đối tượng tài liệu
DTD	Document Type Definition	Định nghĩa kiểu tài liệu
HTML	HyperText Markup Language	Ngôn ngữ đánh dấu siêu văn bản
LAN	Local Area Network	Mạng cục bộ
SGML	Standard Generalized Markup Language	Siêu ngôn ngữ có khả năng sinh ngôn ngữ khác
SQL	Structured Query Language	Ngôn ngữ truy vấn có cấu trúc
URI	Uniform Resource Identifier	Định danh tài nguyên trên Internet
URL	Universal Resource Locator	Địa chỉ định vị tài nguyên toàn cầu
W3C	World Wide Web Consortium	Tập đoàn mạng toàn cầu
WAN	Wide Area Network	Mạng diện rộng
XHTML	Extensible HyperText Markup Language	Ngôn ngữ đánh dấu siêu văn bản mở rộng
XML	eXtensible Markup Language	Ngôn ngữ định dạng mở rộng

DANH MỤC BẢNG BIỂU

Bảng 1.1. Bảng tham chiếu thực thể trong XML.....	14
Bảng 1.2. Các loại nút trong mô hình XML	21
Bảng 1.3. Các phương thức trong mô hình DOM.....	24
Bảng 2.1. Các kiểu phân tử	29
Bảng 2.2. Minh họa định nghĩa kiểu phân tử hợp thành độc giả	30
Bảng 2.3. So sánh các yếu tố trong một tập hợp.....	35
Bảng 2.4. Nội dung so sánh các quan hệ	36
Bảng 2.5. Những kiểu phân tử hợp thành không có thứ tự.....	38
Bảng 2.6. Khi việc bấm nhỏ là có hay không là 1 lựa chọn thích hợp	49

DANH MỤC CÁC HÌNH

Hình 1.1. Cấu trúc tài liệu XML	5
Hình 1.2. Tài liệu XML được biểu diễn theo cấu trúc cây	23
Hình 2.1. Các mức so sánh giữa mô hình quan hệ và XML.....	33
Hình 2.2. Sơ đồ thực thể liên kết.....	40
Hình 2.3. Chuẩn hóa các quan hệ.....	41
Hình 2.4. Phi chuẩn quan hệ	41
Hình 2.5. Những ánh xạ cơ bản từ CSDL quan hệ thành tài liệu XML	42
Hình 2.6. Ánh xạ một CSDL sang một tài liệu XML	43
Hình 2.7. Sơ đồ chuyển CSDL quan hệ thành tài liệu XML	44
Hình 2.8. Chuyển từ bảng CSDL quan hệ thành tài liệu XML.....	45
Hình 2.9. Tài liệu XML được thể hiện theo mô hình DOM	45
Hình 2.10. Sơ đồ chuyển tài liệu XML thành CSDL quan hệ	46
Hình 2.11. Những ánh xạ cơ bản từ tài liệu XML thành CSDL quan hệ	47
Hình 2.12. Ánh xạ một tài liệu XML thành một CSDL quan hệ.....	48
Hình 2.13. Chuyển từ một tài liệu XML thành một CSDL quan hệ.....	48
Hình 2.14. Mô hình XML ở dạng bảng và dạng cây	50
Hình 2.15. Biểu diễn tài liệu XML theo mô hình DOM.....	51
Hình 2.16. Biểu diễn tài liệu XML theo dạng bảng.....	52
Hình 2.17. Chuyển tài liệu XML sang cơ sở dữ liệu quan hệ.....	56
Hình 2.18. Chuyển một CSDL sang tài liệu XML.....	60
Hình 3.1. Giao diện tổng quan của hệ thống.....	62
Hình 3.2. Màn hình đăng nhập vào quản lý hệ thống	63
Hình 3.3. Màn hình đọc giả.....	63
Hình 3.4. Kết quả tra cứu thông tin đọc giả	65
Hình 3.5. Kết quả biểu diễn tài liệu XML theo DOM	65
Hình 3.6. Kết quả biểu diễn tài liệu XML theo dạng bảng	66
Hình 3.7. Kết quả chuyển đổi từ CSDL sang tài liệu XML.....	66
Hình 3.8. Kết quả chuyển đổi từ XML sang CSDL quan hệ	67

MỞ ĐẦU

1. Lý do chọn đề tài

Ngày nay công nghệ thông tin đã được ứng dụng rộng rãi trong mọi lĩnh vực kinh tế, chính trị, xã hội, y tế và giáo dục,.... Đóng vai trò rất quan trọng trong quá trình chuyển dịch nền kinh tế, xã hội, làm thay đổi đáng kể về lối sống, phong cách làm việc và nhận thức trong công việc.

Với sự phát triển vượt bậc trong lĩnh vực công nghệ mạng đã cho ra đời nhiều dịch vụ từ xa, đặc biệt là sự ra đời của dịch vụ “application as a service”. Các giao dịch, tương tác, giao tiếp với nhau đều thông qua môi trường Internet. Điều này đã biến công nghệ thông tin trở thành một công cụ không thể thiếu được trong sự phát triển của mọi lĩnh vực, mọi ngành nghề.

Cùng với sự phát triển đó thì có hàng loạt các ngôn ngữ và các giao thức giao tiếp, thiết kế trang web cũng ra đời để phục vụ cho những mục đích, dịch vụ nhất định. Có nhiều ngôn ngữ đã phát triển phổ biến như HTML, DHTML, Java, PHP, ASP,.... mỗi loại ngôn ngữ cũng đều có ưu và nhược điểm riêng.

XML trở nên phổ biến có rất nhiều nguyên nhân. Điểm quan trọng nhất là XML cho phép dễ dàng xử lý, chuyển tải và trao đổi dữ liệu giữa nhiều ứng dụng và tài liệu, người dùng với các định dạng khác nhau.

Nhận thấy tính thiết thực của vấn đề này và được sự gợi ý của giảng viên hướng dẫn, em đã chọn đề tài “*Nghiên cứu cách chuyển đổi cơ sở dữ liệu quan hệ và XML, ứng dụng vào hệ thống quản lý thư viện*” làm đề tài cho luận văn thạc sĩ của mình.

2. Mục tiêu nghiên cứu

- Tìm hiểu về XML và cách định dạng các tài liệu XML
- Nghiên cứu XML để đơn giản hóa việc chia sẻ dữ liệu giữa các hệ thống khác nhau, đặc biệt là các hệ thống được kết nối với Internet.

- Nghiên cứu các thuật toán chuyển đổi một cơ sở dữ liệu quan hệ thành một cơ sở dữ liệu dạng XML và ngược lại chuyển một cơ sở dữ liệu dạng XML về một cơ sở dữ liệu quan hệ.

3. Phương pháp nghiên cứu

- Nghiên cứu qua các tài liệu như: sách, các bài báo, thông tin trên các website và các tài liệu liên quan.

- Phân tích, tổng hợp lý thuyết và giới thiệu các thuật toán chuyển đổi một cơ sở dữ liệu quan hệ sang dạng tài liệu XML và ngược lại.

- Ứng dụng ngôn ngữ lập trình C# trong môi trường ASP.NET để triển khai xây dựng một chương trình ứng dụng demo về việc “Quản lý thư viện” tại Thư viện Trường Cao đẳng Công nghiệp Nam Định.

4. Tổng quan luận văn

Luận văn được trình bày theo hình thức từ trên xuống. Bắt đầu của mỗi phần đều đưa ra những khái niệm cơ bản và quy định cho phần trình bày tiếp sau nhằm mục đích giúp dễ dàng trong khi đọc, dần dần đi sâu vào để thảo luận rõ hơn những vấn đề liên quan.

Luận văn được trình bày trong 3 chương và phần kết luận

Chương 1: Tổng quan về XML và cơ sở dữ liệu quan hệ

Trình bày tổng quan về cấu trúc một tài liệu XML, định nghĩa kiểu tư liệu DTD, lược đồ, mô hình đối tượng tài liệu DOM, phân tích tài liệu XML theo mô hình DOM và các khái niệm cơ sở về cơ sở dữ liệu (CSDL) quan hệ.

Chương 2: Chuyển đổi giữa cơ sở dữ liệu quan hệ và XML

Giới thiệu phương pháp luận chuyển một tài liệu XML sang một CSDL quan hệ và ngược lại. Đưa ra thuật toán chuyển đổi giữa cơ sở dữ liệu quan hệ và XML.

Chương 3: Thực nghiệm

Trình bày việc cài đặt thuật toán bằng ngôn ngữ lập trình C# trong môi trường ASP.NET, trên cơ sở xây dựng ứng dụng “Quản lý thư viện” tại thư viện trường Cao đẳng Công nghiệp Nam Định

Kết luận: Tóm tắt các nội dung chính, các kết quả đạt được và hướng nghiên cứu tiếp theo của luận văn.