

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

NGÔ VĂN ĐỊNH

PHÉP DỊCH CHUYỂN LƯỢC ĐỘ KHỐI
TRONG MÔ HÌNH DỮ LIỆU DẠNG KHỐI

LUẬN VĂN THẠC SỸ KHOA HỌC MÁY TÍNH

Thái Nguyên - 2015

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

NGÔ VĂN ĐỊNH

**PHÉP DỊCH CHUYỂN LƯỢC ĐỒ KHỐI
TRONG MÔ HÌNH DỮ LIỆU DẠNG KHỐI**

CHUYÊN NGÀNH: KHOA HỌC MÁY TÍNH

MÃ SỐ: 60 48 0101

LUẬN VĂN THẠC SỸ KHOA HỌC MÁY TÍNH

Ng-êi h-íng đến khoa hăc

TS LÊ VĂN PHÙNG

Thái Nguyên - 2015

LỜI CẢM ƠN

Để hoàn thành luận văn này tôi đã nhận được sự giúp đỡ tận tình của thầy hướng dẫn khoa học, của các thầy cô trường Đại học Công nghệ thông tin và truyền thông - Đại học Thái Nguyên. Tôi xin chân thành cảm ơn các thầy cô trường Đại học Công nghệ thông tin và truyền thông - Đại học Thái Nguyên đã tạo điều kiện học tập, nghiên cứu và giúp đỡ tôi rất nhiều trong quá trình làm luận văn. Đặc biệt tôi xin cảm ơn thầy **TS Lê Văn Phùng** đã tận tình hướng dẫn, chỉ bảo tôi trong suốt quá trình học tập, nghiên cứu đề tài và giúp đỡ tôi hoàn thành bản luận văn này.

Thái Nguyên, ngày 15 tháng 5 năm 2015

Học viên

Ngô Văn Định

LỜI CAM ĐOAN

Tôi xin cam đoan đây là kết quả nghiên cứu của tôi dưới sự hướng dẫn khoa học của TS. Lê Văn Phùng.

Các số liệu, kết quả nêu trong luận văn là trung thực và chưa từng được ai công bố trong bất kỳ công trình nào khác.

Học viên

Ngô Văn Định

MỤC LỤC

	Trang
MỤC LỤC.....	i
DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT	iii
DANH MỤC CÁC HÌNH.....	iv
MỞ ĐẦU.....	1
CHƯƠNG 1. MÔ HÌNH DỮ LIỆU DẠNG KHỐI.....	4
1.1. Một số mô hình dữ liệu tiêu biểu	4
1.1.1. Mô hình dữ liệu quan hệ.....	4
1.1.2. Mô hình hướng đối tượng.....	4
1.1.3. Mô hình dữ liệu dạng khối	5
1.2. Khối, lược đồ khối và các đặc trưng cơ bản.....	5
1.2.1. Khái niệm khối và lược đồ khối	5
1.2.2. Các phép tính cơ bản trên khối.....	8
1.2.3. Khái niệm phụ thuộc hàm.....	15
1.2.4. Bao đóng của tập thuộc tính chỉ số.....	16
1.2.5. Khóa của lược đồ khối R đối với tập F trên R.....	19
1.2.6. Các dạng chuẩn, tựa chuẩn và tựa chuẩn hóa trên lược đồ khối ...	22
1.2.7. Khái niệm về phủ và phủ tối thiểu của tập phụ thuộc hàm	31
Kết luận chương 1	33
CHƯƠNG 2. PHÉP DỊCH CHUYỂN LƯỢC ĐỒ KHỐI.....	34
2.1. Phép dịch chuyển lược đồ quan hệ.....	34
2.1.1. Định nghĩa	34
2.1.2. Thuật toán dịch chuyển lược đồ quan hệ.....	35
2.1.3. Bổ đề về siêu khoá trong phép dịch chuyển lược đồ quan hệ	39
2.1.4. Dịch chuyển lược đồ quan hệ về dạng cân bằng	40
2.2. Phép dịch chuyển lược đồ khối	43

2.2.1. Định nghĩa	43
2.2.2. Sự khác biệt giữa phép chuyển dịch lược đồ khối so với phép dịch chuyển lược đồ quan hệ	45
2.2.3 Một số thuật toán dịch chuyển lược đồ khối	46
2.2.4. Biểu diễn bao đóng qua phép dịch chuyển	48
2.2.5. Biểu diễn khóa qua phép dịch chuyển	51
2.2.6. Ví dụ	55
Kết luận chương 2	56
CHƯƠNG 3. CHƯƠNG TRÌNH THỬ NGHIỆM	58
3.1. Bài toán thử nghiệm	58
3.2. Phân tích và thiết kế chương trình thử nghiệm	59
3.2.1. Thủ tục dịch chuyển	59
3.2.2. Biểu diễn khóa qua phép dịch chuyển	60
3.2.3. Thiết kế chương trình	60
3.3. Cài đặt và thực hiện chương trình thử nghiệm.....	60
3.3.1. Yêu cầu hệ thống	60
3.3.2. Hệ thống dữ liệu vào/ra	61
3.3.3. Hệ thống giao diện.....	61
3.3.4. Kết quả thử nghiệm chương trình và đánh giá	62
Kết luận chương 3	67
KẾT LUẬN	68
TÀI LIỆU THAM KHẢO.....	69

DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT

Luận văn này dùng thống nhất các ký hiệu và chữ viết tắt sau:

Ký hiệu	Ý nghĩa
CSDL	cơ sở dữ liệu
$\text{Dom}(A)$	miền giá trị của thuộc tính A
LĐQH	lược đồ quan hệ
r hoặc $r(R)$	khối r trên tập R
R_x	lát cắt của $r(R)$ tại điểm x.
PTH	phụ thuộc hàm
$x^{(i)} = (x, A_i)$	các thuộc tính chỉ số của lược đồ khối ($x \in \text{id}$, $i = 1..n$)
$\text{id}^{(i)} = \{x^{(i)} x \in \text{id}\}$	tập các thuộc tính chỉ số của lược đồ khối.
F_h	tập các phụ thuộc hàm trên R
F_{hx}	tập các phụ thuộc hàm trên R_x

DANH MỤC CÁC HÌNH

	Trang
Hình 1.1 Biểu diễn khối SANPHAM	7
Hình 1.2 Ví dụ về phép hợp	10
Hình 1.3 Ví dụ về phép giao	10
Hình 1.4 Ví dụ về phép trừ	11
Hình 3.1 Giao diện chính của chương trình	61
Hình 3.2 Giao diện nhập dữ liệu	62

MỞ ĐẦU

1. Lí do chọn lựa đề tài

Ngày nay, công nghệ thông tin đã trở thành một nhân tố không thể thiếu được trong mọi lĩnh vực của đời sống xã hội. Sự bùng nổ nhu cầu xây dựng các hệ thống thông tin, mà trước hết là các hệ thống thông tin quản lý đã trở thành hướng nghiên cứu và quan tâm của rất nhiều nhà khoa học cũng như người sử dụng. Các hệ thống cơ sở dữ liệu (CSDL) đã lần lượt xuất hiện. Vào những năm 60, thế hệ đầu tiên của cơ sở dữ liệu ra đời dưới dạng các mô hình thực thể - liên kết (có đặc điểm nhận dạng đối tượng), mạng phân cấp. Tiếp đến là vào những năm 70 thế hệ thứ hai của CSDL ra đời. Đó là mô hình quan hệ do E. F. Codd đề xuất. Loại mô hình này đã đánh dấu mốc quan trọng về cơ sở lý thuyết của các hệ thống CSDL. Sở dĩ mô hình này được đánh giá cao hơn cả vì nó được xây dựng dựa trên một cơ sở toán học chặt chẽ. Tuy nhiên, do các quan hệ có cấu trúc phẳng (tuyến tính) nên mô hình này chưa đủ đáp ứng đối với các ứng dụng phức tạp, các cơ sở dữ liệu có cấu trúc phi tuyến,...

Trong những năm gần đây, việc nghiên cứu nhằm mở rộng mô hình quan hệ được nhiều nhà khoa học quan tâm. Theo hướng nghiên cứu này đã có một số hướng mở rộng mô hình quan hệ được đề xuất nghiên cứu như: mô hình dữ liệu đa chiều; khối dữ liệu; kho dữ liệu; mô hình dữ liệu dạng khối... Trong đó, ở mô hình dữ liệu dạng khối, các khối là khái niệm cơ bản được mở rộng từ các quan hệ trong mô hình quan hệ, các khối này có thể biểu diễn các dữ liệu có tính chất động (biểu diễn các dữ liệu có thuộc tính thay đổi theo thời gian, không gian...) có khả năng đáp ứng tốt đối với nhiều lớp bài toán phức tạp.

Trong mô hình quan hệ, để giảm tính phức tạp của việc xác định bao đóng, khóa trong các cơ sở dữ liệu lớn, phức tạp, người ta đã đề xuất phép dịch chuyển lược đồ quan hệ (LDQH). Trong mô hình cơ sở dữ liệu dạng

khối, việc xác định khóa và bao đóng càng khó khăn hơn, chính vì vậy mà phép dịch chuyển lược đồ khối đã được đề xuất với mục đích tương tự.

Mục tiêu của đề tài là tìm hiểu kỹ thuật thu gọn lược đồ khối dựa trên phép dịch chuyển lược đồ khối và phương pháp biểu diễn bao đóng và khóa của lược đồ khối thông qua phép dịch chuyển với độ phức tạp thấp hơn so với phương pháp tìm bao đóng và khóa thông thường. Đồng thời, nghiên cứu thuật toán và cài đặt chương trình thử nghiệm với thuật toán dịch chuyển lược đồ khối và biểu diễn khóa của lược đồ khối qua phép dịch chuyển.

2. Mục đích nghiên cứu

Mô hình dữ liệu dạng khối là một mở rộng của mô hình quan hệ với các khối cho phép biểu diễn các dữ liệu có tính chất động (biểu diễn các dữ liệu có thuộc tính thay đổi theo thời gian). Tuy nhiên, các nghiên cứu về mô hình dữ liệu này còn chưa nhiều.

Đối với các cơ sở dữ liệu khối lớn và phức tạp, việc xác định khóa và bao đóng là một việc khó. Nhờ việc dịch chuyển lược đồ khối, việc tính bao đóng và khóa có thể trở nên đơn giản hơn.

3. Nhiệm vụ nghiên cứu

- Nghiên cứu tổng quan về mô hình dữ liệu dạng khối; một số khái niệm và thuật toán liên quan.

- Nghiên cứu phép dịch chuyển lược đồ khối trong mô hình dữ liệu dạng khối và một số thuật toán dịch chuyển lược đồ khối.

- Áp dụng thuật toán dịch chuyển trong bài toán thực tế

4. Đối tượng và phạm vi nghiên cứu:

- **Đối tượng nghiên cứu:** Mô hình dữ liệu dạng khối.
- **Phạm vi nghiên cứu:** Phép dịch chuyển lược đồ khối nhằm giảm nhẹ độ phức tạp trong việc tính bao đóng và khóa của khối.

5. Phương pháp nghiên cứu

- Phương pháp nghiên cứu lý luận.