BỘ GIÁO DỤC VÀ ĐÀO TẠO
VIỆN HÀN LÂM KHOA HỌC VÀ CÔNG NGHỆ VIỆT NAM

VIỆN CÔNG NGHỆ THÔNG TIN

BÙI ĐỨC MINH

NGHIÊN CỨU HỆ SINH ÁNH XẠ ĐÓNG VÀ ỨNG DỤNG TRONG THỂ HIỆN NGỮ NGHĨA DỮ LIỆU

LUẬN ÁN TIẾN SĨ TOÁN HỌC

HÀ NỘI – 2014

BỘ GIÁO DỤC VÀ ĐÀO TẠO
VIỆN HÀN LÂM KHOA HỌC
VÀ CÔNG NGHỆ VIỆT NAM

VIỆN CÔNG NGHỆ THÔNG TIN

BÙI ĐỨC MINH

NGHIÊN CỨU HỆ SINH ÁNH XẠ ĐÓNG VÀ ỨNG DỤNG TRONG THỂ HIỆN NGỮ NGHĨA DỮ LIỆU

Chuyên ngành: BẢO ĐẢM TOÁN HỌC CHO MÁY TÍNH VÀ HỆ THỐNG TÍNH TOÁN
Mã số: 62.46.35.01

LUẬN ÁN TIẾN SĨ TOÁN HỌC

NGƯỜI HƯỚNG DẪN KHOA HỌC:

1. PGS.TSKH. NGUYỄN XUÂN HUY
2. TS. HOÀNG QUANG

HÀ NỘI - 2014

LỜI CAM ĐOAN

Tôi cam đoan đây là công trình nghiên cứu của riêng tôi. Các số liệu, kết quả trong luận án là trung thực và chưa từng công bố trong bất kỳ công trình nào khác.

Tác giả luận án

Bùi Đức Minh

LỜI CÁM ƠN

Luận án được thực hiện và hoàn thành tại Viện Công nghệ Thông tin, Viện Khoa học và Công nghệ Việt Nam, dưới sự hướng dẫn khoa học của PGS TSKH Nguyễn Xuân Huy và TS Hoàng Quang. Nhân dịp này, xin cho tôi được gửi đến những người thầy của mình lời cám ơn chân thành về những chỉ dẫn khoa học và những hướng dẫn tận tình trong quá trình thực hiện luận án. Đặc biệt, xin cho tôi được bày tỏ lòng biết ơn sâu sắc nhất đến PGS TSKH Nguyễn Xuân Huy, người Thầy mà tôi đã may mắn được học tập và làm việc trong khoảng thời gian dài, người đã định hướng, động viên và khơi gợi lòng ham mê nghiên cứu khoa học cũng như truyền thụ các kiến thức, kinh nghiệm sâu sắc về chuyên môn cho tôi trong quá trình học tập và thực hiện luận án.

Lời cám ơn chân thành nhất xin gửi đến GS TS Vũ Đức Thi, PGS TS Đoàn Văn Ban, TS Lê Văn Phùng đã có nhiều nhận xét, góp ý quý báu và định hướng cho tác giả trong việc nghiên cứu đề tài đang thực hiện.

Tôi xin trân trọng cám ơn đến lãnh đạo Viện CNTT, PGS TS Thái Quang Vinh, PGS TS Lương Chi Mai, PGS TS Đặng Văn Đức và các Thầy, Cô trong Viện đã tạo điều kiện tốt nhất cho tôi trong quá trình học tập, nghiên cứu và thực hiện luận án tại Viện.

Cuối cùng, xin cho tôi gửi lời cám ơn chân thành đến Ban Giám hiệu, lãnh đạo các phòng ban cùng các đồng nghiệp Khoa CNTT Trường CĐ GTVT Tp. HCM và gia đình đã tạo điều kiện thuận lợi nhất về vật chất cũng như dành nhiều động viên về mặt tinh thần để tôi có thể yên tâm học tập và hoàn thành luận án.

MỤC LỤC

LỜI CAM ĐOAN **1**

LỜI CÁM ƠN **2**

MỤC LỤC **3**

DANH MỤC CÁC HÌNH **6**

DANH MỤC CÁC BẢNG **7**

DANH MỤC TỪ VIẾT TẮT **8**

PHẦN MỞ ĐẦU **9**

CHƯƠNG 1 MỘT SỐ KHÁI NIỆM CƠ BẢN VỀ CƠ SỞ DỮ LIỆU QUAN HỆ VÀ KHAI PHÁ DỮ LIỆU **18**

1.1. Khái niệm về cơ sở dữ liệu quan hệ 19

1.2. Phụ thuộc hàm 19

1.2.1. Khái niệm phụ thuộc hàm 20

1.2.2. Lược đồ quan hệ 21

1.2.3. Bao đóng tập phụ thuộc hàm 21

1.2.4. Định lý tương đương 22

1.2.5. Bao đóng tập thuộc tính 23

1.2.6. Bài toán thành viên 24

1.3. Khóa và phản khóa của lược đồ quan hệ 24

1.3.1. Khóa của lược đồ quan hệ 25

1.3.2. Phản khóa của lược đồ quan hệ 26

1.4. Một số khái niệm trong khai phá dữ liệu 27

1.4.1. Một số khái niệm cơ bản 27

1.4.2. Luật kết hợp và kết nối Galois 29

1.5. Kết luận chương 1 30

CHƯƠNG 2 ÁNH XẠ ĐÓNG&LÝ THUYẾT GIÀN GIAO VÀ ỨNG DỤNG **31**

2.1. Ánh xạ đóng 33

2.1.1. Các khái niệm và tính chất ánh xạ đóng 33

2.1.2. Phép hạn chế trên ánh xạ đóng 35

2.1.3. Điểm bất động(tập đóng) trên ánh xạ đóng 35

2.2. Các phép toán trên ánh xạ đóng 36

2.2.1. Phép toán hội 36
2.2.2. Phép toán hợp thành 36
2.2.3. Ứng dụng phép toán hợp thành 41
2.3. Cơ sở và phản cơ sở ánh xạ đóng 43
2.3.1. Cơ sở ánh xạ đóng 43
2.3.2. Phản cơ sở ánh xạ đóng 44
2.4. Giàn giao 45
2.4.1. Một số khái niệm cơ bản 45
2.4.2. Sự tương quan giữa tập phản cơ sở và tập đối nguyên tử 48
2.5. Ứng dụng giàn giao với bài toán ẩn tập mục nhạy cảm 50
2.5.1. Đặt vấn đề 50
2.5.2. Phát biểu bài toán 51
2.5.3. Cơ sở lý thuyết 53
2.5.4. Thuật toán ẩn tập mục nhạy cảm 56
2.5.5. Kết quả thử nghiệm 60
2.6. Giàn giao và ứng dụng trong khai thác tập phổ biến 61
2.6.1. Cơ sở lý thuyết 62
2.6.2. Thuật toán xác định họ các tập phổ biến tối đại 63
2.7. Kết luận chương 2 65
CHƯƠNG 3 HỆ SINH AXĐ VÀ MỘT SỐ KẾT QUẢ NGHIÊN CỨU 66
3.1. Hệ sinh ánh xạ đóng 68
3.1.1. Khái niệm hệ sinh AXĐ 68
3.1.2. Ánh xạ cảm sinh 69
3.1.3. Thuật toán xác định ảnh một tập con trong hệ sinh 70
3.2. Giản lược tập luật sinh 71
3.2.1. Một số khái niệm cơ sở 71
3.2.2. Tập giản lược tự nhiên 75
3.2.3. Tập giản lược không dư 76
3.3. Thu gọn hệ sinh ánh xạ đóng 78
3.3.1. Các khái niệm và thuật toán thu gọn hệ sinh AXĐ 79
3.3.2. Biểu diễn ảnh tập con theo phép thu gọn hệ sinh AXĐ 80
3.4. Cơ sở và phản cơ sở hệ sinh ánh xạ đóng 81
3.4.1. Cơ sở hệ sinh AXĐ 82
3.4.2. Phản cơ sở hệ sinh AXĐ 83
3.4.3. Một dạng biểu diễn phản cơ sở hệ sinh AXĐ 84
3.4.4. Sự tương quan giữa các đối tượng trong hệ sinh AXĐ 87

3.5. Ứng dụng hệ sinh AXĐ giải bài toán hệ suy dẫn 90
3.5.1. Các khái niệm và quy tắc suy dẫn 90
3.5.2. Một số dạng bài toán suy dẫn 90
3.6. Hệ sinh cân bằng 94
3.6.1. Các khái niệm và một số tính chất 94
3.6.2. Thuật toán thu gọn hệ sinh AXĐ về dạng cân bằng 97
3.7. Ứng dụng hệ sinh AXĐ trong cơ sở dữ liệu 100
3.7.1. Bài toán phân rã và kết nối các quan hệ 100
3.7.2. Một dạng biểu diễn phản khóa của lược đồ quan hệ 103
3.8. Kết luận chương 3 105
PHẦN KẾT LUẬN 106
DANH MỤC CÔNG TRÌNH ĐÃ CÔNG BỐ 109
TÀI LIỆU THAM KHẢO 110

DANH MỤC CÁC HÌNH

Hình 2.1. Đồ thị của giàn các tập mục phổ biến 53
Hình 2.2. Giàn giao đầy đủ của *Poset*(*ABE*).....54
Hình 2.3. Giàn các tập phổ biến sau khi xóa tập mục nhạy cảm 59
Hình 2.4. Giàn các tập phổ biến.....64

DANH MỤC CÁC BẢNG

Bảng 1.1. Bảng *T* với *22* giao tác 29

Bảng 1.2. Các tập mục phổ biến theo ngưỡng $\sigma = 4$ 29

Bảng 2.1. Bảng các tập mục với độ phổ biến và số lần sửa 49

Bảng 2.2. Một số kết quả thử nghiệm 53

Bảng 2.3. Cơ sở dữ liệu giao tác minh họa 61

Bảng 2.4. Các tập con và ảnh tương ứng 64

Bảng 3.1. Danh sách các môn học 91

Bảng 3.2. Quan hệ học trước giữa các môn 92

Bảng 3.3. Tương ứng giữa CSDL và AXĐ 100

DANH MỤC TỪ VIẾT TẮT

AXĐ: Ánh xạ đóng

CSDL: Cơ sở dữ liệu

HSCB: Hệ sinh cân bằng

LĐQH: Lược đồ quan hệ

PTBD: Phụ thuộc Boole dương

PTBDTQ: Phụ thuộc Boole dương tổng quát

PTBDĐT: Phụ thuộc Boole dương đa trị

PTBDTNB: Phụ thuộc Boole dương theo nhóm bộ

PTH: Phụ thuộc hàm