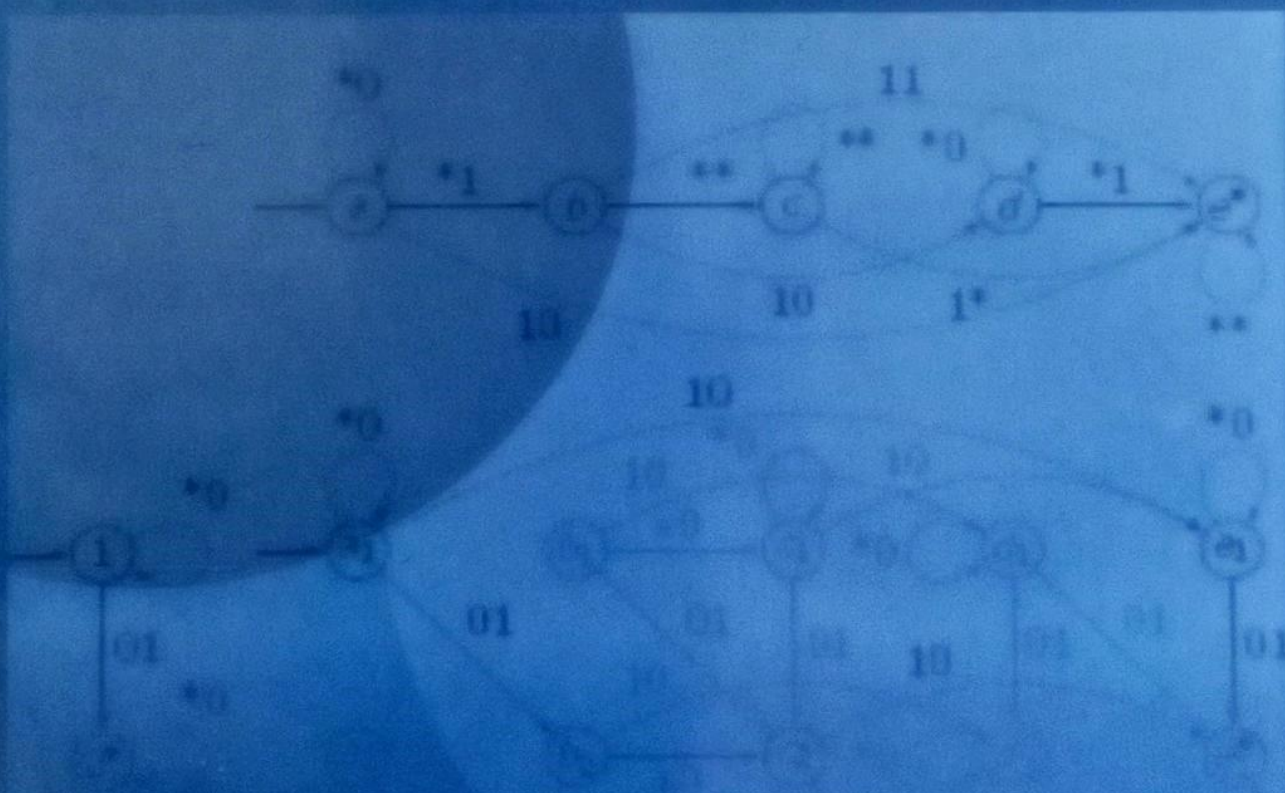




TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP HÀ NỘI

GIÁO TRÌNH

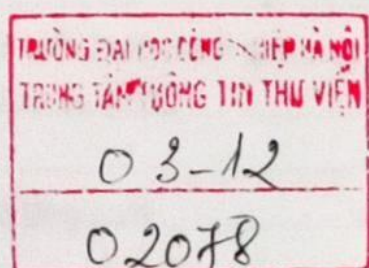
Ngôn ngữ hình thức và Ôtômat



NHÀ XUẤT BẢN GIÁO DỤC VIỆT NAM

TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP HÀ NỘI
TRẦN HÙNG CƯỜNG (chủ biên) - NGUYỄN VĂN THẮNG

GIÁO TRÌNH
**NGÔN NGỮ HÌNH THỨC
VÀ ÔTÔMÁT**



NHÀ XUẤT BẢN GIÁO DỤC VIỆT NAM

MỤC LỤC

Lời nói đầu.....	5
Chương I. Ngôn ngữ hình thức và văn phạm.....	7
1.1. Tập hợp.....	7
1.2. Quan hệ.....	8
1.3. Chứng minh quy nạp.....	10
1.4. Đồ thị và cây.....	10
1.5. Khái niệm chung về ngôn ngữ.....	12
1.6. Văn phạm và sự phân cấp Chomsky.....	13
1.6.1. Văn phạm.....	13
1.6.2. Sự phân cấp Chomsky.....	14
Bài tập.....	15
Chương II. Ôtômát hữu hạn.....	17
2.1. Ôtômát hữu hạn đơn định.....	17
2.2. Ôtômát hữu hạn không đơn định.....	21
2.3. Sự tương đương giữa DFA và NFA.....	23
2.4. NFA với ϵ – dịch chuyển.....	26
Bài tập.....	30
Chương III. Biểu thức chính quy và ngôn ngữ chính quy.....	33
3.1. Biểu thức chính quy.....	33
3.2. Các tính chất đại số của biểu thức chính quy.....	34
3.3. Sự tương đương giữa NFA_{ϵ} và BTCQ.....	35
3.4. Ngôn ngữ chính quy.....	38
3.5. Đồ thị chuyển trạng thái tổng quát.....	39
3.6. Bộ đề bơm.....	41
3.6.1. Nguyên lý lông chim bơm.....	41
3.6.2. Bộ đề bơm.....	42
Bài tập.....	44
Chương IV. Văn phạm chính quy và các tính chất.....	46
4.1. Văn phạm chính quy.....	46
4.1.1 Văn phạm tuyến tính.....	46
4.1.2 Văn phạm chính quy.....	47
4.2. Sự tương đương giữa RG & FA.....	47
4.3. Tính chất đóng của ngôn ngữ chính quy.....	52
Bài tập.....	59

Chương V. Văn phạm phi ngữ cảnh và các tính chất	61
5.1. Xuất xứ và định nghĩa văn phạm phi ngữ cảnh	61
5.2. Thứ tự dẫn xuất và cây dẫn xuất	63
5.2.1. Thứ tự dẫn xuất	63
5.2.2. Cây dẫn xuất	64
5.2.3. Sự nhập nhằng trong văn phạm phi ngữ cảnh	66
5.3. Giảm lược văn phạm phi ngữ cảnh	69
5.3.1. Quy tắc thay thế	69
5.3.2. Loại bỏ các quy tắc rỗng	70
5.3.3. Loại bỏ sản xuất đơn vị	71
5.3.4. Loại ký hiệu vô ích	72
5.4. Dạng chuẩn Chomsky	74
Bài tập	77
Chương VI. Ôtômát đẩy xuống	80
6.1. Các định nghĩa cơ sở	80
6.2. Ngôn ngữ được chấp nhận bởi PDA	86
6.3. Sự tương đương giữa các PDA	88
6.4. Sự tương đương giữa PDA và CFG	89
Bài tập	92
Chương VII. Tính chất của ngôn ngữ phi ngữ cảnh	95
7.1. Tính chất đóng của ngôn ngữ phi ngữ cảnh	95
7.2. Bài toán thành viên	100
7.3. Bổ đề bơm	102
Bài tập	105
Chương VIII. Máy Turing	106
8.1. Mô tả máy Turing chuẩn	106
8.2. Ngôn ngữ và hàm tính được	112
8.3. Luận đề Turing	114
8.4. Giải thuật	114
8.5. Các mô hình khác của máy Turing	115
8.5.1. Máy Turing không đơn định	115
8.5.2. Máy Turing với băng nhiều chiều	116
Bài tập	117
<i>Tài liệu tham khảo</i>	117

Lời nói đầu

Khi con người trao đổi thông tin với nhau, người ta sử dụng ngôn ngữ tự nhiên. Tuy nhiên, ngôn ngữ này đa nghĩa và tùy vào hoàn cảnh khác nhau có thể có cách hiểu khác nhau. Vậy, khi con người trao đổi thông tin với máy tính hoặc giữa các máy tính với nhau liệu có sử dụng được ngôn ngữ này không? Câu trả lời là không thể, mà phải dùng ngôn ngữ có một nghĩa duy nhất. Ngôn ngữ có đặc điểm như vậy được gọi là ngôn ngữ hình thức.

Bất kỳ ngôn ngữ nào kể cả ngôn ngữ hình thức đều hình thành dựa trên một tập hữu hạn các ký hiệu (bảng chữ cái) và tập các quy tắc cho phép hình thành các "câu" trên bảng chữ cái này. Để hình thức hóa các loại ngôn ngữ này người ta sử dụng cơ chế văn phạm. Nói cách khác, văn phạm là cơ chế cho phép sinh ra ngôn ngữ và ngôn ngữ này lại được đoán nhận bởi ô tô-mát (máy trừu tượng) tương ứng. Lý thuyết về ngôn ngữ hình thức và ô tô-mát được hình thức hóa thành các mô hình toán học tương ứng cho các ngôn ngữ lập trình, mô tả các dãy tính toán và khả năng tính toán của máy tính. Ô tô-mát và ngôn ngữ hình thức được áp dụng rộng rãi trong nhiều lĩnh vực khoa học và ứng dụng như: các kỹ thuật dịch, trí tuệ nhân tạo, công nghệ tri thức,...

Đối với người học ngôn ngữ lập trình và để có thể hiểu sâu sắc về ngôn ngữ lập trình, nhất là về mặt cú pháp thì nhất thiết phải có kiến thức về văn phạm, ô tô-mát. Nhằm đáp ứng nhu cầu trên chúng tôi biên soạn giáo trình "**Ngôn ngữ hình thức và Ô tô-mát**". Trong khuôn khổ thời lượng có hạn giáo trình tập trung vào các vấn đề cốt lõi sau: Đó là sự phân lớp văn phạm của Chomsky gồm 4 loại văn phạm cùng với các loại ô tô-mát tương ứng và mối liên hệ giữa chúng. Trong đó đặc biệt là ngôn ngữ phi ngữ cảnh – lớp ngôn ngữ thể hiện các đặc điểm cú pháp của ngôn ngữ lập trình cấp cao.

Nội dung giáo trình bao gồm: Chương 1 trình bày các khái niệm cơ bản về ngôn ngữ và cách xác định loại văn phạm cùng với các loại ô tô-mát tương ứng và mối liên hệ giữa chúng. Trong đó đặc biệt là ngôn ngữ phi ngữ cảnh – lớp ngôn ngữ thể hiện các đặc điểm cú pháp của ngôn ngữ lập trình cấp cao; Chương 2 giới thiệu về ô tô-mát hữu hạn và biểu thức chính quy; Chương 3 trình bày về ngôn ngữ chính quy và các tính chất của nó; Chương 4 giới thiệu về văn phạm chính quy; Chương 5 giới thiệu về văn phạm phi ngữ cảnh; Chương 6 trình bày về ô tô-mát đẩy xuống; Ngôn ngữ phi ngữ cảnh và tính chất được trình bày trong chương 7; Chương 8 đề cập đến máy Turing.

Cuối cùng, chúng tôi xin chân thành cảm ơn khoa Công nghệ Thông tin Trường Đại học Công nghiệp Hà nội đã tạo điều kiện thuận lợi để chúng tôi hoàn thành giáo trình này. Giáo trình xuất bản lần đầu nên khó tránh khỏi thiếu sót, các tác giả mong nhận được những ý kiến đóng góp từ bạn đọc để cuốn sách được hoàn thiện hơn trong lần xuất bản sau.

Mọi ý kiến đóng góp xin gửi về:

Khoa Công nghệ Thông tin - Trường Đại học Công nghiệp Hà Nội

Xin chân thành cảm ơn!

Các tác giả

*Bản quyền tác giả thuộc sở hữu của trường Đại học Công Nghiệp Hà Nội
Công ty CP Sách Giáo dục tại TP. Hà Nội – NXBGD Việt Nam
giữ quyền công bố tác phẩm*

GIÁO TRÌNH NGÔN NGỮ HÌNH THỨC VÀ ÔTÔMÁT

Mã số: 7B014S1-TTS

Số ĐKKH xuất bản: 1048 – 2011/CXB/10-1495/GD

In 1000 cuốn, khổ 17x24cm tại Trung tâm NC&SX Học Liệu

In xong và nộp lưu chiểu tháng 10 năm 2011.