LÝ THUYẾT TÍNH TOÁN

BÀI 2: ÔTÔMAT HỮU HẠN

Phạm Xuân Cường
Khoa Công nghệ thông tin
cuongpx@tlu.edu.vn

Nội dung bài giảng

1. Ôtômat hữu hạn
2. Định nghĩa hình thức
3. Thiết kế Ôtômat hữu hạn
4. Ngôn ngữ chính quy
5. Toán tử chính quy

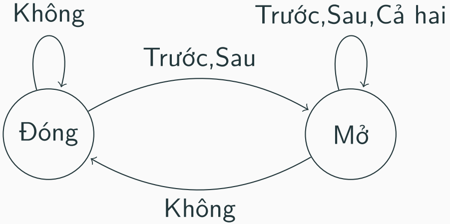
Ôtômat hữu hạn

Ôtômat hữu hạn

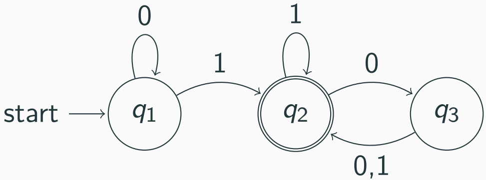
Ôtômat hữu hạn (Finite State Machine - **FSM** hay Finite Automation)

- Là mô hình tính toán đơn giản nhất
- Phù hợp với:
    - Các máy tính hoặc bộ điều khiển nhỏ
    - Có số trạng thái hữu hạn và khá nhỏ

Ví dụ: Bộ điều khiển cửa trượt tự động



Biểu diễn hình học của Ôtômat hữu hạn



- **Trạng thái bắt đầu:** Biểu thị bởi mũi tên chỉ vào nó
- **Trạng thái kết thúc:** Biểu thị bởi vòng tròn kép
- Mũi tên từ trạng thái này sang trạng thái khác được gọi là **chuyển dịch**
- Thông tin đầu ra hoặc là **chấp thuận** hoặc là **bác bỏ**

Ứng dụng của FSM

- Tạo ra các chuỗi tương ứng với mô hình của FSM
- Nhận diện các chuỗi có thỏa mãn mô hình FSM hay không

Ví dụ nhận diện các chuỗi sau:

- 11010101 $\rightarrow$ **Chấp thuận/bác bỏ?**
- 100 $\rightarrow$ **Chấp thuận/bác bỏ?**
- 110000 $\rightarrow$ **Chấp thuận/bác bỏ?**
- 0100 $\rightarrow$ **Chấp thuận/bác bỏ?**
- 101000 $\rightarrow$ **Chấp thuận/bác bỏ?**

$\rightarrow$ **Làm thế nào để biểu diễn các chuỗi chấp thuận bằng 1 ngôn ngữ?**

Định nghĩa hình thức

Định nghĩa hình thức

- Ôtômat hữu hạn $\equiv$ bộ 5 (hay 5 chiều)

$$\mathbf{M = (Q, \Sigma, \delta, q_0, F)}$$

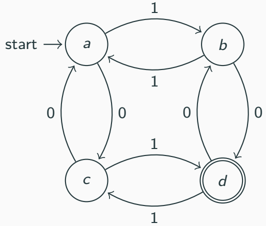
Trong đó:

- **Q:** Tập trạng thái (hữu hạn)
- **Σ:** Bộ chữ, tập hữu hạn các ký tự
- **δ:** Hàm dịch chuyển

$$\boldsymbol{\delta}\mathbf{: Q \ x \ \Sigma} \rightarrow \mathbf{Q}$$

- $\mathbf{q_0}$: Trạng thái bắt đầu ($\mathbf{q_0} \in \mathbf{Q}$)
- **F:** Là tập các trạng thái kết thúc ($\mathbf{F} \subseteq \mathbf{Q}$)

Ví dụ Ôtômat hữu hạn



- **Q:** $\{a,b,c,d\}$
- **Σ:** $\{0,1\}$
- $\mathbf{q_0}$: a
- **F:** $\{d\}$

- **δ:**

		Σ	
		0	1
Trạng thái	a	c	b
	b	d	a
	c	a	d
	d	b	c

Ngôn ngữ của máy M

- Nếu A là tập tất cả các xâu mà máy M chấp nhận $\rightarrow$ A là ngôn ngữ của máy M

$$\mathbf{L(M) = A}$$

- Máy M đoán nhận (recognizes) A
- ~~Máy M chấp thuận (accepts) A~~
  Do một máy có thể chấp thuận vài xâu nhưng nó luôn đoán nhận chỉ một ngôn ngữ
- Nếu máy không chấp thuận một xâu nào thì nó vẫn đoán nhận một ngôn ngữ (**Ngôn ngữ rỗng** - $\emptyset$)

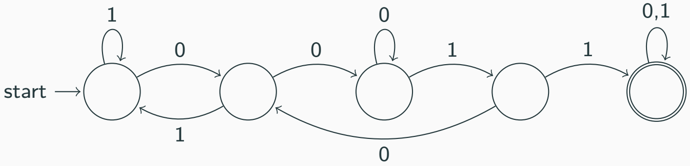
Thiết kế Ôtômat hữu hạn

Thiết kế Ôtômat hữu hạn

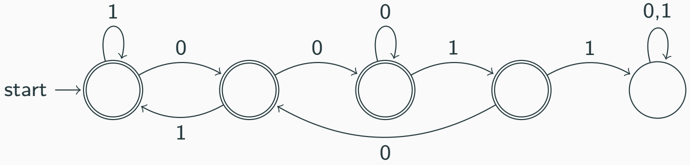
- Cho bộ chữ $\mathbf{\Sigma} = \{0,1\}$. Làm thế nào để đoán nhận tất cả các chuỗi **không** chứa chuỗi 0011?
- Trước tiên, ta thử với bài toán đơn giản hơn: Làm thế nào để đoán nhận tất cả các chuỗi có chứa chuỗi con 0011?

Thiết kế Ôtômat hữu hạn

M_1



M_2



Thiết kế Ôtômat hữu hạn

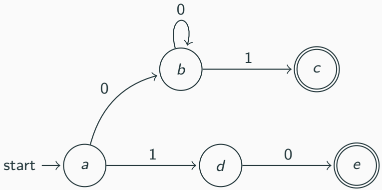
- Thuật ngữ:
  - Một máy trạng thái (FSM) chấp thuận 1 chuỗi nào đó
  - Một máy trạng thái (FSM) đoán nhận 1 ngôn ngữ
- Ký hiệu:

  $L(M_1)$ = Ngôn ngữ mà máy M_1 đoán nhận

  = Tập các chuỗi được xây dựng từ các ký tự $\{0,1\}^*$ mà trong đó có chứa chuỗi 0011 là chuỗi con

  $L(M_2)$ = Tập các chuỗi được xây dựng từ các ký tự $\{0,1\}^*$ mà trong đó không chứa chuỗi 0011 là chuỗi con
- Bản chất ngôn ngữ: **Tập** $\rightarrow L(M_1) = \overline{L(M_2)}$

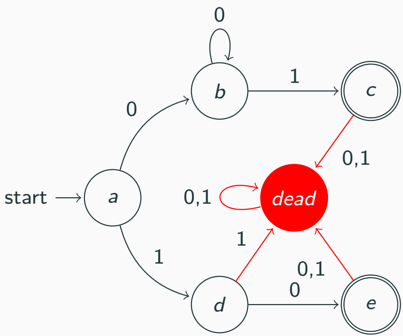
Ví dụ Ôtômat hữu hạn



- FSM trên đoán nhận các chuỗi: 10, 01, 001, 0001, . . . , $0^{+}1$
- L = {w| w là các chuỗi 01,10 hoặc các chuỗi có 1 số 1 liền ngay sau ít nhất 1 số 0}
- Các chuỗi sau điều gì sẽ xảy ra?
  - 111
  - 101010

Điểm chết (Dead states)

$$\mathbf{M = (Q, \Sigma, \delta, q_0, F)}$$



- Để tránh điểm chết → δ cần phải được định nghĩa hết các trường hợp

Ngôn ngữ chính quy

Ngôn ngữ chính quy

- Cho Ôtômat hữu hạn: $\mathbf{M = (Q, \Sigma, \delta, q_0, F)}$ và $w = w_1 w_2 \ldots w_n$ là một xâu trong đó $w_i \in \Sigma$
- M chấp thuận xâu w $\Leftrightarrow \exists$ dãy $r_0, r_2, \ldots, r_{n-1} \in Q$ thỏa mãn điều kiện:
  - $r_0 = \mathbf{q_0}$
  - $\delta(r_i, w_{i+1}) = r_{i+1}$ $(0 \leq i \leq N)$
  - $r_n \in \mathbf{F}$

  $\rightarrow$ **Định nghĩa:** Một ngôn ngữ được gọi là ngôn ngữ chính quy nếu có một Ôtômat hữu hạn nào đó đoán nhận nó
- **Ngôn ngữ nào thì không được coi là ngôn ngữ chính quy?**

Toán tử chính quy

Toán tử chính quy

Giả sử A, B là các ngôn ngữ. Ta có các toán tử chính quy sau:

- Hợp (Union): $A \cup B = \{ x \mid x \in A$ hoặc $x \in B \}$
- Ghép tiếp (Concatenate): $A \circ B = \{ xy \mid x \in A$ và $y \in B \}$
- Sao (Closure): $A^* = \{x_1 x_2 \ldots x_k \mid k \geq 0$ và mỗi $x_i \in A \}$

Ví dụ:

Giả sử ta có bộ chữ $\boldsymbol{\Sigma} = \{a,b,c,\ldots,z\}$

$A = \{aa, b\}$, $B = \{x, yy\}$

$A \cup B = \{aa, b, x, yy\}$

$A \circ B = \{ aax, aayy, bx, byy\}$

$A^* = \{\varepsilon$, aa, b, aaaa, aab, baa, bb, aaaaaa, aaaab, aabaa, aabb,. . . }

Tập đóng

- Tập hợp A + Toán tử $\equiv$ Phần tử của tập A $\rightarrow$ A là tập đóng

Định lý 1

Lớp các ngôn ngữ chính quy là đóng đối với toán tử **hợp**
$\Leftrightarrow$ Nếu A_1 và A_2 là ngôn ngữ chính quy thì $A_1 \cup A_2$ cũng là ngôn ngữ chính quy

Chứng minh

Ý tưởng:

- Giả sử M_1 đoán nhận A_1, M_2 đoán nhận A_2
- Xây dựng M để đoán nhận $A_1 \cup A_2$ $\rightarrow$ **Chứng minh bằng việc xây dựng**

Tập đóng

Chứng minh ĐL 1 (chi tiết)

- $\mathbf{M_1 = (Q_1,\Sigma,\delta_1,q_1,F_1)}$ đoán nhận A_1
- $\mathbf{M_2 = (Q_2,\Sigma,\delta_2,q_2,F_2)}$ đoán nhận A_2
- Xây dựng $\mathbf{M = (Q,\Sigma,\delta,q_0,F)}$ đoán nhận $A_1 \cup A_2$
  Trong đó:
  - $Q = \{(r_1,r_2) \mid r_1 \in Q_1$ và $r_2 \in Q_2\}$
  - $\delta((r_1,r_2),a) = (\delta_1(r_1,a),\delta_2(r_2,a))$ với mỗi $(r_1,r_2) \in Q$, $a \in \Sigma$
  - $q_0 = (q_1,q_2)$
  - $F = \{(r_1,r_2) \mid r_1 \in F_1$ hoặc $r_2 \in F_2\}$

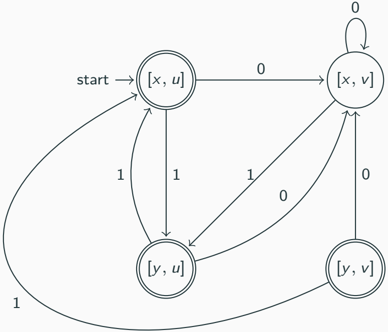
Ví dụ tính đóng của toán tử

$M_1 = (\{x,y\},\{0,1\},\delta_1,x,\{y\})$

$M_2 = (\{u,v\},\{0,1\},\delta_2,\{u\},\{u\})$

$$\mathbf{M = M_1 \cup M_2\ ??}$$

Ví dụ tính đóng của toán tử



$$M = (Q, \Sigma, \delta, q_0, F)$$

$Q =$ **???**

$\Sigma =$ **???**

$\delta =$ **???**

$q_0 =$ **???**

$F =$ **???**

Tập đóng

Định lý 2

Lớp các ngôn ngữ chính quy là đóng đối với toán tử **ghép tiếp**
$\Leftrightarrow$ Nếu A_1 và A_2 là ngôn ngữ chính quy thì $A_1 \circ A_2$ cũng là ngôn ngữ chính quy

Chứng minh

Ý tưởng:

- Giả sử M_1 đoán nhận A_1, M_2 đoán nhận A_2
- Xây dựng M để đoán nhận $A_1 \circ A_2$ $\rightarrow$ **Phần đầu đoán nhận A_1, phần sau đoán nhận A_2**
- Tuy nhiên, ta không biết xâu mà M đoán nhận bị cắt ở đâu $\rightarrow$ **Làm thế nào để biết được?**

Questions?