

Otomat hữu hạn đơn định

Trần Vĩnh Đức



Trường Đại Học Bách Khoa Hà Nội

Ngày 18 tháng 2 năm 2019

Nội dung

Xâu và ngôn ngữ

Otomat hữu hạn đơn định (DFA)

Ví dụ

Định nghĩa hình thức Tính toán

Thiết kế DFA

Các phép toán chính quy

Nội dung

Xâu và ngôn ngữ

Otomat hữu hạn đơn định (DFA)

Ví dụ

Định nghĩa hình thức Tính toán

Thiết kế DFA

Các phép toán chính quy

Bảng chữ cái

- ▶ Một **bảng chữ** là một tập hữu hạn khác rỗng.
- ▶ Mỗi phần tử gọi là chữ cái hoặc ký hiệu.

Ví dụ

- ▶ Bảng chữ nhị phân $\{0, 1\}$,
- ▶ Bảng chữ Roman $\{a, b, \dots, z\}$,
- ▶ ASCII,
- ▶ Unicode.

Xâu = Từ

Xâu trên bảng chữ Σ là một dãy hữu hạn các chữ của Σ .

Ví dụ

- ▶ 01001 là một xâu trên bảng chữ $\{0, 1\}$
- ▶ automata là một xâu trên bảng chữ $\{a, b, c, \dots, z\}$

Ta ký hiệu Σ^* là tập tất cả các xâu trên Σ . Bởi vậy

$$\{0, 1\}^* = \{\epsilon, 0, 1, 00, 01, 10, 11, 000, 001, \dots\}$$

Ký hiệu $|x|$ là độ dài xâu x . Ta có

$$|\text{automata}| = 8$$

Ký hiệu ϵ là **xâu rỗng** (xâu có độ dài 0).

Phép toán trên xâu

Phép ghép xâu hai xâu u và v là xâu được tạo thành bằng cách viết xâu u trước sau đó đến xâu v .

Ví dụ, ghép của `dog` và `house` là xâu `doghouse`.

Ngôn ngữ

Định nghĩa

Các tập con của Σ^* được gọi là các *Ngôn ngữ hình thức* hay ngắn gọn là *Ngôn ngữ* trên Σ .

Ví dụ

- ▶ \emptyset và $\{\epsilon\}$ đều là các ngôn ngữ trên bảng chữ bất kỳ.
- ▶ Các xâu trên bộ chữ $\{0, 1\}$ không chứa liên tiếp hai số 1.
- ▶ Ngôn ngữ L gồm mọi xâu trên bộ chữ $\{0, 1\}$

$$L = \Sigma^* = \{\epsilon, 0, 1, 00, 01, 10, 11, 000, 001, \dots\}.$$

Câu hỏi: Trong L có bao nhiêu xâu độ dài 4? Trong L có bao nhiêu xâu độ dài n ?

Nội dung

Xâu và ngôn ngữ

Otomat hữu hạn đơn định (DFA)

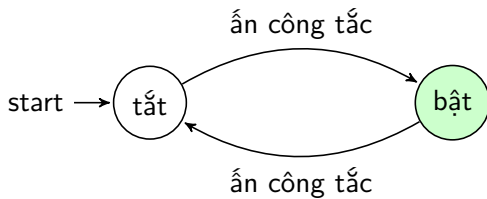
Ví dụ

Định nghĩa hình thức Tính toán

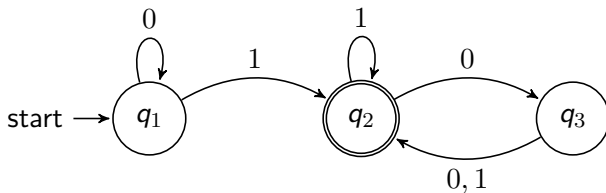
Thiết kế DFA

Các phép toán chính quy

Ví dụ

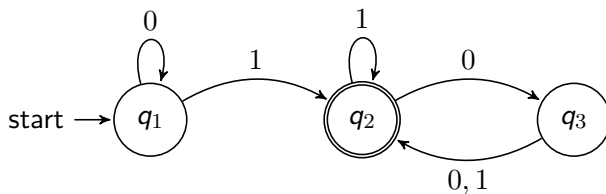


Một otomat hữu hạn đơn định



- ▶ Trạng thái: $\{q_1, q_2, q_3\}$
- ▶ Trạng thái bắt đầu: q_1
- ▶ Trạng thái chấp nhận: $\{q_2\}$
- ▶ Hàm chuyển trạng thái
- ▶ Đầu vào: một chuỗi
- ▶ Đầu ra: chấp nhận hoặc loại bỏ chuỗi vào

Một otomat hữu hạn đơn định



Với chuỗi vào 1101, máy xử lý như sau:

$$\rightarrow q_1 \xrightarrow{1} q_2 \xrightarrow{1} q_2 \xrightarrow{0} q_3 \xrightarrow{1} q_2$$

Máy chấp nhận chuỗi 1101 vì q_2 là trạng thái chấp nhận.

Định nghĩa (Otomat hữu hạn đơn định)

Một otomat hữu hạn đơn định là một bộ năm

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

trong đó

- ▶ Q là một tập hữu hạn gọi là *các trạng thái*;
- ▶ Σ là một bảng chữ;
- ▶ $\delta : Q \times \Sigma \rightarrow Q$ là *hàm chuyển trạng thái*;
- ▶ $q_0 \in Q$ là *trạng thái bắt đầu*;
- ▶ $F \subseteq Q$ là tập *trạng thái chấp nhận*.

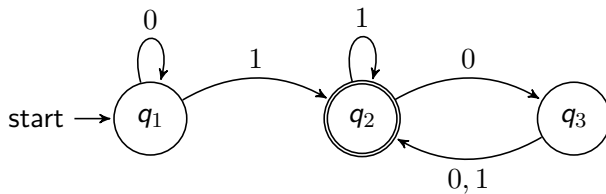
Hàm chuyển trạng thái của DFA

Hàm chuyển

$$\delta : Q \times \Sigma \rightarrow Q$$

- ▶ có hai tham số đầu vào: trạng thái q và ký hiệu vào a .
- ▶ đầu ra $p = \delta(q, a)$ là trạng thái tiếp theo của DFA khi nó đang ở trạng thái q và nhận ký hiệu vào a .

Ví dụ



1. $Q = \{q_1, q_2, q_3\}$,
2. $\Sigma = \{0, 1\}$,
3. δ mô tả bởi

| | 0 | 1 |
|-------|-------|-------|
| q_1 | q_1 | q_2 |
| q_2 | q_3 | q_2 |
| q_3 | q_2 | q_2 |

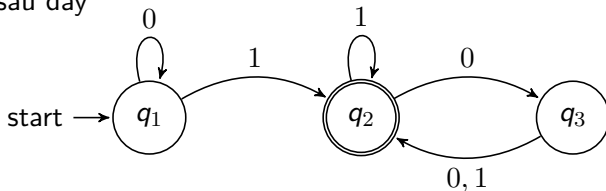
4. q_1 là trạng thái bắt đầu, và
5. $F = \{q_2\}$.

Ngôn ngữ của máy

Máy M *đ đoán nhận* ngôn ngữ A hay A là *ngôn ngữ của máy* M , và viết $L(M) = A$, nếu A là tập tất cả các xâu máy M chấp nhận.

Ví dụ

Xét M_1 sau đây



$$L(M_1) = \{w \mid w \text{ có chứa ít nhất một } 1 \text{ và} \\ \text{có một số chẵn } 0 \text{ sau } 1 \text{ cuối cùng}\}.$$

Nội dung

Xâu và ngôn ngữ

Otomat hữu hạn đơn định (DFA)

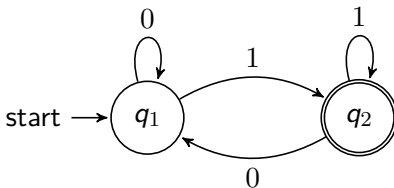
Ví dụ

Định nghĩa hình thức Tính toán

Thiết kế DFA

Các phép toán chính quy

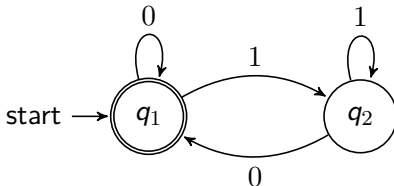
Ví dụ (DFA M_2)



- ▶ DFA M_2 có đoán nhận xâu 1101 không?
- ▶ DFA M_2 có đoán nhận xâu 110 không?
- ▶ Ngôn ngữ của M_2 là gì?

$$L(M_2) = \{w \mid w \text{ kết thúc bởi } 1\}.$$

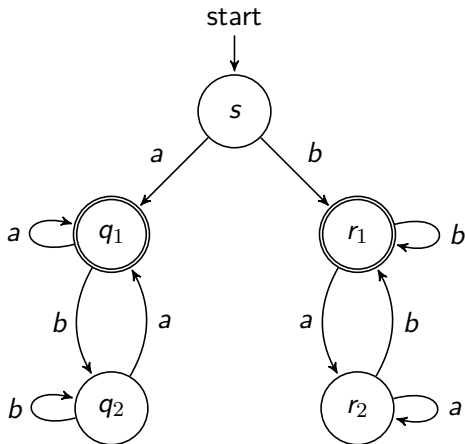
Ví dụ (DFA M_3)



- ▶ DFA M_3 đoán nhận chuỗi ε
- ▶ Ngôn ngữ của M_3 là gì?

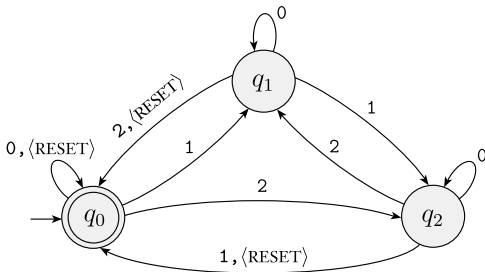
$$L(M_3) = \{w \mid w \text{ là chuỗi rỗng } \varepsilon \text{ hoặc chuỗi kết thúc bởi } 0\}.$$

Ví dụ (DFA M_4)



- Ngôn ngữ của M_4 là gì?
Các xâu bắt đầu và kết thúc bởi cùng một ký hiệu.

Ví dụ (DFA M_5)



- ▶ DFA M_5 tính tổng trên ký hiệu vào theo modun 3.
- ▶ Mỗi lần đọc được $\langle \text{RESET} \rangle$ nó đưa tổng về 0.
- ▶ Nó chấp nhận khi tổng bằng 0 theo modun 3.

Ví dụ

- ▶ Tổng quát hóa ví dụ trước cho modun i thay vì modun 3.
- ▶ DFA $B_i = (Q_i, \Sigma, \delta_i, q_0, \{q_0\})$ trong đó

$$Q_i = \{q_0, q_1, q_2, \dots, q_{i-1}\}$$

và

$$\delta_i(q_j, 0) = q_j,$$

$$\delta_i(q_j, 1) = q_k, \quad \text{với } k = j + 1 \mod i,$$

$$\delta_i(q_j, 2) = q_k, \quad \text{với } k = j + 2 \mod i,$$

$$\delta_i(q_j, \langle \text{RESET} \rangle) = q_0.$$

Nội dung

Xâu và ngôn ngữ

Otomat hữu hạn đơn định (DFA)

Ví dụ

Định nghĩa hình thức Tính toán

Thiết kế DFA

Các phép toán chính quy

Định nghĩa (Tính toán của DFA)

- ▶ Xét DFA $M = (Q, \Sigma, \delta, q_0, F)$
- ▶ và xét chuỗi $w = a_1 a_2 \dots a_n$, với $a_i \in \Sigma$.
- ▶ M **chấp nhận** chuỗi w nếu tồn tại dãy trạng thái

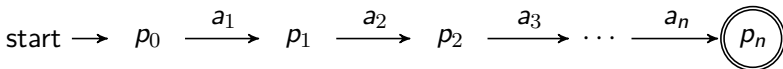
$$p_0, p_1, \dots, p_n \in Q$$

thỏa mãn ba điều kiện:

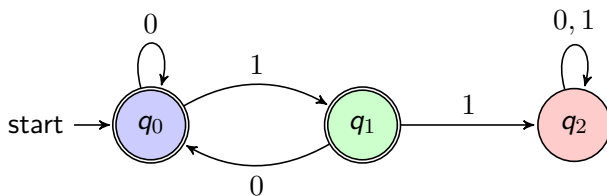
$$p_0 = q_0, \tag{1}$$

$$\delta(p_i, a_{i+1}) = p_{i+1}, \text{ với } i = 0, 1, \dots, n-1, \tag{2}$$

$$p_n \in F. \tag{3}$$



Ví dụ

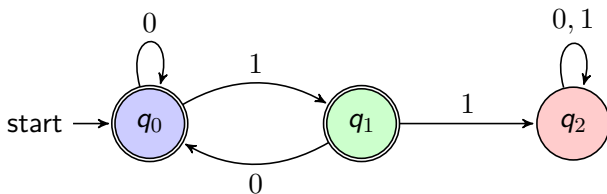


- ▶ chấp nhận chuỗi ε ;
- ▶ chấp nhận chuỗi 101;
- ▶ không chấp nhận chuỗi 10110.

Ngôn ngữ đoán nhận bởi DFA M

$$L(M) = \{w \in \Sigma^* \mid M \text{ chấp nhận } w\}.$$

Ví dụ (DFA M_6)



► Dễ thấy,

$$L(M_6) = \{w \mid w \text{ không có hai số 1 liên tiếp}\}.$$

Ngôn ngữ chính quy

Định nghĩa

Một ngôn ngữ là *chính quy* nếu có một DFA đoán nhận nó.

- ▶ Để chứng minh một ngôn ngữ là chính quy ta chỉ cần xây dựng một DFA đoán nhận nó.

Nội dung

Xâu và ngôn ngữ

Otomat hữu hạn đơn định (DFA)

Ví dụ

Định nghĩa hình thức Tính toán

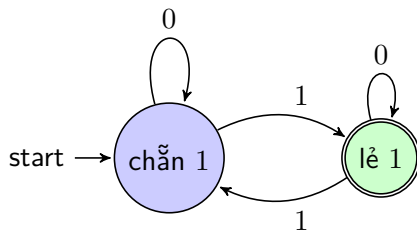
Thiết kế DFA

Các phép toán chính quy

Thiết kế DFA

Ví dụ

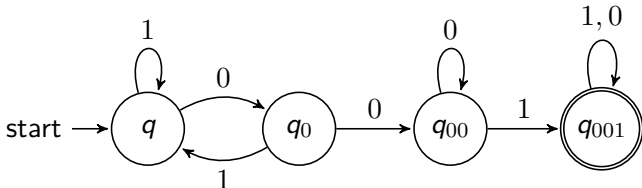
- ▶ Ngôn ngữ trên bảng chữ $\{0, 1\}$ gồm một số lẻ số 1
- ▶ Ta cần phải ghi nhớ gì?



Ví dụ

Xét ngôn ngữ gồm mọi xâu nhị phân chứa xâu con 001. Để nhận dạng mẫu 001, có bốn khả năng:

1. Chưa gặp được ký hiệu nào của mẫu
2. Chỉ gặp một 0
3. Chỉ gặp 00
4. Đã gặp mẫu 001



Nội dung

Xâu và ngôn ngữ

Otomat hữu hạn đơn định (DFA)

Ví dụ

Định nghĩa hình thức Tính toán

Thiết kế DFA

Các phép toán chính quy

Các phép toán chính quy

Định nghĩa

Xét A và B là hai ngôn ngữ. Ta định nghĩa các phép toán **hợp**, **ghép**, và **sao** như sau.

Hợp: $A \cup B = \{x \mid x \in A \text{ hoặc } x \in B\}$.

Ghép: $AB = \{xy \mid x \in A \text{ và } y \in B\}$.

Sao: $A^* = \{x_1 x_2 \dots x_k \mid k \geq 0 \text{ và mỗi } x_i \in A\}$.

Các phép toán Hợp và Ghép là các **phép toán hai ngôi**. Phép toán $*$ là **phép toán một ngôi**.

Ví dụ

- ▶ Xét bộ chữ $\Sigma = \{a, b, \dots, z\}$.
- ▶ Nếu $A = \{\text{good}, \text{bad}\}$ và $B = \{\text{boy}, \text{girl}\}$
- ▶ Ta có

$$A \cup B = \{\text{good}, \text{bad}, \text{boy}, \text{girl}\}$$

$$AB = \{\text{goodboy}, \text{goodgirl}, \text{badboy}, \text{badgirl}\}$$

$$A^* = \{\varepsilon, \text{good}, \text{bad}, \text{goodgood}, \text{goodbad}, \text{badgood}, \text{badbad}, \dots\}$$

Tính đóng

- ▶ Tập các số tự nhiên \mathbb{N} *đóng với phép nhân*: với mọi $x, y \in \mathbb{N}$, ta có $x \times y \in \mathbb{N}$
- ▶ Một cách tổng quát, tập đối tượng là **đóng** với một phép toán nếu áp dụng phép toán này cho các thành viên của tập ta được một đối tượng vẫn thuộc tập này.
- ▶ Ta sẽ chứng minh rằng: Lớp các ngôn ngữ chính quy là đóng với các phép toán chính quy Hợp, Ghép và Sao.

Định lý

Lớp ngôn ngữ chính quy đóng với Phép hợp.

Chứng minh.

- ▶ Xét DFA $M_1 = (Q_1, \Sigma, \delta_1, q_1, F_1)$ đoán nhận A_1
- ▶ Xét DFA $M_2 = (Q_2, \Sigma, \delta_2, q_2, F_2)$ đoán nhận A_2
- ▶ Xây dựng $M = (Q, \Sigma, \delta, q_0, F)$ đoán nhận $A_1 \cup A_2$, với
 - ▶ $Q = \{(r_1, r_2) \mid r_1 \in Q_1 \text{ và } r_2 \in Q_2\}$.
 - ▶ Hàm chuyển δ được định nghĩa như sau: với mỗi $(r_1, r_2) \in Q$ và $a \in \Sigma$, đặt

$$\delta((r_1, r_2), a) = (\delta_1(r_1, a), \delta_2(r_2, a))$$

- ▶ $q_0 = (q_1, q_2)$.
- ▶ $F = \{(r_1, r_2) \mid r_1 \in F_1 \text{ hoặc } r_2 \in F_2\}$.



Chú ý

- ▶ Trong xây dựng M đoán nhận $A_1 \cup A_2$ ta có

$$F = \{(r_1, r_2) \mid r_1 \in F_1 \text{ hoặc } r_2 \in F_2\}$$

có nghĩa rằng: Trạng thái chấp nhận của M là trạng thái chấp nhận của M_1 hoặc trạng thái chấp nhận của M_2 .

- ▶ Để xây dựng M' đoán nhận $A_1 \cap A_2$ ta chỉ cần thay đổi

$$F' = \{(r_1, r_2) \mid r_1 \in F_1 \text{ và } r_2 \in F_2\}$$

có nghĩa rằng: Trạng thái chấp nhận của M phải đồng thời là trạng thái chấp nhận của M_1 và của M_2 .

Định lý

Lớp ngôn ngữ chính quy đóng với phép ghép.

- ▶ Tương tự như phép hợp, bắt đầu từ DFA M_1 và M_2 đoán nhận A_1 và A_2 ta xây dựng DFA M đoán nhận A_1A_2 .
- ▶ với mỗi xâu vào w ta cắt thành $w = w_1w_2$ sao cho:
 - ▶ M_1 chấp nhận w_1
 - ▶ và M_2 chấp nhận w_2 .
- ▶ Vấn đề: Tìm vị trí cắt !
- ▶ Cách giải quyết: dùng kỹ thuật đa định.