

## VÕ TIẾN

Thảo luận kiến thức CNTT trường BK về KHMT(CScience), KTMT(CEngineering)  
<https://www.facebook.com/groups/khmt.ktmt.cse.bku>



## Mô Hình Hóa Toán Học

---

MHH - HK251

## Cuối kì

---

Thảo luận kiến thức CNTT trường BK  
về KHMT(CScience), KTMT(CEngineering)  
<https://www.facebook.com/groups/khmt.ktmt.cse.bku>

# 1 Cơ bản

## Lý thuyết cơ bản về Ngôn ngữ hình thức

### 1. Thuật ngữ cơ bản:

- **Bảng chữ cái**  $\Sigma$ : Tập hữu hạn các ký hiệu, ví dụ  $\{0, 1\}$  hay  $\{a, b, c\}$ .
- **Chuỗi** (String/Word): Dãy hữu hạn ký hiệu từ  $\Sigma$ . Chuỗi rỗng ký hiệu  $\epsilon$ .
- **Ngôn ngữ** (Language): Tập con của  $\Sigma^*$ .
- **Các phép toán trên ngôn ngữ**: Hợp ( $\cup$ ), Tích/Ghép đầu-duôi ( $\cdot$ ), Lũy thừa ( $L^n$ ), Bao đóng Kleene ( $L^*$ ).

### 2. Văn phạm (Grammar):

- Văn phạm  $G = (V, T, P, S)$ :  $V$  là tập biến,  $T$  là tập ký hiệu kết thúc,  $P$  là tập quy tắc sản xuất,  $S$  là biến bắt đầu.
- Ngôn ngữ sinh bởi  $G$ : Tập các chuỗi từ  $T$  có thể dẫn xuất từ  $S$ .

### 3. Biểu thức chính quy (Regular Expressions - RE):

- Dùng để biểu diễn các ngôn ngữ chính quy.
- Ngôn ngữ là **chính quy** nếu có một biểu thức chính quy đại diện.
- Các phép toán chính: + (hợp), · (ghép), \* (lặp/Kleene star).

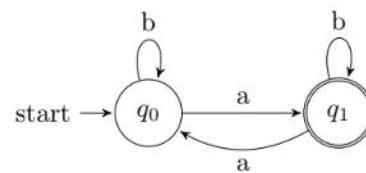
## Automaton

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

- $Q$  : Tập hợp hữu hạn các trạng thái.
- $\Sigma$  : Bảng chữ cái (alphabet), tập các ký hiệu đầu vào.
- $q_0 \in Q$  : Trạng thái khởi đầu.
- $\delta$  : Hàm chuyển trạng thái, tùy loại automaton:
- $F \subseteq Q$  : Tập các trạng thái kết thúc (accepting states).

$$Q = \{q_0, q_1\}, \quad \Sigma = \{a, b\}, \quad q_0 \text{ là trạng thái bắt đầu}, \quad F = \{q_1\}$$

$\delta$	a	b
$q_0$	$q_1$	$q_0$
$q_1$	$q_0$	$q_1$



Đặc điểm	DFA	NFA	$\varepsilon$ -NFA
Hàm chuyển trạng thái	$\delta : Q \times \Sigma \rightarrow Q$	$\delta : Q \times \Sigma \rightarrow \mathcal{P}(Q)$	$\delta : Q \times (\Sigma \cup \{\varepsilon\}) \rightarrow \mathcal{P}(Q)$
Số trạng thái tiếp theo	Chỉ 1 trạng thái	Có thể nhiều trạng thái hoặc 0	Có thể nhiều trạng thái hoặc 0, bao gồm di chuyển $\varepsilon$
Chấp nhận chuỗi	Chỉ 1 đường đi duy nhất	Có thể có nhiều đường đi	Có thể có nhiều đường đi, bao gồm chuyển $\varepsilon$
Trạng thái khởi đầu	1 trạng thái duy nhất	1 trạng thái duy nhất	1 trạng thái duy nhất

Bảng 1: So sánh DFA, NFA và  $\varepsilon$ -NFA

## Trắc nghiệm

1. Cho bảng chữ cái (alphabet)

$$\Sigma = \{a, b\}$$

Tập nào sau đây **không** phải là tập con của  $\Sigma^*$ ?

- a)  $\{\varepsilon, a, ab, bba\}$       b)  $\{a^n b^m \mid n, m \geq 0\}$       c)  $\{ab, ba, abab\}$       d)  $\{a, b, c\}$

*Dáp án: D. Ngôn ngữ vũ trụ  $\Sigma^*$  : Là tập hợp chứa tất cả các chuỗi có thể có trên bảng chữ cái  $\Sigma$ , bao gồm cả chuỗi rỗng.*

2. Cho hai chuỗi  $u$  và  $v$  trên bảng chữ cái  $\Sigma$ . Điều kiện nào sau đây là **đúng** để  $u.v = v.u$ ?

- a)  $u = v$       b)  $u = \varepsilon$       c)  $v = \varepsilon$       d) Tất cả đều đúng

*Dáp án: D*

**8 câu tiếp theo** Cho  $\Sigma = \{a, b, c\}$ ,  $L_1 = \{ab, aa, b\}$  và  $L_2 = \{b, ca, bac\}$ .

3.  $L_1 \cup L_2$  bằng tập nào sau đây?

- a)  $\{ab, aa, b\}$       b)  $\{b, ca, bac\}$       c)  $\{ab, aa, b, ca, bac\}$       d)  $\{ab, aa, ca, bac\}$

*Dáp án: C*

4.  $L_1 \cap L_2$  bằng:

- a)  $\emptyset$       b)  $\{b\}$       c)  $\{ab\}$       d)  $\{aa, b\}$

*Dáp án: B*

5. Hiệu  $L_1 - L_2$  là:

- a)  $\{b\}$       b)  $\{ca, bac\}$       c)  $\{ab, aa\}$       d)  $\{ab, aa, b\}$

*Dáp án: C*

6. Phần bù của  $L_1$  trên  $\Sigma^*$  là:

- a) Tập tất cả các chuỗi chỉ chứa ký tự  $a$  và  $b$       b)  $\Sigma^* \setminus \{ab, aa, b\}$   
c)  $\{ca, bac\}$       d)  $\emptyset$

*Dáp án: B*

7. Tập  $L_1 L_2$  gồm các phần tử nào sau đây?

- a)  $\{abb, abca, abbac, aaca, aabac, bb, bca, bbac\}$   
b)  $\{ab, aa, b, ca, bac\}$   
c)  $\{bab, baa, bb, caab, caaa, cab, bacab, bacaa, bacb\}$   
d)  $\{bb, bca, bac\}$

*Dáp án: A*

8. Tập  $L_2 L_1$  gồm các phần tử nào sau đây?

- a)  $\{abb, abca, abbac, aab, aaca, aabac, bb, bca, bbac\}$   
b)  $\{ab, aa, b, ca, bac\}$   
c)  $\{bab, baa, bb, caab, caaa, cab, bacab, bacaa, bacb\}$   
d)  $\{bb, bca, bac\}$

*Dáp án: C*

9. Tập  $L_1^2$  gồm các phần tử nào sau đây?

- a)  $\{abab, abaa, abb, aaab, aaaa, aab, bab, baa, bb\}$       b)  $\{ab, aa, b\}$   
c)  $\{\varepsilon, ab, aa, b\}$       d)  $\{abb, aab, bb\}$

*Dáp án: A.  $L_1^2 = L_1 \cdot L_1$*

10. Cho  $L_1 = \{ab, aa, b\}$ . Chuỗi nào sau đây **không** thuộc phép sao Kleene  $L_1^*$ ?

- a)  $\varepsilon$       b)  $ab$       c)  $bbbbbabaab$       d)  $aba$

*Dáp án: D*

11. Xét bảng chữ cái  $\Sigma = \{a, b, c\}$  và ngôn ngữ

$$L = \{a, abb, bba, ba, c\}.$$

Chuỗi nào sau đây **thuộc** ngôn ngữ  $L^*$ ?

- a) abaaacbb      b) aaabbbbba      c) aabacabba      d) babacbbbaaa

*Dáp án: C - a.a.ba.c.a.bba thuộc L<sup>6</sup>*

12. Xét bảng chữ cái  $\Sigma = \{a, b, c\}$  và ngôn ngữ

$$L = \{a, aab, bbc, ba\}.$$

Chuỗi nào sau đây **không** thuộc ngôn ngữ  $L^4$ ?

- a) aababbc      b) baaaaab      c) abaaabba      d) abbcaab

*Dáp án: D. aababbc=a.a.ba.bbc, baaaaab=ba.a.a.aab, abaaabba=a.ba.aab.ba*

13. Xét các ngôn ngữ  $L_1 = \emptyset$  và  $L_2 = \{a\}$ . Biểu thức nào sau đây biểu diễn đúng ngôn ngữ

$$L_1 L_2^* \cup L_1^*$$

- a)  $\emptyset$       b)  $\{\varepsilon\}$       c)  $\{a^*\}$       d)  $\{a, \varepsilon\}$

*Dáp án: B - vì L1 là tập rỗng nếu chuỗi rỗng thì đáp án là C*

14. Cho văn phạm  $G = (N, T, S, P)$  với:

$$N = \{S\}, \quad T = \{a, b\}$$

$$P : \quad S \rightarrow aSb \mid \varepsilon$$

Chuỗi nào sau đây **KHÔNG** được sinh bởi văn phạm  $G$ ?

- a)  $\varepsilon$       b)  $ab$       c)  $aabb$       d)  $abab$

*Dáp án: D*

15. Cho hai ngôn ngữ  $S$  và  $T$  trên bảng chữ cái

$$\Sigma = \{a, b\}$$

được biểu diễn bởi các biểu thức chính quy lần lượt là

$$S = (a + b^*)^*, \quad T = (a + b)^*$$

Phát biểu nào sau đây là **đúng**?

- a)  $S \subset T$       b)  $S = T$       c)  $T \subset S$       d)  $S \cap T = \emptyset$

*Dáp án: B - mọi giá trị trong tập S đều có trong tập T và ngược lại*

16. Công thức nào sau đây **KHÔNG** phải là biểu thức chính quy của ngôn ngữ bao gồm mọi chuỗi bit chứa **ít nhất ba bit 0**?

- a)  $(0 + 1)^*0(0 + 1)^*0(0 + 1)^*0(0 + 1)^*$       b)  $1^*01^*01^*0(0 + 1)^*$   
c)  $(0 + 1)^*000(0 + 1)^*$       d)  $(0 + 1)^*01^*01^*01^*$

*Dáp án: C - chuỗi 10100 không được sinh ra bởi câu C*

17. Biết rằng một ngôn ngữ  $L(G)$  được sinh bởi một văn phạm  $G = (V, T, P, S)$  với:

$$V = \{S\}, \quad T = \{a, b\}$$

và

$$L(G) = \{w \in T^* \mid S \Rightarrow^* w\} = \{a^n b^m \mid n + m \text{ là số chẵn}\}.$$

Trong đó  $\Rightarrow^*$  là phép suy diễn trên  $T^*$ . Tập luật sinh  $P$  nào sau đây là **đúng**?

- a)  $P = \{S \rightarrow S_1, S_1 \rightarrow \varepsilon, S_1 \rightarrow aaS_1bb\}$
- b)  $P = \{S \rightarrow S_2, S_2 \rightarrow \varepsilon, S_2 \rightarrow aS_1b\}$
- c)  $P = \{S \rightarrow S_1, S \rightarrow S_2, S_1 \rightarrow \varepsilon, S_1 \rightarrow aaS_1bb, S_2 \rightarrow aS_1b\}$
- d)  $P = \{S \rightarrow S_1, S \rightarrow S_2, S_1 \rightarrow \varepsilon, S_1 \rightarrow a^2S_1b^2, S_2 \rightarrow bS_1a\}$

*Dáp án: C. khi cả 2 cùng chẵn hoặc cả 2 cùng lẻ với s1 là chẵn và s2 là lẻ, câu d sai cho phần lẻ đang bị ngược ba thay vì ab*

18. Biết rằng một ngôn ngữ  $L(G)$  được sinh bởi ngữ pháp

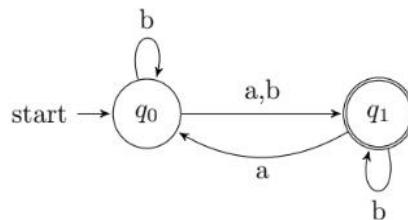
$$G = (V, T, P, S)$$

Tập ký hiệu tận cùng:  $T = \{x, y\}$ ,  $\varepsilon$  là chuỗi rỗng, Tập luật sinh:  $P = \{S \rightarrow xSx, S \rightarrow xyx, S \rightarrow \varepsilon\}$ , Ký hiệu  $\Rightarrow^*$  là phép suy diễn trên  $T^*$ . Ngôn ngữ được sinh bởi ngữ pháp  $G$  là:  $L(G) = \{w \in T^* \mid S \Rightarrow^* w\}$ . Ngôn ngữ nào sau đây được sinh ra bởi ngữ pháp  $G$ ?

- a)  $L(G) = \{x^n y^{n+1} \mid n \geq 1\}$
- b)  $L(G) = \{x^n y^n \mid n > 0\}$
- c)  $L(G) = \{x^n y x^n \mid n \geq 1\} \cup \{x^{2n} \mid n \geq 0\}$
- d)  $L(G) = \{x^{n+1} y^n \mid n > 0\}$

*Dáp án: C.*

19. Xét automaton sau:

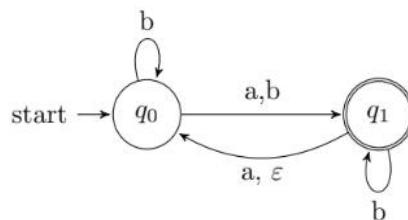


Automaton trên thuộc loại nào?

- a) DFA
- b) NFA
- c)  $\varepsilon$ -NFA
- d) không xác định

*Dáp án: B. Tại trạng thái  $q_0$ , với ký hiệu  $b$  có thể chuyển đến hai trạng thái  $q_0$  và  $q_1$ , nên automaton là NFA.*

20. Xét automaton sau:

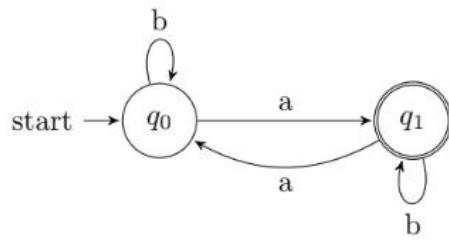


Automaton trên thuộc loại nào?

- a) DFA
- b) NFA
- c)  $\varepsilon$ -NFA
- d) không xác định

*Dáp án: C. Tại trạng thái  $q_0$ , với ký hiệu  $b$  có thể chuyển đến hai trạng thái  $q_0$  và  $q_1$ , nên automaton là NFA kết hợp với  $\varepsilon$ .*

21. Xét automaton sau:



Automaton trên thuộc loại nào?

- a) DFA      b) NFA      c)  $\epsilon$ -NFA      d) không xác định

*Dáp án: A.*

## 2 Các thuật toán

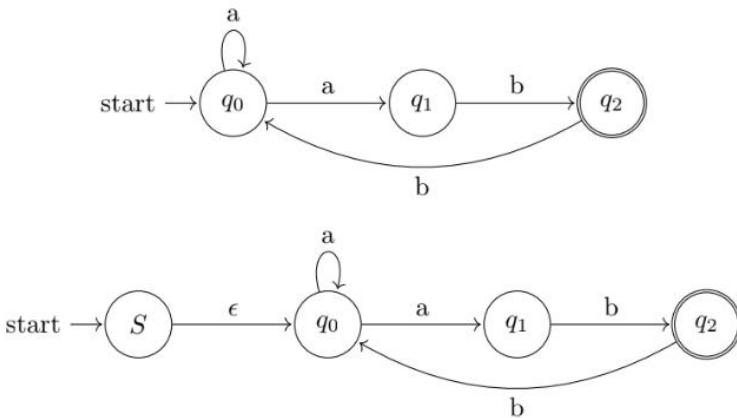
### 2.1 Biểu thức chính quy và Automata

RE	NFA
$A$	start $\rightarrow$ $q_0 \xrightarrow{A} q_1$
$\epsilon$	start $\rightarrow$ $q_0 \xrightarrow{\epsilon} q_1$
$A + B$	start $\rightarrow$ $q_0 \xrightarrow{A} q_1$ $q_0 \xrightarrow{B} q_1$
$AB$	start $\rightarrow$ $q_0 \xrightarrow{A} q_1 \xrightarrow{B} q_2$
$A^*$	start $\rightarrow$ $q_0 \xrightarrow{A} q_0$
$AB^*$	start $\rightarrow$ $q_0 \xrightarrow{A} q_1 \xrightarrow{B} q_1$
$A^*B$	start $\rightarrow$ $q_0 \xrightarrow{A} q_0 \xrightarrow{B} q_1$

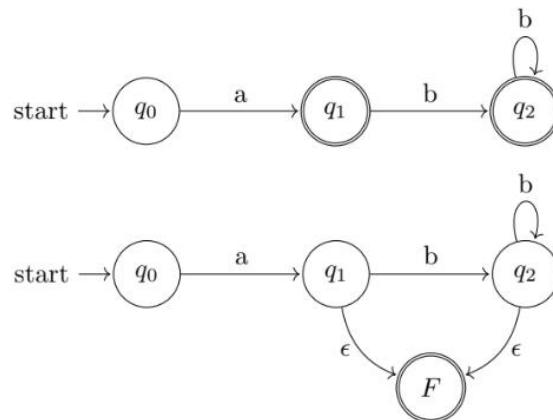
Bảng 2: Quy tắc chuyển đổi biểu thức chính quy sang NFA

### Các bước chuyển đổi Automata sang biểu thức chính quy

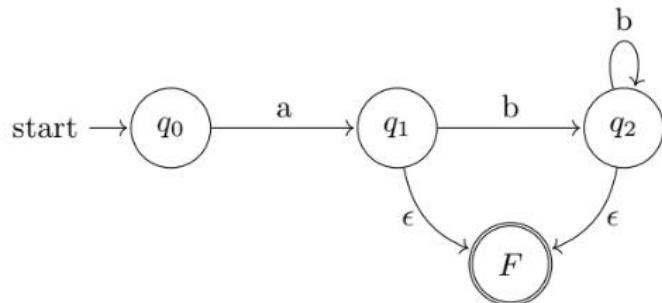
- Không có cạnh vào ở trạng thái khởi đầu (**thêm node S tới q0 bằng  $\epsilon$** ).



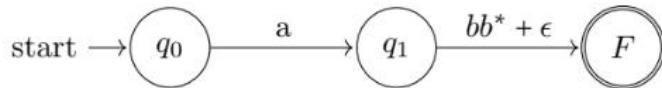
- Không có cạnh ra ở trạng thái kết thúc và chỉ có một trạng thái kết thúc duy nhất. (**thêm node F đi bằng  $\epsilon$  từ các node final**).



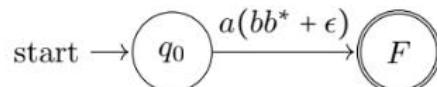
- Loại bỏ trạng thái trung gian một cách lần lượt.



**Loại bỏ  $q_2$**



**Loại bỏ  $q_1$**



## 2.2 Chuyển đổi $\epsilon$ -NFA, NFA, DFA and MinDFA

xem bài giảng <https://www.youtube.com/watch?v=5tVrYj8xtsk>