

LÝ THUYẾT TÍNH TOÁN

BÀI 4: BIỂU THỨC CHÍNH QUY

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

1. Khái niệm
2. Định nghĩa hình thức
3. Sự tương đương với Ôtômat hữu hạn

Khái niệm

- **Biểu thức chính quy:** Sử dụng các toán tử chính quy để biểu diễn một biểu thức mô tả ngôn ngữ
Ví dụ: $(0\cup 1)0^*$
→ Tất cả các chuỗi bắt đầu bằng 1 ký tự 0 hoặc 1 và sau đó là một số nào đó các ký tự 0
- Vai trò của Biểu thức chính quy: Là một phương pháp mạnh để mô tả 1 mẫu văn bản nào đó
→ Trong một số ngôn ngữ lập trình đều ứng dụng kỹ thuật mô tả mẫu bằng biểu thức chính quy (**Regular Expression**)

Định nghĩa hình thức

Định nghĩa hình thức của biểu thức chính quy

Ta nói R là một biểu thức chính quy nếu R là:

1. a với a là ký hiệu nào đó trong bộ chữ Σ
2. ϵ
3. \emptyset
4. $(R_1 \cup R_2)$ trong đó R_1 và R_2 là các biểu thức chính quy
5. $(R_1 \circ R_2)$ trong đó R_1 và R_2 là các biểu thức chính quy
6. (R_1^*) trong đó R_1 là một biểu thức chính quy

Độ ưu tiên của các toán tử chính quy

- Toán tử sao có độ ưu tiên cao nhất
$$ab^* = a(b^*) \neq (ab)^*$$
- Toán tử ghép tiếp có độ ưu tiên cao hơn toán tử hợp
$$a \circ b \cup c = (a \circ b) \cup c \neq a(b \cup c)$$
- Một số ký hiệu khác:
 - Hoặç (Union): $ab|c = (ab)|c \neq a(b|c)$
 - Sao: $a^* = \{a\} = \{a\}^*$
 - 1 hoặc nhiều: $a^+ = aa^* = \{a\}^+$
 - Tùy chọn: $[a] = a|\epsilon = (a \cup \epsilon) = a?$

Ví dụ về độ ưu tiên toán tử chính quy

- $aab \cup caab \cup caa = \text{????}$
- $aab | caab | caa = \text{????}$
- $d \cup ab^* cd^* = \text{????}$
- $d | ab^* cd^* = \text{????}$

Ví dụ về độ ưu tiên toán tử chính quy

- $aab \cup caab \cup caa = (aab) \cup (caab) \cup caa$
- $aab|caab|caa = (aab)|(caab)|(caa)$
- $d \cup ab^* cd^* = d \cup (a(b^*)c(d^*))$
- $d|ab^* cd^* = d|(a(b^*)c(d^*))$

Ví dụ biểu thức chính quy

Giả thiết sử dụng bộ chữ $\Sigma = \{0,1\}$

1. $0^*10^* = \{w|w \text{ chỉ có một ký hiệu } 1\}$
2. $\Sigma^*1\Sigma^* = \{w|w \text{ có ít nhất một ký hiệu } 1\}$
3. $\Sigma^*001\Sigma^* = \{w|w \text{ có chứa xâu con } 001\}$
4. $1^*(01^+)^* = \{w|sau \text{ mỗi ký hiệu } 0 \text{ trong } w \text{ sẽ có ít nhất } 1 \text{ ký hiệu } 1\}$
5. $(\Sigma\Sigma)^* = \{w|w \text{ là xâu có độ dài là một số chẵn}\}$
6. $01 \cup 10 = \{01, 10\}$