

Тема 3: Регулярні множини

- 1. Регулярні множини і регулярні вирази
- 2. Побудова регулярного виразу по праволінійній граматиці
- 3. Алгоритм побудови праволінійної граматики по регулярному виразу

1. Регулярні множини і регулярні вирази

Нехай V - скінченний алфавіт. Рекурсивно *регулярна множина* в алфавіті визначається так:

- 1) \emptyset –регулярна множина в алфавіті V ;
- 2) $\{e\}$ –регулярна множина в алфавіті V ;
- 3) $\{a\}$ –регулярна множина в алфавіті V
для всіх $a \in V$;
- 4) якщо P, Q регулярні множини в V , то
 $P \cup Q, PQ, P^*$ – регулярні множини в
алфавіті V ;
- 5) Ніщо інше не є регулярною множиною в V .

Регулярний вираз в алфавіті V визначається рекурсивно:

- 1) \emptyset - регулярний вираз, що позначає регулярну множину \emptyset ;
- 2) e - регулярний вираз, що позначає регулярну множину $\{e\}$;
- 3) якщо $a \in V$, то a - регулярний вираз, що позначає регулярну множину $\{a\}$;
- 4) якщо p, q - регулярні вирази, що позначають регулярні множини P, Q , то
 - $(p+q)$ регулярний вираз, що позначає регулярну множину $P \cup Q$;
 - (pq) регулярний вираз, що позначає регулярну множину PQ ;
 - $(p)^*$ регулярний вираз, що позначає регулярну множину P^* ;
- 5) ніщо інше не є регулярним виразом.

Приклади регулярних виразів.

- 1) 01 позначає множину $\{01\}$;
- 2) 0^* позначає множину $\{0\}^*$;
- 3) $(0+1)^*$ позначає множину $\{0,1\}^*$;
- 4) $(0+1)^*011$ позначає множину усіх ланцюжків, що складаються з нулів і одиниць і закінчуються ланцюжком 011 ;
- 5) $(a+b)(a+b+0+1)^*$ позначає множину усіх ланцюжків з множини $\{a,b,0,1\}^*$, що починаються з a або b ;
- 6) $(11+00)^*(01+10)(11+00)^*$

Тотожності над регулярними виразами

- 1) $\alpha + \beta = \beta + \alpha$
- 2) $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$
- 3) $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$
- 4) $\alpha e = e\alpha = \alpha$
- 5) $\alpha^* = \alpha + \alpha^*$
- 6) $\alpha + \alpha = \alpha$
- 7) $\emptyset^* = e$
- 8) $\alpha(\beta\gamma) = (\alpha\beta)\gamma$
- 9) $(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$
- 10) $\alpha\emptyset = \emptyset\alpha = \emptyset$
- 11) $(\alpha^*)^* = \alpha^*$
- 12) $\alpha + \emptyset = \alpha$

Доведення. 1) Нехай α, β позначають множини A, B .

Тоді $\alpha + \beta$ позначає $A \cup B$,

$\beta + \alpha$ позначає $B \cup A$.

Оскільки $A \cup B = B \cup A$, то $\alpha + \beta = \beta + \alpha$.

2. Побудова регулярного виразу по праволінійній граматиці

Рівняння з регулярними коефіцієнтами:

$$X = \alpha X + \beta, \quad (1)$$

α, β – регулярні вирази.

$$X = \alpha^* \beta \quad (2)$$

– розв'язок (1)

$$\alpha^* \beta = \alpha \alpha^* \beta + \beta \quad (3)$$

При $\alpha = \epsilon$

$X = \alpha^*(\beta + \gamma)$ - також розв'язок (1), оскільки $\beta + \gamma =^{(6)} \beta + \gamma + \beta$.

$X = \alpha^* \beta$ - найменша нерухома точка рівняння (1).

$$G = (\{S, A, B\}, \{0,1\}, P, S)$$

$P:$

$$S \rightarrow 0A \mid 1S \mid e$$

$$A \rightarrow 0B \mid 1A$$

$$B \rightarrow 0S \mid 1B$$

$$S \rightarrow x_1, A \rightarrow x_2, B \rightarrow x_3$$

$$\begin{cases} x_1 = 0x_2 + 1x_1 + e \\ x_2 = 0x_3 + 1x_2 \\ x_3 = 0x_1 + 1x_3 \end{cases}$$

$$X = \alpha X + \beta \quad X = \alpha^* \beta$$

$$x_1 = \underset{\alpha}{1} x_1 + \underset{\beta}{0x_2 + e}$$

$$x_1 = 1^* (0x_2 + e)$$

$$x_2 = 1^* 0x_3$$

$$x_3 = 01^* (0x_2 + e) + 1x_3$$

$$\begin{aligned} x_3 &= 01^* 0x_2 + 01^* + 1x_3 = \\ &= 01^* 01^* 0x_3 + 01^* + 1x_3 = \\ &= (\underset{\alpha}{01^* 01^* 0 + 1^*}) x_3 + \underset{\beta}{01^*} \end{aligned}$$

$$x_3 = (01^* 01^* 0 + 1)^* 01^*$$

$$x_2 = 1^* 0 (01^* 01^* 0 + 1)^* 01^*$$

$$x_1 = 1^* (01^* 0 (01^* 01^* 0 + 1)^* 01^* + e)$$

$$G = (\{S, A, B\}, \{0, 1\}, P, S)$$


$$l = 1^*(01^*0(01^*01^*+1)^*01^*+e)$$

Завдання додому:



Розв'язати систему рівнянь із регулярними коефіцієнтами:

$$\begin{cases} A_1 = (01^* + 1)A_1 + A_2 \\ A_2 = 11 + 1A_1 + 00A_3 \\ A_3 = e + A_1 + A_2 \end{cases}$$

Програмна реалізація

 **Введіть правила**

Нетермінали ABS	Термінали 0 1 e
Граматика	
S->	<input type="text"/> <input type="text"/> <input type="text"/>
A->	<input type="text"/> <input type="text"/>
B->	<input type="text"/> <input type="text"/>

РЕЗУЛЬТАТ

Граматика

S-> 0A | 1S | e
A-> 0B | 1A
B-> 0S | 1B

Система

$x_1 = 0x_2 + 1x_1 + e$
 $x_2 = 0x_3 + 1x_2$
 $x_3 = 0x_1 + 1x_3$

Розв'язок системи (Регулярний вираз)

$reg = 1^*(01^*0(01^*01^*0+1)^*01^* + e)$

Проміжні результати

$x_1 = 1^*(01^*0(01^*01^*0+1)^*01^* + e)$
 $x_2 = 1^*0(01^*01^*0+1)^*01^*$
 $x_3 = (01^*01^*0+1)^*01^*$



3. Алгоритм побудови праволінійної граматики по регулярному виразу

1. Для регулярного виразу \emptyset : $G = (\{S\}, T, \emptyset, S)$
2. Для регулярного виразу e : $G = (\{S\}, T, \{S \rightarrow e\}, S)$
3. Для регулярного виразу $a, a \in T$: $G = (\{S\}, T, \{S \rightarrow a\}, S)$

Нехай ми маємо регулярні вирази

$$I_1: G_1 = (N_1, T, P_1, S_1) \quad N_1 \cap N_2 = \emptyset.$$

$$I_2: G_2 = (N_2, T, P_2, S_2)$$

4. Для $I_1 + I_2$: $G_3 = (N_1 \cup N_2 \cup S_3, T, P_1 \cup P_2 \cup \{S_3 \rightarrow S_1 \mid S_2\}, S_3)$
5. Для $I_1 I_2$: $G_4 = (N_1 \cup N_2, T, P_4, S_1)$

P_4 :

а) якщо $A \rightarrow xB \in P_1$, то це правило належить P_4

б) якщо $A \rightarrow x \in P_1$, тоді $A \rightarrow xS_2 \in P_4$

в) всі правила з P_2 належать P_4 .

6. Для $(l_1)^* : G_5 = (N_1 \boxtimes S_5, T, P_5, S_5)$

P_5 :

а) якщо $A \rightarrow xB \in P_1$, то це правило належить P_5

б) якщо $A \rightarrow x \in P_1$, то $\begin{cases} A \rightarrow x \\ A \rightarrow xS_5 \end{cases} \in P_5$

в) $S_5 \rightarrow S_1 | e$ належить P_5

Приклад 2. $l=(101)^*(010)^*$

$l_1 = 101$

$l_2 = 010$

$P_1 : S_1 \rightarrow 101$

$P_2 : S_2 \rightarrow 010$

$l_3 = (101)^*$

$l_4 = (010)^*$

P_3 :

P_4 :

$S_3 \rightarrow S_1 | e$

$S_4 \rightarrow S_2 | e$

$S_1 \rightarrow 101$

$S_2 \rightarrow 010$

$S_1 \rightarrow 101S_3$

$S_2 \rightarrow 010S_4$

P_5 :

$S_3 \rightarrow S_1 | S_4$

$S_1 \rightarrow 101S_4$

$S_1 \rightarrow 101S_3$

$S_4 \rightarrow S_2 | e$

$S_2 \rightarrow 010$

$S_2 \rightarrow 010S_4$

$P_5:$

$S_3 \rightarrow S_1 | S_4$

$S_1 \rightarrow 101S_4$

$S_1 \rightarrow 101S_3$

$S_4 \rightarrow S_2 | e$

$S_2 \rightarrow 010$

$S_2 \rightarrow 010S_4$

$S_3 = S$

$S_1 = A$

$S_2 = B$

$S_4 = C$

$P_5 = P$

$P:$

$S \rightarrow A | C$

$A \rightarrow 101C | 101S$

$C \rightarrow B | e$

$B \rightarrow 010 | 010C$

$l=(101)^*(010)^*$

$G=(\{S,A,B,C\},\{0,1\}, P, S)$

Завдання додому: Для регулярного виразу $l=(01)^*(0+1)$ побудувати праволінійну граматику.

Програмна реалізація

