

ФОРМАЛНИ ЕЗИЦИ И ГРАМАТИКИ

Думи и езици.

➤ **Дефиниция:** Дума над азбуката V ще наричаме всяка крайна редица от букви на V . Дума, в която не влиза нито една буква се нарича празна дума, която се означава с ε . С V^* се означава множеството от всички думи над V , включително и празната.

➤ **Дефиниция:** Всяко подмножество L на V^* се нарича **формален език** над азбуката V .

При сечение, обединение, допълнение на два езика ще разбираме езика, който се получава от дадените езици, разглеждани като множества.

Под произведение на езиците L_1 и L_2 ще разбираме езика

$$L_1 L_2 = \{\alpha \in V^* \mid \alpha = \beta\gamma, \beta \in L_1, \gamma \in L_2\},$$

т.е. множеството от всички думи, които се представят като конкатенация на две думи.

Първата от които е от L_1 , а втората от L_2 .

Задача 1. Нека $V = \{0,1,2\}$ е азбука.

а) Образувайте формалния език L_1 , който се състои от всички двубуквени и трибуквени думи, съставени от различни букви.

б) за $L_2 = \{021, 20, 103, 412, 40\}$, намерете:

$$L_1 \cup L_2$$

$$L_1 \cap L_2$$

$$L_1 - L_2$$

в) $L_2 L_3 = ?$, като $L_3 = \{aa, bb, c\}$

г) $L_3^2 = ?$

➤ **Дефиниция:** Под n -та степен на езика L ще разбираме езика

$$L^n = \underbrace{LL \dots L}_n$$

като по дефиниция $L^0 = \varepsilon$.

➤ **Дефиниция:** **Итерация** L^* на произволен език L ще наричаме обединението на всички степени на L , т.е.

$$L^* = \bigcup_{n \geq 0} L^n$$

Итерацията се състои от всички възможни конкатенации на произволен брой думи на L . Ако всички думи са с дължина 1, тогава $L^* = V^*$ - всички думи над V .

Ако $L = \{a, b, c\}$, то $L.L = \{aa, ab, ac, ba, bb, bc, ca, cb, cc\}$, т.е. това са всички двубуквени думи, L^3 – са всички трибуквени и т.н. V^* са всички думи над азбуката V .

Задача 2. Нека $L_1 = \{aba\}$ и $L_2 = \{b\}$ са формални езици над азбуката $V = \{a, b\}$. Опишете думите от езика $L = L_1^*L_2$.

Задача 3. Нека $V = \{0\}$ $W = \{1\}$. Опишете кои думи изграждат следните езици:

А) $(V \cup W)^*$

Б) $V^*.W^*$

В) Вярно ли е, че $(V \cup W)^* = V^*.W^*$

Граматики

➤ **Дефиниция: Граматика-генератор** е наредена четворка от вида: $\Gamma = \langle V, W, S, P \rangle$, където:

- V е азбука на терминалните символи (азбука на пораждащия език)
- W е вътрешна нетерминална азбука от вътрешни състояния на УУ (синтактични категории в езика), като $V \cap W = \emptyset$
- $S \in W$ е начално състояние
- P е крайно множество от правила на пораждащата граматика, които представляват наредени двойки от думи, съставени от терминални и нетерминални символи $\langle \alpha, \beta \rangle$ над $V \cup W$, като в α има поне един нетерминален символ (за β не е задължително).

Задача 4. Кои думи поражда граматиката:

$$\Gamma = \langle \{a, b\}, \{s\}, s, \{s \rightarrow asb, s \rightarrow ab\} \rangle$$

Задача 5. Кои думи поражда граматиката:

$$\Gamma = \langle \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}, \{s\}, s, \{s \rightarrow s0, s \rightarrow s1, s \rightarrow s2 \dots s \rightarrow s9, s \rightarrow 0, s \rightarrow 1, s \rightarrow 2, \dots s \rightarrow 9\} \rangle$$

Задача 6. Кои думи поражда граматиката

$$\Gamma = \langle \{a, b, c\}, \{s\}, s, \{s \rightarrow aca; s \rightarrow asa, s \rightarrow bcb, s \rightarrow bsb\} \rangle$$

Задача 7. Опишете езика, породен от автоматната граматика:

$$\Gamma = \langle \{0, 1\}, \{c, a, b\}, c, \{c \rightarrow 0a, a \rightarrow 0b, b \rightarrow 1\} \rangle$$

Задача 8. Постройте граматика, която да поражда всички думи от нули и единици, започващи с 10, т.е.

$$L = \{10\alpha \mid \alpha \in \{0, 1\}^*\}.$$

Задача 9. Постройте граматика, която да поражда всички думи от следните езици:

$$L_1 = \{abab\}$$

$$L_2 = \{a^n, n \geq 1\}$$

$$L_3 = \{a^{2^n}, n \geq 1\}$$

$$L_4 = \{a^{3^n}b, n \geq 0\}$$

$$L_5 = \{a^n b^n, n \geq 1\} \quad L_6 = \{a^n b^{2^n}, n \geq 1\}$$

$$L_7 = \{\alpha aa, \alpha \in \{a, b\}^*\}$$

Автоматни граматики.

➤ **Дефиниция:** Пораждаща граматика $\Gamma = \langle V, W, S, P \rangle$, чиито правила са от вида : $A \rightarrow aB$, $A \rightarrow a$, като $(A, B \in W, a \in V)$ се нарича **автоматна** или граматика от тип 3.

Задача 1. Да се напише автоматна граматика, която да поражда думите от езика над азбуката $V = \{0, 1\}^*$.

а) $L_1 = \{10\alpha, \alpha \in \{0, 1\}^*\};$

б) $L_2 = \{\alpha 10, \alpha \in \{0, 1\}^*\};$

в) $L_3 = \{\alpha 10\beta, \alpha \in \{0, 1\}^*\};$

Задача 2. Да се напише автоматна граматика, която да поражда думите от езика над азбуката $V = \{a, b\}^*$.

а) $L_1 = \{a^n b, n \geq 0\};$

б) $L_2 = \{a^n b, n \geq 1\};$

в) $L_3 = \{a b^n, n \geq 0\};$

г) $L_3 = \{a b^n, n \geq 1\};$

д) $L_4 = \{a^n b^m, n, m \geq 0\};$

е) $L_5 = \{a\alpha b, \alpha \in \{a, b\}^*\};$

ж) $L_6 = \{ab\alpha ba, \alpha \in \{a, b\}^*\};$

з) $L_7 = \{\alpha, \alpha \in \{a, b\}^* \text{ и започва и завършва с една и съща буква} \}.$

Допълнителни задачи:

Задача 1. Нека $A = \{a\}$ и $B = \{b\}$. Опишете какви думи съдържат следните формални езици:

- | | |
|-------------------|--------------------------|
| а) AB^* ; | е) $A^* \cap B^*$ |
| б) $A^* B^*$; | ж) $(A \cup B)^*$ |
| в) $(AB)^*$; | з) $(A \cap B)^*$ |
| г) $A \cup B$; | и) $(A \cup B)^* ABA$ |
| д) $A^* \cup B^*$ | й) $(A^* \cup AB)^* A$. |

Задача 2. Определете езика породен от граматиката:

- а) $\Gamma = \langle \{0, 1\}, \{S\}, \{S \rightarrow 0S \mid 1S \mid \varepsilon\} \rangle$;
- б) $\Gamma = \langle \{0, 1\}, \{S, A\}, S, \{S \rightarrow 011 \mid 0S1 \mid 01A, A \rightarrow 1A \mid 1\} \rangle$;
- в) $\Gamma = \langle \{0, 1\}, \{S\}, S, \{S \rightarrow 1 \mid 0S1 \mid 1S0 \mid 0S0 \mid 1S1\} \rangle$;
- г) $\Gamma = \langle \{a, b, c\}, \{S, B, C\}, S, \{S \rightarrow aSBC \mid abC, CB \rightarrow BC, bB \rightarrow bb, bC \rightarrow bc, cC \rightarrow cc\} \rangle$;