Теория вычислений

Недели 1-2. Регулярные языки и конечные автоматы

Если в задаче не указан алфавит, то языки в задаче определены над алфавитом $\Sigma = \{a, b\}$.

1. Доказать, что следующие языки являются регулярными:

а) язык, состоящий из пустого слова;

б) язык, состоящий из всех слов.

2. Верно ли, что **а)** $\varepsilon \in \{a, aab, aba\}$? **б)** $\varnothing \in \{a, aab, aba\}$?

3. Построить регулярное выражение для

а) языка из слов, содержащих в качестве подслова слово aab;

б) языка, слова которого не содержат подслово ab;

в) языка всех слов с чётным числом букв a.

4. Построить НКА, распознающий язык из слов, содержащих подслово ab.

5. Являются ли следующие языки регулярными?

а) язык L_1 из слов, в которых подслов ab и ba одинаковое количество: $L_1 = \{w : |w|_{ab} = |w|_{ba}\}$;

б) $L_2 = \{a^{2^n} \mid n \in \mathbb{N}\} \subseteq a^*$;

в) язык L_3 состоит из двоичных записей чисел, взаимно простых с 2016;

г) язык L_4 всех слов в алфавите $\{0, 1\}$, которые представляют числа в двоичной записи, дающие остаток два при делении на три (слово читается со старших разрядов).
Например, $001010\,(1010_2 = 10_{10} = 3 \times 3 + 1) \notin L_4$, а $10001\,(10001_2 = 17_{10} = 5 \times 3 + 2) \in L_4$.

6. Пусть $A, B \in \mathsf{REG}$; $C, D \notin \mathsf{REG}$. Верны ли следующие утверждения (для произвольных A, B, C, D):

а) $A \cap B \in \mathsf{REG}$; **б)** $A \setminus B \in \mathsf{REG}$; **в)** $A \cup C \notin \mathsf{REG}$;
г) $A \cap C \notin \mathsf{REG}$; **д)** $C \cup D \notin \mathsf{REG}$; **е)** $C \cap D \notin \mathsf{REG}$.

7. Определим операцию обращения R. Пусть $w = w_1 w_2 \dots w_n,\ w_i \in \Sigma$, тогда $w^R = w_n w_{n-1} \dots w_1$. Верно ли, что если L регулярный язык, то и язык $L^R = \{w \mid w^R \in L\}$ является регулярным?

8. Пусть $X_1, X_2, \dots, X_n, \dots$ бесконечное семейство регулярных языков. Верно ли, что

а) $\bigcup\limits_{n=1}^{\infty} X_i \in \mathsf{REG}$? **б)** $\bigcap\limits_{n=1}^{\infty} X_i \in \mathsf{REG}$?

9. К языку L_1 добавили конечный язык R и получили язык L ($L = L_1 \cup R$). Язык L оказался регулярным. Верно ли, что язык L_1 мог быть нерегулярным?

10. Покажите, что следующий язык удовлетворяет лемме о накачке для регулярных языков, но сам регулярным не является:

$$L = \{a^k b^{2^i} \mid i, k \geqslant 0\} \cup \{b^j \mid j \geqslant 0\} \cup \{a^m b^n \mid m > 1, n \geqslant 0\}.$$

11. Язык $L_{b,i}$ состоит из слов, в которых на i-ой позиции от конца стоит b: $L_{b,i} = \{w : |w| = n,\ w_{n-i} = b\}$.

а) Постройте НКА, распознающий язык $L_{b,3}$.

б) Постройте минимальный ДКА, распознающий язык $L_{b,3}$.

в) Докажите, что для каждого i существует НКА с i состояниями, распознающий язык $L_{b,i}$, а минимальный ДКА, распознающий язык $L_{b,i}$, имеет 2^i состояний.

12. Верно ли, что если пересечение языков $L_1,\ L_2 \subseteq \{a, b\}^*$ содержит язык $F = \{a^n b^n \mid n \geqslant 1\}$: $F \subseteq L_1 \cap L_2$, то хотя бы один из языков L_1 и L_2 является нерегулярным?

Недели 1-2. Регулярные языки и конечные автоматы

Определение. Множество M называют моноидом, если M замкнуто относительно операции $\circ$ и содержит единицу:

$$\forall x, y \in M : x \circ y \in M;\ \exists 1 \in M : \forall x \in M\ 1 \circ x = x \circ 1 = x.$$

Говорят, что моноид M распознаёт язык $L \subseteq \Sigma^*$, если существует такой морфизм $\varphi : \Sigma^* \to M$ и такое подмножество $S \subseteq M$, что $w \in L \iff \varphi(w) \in S$.

1. Докажите, что регулярные языки – это в точности языки, распознаваемые конечными моноидами.

Определение. Детерминированным двусторонним конечным автоматом (2ДКА) M называют следующую модель вычислений, описываемую набором $\langle \Sigma, Q, \triangleright, \triangleleft, q_0, F\delta \rangle$, где

- Σ — алфавит;
- Q — множество состояний;
- $\triangleright \notin \Sigma$ — левый маркер конца строки;
- $\triangleleft \notin \Sigma$ — правый маркер конца строки;
- $q_0 \in Q$ — начальное состояние;
- q_f – принимающее состояние;
- $\delta : Q \times \{\Sigma \cup \{\triangleright, \triangleleft\}\} \to Q \times \{-1, +1\}$ — функция переходов.

Автомат имеет ленту входа, в ячейках которой написано входное слово, окружённое левым и правым маркерами конца строки: $\triangleright w_1 w_2 \ldots w_n \triangleleft$. В начале работы головка стоит на первом символе w_1 слова w. В зависимости от состояния и символа на ленте, автомат меняет состояние и перемещает головку вправо или влево согласно функции перехода, при этом автомат не перемещает головку за границу маркеров. Слово принимается автоматом, если при обработке слова автомат перешёл в состояние q_f – в этом случае автомат останавливается.

Недетерминированный двустронний конечный автомат получается из недетерминированного как и НКА из ДКА – путём замены функции переходов на отношение переходов.

2. Докажите, что регулярные языки – это в точности языки, распознаваемые двусторонними конечными автоматами.

Определение. Обобщённым регулярным выражением (ОРВ) назовём регулярные выражения, в которых есть операция взятия дополнения $R^c = \Sigma^* \setminus R$. Максимальное количество вложенных операций $*$, используемых в ОРВ, называют его обобщённой звёздочка-высотой (Generalied Star Height). Обобщённой звёздочка-высотой регулярного языка L называют минимальную звёздочка-высоту по ОРВ R, порождающим L.

3. Приведите примеры языков с обобщённой звёздочка-высотой 0 и 1.

Открытые вопросы

4. Существование экспоненциального разрыва между двусторонними автоматами. Существует ли такая последовательность регулярных языков L_n, что некоторый 2НКА, распознающий L_n, имеет n состояний, а любой 2ДКА, распознающий L_n имеет не менее 2^n состояний.

5. Существует ли регулярный язык с обобщённой звёздочка-высотой 2?