О сложности моделирования автоматами регулярных событий

С. С. Морозов (МГУ им. М. В. Ломоносова, Москва)

Изучается свойство автомата содержать во множестве своих выходных слов заданное регулярное множество. Устанавливается, что данное свойство может быть проверено путем изучения строения множества выходных слов автомата ограниченной длины.

Ключевые слова: абстрактный конечный автомат, регулярное событие, регулярное выражение, генератор, алгоритмическая разрешимость.

Определение 1. Назовём конфигурацией автомата любое подмножество множества его состояний.

Пусть дан автомат $V = (A, Q, B, \varphi, \psi)$, $|Q| = n$; выделенная как начальная его конфигурация $K_0 \subset Q, K_0 \neq \varnothing$; регулярное событие $S \subset B^*$, заданное с помощью регулярного выражения R, длина которого равна l, максимальная глубина вложенности итерации в котором равна g, а число вхождений множеств, состоящих из одного однобуквенного слова, равно k.

Определим несколько функций:

$$\varphi(q, a_1 \dots a_{n-1} a_n) = \varphi(\varphi(q, a_1 \dots a_{n-1}), a_n),\ q \in Q, a_i \in B$$

$$\overline{\psi}(q, a_1 \dots a_{n-1} a_n) = \overline{\psi}(q, a_1 \dots a_{n-1}) \psi(\varphi(q, a_1 \dots a_{n-1}), a_n),\ q \in Q, a_i \in B$$

$$\psi(q, \alpha) = [_1(\overline{\psi}(q, \alpha)),\ \alpha = a_1 \dots a_n \in B^*$$

$$\omega(K, \alpha) = \{\varphi(q, \beta) | \overline{\psi}(q, \beta) = \alpha,\ q \in K,\ \beta \in A^*\} \subset Q,\ K \subset Q,\ \alpha \in B^*$$

$$\omega(K, S) = \{\omega(K, \beta) |\ \beta \in S\},\ S \subset B^*,\ K \subset Q$$

$$\omega(K, S, k) = \{\omega(K, \beta) |\ \beta \in S, |\beta| \leqslant k\},\ S \subset B^*,\ K \subset Q,\ k \in \mathbb{N}$$

$$\overline{\psi}(K) = \{\overline{\psi}(q, \alpha)\ |\ q \in K,\ \alpha \in A^*\},\ K \subset Q$$

Определение 2. Будем говорить, что $l_R(n)$ — проверочная длина события S, заданного регулярным выражением R, если для любых V и K_0

выполняется следующее свойство: если множество всех слов длины не большей, чем l_R, из S лежит во множестве $\overline{\psi}(K_0)$, то из этого следует, что и всё S лежит во множестве $\overline{\psi}(K_0)$.

Определение 3. Пусть регулярное событие S, заданное регулярным выражением R, имеет проверочную длину l_R. Будем говорить, что оно l_R-описуемо, если $\forall K \subset Q\ \omega(K,S) = \omega(K,S,l_R)$.

Теорема 1. *$l_R = k(2^n - 1)^g$ является проверочной длиной для S, заданного R.*

В следующих теоремах $K_0 = \{q_0\}$, то есть автоматы инициальные.

Пусть n_i — число состояний автомата, k_i — количество вхождений символов алфавита в регулярное выражение, b_i — количество символов объединения в регулярном выражении, тогда верны следующие 2 теоремы:

Теорема 2. 1. *Пусть множества всех слов длины не большей, чем $l = 2^{2k+2b} + 2^{n+1} - 1$, из множеств, заданного регулярным выражением и генерируемого автоматом, равны. Тогда равны и множества всех слов, порождаемые этими автоматом и регулярным выражением.*
2. *Пусть множества всех слов длины не большей, чем $l = 2^{n_1+1} + 2^{n_2+1} - 1$, из множеств, генерируемых 2 автоматами, равны. Тогда равны и множества всех слов, генерируемые этими двумя автоматами.*
3. *Пусть множества всех слов длины не большей, чем $l = 2^{2k_1+2b_1} + 2^{2k_2+2b_2} - 1$, из множеств, заданных 2 регулярными выражениями, равны. Тогда равны и множества всех слов, задаваемые этими регулярными выражениями.*

Теорема 3. 1. *Пусть для автомата и регулярного выражения выполняется следующее: множество всех слов длины не большей l из множества слов, генерируемого автоматом, лежит во множестве всех слов длины не большей l из множества слов, задаваемого регулярным выражением (либо наоборот), где l равно $2^{n+2k+2g+1}$. Тогда то же вложение выполнено и для множеств вообще всех слов, задаваемых данными автоматом и регулярным выражением.*

2. *Пусть для 2 автоматов выполняется следующее: множество всех слов длины не большей l из множества слов, генерируемого одним автоматом, лежит во множестве всех слов длины не большей l из множества слов, генерируемого другим автоматом, где l равно $2^{n_1+n_2+2}$.*

Тогда то же вложение выполнено и для множеств вообще всех слов, генерируемых данными автоматами.
3. *Пусть для 2 регулярных выражений выполняется следующее: множество всех слов длины не большей l из множества слов, задаваемого одним регулярным выражением, лежит во множестве всех слов длины не большей l из множества слов, задаваемого другим регулярным выражением, где l равно $2^{2k_1+2g_1+2k_2+2g_2}$. Тогда то же вложение выполнено и для множеств вообще всех слов, задаваемых данными регулярными выражениями.*

Обозначим множество слов, которые при всевозможных входных словах может выдать автомат V через $[V]$.

Теорема 4. 1. *Пусть $n_1, n_2 \in \mathbb{N}, n_2 \geqslant 2$, конечные алфавиты B и A удовлетворяют условиям $|B| = p, p \geqslant n_2$ и $|A| = m, m \geqslant p+n_2-2$. Тогда для конечных автоматов вида $V_1 = (A, Q_1, B, \varphi_1, \psi_1)$, где $|Q_1| = n_1$, и вида $V_2 = (A, Q_2, B, \varphi_2, \psi_2)$, где $|Q_2| = n_2$, выполняется $l_{[V_1]}(n_2) \geqslant 2^{n_2} - 1$.*
2. *Пусть $n \in \mathbb{N}$, конечные алфавиты B и A удовлетворяют условиям $|B| = p, p \geqslant n$, и $|A| = m, m \geqslant p+n-2$, натуральные числа k и b — условиям $k \geqslant p, b = k-1$. Тогда для конечных автоматов $V = (A, Q, B, \varphi, \psi)$ таких, что $|Q| = n$, и регулярных множеств S в алфавите B, заданных регулярными выражениями R, количество вхождений символов алфавита B в которые равно k, а символов объединения — b, выполняется $l_R(n) \geqslant 2^n - 1$.*

Список литературы

[1] Кудрявцев В. Б., Алешин С. В., Подколзин А. С. Введение в теорию автоматов. — М.: Наука, 1985.

[2] Подколзин А. С. О сложности распознавания автоматов-генераторов // Дискретный анализ. — Новосибирск: Институт математики СО АН СССР, 1972. — Вып. 21. — С. 31–61.