

АЛГОРИТМ ВЫБОРА СТРУКТУРЫ ДЕРЕВООБРАЗНОГО АВТОМАТА

А. П. Стахов, Н. В. Алипов

Харьков

Ранее нами был описан метод структурного синтеза деревообразных автоматов¹. Настоящая работа посвящена вопросу о построении алгоритма выбора структуры деревообразного автомата (ДА).

Структура ДА определяется входным алфавитом и видом функции $f(s)$, которая зависит от допустимого множества входных слов ДА. Входной алфавит ДА определяет структуру выходного слова системы автоматов. Так, если входной алфавит не содержит символа α , то выходным символом системы автоматов считается состояние автомата Σ . Если входной алфавит содержит символ α , то выходным символом системы автоматов является конъюнкция состояний автоматов Σ и $A_1(m)$, т. е. структура выходного слова системы автоматов зависит только от символа α .

Функция $f(s)$ задает структуру системы автоматов. Так, если $f(s) = \text{const}$, система содержит автоматы $A_k(1)$, Σ , а при $f(s) = \text{var}$ — автоматы $A_1(m)$, $A_k(1)$ и Σ .

Нетрудно видеть, что выбор структуры ДА включает две задачи: определение структуры выходного слова системы автоматов и выбор структуры системы автоматов.

Решение первой задачи не вызывает никаких затруднений, так как она определяется только вхождением символа α во входной алфавит ДА. Решение второй задачи не является тривиальным, поскольку тип функции $f(s)$ невозможно определить из задания ДА. Для определения его разработан алгоритм.

Введем ряд обозначений.

1. Стрелку γ графа M , задающего ДА, назовем $q_i(s)$ -помеченной, если переход ДА, соответствующий стрелке i вызывается символом входного алфавита, весовая функция которого равна $q_i(s)$. При этом предполагается, что стрелки графа M упорядочены.

2. Подмножество $R_i \in R$ будем считать порожденным символом входного алфавита x_i , если оно включает только те слова из множества допустимых слов R , у которых начальным отрезком минимальной длины является символ x_i ($i = 0, 1, 2, \dots, k$).

3. Слово, относящееся к подмножеству R_i , назовем порожденным символом входного алфавита x_i .

Подмножество весовых функций, соответствующих подмножеству входных слов R_i , будем обозначать H_i ($i = 1, 2, \dots, k$).

Докажем истинность следующего утверждения. Для того чтобы никакие два начальных отрезка слов из R и никакие два слова из R при

¹ А. П. Стахов, Н. В. Алипов. Метод структурного синтеза деревообразных автоматов. Сб. «Приборы и системы автоматизации», вып. 12. Изд-во Харьковск. ун-та, 1969.

$f(s) = \text{const}$ не имели равных весовых функций, необходимо соблюдать следующее:

а) весовые коэффициенты символов входного алфавита ДА должны быть связаны соотношениями

$$b_2 = mb_1 + 1,$$

$$b_3 = mb_2 + 1,$$

$$b_k = mb_{k-1} + 1;$$

б) весовые функции входных слов должны иметь вид

$$Q_m = b_{i_1} + b_{j_2} + b_{q_3} + \dots + b_{l_m}$$

при условии, что

$$b_{i_1} > b_{j_2} > \dots > b_{l_m},$$

где $b_{i_1}, b_{j_2}, \dots, b_{l_m} \in \{b_1, b_2, \dots, b_k\}$.

Согласно VIII (см. названную работу) весовая функция входного слова не зависит от символа α , поэтому докажем данное утверждение для случая, когда входной алфавит не содержит символа α .

Доказательство необходимости. Предположим, что никакие два начальных отрезка слов из R и никакие два слова из R при $f(s) = \text{const}$ не имеют равных весовых функций. Принятое предположение равносильно тому, что существует взаимнооднозначное соответствие между множеством весовых функций H и множеством входных слов R . Последнее означает, что существует взаимнооднозначное соответствие между R_i и H_i и для любых i, j при $i \neq j$ выполняется условие

$$H_j \cap H_i = \emptyset \quad (i, j = 1, 2, \dots, k),$$

Пусть условие а) не выполняется, тогда максимальное значение весовой функции слова, порожденного символом x_1 , составляет mb_1 , а минимальное значение весовой функции слова, порожденного символом x_k , равно b_2 . Очевидно, при некотором m возможно неравенство

$$b_2 < mb_1, \quad (1)$$

т. е. $H_1 \cap H_2 \neq \emptyset$, что противоречит принятому предположению.

Пусть условие б) невыполнимо, тогда среди каких-либо подмножеств H_j и H_q всегда найдутся весовые функции вида

$$b_i + b_n; b_n + b_i. \quad (2)$$

Наличие весовых функций вида (2) приводит к тому, что

$$H_i \cap H_q \neq \emptyset, \quad (3)$$

а это противоречит принятому предположению.

Пусть выполняются условия а) и б). Тогда в силу условия а) для максимальной весовой функции порожденного символом x_i слова, равной mb_i , и минимальной весовой функции порожденного символом x_j слова, равной b_j , при $i < j$ выполняется условие

$$mb_i < b_j. \quad (4)$$

Справедливость неравенства (4) легко доказать методом индукции.

Действительно, пусть $j=i+1$, тогда согласно условию а) имеем $b_j = mb_i + 1$.

Поскольку $mb_i + 1 > mb_i$ то истинность (4) очевидна.

Пусть (4) справедливо для $j = i + n$, докажем, что (4) справедливо и для $j = i + n + 1$.

Согласно а) имеем $b_{i+n+1} = mb_{i+n} + 1$. Так как $b_{i+n} > mb_i$, то $mb_{i+n} + 1$ тем более больше mb_i , что и доказывает истинность (4).

В силу условия б) никакие два подмножества H_i и H_j не содержат одновременно слов вида (2). При истинности (4) это означает, что для любых подмножеств H_i и H_j при $i < j$ выполняется условие

$$H_i \cap H_j = 0. \quad (5)$$

Выполнение равенства (5) для любых i, j доказывает необходимость условий а) и б). Утверждение доказано.

Из утверждения вытекают следствия:

Следствие 1. $f(s) = \text{const}$, если среди вершин графа М, задающего ДА, нет ни одной такой вершины a_p из которой выходила хотя бы одна стрелка, помеченная большим числом, чем максимально помеченная стрелка, входящая в вершину a_p .

Следствие 2. Если каждый символ входного алфавита ДА вызывает единственный переход ДА, то $f(s) = \text{const}$.

Утверждение, следствия 1, 2 дают возможность сформулировать алгоритм определения типа функции $f(s)$ (структуры системы автоматов) для следующих случаев.

а) Входной алфавит ДА не содержит символа α .

1. Задать деревообразный автомат графом М, перейти на 2.

2. Приписать символам входного алфавита весовые коэффициенты согласно утверждению 4, перейти на 3.

3. Пометить каждую стрелку графа, перейти на 4.

4. Выяснить, удовлетворяет ли каждая вершина графа М следствию 1 или 2. Если среди вершин графа М нет ни одной такой, которая не удовлетворяла бы следствию 1, то $f(s) = \text{const}$ (система автоматов содержит автоматы $A_k(1), \Sigma$). Если среди вершин графа М есть хотя бы одна вершина, которая не удовлетворяет следствию 1, но удовлетворяет следствию 2, то $f(s) = \text{const}$ (система автоматов содержит автоматы $A_k(1), \Sigma$), иначе $f(s) = \text{var}$ (система автоматов содержит автоматы $A_1(m), A_k(1)$, и Σ).

б) Входной алфавит ДА содержит символ α .

Согласно VIII (см. указанную работу) граф автомата Σ получается из графа ДА отождествлением двух вершин a_j, a_p графа ДА, если переход ДА из состояния a_j в состояние a_p вызывается символом α . Граф автомата Σ назовем нормированным. Приведение графа ДА к графу автомата Σ назовем операцией нормирования. Введенные определения, утверждение следствия 1, 2 дают возможность сформулировать алгоритм для этого случая.

1. Задать деревообразный автомат графом М, перейти на 2.

2. Пронормировать граф ДА, перейти на 3.

3. Приписать символам входного алфавита соответствующий вес согласно утверждению, перейти на 4.

4. Пометить стрелки пронормированного графа ДА, перейти на 5.

5. Выяснить, удовлетворяет ли каждая вершина графа М следствию 1 или 2. Если среди вершин графа М нет ни одной такой, которая не удовлетворяла бы следствию 1, то $f(s) = \text{const}$ (система автоматов содержит автоматы $\Sigma, A_k(1)$). Если среди вершин графа М есть

хотя бы одна вершина, которая не удовлетворяет следствию 1, но удовлетворяет следствию 2, то $f(s) = \text{const}$ (система автоматов содержит автоматы $\Sigma, A_k(1)$), иначе $f(s) = \text{var}$ (система автоматов содержит автоматы $A_1(m), A_k(1), \Sigma$).

Легко видеть, что для случая а) алгоритм определения структуры системы автоматов совпадает с алгоритмом выбора структуры ДА, так как в этом случае структура системы автоматов совпадает с структурой ДА.

Для случая б) алгоритм определения структуры системы автоматов отличается от алгоритма выбора структуры ДА только пунктом 5. Отличие заключается в том, что при $f(s) = \text{const}$ для получения структуры ДА следует к структуре системы автоматов добавить автомат $A_1(m)$ с целью организации выходных символов.

