Доклады Академии наук СССР
1970. Том 190, № 5

КИБЕРНЕТИКА И ТЕОРИЯ РЕГУЛИРОВАНИЯ

Я. М. БАРЗДИНЬ

О РАСШИФРОВКЕ АВТОМАТОВ ПРИ ОТСУТСТВИИ ВЕРХНЕЙ ОЦЕНКИ ЧИСЛА СОСТОЯНИЙ

(Представлено академиком В. М. Глушковым 22 V 1969)

Под автоматами будем понимать конечные инициальные автоматы Мили с нумерованными состояниями $q_1, q_2, q_3, \ldots$ Состояние q_1, если не оговорено противное, будем считать начальным. Далее будем считать, что у всех автоматов один и тот же входной алфавит $X = \{x_1, \ldots, x_a\}$, где $a = \mathrm{const} \geqslant 2$, и один и тот же выходной алфавит $Y = \{y_1, \ldots, y_b\}$, где $b = \mathrm{const} \geqslant 2$. Оператор, который реализует автомат $\mathfrak{M}$ при начальном состоянии q_i, обозначим через $T(\mathfrak{M}, q_i)$.

Пусть $\mathfrak{M}$ — автомат, о внутренней структуре (диаграмме состояний) которого ничего не известно, в том числе и о верхней оценке числа состояний (или известно, но при данной постановке не разрешается пользоваться). Такой автомат будем называть черным ящиком (ч.я.). Предполагается, что с ч.я. можно проводить эксперименты. Как известно, невозможен эксперимент, с помощью которого можно было расшифровать любой ч.я. Однако, как показано в работе [1], существует кратный эксперимент (разумеется, разветвленный), который позволяет расшифровать «большинство» ч.я. Цель данной заметки — получить аналогичный результат в случае простых экспериментов и оценить длину соответствующих простых экспериментов.

Точная постановка проблемы расшифровки связана с уточнением понятия алгоритма над ч.я. (алгоритма расшифровки). Содержательно под алгоритмом Ω над ч.я. будем понимать эффективное предписание, описывающее простой разветвленный эксперимент и указывающее, как по результату этого эксперимента строить соответствующий автомат (который предположительно работает так же, как заданный ч.я.). Точнее, алгоритм Ω, примененный к ч.я. $\mathfrak{M}$, работает по шагам. На каждом шаге алгоритм Ω проделывает с ч.я. $\mathfrak{M}$, находящемся в том состоянии, в котором он остался после предыдущего шага, некоторый простой однородный эксперимент, и в зависимости от результата этого эксперимента, а также результата экспериментов, проделанных на предыдущих шагах, делает одно из двух: а) или выдает результат $\Omega(\mathfrak{M})$ — некоторый автомат (например, в виде диаграммы); б) или строит входное слово, которое определяет простой эксперимент, проводимый на следующем шаге. Состояние автомата $\mathfrak{M}$, в которое он переходит в результате применения алгоритма Ω, обозначим через q_Ω. Будем говорить, что алгоритм Ω расшифровывает ч.я. $\mathfrak{M}$, если $T(\Omega(\mathfrak{M}), q_1) = T(\mathfrak{M}, q_\Omega)$, т. е. если алгоритм Ω однозначно определяет дальнейшее поведение ч.я. $\mathfrak{M}$. Длину простого эксперимента, который проделывает алгоритм Ω, примененный к $\mathfrak{M}$, обозначим через $\Omega'(\mathfrak{M})$. Положим $\Omega^*(k) = \max \Omega'(\mathfrak{M})$, где max берется по всем автоматам $\mathfrak{M}$, имеющим k состояний.

Пусть $\mathscr{L}$ — класс всех попарно неодинаковых* автоматов (ч.я.) $\{\mathscr{L}_\lambda\}$ — некоторое разбиение класса $\mathscr{L}$ на конечные подклассы и $\mathscr{L}_\lambda{}^\Omega$ —

* Два автомата $\mathfrak{M}_1$ и $\mathfrak{M}_2$ мы считаем одинаковыми тогда и только тогда, когда любые два состояния с одинаковыми номерами как в диаграмме $\mathfrak{M}_1$, так и в диаграмме $\mathfrak{M}_2$ связаны между собой ребрами, одинаково ориентированными и с одинаковыми пометками.

множество тех автоматов из $\mathcal{L}_\lambda$, которые алгоритм Ω расшифровывает. Рассмотрим отношение $P_\lambda = |\mathcal{L}_\lambda^\Omega| / |\mathcal{L}_\lambda|$. Будем говорить, что алгоритм Ω расшифровывает автоматы (ч.я.) с частотой $1 - \varepsilon$ при данном разбиении $\{\mathcal{L}_\lambda\}$, если $P_\lambda \geqslant 1 - \varepsilon$ для всех λ. В данной заметке мы рассмотрим разбиение, при котором два автомата относятся к одному и тому же классу, если они отличаются только функциями выхода (такое разбиение будем называть разбиением по графу). Это разбиение является достаточно мелким. Поэтому, если мы покажем, что ч.я. можно расшифровывать с частотой $1 - \varepsilon$ при данном разбиении, то тем более это будет верно и при более крупных разбиениях, например таких, как разбиение по числу состояний, когда любые два автомата относятся к одному подклассу, если они имеют одинаковое число состояний. Впредь, говоря, что алгоритм Ω расшифровывает ч.я. равномерно с частотой $1 - \varepsilon$, мы будем иметь в виду именно разбиение по графу. Под автоматным графом будем понимать граф, который получается из диаграммы автомата, если стереть выходные пометки (а входные оставить). Очевидно, из автоматного графа G с k вершинами, расставляя всевозможными способами выходные пометки, можно получить b^{ak} попарно неодинаковых автоматов. Класс таких автоматов обозначим через $\bar{G}$. Таким образом, если алгоритм Ω расшифровывает ч.я. равномерно с частотой $1 - \varepsilon$, то это означает, что для любого автоматного графа G имеет место $|\bar{G}^\Omega| / |\bar{G}| \geqslant 1 - \varepsilon$. Основным результатом данной заметки является

Теорема 1*. *Для любого $\varepsilon > 0$ существует алгоритм Ω, который расшифровывает ч.я. равномерно с частотой $1 - \varepsilon$ и имеет $\Omega^*(k) \leqslant C_\varepsilon k^C$, где C — константа, не зависящая от k и ε, C_ε — константа, не зависящая от k, но зависящая от ε.*

Будем говорить, что входное слово x остаточно различает автоматы $\mathfrak{M}_1$ и $\mathfrak{M}_2$, если справедливо следующее: или $\mathfrak{M}_1$ и $\mathfrak{M}_2$ при подаче слова x выдают различные выходные слова, или $\mathfrak{M}_1$ и $\mathfrak{M}_2$ после подачи слов x реализуют одинаковые операторы (т. е. $T(\mathfrak{M}_1, q_1 x) = T(\mathfrak{M}_2, q_1 x)$). Под суммарной степенью различимости автоматов $\mathfrak{M}_1$ и $\mathfrak{M}_2$ будем понимать различимости автомата $\mathfrak{M}_1 + \mathfrak{M}_2$, который получается из $\mathfrak{M}_1$ и $\mathfrak{M}_2$, если их рассматривать как один автомат.

Лемма 1**. *Для любого множества U автоматов и любого натурального s существует входное слово длины не более $[4a^{s+1} \ln |U|]$, которое остаточно различает любые два автомата из U, имеющие суммарную степень различимости не более s.*

Докажем лемму. Пусть $\mathfrak{M}_1$ и $\mathfrak{M}_2$ — автоматы, имеющие суммарную степень различимости не более s. Сначала оценим снизу число входных слов длины $2s$, которые остаточно различают $\mathfrak{M}_1$ и $\mathfrak{M}_2$. Очевидно, для любого входного слова v длины $l(v)$ существует входное слово длины $l(v) + s$, начинающееся с v и остаточно различающее $\mathfrak{M}_1$ и $\mathfrak{M}_2$: обозначим его через $(v)\langle s\rangle$ (если таких несколько, то одно из них). Заметим, что если $l(v) \leqslant s$, то число всевозможных продолжений слова $(v)\langle s\rangle$ до длины $2s$ равно $a^{s-l(v)}$, и все эти продолжения будут остаточно различать $\mathfrak{M}_1$ и $\mathfrak{M}_2$.

Опишем теперь одну процедуру выделения слов, остаточно различающих $\mathfrak{M}_1$ и $\mathfrak{M}_2$.

---

* Аналогичное утверждение при более крупном разбиении (по числу состояний) было доказано автором данной заметки совместно с М. П. Василевским.

** Из этой леммы непосредственно вытекают теоремы 1, 2 и 3 работы (2). Для этого достаточно учесть только следующее: а) если U_k — класс всех попарно неодинаковых автоматов с k состояниями, то $|U_k| = (bk)^{ak}$; б) суммарная степень различимости двух автоматов из U_k не больше $2k - 1$, т. е. не больше степени восстановления (для теорем 1 и 2 из (2)); в) суммарная степень различимости двух автоматов из U_k, отличающихся только фиксацией начального состояния, не больше $k - 1$, т. е. не больше обычной степени различимости (для теоремы 3 из (2)). Еще из этой леммы и оценки степени восстановления для почти всех автоматов (3) вытекает следующий факт: почти все автоматы с k состояниями можно расшифровать (в упомянутом выше смысле) простым неразветвленным экспериментом длины k^C, где C — константа.

Шаг 1. Рассмотрим входные слова $x = x_i$ длины 1; число таких слов равно a. Соответственно каждому из них выделим слово $(x_i)\langle s\rangle$. Очевидно, число всевозможных продолжений всех выделенных слов вида $(x_i)\langle s\rangle$ до длины $2s$ равно aa^{s-1}.

Шаг k $(k \leqslant s)$. Рассмотрим входные слова $x = x_{i_1} \ldots x_{ik}$ длины k, отличные от начальных кусков длины k ранее выделенных слов; число таких слов равно $a^k - a^{k-1}$, так как число ранее выделенных слов равно a^{k-1} и все они имеют различные начальные куски длины k (даже длины $k-1$). Соответственно каждому из них выделим слово $(x_{i_1} \ldots x_{i_k})\langle s\rangle$ (таким образом, число выделенных слов вида $(x_{i_1} \ldots x_{i_k})\langle s\rangle$ будет равно $a^k - a^{k-1}$). Очевидно, число всевозможных продолжений всех выделенных слов вида $(x_{i_1} \ldots x_{i_k})\langle s\rangle$ до длины $2s$ равно $(a^k - a^{k-1})a^{s-k}$.

В результате получаем, что общее число входных слов длины $2s$, являющихся продолжением слов, выделенных в течение первых s шагов, равно $aa^{s-1} + \ldots + (a^k - a^{k-1})a^{s-k} + \ldots + (a^s - a^{s-1})a^0 \geqslant sa^{s-1}$ (напомним, что $a \geqslant 2$). Согласно сказанному выше, все эти слова будут остаточно различать $\mathfrak{M}_1$ и $\mathfrak{M}_2$. Таким образом, общее число входных слов длины $2s$, остаточно не различающих $\mathfrak{M}_1$ и $\mathfrak{M}_2$, не превосходит $a^{2s} - sa^{s-1}$. Очевидно, этот результат остается справедливым и при любом другом выборе начальных состояний автоматов $\mathfrak{M}_1$ и $\mathfrak{M}_2$. Поэтому индукцией по p получаем, что число входных слов длины $p \cdot 2s$, остаточно не различающих $\mathfrak{M}_1$ и $\mathfrak{M}_2$, не превосходит $(a^{2s} - sa^{s-1})^p$.

Пусть теперь U — произвольное множество автоматов. Тогда число входных слов длины $2ps$, остаточно не различающих хотя бы два автомата из U, имеющих суммарную степень различимости не более s, не превосходит $C_{|U|}^2 (a^{2s} - sa^{s-1})^p < {}^1/_2 |U|^2 (a^{2s} - sa^{s-1})^p$. С другой стороны, общее число входных слов длины $2ps$ равно a^{2ps}. Поэтому, если при некотором p_0 ${}^1/_2 |U|^2 (a^{2s} - sa^{s-1})^{p_0} \leqslant a^{2p_0 s}$, то среди входных слов длины $2p_0 s$ обязательно найдется слово, остаточно различающее любые два автомата из U, имеющие суммарную степень различимости не более s. Выражая из последнего неравенства p_0 и учитывая, что $|\ln(1 - sa^{-(s+1)})| > sa^{-(s+1)}$, получаем, что в качестве p_0 можно брать, например, $[2s^{-1}a^{s+1} \ln |U|]$. Отсюда вытекает лемма 1.

Пусть $D_G(x)$ — множество вершин графа G (автомата G), которые достижимы из вершины q_1 словом x, и $A_G(x)$ — множество вершин графа G (автомата G), которые достижимы из вершины q_1 словом x или из вершины $q_1 x$ произвольными словами. Пользуясь такими же соображениями, как и выше, можно доказать следующую лемму.

Лемма 2. *Для любого натурального k существует входное слово $b(k)$ длины не более $[2ka^{k+1} \ln 2k]$, обладающее следующим свойством: какой бы автоматный граф G мы не брали, если $|A_G(b(k))| \geqslant k$, то $|D_G(b(k))| \geqslant k$.*

Пусть q_i, q_j — вершины графа G и r — натуральное число. Обозначим через $\widetilde{G}(q_i, q_j, r)$ множество всех тех автоматов из $\widetilde{G}$, у которых состояния q_i и q_j не различимы входными словами длины r, но различимы входными словами бóльшей длины.

Лемма 3*. *Для любого автоматного графа G, любых его вершин q_i, q_j и любого натурального r выполняется неравенство $|\widetilde{G}(q_i, q_j, r)| / |\widetilde{G}| < b^{-r/2}$.*

---

* Из этой леммы, в частности, вытекает следующий факт. Скажем, что равномерно почти все автоматы имеют степень различимости не бóльшую $\varphi(k)$, если $\min\limits_{|G|=k} \frac{|\widetilde{G}_{\varphi(k)}|}{|\widetilde{G}|} \to 1$ при $k \to \infty$, где $|G|$ — число вершин графа G и $\widetilde{G}_{\varphi(k)}$ — множество всех тех автоматов из $\widetilde{G}$, которые имеют степень различимости не бóльшую $\varphi(k)$. Тогда справедливо следующее: равномерно почти все автоматы имеют степень различимости не бóльшую $C \log k$. (Оценка степени различимости просто для почти всех автоматов дана в [3].)

Теперь в общих чертах дадим идею доказательства теоремы 1. Соответствующий простой эксперимент строится по шагам. На каждом шаге, исходя из некоторого числа s, содержательно означающего очередную гипотезу о числе состояний ч.я. $\mathfrak{M}$, и используя леммы 1, 2 и 3, строится эксперимент ζ_s такой, что для «большинства» автоматов $\mathfrak{M}$ справедливо следующее: а) если $\mathfrak{M}$ имеет не больше чем s состояний (точнее, $A_{\mathfrak{M}}(\zeta_s) \leqslant s$), то ζ_s расшифровывает $\mathfrak{M}$; б) если $\mathfrak{M}$ имеет больше чем s состояний (точнее, $A_{\mathfrak{M}}(\zeta_s) > s$), то ζ_s достигает достаточно много состояний автомата $\mathfrak{M}$. В последнем случае строится новая гипотеза $s' > s$, и процедура построения эксперимента продолжается.

В заключение рассмотрим еще одну задачу. Под стратегией будем понимать функцию вида $\Phi(x, y, x_i)$, где x — слово в входном алфавите X; y — слово в выходном алфавите Y; x_i — буква алфавита X; значение функции $\Phi(x, y, x_i)$ — буква алфавита Y. Содержательно стратегию будем интерпретировать как правило «предсказания» той буквы, которую выдаст автомат при подаче буквы x_i, если известно, что до этого было подано входное слово x, которое автомат переработал в выходное слово y. Пусть $\mathfrak{M}$ — автомат, $\omega = (x(1), x(2), \ldots, x(i), \ldots)$ — бесконечная входная последовательность и $(y(1), y(2), \ldots, y(i), \ldots)$ — соответствующая выходная последовательность (которую выдает $\mathfrak{M}$ при подаче ω). Скажем, что стратегия Φ на $\mathfrak{M}$ и ω допускает ошибку в i-й момент, если $\Phi(x(1) \ldots x(i), y(1) \ldots y(i), x(i+1)) \neq y(i+1)$. Число моментов, в которых стратегия Φ допускает ошибку, обозначим через $\Phi_\omega'(\mathfrak{M})$. Положим $\Phi_\omega^*(k) = \max \Phi_\omega'(\mathfrak{M})$, где max берется по всем автоматам $\mathfrak{M}$, имеющим k состояний.

Теорема 2. *Существует эффективная стратегия Φ такая, что для любой бесконечной входной последовательности ω имеет место $\Phi_\omega^*(k) \leqslant Ck \log_2 k$, где C — константа, не зависящая от k* и ω.

Эта теорема сохраняет силу, если вместо конечных автоматов рассматривать машины Тьюринга с входными и выходными каналами, употребляющими соответственно алфавиты X и Y. При этом предполагается, что машины начинают работать с пустой ленты и что внешний алфавит у всех машин один и тот же.

В случае периодических последовательностей ω (с периодом p), как легко убедиться, $\Phi_\omega^*(k) \leqslant C_p k$. Однако в общем случае, как показывает следующая теорема, оценка из теоремы 2 по порядку не может быть понижена.

Теорема 3. *Существуют бесконечная входная последовательность ω и константа C_0 такие, что для любой стратегии Φ имеет место $\Phi_\omega^*(k) \geqslant C_0 k \log_2 k$.*

Вычислительный центр
Латвийского государственного университета им. П. Стучки
Рига

Поступило
19 V 1969

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] Я. М. Барздинь, Проблемы кибернетики, в. 21, М., 1969. [2] А. А. Мучник, Проблемы кибернетики, в. 20, М., 1968. [3] А. Д. Коршунов, Дискретный анализ, в. 10, Новосибирск, 1967, стр. 39.