Автоматные модели защищенных компьютерных систем*

А. В. Галатенко

В работе рассматривается автоматная система, часть состояний которой объявляется безопасной. Исследуются языки, не выводящие автомат из класса безопасных состояний или выводящие из класса безопасных состояний не более, чем на ε.

1. Основные понятия и результаты

Под конечным автоматом мы будем понимать четверку $V = (A, Q, \varphi, q_0)$, где A — конечное множество входных символов, Q — конечное множество состояний, $\varphi : A \times Q \to Q$ — функция переходов, $q_0 \in Q$ — начальное состояние. Пусть $Q = S \cup I$, причем $S \cap I = \varnothing$. Состояния из S назовем безопасными, состояния из I — небезопасными. Далее будем предполагать, что начальное состояние является безопасным.

Обозначим через A^* множество всех конечных слов в алфавите A. Функция φ может быть продолжена на множество $A^* \times Q$ по мультипликативности. Подмножество A^* называется языком. Каждому слову $\alpha \in A^*$ соответствует слово $\varkappa(\alpha) \in Q^*$, $\varkappa(\alpha) = \varphi(\alpha, q_0)$. Назовем слово $\alpha \in A^*$ безопасным, если $\varkappa(\alpha) \in S^*$. Назовем язык $\mathcal{A} \subseteq A^*$ безопасным (S-языком), если все слова, составляющие $\mathcal{A}$, безопасны, и не существует безопасных слов, не принадлежащих $\mathcal{A}$.

Сформулируем ряд утверждений о свойствах S-языков.

Лемма 1. *Если $\mathcal{A}$ является S-языком, то $\mathcal{A}$ регулярен.*

*Работа выполнена при финансовой поддержке РФФИ (грант 06–01–00240).

Пусть $\mathcal{A}$ — некоторый язык. Обозначим через $[\mathcal{A}]$ множество всех начал слов, принадлежащих $\mathcal{A}$.

Лемма 2. *Язык $\mathcal{A}$ является S-языком для некоторого автомата V тогда и только тогда, когда выполнены следующие условия:*

1) $\mathcal{A}$ *непуст;*

2) $\mathcal{A}$ *регулярен;*

3) $[\mathcal{A}] \subseteq \mathcal{A}$.

Лемма 3. *Автоматно перечислимые языки (см. [3]) с добавленным пустым словом являются S-языками, но, вообще говоря, не наоборот.*

Содержательно безопасный язык может интерпретироваться несколькими способами:

- в системах активного аудита (см. например [1]), как описания легального поведения пользователей (в этом случае выход из множества безопасных состояний трактуется как атака);
- при задании моделей злоумышленника (в этом случае выход из множества безопасных состояний может трактоваться как обнаружение злоумышленника);
- при построении моделей гарантированно защищенных систем (см. например [5], [2]; в таких моделях доказательства как правило проводятся индуктивно, а лемма 2 означает, что S-языки как раз описывают все регулярные языки, для которых проходит индуктивное доказательство).

Оценим число S-языков, генерируемых одним автоматом. Будем предполагать, что все состояния автомата достижимы из q_0. Обозначим через NS_V число S-языков, которые можно задать в автомате V.

Лемма 4. *Для конечного автомата V с n состояниями имеет место неравенство $NS_V \geqslant n$, и эта оценка неулучшаема.*

Определим функцию Шеннона $LS(n)$ числа S-языков, задаваемых автоматами с n состояниями, следующим образом. $LS(n) = \max\limits_{\{V:|Q|=n\}}(NS_V)$.

Теорема 1. *Пусть V — автомат с n состояниями и входным алфавитом из l элементов. Тогда*

$$2^{\frac{l-1}{l^2}n} \leqslant LS(n) \leqslant 2^{n-1}.$$

Если $n = \frac{l^k-1}{l-1}$ для некоторого $k \in \mathbb{N}$, то

$$2^{\frac{l}{l+1}n} \leqslant LS(n) \leqslant 2^{n-1}.$$

Рассмотрим произвольное $\varepsilon > 0$. Введем функции $s : Q^* \to \mathbb{N} \cup \{0\}$ и $i : Q^* \to \mathbb{N} \cup \{0\}$ следующим образом. Пусть $\varkappa \in Q^*$. Функция $s(\varkappa)$ равняется числу букв $\varkappa$, содержащихся в S, $i(\varkappa)$ равняется числу букв $\varkappa$, содержащихся в I. Обозначим через $|\varkappa|$ число букв в слове $\varkappa$. Назовем слово $\varkappa$ ε-безопасным, если $\frac{i(\varkappa)}{|\varkappa|} < \varepsilon$. Назовем язык $\mathcal{A}$ ε-безопасным (S_ε-языком), если все слова из $\mathcal{A}$ ε-безопасны, и не существует ε-безопасных слов, не принадлежащих $\mathcal{A}$.

S_ε-языки являются естественным обобщением S-языков. Содержательно рассмотрение возможности выхода из множества безопасных состояний может интерпретироваться как признание ущерба, не превышающего некоторого порога, допустимым. Отметим, что свойства S_ε-языков существенно отличаются от свойств S-языков.

Теорема 2. *Если $\varepsilon \in \mathbb{Q}$, $0 < \varepsilon < 1$, то существует конечный автомат V такой, что определяемый им S_ε-язык является контекстно-свободным, но не регулярным, и существуют конечные языки, не являющиеся S_ε-языками ни при каких $\varepsilon \in \mathbb{Q}$.*

Замечание. При $\varepsilon = 0$ мы остаемся в множестве S-языков. $S_1 = A^*$.

Теорема 3. *Если $\varepsilon_1, \varepsilon_2 \in \mathbb{Q}$, $0 \leqslant \varepsilon_1 < \varepsilon_2 \leqslant 1$, то существует язык, являющийся S_{ε_1}-языком, но не являющийся S_{ε_2}-языком. Если ε_2 не равно 1, то существует язык, являющийся S_{ε_2}-языком, но не являющийся S_{ε_1}-языком.*

Автор выражает глубокую благодарность своему научному руководителю, д.ф.-м.н., проф. В. Б. Кудрявцеву за постановки задач и внимание к работе.

2. Доказательства утверждений

2.1. Доказательство леммы 1

Пусть $V = (A, Q = S \cup I, \varphi, q_0)$ — автомат, задающий S-язык $\mathcal{A}$. Рассмотрим автомат $V' = (A, S \cup q', \varphi', q_0)$, где q' не принадлежит S, $\varphi'(a, q) = \varphi(a, q)$, если $q \in S$ и $\varphi(a, q) \in S$, и $\varphi'(a, q) = q'$ в противном случае. Рассмотрим язык L, задаваемый V' при множестве финальных состояний, равном S. По теореме Клини ([3]) L регулярен. Покажем, что $L = \mathcal{A}$. Если слово α принадлежит L, то при подаче V' на вход α автомат не попадет в состояние q' (так как оно является поглощающим), следовательно переходы будут только в состояния из S, и по определению функции φ' при подаче α на вход V переходы также будут только в S-состояния, то есть $\alpha \in \mathcal{A}$. Обратно, если $\alpha \in \mathcal{A}$, то по определению функции φ' $q_\alpha = \varphi'(\alpha, q) = \varphi(\alpha, q) \in S$, то есть $\alpha \in L$. Лемма доказана.

2.2. Доказательство леммы 2

Пусть $\mathcal{A}$ является $S-$языком для некоторого автомата $V = (A, Q = S \cup I, \varphi, q_0)$. Необходимость условия 1 следует из того, что $q_0 \in S$, следовательно пустое слово $\varepsilon \in \mathcal{A}$. Необходимость условия 2 следует из леммы 1. Докажем необходимость условия 3. Предположим противное. Пусть существует слово $\alpha \in \mathcal{A}$, причем слово α', являющееся префиксом α ($\alpha = \alpha'|\alpha''$), не принадлежит $\mathcal{A}$. Это значит, что $\varphi(\alpha', q_0)$ не принадлежит S. Но $\varphi(\alpha, q_0) = \varphi(\alpha'', \varphi(\alpha', q_0)$, то есть при поступлении V на вход слова α автомат выходит из множества безопасных состояний. Противоречие.

Пусть L — язык, удовлетворяющий свойствам 1–3. Покажем, что L является S-языком для некоторого автомата. Так как L регулярен, существует автомат $V = (A, Q, \varphi, q_0)$, распознающий L множеством финальных состояний Q_F. Из условий 1 и 3 следует, что $q_0 \in Q_F$. Действительно, если L состоит только из пустого слова, $q_0 \in Q_F$. Если же существует $\alpha \in L$, то пустой префикс α также принадлежит L, и $q_0 \in Q_F$. Покажем, что если $\alpha \in L$, то при подаче α на вход V переходы осуществляются только по состояниям из Q_F. Действительно, если на каком-то префиксе α автомат пришел в состояние не

из Q_F, то такой префикс окажется непринадлежащим языку, и условие 3 нарушится. Из доказанного следует, что L является S-языком для автомата V с $S = Q_F$, $I = Q \setminus Q_F$. Лемма доказана.

2.3. Доказательство леммы 3

По теореме о характеризации автоматно перечислимых языков ([3]), язык L является автоматно перечислимым тогда и только тогда, когда выполнены следующие условия:

1) $\mathcal{L}$ непуст;
2) $\mathcal{L}$ регулярен;
3) $[\mathcal{L}] \subseteq \mathcal{L}$ для всех непустых префиксов;
4) Каждое слово L является началом некоторого другого слова из L.

Следовательно, условия 1 и 2 леммы 2 выполнены. Условие 3 леммы 2 следует из условия 3 автоматной перечислимости (для непустых префиксов) и добавления пустого слова (для пустого префикса).

Рассмотрим автомат, у которого не происходит возврат в начальное состояние ни по какому непустому слову (например, автомат с диаграммой на рис. 1). Пусть множество S состоит только из начального состояния. В таком случае S-язык состоит только из пустого слова и не является автоматно перечислимым (например, в силу нарушения условия 4). Лемма доказана.

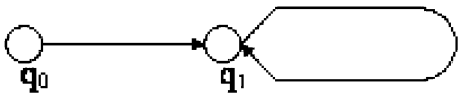

Рис. 1. Пример автомата, задающего S-язык только из пустого слова.

2.4. Доказательство леммы 4

Рассмотрим автомат $V = (A, Q, \varphi, q_0)$. Пусть $|Q| = n$, и все состояния достижимы из q_0. Пусть $c_1, c_2, \ldots, c_{n-1}$ — цепи минимальной длины на диаграмме Мура, ведущие из начального состояния в

состояния, отличные от начального. Такие цепи существуют в силу условия достижимости. Все цепи попарно различны, так как отличаются конечные вершины. Поставим каждой цепи c_i в соответствие Q_i — множество состояний, входящих в цепь. В силу минимальности цепей все Q_i попарно различны. Действительно, в каждой цепи все вершины попарно различны (иначе возник бы цикл, что противоречит определению цепи). Пусть $Q_i = Q_j$; рассмотрим цепь c'_i, являющейся собственной подцепью c_j от q_0 до q_i. Так как все вершины в цепи попарно различны, длина c_j равна длине c_i и строго меньше длины c'_i, что противоречит минимальности цепей. Рассмотрим языки, порождаемые множествами безопасных состояний, равными Q_i. Все такие языки непусты (так как содержат слова, соответствующее цепям c_i, по которым построены множества Q_i) и попарно различны (так как все Q_i попарно различны, и каждое состояние достижимо по состояниям Q_i). Таких языков $n-1$. Добавим язык, порождаемый состоянием q_0. Он отличается от языков, порожденных Q_i, так как не содержит ни одного слова, переводящего автомат в другое состояние. Следовательно, $NS_V \geqslant n$.

Рассмотрим автомат $V' = (A, Q', \varphi', {q_0}')$, диаграмма Мура которого имеет вид цепи с петлей в последнем состоянии цепи (рис. 2). Легко увидеть, что $NS'_V = n$. Действительно, для любого множества $S \subseteq Q'$ достижимы по состояниям из S только те состояния, которые не отделены от ${q_0}'$ состоянием из $Q' \setminus S$, то есть различных S-языков столько же, сколько цепей от начального состояния, то есть n. Лемма доказана.

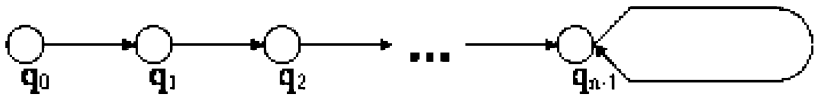

Рис. 2. Автомат с $NS_V = n$.

2.5. Вспомогательные утверждения

В данном разделе будут рассмотрены вспомогательные утверждения, необходимые для доказательства теорем.

Рассмотрим автомат $V = (A, Q, \varphi, q_0)$, $S \subseteq Q$. Будем говорить, что два состояния $q_1, q_2 \in S$ являются S-связными ($q_1 \rho_S q_2$), если в диаграмме Мура автомата V существует путь из q_1 в q_2 такой, что все вершины этого пути лежат в S. Назовем множеством S-связности автомата V все состояния $q \in S$ такие, что $q_0 \rho q$.

Лемма 5 (о связности). *Пусть $V_1 = (A, Q, \varphi_1, q_0^1)$ и $V_1 = (A, Q, \varphi_2, q_0^2)$ — два автомата, S_1 и S_2 — множества безопасных состояний, причем множества S_1 и S_2-связности совпадают. Тогда совпадают S_1 и S_2 языки.*

Доказательство. Пусть $q \in S_1$ не принадлежит множеству связности S_1. Тогда ни одно слово S_1 языка не переводит автомат V в состояние q. Действительно, предположим обратное. Пусть существует слово α, входящее в S_1-язык. Тогда, по определению безопасного языка, при подаче α на вход V все переходы осуществляются в состояния из S_1, то есть $q_0 \rho q$, что противоречит предположению. Следовательно, язык, задаваемый S_1, совпадает с языком, задаваемым множеством S_1-связности. Аналогично, язык, задаваемый S_2, совпадает с языком, задаваемым множеством S_2-связности. Так как множества связности совпадают, лемма доказана.

Следствие 1. *Пусть $V = (A, Q, \varphi, q_0)$. Тогда NS_V совпадает с числом различных S-связных подмножеств Q.*

Лемма 6 (о монотонности). *Пусть $V_1 = (A, Q, \varphi_1, q_0)$, $V_2 = (A, Q, \varphi_2, q_0)$ — два автомата, причем для любого $S \subseteq Q$ множество S-связности V_2 содержит множество S-связности V_1. Тогда $NS_{V_1} \leqslant NS_{V_2}$.*

Доказательство. В силу следствия из леммы о связности достаточно показать, что число различных S-связных подмножеств Q для автомата V_2 больше или равно числа таких подмножеств для V_1. Пусть $S_1 \subseteq Q$ и $S_2 \subseteq Q$ порождают различные связные множества для автомата V_1. Следовательно, связное S_1 множество содержит элемент, не принадлежащий S_2, а связное S_2 множество содержит элемент, не принадлежащий S_1. В силу того, что связные множества для V_2 содержат соответствующие множества для S_1, они также различаются. Лемма доказана.

Лемма 7 (о понижении кратности ребер). *Пусть* $V_1 = (A, Q, \varphi_1, q_0)$, $V_2 = (A, Q, \varphi_2, q_0)$ — *два автомата, диаграммы Мура которых имеют одно отличие: у диаграммы* V_1 *имеется кратное ребро (переход в одно и то же состояние по двум различным входным символам), а у* V_2 *вместо кратного ребра добавлен переход в другое состояние (рис. 3). Тогда* $NS_{V_1} \leqslant NS_{V_2}$.

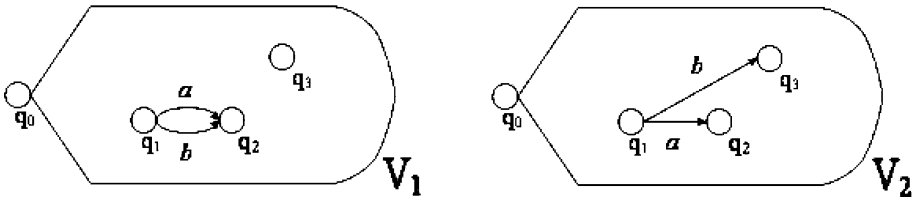

Рис. 3. Операция понижения кратности ребер.

Доказательство. По лемме о монотонности, достаточно показать, что для любого $S \subseteq Q$ множество S-связности V_2 содержит множество S-связности V_1. Если S не содержит состояния, входящие в отличающийся блок диаграммы Мура, из совпадения диаграммы Мура следует совпадение S-связных множеств. Пусть S включает отличающийся блок целиком или частично. Покажем, что и в этом случае S-связность для V_1 влечет S-связность для V_2. Действительно, если путь для V_1 не включает модифицированную вершину, для V_2 имеется такой же путь. Если же модифицированная вершина входит в путь, можно построить путь для V_2, заменив пометку модифицированного кратного ребра на пометку оставшегося ребра. Лемма доказана.

Лемма 8 (о ярусе). *Пусть* $V_1 = (A, Q, \varphi_1, q_0)$, $V_2 = (A, Q, \varphi_2, q_0)$ — *два автомата, причем в диаграмме Мура для* V_1 *имеется неориентированный цикл длины 3* $q_1q_2q_3$, *а в диаграмме Мура для* V_2 *цикл разомкнут (рис. 4). Тогда* $NS_{V_1} \leqslant NS_{V_2}$.

Доказательство. По лемме о монотонности, достаточно показать, что для любого $S \subseteq Q$ множество S-связности V_2 содержит множество S-связности V_1. Если S не содержит состояния, входящие в отличающийся блок диаграммы Мура, из совпадения диаграммы Мура

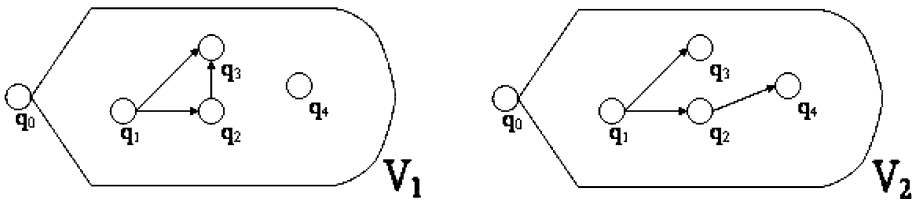

Рис. 4. Операция размыкания яруса.

следует совпадение S-связных множеств. Пусть S включает отличающийся блок целиком или частично. Если в S входит ровно одна вершина цикла, множества связности V_2 содержат множества связности V_1, так как отличие заключается в дополнительном ребре в диаграмме V_2. Если в S входит три вершины, множества связности V_2 содержат множества связности V_1, так как в любом пути для V_1 можно заменить подпути $q_1q_2q_3$ на q_1q_3. Если в S входит две вершины, то либо они обе не попадают в множества связности V_1, и монотонность сохраняется, или обе одновременно входят или не входят в множества связности V_1 и V_2, так как в диаграммах Мура обоих автоматов соединены ребром. Лемма доказана.

2.6. Доказательство теоремы 1

Верхняя оценка вытекает из того, что начальное состояние должно принадлежать S по определению. Всего подмножеств состояний, содержащих начальное, — 2^{n-1}. По лемме о монотонности, для любого автомата V $NS_V \leqslant 2^{n-1}$. Отметим, что оценка достигается на автоматах с $l+1$ состоянием. Действительно, по лемме о размыкании кратных ребер и лемме о ярусе достаточно подсчитать число поддеревьев дерева из корня и l листьев (рис. 5), содержащих корень, а их 2^l.

Нижняя оценка доказывается следующим образом. Рассмотрим автомат, диаграмма Мура которого имеет вид полного сбалансированного дерева с петлями, присоединенными к листьям (рис. 6). Ели последний ярус полон, то $n = \frac{l^k-1}{l-1}$ для некоторого $k \in \mathbb{N}$, как сумма геометрической прогрессии. Вычислим количество $f(k)$ различных

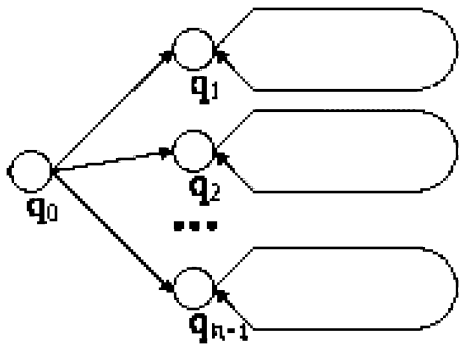

Рис. 5. Автомат, на котором достигается верхняя оценка.

связных множеств такого автомата. Покажем, что $f(k)$ удовлетворяет рекурентному соотношению

$$f(k+1) = (1+f(k))^l, f(1) = 1.$$

Действительно, при $k = 1$ получается автомат с одним состоянием, входящим в любое множество S по определению. Пусть вычислено $f(k)$. Рассмотрим дерево высоты $k+1$. Число множеств связности такого дерева можно посчитать следующим образом. Рассмотрим множества, содержащие m вершин второго яруса, $0 \leqslant m \leqslant l$. Их C_l^m, и мощность каждого равна $f(k)^m$. При суммировании возникает бином Ньютона, что доказывает рекурентное соотношение. По следствию из леммы о связности, $NS_V = f(k)$.

Так как $f(k)+1 > f(k)$, $f(k) <= f(t)^{l^{k-t}}$ для любого натурального t, не превосходящего k. Учитывая, что $k = \log_l(1+n(l-1))$, получаем, что

$$f(t)^{l^{k-t}} = f(t)^{l^{\log_l(1+n(l-1))-t}} = f(t)^{\frac{1+n(l-1)}{l^t}} = 2^{\frac{\log f(t)(1+n(l-1))}{l^t}}.$$

При $t = 2$ получаем оценку из условия теоремы. Действительно, $f(2) = 2^l$, $\log f(2) = l$, а второй сомножитель числителя показателя можно уменьшить на 1 с сохранением неравенства.

Замечание 1. если брать большие t, оценка будет улучшаться. Для $l = 2$, например, легко получается $\frac{37}{64}$ вместо $\frac{1}{2}$.

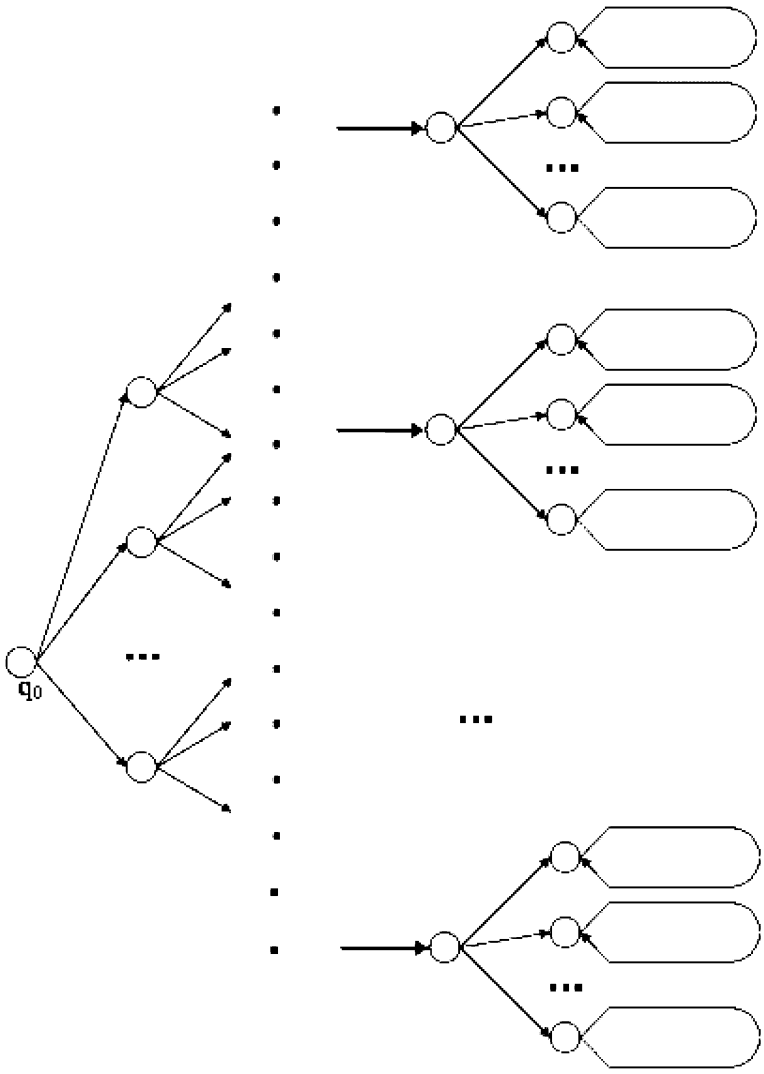

Рис. 6. Полное сбалансированное дерево с петлями.

Замечание 2. При росте мощности алфавита показатель стремится к n, то есть доля безопасных языков велика.

При оставшихся n рассмотрим максимальное сбалансированное

полное поддерево. В этом случае в формуле для k нужно взять целую часть, то есть показатель уменьшится не более, чем на 1, то есть в итоговой формуле показатель поделится еще на l.

2.7. Доказательство теоремы 2

Рассмотрим автомат V_{01} с двумя состояниями и входным алфавитом из 0 и 1, по 1 переходящем в безопасное состояние, по 0 — в небезопасное. Рассмотрим произвольное $\varepsilon \in \mathbb{Q}$, $0 < \varepsilon < 1$. Пусть $\varepsilon = \frac{p}{q}$. Тогда ε-безопасный язык состоит из тех и только тех слов, в которых число единиц, умноженное на $q - p$, больше или равно числа нулей, умноженного на p. Действительно, число безопасных состояний равно числу 1, число небезопасных состояний — числу 0, и отношение числа небезопасных состояний к длине слова равно $|0|/|1| + |0|$, где $|0|$ и $|1|$ — число нулей и единиц соответственно. Это соотношение не превосходит ε тогда и только тогда, когда выполнены приведенные выше условия.

Покажем, что построенный язык не является регулярным. Действительно, предположим противное — существует автомат V, распознающий язык. Пусть число состояний V равно n. Рассмотрим входное слово, состоящее из $(q-p)n$ единиц, за которыми идут pn нулей. Слово принадлежит языку, следовательно переводит V в принимающее состояние. Рассмотрим последовательность состояний, в которые осуществляется переход при подаче нулей. По принципу ящиков Дирихле, среди таких состояний найдутся повторяющиеся, то есть если добавить в конце слова нули, число которых кратно расстоянию между повторяющимися состояниями, слово будет приниматься автоматом, однако оно не будет принадлежать языку. Противоречие.

Покажем, что любой ε-безопасный язык является контекстно-свободным. Пусть $\varepsilon = \frac{p}{q}$. По теореме о распознавании контекстно-свободных языков магазинными автоматами ([4]) достаточно показать, что любой такой язык распознается магазинным автоматом. Автомат будем строить следующим образом. Возьмем ленточный алфавит из трех символов — $+, -, \lambda$. Множество состояний представляет собой прямое произведение множества состояний автомата V, порождающего ε-безопасный язык, двухэлементного множества $\{A, R\}$, со-

держащего принимающий символ и отвергающий символ, и счетчика до q. Начальным состоянием является тройка $(q_0, A, 0)$. Множество принимающих состояний состоит из тех и только тех элементов, в которых присутствует принимающий символ. Переходы осуществляются по следующему правилу. Если вход непустой, то первый элемент нового состояния равен значению функции переходов V на первом элементе старого состояния и пришедшем входе, в противном случае первый элемент не изменяется. Если вход непустой, то значение счетчика полагается равным $q-1$, в противном случае значение счетчика уменьшается на 1. Если значение счетчика равно 0, то принимаются только непустые входы, в противном случае принимаются только пустые входы. Если вход пустой, первая компонента состояния безопасна, а верхний символ ленты равен $+$, на ленту дописывается $+$. Если вход пустой, первая компонента состояния небезопасна, а верхний символ ленты равен $-$, верхний символ стирается. Если вход пустой, первая компонента состояния небезопасна, а верхний символ ленты равен $-$, на ленту дописывается $-$. Если вход пустой, первая компонента состояния безопасна, а верхний символ ленты равен $+$, верхний символ стирается. Вторая компонента состояния равна A либо в начальный момент, либо тогда, когда счетчик равен 0, а верхний символ на ленте равен $+$.

Покажем, что построенный автомат распознает ε-безопасный язык. Поставим в соответствие каждому входному слову число, вычисляющееся по следующему правилу. Число безопасных состояний на слове умножается на $q-p$, затем вычитается число небезопасных состояний, умноженное на p. Очевидно, что слово принадлежит языку тогда и только тогда, когда ему соответствует положительное число. Покажем, что знак числа является верхним символом ленты. Действительно, в каждый момент обработки непустого входного символа на ленте записано число, соответствующее поданному слову (если число отрицательное, его модуль равен числу минусов, в противном случае — числу плюсов). При этом на ленте имеются символы только одного вида. Это легко показывается индукцией по длине входного слова.

Рассмотрим язык L, состоящий из двух слов: 00 и 111. Предположим, L является ε-безопасным для некоторого автомата V. Так как

00 принадлежит L, а 0 — нет, на слове 00 V сначала переходит в состояние из I, затем — в состояние из S, и $\varepsilon \geqslant \frac{1}{2}$. Так как слово 111 принадлежит L, а слова 1 и 11 — нет, на слове 111 сначала V переходит в состояние из I, а в конце — из S. Пусть второе состояние принадлежит I. Следовательно, $\varepsilon \geqslant \frac{2}{3}$, следовательно слово 001 принадлежит L — противоречие. Пусть второе состояние принадлежит S. Так как 11 не принадлежит L, $\varepsilon < \frac{1}{2}$ — противоречие. Теорема 2 доказана.

2.8. Доказательство теоремы 3

Пусть $\varepsilon_1, \varepsilon_2 \in \mathbb{Q}$, $0 \leqslant \varepsilon_1 < \varepsilon_2 \leqslant 1$. Обозначим через L_ε ε-безопасный язык, порожденный автоматом V_{01}.

Если $\varepsilon_2 = 1$, то в качестве ε_1-безопасного языка рассмотрим L_{ε_1}. Так как все 1-безопасные языки состоят из всех конечных слов, а L_{ε_1} — нет, случай $\varepsilon_2 = 1$ доказан.

Если $\varepsilon_1 = 0$, то ε_1-безопасные языки являются просто безопасными языками. Пусть $\varepsilon_2 = \frac{p}{q}$. Рассмотрим язык L, состоящий из слова длины $2q - 1$ и всех его префиксов. По лемме 2, этот язык является безопасным. Предположим, существует автомат V, для которого L будет ε_2-безопасным. Рассмотрим последовательность состояний на слове максимальной длины. Безопасных состояний должно быть $2(q - p)$, так как $2(q - p) - 1 < (2q - 1)\frac{q-p}{q}$. Следовательно, если приписать к слову максимальной длины произвольный символ, оно все равно будет входить в ε_2-безопасный язык автомата V. Противоречие.

Пусть ε_2 не равно 1. Так как все 0-языки регулярны, а L_{ε_2} регулярным не является, случай $\varepsilon_1 = 0$ доказан.

Пусть $0 < \varepsilon_1 < \varepsilon_2 < 1$. Рассмотрим L_{ε_1}. Без ограничения общности можно считать, что $\varepsilon_1 = \frac{p_1}{q}$, $\varepsilon_2 = \frac{p_2}{q}$, причем $p_1 < p_2$. Предположим, существует автомат V, для которого L_{ε_1} будет ε_2-безопасным. Рассмотрим входные слова α_n, определяемые следующим образом. $\alpha_1 = 1$. Предположим, что слово α_n построено. Тогда $\alpha_{n+1} = \alpha_n|0$, если в слове $\alpha_n|0$ доля 0 не превосходит $\frac{p_1}{q}$, в противном случае положим $\alpha_{n+1} = \alpha_n|1$ (символ $|$ означает конкатенацию). Отметим, что все построенные слова принадлежат L_{ε_1}. Рассмотрим последовательность состояний, порождаемую в V построенными словами.

Легко увидеть, что при подаче на вход символа 1 осуществляется переход в безопасное состояние, при подаче на вход 0 — в небезопасное. Действительно, так как $\alpha_1 \in L_{\varepsilon_1}$, первое состояние безопасно. Если последний символ α_{n+1} равен 1, то последнее состояние безопасно, так как в противном случае слово, полученное из α_{n+1} заменой последней 1 на 0, принималось V, что противоречит построению α_{n+1}. Если последний символ α_{n+1} равен 0, то последнее состояние небезопасно, так как в противном случае можно рассмотреть минимальное $n' > n+1$ такое, что последний символ $\alpha_{n'}$ равен 1. Тогда при замене последнего символа $\alpha_{n'}$ на 0 получится слово, принимаемое V, но не принадлежащее L_{ε_1}, что приводит к противоречию.

Если $|\alpha_n| = q$, то по построению число 0 в α_n строго меньше $p_1 + 1 \leqslant p_2$. Следовательно, слово, полученное заменой самой правой 1 на 0 принимается V, что приводит к противоречию.

Пусть $0 < \varepsilon_1 < \varepsilon_2 < 1$. Рассмотрим L_{ε_2}. Без ограничения общности можно считать, что $\varepsilon_1 = \frac{p_1}{q}$, $\varepsilon_2 = \frac{p_2}{q}$, причем $p_1 < p_2$. Предположим, существует автомат V, для которого L_{ε_2} будет ε_1-безопасным. Рассмотрим входные слова α_n, определяемые следующим образом. $\alpha_1 = 0$. Предположим, что слово α_n построено. Тогда $\alpha_{n+1} = \alpha_n|1$, если в слове $\alpha_n|1$ доля 1 не превосходит $\frac{q-p_2}{q}$, в противном случае положим $\alpha_{n+1} = \alpha_n|0$. Отметим, что все построенные слова не принадлежат L_{ε_2}. Рассмотрим последовательность состояний, порождаемую в V построенными словами.

Легко увидеть, что при подаче на вход символа 1 осуществляется переход в безопасное состояние, при подаче на вход 0 — в небезопасное. Действительно, так как $\alpha_1 \notin L_{\varepsilon_2}$, первое состояние небезопасно. Если последний символ α_{n+1} равен 0, то последнее состояние небезопасно, так как в противном случае слово, полученное из α_{n+1} заменой последней 0 на 1, не принималось V, что противоречит построению α_{n+1}. Если последний символ α_{n+1} равен 1, то последнее состояние безопасно, так как в противном случае можно рассмотреть минимальное $n' > n+1$ такое, что последний символ $\alpha_{n'}$ равен 0. Тогда при замене последнего символа $\alpha_{n'}$ на 1 получится слово, не принимаемое V, но принадлежащее L_{ε_2}, что приводит к противоречию.

Если $|\alpha_n| = q$, то по построению число 1 в α_n строго меньше

$q - p_2 > q - p_1$. Следовательно, слово, полученное заменой самого правого 0 на 1 не принимается V, что приводит к противоречию. Теорема доказана.

Замечание 3. При доказательстве теоремы 3 показано, что для любого $\varepsilon \in \mathbb{Q}$, $0 < \varepsilon < 1$, существуют языки, являющиеся ε-безопасными и не являющиеся ε_1-безопасными ни при каких $\varepsilon_1 \in \mathbb{Q}$, $\varepsilon_1 \neq \varepsilon$.

Замечание 4. Если отказаться от требования рациональности ε, мощность множества языков, являющихся ε-безопасными для некоторого $\varepsilon \in [0, 1]$, равна континууму. Действительно, верхняя оценка следует из континуальности множества всех языков, нижняя — из попарной различности языков L_ε.

Список литературы

[1] Галатенко А. В. Активный аудит // JetInfo. № 8. 1999. (http://www.jetinfo.ru/1999/8/1/article1.8.1999.html)

[2] Галатенко А. В. Об автоматной модели защищенных компьютерных систем // Интеллектуальные системы. Т. 4. Вып. 3–4. Москва, 1999. С. 263–270.

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

[4] Маркус С. Теоретико-множественные модели языков. М.: Наука, 1970.

[5] Moskowitz I. S., Costich O. L. A Classical Automata Approach to Noninterference Type Problems. Department of the Navy, Naval Research Laboraory, 1992.