PROCEEDINGS

of the Union of Scientists - Ruse

Book 5

Mathematics, Informatics and Physics

Volume 9, 2012

RUSE

The Ruse Branch of the Union of Scientists in Bulgaria was founded in 1956. Its first Chairman was Prof. Stoyan Petrov. He was followed by Prof. Trifon Georgiev, Prof. Kolyo Vasilev, Prof. Georgi Popov, Prof. Mityo Kanev, Assoc. Prof. Boris Borisov, Prof. Emil Marinov, Prof. Hristo Beloev. The individual members number nearly 300 recognized scientists from Ruse, organized in 13 scientific sections. There are several collective members too – organizations and companies from Ruse, known for their success in the field of science and higher education, or their applied research activities. The activities of the Union of Scientists – Ruse are numerous: scientific, educational and other humanitarian events directly related to hot issues in the development of Ruse region, including its infrastructure, environment, history and future development; commitment to the development of the scientific organizations in Ruse, the professional development and growth of the scientists and the protection of their individual rights.

The Union of Scientists – Ruse (US – Ruse) organizes publishing of scientific and popular informative literature, and since 1998 – the “Proceedings of the Union of Scientists- Ruse".

BOOK 5

"MATHEMATICS, INFORMATICS AND PHYSICS"

VOLUME 9

CONTENTS

Mathematics

Informatics

Physics

δ -CHARACTERISTIC SETS FOR A FINITE STATE ACCEPTOR

Mihail Kirilov

Angel Kanchev University of Ruse

Abstract: *A characteristic set for a specific word* ***u*** *and a finite state acceptor has been considered. The word* ***u*** *consists of string of words that the automaton accepts divided by the symbol* ***s*** *such that when read by the automaton, it makes a transition from the final state to the initial state.*
Keywords: *finite state machine, characteristic set.*

INTRODUCTION

Finite state automata (FSA) play an important role in discrete mathematics. They are used in many different areas including computer science, philosophy, biology, mathematics, logic. In computer science FSA are used in modeling of application behavior, compilers and network protocols. In biology and artificial intelligence research, FSA are sometimes used to describe neurological systems. They are also used to describe fractals.

EXPOSITION

On the base of the unified approach to the dynamical systems and the finite automata [2, 3, 4, 5] and considering the summary papers [2, 3, 4, 5, 7] in this paper the δ-characteristic sets for a word and finite state acceptor has been examined.

Let's recall some fundamental definitions, modifying some of them and introducing certain special notations and suitable general replacements following [1], [5], [6].

Definition 1. A finite automaton acceptor is a 5-tuple $M = \{Q, V, q_0, \delta, q_m\}$ where $Q = \{q_0, q_1, \dots, q_m\}$ is a finite set of control states; *V* is the input alphabet; $\delta : QxV \to Q$ is the state transition function; $q_0 \in Q$ is the initial state; $q_m \in Q$ is the final state.

Note: *T(M)* is the set of all words, accepted by the automaton *M*. According to Cleen's theorem *T(M)* is a regular expression. Put $B(k) \subset T(M)$, where *B(k)* is the set of all words $\beta \in V^*$ accepted by the automaton *M* such that $|\beta| = k$. Then the words in *B(k)* are arranged in an ascending way by alphabetical order. Assume that $u = u_1 s u_2 s u_3 \dots$ is a word, $u \in V^*$, where $u_i \in T(M)$ and $\delta(q_m, s) = q_0$. $u_i = s_{i_1} s_{i_2} s_{i_{m_i}}$, $u_i \in B(m_i)$, $\delta(q_0, u_i) = q_m$.

Let $p(\alpha, n)$ be the string of the first *n* letters of the word α.

Definition 2. The set $C_\alpha(n) = \{\tau \in N \mid \delta(q, p(\alpha, n)) = \delta(q, p(\alpha, n + \tau))\}$ is called δ-characteristic set of semi-dynamical system defined by the FSA *M*.

Example: The FSA $M = \{Q, V, q_0, \delta, q_m\}$ is defined as follows: $Q = \{1,2,3,4\}$, $V = \{a, b, c\}$, $q_0 = 1$, $q_m = 4$ and its function λ is given in fig.1

M	1	2	3	4
a	3	4	2	4
b	4	3	1	3
c	2	1	4	1

Fig.1

The set of words *B(k)* ($k = 1,2,3,\dots$) follows:

$B(1) = \{b\}$

$B(2) = \{ac, ba, ca\}$

$B(3) = \{aaa, abb, aca, baa, bbc, bcb, caa, cbc, ccb\}$

$B(4) = \{aaaa, aabc, aacb, abac, abba, abca, acaa, acbc, accb, baaa, babc, bacb, bbaa,$
$bbbb, bbca, bcac, bcba, bcca, caaa, cabc, cacb, cbaa, cbbb, cbca, ccac, ccba, ccca\}$

For every control state of *M* exists a symbol $s \in V$ such that $\delta(q_i, s) = q_m$. This fact leads to the proposition that the first $(k-1)$ symbols for a word in *B(k)* can be chosen as any of the symbols from V^*.

Therefore for every symbol from the first $(k-1)$ symbols in a word from *B(k)* there are three possibilities – *a*, *b* or *c*. Thus the number of words in *B(k)* is 3^{k-1}.

Experimental results over three words are done. The difference in the words is that the second one consists of words with greater length that the automaton accepts than the first one. The third one consists of shorter words that the automaton accepts. The role of the symbol *s* is assigned to the symbol *c*.

Experimental results 1:

Consider the word

u = ***babc****c****aaaaa****c****abb****c****bacb****c****ac****c****ccccccabb****c****cbc****c*…,

where

$u_1 = babc$, $u_2 = aaaaa$, $u_3 = abb$, $u_4 = bacb$, $u_5 = ac$, $u_6 = ccccccabb$, $u_7 = cbc$, $s = c$.

In the process of reading of the word *u*, the next control states were passed over:

1→4434132444131414441413412121213144234…

$C_u(1) = \{1, 3, 7, 8, 9, 13, 15, 16, 18, 21, 31, 32, 35, \dots\}$;

$C_u(2) = \{2, 6, 7, 8, 12, 14, 15, 17, 20, 30, 31, 34, \dots\}$;

$C_u(3) = \{3, 9, 18, 27, 32, \dots\}$;

$C_u(4) = \{4, 5, 6, 10, 12, 13, 15, 18, 28, 29, 32, \dots\}$;

$C_u(5) = \{6, 8, 10, 13, 15, 18, 20, 22, 24, 26, \dots\}$;

$C_u(6) = \{6, 15, 24, 29, \dots\}$;

$C_u(7) = \{17, 19, 21, 27, \dots\}$;

$C_u(8) = \{1, 2, 6, 8, 9, 11, 14, 24, 25, 28, \dots\}$;

$C_u(9) = \{1, 5, 7, 8, 10, 13, 23, 24, 27, \dots\}$;

$C_u(10) = \{4, 6, 7, 9, 12, 22, 23, 26, \dots\}$;

$C_u(11) = \{2, 4, 7, 9, 12, 14, 16, 18, 20, \dots\}$;

…

Experimental results 2:

Consider the word

u = ***bbbb****c****abb****c****bbaa****c****ac****c****ababababb****c****cac****c****acaa****c****ccb****c*…,

where

$u_1 = bbbb$, $u_2 = abb$, $u_3 = bbaa$, $u_4 = ac$, $u_5 = ababababb$, $u_6 = cac$, $u_7 = acaa$, $u_8 = ccb$, $s = c$.

In the process of reading the word u, the next control states were passed over:

1→43141314143241341313131314121341344412141…
$C_u(1)=\{3,7,9,12,15,25,30,33,34,35,39,...\}$;
$C_u(2)=\{4,9,13,16,18,20,22,28,31,...\}$;
$C_u(3)=\{2,4,6,11,14,16,18,20,22,24,26,29,34,36,38,...\}$;
$C_u(4)=\{8,10,13,16,26,31,34,35,36,40,...\}$;
$C_u(5)=\{2,4,9,12,14,16,18,20,22,24,27,32,34,36,...\}$;
$C_u(6)=\{5,9,12,14,16,18,24,27,...\}$;
$C_u(7)=\{2,7,10,12,14,16,18,20,22,25,30,32,34,...\}$;
$C_u(8)=\{2,5,8,18,23,26,27,28,32,...\}$;
$C_u(9)=\{5,8,10,12,14,16,18,20,23,28,30,32,...\}$;
$C_u(10)=\{3,6,15,20,23,24,25,29,...\}$;
$C_u(11)=\{4,7,9,11,13,19,22,...\}$;
…

Experimental results 3:
Consider the word
u = *abbcbcbacccaccccaccbabccacaccaacbccabcccac*…,
where
$u_1=abb$, $u_2=bcba$, $u_3=ccac$, $u_4=ccac$, $u_5=babc$, $u_6=aca$, $u_7=caa$, $u_8=b$, $u_9=cabc$, $u_{10}=ca$, $s=c$.
In the process of reading the word u, the next control states were passed over:

1→31414144121341213414434134412441412434124I…
$C_u(1)=\{11,16,21,24,36,...\}$;
$C_u(2)=\{2,4,7,9,12,14,17,22,26,30,32,37,40,...\}$;
$C_u(3)=\{2,4,5,10,15,17,18,20,23,24,27,28,30,33,35,38,...\}$;
$C_u(4)=\{2,5,7,10,12,15,20,24,28,30,35,38,...\}$;
$C_u(5)=\{2,3,8,13,15,16,18,21,22,25,26,28,31,33,36,...\}$;
$C_u(6)=\{3,5,8,10,13,18,22,26,28,33,36,...\}$;
$C_u(7)=\{1,6,11,13,14,16,19,20,23,24,26,29,31,34,...\}$;
$C_u(8)=\{5,10,12,13,15,18,19,22,23,25,28,30,33,...\}$;
$C_u(9)=\{2,5,7,10,15,19,23,25,30,33,...\}$;
$C_u(10)=\{5,18,24,29,...\}$;
$C_u(11)=\{3,5,8,13,17,21,23,28,31,...\}$;
…

CONCLUSION
In this paper δ - characteristic set for a word *u* and FSA *M* has been considered. The word *u* consists of string of symbols that the automata accepts and the symbol s between every couple of words, where $s\in V:\delta(q_m,s)=q_0$. The next aim is to consider optimization

problem for a word and FSA over the state transition function with weighting. The numerical experiments have been performed using suitable software, developed by the author [5].

REFERENCES

[1] Manev K., Introduction to Discrete Mathematics, KLMN, Sofia, 2003.
[2] Saperstone S. H., Semidynamical Systems in Infinite Dimensional Spaces, Applied mathematical systems, 1981.
[3] Dochev D. Tr., On some properties of discrete semidynamical systems generated by finite automatons, UAAM, Sofia, V.21, 1985, 53-64.
[4] Sibirskiy K., Introduction to Topological Dynamics, AN MSSP, Chisinau, 1970.
[5] Dochev, D.Tr., M.K. Kirilov, Fundamental properties of discrete semidynamical systems defined by finite automatons. Научни Трудове на Русенския Университет, V.46, B 6, 2007, 23-30.
[6] Kirilov, M.K. Splitting of the sets of words with fixed length. Научни Трудове на Русенския Университет, V.47, B 5.1, 2008, 51-56.
[7] Hopcroft J., Aho A., and Ullman J., Data Structures and Algorithms, Addison-Wesley Series in Computer Science and Information Processing, 1983.

CONTACT ADDRESS

Assist. Mihail Kirilov, PhD
Department of Algebra and Geometry
Angel Kanchev University of Ruse
8 Studentska Str., 7017 Ruse, Bulgaria
E-mail: mkirilov@uni-ruse.bg

δ -ХАРАКТЕРИСТИЧНИ МНОЖЕСТВА ЗА КРАЕН АВТОМАТ РАЗПОЗНАВАТЕЛ

Михаил Кирилов

Русенски университет „Ангел Кънчев”

Резюме: *Разгледани са характеристични множества за специфична дума u и краен автомат разпознавател. Думата u се състои от поредица от думи, които се разпознават от автомата, разделени от символ s, такъв че когато бъде прочетен от автомата, се прави преход от заключителното състояние към началното състояние.*

Ключови думи: *краен автомат разпознавател, характеристични множества.*

ISSN 1314-3077