V. A. Luzhetskyi, Dc. Sc(Eng.)., Prof.; V. M. Mykhailevych, Dc. Sc(Eng.)., Prof.; A. V. Mykhailevych; V. F. Kaplun

EVALUATION OF NUMBER OF THE UNIQUE LINEAR RECURRENT SEQUENCES OF THE SECOND ORDER

Linear recurrent sequences, for which each member, beginning with the third one, equals the sum of the two previous, had been researched from the point of view of possible compression and data enciphering. There had been received the correlation for the determination of all members of such sequences, belonging to the interval on condition of specific restrictions for the first two members. It had been established that the density in completing the sequence of natural numbers with the indicated members is insufficient for obtaining high practical value.

Key words: linear recurrent sequences, m-valued numbers, density, compression.

Introduction

There are approaches to compression, based on using statistic data peculiarities and on using dictionary as well as on specific representations of numerical data [1]. Specialists however are in search of more efficient compression methods. One of the approaches is based upon the use of specific functional data dependences, represented as the whole numbers. Papers [0, 0] suggest an approach to data compression, based upon the compact representation of large numbers by linear homogeneous recurrent sequences of the second order. However there are no theoretical evaluations of number of unique sequences of such type with specific restrictions for the first members, which, in their turn, would allow to receive the evaluation of potential possibilities of the similar way for data compression. The objective of the paper is to study the density in completing with members of the above sequences certain ranges of natural numbers.

The tasks of the paper:

Development of methodic for evaluation of number of the unique linear homogeneous recurrent sequences of the second order at specific restrictions for the first two members;

Receiving correlation for the determination of number of the unique sequences of the above type;

Receiving the evaluation for the determination of density of completion of natural numbers with members of the sequences under research.

The main part

Let us consider the generalized Fibonacci sequences [0], namely: sequences, for which each member, beginning with the third one, equals the sum of the previous:

$$u_1, u_2, \dots, u_n = u_{n-2} + u_{n-1}, \dots \quad n = 3, 4, \dots,$$
 (1)

where u_1 , u_2 – are positive whole numbers.

Paper [0] formulates and proves the following peculiarities: any sequence of number determined by correlation (1), when m>2 contains 4 or 5 *m*-valued numbers. Really, in accordance with the adduced proofs [0], the formulation of this peculiarity must be as follows: for any u_1 , u_2 sequence (1) there is such whole number $M=M(u_1, u_2)$, that for all $m\ge M$ in the sequence there are only 4 or 5 *m*-valued numbers.

The above-mentioned formulation is the generalization of the peculiarities formulated in [0] regarding the Fibonacci sequence of numbers for which M=2.

Basing on closed correlation for the general member of sequence (1),

$$u_n = u_1 \cdot \frac{\alpha^{n-2} - \beta^{n-2}}{\sqrt{5}} + u_2 \cdot \frac{\alpha^{n-1} - \beta^{n-1}}{\sqrt{5}},$$
(2)

where

$$\alpha = \frac{1+\sqrt{5}}{2}, \quad \beta = \frac{1-\sqrt{5}}{2},$$
(3)

allows to receive formula for the calculation of number of k_m *m*-valued numbers of sequence (1) depending on values *m*, u_1 , u_2 :

$$k_{m} = \begin{cases} 5, \ \{M\} \leq \left\{\frac{1}{\lg \alpha}\right\} \approx 0,78972, \\ 4, \ \{M\} > \left\{\frac{1}{\lg \alpha}\right\} \approx 0,78972, \\ 0,5 \cdot \lg 5 + m - \lg \left(\frac{u_{2} - u_{1}}{\alpha} + u_{1}\right) \\ \lg \alpha \end{cases},$$
(4)

where $\{x\}$ designates the fractional part of the number *x*.

Direct calculation in the environment of the system of symbol mathematics Maple allowed to realize the verification of correctness in correlation (4) for the following ranges of arguments variation: $m=8\div20$, $u_1=1\div300$, $u_2=1\div300$.

Fixing the two parameters out of three and changing the third parameter in the indicated range allowed the research the frequency in appearance of values $k_m=4$, $k_m=5$.

On the base of the received data there had been built the table for distribution of possible amount of k_m *m*-valued numbers of the sequence (1).

Table 1

Distribution k_m

$(k_n)_i$	4	5
p_i	0,21	0,79

The same result follows from the correlation (4) and the assumption on equal distribution of mantissa of numbers, determined by correlation $\{(0,5 \cdot \lg 5 + m - \lg (u_2 - u_1/\alpha + u_1))/\lg \alpha\}$.

It is obvious that the density in completing the series of natural numbers of the separate recurrent sequence is pretty small already for m>4 and quickly decreases with the increase in m. At the same time, for the evaluation of efficiency in application of transformation of the whole numbers by means of the sequences (1) it is necessary to evaluate the number of unique sequences with the specific restrictions for the first two members.

We restrict the range of the change of the first two members of sequence by the segment [1, 99]. Then calculate the total number of sequences (1) which may be formed under the given restrictions. This number equals the number of arrangements with repetition form the set N=99 elements in k=2 elements:

$$\overline{A}_N^k = N^k = 99^2 = 9801.$$
⁽⁵⁾

However not all such sequences are unique. By the set of unique sequences we mean such set of sequences for each of which the following condition is satisfied: the given sequence may not be received by truncation of the final number p>0 of the first members of any other sequence of the given set. Let us give an example of two sequences of the type (1):

1) 3, 2, 5, 7, 12, 19, 31, 50, 81, 131, 212, 343, 555, 898, 1453, 2351, 3804, 6155, 9959, 16114, 26073, 42187, ...

2) 50, 81, 131, 212, 343, 555, 898, 1453, 2351, 3804, 6155, 9959, 16114, 26073, 42187, 68260,110447, ...

It is obvious that the second sequence is possible to receive form the first by truncation of the first seven numbers.

So our task is to determine the set of unique sequences and calculate their number with the availability of some restrictions on the range of change in the first two members of sequence. The following algorithm is suggested:

1. There shall be created the void set $M^{(2)}$, elements of which are ordered pairs of numbers.

2. For each of 9801 pairs of numbers $a_1 \in [1, 99]$, $a_2 \in [1, 99]$ the sequence (1) continues to the left up to the smallest positive u_1 . The received pair of numbers $[u_1, u_2]_i$ is being fixed.

3. The existence of the current element in the set $M^{(2)}$ is repeated. If such an element does not exist, then the current pair of numbers is introduced into the indicated set.

As a result the set $M^{(2)}$ of the unique sequences (1), is being formed, each of which is presented by its first two members u_1 and u_2 , satisfying the conditions u_2 - $u_1 \leq 0$.

The picture presents the graphic illustration to the arrangement of the unique sequences for different ranges of possible values of the first two members of the sequence (vertical column of the numbers – values u_1 , horizontal – u_2). Sign "+" designates all cells which are intersection of pair of numbers, causing the unique sequence.

On the base of the reduced data there had been built the correlation for calculation of number of unique sequences q_N , the first two members of which satisfy the condition

$$u_2 - u_1 \le 0; \ u_1 \in [1, N], \ u_2 \in [1, N].$$
 (6)

This correlation looks like:

$$q_N = N^2 - \sum_{i=1}^{\frac{N-1}{2}} (2i-1) - N + 1 - \sum_{i=1}^{\frac{N-3}{2}} 2i.$$
(7)

After simple transformations we receive the correlation for calculating the number of unique sequences q_N depending on the range of change of the first two members

$$q_N = \frac{1}{2} \cdot N \cdot \left(N+1\right). \tag{8}$$

When *N*=99 the effect of compression may be revealed on 8-valued members. According to (4), we assume that the number of k_m =5 *m*-valued number of sequence (1) is met $\frac{0,79}{0,21}\approx3,76$ times more frequently than k_m =4. Then the density of completion of range of natural numbers $10^7 \div 10^8$ -1 by members of sequence (1) may be determined as the correlation of all numbers of sequence which fall into the set range, to the length of this range:

$$p_{2} = \frac{q_{99} \cdot (5 \cdot 0, 79 + 4 \cdot 0, 21)}{(10^{8} - 1) - (10^{7} - 1)} =$$

$$= \frac{4950 \cdot 4, 79}{9 \cdot 10^{7}} \approx 0,00027$$
(9)



INFORMATIONAL TECHNOLOGIES AND COMPUTER ENGINEERING

4

With the increase in the range of changing the first two numbers of the sequence increase the length of numbers, for which there appears the effect of compression. Generalizing (8) in case $N=10^{r}-1$ (r=2, 3, ...), we write

$$p_{r} = \frac{10^{r} \cdot (10^{r} - 1) \cdot 4,79}{2 \cdot (10^{2r+4} - 10^{2r+3})} \approx , \qquad (10)$$
$$\approx \frac{4,79}{18 \cdot 10^{3}} \approx 0,00027$$

From which it follows that the density of completion the sequence of natural numbers by numbers of sequence (1) is not changed along with the increase of the range N, following the above law.

In accordance with the relation (9) the number of *m*-valued digits for all sequences (1), which satisfy the conditions (6) when N=99, equals $4950*4,79\approx23711$. For the verification of accuracy of the received density evaluation p_r following the specially made program in the environment of symbol mathematics Maple, there had been determined the set of *m*-valued digits for different values of *m*. The results of calculations are given in table 2.

Table2

m	n _m	Error δ, %
8	23667	0,19
9	23682	0,12
10	23712	0,004
11	23692	0,08
12	23668	0,18

Number of n_m m-valued digits of the sequence

Obviously, the received density is too small and this stipulates for necessity in further research for the results with high practical value.

Conclusions

1. There had been developed the technique for the determination of the number of the unique recurrent sequences (1), corresponding to conditions (6). The illustration for the determination of the unique recurrent sequences (1), built for some specific case, allowed to receive correlation for calculating the number of unique sequences and total number of their members, belonging to different interval (depending or the range of change of the first two members).

2. There had been suggested the relations for the determination of density of completion of range of natural members with members of sequence (1) in the form of relation of number of all members of the unique sequences, which fall under the set range, to the length of this range.

3. Theoretical evaluation of the above-mentioned density does not exceed $\cdot 10^{-4}$ and under specific conditions does not change during the growth of the varying interval of the first two members of the sequence and corresponding increase of the under research range.

4. Experimental verification by specially composed program in the environment of system of symbol mathematics Maple proved the reliability of conclusions for the specific range of values of the two first members of sequence.

5. Not high value of density stipulates for the density of further research in this direction

REFERENCES

Ватолин Д. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео / Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин. – М.: ДИАЛОГ-МИФИ, 2003. – 384 с.

2. Анисимов А. В. Обратное преобразование Фибоначчи / А. В. Анисимов, Я. П. Рындин, С. Е. Редько // Кибернетика. – 1982. – № 3. – С. 9 – 11.

3. Кшановський О. Д. Арифметичні методи ущільнення цифрової інформації / О. Д. Кшановський, С. В. Тітарчук, В. А. Лужецький // Вісник ВПІ. – 1999. – № 5. – С. 83 – 87.

4. Воробьёв Н. Н. Числа Фибоначчи / Н. Н. Воробьёв. – М.: Наука, 1978. – 144 с.

5. Лужецький В. А. Щільність заповнення ряду натуральних чисел членами лінійних рекурентних послідовностей другого порядку / В. А. Лужецький, В. М. Михалевич, О. В. Михалевич, В. А. Каплун // Вісник Вінницького політехнічного університету. – 2010. – № 4. – С. 41 – 45.

6. Алфутова Н. Б. Алгебра и теория чисел. Сборник задач для математических школ / Н. Б. Алфутова, А. В. Устинов. – М.: МЦНМО, 2002. – 264 с. – ISBN 5-94057-038-0.

Volodymyr Luzhetskyi – Dc.Sc(Eng.)., Professor, Head of Department for Information Protection.

Volodymyr Mykhailevych – Dc.Sc(Eng.), Professor, Head of Department for Higher Mathematics.

Olexiy Mykhailevych – Student.

Valentyna Kaplun – Senior lecturer with the Department of Information Protection. Vinnytsia National Technical University.