Complexity of finite symbolic sequences

geometry of finite functional spaces

Serhi Kozinii

Aston University, department of Engineering and Applied Science, UK

sergey.koziniy@gmail.com

Introduction

Sequence "001 001 001 001" is less complicated than 010 110 101 001. Let's consider mathematical theory which could justify this intuitive statement

Differentiation of sequence of numbers

To explore the complexity of sequence let's define operator $A : M- > M, y = Ax$ by formula $y_i = x_{i+1} - x_i$ and let's define x_{n+1} as x_1 to make our sequence cyclic and easily get y as sequence of length n. In the case of sequences consisting of zeros and ones we can set this operator as graph with 2^n nodes and the each edge in this graph comes from x to Ax

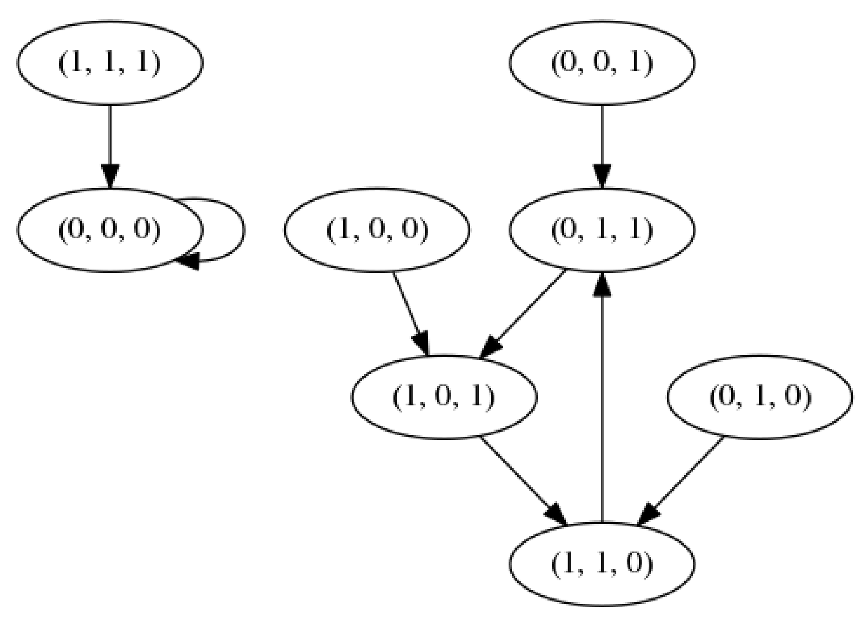

Figure 1: As we can see graph of operator A in case of $n = 3$ contains $2^3 = 8$ edges and two connectivity components. Each of them contains exactly one cycle

Let's designate $(O_m * T_n)$ as cycle equipped with forest of binary trees contains n nodes each. And $A + B$ is just a graph than contains two connectivity components A and B. So the previous graph can be written as

$$(O_1 * T_2) + (O_3 * T_2)$$

Complexity definition

As the measure of complexity of numbers sequence (element of $\mathbb{Z}_2^n$) we will use geometrical properties of A-operator graph and position of node x in this graph. We assume sequence x to be more complex than sequence y if it belongs to connectivity component with bigger cycle. In case of the same cycle length the more complex sequence lies farther from cycle.

Sraightforward computations for different cases of n

Computations of A-operator graphs for different cases of n for $n \leq 1$

n	NoC	Structure
2	1	$(O_1 * T_4)$
3	2	$(O_3 * T_2) + (O_1 * T_2)$
4	1	$(O_1 * T_{16})$
5	2	$(O_{15} * T_2) + (O_1 * T_2)$
6	4	$2(O_6 * T_4) + (O_3 * T_4) + (O_1 * T_4)$
7	10	$9(O_7 * T_2) + (O_1 * T_2)$
8	1	$(O_1 * T_{256})$
9	6	$4(O_{63} * T_2) + (O_3 * T_2) + (O_1 * T_2)$
10	10	$8(O_{30} * T_4) + (O_{15} * T_4) + (O_1 * T_4)$
11	4	$3(O_{341} * T_2) + (O_1 * T_2)$

Here NoC is number of conectivity components.

Complexity of discrete logaritmic function

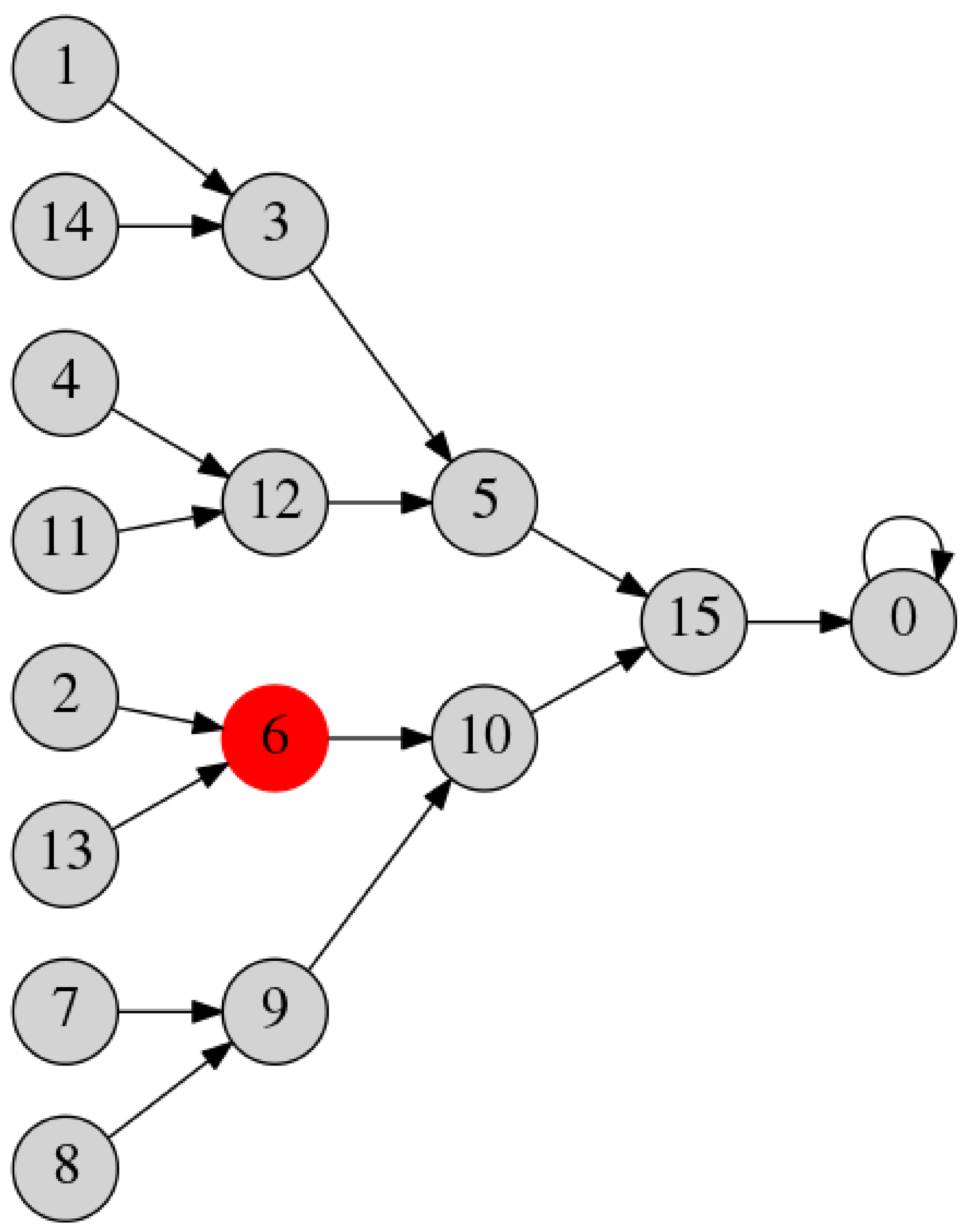

Figure 2: Graph of A operator in case of n = 4. (p = 5, r = 2). Discrette logaritms are: $log_2 1 = 4, log_2 2 = 1, log_2 3 = 3 log_2 4 = 2$ Redusing modulo 2 gives us a secuence $(0, 1, 1, 0)$(binary representation of 6) which is highligted by red in graph. In all cases of n for which graphs were computed such logaritm was related to connectivity component with longest cycle and was on the distance of n or n-1 from this cycle where n is the highest posible distance to the cycle

It was observed that logaritmic function $x(t) = log_r t\,(mod\,p)$ where r is primitive root, and $p = n + 1$ is prime occupies the most, or nearly the most complex point in the A-operator graph. The reasons of such property are unclear, but it probably relates to computational complexity of discrete logarithm which is used in public key cryptography algorithms.

Polynimials

If sequense x is a polynimial of power d or less then $A^{d+1}x = 0$, so x belongs to connectivity component with cycle 0 of period 1. Contrariwise, if sequense belongs to such a component, it's a polynomial.

Polinomial here is sequence generated by function $x(t) = a_1 t^r + ... + a_r t + a_{r+1}$ where $t = 1$, coeficients are rational, ad reulting value is reduced modulo 2

Conclusions

Described theory shows us unexpected properies of elements of finite functional space. Possibly it could be used to estimate complexity of real-world signals.

Forthcoming Research

Researching applicability of the methods based on described theory of complexity to chaotic and stochastic signals.

Acknowledgments

I'm grateful to to Dmitriy Nerukh and Nataly Vaysfeld, my tutors in UK and Ukraine, and all people who guided me through the application process with Erasmus+ programme

References

[1] Vladimir Igorevich Arnold, Complexity of finite sequences of zeros and ones and geometry of finite spaces of functions 1994.