

**Determination of random binary sequence as a combinatorial object.
Calculation of matching fragments in random binary sequences
Filatov O.**

**Определение случайной бинарной последовательности как комбинаторного
объекта. Расчёт совпадающих фрагментов в случайных бинарных
последовательностях
Филатов О. В.**

*Филатов Олег Владимирович / Filatov Oleg Vladimirovich - инженер-программист,
ИТЦ Модуль, г. Москва.*

Аннотация: «Комбинаторика длинных последовательностей» открыла такое «поведение» исследуемых фрагментов в случайных последовательностях, которое начинает напоминать поведение частиц в микромире. Фрагменты последовательности оказывают влияние друг на друга. Фрагменты, как и физические частицы микромира, «чувствуют» когда их исследуют. Дано определение случайной последовательности, использующие явление влияния фрагментов друг на друга. А так же, описано совпадение фрагментов в независимых случайных последовательностях.

Abstract: «Combinatorics long sequences» opened a «behavior» of the test pieces in random sequences, which is beginning to resemble the behavior of particles in microcosm. Fragments of the sequences have an impact on each other. Fragments as physical particles of the microcosm, «feel» when they are exploring. The definition of a random sequence using the phenomenon of fragments of influence on each other. And as described in coincidence fragments of independent random sequences.

Ключевые слова: Комбинаторика длинных последовательностей, составные события, комбинации Пенни, игра Пенни, Пенни, бинарная последовательность, эл, цуга, формула расчёта цуг, случайный бинарный объект, сложность индивидуального объекта, несжимаемые на один.

Keywords: Combinatorics long sequences, composite events, the combination Penny game Penny, Penny, a binary sequence, el, a train, a train of calculation formula, a random binary object, the complexity of the individual object, incompressible one.

Сокращения:

ф.; ф-ла – формула; Эл – элементарное бинарное случайное событие (0; 1);

П-ть – последовательность.

Введение.

В последние годы шли работы по развитию «Комбинаторики длинных бинарных последовательностей». В результате работ найдено новое определение (описание) случайной бинарной последовательности через формулу образующих её цуг. Найдена связь выпадений случайных событий (бросков монеты) с количеством составных событий и цуг в последовательности [1 – 4, 6].

«Комбинаторика длинных последовательностей» открыла такое «поведение» исследуемых фрагментов в случайной бинарной последовательности, которое начинает напоминать поведение частиц в микромире. Фрагменты последовательности оказывают влияние друг на друга. И как физические частицы в микромире, фрагменты «чувствуют» когда их исследуют. Такое «поведение» бинарных фрагментов случайной последовательности бросает вызов устоявшимся интуитивным ожиданиям, развитым за многие годы теорией вероятностей. Например, найдена зависимость последующих составных событий от ранее выпавших составных событий [1, 4]. А так же, найдены правила поиска составных событий, в случайной бинарной п-ти, которые меняют вероятности их обнаружения [9].

Известны пропорции выигрышных комбинаций в парадоксальной игре Пенни, но только применение идей комбинаторики длинных последовательностей позволило количественно рассчитать комбинации Пенни как функцию числа бросков монеты [8].

Рассчитать «сложность индивидуального случайного бинарного объекта» [7].

Формулы «Комбинаторики длинных последовательностей» работают с п-ми значительных длин, например: 10^{10} бинарных событий.

«Комбинаторика длинных последовательностей», вероятно, дальше будет развиваться, изучая не только индивидуальные п-ти но и их взаимодействия друг с другом. В статье исследованы *признаки* независимого, друг от друга, образования случайных бинарных последовательностей (файлов). Они позволяют отделять структуры с параллельным, несвязанным друг с другом, образованием одинаковой информации, от структур в которых информация последовательно мигрировала (передавалась по наследству с разной степенью искажения) от первоначальной родительской структуры к поколениям-потомкам. В статье представлены формулы распознавания спонтанного, независимого возникновения

одинаковой информации в нескольких не взаимодействующих друг с другом процессах. В социальном мире имеются области спонтанного, независимого, образования идентичных индивидуальных объектов. К такому проявлению относится феномен близнецов и феномен одновременности появления открытий в разных географических местах Земли.

Интересны психологические аспекты спонтанного совпадения информации, например феномен «электронного голоса» или «шёпота духов», «спиритического сеанса», «сигнала внеземной цивилизации». Когда спонтанное возникновение физического сигнала во внешнем мире совпадает с пониманием этого сигнала в сознании человека в качестве осмысленных слов.

Основная часть

Сжатый архиватором файл находится в состоянии наивысшей сложности. Его нельзя сжать в меньшее по размеру «зерно», из которого снова можно «вырастить» (развернуть) этот файл. Такие файлы называют: «несжимаемые на один» файлы¹. Амплитудно-частотная характеристика «несжимаемых на один» файлов стремится к амплитудно-частотной характеристике случайной бинарной п-ти [6, 7].

Число цуг и число полувольт (составных событий) в цугах случайной бинарной последовательности зависит от числа элементарных событий п-ти N (бросков монеты). Введём определение случайной бинарной п-ти через понятия «Комбинаторики длинных п-й».

Комбинаторное определение случайной бинарной последовательности

Случайной бинарной п-тью называется п-ть для которой выполняются пункты:

1) Число цуг ${}^n C_w(N)$ в п-ти зависит от N (число событий п-ти) и численно стремится к величине рассчитываемой по ф.1 [1 - 4].

2) Число нулей и единиц последовательности находится в динамическом равновесии, их отношение стремится к единице с ростом N.

3) Для любой подпоследовательности образуемой из событий п-ти в «тёмную» (величины отбираемых элементарных событий (эл) неизвестны и каждый эл отбирается один раз), по любым правилам или без всяких правил, выполняются пункты 1 и 2.

Формула расчёта цуг ${}^n C_w$ моды n, [1 - 4]:

$${}^n C_w = \frac{(2^n - 1)^2}{2^{n(w+2)+1}} N \quad \text{Ф. 1}$$

Где, n – число элементарных событий образующих составные события в цуге; w – число полувольт (составных событий) в цуге; N – число элементарных событий (подбрасываний монеты) в бинарной последовательности.

Пример. Рассмотрим три последовательности: $N_1=1*10^6$ элементарных событий (эл), $N_2=1*10^7$ эл, $N_3=1*10^8$ эл. Рассчитаем для них по ф.1 мат. ожидание цуг:

$${}^{n=1} C_{w=8}(N) = \langle \underline{01010101} \rangle + \langle \underline{10101010} \rangle.$$

$$\text{Подставим значения } n=1 \text{ и } w=8 \text{ в ф.1: } {}^{n=1} C_{w=8}(N) = \frac{(2^1-1)^2}{2^{1 \cdot (8+2)+1}} N = \frac{N}{2^{11}}.$$

$$\text{Получим: } {}^1 C_8(N_1) = 488; {}^1 C_8(N_2) = 4883; {}^1 C_8(N_3) = 48828.$$

В [1] приведён алгоритм генерация идеальной бинарной случайной последовательности (ИБСП) при помощи ф.1. Для облегчения понимания работы алгоритма, из [1] был исключён блок контроля численного равенства нулей и единиц.

Сложность индивидуального случайного бинарного объекта

Индивидуальный случайный бинарный объект является частью (извлечённой или не извлечённой) случайной бинарной п-ти. Определимся, что будем сравнивать друг с другом сложности индивидуальных объектов обладающие одинаковыми длинами.

Сложность индивидуального объекта зависит от распространённости в природе группы (класса), к которой он принадлежит.

Простых объектов много. Сложных объектов мало.

Группы с большим числом членов отнесём к объектам малой сложности (простых объектов много).

Группы с малым числом членов содержат объекты большой сложности (сложных объектов мало).

¹ Используя наработки «Комбинаторики длинных последовательностей» автору этой статьи удалось сжать «несжимаемый» файл на «ритуальный» один байт, и показать существование подмножества из бесконечного числа сжатых п-ей, полученных из считавшихся ранее «несжимаемых на один п-ей»: «Проблемы современной науки и образования» № 9, 12, 2015 г.

Критерием сортировки объектов по группам является число инверсий внутри объекта. Объекты, обладающие одинаковым числом инверсий, относятся к одной группе. Пример объектов одной группы: «011», «001», «110», «100» (все объекты содержат по одной внутренней инверсии).

И так, сложность индивидуального объекта определяется принадлежностью к группе с некоторой частотой встреч в природе, а принадлежность к группе определяется числом инверсных переходов в объекте [7].

Индивидуальные последовательности (объекты) составляющую одну группу имеют одинаковую сложность.

Рассмотрим распределение индивидуальных последовательностей по степеням сложности на примере 1.

Пример 1. Даны четыре п-ти каждая, из которых состоит из восьми элементарных событий (эл): «00000000»; «11111111»; «01010101»; «10101010». Требуется разделить объекты (индивидуальные п-ти) по степени сложности.

Таблица 0. Сложность бинарного индивидуального объекта из восьми элементов

<i>n</i>	<i>i</i> инверсий в слове	0	1	2	3	4	5	6	7
8	${}^n C_0$ - сложность слова	2	14,30	42,48	70,60	70,70	42,78	14,85	2,90
	${}^n C_0^*$ - сложность* слова	0,30	1,15	1,636	1,845	1,846	1,640	1,17	0,36
	${}^n_i p$	$7,81^{-3}$	$5,47^{-2}$	0,16	0,27	0,27	0,16	$5,47^{-2}$	$7,81^{-3}$
	$I(C_i^{n=8})$ - вариантов слов	2	14	42	70	70	42	14	2

Так в примере 1, объекты делятся на две группы. Группа из объектов $I(C_{i=0}^{n=8})$, в которых ноль инверсий ($i=0$): «00000000»; «11111111» - два объекта этой группы обозначены в нижней строке $I(C_i^{n=8})$, в столбце «0». Группа из объектов $I(C_{i=7}^{n=8})$, в которых семь инверсий ($i=7$): «01010101»; «10101010» - два объекта этой группы обозначены в столбце «7».

В остальных столбцах (i : «1» - «6») посчитано, сколько существует объектов с указанным числом инверсий внутри них. Например, существует $I(C_{i=2}^{n=8})=42$ различных объекта с двумя инверсиями ($i=2$) внутри и одинаковой сложностью ${}^n C_0$ для любого объекта этой группы (класса). И существует $I(C_{i=5}^{n=8})=42$ объекта с пятью инверсиями ($i=5$) внутри, с одинаковой сложностью ${}^n C_0$ для любого объекта этой группы (класса).

В слове длиной n событий можно комбинаторным путём разместить $I(C_i^n)$ слов, в каждом слове будет по i инверсии. Формула ф.1.1 для расчета числа инверсионных слов (объектов) $I(C_i^n)$ по i инверсии в каждом слове [1, 7]:

$$I(C_i^n) = \frac{2 \cdot (n-1)!}{i! \cdot (n-1-i)!}; \text{ где } n = 1,2,3, \dots; i \leq n-1 \quad \text{Ф. 1.1}$$

$I(C_i^n)$ – число комбинаций содержащих внутри себя по i инверсий; n – длина слова; i – число инверсий внутри слова.

В таблице 0 в строке « $I(C_i^{n=8})$ - вариантов слов» представлены рассчитанные по ф.1.1 количества комбинаций для каждого возможного числа инверсий i в восьмиразрядном слове. Как видно, реже всего встречаются в природе, по два раза, комбинации $i = 0$ («00000000»; «11111111») и $i = 7$ («01010101»; «10101010»).

Вероятность выпадения объекта группы (класса) $I(C_i^n)$ определяется по ф.1.2:

$${}_i p = \frac{I(C_i^n)}{2^n} \quad \text{Ф. 1.2}$$

В таблицу 0 внесены вероятности встреч в природе групп (классов) $I(C_i^n)$ для восьми разрядного слова.

Колмогоров предложил считать, что комбинации $i = 0$ («00000000»; «11111111») и $i = 7$ («01010101»; «10101010») имеют малую сложность. Но как видно из таблицы 0, как раз эти комбинации принадлежат самым сложным классам. Сложность класса определяется числом его членов. И сложность индивидуального объекта зависит от распространённости в природе его группы (класса), самые малочисленные классы содержат самые сложные индивидуальные объекты. Но по Колмогорову это самые простые объекты. Более сложные объекты по Колмогорову обладают тремя и четырьмя инверсиями внутри себя. В естественно природной классификации, количественной, наоборот, эти

объекты с тремя и четырьмя инверсиями простые (бактерий много – они простые, слонов мало – они сложные).

Любимый вопрос лекторов к аудитории о том, почему людям кажется, что последовательность «10101010101010» встречается редко в естественно природной классификации, имеет простой ответ, что она принадлежит к редко встречаемому классу (от этого и редко встречается).

Возникает противоречие между определением сложности по Колмогорову и естественно природной классификации сложности объектов. То, что по Колмогорову сложно, в естественно природной классификации просто. И наоборот.

В свою очередь, академик Арнольд открыл, что инверсия элементарных событий – это более сложное состояние, чем их повтор [10].

Попробуем достичь в формулах ф.1.2 (а, б) компромисс в понятии сложности индивидуального объекта n_iCo между: колмогоровской сложностью, естественно природной классификацией сложности и выводами академика Арнольда, о том, что инверсия элементарных событий – это более сложное состояние, чем их повтор [10].

Расчёт величины демонстрационной сложности индивидуального объекта n_iCo , по ф.1.2(а, б):

$${}^n_iCo = I(C_i^n) + \lg_{10}(1 + i) \quad \text{Ф. 1.2 (а)}$$

$${}^n_iCo^* = \lg_{10} \left(I(C_i^n) + \lg_{10} \left(1 + \frac{i}{n} \right) \right) \quad \text{Ф. 1.2 (б)}$$

Где: n_iCo - Сложность индивидуального Объекта (the Complexity of the Object); n - длина слова (число элементарных событий, эл); i - инверсий в слове; $I(C_i^n)$ – ф.1.1, число различных вариантов слов содержащих внутри себя по i - инверсии.

Так как сложность является абстрактной величиной, то её масштаб выбирается произвольно, что и отражает ${}^n_iCo^*$ из ф.1.2 (б).

Расчёт значений демонстрационной сложности бинарного индивидуального объекта (слова) из восьми элементов, по ф.1.2(а) и ф.1.2(б), приведён в таблице 0.

Формулы 1.2 (а, б) состоят из двух слагаемых. Формулы расчёта сложности 1.2 (а, б) с помощью члена $I(C_i^n)$ подогнаны под утверждения Колмогорова, что события: «00..000»; «11..11»; «0101..»; «1010..» являются простыми, а более сложными событиями являются сложно-описываемые последовательности. Слагаемое $I(C_i^n)$ вносит основной вклад в n_iCo и обеспечивает колмогоровскую форму графику [7] для распределения сложностей.

Действительно, наибольшими численными значениями (70,60; 1,845 и 70,70; 1,846) обладают самые сложные по Колмогорову классы индивидуальных объектов. А наименьшими значениями (2; 2,90 и 0,30; 0,36) обладают самые простые по Колмогорову индивидуальные классы.

Естественно-природная классификация сложности $I(C_i^n)$, ф.1.1 присутствует в формулах 1.2 (а, б) в виде основного члена.

Второе слагаемое, $\lg_{10}(1 + \dots)$, введено для стыковки с выводами академика Арнольда, о том, что инверсия элементарных событий – это более сложное состояние, чем повтор². Поэтому, сложность инверсионных последовательностей: «0101..»; «1010..» была искусственно увеличена для событий с большим числом инверсий.

То, что пути nC_w («0101..»; «1010..»), ф.1, более сложны, чем последовательные составные события nS («00..000»; «11..11»), [1, 2] подтверждается тем, что комбинаторика длинных последовательностей относит их к разным структурным уровням (составные события nS образуют пути nC_w – сложные объекты nC_w собраны из простых объектов nS), рис.1, 2 в [4, 1]. А так же тем, что при последовательном поиске в n-ти пути nC_w встречаются реже, чем составные события nS такой же длины (в элах). Поэтому, второе слагаемое $\lg_{10}(1 + \dots)$ должно иметь более сложное поведение, чем на данный момент дано в ф.1.2(а, б).

Признаки независимости образования случайных бинарных последовательностей.

² Филатов О. В., Филатов И. О., Статья «Эффект Арнольда – Филатова. Золотое, серебряное сечения.

Альтернативная запись бесконечно сложной последовательности. Аргументация по фундаментальности «Потоковой теории», «Журнал научных публикаций аспирантов и докторантов», № 12, 2014.

В макромире проявлением несвязных между собой совпадений можно считать людей-близнецов одинакового возраста. А так же, одновременность научных открытий разными, технических разработок, не связанными между собой учёными.

В этом подразделе рассматриваются никак не связанные между собой бинарные последовательности (файлы). В этих файлах ищутся спонтанно возникающие объекты (фрагменты) имеющие одинаковые координаты (удаление от начала файла) и содержащие одинаковую информацию.

Под совпадающими данными (фрагментами) в двух и более файлах, п-тях, понимаются биты, имеющие одинаковую координату (удаление) от начала файла и одинаковые значения. Примеры пар совпадающих фрагментов в двух файлах (совпадающие фрагменты подчёркнуты): L1: (101, 000); L2: (1011, 0010); L2: (1001, 0000); L3: (11101, 01100); L4: (011111, 111110).

В приведённых примерах буква L обозначает длину совпадающих (подчёркнутых) участков в файлах. В таблице 1, в столбце «L» указаны длины совпадающих участков.

В таблице 1, в столбце «F=2» представлены числа совпадающих фрагментов в двух файлах со случайными данными. Длина каждого из файлов в элементарных бинарных событиях (элах) равно $2 \cdot 10^7$.

Так, например, участки длиной L=2 встречались 1249838 раз. В качестве примера приведём возможные комбинации одинаковых участков длины L2: (00; 00); (01; 01); (10; 10); (11; 11).

В таблицу 1, в столбец «F=3» внесены численности случайным образом совпавших участков длин L одновременно в трёх файлах. Например, биты «11» одновременно находятся в позициях 33 и 34 в трёх файлах. Всего в трёх файлах (F=3) было найдено 702856 идентичных участков длиной два (L=2) одновременно на одних и тех же позициях.

В столбцах «F=4» и «F=5» внесены численности случайным образом совпавших участков длин L в четырёх и пяти файлах одновременно.

Таблица 1. «Число случайных совпадений подряд в L позициях F файлов»

N = 20 000 000		F=2	F=3	F=4	F=5	
1	L	$L S_N^F$ (экспериментальные данные)				
2	1	2501531	2812068	1912002	1098305	
	2	1249838	702856	239680	68472	
	3	624386	176035	29944	4315	
	4	313120	44176	3744	290	
	5	155981	10849	468	20	
	6	77965	2768	49	1	
	7	38874	684	5		
	8	19586	160	1		
	9	9729	38			
	10	4952	8			
	11	2520	3			
	12	1149	2			
				
22	3					
23	1					
3	$S_N^{OF} = \sum_{L=1}^{\infty} L S_N^F$	5000802	3749647	2185893	1171403	
4	$\Phi 2: S_N^{OF} = \frac{2^{F-1} - 1}{2^{2(F-1)}} N$	5000000	3750000	2187500	1171875	
5	$eL_N^{OF} = \sum_{L=1}^{\infty} L \cdot L S_N^F; \left(eL_N^{OF} = \frac{N}{2^{F-1}} \right)$	9998343 (10000000)	4999989 (5000000)	2498847 (2500000)	1249460 (1250000)	
6	$\Phi 5: S_N^F = \frac{(2^{F-1} - 1)^2}{2^{3(F-1)}} N$	2500000	2812500	1914062,5	1098632,8	
7	$\Phi 6: L S_N^F = \frac{(2^{F-1} - 1)^2}{2^{(F-1)(L+2)}} N$	L=1 L=2 L=3 L=4	2500000 1250000 625000 312500	2812500 703125 175781 43945	1914062 239257 29907 3738	1098632 68664 4291 268

В таблице 1, в строку 4 внесены теоретически рассчитанные суммы S_N^{OF} всех перекрывающихся (совпадающих) участков не зависимо от их длин (всех длин), ф.2:

$$S_N^{OF} = \frac{2^{F-1} - 1}{2^{2(F-1)}} N \quad \Phi. 2$$

Где: OF и F – число перекрывающихся друг другом файлов (Overlaped Files); N – число элементарных событий (эл) в файлах (п-ях).

В таблице 1, в строке 3 суммы S_N^{OF} по столбцам «F=2» - «F=5» для соответствующих значений $L S_N^F$ в столбцах строки 2.

Поскольку нас изначально интересует el_N^{OF} - количество информации (бит) угаданной (совпавшей) лишь в одном файле из F , то в строке 5 посчитаны совпадения бит относительно одного файла. Хотя совпало по L бит в F независимых файлах ($F \cdot L$ элементарных события), учитывается только L из них, ф.3:

$$el_N^{OF} = \sum_{L=1}^{\infty} L \cdot L S_N^F = \frac{N}{2^{F-1}} = N \cdot 2^{(1-F)} \quad \Phi. 3$$

Где, F – число файлов, в которых ищется совпадение информационных участков (индивидуальные бинарные объекты).

Отношение числа угаданных бит el_N^{OF} к числу всех участков S_N^{OF} (без относительно длины каждого участка) дано в ф.4:

$$\frac{el_N^{OF}}{S_N^{OF}} = \frac{2^{F-1}}{2^{F-1} - 1} = \frac{1}{1 - 2^{1-F}} \quad \Phi. 4$$

Из ф.4 получаем ф.2: $S_N^{OF} = \frac{2^{F-1}-1}{2^{F-1}} \cdot el_N^{OF} = \frac{2^{F-1}-1}{2^{F-1}} \cdot \frac{N}{2^{F-1}} = \frac{2^{F-1}-1}{2^{2(F-1)}} N$

Зная формулу расчёта S_N^{OF} (ф.4) получаем формулу расчёта для ${}^1S_N^F$ - единичных совпадений (отдельных единичных угадываний): ${}^1S_N^F = S_N^{OF} \cdot \frac{2^{F-1}-1}{2^{F-1}} = \frac{2^{F-1}-1}{2^{2(F-1)}} N \cdot \frac{2^{F-1}-1}{2^{F-1}} = \frac{(2^{F-1}-1)^2}{2^{3(F-1)}} N$; ф.5:

$${}^1S_N^F = \frac{(2^{F-1} - 1)^2}{2^{3(F-1)}} N \quad \Phi. 5$$

Рассчитанные по ф.5 значения даны в строке 6. Они хорошо совпадают с полученными экспериментальным путём верхними значениями, в строке 1 (2501531; 2812068; 1912002; 1098305). Так в пяти независимых файлах (F=5) в эксперименте было найдено 1098305 одинарных (L=1) совпадений, а по ф.5 получается 1098632 совпадения (строка 6, столбец F=5).

Зная ${}^1S_N^F$ и замечая, что с ростом длины L на единицу отношение $\frac{{}^L S_N^F}{{}^{L+1} S_N^F} = 2^{F-1}$ легко перейти к расчёту любого ${}^L S_N^F$, (экспериментальные данные даны в строке 2), ф.6:

$${}^L S_N^F = \frac{{}^1 S_N^F}{2^{(L-1)(F-1)}} = \frac{(2^{F-1} - 1)^2}{2^{(F-1)(L+2)}} N \quad \Phi. 6$$

Заметим, что при $F=1$ объект сравнения отсутствует, и ${}^L S_N^F = \frac{0}{1} \cdot N = 0$

В строке 7 таблицы даны рассчитанные по ф.6 мат. ожидания совпадений фрагментов длины L в F фалах. Как видно, рассчитанные мат. ожидания из строки 7 хорошо совпадают с соответствующими экспериментально найденными величинами из строки 2 таблицы.

Сравнение с составными событиями в одиночном файле (F=1).

По ф.6, для двух независимых случайных файлов число совпадений ${}^L S_N^2$ длины L будет: ${}^L S_N^2 = \frac{N}{2^{L+2}}$. Сравнивая полученное отношение ${}^L S_N^2$ с числом составных событий ${}^L S_N^1$ [1, 2] длины L в отдельном случайном файле (F=1): ${}^L S_N^1 = \frac{N}{2^{L+1}}$, видим, что: $2 \cdot {}^L S_N^2 = {}^L S_N^1$.

Отношение формулы расчёта числа составных событий ${}^L S_N^1$ для отдельного случайного файла (п-ти) [1, 2], и формулы ф.6 (при F=2), для расчёта числа совпадающих участком длины L равно двум, ф.7:

$$\frac{{}^L S_N}{{}^L S_N^2} = \frac{N}{2^{L+1}} \cdot \frac{N}{2^{L+2}} = 2 \quad \text{Ф. 7}$$

Для двух независимых случайных п-ей (файлов), F=2, под одинаковыми адресами от начала файлов находятся одинаковые фрагменты, например: «111» и «111», «1010» и «1010», «110100» и «110100», «0» и «0». Средняя длина одного такого фрагмента равна средней длине составного события в одинарной бинарной п-ти F=1, и равна двум. Действительно, разделим число пар совпавших элементарных событий (эл), ф.3, на число всех совпадающих пар любой длины S_N^{OF} , ф.2, получим ф.8:

$$\frac{el_N^{OF}}{S_N^{OF}} = \frac{N}{2^{F-1}} \cdot \frac{2^{F-1} - 1}{2^{2(F-1)}} N = \frac{2^{F-1}}{2^{F-1} - 1} \quad \text{Ф. 8}$$

Подставляем F=2 в ф.8: $\frac{2^{2-1}}{2^{2-1}-1} = 2$. Из этого следует, что средняя длина не совпадающей пары фрагментов то же равна двум.

Обсуждения

Алгоритм генерации идеальной бинарной псевдослучайной последовательности (ИБПС) часто добавляет на очередном шаге своей работы, к создаваемой ИБПС, не одно, а группу эл (элементарных случайных событий). Примеры добавляемых на очередном шаге генерации эл: «0», «1», «00», «10101», «11001100», «000000», ...

Алгоритм создания [1] идеальной псевдослучайной последовательности управляет комбинационным размещением цуговых структур во времени (или в одномерном пространстве на оси N). Изменением, настройкой параметров создающего алгоритма достигается изменение в порядке следования цуг в п-ях, что позволяет создать не пустое множество псевдослучайных п-тей отвечающих всем критериям случайности. Причём, созданные п-ти являлись «не сжимаемыми на один» традиционными методами сжатия.

Отличительной особенностью алгоритма создания [1] является добавление псевдослучайных событий группами по несколько элементарных событий зараз. Работа алгоритма создания [1] фактически является комбинаторным размещением цуговых структур. Комбинаторное образование множества индивидуальных идеальных бинарных п-тей означает, что «Потоковая теория» относится больше к комбинаторному разделу, чем к вероятностному. Для отражения новой идентификации «Потоковая теория» получила рабочие название «Комбинаторика Длинных Последовательностей» («код пи»).

В рамках работ 2014 года была опробована попытка восстановления текста после превращения его гласных букв в единицы, а согласных букв в нули. Идея восстановления текста на тот момент оказалась не реализуемой. Но в результате той работы была введена система координат для разных разговорных человеческих языков мира [5]. В этой системе координат все языки сравнивались с естественным шумовым (хаотическим) уровнем. Сейчас появилось много работ по расшифровке языков животных. У автора статьи возникло ощущение, что учёными, возможно, будут созданы и алфавиты для этих языков. Было бы интересно применить методику [5] для языков животных.

В рамках работ по поиску применения комбинаторики длинных последовательностей (кода пи) к объяснениям лженаучных наблюдений (явлений), так же были рассчитаны математические ожидания появления тех или иных фраз в шумовой последовательности – феномен «Электронного голоса». Оказалось, что можно рассчитать число встреч фразы в п-ти длины N в таком феномене как «электронный голос», и число встреч фразы, на длине N, зависит от разговорного языка, на котором хотят услышать фразу. В сборнике «Потоковая теория: из сайта в книгу». Москва, «Век информации», 2014, приводится пример пояса в потоковой последовательности прямоугольного объекта содержащего букву «Я» и просто монотонно заполненного либо «0», либо «1» прямоугольного квадрата. Оказалось, что слово «Я» встречается чаще «чёрного квадрата» (это объясняется эффектами поисковых шаблонов Пенни [8]).

Отличие поисковых шаблонов Пенни от индивидуальных объектов.

Совсем другой случай (задачу) представляет указание внутри случайной бинарной п-ти некоего фрагмента, с предложением посчитать число последовательных вхождений этого фрагмента в п-ть. Этот случай принадлежит к классу задач по расчёту шаблонов Пенни. Решение задач такого рода разобраны в [8]. В задаче о последовательном поиске фрагмента внутри бинарной последовательности (игра Пенни), искомый фрагмент не является индивидуальным объектом независимым от своего окружения. Он

является шаблоном из парадоксальной игры Пенни [8], «интерферирующим» объектом, так как своим существованием влияет на последующие за ним объекты: «возмущает» окружающую среду, «распространяет поле сил» на последующие объекты и т.д. Затруднительно придумать название явлению для взаимодействия соседних областей в случайной бинарной п-ти, считавшейся ранее эталоном независимости, примером полного отсутствия причинно-следственных связей между образующими её событиями. Может быть, эти воздействия спектров (фрагментов п-ти) подчиняются Эйнштейновскому принципу локальности и в этом явлении происходит объединение математики и материи, и случайная п-ть показывает такие же свойства неопределённости, как и вероятностная неопределённость микромира?

Противоречие между частотой встреч объектов в природе и колмогоровской сложностью.

Известно определение сложности природных объектов как функция частоты их встреч в природе: простые объекты встречается часто (водород), сложные объекты встречается редко (золото). Этот принцип, сложность – частота, противоречит понятию колмогоровской сложности, так как объекты малой колмогоровской сложности: «00..000»; «11..11»; «0101..»; «1010..» принадлежат к самым редким инверсным классам $I(C_i^n)$ индивидуальных случайных объектов (по два объекта в каждом (любом) инверсионном классе), таблица 0. Более сложные объекты, по Колмогорову, обладающие тремя и четырьмя инверсиями, наиболее распространены (по 70 объектов в каждой инверсионном классе), для случая в таблице 0 (что противоречит естественной природной частотной сложности).

Выводы

1) Любая случайная бинарная последовательность (достаточной длины) является объединением цуг (мат. ожидание которых рассчитывается по ф.1), что позволяет сконструировать «Идеальную бинарную псевдослучайную п-ть». Алгоритм создания которой основан на ф. 1 представлен в [1].

2) Дан расчёт величины демонстрационный сложности индивидуального бинарного случайного объекта, ф.1.2 (а, б). Введена формула расчёта вероятности выпадений классов случайных объектов, ф.1.2.

3) Частота встреч объектов в природе противоречит понятию колмогоровской сложности, так как объекты малой колмогоровской сложности: «00..000»; «11..11»; «0101..»; «1010..» принадлежат к самым редким, а значит сложным, природным классам.

4) Признаком не связанности двух и более последовательностей (файлов) друг с другом являются описываемые в статье формульные отношения для совпадающих участков последовательностей (файлов). Приведены формулы, позволяющие рассчитывать численность случайно совпадающих фрагментов в независимых последовательностях, как функцию величин: N – длины каждой последовательности, F – числа последовательностей (файлов), L – числа бит (длин) случайно совпадающих фрагментов.

Литература

1. *Филатов О. В., Филатов И. О.* «Закономерность в выпадении монет – закон потоковой последовательности». Германия, Издательский Дом: LAPLAMBERT Academic Publishing, 2015, с. 268.
2. *Филатов О. В., Филатов И. О.* статья «О закономерностях структуры бинарной последовательности», «Журнал научных публикаций аспирантов и докторантов», № 5, 2014.
3. *Филатов О. В., Филатов И. О.* статья «О закономерностях структуры бинарной последовательности (продолжение)», «Журнал научных публикаций аспирантов и докторантов», № 6, 2014.
4. *Филатов О. В., Филатов И. О.* статья «О закономерностях структуры бинарной последовательности (продолжение 2)», «Журнал научных публикаций аспирантов и докторантов», № 7, 2014.
5. *Филатов О. В.* Статья «Методика поиска родства языков по чередованию гласных и согласных букв в письменных источниках», «Журнал научных публикаций аспирантов и докторантов», № 9, 2014.
6. *Филатов О. В.* статья «Теорема «О амплитудно-частотной характеристике идеальной бинарной случайной последовательности», «Проблемы современной науки и образования», № 1 (31), 2015 г.
7. *Филатов О. В.* статья «Числовая оценка Колмогоровской сложности. Определение вероятности через смену событий», «Проблемы современной науки и образования», № 8 (38), 2015 г.
8. *Филатов О. В.* статья «Расчёт численностей поисковых шаблонов в парадоксе Пенни», «Проблемы современной науки и образования», № 11 (41), 2015 г.
9. *Филатов О. В.* статья «Описание схем управления вероятностью выпадения независимых составных событий», «Проблемы современной науки и образования», № 2 (44), 2016 г.
10. *Владимир Игоревич Арнольд*, Математический институт им. В. А. Стеклова, Москва. Публичная лекция 13 мая 2006 г. для школьников и студентов.