Randomness as Absence of Symmetry

Gideon Samid
Department of Electrical Engineering and Computer Science
Case Western Reserve University, Cleveland, OH
BitMint, LLC
Gideon@BitMint.com

THE 17TH INTERNATIONAL CONFERENCE ON INFORMATION & KNOWLEDGE ENGINEERING

(IKE'18: JULY 30 - AUGUST 2, 2018, LAS VEGAS, USA)

Abstract: **Randomness is an intensely intuitive notion which despite its serving as a central pillar for modern cryptography, is not yet well defined. There is no shortage of definitions, to be sure, but there is no objective way to rank them. Which is a situation inviting more definition candidates, given how critical it is to use 'high quality randomness' as a 'high octane' fuel for cryptographic engines. We propose a randomness metric applicable to any string of symbols of any size, resulting in a linear scale ranging from 'zero randomness' to 'full randomness'. Initial examinations indicate that such a metric may have meaningful cryptanalytic applications since ciphers have a characteristic randomization action, allowing the randomized ciphertext to point to likely plaintexts (sorted by their randomization status). More profoundly a good randomization metric will serve as a knowledge acquisition gauge, with applications beyond cryptography. The proposed metric bears some analogy to the decomposition of integers into a list of primes, rating their smoothness by their defining primes. Similarly an arbitrary bit string of any length is viewed as concatenation of prime strings, such that the number of these prime strings reflects the randomness of the concatenated string. Prime strings are 'zero randomness' strings, in as much as they have a prime symmetry. A slew of applications is under investigation.**

I. INTRODUCTION

Randomness may be viewed as the starting point of mathematics, which is the handling of order, the absence of which is the intuitive grasp of randomness. Being then the boundary of mathematics, it cannot very well be defined within its realm. Indeed the stubborn resistance to its definition is its intrinsic characteristics.

Mathematics in many a field is using a working definition of randomness, and builds upon it. It stands to reason then, that additional work in the area of definition might prove promising, which is the motivation for this article.

We will offer below a quick review of the various approaches to randomness, and how they leave room for different takes.

A. Brief Review of Approaches to Randomness

One may discern three points of view for randomness: the writer's point of view, the reader's point of view, and the per-se point of view. The writer is the source that generates a random sequence. This view focuses on the generation process. Accordingly, a contraption like the one manufactured by IDQ in Geneva [7] where a photon is shot towards a slanted half-way mirror and has 50% chance to go through, and 50% chance to bounce off, is considered a source of quantum-grade randomness. Here the randomness claim is based on the state of the art of physics, not on mathematics.

By contrast, the reader's point of view is taken through the ability of an examiner of a random sequence to draw conclusions, to discern pattern, to use for some added predictive value. To the extent that a series of random bits does not teach its reader any useful knowledge, it is considered properly random. This approach hinges on the deductive powers of the reader, not on mathematical rigor.

The third approach is the per-se idea: to ignore who writes the examined randomness, and similarly to ignore what one can learn from it, but rather examine the data per se, and issue a measure of its randomness. This approach is dominated by Kolmogorov test: a string x that can be generated via a general purpose Turing machine from an input y, such that $|y| < |x|$, (y is shorter than x), is not random. It identifies then as random a string z that cannot be generated from a shorter one. This definition drives ambiguity to another corner: how to prove that no shorter string will generate a given string. Also, this definition brings up a question of the relevant alphabet. A string per se does not identify the alphabet from which it is drawn, and different alphabets may give different results.

Both writer and reader approaches resort to probability calculus. The central idea is that every letter in a random

string is independently determined and has a 1/b chance to be of a given identity (where b is the number of letters in the relevant alphabet). This assumption leads to numerous possible randomness criteria. Any string of letters s, may be judged as random if any given proper subset, p ⊂ s, appears in s, "close enough" to its expected value. Only that this is impossible for all possible subsets p.

II. String Symmetry

Let s by a string comprised of n ordered letters: $l_0, l_1,.....l_{n-1}$. We will regard s as 'self symmetric' ,or simply 'symmetric' for the case where n=1. Such that s='A', s='&', s='π' and s='0' are all symmetric strings. For n >1 if $l_i \neq l_j$ for all i,j=0,1,2,...n-1 then s will be regarded as 'asymmetric'.

Let s' = $R_t(s)$ represent a right shift of t rounds of s, such that:

$$l'_i = l_{i-t}$$

for i=0,1,2...(n-1), where s'=$l'_0, l'_1,....l'_{n-1}$, and where for:

$$i = k \bmod n,\ l_i = l_k.$$

If s' = s then s will be regarded as a symmetric string of degree t, or simply 'symmetric' string.

Illustration: s= 'AAAA' is symmetric with respect to t=1,2,....; s='ABABAB' is symmetric of degree t=2, and s='ABCABCABCABC' is symmetric with respect to t=3,6. By contrast s='ABABABABABABABABZ' is asymmetric.

Lemma 1: if s is symmetric of degree t then:

$$l_i = l_{i+t}$$

for i=0,1,2..(n-1).

Proof: per the definition of symmetry $l'_{i+t} = l_i$, and since s'=s, $l'_{i+t} = l_{i+t}$, so we have $l_i = l_{i+t}$

Lemma 2: if a string s is symmetric of degree t under a right shift, it is also symmetric under a left shift.

Proof: if $s'=R_t^{right}(s)$, then $s=R_t^{left}(s')$ and vice versa.

Lemma 3: if s is symmetric of degree t, then it is symmetric of degrees *2t, 3t,....kt*, where *n-kt* ≥ *0*, k is a positive integer.

Definition: s will be regarded as symmetric of basic degree t*, if s is symmetric of degree t*, and there exists a value i=0,1,2...(n-1) such that:

$$l_i \neq l_{i+t^*-1}$$

In other words, s is not symmetric at degree t < t*.

Lemma 4: if s is symmetric with basic degree t*, then t* | n.

Proof: Suppose t* does not divide n. We can then write n-kt* < t* for some integer k. Since t* is the smallest degree of symmetry, then the last segment will be comprised of less than t* letters, and will not be able to replace the t* letters that shifted from the first segment.. We must therefore conclude that t* divides n.

We concern ourselves now with the number of distinct symmetric strings defined over a string comprised of n letters drawn from an alphabet of m letters: $l_1, l_2,....l_m$

Clearly, the m strings $\{l_i\}^n$ for i=1,2,...m are symmetrical with degree t*=1. For n prime, there is no t*>1 with respect to which any string will be symmetric, as indicated by lemma 4. So for n prime the number of constituent shift-symmetric strings is m. We shall define the symmetry ratio, σ, of a string s of size n as the ratio between the number of symmetric strings to all possible strings:

$$\sigma(s \mid |s|=n \ \&\ \text{n-prime}) = m/2^n$$

and clearly:

$$\text{Lim } \sigma(s \mid |s|=n \ \&\text{n-prime})_{\ n\to\infty} = 0$$

If n is composite of the form $n = p_1^{q_1} p_2^{q_2}....p_w^{q_w}$, where $p_1, p_2,... p_w$ are the prime factors of n, by rising order $p_1 < p_2 < ... p_w$, then we can count symmetric strings of degrees $p_i, 2*p_i,....q_i * p_i$ for w>1, and we can count symmetric strings of degrees $p_i, 2*p_i,...(q_i-1) * p_i$ for w=1, for i=1,2,...w.

The largest possible t value is n/p_1. There are $m^{n/p1}$ symmetric strings, and hence the symmetry ratio is given by:

$$\sigma(s) = m^{n/p1} / m^n = 1/ m^{n(1-1/p_1)}$$

The highest symmetry ratio is for even numbers where $p_1=2$:

$$\sigma(s) = 1 / m^{n(1-0.5)} = 1 / n^{0.5}$$

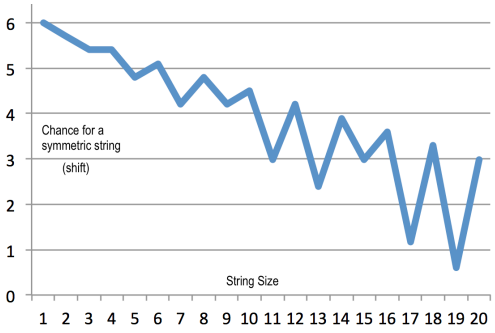
which is quite a small ratio which approaches zero for larger strings:

$$Lim\ \sigma(s)_{\ n\to\infty} = 0$$

And the smallest symmetry ratio happens when $p_1 \sim \sqrt{n}$

$$\sigma(s) \sim \frac{1}{m^{n-\sqrt{n}}}$$

As can be depicted for the binary case, m=2:



The logarithmic graph shows the reverse peaks associated with prime numbers, as well as the general down trend that indicates that for larger and larger strings the probability of encountering a symmetric string is fast diminishing.

A. Rotational Symmetries

In addition to the shift symmetries, we may recognize rotational symmetries: Let a string s comprised of an even number of letters, n, be rotated around an imaginary axis placed between letter 0.5n and letter 0.5n+1. Accordingly, letter $l_{0.5n-i}$ becomes $l'_{0.5n+i-1}$ for i=-0.5n, ...-1, and letter $l_{0.5n+i-1}$ becomes $l'_{0.5n-i}$ for i=0.5n...1

Illustration: s='ABCDXYZW' where n=8, we rotate letter l_0 to become l_7, while l_7 becomes $=l_0$, same for letter l_1 exchanging places with l_6 to generate s'='WZYXDCBA'

For a string comprised of n odd letters, letter $l_{0.5(n-1)}$ will be regarded as the axis of rotation. Rotation will be executed by moving letter $l_{0.5(n-1)+i}$ to letter $l'_{0.5(n-1)-i}$ for i=-0.5(n-1) to 0.5(n-1).

Illustration: s= 'ABCXDEF', n=7, letter X will serve as the axis and letter A and F will exchange places, similarly letters B and E and letter C and D.

Much as the case with shift symmetry, if a string is left unchanged under rotation, it will be left unchanged under any number of rotations.

Examples for symmetries are s='ABCDEEDCBA', s='ABCDEXEDCBA'

B. Symmetric Exchange

Given a string s comprised of n even letters, then one can exchange the two halves (each comprising 0.5n bits). This exchange is identical to a shift operation with t=0.5n. However a similar exchange can be carried out over a string of n odd letters, where letter number 0.5(n-1) is regarded as the axis around which the two parts comprising each of 0.5(n-1) letters will change places. This will be a symmetric exchange, regardless of the identity of axis letter, as long as the two 'wings' (the two substrings of 0.5(n-1) letters) are identical.

Illustration: let s='ABCXDEF', n=7, 'X' will serve as the axis and 'ABC' and 'DEF' will exchange places. Similarly s='ABCDEXABCDE' is symmetric per this exchange.

C. Total Symmetry

We have seen that a string s comprised of n letters l_0, l_1,.....$l_{(n-1)}$ may exhibit shift symmetry $\mathbf{\Sigma}_s$, or a rotational symmetry, $\mathbf{\Sigma}_r$, or an exchange summary, $\mathbf{\Sigma}_e$. The proportion of strings that show shift symmetry is σ_s, the proportion of strings that show rotational symmetry is σ_r, and the proportion of strings that show exchange symmetry is σ_e.

For a string s of size m letters, the chance for an arbitrary string like this to be symmetric is:

$$Pr[s: symmetric] = \sigma_s + \sigma_r + \sigma_e$$

For n even, the exchange option is covered by the shift option over t=0.5n. The rotational symmetry will happen: $m^{0.5n}$, so:

$$\sigma(s) = m^{0.5n} / m^n = 1/ m^{0.5n}$$

where clearly Lim $\sigma(s)_{n \to \infty} = 0$

For an odd n, the exchange symmetry requires an extra computation. Each pattern of the first 0.5(n-1) letters will have two options for the latter 0.5(n-1) letters: one option for rotation and one for exchange. And that is true for all the m options for the axis. Hence:

$$\sigma(s)_r + \sigma(s)_e = 2* m * m^{0.5(n-1)} = 2*m^{0.5(n+1)}$$

For an odd n=|s|.

3.0 Symmetric Strings as Building Blocks

We have identified symmetry within strings of any size, of any alphabet, and have seen that their frequency is diminishing fast as the string grows in size. Yet, even the largest string is associated with symmetric occurrences. This is reminiscent of the situation with prime numbers -- their average frequency diminishes, but primes occur even at the furthest sections of the natural numbers sequence. It may then be promising to take this analogy further. Every natural number is either a prime itself or a composite of primes. Why not view any string as either symmetric itself, or composed of symmetric sub-strings?

Let s be a string comprised of n letters, and let s_1, s_2,... s_p be p substrings of s which concatenate in order to s:

$$s = s_1 \,||\, s_2 \,||\, ... \,||\, s_p$$

If every substring s_i (i=1,2...p) is itself a symmetric string then we shall regard this set of p strings, as a symmetric decomposition of s. Illustration: let s='ABCXCBA101010UVZUV'. Let us decompose s into three concatenated substrings as follows: s_1 = ABCXCBA s_2 = 101010, and s_3 = UVZUV. s_1 is symmetric per rotation, s_2 is symmetric per shift of degree 2, and s_3 is symmetric per exchange. Since all the substrings that concatenate into s are symmetric, we shall regard s → {s_1, s_2, s_3} as a symmetric decomposition of s.

Lemma 5: Every string s may be constructed as a concatenation of symmetric substrings.

Proof: Since we defined a single letter string (n=1) as self symmetric, it is always possible to regard any string s of size n symbols as a concatenation of these very n symbols -- each is self symmetric.

Lemma 6: A string s comprised of n letters L_s = {l_0, l_1,....l_{n-1}} where $l_i \neq l_j$ for any i ≠ j (i,j,=0,1,2,...(n-1)} has only one symmetric decomposition, to n substrings, each comprised of a single letter in s.

Proof: since no letter in s repeats itself there can be no symmetry apart from the trivial self symmetry applicable to all the letters in s, which is the only decomposition.

In general, a string s may be decomposed in several distinct ways to symmetric substrings. Illustration: Let s='XYZXYZXYZX'. s can be decomposed symmetrically

to: s_1 = 'XYZXYZXYZ', s_2 ='X', where s_1 is shift symmetric with degree t=3. But also to: s'_1 = 'XYZXY' and s'_2 = 'ZXYZX', where both s'_1 and s_2 are exchange-symmetric. Another decomposition will be: s''_1 = 'X', s''_2 = 'Y', s''_3 = 'Z', s''_4 = 'XYZXYZ', and s''_5='X'.

The largest number of symmetric sub-strings has been shown to be n=|s|, the number of letters in s. The smallest number q of symmetric substrings that a given string s may be decomposed to is a property of the string irrespective of the alphabet from which the letters of the string were drawn. For a symmetric string q=1. For a string comprised of non-repeat letters, we have seen that q=n.

We consider the set S(m,n) of strings comprised of n letters each, drawn from an alphabet comprised of m letters. S contains m^n strings. Each string $s \in S$ is associated with a q-value, the smallest number of symmetric sub-strings it may be decomposed to. We have seen that $1 \le q \le n=|s|$.

III. Randomness as Lack of Symmetry

The journey of mathematics and science is a hunt for invariables -- that which could serve as anchor for our understanding and mental comfort. Be it natural laws, be it immutable theorems, and be it geometry based symmetry under well specified action. In that light randomness is that which defies the existence of invariables, declares the absence of pattern. It stands to reason then that symmetry should be viewed as a metric of poor randomness.

In particular, striving to produce an objective metric for the randomness of a given string, it stands to reason to approach this pursuit through the value q of the given string s that represents *the smallest* number of symmetric substrings to which s may be decomposed. And since we have $1 \le q \le |s|$, we can normalize the randomness metric of s to be:

$$\rho(s) = \frac{q(s)-1}{|s|-1}$$

So that we have: $0 \le \rho(s) \le 1$. $\rho(s)=0$ implies 'zero randomness' which comes about if s per se is symmetric, and $\rho(s)=1$ implies maximum randomness.

The set S(m,n) is comprised of m^n strings, each of which is associated with its q (and ρ) value. There are exactly |s| ρ (and q) values, and hence we have a mapping of $m^n \rightarrow |n|$, or on average, m^n/n strings share a randomness measure. And each set S(m,n) is associated with a characteristic randomness distribution curve, $R_{dist}(\rho)$ which measures the count of strings per a given ρ (or q) value. Clearly:

$$\Sigma\, R_{dist}(q)_{\ for\ q=1,2,\ldots n} = m^n$$

A. Analysis of the Symmetry Based Metric for Randomness

The first question of interest is to what extent does this definition capture the intuitive notion of randomness. A disorderly string is not likely to have large symmetric sub strings, and hence its randomness metric will be high. An interesting way to appreciate the consistency of this definition with our intuitive grasp is by trying to build a highly randomized string out of a given alphabet.

Let's try with m=4. We consider strings of varying length, comprised of the letters X, Y, Z, and W. For n=4, it is easy, any sequence of the four letters will be rated as fully randomized. Let's use then s=XYZW. Now setting up the 5th letter, we can't use W since it will combine with the letter W at the 4th position to define a symmetric substring. We can't use Z either, since the substring ZWZ will rate as symmetric. So let's choose X (between X and Y). Now s(n=5)='XYZWX' The 6th letter can't be X and can't be W (as argued above). So we have to choose between Y and Z. let's choose Y: s(n=6)='XYZWXY' The 7th letter can't be Y, nor X, but not Z either because it will then create a perfectly symmetric string based on exchange symmetry. So we have to use W: S(n=7)='XYZWXYW' The 8th letter can't be W, can't be Y, but also can't be X because it will then define a substring WXYWX, which is symmetric by exchange. So far we encountered two subsequent instances where we had no choice, only one letter was possible to insure the full randomness of the growing string. It is conjectured that it would be hard to build a fully randomized string of an arbitrary size n, (for m=4) but it is possible.

By contrast, for m=2 and m=3 it is impossible to build a fully randomized string, of arbitrary length, as is readily shown: for m=2, for n=2 we have s='01' or s='10' because '00, or '11' are symmetric. But the third letter will make '01' to '011' or '010' -- both are symmetric. For m=3 we start s='XYZ'. The fourth letter will have to be X: s='XYZX'. The fifth letter can't be X nor Z, but also Y won't do because s='XYZXY is exchange symmetric.

This demonstrates how difficult it is to construct a fully randomized string, which is consistent with the intuitive grasp. On the other hand, envisioned orderly strings will rate as low randomness. E.g. 0000...0, 01010101...., 10011001001...

The second question that comes to mind is *utility*. How useful is this metrics of randomness. This symmetric metric can readily be used in conjunction with any random number generator. The output stream would be appraised symmetry-wise, and failing strings will be discarded.

B. Cryptographic Utility of Symmetric Metric

As mentioned above the simple way to use this symmetric method is as a qualifier. A procedure to dismiss candidates for random strings which don't withstand the test. But the main utility is associated with the rich nature of this measurement. It spans from zero to total randomness. One could study a cipher by comparing the randomness metric of the plaintext, ρ(p), with the randomness metric of the corresponding ciphertext, ρ(c). A stream of ciphertext can be chopped to blocks of size b letters, allowing each block to be symmetrically analyzed for randomness, and then

depicting the randomness distribution curve R_{dist} for cryptanalytic purposes.

In this application the ρ metric seems to challenge the more familiar entropy measurement. There is a fundamental difference. Entropy is based on assessing probabilities which is based on a large set of operational assumptions. The symmetric metric has none of it. In fact one does not even have to identify a reference alphabet, and make do with the letters that appear in the measured string. No further operational assumption is necessary.

The third question of interest is how does the symmetric metric compare with the more established means to measure randomness. As mentioned in the introduction the common metrics for randomness either address the means for generating the randomness candidate string, or to means to decipher it, to learn its pattern, or thirdly, to a Turing machine test -- whether there is a smaller string that can generate the candidate string via a Turing machine. And lastly, the above mentioned entropy which is based on calculated probabilities. By contrast, the symmetric metric is based only on the candidate string per se, not even on the alphabet from which its letters are drawn. It is easy to measure and to deduce from.

IV. Symmetric Metric of Randomness - Mathematical Abstraction Level

We consider a mathematical construct E, and an associated operator α which may operate on it: E' =α(E). If E'=E, we consider E symmetric with respect to α. We define the symmetric index (marker) of E as: $E^s_\alpha = \{0,1\}$, where 1 indicates that E is symmetric with respect to α, and 0, otherwise. We define a set of unit entities: $U = u_1, u_2,.....$ and an addition (+) function, such that E can be expressed as an addition of some or these units: $E = \mathbf{\Sigma}\ u_i$, i=1,2,.....t. We set for every u_i: $u_i^s{}_\alpha = 1$, namely we define a unit entity as inherently symmetric. Hence E may be seen as comprised of t unit entities for which $\mathbf{\Sigma}\ u_i^s{}_\alpha = t$.

Applying the selected addition function over a subset of the t unit entities that add up to E, we define an E subentity $e_j = \mathbf{\Sigma}\ u_i$ for i=1,2,...t' (t' ≤ t). We can then write entity E as comprised of subenttities $E = \mathbf{\Sigma}\ e_i$ i=1,2,...q . We construct this structure such that operator alpha is defined over each subenttity e_i, which can be either symmetric or not: $e_i^s{}_\alpha = \{0,1\}$

There may be many ways to assemble the t' unit entities to comprise E to q subentities. For each of such assemblies we can define the 'symmetric count' of E as: $\sigma^*(E) = \mathbf{\Sigma}\ e_i^s{}_\alpha$ for i=1,2,...q. Clearly $\sigma^*(E)_{max} = t$, which is the case where for all possible subentities e_j that are comprised of two or more unit entities we have $e_j^s{}_\alpha = 0$.

The interesting question here is the smallest value for σ*(E) . For some combination of unit entities there may be subentities of large number of unit entities and for them there is a symmetry with respect to α (all relative to the well defined addition function that dictates how unit entities assemble to subentities, and how subentities assemble to the entity). For E with respect to α, symmetry is a binary choice {0,1}. However if E can be comprised of q subentities such that each of these entities is symmetric with respect to α then to the extent that q < t then E is more 'symmetric' with respect to α. In other words, let E_1 and E_2 be two entities, such that each is comprised of some unit entities $u_1, u_2,....$. E_1 is comprised of t_1 of such units and E_2 is comprised of t_2 of such units. Let's first examine the case where $t=t_1 = t_2$. In that case if $q_1 < q_2$, then E_1 is comprised of larger subentities (and so fewer of them) which are all symmetric with respect to α, or say: $\mathbf{\Sigma}\ e_i^s{}_\alpha < \mathbf{\Sigma}\ e_j^s{}_\alpha$ where i summarizes the subentities for E_1 and j summarizes the subentities for E_2. This implies that E_1 is comprised of fewer (and larger) subentities that are all symmetric with respect to α. And being more symmetric implies being less random. We therefore can define symmetric metric for randomness in the form:

$$\rho_\alpha(E) = (\Sigma\ e_i^s{}_\alpha - 1)/(\Sigma\ u_i^s{}_\alpha - 1)$$

$\rho_\alpha(E) = 0$ is the case where $E^s = 1$ namely E is itself is symmetric with respect to α, and $\rho_\alpha(E)=1$ in the case where q=t, or the case where any addition of the unit entities that comprise E is asymmetric with respect to alpha. Or say, the only way to express E as a combination of symmetric ingredients is to regard E as a combination of its unit entities. Since no subentity larger than the unit entities has symmetry with respect to α. The specific form of the randomness formula is adjusted to map randomness on stretch from 0 to 1.

Justification : The proposed metric ρ(E) for the randomness of an entity E requires some arguments in favor of its validity. We consider a set of mathematical entities of the type of E, ‹ E › of size |‹ E ›|=g. If E is drawn uniformly from ‹ E›, then its randomness value ρ(E) is inconsequential because every member of the set has the same chance (1/g) to be selected. However, if one knows the value of ρ(E), and tries to use this knowledge to guess the identity of E, then to the extent that highly symmetric entities are a minority, then knowledge of a low level or ρ(E) is very consequential.

Normalized Symmetric Metric: The set U of all possible unit entities $u_1, u_2,.....u_m$, counting m units, may be larger than the number of such units needed to construct the subject entity E. |E| < |U|. In that case there would be some such strings for which the randomness metric ρ(E)=1 because then the only de-composition of E to symmetric subentities is through the unit entities, and hence $\Sigma\ e_i^s{}_\alpha = \Sigma\ u_i^s{}_\alpha$ leading to ρ=1. However, in the case where |E| > |U|, there will have to be duplication of unit entities in the construction of E. This might lead to a situation where for a given E expressed through such a set U, it is impossible to compose an E for which ρ(E)=1. For a given size |E| there would be various E entities with different randomness metric, which will have a maximum value $\rho(E)_{max} <= 1$. Such entities will have a corresponding number q_{max} of symmetric substrings, counted for each E as the smallest

possible number for that E. This would lead to a normalized expression of randomness:

$$\rho(E)_{normalized} = \rho^*(E) = (\Sigma\, e_i^s{}_{\alpha} - 1)/(q_{max} - 1)$$

V. Symmetric Binary Strings

Let's consider a string s =

10001110011100011110011000111100001110100101010

comprised of 47 bits |s|=n=47. The smallest number of symmetric sub strings that comprise s (q) is 10:

s =1 ~ 0 ~ 0011100 ~ 1 ~ 11000111100110001111100 ~ 00 ~ 11 ~ 101001010 ~ 1 ~ 0

which evaluates to ρ(s)=20%

The following 30 bits string:

011000010111100111000001010110

breaks down to a smallest number of symmetric substrings of count 10: ~ 0 ~ 1 ~ 100001 ~ 0 ~ 1 ~ 1110011100 ~ 00 ~ 010101 ~ 1 ~ 0 (ρ=31%)

By evaluating the symmetry of all 2^{30} strings of same size one finds out that q=10 as above, is the maximum number of concatenating substrings, as exemplified below: String s_1 = 010111010101110010001100101000 has a smallest number of concatenated substrings 5, so $\rho(s_1)$ = (5-1)(30-1) = 14%. Similarly s_2 = 011110101001001111000110101010 divides to no less than 7 symmetric substrings. s_3 = 100001010011000010011010110001 divides to no less than q=8 symmetric substrings, and s_4 = 110010111011001111101110011001 divides to 9 as the smallest count of symmetric substrings. By contrast s_5 =010101010101010101010101010101 will have q=1. We can thus write down a table of randomness rating both regular or "raw" (ρ) and 'normalized' (ρ*):

	q	ρ	ρ*
s1	5	0.14	0.44
s2	7	0.21	0.67
s3	8	0.24	0.78
s4	9	0.28	0.89
s5	1	0.00	0.00

VI. Summary

Randomness may be viewed as the boundary of mathematical analysis, and hence the profound interest in its pursuit. It was D. E. Knuth who cunningly observed (1998): "*The mathematical theory of probability and statistics scrupulously avoids the issue* (Randomness)" And this observation is still fitting so many decades later. In 1991 the American Journal of Physics (59,700) stated: "*Randomness is a fundamental but elusive concept in mathematics and physics. Even for the elementary case of a random binary sequence, a generally accepted and operational definition is lacking.*" This state of affairs by itself serves as a motivation for efforts like herein -- to propose new and independent ways to define, measure and appraise randomness. Any new insight with respect to randomness will have a significant impact on modern cryptography that relies on randomness as its central pillar. The proposed symmetry based metric for randomness qualifies as a metric of interest. Several immediate practical impacts are now under investigation.

VII. Reference

1. "NIST Randomness Beacon" https://www.nist.gov/programs-projects/nist-randomness-beacon
2. Nies, A. "Computability and Randomness" Oxford University Press * Oxford Logic Guides, 2009
3. Brock, Hommes "Rational Routes to Randomness" https://www.nist.gov/programs-projects/nist-randomness-beacon
4. Chaitin G. J. "Randomness and Mathematical Proof" Scientific American 232, No. 5 (May 1975), pp. 47-52
5. Ahlswede "Common Randomness in Information Theory" IEEE Transactions on Information Theory, Vol 39, No 4 July 1993
6. Downey R. "Algorithmic Randomness and Complexity" School of Mathematics and Computing Sciences, Victoria University, Wellington New Zealand. 2007
7. ID Quantique (IDQ) https://www.idquantique.com