
Received: 12 December 2012, Accepted: 2 February 2013
 Edited by: G. Mindlin
 Licence: Creative Commons Attribution 3.0
 DOI: <http://dx.doi.org/10.4279/PIP.050001>



ISSN 1852-4249

LT²C²: A language of thought with Turing-computable Kolmogorov complexity

Sergio Romano,^{1*} Mariano Sigman,^{2,3†} Santiago Figueira^{1,3‡}

In this paper, we present a theoretical effort to connect the theory of program size to psychology by implementing a concrete language of thought with Turing-computable Kolmogorov complexity (LT²C²) satisfying the following requirements: 1) to be simple enough so that the complexity of any given finite binary sequence can be computed, 2) to be based on tangible operations of human reasoning (*printing, repeating...*), 3) to be sufficiently powerful to generate all possible sequences but not too powerful as to identify regularities which would be *invisible* to humans. We first formalize LT²C², giving its syntax and semantics and defining an adequate notion of *program size*. Our setting leads to a Kolmogorov complexity function relative to LT²C² which is computable in polynomial time, and it also induces a prediction algorithm in the spirit of Solomonoff's inductive inference theory. We then prove the efficacy of this language by investigating regularities in strings produced by participants attempting to generate random strings. Participants had a profound understanding of randomness and hence avoided typical misconceptions such as exaggerating the number of alternations. We reasoned that remaining regularities would express the algorithmic nature of human thoughts, revealed in the form of specific patterns. Kolmogorov complexity relative to LT²C² passed three expected tests examined here: 1) human sequences were less complex than control PRNG sequences, 2) human sequences were not stationary, showing decreasing values of complexity resulting from fatigue, 3) each individual showed traces of algorithmic stability since fitting of partial sequences was more effective to predict subsequent sequences than average fits. This work extends on previous efforts to combine notions of Kolmogorov complexity theory and algorithmic information theory to psychology, by explicitly proposing a language which may describe the patterns of human thoughts.

*E-mail: sgromano@dc.uba.ar

†E-mail: sigman@df.uba.ar

‡E-mail: santiago@dc.uba.ar

¹ Department of Computer Science, FCEN, University of Buenos Aires, Pabellón I, Ciudad Universitaria (C1428EGA) Buenos Aires, Argentina.

² Laboratory of Integrative Neuroscience, Physics Department, FCEN, University of Buenos Aires, Pabellón I, Ciudad Universitaria (C1428EGA) Buenos Aires, Argentina.

³ CONICET, Argentina.

I. Introduction

Although people feel they understand the concept of randomness [1], humans are unable to produce random sequences, even when instructed to do so [2–6], and to perceive randomness in a way that is inconsistent with probability theory [7–10]. For instance, random sequences are not perceived by participants as such because runs appear too long to be random [11,12] and, similarly, sequences pro-

statistical models that incorporate the detection of certain regularities, which are classified in terms of the Chomsky hierarchy [23]. They showed the existence of motifs (repetition, symmetry) and related their probability distributions to Kolmogorov complexity via Levin's coding theorem (cf. section VII. for more details).

The main novelty of our work is to develop a class of specific programming languages (or Turing machines) which allows us to stick to the theory of program size developed by Kolmogorov, Solomonoff and Chaitin. We use the patterns of sequences of humans aiming to produce random strings to fit, for each individual, the language which captures these regularities.

II. Mathematical theory of randomness

The idea behind Kolmogorov complexity theory is to study the length of the descriptions that a formal language can produce to identify a given string. All descriptions are finite words over a finite alphabet, and hence each description has a finite length—or, more generally, a suitable notion of *size*. One string may have many descriptions, but any description should describe one and only one string. Roughly, the Kolmogorov complexity [17] of a string x is the length of the shortest description of x . So a string is 'simple' if it has at least one short description, and it is 'complex' if all its descriptions are long. Random strings are those with high complexity.

As we have mentioned, Kolmogorov complexity uses programming languages to describe strings. Some programming languages are Turing complete, which means that any partial computable function can be represented in it. The commonly used programming languages, like C++ or Java, are all Turing complete. However, there are also Turing incomplete programming languages, which are less powerful but more convenient for specific tasks.

In any reasonable imperative language, one can describe x_2 above with a program like (1), of length 26, which is considerably smaller than 40, the size of the described string. It is clear that x_2 is 'simple'. The case of x_3 is a bit tricky. Although at first sight it seems to have a complete lack of structure, it contains a hidden pattern: it consists of the first forty binary digits of π after the decimal

point. This pattern could hardly be recognized by the reader, but once it is revealed to us, we agree that x_3 must also be tagged as 'simple'. Observe that the underlying programming language is central: x_3 is 'simple' with the proviso that the language is strong enough to represent (in a reasonable way) an algorithm for computing the bits of π —a language to which humans are not likely to have access when they try to find patterns in a string. Finally, for x_1 , the best way to describe it seems to be something like

print '1100101001111000101000110101100110011100',

which includes the string in question verbatim, length 48. Hence x_1 only has long descriptions and hence it is 'complex'.

In general, both the string of length n which alternates 0s and 1s and the string which consists of the first n binary digits of π after the decimal point can be computed by a program of length $\approx \log n$ —and this applies to any computable sequence. The idea of the algorithmic randomness theory is that a truly random string of length n necessarily needs a program of length $\approx n$ (cf. section ii. for details).

i. Languages, Turing machines and Kolmogorov complexity

Any programming language \mathcal{L} can be formalized with a Turing machine $M_{\mathcal{L}}$, so that programs of \mathcal{L} are represented as inputs of $M_{\mathcal{L}}$ via an adequate binary codification. If \mathcal{L} is Turing complete then the corresponding machine $M_{\mathcal{L}}$ is called *universal*, which is equivalent to say that $M_{\mathcal{L}}$ can simulate any other Turing machine.

Let $\{0, 1\}^*$ denote the set of finite words over the binary alphabet. Given a Turing machine M , a program p and a string x ($p, x \in \{0, 1\}^*$), we say that p is an M -description of x if $M(p) = x$ —i.e., the program p , when executed in the machine M , computes x . Here we do not care about the time that the computation needs, or the memory it consumes. The Kolmogorov complexity of $x \in \{0, 1\}^*$ relative to M is defined by the length of the shorter M -description of x . More formally,

$$K_M(x) \stackrel{\text{def}}{=} \min\{|p| : M(p) = x\} \cup \{\infty\},$$

where $|p|$ denotes the length of p . Here M is *any* given Turing machine, possibly one with a very specific behavior, so it may be the case that a given

string x does not have any M -description at all. In this case, $M(x) = \infty$. In practical terms, a machine M is a useful candidate to measure complexity if it computes a surjective function. In this case, every string x has at least one M -description and therefore $K_M(x) < \infty$.

ii. Randomness for finite words

The strength of Kolmogorov complexity appears when M is set to any universal Turing machine U . The invariance theorem states that K_U is minimal, in the sense that for every Turing machine M there is a constant c_M such that for all $x \in \{0, 1\}^*$ we have $K_U(x) \leq K_M(x) + c_M$. Here, c_M can be seen as the specification of the language M in U (i.e., the information contained in c_M tells U that the machine to be simulated is M). If U and U' are two universal Turing then K_U and $K_{U'}$ differ at most by a constant. In a few words, $K_U(x)$ represents the length of the ultimate compressed version of x , performed by means of algorithmic processes.

For analysis of arbitrarily long sequences, c_M becomes negligible and hence for nonpractical aspects of the theory the choice of the machine is not relevant. However, for short sequences, as we study here, this becomes a fundamental problem, as notions of complexity are highly dependent on the choice of the underlying machine through the constant c_M . The most trivial example, as referred in the introduction, is that for any given sequence, say x_1 , there is a machine M for which x_1 has minimal complexity.

iii. Solomonoff induction

Here we have presented *compression* as a framework to understand randomness. Another very influential paradigm proposed by Schnorr is to use the notion of *martingale* (roughly, a betting strategy), by which a sequence is random if there is no computable martingale capable of predicting forthcoming symbols (say, of a binary alphabet $\{0, 1\}$) better than chance [28, 29]. In the 1960s, Solomonoff [21] proposed a universal prediction method which successfully approximates any distribution μ , with the only requirement of μ being computable.

This theory brings together concepts of algorithmic information, Kolmogorov complexity and prob-

ability theory. Roughly, the idea is that amongst all ‘explanations’ of x , those which are ‘simple’ are more relevant, hence following Occam’s razor principle: amongst all hypothesis that are consistent with the data, choose the simplest. Here the ‘explanations’ are formalized as programs computing x , and ‘simple’ means low Kolmogorov complexity.

Solomonoff’s theory, builds on the notion of *monotone* (and prefix) Turing machines. Monotone machines are ordinary Turing machines with a one-way read-only input tape, some work tapes, and a one-way *write-only* output tape. The output is written one symbol at a time, and no erasing is possible in it. The output can be finite if the machine halts, or infinite in case the machine computes forever. The output head of monotone machines can only “print and move to the right” so they are well suited for the problem of inference of forthcoming symbols based on partial (and finite) states of the output sequence. Any monotone machine N has the *monotonicity* property (hence its name) with respect to extension: if $p, q \in \{0, 1\}^*$ then $N(p)$ is a prefix of $N(p \hat{\ } q)$, where $p \hat{\ } q$ denotes the concatenation of p and q .

One of Solomonoff’s fundamental results is that given a finite observed sequence $x \in \{0, 1\}^*$, the most likely finite continuation is the one in which the concatenation of x and y is less complex in a Kolmogorov sense. This is formalized in the following result (see theorem 5.2.3 of [24]): for almost all infinite binary sequences X (in the sense of μ) we have

$$-\lim_{n \rightarrow \infty} \log \mu(y \mid X \upharpoonright n) = \lim_{n \rightarrow \infty} Km_U((X \upharpoonright n) \hat{\ } y) - Km_U(X \upharpoonright n) + O(1) < \infty.$$

Here, $X \upharpoonright n$ represents the first n symbols of X , and Km_U is the monotone Kolmogorov complexity relative to a monotone universal machine U . That is, $Km_U(x)$ is defined as the length of the shortest program p such that the output of $U(p)$ starts with x —and possibly has a (finite or infinite) continuation.

In other words, Solomonoff inductive inference leads to a method of prediction based on data compression, whose idea is that whenever the source has output the string x , it is a good heuristic to choose the extrapolation y of x that minimizes $Km_U(x \hat{\ } y)$. For instance, if one has observed x_2 , it is more likely for the continuation to be 1010

rather than 0101, as the former can be succinctly described by a program like

$$\text{repeat 22 times: print '10'}. \quad (2)$$

and the latter looks more difficult to describe; indeed the shorter program describing it seems to be something like

$$\begin{aligned} &\text{repeat 20 times: print '10'}; \\ &\text{print '0101'}. \end{aligned} \quad (3)$$

Intuitively, as program (2) is shorter than (3), $x_2 \wedge 1010$ is more probable than $x_2 \wedge 0101$. Hence, if we have seen x_2 , it seems to be a better strategy to predict 1.

III. A framework for human thoughts

The notion of thought is not well grounded. We lack an operative working definition and, as also happens with other terms in neuroscience (consciousness, self, ...), the word *thought* is highly polysemic in common language. It may refer, for example, to a belief, to an idea or to the content of the conscious mind. Due to this difficulty, the mere notion of *thought* has not been a principal or directed object of study in neuroscience, although of course it is always present implicitly, vaguely, without a formal definition.

Here we do not intend to elaborate an extensive review on the philosophical and biological conceptions of thoughts (see [30] for a good review on thoughts). Nor are we in a theoretical position to provide a full formal definition of a thought. Instead, we point to the key assumptions of our framework about the nature of thoughts. This accounts to defining constraints in the class of thoughts which we aim to describe. In other words, we do not claim to provide a general theory of human thoughts (which is not amenable at this stage lacking a full definition of the class) but rather of a subset of thoughts which satisfy certain constraints defined below.

For instance, E.B. Titchener and W. Wundt, the founders of structuralist school in psychology (seeking structure in the mind without evoking meta-physical conceptions, a tradition which we inherit and to which we adhere), believed that thoughts

were images (there are not *imageless thoughts*) and hence can be broken down to elementary sensations [30]. While we do not necessarily agree with this propositions (see Carey [31] for more contemporary versions denying the sensory foundations of conceptual knowledge), here we do not intend to explain all possible thoughts but rather a subset, a simpler class which—in agreement with the Wundt and Titchener—can be expressed in images. More precisely, we develop a theory which may account for Boole’s [32] notion of thoughts as propositions and statements about the world which can be represented symbolically. Hence, a first and crucial assumption of our framework is that thoughts are discrete. Elsewhere we have extensively discussed [33–39] how the human brain, whose architecture is quite different from Turing machines, can emerge in a form of computation which is discrete, symbolic and resembles Turing devices.

Second, here we focus on the notion of “prop-less” mental activity, i.e., whatever (symbolic) computations can be carried out by humans without resorting to external aids such as paper, marbles, computers or books. This is done by actually asking participants to perform the task “in their heads”. Again, this is not intended to set a proposition about the universality of human thoughts but, instead, a narrower set of thoughts which we conceive is theoretically addressable in this mathematical framework.

Summarizing:

1. We think we do not have a good mathematical (even philosophical) conception of thoughts, as mental structures, yet.
2. Intuitively (and philosophically), we adhere to a materialistic and computable approach to thoughts. Broadly, one can think (to picture, not to provide a formal framework) that thoughts are formations of the mind with certain stability which defines distinguishable clusters or objects [40–42].
3. While the set of such objects and the rules of their transitions may be of many different forms (analogous, parallel, unconscious, unlinked to sensory experience, non-linguistic, non-symbolic), here we work on a subset of thoughts, a class defined by Boole’s attempt

to formalize thought as symbolic propositions about the world.

4. This states —which may correspond to human “conscious rational thoughts”, the seed of Boole and Turing foundations [34,34]— are discrete and defined by symbols and potentially represented by a Turing device.
5. We focus on an even narrower space of thoughts. Binary formations (right or left, zero or one) to focus on what kind of language better describes these transitions. This work can be naturally extended to understand discrete transitions in conceptual formations [43–45].
6. We concentrate on prop-less mental activity to understand limitations of the human mind when it does not have evident external support (paper, computer...)

IV. Implementing a language of thought with Turing-computable complexity

As explained in section II.i., Kolmogorov complexity considers all possible computable compressors and assigns to a string x the length of the shortest of the corresponding compressions. This seems to be a perfect theory of compression but it has a drawback: the function K_U is not computable, that is, there is no effective procedure to calculate $K_U(x)$ given x .

On the other hand, the definition of randomness introduced in section II.i., having very deep and intricate connections with algorithmic information and computability theories, is simply too strong to explain our own perception of randomness. To detect that x_3 consists of the first twenty bits of π is incompatible with human patterns of thought.

Hence, the intrinsic algorithms (or observed patterns) which make human sequences not random are too restricted to be accounted by a *universal* machine and may be better described by a *specific* machine. Furthermore, our hypothesis is that each person uses his own particular specific machine or algorithm to generate a random string.

As a first step in this complicated enterprise, we propose to work with a specific language LT^2C^2 which meets the following requirements:

- LT^2C^2 must reflect some plausible features of our mental activity when finding succinct descriptions of words. For instance, finding repetitions in a sequence such as x_2 seems to be something easy for our brain, but detecting numerical dependencies between its digits as in x_3 seems to be very unlikely.
- LT^2C^2 must be able to describe any string in $\{0,1\}^*$. This means that the map given by the induced machine $N \stackrel{\text{def}}{=} N_{LT^2C^2}$ must be surjective.
- N must be simple enough so that K_N —the Kolmogorov complexity relative to N — becomes computable. This requirement clearly makes LT^2C^2 Turing incomplete, but as we have seen before, this is consistent with human deviations from randomness.
- The rate of compression given by K_N must be sensible for very short strings, since our experiments will produce such strings. For instance, the approach, followed in [46], of using the size of the compressed file via general-purpose compressors like Lempel-Ziv based dictionary (gzip) or block based (bzip2) to approximate the Kolmogorov complexity does not work in our setting. This method works best for long files.
- LT^2C^2 should have certain degrees of freedom, which can be adjusted in order to approximate the specific machine that each individual follows during the process of randomness generation.

We will not go into the details on how to codify the instructions of LT^2C^2 into binary strings of N : for the sake of simplicity we take N as a surjective total mapping $LT^2C^2 \rightarrow \{0,1\}^*$. We restrict ourselves to describe the grammar and semantics of our proposed programming language LT^2C^2 . It is basically an imperative language with only two classes of instructions: a sort of *print i* , which prints the bit i in the output; and a sort of *repeat n times P* , which for a fixed $n \in \mathbb{N}$ it repeats n times the program P . The former is simply represented as i and the latter as $(P)^n$.

Formally, we set the alphabet $\{0,1,(,),^0,\dots,^9\}$ and define LT^2C^2 over such alphabet with the following grammar:

$$P ::= \epsilon \mid 0 \mid 1 \mid PP \mid (P)^n,$$

where $n > 1$ is the decimal representation of $n \in \mathbb{N}$ and ϵ denotes the empty string. The semantics of LT^2C^2 is given through the behavior of N as follows:

$$\begin{aligned} N(i) &\stackrel{\text{def}}{=} p \quad \text{for } i \in \{\epsilon, 0, 1\} \\ N(P_1P_2) &\stackrel{\text{def}}{=} N(P_1) \wedge N(P_2) \\ N((P)^n) &\stackrel{\text{def}}{=} \underbrace{N(P) \wedge \dots \wedge N(P)}_{n \text{ times}}. \end{aligned}$$

N is not universal, but every string x has a program in N which describes it: namely x itself. Furthermore, N is monotone in the sense that if $p, q \in \text{LT}^2\text{C}^2$ then $N(p)$ is a prefix of $N(p \wedge q)$. In Table 1, the first column shows some examples of N -programs which compute 1001001001.

<i>program</i>	<i>size</i>
1001001001	10
$(100)^2 1 (0)^2 1$	6.6
$(100)^3 1$	4.5
$1((0)^2 1)^3$	3.8

Table 1: Some N -descriptions of 1001001001 and its sizes for $b = r = 1$

i. Kolmogorov complexity for LT^2C^2

The Kolmogorov complexity relative to N (and hence to the language LT^2C^2) is defined as

$$K_N(x) \stackrel{\text{def}}{=} \min\{\|p\| : p \in \text{LT}^2\text{C}^2, N(p) = x\},$$

where $\|p\|$, the *size* of a program p , is inductively defined as:

$$\begin{aligned} \|\epsilon\| &\stackrel{\text{def}}{=} 0 \\ \|p\| &\stackrel{\text{def}}{=} b \quad \text{for } p \in \{0, 1\} \\ \|P_1P_2\| &\stackrel{\text{def}}{=} \|P_1\| + \|P_2\| \\ \|(P)^n\| &\stackrel{\text{def}}{=} r \cdot \log n + \|P\|. \end{aligned}$$

In the above definition, $b \in \mathbb{N}$, $r \in \mathbb{R}$ are two parameters that control the relative weight of the *print*

operation and the *repeat n times* operation. In the sequel, we drop the subindex of K_N and simply write $K \stackrel{\text{def}}{=} K_N$. Table 1 shows some examples of the size of N -programs when $b = r = 1$. Observe that for all x we have $K(x) \leq \|x\|$.

It is not difficult to see that $K(x)$ depends only on the values of $K(y)$, where y is any nonempty and proper substring of x . Since $\|\cdot\|$ is computable in polynomial time, using dynamic programming one can calculate $K(x)$ in polynomial time. This, of course, is a major difference with respect to the Kolmogorov complexity relative to a *universal* machine, which is not computable.

ii. From compression to prediction

As one can imagine, the perfect universal prediction method described in section II.iii. is, again, non-computable. We define a computable prediction algorithm based on Solomonoff's theory of inductive inference but using K , the Kolmogorov complexity relative to LT^2C^2 , instead of Km_U (which depends on a universal machine). To predict the next symbol of $x \in \{0, 1\}^*$, we follow the idea described in section II.iii.: amongst all extrapolations y of x we choose the one that minimizes $K(x \wedge y)$. If such y starts with 1, we predict 1, else we predict 0. Since we cannot examine the infinitely many extrapolations, we restrict to those up to a fixed given length ℓ_F . Also, we do not take into account the whole x but only a suffix of length ℓ_P . Both ℓ_F and ℓ_P are parameters which control, respectively, how many extrapolation bits are examined (ℓ_F many *Future* bits) and how many bits of the tail of x (ℓ_P many *Past* bits) are considered.

Let $\{0, 1\}^n$ (resp. $\{0, 1\}^{\leq n}$) be the set of words over the binary alphabet $\{0, 1\}$ of length n (resp. at most n). Formally, the prediction method is as follows. Suppose $x = x_1 \dots x_n$ ($x_i \in \{0, 1\}$) is a string. The next symbol is determined as follows:

$$\text{Next}(x_1 \dots x_n) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{if } m_0 < m_1; \\ 1 & \text{if } m_0 > m_1; \\ g(x_{n-\ell_P} \dots x_n) & \text{otherwise.} \end{cases}$$

where for $i \in \{0, 1\}$,

$$m_i \stackrel{\text{def}}{=} \min\{K(x_{n-\ell_P} \dots x_n i \wedge y) : y \in \{0, 1\}^{\leq \ell_F}\},$$

and $g : \{0, 1\}^{\ell_P} \rightarrow \{0, 1\}$ is defined as $g(z) = i$ if the number of occurrences of i in z is greater than the number of occurrences of $1 - i$ in z ; in case the number of occurrences of 1s and 0s in z coincide then $g(z)$ is defined as the last bit of z .

V. Methods

Thirty eight volunteers (mean age = 24) participated in an experiment to examine the capacity of LT^2C^2 to identify regularities in production of binary sequences. Participants were asked to produce random sequences, without further instruction.

All the participants were college students or graduates with programming experience and knowledge of the theoretical foundations of randomness and computability. This was intended to test these ideas in a *hard* sample where we did not expect typical errors which results from a misunderstanding of chance.

The experiment was divided in four blocks. In each block the participant pressed freely the left or right arrow 120 times.

After each key press, the participant received a notification with a green square which progressively filled a line to indicate the participant the number of choices made. At the end of the block, participants were provided feedback of how many times the predictor method has correctly predicted their input. After this point, a new trial would start.

38 participants performed 4 sequences, yielding a total of 152 sequences. 14 sequences were excluded from analysis because they had an extremely high level of predictability. Including these sequences would have actually improved all the scores reported here.

The experiment was programmed in ActionScript and can be seen at <http://gamesdata.lafhis-server.exp.dc.uba.ar/azarexp>.

VI. Results

i. Law of large numbers

Any reasonable notion of randomness for strings on base 2 should imply Borel's normality, or the law of large numbers in the sense that if $x \in \{0, 1\}^n$ is random then the number of occurrences of any given string y in x divided by n should tend to $2^{-|y|}$, as n goes to infinity.

A well-known result obtained in some investigations on generation or perception of randomness in binary sequences is that people tend to increase the number of alternations of symbols with respect to the expected value [27]. Given a string x of length n with r runs, there are $n - 1$ transitions between successive symbols and the number of alternations between symbol types is $r - 1$. The *probability of alternation* of the string x is defined as

$$P(x) : \{0, 1\}^{\geq 2} \rightarrow [0, 1]$$

$$P(x) = \frac{r-1}{n-1}.$$

In our experiment, the average $P(x)$ of participants was 0.51, very close to the expected *probability of alternation* of a random sequence which should be 0.5. A t-test on the $P(x)$ of the strings produced by participants, where the null hypothesis is that they are a random sample from a normal distribution with mean 0.5, shows that the hypothesis cannot be rejected as the p -value is 0.31 and the confidence interval on the mean is [0.49, 0.53]. This means that the *probability of alternation* is not a good measure to distinguish participant's strings from random ones, or at least, that the participants in this very experiment can bypass this validation.

Although the *probability of alternation* was close to the expected value in a random string, participants tend to produce n -grams of length ≥ 2 with probability distributions which are not equiprobable (see Fig. 1). Strings containing more alternations (like 1010, 0101, 010, 101) and 3- and 4- runs have a higher frequency than expected by chance. This might be seen as an effort from participants to keep the *probability of alternation* close to 0.5 by compensating the excess of alternations with blocks of repetitions of the same symbol.

ii. Comparing human randomness with other random sources

We asked whether K , the Kolmogorov complexity relative to LT^2C^2 defined in section IV.i., is able to detect and compress more patterns in strings generated by participants than in strings produced by other sources, which are considered *random* for many practical issues. In particular, we studied strings originated by two sources: Pseudo-Random Number Generator (PRNG) and Atmospheric Noise (AN).

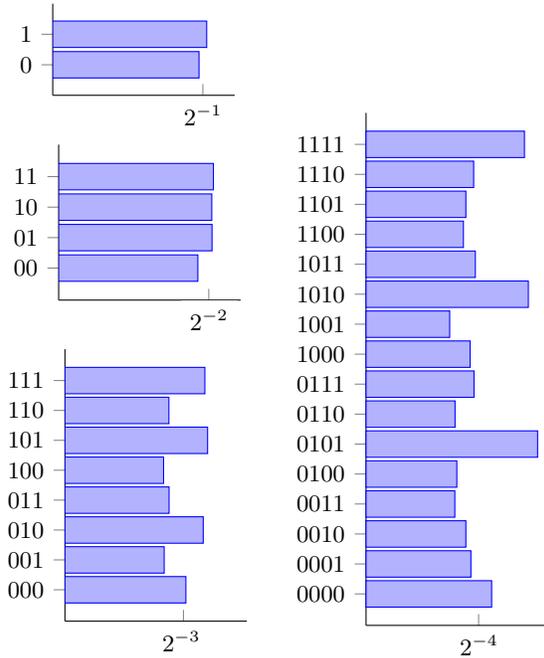


Figure 1: Frequency of sub-strings up to length 4

For the PRNG source, we chose the Mersenne Twister algorithm [47] (specifically, the second revision from 2002 that is currently implemented in *GNU Scientific Library*). The atmospheric noise was taken from *random.org* site (property of Randomness and Integrity Services Limited) which also runs real-time statistic tests recommended by the US National Institute of Standards and Technology to ensure the random quality of the numbers produced over time.

In Table 2, we summarize our results using $b = 1$ and $r = 1$ for the parameters of K as defined in section IV.i.

	<i>Participants</i>	<i>PRNG</i>	<i>AN</i>
Mean μ	48.43	52.99	53.88
Std σ	6.62	3.06	2.87
1 st quartile	45.30	50.42	51.88
Median	49.23	53.15	53.85
3 rd quartile	51.79	55.21	55.79

Table 2: Values of $K(x)$, where x is a string produced by participants, PRNG or AN sources

The mean and median of K increases when comparing participant’s string with PRNG or AN strings. This difference was significant, as confirmed by a t-test (p -value of 4.9×10^{-11} when comparing participant’s sample with PRNG one, a p -value of 1.2×10^{-15} when comparing participant’s with AN and a p -value of 1.4×10^{-2} when comparing PRNG with AN sample).

Therefore, despite the simplicity of LT^2C^2 , based merely on *prints* and *repeats*, it is rich enough to identify regularities of human sequences. The K function relative to LT^2C^2 is an effective and significant measure to distinguish strings produced by participants with profound understanding in the mathematics of randomness, from PRNG and AN strings. As expected, humans produce less complex (i.e., less random) strings than those produced by PRNG or atmospheric noise sources.

iii. Mental fatigue

On cognitively demanding tasks, fatigue affects performance by deteriorating the capacity to organize behavior [48–52]. Specifically, Weiss claim that boredom may be a factor that increases non-randomness [48]. Hence, as another test to the ability of K relative to LT^2C^2 to identify idiosyncratic elements of human regularities, we asked whether the random quality of the participant’s string deteriorated with time.

For each of the 138 strings $x = x_1 \cdots x_{120}$ ($x_i \in \{0, 1\}$) produced by the participants, we measured the K complexity of all the sub-strings of length 30.

Specifically, we calculated the average $K(x_i \cdots x_{i+30})$ from the 138 strings for each $i \in [0, 90]$ (see Fig. 2), using the same parameters as in section VI.ii. ($b = r = 1$), and compared to the same sliding average procedure for PRNG (Fig. 3) and AN sources (Fig. 4).

The sole source which showed a significant linear regression was human generated data (see Table 3) which, as expected, showed a negative correlation indicating that participants produced less complex or random strings over time (slope -0.007 , $p < 0.02$).

The finding of a fatigue-related effect shows that the unropped, i.e., resource-limited, human Turing machine is not only limited in terms of the language it can parse, but also in terms of the amount of time it can dedicate to a particular task.

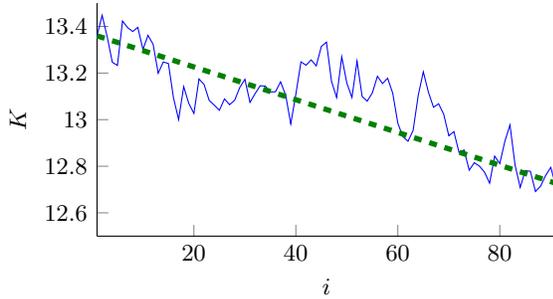


Figure 2: Average of $K(x_i \cdots x_{i+30})$ for participants

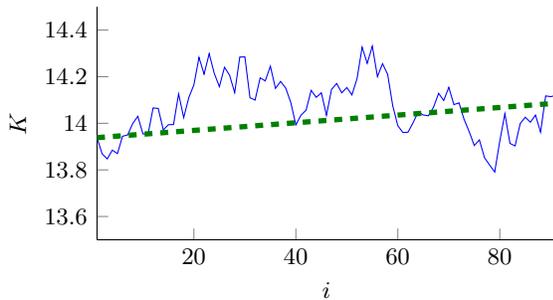


Figure 3: Average of $K(x_i \cdots x_{i+30})$ for PRNG

iv. Predictability

In section IV.ii., we introduced a prediction method with two parameters: ℓ_F and ℓ_P . A predictor based on LT^2C^2 achieved levels of predictability close to 56% which were highly significant (see Table 4). The predictor, as expected, performed at chance for the control PRNG and AN data. This fit was relatively insensitive to the values of ℓ_P and ℓ_F ,

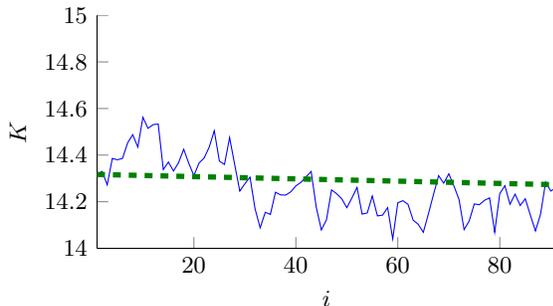


Figure 4: Average of $K(x_i \cdots x_{i+30})$ for AN

	<i>Participants</i>	<i>PRNG</i>	<i>AN</i>
Mean slope	-0.007	0.0016	-0.0005
<i>p</i> -value	0.02	0.5	0.8
CI	[-0.01,-0.001]	[-0.003,0.006]	[-0.005,0.004]

Table 3: Predictability

contrary to our intuition that there may be a memory scale which would correspond in this framework to a given length.

A very important aspect of this investigation, in line with the prior work of [23], is to inquire whether specific parameters are stable for a given individual. To this aim, we optimized, for each participant, the parameters using the first 80 symbols of the sequence and then tested these parameters in the second half of each segment (last 80 symbols of the sequence)

After this optimization procedure, mean predictability increased significantly to 58.14% ($p < 0.002$, see Table 5). As expected, the optimization based on partial data of PRNG and AN resulted in no improvement in the classifier, which remained at chance with no significant difference ($p < 0.3$, $p < 0.2$, respectively).

Hence, while the specific parameters for compression vary widely across each individual, they show stability in the time-scale of this experiment.

	<i>Participants</i>	<i>PRNG</i>	<i>AN</i>
Mean μ	56.16	50.69	49.48
Std σ	0.07	0.02	0.02
1 st quartile	49.97	48.84	48.30
Median	55.02	50.77	49.04
3 rd quartile	59.75	52.21	50.46

Table 4: Average predictability

	<i>Participants</i>	<i>PRNG</i>	<i>AN</i>
Mean μ	58.14	51.20	49.01
Std σ	0.07	0.04	0.03
1 st quartile	52.88	48.56	47.11
Median	56.73	50.72	49.28
3 rd quartile	62.02	53.85	50.48

Table 5: Optimized predictability

VII. Discussion

Here we analyzed strings produced by participants attempting to generate random strings. Participants had a profound understanding of randomness and hence avoided typical misconceptions such as exaggerating the number of alternations. We reasoned that remaining regularities would express the algorithmic nature of human thoughts, revealed in the form of specific patterns.

Our effort here was to bridge the gap between Kolmogorov theory and psychology, developing a concrete language, LT^2C^2 , satisfying the following requirements: 1) to be simple enough so that the complexity of any given sequence can be computed, 2) to be based on tangible operations of human reasoning (*printing*, *repeating*, ...), 3) to be sufficiently powerful to generate all possible sequences but not too powerful as to identify regularities which would be *invisible* to humans.

More specifically, our aim is to develop a class of languages with certain degrees of freedom which can then be fit to an individual (or an individual in a specific context and time). Here, we opted for a comparably easier strategy by only allowing the relative cost of each operation to vary. However, a natural extension of this framework is to generate classes of languages where structural and qualitative aspects of the language are free to vary. For instance, one can devise a program structure for repeating portions of (not necessarily neighboring) code, or considering the more general framework of *for-programs* where the repetitions are more general than in our setting: *for* $i=1$ *to* n *do* $P(i)$, where P is a program that uses the successive values of $i = 1, 2, \dots, n$ in each iteration. For instance, the following program

```
for i=1 to 6 do
  print '0'
  repeat i times: print '1'
```

would describe the string

010110111011110111110111111.

The challenge from the computational theoretical point of view is to define an extension which induces a computable (even more, feasible, whenever possible) Kolmogorov complexity. For instance, adding simple control structures like conditional jumps or

allowing the use of imperative program variables may turn the language into Turing-complete, with the theoretical consequences that we already mentioned. The aim is to keep the language simple and yet include structures to compact some patterns which are compatible with the human language of thought.

We emphasize that our aim here was not to generate an optimal predictor of human sequences. Clearly, restricting LT^2C^2 to a very rudimentary language is not the way to go to identify vast classes of patterns. Our goal, instead, was to use human sequences to calibrate a language which expresses and captures specific patterns of human thought in a tangible and concrete way.

Our model is based on ideas from Kolmogorov complexity and Solomonoff's induction. It is important to compare it to what we think is the closest and more similar approach in previous studies: the work [23] of Griffiths and Tenenbaum's. Griffiths and Tenenbaum devise a series of statistical models that account for different kind of regularities. Each model Z is fixed and assigns to every binary string x a probability $P_Z(x)$. This probabilistic approach is connected to Kolmogorov complexity theory via Levin's famous Coding Theorem, which points out a remarkably numerical relation between the algorithmic probability $P_U(x)$ (the probability that the universal prefix Turing machine U outputs x when the input is filled-up with the results of coin tosses) and the (prefix) Kolmogorov complexity K_U described in section II.i. Formally, the theorem states that there is a constant c such that for any string $x \in \{0, 1\}^*$ such that

$$| -\log P_U(x) - K_U(x) | \leq c \quad (4)$$

(the reader is referred to section 4.3.4 of [24] for more details). Griffiths & Tenenbaum's bridge to Kolmogorov complexity is only established through this last theoretical result: replacing P_U by P_Z in Eq. (4) should automatically give us *some* Kolmogorov complexity K_Z with respect to *some* underlying Turing machine Z .

While there is hence a formal relation to Kolmogorov complexity, there is no explicit definition of the underlying *machine*, and hence no notion of *program*.

On the contrary, we propose a specific *language*

of thought, formalized as the programming language LT^2C^2 or, alternatively, as a Turing machine N , which assigns formal *semantics* to each program. Semantics are given, precisely, through the behavior of N . The fundamental introduction of program semantics and the clear distinction between *inputs* (programs of N) and *outputs* (binary strings) allows us to give a straightforward definition of Kolmogorov complexity relative to N , denoted K_N , which—because of the choice of LT^2C^2 —becomes computable in polynomial time. Once we count with a complexity function, we apply Solomonoff's ideas of inductive inference to obtain a predictor which tries to guess the continuation of a given string under the assumption that the most probable one is the most compressible in terms of LT^2C^2 -Kolmogorov complexity. As in [23], we also make use of the Coding Theorem (4), but in the opposite direction: given the complexity K_N , we derive an algorithmic probability P_N .

This work is mainly a theoretical development, to develop a framework to adapt Kolmogorov ideas in a constructive procedure (i.e., defining an explicit language) to identify regularities in human sequences. The theory was validated experimentally, as three tests were satisfied: 1) human sequences were less complex than control PRNG sequences, 2) human sequences were non-stationary, showing decreasing values of complexity, 3) each individual showed traces of algorithmic stability since fitting of partial data was more effective to predict subsequent data than average fits. Our hope is that this theory may constitute, in the future, a useful framework to ground and describe the patterns of human thoughts.

Acknowledgements - The authors are thankful to Daniel Gorín and Guillermo Cecchi for useful discussions. S. Figueira is partially supported by grants PICT-2011-0365 and UBACyT 20020110100025.

-
- [1] M Kac, *What is random?*, Am. Sci. **71**, 405 (1983).
- [2] H Reichenbach, *The Theory of Probability*, University of California Press, Berkeley (1949).
- [3] G S Tune, *Response preferences: A review of some relevant literature*, Psychol. Bull. **61**, 286 (1964).
- [4] A D Baddeley, *The capacity for generating information by randomization*, Q. J. Exp. Psychol. **18**, 119 (1966).
- [5] A Tversky, D Kahneman, *Belief in the law of small numbers*, Psychol. Bull. **76**, 105 (1971).
- [6] W A Wagenaar, *Randomness and randomizers: Maybe the problem is not so big*, J. Behav. Decis. Making **4**, 220 (1991).
- [7] R Falk, *Perception of randomness*, Unpublished doctoral dissertation, Hebrew University of Jerusalem (1975).
- [8] R Falk, *The perception of randomness*, In: Proceedings of the fifth international conference for the psychology of mathematics education, Vol. 1, Pag. 222, Grenoble, France (1981).
- [9] D Kahneman, A Tversky, *Subjective probability: A judgment of representativeness*, Cognitive Psychol. **3**, 430 (1972).
- [10] A Tversky, D Kahneman, *Subjective probability: A judgment of representativeness*, Cognitive Psychol. **3**, 430 (1972).
- [11] T Gilovich, R Vallone, A Tversky, *The hot hand in basketball: On the misperception of random sequences*, Cognitive Psychol. **17**, 295 (1985).
- [12] W A Wgenaar, G B Kerenm, *Cance and luck are not the same*, J. Behav. Decis. Making **1**, 65 (1988).
- [13] D Budescu, A Rapoport, *Subjective randomization in one-and two-person games*, J. Behav. Decis. Making **7**, 261 (1994).
- [14] A Rapoport, D V Budescu, *Generation of random series in two-person strictly competitive games*, J. Exp. Psychol. Gen. **121**, 352 (1992).
- [15] P D Larkey, R A Smith, J B Kadane, *It's okay to believe in the hot hand*, Chance: New Directions for Statistics and Computing **2**, 22–30 (1989).

- [16] A Tversky, T Gilovich, *The "hot hand": Statistical reality or cognitive illusion?*, *Chance: New Directions for Statistics and Computing* **2**, 31 (1989).
- [17] A N Kolmogorov, *Three approaches to the quantitative definition of information*, *Probl. Inf. Transm.* **1**, 1 (1965).
- [18] G J Chaitin, *A theory of program size formally identical to information theory*, *J. AMC* **22**, 329 (1975).
- [19] L A Levin, A K Zvonkin, *The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms*, *Russ. Math. Surv.* **25**, 83 (1970).
- [20] D G Champernowne, *The construction of decimals in the scale of ten*, *J. London Math. Soc.* **8**, 254 (1933).
- [21] R J Solomonoff, *A formal theory of inductive inference: Part I*, *Inform. Control* **7**, 1 (1964); *ibid. Part II* **7**, 224 (1964).
- [22] G Chaitin, *On the length of programs for computing finite binary sequences: statistical considerations*, *J. ACM* **13**, 547 (1969).
- [23] T L Griffiths, J B Tenenbaum, *Probability, algorithmic complexity, and subjective randomness*, In: *Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society*, Eds. R Alterman, D Hirsh, Pag. 480, Cognitive Science Society, Boston (MA, USA), (2003).
- [24] M Li, P M Vitányi, *An introduction to Kolmogorov complexity and its applications*, Springer, Berlin, 3rd edition (2008).
- [25] J Feldman, *Minimization of boolean complexity in human concept learning*, *Nature London* **407**, 630 (2000).
- [26] N Chater, *The search for simplicity: A fundamental cognitive principle?*, *Q. J. Exp. Psychol.* **52A**, 273 (1999).
- [27] R Falk, C Konold, *Making sense of randomness: Implicit encoding as a bias for judgment*, *Psychol. Rev.* **104**, 301 (1997).
- [28] C P Schnorr, *Zufälligkeit und Wahrscheinlichkeit*, *Lecture Notes in Mathematics* vol. 218. Springer-Verlag, Berlin, New York (1971).
- [29] C P Schnorr, *A unified approach to the definition of a random sequence*, *Math. Syst. Theory* **5**, 246 (1971).
- [30] D Dellarosa, *A history of thinking*, In: *The psychology of human thought*, Eds. R J Sternberg, E E Smith, Cambridge University Press, Cambridge (USA) (1988).
- [31] S Carey, *The origin of concepts*, Oxford University Press, Oxford (USA) (2009).
- [32] G Boole, *An investigation of the laws of thought: on which are founded the mathematical theories of logic and probabilities*, Vol. 2, Walton and Maberly, London (1854).
- [33] A Zylberberg, S Dehaene, G Mindlin, M Sigman, *Neurophysiological bases of exponential sensory decay and top-down memory retrieval: a model*, *Front. Comput. Neurosci.* **3**, 4 (2009).
- [34] A Zylberberg, S Dehaene, P Roelfsema, M Sigman, *The human Turing machine: a neural framework for mental programs*, *Trends. Cogn. Sci.* **15**, 293 (2011).
- [35] M Graziano, P Polosecki, D Shalom, M Sigman, *Parsing a perceptual decision into a sequence of moments of thought*, *Front. Integ. Neurosci.* **5**, 45 (2011).
- [36] A Zylberberg, P Barttfeld, M Sigman, *The construction of confidence in a perceptual decision*, *Front. Integ. Neurosci.* **6**, 79 (2012).
- [37] D Shalom, B Dagnino, M Sigman, *Looking at breakout: Urgency and predictability direct eye events*, *Vision Res.* **51**, 1262 (2011).
- [38] S Dehaene, M Sigman, *From a single decision to a multi-step algorithm*, *Curr. Opin. Neurobiol.* **22**, 937 (2012).
- [39] J Kamienkowski, H Pashler, S Dehaene, M Sigman, *Effects of practice on task architecture: Combined evidence from interference experiments and random-walk models of decision making*, *Cognition* **119**, 81 (2011).

- [40] A Zylberberg, D Slezak, P Roelfsema, S Dehaene, M Sigman, *The brain's router: a cortical network model of serial processing in the primate brain*, PLoS Comput. Biol. **6**, e1000765 (2010).
- [41] L Gallos, H Makse, M Sigman, *A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks*, P. Natl. Acad. Sci. USA **109**, 2825 (2012).
- [42] L Gallos, M Sigman, H Makse, *The conundrum of functional brain networks: small-world efficiency or fractal modularity*, Front. Physiol. **3** 123 (2012).
- [43] M Costa, F Bonomo, M Sigman, *Scale-invariant transition probabilities in free word association trajectories*, Front. Integ. Neurosci. **3** 19 (2009).
- [44] N Mota, N Vasconcelos, N Lemos, A Pieretti, O Kinouchi, G Cecchi, M Copelli, S Ribeiro, *Speech graphs provide a quantitative measure of thought disorder in psychosis*, PloS one **7**, e34928 (2012).
- [45] M Sigman, G Cecchi, *Global organization of the wordnet lexicon*, P. Natl. Acad. Sci. USA **99**, 1742 (2002).
- [46] R Cilibrasi, P M Vitányi, *Clustering by compression*, IEEE T. Inform. Theory **51**, 1523 (2005).
- [47] M Matsumoto, T Nishimura, *Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator*, ACM Trans. Model. Comput. Simul. **8**, 3 (1998).
- [48] R L Weiss, *On producing random responses*, Psychon. Rep. **14**, 931 (1964).
- [49] F Bartlett, *Fatigue following highly skilled work*, Nature (London) **147**, 717 (1941).
- [50] D E Broadbent, *Is a fatigue test now possible?*, Ergonomics **22**, 1277 (1979).
- [51] W Floyd, A Welford, *Symposium on fatigue and symposium on human factors in equipment design*, Eds. W F Floyd, A T Welford, Arno Press, New York (1953).
- [52] R Hockey, *Stress and fatigue in human performance*, Wiley, Chichester (1983).