Unifying Complexity and Information

Da-guan Ke  

1Department of Biomedical Engineering, Wenzhou Medical College, Wenzhou 325035, China, 2Department of Mathematics, Zhejiang University, Hangzhou 310027, China.

Complex systems, arising in many contexts in the computer, life, social, and physical sciences, have not shared a generally-accepted complexity measure playing a fundamental role as the Shannon entropy \( H \) in statistical mechanics. Superficially-conflicting criteria of complexity measurement, i.e. complexity-randomness (C-R) relations, have given rise to a special measure intrinsically adaptable to more than one criterion. However, deep causes of the conflict and the adaptability are not much clear. Here I trace the root of each representative or adaptable measure to its particular universal data-generating or -regenerating model (UDGM or UDRM). A representative measure for deterministic dynamical systems is found as a counterpart of the \( H \) for random process, clearly redefining the boundary of different criteria. And a specific UDRM achieving the intrinsic adaptability enables a general information measure that ultimately solves all major disputes. This work encourages a single framework covering deterministic systems, statistical mechanics and real-world living organisms.

A widely-appropriate complexity measure is needed in numerous areas \(^1\text{-}^4\), for directly applications or for a unified theoretical framework like statistical mechanics \(^5\). To date, except some special cases \(^5\text{-}^8\), most calculable complexity measures are superficially classified into three types (I, II, and III) depending on three C-R relations: a monotonically ascending curve, a convex curve, and a monotonically descending curve, respectively \(^9\). Competition is mainly between the type-I and the type-II. According to many type-II supporters, highly complex systems like human brains evidently exist at a critical transition point between randomness (or deterministic chaos) and regularity called the edge of chaos or weak chaos \(^6\text{-}^9\), and then an ideal type-II measure should regard an object of weak chaos as the most complex and low-periodic objects and completely-chaotic objects the simplest, despite most type-II measures, e.g. the logical depth \(^10\), do not suffice.

There is another deeper but perhaps not strictly accurate classification. Deterministic measures are usually estimations of the incomputable concept Kolmogorov complexity \(^1\) (KC), the uncompressible information account of individual object defined as the length of the minimal computer program that regenerates the object; while statistical measures are chiefly derived from \( H \) for systems describable in probabilistic language \(^1\text{-}^4\). A quantity is called extensive if it scales (asymptotically) with the size \( r_w \) of the random word (string) \( X^{r_w} = X_1X_2\cdots X_{r_w} \) that describes the system under consideration \(^1\text{-}^4\). In an isolated system, \( H \) is extensive, whereas even some statistical measures appear not \(^5\). Then disputes arise.

Amazingly, a recently-introduced deterministic measure lattice complexity \( C_L \) exhibits intrinsic adaptability to various C-R relations \(^5\text{-}^7\), even a degree of sensitivity to weak chaos, implying an ultimate solution of all above-mentioned disputes. For a deterministic \( z \)-nary symbol string \( s \), intrinsic adaptability with a parameter \( r \) needs only treating all (overlapping) length-\( r \) words in \( s \) as \( z \)-nary symbols for specific measures’ calculation; while extrinsic adaptability needs extra variables and operations to show two existing different-type measures’ behavior alone or jointly \(^8\). Traditionally, \( r \) is related to \( r_w \) because \( s \) is often considered a collection of outcomes of a length-\( r \) random word, and a statistical form of intrinsic adaptability may help to find the proper theoretic fundamental.

In this article, \( C_L \), a statistical measure previously known as of type II \(^8\text{-}^9\), is found of somewhat intrinsic C-R adaptability. Further analysis reveals a contradiction between the adaptability and the random UDGM (r-UDGM), i.e. random process, in which entropies are exclusively rooted. With the nonlinear deterministic iterative system being identified as the deterministic UDGM (d-UDGM) that can generate any arbitrary symbol string as the traditional r-UDGM can, the C-R competition is clarified. A particular UDRM containing both r- and d-UDGM is shown to unit major competing ideas of complexity measurement naturally in an estimation of KC.

**Results**

**Deterministic adaptable complexity.** With its widely-used type-I estimation Lempel-Ziv complexity \(^10\text{-}^9\) \( (C_L) \), KC is traditionally considered a measure of randomness \(^10\text{-}^9\). Although this judgment is valid for random objects, the adaptable estimation \( C_L \) reveals more aspects of KC.

---

Received 3 December 2012
Accepted 13 March 2013
Published 5 April 2013

Correspondence and requests for materials should be addressed to D.K. (kdg@wzmc.edu.cn)
Both \( C_{LZ} \) and \( C_L \) simulate a machine reading the given string \( s \) over a finite alphabet \( \Sigma \) continuously into an unlimited memory. Alongside the reading procedure, both algorithms virtually separate \( s \) into uncompressible units and count the unit number as the complexity value of \( s \). The present unit is one such that has a present symbol just being read. It can still be compressible and then extendable.

Compressibility is reduced to duplicability in \( C_{LZ} \). If the present unit can be duplicated from any section of the exhaustive memory including the already-read part of the unit itself, the duplication operation extends the unit simply symbol by symbol until no section of the memory equals the unit. At that time the present symbol is regarded as an insertion making the present unit uncompressible and the next unit, with its first symbol, will become the present (see the example below).

In \( C_L \), a present-unit-extending mechanism prior to the duplication is the deterministic iterative map on \( \Sigma \) following either chaotic (no-symbol-repeating) rule or periodic rule. Iterations of such map are regarded as compressible as duplications, inspired by the fact that short-program-described iterative systems, e.g., logistic maps, can produce any symbolic sequence out of chaotic or periodic orbits.

For example, let \( s = 0010001100110111 \) and the dot and the sign \( "\vee" \) denote the insertion in \( C_{LZ} \) and \( C_L \), respectively. The results are as follows:

\[
s = 01 \cdot 0001 \cdot 0111 \cdot 0111; \\
s = 001 \vee 0001 \vee 1001 \vee 1010111 \vee 1.
\]

According to \( C_{LZ} \), the first symbol is always an insertion without prefix. The second has a duplicable prototype in the memory, but the third makes the unit 01 does not match the exhaustive memory 00. The fourth and fifth symbols 00 together can be duplicated from the previous symbols 0010, while the sixth can not, because 001 has no prototype in 00100. Three following uncompressible units are 11, 00110 and 0111. Since there are 6 separated units, \( C_L(s) = 6 \).

According to \( C_L \), the first two symbols 00 are generated by a 1-period iteration, but the third symbol 1 interrupts the iteration. Because 001 is not duplicable, it is an uncompressible unit. The next unit 001 is identified similarly. Concerning the third unit 1001, since the first two symbols 1 and 0 are different from each other, they should be assumed following chaotic rule; the third symbol 0 implies that a periodic rule is employed with an initial state 1. After a periodic rule is broken, neither 1001 nor its follower unit 1010111 is found duplicable. With the last unit 1 being separated, we see \( C_L(s) = 5 \).

Let parameter \( r = 2 \), any two-symbol word in \( s \) compose a refined symbol and \( C_L(s^r) = 3 \). Let \( r \geq 6 \), \( C_L(s^r) = 1 \). Indeed, as has been shown\(^{22}\), for a finite \( s \), there is a critical order \( r^* \) such that once \( r \geq r^* \), \( s^r \) can be regarded as a single iteration and then \( C_L \) can be roughly approximated by the length of a periodic rule, since \( r^* \) is composed of a type-I measure to a type-II. The objects of highest \( r^* \), with the most difficulty to obtain \( C_L \), are strings of the period-doubling accumulating points known as weak chaos\(^{23}\).

**Adaptable entropy.** In classical information theory\(^{24}\), when the probability \( p_i \) of any event \( x_i \) is obtainable, with a random variable \( X \) representing all \( x \) possible events, the \( \alpha \)-ary Shannon entropy is the mean of the information content \( -\log p_i \) of \( X \)

\[
S = \sum_{i=1}^{\alpha} p_i \log_\alpha p_i.
\]

If \( \alpha = 2 \), the unit of \( S \) is just bit; and if not, since \( \log_2 \alpha = \log_2 \log_\alpha \), one may time \( S \) by \( \log_\alpha \) to get the entropy \( H \) of bit. For an binary independent and identically distributed random word \( X_r^* = X_1 X_2 \ldots X_n \) with 2\(^r\) elemental events and any event \( X_r^* \) having the probability \( p_i^{(r)} \), 2\(^r\)-ary \( S \) becomes the entropy rate (entropy per symbol) of order \( r_w \), \( h_r \):

\[
S(X_r^*) = -\frac{\log_2 \sum_{i=1}^{2^r} p_i^{(r)} \log_2 p_i^{(r)}}{r} = \frac{1}{r_w} \sum_{i=1}^{2^r} p_i^{(r)} \log_2 p_i^{(r)} = \frac{1}{r_w} H(X_r^*) = h_r.
\]

It is well-known that \( S(X_r^*) = 1 \) for uniformly distributed \( X_r^* \). Measuring the mean of bits needed for the shortest description of the random word’s experimental outcome, \( H(X_r^*) \) is extensive because in the limit \( r_w \rightarrow \infty \) the entropy rate \( h_r \) is a constant.

Accompanied by a type-I measure \( C_L \), \( C_L^{16,17} \) is derived from “a hierarchical approach to complexity of infinite stationary strings”\(^{22}\). By **stationary**, we see that statistical properties of the strings in consideration are fixed with time or space changing, a precondition for all statistical measures’ application.

Given a binary string \( s \) of length \( n \), there are totally 2\(^n\) distinct words of length \( r \) (or \( r \)-words for short). Let \( F_r(s) \) denote the frequency of a distinct allowed \( r \)-words that really emerges in \( s \) and \( F_1(r) \) the frequency of a distinct forbidden \( r \)-words in \( s \) counted as follows: if in \( s \) (except its end) a \( (r - 1) \)-word \( s_1 s_2 \ldots s_{r-1} \) emerges \( x \) times, but no \( r \)-word \( s_1 s_2 \ldots s_{r-1} s \) emerges, we account \( s_1 s_2 \ldots s_{r-1} \) a forbidden \( r \)-word emerging \( x \) times. For example, when basic alphabet \( \Sigma = \{0,1\} \), if \( s_1 = 0 \), then we regard the frequency of the allowed word \( s_1 s_2 \ldots s_{r-1} \) as that of the forbidden word \( s_1 s_2 \ldots s_{r-1} 0 \).

With the probability being replaced by relative frequency in \( s \), \( C_L \) is the entropy rate of allowed \( r \)-words and \( C_L \) of forbidden \( r \)-words:

\[
C_1 = -\sum_i P_{ia} \log \frac{P_{ia}}{r}, \quad (3)
\]

\[
C_2 = -\sum_i P_{ib} \log \frac{P_{ib}}{r}. \quad (4)
\]

Here, \( i \) denotes each distinct \( r \)-word; \( P_{ia} = \sum P_{ia}(r) \), and \( P_{ib} = \sum P_{ib}(r) \) are relative frequencies.

Given a finite \( r \), if \( s \) contains all possible (including overlapping) 2\(^r\)-words, no forbidden \( r \)-word occurs. For a completely chaotic (random) case, as \( n \rightarrow \infty \), the number of distinct forbidden \( r \)-words \( N_f(r) \rightarrow 0 \), then \( C_L \rightarrow 0 \).

Since \( s \) is finite, \( C_L \) is a real-world estimate of \( h_r \), while \( C_L \) is roughly adaptable to type I and type II. First, the critical order \( r^* \) still works. If \( s \) is of a level of randomness, every \( r^* \)-word in \( s \) is distinct as well as every \( (r^* - 1) \)-word is. So the number of distinct allowed \( (r^* - 1) \)-word \( N_a(r^* - 1) = N_f(r^* - 1) \), thus \( C_L(r^* - 1) = C_L(r^*) \). Second, there may exist another critical word length \( r_w^* \) such that for \( r \leq r_w^* \), \( N_f(r) = 0 \) and then \( C_L \rightarrow 0 \). By increasing \( r \) from the \( r_w^* \) of a given completely chaotic object to the \( r^* \) of the same object, \( C_L \) roughly achieves a transformation from a type-II measure to a type-I.

If \( s \) is a sequence of minimum period \( m \), when \( r \geq m \), there are \( m \) different \( r \)-words of equal frequency, and then \( C_L \) can simply be predicted.

Functionally, \( C_L \) is composed of a type-I measure \( C_L \) and a type-II measure \( N_f(r) \). As \( N_f(r) \) is a low-precision version of \( C_L \) and \( N_f(r) \geq N_f(r) \), \( N_f(r) \) may cause a precision problem of \( C_L \) in showing type-II behavior.

What is really involved in \( C_L \) calculation is the frequency of every distinct allowed \( r \)-word sharing length-(\( r - 1 \)) prefix with a forbidden \( r \)-word. This means that \( C_L \) is actually applicable to \( \alpha \)-ary strings. For convenience and without loss of generality, let us assume that the strings under consideration are binary hereafter.
Intrinsic adaptability. Let us use the logistic map \( x_t = \mu x_{t-1} (1 - x_{t-1}) \) to exhibit the intrinsic adaptability. The case \( \mu = 4 \) is known as the completely chaotic (pseudo-random) object and the case \( \mu = 3.57 \) a representative sample of weak chaos. After 25000 times iteration deleted as transient, from a trajectory of \( x_t \) we get a binary symbolic sequence of length 8204 by the partition 0.5. With different \( r \), \( C_{1Z} \), \( C_{2} \), \( C_{1} \), and \( C_{2} \) for both \( \mu = 4 \) and \( \mu = 3.57 \) are calculated and shown in Fig. 1.

From Fig. 1, with \( r \) increasing remarkable symmetry can be found in behavior of the two pairs of measures, \( C_{1Z} \) and \( C_{2} \), and \( C_{1} \) and \( C_{2} \). When \( \mu = 4 \), \( r_{c}^* = 10 \), and if we scale the ordinate logarithmically, we will find \( r^* = 26 \). Although \( C_{1Z} \) and \( C_{2} \) of low \( r \) are almost equal, the difference between them increases with \( r \) until \( r \approx r^* \). On the other hand, with \( r \leq r_{c}^* \), \( C_{1} = 1 \) and \( C_{2} = 0 \). They rapidly converge when \( r \) approaching to \( r^* \) and stay equal when \( r \geq r^* + 1 \).

When \( \mu = 3.57 \), \( r_{c}^* = 3 \) [Fig. 1(d)], and \( r^* = 834 \) (not shown in Fig. 1). Actually, in a long range of \( r \) between about 26 and 834, \( C_{1} = C_{2} \). To let \( C_{2} \) act as a type-II measure, \( r \) must be within the range from 4 to 10.

With \( r \leq 3 \), the \( \mu = 3.57 \) and the \( \mu = 4 \) cases are not distinguishable by \( C_{2} \). When \( 26 > r > 10 \), \( C_{1} \) and \( C_{2} \) are both in a type-transition state. The transition range is near the word length \( r = 13 \) [Fig. 1(b)], the solution of the equation \( n = 2^r - r + 1 = 8204 \). It is easy to see that \( r_{c}^* \leq 13 \) and \( r^* \geq 13 \) are valid for any length-\( n \) string and \( r_{c}^* = r^* = 13 \) is valid only if in each overlapping distinct possible \( r \)-word just appears once.

Fixing \( r \) but letting \( \mu \) vary with \( \Delta \mu = 0.0001 \) from 3.5 to 4, results are as shown in Fig. 2. When \( r = 3 \), almost all cases in the region about \( \mu > 3.555 \) obtain \( C_{2} = 0 \) (not shown in Fig. 2); when \( r = 4 \), the zero-\( C_{2} \) region reduces to about \( \mu > 3.907 \) [Fig. 2(b)]; and when \( r = 4 \) to about \( \mu > 3.978 \) (not shown in Fig. 2). Since \( r_{c}^* \) is either 3 or 4 for most of chaotic cases, we see a serious precision problem of \( C_{2} \) of low \( r \) [Fig. 2(b)].

As \( C_{2}(10) = 0 \) for \( \mu = 4 \) [Fig. 2(c)], \( C_{2}(10) \) is roughly type-II. For any chaotic \( s \), \( C_{2} \) always rapidly converge to \( C_{1} \) once the \( r_{c}^* \) of \( s \) has been exceeded [Fig. 1]. Since 10 is certainly larger than 3 or 4, within most chaotic area except a small region very close to the point \( \mu = 4 \), \( C_{1} \) and \( C_{2} \) act similarly.

Highest \( C_{2} \) [as shown in Fig. 2(c)] is not close to the edge of chaos as highest \( C_{L} \) is [see Fig. 8 in Ref. 6]. When \( r \) is only a little higher than the \( r_{c}^* \) of \( \mu = 4 \) case, e.g. \( r = 12 \), \( C_{2} \) becomes definitely a type-I measure [Fig. 2(d)]. Thus, \( C_{2} \) has a significantly smaller range of choice of \( r \) for roughly type-II behavior than \( C_{L} \).

**Symbolic dynamical analysis and UDGM.** Symbolic dynamics for one-dimensional nonlinear iterative systems including logistic maps provides a one-to-one correspondence between any semi-infinite symbol string and the initial point of the trajectory producing the string\(^6\). A finite \( r \)-word represents a deterministic segment enclosing the initial point. Increments of \( r \) will rescale the segment into a shorter one. Therefore, the parameter \( r \) in \( C_{2} \) is also called fine-graining order, while in entropy this name may not be appropriate, as discussed below.

Table 1 shows the distribution of all possible 4-bit words in \( \mu = 3.645 \) case. If two adjacent 3-bit-prefix-sharing words both emerge or not emerge, they are ignored. Hence we get only three distinct 4-bit forbidden words 0110, 0100, and 1100, but ignore 1001, which is also adjacent to an allowed word. Moreover, since one \( r \)-word creates two prefix-sharing \( (r + 1) \)-words, when \( r = 5 \), all (virtual) segments corresponding to 4-bit forbidden words are ignored. It makes \( N_{f}(r) \) fluctuate irrelevantly to the real spatial structure in phase space. For instance, there always exist mid-position adjacent segments not being visited, but when \( r = 4, 5, \) and 6, the number of distinct forbidden words corresponding to such segments equals 2, 0, and 2, respectively. Thus the curves of \( C_{2} \) versus \( r \) can hardly be smooth except some cases of almost complete chaos [Fig. 1(b) and (d)].

Essentially, any entropy is only applicable for a random string emitting \( r \)-words with measureable stationary distribution, or the \( r \)-UDGM of arbitrary given deterministic string \( s \). To apply the \( r \)-UDGM exclusively, one has to ignore all temporal or spatial information of \( s \) unrelated to the distribution, let alone nonstationary objects of no stable distribution. For instance, given \( r = 2 \), the 2-periodic infinite string (01)\(^{\infty} \) has \( C_{2} = 0.5 \) and be regarded as a medium complex case despite its simple temporal structure.

In contrast to the entropy, \( KC \) estimations \( C_{1Z} \) and \( C_{2} \) are in itself designed for single deterministic strings. The successive process in
searching duplicable section ensures the low complexity value for simple regular strings, e.g. (01)\textsuperscript{n} of \(C_{lZ} = 3\) and \(C_l = 1\).

Due to the absence of a perspective of deterministic chaos in its process of dealing with irregular strings, \(C_{lZ}\) is still type-1. Using the terminology of \(C_l\) and \(C_{lZ}\), we may say in \(C_{lZ}\) successively counts up non-overlapping allowed words in assorted lengths, each of which is adaptively increased from 1 to such a value that the word is distinct from any part of the exhaustive memory. The word length adjusting mechanics makes \(C_{lZ}\) be much more fine-grained than \(N_s(r)\) and let the parameter \(r\) become meaningless.

Deterministic iterative systems are considered in \(C_l\) a sort of simple data-generating models. They can be regarded as logistic maps, in which fine-graining order \(r\) represents the size of the segments, or as equivalent one-bit-output binary recurrence equations, in which \(r\) represents the bit number of input. For any given \(s\), when \(r < r^*\), this sort of models is non-universal data-generating model (NUDGM), and when \(r = r^*\), become universal, i.e. the d-UDGM, and leave no space for others.

The d-UDGM can be used alone by regard \(s\) as a single trajectory of a simplest equation, whose operation assignment must be defined as regular as possible. Determining the exact optimal set of arithmetic operations and their assignment in the equation may need countless tentative calculations. What can absolutely not be reduced is the smallest bit number of input, the \(r^*\) of \(s\), representing the system’s uncompressible information.

Henceforth let \(r_w\) denote also the length of assumed random word in the r-UDGM of \(s\) to distinguish it from the fine-graining order \(r\). For a length-\(n\) \(s\), let \(C_l(\{s\}, n)\) replace \(C_l(s')\) and let \(r = 0\) mean that no segment involves iterative mappings, \(C_{lZ}\) can be viewed as a special case of \(C_l\) denoted by \(C_l(0, n)\).

**Discussion**

Logically, a quantity designed for single deterministic strings is unconditionally suitable for stationary random strings because one can calculate the quantity’s probabilistic mean, whereas the mean may not fit any individual string. The quantity’s mean of a length-\(r_w\) random string can be interpreted as the average of \(k\) individual results each computed from a length-\(r_w\) string emitted by the random string with \(k \to \infty\). To arrange such \(k\) emitted strings in a single sample time series, we must assume that the random string, or equivalently the time series, is ergodic, i.e. the relative frequency of any distinct length-\(r_w\) deterministic string in the time series equals the string’s probability: \(P_{aw}(r_w) = p_{aw}(r_w), i = 1, 2, \ldots, 2^r\).

If every \(P_{aw}(r_w)\) is determined by previously known \(p_{aw}(s)\), in an ergodic time series the arrangement of all length-\(r_w\) emitted strings can be overlapping or non-overlapping, in a certain order or disorder, which is pointless for calculating a quantity’s mean: the mean can be theoretically obtained without numerical computation. A theorem of Brudno\textsuperscript{23} states that the KC per symbol of almost all emitted strings of infinite length is equal to the entropy rate \(h\). Likewise, when fine-graining effect is absent or negligible (i.e. \(r = 0\) or 1), maximum \(C_l\) relates to emitted strings of maximum randomness and one can prove that\textsuperscript{1,21}

\[
\lim_{r_w \to \infty} \frac{C_l(0, r_w) \log_2 r_w}{r_w} = \lim_{r_w \to \infty} \frac{C_l(1, r_w) \log_2 r_w}{r_w} = \lim_{r_w \to \infty} \frac{H(r_w)}{r_w} = h. 
\]

with probability 1. Hence, in traditional statistical-mechanics language \(C_l(0, r_w) \log_2 r_w\) or \(C_l(1, r_w) \log_2 r_w\) are not only extensive but also asymptotically equal to \(H\).

In order to apply entropy to a given length-\(n\) time series, we have to assume that \(s\) is ergodic despite its real generating mechanism. Without previously-known \(p_{aw}(s)\), from \(s\) only \(C_l(r_w)\) rather than \(h\) can be obtained. To make \(P_{aw}(r_w) = p_{aw}(s)\) and \(C_l(r_w) = h\), we have to let \(r_w \ll n\). For example, if \(s\) is a given completely random object without

| Table 1 | Distribution of all possible 4-bit words in 8204-points \(\mu = 3.645\) case |
|------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| 4-bit words      | 0000 | 0001 | 0011 | 0010 | 0110 | 0111 | 0101 | 0100 | 1100 | 1101 | 1110 | 1111 | 1111 | 1110 | 1101 | 1100 |
| Types            | F   | F   | A   | F   | F   | A   | A   | A   | A   | A   | A   | A   | A   | A   | A   | A   | A   |
| \(F_{d[4]}\)     | 1   | 0   | 1   | 0   | 0   | 1   | 0   | 1   | 0   | 1   | 0   | 1   | 0   | 1   | 0   | 1   | 0   |

From left to right, all words are enumerated according to their corresponding segments’ positions in the phase space of logistic maps. The frequency of a distinct allowed 4-bit word \(F_{d[4]}\) equals the times the orbit visit the corresponding segment. In the third row, “F” means “forbidden word”, “A” means “allowed word”, and “N” means “neither forbidden word nor allowed word”.

Figure 2 | The \(C_l\) of logistic map. (a) Bifurcation diagram; (b) \(r = 4\); (c) \(r = 10\); (d) \(r = 12\).
forbidden word, we must at least let \( r_w < r_w^* \) to ensure that no distinct \( r_w \)-word has \( P_{udgm}(r_w) = 0 \neq p(i) \).

Here we encounter an unsolvable paradox: to show \( C_1 \)’s type-I behavior, i.e. let \( C_1(r_w) = C_1(r_w^*) \), it has to be valid that \( r_w \rightarrow r_w^* \), or even \( r_w \rightarrow r^* \), for a given completely random object, thus the precondition \( P_{udgm}(r_w) = p(i) \) for any entropy’s applications cannot be satisfied. Moreover, neither an r-UDGM-rooted measure nor a d-UDGM-rooted can embody the intrinsic adaptability, since in the r-UDGM fine-graing process is not only meaningless but also harmful, and in the d-UDGM \( r^* \) of \( s \) is the ultimate point rather than an example of intrinsic adaptability.

Independent of any specific data-generating model, the KC estimation concerns lossless regeneration of given data. The insertion operation in \( C_1 \) fills blanks left by any NUDGM with symbols already known from \( s \) and grants the UDRM containing this NUDGM universality. The duplication operation freely generates repeated words as the r-UDGM does, making the UDRM of \( C_1(0, n) \) a sort of quasi-r-UDGM. Except for low-period objects, \( C_1(0, n) \) shows no noticeable distinctness from \( C_1 \) in its C-R behavior. Without needing to previously set a \( r_w \ll n \) for calculating the average \( C_1(0, r_w) \log r_w \), over all allowed \( r_w \)-words in \( s \), the whole \( s \) is treated as a single emitted object for \( C_1 \), and then \( C_1(0, n) \log n \ per \ symmetric \ appearance \to \) be estimations of \( h \) even better than \( C_1^{1-28} \).

With \( r \) increasing, \( C_1 \) appears a simulator not only of \( H \) but also of \( r^* \). The intrinsic adaptability of \( C_1 \) embodies indeed a general information/complexity measure presenting a smooth transition from the r-UDGM-rooted (superficially type-I) information concept to the d-UDGM-rooted (ideally type-II) complexity concept, all consistent with the principle of KC.

The r-UDGM and the d-UDGM identified here enable us to succinctly redefine the C-R conflict and avoid unnecessary confusions caused by misuse of each UDRM, e.g. about randomness and chaos, not only in complexity measurement. Besides many well-defined deterministic dynamical systems, living organisms\(^{12-22,28} \), e.g. human brain, heart, and economic systems, appear nonstationary, edge-of-chaos, and, strictly speaking, beyond the scope of all types of r-UDGM-rooted statistical mechanics including generalised versions\(^{29} \). For these systems, a d-UDGM-rooted measure or framework is certainly an option and need further studies. However, in living organisms, randomness is not able to be excluded except that noise or free will\(^{30} \) is. Thus, a UDRM-rooted framework that proceeds from a KC-based general information measure, \( C_1 \), or its possible revised version, may have more adaptability to complicated real-world situations than a single-UDGM-rooted.

In brief, though we can separate complexity from information by using the d-UDGM and the r-UDGM alone, it would be more natural to accept a general measure encompassing \( H \) and its d-UDGM counterpart. The behavior of this measure should have been outlined by \( C_1 \), since the d-UDGM always becomes more and more overwhelming with \( r \) increasing.