

Article

Writing, Proofreading and Editing in Information Theory

J. Ricardo Arias-Gonzalez ^{1,2} 

¹ Instituto Madrileño de Estudios Avanzados en Nanociencia, C/Faraday 9, Cantoblanco, 28049 Madrid, Spain; ricardo.arias@imdea.org; Tel.: +34-91-299-8860

² CNB-CSIC-IMDEA Nanociencia Associated Unit “Unidad de Nanobiotecnología”, Cantoblanco, 28049 Madrid, Spain

Received: 5 April 2018; Accepted: 12 May 2018; Published: 15 May 2018



Abstract: Information is a physical entity amenable to be described by an abstract theory. The concepts associated with the creation and post-processing of the information have not, however, been mathematically established, despite being broadly used in many fields of knowledge. Here, inspired by how information is managed in biomolecular systems, we introduce writing, entailing any bit string generation, and revision, as comprising proofreading and editing, in information chains. Our formalism expands the thermodynamic analysis of stochastic chains made up of material subunits to abstract strings of symbols. We introduce a non-Markovian treatment of operational rules over the symbols of the chain that parallels the physical interactions responsible for memory effects in material chains. Our theory underlies any communication system, ranging from human languages and computer science to gene evolution.

Keywords: thermodynamics; stochastic; non-Markovian; memory; proofreading; editing; replication; transcription; translation

1. Introduction

Many concepts related to the treatment of information have ordinarily been used for centuries, but only rigorously defined within the context of information theory by the end of the 1940s [1,2]. Examples are information content, as quantified by the entropy, or communication load, as characterized by the channel capacity. In the same way, error is normally used as a binary concept that refers to either a correct or an incorrect transfer of a symbol. Writing, including the creation of original text in human languages or the generation of genes in biomolecular systems, transcription or translation of information, as well as any kind of evolution of an initial information chain, however, cannot be simply assessed as correct or wrong because many strings of characters may carry similar information. At the level of a single subunit, sometimes, the replacement of a particular symbol, as characterized by the value of a random variable, may be indifferent or simply may not alter the content or meaning of an information chain.

The appropriate introduction of operational rules to construct bit strings is linked to the use of the concept of proofreading and, more importantly, to the establishment of conditions for useful editing [3]. These concepts appear together in biomolecular processes associated with information management [4–6]. Monomers, either nucleotides or amino acids, are chained according to a template sequence. A replicated DNA strand or an RNA transcript thus represents a sequence of symbols that code for proteins; a protein, in turn, comprises a translated gene, an information chain that entails structural and functional fates in a cell. Physical interactions have been evolutionarily adapted in these natural systems generating a direct correspondence between the thermodynamic stability of a biopolymer chain and the transfer of its information [7,8].

Biomolecular engines that process genetic information work under non-equilibrium conditions [9], but not in ensembles in the cell. These systems are one of the most important subjects of stochastic thermodynamics [10], a relatively new extension of classical thermodynamics that deals with small systems [11]. The relations between equilibrium thermodynamic potentials and non-equilibrium work or entropy production through fluctuation theorems [12,13] and the burst of single-molecule biophysics [14], in which individual protein dynamics are experimentally followed [15], demonstrate the importance of studying individual trajectories. In this regard, it is paradigmatic that writing, proofreading and editing of information involve the composition of individual strings of symbols, therefore making the statistical treatment in stochastic thermodynamics useful in this purely informational context [16,17].

We previously formulated a thermodynamic framework for physical systems that carry information in the presence of memory effects [3] with general non-Markovian reach [18]. In this paper, we expand this framework to general systems, both symbolic and physical, conveying information, that is to systems where information content is not determined by thermodynamic laws, but able to respond to ad hoc, imposed operational rules, as is the case for human languages. We then express conditions for both effective proofreading and editing. In the second part, we apply our formalism to the process of copying information. We adapt a toy model that was formulated in the context of DNA replication and transcription and test the universal conditions for effective revision. We end with concluding remarks on our theory, including an outlook toward its alternative exploitation.

2. Analysis

The probability distribution, $\Pr\{X = x\} = p(x)$, of a random variable X with alphabet \mathcal{X} can be written as an exponential of an arbitrary real function, namely, $p(x) \sim a^{f(x)}$, with $a > 1$. For convenience, and without loss of generality, we will introduce probability distributions as:

$$p(x) = \frac{1}{Q} a^{\frac{C(x)}{R}}, \quad (1)$$

where $C(x)$ is a real function of the random variable X and $R > 0$ is a parameter that modulates the weight of function C in the exponent. Q is a normalization factor, as determined from $Q = \sum_{x \in \mathcal{X}} a^{\frac{C(x)}{R}}$. Function C is recovered from the probability distribution by taking the logarithm to base a , namely, $C(x) = R \log_a [Qp(x)]$, and therefore, $C(x)$ can be considered a transform of the probability distribution $p(x)$ that maps the image interval $[0, 1]$ to $(-\infty, R \log_a Q]$.

We next explain the convenience of expressing probabilities as in Equation (1). In particular, we provide meaning to function C in the context of a string of symbols, as characterized by a directional, stochastic chain with memory [19]. Let $v = \{x_1, x_2, \dots, x_i, \dots, x_{n-1}, x_n\}$ be a stochastic chain arranged $1 \rightarrow n$ (say forward or from left to right) by a linear sequence of symbols and/or physical objects, x_i , as values of random variables X_i . We will extend the concept of energy to general linear, stochastic chains, so that it can be useful in information theory.

Definition (Correctness). Let v be a stochastic sequence of symbolic, random variables X_i , $i = 1, \dots, n$. We associate with each X_i a real function called the correctness, C_i , following Equation (1), which, within v , depends on both the symbol x_i and the sequence of previous symbols x_{i-1}, \dots, x_1 , as $C_i \equiv C(x_i; x_{i-1}, \dots, x_1)$. We say that $C(x_i; x_{i-1}, \dots, x_1)$ is the correctness of symbol x_i provided that the previous symbols are x_{i-1}, \dots, x_1 . The correctness C_i will be assigned to symbol x_i on the basis of its precision within the chain v . Specifically, the more correct symbol x_i , the higher its correctness C_i ; the more erroneous x_i , the lower C_i .

Since the probability of a chain can be written as [20]:

$$p_v \equiv \Pr\{X_1 = x_1, \dots, X_n = x_n\} = p(\mathbf{x}) = p(x_1, \dots, x_n) = p(x_1)p(x_2|x_1) \cdots p(x_n|x_{n-1}, \dots, x_1), \quad (2)$$

it is possible to define the total correctness of a sequence, C_ν , as the sum over all the correctness of its symbols:

$$C_\nu \equiv C(\mathbf{x}) = C(x_1, \dots, x_n) = \sum_{i=1}^n C_i(x_i; x_{i-1}, \dots, x_1). \tag{3}$$

The correctness is therefore an additive property of the subunits within the chain.

The interactions, which determine a memory, between subunits in a material chain, as for example in polynucleotide biopolymers [19], may be naturally imposed by the physics of the systems [3]. In these chains, the energy plays the role of the correctness ($C \rightarrow -E$), and parameter R is associated with the thermal energy level, namely $R \rightarrow kT$, where k is the Boltzmann constant and T the temperature. This suggests the modeling of function C as a strategy to introduce ad hoc the operational rules in purely symbolic systems, like in human communication. In this case, each individual contributes differently to the correctness since his/her own capacity to assimilate the operational rules of a specific language entails different protocols for their application, as reflected in the probabilities of Equation (1). The action of a human being in text composition is analogous to those of the DNA and RNA polymerases in replication and transcription, respectively. These nanomachines introduce non-equilibrium protocols with energy dissipation, which adds to the hybridization free energies of the resulting nucleic acids base-pairs [3,19].

Note that replacing C by $C + C_0$, C_0 being a constant, affects the correctness, but not the probabilities. This arbitrariness is also present in Hamiltonian systems, where the reference for potential energies is not universal.

The individual outcomes of a multivariate stochastic process with memory or its time-evolution cannot be rated just as erroneous or correct because several degrees of accuracy may be possible. The correctness function allows extending the concepts of error and certainty (herein used as the opposite of error). Specifically, a stochastic sequence, ν , will be said to be erroneous when its associated correctness, C_ν , is low. The lower C_ν , the more erroneous ν . On the contrary, a chain ν will be said to be certain when its associated correctness, C_ν , is high. The higher C_ν , the more correct ν .

The fact that C_ν is low or high is then the result of a balance between the number of errors (wrong x_i) and certainties (correct x_i) and their impact, as measured by the individual correctness, C_i . A high correctness is such that $C_i/R \gg 1$, and a low correctness, $C_i/R \ll 1$. In any case, it is necessary that $|C_i/R| \gg 1$ for the concepts of both certainty and error to make sense. In the absence of randomness, which can be studied in the limit of very low R , the above formalism leads naturally to what is expected for probabilities, namely that $p_{certain} = 1$ and $p_{error} = 0$.

The probability of a particular configuration in directionally-constructed stochastic chains [19] can now be written as:

$$p_\nu^{(w)} = \frac{a^{\frac{C_\nu}{R}}}{Q_\nu}, \quad Q_\nu \equiv \sum_{\nu'=1}^N a^{\frac{C_{\nu'}'}{R}}, \tag{4}$$

where $\sum_{\nu'=1}^N$ stands for $\sum_{x'_1, \dots, x'_n \in \mathcal{X}}$, Q_ν is the sequence-dependent normalization, N the number of sequences and $C_{\nu'\nu}$ the two-sequence correctness defined as [19]:

$$C_{\nu'\nu} \equiv \sum_{i=1}^n C(x'_i; x_{i-1}, \dots, x_1), \tag{5}$$

with $C_{\nu\nu} = C_\nu$. Such chain-construction statistical treatment was shown to imply an irreversibility, which is inherent to the system as assessed by practical writing constraints (for example, when using a typewriting machine).

In contrast, if we do not restrict how to access each final configuration [3,19], the probability of a certain sequence reads:

$$p_\nu^{(r)} = \frac{a^{\frac{C_\nu}{R}}}{Q}, \quad Q = \sum_{\nu=1}^N a^{\frac{C_\nu}{R}}, \tag{6}$$

being Q the standard normalization. As with partition functions [3,19], the normalizations fulfill $\langle 1/Q_\nu \rangle_r = 1/Q$ and $\langle Q_\nu \rangle_w = Q$, and the probabilities $\langle p_\nu^{(w)}/p_\nu^{(r)} \rangle_r = 1$ and $\langle p_\nu^{(r)}/p_\nu^{(w)} \rangle_w = 1$, where $\langle \dots \rangle_w = \sum_{\nu=1}^N p_\nu^{(w)} \dots$ and $\langle \dots \rangle_r = \sum_{\nu=1}^N p_\nu^{(r)} \dots$.

The formalism given by Equation (6) involves a reversible process and therefore comprises the concept of revision, in which backward recognition of errors (proofreading) and their substitution by new objects (editing) take place [3,19] (as, for example, when using a word processor in a computer). Therefore, we will extend previous results, hence calling the process represented by Equation (4) writing and that represented by Equation (6) revision, thus justifying the use of superindices w and r , respectively.

We now introduce the edition potential, “ A ”, and the entropy, “ H ”, for single chains. Depending on whether we are writing or proofreading and editing, these functions read:

$$A_\nu^{(w)} \equiv R \log_a Q_\nu, \quad A_\nu^{(r)} \equiv R \log_a Q, \tag{7}$$

$$H_\nu^{(w)} \equiv -\log_a p_\nu^{(w)}, \quad H_\nu^{(r)} \equiv -\log_a p_\nu^{(r)}, \tag{8}$$

fulfilling:

$$A_\nu^{(w)} = C_\nu + RH_\nu^{(w)}, \quad A_\nu^{(r)} = C_\nu + RH_\nu^{(r)}, \tag{9}$$

which is the energy conservation analogue. It is noted that the edition potential does not depend on a particular chain, ν , for the case of revision (Equation (7), right). Besides, the correctness of a particular chain, C_ν , does not depend on how it has been constructed (whether with or without revision). Functions “ A ” and “ H ” in Equations (7) and (8) characterize the generation of information and the information content, respectively, of an individual symbolic chain under the rules modeled through function C .

Respective functionals can be constructed by taking ensemble-averages:

$$C^{(w)} \equiv \langle C_\nu \rangle_w, \quad C^{(r)} \equiv \langle C_\nu \rangle_r; \tag{10}$$

$$A^{(w)} \equiv R \langle \log_a Q_\nu \rangle_w, \quad A^{(r)} \equiv R \log_a Q; \tag{11}$$

$$H^{(w)} = -\langle \log_a p_\nu^{(w)} \rangle_w, \quad H^{(r)} = -\langle \log_a p_\nu^{(r)} \rangle_r; \tag{12}$$

where $R \log_a Q = R \langle \log_a Q \rangle$ in Equation (11), right. Equations for writing (left) and revision (right) are formally analogous. Importantly, we recover the definition of entropy in information theory for $a = 2$ [2]. These functionals characterize the average correctness, Equation (10), the generation of information, Equation (11), and information content, Equation (12), of chains constructed under the rules modeled through function C in the absence (superindex w) or presence (r) of revision.

The analogue of the energy conservation for ensemble-average phenomena yields:

$$A^{(w)} = C^{(w)} + RH^{(w)}, \quad A^{(r)} = C^{(r)} + RH^{(r)}, \tag{13}$$

which arise by formally taking expected values in Equations (9). The demonstrations of Equations (9)–(13) are straightforward from their thermodynamic analogues [3]. The following relations can also be deduced from their thermodynamic analogues [3]:

$$C^{(w)} = \left\langle \frac{\partial}{\partial B} \log_a Q_\nu \right\rangle_w, \quad C^{(r)} = \frac{\partial}{\partial B} \log_a Q; \tag{14}$$

$$H^{(w)} = \left\langle \frac{\partial}{\partial R} A_\nu^{(w)} \right\rangle_w, \quad H^{(r)} = \frac{\partial}{\partial R} A^{(r)}; \tag{15}$$

where $B \equiv 1/R$. Since $\frac{\partial}{\partial B} \log_a Q = \left\langle \frac{\partial}{\partial B} \log_a Q \right\rangle$ and $\frac{\partial}{\partial R} A = \left\langle \frac{\partial}{\partial R} A \right\rangle$, expressions on the left (writing), again, are formally analogous to those on the right (revision) in Equations (14) and (15).

The following inequalities hold [3]:

$$A^{(r)} - A^{(w)} \geq 0, \quad (16)$$

$$H^{(r)} - H^{(w)} \geq -\frac{1}{R} (C^{(r)} - C^{(w)}). \quad (17)$$

In addition, it is straightforward to see that $A^{(r)} \geq C^{(r)}$ and that $A^{(w)} \geq C^{(w)}$, which together with Inequality (16), make $A^{(r)} \geq C^{(w)}$.

These expressions allow extending the conditions for effective revision that were found for thermodynamic systems carrying information [3]:

Condition 1 (Effective editing necessary and sufficient condition).

$$\text{Effective editing} \iff H^{(r)} \leq H^{(w)}, \quad (18)$$

$$\text{Ineffective editing} \iff H^{(r)} > H^{(w)}. \quad (19)$$

Condition 2 (Effective proofreading necessary and sufficient condition).

$$\text{Effective proofreading} \iff C^{(r)} \geq C^{(w)}, \quad (20)$$

$$\text{Ineffective proofreading} \iff C^{(r)} < C^{(w)}. \quad (21)$$

The equalities take place when there exists no memory (see the independence limit theorem [19]). Note that the introduction of an arbitrary constant in the definition of the correctness (say C_0 ; see above) does not affect the formulation of these conditions.

Corollary 1 (Effective editing necessary condition and effective proofreading sufficient condition).

$$\text{Effective editing} \Rightarrow \text{Effective proofreading}, \quad (22)$$

Corollary 2 (Ineffective editing sufficient condition and ineffective proofreading necessary condition).

$$\text{Ineffective proofreading} \Rightarrow \text{Ineffective editing} \quad (23)$$

Corollaries 1 and 2 are consequences of Conditions 1 and 2 after the use of Equation (17).

3. Application: Copying

We envision a system, either natural or artificial, that computes information by assembling characters in the same way as a so-called Turing machine and according to adopted rules, as characterized by the correctness. The characters are symbolic elements from an alphabet set, with the independence of the fact that they might have material reality, like atoms or molecules in a physical realization of the system. The alphabet, \mathcal{X} , is conformed by a number $|\mathcal{X}|$ of characters correlated with symbolic operational rules, with independence of whether they might be implemented through existing interactions in a physical realization of the system. In the case of biomolecular systems in cells, intrinsic physical interactions have been accommodated to encode information through evolution. Under the action of protein motors, this information is transferred and subjected to mutations. In the case of human language, symbolic rules have been evolved throughout history, with persons interpreting them and managing information.

In the following, we model the copy of information in a binary-alphabet classical system ($|\mathcal{X}| = 2$, $a = 2$). To this end, we adapt a toy model that was previously developed in the context of DNA

replication and transcription [3,19,21]. We propose that at each memory position, the introduction of a bit value (0/1) adds a contribution $\pm\zeta$ (ζ , a real, positive parameter, which we will name stability contrast) to the correctness of the string and that this contribution is affected through the coupling with the previous bit values (memory), as modeled by a power law (see Supplementary Materials). The coupling strength is controlled by parameter α in the exponent, which increases for decreasing memory effects. Memory effects specifically introduce either a positive or negative feedback, thus increasing or decreasing the amount $\pm\zeta$. For reference, when the feedback is positive or when it is negative, but the memory effects are not too strong, a bit value that contributes with $+\zeta$ to the correctness is assimilated as a correct incorporation and a value with $-\zeta$ as an error.

The correctness in individual bit strings decreases with the number of errors for positive feedback and for negative feedback under weak memory effects (high α), as shown in Figure 1, which is expected for a rationally-designed copying system. When the feedback is negative and memory effects are sufficiently strong (low α), however, the correctness increases with the number of errors, which indicates that the correlations with the previous bit values reverse the effect of the a priori correct/incorrect bit values ($+\zeta / -\zeta$). In both positive and negative feedback cases, the correctness approaches the case of independent variables when memory effects become negligible ($\alpha \rightarrow \infty$). The curves further reveal that the correctness increases/decreases at each memory position, i , in steps of size ζ swayed by the effect of the feedback from the previous bit values in the sequence.

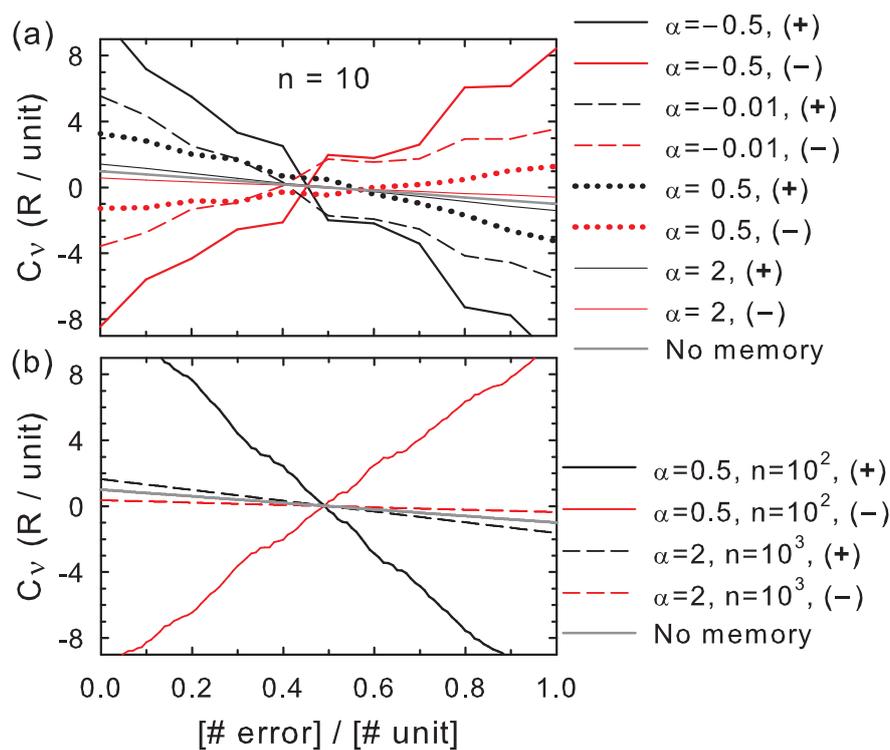


Figure 1. Correctness as a function of the number of errors. Both axes in the graphs are scaled to the number of memory units (n). The stability contrast between correct and wrong insertions is fixed to $\zeta = R$. The memory is modeled by a power law correlation between each incorporated symbol and the previous ones with exponent parameter α (Supplementary Materials). The curves show trends for individual realizations of a copying process for several α , subjected to either positive or negative feedback ((+) or (-) in the legends), for $n = 10$ (a), and $n = 100$ or 1000 (b).

The entropy difference scaled to the number of memory positions, as shown in Figure 2, ranges between $[-1, +1]$, as expected for a binary alphabet. This potential difference exhibits negative values for positive feedback, implying that revision improves the information content of the bit

string (Condition 1). In the case of negative feedback, ΔH becomes positive for weak memory effects, according to an ineffective revision (Condition 1). For strong negative feedback conditions ($\alpha \lesssim 1$), sufficiently long strings and low values of the stability contrast, this potential becomes again negative. In this case, information is generated (Condition 1) by opposites: a bit value that contributes with $+\zeta$ to the correctness generates an overall negative contribution to the correctness due to the strong memory effects (and may be now assumed as an error) and that with $-\zeta$ an overall negative contribution (and may be now considered as a correct incorporation). The information generation in this scenario is, however, of less quality than in the positive-feedback scheme: this strongly-correlated information system with negative feedback rewards the introduction of the above termed a priori errors.

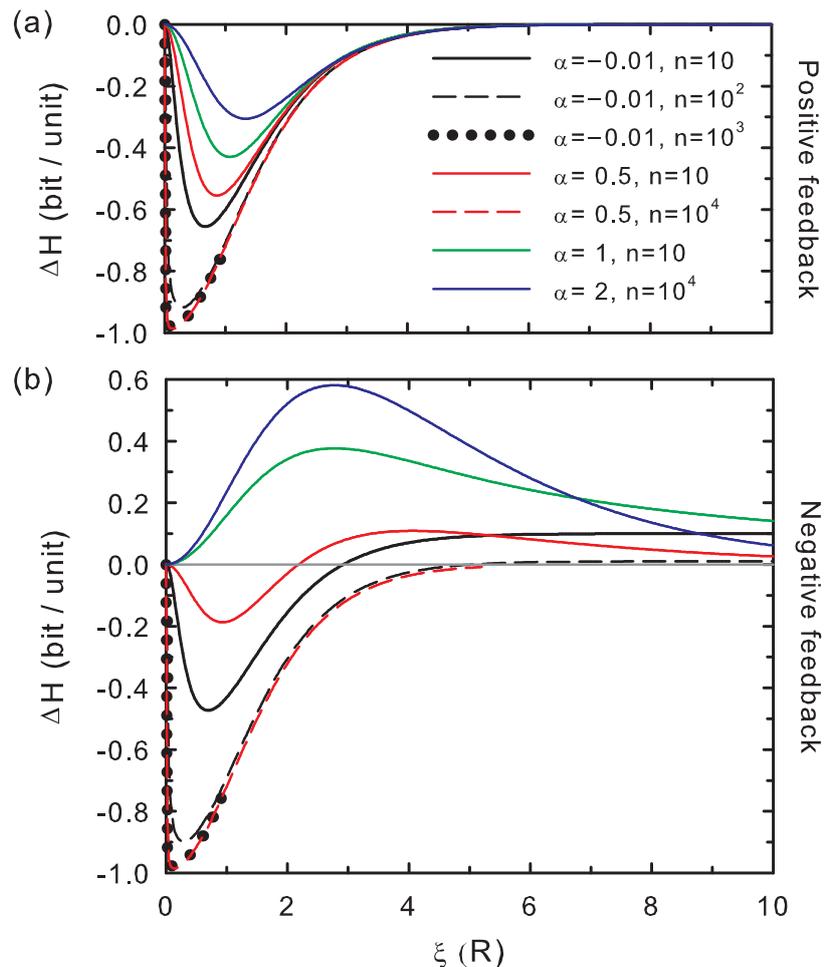


Figure 2. Entropy difference per memory unit as a function of the stability contrast between correct and wrong insertions. The memory is modeled by establishing a power law correlation between each incorporated symbol and the previous ones (Supplementary Materials). The curves show the behavior of $\Delta H = H^{(r)} - H^{(w)}$ for several chain lengths (n) and coupling strengths (α). (a) Positive feedback; (b) negative feedback.

The correctness potential difference scaled to the number of memory positions, (Figure 3) is positive for positive feedback, which indicates that the proofreading of the strings generated this way is effective (Condition 2). As explained above for Figure 2, the fidelity of the copy increases (all in all, complying with Corollary 1). For negative feedback, there are again two regimes: for weak memory effects (high α), the correctness difference becomes negative, which indicates that proofreading is ineffective (Condition 2) and makes the editing ineffective, as well (Corollary 2). In the strongly-correlated limit (low α) and negative feedback, the correctness becomes again positive. This is

compatible with revision being effective and ineffective, which, as explained for Figure 2, depends on the string length and stability contrast regime. Information, as mentioned, is generated by a counter-balance effect of the memory. This is, again, coherent with the formulated revision Condition 1 and Corollary 1.

The edition potential difference (Figure 4) is always positive, as described by Equation (16); hence, it is also useful as a control check for the simulations. The larger its values, the more distinct the bit string after editing. The curves show that ΔA increases for increasing interaction strength (i.e., decreasing coupling parameter, α) and for increasing chain length (n).

These appear to be the general trends for the absolute value of the correctness potential difference: $|\Delta C|$ overall increases for both increasing interactions (decreasing α) and chain length (n), Figure 3. Both trends are overall followed by the absolute value of the entropy potential difference (Figure 2), although for the negative feedback regime, $|\Delta H|$ must be analyzed with care along the stability contrast coordinate, ζ , for $\alpha \approx 1$, which separates the strong and weak coupling regimes.

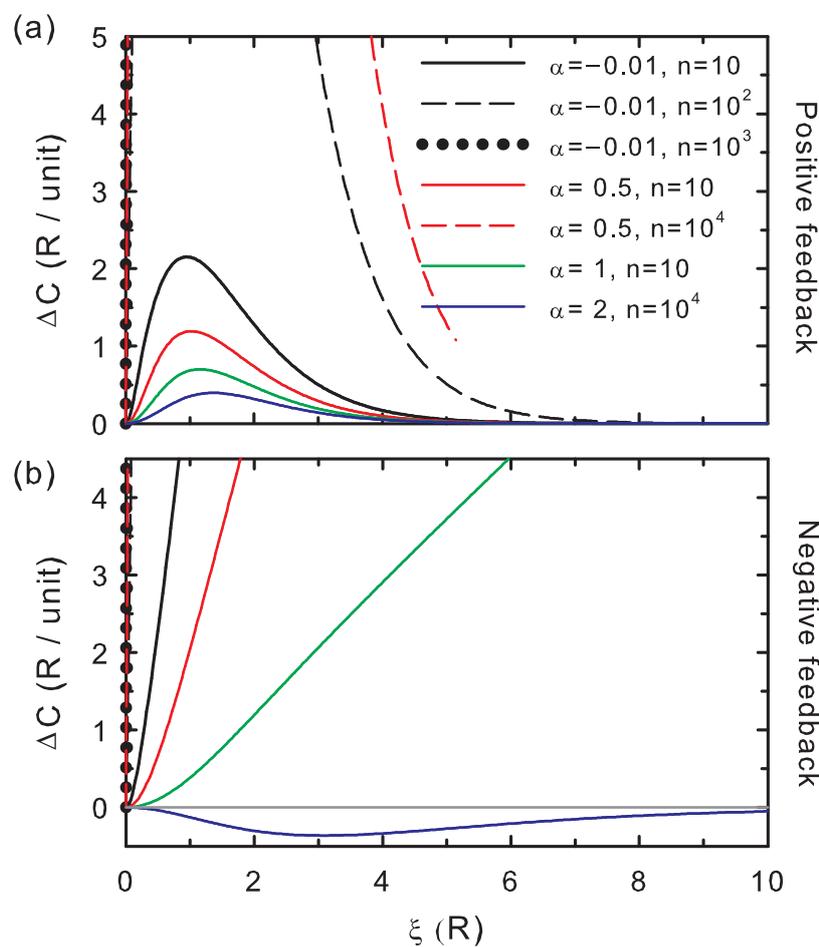


Figure 3. The same as in Figure 2 for the correctness potential difference per unit memory. (a) Positive feedback; (b) negative feedback.

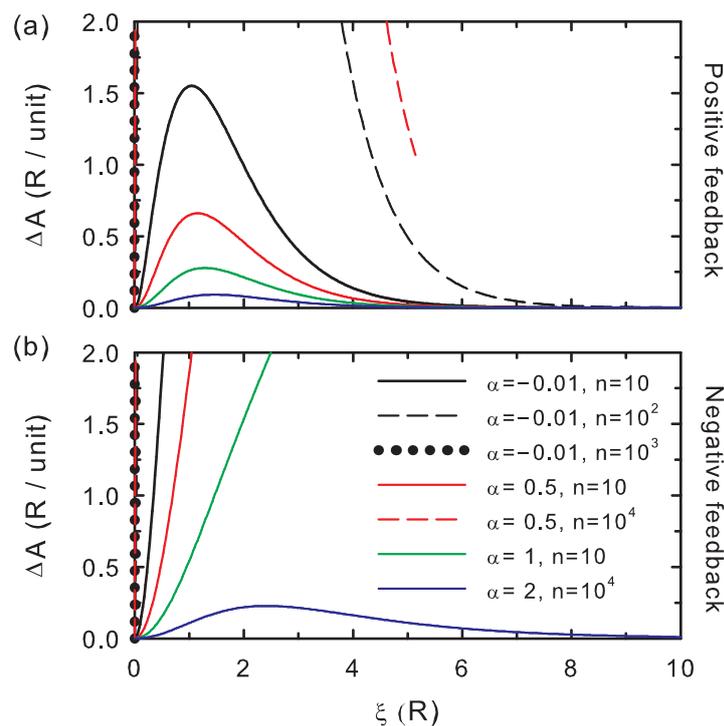


Figure 4. The same as in Figure 2 for the edition potential difference per unit memory. (a) Positive feedback; (b) negative feedback.

4. Conclusions

We have shown that information strings can be studied in analogy with the Hamiltonian treatment of stochastic chains with memory. We have introduced the correctness, which parallels the concept of energy, and have formulated conditions for effective proofreading and editing. When considering thermodynamics in a physical realization of an information system, it is important to note that time is implicitly included in our formalism, either by conceiving of indefinitely long intervals or by addressing this coordinate in the evolution of the correctness function.

Revision, as introduced in this report, entails an equilibrium thermalization in the physical analogy, hence comprising an infinite number of proofreading and editing rounds. The problem of considering one or a finite number of finite-time revisions involves a non-equilibrium modeling analogue.

This work represents a round trip: from information theory to stochastic thermodynamics, non-Markovian dynamics and molecular biophysics, and back. This trip affords for the first time the introduction of writing, proofreading and editing in information theory, with a sufficient degree of conceptualization to be used in social and natural sciences, including the physics of stochastic systems and nanoscale engineering. Our formalism enables semantic roots to bit sequences, which are important in gene evolution analysis and in context-sensitive error correction algorithms for computational linguistics, currently based on statistical occurrences.

Supplementary Materials: The supplementary materials are available online at www.mdpi.com/1099-4300/20/5/368/s1.

Author Contributions: J.R.A.-G. conceived of, performed and interpreted the study and wrote the manuscript.

Acknowledgments: This work was supported by the Spanish Ministry of Economy and Competitiveness (MINECO, Grant MAT2015-71806-R). IMDEANanociencia acknowledges support from the ‘Severo Ochoa’ Programme for Centres of Excellence in R&D (MINECO, Grant SEV-2016-0686). These funds covered the costs to publish in open access.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 623–656. [[CrossRef](#)]
2. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*, 1st ed.; Wiley-Interscience: Hoboken, NJ, USA, 1991.
3. Arias-Gonzalez, J.R. Thermodynamic framework for information in nanoscale systems with memory. *J. Chem. Phys.* **2017**, *147*, 205101. [[CrossRef](#)] [[PubMed](#)]
4. Bustamante, C.; Cheng, W.; Mejia, Y.X. Revisiting the Central Dogma One Molecule at a Time. *Cell* **2011**, *144*, 480–497. [[CrossRef](#)] [[PubMed](#)]
5. Ninio, J. Kinetic amplification of enzyme discrimination. *Biochimie* **1975**, *57*, 587–595. [[CrossRef](#)]
6. Hopfield, J.J. Kinetic Proofreading: A new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. Natl. Acad. Sci. USA* **1974**, *71*, 4135–4139. [[CrossRef](#)] [[PubMed](#)]
7. SantaLucia, J.; Hicks, D. The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 415–440. [[CrossRef](#)] [[PubMed](#)]
8. Arias-Gonzalez, J.R. Single-molecule portrait of DNA and RNA double helices. *Integr. Biol.* **2014**, *6*, 904–925. [[CrossRef](#)] [[PubMed](#)]
9. Martínez, I.A.; Roldán, E.; Dinis, L.; Rica, R.A. Colloidal heat engines: A review. *Soft Matter* **2017**, *13*, 22–36. [[CrossRef](#)] [[PubMed](#)]
10. Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* **2012**, *75*, 126001. [[CrossRef](#)] [[PubMed](#)]
11. Bustamante, C.; Liphardt, J.; Ritort, F. The Nonequilibrium Thermodynamics of Small Systems. *Phys. Today* **2005**, *58*, 43–48. [[CrossRef](#)]
12. Ritort, F. Nonequilibrium fluctuations in small systems: From physics to biology. In *Advances in Chemical Physics*; Rice, S.A., Ed.; Wiley: Chichester, UK, 2008; Chapter 2, Volume 137, pp. 31–123.
13. Hänggi, P.; Talkner, P. The other QFT. *Nat. Phys.* **2015**, *11*, 108–110. [[CrossRef](#)]
14. Bustamante, C. In *singulo Biochemistry: When Less Is More*. *Annu. Rev. Biochem.* **2008**, *77*, 44–50. [[CrossRef](#)] [[PubMed](#)]
15. Ha, T. Probing Nature’s Nanomachines One Molecule at a Time. *Biophys. J.* **2016**, *110*, 1004–1007. [[CrossRef](#)] [[PubMed](#)]
16. Parrondo, J.M.R.; Horowitz, J.M.; Sagawa, T. Thermodynamics of information. *Nat. Phys.* **2015**, *11*, 131–139. [[CrossRef](#)]
17. Strasberg, P.; Schaller, G.; Brandes, T.; Esposito, M. Quantum and Information Thermodynamics: A Unifying Framework Based on Repeated Interactions. *Phys. Rev. X* **2017**, *7*, 021003.
18. Breuer, H.-P.; Laine, E.-M.; Piilo, J.; Vacchini, B. Colloquium: Non-Markovian dynamics in open quantum systems. *Rev. Mod. Phys.* **2016**, *88*, 021002. [[CrossRef](#)]
19. Arias-Gonzalez, J.R. Information management in DNA replication modeled by directional, stochastic chains with memory. *J. Chem. Phys.* **2016**, *145*, 185103. [[CrossRef](#)] [[PubMed](#)]
20. Fisz, M. *Probability Theory and Mathematical Statistics*, 3rd ed.; Krieger Publishing Company: Malabar, FL, USA, 1980.
21. Arias-Gonzalez, J.R. A DNA-centered explanation of the DNA polymerase translocation mechanism. *Sci. Rep.* **2017**, *7*, 7566. [[CrossRef](#)] [[PubMed](#)]

