Characterizing the Asymmetry in Hardness between Synthesis and Destruction of Heteropolymers

Ikumi Kobayashi and Shin-ichi Sasa
Department of Physics, Kyoto University, Kyoto 606-8502, Japan

We present a simple model describing the assembly and disassembly of heteropolymers consisting of two types of monomers A and B. We prove that no matter how we manipulate the concentrations of A and B, it takes longer than the exponential function of d to synthesize a fixed amount of the desired heteropolymer, where d is the number of A-B connections. We also prove the decomposition time is linear for chain length n. When d is proportional to n, synthesis and destruction have an exponential asymmetry. Our findings may facilitate research on the more general asymmetry of operational hardness.

Introduction.—It is hard to construct a complex system, while it is easy to destroy it [1]. A typical example of such a system is a living organism [2]. Even though various kinds of damage may stop biological activity, it is almost impossible to bring an organism back to life once it has died [3]. One possible cause of the irreversibility of biological death is the asymmetry in the difficulty of assembling and destroying structures. To build a structure with some desirable function or property, each component must be carefully arranged in a specific order but no such care is required to destroy it. In this Letter, we attempt to characterize such asymmetry.

A similar asymmetry is formulated in computational complexity theory, where problems are classified by the amount of computational time and memory required to solve a given problem using a model computer, such as a Turing machine [4]. Concretely, the concept of NP and P characterizes the asymmetry of computation. If a problem is in NP and not P, we can efficiently (i.e., in polynomial time of the length of the input bit sequence) determine whether a given candidate is a suitable solution to the problem or not, but it takes an exponentially long time in the worst case to find the solution itself [5]. To the present, there have been many attempts to link the concepts of the theory of computation to the description of natural phenomena [6–18] and this would be an interesting research direction.

Motivated by these studies, in this Letter, we formulate the operational asymmetry between synthesis and destruction of one-dimensional heteropolymers. Because we restrict our analysis to a class of simple systems with an experimentally accessible setting, there is no direct correspondence to computational complexity theory. Nevertheless, we can show that there exists a fundamental limit to the synthesis rate that cannot be exceeded by any external operations in that setting. The results obtained for the specific model may provide useful insights into more general studies on the cost of polymer synthesis or protein synthesis under nonequilibrium environments.

We introduce an idealized description of the synthesis and decomposition of a one-dimensional molecular chain comprising two types of molecular subunits. The concentrations of molecular subunits are controlled by an external operator. The difficulty of synthesis (destruction) is quantified by the time T it takes to synthesize (destruct) a fixed amount of molecular chains with the desired sequence. In the synthesis process, we prove that T is greater than an exponential function of d no matter what external operations are performed, where d is the number of connections between different types of molecular subunits. The difficulty of synthesizing heteropolymers in this system is characterized by this quantity d. We also prove that the time required to decompose a fixed amount of molecular chains is smaller than a linear function of the chain length n. For molecules where d is proportional to n, there is an exponential asymmetry between the hardness of synthesis and destruction.

Synthesis process.—We consider the synthesis of a one-dimensional molecular chain by the sequential binding of molecules A and B to reaction nucleus X as displayed in Fig. 1. An external operator aims to synthesize the desired molecular chain as efficiently as possible by controlling the concentrations of A and B over time.

We assume the only possible reaction is the binding of molecule $Z \in \{A, B\}$ to a molecular chain M to form MZ,

$$M + Z \xrightarrow{k_Z} MZ, \tag{1}$$

where the rate constant k_Z depends only on Z [19]. For simplicity, we ignore the reverse reaction in which a bound molecule is detached. The reaction tank is connected to particle reservoirs that separately supply X, A, and B
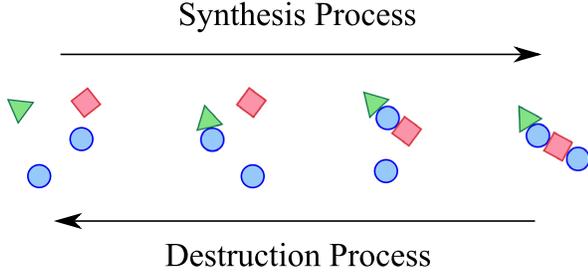
FIG. 1. Schematic illustration of the synthesis and destruction of a one-dimensional molecular chain $XABA$. The green triangle, blue circle, and red square represent the reaction nucleus X, molecule A, and molecule B, respectively. We assume that the only possible reactions are the sequential binding or detachment of the molecules.

[see Fig. 2(a) for an illustration]. The concentration of each chemical species and the temperature in the reaction chamber are spatially homogeneous, and the numbers of each molecule are large enough that the behavior of the
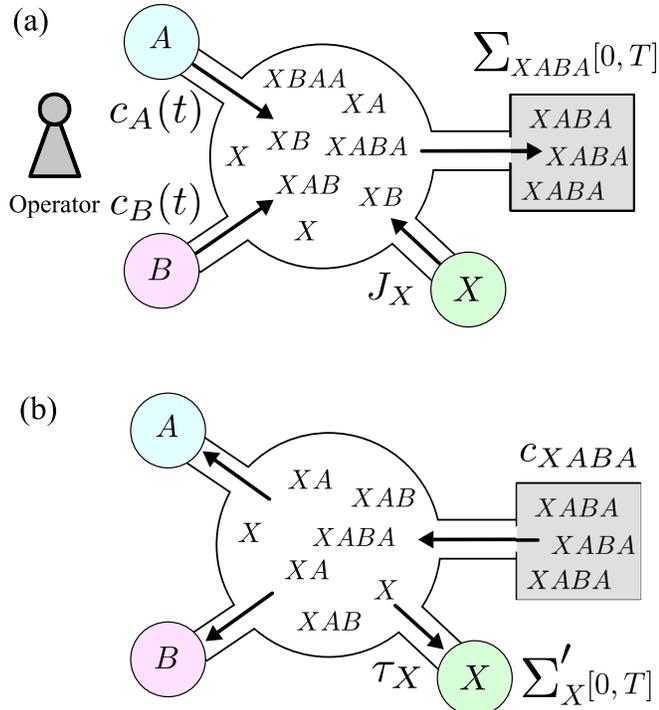

FIG. 2. Schematic illustration of the setup. (a) Setup for synthesis process. The reaction tank is connected to particle reservoirs that supply reaction nucleus X and material molecules A and B. Molecule X is supplied at a constant rate J_X. The operator's task is to synthesize the desired molecule ($XABA$ in this example) as efficiently as possible by controlling the concentrations $c_A(t)$ and $c_B(t)$. (b) Setup for destruction process. The concentration of the molecular chain to be decomposed ($XABA$ in this example) is kept constant and X is recovered to the particle bath at rate $\tau_X^{-1} c_X(t)$. The total amount of X recovered is $\Sigma_X'[0, T]$.

system is well described by the following deterministic rate equations:

$$\frac{d}{dt} c_{MZ}(t) = -k_A c_A(t) c_{MZ}(t) - k_B c_B(t) c_{MZ}(t)$$
$$+ k_Z c_Z(t) c_M(t), \quad (2)$$

$$\frac{d}{dt} c_X(t) = -k_A c_A(t) c_X(t) - k_B c_B(t) c_X(t) + J_X, \quad (3)$$

where $c_M(t)$ is the concentration of molecule M at time t. The first and second terms on the right-hand side of Eq. (2) correspond to the reactions $MZ + A \to MZA$ and $MZ + B \to MZB$, respectively; the third term represents the contribution from the reaction $M + Z \to MZ$.

The concentration dynamics in the reaction chamber can be divided into three types. First, the concentrations of molecules other than X, A, B evolve in time according to Eq. (2). Second, the dynamics of $c_X(t)$ are described by Eq. (3), where the last term represents the constant supply of reaction nucleus X from the particle reservoir to the reaction tank. Third, we assume that the concentrations of molecules A and B are controllable in time by an external operator.

At time $t = 0$, the reaction chamber is empty. The total amount of molecule MZ synthesized by reaction (1) from $t = 0$ to $t = T$ is

$$\Sigma_{MZ}[0, T] \equiv \int_0^T dt \, k_Z c_Z(t) c_M(t). \quad (4)$$

Let d be the number of connections between different types of molecular subunits. For example, $d = 2$ for $XABBA$ and $d = 5$ for $XAABABBAB$.

The first main result of this Letter is that the synthesis of a constant amount of the molecular chain takes an exponentially long time to the molecular length n unless $d = \mathcal{O}(\log n)$, no matter how the concentrations of material molecules are manipulated. Specifically, we can prove that the time T required to synthesize a certain amount $\Sigma_M[0, T]$ of molecular chain M satisfies

$$T > \frac{\Sigma_M[0, T]}{J_X} \times 2^{d/3} \quad (5)$$

for any synthesis protocol $[c_A(t), c_B(t)]_{0 \le t \le T}$.

Destruction process.—Here, we consider processes that fragment the molecular chain into monomers X, A, and B by sequential detachment of $Z \in \{A, B\}$ as displayed in Fig. 1. The possible reactions are desorption of material molecule Z from molecular chain MZ,

$$MZ \xrightarrow{k_Z'} M + Z. \quad (6)$$

Similar to synthesis process, we assume that the rate constant k_Z' depends only on Z and that the reverse reaction

(recombination) does not occur. The rate equations are then

$$\frac{d}{dt}c_{MZ}(t) = -k'_Z c_{MZ}(t) + k'_A c_{MZA}(t) + k'_B c_{MZB}(t), \quad (7)$$

$$\frac{d}{dt}c_X(t) = -\tau_X^{-1}c_X(t) + k'_A c_{XA}(t) + k'_B c_{XB}(t), \quad (8)$$

where τ_X^{-1} is the recovery rate constant of X from the reaction vessel to the particle bath [see Fig. 2(b) for an illustration]. The concentration of the molecular chain to be decomposed c_M is assumed to be constant.

At $t = 0$, the reaction chamber is empty. The total amount of molecular chains completely disassembled from $t = 0$ to $t = T$ is equal to the total amount of reactant nuclei X recovered in the particle reservoir:

$$\Sigma'_X[0, T] \equiv \int_0^T dt \, \tau_X^{-1} c_X(t). \quad (9)$$

We can then show that a linear time with respect to the molecular length n is sufficient to decompose a constant amount $\Sigma'_X[0, T]$ of the molecular chain. That is,

$$T < 2\bar{\tau} \times n + \frac{2\tau_X \Sigma'_X[0, T]}{c_M}, \quad (10)$$

where $\bar{\tau}$ is the average molecular detachment time.

Outline of the derivation.—First, we explain the outline of the proof that the synthesis reaction takes an exponentially long time to the length of the molecular chain. We focus on the processes where different types of molecules are connected. Given a specific synthesis protocol $[c_A(t), c_B(t)]_{0 \le t \le T}$, we can express $c_{MZAB}(t)$ as a function of $c_M(t)$ for any $Z \in \{A, B\}$ and show that $\Sigma_{MZABZ^*}[0, T] < \Sigma_{MZ}[0, T] \times 1/2$ for any $Z, Z^* \in \{A, B\}$. Because this relation applies to each connection [20] between A and B in the molecular chain, the upper limit to the amount synthesized becomes exponentially smaller for larger values of d and the corresponding time required to synthesize a fixed amount of the molecular chain becomes exponentially longer [21].

Next, we outline the proof that the decomposition time is shorter than a linear function of the molecular chain length. Although it is possible to directly solve Eqs. (7) and (8), we instead consider the probability that the molecular chain M is completely degraded within time t. Using Markov's inequality, we find that $c_X(t)$ has a lower bound, which leads to Eq. (10) [21].

Mechanism behind the hard synthesis and easy destruction.—The asymmetry exhibited by our model can be understood from its reaction network structure, which is a Cayley graph of depth n (see Fig. 3). Because of by-product formation, the proportion of molecules that can reach the desired vertex ($XABA$ in this case) becomes
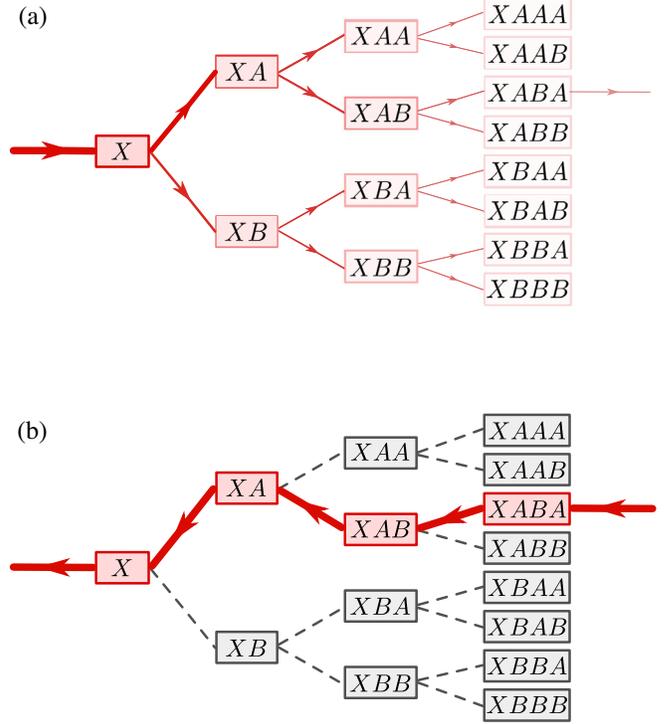
(a)

(b)

FIG. 3. Reaction network in our model. (a) Schematic illustration of the chain synthesis. The color intensity represents the number of molecules. The number of molecules decreases exponentially with the chain length because of the inevitable formation of byproducts. (b) Schematic illustration of chain decomposition. In this case, there is no by-product, and the decomposition proceeds along a single path (red solid line).

exponentially smaller than the initial amount. As a result, the synthesis rate is markedly reduced. In contrast, during chain decomposition, no by-products are produced and the reaction proceeds along a single path. Thus, decomposition only requires a time that is linear to the chain length.

Although we have focused on a specific example, the scheme shown in Fig. 3—an exponentially branched maze with only one exit—may capture a universal aspect that can be applied to more complex and general asymmetries of operational difficulty.

Concluding remarks.—We have described a simple model of the synthesis and decomposition of molecular chains composed of two types of molecular subunits. As shown in Eq. (5), the difficulty of synthesizing heteropolymers in this system is characterized by the number of connections d between different types of monomers. Equations (5) and (10) highlight the distinct asymmetry between assembly and destruction. When d is proportional to n, the time required to assemble a fixed amount of molecular chains is an exponential function of the chain length regardless of how the concentrations of the monomers are manipulated, while the corresponding function for chain destruction is linear. The result in this Letter is located at the starting point for the study of a general

question about the hardness of operations. We thus expect that the exact form of the bound given in this Letter would be useful for seeking a general principle.

In closing, we present five future challenges. First, this study ignored the reverse reaction for simplicity, but we expect from the following naive discussion that explicitly considering the reverse reaction would not affect the results. In the present model, molecular chain MAA is inevitably generated as a byproduct of MAB creation and is the cause of the slow synthesis. If we included the reverse reaction (detachment of A), we would reduce the amount of byproducts but also reduce the amount of the desired product, MAB. Therefore, we would not accelerate the rate of molecular synthesis. The analysis of the model with a reverse reaction is left as a future task.

A second problem is related to stochastic thermodynamics, where an upper bound of the current in nonequilibrium steady-state or a lower bound of the operation time to convert a probability distribution to another one is intensively studied as thermodynamic uncertain relation [22,23] and thermodynamic speed limit [24,25]. Although the settings are different in these relations and the results presented in this Letter, they share the same motivation of searching for a limit that cannot be exceeded regardless of the operational protocol. Thus, it is natural for future research to explore the relationship among them by extending the present results to stochastic systems.

A third problem is related to entropy. One may expect that our result is obtained from a fact that the thermodynamic entropy of a molecular chain is lower than that of disconnected monomer units. While it is impossible to make a lower entropy state from a higher one in thermally isolated systems, the system we study operates out of equilibrium. Thus, the thermodynamic entropy is not directly related to the asymmetry of operational hardness. Rather, our results suggest that it is not the entropy of the molecular chain but its blockiness that characterizes the efficiency of heteropolymer synthesis. When d is small, i.e., the heteropolymer contains large $A(B)$ clumps, the desired molecular chain can be synthesized relatively efficiently by controlling the concentrations of the monomers. Conversely, when d is large, i.e., the heteropolymer does not contain large clumps of $A(B)$, efficient synthesis cannot be achieved no matter how the concentrations of the monomers are varied over time. Although our results are limited to one-dimensional heteropolymers, this observation may be useful to explore the fundamental limits on the synthesis of more general polymers or proteins under nonequilibrium environments.

A fourth problem is related to feedback control. Information thermodynamics [26–28], which extends thermodynamics to include the effects of measurement and feedback control by using information-theoretic quantities, was formulated in the last decade. The system discussed in this Letter is described by deterministic rate equations and therefore does not include the concept of feedback. However, in small fluctuating systems, error correction by feedback mechanisms is expected to be important to generate complex objects [29,30]. Extending our results to stochastic systems where feedback effects become significant is an important research direction to discuss structure formation at small scales.

Finally, we remark that the synthesis of one-dimensional molecular chains has been experimentally studied [31–35]. Since the base of the exponential function in Eq. (5), $2^{1/3}$, is a model-specific parameter, its experimental determination may yield information about the underlying chemical reactions. Formulating a general relationship between chemical reaction networks and operational hardness represents the most significant future challenge.

---

[1] D. Deutsch, Constructor theory, Synthese **190**, 4331 (2013).

[2] C. Marletto, Constructor theory of life, J. R. Soc. Interface **12**, 20141226 (2015).

[3] J. L. Bernat, The definition and criterion of death, Handbook of clinical neurology **118**, 419 (2013).

[4] S. Arora and B. Barak, *Computational Complexity: A Modern Approach* (Cambridge University Press, Cambridge, England, 2009).

[5] C. Moore and S. Mertens, *The Nature of Computation* (Oxford University Press, Oxford, 2011).

[6] F. Barahona, On the computational complexity of Ising spin glass models, J. Phys. A **15**, 3241 (1982).

[7] C. H. Bennett, The thermodynamics of computation—a review, Int. J. Theor. Phys. **21**, 905 (1982).

[8] M. Mézard and G. Parisi, The Bethe lattice spin glass revisited, Eur. Phys. J. B **20**, 217 (2001).

[9] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky, Determining computational complexity from characteristic 'phase transitions', Nature (London) **400**, 133 (1999).

[10] R. Unger and J. Moult, Finding the lowest free energy conformation of a protein is an NP-hard problem: Proof and implications, Bull. Math. Biol. **55**, 1183 (1993).

[11] W. H. Zurek, Algorithmic randomness and physical entropy, Phys. Rev. A **40**, 4731 (1989).

[12] Y. Fu and P. W. Anderson, Application of statistical mechanics to NP-complete problems in combinatorial optimisation, J. Phys. A **19**, 1605 (1986).

[13] T. S. Cubitt, D. Perez-Garcia, and M. M. Wolf, Undecidability of the spectral gap, Nature (London) **528**, 207 (2015).

[14] A. Hjelmfelt, E. D. Weinberger, and J. Ross, Chemical implementation of finite-state machines, Proc. Natl. Acad. Sci. U.S.A. **89**, 383 (1992).

[15] M. O. Magnasco, Chemical Kinetics is Turing Universal, Phys. Rev. Lett. **78**, 1190 (1997).

[16] D. S. Abrams and S. Lloyd, Nonlinear Quantum Mechanics Implies Polynomial-Time Solution for NP-Complete and #P Problems, Phys. Rev. Lett. **81**, 3992 (1998).

[17] A. Lucas, Ising formulations of many NP problems, Front. Phys. **2**, 5 (2014).

[18] L. Kari and G. Rozenberg, The many facets of natural computing, Commun. ACM **51**, 72 (2008).

[19] Physically, this assumption corresponds to a situation in which the structure of the binding site at the end of the molecular chain does not change, regardless of the type of the previously bound molecule.

[20] Strictly speaking, this inequality cannot be applied to all connections, which results in a factor of $1/3$ in Eq. (5). Please refer to Supplemental Material [21] for details.

[21] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.128.247801 for a detailed derivation.

[22] A. C. Barato and U. Seifert, Thermodynamic Uncertainty Relation for Biomolecular Processes, Phys. Rev. Lett. **114**, 158101 (2015).

[23] J. M. Horowitz and T. R. Gingrich, Thermodynamic uncertainty relations constrain non-equilibrium fluctuations, Nat. Phys. **16**, 15 (2020).

[24] N. Shiraishi, K. Funo, and K. Saito, Speed Limit for Classical Stochastic Processes, Phys. Rev. Lett. **121**, 070601 (2018).

[25] S. Ito and A. Dechant, Stochastic Time Evolution, Information Geometry, and the Cramér-Rao Bound, Phys. Rev. X **10**, 021056 (2020).

[26] T. Sagawa and M. Ueda, Second Law of Thermodynamics with Discrete Quantum Feedback Control, Phys. Rev. Lett. **100**, 080403 (2008).

[27] J. M. Parrondo, J. M. Horowitz, and T. Sagawa, Thermodynamics of information, Nature (London) **11**, 131 (2015).

[28] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki, and M. Sano, Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality, Nature (London) **6**, 988 (2010).

[29] J. J. Juárez and M. A. Bevan, Feedback controlled colloidal self-assembly, Adv. Funct. Mater. **22**, 3833 (2012).

[30] P. W. K. Rothemund, N. Papadakis, E. Winfree, and A. Condon, Algorithmic self-assembly of DNA Sierpinski triangles, PLoS Biol. **2**, e424 (2004).

[31] M. Rubinstein, R. H. Colby *et al.*, *Polymer Physics* (Oxford University Press, New York, 2003), Vol. 23.

[32] C. J. Hawker and K. L. Wooley, The convergence of synthetic organic and polymer chemistries, Science **309**, 1200 (2005).

[33] C. B. Mast, S. Schink, U. Gerland, and D. Braun, Escalation of polymerization in a thermal gradient, Proc. Natl. Acad. Sci. U.S.A. **110**, 8030 (2013).

[34] S. I. Cohen, S. Linse, L. M. Luheshi, E. Hellstrand, D. A. White, L. Rajah, D. E. Otzen, M. Vendruscolo, C. M. Dobson, and T. P. Knowles, Proliferation of amyloid-β42 aggregates occurs through a secondary nucleation mechanism, Proc. Natl. Acad. Sci. U.S.A. **110**, 9758 (2013).

[35] P. A. Korevaar, S. J. George, A. J. Markvoort, M. M. Smulders, P. A. Hilbers, A. P. Schenning, T. F. De Greef, and E. Meijer, Pathway complexity in supramolecular polymerization, Nature (London) **481**, 492 (2012).