

# A Use of Variety as a Law of the Universe

**Furkan Semih Dündar**

*Department of Mechanical Engineering*

*Amasya University*

*Amasya, 05100, Turkey*

---

This paper explores the idea of a toy model universe as a character string where each character is either X or -. The string makes a transition between states that are of maximal variety. The definition of variety given by Barbour and Smolin is used in this paper. An interpretation of the toy model universe is given in terms of Everett's many-worlds interpretation. The paper also discusses whether new maximal variety strings can be obtained by addition or subtraction of maximal variety strings. A few comments are included about the use of quantum computers, which may help find the maximal variety strings more quickly.

---

*Keywords:* maximal variety; character strings; Leibniz

---

## 1. Introduction

In this paper, we consider a model universe that consists of character strings of two distinct letters, X and -. We employ this approach as a toy model. More realistic examples may include a graph or a directed hypergraph, as is the case with the Wolfram Model [1–3].

The structure of our toy universe has the concept of variety, that is, a measure of how distinctive the universe is. The idea of variety that we use is given by Barbour and Smolin [4]. It is argued that a universe should evolve in such a way as to attain maximal variety [4, 5]. Recently the idea has begun to attract more attention [6]. However, the approach of [6] is different from ours in the way that they consider an unsupervised algorithm that discovers the laws of the universe. Our approach is not based on a machine learning algorithm, and instead is based very much on the concept of variety.

Leibniz's identity of indiscernibles is a philosophical viewpoint that if two entities are equal in all of their properties, then these two entities are actually the same thing [7] (as regards the case with quantum mechanics, consult [8]). In this paper, we deal with strings that may be composed of two distinct letters, for example, X and -, that have periodic boundary conditions. Due to the periodic boundary condition, we can think of the string of characters as a discrete circle. In

our study, we consider maximal variety strings that are also Leibnizian. We define Leibnizian string configurations in Section 2.2.

The paper is organized as follows: Section 2 gives basic concepts (the idea of a time capsule and the variety of a string) that are used in the rest of the paper. Section 3 gives the *laws of motion* of the universe. Section 4 discusses the addition or subtraction of strings of maximal variety and string substitution systems. Section 5 gives an interpretation of our model in terms of Everett's many worlds. Finally, Section 6 concludes the paper.

## 2. Preliminaries

---

This section gives some background material.

### 2.1 Time Capsules

It is appropriate to begin this subsection with a quote from Barbour [9]: “By a time capsule, I mean any fixed pattern that creates or encodes the appearance of motion, change or history.” So in essence, a time capsule is a configuration that, when observed, creates the illusion of past or motion.

Fossil records are good examples of time capsules, as mentioned in [9]. When a scientist looks at a fossil and investigates its properties, the conclusion is that such and such a creature lived a long time ago. In our model, time capsules will be persistent structures (substrings) shared among states of the universe. We find that some time capsules are present in all the states, whereas others may be present only in a subset of states of the universe.

### 2.2 Variety of a String

The variety of a character string refers to how distinguishable the string is. For example, the string XXXXX has zero variety and the string X-X-X has more variety. The definition we use, mentioned as a suggestion by David Deutsch, is included in [4]. We briefly summarize the main points of variety given in [4]. For clarity, the notation is kept mostly the same.

Let  $s$  be a string of length  $N$ . We consider that the string is circular; that is, the next character after the last character is the first character. Also let us index the first element of  $s$  by 0. We denote by  $\mathcal{N}_i^m$  the substring that is constructed of letters at index positions between  $i - m$  and  $i + m$ . If the index is not in the range  $0 \dots N - 1$ , we take the given index modulo  $N$ . For example, for  $s = \text{XX-X--}$ ,  $\mathcal{N}_1^1 = \text{XX-}$  and  $\mathcal{N}_0^1 = \text{-XX}$ .

Since the marked point, that is  $i$  in  $\mathcal{N}_i^m$ , is always at the center, the notion of string isomorphism given in [4] is that if  $s_1$  and  $s_2$  are two

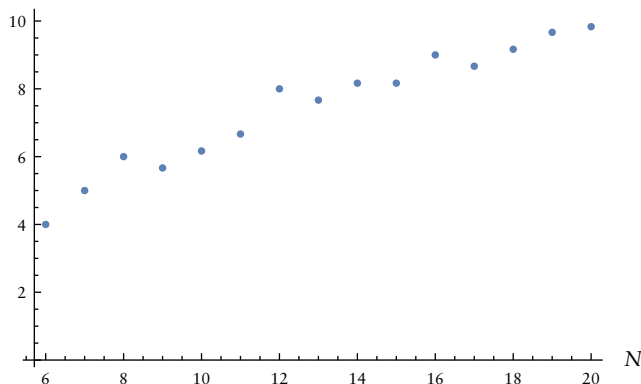
strings, they are isomorphic if  $s_1 = s_2$  modulo cyclic rotations or  $s_1$  is the mirror image of  $s_2$  modulo cyclic rotations. Then we define “ $K_{ij}$  as the smallest  $m$  such that  $N_i^m$  is *not* isomorphic to  $N_j^m$ ” [4] (emphasis in the original). The maximum value of  $m$  is  $m^*$  defined as: if  $N$  is even,  $m^* = N/2 - 1$ , and if  $N$  is odd,  $m^* = (N - 1)/2$ . If for all  $i \neq j$   $K_{ij} \leq m^*$ , we call these configurations *Leibnizian*, and *non-Leibnizian* otherwise.  $K_{ij}$  is called *relative indifference*. The higher it is, the less distinct  $s(i)$  and  $s(j)$ . Using  $K_{ij}$ , we can also define the *absolute indifference*  $R_i$ , which is the maximum of  $K_{ij}$  where  $i$  is kept fixed and  $j$  runs through all indices that are not equal to  $i$ . Then we can give the definition of variety:

$$V' = \sum_i \frac{1}{R_i}. \quad (1)$$

The free and open source program [10] written in Haskell by the author has the function `varietyP` to calculate the variety of a string, among other useful functions. This definition is not unique. The key point that the variety function must satisfy is that it should increase as  $R_i$  decreases and should decrease as  $R_i$  increases. In order to compare the variety of different strings, we may also use a *scaled* version of variety, that is, variety per character, defined as  $V'/N$ .

Table 1 lists the maximum variety for strings with  $N = 6, \dots, 20$ , and Figure 1 shows a plot of these values. Note that there are no Leibnizian strings for  $N < 6$ , thus we started with  $N = 6$ . Figure 2 shows a

Maximum Variety

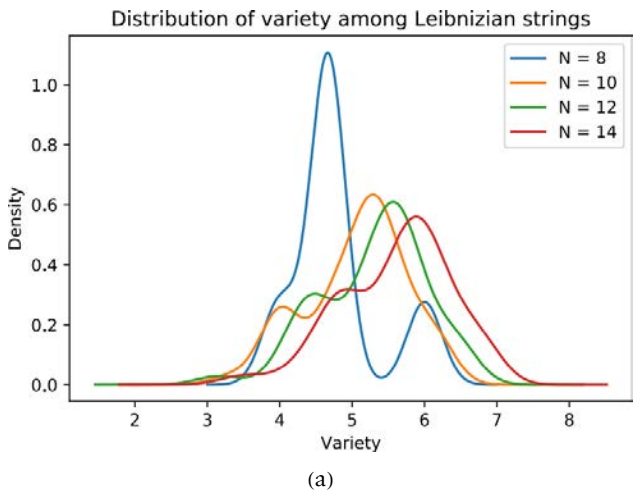


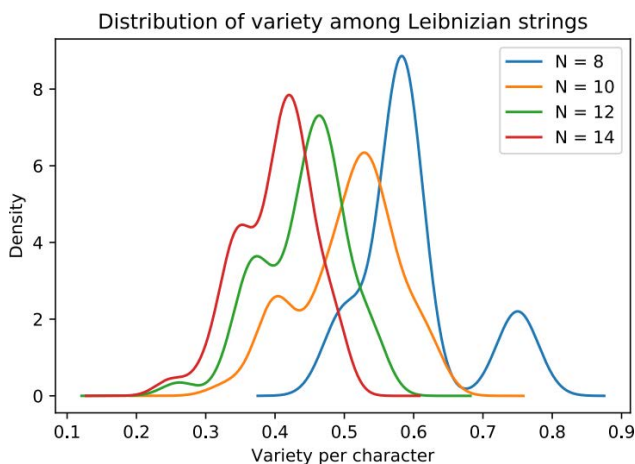
**Figure 1.** List plot of maximum variety values for Leibnizian string of length  $N$ , for rather small  $N$ . As a trend, we see that maximum variety approximately increases as  $N$  increases. Table 1 has the exact values.

$N$	Maximum Variety
6	4
7	5
8	6
9	$17/3$
10	$37/6$
11	$20/3$
12	8
13	$23/3$
14	$49/6$
15	$49/6$
16	9
17	$26/3$
18	$55/6$
19	$29/3$
20	$59/6$

**Table 1.** Maximum variety values for Leibnizian strings of length  $N$ .

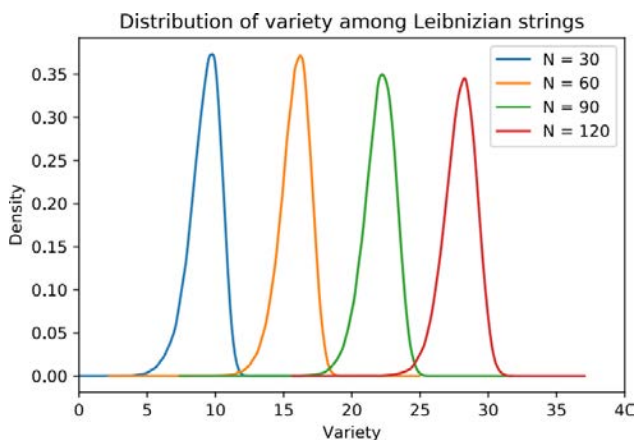
plot of the exact distribution of variety among Leibnizian strings of lengths  $N = 8, 10, 12, 14$ . The plot shows that as  $N$  increases, the mean of the density of variety increases and variety per character decreases. Figure 3 shows a random sample of about  $10^5$  entries of Leibnizian strings of lengths  $N = 30, 60, 90, 120$  since exact calculations could not be done. In this case, the mean of the distribution of variety moves to larger values, whereas the variety per character decreases and has sharper peaks as  $N$  increases.





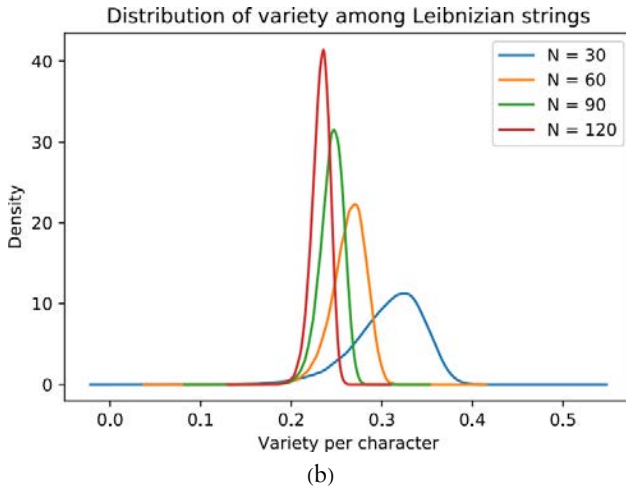
(b)

**Figure 2.** Two plots showing distribution of variety for Leibnizian character strings of lengths  $N = 8, 10, 12, 14$ . The results are exact in the sense that we considered all of the Leibnizian character strings of before-mentioned lengths. The value of maximal variety for  $N = 8$  is 6, for  $N = 10$  is  $37/6$ , for  $N = 12$  is 8 and for  $N = 14$  is  $49/6$ .



(a)

**Figure 3.** (*continues*)



**Figure 3.** Two plots showing distribution of variety for Leibnizian character strings of lengths  $N = 30, 60, 90, 120$ . We could not exhaust the search space and rather made a statistical analysis by taking about  $10^5$  Leibnizian character strings for each  $N$ .

### 3. History as a Maximal Variety String Tree

A maximal variety character string is defined as a string of length  $N$  such that it is Leibnizian and has the maximum-value variety in the state space, which is the space of strings of length  $N$  where each letter is either  $X$  or  $-$ . For instance, for  $N = 7$  modulo symmetries (cyclic and mirror), there is only one maximal variety string  $XX-X--$ , which is found by our code [10], and it is given in [4] as well. As is done in [4], we can think of the letter  $X$  as matter and the letter  $-$  as space, as a rough analogy.

In our model, we begin with a maximal variety string such as given above for  $N = 7$  and find the next strings by a best-matching procedure, similar to the one given in [11] for a system of  $N$  particles in Euclidean space.

We now define the best matching for strings. Let  $s_1, s_2$  be two cyclic strings of lengths  $N$  and  $N + 1$  (or  $n$  and  $m$  in general). The *distance* between them is defined as: keep one string fixed and rotate the other and calculate at each step the Levenshtein edit distance. The minimum value of the Levenshtein edit distance while rotating one string is defined as the distance between two cyclic strings  $s_1$  and  $s_2$ . The Levenshtein edit distance [12] between two strings is the minimum number of operations (deletion, addition and substitution of

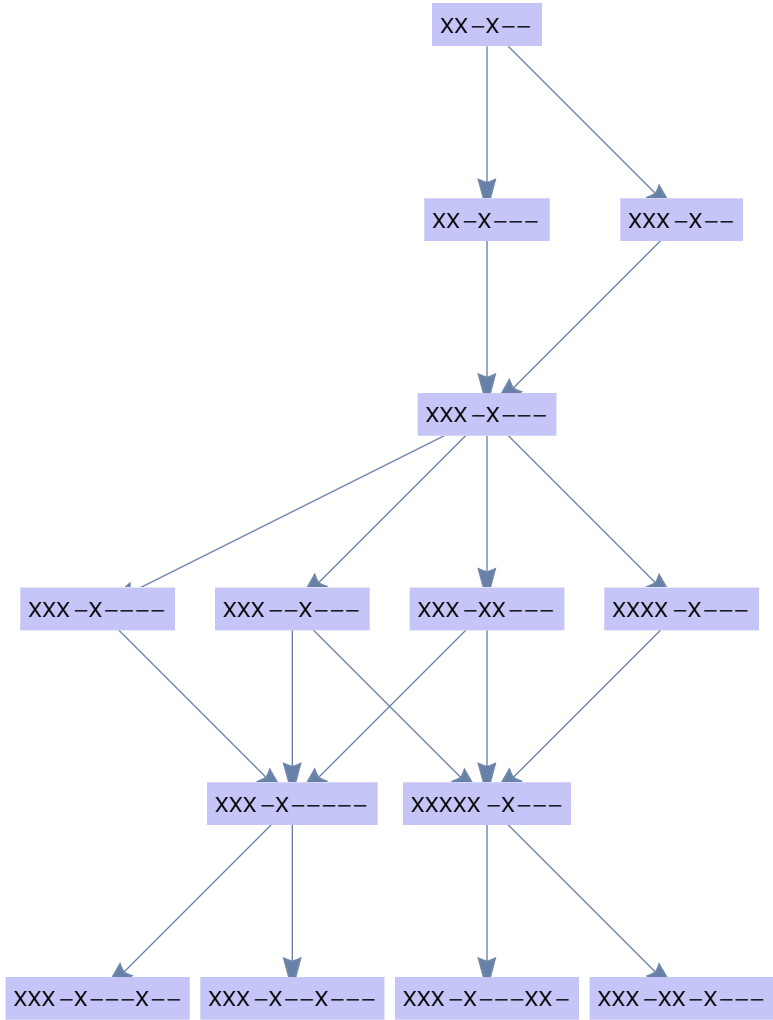
characters) to obtain one string from another. Our best-matched distance is known in the literature as the edit distance for cyclic strings. The algorithm used in [10] has complexity  $O(N^3)$ ; however, since we do not deal with large strings in this paper, the main idea in [10] fits well for our purposes. For example, the Maes algorithm [13] has complexity of  $O(N^2 \log N)$ . For more efficient algorithms on cyclic edit distance, consult [14–16]. See Figure 4 for a simple illustration of the idea. On the other hand, there are applications of cyclic edit distance to biology; for references consult [16].

XX -	-XX	X - X
X -	X -	X -
(a)	(b)	(c)

**Figure 4.** The best-matching procedure of two strings, X- and XX-. (a) The Levenshtein edit distance is 1. The first string is obtained by adding X in the middle of X-. (b) The distance is 2; first add - at the beginning of the second string and change the last character to X. (c) Finally, the distance in this case is 1; just add X at the end of the second string. Hence we conclude that the cyclic edit distance between the strings XX- and X- is 1. It is the minimum of edit distances, as one string is rotated and the second string is kept the same, and is called the cyclic edit distance.

The time evolution of our toy model universe is as follows. Fix a maximal variety string of length  $N$  and name it  $s_N$ . Then let  $S_{N+1}$  be the set of maximal variety strings of length  $N + 1$ . Then the states that come after  $s_N$  are those elements of  $S_{N+1}$  such that they best-match with  $s_N$ ; that is, they have the lowest cyclic Levenshtein edit distance to  $s_N$ . It is certain that at least one string of  $S_{N+1}$  will best-match, but there may be others as well, depending on  $N$ , the definition of variety and the definition of best-matching if other definitions are used.

The history of a physical system in our model is a tree of maximal variety character strings that are best-matched to the previous string. Figure 5 shows an illustration of the tree beginning with  $N = 6$ .



**Figure 5.** The history of a model universe from  $N = 6$  to  $N = 11$ .

**4. String Algebra and String Substitution Systems**

This section considers adding and subtracting maximal variety strings (string algebra) and string substitution systems.

**4.1 String Algebra**

We first consider whether adding or subtracting strings of maximal variety will again produce maximal variety states. It is easy to calculate variety for addition. When the string  $XX-X-$  is added to the string



--XX-X, we obtain the string XX--X---XX-X, which is of maximal variety. However, there is a counterexample as well. For  $N = 6$ , the string XX-X- is a maximal variety string. When we consider injecting this string into itself at all possible places, we obtain at most two Leibnizian strings of length 12. They are XXXX-X---X- and XX-XXX-X---- and have variety  $19/3$ . The maximum variety for  $N = 12$  is 8. The addition of two maximal variety strings did not produce a maximal variety string. However, this variety is still larger than the mean variety of distribution of strings of length 12 (see Figure 2).

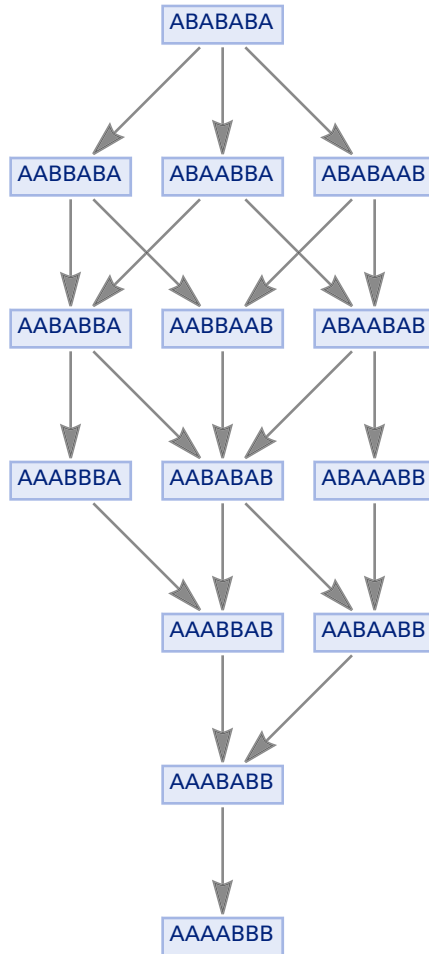
The case for string subtraction is more complicated. For that purpose, we need to find a maximal variety substring of a longer maximal variety string. X-XX---X-XX is a maximal variety string for  $N = 12$ ; it begins with the substring X-XX--, which is a maximal variety substring of length 6. After the subtraction, we obtain -X-XX, which is also a maximal variety string of length 6. On the other hand, there is the following counterexample. XX-X----XX-X is a maximal variety string. The initial XX-X- is also of maximal variety; however, the subtracted part, that is, the string ----XX-X, has variety  $14/3$ , which is not of maximum variety. Hence we conclude that although there might be cases where adding or subtracting two maximal variety strings may provide a maximal variety string, this is not always the case, as counterexamples are provided for both situations.

## ■ 4.2 String Substitution Systems

String substitution systems, sometimes referred to as string rewriting systems or semi-Thue systems [17], are a class of models such that a transformation rule is applied to a string and more strings are produced if a rule matches a substring. As a simple example, consider the following: by starting from the string ABABABA, apply the rule  $BA \rightarrow AB$  and at the end obtain the sorted string. See Figure 6 for an illustration. If the rule matches more than one substring in a string, branching occurs. This is what is called a *multiway system* [18]: the time evolution of a string substitution system is a tree.

Let  $R$  be a rewrite rule that includes these operations:  $\varepsilon \rightarrow X$ ,  $\varepsilon \rightarrow -$ ,  $X \rightarrow \varepsilon$ ,  $- \rightarrow \varepsilon$ ,  $X \rightarrow -$ ,  $- \rightarrow X$ , where  $\varepsilon$  is the empty string. These operations are used in determining the Levenshtein edit distance. Hence we can start from the empty string  $\varepsilon$  and obtain all possible strings that are composed of characters  $X$  and  $-$ . Maximal variety states are also produced. Because in our time evolution of the universe the length of the maximal variety strings is increased by one at each step, there is a possibility that the tree of strings that is obtained by starting from  $\varepsilon$  and applying the rule  $R$  contains the time evolution of the universe as a subtree. Since the edit cost that occurs due to application of  $R$  is one and the minimum edit distance between two consecutive maximal

variety strings in the universe is at least one (because a string of length  $N + 1$  is produced after a string of length  $N$  the minimum possible edit distance is one), this is not certain. However, we can surely say that the time evolution of the universe is contained in the tree generated by an infinite number of applications of  $R$  to  $\varepsilon$ . The next section has more about our interpretation of the universe.



**Figure 6.** An example string substitution system. The rule is  $BA \rightarrow AB$ . After six steps, the initial string is sorted. In this case, there is no periodic boundary condition.

## 5. How to Interpret the History of the Universe

In our model, the size of the universe (i.e., the length of the string) increases by one at each step. We could also start from a bigger universe and then contract the string length. So this model, according to how the increment or decrement of string lengths is handled, may give rise to linear expansion, accelerated expansion and collapse scenarios of various sorts. However, if  $N$  is constant, there can be no dynamics in the current model if there is only one maximal variety string of length  $N$ , and possibly a simple time evolution if there is more than one maximal variety string of length  $N$ .

As mentioned earlier, the definition of variety may be changed, subject to certain conditions. Such a change alters the state of maximal variety strings for a fixed  $N$ . Hence it is a kinematical change. The definition of cyclic edit distance between two cyclic strings may also change, and this gives rise to a different time evolution. So this change is dynamical. Although we tried to give canonical choices for kinematics and dynamics, the definitions can change and move with different choices to model another universe with different kinematics and dynamics.

### 5.1 Everett's Many Worlds Interpretation

One possible way of interpreting the universe in this model is by the Everett's many-worlds interpretation [19]. According to this interpretation, every way of traversing the string tree is realized. In the usual depiction of the many-worlds idea, the universe splits into possibly innumerable different worlds, which then split into many others. This is true in our model, and additionally note that sometimes the paths may converge, as can be seen in Figure 5.

### 5.2 A Few Time Capsules in Our Model

Since any maximal variety strings cannot be uniform—for example,  $XX...XX$  or  $--...--$  have zero variety—all of them will share the substring  $X-$ . This substring can be interpreted as an atom or the cosmic microwave background. In Figure 5, it can be seen that the pattern  $XX-$  is a time capsule that is shared by all the strings in the history of the model universe, as far as it is calculated. This pattern can be regarded as some sort of molecule. As the fossil analogy suggests, a time capsule may persist for some time and then degrade. It is an open problem to find and classify time capsules in our model.

## 6. Conclusion

In this paper, we considered a toy model universe that is a character string of two distinct letters. We considered the idea of maximal variety of character strings and using a best-matching procedure, we put forward the law of motion of the universe that it hops from one maximal variety state to other maximal variety states. We then gave an interpretation of our model in terms of Everett's many worlds.

On a computer with an i5 processor and 8 GB of RAM, the maximum variety of strings of length 6 to 20 can be calculated. Perhaps this bound may be surpassed by writing an optimized C code program; however, for our illustrative purpose, what has been done is sufficient. On the other hand, if a suitable algorithm is found, quantum computers may bring an advantage. When the qubits are initialized as  $|0\rangle$  and Hadamard gates are applied to each qubit, the final state is the superposition of each possible string configuration with equal coefficients. If we can find a gate that maps  $|\psi\rangle$  to  $V'(\psi)|\psi\rangle/A$ , where  $A$  is some complex number that stands to normalize the wavefunction and  $V'(\psi)$  is the variety of the configuration  $\psi$ , then we can also find the maximal variety strings among all states. (Due to unitarity of quantum gates, the variety of uniform strings cannot be zero; we may assign uniform strings the lowest allowed value.) The maximal variety strings will be the ones for which there are many observations. If a quantum algorithm, similar to amplitude amplification, is found that increases the amplitudes of higher-variety strings, there may be a significant speedup for finding the value of maximum variety and the maximal variety strings. The only caveat is that the length of the strings must be less than or equal to the number of qubits on a quantum computer.

## Acknowledgments

The author thanks Enes Burak Dündar for useful discussions and the anonymous referees whose comments improved the paper. Figures 1, 5 and 6 are plotted with Mathematica 12.1 [20], Figures 2 and 3 are plotted with Pandas [21] and Figure 4 is plotted with TikZ&PGF [22].

## References

---

- [1] S. Wolfram, “A Class of Models with the Potential to Represent Fundamental Physics,” *Complex Systems*, 29(2), 2020 pp. 107–536.  
doi:10.25088/ComplexSystems.29.2.107.
- [2] J. Gorard, “Some Quantum Mechanical Properties of the Wolfram Model,” *Complex Systems*, 29(2), 2020 pp. 537–598.  
doi:10.25088/ComplexSystems.29.2.537.
- [3] J. Gorard, “Some Relativistic and Gravitational Properties of the Wolfram Model,” *Complex Systems*, 29(2), 2020 pp. 599–654.  
doi:10.25088/ComplexSystems.29.2.599.
- [4] J. Barbour and L. Smolin, “Extremal Variety as the Foundation of a Cosmological Quantum Theory.” [arxiv.org/abs/hep-th/9203041](https://arxiv.org/abs/hep-th/9203041).
- [5] J. Barbour, “The Deep and Suggestive Principles of Leibnizian Philosophy,” *The Harvard Review of Philosophy*, 11(1), 2003 pp. 45–58.  
doi:10.5840/harvardreview20031112.
- [6] S. Alexander, W. J. Cunningham, J. Lanier, L. Smolin, S. Stanojevic, M. W. Toomey and D. Wecker, “The Autodidactic Universe.”  
[arxiv.org/abs/2104.03902](https://arxiv.org/abs/2104.03902).
- [7] P. Forrest, “The Identity of Indiscernibles,” *The Stanford Encyclopedia of Philosophy*, Winter 2020 Edition (E. N. Zalta, ed.), Metaphysics Research Lab, Stanford University, 2020.  
[plato.stanford.edu/archives/win2020/entries/identity-indiscernible](https://plato.stanford.edu/archives/win2020/entries/identity-indiscernible).
- [8] S. French, “Identity and Individuality in Quantum Theory,” *The Stanford Encyclopedia of Philosophy*, Winter 2020 Edition (E. N. Zalta, ed.), Metaphysics Research Lab, Stanford University, 2020.  
[plato.stanford.edu/archives/win2019/entries/qt-idind](https://plato.stanford.edu/archives/win2019/entries/qt-idind).
- [9] J. Barbour, *The End of Time: The Next Revolution in Physics*, New York: Oxford University Press, 2000.
- [10] F. S. Dünder. “String Variety.” (Apr 7, 2022)  
[github.com/fsdunder/string-variety](https://github.com/fsdunder/string-variety).
- [11] J. Barbour, “Shape Dynamics: An Introduction.”  
[arxiv.org/abs/1105.0183](https://arxiv.org/abs/1105.0183).
- [12] V. I. Levenshtein, “Binary Codes Capable of Correcting Deletions, Insertions, and Reversals,” *Soviet Physics Doklady*, 10(8), 1966 pp. 707–710.
- [13] M. Maes, “On a Cyclic String-to-String Correction Problem,” *Information Processing Letters*, 35(2), 1990 pp. 73–78.  
doi:10.1016/0020-0190(90)90109-B.
- [14] J. Gregor and M. G. Thomason, “Dynamic Programming Alignment of Sequences Representing Cyclic Patterns,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(2), 1993 pp. 129–135.  
doi:10.1109/34.192484.

- [15] A. Marzal and S. Barrachina, “Speeding Up the Computation of the Edit Distance for Cyclic Strings,” in *Proceedings 15th International Conference on Pattern Recognition (ICPR-2000)*, Barcelona, Spain (A. Sanfeliu, J. J. Villanueva, M. Vanrell, R. Alquézar, J. Crowley and Y. Shirai, eds.), Los Alamitos, CA: IEEE Computer Society, 2000 pp. 891–894. doi:10.1109/ICPR.2000.906217.
- [16] L. A. K. Ayad, C. Barton and S. P. Pissis, “A Faster and More Accurate Heuristic for Cyclic Edit Distance Computation,” *Pattern Recognition Letters* 88, 2017 pp. 81–87. doi:10.1016/j.patrec.2017.01.018.
- [17] “Semi-Thue system.” *Oxford Reference*. (Apr 7, 2022) [www.oxfordreference.com/view/10.1093/oi/authority.20110803100454209](http://www.oxfordreference.com/view/10.1093/oi/authority.20110803100454209).
- [18] S. Wolfram, *A New Kind of Science*, Champaign, IL: Wolfram Media, Inc., 2002.
- [19] H. Everett, III, “‘Relative State’ Formulation of Quantum Mechanics,” *Reviews of Modern Physics*, 29(3), 1957 pp. 454–462. doi:10.1103/RevModPhys.29.454.
- [20] Mathematica, Release Version 12.1, Champaign, IL: Wolfram Research, Inc., 2020.
- [21] Pandas, Version 1.0.1, 2020. [pandas.pydata.org](http://pandas.pydata.org).
- [22] “PGF–Create PostScript and PDF graphics in TeX,” Version 3.1.9a, 2021. [www.ctan.org/pkg/pgf](http://www.ctan.org/pkg/pgf).