

Extending The Thue-Morse Sequence

Olivia Appleton

TMW Center for Early Learning + Public Health
University of Chicago
Chicago, Illinois, United States
ORCID: 0009-0004-2296-7033

Dan Rowe

Department of Math and Computer Science
Northern Michigan University
Marquette, Michigan, United States
Email: darowe@nmu.edu

Abstract—In this paper, we discuss various ways to extend the Thue-Morse Sequence [1] when used as a fair-share sequence. Included are N definitions of the original sequence, M extensions to n players, and proofs of equality for all definitions. In the appendix are several complexity analyses for both space and time of each definition.

Index Terms—TBA

I. INTRODUCTION

Make sure to add that while this paper does not deal with negative or fractional bases, many of the definitions are trivially extendable to that domain, and at least one of them already has been in another paper.

II. THE ORIGINAL SEQUENCE

A. Definition 1 - Parity of Hamming Weight

This definition appears in [1–3]

The Hamming Weight, as typically defined, is the digit sum of a binary number. In other words, it is a count of the high bits in a given number. A common way to generate the Thue-Morse Sequence is to take the parity of the Hamming Weight for each natural number. We can use that as follows:

$$\begin{aligned} p(0) &= 0 \\ p(n) &= n + p\left(\left\lfloor \frac{n}{2} \right\rfloor\right) \pmod{2} \end{aligned} \quad (1)$$

Under this definition, you can construct the Thue-Morse Sequence using the following, starting at 0:

$$T_{2,1}(n) = p(n) \quad (2)$$

The subscript indicates that we are using 2 players (writing in base 2) and that we are using the first definition laid out in this paper. Note that when we extend to n players, the T function will get a second parameter for the number of players, so it will look like $T_{n,d}(x, s)$, where s is the size of the player pool, and therefore the base we use to define the sequence.

B. Definition 2 - Invert and Extend

Appears in [1]

This definition is more natural to think about as extending a tuple that contains the sequence. We will give a recurrence relation below, but to build an intuition we will work in this framework first.

Let $t(n)$ be the first 2^n elements of the Thue-Morse Sequence. Given this, we can define:

$$\text{inv}(\mathbf{x}) = \begin{cases} 0, & \text{if } x_i = 1 \\ 1, & \text{if } x_i = 0 \end{cases} \quad \text{for } \mathbf{x} = (x_0, x_1, \dots, x_{n-1}) \quad (3)$$

$$\begin{aligned} t(0) &= \langle 0 \rangle \\ t(n) &= t(n-1) \parallel \text{inv}(t(n-1)) \end{aligned} \quad (4)$$

Given the above, we can define a recurrence relation that will give us individual elements. It will be less efficient to compute, but will allow proofs of equivalence to be easier.

$$\begin{aligned} T_{2,2}(0) &= 0 \\ T_{2,2}(n) &= T_{2,2}\left(n - 2^{\lfloor \log_2(n) \rfloor}\right) \pm 1 \pmod{2} \end{aligned} \quad (5)$$

C. Definition 3 - Substitute and Flatten

This definition appears in [1, 2, 4]

$$\begin{aligned} s(n) &= \begin{cases} \langle 0, 1 \rangle, & \text{if } n = 0 \\ \langle 1, 0 \rangle, & \text{if } n = 1 \end{cases} \\ t(0) &= \langle 0 \rangle \\ t(n) &= \bigparallel_{i=0}^{2^{n-1}-1} s(t(n-1)_i) \\ T_{2,3}(n) &= t(\lceil \log_2(n+1) \rceil)_n \end{aligned} \quad (6)$$

So for example, calculating $T_{2,3}(3)$ would look like:

$$\begin{aligned} t(0) &= \langle 0 \rangle \\ t(1) &= \bigparallel_{i=0}^0 s(t(0)_i) = \langle 0, 1 \rangle \\ t(2) &= \bigparallel_{i=0}^1 s(t(1)_i) = \langle 0, 1, 1, 0 \rangle \\ T_{2,3}(3) &= t(\lceil \log_2(3+1) \rceil)_3 \\ &= t(2)_3 \\ &= \langle 0, 1, 1, 0 \rangle_3 \\ &= 0 \end{aligned}$$

D. Definition 4 - Recursion

This definition appears in [1, 4]

$$\begin{aligned} T_{2,4}(0) &= 0 \\ T_{2,4}(2n) &= T_{2,4}(n) \\ T_{2,4}(2n+1) &= 1 - T_{2,4}(n) \pmod{2} \end{aligned} \quad (7)$$

E. Definition 5 - Highest Bit Difference

This definition appears in [5]

The text below is from Wiki and needs to be entirely rewritten. I was able to derive the formula on my own from translating their code. This method leads to a fast method for computing the Thue–Morse sequence: start with $t_0 = 0$, and then, for each n , find the highest-order bit in the binary representation of n that is different from the same bit in the representation of $n - 1$. If this bit is at an even index, t_n differs from t_{n-1} , and otherwise it is the same as t_{n-1} .

$$\begin{aligned} T_{2,5}(0) &= 0 \\ T_{2,5}(n) &= \left[\log_2(n \oplus (n-1)) \right] + T_{2,5}(n-1) \pm 1 \pmod{2} \end{aligned} \quad (8)$$

F. Definition 6 - Odious Number Derivation

Definition appears in [1]

Another way to generate the Thue-Morse Sequence is to take the sequence of Odious Numbers [6] mod 2. Odious numbers are those with an odd number of 1s in their binary representation. Note that the player numbers in this derivation are swapped, so when generating this for testing and extension, we add 1 to the result. Some simple generating code for this is as follows:

```
from itertools import count

def seq_p2_d06():
    for i in count():
        if bin(i).count('1') % 2:
            yield (i + 1) % 2
```

Aren't Odious Numbers exactly the numbers where the parity of the hamming weight is 1? So doesn't that mean that the Thue-Morse Sequence selects which numbers are Odious? From cursory testing, it seems to. There's something to be had there.

A possible way to extend this would be to reinterpret this as where the digit sum is not n-even

G. Summary

III. PROVING EQUIVALENCE BETWEEN STANDARD DEFINITIONS

A. Correlating Definition 1 and Definition 4

Note that in Equation 7 where we define $T_{2,4}$, we are working in mod 2, where +1 and -1 are logically equivalent. We can therefore simplify its definition to be:

$$\begin{aligned} T_{2,4}(0) &= 0 \\ T_{2,4}(n) &= n + T_{2,4}\left(\left\lfloor \frac{n}{2} \right\rfloor\right) \pmod{2} \end{aligned} \quad (9)$$

This is identical to Equation 2, where we define $T_{2,1}$.

IV. THE EXTENSIONS

A. Definition 1 - Modular Digit Sums

Definition appears in [7, 8]

To extend definition 1 from 2 to n players, we must first map our concept of parity to base n . We can do this by taking the parity equation defined above and replacing the 2s with n , for $n \in \mathbb{Z}_{\geq 2}$.

$$\begin{aligned} p_n(0) &= 0 \\ p_n(x) &= x + p_n\left(\left\lfloor \frac{x}{n} \right\rfloor\right) \pmod{n} \end{aligned} \quad (10)$$

Under this definition, you can construct the Thue-Morse Sequence using the following, starting at 0:

$$T_{n,1}(x, s) = p_s(x) \quad (11)$$

1) *Proof of Equivalence with Original Definition 1:* It is clear from visual inspection that p_2 is identical to our original definition of p .

$$\begin{aligned} p_2(x) &= p(x) \\ x + p_2\left(\left\lfloor \frac{x}{2} \right\rfloor\right) &= x + p\left(\left\lfloor \frac{x}{2} \right\rfloor\right) \\ x + \left\lfloor \frac{x}{2} \right\rfloor + p_2\left(\left\lfloor \frac{x}{2^2} \right\rfloor\right) &= x + \left\lfloor \frac{x}{2} \right\rfloor + p\left(\left\lfloor \frac{x}{2^2} \right\rfloor\right) \\ x + \left\lfloor \frac{x}{2} \right\rfloor + \left\lfloor \frac{x}{2^2} \right\rfloor + \dots &= x + \left\lfloor \frac{x}{2} \right\rfloor + \left\lfloor \frac{x}{2^2} \right\rfloor + \dots \end{aligned} \quad (12)$$

B. Definition 2 - Increment and Extend

In the original version of this definition, we inverted the elements. In base 2, this is the same thing as adding 1 (mod 2). Given that, let $t(x, n)$ be the first n^x elements of the Extended Thue-Morse Sequence, for $n \in \mathbb{Z}_{\geq 2}$.

$$\text{inc}(\mathbf{x}, n) = \begin{matrix} x_i + 1 \pmod{n} \\ \text{for } \mathbf{x} = (x_0, x_1, \dots, x_{n-1}) \end{matrix} \quad (13)$$

$$\begin{aligned} t(0, n) &= \langle 0 \rangle \\ t(1, n) &= \langle 0, 1, \dots, n-1 \rangle \\ t(x, n) &= t(x-1, n) \cdot \text{inc}(t(x-1, n), n) \end{aligned} \quad (14)$$

Given the above, we can define a recurrence relation that will give us individual elements. It will be less efficient to compute, but will allow proofs of equivalence to be easier.

$$\begin{aligned} T_{n,2}(0, s) &= 0 \\ T_{n,2}(x, s) &= T_{n,2}\left(x - s^{\lfloor \log_s(x) \rfloor}, s\right) + 1 \pmod{s} \end{aligned} \quad (15)$$

1) *Proof of Equivalence with Original Definition 2:*

C. Definition 3 - Substitute and Flatten

There's a bit of a leap here, since we have to explain why the rotation is equivalent to the binary choice presented in the original. There also might be a better syntax to define the rotation, perhaps using the format used in *inv* and *inc*.

$$\begin{aligned}
 b(s) &= \langle 0, 1, \dots, s-2, s-1 \rangle \\
 r(a, x) &= \prod_{i=0}^{|a|-1} \langle a_{i+x \bmod |a|} \rangle \\
 s(x, s) &= r(b(s), x) \\
 t(0) &= \langle 0 \rangle \\
 t(x, s) &= \prod_{i=0}^{2^{x-1}-1} s(t(x-1)_i, s) \\
 T_{n,3}(x, s) &= t(\lceil \log_s(x+1) \rceil, s)_x
 \end{aligned} \tag{16}$$

1) Proof of Equivalence with Original Definition 3:

D. Definition 4 - Recursion

Need to reference eq 9 in definition

1) Proof of Equivalence with Original Definition 4:

E. Definition 5 - Highest Digit Difference

This is speculative, and testing needs to be done

1) Proof of Equivalence with Original Definition 5:

F. Definition 6 - Odious Extension

This is speculative, and testing needs to be done

1) Proof of Equivalence with Original Definition 6:

G. Summary

V. PROVING EQUIVALENCE BETWEEN EXTENDED DEFINITIONS

A. Summary

VI. PROVING PERSISTENCE OF ORIGINAL PROPERTIES

VII. ACKNOWLEDGMENT

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

VIII. APPENDIX

A. Complexity of Original Definition 1

- 1) Time Complexity:
- 2) Space Complexity:

B. Complexity of Original Definition 2

- 1) Time Complexity:
- 2) Space Complexity:

C. Complexity of Original Definition 3

- 1) Time Complexity:
- 2) Space Complexity:

D. Complexity of Original Definition 4

- 1) Time Complexity:
- 2) Space Complexity:

E. Complexity of Original Definition 5

- 1) Time Complexity:
- 2) Space Complexity:

F. Complexity of Original Definition 6

- 1) Time Complexity:
- 2) Space Complexity:

G. Complexity of Extension Definition 1

- 1) Time Complexity:
- 2) Space Complexity:

H. Complexity of Extension Definition 2

- 1) Time Complexity:
- 2) Space Complexity:

I. Complexity of Extension Definition 3

- 1) Time Complexity:
- 2) Space Complexity:

J. Complexity of Extension Definition 4

- 1) Time Complexity:
- 2) Space Complexity:

K. Complexity of Extension Definition 5

- 1) Time Complexity:
- 2) Space Complexity:

L. Complexity of Extension Definition 6

- 1) Time Complexity:
- 2) Space Complexity:

REFERENCES

- [1] OEIS Foundation Inc. The on-line encyclopedia of integer sequences (oeis). Accessed: 2024-10-24, 2024. OEIS Sequence ID: A010060.
- [2] Lukas Spiegelhofer. The level of distribution of the thue–morse sequence. *Compositio Mathematica*, 156(12):2560–2587, 2020.
- [3] Jean-Paul Allouche and Jeffrey Shallit. The ubiquitous prouhet-thue-morse sequence. In C. Ding, T. Hellesteth, and H. Niederreiter, editors, *Sequences and their Applications*, pages 1–16, London, 1999. Springer London.
- [4] M. Kolář, M. K. Ali, and Franco Nori. Generalized thue-morse chains and their physical properties. *Phys. Rev. B*, 43:1034–1047, Jan 1991.
- [5] Jorg Arndt. *Matters computational*. Springer, Berlin, Germany, October 2010.
- [6] OEIS Foundation Inc. The on-line encyclopedia of integer sequences (oeis). Accessed: 2024-10-24, 2024. OEIS Sequence ID: A000069.
- [7] Ricardo Astudillo. On a class of thue-morse type sequences. *Journal of Integer Sequences*, 6, 2003.
- [8] F. M. Dekking. The thue-morse sequence in base 3/2. *Journal of Integer Sequences*, 26, 2023.