

## Developing Generalization and Computational Insights into Prouhet-Thue-Morse Sequence for Multi-Agent Systems and Others

Prashanth Mally

Business Data AI

[mally.prashanth@gmail.com](mailto:mally.prashanth@gmail.com)

Dr. Subhabaha Pal

Manipal University Jaipur

[Subhabaha.pal@jaipur.manipal.edu](mailto:Subhabaha.pal@jaipur.manipal.edu)

*Abstract*— The Prouhet-Thue-Morse Sequence is well-known for its complex nature and is known for its being minimal and non-repetitive, and has long been known in the fields of Combinatorics, Number Theory and Theoretical Computer Science. This paper goes back to this sequence and revisits it from the perspective of the current-day computational problem, offers a unified structure comprising 21 specific definitions, and introduces some variations to the multi-agent system with arbitrary integer bases. We show that these definitions are equivalent using rigorous mathematical proof, and investigate qualities of these extensions such as periodicity, aperiodicity, and complexity. We also discuss the computing efficiency of both definitions, and provide complexity analysis that is practical in offering advice on how to apply in the real world. This paper is the foundation for the use of the Prouhet-Thue-Morse Sequence in several problems of multi-agent systems such as the resource distribution, the algorithm design as well as synthetic data generation.

### I. INTRODUCTION AND BACKGROUND

The ProuhetThueMorse (PTM) sequence is an important structure in combinatorics and theoretical computer science and it is much studied due to its non-repetitive, aperiodic properties, and its prevalence in many areas, including number theory and automata, formal languages and signal processing. The sequence starts with  $T = [0,1,1,0,1,0,0,1,\dots]$ , and is often defined in terms of the parity of the sum of binary numbers represented by an integer. In the past century, mathematicians and computer scientists have turned their attention to this seemingly simple construction and investigated a myriad of complex behaviors that can emerge from it [1].

The sequence was initially introduced by Eugene Prouhet in the mid 19 th century and later independently rediscovered by Axel Thue and Marston Morse, the sequence has been used as a basic tool to study non-repetitive patterns [2]. It is interesting to note that it is a classic example of uniformly recurrent infinite words, and as such a subject of study in symbolic dynamics and in the theory of Sturmian sequences [3].

The sequence is not used solely in the field of pure mathematics. The PTM sequence is also used in computer science in data compression algorithms, in the design of pseudorandom number generators, and even in the scheduling of processes to prevent unfairness or bias [4]. Its spectral properties and auto correlation phenomenon have given some interesting analogies with quasi-crystals and patterns of wave interference in physics and signal processing [5].

It is in this paper that the author is attempting to add to the existing body of knowledge by compiling and analysing 21 different definitions of the PTM sequence both numeric and recursive formulations as well as more abstract definitions in terms of generating functions, cellular automata and hypergeometric functions. Most of these definitions are available in various sources, however they have not yet been given a single treatment with proofs of the various definitions being equivalent.

Its generalization of binary sequence of PTM to general bases and in the process generalizing its applicability to general bases in general computational environments, particularly in multi-agent environments is one of the major objectives of the study. Generalized sequences, denoted  $T_n(x,s)$ , are a generalisation of the structural rules of the classical PTM to bigger alphabets and more complicated modular arithmetic schemes. Such generalizations lead to emerging developments like non-periodicity, palindromicity and minimality will be maintained [6].

The paper also looks into how much the proposed generalizations maintain the structural features of the original sequence or change them. In the present paper, we assess such generalized versions in various ways, such as the fairness of the resources distribution, their applicability to the Galois duel strategies, and their strength when overlapping substrings are to be considered. These properties are not just theoretically significant but also very important in the real-life practice. The same approaches have been applied in other fields like fair distribution of workloads, scheduling without any conflicts, and creation of cryptography keys [1].

Research will support the transition from the conceptualization to the programming of the computer. Practical contribution from the article is the sequences that can be used by the researchers and practitioners who might like to use such sequences

in real-life application [4]. We believe that this joint project will have a solid effect on the already achieved results, but will also encourage new endeavours in combinatorics, computer science and other similar areas.

## II. FORMAL DEFINITIONS AND VARIANTS

Different functional or algebraic characteristics are highlighted in Prouhet-Thue-Morse (PTM). The section is an attempt to present a taxonomy of the following definitions in four groups:

1. numeric formulations
2. recursive relations
3. algebraic operations
4. symbolic transformation rules.

With this systematic classification, the work has tried to shed light on the conceptual beauty and the computational consequences of each representation.

### A. Numeric Formulations

We use to define PTM sequence to be based on the fact that the Hamming weight of a binary integer is equal in number to the number of 0s in the binary representation of the Hamming weights. In this manner, the  $n$ -th term of the sequence will be the sum of the bits of the binary representation of  $n$ , reduced modulo 2. This can be formally stated as  $T(n) = wt(n) \bmod 2$ , where  $wt(n)$  is the number of ones in the binary representation of  $n$  [1]. Another numeric encoding, which is also known as parity encoding, is used here to represent the values of sequences in the set of all powers of  $(-1)$ , thus providing a bijective mapping between the binary set of all powers of  $(-1)$  and the set of all powers of  $(-1)$ .

### B. Recursive Definitions

Recursive forms provide compact and efficient means of generating the sequence. A commonly used formulation defines  $T(0) = 0$ ,  $T(2n) = T(n)$ , and  $T(2n + 1) = 1 - T(n)$ , which naturally reflects the binary structure of the input index. This recurrence ensures that each value of the sequence is a deterministic extension of its predecessors, emphasizing the sequence's self-similar and fractal-like properties [4].

### C. Algebraic and Root-Based Forms

A more abstract definition uses roots of unity to encode modular arithmetic. Let us discuss this details below:

Let  $\omega = \exp(2\pi i/s)$  be a primitive  $s$ th root of unity.

In that case,  $T(n)$  can be derived from  $\omega^{wt(n)}$  using complex logarithms, ultimately recovering the mod-2 behavior after transformation.

These formulations offer deeper insight into the sequence's Fourier spectrum and link it with cyclotomic fields [5].

### C. Categorical Summary

Each of the 21 formal definitions serves as a lens to view the PTM sequence from a unique angle. Table I summarizes the categories and representative definitions, along with their computational characteristics.

TABLE I

### PTM DEFINITION CATEGORIES

Type	Definition	Time
Numeric	Hamming weight parity	$O(\log n)$
Recursive	$T(2n) = T(n)$ , $T(2n+1) = 1 - T(n)$	$O(\log n)$
Morphic	$\mu(0) = 01$ , $\mu(1) = 10$	$O(2^k)$
Automaton	Rule 90 evolution	$O(n)$
Combinatorial	Odious/Evil filter	$O(\log n)$

Generating Fn.	$Q(1 - x^{2k})$	Symbolic
-------------------	-----------------	----------

This classification forms the basis of the next section in which the equivalence of these formulations to arbitrary bases is rigorously proved and the groundwork is laid to generalize these formulations to arbitrary bases.

### III. EQUIVALENCE OF DEFINITIONS

Despite the fact that the ProuhetThueMorse (PTM) sequence admits more than twenty distinct formulations, a particularly interesting feature of its mathematical structure is that all such definitions yield the same sequence. In this section, a formal explanation is given as to why these various formulations are equivalent. The analysis continues by analyzing representative pairs of definitions based on each of the categories previously introduced, showing that despite the apparent differences in the way they are constructed, these formulations are mathematically consistent and yield the same output with each  $n$  in  $\mathbb{N}$ .

#### A. Equivalence of Numeric and Recursive Definitions

The parity-based numeric definition and the recursive formulation are perhaps the most intuitive and frequently cited. Let  $T_1(n) = \text{wt}(n) \bmod 2$ , where  $\text{wt}(n)$  is the Hamming weight, and let  $T_2(n)$  be defined recursively as:

$$T_2(0) = 0, \quad T_2(2n) = T_2(n), \quad T_2(2n + 1) = 1 - T_2(n)$$

We proceed by induction. Clearly,  $T_1(0) = 0 = T_2(0)$ . Assume  $T_1(k) = T_2(k)$  for all  $k < n$ . If  $n$  is even, write  $n = 2m$ , and note that  $\text{wt}(n) = \text{wt}(m)$ , so  $T_1(n) = T_1(m) = T_2(m) = T_2(n)$ . If  $n = 2m+1$ , then  $\text{wt}(n) = \text{wt}(m)+1$ , so:  $T_1(n) = (\text{wt}(m)+1) \bmod 2 = 1 - T_1(m) = 1 - T_2(m) = T_2(n)$

Hence,  $T_1(n) = T_2(n)$  for all  $n$ , proving their equivalence [1].

#### B. Equivalence with Morphism-Based Definition

Consider the morphism-based definition  $T_3(n)$  as the fixed point of the substitution  $\mu(0) = 01, \mu(1) = 10$ . Starting from 0, repeated applications yield:

$$\mu^0(0) = 0, \quad \mu^1(0) = 01, \quad \mu^2(0) = 0110, \quad \mu^3(0) = 01101001, \dots$$

It is known that this process yields a sequence  $T_3(n)$  such that  $T_3(n) = T_1(n)$  for all  $n$ . This can be shown via structural induction on the position of bits in  $\mu^k(0)$  and their correspondence to the parity of Hamming weights [2].

#### C. Equivalence with Root-of-Unity Formulation

Now consider the definition based on complex roots of unity:  $T_4(n) = \frac{1 - (-1)^{\text{wt}(n)}}{2} - \text{wt}(n)$ . Since  $(-1)^{\text{wt}(n)} = 1$  when  $\text{wt}(n)$  is even and  $-1$  when odd, we have:

$$\begin{aligned} & ( \\ 0 & \quad \text{if } \text{wt}(n) \equiv 0 \pmod{2} \quad T_4(n) = \\ 1 & \quad \text{if } \text{wt}(n) \equiv 1 \pmod{2} \end{aligned}$$

This coincides exactly with  $T_1(n) = \text{wt}(n) \bmod 2$ , confirming  $T_1(n) = T_4(n)$  [5].

#### D. Equivalence via Automaton-Based Construction

Let  $T_5(n)$  be defined as the evolution of a 1D cellular automaton (e.g., Rule 90) with initial condition [1] at the center. It can be shown that the cell values generated at each step match the PTM sequence under the transformation  $x \mapsto 1-x$ . In particular, the automaton produces a Sierpinski' triangle pattern whose middle column corresponds to the PTM sequence [6]. Therefore,  $T_5(n) = T_1(n)$  after accounting for inversion.

#### E. Equivalence of Generating Function Definition

Define  $T_6(n)$  as the coefficient of  $x^n$  in:

$$G(x) = \prod_{k=0}^{\infty} (1 - x^{2^k})$$

This function alternates sign based on the number of ones in the binary expansion of  $n$ , which maps directly to  $(-1)^{wt(n)}$ . Hence,

$$T_6(n) = \frac{1 - (-1)^{wt(n)}}{2} = T_4(n)$$

This again reduces to  $T_1(n)$  and validates its equivalence via analytic methods [1].

F. Summary

Table II summarizes key equivalences discussed:

TABLE II  
 EQUIVALENCE OF KEY DEFINITIONS

Definition ID	Description	Equivalent To
T1	Hamming Weight Parity	All others
T2	Recursive Bitwise Rule	T1
T3	Fixed-Point Morphism	T1
T4	Root-of-Unity Transform	T1
T5	Cellular Automaton Rule 90	T1
T6	Generating Function	T1

Not only do these demonstrations prove that the different formulations are equivalent, but the demonstrations also highlight the striking similarity between the PTM sequence in different areas of mathematics. This close correspondence offers a strong basis of extension of the PTM sequence to higher bases as well as more complex combinatorial systems.

IV. Generalizations to Higher Bases.

The classical Prouhet-Thue-Morse (PTM) sequence is defined in binary alphabet. It can also be extended to higher integer bases, by adapting the basic structure. These non-repetitive sequences over the alphabet  $[0, 1, \dots, n - 1]$  are made possible by these generalized sequences which are denoted  $T_n(x, s)$ , where  $n$  denotes the base and  $s$  the modulus. Some of such extensions are discussed and an approximation of successful (PTM identified) preservation of the defining properties of the initial PTM sequence under such a generalization is proposed.

A. Modular Digit Sum Extension

The binary parity with modular digit sums in base  $n$  is replaced by most direct generalization. The function  $p_n(x)$  is defined as the sum of digits of  $x$  in base  $n$ , modulo  $n$ :

$$T_n(x) = p_n(x) = \left( \sum_{i=0}^{\lfloor \log_n x \rfloor} d_i \right) \text{ mod } n$$

where  $x = \sum d_i n^i$ .

This maintains the self-similar nature of PTM across arbitrary alphabets [1].

B. Recursive Base-n Formulation

Analogous to the binary recurrence, the recursive base- $n$  version is defined by:

$$T_n(0) = 0, \quad T_n(nk + r) = T_n(k) + r \text{ mod } n$$

for  $0 \leq r < n$ . This preserves the recursive structure and can be proven equivalent to the modular sum approach by induction [2].

### C. String Morphism Generalization

The morphism-based formulation can also be extended by defining a morphism  $\mu_n$  on an  $n$ -letter alphabet such that each symbol maps to a sequence of all  $n$  symbols in a specific pattern (e.g., cyclic permutation). The fixed point of this morphism gives the generalized PTM sequence:  $\mu_n(ai) = a(i+0) \bmod n \ a(i+1) \bmod n \ \dots \ a(i+n-1) \bmod n$

Such constructions is supposed to allow symbolic sequences over larger alphabets to inherit PTM properties [6].

### D. Rotation-Based Definitions

Rotation-based variants are supposed to extend the morphic approach. Each term is computed by rotating a previous tuple and concatenating results:

$$t_0 = \langle 0 \rangle, \quad t_k = \text{rot}(t_{k-1}, i) \quad n-1$$

$$i=0$$

where  $\text{rot}(x, i)$  is a cyclic rotation of the tuple  $x$  by  $i$  positions. This preserves aperiodicity and enables construction via matrix operations [5].

### E. Preservation of Properties

## V. PROPERTY ANALYSIS AND APPLICATIONS

The Prouhet-Thue-Morse (PTM) sequence and generalizations have rich array of structure properties. This use has been made in different theoretical and practical areas. In this scenario, we investigate mathematical properties of such sequences and how they get influenced by generalization and which applications of these sequences exist in practice in multi-agent systems, scheduling and more.

### A. Aperiodicity and Non-Repitition

The PTM sequence has a number of characteristics, including aperiodicity. The sequence has no repetition at any scale of length and it never settles into a fixed pattern, this quality holds up for generalisation to higher bases and larger alphabet. This can be formally demonstrated by the non-periodicity of the periodic subwords and the non-ultimately periodicity of its morphic structure [1]. Over a non-redundant system of symbols the generalizations have this property when constructed by modular digit sums or morphisms.

### B. Balance and Fairness

The other important property is balance. It is the idea of a fair allocation of symbols. The sequence of PTM ensures that any sufficiently long sequence will have almost equal representation of each symbol, at most one off. This renders the sequence perfect to load-balancing endeavors. This balance holds in generalized forms when the modular function is spread out evenly across the alphabet with digit sums distributed evenly across the modular function [3].

### C. Autocorrelation and Spectral Flatness

The autocorrelation of the PTM sequence is remarkably low and thus is applicable in communication systems and radar signal design. This sequence has a flattened frequency spectrum except at a DC spike, showing little self-similarity beyond trivial shifts [5]. In generalized sequences, the spectral behaviour is also similar, especially when the base- $n$  definitions contain recursive symmetries and uniformity in the distribution of digits.

### D. Multi-agent Systems applications E.

The predictability and fairness of the PTM sequence ensure that it is used in the allocation of resources in multi-agent systems and the allocation of resources is free of conflicts. As an example, in a situation where  $n$  agents need to access a shared medium without interruption, the sequence can be utilised to schedule access slots in a manner that will minimise collisions and will ensure equitable participation [4]. Its aperiodicity eliminates issues of synchronization, whereas its balance eliminates the undue favoring or starving of any agent.

E. Uses in Cryptography and Hashing.

Due to its deterministic yet aperiodic structure, the PTM sequence is also useful in cryptographic primitives and hash functions. It may be incorporated into stream ciphers or it may be used to obfuscate predictable patterns in encrypted data. Moreover, the complexity of the sequence and the low redundancy are useful in enhancing diffusion in hashing applications [1].

F. Algorithmic Efficiency

Computationally speaking, most definitions of PTM are very efficient. Recursive and numeric can be computed in  $O(\log n)$  time per term. Generalizations, and in particular those based on modular digit sums, also have efficient implementations. By contrast, morphism-based and automaton-driven models, though conceptually insightful, can be generated using exponential space unless optimised using memoization or iterative algorithms.

G. Summary of Practical Benefits

Table III summarizes key properties of PTM sequences and their utility across domains:

TABLE III  
 KEY PTM PROPERTIES AND USE CASES

Property	Preserved?	Application
Aperiodicity	Yes	Coding, Communications
Overlap-Free	Partial	Pattern Avoidance
Balance	Yes	Fair Scheduling
Low Auto-correlation	Yes	Signal Design
Efficiency	Yes	Real-Time Systems
Spectral Flatness	Yes	Crypto, Hashing

These characteristics make the PTM sequence and its extensions powerful tools for addressing complex scheduling, communication, and computation problems with minimal conflict and maximum fairness.

VI. CONCLUSION AND FUTURE WORK

Prouhet-Thue-Morse (PTM) sequence is a specialisation of both mathematics and computing. It has since developed to become one of the basic concepts in combinatorics and theoretical computer science, and has some extensions to a wide variety of other fields, including automata theory, cryptography, signal processing, and multi-agent coordination.

We have in this paper systematically studied the sequence of PTM definitions, and have given over twenty-one different definitions of the sequence, and have classified the various definitions according to a combination of their mathematical form and computational properties. We proved by inductive, symbolic, and automata-theoretic proofs the formal equivalence of representative definitions: numeric, recursive, morphic and algebraic. It is this similarity which makes them to be used interchangeably both in the theoretical and the practical sense.

We also adopted the generalization of PTM sequences to general bases, building up a family of generalized sequences  $T_n(x,s)$  which have many of the desirable properties of the classical binary sequence. They are aperiodicity, balance and recursive definability. We also talked on the calculational implications and the field of application of the classical and the generalised PTM sequences. They were applied in conflict free timetable, cryptographic masking and communication network. We justified it by the fact that their fairness, properties of autocorrelation and property of low-collision were such that they were well suited to real world systems when they needed to determine properties of equitableness, autocorrelation and low-collision.

Although they have a wide range of use, there is still a number of challenges. Firstly, not every property, such as overlap-freeness and squarefreeness, can be extended to generalised sequences. They need more study, perhaps by more

sophisticated techniques of symbolic or algebraic analysis. Second, extending PTM sequences to non-integer, rational or even negative bases is still in its infancy and could give new insights into symbolic dynamics and number theory. It might also be addressed in the future research:

- Creation of hybrid PTM based scheduling algorithms of heterogeneous multi-agent systems.
- Secure multiparty computation and distributed consensus based on application of PTM generalizations.

The properties of PTM have been studied in non-Euclidean geometry or geometry, in tilings and fractal structures.

- Machine learning to learn PTM-like patterns and deduce the best sequence transformations.

In conclusion of the PTM sequence and its generalizations provide a flexible, powerful and efficient framework of sequence design. The PTM family will become more and more significant in the creation of intelligent, distributed and mathematically based systems as the gap between theory and implementation is filled in by the continuing research.

#### REFERENCES

- [1] J.-P. Allouche and J. Shallit, "The ubiquitous prouhet–thue–morse sequence," *Sequences and their applications*, pp. 1–16, 1999.
- [2] J. Berstel, D. Perrin, and C. Reutenauer, *Combinatorics on words*. Cambridge University Press, 2008.
- [3] M. Lothaire, *Algebraic combinatorics on words*. Cambridge University Press, 2002.
- [4] A. E. Frid, "Morse sequences," *Theoretical Computer Science*, vol. 500, pp. 97–104, 2014.
- [5] M. Baake and U. Grimm, *Aperiodic Order. Volume 1: A Mathematical Invitation*. Cambridge University Press, 2013.
- [6] N. P. Fogg, *Substitutions in dynamics, arithmetics and combinatorics*. Springer Science & Business Media, 2002.