TRANSFORMATICS 101 - explained

Joseph Willrich Lutalo[*]
Nuchwezi Research
joewillrich@gmail.com, jwl@nuchwezi.com

October 6, 2025

Abstract

In only **10 general foundational results**[1][2] — all **based on earlier mathematical research** by the author, as well as **10 extra abstractions** spanning **how transformatics might be leveraged** as **a basis mathematics** `for` `problems` in; `number theory` and `cryptography` (o-SSIs); `mathematical statistics` (ADM, TCR, PCR, MSS, SC, PC); `linear algebra` (super sequences); `computer science` (TEA language, transformer-chains/sequence-transformers) and `information theory` (entropy of a sequence) among others, we distill and establish, in a summarized form, the **core foundations of Transformatics**[3] as its own mathematical theory, branch, or **discipline**. In this **mini-thesis**. Each foundational result presented is likewise accompanied by one or **more references** to some earlier work where the idea or presented concept was first developed, presented, or applied by the author during their research. **Useful notes and commentary** are also included where necessary, and in the extension decade of **proposed abstractions**, also commentary about how the proposals relate to the decade of foundational results likewise catered for.

Keywords: Foundations, Transformatics, Sequence Analysis, Sequence Transformers, Thesis

1 Foundations of a New Mathematics of Sequences

1.1 Result 1: The Empty Sequence[4]

$$\Theta = \langle \rangle \implies |\Theta| = 0 \quad \wedge \quad \Theta \equiv \emptyset$$

1.2 Result 2: A Symbol Set[5]

$$\psi_\beta^n = \langle \beta_{i \in [1,n]} \rangle : n \in \mathbb{N} : \mathbb{L}(\beta_i \in \psi_\beta) = 1 \quad \forall \beta_i \in \psi_\beta$$

1.3 Result 3: A Sequence[4]

$$\Theta^n = \langle \theta_{i \in [1,n]} \rangle : \mathbb{N} \times \psi_\beta : n \geq 1$$

---

[*]Principal Investigator at Nuchwezi Research, GARUGA, Uganda.
ORCID: `https://orcid.org/0000-0002-0002-4657`

1.4 Result 4: Sequence Cardinality[6][5]

$$\forall \Theta^n = \langle \theta_{i\in[1,n]} \rangle : \mathbb{N} \times \psi_\beta \quad \implies \quad \unrhd(\Theta^n) = |\Theta^n| = n \in \mathbb{N}$$

1.5 Result 5: A Sequence Symbol Set[7][5]

$$\psi(\Theta^n) = \langle \theta_{i\in[1,k]} \rangle : k, n \in \mathbb{N} \wedge k \geq 1 : \unrhd(\theta_i \in \psi(\Theta^n)) = 1 \quad \forall \theta_i \in \Theta^n : k \leq n : \unrhd(\psi(\Theta^n)) = k$$

1.6 Result 6: A Sequence Transformation[3][4]

$$\Theta^n = \langle \theta_{i\in[1,n]} \rangle : \mathbb{N} \times \psi_\beta \to \Theta^* = \langle \theta^*_{i\in[1,k]} \rangle : \mathbb{N} \times \psi_* : \psi_\beta \neq \psi_* \vee \psi(\Theta^n) \neq \psi(\Theta^*)$$

1.7 Result 7: A Sequence Transformer[8][3]

$$\Theta^n = \langle \theta_{i\in[1,n]} \rangle : \mathbb{N} \times \psi_\beta \xrightarrow{f(\Theta)=\Theta^*} \Theta^* = \langle \theta^*_{i\in[1,k]} \rangle : \mathbb{N} \times \psi_*;$$
$$\psi_\beta \neq \psi_* \vee \psi(\Theta^n) \neq \psi(\Theta^*)$$

1.8 Result 8: A Sequence Filter[3]

$$\Theta^n = \langle \theta_{i\in[1,n]} \rangle : \mathbb{N} \times \psi_\beta \xrightarrow{f(\Theta,\psi_{\beta*})=\Theta^*} \Theta^* = \langle \theta^*_{i\in[1,k]} \rangle : \mathbb{N} \times \psi_{\beta*};$$
$$k \in \mathbb{N} : \psi_{\beta*} \subseteq \psi_\beta \implies \unrhd(\psi_{\beta*}) = k \leq \unrhd(\psi_\beta) \implies \unrhd(\Theta^*) \leq \unrhd(\Theta^n)$$

1.9 Result 9: A Sequence Generator[3][9]

$$\Theta^n | \emptyset \xrightarrow{f(\psi_\beta,k))=\Theta^*} \Theta^k = \langle \beta^*_{i\in[1,k]} \rangle : \mathbb{N} \times \psi_\beta : k \in \mathbb{N} : \forall \beta_i \in \Theta^k \implies \beta_i \in \psi_\beta \quad \wedge \quad k \geq 1$$

1.10 Result 10: A Sequence Sampler[10]

$$\Theta^n \xrightarrow{f(\Theta,k))=\Theta^k} \Theta^k = \langle \theta_{i\in[1,k]} \rangle : \mathbb{N} \times \psi(\Theta^n) : k, n \in \mathbb{N} \wedge k \geq 1$$

2 10 Proposed Abstractions Leveraging Transformatics

2.1 Proposal 1: The o-SSI Sequence from any Sequence[7]

Assume we consider the sequence $\Theta : \mathbb{N} \times \psi_*$ of any length and spanning an arbitrary symbol set ψ_*. Then, in case we wish to reduce Θ to a representative sequence Θ_{ossi} that only contains elements of Θ spanning the symbol set, ψ_β, of some particular base, β; in which case, Θ_{ossi} would be an **orthogonal symbol set identity** (o-SSI) sequence under base-β, then, the following o-SSI-generator — also **an extractor**[1] and **a reducer** transformer, defined using the calculus and semantics of transformatics, would suffice[2]:

Transformer 1 (The **o-SSI Extractor**). $\Theta \langle \theta_{i\in[1,n]} \rangle : \mathbb{N} \times \psi_* \xrightarrow{O_{filter}(\Theta,\psi_\beta)} \Theta^k \langle \theta_i \rangle : \mathbb{N} \times \psi_\beta = \Theta_{ossi};$
$0 \leq k \leq \unrhd(\psi_\beta) \leq \unrhd(\Theta) = n \in \mathbb{N} : \forall \theta_i \in \Theta^k \quad \exists \theta_i \in \psi_\beta$
$\wedge \quad \forall i,j \in \mathbb{N} \wedge \theta_i, \theta_j \in \Theta^k : I(\theta_i,\Theta) \leq I(\theta_j,\Theta) \implies I(\theta_i,\Theta^k) \leq I(\theta_j,\Theta^k)$

Note that, the **o-SSI Sequence**, Θ_{ossi}, of any sequence Θ under β would likewise be related to the concept of the **Specific Symbol Set**[3], $\hat{\psi}(\Theta)$, of that sequence, in the sense that, at most, Θ_{ossi} is equivalent to $\hat{\psi}_\beta(\Theta)$ or that their cardinalities and memberships are related as:

$$\unrhd(\Theta_{ossi}) \leq \unrhd(\hat{\psi}_\beta(\Theta)) \leq \unrhd(\psi_\beta) \leq \unrhd(\Theta) \tag{1}$$

---

[1]In the sense that, where a sequence already contains the essential symbols spanning some particular symbol set, and even though they might be scattered about or 'buried' around in the sequence surrounded by 'noise', the function would merely extract those essential symbols [only,] from that sequence in their natural order of first occurrence.

[2]Say, and as almost all transformatics might be in practice — readily translatable into a working or equivalent computer program based on or defined using the chaining of sequence transformers such as in a language as TEA[11].

[3]Refer to **Definition 3** in [7]

and as for membership:

$$\psi(\Theta_{ossi}) \subseteq \hat{\psi}_\beta(\Theta) \subseteq \psi_\beta \subseteq \Theta \tag{2}$$

2.2 Proposal 2: The Anagram Distance Between Any Two Sequences[6]

Occasionally, one shall find that they need to quantify how different some two or more sequences are. In the simplest case, we might merely sort a collection of sequence by the relative size of each sequence, and thus by their sequence cardinality for example — $\underline{\mathfrak{u}}(\Theta)$ Vs $\underline{\mathfrak{u}}(\Theta^*)$ for some two sequences involved in a general transformation such as:

Transformation 1. $\Theta \to \Theta^* \quad \Longrightarrow \quad \nabla_{\Theta,\Theta^*} = (\underline{\mathfrak{u}}(\Theta) - \underline{\mathfrak{u}}(\Theta^*));$

$$\nabla_{\Theta,\Theta^*} = \begin{cases} < 0 & compression \\ 0 & conservation \\ > 0 & contraction \end{cases}$$

However **Transformation 1** only caters for comparisons or analysis relative to size of the sequences being compared. And such that ∇_{Θ,Θ^*} shall work just fine where the purpose of the sequence analysis is just concerned with the relative sizes of the sequences. But, and as we established in the **Anagram Distance Theory**(ADT) paper[6], there are many important cases when it is necessary to compare two or more sequences relative to the relative ordering of members in them and not just by their member composition or frequencies. Thus the **Anagram Distance Measure**(ADM), and which, for any two sequences as presented above, the essential measure would be:

$$\tilde{A}(\Theta \to \Theta^*) = \frac{1}{\underline{\mathfrak{u}}(\Theta)} \times \sum_{\forall \theta_i \, \in \Theta}^{\underline{\mathfrak{u}}(\Theta)} |I(\theta_i, \Theta) - I(\theta_i, \Theta^*)| \tag{3}$$

Which, as we saw in [6], has the interesting interpretation as:

$$\tilde{A}(\Theta \to \Theta^*) = \begin{cases} 0, & \Theta = \Theta^*, \\ \leq 1, & \Theta \approx \Theta^*, \\ > 1, & \Theta \ll \Theta^*. \end{cases} \tag{4}$$

And thus can help us observe/compute, quantify and appreciate, the **relative disorder** resulting from one sequence being transformed into another (of potentially similar cardinality and member composition/distribution). A very important formalism in mathematical statistics dealing with measures of variability where traditional measures such as the range, mean deviation, variance and standard deviation statistics can't catch any actual differences between two or more sequences as we quantifiably/provably demonstrated in [3], but also, for simpler difference measures such as just counts/cardinality as we have seen in **Transformation 1**, and earlier on in [10].

2.3 Proposal 3: The Transformer Compression Ratio of any Sequence Transformation[3]

blah blah...

2.4 Proposal 4: The Piecemeal Compression Ratio of Any Sequence Transformation[3]

blah blah...

2.5 Proposal 5: The Modal Sequence Statistic of Any Sequence[3]

blah blah...

2.6 Proposal 6: The Characteristic of Any Sequence[10]

blah blah...

3

2.7 Proposal 7: The Population Characteristic of Any Collection of Sequences[10]

blah blah...

2.8 Proposal 8: A Super Sequence for One or More Sequences[10]

blah blah...

2.9 Proposal 9: Programs as Chains of Sequence Transformers (Transformer-Chains)[10][11]

blah blah...

2.10 Proposal 10: Entropy of a Sequence[4]

blah blah...

References

[1] Joseph Willrich Lutalo. Transformatics 101. *Thesis*, 10 2025. A first distillation of the essential foundations of Transformatics as its own field of mathematics. Accessible via: `https://doi.org/10.6084/m9.figshare.30285106.v1`.

[2] Joseph Willrich Lutalo. **Transformatics 101 (1-Page Executive Summary)**, Oct 2025. *The 1-pager PDF is available at:* `https://doi.org/10.6084/m9.figshare.30284035`.

[3] Joseph Willrich Lutalo. **The Theory of Sequence Transformers & their Statistics**: The 3 information sequence transformer families (anagrammatizers, protractors, compressors) and 4 new and relevant statistical measures applicable to them: Anagram distance, modal sequence statistic, transformation compression ratio and piecemeal compression ratio. *Academia*, 2025. `https://doi.org/10.6084/m9.figshare.29505824.v3`.

[4] Joseph Willrich Lutalo. **Philosophical and Mathematical Foundations of A Number Generating System: The Lu-Number System**. *Academia.edu*, 2025. *Accessible via* `https://doi.org/10.6084/m9.figshare.29262749`.

[5] Joseph Willrich Lutalo. **A general theory of number cardinality**. *Academia.edu*, Jan 2024. *Accessible via* `https://www.academia.edu/43197243/A_General_Theory_of_Number_Cardinality`.

[6] Joseph Willrich Lutalo. **Introducing The Anagram Distance Statistic, $\tilde{A}$, A Quantifier of Lexical Proximity For Base-10 o-SSI And Any Arbitrary Length Ordered Sequences, Its Relevance And Proposed Applications in Computer Science, Engineering And Mathematical Statistics**. *Academia.edu*, 2025. *Accessible via* `https://doi.org/10.6084/m9.figshare.29402363`.

[7] Joseph Willrich Lutalo. **Concerning A Special Summation That Preserves The Base-10 Orthogonal Symbol Set Identity In Both Addends And The Sum**. *Academia*, 2025. *Accessible via* `https://www.academia.edu/download/122499576/The_Symbol_Set_Identity_paper_Joseph_Willrich_Lutalo_25APR2025.pdf`.

[8] Joseph Willrich Lutalo. Transformatics for PSYCHOLOGY. *FigShare*, 9 2025. *Presented at the 13th International East African Psychology Conference, a copy is accessible via:* `https://doi.org/10.6084/m9.figshare.30231748`.

[9] Joseph Willrich Lutalo. **Applying TRANSFORMATICS: Sequence Generators**. `https://doi.org/10.6084/M9.FIGSHARE.29654645`, 2025. FigShare.

[10] Joseph Willrich Lutalo. **Applying TRANSFORMATICS in GENETICS**. *I*POW*, pages 1–206, Aug 2025. A mini-treatise on the mathematics, philosophy and practice of Transformatics for Geneticists. `https://doi.org/10.6084/m9.figshare.29941127`.

[11] Joseph Willrich Lutalo. **TEA GitHub Project — The Reference Implementation of TEA (Transforming Executable Alphabet) computer programming language**. `https://github.com/mcnemesis/cli_tttt/`, 2024. *Single Source of Truth concerning the reference standards of TEA on the web and native operating systems.*