

Contents

Nonlinear Tree-Based Numeration Systems: Consolidated Synthesis	1
0. Abstract	1
1. The Core Theorem: Positional Notation as an Ultrametric Tree	1
2. Five Research Threads — Deep Dive	2
3. Extensions & Future Directions	3
4. Production Infrastructure: Ultrametric Engine	4
5. The Gap Analysis	6
6. Research Roadmap	6
7. Key References	6

Nonlinear Tree-Based Numeration Systems: Consolidated Synthesis

Author: QNFO Research Collective | **Date:** 2026-06-29 **License:** QNFO Unified License Agreement (QNFO-ULA): <https://legal.qnfo.org/>

0. Abstract

Positional notation — the ordinary decimal and binary systems that underlie all computation — is natively an ultrametric tree. The digit string is a path from root to leaf; the base is a branching factor; the distance between two numbers is the depth of their lowest common ancestor. This insight, formalized through the b -adic valuation $d(x, y) = b^{-v_b(x-y)}$, unifies five research threads — distance-dependent weighting (Rofa 2013), hereditary binary compression (Tarau 2013), rational base representation trees (Akiyama et al. 2018), automata-theoretic formalism (Charlier & Kreczman 2025, Lecomte & Rigo 2001), and historical/biological parallels (Chrisomalis 2010) — and opens extensions into tropical geometry, perfectoid spaces, quantum computing, and interface design. This document synthesizes the full investigation.

1. The Core Theorem: Positional Notation as an Ultrametric Tree

Theorem (Native Ultrametric of Positional Notation). [established] In any integer base $b \geq 2$, the distance between two integers measured by the highest power of b dividing their difference is exactly the b -adic valuation:

$$d(x, y) = b^{-v_b(x-y)}$$

where $v_b(n)$ is the exponent of the highest power of b dividing n . This metric satisfies the strong triangle inequality:

$$d(x, z) \leq \max\{d(x, y), d(y, z)\}$$

Proof (sketch). Two numbers share k trailing digits in base b iff $b^k \mid (x - y)$ but $b^{k+1} \nmid (x - y)$, giving $d(x, y) = b^{-k}$. For any three integers, if x shares k_1 trailing digits with y and y shares k_2 with z , then x and z share at least $\min(k_1, k_2)$ trailing digits. Hence $d(x, z) \leq b^{-\min(k_1, k_2)} = \max(d(x, y), d(y, z))$. $\square$

The Tree Structure: - Level 0 (rightmost digit) branches b ways per digit $a_0 \in \{0, \dots, b-1\}$ - Each integer corresponds to a unique root-to-leaf path - Ultrametric distance = depth of lowest common ancestor = count of trailing matching digits

The Inversion: The Archimedean “number line” is a projection. The native geometry is the discrete ultrametric tree — a Cantor set, clopen balls, all triangles isosceles.

Ostrowski’s Theorem [established] states that $\mathbb{Q}$ has exactly two types of completions: the Archimedean $\mathbb{R}$ and the p -adic $\mathbb{Q}_p$ for each prime p . Both are valid. Only one is canonized.

2. Five Research Threads — Deep Dive

Thread 1: Distance-Dependent Weighting (Rofa, 2013)

Core innovation: A digit’s value depends on its vertex-to-root distance in the tree, not on its string position. This activates the digit “0” — it is no longer a passive placeholder. “012” and “12” yield distinct numerical values.

Generalized valuation:

$$val(\text{path}) = f(\text{depth of node, digit value})$$

where f need not factor as $digit \times base^{\text{position}}$.

Implications: Redundancy (multiple representations for the same integer), applications in error detection/correction, and a new class of computational arithmetic.

Thread 2: Hereditary Binary Arithmetic (Tarau, 2013)

Problem: Hyper-large numbers (Knuth up-arrow scale) cannot even have their *length* written in flat notation.

Solution: Recursively compress binary strings via run-length encoding, then encode the run lengths in the same compressed format. This yields a **tree of trees** where each node compresses the binary expansion of its subtree.

Data type:

data Tree = Zero | One | Cons Tree Tree

Result: Numbers requiring 10^{100} decimal digits compress to a tree of depth ~ 20 . Addition, multiplication, and exponentiation remain computable *directly on the compressed hereditary form* — complexity shifts from $O(\log N)$ to $O(\text{depth of hereditary tree})$, which can be exponentially smaller for hyper-large numbers.

Thread 3: Rational Base Representation Trees (Akiyama et al., 2008/2018)

Surprising fact: In integer bases, not every rational has a finite expansion ($1/3 = 0.\bar{3}$). In rational base p/q (coprime, $p > q > 1$), **every integer has a unique finite expansion** — the opposite behavior.

Conversion algorithm:

$$N = \frac{p}{q} \cdot Q + r, \quad 0 \leq r < p/q$$

which is equivalent to $Nq = pQ + rq$. Iterating produces the finite expansion.

The fractal tree: Unlike integer bases where every subtree is isomorphic, rational base trees vary by modular residue class. Some subtrees appear infinitely often; others finitely often. The language of representations is typically **not regular** — a fundamentally more complex object.

Connection: Rational bases are a special case of **beta-expansions** (Parry, 1960), linking to symbolic dynamics and subshifts of finite type.

Thread 4: Automata-Theoretic Formalism (Lecomte & Rigo, 2001; Charlier & Kreczman, 2025)

Abstract numeration system: A regular language L over Σ , ordered lexicographically, where the n -th word represents integer n .

Key result (Charlier & Kreczman, 2025): The set of valid representations $\mathcal{R}(U)$ for a weight sequence U_n is **regular** iff U_n satisfies a linear recurrence with integer coefficients — equivalently, the base is a **Pisot number** (algebraic integer > 1 with all conjugates < 1 in modulus).

Why regularity matters: - Normalization computable by a finite transducer - Addition recognizable by a finite automaton (digit-by-digit, 3-track) - Properties decidable in linear time

Buchi’s theorem [established]: The first-order theory of $\langle \mathbb{N}, + \rangle$ (Presburger arithmetic) is decidable *because* addition is automaton-recognizable. The open question: which nonlinear weight sequences preserve recognizability?

Dumont-Thomas connection: Substitutions $\sigma : \Sigma^* \rightarrow \Sigma^*$ generate infinite words; prefix trees form numeration systems in non-standard bases (the dominant eigenvalue of the substitution’s incidence matrix).

Thread 5: Historical Context & Biological Parallels

Chrisomalis (2010) catalogued 100+ numerical notation systems across 5,500 years. The transition to positional notation was convergent evolution — multiple civilizations independently built tree structures.

System	Native Tree Feature	Suppressed By
Babylonian sexagesimal	Floating radix point = undetermined tree depth	Decimal point
Mayan calendar numerals	Explicit vertical stacking = visible tree	Linear timeline
Modern binary/decimal	bit/digit strings are tree paths	“Number line” projection

Three historical erasures: 1. **ZFC set theory:** Numbers constructed from $\emptyset$ and $\in$ — no tree structure encoded 2. **Dedekind cuts:** $\mathbb{R}$ as connected continuum canonized; $\mathbb{Q}_p$ marginalized as “exotic” 3. **Digital computation:** Hardware enforces tree structure (bit-trees) but software presents linear abstraction

Laws of Form (Spencer-Brown, 1969): An alternative foundation where “drawing a distinction” is primitive. A positional numeral is nested marks — a tree. The base is grouping b distinctions into a higher-order distinction. “10” is a **re-entry** of the form into its own boundary.

DNA parallel: The genetic code uses a 4-letter alphabet in codons (triplets) — structurally a 4-ary tree of depth 3 with 64 leaves mapping to ~ 20 amino acids. The code is degenerate and non-power positional — structurally analogous to a redundancy-allowing nonlinear tree-based numeration system.

3. Extensions & Future Directions

3.1 Theoretical Extensions

Extension	Core Question	Status
Weight sequence classification	For which U_n is arithmetic polynomial-time? Regular?	Partial: Pisot = regular; Salem = open
Tropical geometry	Max-plus algebra is natively ultrametric; what are tropical numeration systems?	Unexplored

Extension	Core Question	Status
Perfectoid spaces (Scholze, 2012)	Can $\mathcal{O}_K$ where K is perfectoid serve as a numeration ground ring?	Speculative
Bruhat-Tits buildings	Generalized trees for $\mathrm{SL}_n(\mathbb{Q}_p)$ — building-based numeration for algebraic groups?	Open
Non-commutative trees	Free group Cayley graphs as non-commutative numeration	Unexplored

3.2 Computational Frontiers

Silent-Radix Cryptography: The fact that a digit string cannot internally specify its base is a cryptographic primitive. A secret large base b makes numbers ambiguous without it — reducing to an integer knapsack problem.

Ultrametric Quantum Computing [speculative]: - Thesis I: Tree topology confines errors geometrically — $p_L \propto (p/p_{\mathrm{th}})^{d_{\mathrm{tree}}}$ - Thesis II: 3D-integrated transmon qubits in coaxial cavities via TSVs — manufacturable - Thesis III: Prime-frequency spectral engineering for frequency-domain multiplexing - Disconfirmed if: depth-2 tree (8 transmons) shows no improvement over 2D layout

p-Adic Neural Networks: Replace real weights with p -adic valued weights. Non-Archimedean metric enables exact computation (no floating-point rounding).

Tree-Native Programming Language: First-class numeration bases, active zero, base-polymorphic arithmetic, ultrametric distance as primitive.

3.3 The p-Adic Curriculum

Age	Content
5-6	The tree as complete representation of $\mathbb{Z}_2$
6-7	Expansion to base p ; “how far apart in the tree?”
7-8	p -adic metric formalized: $ x _2 = 2^{-\mathrm{depth}}$; clopen sets
8-9	Sequences and convergence: “climbing the tree together”
9-10	Differentiation: difference quotients, strict differentiability
10-11	Integration: Volkenborn integral, Mahler basis, Fundamental Theorem
11+	Introduce $\mathbb{R}$ as the “other” completion of $\mathbb{Q}$

Key principle: Build ultrametric intuition before Archimedean. The line becomes a specialized tool, not the default geometry.

3.4 Interface Design: The Archimedeanized Tree

All interfaces degrade an ultrametric catalog (file system, knowledge graph) into an Archimedean line (scrollbar, feed, search results). Design principle: the tree is ground truth; lateral edges are shortcuts; both views simultaneously available.

4. Production Infrastructure: Ultrametric Engine

The QNFO ultrametric engine (`ask-qwav.q08.workers.dev`, v2.8) implements the core discovery principles in production:

4.1 Deployed State

Metric	Value
Papers indexed	109
Vectors in Vectorize	170 (BGE-M3, 1024-dim)
AI model	@cf/deepseek-ai/deepseek-r1-distill-qwen-32b
Embedding model	@cf/baai/bge-m3
Total Cloudflare Workers	28

4.2 Working Endpoints

Endpoint	Purpose	Mathematical Principle
/api/search?q=X	Semantic search via vector embeddings	Cosine similarity in 1024-dim embedding space — the ultrametric tree is implicit in the nearest-neighbor structure of the vector index
/api/papers	List all papers	1D linear listing — the Archimedean projection of the tree
/api/ask?q=X	AI-powered Q&A over corpus	Retrieval-augmented generation using embedding search
/api/stats	Paper metadata statistics	Quantitative audit of the corpus
/api/health	System health	Operational monitoring

4.3 Theory-to-Engine Mapping

Mathematical Concept	Engine Implementation
Ultrametric distance $d(x, y) = b^{-v_b(x-y)}$	Cosine similarity on 1024-dim BGE-M3 embeddings — an ultrametric in the limit of single-linkage clustering
Tree structure of positional notation	Vectorize index partitions the embedding space into a hierarchical tree (nearest-neighbor graph)
p-adic cache TTL $= 15s \times 2^{\text{ord}_2}$	Documented in skill but not exposed as an endpoint in v2.8
3-phase discovery (word $\rightarrow$ cluster $\rightarrow$ tree)	v2.8 uses vectorize-only search; cluster expansion and tree-based search not deployed
Spectral analysis (Tate, Amice, Intrinsic Amice)	Not deployed in this instance
Bruhat-Tits building	Not deployed — theoretical only

4.4 Architecture

Workers (28 total)

- ask-qwav - ultrametric discovery engine (v2.8)
- ultrametric-tree-api - tree API (separate worker)
- graph-api - knowledge graph (graph-api.qnfo.org)
- qnfo-data-api - unified data API (v2 endpoints)
- living-paper-api - living paper system
- cms-api - content management

22 other operational workers

R2: qnfo/ namespace (canonical storage)

D1: qnfo-audit (18 tables, FTS5)

Vectorize: BGE-M3 embeddings, 170 vectors, 1024-dim
 Pages: deep.qwav.tech, papers.qnfo.org

5. The Gap Analysis

Gap	Description	Impact
External literature search	arXiv and Semantic Scholar APIs unreachable from this environment	Literature discovery limited to internal QNFO corpus (109 papers)
Advanced engine features	Spectral analysis, Bruhat-Tits, Berkovich, dendrogram endpoints not deployed	Engine operates as basic semantic search, not full ultrametric tree explorer
Engine-skill drift	Skill documents 27+ endpoints; v2.8 has 11	Skill documentation needs updating
DNA connection	The genetic code's tree structure is noted as an analogy but not mathematically formalized	Open research question
Empirical validation	Quantum computing thesis, p-adic curriculum, silent-radix crypto — all [speculative] or [my conjecture]	Requires experimental verification

6. Research Roadmap

Timeframe	Milestone	Certainty
1-5 years	Pilot p-adic curriculum vs. control; depth-2 tree quantum chip benchmarked	Feasible
5-10 years	Full weight sequence classification; tree-native programming language prototype	Likely
10-20 years	Tree-based cryptographic standards; DNA-numeration isomorphism proven/refuted	Speculative
20-50 years	Bruhat-Tits numeration for n -dimensional groups; reflective numeration systems	Highly speculative

7. Key References

1. Rofa (Quni-Gudzinas, 2013). *On Number Representation*. arXiv:2005.10207. [EXTERNAL-SOURCE]
2. Tarau, P. (2013). *Arithmetic Algorithms for Hereditarily Binary Natural Numbers*. arXiv:1310.8277. [WEB-SEARCH]
3. Akiyama, S., Frougny, C., & Sakarovitch, J. (2008). On subtrees of the representation tree in rational base numeration systems. *Discrete Mathematics & Theoretical Computer Science*. [WEB-SEARCH]

4. **Charlier, E. & Kreczman, S. (2025).** Positional numeration systems and regular languages. [WEB-SEARCH]
5. **Lecomte, P. & Rigo, M. (2001).** Abstract numeration systems. In *Words, Languages & Combinatorics III*. [LLM-INFERRED]
6. **Chrisomalis, S. (2010).** *Numerical Notation: A Comparative History*. Cambridge. [WEB-SEARCH]
7. **Spencer-Brown, G. (1969).** *Laws of Form*. Allen & Unwin. [established]
8. **Holly, J. (2001).** Pictures of Ultrametric Spaces, the p -adic Numbers, and Valued Fields. *Amer. Math. Monthly*, 108(8), 721-728. [established]
9. **QNFO Research Collective (2026).** *The Ultrametric Foundation*. deep.qwav.tech/papers/ultrametric-foundation/. [EXTERNAL-SOURCE]
10. **Scholze, P. (2012).** Perfectoid Spaces. *Publ. Math. IHES*, 116, 245-313. [established]

Synthesized from: five-thread deep-dive, extensions & future directions analysis, live Cloudflare infrastructure audit (28 Workers, ask-qwav v2.8, 109 papers), and the Ultrametric Foundation thesis.