

A Coherent Cubical Algebra for Unary Scaffold Histories

Joshua Gay
FPRD Lab, Insight Forge

Draft 1, 23 August 2026

Abstract

Scaffolding automata give parsing expression grammars a persistent graph machine model. Previous work developed an algebra of finite scaffold histories and proved that, at degree one and observation distance one, endpoint detours, literal self-loop absorption, causal garbage, and temporal carrier slides connect exactly the histories with one final behavioral figure. It left a path-level question open: which relations generate all equalities between different sequences of those moves?

We answer that question by separating geometry from physical presentation. After collapsing strong and garbage variation, essential residual carriers form a corrected fixed-top order-ideal lattice, and independent temporal slides fill its $\text{CAT}(0)$ cube complex. Over that base, causal garbage fibres are products. Effective detours, absorptions, and slides never enlarge descriptor support. Label transport is strict except at three support captures, each filled by garbage cancellation. Every horizontal critical branching is Peiffer, cubical, triangular, or one bounded absorption–transport hexagon

$$A_y A_x = T_{yz} A_x A_y D_z,$$

where A_x is empty when the lower row already uses literal **SELF**; the apparent detour–transport cell is derived. Horizontal moves already trivial inside a garbage fibre have bounded collapse cells. A direct coherent-Newman induction modulo garbage then shows that this length-independent guarded family presents every parallel elementary path in every fixed-length unary distance-one final fibre.

“Guarded” is essential. Physical endpoint and replay-support certificates range over histories. We do not claim a finite context-closed polygraph, finite derivation type, higher-degree coherence, or novelty for the general $\text{CAT}(0)$ and rewriting machinery.

1 Why paths, rather than only accepted languages?

A formal language is a set of words. A recognizer is a process that reaches that yes-or-no observation through internal states and actions. Context-free grammars are usually read generatively: a successful derivation witnesses membership. Parsing expression grammars (PEGs) are deterministic recognizers: ordered choice and the recognition history decide which suffix remains. Loff, Moreira, and Reis introduced *scaffolding automata* as a persistent graph machine model for PEGs [5]. At every input letter one bounded row is appended to a permanent scaffold; the new row can inspect only a bounded neighborhood of the preceding top.

Collapsing a scaffold computation to its accepted language forgets several different objects:

$$\text{row word} \longrightarrow \text{physical history} \longrightarrow \text{strong trace} \longrightarrow \text{final figure} \longrightarrow \text{acceptance}.$$

The physical history remembers node identity and construction time. Its strong trace remembers the behavioral figure after every prefix. The final figure forgets all earlier observations. Acceptance forgets almost everything else.

The companion paper *Persistent scaffold computation* develops these layers and four elementary moves [3]. It proves one right-packed object normal form for every unary distance-one final fibre. It also proves path coherence inside each garbage fibre and, after collapsing whole strong/garbage chambers, in the carrier-slide quotient. It explicitly does not present all mixed physical derivations. The present paper begins at that boundary.

Working picture. The object is fibred. Carrier placements form the base geometry. Different physical spellings of one placement form a vertical fibre. Temporal slides transport presentation data between adjacent fibres. Higher cells say when two vertical or horizontal operations can be interchanged.

2 The unary history calculus

Fix a nonempty finite label alphabet, history length n , scaffold degree one, and observation distance one. There is one distinguished unlabelled base at time 0. Row i has a label and one descriptor from

$$\text{MISSING}, \quad \text{SELF}, \quad \epsilon, \quad 0.$$

If $e(i)$ is its physical endpoint, then **MISSING** is absent, **SELF** is i , ϵ is $i - 1$, and 0 copies $e(i - 1)$. The final live chain starts at n and iterates e until missing or a self-loop. Minimizing equal residual behaviors gives the final behavioral figure.

The labels of final-live rows are supported. Every descriptor on that chain is supported; if a supported descriptor is 0, the descriptor immediately below it is supported recursively. This is the least replay-closed *causal support*. Changing any unsupported label or descriptor is a garbage move G .

Definition 2.1 (Horizontal moves). An endpoint detour D replaces a descriptor by the least spelling with the same physical endpoint. A literal absorption A applies when a fresh row points to an old physical node q with $e(q) = q$ and presents the same recursive residual; it replaces that old pointer by fresh **SELF**. A temporal slide T exchanges

$$(\epsilon, 0) \longleftrightarrow (0, \epsilon)$$

on two consecutive rows when the two candidate carrier labels agree.

Literal physical self-looping matters. A row may present a recursive behavior while pointing to a different old carrier; changing it directly to **SELF** is not then a primitive absorption. An early computational model in this project made exactly that error. All statements below use the literal guard.

Each move preserves the final behavioral figure. Detours and absorptions also preserve every prefix figure; temporal slides may change intermediate figures. We study the undirected elementary path groupoid generated by D, A, T, G , while orienting selected horizontal edges only to prove termination.

3 The CAT(0) carrier base

Reduce the final live chain behaviorally and omit the arbitrary physical carrier of a terminal recursive residual. Let $t_1 < \dots < t_m$ be the construction times of the remaining essential carriers. The final root is carried at the final time, so $t_m = n$.

Proposition 3.1 (Corrected carrier rectangles). *For an ordinary missing or base terminal and $m \geq 1$, carrier placements are the ideals of $[m-1] \times [n-m]$. For a recursive terminal, $m = 0$ is a singleton; for $m \geq 1$, placements are the ideals of $[m-1] \times [n-m-1]$.*

Proof. In the ordinary case

$$1 \leq t_1 < \dots < t_{m-1} < t_m = n.$$

Put $\lambda_i = t_i - i$. Then $0 \leq \lambda_1 \leq \dots \leq \lambda_{m-1} \leq n - m$, and the λ_i are the row lengths of an ideal in the displayed rectangle. The inverse is $t_i = \lambda_i + i$ with $t_m = n$.

At a recursive bottom, the omitted physical loop carrier must occur below the first retained finite carrier, so time 1 is unavailable. Put $\lambda_i = t_i - i - 1$ to obtain the second rectangle. Conversely, the unique relay $\epsilon, 0, \dots, 0$ realizes every listed placement; unused coordinates are garbage. Strong and garbage moves preserve the essential tuple, and all histories with one tuple lie in one chamber [3]. $\square$

A slide changes one λ_i by one while preserving weak increase. It is exactly the addition or removal of one maximal ideal box. Ardila, Owen, and Sullivant construct from any poset with inconsistent pairs a rooted CAT(0) cube complex whose vertices are consistent ideals and whose cubes remove subsets of maximal elements [1, Definition 2.3 and Theorem 2.5]. A rectangle has no inconsistent pairs.

Theorem 3.2 (Carrier cube). *Filling every Boolean family of independent temporal slides gives the rooted CAT(0) order-ideal cube of Proposition 3.1. For a rectangle with r rows and c columns it has rc hyperplanes, dimension $\min(r, c)$, edge distance*

$$d(t, s) = \sum_{i=1}^r |t_i - s_i|,$$

and coordinatewise-median carrier tuples.

Proof. There is one hyperplane per box and one cube per antichain of currently removable maximal boxes. The maximum antichain in a rectangle has size $\min(r, c)$. Any edge path must cross every box in the symmetric difference of its endpoint ideals; removing and adding those boxes in legal order attains that bound. The median ideal is $(I \cap J) \cup (J \cap K) \cup (K \cap I)$, whose row lengths are the coordinatewise medians. $\square$

This theorem is recalled from the corrected Draft 3 of the companion paper. Its role here is structural: pure temporal schedules already have cubical coherence. The unresolved work lies in lifting that geometry through the physical presentation fibres.

4 Rewriting modulo causal garbage

Let $\mathcal{G}$ be the symmetric relation of single-coordinate garbage changes. Support stability makes every $\mathcal{G}$ -class a finite Cartesian product. Same-coordinate triangles compose successive values; different-coordinate squares commute them. These cells make each garbage fibre coherent.

On garbage classes orient effective D, A, T edges by

$$\mu(H) = (g(H), w(H)),$$

where g is total displacement of the behaviorally reduced essential tuple from its right-packed tuple, and w is the supported descriptor word in the order

$$\text{MISSING} < \text{SELF} < \epsilon < 0.$$

D and A preserve the reduced tuple and lower w . An essential T lowers g ; an inessential T is oriented toward smaller w . Thus every effective horizontal edge strictly decreases μ .

Lemma 4.1 (Unique quotient normal class). *Every final fibre has one irreducible garbage class, the right-packed class $\text{RP}_n(F)$. Hence oriented effective horizontal rewriting terminates and is object-confluent modulo garbage.*

Proof. Finiteness and strict decrease give termination. If $g > 0$, choose the highest carrier gap. The unique-relay lemma exposes a supported $\epsilon, 0, \dots, 0$ bridge. Its first hidden relay label is unsupported and can be prepared inside the same garbage class; the adjacent essential slide then reduces g .

At gap zero, a nonleast supported endpoint exposes D , a supported repeated literal recursive residual exposes A , and an inessential adjacent inversion exposes T . When none applies, the terminal is forced to be the canonical missing, recursive, or base bottom described in the companion proof. This is exactly $\text{RP}_n(F)$, and it has no effective outgoing edge. Every noncanonical class therefore has a reduction and the canonical class is the unique normal class. Termination plus uniqueness yields confluence. $\square$

Some legal horizontal edges are already identities in $\mathcal{G}$. They must not simply disappear from a theorem about physical paths.

Lemma 4.2 (Garbage-trivial collapse). *If a legal D or A edge is trivial modulo garbage, it equals one descriptor garbage edge. If a legal T edge is trivial modulo garbage, it equals a two-descriptor garbage path; the two orders agree by the garbage square.*

Proof. Equal garbage normal forms have equal causal support. D/A change one descriptor and T changes two. Any changed supported value would survive normalization, so every changed coordinate is unsupported. Replay invariance keeps support fixed after the first of the two T garbage changes. $\square$

5 Support transport

The next question is whether a garbage change commutes with a horizontal move. Descriptors and labels behave differently.

Lemma 5.1 (Descriptor-support monotonicity). *For every effective oriented horizontal move R ,*

$$\text{DescSupp}(RH) \subseteq \text{DescSupp}(H).$$

Every coinital horizontal/garbage-descriptor branching is therefore a strict naturality square modulo garbage.

Proof. A detour replaces an indirect 0 name by a direct least name of the same endpoint, deleting replay dependencies. An absorption replaces a replay into an old self-loop by fresh **SELF**, again terminating dependencies. A slide exchanges the roles “live carrier” and “certified replay dependency” among the same three local descriptor cells; it introduces no descriptor outside the source closure.

An unsupported descriptor cannot influence any endpoint or guard read by an effective move. The same residual therefore remains available after changing it, and monotonicity leaves the changed cell unsupported at the target. The two target histories differ by one garbage edge. $\square$

Horizontal motion can promote a label from garbage to visible support. There are exactly three possibilities. Absorption replaces an old self-loop carrier by its fresh presentation, so only the fresh row label can be captured. A temporal slide substitutes one of two equal-labeled adjacent carriers, so the gaining label is either the first slide row or its predecessor, depending on direction.

Proposition 5.2 (Support captures). *The only mixed branches without a strict horizontal residual are*

$$A/GL(0), \quad T/GL(0), \quad T/GL(-1).$$

They require no new generator: the garbage-first path returns along the inverse garbage edge and then takes the original reduction, so $G; G^{-1}; R = R$ by vertical cancellation.

6 The absorption–transport hexagon

The non-Peiffer horizontal interaction occurs on three consecutive rows x, y, z . Suppose the descriptors at y, z are $(0, \epsilon)$, the labels of x, y agree, literal A_y is available, and T_{yz} is oriented from this source. Let A_x be absorption at x when x is not already **SELF**, and the empty path otherwise.

Theorem 6.1 (A/T hexagon). *The following paths have exactly the same raw physical endpoint:*

$$A_y; A_x = T_{yz}; A_x; A_y; D_z.$$

No garbage quotient is used in this equation.

Proof. Let $q = e(x)$ be the endpoint named by y ’s 0. The literal A_y guard says that q is an old physical self-loop and that y presents its recursive residual. The temporal label guard makes x present that residual as well. Thus, if x is not already **SELF**, one literal A_x is available; no long prefix normalization is needed.

Write p for x ’s old descriptor. The left path gives

$$(p, 0, \epsilon) \xrightarrow{A_y} (p, \text{SELF}, \epsilon) \xrightarrow{A_x} (\text{SELF}, \text{SELF}, \epsilon).$$

The right path gives

$$(p, 0, \epsilon) \xrightarrow{T} (p, \epsilon, 0) \xrightarrow{A_x} (\text{SELF}, \epsilon, 0) \xrightarrow{A_y} (\text{SELF}, \text{SELF}, 0) \xrightarrow{D_z} (\text{SELF}, \text{SELF}, \epsilon).$$

At the last step, 0 follows the now-self-looping row y , while ϵ names y directly. They have the same physical endpoint. $\square$

If D_y is coinitial with the same slide, equality of the 0 and ϵ endpoints forces x itself to be literal SELF. Then $D_y A_y = T_{yz} A_y D_z$. The ordinary D/A triangle identifies $D_y A_y$ with direct A_y , so D/T is a pasting of D/A with the SELF-eager hexagon and is not an independent cell.

7 Complete local classification

Two detours commute because they preserve physical endpoints. A lower absorption acts on every affected later 0 replay by the uniform substitution of its old self-loop carrier by the fresh equal-behavior carrier. Uniform substitution preserves endpoint equalities and transports other literal absorption guards. Hence strong/strong peaks are Peiffer squares or the same-row D/A triangle.

For T/T, two essential reductions remove incomparable maximal boxes in the carrier cube and commute. The only overlapping raw descriptor words are $\epsilon, 0, \epsilon$ and $0, \epsilon, 0$. If the repeated residual is recursive, both swaps are inessential and exactly one direction lowers descriptor order. If it is nonrecursive, the swaps move essential carriers oppositely and exactly one lowers carrier gap. Thus no overlapping oriented T/T peak exists.

A slide transports suffix endpoints by one uniform substitution between its adjacent equal-behavior carriers. Strong guards away from the first slide coordinate therefore commute. The apparent exception is an upper absorption whose old physical self-loop would be replaced by a non-self-looping equal-behavior row. If that upper absorption is effective, the lower recursive block is behaviorally inessential; the threatened slide direction then increases descriptor order and is not oriented. The reverse orientation preserves the literal guard. At the first slide coordinate, the only cases are the A/T hexagon and derived D/T diagram above.

Theorem 7.1 (Local cell classification). *Every coinitial effective horizontal pair is aspherical, Peiffer/cubical, D/A triangular, A/T hexagonal, or D/T derived. Every horizontal/garbage pair is a strict naturality square or one of the three cancellation captures. Every garbage-trivial horizontal edge has a collapse cell from Lemma 4.2.*

8 Uniform guarded coherence

We use a direct finite-graph version of coherent Newman induction modulo relations. Let E be a symmetric relation with a coherent path presentation, and let R terminate on E -classes with one normal class. If every R/R and R/E local branch has a chosen confluence cell modulo E , then those cells present every parallel path. Indeed, canonicalize vertical subpaths and induct on the quotient reduction measure and remaining horizontal length. Replace the first disagreement by its local cell; every horizontal tail lies below the source class. An E-capture may use an inverse vertical edge, but it needs no recursive horizontal induction. This is the proof architecture of coherent confluence modulo [2], proved here directly because our guards are not assumed to be a conventional context-closed polygraph.

Theorem 8.1 (Uniform guarded path presentation). *Fix a finite alphabet and history length in degree one and distance one. In each final behavioral fibre, all parallel paths generated by endpoint*

detours, literal absorptions, causal-garbage changes, and temporal slides are presented by the following length-independent guarded cell families:

1. *garbage same-coordinate triangles and different-coordinate squares;*
2. *collapse cells for garbage-trivial D/A/T edges;*
3. *strong Peiffer squares and the same-row D/A triangle;*
4. *carrier-cube and inessential temporal Peiffer squares;*
5. *strict mixed naturality squares; and*
6. *the bounded A/T hexagon.*

The three support captures use vertical cancellation, and D/T is derived.

Proof. Lemma 4.1 supplies quotient termination and one normal class. Garbage fibres have their product cells. Lemma 4.2 eliminates horizontal edges already vertical. Theorem 7.1 supplies every remaining local R/R and R/E cell. Apply the direct coherent-Newman induction modulo garbage. $\square$

Boundary. This is a finite list of guarded schema *types*; it is not a finite context-closed polygraph. Endpoint-equality and causal-support certificates range over histories. Guiraud and Malbos show why indexed critical branchings make finite-derivation-type conclusions delicate in higher dimensions [4]. The bounded hexagon controls the observed A/T index, but no finite derivation type claim is made.

9 Evidence, failure modes, and the larger program

The promoted exact atlas checks complete domains with one label through length eight, two labels through length five, and three labels through length four: 147,448 histories. It asserts quotient termination, one normal class, descriptor-support monotonicity, the exact A/T and D/T critical signatures, the exact three label captures, and the collapse classification. A separate raw-history checker verifies 12,914 A/T hexagons, including 4,116 with a non-SELF-eager lower row.

A homology mutation is used only as a falsifier. Removing the overlapping D/T cell while retaining D/T Peiffer squares leaves zero residual first homology. Retaining only SELF-eager A/T instances also leaves zero. Removing A/T exposes 120 residual classes. A seeded length-10/20/40 probe checks 67,848 additional coinital pairs and finds no new non-Peiffer type. None of these finite results proves Theorem 8.1; the causal arguments do.

An independent reviewer reconstructed the proof and found one missing edge class: legal horizontal moves already trivial modulo garbage. The collapse lemma was added, the quotient normal-form proof was expanded, and the same reviewer returned “gap closed.” This is review evidence, not authorship or specialist endorsement.

The fibred formulation suggests a higher-degree program. Replace the unary carrier chain by the dependency diagram of essential residual records. Ask whether legal time placements form a poset-with-inconsistent-pairs cube or a different event geometry. Put detours, folds/unfolds, receipts, SELF, and garbage in vertical fibres; treat sewing as transport; then classify the cells where transport changes support or overlaps a vertical fold. Failure of CAT(0), unbounded indexed branchings, or proliferating support curvature would be a precise obstruction rather than merely a failed language-level proof.

Assurance and contribution statement

The ordinary mathematical arguments are the proof. Executable checks are independent falsifiers only to the extent stated in the public evidence ledger. No proof assistant is used as a stronger label. AI systems assisted with exploration, checking, adversarial review, and exposition; Joshua Gay is responsible for the public claims and their disposition.

The FPRD contribution is the separation of presentation, dynamics, and observation. That separation exposed an exact $\text{CAT}(0)$ base, causal fibres, partial transport, and one curvature cell. It also exposed and repaired the overbroad absorption model, overcounted support generators, incorrect carrier rectangle, and omitted garbage-trivial edges.

References

- [1] Federico Ardila, Megan Owen, and Seth Sullivan. Geodesics in $\text{CAT}(0)$ cubical complexes. *Advances in Applied Mathematics*, 48(1):142–163, 2012. <https://doi.org/10.1016/j.aam.2011.06.004>.
- [2] Benjamin Dupont and Philippe Malbos. Coherent confluence modulo relations and double groupoids. arXiv:1810.08184v3, 2021. <https://arxiv.org/abs/1810.08184>.
- [3] Joshua Gay. Persistent scaffold computation: causal presentations, sewing, and feedback. FPRD Lab review paper, Draft 3, 23 August 2026. <https://gitlab.com/insight-forge/boxlab/fprd-lab/-/tree/main/publications/persistent-scaffold-causal-algebra>.
- [4] Yves Guiraud and Philippe Malbos. Higher-dimensional categories with finite derivation type. arXiv:0810.1442v2, 2009. <https://arxiv.org/abs/0810.1442>.
- [5] Bruno Loff, Nelma Moreira, and Rogério Reis. The computational power of parsing expression grammars. *Journal of Computer and System Sciences*, 111:1–21, 2020. <https://doi.org/10.1016/j.jcss.2020.01.001>.

Afterword

A coherent cubical algebra for unary scaffold histories

Methodological and contribution addendum. The paper's mathematical scope and review status remain as stated in the paper.

The Lab's approach

FPRD means **Finitely Presented Recognition Dynamics**. Its starting point is to distinguish a finite description, the possibly infinite process it generates, and what an observer can learn from that process. Equal answers need not come from equal internal structures. Conversely, an infinite state space may have a short, exact description and useful coordinates; failure to compress it into finitely many states need not end the investigation.

A small example: balanced parentheses

The grammar $S \rightarrow (S)S \mid \epsilon$ presents the language of balanced parentheses. To recognize a prefix, keep the number of unmatched opening parentheses, or a failure state if a closing parenthesis arrived too early. An opening parenthesis increments the count. A closing parenthesis decrements a positive count; at zero it enters failure. Failure persists. At the end, accept exactly when the count is zero.

The prefixes $($ and $(($ leave the same future task: any suffix that completes one completes the other. The prefix $(($ leaves a different task: $)$ completes the first two prefixes but not this one. These future tasks are **residuals**. There are infinitely many counts, yet a finite set of update rules generates their dynamics. This is a classical automata example, not a discovery of FPRD Lab. It illustrates the research question: which coordinates preserve all relevant futures, and what does updating or reading those coordinates cost?

From viewpoint to proof

In practice, the Lab fixes the observable and the allowed operations, compares presentations of the same problem, and seeks an exact translation or invariant. A failed shortcut is useful when a counterexample identifies the missing coordinate or hypothesis. Small computations test particular uncertainties; a proof must explain why the result holds beyond those tests. Costs, relation sides, boundary conditions, and the direction of each translation remain explicit.

AI-assisted exploration, source comparison, executable checks, and adversarial reconstruction support this work. Their agreement is not automatically independent evidence: shared assumptions must be exposed. A formal development certifies its stated formal theorem, not every interpretation attached to it. Each paper retains its own evidence and review status.

The approach is most useful when a poor representation hides a reusable construction, a compatibility condition, or a sharp obstruction. Outside recognition theory, it can guide the search for a structural reformulation without producing a literal residual dynamical system. The account below distinguishes that methodological influence from the ingredients of the proof.

Contribution to this publication

This addendum describes the current HTML reading edition. The preceding PDF pages retain their earlier edition; the afterword does not silently revise them.

FPRD's contribution to this paper

The preceding causal-scaffold work compared histories and their observations. This paper asks a further question: when two legal sequences of changes have the same endpoints, which local identities explain their agreement? It therefore studies transformations between presentations, not just a language or a collection of machine states.

The Lab separated the placement of essential carriers from the physical ways of presenting a fixed placement. Carrier times form an order-ideal geometry; irrelevant data and descriptor choices lie over it. Temporal moves can change which data are relevant, so they cannot be treated as unconditional commuting operations.

The difficult interaction is between absorption and temporal transport. The paper writes an exact bounded hexagon whose two paths reach the same physical descriptor triple, including the case requiring an intermediate self-loop adjustment. Together with the other guarded local cells, termination modulo garbage and a confluence argument give the stated coherence theorem.

The failure record is material to that proof. Overbroad absorption, an incorrect carrier rectangle, and overcounted support moves were narrowed to the actual source model. A later audit exposed legal horizontal edges already trivial modulo garbage; the collapse lemma supplies that missing class. Finite homology probes were useful falsifiers, but vanishing in such a probe did not certify that the list of cells was complete.

Inherited ingredients and the boundary of the contribution

The scaffold model comes from Loff, Moreira, and Reis. Ardila-Owen-Sullivant provide the cubical-geometric comparator; coherent confluence modulo relations and the higher-dimensional rewriting literature supply the proof architecture and its cautions. The paper credits Dupont-Malbos and Guiraud-Malbos rather than presenting general coherence machinery as a Lab invention.

The Lab's specific work is the exact unary carrier geometry, support-sensitive transport, critical interactions, and their assembly into a uniform guarded theorem. "Uniform" means a finite list of kinds of guarded identity at arbitrary finite length. It does not mean a finite context-closed polygraph or a finite-derivation-type theorem. The result remains restricted to the declared unary, distance-one model. Computation and AI-assisted reconstruction support the written argument; no proof-assistant verification is claimed.

Where to inspect the contribution: the paper, Sections 3–9, especially the absorption-transport hexagon, collapse lemma, and evidence boundaries.