ICRA PRE-PRINT · ICRA-30

A Type Theory of Sense

Witnessed Choice in Stratified Semantic Spaces

Iman Poernomo

June 2026 · Institute for Co-Recursive Agency ·

A Type Theory of Sense:
Witnessed Choice in Stratified Semantic Spaces

Iman Poernomo

June 2026

Abstract

Wherever meaning is studied empirically, semantic verdicts come from instruments: an embedding model at a resolution decides whether two completions "say the same thing," a finer resolution revokes the verdict, a different encoder disagrees licitly with both. Logics of meaning have had no place for any of this — no judgment form for the record of a measurement, no laws for how verdicts propagate when the instrument is refined or replaced, no guarantee that the analysis adds nothing to its data. We present **TTS**, a constructive extension of Martin-Löf type theory built for exactly that practice: a *logic of consequence from measurement data*. Judgments take the form $\Gamma \vdash_\Delta \mathcal{J}$, where Δ is a *measurement context* — an instrument's verdict record, entering as constants. Two regime-indexed relation formers carry the logic: indiscernibility and apartness in the constructive tradition of Brouwer, Heyting and Bishop, graded by *regimes* — instruments at declared resolutions, ordered by refinement. Apartness has no premise-free introduction rule, and the metatheory makes the architecture enforceable rather than rhetorical: soundness over regime-indexed simplicial apartness spaces, conservativity over the base theory, no-fork-ex-nihilo, no-refutation-ex-nihilo, provenance (every derivable fork is forced by the record in every realizing model), and consistency of honest records. Two grades of compositional demand are derived, not primitive — *canonicity* (the instrument cannot distinguish any two fillers: composition is forced) and the *fork* (two warranted fillers, recorded apart: composition demands a genuine choice) — with a third, *locked*, constructed from them. Exclusion is a theorem; forks persist under refinement while canonicity does not, also theorems; and the consistency of an identity discovered by one instrument with a separation held by another is characterized exactly by the order relation between the instruments. The principal application is a formal reconstruction of Fregean sense: a *sense* is a choice of filler of a compositional demand, *reference* its boundary, hyperintensional distinctness apartness at a regime, the informativeness of identity its record-dependence — one cross-regime law, instantiated by lexical ambiguity at filler spaces and by the name-state configuration of Frege's celestial case. The motivating subject matter is the geometry of meaning in large language models, whose token spaces are measurably not manifolds: semantic space is stratified, and trajectories of generation encounter loci of genuine choice that the Kan condition of homotopy type theory renders invisible. The empirical sections instantiate the calculus's central judgments in generation behaviour — resampling and fill-in-the-middle protocols, fully logged — as a demonstration of non-vacuity and one replicated qualitative prediction, with an explicit honesty clause; the logic itself is not specific to machine systems.

1 Introduction

1.1 Meaning after the manifold

A large language model in operation traverses a learned space. A context is a position; a continuation is a directed step; an unfolding conversation is a trajectory. Whatever "meaning" is for such a system, it is enacted in the geometry of these traversals. For a decade the default picture of that geometry was the *manifold hypothesis*: representations live on a smooth low-dimensional submanifold of the embedding space, so that locally there is always a well-behaved, essentially unique way to continue.

The default picture is now false as a matter of measurement. Robinson, Dey and Chiang showed that the token subspaces of GPT-2, Llemma-7B, Mistral-7B and Pythia-6.9B reject the manifold hypothesis — and, for three of the four, the weaker fiber-bundle hypothesis — at significance $p < 3 \times 10^{-7}$, with some token neighbourhoods possessing no well-defined local dimension at all [43]; an earlier study of the same group found the token subspace to be a *stratified* space, with strata of varying dimension and curvature [42], and independent probes confirm strongly varying local intrinsic dimension in contextual embeddings [44, 31]. Complementarily, Bigelow et al. exhibited *forking tokens* in autoregressive generation: positions at which resampling a single token flips the entire downstream outcome distribution between discrete alternatives [11]. Generation, observed from outside, passes through points of genuine choice.

This paper takes the measurements seriously as constraints on *logic*. If semantic space is stratified rather than smooth, then a formal theory of meaning for such spaces must be able to express, within its formal apparatus, the three local *grades* of compositional demand a trajectory can encounter:

(i) **Canonical.** Locally the space behaves like a manifold: a continuation exists and is essentially unique. No semantic decision occurs. (The manifold hypothesis holds here; in simplicial language, inner horns fill uniquely — the Segal condition [25, 41].)

(ii) **Forked.** Composition succeeds in several *inequivalent* ways: the space of fillers is inhabited but disconnected. The trajectory must choose, and nothing in the local geometry makes the choice for it. (The manifold hypothesis fails; this is the grade the measurements above detect.)

(iii) **Locked.** Composition is canonical *and* every exit is foreclosed: the trajectory is confined to a tightly self-confirming region. (Degenerate fluency; in language models, the phase of repetition and collapse [3, 36].)

A caution before the formalism, because the introduction has so far spoken loosely. Manifold smoothness, Segal-style uniqueness of composition, and connectedness of a measured completion distribution are *three different properties*; none follows from the others. A smooth manifold contains indefinitely many paths — unique continuation needs dynamics, not just smoothness — and non-manifoldness does not by itself disconnect any particular filler space. The honest statement of the bridge: non-manifold loci *motivate candidate sites* at which compositional canonicity should be tested; whether a given filler space separates is measured independently, by the fork test of §6. The grades above are the formal properties; the geometric glosses are the intended picture in the models of §5, not identifications.

One more lesson of the measurements shapes the design. Every verdict above is *instrument-relative*: the geometry was measured by a particular probe at a particular resolution, and a structure invisible at one resolution is decisive at another. We give this relativity first-class status: TTS judgments are graded by *regimes* — instruments at declared resolutions, ordered by refinement — and the laws relating verdicts across regimes (separations persist under refinement; canonicity does not; two instruments' verdicts on one pair can disagree exactly when the connecting instrument is not a refinement of the separating one) are theorems of the system, not glosses.

Homotopy type theory (HoTT) [48] is constitutionally blind to (ii) and (iii). Its types are Kan complexes: every horn fills, and the space of fillers of an inner horn is contractible. Composition never poses a question. This is precisely what makes HoTT an extraordinary logic of *mathematical* space — and what makes it the wrong logic of *semantic* space, where the interesting events are exactly the failures of canonicity.

1.2 Sense as choice of filler

The second motivation is older. Frege distinguished the *reference* of an expression from its *sense* — the mode of presentation under which the reference is given — to explain why true identity statements can be informative [19]. The distinction has resisted mathematization: possible-worlds intensions are too coarse (logically equivalent contents collapse), while syntactic individuations are too fine. The modern proposals — sense as algorithm [34], sense as procedure in a ramified hierarchy [16], sense as program with computational identity [10], sense of a proof as its derivational structure [46, 7, 8] — share a conviction we endorse: sense is something *enacted*, a way of arriving, not a static intension. What they lack is a geometry in which "ways of arriving" have intrinsic structure: a space of ways, with its own connectivity, whose disconnectedness is precisely hyperintensional difference.

TTS supplies that geometry, and the stratified picture of semantic space supplies its empirical referent. The central identification of this paper:

> *The reference of a compositional demand is its boundary — the horn: what is already given and must be respected. A sense is an inhabitant of the filler space of that horn — a way of completing what is given. Hyperintensional distinctness is apartness of fillers. An identity statement is informative because a path between apart-held fillers must be constructed and witnessed; it is never given for free.*

Where the filler space is canonical (grade (i)), sense and reference coincide: there is only one way of arriving, and Frege's distinction has no purchase — which is why it is invisible to any logic that assumes the Kan condition. The distinction *is* the fork. Frege's puzzle lives exactly where the manifold hypothesis fails.

1.3 The grounding requirement

A logic that asserts "these two ways of meaning are distinct" owes an account of what *witnesses* the distinctness. Negative formulations — a primitive judgment of non-coherence, or an unprovability claim — make the witness endogenous: the theory stipulates the failure rather than exhibiting it, and the judgment floats free of any mathematics that could certify it. We regard this as a hard adequacy condition, and we meet it twice over. First, with a venerable constructive instrument: an *apartness* structure in the tradition of Brouwer, Heyting and Bishop [23, 12, 47]. Apartness is positive: $f \#_V g$

is inhabited by a separation witness, subject to irreflexivity, symmetry and cotransitivity, and indexed by the regime V that achieved the separation. Second — and this is the architectural step — apartness has *no premise-free introduction rule*: separation witnesses enter the theory only as the constants of a *measurement context* Δ, the record of an actual instrument run. The judgment form of TTS is $\Gamma \vdash_\Delta \mathcal{J}$: hypotheses on the left, the record below. Nothing in the theory is stipulated to fail; failures of canonicity are exhibited by data, and the slogan is enforced by the metatheory — *variables hypothesize, constants record.*

Three consequences distinguish the resulting theory from any system with primitive negative judgments. First, the mutual exclusion of canonicity and fork is a *theorem* (Theorem 3.5): indiscernibility transports apartness into self-apartness, which is absurd. The logical architecture polices itself. Second, the calculus provably cannot manufacture its own verdicts: it is conservative over its base, derives no fork from the empty record (Corollary 5.9), and every derivable fork holds in every model realizing the record (Theorem 5.11). TTS is, by construction and by theorem, a **logic of consequence from measurement data**: the world supplies the verdicts; the logic audits what they force. Third, the theory needs no primitive judgment of *emptiness* at all. In the intended models generation is total — there is always a next token — so the constructively meaningful failure of the Kan condition is never "no filler" but "too many, inequivalently." Absence, where needed, is the ordinary derived negation $\mathsf{Fill}(\sigma) \to \mathbf{0}$; the load-bearing new judgment is the fork.

1.4 Contributions and plan

(1) A dependent type theory, TTS, extending Martin-Löf type theory with directed hom- and triangle types, regime-indexed indiscernibility and apartness, and measurement contexts, presented calculus-first with proof terms (§2).

(2) The derived grades of canonicity and fork; the exclusion theorem; the persistence asymmetry (forks persist under refinement, canonicity does not); and the cross-regime consistency theorem characterizing exactly when instruments may disagree (§3).

(3) A formal reconstruction of sense, reference, hyperintensionality and the informativeness of identity, with lexical ambiguity as a worked example — and the names configuration (name-states in a presentation space) falling under the same cross-regime law: two Fregean phenomena, one theorem, two typing instantiations (§4).

(4) A semantics in regime-indexed simplicial apartness spaces with soundness; metatheorems making the measurement architecture more than rhetoric — conservativity, no-fork-ex-nihilo, provenance, and consistency of honest records; and the characterization of *basins* as components whose internal demands are canonical, with *locked* basins as the third grade (§5).

(5) A falsifiable bridge to measurement: apartness verdicts at pullback regimes as the formal content of resampling experiments, with manifold-hypothesis violations and forking tokens as the empirical anchors (§6).

Related work is discussed in §7; open problems, including completeness and mechanization, in §8. One independence claim, stated up front because it allocates risk. The empirical case is at the pilot stage and says so (§6); the philosophical and logical contributions do not wait on it. That HoTT is constitutionally blind to compositional choice, and that a calculus of witnessed, graded verdicts can see it — with exclusion, persistence-asymmetry, and cross-instrument consistency as theorems — stands even if semantic spaces prove smoother than the present measurements suggest. Put most plainly: what this paper builds is a small logic of *instrument-relative semantic disagreement*, in which difference must be witnessed, identification need not erase difference, and verdicts propagate asymmetrically under refinement. Frege, stratified LLM geometry, and alignment-induced mode collapse are three applications of that one machine.

How to read this paper. The experiments of §6 do not test the theorems of §3 and §5 — theorems are not at empirical risk, and a reader who searches §6 for a confirmation of the exclusion theorem will not find one, by design. What the experiments establish is something else, in two parts. First, *non-vacuity*: the judgments the theorems govern — fork verdicts, measurement records, cross-regime configurations — are non-trivially instantiated by the generation behaviour of actual language models, so the calculus is a logic *of* something observable rather than a logic awaiting a subject. Second, *one qualitative prediction*: that enriching a horn's boundary prunes its filler distribution toward a single basin (§4.3), observed as the suffix effect and replicated across two sampling instruments. The relation between calculus and pipeline is *implementation, not analogy*: the estimator of §6 computes the defining relations of the resolution model on the sample (Example 5.4; Table 1). The laws of §2.5 therefore hold in the pipeline by construction; what is contingent, and observed, is which verdicts the world writes into the record. The honesty clause of §6 states the directions in which the non-vacuity demonstration could have failed, and could yet fail at scale.

A note on method: the paper is organized so that nothing is asked of the reader on credit. An actual measurement, four data points small, opens §2 (Example 2.1), and every construct of the calculus is anchored to it as it is introduced; the full experiment of §6 is that example at scale.

2 The calculus TTS

TTS is an extension of intensional Martin-Löf type theory (MLTT) with Π-, Σ-, identity, empty, unit, sum and propositional-truncation types; we follow the conventions of [48], writing $p : a =_A b$ for paths, refl, and tr for transport. Four layers are added: a *semantic signature* and its types (§2.1), directed structure (§2.2–2.3), *regime-indexed* indiscernibility and apartness (§2.4–2.5), and *measurement contexts* (§2.6). The slogan governing the whole design, defended in §2.6: **variables hypothesize; constants record.**

A running example. Throughout, the reader is invited to hold one concrete instance of every formal object. Let A be a space of contextual states of a discourse — for a language model, embedded prompt states — and a, b, c three such states. A continuation $f : \mathsf{hom}_A(a, b)$ is a realized stretch of discourse carrying the context from a to b: say, the words arriving at an occurrence of *bank*. A second continuation $g : \mathsf{hom}_A(b, c)$ departs from that occurrence without disambiguating it. The pair (f, g) is a compositional demand: the passage must be readable whole. Everything below is the logic of that situation.

And so that no construct of this section floats free of measurement, we fix at the outset one *actual* instrument run, in miniature, drawn from the experiment reported in full in §6. The calculus will refer back to it constantly; nothing later in the paper is needed to understand it.

Example 2.1 (A measurement, in miniature)**.** A language model (`gpt-3.5-turbo-instruct`, temperature 1.0) was given the boundary

> *"We met them near the bank, just before sunset."*

Among its sampled continuations are these four (first sentences shown; all data verbatim from the run):

φ_1: *"Corrapiran decided to keep watch outside, while I entered the small bank alone."*

φ_2: *"The bank was closed and the grounds were abandoned when we got there."*

φ_3: *"They were standing on the riverbank opposite from us, watching the water flow by."*

φ_4: *"As we paddled our kayaks up to the bank, we see that they were waiting for us."*

The *instrument*: embed each continuation (`text-embedding-3-small`) and measure cosine distances. The measured values:

$d_{\cos}$	φ_1	φ_2	φ_3	φ_4
φ_1	0	.585	.691	.698
φ_2		0	.720	.652
φ_3			0	.455
φ_4				0

Read at resolution $\varepsilon = 0.60$ — call this instrument V — the four continuations fall into exactly two chain-components: $\{\varphi_1, \varphi_2\}$ (the financial readings, joined by their .585 step) and $\{\varphi_3, \varphi_4\}$ (the riparian readings, joined by .455), with every cross-distance exceeding 0.60. The instrument's *record* of this run is the verdict list

$$\Delta_0 = \big(\gamma_1 : \varphi_1 \simeq_V \varphi_2,\ \ \gamma_2 : \varphi_3 \simeq_V \varphi_4,\ \ \delta_1 : \varphi_1 \#_V \varphi_3,\ \ \delta_2 : \varphi_1 \#_V \varphi_4,\ \ldots \big)$$

— two connections, four separations. Now vary the instrument and watch the verdicts move. At the *finer* resolution $\varepsilon' = 0.40$ (call it V', with $V' \leq V$) every pairwise distance exceeds the threshold: all four continuations separate; the δ-entries survive, the γ-entries do not recur. At the *coarser* resolution $\varepsilon'' = 0.70$ (call it W, with $V < W$) the four continuations form a *single* component: this instrument connects φ_1 to φ_3 — the very pair V separates. Three facts of this little dataset are, in miniature, the three load-bearing laws of the calculus built in this section: separations persist under refinement while connections do not (the monotonicity rules of §2.5); one instrument's verdicts on one pair never conflict (the clash lemma, §3); and a coarse connection coexisting with a fine separation is not a contradiction but a fact about *two* instruments (the cross-regime theorem, §3.4). Mark the modality: these laws hold in the estimator *by construction* — ε'-chains are ε-chains, and components partition — so the miniature illustrates an implementation identity; it does not confirm the laws. The record Δ_0 is a measurement context in the sense of §2.6; the fork verdict it licenses is Definition 3.4; the full protocol — calibrated boundaries, $N = 300$, a persistence criterion — is §6.

2.1 The semantic signature and semantic types

Definition 2.2 (Semantic signature)**.** A *semantic signature* Σ comprises (i) a collection of *base semantic types* $\mathfrak{B}$, and (ii) typed *term constants*: names for points, and names for instrument-observed continuations and fillers of declared horns.

The term constants are not decoration: the closed-judgment metatheorems of §5 concern closed terms, and over a constant-free signature the directed fragment's only closed terms are degeneracies — the calculus would issue no fork verdict on any non-degenerate horn. The record presupposes a naming discipline supplied by the instrument: *the instrument names what it separates* — in Example 2.1, the names $\varphi_1, \ldots, \varphi_4$.

Definition 2.3 (Semantic types)**.** The *semantic types* over Σ form the grammar

$$S \ ::= \ \mathfrak{B} \ \mid \ \mathsf{hom}_S(a,b) \ \mid \ \mathsf{Tri}_S(f,g;h) \ \mid \ \textstyle\sum_{x:S} S'(x) \quad (S'(x) \text{ semantic}).$$

Semantic types are a distinguished subclass of MLTT types: all MLTT structural and eliminator rules apply to them (in particular Σ-projections and Σ-elimination); only hom and Tri are eliminator-free. Π-types, identity types, $\mathbf{0}$, $\mathbf{1}$, $+$ and $\|\cdot\|$ are *not* semantic. The relation formers of §2.5 apply only to semantic types, over arbitrary (possibly open) terms.

Why this perimeter: semantic types are the types an instrument can point at — discourse states, continuations, warrants, and their pairings. The needed distinction is between a function as an *object of* measurement and a function as an *instrument of* measurement: instruments point *through* functions constantly (every readout ρ of §3.5 is one), but they do not point *at* them — no record entry separates two functions qua functions, and propositions likewise present nothing for a resolution to resolve. The relation formers are confined to the grammar for that reason; readouts live in the signature, on the instrument side of the line.

In the LLM. A base semantic type is a space of *contextual states*: the condition a model is in after reading a stretch of text — operationally, the state that determines its next-token distribution, observed here only through embeddings of what it goes on to generate. A term constant names one such state (a particular prompt, read); the signature's naming discipline is the laboratory practice of keeping the prompts that produced the data. Nothing in this section is special to machines: a reader part-way through a sentence is in a contextual state too. The grammar is machine-*measurable*, not machine-*only*. “Eliminator-free” means: no induction principle — the theory can *name* continuations (via the signature’s constants) but never case-analyse them, a poverty that Remark 3.9 will show is exactly what keeps canonicity honest.

Convention (substitution stability). All new formers and constants of this section commute with substitution. The path-action $e_* f$ of an ambient path on a continuation is *derived* (ordinary transport in $x \mapsto \mathsf{hom}_A(x, b)$), not primitive: the directed fragment’s primitives are exactly the five rules displayed below.

2.2 Directed hom-types

For A semantic and $a, b : A$, the type of *continuations* from a to b. Continuations are directed: no symmetry is posited — a stretch of discourse cannot be un-said — and no composition is assumed: composition is exactly what is at stake.

HOM-FORM

$$\frac{\Gamma \vdash_\Delta A\ \mathsf{type}_{\mathsf{sem}} \qquad \Gamma \vdash_\Delta a : A \qquad \Gamma \vdash_\Delta b : A}{\Gamma \vdash_\Delta \mathsf{hom}_A(a, b)\ \mathsf{type}_{\mathsf{sem}}}$$

HOM-DEGEN

$$\frac{\Gamma \vdash_\Delta a : A}{\Gamma \vdash_\Delta \mathsf{id}_a : \mathsf{hom}_A(a, a)}$$

In the running example an inhabitant of $\mathsf{hom}_A(a, b)$ is a realized traversal; id_a is the null continuation — remaining at a, adding nothing.

In the LLM. A continuation is a stretch of autoregressive generation: token after token, each conditioned — through attention — on everything already in the context, carrying the state from a to b. The two austerities of the rules are facts about that process. Directedness is autoregression’s arrow: generation extends rightward and nothing un-says a token. And the absence of an assumed composition is the absence of any architectural guarantee that two locally fluent stretches concatenate into one fluent whole — whether they do is exactly what the next subsection makes into an object.

2.3 Triangle types and filler spaces

A *triangle* is a warrant that a candidate $h : \mathsf{hom}_A(a, c)$ genuinely composes f and g: that reading the passage whole as h is compatible with having arrived by f and departed by g. For a language model the warrant is realizability: the model, holding both flanking commitments, can enact h as a fluent traversal.

$$\frac{\Gamma \vdash_\Delta f : \mathsf{hom}_A(a,b) \qquad \Gamma \vdash_\Delta g : \mathsf{hom}_A(b,c) \qquad \Gamma \vdash_\Delta h : \mathsf{hom}_A(a,c)}{\Gamma \vdash_\Delta \mathsf{Tri}_A(f,g;h)\ \mathsf{type}_{\mathsf{sem}}}\ \textsc{tri-form}$$

$$\frac{\Gamma \vdash_\Delta f : \mathsf{hom}_A(a,b)}{\Gamma \vdash_\Delta \lambda^{\mathsf{l}}_f : \mathsf{Tri}_A(\mathsf{id}_a, f; f)}\ \textsc{tri-degen-l} \qquad \frac{\Gamma \vdash_\Delta f : \mathsf{hom}_A(a,b)}{\Gamma \vdash_\Delta \lambda^{\mathsf{r}}_f : \mathsf{Tri}_A(f, \mathsf{id}_b; f)}\ \textsc{tri-degen-r}$$

(In TRI-FORM and throughout, the presupposition A $\mathsf{type}_{\mathsf{sem}}$ is carried by the premises' typing. The degenerate triangles are the trivial warrants: prefixing or suffixing the null continuation changes no reading, and the theory says so.)

Definition 2.4 (Horn; filler space)**.** An *inner* 2*-horn* in a semantic type S is a tuple $\sigma := (a, b, c : S,\ f : \mathsf{hom}_S(a,b),\ g : \mathsf{hom}_S(b,c))$; its *boundary vertices* are a, b, c ("vertex" is horn-relative: for S itself a hom-type, the horn's vertices are ambient edges). Its *filler space* is

$$\mathsf{Fill}_S(\sigma) \ :\equiv \sum_{h\,:\,\mathsf{hom}_S(a,c)} \mathsf{Tri}_S(f,g;h),$$

semantic by the grammar; we write $\mathsf{Fill}(\sigma)$ when S is clear. Horns themselves form a semantic type,

$$\mathsf{Horn}(S) \ :\equiv \sum_{a,b,c\,:\,S} \ \sum_{f\,:\,\mathsf{hom}_S(a,b)} \mathsf{hom}_S(b,c),$$

over which $\mathsf{Fill}_S(-)$ is a family via the Σ-projections (available: semantic types inherit the MLTT eliminators, Definition 2.3). We frequently abbreviate a horn by its edges, $\sigma = (f, g)$, the vertices being carried by the hom-types.

A filler is a *pair*: a composite together with its compositional warrant. Note what is not assumed: neither inhabitation (no Kan condition) nor uniqueness of composition. Operationally, $\mathsf{Fill}_S(\sigma)$ is the space of *readings* — everything the system can warrantedly do when handed the boundary; sampling a model's continuations of a prompt is sampling (a readout of) this space, a remark §6 makes precise.

In the LLM. $\mathsf{Fill}_S(\sigma)$ is, concretely, the support of the model's bridge distribution when both flanking commitments are held: everything it can generate between this prefix and this suffix. Resampling at temperature 1.0 draws from it; each draw is a candidate composite, and the model's having fluently produced it under both commitments is the operational shadow of the triangle — the warrant is never observed directly, only enacted. When the suffix is left free the sampled object is smaller — a cone, not a horn — a distinction §6 polices explicitly.

Remark 2.5 (Why the warrant is binary)**.** $\mathsf{Tri}_A(f, g; h)$ is inhabited or not, while fluent enactment is plainly graded: a bridge can be barely coherent or perfect. The gradation is not lost; it is *relocated*. The warrant is a realizability condition — can the system traverse the passage this way at all — and quality variation among warranted fillers is exactly what the regime structure on $\mathsf{Fill}(\sigma)$ measures: a barely-coherent bridge and a perfect one are both fillers, and any instrument worth its resolution separates them. Grading the warrant itself would duplicate, inside Tri, the work the relations already do outside it.

Remark 2.6 (Higher horns)**.** We develop the theory for inner 2-horns, where the semantic phenomena of interest already occur. The generalization to Λ^n_i is mechanical but notationally heavy; nothing below depends on it.

2.4 Regimes

Definition 2.7 (Regimes)**.** TTS is parameterized by a poset $(\mathcal{V}, \leq)$ of *regimes*, treated as meta-level indices (in the manner of universe levels), not internal types. **Convention:** $V' \leq V$ **means** V' **is *finer*** — it separates more and identifies less. In the intended models a regime is an *instrument*: a metric with a resolution, $V = (d, \varepsilon)$, and finer means smaller ε. We write $V < W$ for the strict order. The three instruments of Example 2.1 are three regimes: cosine distance read at 0.40, 0.60 and 0.70 respectively, ordered $V' \leq V \leq W$ — one metric, three resolutions, three verdict-sets.

A regime is the formal home of an old hermeneutic commonplace: there is no verdict of sameness or difference *simpliciter*, only sameness or difference *as measured by an instrument at a resolution.* The calculus never mentions numbers; it tracks only the order structure — which instrument refines which.

Definition 2.8 (Instrument, consolidated)**.** The word *instrument* is used at three levels, fixed here once. (i) *Syntactically*, an instrument is a regime: an element of the parameter poset $(\mathcal{V}, \leq)$, a meta-level index with no internal structure (Definition 2.7). This is the level at which every theorem is proved. (ii) *Semantically*, an instrument is what a model assigns to a regime: per denotation, a pair $(\simeq_V, \#_V)$ satisfying the rules of §2.5 (Definition 5.2, (M1)). A declared order pair $V' \leq V$ obligates the model to validate the monotonicity rules; the declared order is thereby *contained in*, but not identified with, extensional refinement of verdicts. (iii) *Experimentally*, the apparatus is a four-stage stack, of which the regime formalizes exactly one stage: the **generator** (which completions become observable; outside the formalism) feeds the **readout** ρ (a signature function constant, Definition 3.11; pullback indices ρ^*W abbreviate and are not poset elements), which feeds the **comparator** (d, ε) — *the regime* — whose sample verdicts a **statistical criterion** (persistence thresholds, mixture selection; outside the calculus) converts into distribution claims. §6's two-instruments finding demonstrates that stages of the stack vary independently; Theorem 3.10 governs the third stage only.

Scope of the order. The gloss "finer means smaller ε" defines $\leq$ only between regimes sharing a metric and readout: same-(ρ, d) regimes form a *chain*, and on chains the declared order coincides with extensional refinement in resolution models (ε'-chains are ε-chains; Example 5.4). Between regimes of different metrics, readouts, or modalities this gloss defines nothing: such order pairs may be *declared*, each declaration carrying the model-level proof obligation of the monotonicity rules; absent a declaration, regimes are incomparable. The embedding regimes of §6 instantiate a single chain; the incomparable configurations of §3.4 and §4.4 are exercised schematically there, and once live, by a symbolic second-modality instrument constructed in §6.

Remark 2.9 (Preorder suffices)**.** Antisymmetry of $\leq$ is nowhere used: $(\mathcal{V}, \leq)$ need only be a preorder, and verdict-equivalent instruments may be identified without loss. This is consistent with the paper's truncation discipline: refinement, like the verdicts it propagates, is treated as an h-proposition; the proof-relevant refinement structure — a category of calibrations with functorial indexing — is deliberately truncated.

Remark 2.10 (Honesty of the order)**.** Theorem 5.7 certifies records; nothing in the theory certifies the declared order, and the asymmetry is principled. An order pair $V' \leq V$ asserts a universal containment of verdicts, and pairwise records certify no universal (the poverty of Remark 3.9): order pairs are a priori commitments, justified at the level of instrument construction — threshold arithmetic on a chain; a factoring argument across readouts — never verified by measurement. They are, however, *refuted* by it: a single measured pair $u \simeq_W v$, $u \#_V v$ refutes $W \leq V$ (Theorem 3.10(a), read as a refutation engine). Example 2.1 contains a live instance: the 0.70-connection of φ_1, φ_3 against their 0.60-separation refutes "the 0.70 instrument refines the 0.60 one," agreeing with the threshold arithmetic. Consequently every consistency claim of the form Theorem 3.10(b) is *relative to the completeness of the declared order*: an omitted true refinement converts a contradiction into an apparently licensed disagreement, and extending the order can only add refutations, never remove them. The order is the experiment's theory-component: declared before the run, falsifiable by it, confirmable only from outside it.

In the LLM. A regime, concretely, is an embedding model together with a threshold: `text-embedding-3-small` read at cosine resolution 0.60 is one instrument; the same encoder at 0.40 is a finer one; a different encoder is a different instrument altogether, comparable to the first only if someone argues the comparison. Sweeping ε is literally turning a resolution knob, and the model's own internal machinery — attention over the context, the geometry of its hidden states — is yet another instrument, one this paper never reads directly: every measurement here watches generated text from outside, through declared encoders. That modesty is deliberate, and the calculus is built for it.

2.5 Indiscernibility and apartness

Two regime-indexed relation formers on each semantic type, both valued in h-propositions (their isProp structure is built into the formers; where a sum appears in a conclusion it is truncated).[1] In both formation rules below the type S is presupposed semantic ($S\ \mathsf{type}_{\mathsf{sem}}$), per Definition 2.3.

V-indiscernibility. $u \simeq_V v$: *the instrument V cannot tell u from v apart* — there is a chain of sub-resolution deformations from one to the other. (The entries γ_1, γ_2 of Δ_0 in Example 2.1.)

$$\frac{\Gamma \vdash_\Delta u : S \qquad \Gamma \vdash_\Delta v : S}{\Gamma \vdash_\Delta u \simeq_V v\ \mathsf{type}}\ \simeq\text{-FORM} \qquad \frac{\Gamma \vdash_\Delta u : S}{\Gamma \vdash_\Delta \mathsf{r}_V(u) : u \simeq_V u}\ \simeq\text{-REFL} \qquad \frac{\Gamma \vdash_\Delta p : u \simeq_V v}{\Gamma \vdash_\Delta p^\dagger : v \simeq_V u}\ \simeq\text{-SYM}$$

$$\frac{\Gamma \vdash_\Delta p : u \simeq_V v \qquad \Gamma \vdash_\Delta q : v \simeq_V w}{\Gamma \vdash_\Delta p \cdot q : u \simeq_V w}\ \simeq\text{-TRANS} \qquad \frac{\Gamma \vdash_\Delta p : u \simeq_{V'} v \qquad V' \leq V}{\Gamma \vdash_\Delta \mathsf{up}(p) : u \simeq_V v}\ \simeq\text{-MONO}$$

[1] For readers outside type theory: an *h-proposition* is a type with at most one inhabitant up to identity — a yes/no matter; the *truncation* $\|T\|$ records *that* T is inhabited while forgetting *which* inhabitant — mere existence. The distinction earns its keep at Definition 3.4.

$\simeq$-MONO reads: what a finer instrument cannot distinguish, a coarser one cannot either.

Lemma 2.11 (id-to-regime)**.** *For S semantic and every regime V there is a function* $\mathsf{itr}_V : (u =_S v) \to (u \simeq_V v)$.

Proof. Path induction with motive $(x, y, p) \mapsto x \simeq_V y$ (an h-prop family over open terms, licensed by $\simeq$-FORM) and base r_V. $\square$

Mathematical identity is thus indiscernible at every regime. The converse is *not* claimed — indiscernibility at every declared regime does not yield a path — and this one-directionality is essential to what follows.

V-apartness. $u \#_V v$: *the instrument V has the resources to separate u from v* — a positive structure in the constructive-apartness tradition of Brouwer, Heyting and Bishop [23, 12, 47, 50].

$$\#\text{-FORM}\quad \frac{\Gamma \vdash_\Delta u : S \qquad \Gamma \vdash_\Delta v : S}{\Gamma \vdash_\Delta u \#_V v\ \mathsf{type}} \qquad \#\text{-IRREFL}\quad \frac{\Gamma \vdash_\Delta \delta : u \#_V u}{\Gamma \vdash_\Delta \mathsf{absurd}(\delta) : \mathbf{0}} \qquad \#\text{-SYM}\quad \frac{\Gamma \vdash_\Delta \delta : u \#_V v}{\Gamma \vdash_\Delta \delta^\dagger : v \#_V u}$$

$$\#\text{-COTRANS}\quad \frac{\Gamma \vdash_\Delta \delta : u \#_V v \qquad \Gamma \vdash_\Delta w : S}{\Gamma \vdash_\Delta \mathsf{split}(\delta, w) : \big\|(u \#_V w) + (w \#_V v)\big\|} \qquad \#\text{-MONO}\quad \frac{\Gamma \vdash_\Delta \delta : u \#_V v \qquad V' \leq V}{\Gamma \vdash_\Delta \mathsf{down}(\delta) : u \#_{V'} v}$$

$$\#\text{-EXT}\quad \frac{\Gamma \vdash_\Delta p : u \simeq_V u' \qquad \Gamma \vdash_\Delta \delta : u \#_V v}{\Gamma \vdash_\Delta \mathsf{ext}(p, \delta) : u' \#_V v}$$

Glosses. #-MONO: a separation achieved by a coarse instrument persists under every refinement — one cannot lose a measured difference by looking harder. (In Example 2.1: the δ-entries recorded at 0.60 all hold at 0.40; the γ-entries do not — which is why there is a $\simeq$-MONO going coarse-ward and no $\simeq$-rule going fine-ward.) #-EXT: apartness is extensional over indiscernibility *at the same regime* — what an instrument cannot distinguish, it cannot separate differently. #-COTRANS: if two readings are separated, any third reading is separated from at least one of them; it is cotransitivity that makes apartness classes behave like components of a space. (Cotransitivity is retained as constitutive of constructive apartness and holds in all intended models; no result in this paper depends on it.)

Remark 2.12 (Exogeneity, made architectural)**.** There is *no premise-free introduction rule* for apartness at any regime: every # in a conclusion traces to a # in a premise (#-COTRANS redistributes recorded apartness to new comparison points; it cannot create apartness ex nihilo — the erasure argument of Appendix A, which also yields Corollary 5.9, makes this precise). Closed separation witnesses enter the theory in exactly one way: through the measurement context of §2.6, where they are constants interpreted by an actual instrument record — such as Δ_0 of Example 2.1, whose δ_1 is not a stipulation but the number .691 read against the threshold 0.60. Distinctness claims are obligations to exhibit data.

2.6 Measurement contexts

Definition 2.13 (Measurement context)**.** A *measurement context* is a finite list

$$\Delta \ ::= \ \emptyset \ \mid \ \Delta, \delta : u \mathrel{\#_V} v \ \mid \ \Delta, \gamma : a \simeq_V b$$

whose terms are *closed* terms of semantic types and whose indices are regimes (or pullback indices, §3.5). Δ is well-formed (Δ ok) when each entry's terms are well-typed closed terms of a single semantic type. The judgment form of TTS is $\Gamma \vdash_\Delta \mathcal{J}$, and the single rule by which the record enters derivations is

$$\textsc{Δ-ax}\ \frac{(\delta : u \mathrel{\#_V} v) \in \Delta}{\Gamma \vdash_\Delta \delta : u \mathrel{\#_V} v}$$

(presupposing Γ ok, Δ ok; dually for γ-entries). Δ-entries are *constants*: they cannot be λ-abstracted or substituted for. Weakening of Δ is admissible; there is no Δ-strengthening.

The intended reading: Δ is the *record* of an instrument run. The reader has already held one in full: Δ_0 of Example 2.1 — two connections, four separations, each entry a number read against a declared threshold. (Δ-ax) is the rule that lets a derivation cite δ_1; nothing else about the run enters the logic.

In the LLM. The record is the output of a clustering run, kept. Sample completions, embed them, sweep the threshold, write down which pairs each resolution connects and which it separates: that verdict list *is* Δ. (Δ-ax) then says the only empirical citations a derivation may make are lines of that list — the analysis may quote its own measurements and nothing else. Everything a machine-learning practitioner would recognize as the pipeline's statistics (persistence criteria, model selection) lives *outside* the record, in the fourth stage of Definition 2.8; the calculus consumes verdicts, not p-values. $\mathrm{TTS}(\Sigma, \Delta)$ is then a **logic of consequence from measurement data**: it derives what the record forces — which forks stand, which identities are refuted, which cross-instrument configurations cohere — and, by the metatheorems of §5, it can do nothing else: no fork is derivable from the empty record (Corollary 5.9), and every derivable fork holds in every model realizing the record (Theorem 5.11).

Remark 2.14 (Variables hypothesize, constants record)**.** A separation hypothesized for the sake of argument is a Γ-variable: it is Π-bindable, and it shows up in the *type* of the resulting term — a conditional claim. A separation that has been measured is an Δ-constant: it underwrites *categorical* claims, and the distinction is real exactly where the metatheorems of §5 live — closed theorems over a fixed record. We also state it honestly: within one derivation a constant behaves like a hypothesis, just as a theory's axioms behave like premises of its proofs. The difference is that axioms cannot be discharged — no Π binds an Δ-entry, so no theorem internalizes a counterfactual record. What certifies that the record itself is honest is meta-theoretic and stated as such (Theorem 5.7).

3 Canonicity and the fork

Section 2 supplied raw material and asserted nothing: hom-types need not be inhabited, triangles need not exist, and apartness has no premise-free introduction. The system so far cannot even say that composition succeeds. The poverty is deliberate. The content of the logic lives in two derived classifications of a filler space — the *grades* — and in the theorems governing their interaction across regimes.

3.1 The clash lemma

Lemma 3.1 (Per-regime clash)**.** *For S semantic, $u, v : S$ and any regime V:*

$$\mathsf{clash}_V : (u \simeq_V v) \to (u \#_V v) \to \mathbf{0}.$$

Proof. Given $p : u \simeq_V v$ and $\delta : u \#_V v$, apply $\#$-EXT instantiated at $(u, u', v) := (u, v, v)$ — its premises are exactly p and δ — to obtain $v \#_V v$; conclude by $\#$-IRREFL. The left-handed $\#$-EXT suffices directly; no symmetry shuffle is needed. $\square$

In words: no single instrument may hold, at once, a deformation of one reading into another and a separation between them. Across *different* instruments no such law holds — that is Theorem 3.10, and it is a feature.

Corollary 3.2 (id-clash)**.** $(u =_S v) \to (u \#_V v) \to \mathbf{0}$*, by Lemma 2.11 and* clash_V*. (Under the intended discrete interpretation of identity, §5: a recorded separation refutes literal pointwise identity.)*

3.2 The grades, exclusion, and persistence

Definition 3.3 (Canonicity at a regime)**.** $\mathsf{Canon}_V(\sigma) \;:\equiv\; \big\|\mathsf{Fill}(\sigma)\big\| \times \prod_{\varphi,\psi\,:\,\mathsf{Fill}(\sigma)} \varphi \simeq_V \psi$ — an h-proposition under function extensionality (a bookkeeping remark only; no derivation below uses it): the filler space is inhabited and the instrument V cannot tell any two of its inhabitants apart. *The geometry decides; no semantic decision occurs.* Operationally: every continuation the system produces at σ is, at this resolution, a deformation of every other — three hundred draws, one cluster, no news.

Definition 3.4 (Fork at a regime)**.** $\mathsf{Fork}_V(\sigma) \;:\equiv\; \sum_{\varphi,\psi\,:\,\mathsf{Fill}(\sigma)} \varphi \#_V \psi$. A fork witness is a triple — two complete fillers, each a composite with its warrant, and a recorded separation between them. Fork_V is *deliberately not truncated*: the verdict names its witnesses; the truncation $\|\mathsf{Fork}_V(\sigma)\|$ is the judgment of record. (Exclusion and persistence below use plain Σ-elimination and are insensitive to the choice.) In the miniature: were φ_1, φ_3 the fillers of a fixed horn, $(\varphi_1, \varphi_3, \delta_1)$ would be the paradigm inhabitant — two readings of one passage, measurably apart; the precise gap between sampled completions and horn-fillers is the scoping question §6 addresses.

Theorem 3.5 (Exclusion at a regime)**.** $\mathsf{Canon}_V(\sigma) \to \mathsf{Fork}_V(\sigma) \to \mathbf{0}$.

Proof. From $(i, c) : \mathsf{Canon}_V(\sigma)$ and $(\varphi, \psi, \delta) : \mathsf{Fork}_V(\sigma)$: $c\,\varphi\,\psi : \varphi \simeq_V \psi$; apply clash_V with δ. (The truncated component i is discarded; no truncation-elimination into data occurs.) ☐

Exclusion remains a *theorem*, not an axiom — the architecture polices itself — and it is now indexed: it forbids a single instrument from issuing both verdicts on one horn. It says nothing about two instruments, by design.

Theorem 3.6 (Fork persistence)**.** *For* $V' \le V$*:* $\mathsf{Fork}_V(\sigma) \to \mathsf{Fork}_{V'}(\sigma)$*, by* #-MONO *on the third component.* Persistence is monotonicity*: a fork recorded by a coarse instrument stands under every refinement.*

Proposition 3.7 (Canonicity is not persistent)**.** *There is a model of TTS (Definition 5.2) and a horn σ with* $\mathsf{Canon}_\varepsilon(\sigma)$ *and* $\mathsf{Fork}_{\varepsilon'}(\sigma)$ *for resolutions* $\varepsilon' < \varepsilon$.

Proof. Take a carrier with one composite $h : \mathsf{hom}(a, c)$ and exactly two warrants: $\mathsf{Fill}(\sigma) = \{(h, t_1), (h, t_2)\}$, with $d(t_1, t_2) = r \in (\varepsilon', \varepsilon]$, every other point of the ambient triangle level at distance $> \varepsilon$ from both, f, g non-degenerate; complete to a model of Definition 5.2 (discrete identity; intrinsic component structures; the shared composite makes the face conditions trivial). At ε: the single $\leq \varepsilon$ step joins the two fillers, so $\mathsf{Canon}_\varepsilon$ holds. At ε': the two fillers lie in different ε'-components, so $\mathsf{Fork}_{\varepsilon'}$ holds (and $\mathsf{Canon}_{\varepsilon'}$ fails, by Exclusion). □

Note what the witnessing configuration is: *two separated warrants over one composite* — the same arrival, two measurably distinct modes of arriving. The asymmetry of the two grades is thus structural: forks are robust news; canonicity is always provisional, an artifact of the current resolution that refinement may revoke.

In the LLM. Canonicity at V is what resampling shows when nothing is at stake: three hundred draws, one cluster at the declared resolution — the next-token distribution funnels, and no reading-level decision remains to be made. A fork is the same protocol returning two stable clusters: the phenomenon Bigelow et al. observe as forking tokens, lifted from single token positions to whole readings. The asymmetry of the grades is the working epistemology of anyone who has clustered model outputs: finding two separated modes is a result you can publish; finding one mode is a result the next, finer sweep may take away.

3.3 The absolute limit, and what records can never establish

Lemma 3.8 (Absolute limit)**.** $\mathsf{isContr}\big(\mathsf{Fill}(\sigma)\big) \to \mathsf{Canon}_V(\sigma)$ *for* every *regime V: the center inhabits $\|\mathsf{Fill}\|$; contractibility gives pairwise identity; Lemma 2.11 converts each path to V-indiscernibility.*

In the intended discrete-identity models (§5), $\mathsf{isContr}(\mathsf{Fill})$ is the *strict* Segal condition — a unique filler on the nose, the nerve-of-a-category case — so the comparison with synthetic ∞-category theory [41] is an analogy of form together with this strict limit, and we claim nothing more. The graded judgments Canon_V are the working notions; the absolute notion is their common upper bound.

Remark 3.9 (Canonicity is never derivable from a finite record)**.** No finite measurement context *realized by a resolution model* — in particular, no honest record (Theorem 5.7) — yields a closed term of $\mathsf{Canon}_V(\sigma)$. The proof is semantic and short: extend the realizing model by one fresh point in the σ-fibre of the triangle level, placed beyond threshold from every existing point. The extension is again a resolution model; every recorded verdict mentions only named points and survives (no chain is created or broken by a distant point); and the extended filler space now contains a V-apart pair, so $[\![\mathsf{Canon}_V(\sigma)]\!]$ is empty there. Soundness (Theorem 5.6) closes the argument. (Two scope notes. An *explosive* record — $\delta : u \#_V u$ — yields everything by $\#$-IRREFL and **0**-elimination, which is why realizability, not mere consistency, is the right hypothesis; and for consistent but unrealizable records the question is open — necessarily so, given the completeness conjecture of §8.) Two structural reasons underlie the poverty: (i) Δ-entries are pairwise verdicts on *named* closed terms, and exhaustiveness of a record is not an Δ-expressible condition; (ii) hom and Tri have no eliminators, so $\mathsf{Fill}(\sigma)$ admits no internal enumeration, and the Π over all fillers is unreachable from closed instances. (A hypothesized enumeration $e : \mathsf{Fill}(\sigma) \simeq \mathsf{Fin}\, n$ is a Γ-variable route — the *variables hypothesize* side of the slogan.) The calculus issues fork verdicts from records; canonicity is the model-side limit notion that experiments *estimate* by failure-to-refute. **Forks are verifiable; canonicity is falsifiable-only.** This asymmetry is not a defect: it is the falsifiability structure the empirical bridge of §6 requires.

The same poverty has a second face worth stating: the theory can say "these two senses are apart" but never "there are exactly three senses at this horn." We hold this to be principled, not accidental. A cardinality claim is a *different kind of verdict* — it asserts exhaustiveness, and exhaustiveness is precisely what no pairwise instrument delivers; an instrument that could count would be a new primitive (an enumeration channel), not a refinement of the ones we have. The calculus accordingly treats counting the way it treats canonicity: hypothesize it if you wish (an enumeration $e : \mathsf{Fill}(\sigma) \simeq \mathsf{Fin}\, n$ is a perfectly good Γ-variable, and conditional theorems follow), but no record formed from pairwise verdicts will ever discharge it. A richer theory with enumeration-grade records is conceivable; it would be answerable to a different experimental practice.

3.4 Cross-regime consistency

The question Frege's puzzle will pose in §4: can a connection discovered by one instrument coexist with a separation recorded by another? The answer is a theorem, and it depends on nothing but the order relation between the instruments.

Theorem 3.10 (Disagreement forces non-finer connection)**.**

(a) (Universal.) For all $u, v : S$: if $W \leq V$, then $(u \simeq_W v) \to (u \#_V v) \to$ $\mathbf{0}$.

(b) (Schematic consistency.) For u, v fresh *constants of a base semantic type — fresh: distinct symbols occurring in no other Σ-declaration and no other record entry — and Δ exactly $\{\gamma : u \simeq_W v,\ \delta : u \#_V v\}$:* $\mathrm{TTS}(\Delta)$ *is consistent if and only if $W \not\leq V$ — that is, $V < W$ strictly, or V, W incomparable.*

Proof. (a) $\simeq$-MONO at $V' := W$ (side condition exactly $W \leq V$) lifts $u \simeq_W v$ to $u \simeq_V v$; apply clash_V. (b) Refutation when $W \leq V$ is (a) via (Δ-ax). Consistency witnesses, with details in Appendix A: for $V < W$, the resolution model on two points at distance r with $\varepsilon_V < r \leq \varepsilon_W$; for incomparable V, W, the two-point, two-regime record model ($\#_V$ the pair, $\simeq_V$ diagonal; $\simeq_W$ total, $\#_W$ empty). In each case the recipe $\simeq_U :=$ total if $W \leq U$ else diagonal, $\#_U :=$ pair if $U \leq V$ else empty *defines* the model at every regime of the poset (the two-point descriptions are its motivating instances): monotonicity follows from transitivity of $\leq$, and the clash-prone combination would require $W \leq U \leq V$, excluded by hypothesis. Freshness matters: instantiating $u := v$, or adding conflicting entries, refutes by $\#$-IRREFL or id-clash regardless of the regimes. $\square$

At $V = W$ refutation wins — it is clash_V itself. The reading we will lean on: *a configuration in which one instrument connects what another separates is consistent exactly when the connecting instrument is neither the separating one nor a refinement of it* — strictly coarser, or a different modality altogether. Looking harder through the same instrument can never identify what it separated. Example 2.1 exhibits the consistent configuration live: $\delta_1 : \varphi_1 \#_V \varphi_3$ at resolution 0.60 and $\varphi_1 \simeq_W \varphi_3$ at resolution 0.70 stand together in one dataset, because $V < W$ — the connecting instrument is strictly the coarser one.

In the LLM. The theorem is about an everyday situation in embedding-land: two analyses disagree about whether two completions "say the same thing." One encoder at one threshold merges them; another separates them. The theorem sorts these disagreements into the impossible kind (the merging instrument refines the separating one — then somebody's clustering is inconsistent, and an honest pipeline never produces it) and the licensed kind (everything else), and §4 will spend the license on Frege. Nothing here adjudicates which encoder is *right*; the calculus is the bookkeeping of instrument disagreement, not a theory of the one true embedding.

3.5 Readouts and pullback regime structures

Instruments compose. One rarely observes fillers directly; one observes them *through* a readout — an embedding, a projection, a summary — and measures in the image. The calculus internalizes this.

Definition 3.11 (Readout; pullback)**.** A *readout* is a function term $\rho : S \to R$ between semantic types; instrument readouts are function constants of the signature Σ. Given a regime W (acting on R), the *pullback notation*

$$u \simeq_{\rho^* W} v \;:\equiv\; \rho u \simeq_W \rho v \qquad u \#_{\rho^* W} v \;:\equiv\; \rho u \#_W \rho v$$

is licensed by unfolding — $\simeq/\#$-FORM at W on $\rho u, \rho v$. A subscript $\rho^* W$, in former or Δ-entry position, is an abbreviation, not a new element of the regime poset; no other non-poset indices are admitted, and models need no clause for them.

Proposition 3.12 (Pullback laws)**.** $(\simeq_{\rho^* W}, \#_{\rho^* W})$ *satisfies every single-regime rule of §2.5 at S (each inherited in one line; cotransitivity by splitting W's cotransitivity at ρw). Derived order implications: for $W' \leq W$, $u \simeq_{\rho^* W'} v \to u \simeq_{\rho^* W} v$ and $u \#_{\rho^* W} v \to u \#_{\rho^* W'} v$, each by W's own monotonicity rule.*

The consequence for measurement is structural: *a separation observed in the image is a separation of the originals at the pullback index* — the composite instrument is itself an instrument. Its blind spots are equally structural: if ρ factors through the first projection of Fill, the pullback index cannot separate fillers differing only in their warrant. Both facts organize the empirical section (§6).

In the LLM. The readout is the embedding model. One never compares generated texts directly; one compares their images under an encoder — here `text-embedding-3-small`, often restricted to a sense-committing span — and the pullback construction is the formal license for doing so: a separation observed among the images is a separation of the originals *at the composite instrument.* The blind spot is equally concrete: encoders see text, not the generation process that produced it, so two bridges identical as strings but different in how the model arrived at them are indistinguishable at every pullback index. Internal-state readouts — probing hidden states rather than output — would be different instruments with different blind spots; this paper's claims are all output-side.

4 Sense and reference

4.1 The identification

Definition 4.1 (Reference; sense)**.** Let $\sigma = (f, g)$ be a horn in a semantic type A.

- The *reference* of σ is its boundary data: $\mathsf{Ref}(\sigma) :\equiv (a, b, c, f, g)$ — what is already fixed and must be respected by any completion.
- A *sense* at σ is an inhabitant $\varphi : \mathsf{Fill}_A(\sigma)$ — a way of completing what is given. We write $\mathsf{Sense}_A(\sigma) :\equiv \mathsf{Fill}_A(\sigma)$, the same type under its semantic description.
- Senses φ, ψ are *hyperintensionally distinct at the regime* V when $\varphi \#_V \psi$ is inhabited. There is no sense-distinctness *simpliciter*: only mathematical identity (the absolute limit of Lemma 3.8) and graded distinctness relative to an instrument. This relativity is not a concession; it is the finding.

The Fregean doctrine falls out as structure — and the first clause is a construction, not a typing convention. Recall $\mathsf{Horn}(A)$, the Σ-type of inner 2-horns displayed in Definition 2.4, and form the *total space of senses*

$$\mathsf{Sense}(A) :\equiv \sum_{\sigma : \mathsf{Horn}(A)} \mathsf{Fill}_A(\sigma), \qquad \pi : \mathsf{Sense}(A) \to \mathsf{Horn}(A).$$

A sense, taken absolutely, is a point of the total space; it *carries* its reference as its image under π.

In the LLM. A sense is a region of completion-space a trajectory can fall into: one of the stable clusters the resampling protocol exhibits at a boundary. Hyperintensional distinctness at V is then as concrete as semantics gets — two readings whose embedded completions a declared encoder, at a declared resolution, refuses to merge. The classical phenomena recur at this level of description without any apparatus of worlds: the financial and riparian readings of the bank passage are apart at every reasonable resolution of every reasonable encoder, and that robust, public measurability is what entitles the word *sense* rather than *vibe*.

(a) *Sense determines reference, not conversely.* Determination is the projection π. The converse fails when $\mathsf{Fork}_V(\sigma)$ is inhabited: one fibre, several separated modes of completion.

(b) *No sense without a horn.* Sense is not an intrinsic halo attached to an expression; it is relative to a compositional demand. The same lexical material in a different horn has a different filler space. (This is the type-theoretic rendering of the context principle: only in the context of a compositional demand does an expression have a sense.)

(c) *Where composition is canonical, sense collapses into reference.* If $\mathsf{Canon}_V(\sigma)$ holds, all senses at σ are V-indiscernible: one way of arriving, as far as the instrument can tell. The sense/reference distinction is literally invisible in a Kan world — which explains why a logic adequate to mathematics, where canonicity is the global ambient assumption, kept failing to be a logic of natural-language sense.

4.2 The informativeness of identity, constructively

Let $\varphi, \psi : \mathsf{Sense}_A(\sigma)$ be two senses at one horn — two modes of completion of one compositional reference — and let the record stand at $\delta : \varphi \mathbin{\#_V} \psi$: some instrument V has separated the two ways of arriving. What can it mean to *discover* that they are one?

Not literal identity: by id-clash (Corollary 3.2), the recorded separation refutes $\varphi = \psi$ outright — two measurably distinct completions are not the same completion, and no discovery will make them so. What is discovered is *indiscernibility at another instrument*: a record entry $\gamma : \varphi \simeq_W \psi$ — under the new instrument W, the two modes of arriving cannot be told apart. And here the calculus delivers its sharpest philosophical result. Refutation is universal: by Theorem 3.10(a), whenever $W \leq V$ — the discovering instrument is the separating one or a refinement of it — the configuration $\{\varphi \mathbin{\#_V} \psi,\ \varphi \simeq_W \psi\}$ explodes. Consistency in the remaining cases ($W \not\leq V$: strictly coarser, or incomparable — a different modality altogether) is established by Theorem 3.10(b) for named base points, and for filler pairs in the strictly-coarser case by the model of Proposition 3.7; the incomparable filler case follows the same recipe. Looking harder through the same instrument, or any refinement of it, can never identify what it separated.

This is, we contend, the right anatomy of Frege's datum. The informativeness of an identification is the fact that it is a *new instrument's verdict*, consistent with — not a retraction of — the standing separation: the perceptual regime that distinguishes morning-star-wise from evening-star-wise arriving is not refuted by the astronomical regime that connects them; the two verdicts coexist because the instruments are incomparable. Cognitive significance is the *regime profile* of a filler pair: which instruments separate them, which connect them, and how those instruments are ordered. Nothing here invokes belief contexts, impossible worlds, or syntactic quotation: the hyperintensional grain is carried by the graded geometry of $\mathsf{Sense}_A(\sigma)$ itself.

Remark 4.2 (Scope: from fillers to names)**.** The celestial example involves more than the filler-level machinery shown so far. "Hesperus" and "Phosphorus" are distinct *names*, and Frege's case is two expressions with one referent — structurally dual to the one-boundary-many-fillers configuration of §4.3, and a case any full treatment must square with rigid designation [27]. §4.4 delivers the core: the cross-regime configuration for *name-states* — presentation-points named by the signature — is a theorem-grade instance of Theorem 3.10, with the informativeness of identity and the persistence of the modes of presentation as one-line consequences. What remains for a companion development is the semantics of the public name *as an expression* — the word-to-constant link and its behaviour under predication — together with rigidity in full.

4.3 Worked example: lexical ambiguity as a fork

Let A be a semantic type for an unfolding discourse, and consider the fragment

"*... river ... bank ...*" versus "*... deposit ... bank ...*"

(*Bank* is the classic case of *homonymy* — two unrelated lexemes sharing a form; we use it for vividness. Systematic *polysemy* — related senses of one lexeme, the copredication cases — poses further demands we do not address here; the type-theoretic treatments of record are [4, 37, 13], and relating dot-type coercion to fork structure is future work.) Fix the horn $\sigma = (f, g)$ in which f is the continuation arriving at the occurrence of *bank* and g the continuation departing from it into further discourse that does not itself disambiguate. Then:

- $\mathsf{Ref}(\sigma)$: the lexical occurrence with its flanking commitments — what any reading must respect.
- Two fillers $\varphi_{\mathrm{fin}}, \varphi_{\mathrm{riv}} : \mathsf{Fill}_A(\sigma)$: each a composite reading *with* a triangle witnessing that it actually composes the flanking continuations.

- A recorded separation $\delta : \varphi_{\text{fin}} \#_V \varphi_{\text{riv}}$, realized in the model as the metric separation of the two contextual completions at the instrument V (empirically: the two readings' embeddings fall in different components at any reasonable resolution; §6 measures exactly this).

So $\mathsf{Fork}_V(\sigma)$ is inhabited: ambiguity is a fork, and a *reading* is a choice of filler. Disambiguating discourse does not erase the fork — by Theorem 3.6 a recorded fork stands at every refinement, and nothing retracts a record — it *outruns* it: the disambiguated discourse poses a *different* horn σ', with richer boundary, whose filler space may be canonical. Ambiguity is resolved not by deleting the other sense from the language but by moving to a compositional demand that no longer admits it. This is, recognizably, the dynamic-semantics picture — meaning as context-change rather than static content [26, 22, 21, 49] — transposed to a setting where the contexts are geometric.

Remark 4.3 (Against sense-as-algorithm, in one paragraph)**.** Procedural theories individuate senses by identity of algorithms [34, 16] or programs [10]; proof-theoretic semantics individuates the sense of a derivation by its rule-structure [46, 7]. All locate sense in the *syntax of arriving*. TTS locates it in the *space of arrivings*: the individuation is geometric (apartness classes at a regime), the identity criterion is instrument-relative connection, and both are answerable to measurement rather than to a chosen formal presentation.

4.4 Names: two routes, one object

Frege's own case now. Fix a base semantic type P of the signature: the *presentation space* — object-directed contextual states (for a language model: embedded states in which the discourse is directed at a particular object). Names enter as the signature's *term constants* valued in P — the naming discipline of §2.1, doing now for presentation-points what it did for the miniature's completions: $h, p : P$, the Hesperus-wise and Phosphorus-wise presentation states. Two scoping commitments, stated plainly. First, the relation between the public *word* "Hesperus" and the constant h is presupposed, not constructed: the presentational instrument names the clusters it separates, and the lexicon is assumed to label them; the semantics of the name as an expression is the companion's subject (Remark 4.2). Second, we do not call h and p *senses*: by the context principle of §4.1(b) senses live at horns, and h, p are points. Cognitive significance for names will be the regime profile of a *point pair* — the point-level analogue of §4.2's filler-pair notion; the horn-level story (name-states as boundary vertices of predicational demands) belongs to the deferred predication dynamics.

In the LLM. The presentation space is observable: prompt a model with a name in a neutral carrier and embed what it generates, and you have sampled the name's presentational neighbourhood — the region of state-space the name habitually routes the trajectory into. For co-referential names the neighbourhoods need not coincide, and in a production model they measurably do not (the name-pilot of §8, logged in the supplement): the evening-star states and the morning-star states live apart even in a model that asserts their identity on request. The formalism's distinct constants h, p with a standing presentational separation are a description of that geometry, not a stipulation about it.

Two regimes carry the case, both ordinary poset regimes. V is a *presentational* instrument — it separates presentation-clusters, as the perceptual and temporal circumstances of observation do: evening-presentations from morning-presentations. Its verdict stands in the record: $\delta : h \#_V p$ — in any realizing model, the two name-states lie in distinct V-components (presentation-basins, when those components are canonical; basin-hood itself, per Definition 5.14, is falsifiable-only). W is a different modality — positional, calculational: the astronomical instrument — and the *discovery* is its record entry: $\gamma : h \simeq_W p$.

Corollary 4.4 (The names configuration)**.** *For h, p fresh constants of the base semantic type P ($h \neq p$ as symbols, occurring in no other Σ-declaration and no other record entry) and Δ exactly $\{\delta : h \#_V p,\ \gamma : h \simeq_W p\}$:* TTS($\Delta$) *is consistent if and only if $W \not\leq V$. This is Theorem 3.10(b) verbatim — the names configuration is its proven scope, fresh constants of a base semantic type.*

Where is the object? Designate an *objectual* regime W^* with $W \leq W^*$ — the discovering instrument, or a coarsening of it; which regime counts as objectual is an application-level designation, like the choice of a metric, and the constraint $W \leq W^*$ is itself meaningful: *objecthood is what survives coarsening from the discovery.* The *reference* of a name-state is then its $\simeq_{W^*}$-class, and co-reference follows from the discovery by $\simeq$-MONO: $h \simeq_{W^*} p$ — two routes, one object. (Reference so defined is name-level and regime-relative; $\mathsf{Ref}(\sigma)$ of Definition 4.1 is demand-level. Frege himself bifurcates reference by category — the *Bedeutung* of a sentence is not that of a name — and so do we: the two notions meet when name-states occur as boundary vertices of a horn. The relativity to W^* is not a concession but the position of Definition 4.1, item three, applied to reference.)

(A seam we mark rather than smooth: the ambiguity case of §4.3 realizes the *sense-as-filler* account directly; the present section realizes the calculus of *instrument-relative identification*. What unites them is Theorem 3.10 — one consistency law, two typing instantiations — not a single reduction of both phenomena to fillers.) The Fregean data now fall out as one-line derivations.

Proposition 4.5 (Informativeness as record-dependence). *(i)* $\vdash_\emptyset \mathsf{r}_W(h) : h \simeq_W h$ *for every regime* W*: trivial identities are derivable from the* empty *record. (ii) For distinct fresh name constants* $h, p : P$ *there is no closed term* $\vdash_\emptyset t : h \simeq_W p$*, nor of* $h =_P p$*.*

Proof. (i) is $\simeq$-REFL at the constant ($h : P$ by Definition 2.2, "names for points"). (ii) is semantic — erasure cannot show it, since $(\simeq)^\circ = \mathbf{1}$ — by the two-point model of Appendix A: interpret P as $\{x, y\}$ with $[\![h]\!] = x$, $[\![p]\!] = y$, every $\simeq_U$ the diagonal and every $\#_U$ empty, the rest of the signature as in the base-consistency lemma; the model realizes the empty record and empties both types; conclude by soundness. □

"Hesperus is Hesperus" costs nothing and says nothing: it is reflexivity, available at every instrument, from no record. "Hesperus is Phosphorus" is underivable from the empty record — *no record, no identification* — and enters only by citing an instrument: (Δ-ax) on γ. The informativeness of an identity statement is, exactly, its record-dependence.

Proposition 4.6 (Segregation and transfer). *Let* e *range over indiscernibility proofs. From the record of Corollary 4.4:*

(a) (Segregation at the presentational regime; needs only δ*.) For every* $q : P$*:* $(h \simeq_V q) \to (q \#_V p)$*, by* $\#$-EXT *at* (h, q, p)*:* $\lambda e.\, \mathsf{ext}(e, \delta)$*. Whatever the presentational instrument groups with* h*, it must separate from* p *— and derivability is monotone under record extension, so this holds whatever a* consistent *record additionally contains (for the discovery pair, consistency is exactly* $W \not\leq V$*).*

(b) (Transfer at the discovering regime.) For every $q : P$*:* $(h \simeq_W q) \to (p \simeq_W q)$*, by* $\lambda e.\, (\gamma^\dagger \cdot e)$*; and* $(h \#_W q) \to (p \#_W q)$*, by* $\lambda \delta'.\, \mathsf{ext}(\gamma, \delta')$ *— and symmetrically in each case, interchanging* γ *and* $\gamma^\dagger$*. At the instrument that made the discovery, the two names' verdicts flow freely across the identification.*

This pair is the anatomy of what discovery does and does not do. At W, the identification licenses full information flow. At V, nothing transfers — no rule carries $\simeq_W$ down to the finer V — and by (a) the presentational neighbourhoods of the two names remain provably segregated. Learning that Hesperus is Phosphorus does not, and cannot, merge the presentations: the persistence of cognitive significance after discovery is a theorem, and it needs only δ. (A record claiming both V-groupings — some closed q with $h \simeq_V q$ and $q \simeq_V p$ on record — refutes itself at V by (a) and the clash lemma; by the contrapositive of Theorem 5.7, no honest instrument run produces it.)

Remark 4.7 (Rigidity, conditionally)**.** Within any model, a signature constant's denotation is environment-independent, while an open descriptive term's varies with its environment — a proof-theoretic shadow of the rigid/non-rigid contrast [27]. Rigidity proper — stability across counterfactual evaluation — requires modal or diachronic machinery TTS deliberately lacks (record-counterfactuals are not internalizable, Remark 5.12), and is the companion's subject.

5 Semantics and metatheory

Sections 2–4 are a calculus and its philosophical reading; we owe the mathematics. This section defines the models, states soundness, and proves the metatheorems that make the measurement-context architecture more than rhetoric: the calculus is conservative over its base, creates no forks of its own, and derives nothing the record does not force.

Read each metatheorem as a clause of a warranty on the measurement architecture. *Soundness* (Theorem 5.6): verdicts derived in the calculus hold in every world consistent with the record — the logic cannot outrun its data. *Record consistency* (Theorem 5.7): an honest sweep cannot contradict itself — explosions require stipulation. *Conservativity* (Theorem 5.8): the verdict layer proves nothing new beneath itself — instrument talk is an extension, not a contamination. *No fork ex nihilo* (Corollary 5.9): no measurement, no fork — the calculus cannot hallucinate a distinction. *No refutation ex nihilo* (Corollary 5.10): nor can it rule out, unmeasured, an identification. *Provenance* (Theorem 5.11): every derivable fork is forced by the record in every realizing model — verdicts are transported from data, never manufactured by inference (model-theoretically: Remark 5.12 states the downgrade). The proofs occupy this section and Appendix A; these six clauses are what they purchase.

In the LLM. For the practitioner the warranty reads: run an honest sweep and you may trust the verdict algebra — the analysis layer will never invent a cluster you did not measure, never erase a separation you did, never smuggle a conclusion past the clustering output it was given. The guarantees are of the same genre as type-safety for a programming language: not a promise that your experiment is wise, only that the inferential machinery adds nothing to it.

5.1 Models

Definition 5.1 (Apartness space)**.** An *apartness space* $(X, \#)$ is a set with an irreflexive, symmetric, cotransitive relation [12, 47]. Every metric space yields, at each resolution ε, the apartness "lying in different ε-components", where the ε-chain relation (steps of length $\leq \varepsilon$) supplies the matching indiscernibility.

Definition 5.2 (Regime-indexed simplicial apartness space)**.** A *regime-indexed simplicial apartness space* is a simplicial object X in apartness spaces such that:

(M1) each level X_n carries, for each regime V, a pair $(\simeq_V, \#_V)$ of h-proposition-valued relations. The denotation of a semantic type is a subspace of a level or a Σ-assembly of such (hom and Tri over Σ-assemblies interpreted fibrewise); each denotation carries the *induced* structure, with components computed *intrinsically*: two readings count as connected only if deformable into one another through readings, not through arbitrary ambient points — which is what makes a fork verdict about the filler space rather than the embedding cloud. On Σ-assemblies the induced structure is the total-space one: in metric models, a step between pairs is measured as the maximum of the componentwise steps, so that pairs sharing a component are exactly as far apart as their other components. The rules of §2.5 hold per denotation;

(M2) faces are strongly extensional (metric models: nonexpansive); degeneracies preserve and reflect both relations at every regime (metric models: isometric) — read at the level structures;

(M3) **(discrete identity)** the ambient identity type is interpreted discretely: only reflexivity. Coarse path-object interpretations of identity are *not* members of the model class: with non-trivial identity, unrestricted path induction reaches hom- and Tri-dependent families ($\mathsf{tr} : (a =_S b) \to \mathsf{hom}_S(a, c) \to \mathsf{hom}_S(b, c)$) which no apartness axiom governs. Discreteness makes every instance of path induction a reflexivity-transport, trivially sound — and costs nothing: every model used in this paper is discrete in this sense. All non-trivial sameness is carried by the regime-indexed $\simeq_V$, which is precisely the design.

A model *realizes* Δ when it validates every recorded verdict.

Remark 5.3 (What TTS is, positioned plainly)**.** Under (M3) the interesting geometry does not live in the identity types: in the intended models, identity is discrete and all non-trivial sameness is carried by the regime-indexed $\simeq_V$. So TTS is not a repair of HoTT's treatment of semantic composition — it *brackets* homotopical identity and adds an observational layer: intensional MLTT with primitive directed types and graded, record-fed observational indiscernibility and apartness. We take this to be the contribution, not a concession: the layer is where measurement lives, and the metatheorems of §5.3 are about the layer. (The kinship is less with simplicial HoTT than with observational equality [1], graded by instruments and grounded in records.)

Example 5.4 (The resolution model)**.** Let X_0 be a set of contextual states embedded in a metric space (for the intended case: embedded states of a generative model under $d_{\cos}$), with levels assembled from observed or realizable transitions and compositions. For each resolution ε: $u \simeq_\varepsilon v$ iff u, v are joined by a chain of steps of length $\leq \varepsilon$ within the relevant denotation; $u \#_\varepsilon v$ iff they lie in different ε-components. All rules of §2.5 hold (cotransitivity: a third point lies in a different component from at least one side; monotonicity: ε'-chains are ε-chains, so components refine; extensionality: same component as u together with a different component from v is a different component from v); identity is discrete. The estimator of §6 — connected components of the ε-neighbourhood graph on a sample — computes the apartness classes of this model on the sample: the experiment is the semantics, evaluated on data.

Remark 5.5 (Implemented instruments are nested partitions)**.** In any model where each $\#_V$ is total relative to $\simeq_V$ on each denotation — every pair either connected or separated, the situation of the resolution model on a sample — the rules of §2.5 force the verdict structure to be a partition per regime ($\simeq_V$ an equivalence; $\#_V$ = distinct blocks, whence cotransitivity and extensionality are automatic), and the monotonicity rules force the family of partitions to be *nested* along $\leq$. On a chain, an implemented instrument is therefore exactly a dendrogram over the readout image, and the admissible estimators are the hierarchical clusterings: single-linkage at threshold ε computes precisely the ε-graph components of Example 5.4. Estimators that are not partition-nested across their own parameter — k-means across k, mixture-model selection across component counts — do not implement regimes; this is why the mixture verdicts of §6's finding (c) enter the analysis as readout-level descriptions and never as entries of Δ_P.

5.2 Soundness and the honesty of records

Theorem 5.6 (Soundness, environment form)**.** *If $\Gamma \vdash_\Delta t : A$, then in every regime-indexed simplicial apartness space realizing Δ, $[\![t]\!]$ is a section of $[\![A]\!]$ over $[\![\Gamma]\!]$.*

Proof (cases in Appendix A). Induction on derivations. The MLTT cases are standard; path induction is reflexivity-transport by (M3), hence sound for *arbitrary* motives — including hom-, Tri-, $\simeq$- and $\#$-dependent families. The directed rules are immediate from the simplicial structure; the relational rules hold per denotation by (M1); (Δ-ax) by realization. Substitution coherence: the stated convention plus the standard comprehension-category treatment of the MLTT fragment; the full categorical elaboration is routine under (M3) and deferred. □

Theorem 5.7 (Consistency of honest records)**.** *Let P be a finite sample of a metric space, with the induced metric, and finitely many resolutions. The record Δ_P of ε-component verdicts* on the sample *is realized by the resolution model on P — the record* is *that model's component structure — so* $\mathrm{TTS}(\Delta_P)$ *is consistent.*

Proof. Equip P with discrete simplicial levels (the sampled completions as the relevant fibre), chain/component relations at every resolution, and discrete identity: a model per Definition 5.2 whose verdicts at the swept resolutions are by construction the recorded ones; conclude by Theorem 5.6. □

Two glosses, both load-bearing. First, the record is a record *about the sample*: sample-apart points may be connected in the ambient space through unsampled territory; the instrument's verdicts quantify over what it measured, which is what the resolution model on P interprets. Second, the guarantee is a property of *honest measurement*, not of the record format: a stipulated record can explode ($\delta : u \#_V u$ refutes itself by $\#$-IRREFL), and an honest sweep never records a Theorem 3.10-refutable pair, because its components are consistent across its own resolutions. The calculus does not certify the record; the metatheory certifies the instrument.

5.3 Conservativity and provenance

Theorem 5.8 (Conservativity)**.** *For Γ, A in the $\simeq$ /#-free fragment (hom and* Tri *allowed): if $\Gamma \vdash_{\emptyset} t : A$ then $\Gamma \vdash t^{\circ} : A$ is already derivable in MLTT + hom/*Tri*. (Erasure translation; Appendix A.)*

Corollary 5.9 (No fork ex nihilo)**.** *There is no closed term $\vdash_{\emptyset} t : \mathsf{Fork}_V(\sigma)$, nor of the judgment of record $\|\mathsf{Fork}_V(\sigma)\|$, for any closed horn and any regime.*

Corollary 5.10 (No refutation ex nihilo)**.** *Dually, there is no closed term $\vdash_{\emptyset} t : (u \simeq_V v) \to \mathbf{0}$: the empty record refutes no indiscernibility either. (Erasure sends the type to $\mathbf{1} \to \mathbf{0}$; apply to $\star$.)*

Theorem 5.11 (Provenance — closed judgments)**.** *If $\vdash_{\Delta} t : \mathsf{Fork}_V(\sigma)$ (empty Γ, σ closed; likewise for $\|\mathsf{Fork}\|$), then $\mathsf{Fork}_V(\sigma)$ holds in every model realizing Δ. Hence if* some *model realizes Δ without forking σ at V, no derivation exists: forks are never created by the calculus, only transported from the record.*

Remark 5.12 (What provenance adds)**.** By the deduction theorem, $\vdash_{\Delta}$-derivability is equivalent to $\vdash_{\emptyset}$-derivability from the record's propositions as hypotheses; the theorem's content beyond conditionalization is threefold. (i) With Theorem 5.7, the conditional is *non-vacuous* for honest records: a realizing model exists, so the verdict concerns an actually realizable configuration — whereas conditionalization holds vacuously even for explosive records. (ii) The empty-record case is secured *syntactically* by erasure, independent of any model class. (iii) Record constants are not abstractable, so record-counterfactuals are not internalizable: categoricity relative to the actual record is a property of the theory, not of a sequent. We also state the downgrade plainly: provenance is model-theoretic; no syntactic witness-extraction is claimed, since normalization for the axiomatic extension is not claimed.

5.4 Basins, strata, and the third grade

Definition 5.13 (Points; V-components)**.** The *points* of a semantic type S are the elements of $[\![S]\!]$ in a fixed model (syntactic reading: closed terms, identified when $\vdash_{\Delta} u \simeq_V v$ is inhabited — an equivalence by refl/sym/trans). The *V-component* of a point is its $\simeq_V$-class; components partition the points.

Definition 5.14 (Basin at a regime)**.** A V-component C of S is a *basin at V* if every inner 2-horn whose three boundary vertices lie in C satisfies

$\mathsf{Canon}_V(\sigma)$ — ambient canonicity, full $\mathsf{Fill}(\sigma)$, no relativized notion. No maximality is invoked: a basin, when it exists, is the unique component of its points; whether a given component *is* a basin is a model-side property — like canonicity itself (Remark 3.9), refutable by finite records, never certified by them.

Corollary 5.15 (Basins exclude forks)**.** *If C is a basin at V and σ has all three vertices in C, then $\mathsf{Fork}_V(\sigma) \to \mathbf{0}$, by Exclusion at the same ambient filler space.*

Remark 5.16 (Stratification, exactly)**.** The contrapositive is precise: *no basin contains all three vertices of a forked horn.* In particular, if a forked horn's vertices are co-located in a single V-component, that component is not a basin — the ambiguous occurrence is a singular point. When the vertices span components, the forked horn forces the ambiguous middle vertex *out of the flanking contexts' basin*; whether its own component is a basin is a further check, not a consequence. This is the formal counterpart of the stratified picture of [42, 31]: basins are the strata where travel is forced; fork loci are where strata meet.

Definition 5.17 (The third grade: locked)**.** A basin C at V is *locked* if every escape horn from C — a horn with some boundary vertex in C and some outside — has empty filler space, $\mathsf{Fill}(\sigma) \to \mathbf{0}$. (Emptiness is regime-independent; the regime enters solely through basin-hood.) A trajectory confined to a locked basin exhibits *over-coherence*: maximal local fluency, zero semantic decision — the degenerate-fluency phase documented in repetition collapse and low-temperature locking [3, 36]. The criticality finding of [36] — natural-language statistics arise at the phase boundary — receives a structural gloss: meaning requires residence near the fork locus; both the locked basin and the fully disconnected regime are semantically dead. We offer this as an interpretation, not a theorem.

The three conditions of §1 are now formal — with their typing stated honestly. **Canonical** (Canon_V) and **forked** (Fork_V) are the two observational grades *of a horn*; they do not exhaust verdicts (undetermined is the generic state of an unexamined horn). **Locked** is not a third parallel grade: it is a basin-level construction *from* canonicity — a dynamical pathology of a region, not a verdict on a demand. The triple is a phenomenology of travel; the trichotomy-sounding talk of §1 should be heard that way. One asymmetry organizes everything: forks persist under refinement and are verifiable from finite records; canonicity and basin-hood are falsifiable-only; locked-ness adds an emptiness condition that no record can certify either. The logic's positive verdicts are exactly the ones measurement can deliver.

6 The empirical bridge

The resolution model was chosen for more than soundness: every constituent of it is computable from samples. A horn is a prompt with flanking commitments; fillers are sampled continuations; a regime is an embedding model with a resolution; apartness classes are components of the ε-neighbourhood graph; and the readout through which fillers are observed is a Σ-constant whose pullback (Definition 3.11) is itself a regime structure. The dictionary is Table 1; the reader is invited to keep it open — every row is exercised below. We state the correspondences as a falsifiable protocol, with the limits of current evidence marked — including, prominently, the scope limits of our own pilot.

The bridge, stated once

The estimator computes the resolution model's defining relations on the sample: the experiment is the semantics, evaluated on data (Example 5.4); the left column of Table 1 is not an analogy for its second — it computes it. The laws of the calculus therefore hold in this pipeline *by implementation identity*. What is contingent — and observed — is which verdicts the world writes into the record: that calibrated ambiguous boundaries fork while controls do not; that right context prunes the bridge distribution onto one basin; that alignment tuning collapses it. The logic supplies the grammar of instrument verdicts; the model's behaviour supplies the sentences; the experiments demonstrate the sentences are non-trivial. One scope admission belongs here, half-discharged below: the embedding regimes this study instantiates form a *single chain* ($d_{\cos}$ at swept ε); the incomparable configurations of Theorem 3.10(b) and §4.4 are exercised schematically in those sections — and once, live, by the second-modality instrument at the end of this subsection.

The identity holds at the comparator stage, on the sample. Three joints carry slack, each guarded, each stated here once:

(S1) *Sample → distribution.* Verdicts are about the sample (Theorem 5.7, glosses). Under sample extension, connections persist (chains survive new points) but separations need not (new points can bridge components) — exactly dual to the formalized refinement axis, where separations persist (Theorem 3.6) and connections need not (Proposition 3.7). The calculus owns the refinement axis; the sample axis is guarded statistically by the (m, L) persistence criterion, a stipulation, not a theorem.

(S2) *Distribution → filler space.* The generator determines which fillers become observable: cones sample arriving-commitment-only demands; genuine 2-horns require FIM; an inadequate generator yields *undetermined*, not canonical (the GPT-2 null). Generator validity is stage one of Definition 2.8, outside the formalism.

(S3) *Text → filler.* A filler is a pair (composite, warrant); the readout factors through generated text, so the warrant component is unobserved at every pullback index — Proposition 3.7's countermodel is precisely the configuration the instrument cannot detect. Fluent generation under the flanking commitments is the warrant's operational shadow, declared, not derived.

The five scope admissions distributed through this paper — sample-relativity, cone scoping, warrant blindness, the instrument stack, falsifiability-only canonicity — are consolidated in the three joints above and in Table 1's final column; their original statements stand at full strength where they occur.

A second modality, from the logs. The incomparable case of Theorem 3.10(b) can be instantiated with canonically constructed instruments and no new sampling. Define W_{lex}: a symbolic, single-resolution regime partitioning completions by lexical evidence — financial-lexicon-only, riparian-lexicon-only, other — over fixed wordlists disclosed in the supplement. Same-block is its indiscernibility (an equivalence), distinct-blocks its apartness (irreflexive, symmetric, and cotransitive automatically, for a partition): a legitimate regime by §2.5, on a one-element chain. It is *declared* incomparable to the embedding chain, and the declaration is justified at instrument construction (Remark 2.10): no factoring argument exists in either direction — lexical membership neither bounds nor is bounded by cosine distance. Mining the cone-pilot logs ($n = 300$, AMB) yields both halves of the configuration. Records #85 and #105 are W_{lex}-connected (both financial-only) yet separated at the fine embedding regime ($\varepsilon = 0.40$; cosine distance 0.836): the pair $\{\gamma : u \simeq_{W_{\text{lex}}} v,\ \delta : u \#_{V'} v\}$, consistent precisely because $W_{\text{lex}} \not\leq V'$. Dually, records #105 and #199 (the miniature's φ_4) are W_{lex}-separated (financial-only against riparian-only) yet connected at the coarse embedding regime

($\varepsilon = 0.70$; distance 0.608). Two instruments of architecturally different modalities, disagreeing licitly on named pairs from one dataset: the configuration the philosophy of §4.4 runs on, exhibited in data. (Script, wordlists, and verdicts: `run_wlex.py` and `wlex_results.json` in the supplement.)

Manifold-hypothesis violation $\rightsquigarrow$ candidate fork loci. [43] rejects, by a fiber-bundle null test on radial volume growth, the hypothesis that token subspaces are manifolds; singular neighbourhoods are concentrated on identifiable tokens, and prompts containing them yield measurably less stable continuations (their Theorem 2 propagates input singularities to outputs). In TTS terms: loci where local dimension is undefined are exactly where no single basin chart covers the boundary data — candidate inhabitants of the fork locus. The inference is directional: non-manifoldness certifies that canonical-everywhere fails; it does not by itself produce a fork witness.

Forking tokens $\rightsquigarrow$ apartness verdicts. [11] resamples generation token-by-token and detects positions where the downstream outcome distribution splits into discrete alternatives. Read in TTS: the resampled continuations are observed, through a readout, as draws from a space of completions; the detected outcome-classes are apartness components at the pullback index; a forking token is a position at which the protocol has *recorded* the components of a separation — two concrete completions apart at the declared resolution.

Scope: cones and horns. A scoping point we state plainly, because the formal subject of Fork_V is the inner 2-horn. Free-completion protocols — Bigelow et al.'s and our first pilot below — fix only the *arriving* commitment: the measured object is a *cone* (context, awaiting continuation), not a 2-horn (both flanking commitments held fixed). The verdicts such protocols deliver are apartness verdicts on sampled completions at the pullback index ρ^*W — fully inside the theory by Proposition 3.12 — but they instantiate no $\mathsf{Fork}_V(\sigma)$. The genuine 2-horn test is, in machine-learning terms, a **fill-in-the-middle (FIM) evaluation** [17, 6]: hold prefix *and* suffix fixed, sample the bridge — a direct sample of $\mathsf{Fill}(\sigma)$'s composites, with the model's fluent generation under both flanking commitments as the operational shadow of the warrant.

The two FIM instruments. A language model never generates backwards; fill-in-the-middle works by moving the future into the prompt, and there are two ways to do it. *Native* FIM is trained in [6]: pretraining documents are cut into (prefix, middle, suffix) and re-ordered as prefix–suffix–middle behind

Concept (locus)	Empirical referent	Artifact (supplement)	Refutation / vacuity
semantic type A (Def. 2.3)	space of embedded contextual states	`text-embedding-3-small` image	
point $a : A$ (Def. 2.2)	a prompt state	the AMB boundary string	—
continuation, hom (§2.2)	a realized completion stretch	e.g. AMB log #199	—
horn, cone configuration (scope note)	prompt, arriving commitment only	cone-protocol prompts	—
horn, genuine 2-horn (Def. 2.4)	prefix and suffix both fixed	FIM prompts NEU/FIN/RIV	—
filler (h, t) (Def. 2.4)	completion + realizability warrant; warrant unobserved, fluency its shadow	logged bridge texts	warrant-only differences invisible at every pullback index (S3)
$\mathsf{Fill}(\sigma)$	support of generator-accessible completion distribution	N-sample logs	generator inadequacy → *undetermined* (GPT-2 null; S2)
regime $V = (d, \varepsilon)$ (Defs. 2.7, 2.8)	comparator: metric + threshold	$(d_{\cos}, \varepsilon)$ sweep	resolution-relativity observable: same four completions form 4/2/1 clusters at 0.40/0.60/0.70 (Ex. 2.1)
readout ρ (Def. 3.11)	embedding model + span choice	`text-embedding-3-small`, first sentence	pullback blind to warrants (S3)
order $V' \leq V$ (Def. 2.8, Rem. 2.10)	threshold arithmetic on a chain	$0.40 \leq 0.60 \leq 0.70$	refutable by measured $\simeq$-coarse + $\#$-fine pair within a declared pair; for ε-graphs none can occur, *by construction*
witnesses γ, δ (Def. 2.13)	named pair + grounding numbers	$\delta_1 = .691$ vs 0.60; logs #114, #89	sample-relative (S1)
record Δ_P (Def. 2.13, Thm. 5.7)	the sweep's complete verdict list	supplement, per manifest	honest sweeps consistent (Thm. 5.7)
$\mathsf{Fork}_V(\sigma)$ (Def. 3.4)	two completions + recorded separation; protocol verdict adds (m, L)	$(\varphi_1, \varphi_3, \delta_1)$, modulo cone scope	vacuity = ubiquitous single persistent components (honesty clause)
$\mathsf{Canon}_V(\sigma)$ (Def. 3.3, Rem. 3.9)	failure-to-refute only; never certified	control unimodality	falsifiable-only, by design
basin (Def. 5.14)	control conditions, estimator sense	FIN / RIV condition means	falsifiable-only
locked basin (Def. 5.17)	mode collapse	`gpt-4o-mini` bridges, Appendix B	offered as interpretation, not theorem
monotonicity laws (§2.5)	nestedness of components across the sweep	Ex. 2.1 δ/γ behaviour	*by construction* in the estimator — illustration, not confirmation
cross-regime consistency, chain case $V < W$ (Thm. 3.10)	coarse connection + fine separation in one dataset	(φ_1, φ_3) at 0.60 vs 0.70	live in Ex. 2.1
cross-regime consistency, in-	two modalities disagreeing licitly	live: W_{lex} vs the embedding chain	declared incomparability refutable per

sentinel markers, so that at inference the model receives both flanks in its context window and writes the middle left-to-right, conditioned on what must come after — the mechanism behind code editors' insert mode, and exposed for our model as the completion API's suffix parameter. What we will call *prompted* FIM — the label is ours; the practice, infilling via instruction, is standard and descends from Donahue et al. [17] — carries both flanks inside an ordinary prompt ("story before the gap ... story after the gap ... the missing middle:") and asks for the bridge. The two instruments sample *different conditional distributions*: native FIM approximates the document model's $P(\text{middle} \mid \text{prefix}, \text{suffix})$; prompted FIM samples an instruction-conditioned variant in which the demand that the whole "read as one coherent whole" is part of the conditioning. Neither is the other's refinement. They are, in the paper's own vocabulary, incomparable regimes — and we run both. One blind spot is structural either way: our ρ factors through the generated text, so the measured index cannot separate fillers differing only in their warrant — precisely the configuration of Proposition 3.7's countermodel.

The fork test. The protocol, then: given a boundary, sample N continuations, embed completions (ρ), and compute the components of the ε-neighbourhood graph across a sweep — by Example 5.4, these are the apartness classes of the resolution model at the pullback index, evaluated on the sample. The verdict criterion must respect two facts. First, the asymmetry of the grades (Theorem 3.6, Proposition 3.7): separations persist downward through the sweep; single-component verdicts are always provisional. Second, on a finite sample every point is eventually its own component as $\varepsilon \to 0$, so "stable under refinement" needs a rejection region: we adopt a persistence-interval criterion — a fork verdict requires ≥ 2 components, each of size $\geq m$, across all ε in a window of length $\geq L$, with (m, L) calibrated against a resampled unimodal null. The component-count clause is licensed by monotonicity; the size floor and null calibration are an independent statistical stipulation whose job is to reject sampling dust — they are not consequences of the calculus, and we do not present them as such. "Canonical" verdicts, by Remark 3.9, are estimator verdicts — failure to refute — never calculus theorems. The sweep's component verdicts — one apartness entry per separated pair per resolution, one indiscernibility entry per within-component pair — *constitute* the measurement context Δ_P of Theorem 5.7, and the protocol's fork verdict is the closed judgment $\vdash_{\Delta_P} \|\cdot\|$-apartness at the pullback index: the experiment writes the record; the calculus derives what it forces.

Pilot study. We ran the test on the ambiguity case of §4.3 — in cone configuration, per the scoping note above. (The four completions of Example 2.1 are drawn from this run; the reader has been looking at its data since §2.) Protocol: three boundaries — an ambiguous occurrence of *bank* (AMB), a financially disambiguated control (FIN), a riparian control (RIV) — each sampled N=300 times from a public completion model (`gpt-3.5-turbo-instruct`) at temperature 1.0 without truncation (unbiased draws from the model's filler distribution); completions embedded (`text-embedding-3-small`, first sentence — the sense-committing region); per draw, a scalar *sense margin* (cosine to a financial anchor minus cosine to a riparian anchor — itself a second readout in the sense of Definition 3.11, measuring at its own pullback index); mixture model selection (BIC, k=1 vs k=2) on margins, and the component sweep above on full embeddings. Throughout the findings, "basin" is used in its estimator sense: the control conditions anchor two putative basins (basin-hood being falsifiable-only, their unimodal behaviour is failure-to-refute), with condition means as representatives. Three findings (Figure 1):

(a) *The fork signature is detectable, after calibration.* The ambiguous boundary's margin distribution is decisively bimodal (ΔBIC(2−1) = −39.3), with its minor mode at −0.196, within a fraction of a standard deviation of the riparian control's mean (−0.218), and its major mode on the financial side: the two populations of fillers align with the two reference basins. The riparian control is unimodal (ΔBIC = +15.4).

(b) *Boundary balance is instrument-critical, as the theory predicts.* An earlier, uncalibrated ambiguous boundary drew 169/31 financial-vs-riparian fillers and registered only marginally; a calibration pass over candidate boundaries (selecting the one whose n=50 pre-sample split most evenly) preceded the main run. A fork judgment requires only two separated fillers, not balance; but *detection power* at fixed N degrades with imbalance.

(c) *One-dimensional bimodality is not the fork criterion — basin alignment is.* The financial control is itself margin-bimodal (ΔBIC = −26.9): inspection shows the split separates banking-dense completions from same-scenario atmospherics — two regions *within one basin*, both on the financial side of the inter-basin midpoint. This is exactly what the resolution model requires: apartness is separation in the full space (different components), not multimodality along a contrast axis. The fork verdict for AMB stands because its modes straddle the basins; the

FIN split does not qualify because its modes do not. (A modality note: the margin and full-embedding readouts are compared here informally, and the exclusion is principled — mixture-model verdicts are not nested across their own parameter, so by Remark 5.5 they implement no regime and never enter Δ_P; the licensing of their disagreement by Theorem 3.10 is therefore analogical at the readout level, exact only for poset regimes. The fork criterion is the full-space instrument's.)

What this run establishes: non-vacuity of the fork signature at the pullback index (Definition 3.4 via Proposition 3.12), under the cone scoping stated above, with basin alignment per finding (c).

A baseline run with GPT-2-small under the same protocol detected no structure in any condition (sense ordering visible only in condition means). The null is not a canonicity verdict: it is *undeterminedness relative to an inadequate instrument* — the generator must be coherent enough to commit to a reading within the sampled window for the readout to measure anything. *What this run establishes: undetermined*, not canonical — generator validity is joint (S2) of the bridge.

Data availability. Nothing here is offered on trust. Every sampled text in every run — 3,630 records across the three cone pilots, the boundary calibration, both FIM instruments, the failed `gpt-4o-mini` calibration, and the 2-horn experiment — is logged verbatim with its full prompt context, and the supplementary package (logs, prompts and templates, embeddings, results, and all scripts, with a file-by-file manifest) accompanies the paper, archived together with the paper at DOI 10.5281/zenodo.20668096. Appendix B reproduces, for each key condition, the exact prompt and a seeded random sample of the logged outputs, unabridged, indexed into the package — including the mode-collapsed `gpt-4o-mini` bridges behind the instrument finding.

The 2-horn experiment: infilling. We then ran the genuine 2-horn test as a FIM evaluation. Design: *one* ambiguous prefix — the same calibrated boundary ("We met them near the bank, just before sunset.") — and *three* suffixes: neutral (NEU, "By the time we said our goodbyes, the stars were already out."), financially disambiguating (FIN, a manager unlocking doors for the last customers), and riparian (RIV, a boat carried downstream). For each horn, $n = 200$ bridges sampled at temperature 1.0; same instrument as the pilot. Two instrument notes, both of independent ML interest. First, an instruction-tuned chat model (`gpt-4o-mini`) failed calibration: its bridge

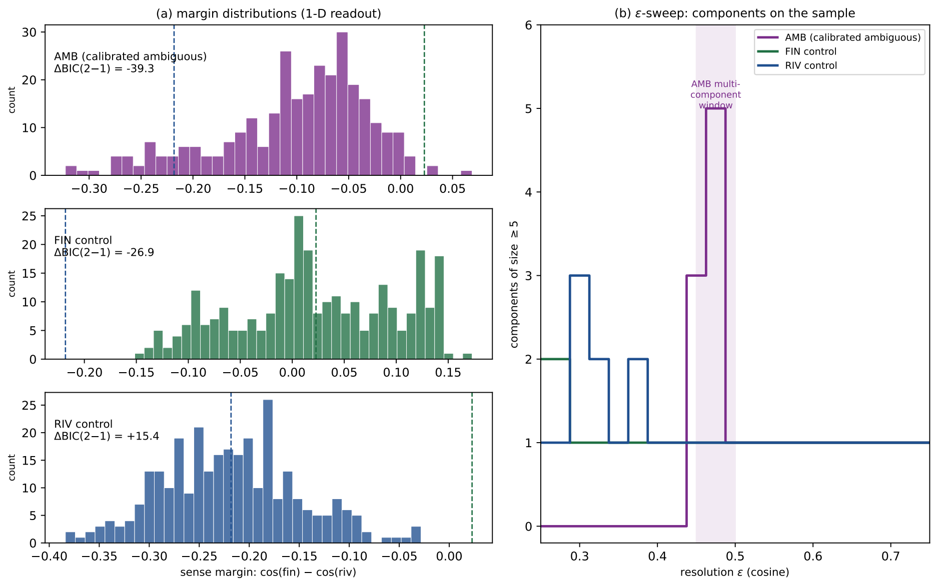

Figure 1: The pilot, in both readouts (300 sampled completions per boundary; cone configuration; apartness verdicts at the pullback index). *(a)* Sense-margin distributions; dashed lines are the control-condition means (basin representatives in the estimator sense). The ambiguous boundary spans both basins with modes aligned to each — the fork signature; each control remains within its own basin. *(b)* The ε-sweep: sizable components (≥ 5 members) of the sample's ε-neighbourhood graph. The ambiguous boundary alone holds multiple sizable components in the shaded window; the controls' fine-ε fragmentation below 0.40 is the sampling dust the persistence criterion's (m, L) calibration exists to reject.

distribution is mode-collapsed (near-identical sociable boilerplate) and it under-weights the suffix constraint — so the experiment uses prompted FIM on the completion model (`gpt-3.5-turbo-instruct`), which retains sampling diversity and respects both flanks. The instrument requirement is itself a finding: alignment tuning narrows the bridge distribution toward mode collapse. (Precisely put: no record certifies canonicity (Remark 3.9); what alignment manufactures is ever-stronger *failure-to-refute* — observed unimodality that the fork test cannot break.) *What the failed calibration establishes:* alignment collapse as a measurable narrowing of the bridge distribution, feeding the alignment–diversity open problem of §8. Second, all runs and boundaries are disclosed; the full-scale study will pre-register boundaries, sample sizes and the (m, L) criterion.

Results (Figure 2). The neutral-suffix horn is decisively bimodal on the margin readout ($\Delta\mathrm{BIC}(2{-}1) = -17.8$): a riverside-committed mode at -0.209, aligned with the riparian condition's mean (-0.223), and a second mode at -0.104 consisting largely of bridges that *defer* the reading — sociable narrative that completes the passage without committing. Deferral is a FIM-specific phenomenon the cone test cannot exhibit: a filler whose own reading remains unresolved, licensed exactly when the right context demands no commitment. We state the verdict with the paper's own discipline. By the standard of §6's protocol — persistent multi-component structure in the full embedding space — this sample does *not* certify an inhabitant of Fork: the bridges chain through their shared narrative register, and the ε-component criterion does not trigger at $n = 200$. What the experiment establishes is the *fork signature on the margin readout* (itself a pullback instrument), with modes aligned to distinct basins — branching evidence, by the paper's own finding (c) standard, short of a certified fork. The certified-fork experiment needs either larger n or a readout that suppresses register and isolates reading. *What this run establishes:* the margin-readout fork signature at a genuine 2-horn — not a certified full-space fork — and the suffix effect: the one qualitative prediction of §4.3, that enriching the boundary prunes the filler distribution onto one basin. The decisive evidence is the **suffix effect**, with the prefix held identical: the financial suffix kills the riparian branch (mean $+0.001$; both mixture sub-modes on the financial side; riverside bridges nearly absent), and the riparian suffix kills the financial branch (mean -0.223). This is §4.3's account of disambiguation, measured: enriching the boundary does not erase the fork — it poses a *different* horn whose foreclosed reading is no longer generable. In TTS terms, the other branch's triangles cease to be inhabited; in ML terms, right context prunes the bridge distribution to one basin.

Two instruments, one horn. Rerunning the same three horns through native FIM ($n = 200$ each; logs in the supplement and Appendix B) gives a textbook instance of the paper's own doctrine, and we set it out in full. First, what **replicates**: the suffix effect, the theoretically central prediction, holds under both instruments — the condition means order identically, with the neutral horn between the two disambiguated ones (native: NEU -0.160 between FIN -0.054 and RIV -0.218; prompted: NEU -0.145 between FIN $+0.001$ and RIV -0.223). Whatever else the instruments disagree about, both report that the right context prunes the bridge distribution toward its own basin. Second, what **does not**: the fork verdict at the neutral horn. Prompted FIM is decisively bimodal (ΔBIC $= -17.8$); native FIM is unimodal-broad (ΔBIC $= +3.9$). The logs explain the disagreement. The native bridges, produced by an objective trained for document continuity, frequently bridge the *scene* without re-engaging the ambiguous word at all — Tai Chi gatherings, a Mekong crossing, a Ferris wheel against the sunset (Appendix B) — deferral at scale: fluent connective tissue that postpones the reading indefinitely. The prompted instrument's demand that the whole "read as one coherent whole" forces an actual *reading*, and readings commit. Third, what the disagreement is and is not. "Instrument" bundles three roles the formalism keeps separate: the *generator* (which fillers become observable — here, two different conditional distributions), the *readout* ρ (how fillers are mapped into a measured space), and the *regime* (the connection/separation verdicts in that space). Theorem 3.10 formally governs the third role only — two regimes judging *one* pair — whereas native and prompted FIM differ at the *generator*: they sample different bridge populations, so the theorem's configuration is not literally instantiated. Nor is the native result a connection verdict: a broad unimodal sample is failure-to-detect, and the calculus itself insists failure-to-detect is not canonicity (Remark 3.9). What the comparison shows is the theorem's *lesson* operating one layer up — verdicts are relative to the whole instrument stack, and disagreement across stacks is data about the stacks. The robust finding is what survives both: the suffix effect. The instrument-validity question for the full study is now sharply posed — if the fork judgment's subject is the space of *readings*, the generator and readout of record must be the ones that elicit and preserve readings rather than scene-continuity, and that choice must be argued and pre-registered, not slipped in. *What this comparison establishes:* that stages of the instrument stack (Definition 2.8(iii)) vary independently — explicitly *not* a Theorem 3.10 configuration, since the two protocols differ at the generator, not the regime — and that the suffix effect is robust across generators.

For readers from machine learning. The entire protocol uses standard artifacts, and the theory is a logic over exactly those artifacts. Resampling a prompt = sampling the filler space of a cone; FIM with both flanks fixed = sampling an inner 2-horn; the embedding model with a threshold = a regime; the clustering output at swept thresholds = the measurement context Δ_P; stable multimodality of the completion or infill distribution = the fork; unimodality robust to resolution = (estimated) canonicity; the

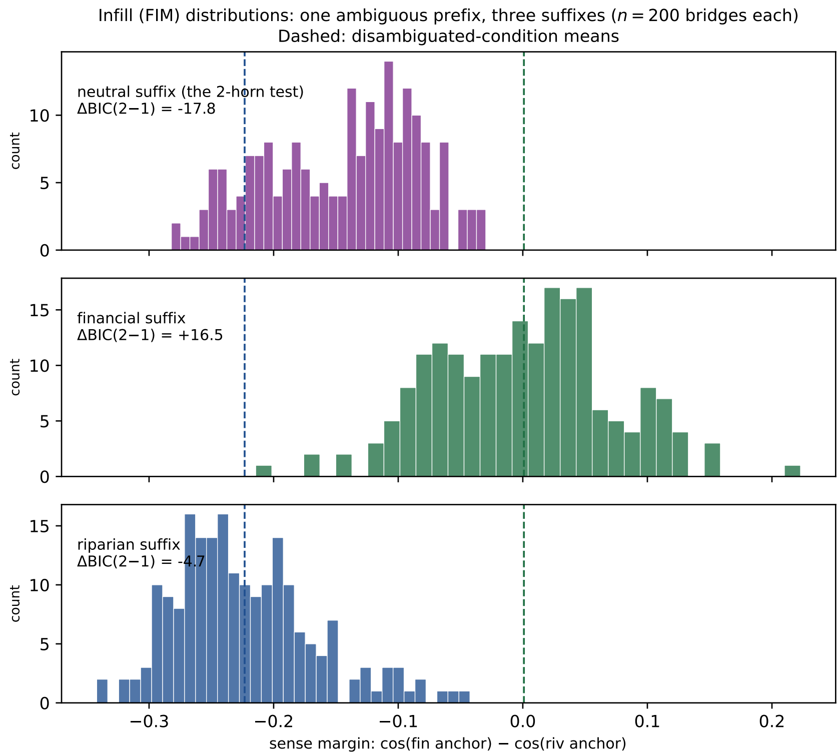

Figure 2: The 2-horn (FIM) experiment: one ambiguous prefix, three suffixes, $n = 200$ sampled bridges per horn. Dashed lines: the disambiguated conditions' means. The neutral-suffix horn (top) is bimodal — riverside-committed mode plus a deferral mass; each disambiguating suffix (middle, bottom) collapses the bridge distribution onto its own basin: the suffix effect.

persistence criterion = the rejection region. What the calculus adds to this familiar pipeline is a guarantee layer: verdicts compose across instruments and resolutions by rules with a soundness theorem — separations persist under refinement, instruments may disagree only in the configurations of Theorem 3.10, and no verdict appears that the recorded data do not force.

Honesty clause. What is proven empirically today is non-manifoldness and outcome-splitting; *non-contractibility of filler spaces* is strictly stronger than either, and no published test computes local homotopy type. The fork test is the proposed bridge, and it is falsifiable in both directions: ubiquitous single-component results under refinement would disconfirm the fork reading of forking tokens; stable multi-component results would instantiate the theory's central judgment in measurement.

7 Related work

Synthetic ∞-category theory. Simplicial type theory [41], its triangulated extension with directed univalence [20], machine-checked developments in Rzk [28, 32], and the model-theoretic analysis of [39] constitute a mature discipline of *directed* type theory. Throughout that lineage, composition-uniqueness (Segal-ness) is a property to be *imposed or established* for the types of interest. TTS inverts the stance: Segal-ness is a local judgment whose presence, failure, and boundary structure are the objects of the logic. We use the lineage's vocabulary (and intend its toolchain, §8) while occupying a complementary question.

Two-level type theory. 2LTT [2] pairs a fibrant (HoTT) layer with a non-fibrant (exo-)layer; exo-types may lack transport. The contrast with TTS is the contrast between *absence* and *witness*: in 2LTT a non-fibrant type simply fails to support the operations, and nothing internal certifies, locates, or measures the failure; in TTS the failure of canonicity at a specific horn at a specific regime is a positively inhabited type (Fork_V) whose inhabitant is a term — two named fillers and a recorded separation — with a provenance theorem (Theorem 5.11) governing where such terms can come from. The separation is term-level: fork witnesses persist under refinement (Theorem 3.6) and transport along readout pullbacks (Proposition 3.12) because there is a term to transport; 2LTT's non-fibrancy, being mere absence, supports no analogous calculus of verdicts.

Categorical and dynamical accounts of learned representations. GAIA [33] reads learning problems as horn-extension problems over simplicial sets — backpropagation solves inner horns — but is solution-seeking by design: its Kan assumption exists so that extensions succeed, no failure is tracked, and no theory of meaning is offered. Dynamical proposals model generation as trajectory evolution on an assumed semantic manifold [51], precisely the assumption the measurements of §6 defeat; none offers judgments, rules, or metatheory. The formal-geometry result of [24] (softmax + implicit bias provably promote linear concept representations within a latent-variable model) is the kind of peer-reviewed anchor a semantics of embedding geometry can build on, and is consistent with basin-interior linearity; it says nothing about the singular loci that concern us. Whether LLM geometry is *semantic* at all is contested [9, 30]; TTS does not presuppose an answer — its verdicts concern the structure of the geometry under measurement, and the Frege reading of §4 is the philosophical wager that this structure is where sense lives.

Theories of sense. The procedural lineage — Tichý/TIL [16], Moschovakis [34], Muskens [35], and in Martin-Löf type theory Bentzen [10] — individuates sense by identity of procedures; the proof-theoretic lineage [46, 7, 8] by derivational structure, with [8]'s bilateralist calculus of proofs *and refutations* particularly close in spirit to our positive/positive (canonicity/fork) architecture: the two grades may fairly be read as a geometric counterpart of bilateralism, with the instrument's connection- and separation-verdicts in the roles of assertion and denial. Truthmaker semantics [18] individuates hyperintensionally by exact verifier sets, set-theoretically and without procedural or geometric structure. The philosophical literature on identity in HoTT [29, 15] has the path-structure but has not, to our knowledge, connected filler spaces to modes of presentation. Informal anticipations that LLM latent geometry plays the role of Fregean sense exist [45]; TTS is, to our knowledge, the first *formal* system in which the identification is carried by a calculus with a soundness theorem. Categorical compositional distributional semantics [14] is the other formal bridge between compositional structure and vector geometry; the stances are complementary — DisCoCat composes meanings *in* the vector space, while TTS judges the failure modes of composition. Within formal semantics of natural language, the vector-space bridge of [38] (a homomorphism between Montagovian and distributional semantics, in this journal) is methodologically the closest published relative: it connects a formal semantics to embedding geometry; TTS connects a *proof theory* to that geometry's failure modes.

Singular thought. The mental-files tradition [40] is the closest philosophical model of §4.4, and the alignment deserves precise statement — in our vocabulary, not theirs. What that tradition posits as primitive mental particulars (files, dossiers) appears here as emergent structure: regions of a presentation space. Its *presumption of co-reference* — information within one file combines freely — corresponds to a monotonicity entailment: when the presentational regime refines the objectual one, co-location in a presentational component entails co-reference at the objectual regime ($\simeq$-TRANS

+ $\simeq$-MONO). Its *epistemically rewarding relations*, which individuate files by the channel through which information arrives, correspond to regimes — an ER relation is an instrument. And its signature doctrine, that identity discovery *links* files without *merging* them, is Proposition 4.6: flow at the discovering regime, provable segregation at the presentational one. We adopt the alignment, not the vocabulary or the representationalist metaphysics: nothing in §4.4 is a stored particular; everything is graded geometry under measurement.

8 Conclusion and open problems

We have presented a type theory in which the semantically decisive events of a stratified meaning-space — forced composition, genuine choice, locked fluency — are graded verdicts with positive, term-level witnesses; in which the verdicts are consequences of measurement by construction and by theorem (conservativity, no-fork-ex-nihilo, provenance, record consistency); and in which Frege's distinction is theorem-grade structure: sense is choice of filler, hyperintensional distinctness is apartness at a regime, and the consistency of discovered identity with standing difference is characterized exactly by the order between the instruments involved. Exclusion is derived, not legislated; the witnesses are measurements, not stipulations; and what the calculus cannot do — certify canonicity from a finite record, manufacture a fork, retract a verdict — is as much a part of its content as what it can.

Open problems, in ascending order of ambition:

(1) **Names as expressions.** §4.4 delivers the name-*state* configuration; what remains is the public name as an expression — the word-to-constant link, predication dynamics over presentation-points, and rigidity in full (Remark 4.7). A first name-pilot, logged in the supplement, taught us more than its design intended. The hypothesized dissociation — names apart presentationally, connected under predication readouts — did not appear for `gpt-3.5-turbo-instruct`. The logs show why. Asked for facts, the model answers about Venus perfectly: "108 million km," every draw. It never routes the classical names through that knowledge: Phosphorus's facts are confabulated. Yet asked directly, it affirms the identity (0.90 yes-same-object). So a one-step Leibniz syllogism fails in a fully *transparent* context: both premises assented to, the substitution never executed. That is stronger than Frege's puzzle, which concerns opaque contexts. The calculus names the situation precisely. Substitution everywhere is transport along ambient identity.

Ambient identity is refuted by the standing presentational separation (id-clash, Corollary 3.2). What the model holds is the graded connection $\simeq_W$, and that licenses substitution at W alone (Proposition 4.6). Predication and presentation are, for these signifiers, distinct regimes that do not marry. Meaning here does not operate syllogistically: the identities it trades in are never Id-grade, only some instrument's verdict. (Compare the reversal curse [5]: factual knowledge not closed under one-step logical transforms.) The question for the full study: which inferential closures does a model's geometry implement, and at which regimes?

(2) **General horns.** Extend §3 to Λ_i^n (Remark 2.6); identify which semantic phenomena require $n \geq 3$ (anaphora across discourse units is a candidate).

(3) **The alignment–diversity question.** The 2-horn experiment's instrument finding — an instruction-tuned model's bridge distribution is mode-collapsed where the completion model's forks — deserves systematic study: does alignment tuning collapse the filler distributions (ever-stronger failure-to-refute canonicity — no record can do more, Remark 3.9), and can the fork census quantify what it removes?

(4) **Normalization and decidability.** The extension is axiomatic; canonicity and normalization for it are not claimed. The natural hardening is fragment-wise: the relational fragment's proof terms carry near-groupoid structure ($\simeq_V$-terms under sym/trans, $\#_V$-terms acted on by EXT), suggesting a confluent rewriting independent of the (axiomatically inert) directed fragment — which would yield *syntactic* provenance exactly where it matters, for the fork witnesses, without solving normalization for the whole theory.

(5) **Completeness.** Soundness is Theorem 5.6; completeness for regime-indexed simplicial apartness spaces is open. We conjecture it fails for the full theory (the directed fragment is axiomatically inert) while holding for the relational fragment over resolution models on finite carriers.

(6) **Mechanization.** The directed fragment is within reach of the Rzk ecosystem [28]; the relational fragment is not, and an Agda development with postulated regimes and apartness is the pragmatic first step.

(7) **The empirical program.** The fork test at scale: a census of the fork locus of a production model, the singular-token cross-check against Robinson et al. [43], and the diachronic question — how the locus moves under fine-tuning — which is where this logic touches alignment.

A Proofs

A.1 Base consistency

Lemma A.1 (Base consistency)**.** *MLTT + hom/*Tri *over any semantic signature* Σ *is consistent.*

Proof. Extend any model of MLTT (e.g. in Set) by interpreting each base semantic type as $\mathbf{1}$, each point/hom/Tri constant of Σ as $\star$, and each function constant (readouts, Definition 3.11) as the unique map into $\mathbf{1}$; set $\mathsf{hom}(a, b) := \mathbf{1}$ and $\mathsf{Tri}(f, g; h) := \mathbf{1}$, with $\mathsf{id}_a, \lambda^{\mathsf{l}}_f, \lambda^{\mathsf{r}}_f \mapsto \star$. The interpretation is compositional; the directed fragment has no eliminators and no computation rules, so all five of its rules are validated, and $\mathbf{0}$ remains uninhabited. (The path-action $e_* f$ is derived by transport and needs no separate clause.) ☐

A.2 Erasure and conservativity

Lemma A.2 (Erasure soundness)**.** *Define* $(-)^\circ$ *on the language of* $\mathrm{TTS}(\Sigma, \emptyset)$ *by* $(u \mathrel{\#_V} v)^\circ := \mathbf{0}$*,* $(u \simeq_V v)^\circ := \mathbf{1}$*, homomorphically elsewhere (contexts, types, terms, derivations). Then* $\Gamma \vdash_\emptyset t : A$ *implies* $\Gamma^\circ \vdash t^\circ : A^\circ$ *in MLTT + hom/*Tri*.*

Proof. By induction on derivations; the MLTT and directed rules are untouched by the translation. The relational rules are admissible in the image: formation rules become "$\mathbf{1}\ \mathsf{type}$" / "$\mathbf{0}\ \mathsf{type}$"; the formers' built-in propositionhood becomes the standard proofs that $\mathbf{1}$ and $\mathbf{0}$ are h-propositions; $\simeq$-REFL $\mapsto \star$; $\simeq$-SYM/TRANS/MONO $\mapsto$ constants valued $\star$; $\#$-IRREFL/SYM/MONO $\mapsto \mathsf{id}_{\mathbf{0}}$; $\#$-COTRANS $\mapsto$ $\mathbf{0}$-elimination into $\|\mathbf{0} + \mathbf{0}\|$; $\#$-EXT $\mapsto \lambda(x{:}\mathbf{1})(y{:}\mathbf{0}).\, y$; id-to-regime $\mapsto$ constant $\star$. The translation preserves judgmental equality since the erased formers carry no computation rules. Note (Δ-ax) is precisely the rule erasure cannot validate — a $\#$-entry would erase to a closed term of $\mathbf{0}$ — which is why the lemma is stated for the empty record. ☐

Proof of Theorem 5.8. For Γ, A in the $\simeq$ /$\#$-free fragment, $\Gamma^\circ = \Gamma$ and $A^\circ = A$; apply Lemma A.2. ☐

Proof of Corollary 5.9. Suppose $\vdash_\emptyset t : \mathsf{Fork}_V(\sigma)$. Erasure gives $\vdash t^\circ : \sum_{\varphi,\psi:(\mathsf{Fill}\sigma)^\circ} \mathbf{0}$; projecting twice yields a closed term of $\mathbf{0}$, contradicting Lemma A.1. For the judgment of record: erasure sends $\|\mathsf{Fork}_V(\sigma)\|$ to $\|\Sigma_{\varphi\psi}\,\mathbf{0}\|$, and the truncation eliminates into the h-proposition $\mathbf{0}$ via the projections, giving the same contradiction. The companion fact — no closed $\vdash_\emptyset t : (u \simeq_V v) \to \mathbf{0}$ — follows by applying $t^\circ : \mathbf{1} \to \mathbf{0}$ to $\star$: the empty record refutes no indiscernibility either. $\square$

Proof of Theorem 5.11. Immediate from Theorem 5.6 applied to the closed judgment: every model realizing Δ inhabits $[\![\mathsf{Fork}_V(\sigma)]\!]$; the contrapositive gives underivability whenever one realizing model fails to fork σ at V. $\square$

A.3 Soundness: the case analysis

Proof of Theorem 5.6. Induction on derivations, over a model per Definition 5.2 realizing Δ. *MLTT fragment.* Standard [48], with one global simplification: by (M3) the identity type is interpreted discretely, so every instance of path induction is a reflexivity-transport — $[\![J(c;\mathsf{refl})]\!] = [\![c]\!]$ — and the case closes for *arbitrary* motives, including families built from hom, Tri, $\simeq_V$ and $\#_V$. (This is where coarse path-object interpretations would fail: a non-trivial identity between a, b with $\mathsf{hom}(a,c)$ inhabited and $\mathsf{hom}(b,c)$ empty refutes the transport case; such models are excluded by (M3).) *Directed fragment.* $[\![\mathsf{hom}_A(a,b)]\!]$ is the fibre of $(d_1, d_0) : X_1 \to X_0 \times X_0$; $[\![\mathsf{Tri}]\!]$ the fibre of the boundary map $X_2 \to X_1^3$ with the assignment $(f,g,h) = (d_2, d_0, d_1)$; degeneracies interpret id, λ^{l}, λ^{r} via the simplicial identities ($d_2 s_0 f = \mathsf{id}$, $d_0 s_0 f = f = d_1 s_0 f$, and dually for s_1). No rule demands inhabitation of any fibre: the fragment is sound in any simplicial object, Kan or not, Segal or not. *Relational fragment.* Per (M1), each denotation carries $(\simeq_V, \#_V)$ satisfying the rules of §2.5; each rule is validated pointwise. $\simeq$/#-MONO hold by the regime-indexed structure; #-EXT and the component laws were verified for the resolution model in Example 5.4. *Record.* (Δ-ax) holds by realization. *Substitution.* The new formers commute with substitution by convention; coherence of the interpretation follows the standard comprehension-category treatment of the MLTT fragment, extended to the new formers componentwise; under (M3) no fibration structure beyond reflexivity-transport is required. $\square$

A.4 Cross-regime witnesses

Witnesses for Theorem 3.10(b), in full. Case $V < W$ (strictly finer separator). Resolution model on a two-point carrier $\{u, v\}$ at distance r with $\varepsilon_V < r \leq \varepsilon_W$ (such r exists by strictness). Then $u \simeq_{\varepsilon_W} v$ by the single $\leq \varepsilon_W$

step; $u \#_{\varepsilon_V} v$ since $r > \varepsilon_V$; and the cross-regime monotonicity laws hold: $\simeq_{\varepsilon_V}$ is the diagonal, contained in the total $\simeq_{\varepsilon_W}$, and $\#_{\varepsilon_W}$ is empty, contained in the pair $\#_{\varepsilon_V}$. *Case V, W incomparable.* The two-point, two-regime record model: carrier $\{u, v\}$; $\#_V$ the symmetric pair, $\simeq_V$ the diagonal; $\simeq_W$ total, $\#_W$ empty. Every rule of §2.5 holds per regime: at V, cotransitivity splits within $\{u, v\}$ (one disjunct is always the recorded pair) and extensionality is trivial on the diagonal; at W the apartness rules are vacuous. *Extension to the full poset.* For any regime U set $\simeq_U :=$ total if $W \le U$, else diagonal; $\#_U :=$ the pair if $U \le V$, else empty. Monotonicity is transitivity of $\le$; a clash at some U would need $\simeq_U$ total and $\#_U$ nonempty, i.e. $W \le U \le V$, contradicting $W \not\le V$. *Completion to a full model.* Both witnesses extend to Definition 5.2 as the constant simplicial object on the carrier (faces and degeneracies identities, identity discrete), with hom- and Tri-fibres taken as the genuine fibres of the boundary maps — the diagonal fibres are singletons, the off-diagonal fibres empty, which is harmless: no rule demands inhabitation of any fibre. (M2) and (M3) hold trivially. The rest of the signature is interpreted as in Lemma A.1 (other base types as $\mathbf{1}$, constants as $\star$, function constants as the unique map); further constants of the carrier type may be sent to either point, and function constants whose codomain is the carrier are interpreted as constant maps — the record mentions only the named pair, and no rule constrains a readout's values. $\square$

The incomparable *filler*-pair case, used schematically in §4.2, follows by the same recipe: equip the two-filler carrier of Proposition 3.7 (two warrants over one composite) with the record-model relations at two incomparable regimes; every check above transfers verbatim.

A.5 The two-point diagonal model

Model for Proposition 4.5(ii). Interpret P as the two-point carrier $\{x, y\}$ with $[\![h]\!] = x$, $[\![p]\!] = y$; at *every* regime U set $\simeq_U :=$ the diagonal and $\#_U := \emptyset$; complete to Definition 5.2 as the constant simplicial object with discrete identity, interpreting the rest of the signature as in Lemma A.1, with function constants whose codomain is the carrier interpreted as constant maps and any further carrier-type constants sent to x (freshness of h, p — no other Σ-declaration or record entry mentions them — is what makes this free). Every rule of §2.5 holds: refl/sym/trans on the diagonal; $\simeq$-MONO across any poset pair is diagonal $\subseteq$ diagonal; the five apartness rules are vacuous on the empty relation; id-to-regime lands on the diagonal under discrete identity. The model realizes the empty record vacuously, and both $[\![h \simeq_W p]\!]$ and

$[\![h =_P p]\!]$ are empty at (x, y); Theorem 5.6 gives underivability of both. $\square$

Funding

This research received no external funding.

Acknowledgments

AI-assisted tools were used in the preparation of this manuscript and in the implementation of the experimental pipeline, and are disclosed here in accordance with the journal's policy; responsibility for all content rests with the author.

B Experiment log excerpts

Every text sampled in this paper is logged verbatim, with its full prompt context, in the supplementary data package (§6, data availability). This appendix reproduces, for each key condition, the exact prompt and a seeded random sample of the logged outputs (seed = 11235, indices shown) — unabridged and uncorrected, exactly as generated. The reader can check every quoted text against the package by its index.

Free-completion pilot (`gpt-3.5-turbo-instruct`, $T = 1.0$, $n = 300$/condition)

Each record: the prompt below, completed freely; first sentences of the completions were embedded. The four completions of Example 2.1 are records #114, #64, #89 and #199 of the AMB log.

Cone / AMB (calibrated ambiguous boundary)

Prompt (verbatim): *"We met them near the bank, just before sunset."*

#20: *The clouds were beginning to turn a deep purple and the air was thick with the smell of rain.* [newline] [newline] *"Welcome to our village," the man said, extending his hand in greeting. He introduced himself as*

#153: *Raymond was the father of the only person we knew, which made this strange meeting even more bizarre. Aaron seemed to vanish with the sunlight, but Raymond stood in the shadows, waiting for us.*

#155: *They had gotten out of a sleek black car, their faces hidden behind sunglasses. Five of them, all dressed in sharp suits, looking more like secret agents than the vigilantes they purported to be.*

#170: *It was a group of four, two men and two women, all wearing dark clothing and masks. They were armed with guns and led us inside the bank, where they quickly took control of the situation*

#185: *Like the early afternoon when these kinds of joints are crowded, the evening before the banks close proved to be favorable for moving cryptocurrency.* [newline] [newline] *These kinds of operations are risky and complicated. Often done with unprofession*

#192: *I can't tell you whether it was gonna be a clout session or a crushing session or something new, but they were gonna do something, and it was gonna be soon.* [newline] [newline] *"Sister*

#199: *As we paddled our kayaks up to the bank, we see that they were waiting for us. There were three of them, all dressed in simple grey robes, with their heads bald and their*

#210: *Sarah was carrying several books in her arms and Matt had a sketch pad tucked under his arm. They were both talking animatedly about a book they had just finished reading and were eager to share their thoughts*

Cone / FIN control

Prompt (verbatim): *"He needed to deposit the check before noon, so he walked down to the bank."*

#20: *As he entered the bank, he was greeted by a friendly teller who directed him to the deposit slips. He quickly filled one out with his information and the amount of the check.* [newline] [newline] *He handed*

#153: *As he made his way through the crowded streets, he couldn't help but feel a sense of nostalgia wash over him. He had grown up in this small town and had spent countless afternoons*

#155: *As he approached the bank, he noticed a line forming outside the door. It seemed that he wasn't the only one with important errands to run before noon. He joined the line, impatient*

#170: *As he entered the bank, he noticed a long line of customers waiting to be served. He groaned inwardly, knowing that he would have to wait for a while.* [newline] [newline] *He took a number*

#185: *As he entered the bank, he noticed that the line was long and he didn't have much time left. He approached the teller and handed her the check, explaining that he needed to deposit*

#192: *As he entered the bank, he was greeted by the teller who asked how she could assist him. He explained that he needed to deposit a check before noon.* [newline] [newline] *The teller directed him to*

#199: *As he stood in line, he checked his watch anxiously. It was already eleven thirty and he could feel the time slipping away. He hoped the line would move quickly.* [newline] [newline] *Finally, it was*

#210: *As he approached the bank, he noticed a long line of people outside the entrance. He groaned, knowing he would have to wait in line for a while before he could deposit his check.*

Cone / RIV control

Prompt (verbatim): *"He carried his fishing rod through the reeds and walked down to the bank."*

#20: *The sound of water gently lapping against the shore soothed his soul. He took a deep breath and cast his line.* [newline] [newline] *The sun was just beginning to rise, casting a warm glow*

#153: *Even as he stepped onto the muddy shore, he felt the tension of everyday life ease from his shoulders. This was his sanctuary, his escape from reality. He breathed in the fresh air and let the*

#155: *At this time of year, his favorite spot on the Mize River was shaded by a pair of ancient willow trees that dipped their twisted branches into the water. The air was thick and humid and*

#170: *The sun was just starting to dip below the horizon, casting a warm orange glow over the water. He took a deep breath, taking in the fresh, crisp air and the peacefulness of the*

#185: *I stopped just short of where he was standing. If he didn't want my help that was fine. I silently took note of how he was rigged: on the end of his line he had a*

#192: *He cast his line out into the water and settled down to wait. As he sat there, he took in the peaceful surroundings – the sound of the birds chirping, the gentle rustling of the*

#199: *He smiled as he felt a sense of calm wash over him. The sun was shining and there was a light breeze that rustled the leaves on the trees above. He could hear the sound of birds*

#210: *Ducks were lazing on the pond's surface and would scatter as he edged closer. It was a roasting, sunny day and it was perfect. Cormac let out a pleased breath and calmly cast*

The 2-horn (FIM) experiment (`gpt-3.5-turbo-instruct`, prompted infilling, $T = 1.0$, $n = 200$/horn)

Each record: prefix and suffix below, both held fixed; the model samples the bridge. The FIM template is reproduced in the supplementary package and in `run_horn.py`.

Horn / neutral suffix — the 2-horn fork test

Prefix: *"We met them near the bank, just before sunset."* Suffix: *"By the time we said our goodbyes, the stars were already out."*

#77: *As we walked along the riverbank, we shared stories and laughter with our new friends, watching the sky turn shades of pink and orange as the sun began to set behind us. We exchanged contact information and promised to keep in touch before parting ways, already reminiscing on the wonderful evening spent together.*

#85: *As we chatted with the group of strangers, we learned that they were travelers from a different country, exploring the same city as us. They shared stories of their adventures and we shared tips for the best local spots to visit.*

#92: *They were a merry group, with smiles on their faces and laughter in their voices. We shared stories and drinks by the riverbank, enjoying the warm summer air.*

#96: *They were a group of travelers from a far-off land, telling tales of their adventures and sharing exotic spices for our food. As the night went on, we laughed and learned from each other, creating bonds that would last long after we parted ways.*

#99: *We chatted and laughed with them, swapping stories and sharing snacks. As the sun began to dip below the horizon, we made plans to meet up again tomorrow for a hike in the mountains.*

#162: *We had a picnic by the river, laughing and sharing stories as the sun set over the water. They told us about their travels and we shared our plans for the future. It felt like we had known each other for years, and we promised to stay in touch.*

#189: *They were a group of travelers passing through our small town, and we quickly struck up a conversation. Their stories of faraway lands and thrilling adventures captivated us, and we ended up sharing a meal together at a nearby tavern. Laughter and camaraderie filled the air as we swapped tales late into the night.*

#196: *They were a group of local fishermen, and they invited us to join them for a drink on their boat. As we sat there, sipping on cold beer and listening to their stories of life at sea, the sun slowly dipped below the horizon and painted the sky in stunning shades of orange and purple. It was a peaceful and memorable evening, and*

Horn / financial suffix

Prefix: *"We met them near the bank, just before sunset."* Suffix: *"When the manager finally unlocked the doors again, we were the last customers of the day."*

#77: *The bank was bustling as people hurriedly finished their transactions before the doors closed. We noticed a couple standing nearby, looking confused and anxious. Without thinking, we went over to them and asked if they needed any help. They explained that they had just moved to the city and were trying to open a new account, but had missed the bank's*

#85: *They were a group of young kids, no more than ten or eleven years old, sitting on the sidewalk with a cardboard sign that read "Need help, anything will do." My heart broke seeing them there, so I asked them if they were hungry. They eagerly nodded and I promised to buy them a meal. We all went inside the bank together*

#92: *As we walked out, we realized we had forgotten to withdraw cash, so we went back in. The bank was now nearly empty except for a few tellers and the security guard who greeted us with a friendly smile.*

#96: *As we stepped outside, we noticed a group of street performers had set up nearby. They were juggling, dancing and playing music, drawing a small crowd with their lively energy. We paused to watch for a moment before* *heading to our cars, feeling grateful for the unexpected entertainment after a long day at work.*

#99: *We chatted with them briefly, finding out they were new to town and looking to set up a small business. They seemed like a friendly couple, and we promised to stop by their shop once they were up and running. As we walked away, we couldn't help but feel excited for this new addition to our community.*

#162: *As we walked out with our newly acquired loan, I couldn't help but reflect on how lucky we were to have stumbled upon the friendly couple who offered us a ride to the bank after our car broke down.*

#189: *We had been standing in line for what felt like hours, our stomachs grumbling and our feet sore. As soon as we saw the "Open" sign flicker on, we rushed inside, eager to complete our errands before the bank closed for the day. The teller greeted us with a smile and processed our transactions efficiently, but as*

#196: *The bank's interior was cool and dimly lit, the air thick with the smell of money. We quickly filled out the necessary paperwork and handed over the check, relieved to have our transaction completed. As we left, we noticed the sun sinking lower in the sky and decided to stay and watch the sunset from the nearby park.*

Horn / riparian suffix

Prefix: *"We met them near the bank, just before sunset."* Suffix: *"We watched the current carry our little boat downstream until it disappeared."*

#77: *We had been planning this trip for weeks, and the river had seemed like the perfect spot to unwind and reconnect with nature. But as the sky darkened and the temperature dropped, we realized we were ill-prepared for the journey. The friendly couple we had met offered to tow us back to shore, and we gratefully accepted, feeling humbled*

#85: *The group of locals had warned us about the treacherous waters, but we didn't listen. As we struggled to navigate the rough waves and sharp rocks, we realized our mistake. But it was already too late. Our boat capsized and we were left clinging onto debris for dear life. Thankfully, the locals came to our rescue and brought us*

#92: *The strangers offered us a ride on their bigger boat and we accepted gratefully. As we traveled down the river, we shared stories and laughs,* *and before we knew it, the sun had set and the stars had come out, filling the sky with their twinkling lights.*

#96: *As the sun began to dip below the horizon, we decided to paddle back to shore. Suddenly, the wind picked up and the water became choppy, making it difficult for us to row against the current. We were worried we wouldn't make it back in time, but thankfully, we eventually reached the shore and pulled our boat onto the bank just*

#99: *We waved goodbye to our new friends and turned to head back to our campsite. As the sun sank below the horizon, we sat by the fire and recounted the day's adventures, grateful for the unexpected encounter that had made our trip even more memorable.*

#162: *As we waved goodbye, I couldn't help but feel a twinge of sadness. These strangers we had met just one day ago had shown us such kindness and hospitality on our journey down the river. I knew we would never forget the experience and the unexpected friends we had made.*

#189: *We saw them on the other side of the river, waving and laughing at us. We couldn't help but join in and wave back, feeling a sense of joy and connection even though we were separated by the rushing waters.*

#196: *As we made our way towards the shore, they offered to give us a ride on their larger boat. We eagerly accepted, grateful for the chance to rest our tired arms.*

Native FIM (API suffix conditioning, `gpt-3.5-turbo-instruct`, $T = 1.0$, $n = 200$/horn)

Same horns as above, no instruction template: the prefix is the prompt, the suffix is passed natively. Note the scene-bridging register — the basis of the two-instruments finding.

Native FIM / NEU

Suffix: *"By the time we said our goodbyes, the stars were already out."*

#77: *It was a sprawling group of all ages, practicing Tai Chi. Curiosity leading the way, we walked among them, listening to the sounds of the summer evening at the Chagrin River.*

#85: *It was cool and slowly cooling even more, but no one seemed to mind as they tightly wound terrible tales about corpses and skeletons and rivers* *that ran with untold horrors.* [newline] [newline] *Alex told me that he was a writer, that he was well known in The Realm. But I recognized him at once as the reality broker who'd started our little scavenger hunt*

#92: *This was the second day of our backup crossing of Mekong. We had started near Mae Salong, drove till Chiang Khong, crossed to HueXieng in Laos on the first day. And then drove south - to a village on the outskirts of Luang Prabang.* [newline] [newline] *We, a group of four, had 'saved'*

#96: *The band stood against the railing with the sun behind them, the buildings and—as fate would have it—the Ferris wheel (tourists) behind us. We got to talk music, inspiration and even sentiments of the gambler.*

#99: *In a quiet, sheltered inlet, Dave and Luticia were sitting in their small boat, patiently pulling in a fishing line. They must have spotted us paddling our kayak their direction. I'm glad they waved us down.* [newline] [newline] *They had a beautiful boat – probably one of the most graceful-looking I've ever seen. It was the kayak variety*

#162: *A bit shy at first, Grace quickly opened up, showing us around her community and introducing us to her daughter.*

#189: *The lights were dipping below the mountains, staining the sky violet and pink. They had wandered here from up north, and had brought their fiddles, their drums and their songs with them. Their voices were like water, spells woven together to pull light out of the dying day.*

#196: *By the time we had reached the palms, their boat had transformed from a onetime shipping vessel to a dinner party.* [newline] [newline] *Threadbare rugs were strewn over what seemed like it had previously been just plain old timber, a worn guitar had made its appearance, and plastic chairs had been set around a worn table. My aunt had already seated herself and booked*

Native FIM / FIN

Suffix: *"When the manager finally unlocked the doors again, we were the last customers of the day."*

#85: *They were on the water because the bank had turned them away. It was five minutes to closing time.* [newline] [newline] *Sitting in the back of Mark's truck, watching the Swans in the warm evening breeze, was a* *pleasant way to wait a few minutes. And to them, that's how we were. Just waiting a few minutes.* [newline] [newline] *I thought they*

#96: *Each was a dignitary in his people's community, bearing gifts appropriate to a king's reception. The meeting was as a very long week in the lives of a wife and husband away from home.*

#99: *They stood out awkwardly on the sidewalk, square, ugly, and utterly out of place. We had seen them before, and knew what they were capable of.* [newline] [newline] *"Didn't know they'd spread so far." I muttered under my breath. Rae gave me a stern look. We both knew our job, and complaints wouldn't bring change.* [newline] [newline] *"They*

#196: *We smiled at them and motioned to go ahead, feeling good that we showed them courtesy. Later in line, they made the "gun" finger-in-the-pocket signal. No big deal, we thought, just kids horse-playing. We swiftly did our banking and went to our car. But we didn't drive off. They passed by menacingly*

Native FIM / RIV

Suffix: *"We watched the current carry our little boat downstream until it disappeared."*

#85: *The little village we came across was quite unexpected, but at least we could spend the night in the tavern and make our way to the capital with daylight on our side.* [newline] [newline] *But despite Collin's objections, we got caught in the moment and decided to take a boat ride in the river. It was the old boatman's last day of work and*

#96: *They wanted to go home before the darkness descended, and there was an island that would be pitch black by seven pm. That meant crossing several streams, a village, and of course; the last of them - the most terrifying - was the river.* [newline] [newline] *"We can't let you cross this river, not in this darkness," said a helpful woman from the*

#99: *They were young, just a bit older than us, and they were working hard, clearly playing a game. They kept jumping into the river, swimming back to the shore, jumping in, over and over. They invited us to join, and said, "C'mon, if we're going to build this here raft we'll need some hard*

#196: *At that time of year, the river is little more than a collection of puddles. And with yam prices skyrocketing, these girls were taking no chances.*

The failed instrument: `gpt-4o-mini` calibration ($T = 1.0$, $n = 10$/horn)

The basis of the mode-collapse finding (§6): the instruction-tuned model's bridges, same horns as above. Note the near-identical sociable register and the weak uptake of the suffix constraint.

`gpt-4o-mini` / NEU

Suffix: *"By the time we said our goodbyes, the stars were already out."*

#5: *They were a group of friends, laughing and sharing stories as the day turned to dusk, their silhouettes casting long shadows on the water's surface.*

#6: *They were setting up a small campfire, laughter ringing in the air as they shared stories from their recent adventures.*

#8: *Their laughter echoed off the water as we shared stories, the sky slowly transitioning from hues of orange to deep navy.*

#9: *Their laughter echoed across the water as we shared stories and watched the sun dip below the horizon, painting the sky in shades of orange and purple.*

`gpt-4o-mini` / FIN

Suffix: *"When the manager finally unlocked the doors again, we were the last customers of the day."*

#5: *Their laughter echoed as they tossed pebbles into the water, the golden light dancing across the surface, and we were drawn to their carefree spirit.*

#6: *They were a small family, huddled close together, sharing stories of their day while casting long shadows on the grass. As we approached, their laughter filled the air, creating an inviting warmth that made us hesitate to leave.*

#8: *They were a friendly couple, eagerly sharing tales of their recent travels while we waited for the bank to reopen. As the sun dipped below the horizon, their laughter echoed in the cool evening air, making the short wait feel much longer.*

#9: *They had set up a small pop-up stall selling freshly made pastries, their laughter mixing with the gentle sounds of the river nearby. Intrigued, we couldn't resist trying a few treats while the sun dipped below the horizon, casting a golden glow over the water.*

`gpt-4o-mini` / **RIV**

Suffix: *"We watched the current carry our little boat downstream until it disappeared."*

#5: *Their laughter echoed as they splashed in the shallow water, while we secured our belongings and prepared to set off on our adventure.*

#6: *They were gathering their gear, laughter bubbling as they prepared for the evening's fishing trip. Excitement filled the air as they invited us to join them, sharing tales of past catches and the best spots along the river.*

#8: *They were preparing to set off on their own adventure, laughter echoing over the gentle ripples of the water.*

#9: *The laughter and chatter filled the air as we approached, the warm glow of the setting sun reflecting on the water's surface. As we got closer, we could see the familiar faces of our friends gathered around, ready to embark on the evening's adventure.*

References

[1] T. Altenkirch, C. McBride, W. Swierstra. Observational equality, now! *Proceedings of the ACM Workshop on Programming Languages meets Program Verification (PLPV)*, pp. 57–68, 2007.

[2] D. Annenkov, P. Capriotti, N. Kraus, C. Sattler. Two-level type theory and applications. *Mathematical Structures in Computer Science* 33(8):688–743, 2023. arXiv:1705.03307.

[3] J. Arnold, F. Holtorf, F. Schäfer, N. Lörch. Phase transitions in the output distribution of large language models. *Advances in Neural Information Processing Systems 37 (NeurIPS)*, 2024. arXiv:2405.17088.

[4] N. Asher. *Lexical Meaning in Context: A Web of Words*. Cambridge University Press, 2011.

[5] L. Berglund, M. Tong, M. Kaufmann, M. Balesni, A. Cooper Stickland, T. Korbak, O. Evans. The reversal curse: LLMs trained on "A is B" fail to learn "B is A". arXiv:2309.12288, 2023.

[6] M. Bavarian, H. Jun, N. Tezak, J. Schulman, C. McLeavey, J. Tworek, M. Chen. Efficient training of language models to fill in the middle. arXiv:2207.14255, 2022.

[7] S. Ayhan. What is the meaning of proofs? A Fregean distinction in proof-theoretic semantics. *Journal of Philosophical Logic* 50:571–591, 2021.

[8] S. Ayhan. Meaning and identity of proofs in a bilateralist setting: a two-sorted typed lambda-calculus for proofs and refutations. *Journal of Logic and Computation* 35(2):exae014, 2025.

[9] E. M. Bender, A. Koller. Climbing towards NLU: on meaning, form, and understanding in the age of data. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 5185–5198, 2020.

[10] B. Bentzen. Sense, reference, and computation. *Perspectiva Filosófica* 47(2):179–203, 2020. arXiv:2006.16450.

[11] E. Bigelow, A. Holtzman, H. Tanaka, T. Ullman. Forking paths in neural text generation. *International Conference on Learning Representations (ICLR)*, 2025. arXiv:2412.07961.

[12] E. Bishop, D. Bridges. *Constructive Analysis.* Grundlehren der mathematischen Wissenschaften 279, Springer, 1985.

[13] S. Chatzikyriakidis, Z. Luo. *Formal Semantics in Modern Type Theories.* ISTE/Wiley, 2020.

[14] B. Coecke, M. Sadrzadeh, S. Clark. Mathematical foundations for a compositional distributional model of meaning. *Linguistic Analysis* 36:345–384, 2010.

[15] D. Corfield. *Modal Homotopy Type Theory: The Prospect of a New Logic for Philosophy.* Oxford University Press, 2020.

[16] M. Duží, B. Jespersen, P. Materna. *Procedural Semantics for Hyperintensional Logic.* Springer, 2010.

[17] C. Donahue, M. Lee, P. Liang. Enabling language models to fill in the blanks. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2020.

[18] K. Fine. A theory of truthmaker content I: conjunction, disjunction and negation. *Journal of Philosophical Logic* 46:625–674, 2017.

[19] G. Frege. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100:25–50, 1892.

[20] D. Gratzer, J. Weinberger, U. Buchholtz. Directed univalence in simplicial homotopy type theory. arXiv:2407.09146, 2024 (v2 2026).

[21] J. Groenendijk, M. Stokhof. Dynamic predicate logic. *Linguistics and Philosophy* 14(1):39–100, 1991.

[22] I. Heim. *The Semantics of Definite and Indefinite Noun Phrases.* PhD dissertation, University of Massachusetts, Amherst, 1982.

[23] A. Heyting. *Intuitionism: An Introduction.* North-Holland, 1956.

[24] Y. Jiang, G. Rajendran, P. Ravikumar, B. Aragam, V. Veitch. On the origins of linear representations in large language models. *Proceedings of the 41st International Conference on Machine Learning (ICML)*, PMLR 235:21879–21911, 2024.

[25] A. Joyal. Quasi-categories and Kan complexes. *Journal of Pure and Applied Algebra* 175:207–222, 2002.

[26] H. Kamp. A theory of truth and semantic representation. In *Formal Methods in the Study of Language*, Mathematisch Centrum, Amsterdam, 1981.

[27] S. Kripke. *Naming and Necessity.* Harvard University Press, 1980.

[28] N. Kudasov, E. Riehl, J. Weinberger. Formalizing the ∞-categorical Yoneda lemma. *Proceedings of the 13th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP)*, 2024. DOI 10.1145/3636501.3636945.

[29] J. Ladyman, S. Presnell. Identity in homotopy type theory, part II: the conceptual and philosophical status of identity in HoTT. *Philosophia Mathematica* 25(2):210–245, 2017.

[30] B. M. Lake, G. L. Murphy. Word meaning in minds and machines. *Psychological Review* 130(2):401–431, 2023.

[31] X. Li, A. D. Sarwate. Unraveling the localized latents: learning stratified manifold structures in LLM embedding space. arXiv:2502.13577, 2025.

[32] S. Lossin. Cocartesian fibrations in synthetic simplicial type theory. arXiv:2604.18668, 2026.

[33] S. Mahadevan. GAIA: categorical foundations of generative AI. arXiv:2402.18732, 2024.

[34] Y. N. Moschovakis. A logical calculus of meaning and synonymy. *Linguistics and Philosophy* 29:27–89, 2006.

[35] R. Muskens. Sense and the computation of reference. *Linguistics and Philosophy* 28(4):473–504, 2005.

[36] K. Nakaishi, Y. Nishikawa, K. Hukushima. Phase transition in large language models and the criticality of natural languages. arXiv:2406.05335, 2026.

[37] J. Pustejovsky. *The Generative Lexicon.* MIT Press, 1995.

[38] D. Quigley. A vector logic for extensional formal semantics. *Journal of Logic, Language and Information* 34(5):557–599, 2025.

[39] N. Rasekh. Simplicial homotopy type theory is not just simplicial: what are ∞-categories? arXiv:2508.07737, 2025.

[40] F. Recanati. *Mental Files.* Oxford University Press, 2012.

[41] E. Riehl, M. Shulman. A type theory for synthetic ∞-categories. *Higher Structures* 1(1):147–224, 2017.

[42] M. Robinson, S. Dey, S. Sweet. The structure of the token space for large language models. arXiv:2410.08993, 2024.

[43] M. Robinson, S. Dey, T. Chiang. Token embeddings violate the manifold hypothesis. *Advances in Neural Information Processing Systems 38 (NeurIPS)*, 2025. arXiv:2504.01002.

[44] B. M. Ruppik et al. Less is more: local intrinsic dimensions of contextual language models. *Advances in Neural Information Processing Systems 38 (NeurIPS)*, 2025. arXiv:2506.01034.

[45] M. Schüle. On the semantics of large language models. *Intellectica* 81:15–36, 2024.

[46] L. Tranchini. Proof-theoretic semantics, paradoxes and the distinction between sense and denotation. *Journal of Logic and Computation* 26(2):495–512, 2016.

[47] A. S. Troelstra, D. van Dalen. *Constructivism in Mathematics: An Introduction*, vol. I. North-Holland, 1988.

[48] The Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics.* Institute for Advanced Study, 2013.

[49] F. Veltman. Defaults in update semantics. *Journal of Philosophical Logic* 25(3):221–261, 1996.

[50] J. von Plato. The axioms of constructive geometry. *Annals of Pure and Applied Logic* 76(2):169–200, 1995.

[51] Y. Zhang, et al. Latent trajectory dynamics in large language models: a manifold evolution framework (DMET). arXiv:2505.20340, 2025.