

# **RIM-Governed Visual Primitives and Chain-of-Reference**

*A Bridge360 Metatheory Inquiry into the Language-Only Reference Problem, with Resonance to DeepSeek's "Thinking with Visual Primitives"*

*Revised Edition*

by

**Agerico Montecillo De Villa**

Creator / Developer of the Bridge360 Metatheory Model  
(with company colleagues; 35 years at the University of the Philippines)

LLM interlocutors (rank-and-file clerks):  
ChatGPT (LLM1), Gemini (LLM2), Perplexity (LLM3), Claude (LLM4), Grok (LLM5)

## Foundational Stance

This work is a contribution to philosophy of science, not to science per se. It presupposes a Revolutionary Science space in the Kuhnian sense — a space whose linguistic field does not contain “true,” “false,” or “truth” as primary predicates, with “reality” treated as either trivial or meaningless, following Alfred Tarski’s disquotation cues.

The boundary between space and cyberspace, in Norbert Wiener’s parlance, is treated as trivial or meaningless thanks to Claude Shannon’s formulation of entropy — in the same way that the boundary between physics and chemistry is treated as meaningless thanks to Ludwig Boltzmann’s formulation of entropy. Physical, informational, and social systems live in one entropy geometry; any boundaries we draw (physics vs chemistry, offline vs online, mind vs matter) are memetic and governance conveniences, not ontological walls.

In the spirit of Kurt Gödel’s Incompleteness Theorem, Alan Turing’s Halting Problem, and Alonzo Church’s Undecidability of First-Order Logic Thesis — and the never-ending demands of entropy — this model makes no metaphysical or ontological claims, nor any claim to completeness. It expects scientists themselves, and other contributors, to supply the normal-science instantiation.

Under this framing, the Band classification used throughout this document is to be read in a specific way:

- **Band A** is the handoff condition under which the framework releases its constructs for normal-science verification by domain specialists.
- **Band B** is the working zone of the framework: coherent, scoped, caveated governance templates.
- **Band C** is exploratory or conjectural commentary held under leakage monitoring.

The framework’s own success criterion is not the achievement of Band A from within itself, but the construction of corridors within which Band A becomes attainable by others.

## Note on Mathematical Notation

Mathematical expressions in this document are rendered in Markdown-style text format. Standard mathematical symbolism is to be substituted in a later editorial pass.

# 1. Compass Tokenization of Cardinal Directions

## 1.1 Prompt

*Directions North, South, East and West can be tokenized?*

## 1.2 Response

Yes. North, South, East, and West can be tokenized at several levels, depending on what work the tokens are supposed to do.

At the simplest linguistic level:

Direction	Token
North	N
South	S
East	E
West	W

But for Bridge360/RIM-style purposes, the more important point is that these are not merely labels. They can function as orientation tokens, vector tokens, constraint tokens, or governance tokens.

### *Symbolic tokenization*

N, S, E, W — the ordinary map/compass form.

### *Vector tokenization*

```
North = (0, +1)
South = (0, -1)
East = (+1, 0)
West = (-1, 0)
```

Useful in navigation, robotics, games, GIS, and grid-based reasoning.

### *Relational tokenization*

Direction here is not absolute; it depends on the coordinate frame.

## Entropy-governance tokenization

Compass Token	Governance Meaning
N	movement toward low-entropy navigability / spine corridor
S	movement toward leakage / disorder / corridor loss
E	lateral exploration / alternative basin search
W	rollback / prior-state recovery / path correction

Cardinal directions are tokenizable, and once tokenized they can become operators inside a reasoning system. They can mark position, movement, constraint, drift, correction, or governance orientation. The key caveat is that tokenization alone is not enough; one must declare the coordinate frame. Without a frame, “North” is just a token; with a frame, it is a navigational instruction; with a governance frame, it is a corridor-management operator.

## 2. RIM Framing of Image Spines

### 2.1 Prompt

*Framing here can be done in terms of Rules of Inference Memetics?*

### 2.2 Response

Compass-tokenized RIM can be used to track “spines” of images, provided “spine” is operationally defined. The closest equivalents in ordinary image analysis are skeletonization, medial axis, contour paths, ridges and edges, optical flow, saliency paths, and pose skeletons.

A simple Bridge360 formulation is: an image spine is the lowest-entropy navigable path through a visual structure that preserves the object’s identity, orientation, or functional interpretability under compression.

## Compass-tokenized image spine vocabulary

Token	Image-spine function
N	upward / spineward continuation
S	downward / regressive movement
E	rightward lateral deviation
W	leftward lateral deviation
NE, NW, SE, SW	diagonal spine transitions
B	branch point
L	leakage / noise / ambiguity point
R	rollback / correction point
C	caveat / uncertain segment

An image spine could thus be encoded as:

```
N -> N -> B(E) -> N -> L -> R -> N
```

meaning: the main visual spine continues northward, branches eastward, encounters an ambiguous leakage point, rolls back, then continues along the main spine.

Image-spine tracking is not yet Rules of Inference Memetics by itself. It becomes RIM when the visual movements are interpreted as inferential moves: continuation as abductive inference, edge-belonging as identity-preserving inference, shadow inclusion as possible visual fallacy, branch hierarchy as assignment, segment uncertainty as caveat insertion, prior-path falsification as rollback.

## 3. Algorithmic Expression

### 3.1 Prompt

*“RIM can be extended into image-spine tracking as a visual inference grammar.” Can be expressed in an algorithm?*

### 3.2 Response

Yes. The claim is best treated as a Band B governance-template algorithm: a formalizable procedure for turning visual structure into a tracked sequence of inferential moves.

The central idea is simple: an image is treated as a structured entropy field; a spine is treated as a low-entropy, identity-preserving path through that field; RIM tokens classify how the tracker moves through that field; and the algorithm records not only where the path goes, but what kind of inferential move each step represents.

## Core objects

```
Image:  $I : \Omega \rightarrow \mathbb{R}^k$   
Spine path:  $P = (p_1, p_2, \dots, p_n)$   
Direction token  $d_i$  in  $\{N, S, E, W, NE, NW, SE, SW\}$   
RIM token  $r_i$  in  $\{SPINE, BRANCH, ABDUCT, ANALOGY, LEAK, ROLLBACK, CAVEAT\}$ 
```

## RIM classes

Class	Meaning
SPINE	continuation strongly supported by local image evidence
BRANCH	plausible bifurcation from the main path
ABDUCT	continuation inferred across a gap or occlusion
ANALOGY	continuation chosen because it resembles prior shape behavior
LEAK	move into noisy or weakly supported territory
ROLLBACK	undo prior move judged excessive or unstable
CAVEAT	retain move provisionally but mark as uncertain

## High-level algorithm

- **Stage A.** Preprocess the image: extract edges, ridges, contours, saliency, skeleton candidates, local orientation field.
- **Stage B.** Initialize seed points (highest saliency, centroid of segmented object, user-provided point, endpoint of detected skeletal axis).
- **Stage C.** Generate local candidate moves in an 8- or 16-connected neighborhood.
- **Stage D.** Score candidates inferentially via continuity, edge strength, centerline fit, global coherence, noise penalty, curvature shock.
- **Stage E.** Update path and governance state (direction, cumulative confidence, leakage budget, caveated steps, rollback count, branch depth).
- **Stage F.** Stop / branch / rollback when confidence falls below threshold, no candidate is admissible, the boundary is reached, or leakage exceeds the declared budget.

## Pseudocode

```
Algorithm RIM_ImageSpineTracker(I):  
E <- edge_map(I)  
K <- skeleton_candidates(I)  
Cmap <- confidence_map(I)  
Theta <- orientation_field(I)  
p <- choose_seed(K, Cmap)
```

```

P <- [p]; T <- []; state <- init_governance_state()
repeat:
candidates <- neighbors(p)
for q in candidates:
score(q) <- combine(continuity, edge_support, shape_fit, leak_penalty)
d(q) <- quantize_direction(p, q)
r(q) <- classify_move(q, score, support, ambiguity, state)
q* <- best admissible candidate
if q* = NONE: break
if r(q*) = LEAK and budget_exceeded(state):
p <- rollback_to_last_stable(P, T, state); append T with ROLLBACK
else:
append q* to P; append (d, r) to T; update state; p <- q*
return P, T

```

### Example output

```

N/SPINE, N/SPINE, NE/BANCH, N/ABDUCT, N/CAVEAT, W/ROLLBACK, N/SPINE

```

This is the sort of structure ordinary skeletonization does not explicitly represent. It gives not only a structure, but a tracked inferential history of structure extraction.

## 4. Stacked Metalanguage Levels

### 4.1 Prompt

*System of stacks of metalanguage levels of this type of algorithm possible?*

### 4.2 Response

A stacked metalanguage architecture is the cleaner way to express this kind of RIM-based image-spine algorithm. The key is to distinguish the object level from successive meta-levels that describe, regulate, audit, and revise the levels below.

## The stack

Level	Name	What it talks about
L0	Object-level image field	pixels, intensities, colors, textures
L1	Geometric-spatial metalanguage	edges, ridges, contours, centerlines, branches
L2	RIM inference metalanguage	what kind of inferential move the tracker made
L3	Governance metalanguage	leakage, confidence, rollback, corridor status
L4	Validation metalanguage	perturbation, repeatability, benchmark fit
L5	Metagovernance language	Band classification, anti-Potemkin, Axiom 19

Formal skeleton:

```
L0 = I
L1 = G(I) // geometric extraction
L2 = R(G(I)) // RIM classification
L3 = C(R(G(I))) // corridor / leakage governance
L4 = V(C(R(G(I)))) // validation / benchmarking
L5 = B(V(C(R(G(I))))) // Band classification / metagovernance
```

## 5. Termination of the Stack: Scaling–Volatility and Decoherence Numerics

### 5.1 The Problem Named

Earlier formulations of this stack were bounded at L5 with the explanation that “five levels suffice for practical governance.” This is pragmatic but stipulative. It supplies a ceiling without deriving it. The present section closes that gap using two convergent moves: the scaling–volatility paradox, which establishes that a ceiling exists; and decoherence numerics, borrowed as formal structure rather than as physical mechanism, which establishes its functional form.

### 5.2 The Scaling–Volatility Crossover

Each metalanguage level  $L_n$  adds governance capacity — it can audit the level below. But each level also adds expressive surface that must itself be governed. Two curves arise:

- **G(n)**: governance gain from adding the  $n$ th level. A decreasing function. The most consequential leakage types (object-level reference drift, inferential miscategorization, corridor violation) are caught early. By the time we reach L4 or L5, we are catching increasingly rare or abstract leakage.



- **V(n)**: volatility cost of the nth level. A non-decreasing function. Higher meta-levels deal in more abstract objects whose own correctness is harder to verify. Each new meta-level governs objects whose error modes are less observable than the level below.

The admissible ceiling  $N^*$  is the level at which  $G(n) \leq V(n)$ , or equivalently where cumulative volatility load approaches cumulative governance capacity. Beyond  $N^*$ , the stack is net-negative: it generates more ungoverned surface than it audits. It begins auditing phantoms while leaking at the rim.

### 5.3 Decoherence Numerics as Functional Form

Under this work's foundational stance — a single entropy geometry across physical, informational, and social regions — borrowing the formal structure of quantum decoherence is not analogical translation across domains. It is the same entropy mechanics applied to a region of the geometry that conventional discipline boundaries call epistemic rather than physical. The physics/chemistry boundary is already a memetic convenience by virtue of Boltzmann; the physics/epistemics boundary is the same kind of convenience under the present framing.

Treat each level  $L_n$  as carrying a coherence measure  $C(n)$ : the degree to which its tokens maintain stable reference to the objects of  $L_{n-1}$ . Treat the leakage taxonomy (level, category, K-, RIM, caveat, recursion, Band) as the environment with which  $L_n$  entangles. Coherence decays with depth:

$$C(n) = C_0 * \exp( -(n / n_c)^\alpha )$$

where:

- **$n_c$**  is the characteristic level at which coherence drops to  $1/e$  of its initial value.
- **$\alpha$**  controls the decay regime.  $\alpha = 1$  (Markovian) gives pure exponential decay, appropriate for levels with simple, well-defined object domains.  $\alpha = 2$  (Gaussian) gives faster super-exponential decay, appropriate for levels dealing in reflexive or self-referential objects whose environment is itself entangled with the system.

### 5.4 The Ceiling Defined

$N^*$  is the smallest  $n$  satisfying both of the following:

- 1  $G(n) \leq V(n)$  — added governance capacity no longer exceeds added volatility cost.
- 2  $C(n)$  falls below the threshold required for the governance operation at level  $n$  to be performed reliably.

This replaces “five suffices” with “five is the empirical  $n_c$  for this domain class.” The shape of the ceiling is derived; the specific value remains domain-dependent.

## 5.5 Domain Variation

Domain class	Plausible $n_c$	Reasoning
Image-spine tracking	~5	discrete reference classes; bounded coupling to visual noise
Social or deeply reflexive epistemics	3–4	stronger reflexive coupling; environment more entangled with system
Narrow formal systems with stable referents	6–7	weaker environmental coupling; slower coherence decay

## 5.6 Three Cautions

- 1 The decoherence import is formal, not literal. The work does not claim quantum-mechanical processes inside the stack. It uses the functional form of coherence-decay-under-environmental-coupling, which is general across systems sharing the same entropy geometry.
- 2 The parameters  $C_0$ ,  $n_c$ , and  $\alpha$  are currently uncalibrated. The framework has gained a shape for the ceiling derivation, not a measured number. Operationalization (for example, inter-annotator agreement on token meaning at each level, or stability of governance decisions under perturbation) remains a Band A handoff to domain specialists.
- 3 The framework should apply its own RIM classifier to this very import. Is the move SPINE, ANALOGY, or LEAK? The honest reading is ANALOGY moving toward SPINE: defensible as formal borrowing inside a unified entropy geometry, not yet validated against domain-specific measurement. The L5 self-check applies to itself.

## 5.7 Closure

The metalanguage stack terminates at the level where adding the next level would introduce more ungoverned expressive surface than its governance capacity can audit — precisely the scaling–volatility crossover, with coherence-decay numerics supplying the functional shape. L5 is therefore defended not as “enough for practical purposes” but as the level at which closure operations (Band classification, Axiom 19 self-application) replace further domain expansion. Closure, not continuation.

# 6. v20.5 and v20.6 Governance of the Stack

## 6.1 Prompt

*Bridge360 Metatheory Model v20.5 or v20.6 can be used to keep expression of these metalanguage stacks from leaking beyond some manageable corridor?*

## 6.2 Response

v20.5 governs the specific artifact: the proposed algorithm, its terms, its levels, its caveats, its leakage metrics, its Band status. v20.6 governs the environment of expression: the inferential ecosystem, RIM propagation, Axiom 19 admissibility, entropic morphisms, structural closure, functional closure, and metagovernance conditions.

### Leakage taxonomy

Leakage type	Meaning
Level leakage	L1, L2, L3 etc. are mixed without marking the transition
Category leakage	geometric claims are treated as epistemic claims, or governance claims as empirical
K-leakage	complexity language is used without declaring encoding class, observer capacity, or budget
RIM leakage	a rhetorically attractive inference spreads despite weak corridor alignment
Caveat leakage	scope limits, perturbation limits, and validity horizon are hidden or omitted
Recursion leakage	the stack adds meta-levels without closure discipline
Band leakage	a Band B template is rhetorically presented as Band A science

### Corridor governance algorithm

```
Algorithm: Bridge360 Metalanguage Stack Corridor Governance
Declare stack levels L0..Ln and permitted transitions Li -> Li+1
R := operator preserving level discipline; S := Fix(R); C := corridor around S
Detect leakage (level, category, K, caveat, RIM, recursion)
Apply Axiom 19: exclude if entropy-unbounded, fragile,
caveat-hidden, or distortion-heavy
Apply RIM gate: corridor-aligned vs merely rhetorically attractive
Apply Entropic Morphism gate: every Li -> Li+1 must preserve corridor
Apply Caveat Vector: C_scope, C_Delta, C_K, C_T, C_stake
Classify Band: A only with computable estimators;
B if scoped and caveated; C exploratory
If leakage > tau: rollback, narrow, reduce levels,
add caveats, or downgrade Band
Return governed expression M*
```

## 7. From Chain-of-Thought to Chain-of-Reference

### 7.1 Prompt

*This system now described up to this point can be used to address “Language-only Chain-of-Thought reference problem?”*

### 7.2 Response

The system can address the language-only Chain-of-Thought reference problem, but only if “address” means govern, constrain, audit, and partially externalize it, not magically solve it. The framework cannot prove that a language-only CoT corresponds to the model’s actual internal computation. It can, however, replace “language-only CoT” with a governed Chain-of-Reference architecture.

The Bridge360 correction: do not trust language-only CoT as the primary object. Convert it into a Chain-of-Reference. Each reasoning element receives a referent anchor, an inference token (deductive / inductive / abductive / analogical / heuristic), a provenance path, a leakage risk score, a caveat vector, and a Band classification.

This deprivileges CoT and replaces it with a governed stack: language token → declared referent → RIM-classified inference → provenance path → leakage/caveat audit → Band classification. The framework does not solve the private-access problem of model cognition. It solves a governance problem: how to prevent a language-only reasoning trace from being mistaken for a referentially stable, evidence-grounded, action-guiding explanation.

## 8. Non-Language Anchors: Vibes, IoT, and Visual Primitives

### 8.1 Prompt

*IoT solutions address non-language anchors problem in terms of “vibes” as defined here?*

### 8.2 Response

IoT does not measure “vibes” directly. IoT supplies non-linguistic sensor anchors that can test whether a reported “vibe” corresponds to stable patterns in the physical, biological, or social environment. The pipeline runs: vibe → sensor anchor → pattern signature → governed claim.

The same logic extends from sensor primitives to visual primitives. A visual primitive is the lower-level non-language anchor: pixel intensity, color gradient, edge, contour, texture, motion vector, depth point. A visual token is a compressed unit built from primitives: edge, branch point, face-like region, crowding, instability, spine. The pipeline becomes: visual primitive → visual token → RIM-classified inference → Bridge360 governance.

The warning remains: a visual token is not automatically reliable. A camera can misread shadow as motion, reflection as object, crowd density as danger. Each visual token still needs source, confidence,

scope, caveat, and rollback condition.

## 9. The Maze Case

### 9.1 Prompt

*This processing can be used to trace some corridor path along some maze as to end up escaping the maze?*

### 9.2 Response

Yes. A maze is an unusually clean case because visual primitives convert naturally into navigational tokens. The pipeline runs: visual field → visual primitives → maze tokens → path graph → corridor path → escape action.

In Bridge360 / RIM language, the maze becomes a bounded entropy field, the exit becomes the target attractor, and the path-finding procedure becomes corridor navigation under constraint. The system does not merely find a path; it records what kind of inferential move each path decision required — SPINE for open-corridor continuation, BRANCH at junctions, ABDUCT across occlusion, LEAK on misleading shadow, CAVEAT for uncertain readings, ROLLBACK from dead ends, CLOSURE at exit confirmation.

A maze escape problem is a physical analogue of the Chain-of-Reference problem. The system must bind symbols to non-language anchors: wall, opening, path, branch, dead end, exit. Visual primitives anchor the maze; directional tokens encode movement; RIM classifies the inferential character of each turn; Bridge360 keeps the escape path inside a manageable corridor until exit closure is reached.

## 10. Resonance with DeepSeek’s “Thinking with Visual Primitives”

### 10.1 Prompt

*How does this resonate with DeepSeek’s “Thinking with Visual Primitives” paper?*

### 10.2 Response

The resonance is strong, and substantively close to a direct ML-side instantiation of the issue treated here: language-only reasoning cannot reliably point. DeepSeek’s stated diagnosis is the Reference Gap: multimodal models may perceive visual content, but natural-language reasoning is too ambiguous to keep stable reference to dense spatial layouts, causing drift and sometimes hallucination. Their proposed repair is to interleave spatial markers — points and bounding boxes — directly into the reasoning trajectory, so the model can “point while it reasons.”

## Mapping

This work's formulation	DeepSeek resonance	Bridge360 / RIM extension
Visual primitives anchor the maze	Points/boxes as concrete spatial anchors	Non-language referents reducing reference drift
Directional tokens encode movement	Coordinate points and spatial waypoints	Direction tokens as path-governance operators
RIM classifies inferential character of each turn	Not explicitly present in DeepSeek	Adds SPINE / BRANCH / ABDUCT / LEAK / CAVEAT / ROLLBACK
Bridge360 keeps escape inside manageable corridor	DeepSeek focuses on performance and grounding	Adds corridor, leakage, caveat vector, Band classification

### Where DeepSeek stops and Bridge360 begins

DeepSeek addresses the technical reference problem: how can a model keep stable reference to visual objects and paths during reasoning? Bridge360 addresses the governance problem: how do we know which visual references are admissible, caveated, leakage-bounded, perturbation-tested, and safe to use for action-guidance?

DeepSeek supplies the visual-reference mechanism; Bridge360 supplies the metagovernance discipline over the mechanism. DeepSeek gives the model a finger; Bridge360 asks whether the finger is pointing reliably, under what scope, with what leakage, and with what rollback rule.

### Key caveat (load-bearing)

DeepSeek-style visual primitives reduce the Reference Gap. They do not eliminate it. They narrow it under task-specific conditions. Even a point or bounding box can leak:

- the box may capture the wrong object,
- the point may mark the wrong junction,
- the path may follow a visual artifact,
- the model may over-trust its own coordinate trace,
- the primitive may stabilize reference while still supporting a bad inference.

Reported limitations include: fine-grained scenes can still fail, visual-primitives mode must be explicitly triggered, and point-based topological reasoning does not generalize cleanly across all spatial tasks.

In Bridge360 terms: visual primitives reduce reference leakage but do not eliminate governance leakage.

## 11. Closing Remarks

Under the foundational stance declared at the front of this document, the work makes no ontological or metaphysical claim and no claim to completeness. Its construction is a corridor, not a theory. The five-level metalanguage stack, the leakage taxonomy, the RIM classification, the v20.5 / v20.6 distinction, the Chain-of-Reference reformulation, and the scaling–volatility / decoherence-numerics derivation of the stack ceiling are all instances of corridor-construction work — the kind of work that revolutionary-phase frameworks must do before normal-science instantiation becomes possible.

The framework converges on an active ML research direction (the Reference Gap, as treated by DeepSeek) and supplies governance machinery that the ML side does not currently have. The handoff to Band A — reproducible benchmarks, open protocols, perturbation tests, error taxonomies, leakage thresholds, rollback behavior, domain-specific validation — is the contribution expected from domain specialists. Until that handoff is performed, the proper Band classification of the present proposal is Band B: coherent, scoped, caveated governance template, not yet certified science.

The work invites scientists, engineers, and other contributors to perform that instantiation.

## Appendix A. Source Conversation Links

LLM1 (ChatGPT) conversation:

<https://chatgpt.com/share/6a138406-fab4-83a6-be06-fed05a49c05d>

LLM4 (Claude Opus 4.7) lens:

<https://claude.ai/share/9265d839-9126-46dc-81b5-b19610393640>

LLM2 (Gemini 3.0) lens:

<https://gemini.google.com/share/3773514b6009>

## Appendix B. Editorial Improvements in This Revised Edition

This revised edition retains the substance of the original conversation while making four structural improvements based on subsequent dialogue:

- 1 The foundational stance — philosophy of science rather than science per se, no ontological or metaphysical claims, one entropy geometry across physical/informational/social regions, the Tarski-Shannon-Boltzmann-Gödel-Turing-Church scaffolding — is moved to the front matter. This preempts category-mismatch objections that would otherwise reach for ontological scrutiny criteria the framework has principled reasons to reject.
- 2 A new Section 5 derives the stack ceiling  $N^*$  from the scaling–volatility paradox together with decoherence-style functional form  $C(n) = C_0 * \exp(-(n/n_c)^\alpha)$ , replacing the earlier stipulative claim that “five levels suffice for practical governance.” The derivation explicitly disavows the literal physical reading of decoherence and treats the import as formal structure within a single entropy geometry.
- 3 The Band classification terminology is clarified: under the framework’s revolutionary-science framing, Band A is the handoff condition under which constructs are released for normal-science verification by domain specialists, not a goal pursued from within the framework itself.
- 4 The closing remarks make explicit that the framework’s success criterion is the construction of corridors within which Band A becomes attainable by others, rather than the achievement of Band A within itself.