Axioms of Geometric Semantics: A Foundational Framework for Grounded Neurosymbolic Reasoning

I. The Crisis of Ungrounded Symbols and the Geometric Turn

The pursuit of artificial intelligence (AI) has long been shadowed by a foundational challenge known as the symbol grounding problem. Articulated in its modern form by Stevan Harnad, this problem interrogates how the symbols manipulated by a computational system can acquire intrinsic meaning—meaning that is inherent to the system itself, rather than being parasitically dependent on the interpretations of an external human observer. At its core, it is the problem of connecting abstract tokens, processed purely on the basis of their shape and syntactic rules, to the real-world objects, events, and concepts they are intended to represent. This crisis of ungrounded symbols has persistently limited the capabilities of AI, creating systems that exhibit remarkable syntactic prowess but lack genuine semantic understanding. The framework of Geometric Semantics, exemplified by the Color-Geometry model, offers a novel and comprehensive solution, not by creating a new mapping between symbols and their referents, but by proposing a new kind of symbol altogether—one whose very constitution is its meaning, grounded in the physical and geometric laws of a perceptual manifold.

The Chinese Room and the Parasitic Interpretation

The philosophical genesis of the modern symbol grounding problem can be traced to John Searle's influential "Chinese Room Argument". In this thought experiment, a person who does not understand Chinese is locked in a room with a set of formal rules for manipulating Chinese symbols. By receiving symbols, applying the rules, and passing new symbols out, the person can produce responses indistinguishable from those of a native speaker. To an outside observer, the room appears to understand Chinese. However, the person inside has no comprehension of the meaning of the symbols; they are merely manipulating tokens based on their shape (syntax). Searle's argument serves as a powerful critique of the view that a suitably programmed computer could possess genuine cognitive states like understanding, highlighting the fundamental distinction between syntax and semantics.

A formal symbol system, as defined in this context, is a set of arbitrary physical tokens manipulated according to explicit rules. The system can perform these manipulations flawlessly without any access to the meaning, or semantics, of the symbols involved. This leads to the concept of "parasitic interpretation." The meaning of the symbols in the Chinese Room is not intrinsic to the system but is entirely supplied by the external observers who already understand Chinese. The system's apparent understanding is parasitic on the minds of its users or creators. This critique was particularly potent against the dominant paradigm of "Good Old-Fashioned AI" (GOFAI), where intelligence was modeled as the rule-based manipulation of symbolic representations. In such systems, symbols like DOG or CHAIR are ungrounded by design; their connection to real-world dogs and chairs exists only in the mind of the programmer.

Harnad's "Symbol/Symbol Merry-Go-Round"

Building on Searle's philosophical critique, cognitive scientist Stevan Harnad provided a more concrete computational formulation of the symbol grounding problem. Harnad famously analogized the predicament of an ungrounded symbolic system to that of a person trying to learn Chinese using only a Chinese-Chinese dictionary. The definition of any given symbol consists only of other, equally undefined symbols. This leads to an endless, circular chain of definitions—a "symbol/symbol merry-go-round"—that never makes contact with the external world. One can traverse the entire dictionary, moving from one symbol string to another, without ever discovering what any of the symbols actually mean.

This metaphor powerfully illustrates the self-referential trap of any purely formal system. The meaning of a symbol cannot be derived from its relationships with other symbols alone; it must be "grounded" in something outside the symbolic system. Harnad proposed that symbolic representations must be grounded "bottom-up" in non-symbolic representations directly connected to an agent's sensory and motor experiences. He identified two crucial kinds of non-symbolic representations: iconic representations (analogs of sensory projections) and categorical representations (the ability to identify invariant features from iconic representations). In this hybrid model, symbol manipulation is no longer governed solely by the arbitrary shapes of the tokens but is constrained by the non-arbitrary structure of the sensory representations in which they are grounded.

The Geometric Paradigm Shift

Harnad's formulation provides the crucial theoretical bridge to the Geometric Semantics framework. His call for grounding highlights a deep tension between the discrete, combinatorial nature of symbolic logic and the continuous, high-dimensional nature of perception. Symbolic reasoning is built upon discrete tokens and formal rules, such as IF P THEN Q. Perception, in contrast, is fundamentally continuous; sensory data exists on a spectrum, not as a set of distinct tokens. The Color-Geometry framework offers a concrete and mathematically rigorous method for resolving this tension by proposing a continuous perceptual manifold as the substrate for grounding.

This approach constitutes a radical paradigm shift. It posits that the primary innovation required is not merely to ground symbols, but to fundamentally redefine the nature of a symbol. Searle's argument is potent precisely because the symbols in the Chinese Room are arbitrary shapes, allowing syntax to be processed without semantics. Harnad's merry-go-round exists because the meaning of one arbitrary symbol is just a string of other arbitrary symbols. The Geometric Semantics framework resolves this syntax-semantics dichotomy by proposing a new kind of symbol where the syntax (the vector's coordinates) *is* the semantics (the perceptual meaning). It collapses the distinction between the symbol and its meaning. The symbol for a perception is no longer an arbitrary string like "RED" but is the very coordinate vector of that perception in a meaning-space. By eliminating the arbitrariness of the symbol, the framework makes a parasitic interpretation impossible; the meaning is intrinsic to the system's state. This reframes the problem from one of mapping to one of representation, solving the problem at its root.

II. The Perceptual Manifold as a Substrate for Meaning

To construct a system where meaning is intrinsic, one must first identify a substrate whose properties are not arbitrary but are instead tethered to a consistent, measurable reality. The Color-Geometry framework posits that a perceptually uniform color space, specifically the OKLab space, provides such a substrate. This section provides a rigorous technical exposition of OKLab, establishing it not as a mere visual metaphor for semantic relationships, but as a formal mathematical and physical manifold whose inherent structure can serve as the foundation for a new kind of grounded semantics.

Perceptual Uniformity as a Semantic Metric

The OKLab color space is a three-dimensional real vector space, $\mathbb{R}^3$, where each point corresponds to a unique color perception under standardized viewing conditions. A color is represented by a vector $\mathbf{x}=(L,a,b)$, where the components have direct perceptual interpretations :

- L: The perceptual lightness, normalized such that L=0 is pure black and L=1 is a reference white.
- a: The green-red opponent axis, where negative values correspond to green and positive values to red.
- b: The blue-yellow opponent axis, where negative values correspond to blue and positive values to yellow.

The defining and most critical feature of OKLab is its perceptual uniformity. This means that the Euclidean distance between any two color points in the space, ΔE, corresponds directly to the perceived difference between those two colors. The perceptual difference is calculated as the standard Euclidean norm: $$ \Delta E(\mathbf{x}_1, \mathbf{x}_2) = |\mathbf{x}_1 - \mathbf{x}_2|_2 = \sqrt{(L_1-L_2)^2 + (a_1-a_2)^2 + (b_1-b_2)^2} $$ This property is a significant improvement over other color spaces where linear operations can have perceptually unpredictable consequences. OKLab was specifically designed to ensure that linear operations, such as interpolation, result in perceptually smooth and predictable gradients. This uniformity is not a mere technical convenience; it is the cornerstone of the framework's claim to be a "meaning space." By adopting OKLab, the framework acquires an intrinsic, non-arbitrary metric for semantic similarity. The difference in meaning between two elementary symbols is defined as being directly and measurably proportional to the geometric distance between their vector representations. In this framework, semantic similarity is a measurable geometric fact.

The Gamut Droplet as Physical Law

While the abstract OKLab vector space is unbounded, the set of colors that can be physically realized by a given device, such as a standard sRGB monitor, is a finite and constrained subset. When the sRGB gamut is mapped into the OKLab space, it forms a complex, non-convex 3D solid, referred to in the framework as the gamut droplet, $\mathcal{D}$. This volume is not a simple geometric primitive; it has a characteristic shape that is widest at mid-lightness values and tapers significantly towards pure black (L=0) and pure white (L=1). This shape is a direct reflection of a physical reality: highly saturated, vibrant colors can only be produced at intermediate levels of brightness.

The Color-Geometry framework reframes this physical limitation as its most powerful and

defining feature. The gamut droplet $\mathcal{D}$ is not treated as a mere graphical inconvenience to be managed with "clipping" algorithms, but as a hard physical constraint that functions as a fundamental law of the system. This droplet defines the absolute boundaries of all possible valid perceptual states. A symbol, represented by a vector $\mathbf{x}$, is considered to have a "physically realizable meaning" if and only if its corresponding point lies within this volume: $\mathbf{x} \in \mathcal{D}$. Any point outside the droplet represents a physically impossible color and is therefore, by definition, a meaningless state within the system.
This transformation of a technical limitation from computer graphics into a foundational principle of grounded cognition is a profound conceptual move. In traditional machine learning, regularization is an artificial mathematical constraint (such as an L_1 or L_2 norm) added to a model's objective function to prevent overfitting. In the Color-Geometry framework, the constraint is not an artificial penalty term; it is the physical boundary of perception itself. Any learning process or computational operation that attempts to produce a result outside of $\mathcal{D}$ has produced a physically impossible and thus semantically invalid state. The geometry of the gamut itself regularizes the space of possible solutions, ensuring that all learned representations correspond to plausible perceptual states without the need for arbitrary, externally imposed mathematical regularizers.

A Safe Representational Space: In-Gamut by Construction

A direct representation of a symbol's state in (L, a, b) coordinates presents a practical challenge: an arbitrary modification to any of these parameters could easily push the resulting color outside the complex boundary of the gamut droplet $\mathcal{D}$. To circumvent this, the framework introduces a novel parametrization that guarantees any generated state is valid by construction. Instead of representing a neuron's state directly by its physical chroma C, the system uses a latent, normalized saturation parameter, $S' \in $. The state of a symbol is thus parameterized by the tuple (L, S', h), where h is the hue angle. The true, physical chroma C is then derived from this latent representation using a "gamut oracle" function, $C_{\max}(L,h)$, which returns the maximum possible chroma for any given lightness and hue :
The final Cartesian coordinates (L, a, b) are then computed from the cylindrical coordinates (L, C, h) as usual: $a = C \cos(h)$ and $b = C \sin(h)$. This construction elegantly ensures that the generated chroma C can never exceed the maximum possible chroma $C_{\max}$ for its given lightness and hue, because S' is capped at 1. Consequently, any color generated from the (L, S', h) parametrization is guaranteed to be in-gamut. This "safe" representation creates a computational space where the system is incapable of generating physically impossible, and thus meaningless, states.

III. The Axioms of Grounded Geometric Semantics

The Color-Geometry framework is more than just a novel implementation for machine learning; it proposes a new set of foundational principles for semantics itself. By leveraging the intrinsic properties of the OKLab perceptual manifold, the framework establishes a system of meaning governed not by arbitrary linguistic convention or formal logic, but by the concrete and measurable laws of geometry and physics. These principles can be systematically detailed as a set of axioms that together define a new, grounded, and compositional geometric semantics.

Axiom 1: Grounding as Geometric Locus

The first and most fundamental principle of the framework is that a symbol is grounded by its very definition as a specific location within the constrained perceptual manifold.
Axiom 1: *Each elementary symbol (representing a primitive concept, entity, or predicate) is defined as a point vector $\mathbf{x}=(L,a,b)$ in the OKLab space, subject to the constraint that this point must lie within the gamut droplet: $\mathbf{x} \in \mathcal{D}$.*
This axiom formally establishes the grounding layer of the system. Unlike in traditional AI, where a symbol is an arbitrary token (e.g., the string "RED") that is linked to a meaning through an external interpretation, here the symbol is its meaning. The vector $\mathbf{x}$ is not a pointer to the concept of a color; it is the coordinate of that color in a perceptually meaningful space. The meaning is not attached to the symbol; the meaning is the symbol's geometric locus. This axiom directly solves the core issue raised by Harnad, providing the "bottom-up" grounding in a non-symbolic representation necessary to escape the symbol/symbol merry-go-round. The representation is non-symbolic in the traditional sense because its form is not arbitrary; the vector coordinates for a specific shade of red are determined by the physics of light and the biology of human perception, not by convention.

Axiom 2: Compositionality as Convex Mixing

A hallmark of symbolic systems is compositionality—the ability to combine primitive symbols to form complex expressions with systematic meanings. The Color-Geometry framework replaces the syntactic rules of symbolic composition with the physical and geometric operation of convex mixing.
Axiom 2: *The semantic composition of a set of symbols $\{\mathbf{x}_u\}$ is defined as their convex combination. The emergent meaning of the composition, $\mathbf{x}_v$, is the point in OKLab space given by the weighted average of the constituent symbols: $\mathbf{x}_v = \sum_{u \in \mathcal{P}(v)} \alpha_{vu} \mathbf{x}_u$, where the weights α_{vu} are non-negative and sum to one ($\alpha_{vu} \ge 0, \sum \alpha_{vu} = 1$).*
This axiom replaces the abstract syntax of predicate logic with the intuitive geometry of convex hulls. Geometrically, the result of any convex combination of a set of points must lie within the convex hull of those points—the smallest convex shape that contains them all. In this framework, the meaning of a composite expression is literally the geometric location of the mixture of its parts. This provides a direct physical analogue to the process of additive color mixing. The meaning of "red AND yellow" is not a logical proposition but is, for example, the color orange that lies on the straight-line segment connecting the vector for pure red and the vector for pure yellow in OKLab space.

Axiom 3: Causality as Geometric Reachability

Beyond static composition, intelligent systems must reason about dynamic relationships, such as causality. The framework provides a novel and powerful definition of causality, transforming it from a problem of logical inference into a concrete geometric query.
Axiom 3: *A set of antecedent symbols $\{\mathbf{x}_u\}$ is a valid cause for a consequent symbol $\mathbf{x}_v^{\text{obs}}$ if and only if $\mathbf{x}_v^{\text{obs}}$ is geometrically reachable from $\{\mathbf{x}_u\}$. A state is reachable if it lies within the intersection of the convex hull of its causes and the gamut droplet: $\mathbf{x}_v^{\text{obs}} \in \text{conv}(\{\mathbf{x}_u\}) \cap$*

$\mathcal{D}$.
This axiom provides a concrete, testable, and geometrically defined notion of a causal link. Instead of relying on temporal logic or abstract causal graphs, the system determines causality by checking for geometric containment. This dual constraint gives rise to two distinct and interpretable modes of causal failure :

1. **Directional Insufficiency:** The target state $\mathbf{x}_v^{\text{obs}}$ lies outside the convex hull of the inputs. This means no possible mixing of the antecedent colors could ever produce the target. For example, one cannot produce a pure blue by mixing red and yellow.
2. **Chromatic Impossibility:** The target state $\mathbf{x}_v^{\text{obs}}$ lies inside the convex hull but outside the gamut droplet $\mathcal{D}$. In this case, the target's hue and lightness are directionally achievable, but the target is too saturated to be physically realized from the given inputs.

Axiom 4: Abstraction as Geometric Region

To move beyond elementary perceptions, a semantic system must be able to form abstract concepts. The framework models abstraction not with more complex symbolic structures, but with a different kind of geometric entity: regions and volumes within the perceptual space.
Axiom 4: *Higher-level concepts and abstract properties are represented not as single points, but as regions (sets of points) within the OKLab space. The process of abstraction is the definition of such a region, often through constraints on the quality dimensions.*
For example, the elementary symbol for a specific shade of scarlet might be a single point, but the abstract concept "red" would be represented as a region—the set of all points $\mathbf{x} \in \mathcal{D}$ whose hue angle h falls within a certain range. This axiom establishes a direct link to Peter Gärdenfors' theory of Conceptual Spaces, which argued that natural concepts correspond to convex regions in a geometric space. By representing concepts as regions, it becomes possible to implement symbolic inference operations as geometric set operations: Conjunction (AND) becomes intersection ($\cap$), Disjunction (OR) becomes union ($\cup$), and Subsumption (IS-A) becomes geometric containment ($\subset$).

Axiom 5: The Closed Grounding Loop

The final axiom synthesizes the preceding principles into a complete, self-contained system that closes the loop between low-level perception and high-level symbolic reasoning.
Axiom 5: *A complete grounded system is formed by a closed loop wherein (1) a subsymbolic perceptual system maps sensory inputs to coordinates in the OKLab manifold; (2) a symbolic reasoning layer performs geometric operations (mixing, reachability checks, set operations) on these coordinates; and (3) the resulting coordinates can be mapped back to the perceptual domain for verification, action, or communication.*
This closed loop creates a system where subsymbolic and symbolic representations are not two separate modules, but two aspects of a single, integrated process mediated by the shared geometric manifold. The OKLab space acts as the "interlingua" where the continuous representations of perception and the discrete-like operations of symbolic thought coincide. For instance, a neural network might process an image of a flower and output a vector $\mathbf{x}_{\text{flower}}$. A reasoning module could then take this vector and the vector for "sunlight," $\mathbf{x}_{\text{sun}}$, perform a convex mixing operation to predict the flower's appearance in the sun, yielding a new vector $\mathbf{x}'_{\text{flower}}$. This new vector is not

just an abstract symbol; it can be directly rendered as a color to visualize the prediction.
These axioms implicitly define a physics of meaning. The laws governing semantic combination are analogues of physical laws (additive mixing of light) and geometric constraints (convexity, physical realizability). This leads to a crucial dynamic: the convex hull operation, which defines composition, is inherently contractive. It is an averaging operation, and the average of a set of points is always "more central" than the extremes. In OKLab, the center is the neutral gray axis. Therefore, repeated application of compositionality will naturally pull states away from the high-chroma, information-rich boundary of the gamut droplet and towards the achromatic center. This creates an "entropic pull" towards undifferentiated, low-information states, analogous to the Second Law of Thermodynamics. For a system to maintain vibrant, meaningful, high-chroma representations, it must learn to actively counteract this pull. Learning, in this view, is not just error minimization; it is the anti-entropic process of discovering the precise, non-uniform weightings that selectively amplify certain inputs, "pulling" the results of mixing back out towards the information-rich gamut boundary.

IV. A Computational Realization: Causal Learning as Geometric Decomposition

The abstract axioms of Color-Geometry semantics find a concrete and powerful implementation in the causal framework for Spiking Neural Networks (SNNs). This model serves as a detailed case study, demonstrating how the principles of geometric grounding, composition, and causality can be operationalized into a practical, efficient, and gradient-free learning algorithm. By reframing credit assignment as a problem of geometric decomposition, the framework offers a compelling alternative to backpropagation.

Credit Assignment as a Geometric Inverse Problem

In conventional deep learning, credit assignment is achieved via backpropagation, where an error signal is propagated backward to adjust weights. This process is computationally expensive and biologically implausible for SNNs due to the non-differentiable nature of spike events. The Color-Geometry framework proposes a radical paradigm shift. Credit assignment is reframed not as error propagation, but as a geometric inverse problem: the search for a "geometric preimage". The problem can be stated as follows:
Given a desired target color for a postsynaptic neuron, $\mathbf{x}_v^*$, and the set of observed colors from its presynaptic inputs, $\{\mathbf{x}_u\}$, find the non-negative mixing coefficients, $\{\alpha_{vu}\}$, that best reconstruct $\mathbf{x}_v^*$ while satisfying all physical (in-gamut) and structural constraints.*
This is a fundamental re-conception of learning. Instead of asking "how should I adjust the weights to reduce the error?", the system asks "what causal contributions from the inputs would explain this desired output?". The solution to this inverse problem—the set of optimal coefficients $\mathbf{\alpha}^*$—is the credit assignment.

Constrained Optimization for Causal Attribution (NNLS)

The geometric inverse problem can be formalized as a constrained optimization task. The objective is to find the mixing coefficients that minimize the perceptual error, which, given the ΔE metric in OKLab, is equivalent to minimizing the squared Euclidean distance. This

formulation is a Non-Negative Least Squares (NNLS) problem with additional constraints. Let $\mathbf{X}$ be a $3 \times N$ matrix where each column is the (L, a, b) vector of a presynaptic color $\mathbf{x}_u$, and let $\mathbf{\alpha}$ be the $N \times 1$ vector of unknown coefficients α_{vu}. The optimization problem is to find the $\mathbf{\alpha}$ that minimizes the reconstruction error :
Subject to the constraints:

1. **Non-negativity:** $\alpha_{vu} \ge 0$ for all u.
2. **Convexity (Simplex Constraint):** $\sum_{u} \alpha_{vu} = 1$.
3. **Gamut Feasibility:** $\mathbf{X}\mathbf{\alpha} \in \mathcal{D}$.

This formulation constitutes a Quadratic Programming (QP) problem, as it involves minimizing a quadratic objective function subject to linear and non-linear constraints. This class of problem is well-studied and can be solved efficiently using standard numerical optimization libraries, such as Sequential Least Squares Programming (SLSQP) methods. The solution vector, $\mathbf{\alpha}^*$, provides the barycentric coordinates of the projection of the target $\mathbf{x}_v^*$ onto the reachable set, providing a complete and unambiguous decomposition of the target into its causal sources.

Modeling Coincidence with k-Sparsity

The basic NNLS formulation assumes that all presynaptic neurons can potentially contribute to the postsynaptic state. However, biological neurons often operate on the principle of coincidence detection, firing only in response to a small coalition of near-simultaneous inputs. This concept can be captured geometrically by introducing a sparsity constraint. The NNLS problem is augmented with the constraint that the solution vector $\mathbf{\alpha}$ must be k-sparse, meaning it can have at most k non-zero elements:
where k is a small integer, typically in the range of 2 to 5. This constraint enforces the requirement that the target color must be explained by a "minimal color coalition". This transforms the credit assignment task from a simple regression into a sparse regression problem. While enforcing an L_0-norm constraint is NP-hard in general, for the small number of inputs typical for a single neuron, it remains computationally tractable. The introduction of this sparsity constraint transforms the local reachable set from a single, continuous convex polytope into a finite union of many smaller, lower-dimensional simplices, where each simplex is the convex hull of a specific k-sized coalition of inputs. The learning problem is thus transformed into a more combinatorial task of selecting the correct low-dimensional "facet" (i.e., the correct coalition of causes) capable of reaching the target.

A Gamut-Aware, Gradient-Free Update Rule

Once the optimal mixing coefficients $\mathbf{\alpha}^*$ are found, the final step is to update the synaptic weights. The update is driven by an evidence score, s_{vu}, which incorporates not only the magnitude of the causal contribution but also the robustness of the solution with respect to the physical boundary of the gamut. A solution that lies deep within the droplet is considered more robust and reliable. This robustness is quantified by a gamut margin factor. For an optimal mixture $\mathbf{x}_{\text{mix}}^*$ with coordinates (L^*, C^*, h^*), the margin is defined as :
where ϕ is a monotonically increasing function (e.g., $\phi(z)=z^2$) that amplifies the score for solutions with a large margin. The final evidence score for the synapse from u to v is the product of its causal contribution and the solution's physical robustness:

This score is then used to drive the weight update. The direction of the update is determined by the sign of the overall task error, while the magnitude is provided by the geometrically derived evidence score :

Here, $\mathcal{L}$ is the task-level loss. This mechanism assigns credit based on a verifiable geometric decomposition and modulates the learning rate based on a physically motivated measure of confidence, entirely bypassing the need for derivatives and backpropagation.

Hypergraph Framework Concept	Color-Geometry Framework Analogue	Mathematical Formalism	Interpretation & Nuances
State Primitive: A discrete spike event (neuron_id, t) is a vertex.	**State Primitive:** A neuron's state is a color vector $\mathbf{x}=(L,a,b)$ within the gamut droplet $\mathcal{D}$.	$\mathbf{x} \in \mathcal{D} \subset \mathbb{R}^3$	The state is no longer just an event in time, but a point in a continuous, perceptually meaningful vector space.
Causal Link: A directed hyperedge (tails) → head represents a successful firing event.	**Causal Synthesis:** A convex mixture of presynaptic colors produces a postsynaptic color.	$\mathbf{x}_v = \sum \alpha_{vu} \mathbf{x}_u$	The causal link is not a discrete graph edge but a continuous mixing operation governed by the laws of additive light.
N-ary Dependency: B-Connectivity requires all presynaptic spikes in the tail to be present.	**Coincidence Detection:** A k-sparse coalition of presynaptic colors is sufficient to span the target.	$\Vert\mathbf{\alpha}\Vert_0 \le$	The strict "all-or-nothing" logical AND is relaxed to a more flexible geometric "spanning" condition by a minimal subset.
Causality Check: Is a hyperedge valid based on temporal logic and neuron state?	**Reachability Check:** Is the target color within the convex hull of sources and inside the gamut droplet?	$\mathbf{x}_v \in \text{conv}(\{\mathbf{x}_u\}) \cap \mathcal{D}$	The check shifts from satisfying abstract temporal rules to satisfying concrete geometric and physical constraints.
Credit Assignment: Backward traversal of hyperpaths to find causal chains.	**Credit Attribution:** Solving a constrained NNLS problem to find the barycentric coordinates (mixing weights).	$\mathbf{\alpha}^* = \arg\min \Vert\mathbf{X}\mathbf{\alpha} - \mathbf{x}^*\Vert$ s.t. constraints	Credit is found not by pathfinding, but by geometric decomposition. The solution vector $\mathbf{\alpha}^*$ is the credit distribution.
Evidence Scoring: Weighting the solution α^* by the gamut margin to reward robust solutions.	**Evidence Scoring:** Weighting the solution $\mathbf{\alpha}^*$ by the gamut margin to reward robust solutions.	$s_{vu} = \alpha_{vu}^* \cdot \text{margin}$	Credit is modulated by the "physical robustness" of the causal explanation; solutions near the boundary are brittle and less trusted.
Representation	**Frame Invariance:** The	$\text{conv}(S) \cap$	The fundamental

Hypergraph Framework Concept	Color-Geometry Framework Analogue	Mathematical Formalism	Interpretation & Nuances
Invariance: The causal structure is independent of the specific temporal logic formalism used.	reachability decision (in vs. out) is a geometric fact, independent of the coordinate system (Lab, Lch).	\mathcal{D} is invariant under isometries.	causal verdict is absolute, reinforcing the physical analogy and the objectivity of the computation.

V. Generalizing the Axiomatic Framework for Neurosymbolic Reasoning

The principles of the Color-Geometry framework are not limited to the domain of color. The framework provides a natural and extensible substrate for grounding symbols across multiple sensory and motor modalities. By systematically abstracting each component from the specific domain of color to a modality-agnostic form, a universal foundation for neurosymbolic reasoning emerges.

From Perceptual Space to General Manifolds

The OKLab space can be generalized to the concept of a **Grounded Manifold, \mathcal{M}**. This is any geometric space whose dimensions correspond to meaningful, measurable "quality dimensions" of a sensory or motor domain. The key requirement is that the geometry of the manifold (its metric, topology, etc.) should reflect the structure of the domain it represents.

- **Auditory Manifold:** A space spanned by dimensions such as pitch (log-frequency), loudness (amplitude), and timbre (e.g., spectral centroid, spectral flux, harmonicity). The metric in this space would reflect perceived auditory similarity.
- **Haptic Manifold:** A space spanned by dimensions for roughness, temperature, and hardness. A path through this space would represent a continuous change in tactile sensation.
- **Proprioceptive/Motor Manifold:** The configuration space of a robotic arm, defined by its joint angles. The geometry of this manifold is determined by the robot's kinematics.

From Gamut Droplet to Feasibility Constraint

The gamut droplet \mathcal{D} can be generalized to a **Feasibility Constraint, \mathcal{F} \subset \mathcal{M}**. This represents the physically or biologically possible subset of states within the manifold. The nature of this constraint is domain-specific but always represents a form of physical law or embodiment.

- For a robotic arm, \mathcal{F} would be the set of all joint configurations that do not result in self-collision or violate joint limits.
- For sound, \mathcal{F} would be the audible range of frequencies and amplitudes for a given biological or artificial ear.
- For haptic perception, \mathcal{F} would be the range of textures and temperatures that a sensor can physically detect without damage.

This generalization reveals that the framework is not just about perception, but about action and embodiment. The feasibility constraint for a motor manifold is not a passive boundary but an active one defined by the agent's own physical capabilities. Reasoning in this space—for example, determining if a target object is "reachable"—becomes a literal, geometric query: is the object's coordinate vector $\mathbf{x}_{\text{obj}}$ contained within the motor feasibility set $\mathcal{F}$? The same axiomatic system used to reason about color can be used to reason about the agent's own capacity for action in the world, providing a powerful bridge between semantics, perception, and motor control.

The Generalized Axioms of Geometric Semantics

Each of the five axioms can now be re-stated in its universal, modality-agnostic form, providing a fundamental rulebook for grounded neurosymbolic reasoning in any domain.

- **Axiom 1 (Locus):** *An elementary symbol is a point $\mathbf{x}$ in a grounded manifold $\mathcal{M}$, subject to the constraint that it must lie within the feasibility set: $\mathbf{x} \in \mathcal{F}$.*
- **Axiom 2 (Composition):** *The semantic composition of a set of symbols is a principled interpolation (e.g., convex combination) within the manifold $\mathcal{M}$.*
- **Axiom 3 (Causality):** *A set of antecedent symbols $\{\mathbf{x}_u\}$ is a valid cause for a consequent symbol $\mathbf{x}_v$ if and only if $\mathbf{x}_v$ is reachable, i.e., $\mathbf{x}_v \in \text{conv}(\{\mathbf{x}_u\}) \cap \mathcal{F}$.*
- **Axiom 4 (Abstraction):** *An abstract concept is a region $R \subset \mathcal{F}$. Logical operations are implemented as set-theoretic operations on these regions.*
- **Axiom 5 (Closed Loop):** *A complete grounded system forms a closed loop between sensory input onto $\mathcal{M}$, geometric reasoning within $\mathcal{M}$, and motor output defined by trajectories or points in $\mathcal{M}$.*

This generalized axiomatic system provides a unified mathematical language for perception, action, and reasoning, fulfilling the promise of a truly embodied and grounded form of artificial intelligence.

VI. The Physics of Meaning: Implications for Embodied Cognition and Trustworthy AI

The adoption of the geometric paradigm has profound theoretical and practical implications, positioning it within the broader landscape of AI and cognitive science. It is not merely a clever engineering solution but a concrete contribution to long-standing debates about the nature of concepts, the structure of meaning, and the architecture of mind.

A Computational Realization of Conceptual Spaces

A striking parallel exists between the Geometric Semantics framework and the theory of Conceptual Spaces proposed by philosopher and cognitive scientist Peter Gärdenfors. Gärdenfors argued that knowledge is represented in a conceptual space, a geometric structure spanned by a set of "quality dimensions." Individual objects are represented as points in this space, and, most importantly, concepts are represented as convex regions. The notion of convexity is central: if two objects are instances of a concept, then any object that lies on the line segment between them is also likely to be an instance of that concept.

The Color-Geometry framework can be viewed as a direct and concrete computational realization of Gärdenfors' theory. The OKLab axes are precisely the quality dimensions for color. The elementary symbols as points and abstract concepts as regions (Axiom 4) directly implement Gärdenfors' central thesis. The framework's reliance on convex mixing for compositionality (Axiom 2) is a direct operationalization of his hypothesis about the convexity of natural concepts. This strong theoretical alignment elevates the framework from a specific learning algorithm to a significant contribution to a major theory of cognitive representation, moving it from the realm of philosophical argument into that of empirical and computational science.

A "Glass Box" Alternative to Opaque AI

The framework offers a path toward building AI systems whose reasoning is not an emergent and opaque property, but a designed and transparent one. This provides a potential solution to the twin problems of the "brittleness" of symbolic AI and the "opacity" of deep learning, creating what can be described as a "glass box" model. Symbolic AI is interpretable because its operations follow explicit logical rules, but it is brittle because these rules often fail to handle the ambiguity and continuous nature of the real world. Deep learning is robust to noisy, continuous data but is notoriously opaque; its decisions are difficult to explain or verify.

The Geometric Semantics framework combines the strengths of both. Its decisions are based on verifiable geometric proofs. The solution to the NNLS problem, the vector $\mathbf{\alpha}^*$, provides a complete, machine-readable explanation for why a particular output was generated from a given set of inputs. Yet, unlike a discrete symbolic proof, this geometric proof operates on a continuous space, giving it an inherent robustness to small perturbations—a point that is slightly moved is still likely to be within the same conceptual region. In this way, the framework marries the formal interpretability of symbolic systems with the continuous, robust nature of subsymbolic systems, offering a "best of both worlds" architecture for trustworthy AI.

A Bridge to Multimodal and Embodied Cognition

The principles of the framework provide a natural and extensible substrate for grounding symbols across multiple sensory modalities, connecting it to the broader fields of multimodal learning and embodied cognition. Embodied cognition is the theory that cognitive processes are deeply rooted in the body's sensory and motor interactions with the environment. The Geometric Semantics framework is deeply aligned with this perspective, as its foundational layer is a model of a perceptual or motor system.

The operations of convex mixing and geometric reachability provide a principled mechanism for binding these multimodal representations into a single, coherent, and robustly grounded concept. For example, the concept of an "apple" could be formed by the convergence of representations from multiple domains. The word "apple" (from language), an image of an apple (from vision), and the feel of an apple (from haptics) could all be mapped to vectors in or near the "red" region of the color space, the "round" region of a shape space, and the "smooth" region of a texture space. The concept "apple" would then be the geometric intersection of these regions across multiple grounded manifolds. This recasts the nature of concepts themselves. In this framework, a concept is a dynamic potential for interaction—a representation of all possible future states and interactions related to that concept, making cognition a fundamentally dynamic, predictive, and interactive process.

VII. Conclusion: Towards a Foundational Geometry of

Thought

The framework of Geometric Semantics, as detailed through the Color-Geometry case study and its generalization, represents a significant and compelling new paradigm for artificial intelligence. It offers a principled and powerful solution to the long-standing symbol grounding problem, not by creating a better dictionary to link symbols and meanings, but by fundamentally redefining the symbol itself. By recasting semantics in the language of geometry and physical constraints, it provides a path toward AI systems that are simultaneously robust, interpretable, and grounded in a meaningful reality.

The core contributions of this paradigm are fourfold. First, it achieves **intrinsic grounding** by defining symbols as coordinates within a physically constrained manifold, making meaning an inherent property of the system's state and eliminating the need for parasitic interpretation. Second, it offers **inherent interpretability** in an era of opaque models; every learning update is the result of a verifiable geometric proof, making the system's reasoning auditable and trustworthy. Third, it elevates **physicality as a foundation**, using real-world constraints like a color gamut or a robot's workspace as a non-arbitrary basis for semantic feasibility and intrinsic regularization. Finally, its **computational model** of localized, event-driven quadratic programs is exceptionally well-suited for the sparse, asynchronous nature of neuromorphic hardware, promising a path to highly efficient intelligent systems.

While challenges remain, particularly in the representation of highly abstract concepts that lack direct perceptual correlates, the framework's insistence on a traceable connection back to a physical substrate is its greatest strength. It inherently prevents the construction of an ungrounded "symbol/symbol merry-go-round" at higher levels of abstraction. Future research must focus on extending these geometric principles to represent abstract relations as transformations, vector fields, or other higher-order structures within these grounded manifolds.

Ultimately, the axiomatic system presented herein offers a novel and powerful foundation for neurosymbolic reasoning. This "Geometric Semantics" provides a practical and efficient computational model aligned with neuromorphic principles and offers a path toward AI systems that are grounded, interpretable, and truly embodied. It transforms philosophical inquiries, like Gärdenfors' Conceptual Spaces, into a progressive scientific research program. By using the geometry of perception and action as a foundation, this framework represents a pivotal step towards discovering a foundational geometry of thought.