A General Framework for Grounded Semantics: From Perceptual Manifolds to a Physics of Meaning

Introduction: The Geometric Turn in Grounded Semantics

The Crisis of Arbitrary Representation

The pursuit of artificial intelligence has been shadowed for decades by a foundational challenge known as the symbol grounding problem. First articulated in its modern form by Stevan Harnad, this problem questions 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 how abstract tokens, processed purely on the basis of their shape and syntactic rules, can connect 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 can exhibit remarkable syntactic prowess but lack genuine semantic understanding.

The philosophical genesis of this challenge 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, 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, highlighting the fundamental distinction between syntax and semantics. This leads to the concept of "parasitic interpretation," where the meaning of the symbols is not intrinsic to the system but is entirely supplied by the external observers who already understand the language.

Building on Searle's philosophical critique, cognitive scientist Stevan Harnad provided a more concrete computational formulation of the 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. This metaphor powerfully illustrates the self-referential trap of any purely formal system. Harnad proposed that to escape this trap, symbolic representations must be grounded "bottom-up" in non-symbolic representations that are directly connected to an agent's sensory and motor experiences. Despite this clear articulation of the problem, traditional solutions have fallen short. Connectionist models, while learning from data, often lack the systematic, compositional structure of thought, while hybrid systems that couple a neural network to a symbolic engine often create an arbitrary and ungrounded link between the two modules. The token "cup" produced by a network is still just an arbitrary string whose meaning is only established by the system's designer, failing to make the meaning intrinsic to the system's architecture.

Gärdenfors' Conceptual Spaces: A Theoretical Solution

A powerful theoretical framework for bridging this gap was proposed by the philosopher and cognitive scientist Peter Gärdenfors in his theory of Conceptual Spaces. Gärdenfors sought to create an intermediate, conceptual level of representation with a geometric structure, positioned between the symbolic and connectionist paradigms. In his theory, knowledge is represented in a conceptual space, a geometric structure spanned by a set of "quality dimensions". These dimensions correspond to the basic features by which objects and concepts can be compared, such as color, shape, weight, temperature, or pitch.

The theory is built on several central hypotheses. First, individual objects or stimuli are represented as points in this multidimensional space. Second, and most importantly, concepts are represented as **convex regions** within this space. The notion of convexity is critical: if two points are instances of a concept, then any point that lies on the line segment between them is also likely to be an instance of that concept. For example, if a specific lime green and a specific forest green are both instances of the concept "green," then any color on the perceptual path between them should also be considered "green". This convexity criterion allows for the natural emergence of prototypes, which can be interpreted as the focal points or centers of these conceptual regions. Finally, similarity is not an arbitrary relation but is defined as the metric distance between points in the space; the closer two points are, the more similar the objects they represent. Gärdenfors' theory thus provides a rich, geometrically-grounded framework for concept learning, induction, and semantics, offering a theoretical solution to the grounding problem by postulating that meaning is structured by the geometry of perception.

Thesis: A Computable, Physically-Grounded Realization

While Gärdenfors' theory provided a compelling philosophical and cognitive framework, it has historically lacked a concrete, operational implementation. This report details such a system, presenting a fully specified, physically-grounded, and computationally tractable realization of Conceptual Spaces theory. The central conceptual move of this framework is a radical one: it resolves the symbol grounding problem not by creating a better mapping between arbitrary symbols and their real-world referents, but by fundamentally redefining the nature of the symbol itself.

In this paradigm, the symbol's state—its coordinate vector in a perceptual manifold—*is* its meaning. The syntax (the vector's coordinates) is inseparable from the semantics (the perceptual experience the coordinates encode). By collapsing this distinction, the framework makes a parasitic interpretation impossible; the meaning is intrinsic to the system's state, solving the grounding problem at its root. The specific implementation in the domain of color, governed by the physical constraints of a display's color gamut, serves as a powerful case study. However, the core argument of this report is that this mechanism can be generalized into a universal principle: a **domain-defined feasibility constraint (\mathcal{F})**. This principle allows the axiomatic system derived from the color domain to be extended across sensory and motor modalities, enabling a unified theory of multimodal grounded semantics where meaning is defined as the causal reachability of grounded states within a physically constrained manifold.

An Axiomatic Framework for Geometric Semantics

The principles of the framework can be systematically detailed as a set of axioms that together define a new, grounded, and compositional geometric semantics. These axioms are presented here in their generalized, modality-agnostic form, establishing a foundational logic governed not by arbitrary convention but by the concrete and measurable laws of geometry and physics.

Foundational Definitions: Manifolds, Dimensions, and Constraints

Before stating the axioms, two foundational concepts must be defined. These concepts serve as the building blocks for the entire system, abstracting the specific properties of the color domain into universal principles.

- **The Grounded Manifold ($\mathcal{M}$):** A grounded manifold is a geometric space, typically a real vector space $\mathbb{R}^n$, whose dimensions correspond to meaningful, measurable "quality dimensions" of a sensory or motor domain. These dimensions are the fundamental features by which stimuli in that domain can be compared, such as pitch and loudness in audition, or roughness and hardness in touch. The key requirement is that the geometry of the manifold—its metric, topology, and other structural properties—should reflect the structure of the domain it represents. For instance, the distance between two points in the manifold should correspond to the perceived similarity between the stimuli they represent.
- **The Feasibility Constraint ($\mathcal{F}$):** The feasibility constraint is the subset of the grounded manifold, $\mathcal{F} \subset \mathcal{M}$, that represents all physically or biologically possible states within that domain. This is the core generalization of the "gamut droplet" from the color-geometry implementation. The nature of this constraint is domain-specific but always represents a form of physical law or embodiment. For a perceptual system, it defines the boundaries of what can be sensed (e.g., the audible range of frequencies); for a motor system, it defines the boundaries of what can be done (e.g., a robot's reachable workspace). This constraint is not an artificial construct but a hard boundary imposed by physics or biology, and it serves as the ultimate arbiter of what can and cannot be represented meaningfully within the system.

The Five Axioms of Grounded Semantics

With these definitions in place, the five core axioms of the framework can be stated in their universal form.

- **Axiom 1 — Grounding as Geometric Locus:** An elementary symbol is defined as a point vector $\mathbf{x}$ in a grounded manifold $\mathcal{M}$, subject to the constraint that this point must lie within the feasibility set: $\mathbf{x} \in \mathcal{F}$. This axiom formally establishes the grounding layer. Unlike in traditional AI, where a symbol is an arbitrary token linked to a meaning through external interpretation, here the symbol *is* its meaning. The vector $\mathbf{x}$ is not a pointer to a concept; it is the coordinate of that concept in a meaningful space. The meaning is not attached to the symbol; the meaning is the symbol's geometric locus.
- **Axiom 2 — Compositionality as Principled Interpolation:** The semantic composition of a set of symbols is defined as a principled interpolation within the manifold $\mathcal{M}$. In the case of vector spaces, this is most naturally realized as a convex combination. The

emergent meaning of a composition, $\mathbf{x}_v$, is the point given by the weighted average of its constituents: $\mathbf{x}_v = \sum_{u} \alpha_{vu} \mathbf{x}_u$, where the weights are non-negative and sum to one. This axiom replaces abstract syntactic rules with the intuitive geometry of convex hulls. The meaning of a composite expression is literally the geometric location of the mixture of its parts.

- **Axiom 3 — Causality as Geometric Reachability:** 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 geometrically reachable from $\{\mathbf{x}_u\}$. A state is defined as reachable if it lies within the intersection of the convex hull of its causes and the feasibility set: $\mathbf{x}_v \in \text{conv}(\{\mathbf{x}_u\}) \cap \mathcal{F}$. This provides a concrete, testable, and geometrically defined notion of a causal link, transforming causality from a problem of logical inference into a geometric query.
- **Axiom 4 — Abstraction as Geometric Region:** A higher-level abstract concept is represented not as a single point, but as a region R within the feasibility set, $R \subset \mathcal{F}$. The process of abstraction is the definition of such a region, often through constraints on the quality dimensions. This axiom directly implements Gärdenfors' central thesis that natural concepts correspond to convex regions in a conceptual space. By representing concepts as regions, it becomes possible to implement logical operations as geometric set-theoretic operations: conjunction (AND) becomes intersection ($\cap$), disjunction (OR) becomes union ($\cup$), and subsumption (IS-A) becomes geometric containment ($\subseteq$).
- **Axiom 5 — The Closed Grounding Loop:** A complete grounded system is formed by a closed loop wherein (1) a subsymbolic perceptual system maps sensory inputs to coordinates in a manifold $\mathcal{M}$; (2) a symbolic reasoning layer performs geometric operations on these coordinates; and (3) the resulting coordinates can be mapped back to the perceptual or motor domain for verification, action, or communication. This closed loop creates a system where subsymbolic and symbolic representations are not separate modules but two aspects of a single, integrated process mediated by the shared geometric manifold.

A "Physics of Meaning"

Taken together, this axiomatic system is more than just a set of computational rules; it implicitly defines a self-contained "physics of meaning." This is not a mere metaphor. The laws governing semantic combination are direct analogues of physical laws and geometric constraints, giving rise to a predictable and lawful system dynamic.

The first law of this physics is established by the Feasibility Constraint, $\mathcal{F}$. It acts as a **conservation law**, defining the space of all possible valid states. Any operation within the system must conserve "meaningfulness" by producing a result that remains within this physically-defined boundary. Any state outside of $\mathcal{F}$ is, by definition, impossible and therefore meaningless.

The second law emerges from the nature of compositionality defined in Axiom 2. The convex combination operation, $\mathbf{x}_v = \sum \alpha_{vu} \mathbf{x}_u$, is inherently contractive. It is an averaging function, and the average of a set of points is always more "central" than the extremes from which it was formed. In any grounded manifold, this creates a dynamic where repeated application of compositionality will naturally pull states away from the information-rich boundary of the feasibility set $\mathcal{F}$ and towards the less differentiated center (e.g., the neutral gray axis in the color manifold). This creates a powerful **"entropic pull"** towards undifferentiated, low-information states, a dynamic analogous to the Second Law of Thermodynamics.

This leads to a profound re-conception of the nature of learning. For a system to maintain vibrant, meaningful, high-information representations in its deeper layers, it must learn to actively counteract this entropic collapse. Learning, in this view, is not merely error minimization; it is the **anti-entropic process** of discovering the precise, non-uniform weightings that selectively amplify certain inputs, actively "pulling" the results of mixing back out towards the information-rich boundary of $\mathcal{F}$. This reframes learning as a dynamic struggle against informational entropy within a constrained physical system, a perspective that will recur throughout this analysis.

Case Study: The Color-Geometry Implementation

To instantiate the generalized axioms and demonstrate their practical application, this section presents a comprehensive case study of the framework's implementation in the domain of human color perception. This specific realization serves as the foundational example from which the general principles were abstracted.

The Perceptual Manifold as Semantic Substrate ($\mathcal{M}_{color}$)

The grounded manifold for this implementation, $\mathcal{M}_{color}$, is the OKLab color space, a modern, perceptually uniform three-dimensional real vector space, $\mathbb{R}^3$. Each point in this space 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 are its quality dimensions with 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 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, a profound departure from traditional symbolic systems where such relationships are either undefined or must be specified externally in an ad-hoc knowledge base.

The Feasibility Constraint as Physical Law ($\mathcal{F}_{color} = \mathcal{D}$)

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 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 but as a fundamental **law of semantic feasibility**. 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 physical boundary becomes the ultimate arbiter of what can and cannot be represented. The condition can be formalized by defining a "gamut oracle" function, $C_{\max}(L,h)$, which returns the maximum possible chroma (saturation) for any given lightness L and hue angle h. The gamut droplet $\mathcal{D}$ is then the set of all points (L, a, b) such that their chroma, $C = \sqrt{a^2+b^2}$, is less than or equal to this maximum possible value :

$$\mathcal{D} = \{ \mathbf{x}=(L,a,b) \in \mathbb{R}^3 \mid \sqrt{a^2+b^2} \le C_{\max}(L, \operatorname{atan2}(b, a)) \}$$

This physical constraint provides a form of **intrinsic, non-symbolic regularization**. 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 this framework, the constraint is not an artificial penalty term; it is the physical boundary of perception itself. Any learning process 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.

Geometric Semantics in Practice

With the manifold and feasibility constraint defined, the generalized axioms can be instantiated with the specific formulas of the color domain.

- **Composition (Axiom 2):** The composition of a set of color vectors $\{\mathbf{x}_u\}$ is their convex combination, $\mathbf{x}_v = \sum \alpha_{vu} \mathbf{x}_u$. 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.
- **Causality (Axiom 3):** A set of antecedent colors $\{\mathbf{x}_u\}$ is a valid cause for an observed consequent color $\mathbf{x}_v^{\text{obs}}$ if and only if $\mathbf{x}_v^{\text{obs}} \in \text{conv}(\{\mathbf{x}_u\}) \cap \mathcal{D}$. This dual constraint gives rise to two distinct and, crucially, interpretable modes of causal failure :
    1. **Directional Insufficiency:** Occurs when $\mathbf{x}_v^{\text{obs}}$ lies outside the convex hull of the inputs. The input palette is fundamentally insufficient; for example, one cannot produce a pure blue by mixing red and yellow.

   2. **Chromatic Impossibility:** Occurs when $\mathbf{x}_v^{\text{obs}}$ lies inside the convex hull but outside the gamut droplet $\mathcal{D}$. The target's hue and lightness are directionally achievable, but the target is too saturated to be physically realized from the given inputs.
- **Abstraction (Axiom 4):** Abstract concepts are represented as regions. For example, the abstract concept "red" is represented as the region of all points within the gamut whose hue angle falls within a certain range: $C_{red} = \{\mathbf{x} \in \mathcal{D} \mid h(\mathbf{x}) \in [h_1, h_2]\}$. Logical operations on these concepts become set operations on their corresponding regions: $C_{A \wedge B} = C_A \cap C_B$.

This direct, computational instantiation of geometric principles provides a powerful validation of Gärdenfors' original theory, moving it from philosophical argument to empirical science. The alignment between the theoretical concepts and their computational equivalents is made explicit in the following table.

Table 1: Alignment with Gärdenfors' Conceptual Spaces

Gärdenfors' Concept	Color-Geometry Implementation Equivalent
Quality Dimensions	OKLab perceptual axes (L, a, b)
Similarity	Euclidean distance (ΔE) in OKLab
Concepts as Convex Regions	Abstraction via region formation (Axiom 4)
Prototypes	Centers of conceptual regions (e.g., Voronoi centers)
Compositionality	Convex mixing of state vectors (Axiom 2)
Induction / Learning	Learning the boundaries of conceptual regions

A Computational Realization: Causal Learning as Geometric Decomposition

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

Credit Assignment as a Geometric Inverse Problem

In conventional deep learning, credit assignment is achieved via backpropagation, where an error signal (gradient) is propagated backward through the network to incrementally adjust synaptic weights. This process, while immensely successful, is computationally expensive and biologically implausible, and it is fundamentally incompatible with the discrete, non-differentiable nature of spiking neurons. The geometric framework proposes a radical paradigm shift. Credit assignment is reframed not as error propagation, but as a **geometric inverse problem**.

The problem is 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 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. It is not an infinitesimal update direction but a complete, holistic decomposition of the target state into its constituent causal sources.

Formalization and Solution (NNLS/QP)

This geometric inverse problem can be formalized as a constrained optimization task that is both mathematically precise and computationally tractable. The objective is to find the mixing coefficients that minimize the perceptual error between the synthesized color and the target color, which, given the ΔE metric in OKLab, is equivalent to minimizing the squared Euclidean distance.

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 is a Non-Negative Least Squares (NNLS) problem with additional constraints. The full problem, including the non-linear gamut feasibility constraint, 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.

Refinements for Plausibility and Robustness

The basic formulation can be refined to better model biological principles and to create a more robust learning signal.

- **k-Sparsity:** Biological neurons often operate on the principle of coincidence detection, firing only in response to a small coalition of near-simultaneous inputs. This 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: $\|\mathbf{\alpha}\|_0 \le k$, where k is a small integer. This enforces the requirement that the target color must be explained by a "minimal color coalition". This transforms the local reachable set from a single convex polytope into a finite union of many smaller, lower-dimensional simplices, creating a more structured, combinatorial landscape that is a closer geometric analogue to discrete logical reasoning.
- **Gamut-Aware 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 that incorporates not only the magnitude of the causal contribution but also the physical robustness of the solution. A solution that lies deep within the gamut droplet is considered more reliable than one that lies perilously close to the boundary of chromatic impossibility. This robustness is quantified by a gamut margin factor. For an

optimal mixture $\mathbf{x}_{\text{mix}}^*$ with chroma C^*, the margin is defined as: where ϕ is a monotonically increasing function that amplifies the score for solutions with a large margin. The final weight update is then modulated by this 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.

Learning as Geometric Proof

This computational model represents a fundamental shift in the nature of learning and reasoning in neural systems. A standard learning step in deep learning via backpropagation answers the question: "In which direction should I move the weights to reduce the error?" The answer is the gradient, an infinitesimal vector pointing towards a local improvement. In contrast, a learning step in this framework answers a different, more structured question: "Is the target state causally constructible from the input states, and if so, what is the exact recipe for its construction?"

The QP solver attempts to find a constructive **geometric proof** for this proposition. The solution, the vector $\mathbf{\alpha}^*$, is the proof itself—a complete, explicit recipe for the construction of the target from the inputs. If no solution exists that satisfies the constraints, the proof fails, indicating that the target is unreachable. The update rule then applies credit based not on an opaque error signal, but on the "quality" of this proof, as measured by the physical robustness or margin term.

This makes the system a **"glass box"** architecture. Each step of reasoning is auditable and interpretable. One can inspect the vector $\mathbf{\alpha}^*$ and the margin to get a complete, machine-readable explanation for why credit was assigned to a particular connection. This process marries the formal interpretability of symbolic systems, which operate on explicit rules, with the continuous, robust nature of subsymbolic systems, which operate on geometric spaces.

This approach has a clear intellectual lineage, evolving from earlier work on causal learning in SNNs based on temporal hypergraphs, as shown in the table below.

Table 2: Isomorphism between Temporal Hypergraph and Color-Geometry Models

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	**Coincidence Detection:** A k-sparse coalition of presynaptic	$\|\mathbf{\alpha}\|_0 \le$	The strict "all-or-nothing" logical AND is relaxed to a

Hypergraph Framework Concept	Color-Geometry Framework Analogue	Mathematical Formalism	Interpretation & Nuances
the tail to be present.	colors is sufficient to span the target.		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.

Extensibility: Grounding Semantics Across Perceptual and Motor Modalities

The principles of the Color-Geometry framework, while demonstrated in a single domain, are not limited to vision. The axiomatic system provides a natural and extensible substrate for grounding symbols across multiple sensory and motor modalities. By systematically identifying the appropriate grounded manifold ($\mathcal{M}$) and feasibility constraint ($\mathcal{F}$) for each domain, a unified architecture for embodied cognition emerges.

The Auditory Manifold ($\mathcal{M}_{audio}$)

The domain of auditory perception can be mapped to a geometric space with its own unique structure and constraints, forming an auditory manifold, $\mathcal{M}_{audio}$.

- **Quality Dimensions:** Based on extensive psychoacoustics research, the primary dimensions of auditory sensation are **pitch**, **loudness**, and **timbre**. Pitch corresponds to the logarithm of the sound's fundamental frequency, while loudness corresponds to its amplitude, typically measured on a logarithmic decibel (dB) scale. Timbre, the quality that distinguishes different musical instruments, is itself a complex sub-manifold. Its dimensions can be characterized by features such as **spectral centroid** (perceived brightness), **spectral flux** (textural change over time), and **roughness** or **harmonicity** (the degree of consonance in the sound's partials). A point in this multidimensional space represents a unique, perceivable sound.
- **Feasibility Constraint ($\mathcal{F}_{audio}$):** The "auditory droplet" is defined by the hard biological limits of the human auditory system. This feasibility constraint, $\mathcal{F}_{audio}$, is a bounded volume in the manifold. Its boundaries are determined by:
    1. **Frequency Range:** The range of human hearing, approximately from 20 Hz to 20,000 Hz, defines the limits along the pitch dimension.

    2. **Amplitude Range:** The range from the absolute threshold of hearing (the quietest sound that can be detected, defined as 0 dB SPL) to the threshold of pain (around 120-140 dB SPL) defines the limits along the loudness dimension.
    3. **Equal-Loudness Contours:** The shape of this volume is not a simple rectangle. Human sensitivity to loudness varies significantly with frequency, as described by the equal-loudness contours (ISO 226). The auditory system is most sensitive in the 2-4 kHz range, meaning the floor of the feasibility volume (the threshold of hearing) is lowest in this region. This non-linear boundary shapes the geometry of all possible auditory experiences.

The Haptic Manifold ($\mathcal{M}_{haptic}$)

The sense of touch, or haptic perception, provides information about the material properties of objects and can be modeled with its own grounded manifold, $\mathcal{M}_{haptic}$.

- **Quality Dimensions:** Psychophysical studies have identified several prominent dimensions of tactile perception. The most consistently reported dimensions are **roughness/smoothness**, **hardness/softness** (compliance), and **temperature** (warmness/coldness). Further research suggests that the roughness dimension can be subdivided into **macro-roughness** (processed spatially) and **fine-roughness** (processed via vibration), and that a distinct **friction** dimension (related to stickiness/slipperiness) also exists. A point in this haptic manifold represents a specific tactile texture.
- **Feasibility Constraint ($\mathcal{F}_{haptic}$):** The "haptic droplet" is defined by the physical limits of real-world materials and the physiological limits of the human sensory system. This feasibility constraint, $\mathcal{F}_{haptic}$, is bounded by:
    1. **Material Properties:** The finite range of physical properties found in nature. For example, the hardness dimension is bounded by materials that are maximally compliant (like gases) and those that are effectively infinitely rigid from a human perspective. Thermal conductivity has a finite range across known materials.
    2. **Physiological Limits:** The operational range of mechanoreceptors and thermoreceptors in the skin. Temperatures outside a certain range are not perceived as warm or cold but as pain, defining a hard boundary on that dimension. Similarly, the ability to discriminate compliance follows Weber's law with a Weber fraction of approximately 15%, which establishes a fundamental resolution or "granularity" for that dimension of the space. The human finger can distinguish features with amplitudes as small as 10 nm, setting a lower bound on the roughness scale.

The Motor Manifold ($\mathcal{M}_{motor}$)

The principles of geometric grounding are not limited to perception; they extend naturally to action and motor control. For an embodied agent like a robot, its capacity for action can be described by a motor manifold, $\mathcal{M}_{motor}$.

- **Quality Dimensions:** For a robotic manipulator, the dimensions of its state space are defined by its kinematics. The most fundamental dimensions are the **joint angles** (θ), which define the robot's configuration or posture. The dynamic state of the robot also includes the first and second time derivatives of these angles: the **joint rates** or velocities ($\dot{\theta}$) and the **joint accelerations** ($\ddot{\theta}$). A point in this high-dimensional configuration space, often called the C-space, represents the complete kinematic state of the robot at an instant in time.

- **Feasibility Constraint (\mathcal{F}_{motor}):** The "motor droplet" is the robot's valid operational space, a volume in its state space defined by hard physical and engineering constraints. This feasibility constraint, \mathcal{F}_{motor}, is bounded by:
  1. **Joint Limits:** Each joint has a limited range of motion (minimum and maximum angles), defining a hyperrectangle in the configuration space.
  2. **Velocity and Acceleration Limits:** Actuators have maximum speeds and accelerations, imposing bounds on the \dot{\theta} and \ddot{\theta} dimensions.
  3. **Workspace and Collision Avoidance:** The robot's end-effector is limited to a finite reachable workspace. Furthermore, the robot must avoid configurations that result in self-collision or collision with obstacles in the environment. These constraints carve out complex, non-convex regions from the state space, defining the final shape of the feasibility set.

The Feasibility Constraint as a Universal Principle of Embodiment

This generalization across domains reveals a profound unifying principle. The "droplet" concept, initially derived from the technical limitations of a computer monitor, is not specific to perception. It is a universal geometric representation of an agent's **embodiment**—the set of physical constraints on its ability to sense and act in the world.

In the color domain, \mathcal{D} represents the limits of a physical device (the monitor). In the auditory and haptic domains, \mathcal{F} represents the limits of a biological sensorium (the ear, the skin). In the motor domain, \mathcal{F} represents the limits of a physical body (the robot's kinematic chain). In every case, the feasibility constraint \mathcal{F} defines the absolute boundary between what is possible and what is impossible for that specific, embodied agent.

This insight unifies perception and action under a single geometric formalism. Reasoning about causality, as defined in Axiom 3, becomes a universal test for embodied plausibility. For example, a planned motor action is "causally valid" if the trajectory of target states lies entirely within the reachable set inside the motor feasibility constraint, \mathcal{F}_{motor}. This provides a powerful and direct bridge between high-level semantics, low-level perception, and real-world motor control, all within a shared mathematical language. The following table summarizes this generalization.

Table 3: Generalization of the Geometric Framework Across Modalities

Modality	Grounded Manifold (\mathcal{M})	Example Quality Dimensions	Feasibility Constraint (\mathcal{F})
Vision (Color)	OKLab Perceptual Space	Lightness (L), Red-Green (a), Blue-Yellow (b)	sRGB Gamut Droplet (\mathcal{D})
Audition	Psychoacoustic Space	Pitch (log-freq), Loudness (dB), Timbre (spectral centroid, flux)	Audible range of frequency and amplitude; Equal-loudness contours
Haptics	Tactile Property Space	Roughness, Hardness, Temperature, Friction	Physical limits of materials; Physiological limits of mechanoreceptors

Modality	Grounded Manifold ($\mathcal{M}$)	Example Quality Dimensions	Feasibility Constraint ($\mathcal{F}$)
Motor Control	Robot Configuration Space	Joint Angles (θ), Joint Velocities ($\dot{\theta}$), Joint Accelerations ($\ddot{\theta}$)	Kinematic workspace, joint limits, max velocity/acceleration, self-collision avoidance

Unified Multimodal Grounding and the Emergence of Abstract Concepts

With individual sensory and motor domains grounded in their respective geometric manifolds, the framework provides the tools to address two of the most challenging problems in cognitive science and AI: how information from different modalities is bound into coherent concepts, and how highly abstract concepts can be grounded in perceptual experience.

The Binding Problem as Geometric Intersection

The classic "binding problem" asks how the brain integrates disparate features—such as the color, shape, and texture of an object—into a single, unified percept. Standard multimodal AI approaches this by learning mappings between modalities, often using opaque "adapter" or "fusion" modules within a deep neural network to create a shared embedding space. While effective, these methods lack transparency and a clear compositional structure.

The geometric framework offers a principled, compositional, and transparent alternative. A multimodal concept is grounded as the **geometric intersection of convex concept regions across multiple, independent manifolds**. This approach follows directly from the axioms. Axiom 4 defines concepts as regions within a single manifold (e.g., the concept "red" is a region $R_{\text{red}} \subset \mathcal{M}_{\text{color}}$). Axiom 4 also defines the logical AND operator as set intersection ($\cap$). Therefore, a complex, multimodal concept like "apple" can be defined compositionally and interpretably as the intersection of its constituent properties, each grounded in its own domain :

$$ \text{Concept}(\text{apple}) = \text{Region}(\text{"red"}) \cap \text{Region}(\text{"round"}) \cap \text{Region}(\text{"smooth"}) $$

where each region exists in its respective manifold: $R_{\text{red}} \subset \mathcal{M}_{\text{color}}$, $R_{\text{round}} \subset \mathcal{M}_{\text{shape}}$, and $R_{\text{smooth}} \subset \mathcal{M}_{\text{haptic}}$. This provides a powerful and explicit mechanism for solving the binding problem. The concept of an apple is not an arbitrary token but is the geometric locus of points that simultaneously satisfy the constraints of being red, round, and smooth, each defined within a physically grounded space. This represents a significant architectural advantage over black-box fusion models, offering a path to multimodal systems with verifiable and compositional semantics.

Bootstrapping Abstraction

A critical challenge for any theory of grounded cognition is to explain the representation of abstract concepts that have no direct or obvious perceptual correlate, such as "justice," "democracy," or "causality". The framework, in its current form, provides a robust mechanism for grounding elementary perceptual properties like "red" or "bright," but it must also provide a pathway for bootstrapping from this foundation to the highest levels of human abstract thought.

The proposed solution is that such concepts are not grounded as simple regions but as **higher-order geometric structures** operating on the foundational manifolds. This approach is inspired by how human cognition often uses physical and spatial metaphors to reason about abstract domains (e.g., "grasping an idea," "a weighty decision," "a close relationship"). The framework provides the computational substrate for these metaphors.

- **Relations as Vectors or Transformations:** A comparative relation like "more than" or "brighter than" can be represented as a vector field pointing in the direction of increasing magnitude along a specific quality dimension (e.g., a vector field on the L axis in the color manifold). Applying this transformation to a point would move it to a "brighter" location.
- **Events as Trajectories:** Dynamic events and actions can be represented as trajectories through a joint manifold. For example, the action of "giving" could be represented as a characteristic trajectory through a combined manifold of motor space ($\mathcal{M}_{\text{motor}}$) and object space, describing the movement of an object from one agent's possession-space to another's.
- **Metaphorical Grounding:** The framework allows operations from one manifold to be applied to another, providing a basis for metaphorical reasoning. The motor-control concept of "grasping," a trajectory in $\mathcal{M}_{\text{motor}}$, could be metaphorically applied to a conceptual space of ideas to represent "understanding" or "comprehending."

The crucial constraint is that all such abstractions must ultimately be traceable back to a sequence of geometric operations on the foundational, grounded manifolds. This ensures a complete chain of traceability from the most abstract thought back to concrete perception and action. By insisting on this physical foundation, the framework inherently prevents the construction of a new, ungrounded "symbol/symbol merry-go-round" at higher levels of abstraction.

Theoretical Implications and Conclusion

The framework of Geometric Semantics, as detailed through the Color-Geometry case study and its generalization, represents a significant new paradigm for artificial intelligence. 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.

A "Glass Box" Architecture for Trustworthy AI

In an era dominated by opaque "black box" models, the framework offers a "glass box" alternative. Traditional 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, encoded in millions of weights, are difficult to explain or verify.

The geometric framework combines the strengths of both paradigms. Its decisions are based on verifiable geometric proofs. The solution to the constrained optimization 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.

The Physics of Meaning Revisited

The dynamics of the system—governed by the conservation of feasibility within the constraint \mathcal{F} and the entropic pull of compositional averaging—reinforce the conclusion that meaning in this model is not a purely abstract or linguistic phenomenon. Instead, it is subject to its own kind of "physical law". The constant tension between the contractive nature of composition and the anti-entropic process of learning, which must discover specific weightings to push representations back towards the information-rich boundary of the feasibility set, frames cognition as a dynamic process governed by principles analogous to those in thermodynamics. This physicalist perspective provides a powerful, non-arbitrary foundation for the structure and evolution of meaning within an intelligent system.

Summary and Future Horizons

This report has detailed a novel paradigm for grounded AI, reframing semantics, causality, and learning in the language of geometry. Its most profound contribution is its direct solution to the symbol grounding problem: by defining symbols as coordinates within a physically constrained perceptual manifold, it makes meaning an intrinsic property of the system's state. The framework's core principle can be summarized by a single equation: $$ \text{Meaning} = \text{Causal Reachability of Grounded States within a Physical Manifold } \mathcal{F}. $$ The introduction of this geometric paradigm opens several fertile and promising avenues for future research that could address its current limitations and expand its scope and power.

1. **Exploring Alternative Geometries:** While OKLab provides an excellent substrate, a systematic investigation into other advanced perceptual spaces (such as JzAzBz for color, or standardized timbre spaces for audio) is a critical next step. Each space possesses a unique geometry and a unique "feasibility droplet," and understanding how these different structures affect learning dynamics will be crucial for developing a general theory of geometric grounding.
2. **Dynamic Basis and Learned Primitives:** The current framework utilizes a static bias vector. Future work should explore methods that allow the network to learn its own set of "basis vectors" or perceptual primitives. This would enable the system to dynamically construct an optimal, task-specific conceptual palette, increasing its expressive power and learning efficiency.
3. **Hardware Co-Design:** The computational primitives of this framework—solving small, parallel QP problems and performing fast lookups in a feasibility oracle—are fundamentally different from the multiply-accumulate operations that dominate conventional deep learning hardware. This invites a new research direction in co-designing neuromorphic accelerators with dedicated hardware units optimized for these specific geometric operations, which could fully leverage the algorithm's event-driven nature and inherent parallelism.
4. **From Philosophy to Science:** Perhaps most importantly, this framework transforms philosophical theories like Gärdenfors' Conceptual Spaces, which have remained largely theoretical, into concrete, falsifiable, and runnable models. This allows researchers to move from philosophical argument to computational experiment. The psychological plausibility of hypotheses like "concept convexity" can be directly tested by building

models to learn human categories and then examining the geometric properties of the resulting learned regions. In this way, the framework has the potential to transform a philosophical inquiry into a progressive scientific research program, using the geometry of perception to unlock the geometry of thought.

Works cited

1. pmc.ncbi.nlm.nih.gov, https://pmc.ncbi.nlm.nih.gov/articles/PMC11792772/#:~:text=2.-,Conceptual%20spaces,a%20geometric%20or%20topological%20structure. 2. Conceptual space - Wikipedia, https://en.wikipedia.org/wiki/Conceptual_space 3. The Geometry and Dynamics of Meaning - PMC - PubMed Central, https://pmc.ncbi.nlm.nih.gov/articles/PMC11792772/ 4. Conceptual Spaces: The Geometry of Thought by Peter Gärdenfors | Goodreads, https://www.goodreads.com/book/show/1877443.Conceptual_Spaces 5. Conceptual spaces : the geometry of thought - Lund University Research Portal, https://portal.research.lu.se/en/publications/conceptual-spaces-the-geometry-of-thought 6. (PDF) Conceptual Spaces as a Framework for Knowledge Representation - ResearchGate, https://www.researchgate.net/publication/233707793_Conceptual_Spaces_as_a_Framework_for_Knowledge_Representation 7. What are “conceptual spaces”? - Lucas Bechberger's Website, https://lucas-bechberger.de/2017/02/14/what-are-conceptual-spaces/ 8. Reasoning about Categories in Conceptual Spaces - NYU Computer Science, https://cs.nyu.edu/faculty/davise/commonsense01/final/Gardenfors.pdf 9. Review of Conceptual Spaces--The Geometry of Thought, https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1554/1453 10. Dimension analyses of perceived sound quality of sound ..., https://www.researchgate.net/publication/229618143_Dimension_analyses_of_perceived_sound_quality_of_sound-reproducing_systems 11. (PDF) Psychophysical Dimensions of Tactile Perception of Textures, https://www.researchgate.net/publication/255704085_Psychophysical_Dimensions_of_Tactile_Perception_of_Textures 12. Perceiving sound: our sense of hearing – Introduction to Biological Psychology, https://openpress.sussex.ac.uk/introductiontobiologicalpsychology/chapter/perceiving-sound-our-sense-of-hearing/ 13. Kinematic Control of Robots With Joint Constraints - ResearchGate, https://www.researchgate.net/publication/235347844_Kinematic_Control_of_Robots_With_Joint_Constraints 14. AUDITORY PERCEPTION AND COGNITIVE PERFORMANCE - USAARL, https://usaarl.health.mil/assets/docs/hmds/Section-19-Chapter-11-Auditory-Perception-and-Cognitive-Performance.pdf 15. 10. What drives perception - Understanding Sound - Pressbooks.pub, https://pressbooks.pub/sound/chapter/measurement-and-perception/ 16. Evaluating the Effect of Voice Quality Covariance on Auditory-Perceptual Evaluation Using a Novel Two-Dimensional Magnitude Estimation Task - ASHA Journals, https://pubs.asha.org/doi/abs/10.1044/2023_JSLHR-23-00226 17. How We Hear: The Perception and Neural Coding of Sound - PMC - PubMed Central - NIH, https://pmc.ncbi.nlm.nih.gov/articles/PMC5819010/ 18. 4. The human auditory system - Audio quality in networked systems - Yamaha Corporation, https://usa.yamaha.com/products/contents/proaudio/docs/audio_quality/04_audio_quality.html 19. Potential Haptic Perceptual Dimensionality of Rendered Compliance - PMC - NIH, https://pmc.ncbi.nlm.nih.gov/articles/PMC9944077/ 20. Psychophysical Dimensions of Material Perception and Methods to Specify Textural Space - TMU Haptics lab.,

https://hci.fpark.tmu.ac.jp/shogo_okamoto/papers/PervasiveHaptics-SO-HN-HNH.pdf 21. Haptic Perception of Material Properties and Implications for Applications - ResearchGate, https://www.researchgate.net/publication/260686146_Haptic_Perception_of_Material_Properties_and_Implications_for_Applications 22. Organic Haptics: Intersection of Materials Chemistry and Tactile Perception - PMC, https://pmc.ncbi.nlm.nih.gov/articles/PMC8281818/ 23. Cues for Haptic Perception of Compliance - UU Research Portal, https://research-portal.uu.nl/files/605699/tth2009040189.pdf 24. Feeling Small: Exploring the Tactile Perception Limits - PMC - NIH, https://pmc.ncbi.nlm.nih.gov/articles/PMC3771396/ 25. Robot kinematics - Wikipedia, https://en.wikipedia.org/wiki/Robot_kinematics 26. Inverse Kinematics and Motion Control of Robots in Aerospace Manufacturing - Clemson University Open Textbooks, https://opentextbooks.clemson.edu/me8930/chapter/inverse-kinematics-and-motion-control-of-robots-in-aerospace-manufacturing/ 27. Chapter 16. Control of Articulated Robots - Intelligent Motion Lab, http://motion.cs.illinois.edu/RoboticSystems/RobotControl.html 28. Ch. 6 - Motion Planning - Robotic Manipulation, https://manipulation.csail.mit.edu/trajectories.html 29. Globally Optimal Inverse Kinematics Method for a Redundant Robot Manipulator with Linear and Nonlinear Constraints - MDPI, https://www.mdpi.com/2218-6581/9/3/61 30. Motion Tasks for Robot Manipulators Subject to Joint Velocity Constraints, http://users.ntua.gr/xpapag/Papers/0861.pdf 31. The Revolution of Multimodal Large Language Models: A Survey - arXiv, https://arxiv.org/html/2402.12451v2 32. From Concrete to Abstract: A Multimodal Generative Approach to Abstract Concept Learning - arXiv, https://arxiv.org/html/2410.02365v1