The Mathematical Structure of Semantic Space for Human-AI Collaboration: Modeling Convergence, Divergence, and Meaning Gravity

Fumio Sagawa
Ashiras, inc.
fumio.sagawa@gmail.com

April 2025

Abstract

We propose a mathematical framework for modeling semantic spaces to enable dynamic meaning alignment between humans and AI systems. By introducing the concepts of semantic distance, convergence, divergence, and meaning gravity, and formalizing communication processes via set theory, metric spaces, and topology, we redefine semantic interaction as a dynamic, quantifiable process. This approach provides a theoretical basis for convergence-driven design in human-AI collaboration.

1 Introduction

Despite the rapid advancement of artificial intelligence, a critical gap remains between technological capabilities and human conceptual understanding. Many assume that AI systems can fulfill arbitrary commands seamlessly, yet overlook the fundamental nature of AI as a structure for dynamic conversation and meaning negotiation. Rather than simply executing instructions, AI engages in the continuous construction and refinement of semantic alignments. Thus, to truly collaborate with AI — to connect seamlessly and meaningfully — we must first redefine what conversation itself is, and how its dynamics can be mathematically modeled.

This paper addresses this challenge by proposing a mathematical framework for semantic space, introducing the notions of convergence, divergence, semantic distance, and meaning gravity to formalize the evolution of shared meaning.

2 Related Work

2.1 Decision-Making Models

Traditional decision-making models, such as game theory and Bayesian decision frameworks, provide structured approaches to optimizing choices under uncertainty. Game theory focuses on strategic interactions among rational agents, while Bayesian decision theory emphasizes probabilistic reasoning based on prior and updated information. However, these models primarily address the optimization of decisions rather than the dynamic evolution of shared semantic understanding during interaction. Thus, a critical gap remains in modeling the formation, negotiation, and adjustment of meaning itself throughout collaborative processes, especially in human-AI contexts.

2.2 Semantic Similarity in NLP

In the field of natural language processing (NLP), various methods for measuring semantic similarity, such as cosine similarity of word embeddings (e.g., Word2Vec, GloVe, BERT), have been extensively developed to quantify the closeness of meanings between linguistic units. While effective for capturing static semantic relationships, these methods often fall short in modeling the dynamic and context-dependent shifts in meaning that arise during dialogue or interaction. Moreover, they typically lack a topological framework that could represent the continuous deformation of semantic structures over time.

2.3 Topological Approaches to Data

Topological data analysis (TDA) has emerged as a powerful technique for uncovering the global shape and structure of high-dimensional data, utilizing concepts such as persistent homology. While TDA offers valuable insights into data connectivity and clustering, its application to the modeling of dynamic semantic spaces—where meanings shift, converge, or diverge through ongoing interaction—remains largely unexplored. Our proposed framework aims to bridge this gap by applying topological modeling directly to semantic dynamics in communication, offering a new approach to understanding meaning negotiation and alignment over time.

3 Mathematical Foundations of Semantic Space

3.1 Sets, Domains, and Elements

We begin by formalizing the fundamental elements of semantic space. Let A denote a set of semantic elements, where each element $x \in A$ represents a discrete unit of meaning—a concept, term, or idea as understood within a particular context.

Definition 3.1 (Set) *A set A is a collection of distinct semantic elements.*

Definition 3.2 (Element) *An element $x \in A$ denotes a basic unit of meaning in the semantic space.*

Definition 3.3 (Domain) *A domain $D \subseteq A$ specifies the currently active set of semantic elements under consideration in a dialogue or decision-making process.*

Crucially, communication operates within a domain D, which may dynamically expand or contract as the interaction progresses. Tracking changes in D is essential for modeling convergence and divergence.

3.2 Definition of Semantic Distance

To quantify the relationship between semantic elements, we define a semantic distance function $d : A \times A \to \mathbb{R}_{\geq 0}$.

Definition 3.4 (Semantic Distance) *For any two elements $x, x' \in A$, the semantic distance $d(x, x')$ satisfies the following properties:*

1. ***Non-negativity**: $d(x, x') \geq 0$*
2. ***Identity of indiscernibles**: $d(x, x') = 0$ if and only if $x = x'$*
3. ***Symmetry**: $d(x, x') = d(x', x)$*
4. ***Triangle inequality**: $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in A$*

Thus, (A, d) forms a metric space, providing a rigorous mathematical foundation for modeling the "closeness" or "distance" of meanings within communication.

In subsequent sections, we build upon these foundations to define dynamic processes such as convergence, divergence, and the influence of meaning gravity within semantic spaces.

4 Dynamics of Convergence and Divergence

4.1 Convergence Structuring

In the course of dialogue or collaborative decision-making, semantic elements can evolve towards greater alignment, a process we term **convergence**.

Definition 4.1 (Convergence Structuring) *A communication process exhibits convergence if, for a given domain $D \subseteq A$, the semantic distances between elements decrease over time:*

$$\forall x, x' \in D, \quad d_{t+1}(x, x') \leq d_t(x, x')$$

where t denotes the time step or dialogue iteration.

Convergence structuring reflects the gradual harmonization of meanings, which is critical for successful mutual understanding and decision-making.

4.2 Divergence and Divergence Degree

Conversely, communication may experience **divergence**, where meanings drift apart.

Definition 4.2 (Divergence) *A communication process exhibits divergence if, during its progression, either:*

- *the domain expands ($D \subset D'$), introducing new semantic elements, or*
- *semantic distances between elements increase:*

$$\exists x, x' \in D', \quad d_{t+1}(x, x') > d_t(x, x')$$

Definition 4.3 (Divergence Degree D_g) *To quantify divergence, we define the divergence degree D_g as:*

$$D_g = \frac{|D'| - |D|}{|D|} + \frac{1}{n} \sum_{(x,x') \in D \times D} (d_{t+1}(x, x') - d_t(x, x'))$$

where:

- $|D|$ *is the cardinality of the domain at time* t
- $|D'|$ *is the cardinality at time* $t + 1$
- $n = |D| \times (|D| - 1)/2$ *is the number of element pairs*

The first term captures domain expansion, and the second term captures semantic drift. A larger D_g indicates a stronger divergence, signaling potential breakdowns in mutual understanding.

4.2.1 Motivation and Interpretation of Divergence Degree

The Divergence Degree D_g quantitatively captures the extent to which semantic alignment is lost over time. The first term, $\frac{|D'|-|D|}{|D|}$, measures domain expansion, reflecting how new concepts or terms enter the dialogue and potentially introduce semantic drift. The second term, involving the average increase in semantic distances, directly quantifies how meanings of existing elements grow apart. Together, these two factors provide a dual lens for assessing communication stability: expansion of the semantic domain and internal semantic dispersion. A larger D_g indicates a higher risk of misalignment, making it an essential indicator for monitoring collaborative interactions, especially in dynamic or uncertain environments.

4.3 Gravity of Meaning

Beyond convergence and divergence, semantic evolution can also be influenced by attractive forces between elements—what we term **meaning gravity**.

Meaning gravity models the tendency of related semantic elements to be drawn closer over time, independent of external intervention. Formally, we define a gravity potential $g : A \times A \to \mathbb{R}$ such that:

$$g(x, x') = -\gamma \cdot s(x, x')$$

where $\gamma > 0$ is a gravity constant and $s(x, x')$ is a semantic similarity measure.

The overall effect of meaning gravity on semantic dynamics can be seen as an additional force promoting convergence, counteracting divergence due to noise, misunderstanding, or external complexity.

Meaning gravity thus serves as a foundational concept for designing communication processes that are naturally robust and self-stabilizing.

Formally, the gravity potential $g(x, x')$ depends on the semantic similarity function $s(x, x')$, which quantifies how closely related two semantic elements are, ranging from 0 (completely unrelated) to 1 (identical meaning). The gravity constant γ regulates the overall strength of attraction across the semantic space. This formulation ensures that elements with higher semantic affinity exert a stronger pull toward each other, promoting local convergence even without explicit guidance.

4.3.1 Intuitive Understanding of Meaning Gravity

Meaning gravity can be intuitively understood as a natural tendency for semantically related concepts to cluster together, similar to how gravitational forces in physics pull masses toward each other. In communication, this ”semantic pull” promotes cohesion even without deliberate coordination. For example, in a conversation about travel, words like ”airport,” ”ticket,” and ”passport” naturally group without explicit negotiation. Mathematically, meaning gravity acts as a stabilizing field that counteracts divergence, encouraging dialogue trajectories to remain within coherent semantic regions. By embedding this gravitational structure into human-AI systems, it becomes possible to design interactions that are more robust, context-aware, and self-repairing.

5 Topological Structure of Meaning

5.1 Modeling Semantic Space as a Topological Space

While the metric space structure provides a means to measure semantic distances, it does not fully capture the subtleties of how meanings can continuously deform, merge, or bifurcate in real-world communication. To model these subtler dynamics, we introduce a topological perspective on semantic space.

Definition 5.1 (Topological Space of Meaning) *A topological space $(A, \mathcal{T})$ consists of a set A of semantic elements and a collection $\mathcal{T} \subseteq 2^A$ of subsets, called open sets, satisfying:*

1. *$\emptyset \in \mathcal{T}$ and $A \in \mathcal{T}$*

2. *The intersection of any two open sets is also an open set.*

3. *The union of any collection of open sets is also an open set.*

Here, open sets represent neighborhoods of meaning, capturing flexible, context-sensitive associations that may not correspond to strict distance thresholds.

5.2 Semantic Convergence and Divergence in Topology

Using the topological structure, we can generalize convergence and divergence beyond metric strictness.

Definition 5.2 (Topological Convergence) *A sequence $\{x_n\} \subset A$ converges to $x \in A$ if, for every open set $U \in \mathcal{T}$ containing x, there exists an N such that for all $n \geq N$, $x_n \in U$.*

This captures the idea that meanings can gradually "enter" the same neighborhood without requiring exact identity, reflecting how human communication often operates.

Definition 5.3 (Topological Divergence) *If no such x exists where $\{x_n\}$ converges, or if meanings escape from common neighborhoods over time, the process exhibits divergence.*

5.2.1 Motivation and Interpretation of Topological Modeling

The move to topology allows us to capture the "soft" nature of human communication where meanings do not necessarily converge to a single point but gather in overlapping regions. Convergence in topology does not require perfect identity but proximity within a flexible neighborhood. This models:

- Partial or fuzzy convergence of meanings.
- Continuous deformations without discrete jumps.
- Local coherence even amid global divergence.

Such flexibility is essential for describing the dynamic, context-sensitive negotiation of meaning in human and human-AI interactions. It also paves the way for sophisticated monitoring of communication dynamics, where the goal is not absolute consensus but the maintenance of topological proximity.

6 Redefining Communication through Semantic Models

6.1 Communication as Dynamic Semantic Structuring

Traditional models often treat communication as the transmission of static information from sender to receiver. However, when examined through the lens of semantic space modeling, communication reveals itself as a dynamic restructuring of meanings within an evolving topological space. Each communicative act reshapes the underlying semantic topology, adjusting the relative positions of concepts, bridging previously distant ideas, or reinforcing existing neighborhoods of meaning.

Rather than viewing understanding as a binary outcome (message received or not), communication should be regarded as the continuous transformation of semantic configurations toward mutual coherence.

6.2 Rethinking Meaning as a Topological Construct

Meaning is not a static property of symbols but a relational structure situated within a semantic topology. In this view:

- Meanings are positions within a structured semantic space.
- Communication is the movement of meanings through this space.
- Understanding corresponds to the convergence of meaning trajectories into overlapping topological regions.

This topological conception allows for flexible, context-sensitive representations of meaning that better capture the complexities of human dialogue and human-AI interaction.

6.3 Semantic Dynamics and Mutual Alignment

The dynamics of semantic spaces naturally introduce concepts such as:

- **Convergence**: The progressive narrowing of semantic distances, leading to shared understanding.
- **Divergence**: The expansion of semantic distances, indicating emerging misunderstanding or conceptual drift.
- **Meaning Gravity**: The intrinsic tendency for semantically related concepts to cluster, fostering natural local convergence even in the absence of explicit guidance.

Effective communication strategies should therefore not merely transmit information but actively manage semantic convergence processes, detect divergence early, and exploit meaning gravity to stabilize and reinforce shared understanding.

6.4 Toward a Dynamic Model of Understanding

We propose modeling understanding itself as a dynamic convergence process within a topological semantic space:

- **Partial understanding** corresponds to partial overlaps between semantic neighborhoods.
- **Stable understanding** corresponds to persistent convergence across dialogue iterations.
- **Misunderstanding** corresponds to divergence or semantic escape from shared topological regions.

This dynamic model enables more nuanced, real-time assessment of human-AI communication quality and provides a theoretical foundation for designing interaction protocols that promote robust and adaptive shared meaning.

7 Applications to Human-AI Collaboration

7.1 Semantic Distance Monitoring

Embedding semantic elements in a structured space enables real-time monitoring of semantic distances between communicating entities. In the context of human-AI collaboration, this approach allows for dynamic feedback loops:

- **Prototyping Example**: A live semantic alignment meter can be displayed during human-AI conversations, showing how "close" or "divergent" the meanings currently are.
- **Practical Use**: Alert systems can trigger when divergence exceeds a critical threshold, prompting clarification or additional prompts.

7.2 Real-Time Convergence Design

Instead of treating convergence as an incidental outcome, dialogues and interfaces can be explicitly designed to foster convergence:

- **Dialogue Strategy**: After every major user input, the system summarizes or paraphrases to test convergence.
- **Design Pattern**: Insert checkpoints where the AI explicitly asks "Are we aligned so far?" and adjust dynamically based on feedback.

7.3 Meaning Homotopy and Compactness

Advanced semantic modeling enables the application of topological concepts like homotopy and compactness to communication design:

- **Meaning Homotopy**: Ensure that topic transitions are smooth. For example, when shifting from "travel" to "visa applications," guide the conversation naturally by intermediate concepts like "immigration."
- **Semantic Compactness**: Limit the semantic search space in AI responses to a compact subset, ensuring that even creative responses remain coherent with the topic.

This approach helps systems stay robust, maintaining clear and coherent communication even as dialogue topics evolve dynamically.

8 Conclusion

This paper proposed a mathematical framework for modeling semantic spaces to address fundamental challenges in human-AI collaboration. By introducing and formalizing key concepts—semantic distance, convergence, divergence, topological structuring, and meaning gravity—we redefined communication not as a static transmission of information, but as a dynamic negotiation of meaning within evolving semantic topologies.

Through set theory, metric spaces, and topological constructs, we provided a formal basis for modeling the fluid evolution of shared understanding. Convergence and divergence dynamics were formalized to capture the natural processes of semantic alignment and drift, while the notion of meaning gravity introduced an intrinsic mechanism for stabilizing communication without explicit external correction.

This perspective reveals that successful collaboration with AI systems depends not merely on command execution, but on the continuous co-evolution of meaning. Monitoring semantic distances, designing interactions to encourage convergence, and leveraging meaning gravity become essential strategies for maintaining coherent, adaptive communication.

Future research will aim to refine the mathematical modeling of real-time semantic dynamics, develop practical tools for semantic monitoring and convergence-driven design, and extend the framework to multi-agent systems where complex semantic negotiations unfold.

Ultimately, to collaborate with AI effectively, humans must transcend static message-passing paradigms and embrace a vision of communication as dynamic, topological meaning construction—a process that is at once mathematical, cognitive, and profoundly human.

Author's Note

AI-assisted refinement was employed to enhance linguistic clarity, not to generate original content. The author regards AI not as a replacement for human thinking, but as a collaborator in the realization of precise articulation.The theoretical framework, mathematical modeling, and conceptual development presented in this paper are entirely the author's original work.

References

[1] T. Mikolov, K. Chen, G. Corrado, and J. Dean, *Efficient Estimation of Word Representations in Vector Space*, arXiv preprint arXiv:1301.3781, 2013.

[2] J. Li, W. Monroe, and D. Jurafsky, *A Simple, Fast Diverse Decoding Algorithm for Neural Generation*, arXiv preprint arXiv:1611.08562, 2016.

[3] P. Gärdenfors, *Conceptual Spaces: The Geometry of Thought*, MIT Press, 2000.