Physical Interpretations of Emergent Behaviors

Bentley Yusen Lin

September 8, 2025

Abstract

This paper presents a novel mathematical framework for interpreting emergent behaviors in complex systems, with a particular focus on modern artificial intelligence (AI). We posit that emergence is not a mystical phenomenon but a predictable consequence of a system's exploration of a high-dimensional state space. By defining the relationship between the microscopic rules of a system and its macroscopic properties through the lens of exploration capacity, we provide a model that explains why low-probability, "miraculous" behaviors become reliable and in sufficiently large systems, such as large language models. The core of our thesis is the derivation of an inverse exponential relationship between the probability of an emergent behavior and the size of the exploration space, offering a quantitative basis for this previously qualitative concept.

1 Introduction

The concept of *emergence* describes the phenomenon where complex behaviors and patterns arise from the interaction of simpler components. These macroscopic behaviors are often difficult to predict from the knowledge of the microscopic rules alone [1]. In fields ranging from thermodynamics to biology, emergence is a cornerstone principle. The recent and rapid advancement of AI, particularly with the rise of deep learning, has brought this concept to the forefront of computer science. Capabilities such as in-context learning, reasoning, and creativity *emerge* in large-scale neural networks without being explicitly programmed [2].

While often described qualitatively, a pressing need exists for a formal, mathematical framework to describe the mechanics of emergence. This paper aims to bridge that gap by proposing a physical and mathematical interpretation of emergence, framing it as a function of a system's exploration capacity within a constrained high-dimensional space.

2 A Model of Exploration and Emergence

2.1 Definitions

Consider a system whose state can be represented as a point in an N-dimensional space. At any discrete time step t, the system can transition to a new state by moving in a chosen direction.

- **Theoretical Exploration Space ($\mathbf{\Omega_T}$)**: In an unconstrained system, the number of possible directional choices per time step is 2^N (positive or negative along each axis).
- **Constrained Exploration Space ($\mathbf{\Omega_C}$)**: Physical and mathematical constraints (e.g., network architecture, loss landscape geometry) limit the system's realistic choices. We model this reduction by defining an *effective exploration dimensionality* K, where $K < N$. The number of available directions per time step is thus reduced to 2^K.

2.2 The Exploration Space Size

Over a sequence of T time steps (e.g., training steps or inference steps), the total number of possible paths S the system can take is given by:

$$S = (2^K)^T = 2^{KT} \tag{1}$$

S represents the total **exploration space size**. The logarithm of S is proportional to the system's informational entropy, representing its potential diversity of behaviors.

2.3 Probability of Emergent Behaviors

Let P_L be a low-probability macroscopic state (e.g., a coherent long-form answer to a complex query). This state is accessible via a set of paths $A \subset S$. If the paths are equally probable, the probability of the system exhibiting the emergent behavior P_L is:

$$P(P_L) = \frac{|A|}{S} = \frac{|A|}{2^{KT}} \tag{2}$$

For many non-trivial behaviors, $|A|$ is constant or grows sub-exponentially with S (i.e., $|A| \in o(2^{KT})$). This leads to our central proposition:

$$P(P_L) \propto \frac{1}{S} = 2^{-KT} \tag{3}$$

Equation 3 establishes an **inverse exponential relationship** between the probability of an emergent behavior and the size of the exploration space. This explains why such behaviors are vanishingly rare in small systems (S is small, $P(P_L)$ is high but $|A|$ is effectively zero).

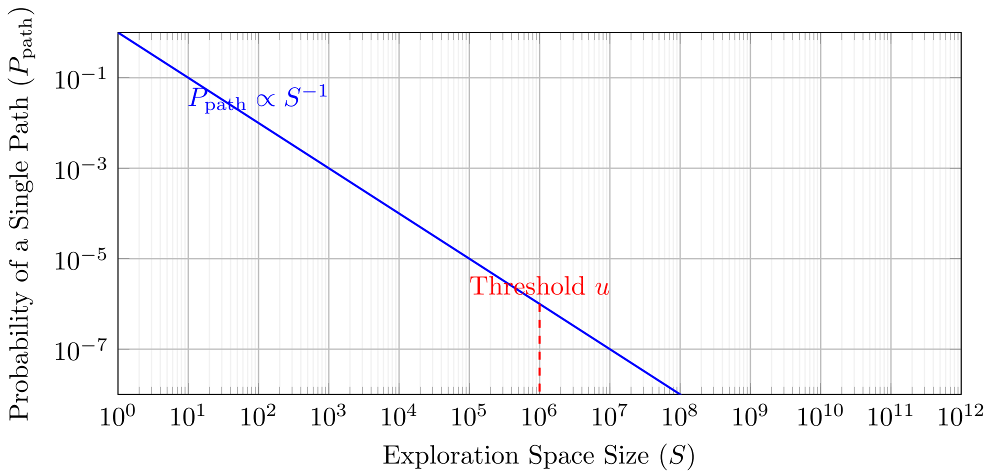

Figure 1: **Probability of a Single Path Decreases with Exploration Space:** The probability P_{path} of any single, specific path (e.g., one that leads to a complex emergent behavior P_L) decays inversely with the total size of the exploration space S. This explains why complex behaviors reside in a low-probability space. The threshold u marks the point where S becomes large enough for the absolute number of paths to any given P_L to become non-negligible, even though the probability for each individual path continues to decay.

3 The Mechanism of Miracle-to-Normality Transition

The transition of an emergent behavior from a "miracle" to a "normality" is a direct consequence of scaling S [3].

3.1 The Threshold of Emergence

There exists a threshold u such that when $S > u$, the absolute number of paths leading to P_L becomes non-negligible:

$$|A| = P(P_L) \cdot S$$

Although $P(P_L)$ decays exponentially, S grows exponentially. Once S is sufficiently large, the product $|A|$ can become large enough that the system frequently encounters state P_L. The "miracle" is observed relative to a system with $S < u$.

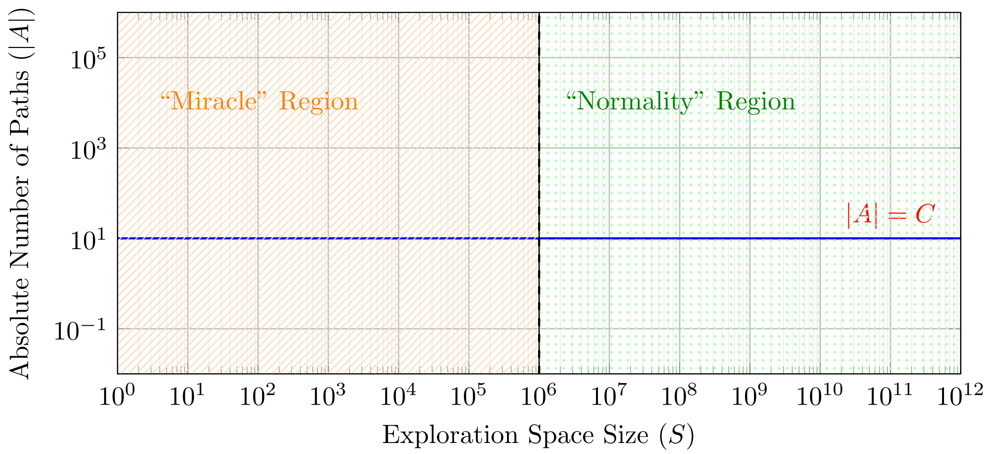

Figure 2: **The Miracle-to-Normality Flip:** This plot illustrates the central mechanism of emergence. Although the probability of any single path decays (as in Figure 1), the absolute number of paths $|A|$ to a specific emergent state is constant (in this simple model where $|A|$ is fixed). For $S < u$ (left of the threshold), the system has a low probability of finding, executing, and completing one of these paths. For $S > u$ (right of the threshold), the exploration space is so vast that it encompasses these paths many times over, making the emergent behavior a common, reliable occurrence. The behavior transitions from a rare "miracle" to a common "normality."

3.2 Engineering Emergence through Scale

Modern AI provides a quintessential example of this principle. The "technical means" to enlarge S are:

- **Increasing K (Effective Dimensionality)**: This is achieved by increasing model parameter count. A larger model has a higher-dimensional, richer space of possible internal representations and transformations [4].
- **Increasing T (Exploration Steps)**: This corresponds to increasing training compute (FLOPs) and allowing longer reasoning chains during inference (e.g., Chain-of-Thought prompting) [5].

By scaling K and T, engineers effectively force S far beyond the threshold u for desired capabilities like reasoning or code generation. Consequently, these once-emergent "miracles" become reliable, marketable features.

4 Discussion and Conclusion

We have formulated a mathematical model where emergence is redefined not as an anomaly but as an inevitable outcome of scale. The inverse exponential relationship $P(P_L) \propto S^{-1}$ provides a quantitative foundation for this phenomenon.

This framework demystifies the behaviors observed in large-scale AI systems. It suggests that the pursuit of Artificial General Intelligence (AGI) may be less about discovering new algorithms and more about the strategic scaling of existing architectures to navigate the exploration space efficiently [6]. The challenge shifts from *whether* a capability can emerge to *how we can guide* the system to find the paths A that lead to desirable and safe macroscopic states P_L.

Future work will focus on refining this model, particularly on relaxing the assumption of path uniformity and formally characterizing the growth function of $|A|$ for specific emergent capabilities.

References

References

[1] Mitchell, M. (2009). *Complexity: A Guided Tour.* Oxford University Press.

[2] Wei, J., et al. (2022). *Emergent Abilities of Large Language Models*. Transactions on Machine Learning Research.

[3] Kaplan, J., et al. (2020). *Scaling Laws for Neural Language Models*. arXiv:2001.08361.

[4] LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep learning*. Nature, 521(7553), 436–444.

[5] Wei, J., et al. (2022). *Chain-of-Thought Prompting Elicits Reasoning in Large Language Models*. Advances in Neural Information Processing Systems.

[6] Hutter, M. (2019). *Human Knowledge: Foundations and Limits*. Online. `https://www.hutter1.net/ai/humknow.htm`