The Attralucian Essays:
Exploring the Finite

First Edition

 Typeset in LaTeX

The Attralucian Essays

Kolmogorov Complexity: A Geofinitist Reinterpretation

Kevin R. Haylett

Kolmogorov Complexity: A Geofinitist Reinterpretation

Overview

Kolmogorov complexity provides a foundational notion of information content: the length of the shortest program that produces a given object. Formally elegant and deeply influential, it underpins areas ranging from data compression to machine learning. However, it is also uncomputable in the general case, placing it beyond direct operational use.

This paper does not attempt to resolve that uncomputability. Instead, it reframes the concept within a Geofinitist framework, where all quantities are finite, measured, and accompanied by uncertainty. In this view, Kolmogorov complexity is replaced not by an exact value, but by a bounded, reproducible interval derived from compression procedures, statistical models, and resource constraints.

Classical Definition and Limitation

For a string $x \in \Sigma^\star$ and a universal Turing machine U, Kolmogorov complexity is defined as:

$$K_U(x) = \min\{|p| : U(p) = x\}.$$

This definition captures the minimal description length of x. A fundamental result, however, is that $K_U(x)$ is uncomputable: no algorithm can determine the shortest program for arbitrary x, a consequence of the Halting Problem.

Despite this, the theory provides powerful insights. It is invariant up to an additive constant across universal machines and serves as a conceptual benchmark for compression. Yet, as a measurable quantity, it remains inaccessible.

A Measured Perspective

In practice, one never computes $K_U(x)$. Instead, one constructs:

- **Upper bounds** via compression algorithms,
- **Lower bounds** via entropy estimates or incompressibility tests.

Geofinitism formalises this practical reality. Rather than treating $K_U(x)$ as an unattainable exact value, it introduces a measured surrogate: a bounded interval derived from reproducible procedures under finite constraints.

Geofinitist Principles Applied

1. Computation as a Resource-Bounded Process

Program search is not infinite but constrained by budgets in time, memory, and energy. Complexity is defined relative to these bounds.

2. Measurement and Uncertainty

Program length, runtime, and output correctness are measured quantities:

$$|p| \pm \varepsilon_p, \quad T \pm \varepsilon_T, \quad y \pm \varepsilon_y.$$

3. Layered Structure

Descriptions span multiple layers: encoding schemes, interpreters, and execution environments. These contribute to total description length and variability.

4. Useful Fiction

The exact $K_U(x)$ is a limiting construct. Its practical role is as a reference point for bounding procedures.

5. Finite Constraints

All measurements occur under finite computational budgets, replacing infinite search with bounded exploration.

Measured Kolmogorov Complexity

Let U be a fixed reference machine with provenance P_U. Define a measured execution:

$$\mathsf{Run}_U(p) = (y, \varepsilon_y, P_U; T, \varepsilon_T; S, \varepsilon_S).$$

For tolerance τ and resource budget $B = (T_{\max}, S_{\max})$, define:

$$K^{\mathbb{M}}_{U,\tau,B}(x) = \min\Big\{|p| : d_{\mathbb{M}}(\mathsf{Run}_U(p), x) \leq \tau,\ \text{resources} \leq B\Big\}.$$

This defines a *measured description length* rather than an exact minimum.

Finite Invariance

For admissible machines U and V:

$$\left|K^{\mathbb{M}}_{U,\tau,B}(x) - K^{\mathbb{M}}_{V,\tau',B'}(x)\right| \leq c_{UV} \pm \varepsilon_{UV},$$

where c_{UV} reflects encoding overhead and ε_{UV} captures measurement variability.

Thus invariance becomes a bounded empirical property.

Operational Bounds

Upper Bounds. Given a compressor C:

$$K^{\mathbb{M}}(x) \leq L_{\mathsf{C}}(x) + c_{\mathsf{C}\to U} \pm \varepsilon_{\mathsf{C}}.$$

Lower Bounds. Using empirical entropy:

$$K^{\mathbb{M}}(x) \gtrsim |x| \cdot \widehat{H}_k(x) - \text{pen}_k \pm \varepsilon_{\text{fit}}.$$

Two-Part Codes and MDL

For a model class $\mathcal{M}$:

$$\mathrm{MDL}_{\mathcal{M}}(x) = \min_{m \in \mathcal{M}} \big(L(m) + L(x|m)\big).$$

This provides a finite surrogate:

$$K^{\mathbb{M}}(x) \approx \mathrm{MDL}_{\mathcal{M}}(x) \pm \varepsilon_{\mathcal{M}}.$$

Robust (Smoothed) Complexity

Define:

$$K^{\mathbb{M}}_{\eta}(x) = \mathbb{E}[K^{\mathbb{M}}(\mathsf{P}_{\eta}(x))].$$

Stability under perturbation distinguishes structural simplicity from brittle encodings.

Comparison with Abstention

Define:

$$\Delta K = K^{\mathbb{M}}(x) - K^{\mathbb{M}}(y).$$

Decide only when $|\Delta K|$ exceeds uncertainty thresholds; otherwise report indeterminacy.

Interpretation

Kolmogorov complexity becomes an operational protocol:

- upper bounds from compressors,
- lower bounds from statistical models,
- uncertainty from measurement and variability,
- decisions based on reproducible margins.

The classical quantity $K_U(x)$ remains a limiting ideal. The practically meaningful object is the *measured description length interval.*

Conclusion

Kolmogorov complexity is not replaced but finitised. Its uncomputability remains in the asymptotic limit, but within finite regimes it yields to a structured, measurable framework.

This transformation aligns theory with practice: information content is no longer an inaccessible minimum, but a reproducible, bounded quantity grounded in real computation.