

Short Communication

The Inexhaustible Frontier: Knowledge, Ignorance, and Boolean Hypercubes

Ramon Carbó-Dorca^{1,2}

1. Institute of Computational Chemistry and Catalysis (IQCC), University of Girona, Girona, Spain; 2. Ronin Institute for Independent Research, USA

We propose a formal framework for understanding the structural relationship between knowledge and ignorance, grounded in the geometry of Boolean hypercubes. Any finite, digitally describable piece of knowledge can be represented as a vertex of a Boolean hypercube: $\mathbf{H}_N \equiv \{0, 1\}^N$, where N encodes the complexity of the description. We demonstrate that extending this hypercube by a single bit — a natural and necessary step as knowledge grows — doubles the vertex space, adding a body of new, unexplored vertices roughly equal in size to the entire prior knowledge base. This process, iterable without bound, implies that the frontier of ignorance expands at least as fast as knowledge itself. We connect this result to Gödel's incompleteness theorems and discuss its implications for the fundamental limits of artificial intelligence systems, including state-of-the-art protein structure predictors such as AlphaFold 2. The argument applies equally to human scientific knowledge, suggesting that inexhaustible ignorance is not a failure of intellect but rather a structural feature of any finite system of knowing operating in an unbounded knowledge space.

Corresponding author: Ramon Carbó-Dorca, ramoncarbodorca@gmail.com

"As our circle of knowledge expands, so does the circumference of darkness surrounding it." —

Attributed to Albert Einstein

1. Introduction

A recurrent theme in epistemology is the asymmetry between what is known and what remains unknown. Scientists across disciplines have observed that answered questions generate new ones, and that progress in a field often reveals previously invisible depths rather than approaching closure. The celebrated Socratic dictum — that wisdom begins in recognizing one's ignorance — has intuitive appeal

but lacks formal grounding. Here we offer such a grounding through the language of Boolean algebra and combinatorics.

Based on previous work on Boolean hypercubes ^[1], and its connection with the limits of artificial intelligence ^[2], our starting point is simple: any piece of knowledge that can be described in a finite, discrete manner — a proposition, a molecular structure, a physical law, a trained neural network, a photograph, any printed book or journal ... — can be encoded as a binary string, and therefore identified with a vertex of a Boolean hypercube. This encoding is not merely metaphorical; it follows from the Church-Turing thesis and Shannon's foundational work on information ^{[3][4]}. The question we pursue is: what does the geometry of the hypercube tell us about the relationship between accumulated knowledge and the space of the unknown?

2. Boolean Hypercubes as a Model of Knowledge

Let $\mathbf{K}$ denote the current totality of human (or artificial) knowledge about a given domain. Assume $\mathbf{K}$ can be encoded digitally. Then $\mathbf{K}$ corresponds to a subset $\mathbf{S}(\mathbf{K})$ of the vertices, in number equal to $|\mathbf{S}(\mathbf{K})|$, of some Boolean hypercube $\mathbf{H}_N \equiv \{0, 1\}^N$, where N is sufficiently large to encode all distinguishable states of knowledge in that domain. The dimension N need not be fixed; it grows as the domain's descriptive complexity increases.

The total number of vertices in $\mathbf{H}_N$ is $|\mathbf{H}_N| = 2^N$. The explored subset $\mathbf{S}(\mathbf{K})$, however large in absolute terms, occupies an exponentially shrinking fraction of the total vertex space as N grows. It can be just one of the vertices of the hypercube $\mathbf{H}_N$, provided the hypercube's dimension is sufficiently large. Even for modest N , the ratio $\frac{|\mathbf{S}(\mathbf{K})|}{2^N}$ rapidly approaches zero. This observation alone establishes a quantitative sense in which knowledge is always a small island: the combinatorial explosion of the full space outpaces any finite accumulation.

3. The Doubling Lemma

Consider the effect of adding a single bit to the description of every vertex in $\mathbf{H}_N$. This operation — concatenating either a 0 or a 1 to each existing vertex — is a natural consequence of any genuine increase in descriptive resolution: a new variable, a new observable, a new dimension of inquiry. The result is the hypercube $\mathbf{H}_{N+1} \equiv \{0, 1\}^{N+1}$, which now contains $2^{N+1} = 2(2^N)$ vertices.

The known number of vertices $[\mathbf{S}(\mathbf{K})]$, extended by one bit, now occupy at most $2[\mathbf{S}(\mathbf{K})]$ positions in $\mathbf{H}_{N+1}$. But the new space contains 2^{N+1} vertices in total. The number of newly introduced, unexplored vertices is therefore: $2^{N+1} - 2[\mathbf{S}(\mathbf{K})] = 2(2^N - [\mathbf{S}(\mathbf{K})])$.

For any realistic $\mathbf{S}(\mathbf{K})$ (with $[\mathbf{S}(\mathbf{K})] \ll 2^N$), the newly unknown region is roughly equal in size to the entire previous hypercube $\mathbf{H}_N$. In other words:

Each step of genuine knowledge expansion creates an unknown frontier at least as large as the prior knowledge base.

This process can be iterated indefinitely — or more precisely, as many times as the domain admits new dimensions of inquiry. In practice, both human science and artificial systems are finite, so the knowledge iteration cycle terminates somewhere or at some point. But the conclusion is robust: at no point in the iteration does the ratio of known to unknown vertices approach unity. The frontier of ignorance grows with knowledge, not against it.

4. Resonance with Gödel's Incompleteness

The above result naturally resonates with Gödel's first incompleteness theorem ^[5], which states that any consistent formal system capable of expressing elementary arithmetic contains true statements that cannot be proved within the system. Gödel's result is syntactic: it concerns provability within a fixed formal language. Our point of view is combinatorial: it concerns the relative size of explored and unexplored knowledge spaces.

The connection is more than analogical. Both results point to the same structural fact: no finite, self-contained system — whether a formal logic, a scientific discipline, or a trained machine learning model — can exhaust the space of truths relevant to its own domain. Where Gödel identifies specific unprovable truths, the Boolean hypercube argument shows that such truths are not exceptional but numerically dominant.

5. Implications for Artificial Intelligence

Modern AI systems, including large language models and structure-prediction networks such as AlphaFold 2 ^[6], are trained on finite corpora of human-generated knowledge. Their internal representations can be understood as learned mappings over a subset of the knowledge hypercube —

precisely $\mathbf{S}(\mathbf{K})$ at the time of training. Such systems perform remarkably within the convex hull of their training distribution, effectively interpolating between known vertices.

However, the doubling lemma implies that any domain in which these systems operate contains an unexplored vertex space that is structurally inaccessible to interpolation alone. AlphaFold 2 illustrates this concretely: for proteins with sparse evolutionary records — few homologs in the sequence databases that constitute its training $\mathbf{S}(\mathbf{K})$ — the model's confidence scores collapse and its predictions become unreliable [6]. The system does not fail arbitrarily; it fails precisely where the target vertex lies far from the explored region.

This is not a correctable engineering limitation. It is a consequence of the geometry of knowledge itself. Expanding the training corpus extends $\mathbf{S}(\mathbf{K})$, but by the doubling lemma, doing so simultaneously extends the unknown frontier.

A more knowledgeable AI is not closer to omniscience — it is adjacent to a larger unknown.

6. The Human Parallel

The same argument applies, without modification, to human scientific knowledge. Each major theoretical advance — Newtonian mechanics, quantum theory, molecular biology — expanded $\mathbf{S}(\mathbf{K})$ substantially while simultaneously revealing that the new $\mathbf{H}_{N+1}$ was vastly larger than the old $\mathbf{H}_N$. This is not a coincidence or a failure of ambition; it is the expected consequence of genuine discovery in a space that grows faster than any exploration can track.

The difference between human scientists and current AI systems may lie not in the Boolean hypercube geometry — which constrains both equally — but in the mechanism of navigation. Human science reaches new vertices through experiment, serendipity, and abductive reasoning: processes that are not purely interpolative. Whether artificial systems can develop analogous mechanisms remains an open question of both technical and philosophical importance.

7. Conclusions

We have shown that Boolean hypercubes provide a natural and rigorous setting for formalizing the relationship between knowledge and ignorance. The doubling lemma establishes that the frontier of ignorance expands at least as fast as knowledge, at every single step of inquiry. This result resonates with Gödel's incompleteness theorems, extends naturally to artificial intelligence, and formalizes what

scientists have long known intuitively: that depth of knowledge and awareness of ignorance grow together.

The practical implication for AI development is clear: evaluating an AI system only within its training distribution is insufficient. The relevant question is always the size and structure of the unknown region adjacent to $\mathbf{S}(\mathbf{K})$ — a question the system itself cannot answer. External validation, experimental science, and human judgment remain indispensable precisely because they can, in principle, reach Boolean hypercube vertices that no purely interpolative system can access.

Knowledge is not a converging series. It is an expanding frontier surrounded, always, by a larger dark pool of unknown material, waiting to be illuminated by the light of exploration.

Statements and Declarations

Potential Competing Interests

The authors state that there are no competing interests in the present work.

Author Contributions

R. Carbó-Dorca conceived the theoretical framework, provided the foundational work on Boolean hypercubes and AI limits ^{[1][2]}, revised and polished the manuscript, and directed the overall intellectual development of the paper. Claude (Anthropic AI) contributed to drafting the initial text, structuring the argument, and expanding the connections to Gödel's incompleteness theorems and AlphaFold 2. The paper arose from an extended human-AI dialogue, and is itself an illustration of its central thesis.

References

1. ^a ^bCarbó-Dorca R (2020). "Cantor-like Infinity Sequences and Gödel-like Incompleteness Revealed by Means of Mersenne Infinite Dimensional Boolean Hypercube Concatenation." *J Math Chem.* 58:1–5.
2. ^a ^bCarbó-Dorca R (2021). "Boolean Hypercubes: The Origin of a Tagged Recursive Logic and the Limits of Artificial Intelligence." *Univ J Math Appl.* 4:41–49.
3. ^ΔShannon CE (1948). "A Mathematical Theory of Communication." *Bell Syst Tech J.* 27:379–423.
4. ^ΔTuring AM (1936). "On Computable Numbers, With an Application to the Entscheidungsproblem." *Proc Lond Math Soc.* 42:230–265.

5. ^aGödel K (1931). "Über Formal Unentscheidbare Sätze Der Principia Mathematica Und Verwandter Systeme I" [On Formally Undecidable Propositions of Principia Mathematica and Related Systems I]. *Monatsh Math Phys.* **38**:173–198.
6. ^a^bJumper J, et al. (2021). "Highly Accurate Protein Structure Prediction With AlphaFold." *Nature.* **596**:583–589.

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.