Intuition Is Almost All You Need

Danil Kutnyy
danil.kutny@gmail.com

Abstract

What is reasoning? What is logic? What is math? Common sense tells us that concepts such as numbers, relations, and logical structures feel inherently familiar—almost intuitive. They seem so obvious, but why? Do they have deeper origins? What is the number? What is addition? Why do they work in this way? Basic axioms of math, their foundation seems to be very intuitive, but absolutely mysteriously appear to the human mind out of nowhere. In a way their true essence magically slips away unnoticed from the consciousness, in an attempt to pinpoint its foundation. This paper delves into the fundamental nature of mathematics, logic, and rational reasoning, examining their "unreasonable effectiveness" in understanding and shaping the world. By drawing parallels between advancements in artificial neural networks and human cognition, the work introduces a novel perspective on intuition as the cornerstone of higher cognitive systems. Intuition is presented not only as a tool for pattern recognition but as the foundation upon which complex reasoning processes, such as mathematical abstraction and logical frameworks are constructed. This perspective redefines the role of consciousness, highlighting its capacity for innovation and exploration within abstract domains. Beyond theoretical insights, the paper outlines potential pathways for advancing artificial general intelligence by leveraging the interplay between intuition and conscious reasoning.

1. Mathematics, logic and reason

There exist many philosophical views on how to understand these ideas, though consensus remains elusive. However, it is safe to say that they can be seen as interconnected concepts (Quine, 1951)[1], or even as a structure within an abstraction ladder, where reasoning crystallizes logic, and logical principles lay the foundation for mathematics. Regardless of one's philosophical stance, we can be certain that all these concepts have proven to be immensely useful for humanity. Let us focus on mathematics for now, as a clear example of a mental tool used to explore, understand, and solve problems that would otherwise be beyond our grasp

Philosophical perspectives vary greatly on the origins and nature of these concepts. **Platonism**, for instance, posits that mathematical and logical truths exist independently in an abstract realm, waiting to be discovered [8]. On the other hand, **formalism**, **nominalism**, and **intuitionism** suggest that mathematics and logic are human constructs or mental frameworks, created to organize and describe our observations of the world [9][10][11]. Despite their differing perspectives, all of these theories acknowledge the utility and undeniable importance of mathematics and logic in shaping our understanding of reality. However, this very importance brings forth a paradox. While these concepts seem intuitively clear and integral to human thought, they also appear unfathomable in their essence. Their ultimate origins remain shrouded in mystery. Why do they exist in the form they do? What accounts for their effectiveness in understanding the world, and why do they so closely align with the structure of reality? These questions continue to haunt philosophers, mathematicians, and scientists alike.

Common sense tells us that concepts such as numbers, relations, and logical structures feel inherently familiar—almost intuitive. This aligns with many philosophical perspectives, from **intuitionism in**

mathematics [12] to **cognitive psychology** in understanding human reasoning [13]. Whether mathematics and logic are part of the very fabric of the universe or human inventions is still a topic of debate. What is certain, however, is that **intuition** plays a pivotal role in all approaches. Even within frameworks that emphasize the formal or symbolic nature of mathematics, intuition remains the cornerstone of how we build our theories and apply reasoning. Intuition allows us to recognize patterns, make judgments, and connect ideas in ways that might not be immediately apparent from the formal structures themselves. But the deeper we explore these formal structures, the more their foundations appear shrouded in vagueness and abstraction. Numbers and logical principles may seem clear and precise when applied, yet their fundamental nature resists easy explanation—why they "work" as they do is a question that remains stubbornly unresolved. This **elusiveness** is not merely a limitation of understanding but may hint at something deeper: that math and logic are tools shaped by our minds to make sense of a reality whose true nature might be **beyond full comprehension**. Intuition guides us through these murky waters, offering a bridge between the abstract formalism of mathematics and the practical reasoning we apply in the real world. Even when we lack a complete understanding of its underlying principles, intuition pushes us forward, enabling progress while leaving the ultimate truths of these structures tantalizingly out of reach.

2. Unreasonable Effectiveness..

Given such an important and mysterious topic, many great minds have independently touched on a fundamental truth about mathematics. The extraordinary usefulness of mathematics in human endeavors raises profound philosophical questions. Mathematics allows us to extend our intuitive understanding, solve problems beyond our mental capacity, and unlock insights into the workings of the universe. Yet, this remarkable utility gives rise to the "Unreasonable Effectiveness" problem: Why should abstract mathematical constructs, often developed with no practical application in mind, prove so indispensable in describing natural phenomena?

Eugene Wigner's seminal essay, *The Unreasonable Effectiveness of Mathematics in the Natural Sciences* (1960) [2], formalized this problem and explored possible answers. Wigner marveled at how mathematical concepts, often developed without practical intention, later proved essential for understanding the universe. This suggests that mathematics connects us to truths far beyond our intuitive grasp.

While our intuition, shaped by evolutionary constraints, provides a useful but limited framework for understanding the world, mathematics offers a way to formalize relationships and explore concepts that intuition alone cannot grasp. Real-world examples lend weight to Wigner's observations. For instance, non-Euclidean geometry, originally a purely theoretical construct, became foundational for Einstein's theory of general relativity, which redefined our understanding of spacetime [3]. Likewise, complex numbers, initially dismissed as "imaginary," are now indispensable in quantum mechanics and electrical engineering [4]. These cases exemplify how seemingly abstract mathematical frameworks can later illuminate profound truths about the natural world, reinforcing the idea that mathematics bridges the gap between human abstraction and universal reality. Galileo's observation that "the book of nature is written in the language of mathematics" [5] further highlights this unique capacity of mathematics to bridge the gap between human cognition and the natural world. Philosophers such as Henri Poincaré [6] have emphasized the creative role of mathematics, describing it as a way to synthesize relationships that intuition alone could not reveal. This creative synthesis, in turn, expands our capacity to engage with the universe, enabling breakthroughs that redefine our understanding of reality. Roger Penrose [7] takes this idea further, suggesting that mathematics exists independently of human cognition, offering access to truths that are neither constructed nor purely invented but instead discovered.

By systematically extending the reach of intuition and exposing us to truths beyond immediate comprehension, mathematics positions itself not merely as a tool but as a lens through which we can explore the limits of our understanding. This ability to formalize problems and derive solutions in ways that defy intuitive reasoning sets the stage for deeper inquiries into the relationship between abstract thought and the human capacity for innovation. Mathematics, logic, and reasoning occupy an essential place in our mental toolbox, yet their true nature remains elusive. Despite their extraordinary usefulness and centrality in human thought, these concepts often defy easy understanding or conceptualization within a single, universally accepted philosophical framework. While mathematics, logic, and reasoning are universally regarded as indispensable tools for problem-solving and innovation, reconciling their nature with a coherent philosophical theory presents a challenge.

3. Lens of Machine Learning

Let us turn to the emerging boundaries of the machine learning (ML) field to approach the philosophical questions we have discussed. In a manner similar to the dilemmas surrounding the foundations of mathematics, ML methods often produce results that are effective, yet remain difficult to fully explain or comprehend. While the fundamental principles of AI and neural networks are well-understood, the intricate workings of these systems—how they process information and arrive at solutions—remain elusive. This presents a symmetrically opposite problem to the one faced in the foundations of mathematics. We understand the underlying mechanisms, but the interpretation of the complex circuitry that leads to insights is still largely opaque. This paradox lies at the heart of modern deep neural network approaches, where we achieve powerful results without fully grasping every detail of the system's internal logic.

For a clear demonstration, let's consider a deep convolutional neural network (CNN) trained on the ImageNet classification dataset. ImageNet contains more than 14 million images, each hand-annotated into diverse classes (Deng, J. et al., 2009) [14]. The CNN is trained to classify each image into a specific category, such as "balloon" or "strawberry." After training, the CNN's parameters are fine-tuned to take an image as input. Through a combination of highly parallelizable computations, including matrix multiplication (network width) and sequential data processing (layer-to-layer, or depth), the network ultimately produces a probability distribution. High values in this distribution indicate the most likely class for the image. (Krizhevsky, A. et al., 2012) [15]

These network computations are rigid in the sense that the network takes an image of the same size as input, performs a fixed number of calculations, and outputs a result of the same size. This design ensures that for inputs of the same size, the time taken by the network remains predictable and consistent, reinforcing the notion of a "fast and automatic" process, where the network's response time is predetermined by its architecture. This means that such an intelligent machine cannot sit and ponder. This design works well in many architectures, where the number of parameters and the size of the data scale appropriately. A similar approach is seen in newer transformer architectures, like OpenAI's GPT series (Vaswani et al., 2017) (Radford, A et al., 2018) (Brown, T. B., A et al., 2020) [16][17][18]. By scaling transformers to billions of parameters and vast datasets, these models have demonstrated the ability to solve increasingly complex intelligent tasks.

With each new challenging task solved by such neural networks, the interoperability gap between a single parameter, a single neuron activation, and its contribution to the overall objective—such as predicting the next token—becomes increasingly vague. This sounds similar to the way the fundamental essence of math, logic, and reasoning appears to become more elusive as we approach it more closely.

To explain why this happens, let's explore how a neural network distinguishes between a cat and a dog in an image. Cat and dog images can be represented as single points in a vast, multidimensional space, with thousands to millions of dimensions, depending on the resolution. To distinguish between a cat and a dog, the neural network must process all these pixels simultaneously to identify key features. With wider and deeper neural networks, these pixels can be processed in parallel, enabling the network to perform enormous computations simultaneously to extract diverse features. As the network passes through multiple layers, it ascends the abstraction ladder—from recognizing basic elements like corners and lines to more complex shapes and gradients, then to textures (Zeiler, M. D., et al., 2014)[19]. In the upper layers, the network can work with high-level abstract concepts, such as "paw," "eye," "hairy," "wrinkled," or "fluffy."

The transformation from concrete pixel data to these abstract concepts is profoundly complex. Each group of pixels is weighted, features are extracted, and then summarized layer by layer for billions of times. Consciously deconstructing and grasping all the computations happening at once can be daunting. This gradual ascent from the most granular, concrete elements to the highly abstract ones using billions and billions of simultaneous computations is what makes the process so difficult to understand. The exact mechanism by which simple pixels are transformed into abstract ideas remains elusive, far beyond our cognitive capacity to fully comprehend.

4. Elusive foundations

This process surprisingly mirrors the challenge we face when trying to explore the fundamental principles of math and logic. Just as neural networks move from concrete pixel data to abstract ideas, our understanding of basic mathematical and logical concepts becomes increasingly elusive as we attempt to peel back the layers of their foundations. The deeper we try to probe, the further we seem to be from truly grasping the essence of these principles. This gap between the concrete and the abstract, and our inability to fully bridge it, highlights the limitations of both our cognition and our understanding of the most fundamental aspects of reality.

In addition to this remarkable coincidence, we've also observed a second astounding similarity: both neural networks processing and human foundational thought processes seem to operate almost instinctively, performing complex tasks in a rigid, timely, and immediate manner (given enough computation). Even advanced models like GPT-4 still operate under the same rigid and "automatic" mechanism as CNNs. GPT-4 doesn't pause to ponder or reflect on what it wants to write. Instead, it processes the input text, conducts N computations in time T and returns the next token, as well as the foundation of math and logic just seems to appear instantly out of nowhere to our consciousness.

This brings us to a fundamental idea that ties all the concepts together: intuition. Intuition, as we've explored, seems to be not just a human trait but a key component that enables both machines and humans to make quick and often accurate decisions, without consciously understanding all the underlying details. In this sense, Large Language Models (LLMs), like GPT, mirror the way intuition functions in our own brains. Just like our brains, which rapidly and automatically draw conclusions from vast amounts of data through what Daniel Kahneman calls *System 1* in *Thinking, Fast and Slow* (Kahneman, 2011)[20], LLMs process and predict the next token in a sequence based on learned patterns. These models, in their own way, are engaging in fast, automatic reasoning, without reflection or deeper conscious thought. This behavior, though it mirrors human intuition, remains elusive in its full explanation—just as the deeper mechanisms of mathematics and reasoning seem to slip further from our grasp as we try to understand them.

One more argument in support of such a position can be made. Obviously, natural neurons are vastly more complex than artificial ones, and this holds true for each layer of abstraction in both artificial and biological neural networks. Despite these differences, artificial neurons were developed specifically to model the computational processes of real neurons. The efficiency and success of artificial neural networks suggest that we have indeed captured some key features of their natural counterparts. Historically, our understanding of the brain has evolved alongside technological advancements. Early on, the brain was conceptualized as a simple stem mechanical system, then later as an analog circuit, and eventually as a computational machine akin to a digital computer. (Descartes, R., 1641)(Wiener, N., 1948)(Von Neumann, J., 1958)[21][22][23]. This shift in thinking reflects the changing ways we've interpreted the brain's functions in relation to emerging technologies. Despite the shortcomings in the evolution of these models, the striking similarities between artificial and natural neural networks make it hard to dismiss as coincidence. Given the efficiency and success of artificial networks in solving intelligent tasks, along with their ability to perform tasks similar to human cognition, it seems increasingly likely that both artificial and natural neural networks share underlying principles. While the details of their differences are still being explored, their functional similarities suggest they represent two variants of the single class of computational machines.

5. Limits of Intuition

Intuition is often celebrated as a mysterious tool of the human mind—an ability to make quick judgments and decisions without the need for conscious reasoning. It serves us well in everyday life, guiding us through complex social interactions, quick problem-solving, and the vast complexities of the world (Gigerenzer, G., et al., 1999) (Plantinga, A., 2000) (Haidt, J., 2001)[24][25][26]. However, as we explore increasingly sophisticated intellectual tasks—whether in mathematics, abstract reasoning, or complex problem-solving—intuition seems to reach its limits. While intuitive thinking can help us process patterns and make sense of known information, it falls short when faced with tasks that require deep, multi-step reasoning or the manipulation of abstract concepts far beyond our immediate experience (Kahneman, 2011)[20]. If intuition in humans is the same intellectual problem-solving mechanism as LLMs, then let's also explore the limits of LLMs. Can we see another intersection in the philosophy of mind and the emerging field of machine learning?

Despite their impressive capabilities in text generation, pattern recognition, and even some problem-solving tasks, LLMs are far from perfect and still struggle with complex, multi-step intellectual tasks that require deeper reasoning. While LLMs like GPT-3 and GPT-4 can process vast amounts of data and generate human-like responses, research has highlighted several areas where they still fall short. These limitations expose the weaknesses inherent in their design and functioning, shedding light on the intellectual tasks that they cannot fully solve or struggle with (Brown et al., 2020)[18].

1. **Multi-Step Reasoning and Complex Problem Solving:** One of the most prominent weaknesses of LLMs is their struggle with multi-step reasoning. While they excel at surface-level tasks, such as answering factual questions or generating coherent text, they often falter when asked to perform tasks that require multi-step logical reasoning or maintaining context over a long sequence of steps. For instance, they may fail to solve problems involving intricate mathematical proofs or multi-step arithmetic. Research by (Wei et al., 2022) [27] on the "chain-of-thought" approach, aimed at improving LLMs' ability to perform logical reasoning, shows that while LLMs can follow simple, structured reasoning paths, they still struggle with complex problem-solving when multiple logical steps must be integrated. In such cases, their performance often deteriorates, highlighting a significant gap in their cognitive capabilities compared to human reasoning.

2. **Abstract and Symbolic Reasoning:** Another significant challenge for LLMs lies in abstract reasoning and handling symbolic representations of knowledge. While LLMs can generate syntactically correct sentences and perform pattern recognition, they struggle when asked to reason abstractly or work with symbols that require logical manipulation outside the scope of training data. Tasks like proving theorems, solving high-level mathematical problems, or even dealing with abstract puzzles often expose LLMs' limitations. Research by Bender et al. (2021)[28] underscores this point, noting that while LLMs are adept at tasks grounded in large corpora of existing knowledge, they struggle with tasks that require the construction of new knowledge or systematic reasoning in abstract spaces.

3. **Understanding and Generalizing to Unseen Problems:** LLMs are, at their core, highly dependent on the data they have been trained on. While they excel at generalizing from seen patterns, they struggle to generalize to new, unseen problems that deviate from their training data. Yuan LeCun [29] argues that LLMs cannot get out of the scope of their training data. They have seen an enormous amount of data and, therefore, can solve tasks in a superhuman manner. But they seem to fall back with multi-step, complex problems. This lack of true adaptability is evident in tasks that require the model to handle novel situations that differ from the examples it has been exposed to. A 2023 study by Brown et al. examined this issue and concluded that LLMs, despite their impressive performance on a wide array of tasks, still exhibit poor transfer learning abilities when faced with problems that involve significant deviations from the training data.

4. **Long-Term Dependency and Memory:** LLMs have limited memory and are often unable to maintain long-term dependencies over a series of interactions or a lengthy sequence of information. This limitation becomes particularly problematic in tasks that require tracking complex, evolving states or maintaining consistency over time. For example, in tasks like story generation or conversation, LLMs may lose track of prior context and introduce contradictions or incoherence. The inability to remember past interactions over long periods highlights a critical gap in their ability to perform tasks that require dynamic memory and ongoing problem-solving. Raffel et al. (2020) [30] explored how transformer models, which underpin most LLMs, struggle with maintaining coherence over extended sequences.

Here, we can draw a parallel with mathematics and explore how it can unlock the limits of our mind and enable us to solve tasks that were once deemed impossible. For instance, can we grasp the Pythagorean Theorem? Can we intuitively calculate the volume of a seven-dimensional sphere? We can, with the aid of mathematics. One reason for this, as Searle and Hidalgo argue, is that we can only operate with a small number of abstract ideas at a time—fewer than ten (Searle, 1992)(Hidalgo, 2015)[31][32]. Comprehending the entire proof of a complex mathematical theorem at once is beyond our cognitive grasp. Sometimes, even with intense effort, our intuition cannot fully grasp it. However, by breaking it into manageable chunks, we can employ basic logic and mathematical principles to solve it piece by piece. When intuition falls short, reason takes over and paves the way. Yet, it seems strange that our powerful intuition, capable of processing thousands of details to form a coherent picture, cannot compete with mathematical tools. If, as Hidalgo posits, we can only process a few abstract ideas at a time, how does intuition fail so profoundly when tackling basic mathematical tasks?

6. Abstraction exploration mechanism

The answer may lie in the limitations of our computational resources and how efficiently we use them. Intuition, like large language models (LLMs), is a very powerful tool for processing familiar data and recognizing patterns. However, how can these systems—human intuition and LLMs alike—solve novel tasks and innovate? This is where the concept of abstract space becomes crucial. Intuition helps us create an abstract representation of the world, extracting patterns to make sense of it. However, it is not an all-powerful mechanism. As Fodor explains, human cognition relies on modular structures, each limited by its capacity for representation and reasoning (Fodor, 1983)[33]. Some patterns remain elusive even for intuition, necessitating new mechanisms, such as mathematical reasoning, to tackle more complex problems.

Similarly, LLMs exhibit limitations akin to human intuition. Bender et al. highlights that LLMs, despite their remarkable capabilities, fail to generate novel or abstract representations outside their training data (Bender et al., 2021)[34]. Moreover, the cognitive limits described by Wittgenstein remind us that the rules governing reasoning are not always intuitive and often require external structures like mathematical systems to navigate complex abstractions (Wittgenstein, 1953)[35]. Ultimately, the gap between intuition and mathematical tools illustrates the necessity of augmenting human cognition with external mechanisms. As Kant argued, mathematics provides the structured framework needed to transcend the limits of human understanding (Kant, 1781)[36]. By leveraging these tools, we can push beyond the boundaries of intuition to solve increasingly intricate problems.

What if, instead of trying to search for solutions in a highly complex world with an unimaginable degree of freedom, we could reduce it to essential aspects? Abstraction is such a tool. As discussed earlier, the abstraction mechanism in the brain (or an LLM) can extract patterns from patterns and climb high up the abstraction ladder. In this space of high abstractions, created by our intuition, the basic principles governing the universe can be crystallize. Logical principles and rational reasoning become the intuitive foundation constructed by the brain while extracting the essence of all the diverse data it encounters. These principles, later formalized as mathematics or logic, are actually the map of a real world. Intuition arises when the brain takes the complex world and creates an abstract, hierarchical, and structured representation of it, it is the purified, essential part of it—a distilled model of the universe as we perceive it. Only then, basic and intuitive logical and mathematical principles emerge. At this point simple scaling of computation power to gain more patterns and insight is not enough, there emerges a new more efficient way of problem-solving from which reason, logic and math appear.

When we explore the entire abstract space and systematize it through reasoning, we uncover corners of reality represented by logical and mathematical principles. This helps explain the "unreasonable effectiveness"(Wigner, 1960)[2] of mathematics. No wonder it is so useful in the real world, and surprisingly, even unintentional mathematical exploration becomes widely applicable. These axioms and basic principles, manipulations themselves represent essential patterns seen in the universe, patterns that intuition has brought to our consciousness. Due to some kind of computational limitations or other limitations of intuition of our brains, it is impossible to gain intuitive insight into complex theorems. However, these theorems can be discovered through mathematics and, once discovered, can often be reapplied in the real world. This process can be seen as a top-down approach, where conscious and rigorous exploration of abstract space—governed and grounded by mathematical principles—yields insights that can be applied in the real world. These newly discovered abstract concepts are in fact rooted in and deeply connected to reality, though the connection is so hard to spot that it cannot be grasped, even the intuition mechanism was not able to see it.

7. Reinterpreting of consciousness

The journey from intuition to logic and mathematics invites us to reinterpret the role of consciousness as the bridge between the automatic, pattern-driven processes of the mind and the deliberate, structured exploration of abstract spaces. Latest LLMs achievement clearly show the power of intuition alone, that does not require resigning to solve very complex intelligent tasks(Bengio, 2019)[37].

Consciousness is not merely a mechanism for integrating information or organizing patterns into higher-order structures—that is well within the realm of intuition. Intuition, as a deeply powerful cognitive tool, excels at recognizing patterns, modeling the world, and even navigating complex scenarios with breathtaking speed and efficiency. It can uncover hidden connections in data often better and generalize effectively from experience (Kahneman, 2011)[38]. However, intuition, for all its sophistication, has its limits: it struggles to venture beyond what is already implicit in the data it processes. It is here, in the domain of exploring abstract spaces and innovating far beyond existing patterns where new emergent mechanisms become crucial, that consciousness reveals its indispensable role.

At the heart of this role lies the idea of **agency**. Consciousness doesn't just explore abstract spaces passively—it creates agents capable of acting within these spaces. These agents, guided by reason-based mechanisms, can pursue long-term goals, test possibilities, and construct frameworks far beyond the capabilities of automatic intuitive processes. This aligns with Dennett’s notion of consciousness as an agent of intentionality and purpose in cognition (Dennett, 1991)[39]. Agency allows consciousness to explore the landscape of abstract thought intentionally, laying the groundwork for creativity and innovation. This capacity to act within and upon abstract spaces is what sets consciousness apart as a unique and transformative force in cognition.

Unlike intuition, which works through automatic and often subconscious generalization, consciousness enables the deliberate, systematic exploration of possibilities that lie outside the reach of automatic processes. This capacity is particularly evident in the realm of mathematics and abstract reasoning, where intuition can guide but cannot fully grasp or innovate without conscious effort. Mathematics, with its highly abstract principles and counterintuitive results, requires consciousness to explore the boundaries of what intuition cannot immediately "see." In this sense, consciousness is a specialized tool for exploring the unknown, discovering new possibilities, and therefore forging connections that intuition cannot infer directly from the data.

Philosophical frameworks like **Integrated Information Theory (IIT)** can be adapted to resonate with this view. While IIT emphasizes the integration of information across networks (Tononi, 2004)[40], such new perspective would argue that integration is already the forte of intuition. Consciousness, in contrast, is not merely integrative—it is exploratory. It allows us to transcend the automatic processes of intuition and deliberately engage with abstract structures, creating new knowledge that would otherwise remain inaccessible. The power of consciousness lies not in refining or organizing information but in stepping into uncharted territories of abstract space.

Similarly, **Predictive Processing Theories**, which describe consciousness as emerging when the brain's predictive models face uncertainty or ambiguity (Clark, 2013)[41], can align with this perspective when reinterpreted. Where intuition builds models based on the data it encounters, consciousness intervenes when those models fall short, opening the door to innovations that intuition cannot directly derive. Consciousness is the mechanism that allows us to work in the abstract, experimental space where logic and reasoning create new frameworks, independent of data-driven generalizations.

Other theories, such as **Global Workspace Theory (GWT)** (Baars, 1988)[42] and **Higher-Order Thought Theories** (Rosenthal, 2005)[43], may emphasize consciousness as the unifying stage for subsystems or the reflective process over intuitive thoughts, but again, powerful intuition perspective shifts the focus. Consciousness is not simply about unifying or generalize—it is about transcending. It is the mechanism that allows us to "see" beyond the patterns intuition presents, exploring and creating within abstract spaces that intuition alone cannot navigate.

Agency completes this picture. It is through agency that consciousness operationalizes its discoveries, bringing abstract reasoning to life by generating actions, plans, and make innovations possible. Intuitive processes alone, while brilliant at handling familiar patterns, are reactive and tethered to the data they process. Agency, powered by consciousness, introduces a proactive, goal-oriented mechanism that can conceive and pursue entirely new trajectories. This capacity for long-term planning, self-direction, and creative problem-solving is a part of what elevates consciousness from intuition and allows for efficient exploration.

In this way, consciousness is not a general-purpose cognitive tool like intuition but a highly specialized mechanism for innovation and agency. It plays a relatively small role in the broader context of intelligence, yet its importance is outsized because it enables the exploration of ideas and the execution of actions far beyond the reach of intuitive generalization. Consciousness, then, is the spark that transforms the merely "smart" into the truly groundbreaking, and agency is the engine that ensures its discoveries shape the world.

8. Predictive Power of the Theory

This theory makes several key predictions regarding cognitive processes, consciousness, and the nature of innovation. These predictions can be categorized into three main areas:

1. **Predicting the Role of Consciousness in Innovation:**

   The theory posits that high cognitive abilities, like abstract reasoning in mathematics, philosophy, and science, are uniquely tied to conscious thought. Innovation in these fields requires deliberate, reflective processing to create models and frameworks beyond immediate experiences. This capacity, central to human culture and technological advancement, eliminates the possibility of philosophical zombies—unconscious beings—as they would lack the ability to solve such complex tasks, given the same computational resource as the human brain.

2. **Predicting the Limitations of Intuition:**

   In contrast, the theory also predicts the limitations of intuition. Intuition excels in solving context-specific problems—such as those encountered in everyday survival, navigation, and routine tasks—where prior knowledge and pattern recognition are most useful. However, intuition's capacity to generate novel ideas or innovate in highly abstract or complex domains, such as advanced mathematics, theoretical physics, or the development of futuristic technologies, is limited. In this sense, intuition is a powerful but ultimately insufficient tool for the kinds of abstract thinking and innovation necessary for transformative breakthroughs in science, philosophy, and technology.

3. **The Path to AGI: Integrating Consciousness and Abstract Exploration**

There is one more crucial implication of the developed theory: it provides a pathway for the creation of Artificial General Intelligence (AGI), particularly by emphasizing the importance of consciousness, abstract exploration, and non-intuitive mechanisms in cognitive processes. Current AI models, especially transformer architectures, excel in pattern recognition and leveraging vast amounts of data for tasks such as language processing and predictive modeling. However, these systems still fall short in their ability to innovate and rigorously navigate the high-dimensional spaces required for creative problem-solving. The theory predicts that achieving AGI and ultimately superintelligence requires the incorporation of mechanisms that mimic conscious reasoning and the ability to engage with complex abstract concepts that intuition alone cannot grasp.

The theory suggests that the key to developing AGI lies in the integration of some kind of a recurrent, or other adaptive computation time mechanism on top of current architectures. This could involve augmenting transformer-based models with the capacity to perform more sophisticated abstract reasoning, akin to the conscious, deliberative processes found in human cognition. By enabling AI systems to continually explore high abstract spaces and to reason beyond simple pattern matching, it becomes possible to move towards systems that can not only solve problems based on existing knowledge but also generate entirely new, innovative solutions—something that current systems struggle with

9. Conclusion

This paper has explored the essence of mathematics, logic, and reasoning, focusing on the core mechanisms that enable them. We began by examining how these cognitive abilities emerge and concentrating on their elusive fundamentals, ultimately concluding that **intuition** plays a central role in this process. However, these mechanisms also allow us to push the boundaries of what intuition alone can accomplish, offering a structured framework to approach complex problems and generate new possibilities.

We have seen that **intuition** is a much more powerful cognitive tool than previously thought, enabling us to make sense of patterns in large datasets and to reason within established frameworks. However, its limitations become clear when scaled to larger tasks—those that require a departure from automatic, intuitive reasoning and the creation of new concepts and structures. In these instances, **mathematics** and **logic** provide the crucial mechanisms to explore abstract spaces, offering a way to formalize and manipulate ideas beyond the reach of immediate, intuitive understanding.

Finally, our exploration has led to the idea that **consciousness** plays a crucial role in facilitating non-intuitive reasoning and abstract exploration. While intuition is necessary for processing information quickly and effectively, **consciousness** allows us to step back, reason abstractly, and consider long-term implications, thereby creating the foundation for **innovation** and **creativity**. This is a crucial step for the future of **AGI** development. Our theory predicts that consciousness-like mechanisms—which engage abstract reasoning and non-intuitive exploration—should be integrated into AI systems, ultimately enabling machines to innovate, reason, and adapt in ways that mirror or even surpass human capabilities.

References:

1. *W. V. O. Quine (1951), Two Dogmas of Empiricism, The Philosophical Review, vol. 60, no. 1, pp. 20–43*

2. Wigner, E. P. (1960). *The unreasonable effectiveness of mathematics in the natural sciences*. Communications on Pure and Applied Mathematics, 13(1), 1–14.

1. Einstein, A. (1915). *Die Feldgleichungen der Gravitation*. Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften, 844–847.

2. Niven, I., Zuckerman, H., & Montgomery, H. L. (1991). *An Introduction to the Theory of Numbers* (5th ed.). Wiley.

3. Galileo Galilei. (1620). *Il Saggiatore* (The Assayer).

3. Poincaré, H. (1905). *Science and Hypothesis*. Dover Publications.

4. Penrose, R. (2004). *The Road to Reality: A Complete Guide to the Laws of the Universe*. Alfred A. Knopf.

5. *Plato. (380 B.C.). The Republic (trans. B. Jowett). Oxford University Press. (Original work published c. 380 B.C.).*

6. *Hilbert, D. (1900). Mathematical Problems. Bulletin of the American Mathematical Society, 8(1), 1–10.*

7. Carnap, R. (1934). *The Logical Syntax of Language*. Routledge & Kegan Paul.

8. Wigner, E. P. (1960). *The Unreasonable Effectiveness of Mathematics in the Natural Sciences*. Communications on Pure and Applied Mathematics, 13(1), 1–14.

9. Heyting, A. (1956). *Intuitionism: An Introduction*. North-Holland Publishing.

10. Piaget, J. (1952). *The Origins of Intelligence in Children*. International Universities Press.

11. *Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. In 2009 IEEE Conference on Computer Vision and Pattern Recognition (pp. 248–255). IEEE.*

12. *Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems (NeurIPS), 1097-1105. Link to Paper*

13. *Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. A., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems (NeurIPS), 30.*

14. *Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training. OpenAI.*

15. *Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., & Amodei, D. (2020). Language models are few-shot learners. In Proceedings of the 34th Conference on Neural Information Processing Systems (pp. 1877–1901).*

16. Zeiler, M. D., & Fergus, R. (2014). *Visualizing and understanding convolutional networks*. In *European Conference on Computer Vision* (pp. 818–833). Springer.

17. *D. Kahneman, Thinking, Fast and Slow, Farrar, Straus and Giroux, 2011.*

18. Descartes, R. (1641). *Meditations on First Philosophy*.

19. Wiener, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*

20. Von Neumann, J. (1958). *The Computer and the Brain*.

21. *Gigerenzer, G., & Todd, P. M. (1999). Simple Heuristics That Make Us Smart. Oxford University Press.*

22. *Plantinga, A. (2000). Warranted Christian Belief. Oxford University Press.*

23. Haidt, J. (2001). *The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment*. *Psychological Review*, 108(4), 814–834.

24. *Wei, J., et al., "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models," Proceedings of the 2022 Conference on Neural Information Processing Systems (NeurIPS 2022).*

25. *Bender, E. M., et al., "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?", Proceedings of the 2021 Conference on Fairness, Accountability, and Transparency.*

26. *LeCun, Y., "A Path Toward Autonomous Machine Intelligence," Frontiers in AI, vol. 3, pp. 1-12, 2020.*

27. *Raffel, C., et al., "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer," Journal of Machine Learning Research, vol. 21, no. 140, pp. 1-67, 2020.*

28. Searle, J. R., *The Rediscovery of the Mind*, MIT Press, 1992.

29. Hidalgo, C. A., *Why Information Grows: The Evolution of Order, from Atoms to Economies*, Basic Books, 2015.

30. Fodor, J. A., *The Modularity of Mind: An Essay on Faculty Psychology*, MIT Press, 1983.

31. *Bender, E. M., et al., "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?", Proceedings of the 2021 Conference on Fairness, Accountability, and Transparency.*

32. Wittgenstein, L., *Philosophical Investigations*, Blackwell, 1953.

33. Kant, I., Critique of Pure Reason, Cambridge University Press, 1781.

34. *Bengio, Y., "The Consciousness Prior," arXiv preprint arXiv:1709.08568, 2019.*

35. *Kahneman, D., Thinking, Fast and Slow, Farrar, Straus and Giroux, 2011.*

36. Dennett, D. C., *Consciousness Explained*, Little, Brown, 1991.

37. *Tononi, G., "An information integration theory of consciousness," BMC Neuroscience, vol. 5, no. 42, 2004.*

38. Clark, A., *Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science*, Behavioral and Brain Sciences, vol. 36, no. 3, pp. 181-204, 2013.

39. *Baars, B. J., A Cognitive Theory of Consciousness, Cambridge University Press, 1988.*

40. *Rosenthal, D. M., Consciousness and Mind, Oxford University Press, 2005.*