On the Morphic Problem in Artificial Neural Networks

1[0000−1111−2222−3333]

[1] University of xxx@unixx.com
http://www.springer.com/gp/computer-science/lncs
[2] ABC Institute, xxxUniversity xxx, xxx
{abc,lncs}@unixxx

Abstract. The paper focuses on the convergence between artificial neural networks (ANNs) and biological neural systems, addressing the challenges of establishing a "morphic relation" between the two. The central problem lies in replicating biological neural networks' dynamic, adaptive, and self-organizing properties within artificial constructs. Neuromorphic engineering (NE), an interdisciplinary field at the intersection of neuroscience and computer science, seeks to design ANNs that emulate biological systems' structure, function, and temporal dynamics. While ANNs have succeeded in areas like pattern recognition and natural language processing, they often need more fluid adaptability and robustness of biological systems. The paper explores recent advances in deep learning models, particularly deep neural networks (DNNs), and their ability to capture structure-sensitive cognitive properties. Challenges remain despite promising findings, such as meta-learning techniques and systematic generalization. DNNs, though efficient, often exhibit opaque and fragile learning mechanisms. The paper advocates further exploring the criteria to establish a genuine morphic relation between ANNs and biological neural systems, emphasizing structural, functional, and dynamic correspondences to advance the neuromorphic engineering field.

Keywords: Artificial Neural Networks · Vector Grounding Problem · Morphic Problem · Neuromorphic Engineering.

1 Introduction

The development of the connectionist paradigm, especially through modern deep neural networks (DNNs), has profoundly impacted long-standing debates in cognitive science and the philosophy of mind. Historically, connectionism faced sharp critiques from proponents of the language of thought hypothesis. Critics argued that connectionist models either failed to represent cognition adequately or merely implemented classical symbol manipulation without offering a genuine alternative (Fodor and McLaughlin, 1990; Fodor and Pylyshyn, 1988; Pinker and Prince, 1988). They contended that connectionist models could not capture essential structure-sensitive properties of cognition, such as systematicity (the capacity to understand and produce an infinite number of sentences) and productivity (the ability to generate and comprehend novel sentences), largely due to the absence of compositional representations in these models.

In the current era of deep learning, this debate has resurfaced with renewed vigor. Critics argue that if DNNs can match human-level performance across a range of cognitive tasks, they must possess features central to language-of-thought architectures (Marcus, 2018; Quilty-Dunn et al., 2023). However, proponents of connectionism suggest that human cognition may not be as strictly rule-governed as classicists claim. They argue that connectionist models can account for structure-sensitive cognitive properties without fully replicating the architecture of symbolic models (Johnson, 2004; Smolensky, 2022; Elman, 2019).

Recent advancements in DNNs, mainly through studies on systematic compositional generalization, offer promising evidence that connectionist models can bridge the gap with classical theories of cognition. For instance, Lake and Baroni (2023) demonstrated that a transformer-based network, trained via meta-learning on various tasks, achieved systematic generalization akin to human-like learning in few-shot tasks. This meta-learning network involved training on a stream of artificial tasks with an underlying "interpretation grammar", achieving high accuracy and human-like error patterns without explicit compositional rules. Similarly, Murty et al. (2023) found that training a network beyond optimal accuracy on training data could lead to better generalising hierarchical rules. These findings suggest that modern DNNs may approximate the structure-sensitive properties of cognition without relying on explicit symbolic manipulation. However, this interpretation depends on assumptions about implementing cognitive properties in DNNs and their relevance to broader cognitive theories (McGrath et al., 2023; Pavlick, 2023). Moreover, mechanistic interpretability research reveals that while DNNs can acquire variable binding mechanisms, these remain "fuzzy" and are not functionally equivalent to the discrete symbol manipulation found in classical systems (Olsson et al., 2022). Consequently, while DNNs can compute over compositional representations with constituent structure, this structure is non-classical and exhibits degrees of role-filler independence, challenging classical notions of cognition.

Now, the paradigm seems to shift again towards Neuromorphic Engineering (NE), an interdisciplinary field at the intersection of neuroscience, computer science, and electrical engineering, dedicated to designing artificial systems that emulate the architecture and dynamics of biological neural networks. Originating from the desire to overcome the limitations of traditional computing paradigms, neuromorphic engineering seeks to create hardware and algorithms inspired by the nervous system's efficiency, adaptability, and robustness. By replicating neural structures and processes, this field aims to develop artificial neural networks (ANNs) that perform complex computations with greater efficiency and capture the essence of biological cognition. The morphic relation between ANNs and biological neural systems is central to neuromorphic engineering. This relation involves establishing correspondences in structure, function, and dynamics between artificial and natural neural networks. The goal is not merely to mimic superficial aspects of the brain but to embody the fundamental principles that underlie neural computation and learning. Neuromorphic systems aspire to replicate features such as synaptic plasticity, spike-timing-dependent processing, and energy-efficient information transmission: these are properties that give biological neural networks their remarkable capabilities. Neuromorphic engineering also confronts the morphic problem: the difficulty of fully capturing the adaptive, context-dependent, and self-organizing properties of biological neural networks within artificial constructs. While ANNs have achieved impressive feats in pattern recognition, language processing, and even game playing, they often lack the fluid adaptability and resilience of their biological counterparts. Biological neurons operate within a rich milieu of biochemical signals, structural plasticity, and environmental interactions that are challenging to replicate artificially.

The pursuit of a morphic relation necessitates satisfying specific criteria, including structural correspondence (mirroring neural architectures), functional equivalence (replicating neural operations), dynamic coherence (aligning temporal behaviors), contextual integration (embedding within larger systems), and robustness (maintaining functionality under stress). Neuromorphic engineering endeavors to meet these criteria, pushing the boundaries of how closely artificial systems can approximate biological ones. So, this paper aims to foster the analysis of the conceptual challenges posed by the potential morphic relation between artificial neural networks (ANNs) and biological neural systems. Specifically, it focuses on identifying and articulating the criteria that are necessary and sufficient to establish such a morphic relation.

2 Deep learning models and cognition

Deep learning models have emerged as a significant focus in cognitive science, sparking optimism and scepticism about their potential to serve as cognitive models. These models, notably deep neural networks (DNNs), offer a new lens through which to examine cognitive processes, primarily due to their ability to learn and generalise across vast amounts of data, sometimes in ways that parallel human cognition. However, this promise is met with considerable debate as the capabilities and limitations of DNNs in truly capturing cognitive phenomena are scrutinised. Understanding the strengths and weaknesses of deep learning models as cognitive models requires a deep detour into their underlying mechanisms, their performance in cognitive tasks, and the philosophical implications of their use in modelling cognition.

One of the primary advantages of deep learning models as cognitive models is their ability to handle vast amounts of data and learn complex patterns without requiring explicit programming of rules or symbolic structures. This characteristic is particularly significant in cognitive science, where human cognition involves processing and integrating vast amounts of sensory and experiential data. For instance, DNNs have shown remarkable success in visual recognition, natural language processing, and game playing, often achieving or surpassing human-level performance. These successes suggest that DNNs can, in some respects, model human cognition's flexible, adaptive, and non-linear nature. Unlike classical symbolic models, which often rely on pre-defined rules and structures, DNNs learn representations and associations directly from data, which more closely mirrors the learning processes observed in humans. This ability to model complex, non-linear relationships makes DNNs powerful tools for simulating cognitive processes such as perception, language understanding, and decision-making.

Moreover, DNNs have demonstrated an ability to generalise from limited data, capturing human-like reasoning and learning aspects. For example, research has shown that with appropriate training regimes, such as meta-learning, DNNs can achieve systematic generalisation—an ability previously thought to be a hallmark of symbolic cognitive architectures. The work of Lake and Baroni (2023) is particularly illuminating in this regard, as they demonstrated that transformer-based networks could achieve human-like generalisation in few-shot learning tasks. It suggests that, under certain conditions, DNNs can mimic human cognition's compositional and systematic properties, challenging the long-held belief that such properties are exclusive to symbolic models. The potential for DNNs to model cognitive tasks without relying on pre-defined rules opens new avenues for understanding how cognitive processes might be implemented in the brain, offering a more biologically plausible alternative to classical models.

However, the advantages of DNNs as cognitive models come with significant caveats. One of the primary criticisms is that while DNNs can achieve human-like performance on specific tasks, the mechanisms by which they do so are often opaque and difficult to interpret. This "black box" nature of deep learning models contrasts sharply with the transparency of classical symbolic models, where the rules and operations are explicitly defined and easily understood. The lack of interpretability in DNNs raises concerns about their utility as cognitive models, as it becomes challenging to map the learned representations and processes in DNNs onto known cognitive and neural mechanisms. If cognitive modelling aims to replicate human behaviour and explain the underlying processes, then the opacity of DNNs presents a significant obstacle. Understanding how DNNs process information, make decisions and generalise requires complex techniques from mechanistic interpretability, which is still in its infancy and far from providing a comprehensive understanding of these models.

Another significant limitation of DNNs as cognitive models is their reliance on vast amounts of data and computational resources, which may not reflect how human cognition operates. Human learners often generalise from a few examples, demonstrating a remarkable ability to learn efficiently and robustly in environments where data is scarce or noisy. In contrast, DNNs typically require large datasets to achieve high performance, and their generalisation abilities can be fragile, especially when faced with out-of-distribution examples or adversarial attacks. This discrepancy raises questions about the ecological validity of DNNs as models of human cognition. If human cognition is characterised by its efficiency and robustness in learning from limited data, then models that require extensive data and careful tuning may not accurately capture the essence of human cognitive processes.

Additionally, while DNNs have shown promise in modelling specific cognitive tasks, they may fall short in capturing higher-order cognitive processes such as abstract reasoning, theory of mind, and complex decision-making. These processes often involve symbolic manipulation, logical reasoning, and the ability to understand and generate explanations—capacities that are traditionally associated with classical cognitive architectures. While there have been advances in integrating symbolic reasoning with deep learning (e.g., neural-symbolic systems), these approaches are still in their early stages and have yet to fully demonstrate that DNNs can model the full range of human cognitive abilities. The challenge lies in bridging the gap between the associative, pattern-based learning that DNNs excel at and the more structured, rule-based reasoning that characterises much of human cognition.

Thus, using DNNs as cognitive models raises essential questions about the nature of representation and computation in cognitive systems. Critics argue that DNNs, even when they achieve human-like performance, may do so in fundamentally different ways from how the human brain operates. For instance, while DNNs can learn compositional representations, these representations are often "fuzzy" and lack the discrete, rule-governed structure that characterises classical symbolic representations. This fuzziness challenges the classical notion of cognitive architectures based on clear, well-defined symbols and rules, leading to debates about whether DNNs can be considered cognitive models or represent a fundamentally different kind of computation. Whether DNNs implement a form of the "language of thought" or offer a new paradigm for understanding cognition remains an open and contentious issue in cognitive science.

In conclusion, deep learning models, particularly DNNs, offer exciting possibilities and significant challenges as cognitive models. Their ability to learn from data and generalise across tasks aligns them with crucial aspects of human cognition, making them valuable tools for exploring cognitive processes. However, their opacity, data requirements, and potential misalignment with higher-order cognitive functions highlight the limitations of these models. As research advances, DNNs will likely play an increasingly important role in cognitive science, but whether they can fully capture the complexities of human cognition or need to be integrated with other approaches remains a critical question. The ongoing dialogue between proponents of connectionist and symbolic models will likely shape the future of cognitive science as researchers strive to develop models that not only replicate human behaviour but also provide insights into the underlying mechanisms of the mind.

3 The neuromorphic engineering project

The Neuromorphic Engineering (NE) project explores the dynamic interplay between natural neural networks (NNNs) and artificial neuromorphic networks (ANNs). Central to this project is the "morphing relation" between NNNs and ANNs, which is not merely technical but profoundly philosophical, raising questions about the nature of intelligence, computation, and consciousness.

According to Tsur (2022), at the heart of NE lies the comparison between biological neural networks and artificial spiking neural networks (SNNs). NNNs operate on principles of real-time adaptive learning, energy efficiency, and decentralised processing. Tsur emphasises that while these biological systems are optimised for sparse communication and energy efficiency, ANNs, particularly the spiking variety, are striving to replicate this in artificial hardware. The challenge, however, is more than replicating neural architecture—it is about understanding how the mechanisms underlying NNNs can inspire innovations in ANNs.

The philosophical problem arises when we consider the morphing of concepts from biology into technology. NNNs evolved over millions of years, optimising for survival, adaptability, and energy efficiency. By contrast, ANNs are designed by humans for human purposes—typically to maximise performance on specific computational tasks. In attempting to bridge this gap, Tsur notes that while artificial networks like SNNs can mimic certain aspects of their biological counterparts (such as spike-timing-dependent plasticity or STDP), they do so in fundamentally different contexts. The context of biological intelligence is survival and adaptability in a dynamic world; the context of artificial intelligence is task optimisation within a designed framework.

One of the philosophical issues Tsur (2022) raises is whether the two systems—biological and artificial—are converging towards the same end. If artificial networks continue to improve, will they one day exhibit the same kind of intelligence that biological systems possess? Or is the difference between them irreconcilable? Tsur does not provide a definitive answer but leaves the question open, inviting readers to reflect on the nature of intelligence and the future of neuromorphic engineering.

The technological side of the debate focuses on practical implementations. Tsur delves into how SNNs are constructed using principles of neuromorphic hardware, including projects like IBM's TrueNorth chip or Intel's Loihi chip. These systems use spiking neurons to mimic biological neurons ' asynchronous, event-driven nature. However, the efficiency of such systems is still far from matching biological systems, particularly in terms of power consumption and adaptability. Tsur points out that while systems like Loihi have shown promise, they are still in the early stages of development compared to the complexity of biological neural systems (Tsur, 2022). He also explores the limitations of neuromorphic computing, particularly regarding energy efficiency and scalability. He acknowledges that while SNNs offer advantages for certain types of computation—such as real-time pattern recognition and low-power operations—they are not suitable for all computational tasks. The morphing relation between NNNs and ANNs, then, is not one of seamless integration but one of careful balancing, where the strengths of each system are leveraged for specific tasks. Biological systems may never be fully replicable in artificial hardware, but their inspiration is invaluable for advancing both fields (Tsur, 2022).

4 From the vector grounding problem to the morphic problem

This section discusses a problem similar to the Symbol Grounding Problem in AI but in the context of connectionist systems like Large Language Models (LLMs) that use vectors instead of symbols, and its connection to the morphic problem. This issue is called the Vector Grounding Problem (Coehlo Mollo and Millière, 2023).

The Vector Grounding Problem refers to the challenge faced by LLMs, which use vectors as numerical representations of text tokens based on statistical relationships. Despite their advanced capabilities, these models manipulate these vectors without any intrinsic connection to real-world entities or meanings, leading to outputs that appear meaningful to humans but are inherently ungrounded in reality. The problem highlights that LLMs, much like symbolic AI systems, struggle to connect their internal representations (vectors) to the external world, rendering their outputs devoid of intrinsic meaning. As artificial neural networks (ANNs), LLMs like GPT-3 operate by processing vectors, yet questions arise regarding whether these systems can produce genuinely meaningful outputs despite lacking direct interaction with the real world. This problem mirrors the earlier "Symbol Grounding Problem", which questioned the capacity of classical AI systems to ground symbols in real-world referents.

Coehlo Mollo and Millière differentiate five distinct types of grounding discussed in the literature: referential, sensorimotor, relational, communicative, and epistemic grounding. While these terms are often used interchangeably, the paper argues that referential grounding, i.e. the capacity of internal representations to refer to actual entities in the world, is central to the Vector Grounding Problem.

The authors suggest that some LLMs, especially those fine-tuned with Reinforcement Learning from Human Feedback (RLHF), possess the necessary features to overcome this challenge. Despite not being embodied or directly multimodal, these models maintain causal-historical relationships with the world through their training data, which enables them to generate outputs that bear intrinsic meaning. This claim challenges the assumption that multimodality (using inputs beyond text) or embodiment is required for an AI system to ground its representations meaningfully.

Ultimately, the paper argues that current advancements in LLMs show promise in addressing the Vector Grounding Problem, implying that artificial systems can potentially develop meaningful representations of the world despite their disembodied nature.

The Morphic Problem of Artificial Neural Networks (MPANNs) emerges as a compelling challenge in the pursuit of optimising ANNs by mimicking the structure and functions of biological neural systems. At its core, the morphic program seeks to enhance the power and efficiency of ANNs by drawing inspiration from the remarkable capabilities of the human brain, which processes information with unparalleled efficiency, adaptability, and generalisation. However, this approach introduces a paradox: If ANNs, like their symbolic AI predecessors, lack a direct relationship with the real world, i.e. a relationship that is fundamental to the functioning of biological brains, how can we effectively optimise these networks through a morphic program? This question strikes at the heart of both the potential and the limitations of biologically inspired AI.

Biological neural networks, such as the human brain, are deeply embedded in and influenced by their environment. They develop and refine their connections through continuous interactions with the real world, gaining meaning and understanding through sensory experiences, embodied action, and evolutionary processes. The brain's ability to ground its representations in the real world is not just a byproduct of its complex structure but also a result of millions of years of evolution, during which neural architectures were honed to solve specific survival-related problems in dynamic environments. This grounding is what enables humans to perceive, understand, and respond to their surroundings in a meaningful way. Thus, the brain's architecture is not merely a computational tool but a deeply integrated system that connects abstract representations to concrete experiences.

In contrast, ANNs are artificial constructs that, while inspired by the brain's architecture, operate within fundamentally different contexts. Unlike biological systems, ANNs are trained on data sets that are often divorced from the sensory-rich and dynamic environments that shape human cognition. These networks learn to recognise patterns and make predictions based on vast amounts of data. However, they do so mainly in statistical and syntactic ways, lacking the semantic grounding that biological neural networks possess. The ANNs' vectors, representing data points or features, are mathematical abstractions that may correspond to patterns in the training data but do not inherently link to real-world entities or experiences. As a result, the outputs generated by ANNs, while sometimes impressively accurate or human-like, are ultimately ungrounded like biological neural responses.

Therefore, the morphic program, which aims to optimise ANNs by mimicking the structure and functionality of biological brains, must grapple with this fundamental disconnection. If the goal is to create ANNs that approach the versatility and adaptability of biological neural networks, merely copying the brain's brain's structure may not suffice. The brain's power lies not only in its architecture but also in its deep integration with the sensory and experiential world—a connection that ANNs currently lack. This raises a crucial question: Can the morphic program succeed in optimising ANNs without somehow embedding them in a world that allows for meaningful interactions and grounding of representations?

One potential approach to resolving this issue is to focus on developing ANNs that incorporate embodied cognition elements. This would involve designing systems where artificial neural networks are modelled after the brain's architecture and embedded in environments that allow real-world interaction. For instance, integrating ANNs with robotics or other forms of embodied agents could provide a platform for these networks to develop more grounded and contextually relevant representations. By engaging with the physical world, these networks could begin to form associations between their internal vectors and the external entities they represent, much like how a child learns about the world through sensory experiences and interaction.

Moreover, the morphic program could explore the incorporation of multimodal learning, where ANNs are trained on diverse data types that simulate the sensory inputs received by biological organisms. This could involve integrating visual, auditory, tactile, and proprioceptive data, allowing ANNs to develop more prosperous and interconnected representations. Such an approach might help bridge the gap between the statistical patterns recognised by ANNs and the meaningful, grounded representations seen in biological systems.

However, even with these strategies, the question remains whether ANNs can genuinely achieve the kind of grounding that biological neural networks possess. The human brain is a product of evolution, with its structures and functions finely tuned for survival in a complex and dynamic world. ANNs, on the other hand, are designed and trained by humans, often for specific tasks that do not require the broad adaptability and understanding that biological systems have. While the morphic program can undoubtedly lead to improvements in ANN performance—making them more efficient, flexible, or capable of generalisation—the extent to which these improvements can replicate the true grounding of biological systems is uncertain.

The morphic problem also touches on more crucial questions about the nature of meaning and understanding in artificial systems. Even if ANNs could be designed to mimic the brain's structure and trained in environments that provide rich sensory data, would this lead to genuine understanding or merely the appearance of understanding? This echoes debates in AI regarding whether machines can ever truly "understand" or whether they simulate understanding through sophisticated pattern recognition. The grounding problem, whether in the context of symbols or vectors, suggests that without a connection to the real-world entities they are meant to represent, any "understanding" in ANNs may be superficial at best.

Ultimately, the morphic problem challenges researchers to think beyond simple structural mimicry and consider the broader context in which biological intelligence operates. It suggests that for ANNs to approach the human brain's brain's capabilities, they must not only mirror its architecture but also engage with the world in a manner that allows for genuine grounding of their internal representations. This may require a fundamental shift in how we design and train ANNs, moving from data-driven approaches to more holistic systems that integrate learning, embodiment, and interaction with the real world.

In conclusion, the morphic problem of artificial neural networks highlights a critical challenge in optimising AI through biological inspiration. At the same time, the brain's architecture provides a powerful model for ANN design; its true strength lies in its deep connection to the real world—a connection that current ANNs lack. To address this, the morphic program must go beyond structural mimicry and seek ways to embed ANNs in environments that allow for meaningful interactions and grounding. Whether through embodied cognition, multimodal learning, or other innovative approaches, the goal should be to create systems that not merely simulate understanding but achieve a form that is as close as possible to that seen in biological systems. Only then can we hope to optimise ANNs in a way that genuinely mirrors the power of the human brain.

5 The definition of morphic relationship in neuromorphic engineering

Neuromorphic engineering is an interdisciplinary field that aims to bridge the gap between artificial neural networks (ANNs) and biological neural networks by capturing and replicating the underlying morphic relations—that is, the structural, functional, and dynamic relationships that define how biological neural systems operate. These morphic relations encompass the intricate interactions of neurons and synapses in the brain to process information, adapt to new stimuli, and enable complex behaviors. The goal of neuromorphic engineering is to design artificial systems that not only mimic the architecture of the brain but also replicate its operational principles. By doing so, it seeks to create more efficient, adaptable, and powerful computing systems that closely resemble human cognition. At the core of the morphic relations that neuromorphic engineering seeks to capture is the fundamental structure of neural networks in the brain. Biological neural networks consist of neurons interconnected by synapses, forming complex circuits that transmit and process information. These neurons exhibit a wide range of behaviors depending on the strength and timing of synaptic inputs, the properties of the neuronal membrane, and the broader network context. Neuromorphic engineering attempts to emulate this complexity by designing ANNs that are deeply rooted in the biological principles governing neural activity. This includes developing artificial neurons that replicate the nonlinear dynamics of real neurons and creating synapses that mimic plasticity—the ability to strengthen or weaken over time—which is essential for learning and memory in the brain. Plasticity is one of the most critical aspects of morphic relations—the brain's ability to adapt its structure and function in response to experience. In biological systems, synaptic plasticity allows the strength of connections between neurons to change based on activity, which is crucial for learning and memory formation. Neuromorphic engineering seeks to incorporate similar mechanisms into ANNs, enabling them to learn and adapt in a manner more akin to human cognition. This involves developing learning algorithms inspired by synaptic plasticity processes observed in the brain. For example, spike-timing-dependent plasticity (STDP), where the timing of neuronal spikes influences synaptic strength, has been a critical focus in neuromorphic research. By implementing STDP in artificial systems, engineers aim to create networks that can learn and adapt in real time, just as the human brain does. Another essential morphic relation that neuromorphic engineering aims to capture is the energy efficiency of the brain. Despite its incredible computational power, the human brain operates on only about 20 watts of power—a fraction of what traditional computers consume. This efficiency arises from several factors, including sparse coding of information (where only a small subset of neurons is active at any given time), parallel processing across multiple neural circuits, and the ability to perform complex computations with low-precision signals. Neuromorphic systems seek to emulate these energy-efficient strategies by developing hardware that performs computations in a manner more similar to the brain. For instance, neuromorphic chips are often designed to use event-driven processing, where computations are performed only when certain conditions are met, much like how neurons fire only when they receive sufficient input. This approach can significantly reduce power consumption, making neuromorphic systems more viable for applications where energy efficiency is crucial. The temporal dynamics of neural activity are another critical aspect of the morphic relations that neuromorphic engineering seeks to replicate. In the brain, neurons communicate through spikes—brief, all-or-nothing electrical impulses—that convey information through their rate and precise timing. The brain's ability to process information rapidly and in real time is largely due to these temporal dynamics, where the timing of spikes can encode complex patterns and relationships. Neuromorphic engineering captures this aspect by designing ANNs that operate on similar temporal principles. Spiking neural networks (SNNs), for example, are a type of ANN where information is encoded in the timing of spikes rather than continuous values, allowing for more biologically realistic models of information processing. By capturing the temporal dynamics of the brain, neuromorphic systems can potentially achieve faster and more efficient processing, particularly in tasks that require real-time responses. Moreover, the connectivity patterns within the brain represent another crucial morphic relation that neuromorphic engineering aims to emulate. The human brain is characterized by a highly complex and non-uniform network of connections, with different regions specialized for various functions and connected in intricate ways that allow for both localized and distributed processing. This connectivity reflects the brain's evolutionary optimization for specific cognitive tasks like vision, language, and motor control. Neuromorphic engineering seeks to replicate these specialized connectivity patterns in ANNs by designing hierarchically organized networks capable of dynamic reconfiguration based on the task at hand. This might involve developing modular networks where different subsets of artificial neurons are dedicated to different tasks or creating networks that can rewire themselves in response to new information, much like how the brain's connectivity can change over time. In addition to these structural and functional aspects, the robustness of biological neural networks is another key morphic relation that neuromorphic engineering aims to capture. The brain is remarkably resilient, capable of functioning effectively despite noise, damage, or variability in its components. This robustness arises from factors like redundancy in neural connections, the ability of neurons to compensate for damaged areas, and the brain's capacity to filter out irrelevant noise. Neuromorphic systems aim to replicate this robustness by designing ANNs that can continue to function even when parts of the network fail or when faced with noisy input. This might involve developing fault-tolerant architectures with multiple pathways for information flow or creating networks that can adapt to changes in their environment without extensive retraining. One of the broader goals of capturing these morphic relations is to move beyond the limitations of current ANNs, which, despite their success in many applications, are still far from replicating the full range of capabilities exhibited by the human brain. While traditional ANNs can be powerful tools for pattern recognition and prediction, they often struggle with tasks requiring real-time processing, adaptability, or understanding of complex, dynamic environments. By capturing the morphic relations that define how the brain processes information, neuromorphic engineering aims to create systems that are not only more efficient and adaptable but also capable of performing a broader range of cognitive tasks in a manner more akin to human intelligence. However, pursuing these goals presents challenges. Capturing the full complexity of the brain's morphic relations requires a deep understanding of both neuroscience and engineering, as well as the ability to translate biological principles into computational models. Moreover, much about how the brain functions—particularly at the level of large-scale networks and their interactions—remains unknown. As such, neuromorphic engineering is as much a process of discovery as it is of design, with researchers constantly refining their models as new insights into brain function emerge. In conclusion, the morphic relations that neuromorphic engineering seeks to capture represent the intricate interplay of structure, function, and dynamics that characterize biological neural networks in the human brain. By emulating these relations, neuromorphic systems aim to replicate the brain's architecture and operational principles, leading to artificial systems that are more efficient, adaptable, and capable of human-like cognition. This approach holds the promise of overcoming the limitations of traditional ANNs, opening up new possibilities for artificial intelligence that more closely mirrors the power and flexibility of the human mind. A relation is morphic if it captures and embodies the structural, functional, and dynamic correspondences between two systems, allowing one system to be understood, analyzed, or replicated in terms of the other. Morphic relations go beyond superficial similarities or mere analogies; they reflect a deep correspondence in the organization, function, and evolution of the systems over time. In the context of neuromorphic engineering, morphic relations specifically refer to the connections between ANNs and biological neural networks in the human brain, where the aim is to replicate not just observable behaviors but the underlying operational principles that govern the brain's function.

6 The criteria for a morphic relation in ANNs

The concept of a morphic relation in the context of artificial intelligence refers to a correspondence between artificial neural networks (ANNs) and biological neural systems, encompassing structural, functional, and dynamic aspects. This relation implies that ANNs can replicate or emulate key properties of biological neural systems to a significant extent. Establishing this relationship requires satisfying specific criteria that capture the fundamental characteristics of both systems.

6.1 Structural Correspondence

Necessary Condition: A morphic relation requires a fundamental structural correspondence between ANNs and biological neural systems. This means that the basic components (neurons, synapses) and their interconnections in ANNs must reflect those found in biological systems.

Sufficient Condition: Structural correspondence is sufficient when the ANN's architecture not only mirrors the components and connections but also allows for the mapping of structural changes from one system to the other while preserving network integrity and functionality.

6.2 Functional Equivalence

Necessary Condition: For a morphic relation to exist, ANNs must replicate key functional operations of biological neural systems, such as learning through synaptic plasticity and processing information via spike-timing dynamics.

Sufficient Condition: Functional equivalence is achieved when ANNs can perform the same tasks as biological systems with comparable efficiency and adaptability, producing outputs that are functionally indistinguishable from those of biological neural networks given the same inputs.

6.3 Dynamic Coherence

Necessary Condition: ANNs must exhibit temporal dynamics that mirror those of biological systems, including adaptation over time in response to stimuli.

Sufficient Condition: Dynamic coherence is sufficient when ANNs not only adapt over time but do so in ways that parallel the timing and sequence of changes observed in biological systems.

6.4 Contextual Integration

Necessary Condition: For a fully morphic relation, the systems must integrate similarly within their broader contexts. This means that the role each system plays in its respective environment—such as how a neural network interacts with other cognitive systems—must be analogous between ANNs and biological networks.

Sufficient Condition: Contextual integration is sufficient when ANNs can be embedded into larger frameworks or ecosystems and continue to function effectively. For example, an ANN with morphic relations to the brain should interact with other AI systems or datasets in a manner similar to how the brain integrates sensory inputs and coordinates responses with other cognitive systems.

6.5 Robustness and Resilience

Necessary Condition: Morphic relations must ensure robustness and resilience in analogous ways. If a biological network can withstand noise, damage, or variability, then the ANN must demonstrate similar robustness to maintain its functional capacity.

Sufficient Condition: Sufficient robustness and resilience are demonstrated when both systems maintain functionality under comparable conditions of stress or perturbation. This includes the ability to recover from disruptions and continue performing essential tasks, reflecting the stability of the morphic relation.

In conclusion, a relation is morphic if it profoundly captures the structural, functional, dynamic, contextual, and resilient aspects of two systems, enabling one to be understood or replicated in terms of the other. These criteria—structural correspondence, functional equivalence, dynamic coherence, contextual integration, and robustness—are both necessary and sufficient to establish a true morphic relation, ensuring that the systems are not just similar on the surface but are fundamentally interconnected in their nature.

7 Metodological challenges

Comparing deep neural networks (DNNs) to human cognition presents significant methodological challenges, particularly when considering the concept of morphic relations. The "morphic problem" underscores the difficulty of optimizing artificial neural networks by merely mimicking the biological structures of the human brain. Meaningful comparisons between DNNs and humans are complex; they require meticulously matched conditions and rigorous experimental designs that account for the profound differences in how artificial and biological systems process information and interact with the world. One primary methodological challenge is the fundamentally different ways these systems are trained and operate. Human cognition develops through a lifetime of sensory experiences, embodied interactions, and social learning within a dynamic environment. In contrast, DNNs are typically trained on static datasets, often consisting of labeled examples that lack the rich, multimodal experiences shaping human understanding. This disparity leads to significant differences in how DNNs and humans represent and process information. To compare these systems meaningfully, researchers must ensure that the conditions under which DNNs are tested closely mirror those under which humans perform similar tasks. Achieving such parity is challenging. The data used to train and test DNNs often lacks the complexity and contextual richness of stimuli humans encounter. For instance, while a DNN might be trained to recognize objects in images based on pixel patterns and statistical correlations, it does so without the broader contextual understanding humans possess. A human recognizes a chair not just by its appearance but also by its function, its relationship to other objects, and its cultural significance. Experimental designs must account for these contextual factors, which are often difficult to quantify and incorporate into the data used for training DNNs. Additionally, DNNs and humans often operate under different constraints and objectives. Humans balance a wide range of cognitive functions—including memory, attention, and emotion—while interacting with a complex environment. DNNs are typically optimized for a single task or a narrow set of tasks, often without consideration for the broader cognitive processes involved in human performance of the same task. This focused optimization can lead DNNs to achieve high performance under controlled conditions, but it does not necessarily mean that their underlying cognitive processes are comparable to those of humans. To address these challenges, experimental designs must carefully control for differences in how DNNs and humans approach tasks. This might involve creating testing environments for DNNs that better simulate human experiences. For example, instead of training a DNN on a static dataset, researchers could use dynamic, interactive environments requiring the DNN to engage with the world more like humans do. Such environments could provide feedback based on the DNN's actions, encouraging the development of representations grounded in real-world interactions. While this approach requires significant advancements in how DNNs are trained and tested, it could lead to more meaningful comparisons with human cognition. Interdisciplinary collaboration is crucial in designing experiments that compare DNNs to humans. Insights from psychology, neuroscience, and cognitive science are essential for capturing the nuances of human thought processes. Similarly, expertise in machine learning and artificial intelligence is vital for understanding the capabilities and limitations of DNNs. By combining these fields, researchers can design experiments that more accurately reflect the conditions under which both DNNs and humans operate, leading to more valid and reliable comparisons. Interpreting results from such comparisons poses its own challenges. Even when DNNs and humans perform similarly on a given task, it does not necessarily mean they use the same underlying processes. A DNN might classify objects based on statistical regularities, while a human relies on visual features, contextual information, and prior knowledge. Without understanding the mechanisms underlying each system's performance, drawing parallels can be misleading. This underscores the importance of not only comparing performance outcomes but also investigating the processes leading to those outcomes. Techniques like neuroimaging in humans and feature visualization in DNNs can provide insights into internal representations and processes, but interpreting these results requires careful consideration of the differences between artificial and biological systems. Another methodological challenge arises from the inherent differences in scale and complexity between human brains and DNNs. The human brain is vastly more complex, with approximately 86 billion neurons interconnected in intricate networks supporting a wide range of cognitive functions. DNNs, while inspired by the brain, are simplified models with far fewer neurons (or units) and connections. This difference in scale means that even if a DNN replicates a specific human behavior, it might do so in a way that is not scalable to the broader, more complex tasks humans perform. Experimental designs must, therefore, be cautious in extrapolating findings from specific tasks to broader claims about human-like cognition in DNNs. The issue of interpretability also plays a crucial role. Human cognition, though not fully understood, is generally interpretable to some extent through introspection, behavioral analysis, and neuroscience. DNNs often function as "black boxes," where the internal processes leading to a particular output are difficult to decipher. This opacity makes it challenging to determine whether a DNN's behavior truly mirrors human cognition or is simply a coincidental outcome of its architecture and training data. Developing methods to improve the interpretability of DNNs, such as analyzing activation patterns or using techniques like saliency mapping, is essential for making more meaningful comparisons. In conclusion, comparing DNNs to humans presents significant methodological challenges stemming from differences in training, information processing, and environmental interactions. The morphic problem highlights the limitations of optimizing DNNs solely by mimicking biological structures, and these limitations extend to experimental designs used in comparisons. To make these comparisons meaningful, researchers must carefully match conditions between the two systems, considering contextual richness, cognitive complexity, and result interpretability. This requires sophisticated experimental designs and collaborative efforts across disciplines to ensure that comparisons between DNNs and human cognition are valid and insightful. By addressing these methodological challenges, we can advance our understanding of both artificial and human intelligence, moving closer to creating AI systems that reflect the capabilities of the human mind.

8 Conclusion

This paper on the assumption of the NE project examines the convergence between natural neural networks (NNNs) and artificial neuromorphic systems (ANNs), presenting an exploration of the "morphic problem" This issue, central to NE, underscores the challenge of translating the flexible, adaptive nature of biological neurons into artificial systems, which often rely on static, rigid architectures. Neuromorphic systems, inspired by the brain's neural architecture, aim to replicate its efficiency, adaptability, and energy usage while addressing the philosophical and technical complexities that arise from this convergence.

One core aspect of the morphic problem involves the difficulty in mapping biological neural networks' dynamic plasticity and continuous learning abilities onto artificial networks. Natural neural networks possess dynamic synapses, nonlinear interactions, and the ability to reorganise themselves in response to stimuli, creating a level of flexibility that artificial systems struggle to mirror. Artificial networks, especially early ANNs, often rely on pre-defined architectures and rules, limiting their adaptability. This fundamental gap represents a critical element of the morphic problem: the tension between the static nature of artificial systems and the fluid, evolving properties of natural neurons.

In addressing the morphic problem, the paper explores the role of spiking neural networks (SNNs), which attempt to bridge this gap by emulating biological neurons' time-dependent, spike-based communication system. SNNs represent a significant advancement in neuromorphic computing by mimicking the asynchronous, event-driven processing found in the brain, such as how neurons fire in response to stimuli or how the brain processes sensory information. However, despite these innovations, challenges persist. SNNs, while capable of performing specific tasks such as pattern recognition and cognitive functions, still fall short of the biological system's scalability, adaptability, and efficiency. The computational power of biological neurons, honed through millions of years of evolution, surpasses what current artificial systems can replicate.

To push beyond these limitations, the paper discusses advanced neuromorphic hardware solutions, such as memristors and crossbar architectures, which offer promising avenues for creating more flexible, scalable, and energy-efficient systems. These technologies can potentially address the morphic problem by enabling artificial systems to emulate the brain's computational efficiency more closely. By moving beyond traditional computing paradigms, neuromorphic engineering aims to create systems that simulate and extend biological neural networks' capabilities, offering a hopeful future for the field.

The morphic problem extends beyond mere technical limitations; it focuses on the nature of intelligence and computation and the possibility of creating genuinely neural-like computation in machines. While neuromorphic systems can approximate the brain's architecture and functions, they remain fundamentally artificial constructs constrained by human design and technological boundaries. The paper argues that solving the morphic problem may require a more profound rethinking of what it means to compute, challenging us to move away from direct imitation of nature toward developing systems that merge biological principles with new computational paradigms.

In conclusion, the morphic problem in neuromorphic engineering is both a technical and conceptual challenge, highlighting the difficulties inherent in replicating the adaptability and plasticity of natural neural networks within artificial systems. By addressing this problem, neuromorphic engineering has the potential to revolutionise computing, paving the way for breakthroughs in fields such as artificial intelligence, robotics, and bioinformatics. For instance, solving the morphic problem could lead to the development of more adaptive and efficient learning algorithms. In robotics, it could enable the creation of more human-like and adaptable robots. In bioinformatics, it could enhance our understanding of biological systems and their computational principles. The paper also underscores that while technology can approximate the workings of the brain, achieving proper neural-like computation may require a fundamental shift in our understanding of biology and computation.

References

1.
2. Fodor J. A. and Pylyshyn, Z. W., 1988, Connectionism and cognitive architecture: A critical analysis. Cognition, 28(1-2), pp. 3-71.
3. Pinker, S. and Prince, A., 1988, On language and connectionism : Analysis of a parallel distributed processing model of language acquisition, Cognition 28(1-2), pp. 73-193.

4. Mandelbaum E., et al., 2022; Problems and Mysteries of the Many Languages of Thought, Cognitive Science 46 (2022), pp. 1-6.
5. Marcus G., 2001, The Algebraic Mind, Cambridge: MIT Press.
6. Marcus, G., 2018; Deep learning : A Critical Appraisal, arXiv:1801.00631v1
7. Quilty-Dunn et al., 2023, Quilty-Dunn J, Porot N, Mandelbaum E. The best game in town: The reemergence of the language-of-thought hypothesis across the cognitive sciences. Behavioral and Brain Sciences. 2023;46:e261.
8. Lake, B.M., Baroni, M. Human-like systematic generalization through a meta-learning neural network. Nature 623, 115–121 (2023)
9. Shikhar Murty, Pratyusha Sharma, Jacob Andreas, and Christopher Manning, 2023, Grokking of Hierarchical Structure in Vanilla Transformers. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), pages 439–448, Toronto, Canada. Association for Computational Linguistics.
10. Tsur, Elishai Ezra, Neuromorphic Engineering The Scientist's, Algorithms Designer's and Computer Architect's Perspectives on Brain-Inspired Computing, Abingdon: CRC Press.